RETINA

REal time Tracking Innovative Approach

A specialized processor for track reconstruction at the LHC crossing rate

Giovanni Punzi

Presentazione nuova iniziativa CSN5

Ferrara, 30/9/2014

Che cosa e' RETINA

- Progetto nato entro i gruppi LHCb-Pisa e LHCb-Milano.
- Motivazione: ricostruzione di tracce in tempo reale a 40MHz.
- Come: utilizzare un nuovo algoritmo di tracking, completamente parallelo, per realizzare tracking a bassa latenza e alto throughput.
- Target iniziale: migliorare le prestazioni del trigger di LHCb.
- Simulazioni software promettenti → **<u>occorre prototipo hardware</u>**
- Perche' una iniziativa di CSN5:

- LHCb ha chiuso lo R&D per l'upgrade (2020), con la scelta di soluzioni piu' convenzionali e gia' ben sviluppate, prima che il nostro progetto potesse arrivare a maturazione.
- Il nostro progetto e' generale e non finalizzato a un esperimento

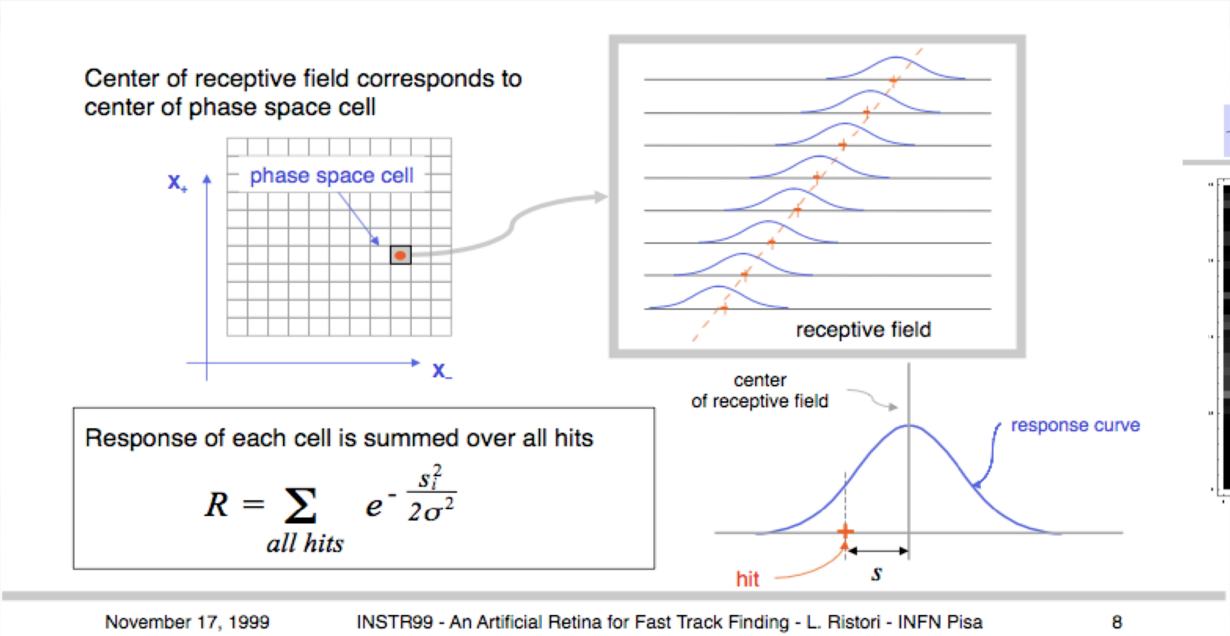
Perche' e' un challenge

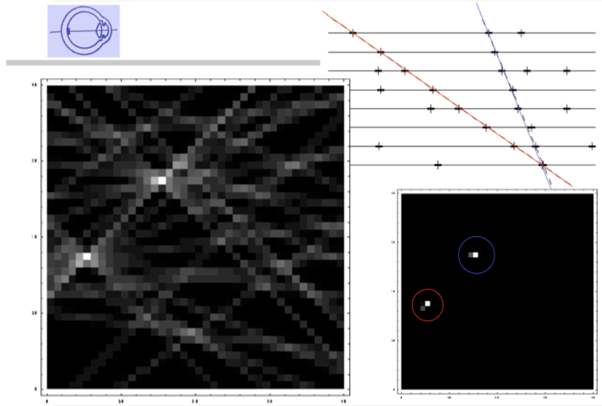
Name	Tech.	Exp.	Year	Event rate	clock	cycles/event	latency
XFT	FPGA	CDF-L1	2000	2.5 MHz	200 MHz	80	4μs
SVT	AM	CDF-L2	2000	0.03 MHz	40 MHz	~1600	<20μs
FTK	AM	ATLAS-L2	2014	0.1 MHz	~200 MHz	~2000	O(10μs)
Compare with the requirements of a L0@LHC:							
?	***?***	***LHC-L0***	***~2018***	***40MHz***	***~1GHz***	***~25***	***few μs***

- The task of L0 tracking at LHC appears daunting despite the progress of electronics – cannot just count on Moore...
- Any complex tracking has required > $O(10^3)$ clock cycles/event (both in latency and throughput)
- No known example of a system making a non-trivial pattern reconstruction in O(25) time units

The Way Forward: a "cellular" tracking algorithm

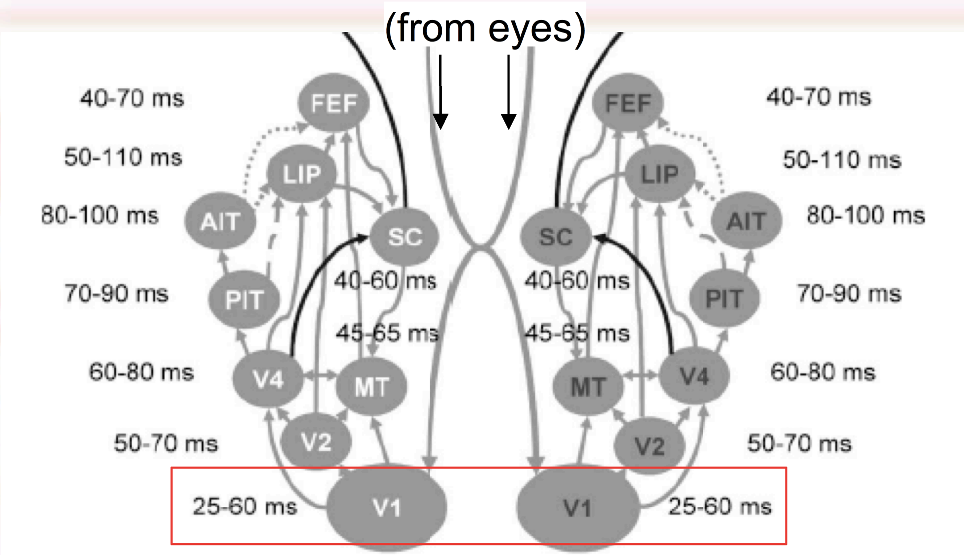
Inspired by mechanism of visual receptive fields [D.H. Hubel, T.N. Wiesel, J. Physiol. 148 (1959) 574],



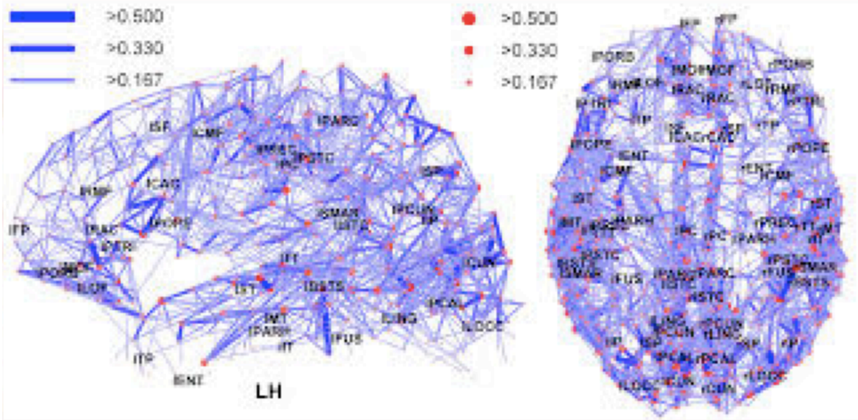


- Basic algorithm proposed by Luciano Ristori back in 2000: *"An artificial retina for fast track finding"* [NIM A453 (2000) 425-429]
  Related to even older "Hough transform" [P.V.C. Hough, Conf.Proc. C590914 (1959) 554].
  There are however, significant differences.
- Totally parallelize everything, no serialization until tracks are found
- Interpolation of analog responses saves internal storage. Also, makes it easier to deal with "missing layers".
- Only a toy model in 2000 → Needed serious developments

Where did the name "RETINA" come from ?



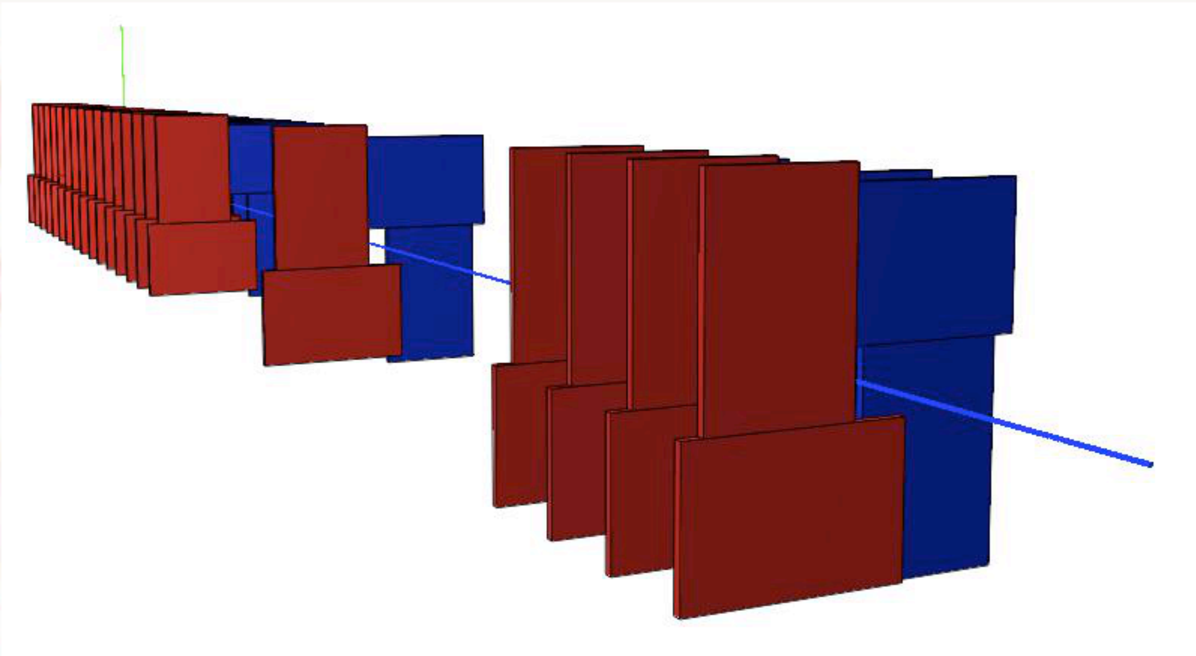
Adapted from H. Kirchner, S.J. Thorpe / Vision Research 46 (2006) 1762–1776



- Analogy with neural organization of retina and associated circuitry
- Early visual areas (V1) in human brain produce a recognizable sketch of the image in ~30ms.
- THE HOPE: Maximum neuron firing frequency is ~1kHz → Adimensionally, same speed of LHC L0 **~30 t.u:** **what we want.** Experimental evidence that V1 functionality can be quantitatively modeled as a "trigger".

[Del Viva MM, Punzi G, Benedetti D PloS one (2013) - DOI:10.1371/journal.pone.0069154]

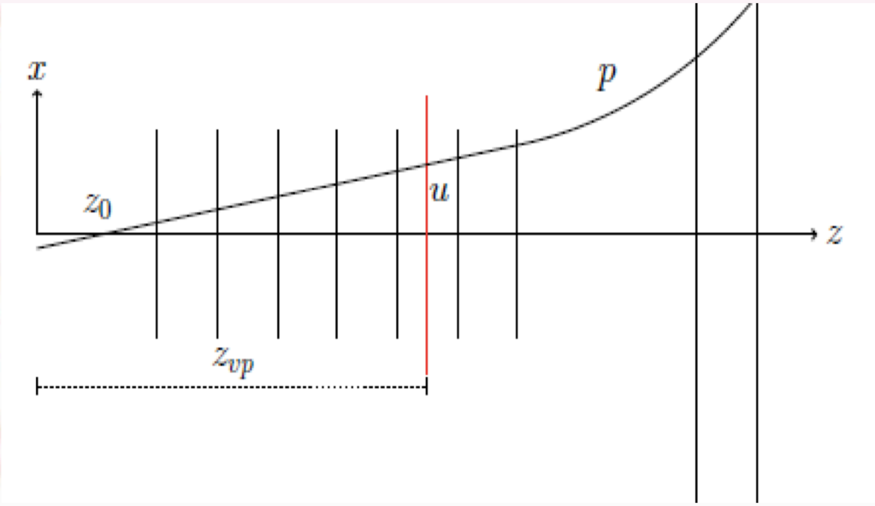
Back to today's reality: Tested on a realistic LHC detector geometry



- Benchmark based on LHCb upgraded VELOPIX detector
- Picked a 6-layer telescope + 2-layer strip detector (UT)
- Motivation: a "significant piece" of the tracking process

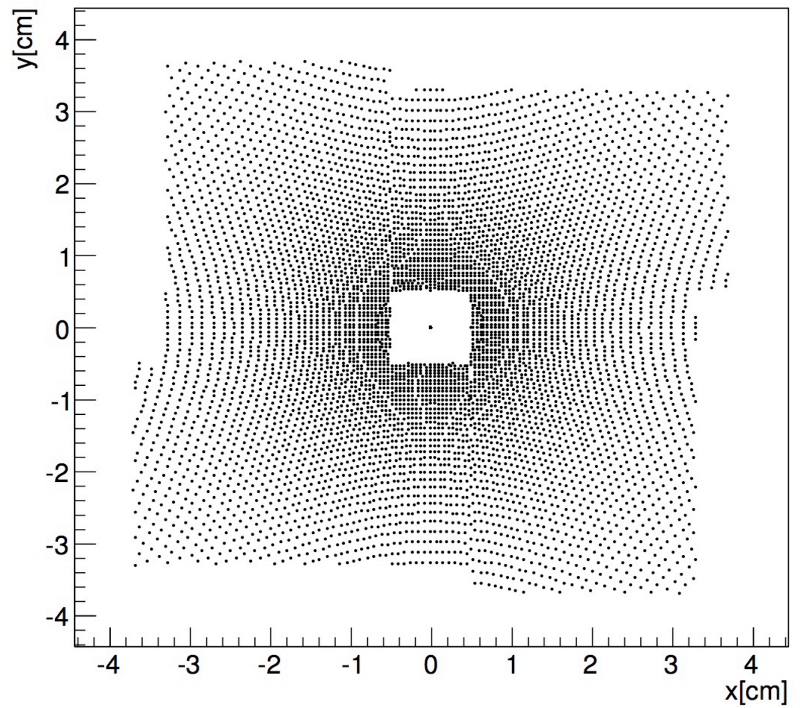
Retina-Tracking in a 3D geometry

- An array of pixel detectors
- Each detector plane provides a (x,y) point at fixed z
- Measure straight tracks in 3D (4 parameters)
- e.g.: θ , θ , z_0, d (impact parameter)
- In case of presence of magnetic field, an additional pamater *p* is sufficient
- Does not need to assume B uniform, or perfect alignment

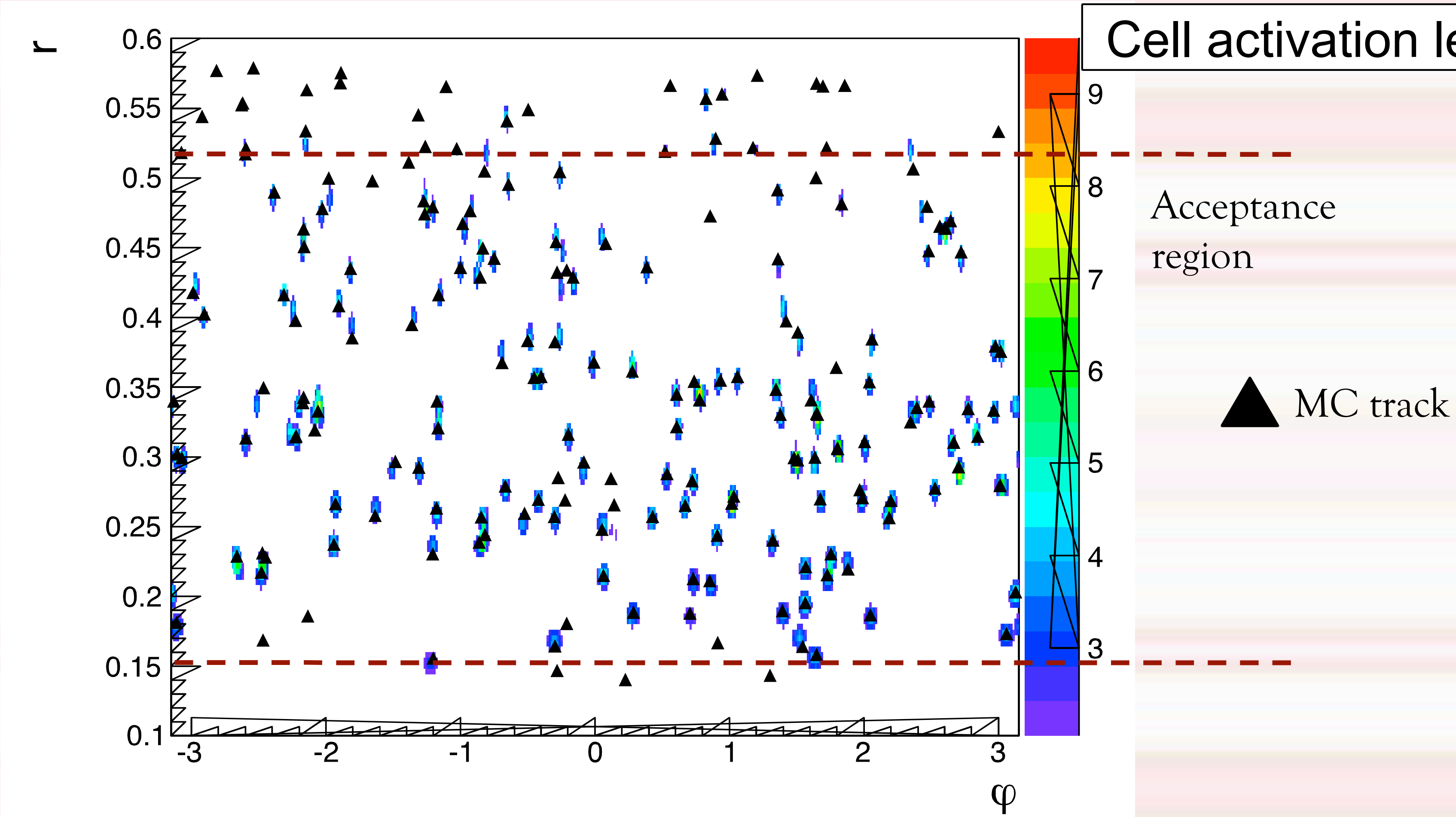


Mapping of detector to a receptor cell array

- Intersection of “base tracks” with detectors gives a map of “nerve endings”
- Every hit on the detector produces a signal on nearby receptors, depending on distance
- I skip on several subtleties. For instance, effective operation require distribution to be non-uniform.
- (not unlike the distribution of photoreceptors in visual system – but it is all virtual in our case, that is, implemented in the electronic network connections)



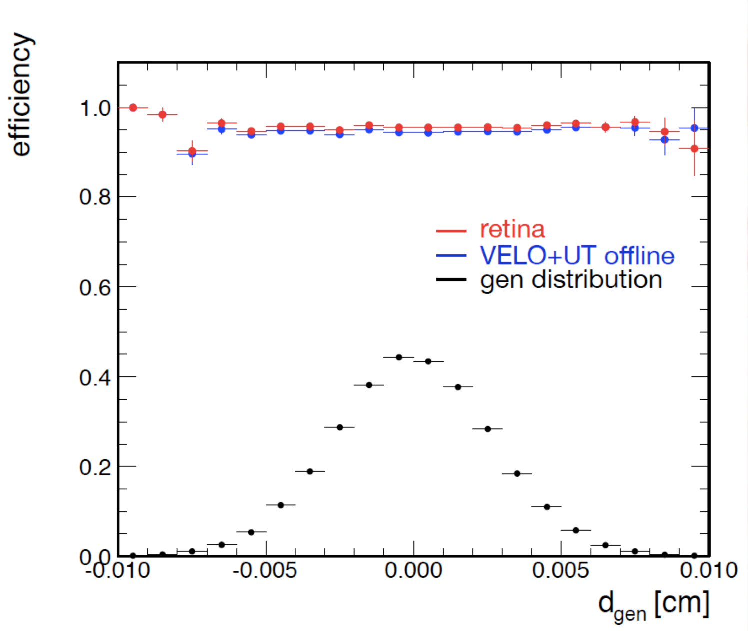
Simulated retina response on LHCb full-MC

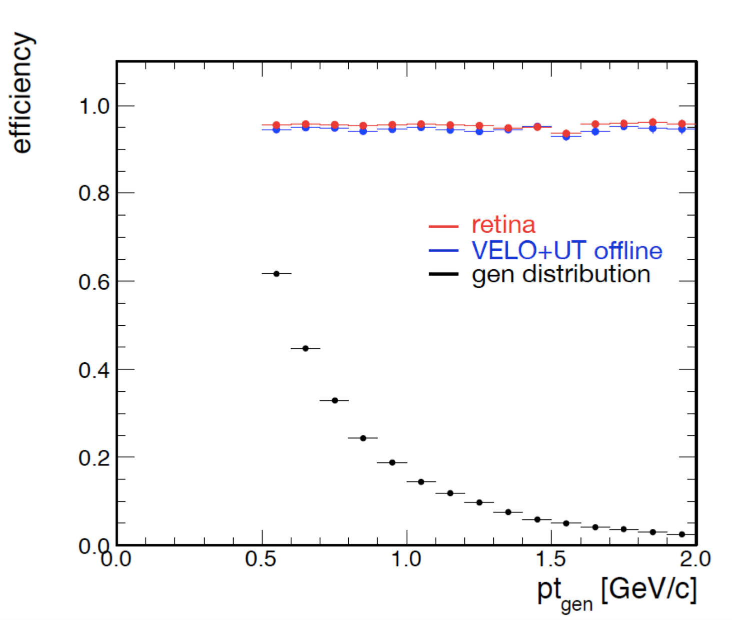


- Colors show activation levels of the 50k retina cells

Performance evaluated on Realistic LHCb MC at full upgrade luminosity

[study available as LHCb public note: LHCb-PUB-2014-026]





- Accounts for all detector effects. Average 7.6 interactions/crossing.
- Efficiency/ghost rate performance comparable to offline reconstruction.

Performance evaluated on Realistic LHCb MC at full upgrade luminosity

[available as LHCb public note: LHCb-PUB-2014-026]

$\nu = 7.6$	ϵ_{ret}	ϵ_{off}	$\eta_{ret\|off}$	$\eta_{off\|ret}$
Longable	0.95	0.95	0.04	0.03
$B_s^0 \to \phi\phi$ signal tracks, longable	0.97	0.97	0.02	0.02
$D^{*+} \to D^0\pi^+$ signal tracks, longable	0.97	0.97	0.03	0.02
$B^0 \to K^*\mu\mu$ signal tracks, longable	0.98	0.98	0.01	0.01

$\nu = 11.4$	ϵ_{ret}	ϵ_{off}	$\eta_{ret\|off}$	$\eta_{off\|ret}$
Longable	0.95	0.94	0.05	0.04
$B_s^0 \to \phi\phi$ signal tracks, longable	0.97	0.97	0.02	0.02

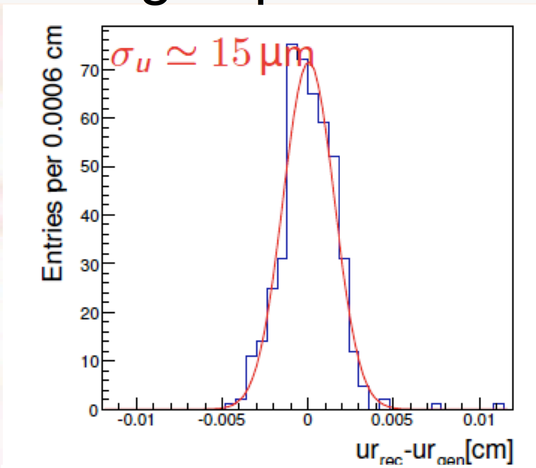
Table 2: Performances of TPU and VELO-UT offline algorithms averaged on 10^4 minimum-bias events and benchmark signal samples generated at $\nu = 7.6$ at $L = 2 \times 10^{33}$ cm^{-2}s^{-1} (top) and $\nu = 11.4$ at $L = 3 \times 10^{33}$ cm^{-2}s^{-1} (bottom). Momentum criteria of $p > 3.0$ GeV/c and $p_T > 0.5$ GeV/c are applied on generated particles only. Generated particles are required not to be electrons and be associated with at least three hits on selected VELO layers and one hits on both the axial UT layers. Only tracks in the fiducial region of $\max(|u|, |v|) < 0.35$ are considered; a fiducial requirement of $|z| < 150$ mm is applied on generated tracks.

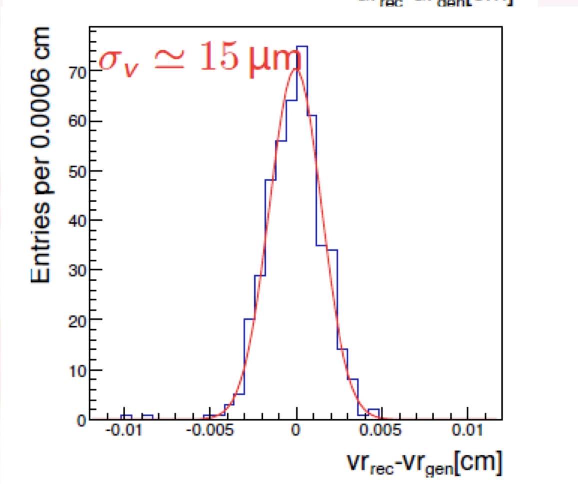
	g_{ret}	g_{off}
$\nu = 7.6$	0.08	0.06
$\nu = 11.4$	0.12	0.08

Table 3: Ghost rates of TPU and VELO-UT offline algorithms, averaged on 10^4 minimum-bias events generated with $\nu = 7.6$ at $L = 2 \times 10^{33}$ cm^{-2}s^{-1} and $\nu = 11.4$ at $L = 3 \times 10^{33}$ cm^{-2}s^{-1}.

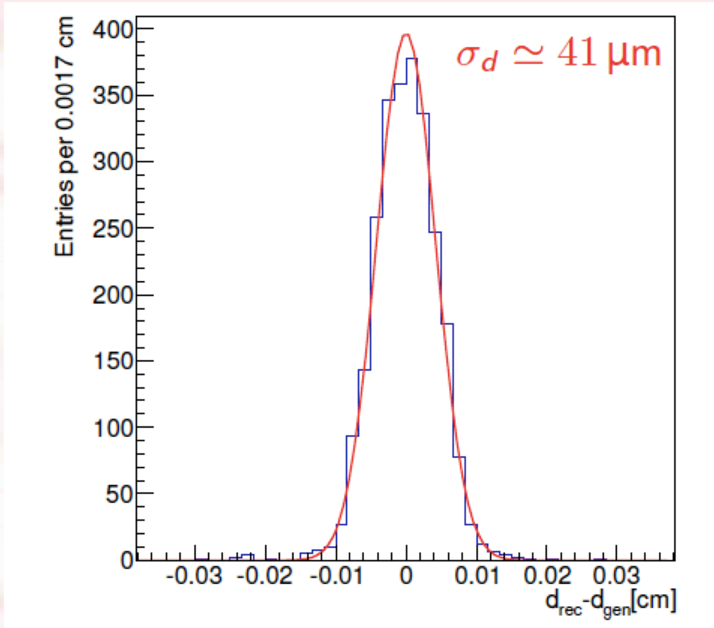
Offline-like Resolutions are obtainable [preliminary]

Main grid parameters





Impact parameter resolution

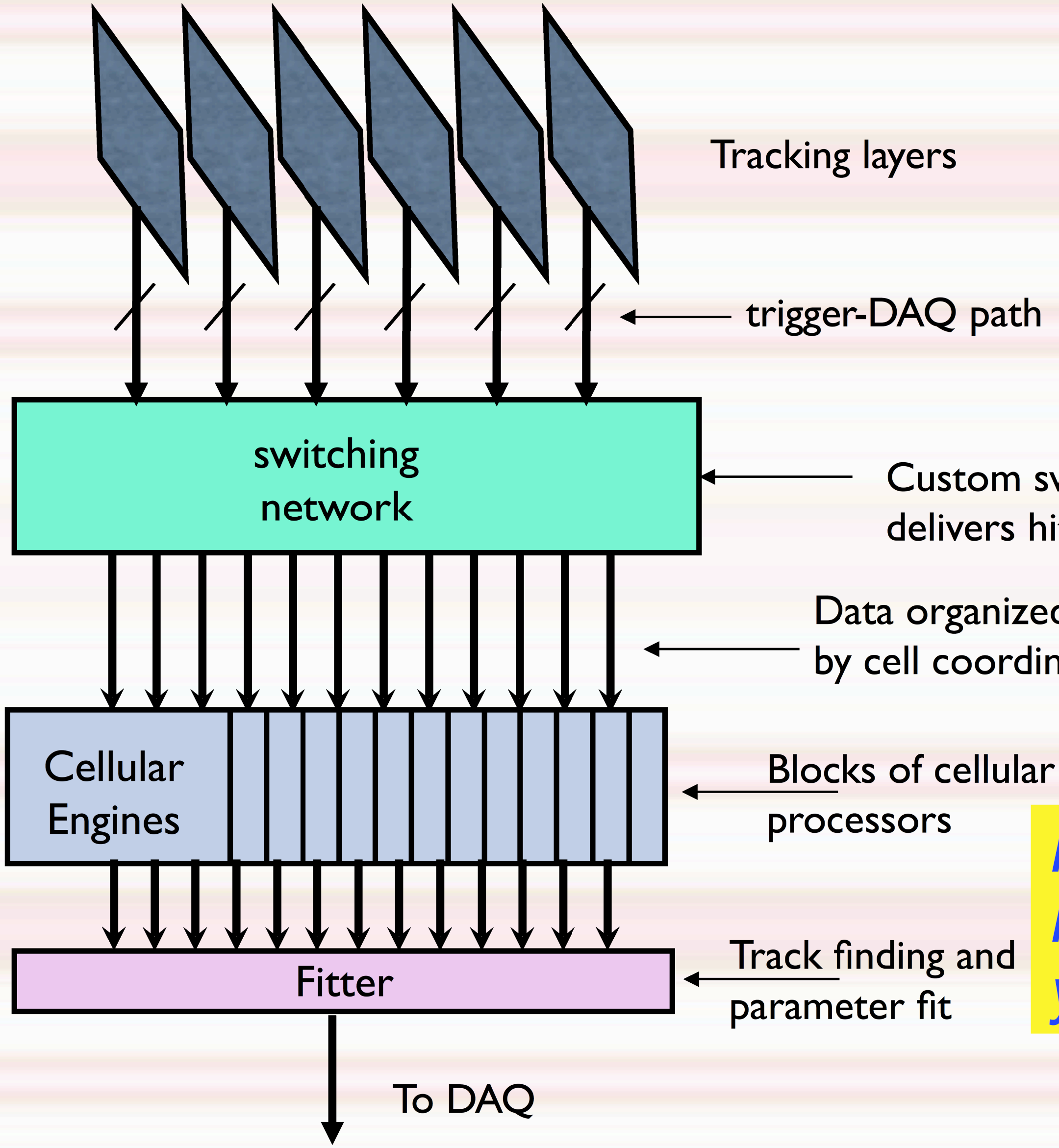


(Compare with offline-style fit: σ = 40μm)

Simulation parameters:
- Pixel size 55μm, res.~12μm
- Hit eff. ~95%
- # of cells 50,000

IMPLEMENTATION STUDIES

System Architecture



Low latency

Task	Latency (cycles)
Switch in readout board	15
Switch in TPU – dispatcher	15
Switch in TPU – fanout	6
Engine processing	70
Clustering	11
Output data	10
Total	< 150

Hi throughput: Preliminary studies yield ~30 cycles/event

FPGA Implementation Studies

- Target high-end FPGA devices.
  - Large I/O capabilities: now O(Tb/s) with optical links !
  - Large internal bandwidth – a must !
  - Fully flexible, easy to program and simulate
  - Steep Moore's slope, and easy to upgrade
  - Highly reliable, easy to maintain and update
  - Industry's method of choice for complex project with a small number of pieces (CT scanners, high-end radars...)

- Start with current, modest cost FPGAs:
  - Altera Stratix III devices, with 200 LE
  - Chose to use NA62 DAQ boards “TEL62”, developed by NA62-Pisa
  - Use ***both*** for *demonstrator* AND *data feeder* at high speed !
    - Easy to procure, and expertise on hand: installed in next run!
    - Produced by INFN-Pisa: software, expertise, etc
  - Build a silicon telescope, to demonstrate a real detector with cosmic
  - After gaining experience, move to new generation (Altera gen 10)

  → more expensive, but more powerful: run clocks close to 1GHz.

FPGA Implementation Studies

- Target high-end FPGA devices.
  - Large I/O capabilities: now O(Tb/s) with optical links !
  - Large internal bandwidth – a must !
  - Fully flexible, easy to program and simulate
  - Steep Moore's slope, and easy to upgrade
  - Highly reliable, easy to maintain and update
  - Industry's method of choice for complex project with a small number of pieces (CT scanners, high-end radars...)

- Start with current, modest cost FPGAs:
  - Altera Stratix III devices, with 200 LE
  - Chose to use NA62 DAQ boards “TEL62”, developed by NA62-Pisa
  - Use ***both*** for *demonstrator* AND *data feeder* at high speed !
    - Easy to procure, and expertise on hand: installed in next run!
    - Produced by INFN-Pisa: software, expertise, etc

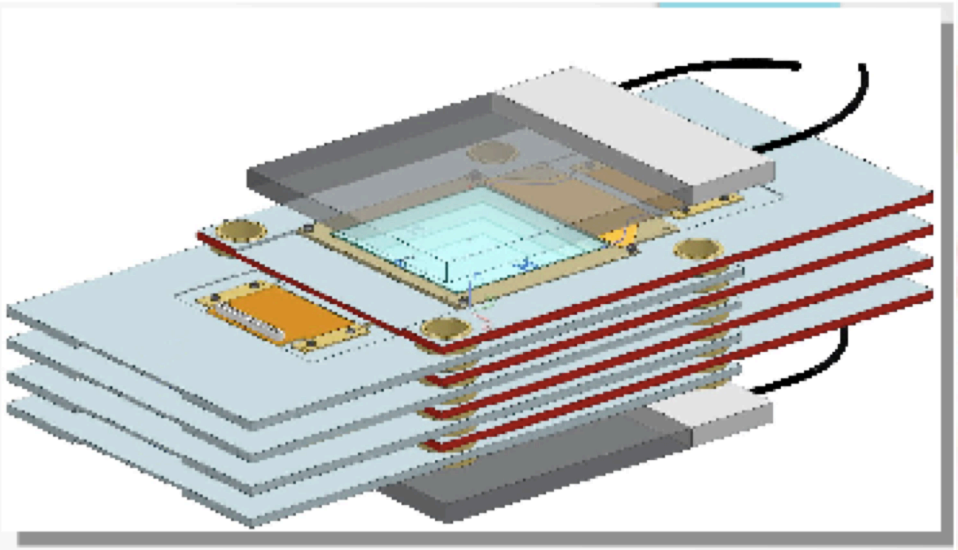
  - Build a silicon telescope, to demonstrate a real detector with cosmic

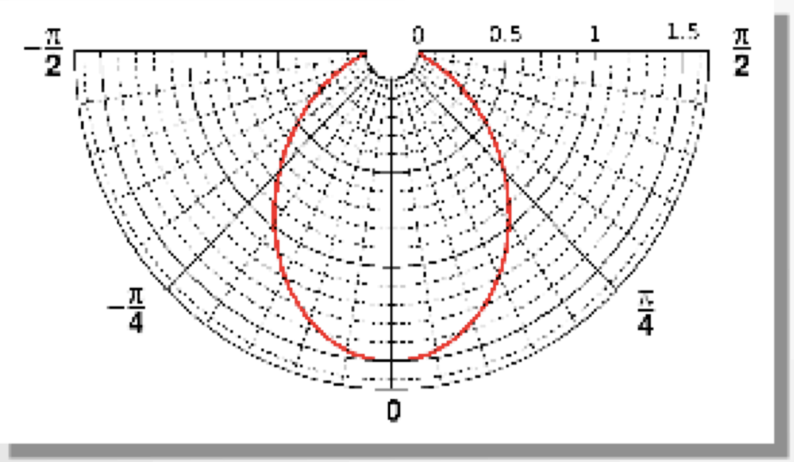
  - After gaining experience, move to new generation (Altera gen 10)

  → more expensive, but more powerful: run clocks close to 1GHz.

Silicon Telescope design

- **8** single-sided strip detectors
- STM OB2 sensors:
  ~10cmx10cm active area
  512 strips
  183 µm pitch
  500 µm thickness
- **0.8cm** distance between planes

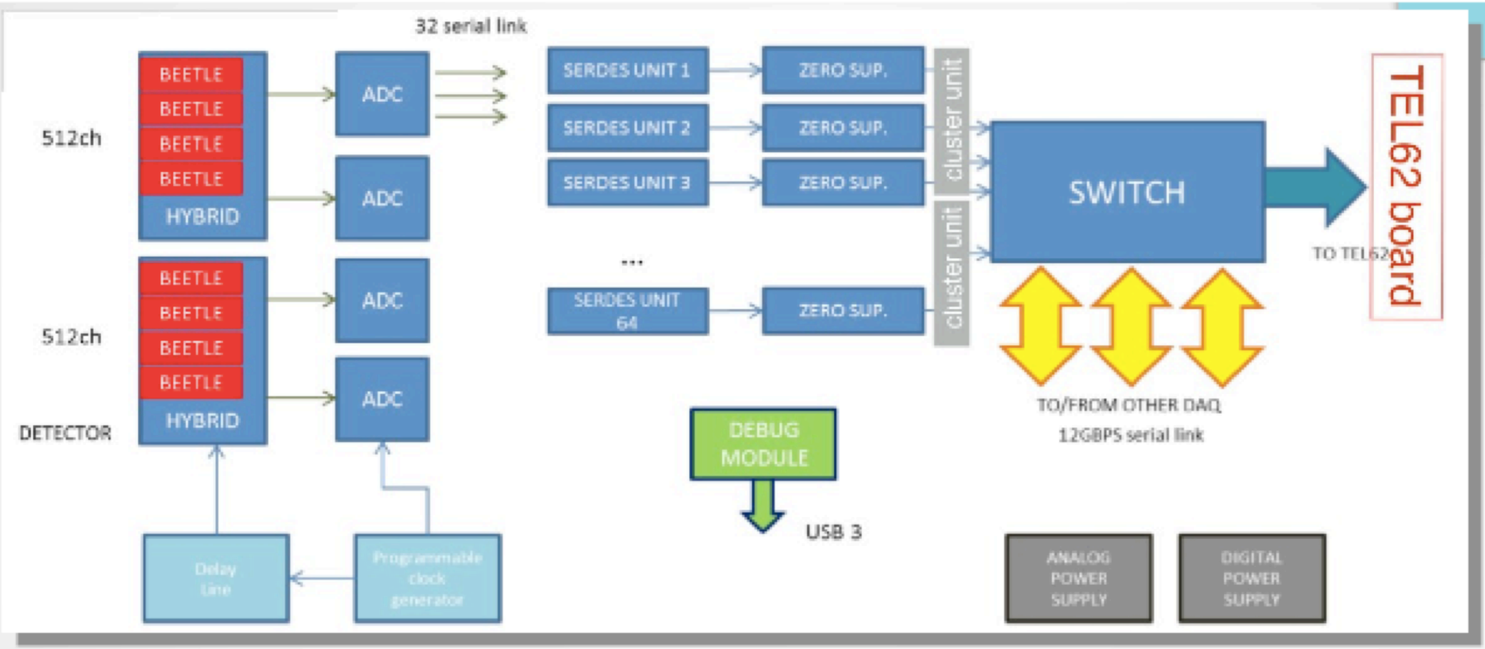




- **1 Hz** expected rate
- Angular distribution:

$$\frac{1}{N_0}\frac{dN}{dcos(\theta)} = \frac{3}{2}cos^2(\theta)$$

Silicon Telescope DAQ



- **4 Beetle chip** for 1 detector readout (4x128=512chan.s)

- **Custom DAQ** board based on Xilinx Kintex 7 FPGA: 1 board for **2 detectors**
  - Maximal readout rate 1.25 Mhz
  - **12-bit ADC** and **zero suppression** digitization
  - **Clustering** for adjacent strips

- Data **delivered** to 4 Altera Stratix III FPGAs of the TEL62 board using a **full mesh switch**

Plan

- Fase 1 (2015) Preparazione setup di test e prototipo con schede TEL62 a PI, e realizzazione di dimostratore basato su rivelatori a silicio e TEL62 e sistema di test a MI.

- Fase 2 (2016) Test di prototipo a PI con dati simulati a 1MHz. Test di dimostratore con dati simulati e con raggi cosmici a MI.

- Fase 3 (2017) Costruzione di piccolo prototipo a full-speed (40MHz) a PI. In caso di risultati positivi, assemblaggio e test di prototipo di maggiori dimensioni per run a test-beam, o parassitico in presa dati di un esperimento LHC.

Person Power

Ricercatori					
Nome	Età	Contratto	Qualifica	Aff.	%
1 Bedeschi Franco		Dipendente	Dirigente di Ricerca	CSN I	20
2 Cenci Riccardo		Associato	Assegnista	CSN I	45
3 Morello Michael Joseph		Associato	Ricercatore	CSN I	30
4 Punzi Giovanni		Associato	Prof. Associato	CSN I	20
5 Ristori Luciano Francesco		Dipendente	Dirigente di Ricerca	CSN I	80
6 Stracka Simone		Associato	Assegnista	CSN I	30
7 Walsh John Joseph		Dipendente	Primo Ricercatore	CSN I	20
			Numero Totale Ricercatori	7	FTE: 2.5

Tecnologi					
Nome	Età	Contratto	Qualifica	Aff.	%
1 Spinella Franco		Dipendente	Tecnologo	CSN I	20
			Numero Totale Tecnologi	1	FTE: 0.2

PISA (RL M. Morello) 2.7 FTE

Strong experience in
- Trigger/DAQ (SVT)
- Tracking

Ricercatori					
Nome	Età	Contratto	Qualifica	Aff.	%
1 Fu Jinlin		Associato	Assegnista	CSN I	20
2 Geraci Angelo		Associato	Prof. Associato	CSN I	20
3 Neri Nicola		Dipendente	Ricercatore	CSN I	20
			Numero Totale Ricercatori	3	FTE: 0.6

Tecnologi					
Nome	Età	Contratto	Qualifica	Aff.	%
1 Abba Andrea		Associato	Assegnista	CSN I	30
2 Caponio Francesco		Associato	Assegnista	CSN I	30
			Numero Totale Tecnologi	2	FTE: 0.6

MILANO (RL N.Neri) 1.2 FTE

Strong experience in
- Electronics (Politecnico)
- Silicon detectors (UT)