



First look at data compression

Konstantin Androsov – INFN Pisa & University of Siena

Massimo Minuti – INFN Pisa

Fabrizio Palla – INFN Pisa

Stamatios Poullos - INFN Pisa & University of Siena

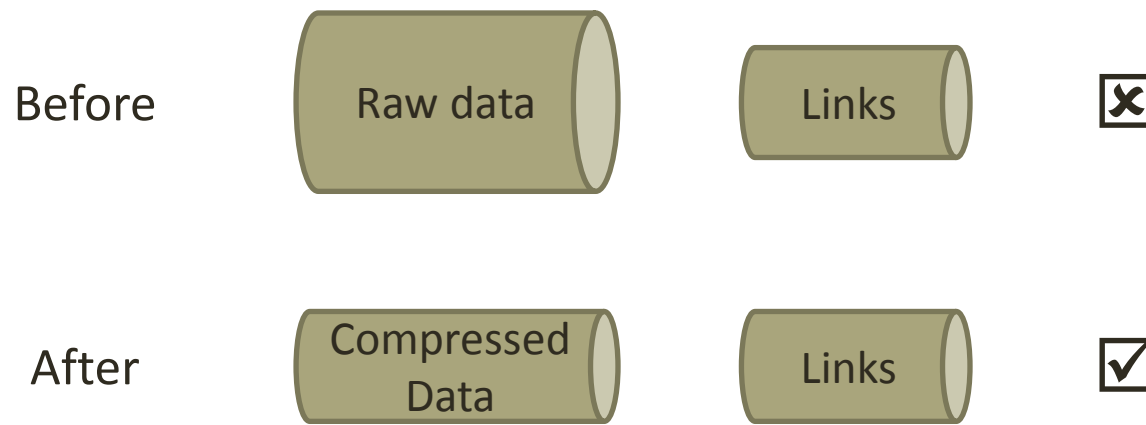
Vth INFIERI workshop

April 2015 – CERN, Geneva

- Target: CMS pixel detector
- Hit rate estimation for inner barrel layer at 140 Pile-Up is about 2 GHz/cm²
- Expected high readout rate (~ 4.8 Gbits/s per chip for 1MHz trigger rate)
- 2.4 Gbits/s max bandwidth per chip (2 E-links)



- An efficient transfer protocol should be developed to collect all measurements and ensure good detector performance
- Lossless compression with decreased output rate to reduce the usage of links
- For 1 MHz L1 trigger accept rate and 320 MHz working clock: 320 clock cycles available!!!





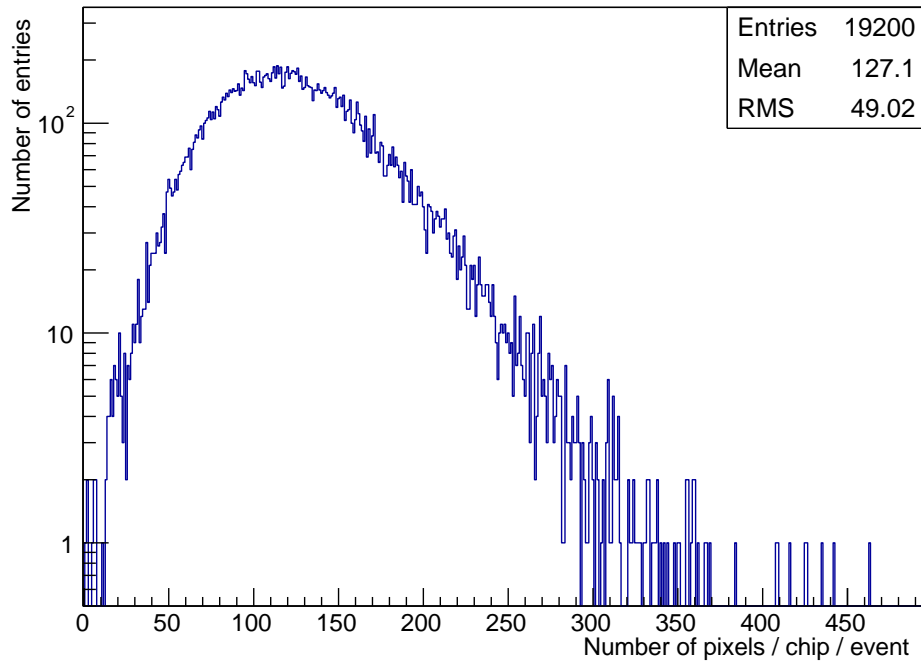
Simulation



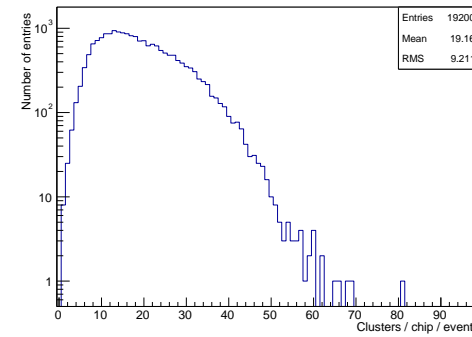
- BPIX Layer1: 3.0 cm, $\text{abs}(z) < 26$ cm
- Cell geometry: 25(bending)x100(beam)x150(thickness) μm^3
- 1 module = 1x4 chips
 - Number of columns per module (chip) : 648 (162)
 - Number of rows per module (chip) : 648 (648)
- Digitizer threshold: 1500 electrons
- Total number of simulated events: 50
- Used 5bit ADC for 560 electrons per ADC count.

SIM results provided by E. Migliore.

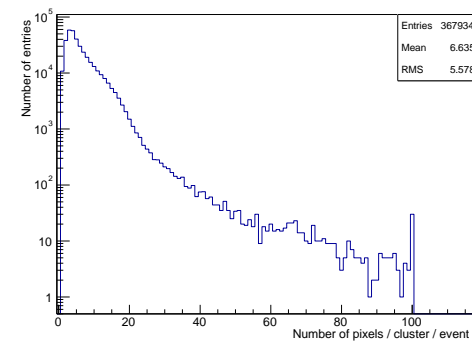
Number of active pixels per chip



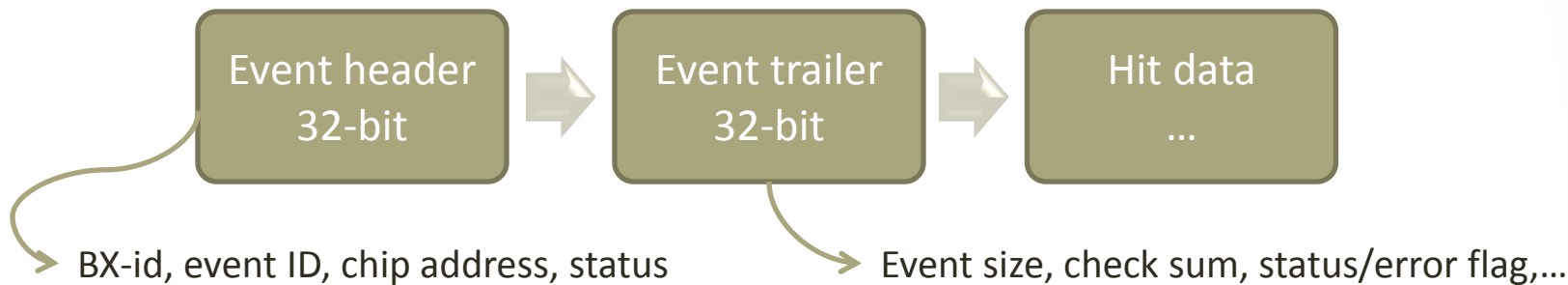
Number of clusters per chip



Number of pixels per cluster



- Using readout format proposed in the draft of Phase 2 pixel system and read-out chip



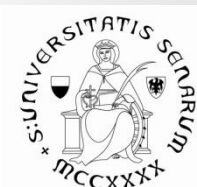
- Default hit data representation:

- (30 bits/pixel)





Default representation



Expected (Optimal encoding)

Alphabet entropy

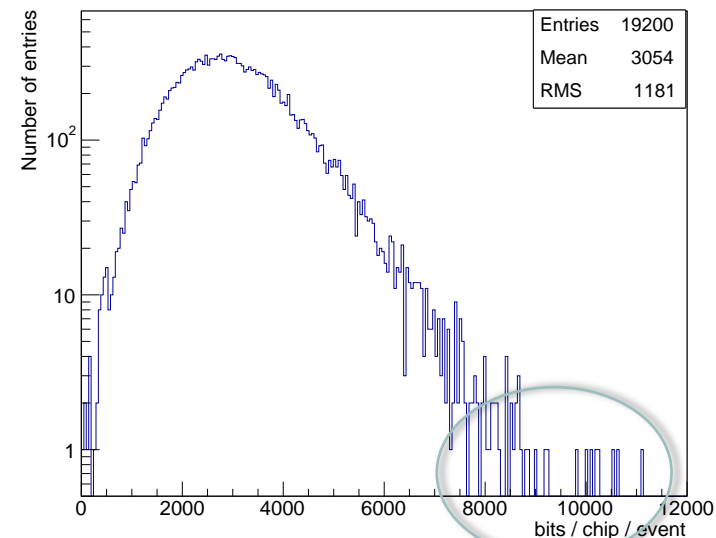
$$H = - \sum_i p_i \log_2(p_i)$$

$$\langle N_{bits} \rangle = (H_x + H_y + H_{ADC}) \langle N_{pixels} \rangle$$

$$\begin{aligned} \langle N_{bits} \rangle &\approx (9.34 + 7.34 + 4.43) \\ &\quad \times 127.1 \\ &\approx 2683.13 \text{ bits/chip/event} \end{aligned}$$

- X and Y entropy is high that makes costly to send unprocessed coordinates per each pixel

Results (No compression)



big tails...

Vth INFIERI workshop

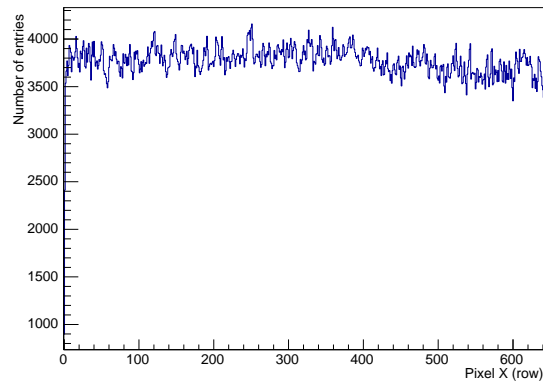
April 2015 – CERN, Geneva



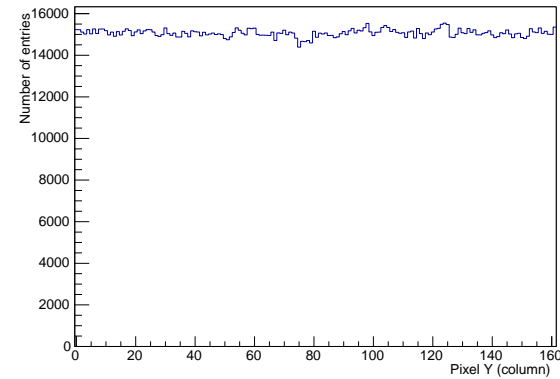
Characteristic distribution



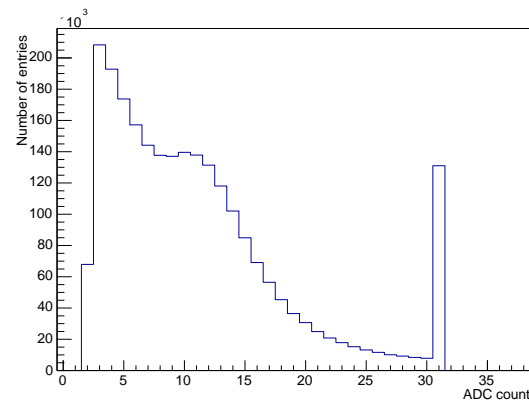
Pixel X



Pixel Y



Active ADC



Vth INFIERI workshop

April 2015 – CERN, Geneva



Arithmetic compression



- 2 approaches
 - Single pixel representation (without zero suppression)
 - Delta representation (zero suppressed data)
- Arithmetic encoding is a form of entropy encoding used in lossless data compression

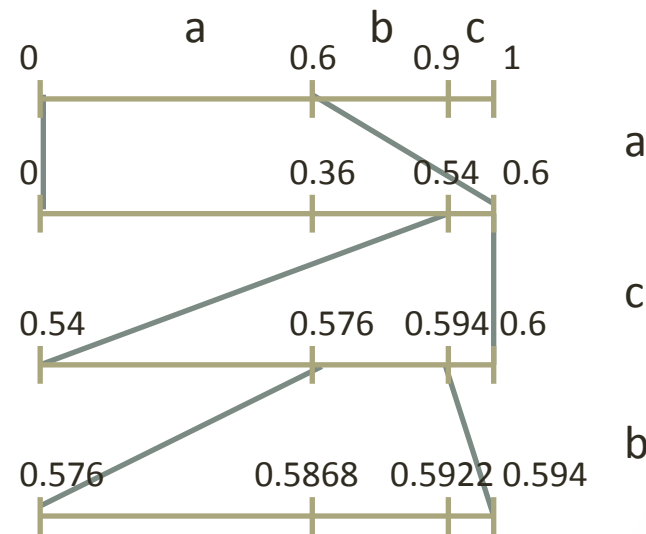
Example:

$P(a)=0.6$

$P(b)=0.3$

$P(c)=0.1$

Phrase: “acb”

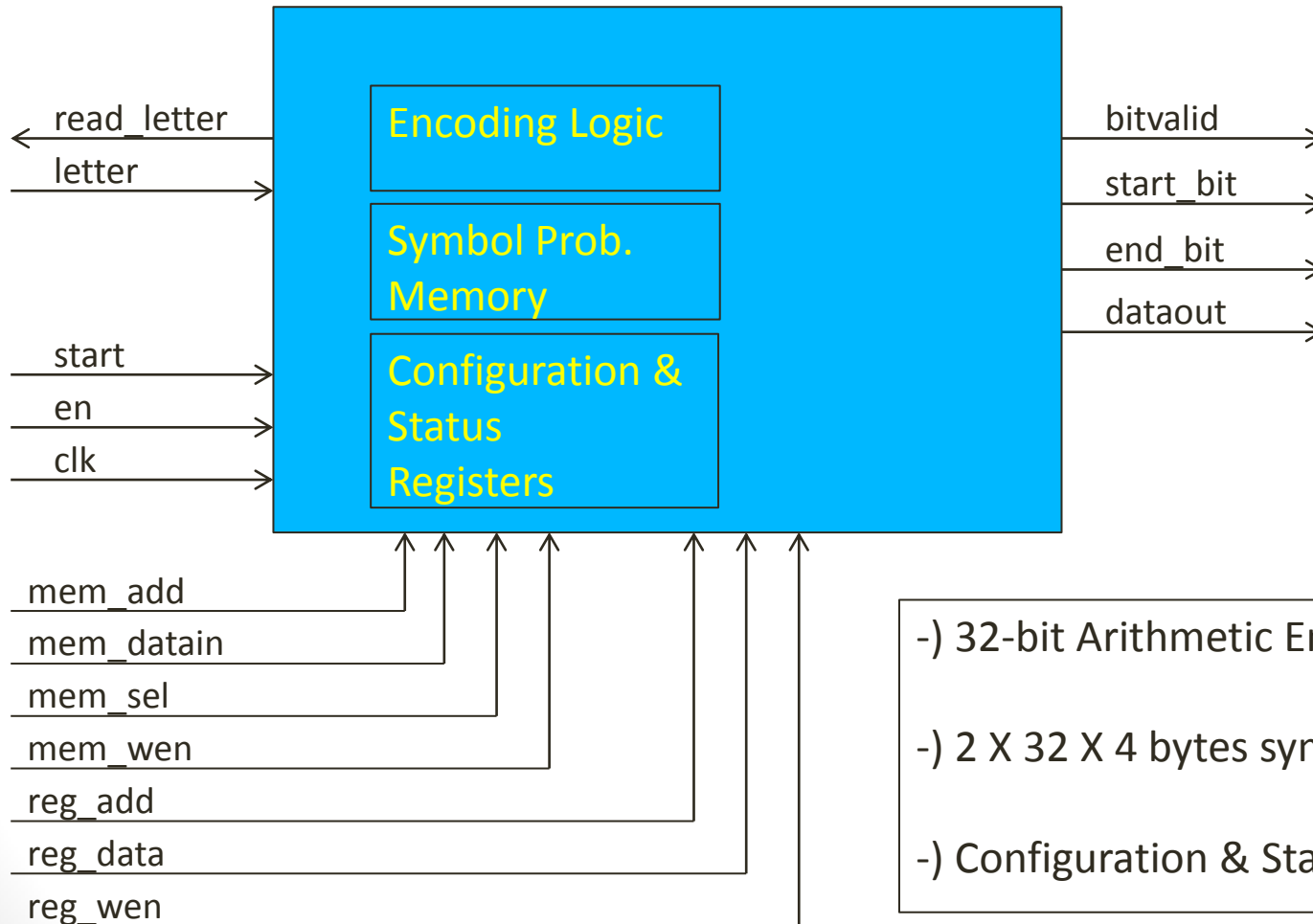


Vth INFIERI workshop

April 2015 – CERN, Geneva



Arithmetic Compressor – HDL implementation



-) 32-bit Arithmetic Encoding Logic;
-) 2 X 32 X 4 bytes symbol Mem.;
-) Configuration & Status regs;



Synthesis/Simulation results



Arithmetic Compressor logic synthesis (IBM 130nm SC library):

cells	49918	
Area	95843	μm^2
Power	75,05	mW@320MHz

Arithmetic Compressor Logic HDL simulation

- performed on L1 simulated data.
- Current performance: 4 clock cycles per pixel(avg).
- Further optimizations expected.



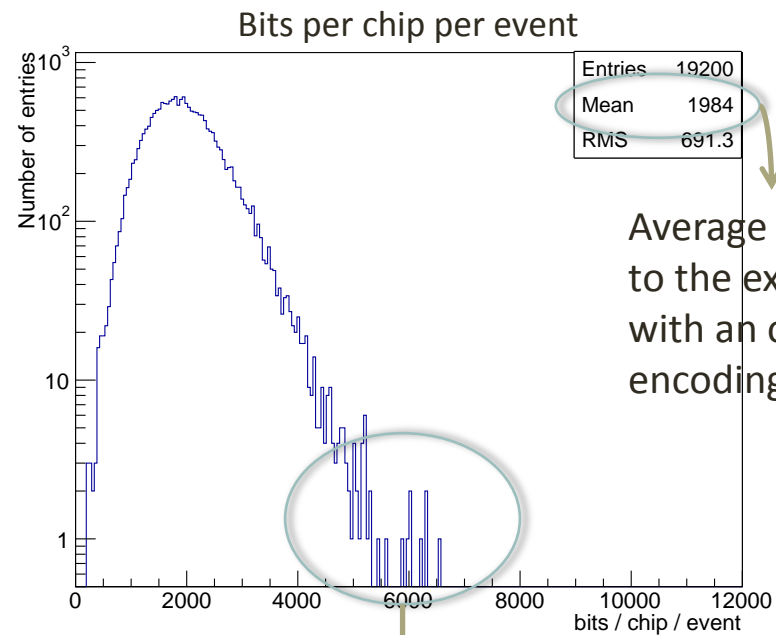
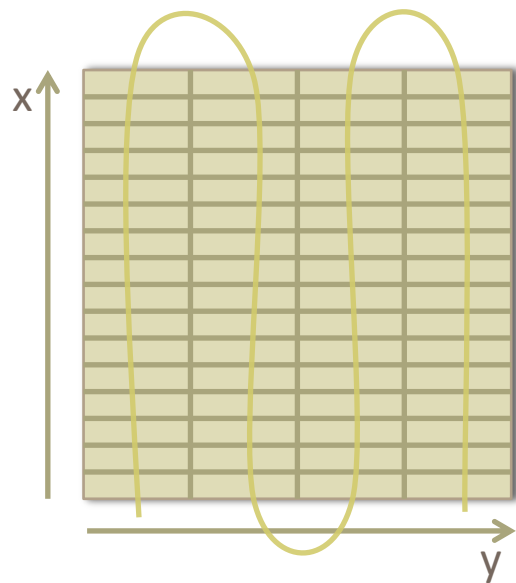
Single pixel representation



$$\langle N_{bits} \rangle = H_{ADC} N_{pixels}$$

$$\langle N_{bits} \rangle \approx 0.019 \times 648 \times 162 \approx 1978.88 \text{ bits/chip/event}$$

- Compression ratio ≈ 1.5
- 1ms for encoding



tails are reduced

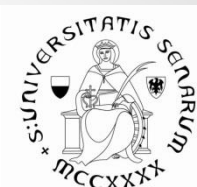
Average is close to the expected with an optimal encoding

Vth INFIERI workshop

April 2015 – CERN, Geneva



Delta representation



- Sending $(\Delta x = x_n - x_{n-1}, \Delta y = y_n - y_{n-1}, adc)$ for each pixel
- Position entropy depends on pixel ordering:

Pixel ordering	H Δx	H Δy	H pixel position
By columns	5.46	1.67	7.13
By arrival	6.70	3.01	9.71
By rows	2.25	3.70	5.95

Encoding column by column

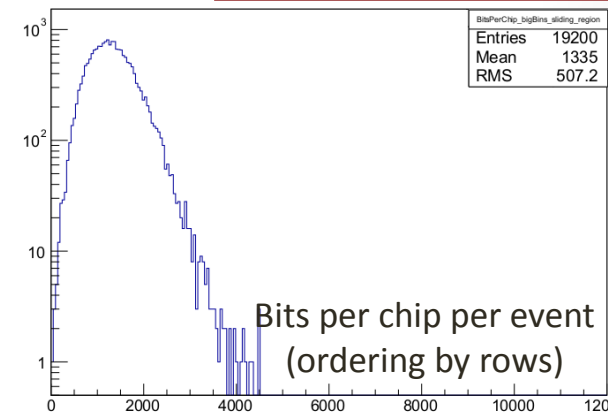
Encoding in horizontal direction by pixel arrival

Encoding row by row

$$\langle N_{bits} \rangle = (H_{\Delta x} + H_{\Delta y} + H_{ADC}) \langle N_{pixels} \rangle$$

$$\langle N_{bits/pixel} \rangle \approx (2.25 + 3.70 + 4.43) \times 127.1 \approx 1320.45 \text{ bits/chip/event}$$

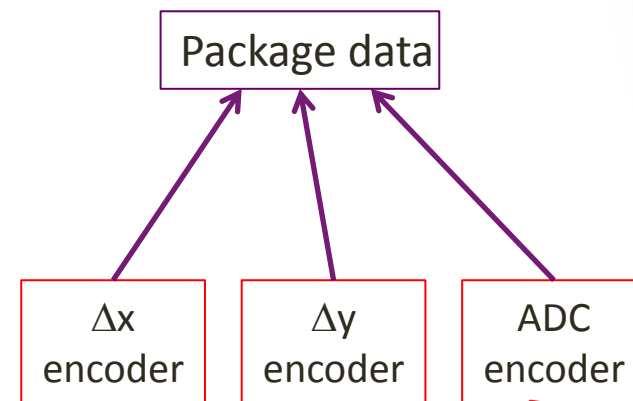
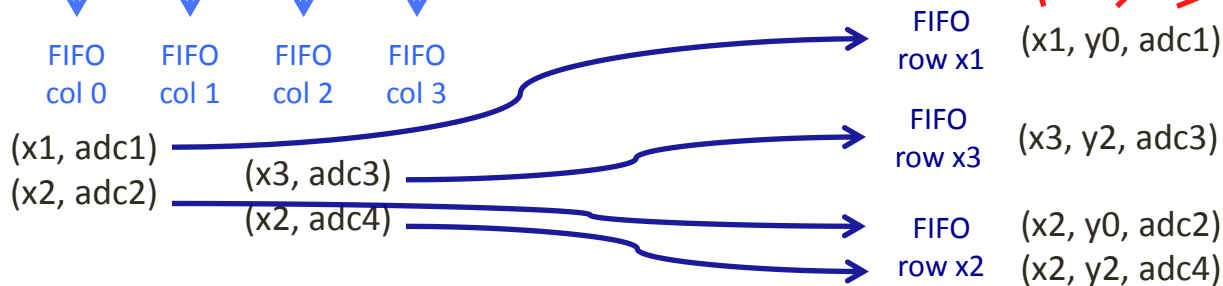
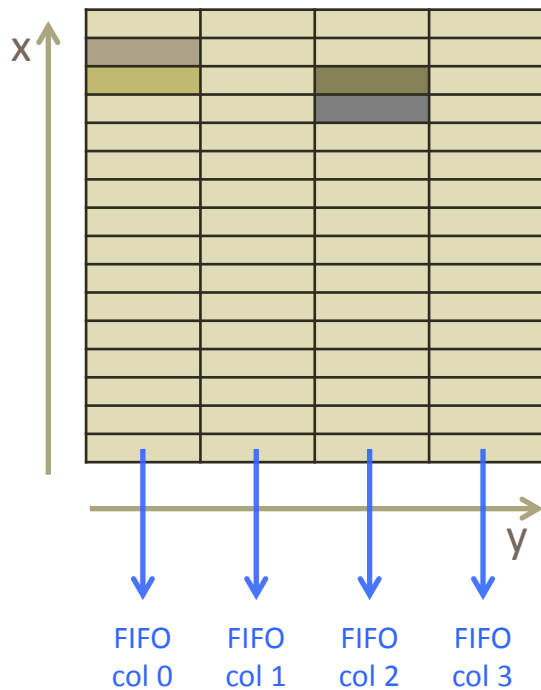
Compression ratio ≈ 2.3



Vth INFIERI workshop

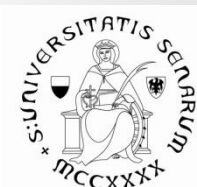
April 2015 – CERN, Geneva

Delta Representation (ordering by rows)





Conclusions and next steps



- Best solution: Delta representation.
- Arithmetic encoding with delta representation: doable for splitting the chip in four regions and encode in parallel
 - Max number of active pixel per chip ≈ 400
 - 100 pixels per region
 - 100 pixels at 320MHz \rightarrow Need 3 clk cycles/pixel
 - 3 sequences x 4 regions = 12 compressors
 - Area estimation for 65nm: $12 \times 0.095(\text{mm}^2)/2 = 0.57 \text{ mm}^2$
 - Dynamic power consumption estimation for 65nm:
 $P_{65} = 12 * (65/130) * (1/1.4)^2 * P_{130} = 230 \text{ mW}$
- 1st task: Fully debug and optimize current arithmetic compressor.
- Next steps:
 - Improve power consumption
 - Synthesize/simulate on 65 nm
 - Verify the functionality of the algorithm for delta representation



Thank you !

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 317446

Vth INFIERI workshop
April 2015 – CERN, Geneva