Development of a data compressor for the CMS phase II pixel detector

Konstantin Androsov [1,2], Massimo Minuti [1], Fabrizio Palla [1], Stamatios Poulios [1,2,*]

1. INFN Pisa, 2. University of Siena

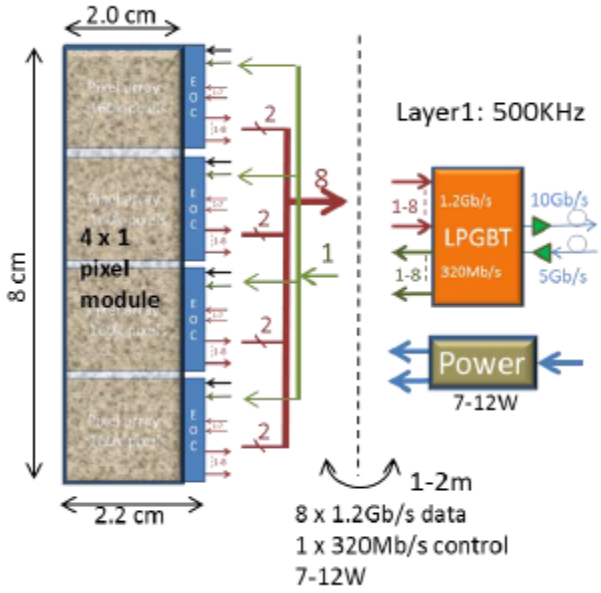
*Supported by the EU FP7-PEOPLE-2012-ITN project nr 317446, INFIERI, "Intelligent Fast Interconnected and Efficient Deices for Frontier Exploitation in Research and Industry"

VIIth INFIERI workshop

April 2016 – Lisbon

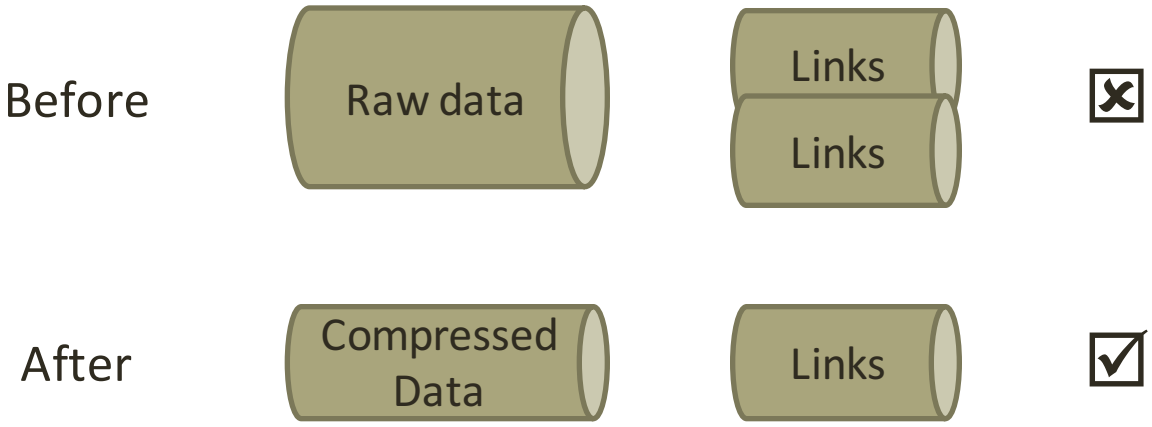
Introduction (1/2)

- 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)



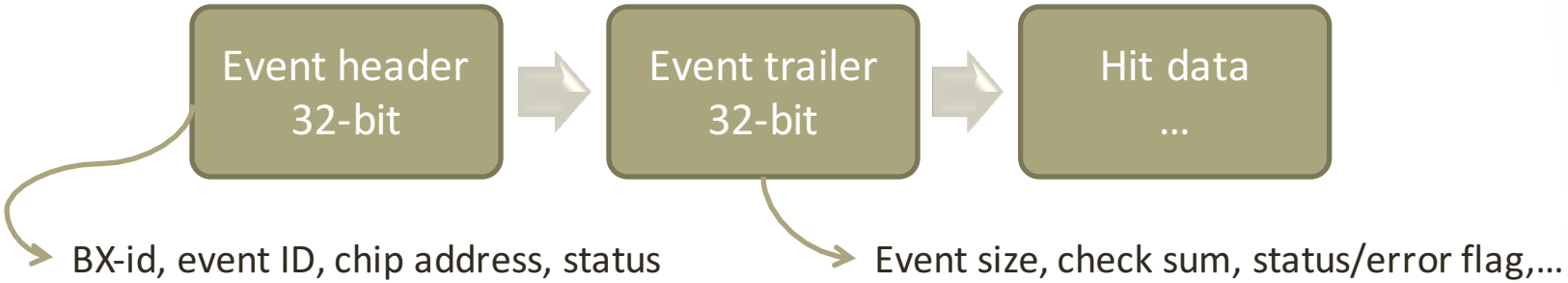
Introduction (2/2)

- 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



Readout format

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



- Default hit data representation:
  - (30 bits/pixel)

Compression scheme

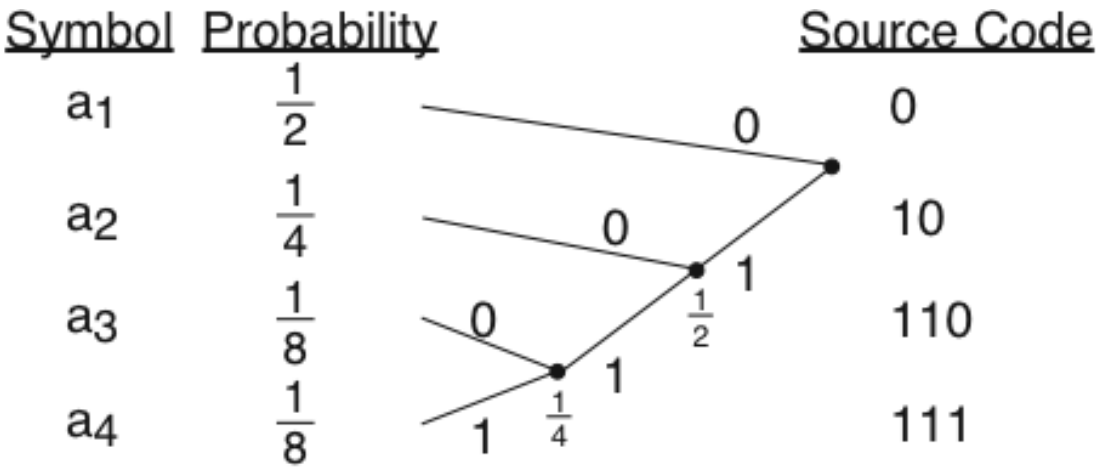
- Data representation
  - How the data are accessed
- Encoding
  - Actual data encoding

Encoders

- Simple compression technique
  - Reduced logic
  - Reduced need for power/area
  - Minimum computation time
- Huffman
  - Simple
  - Fast
- Arithmetic
  - Not very complex
  - Slow

Huffman coding

- Prefix coding
- Assigns shorter codes for more probable symbols

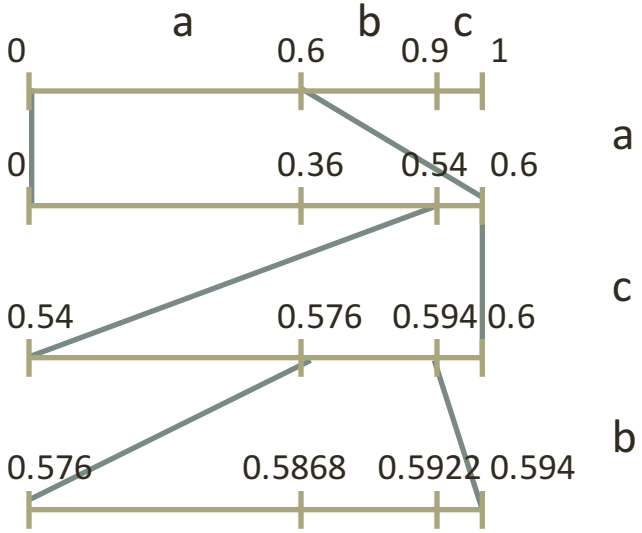


Arithmetic coding

- Arithmetic encoding is a form of entropy encoding used in lossless data compression

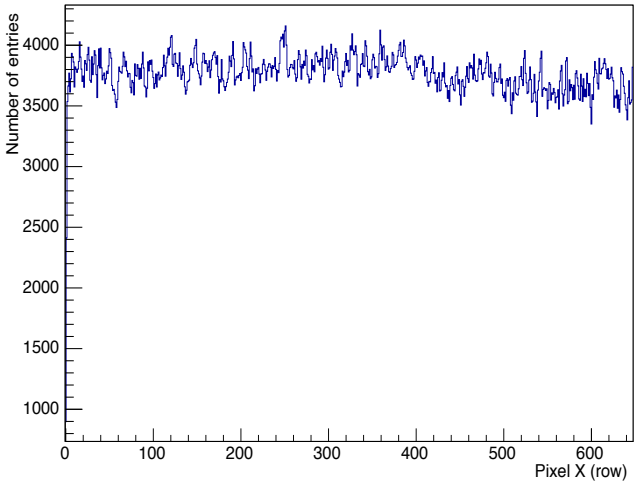
Example:
P(a)=0.6
P(b)=0.3
P(c)=0.1

Phrase: “acb”

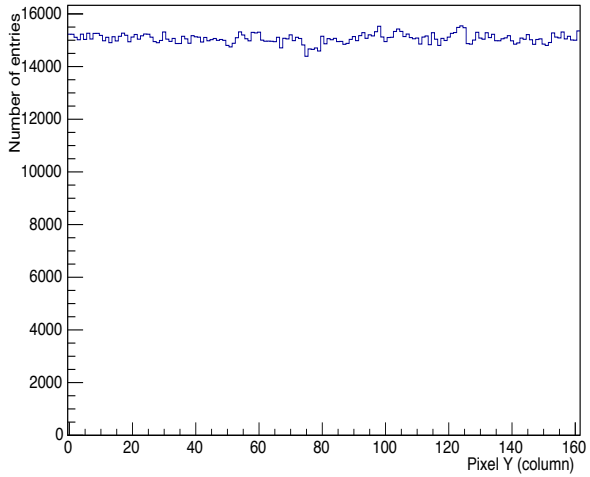


Characteristic distribution

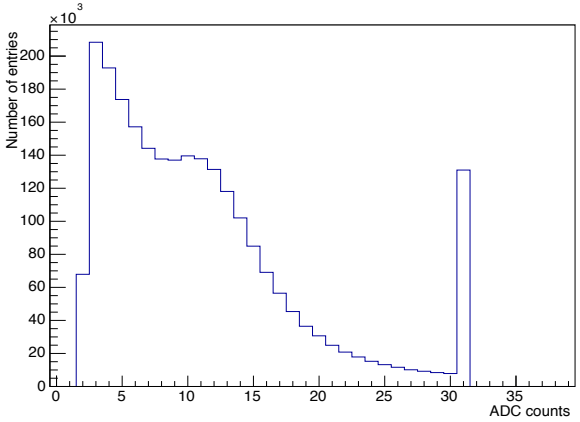
Pixel X



Pixel Y



Active ADC



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:

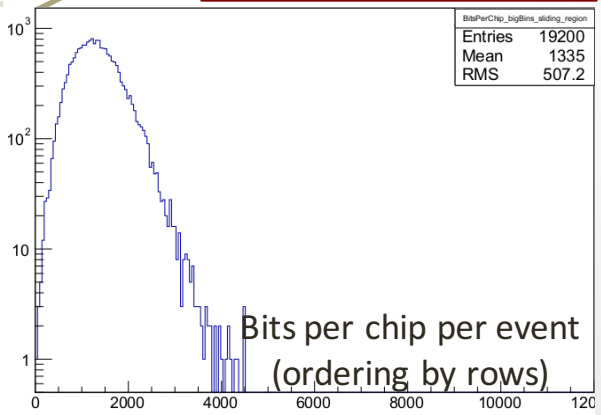
Pixel ordering	H Δy	H Δx	H pixel position
By columns	5.46	1.67	7.13
By arrival	6.70	3.01	9.71
By rows	2.25	3.70	5.95

Compression ratio ≈ 2.3

Encoding column by column

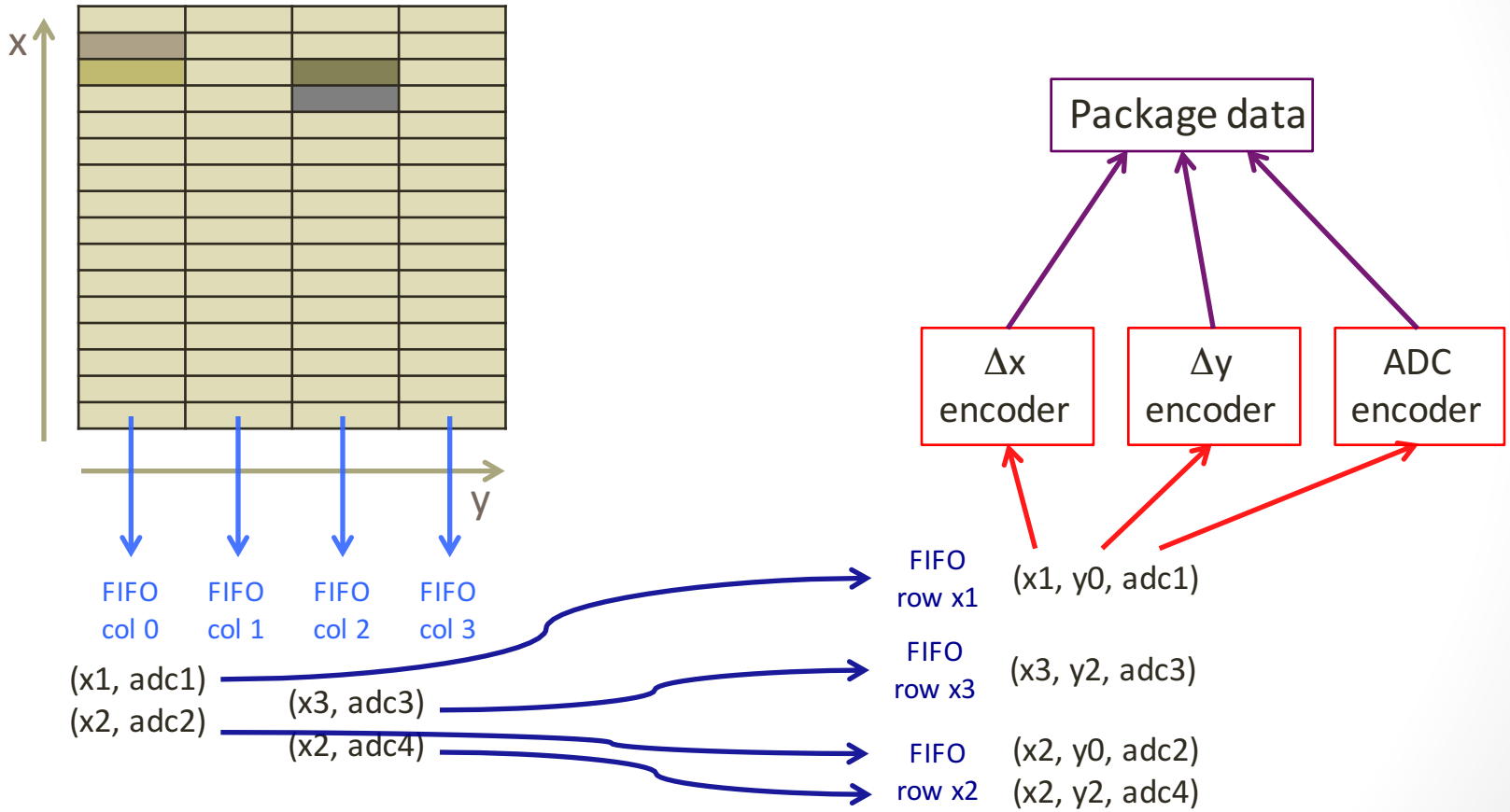
Encoding in horizontal direction by pixel arrival

Encoding row by row



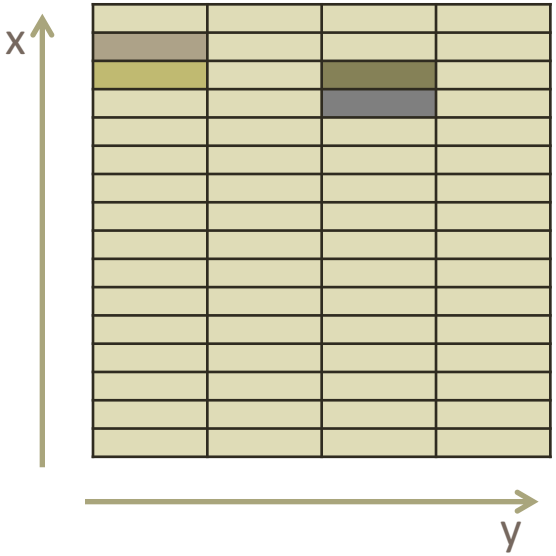
Bits per chip per event (ordering by rows)

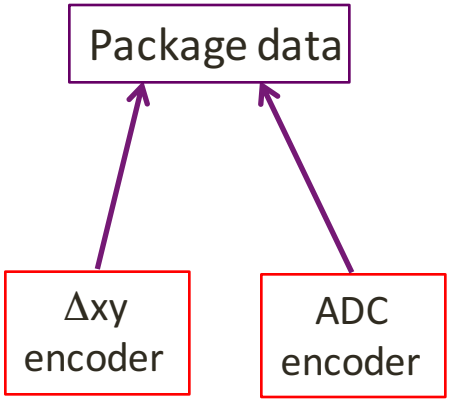
Delta Representation (ordering by rows)



Combined delta

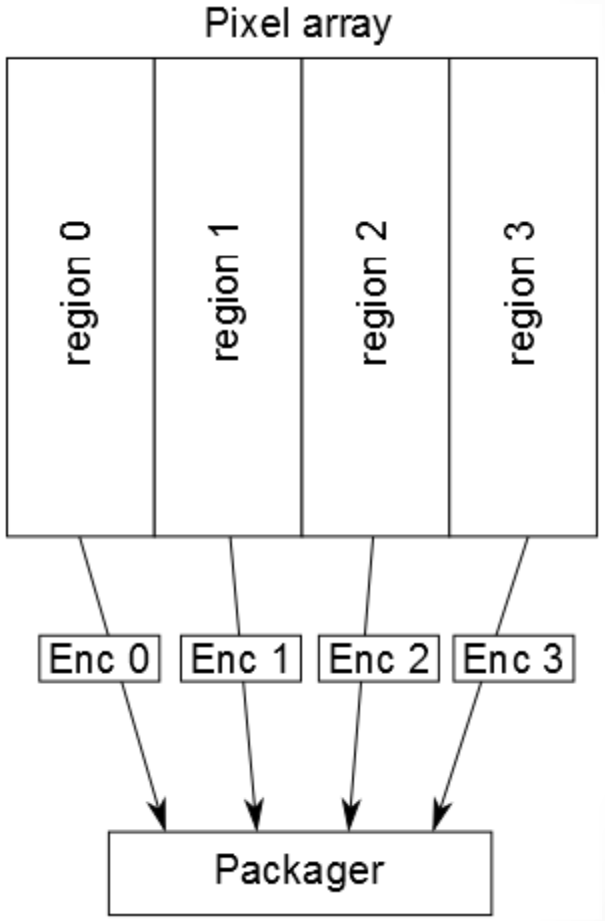
- Combined delta
  - (dx, dy, adc) → (d_{xy}, adc)
  - e.g. dx = 1 , dy = 0 → d_{10}
  - Clusters: d_{10}, d_{01}, d_{11}





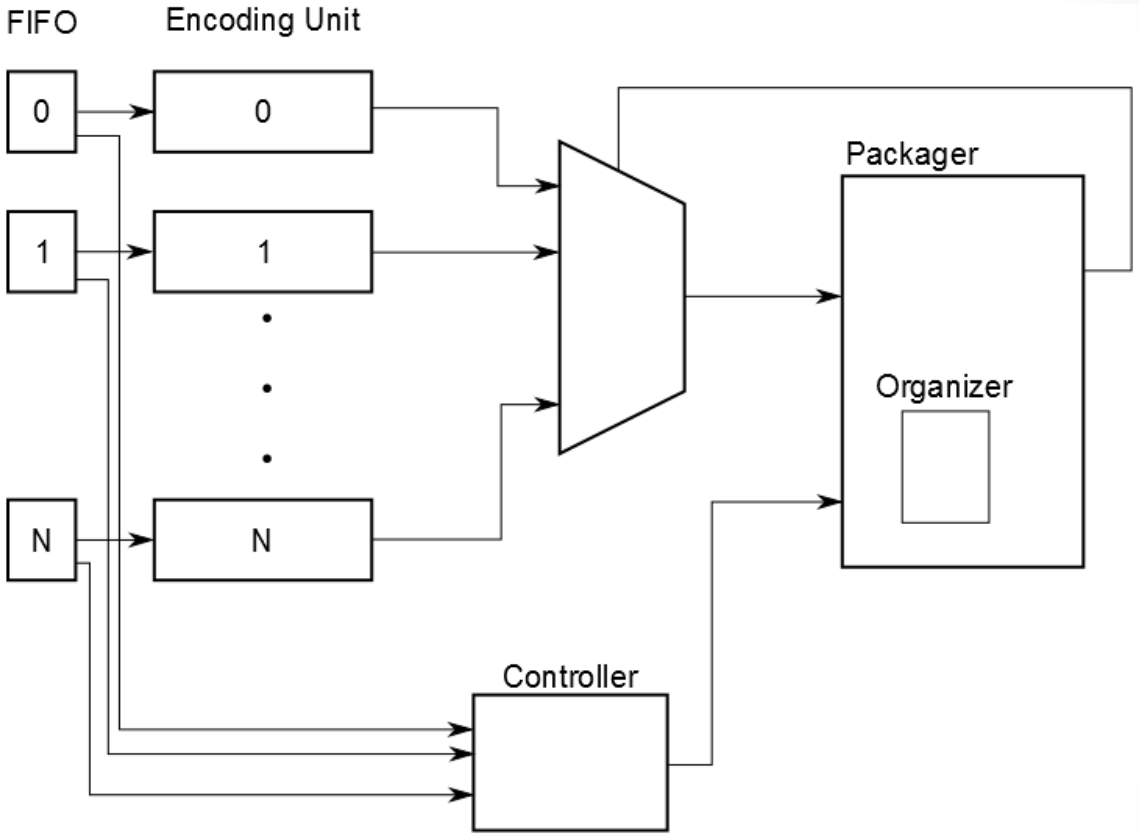
Data flow

- Timing requirements
  - 160 clk cycles available
- Max number of active pixels
  - 400
- Pixel array divided in regions
- Encode in parallel



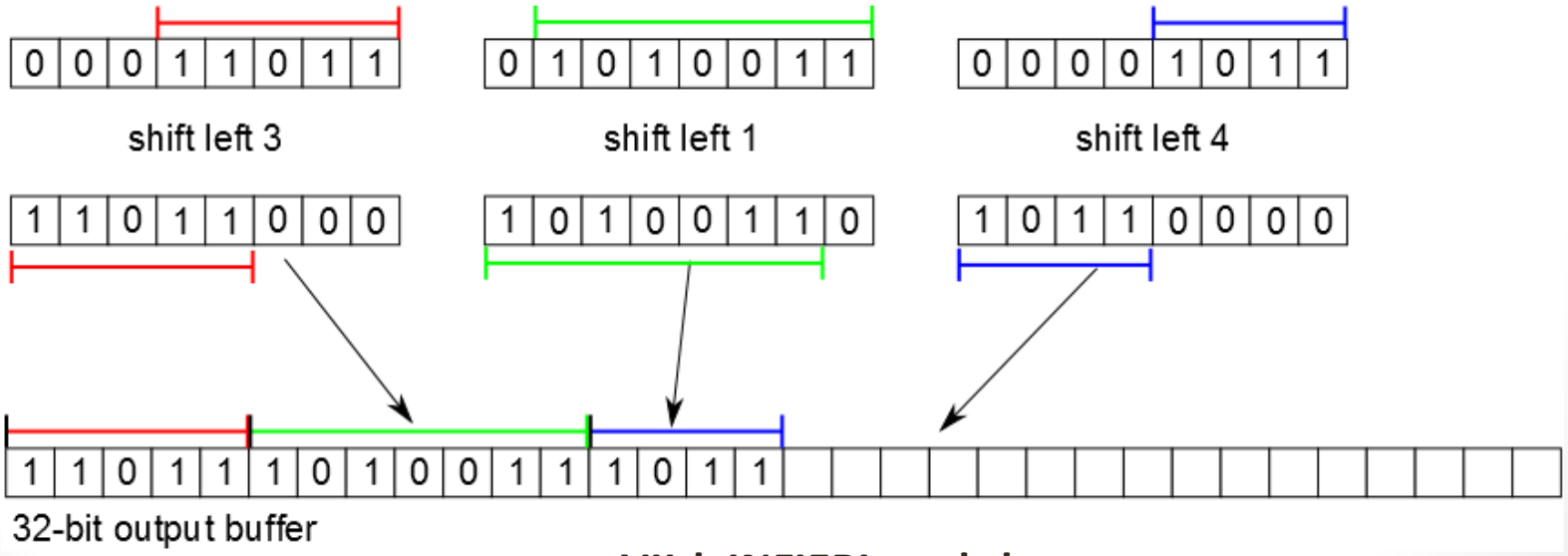
Compressor architecture

- N Encoding Units for N regions
  - encoder_adc
  - encoder_dx
  - encoder_dy
- Packager
  - Final packet
- Organizer
  - Organizes the data in to 32-bit words



Organizer

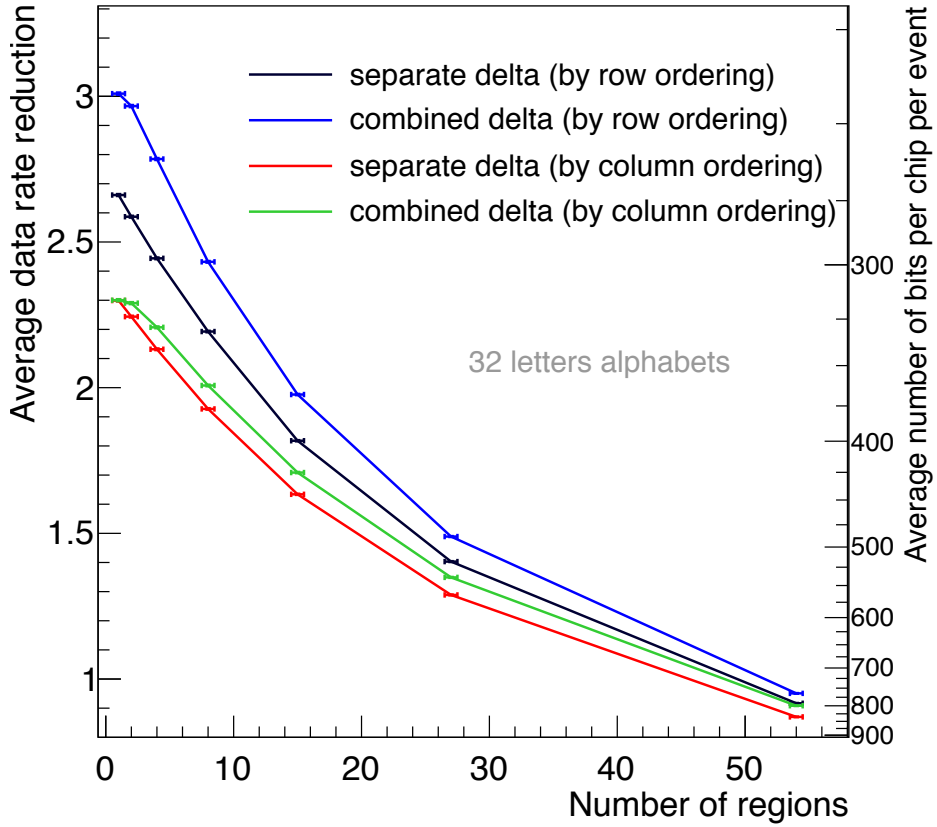
- Huffman codes: variable size
- Output: 32-bit words
- Align data to fit 32-bit
- Support more than one input
  - 3 in case of delta
  - 2 in case of combined delta (future work)



Performance

- Huffman encoder processes 1 symbol in 1 clk cycle
- Sequential encoding
  - \+ requires one “symbol memory” element
  - \- slow
  - ➢ Max # of clk cycles: $\sum n_i + j + 1$,
    n = # of symbols, j = # of active regions, i = region
- Parallel encoding
  - \+ fast
  - \- requires n “symbol memory” elements
  - ➢ Max # of clk cycles: ~ $\max\{n_i\}$, (to be tested)

Compression ratio



Thank you !

The research leading to these results has received funding from the People Program (Marie Curie Actions) of the European Union's Seventh Framework Program FP7/2000-2013/ under REA grant agreement no [317446] INFIERI "INtelligent Fast Interconnected and Efficient Devices for Frontier Exploitation in Research Industry"

VIIth INFIERI workshop

April 2016 – Lisbon