Dynamic Vision Sensor

The Road to Market

June 2, 2017

Yoel Yaffe[*], Nathan Levy, Evgeny Soloveichik, Sebastien Derhy, Ayal Keisar, Elad Rozin, Liron Artsi
Jun-Seok Kim[†], Keunju Park, Bongki Son, Yunjae Suh, Heejae Jung, Changwoo Shin, Jooyeon Woo, Yohan Roh, Hyunku Lee, Hyunsurk (Eric) Ryu

[*]*Samsung Israel Research Center, Israel*
[†]*System LSI , South Korea*
Samsung Electronics

Contact: yoel.yaffe@samsung.com

The Road Less Traveled

image by aosleading

Market Driven Requirements VS Early DVS Prototypes

Requirements	Early DVS Prototypes
Low Cost	Large Pixel No standard for Quality / Die sorting
Minimal Module Size	Large, due to large Pixel -> Large optical format
Ultra Low Power	Low power
Good Event Quality	Redundant events (especially at low light) Motion artifacts, Timing accuracy Noise, Flicker
Low Data Rate	Scene dependent; In some cases may exceed CIS typical throughput
Vision processing at Edge	Not available

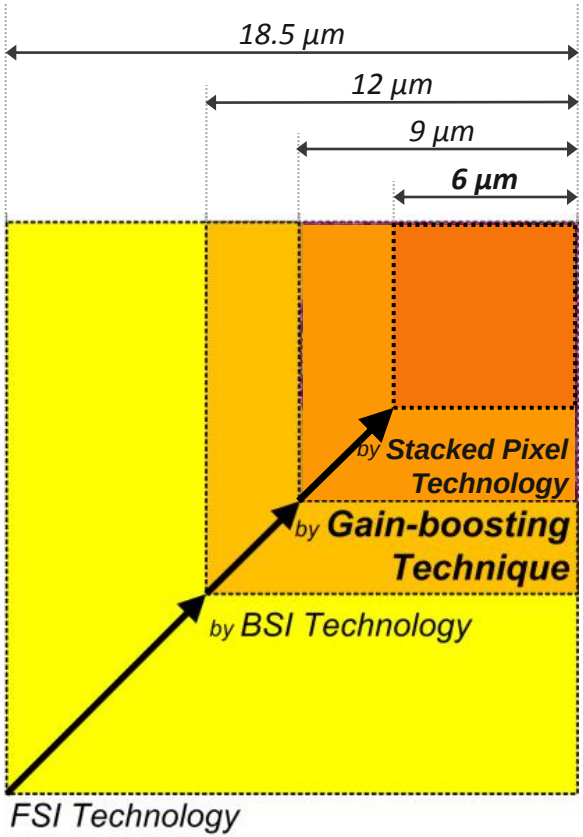
The Road to Market - Overview

- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

The Road to Market - Overview

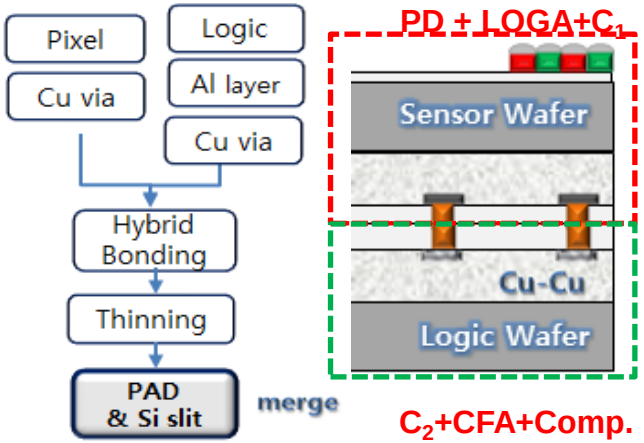
- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

Reducing Pixel Size and Optical format



Pixel Shrink

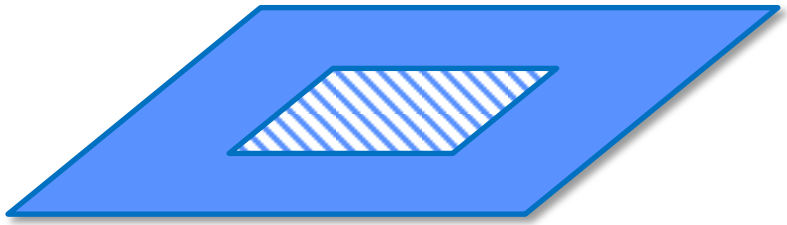
- BSI process
- Downsizing MIMCAP
  Without a decrease in minimum contrast sensitivity
- Stacked Cu-Cu Technology



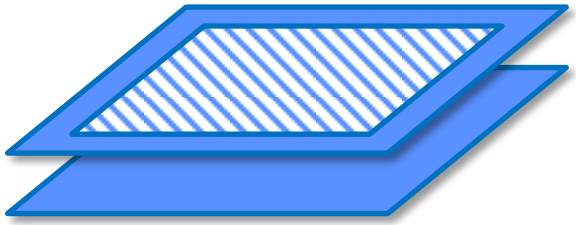
18.5 x 18.5 μm² → 6 x 6 μm2
~90% pixel area reduction

Reducing Pixel Size and Optical format

- Photo Enhancement:



3µm x 3µm Photodiode */ 9µm x 9µm Pixel*



5µm x 5µm Photodiode */ 6 µm x 6 µm Pixel / Wafer-stacking*

Delivering Product to Market - Overview

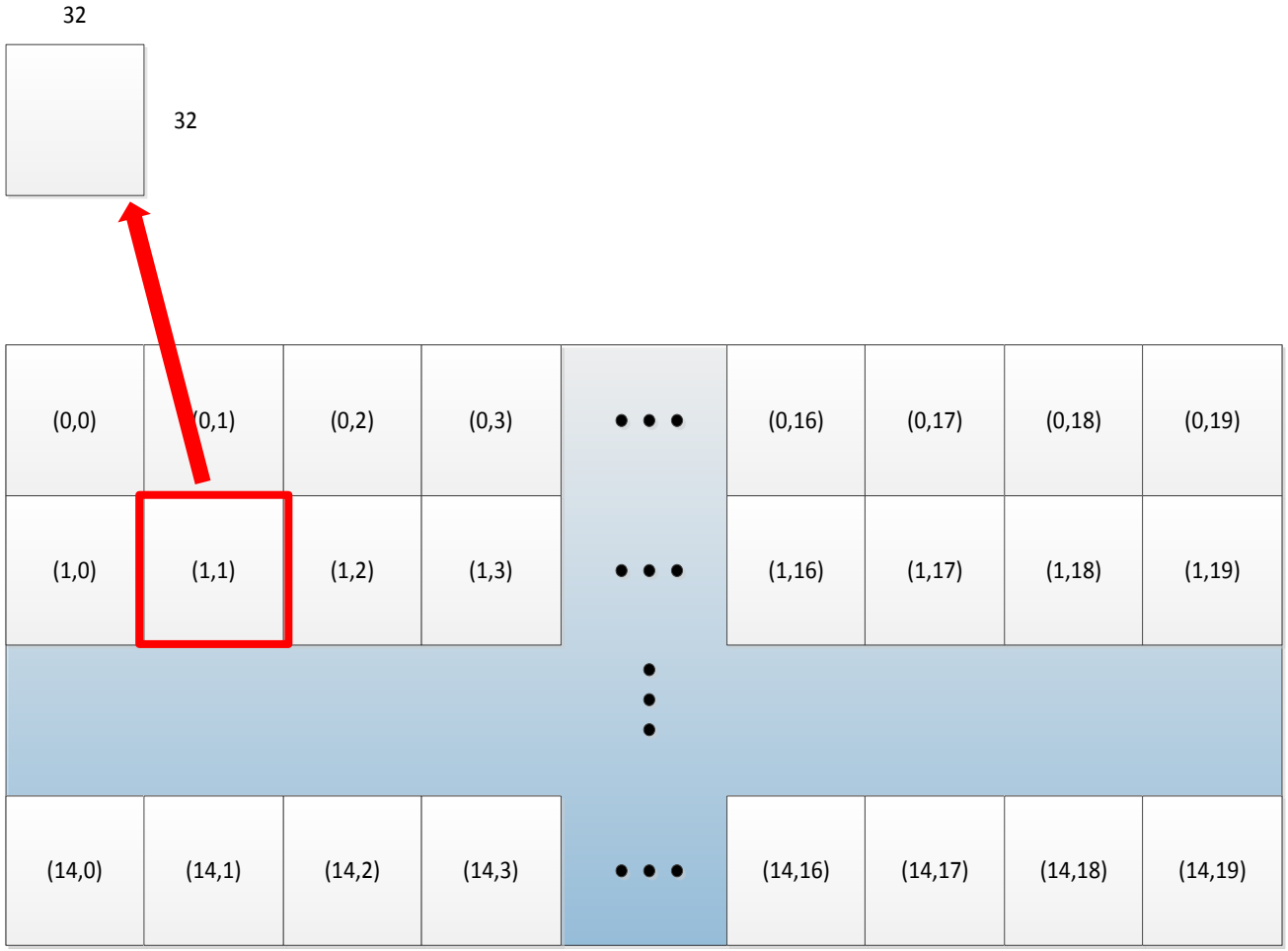
- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

Reducing Power

- Use Samsung's advanced process
- Subsampling modes
- Dynamic wakeup, window of interest

Reducing Power: Statistics for Dynamic Decisions

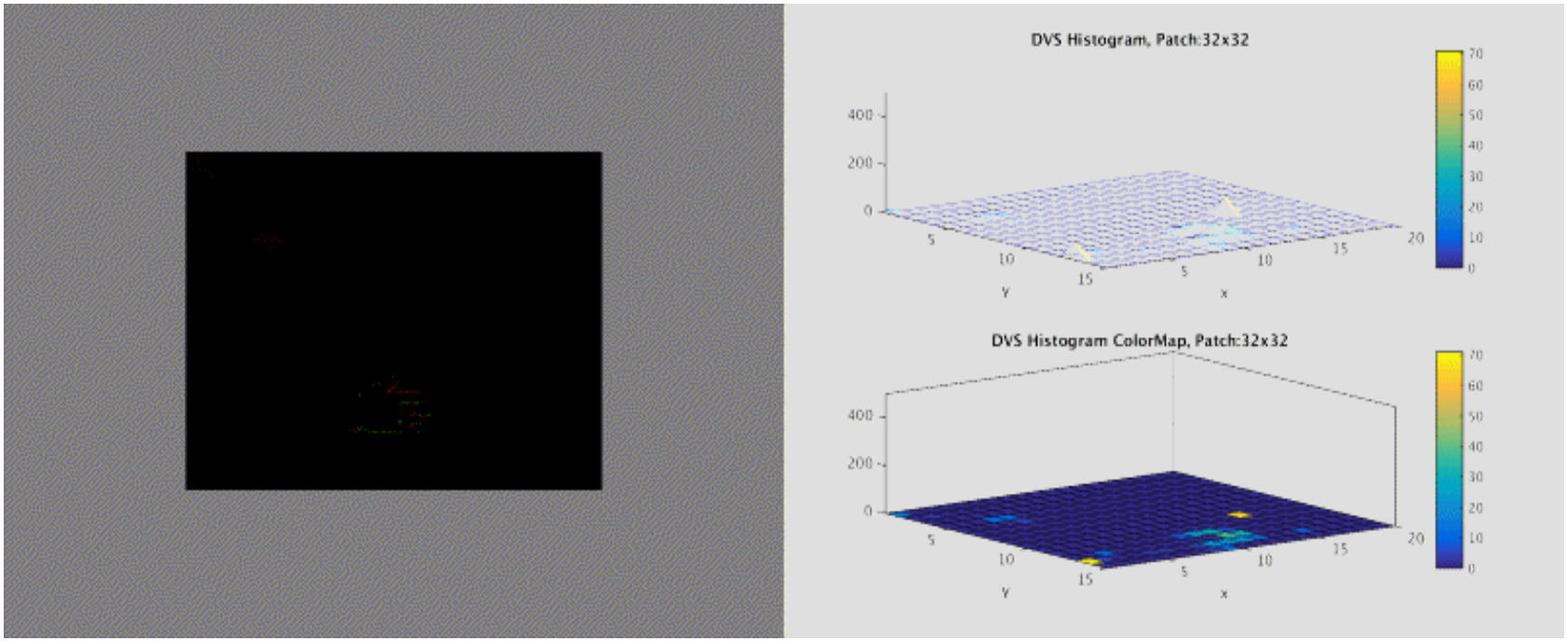
- Spatial Histogram & Blocking Area:



640x480 Image Plane

ISSCC'17

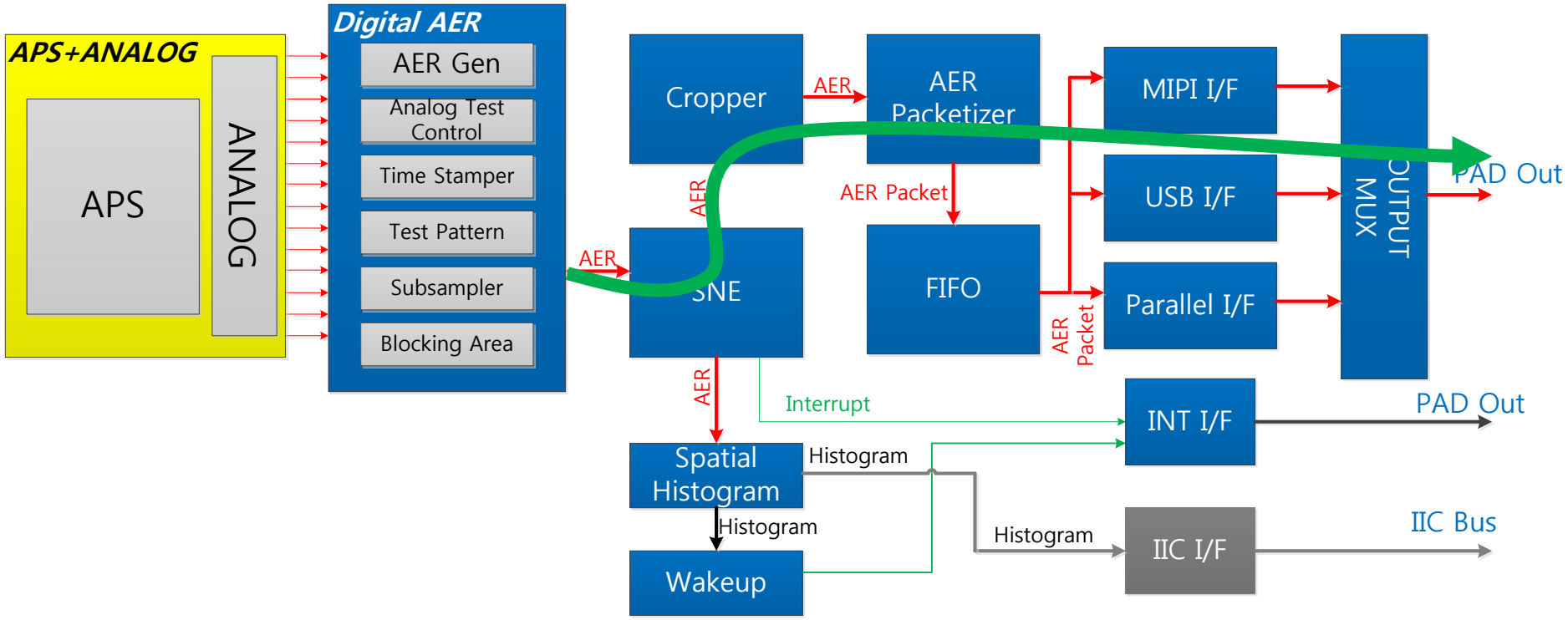
Reducing Power: Statistics for Dynamic Decisions



Spatial Histogram Data

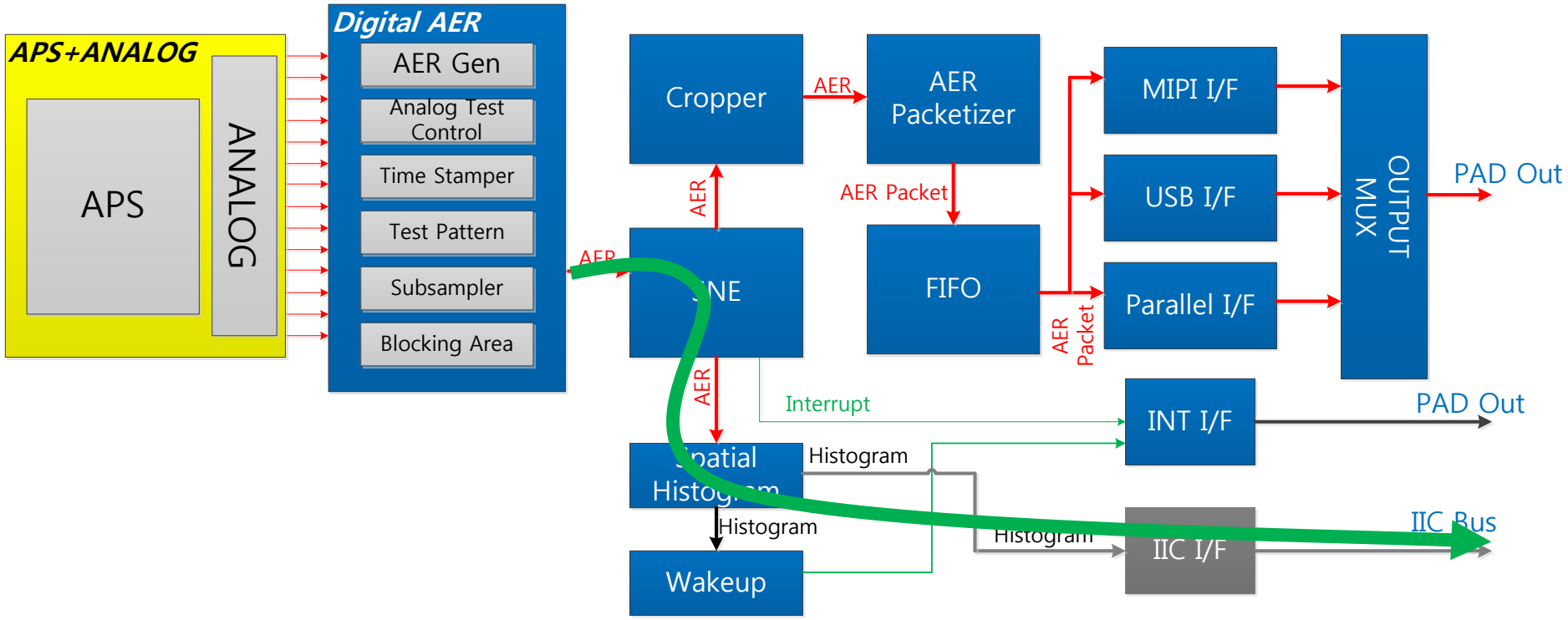
ISSCC'17

Reducing Power: Blocking, Subsampling



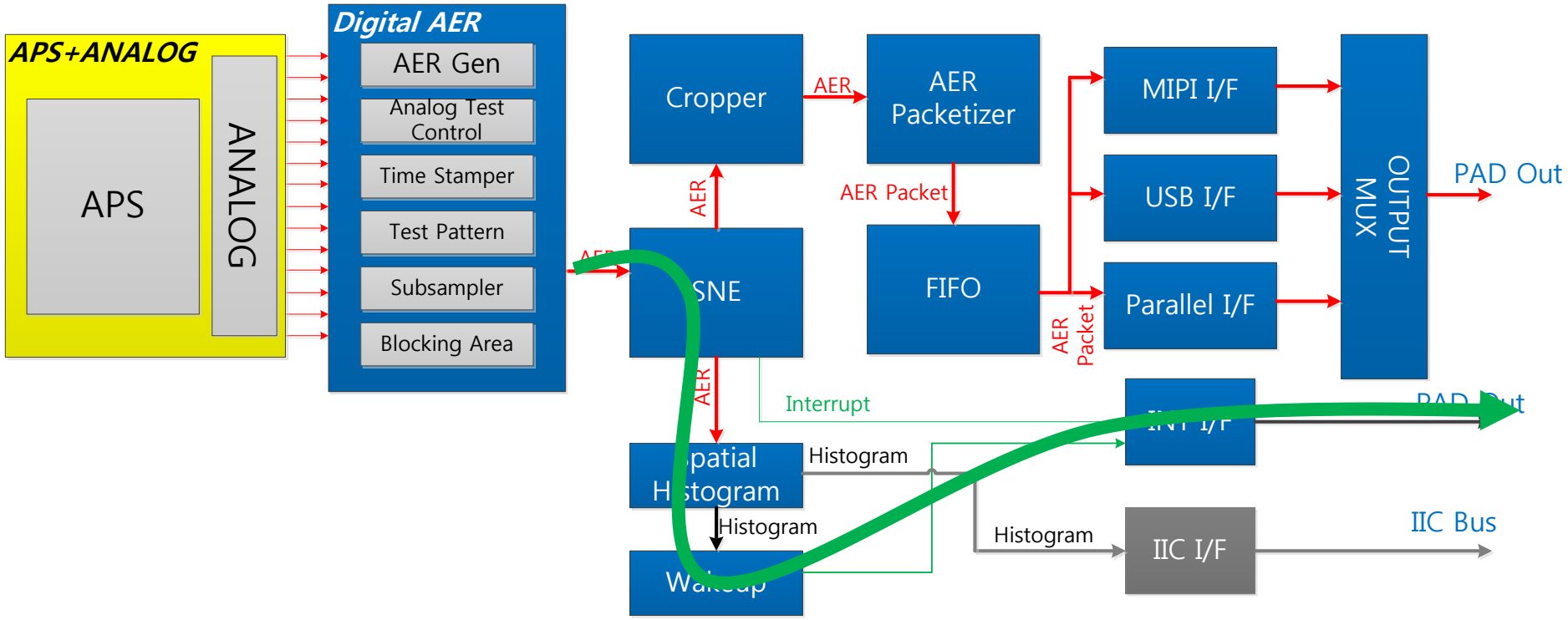
ISSCC'17

Reducing Power: Statistics Collection



ISSCC'17

Reducing Power: Wakeup



ISSCC'17

The Road to Market - Overview

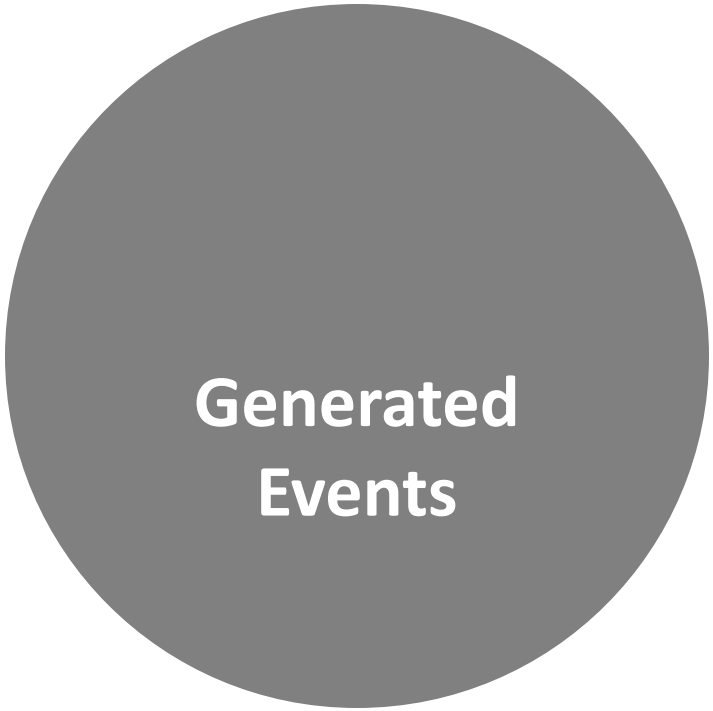
- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

Good Event Quality: Objective

Clean and perfect edge map

→ Obtained by subtracting frames on a High Frame-Rate CIS camera

Good Event Quality



Good Event Quality

Redundant Events
- Falsely Generated
- Multiple detections
- Not applicable

Missing Events
- Not/badly detected
- Discarded

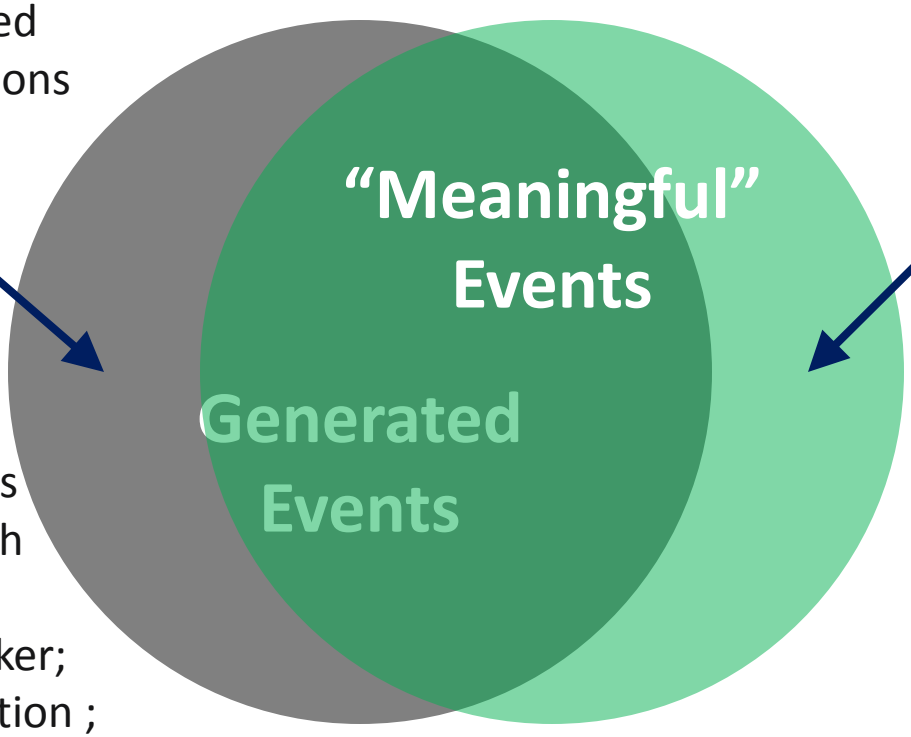
Examples:
- Noise; Bad Pixels
- Response to high contrast change
- Light source flicker; Background motion ; Too many events

Examples:
- Overflow, Timestamp errors
- Filtered (mitigating redundant events)

Good Event Quality

Redundant Events
- Falsely Generated
- Multiple detections
- Not applicable

Missing Events
- Not/badly detected
- Discarded



Examples:
- Noise; Bad Pixels
- Response to high contrast change
- Light source flicker; Background motion ; Too many events

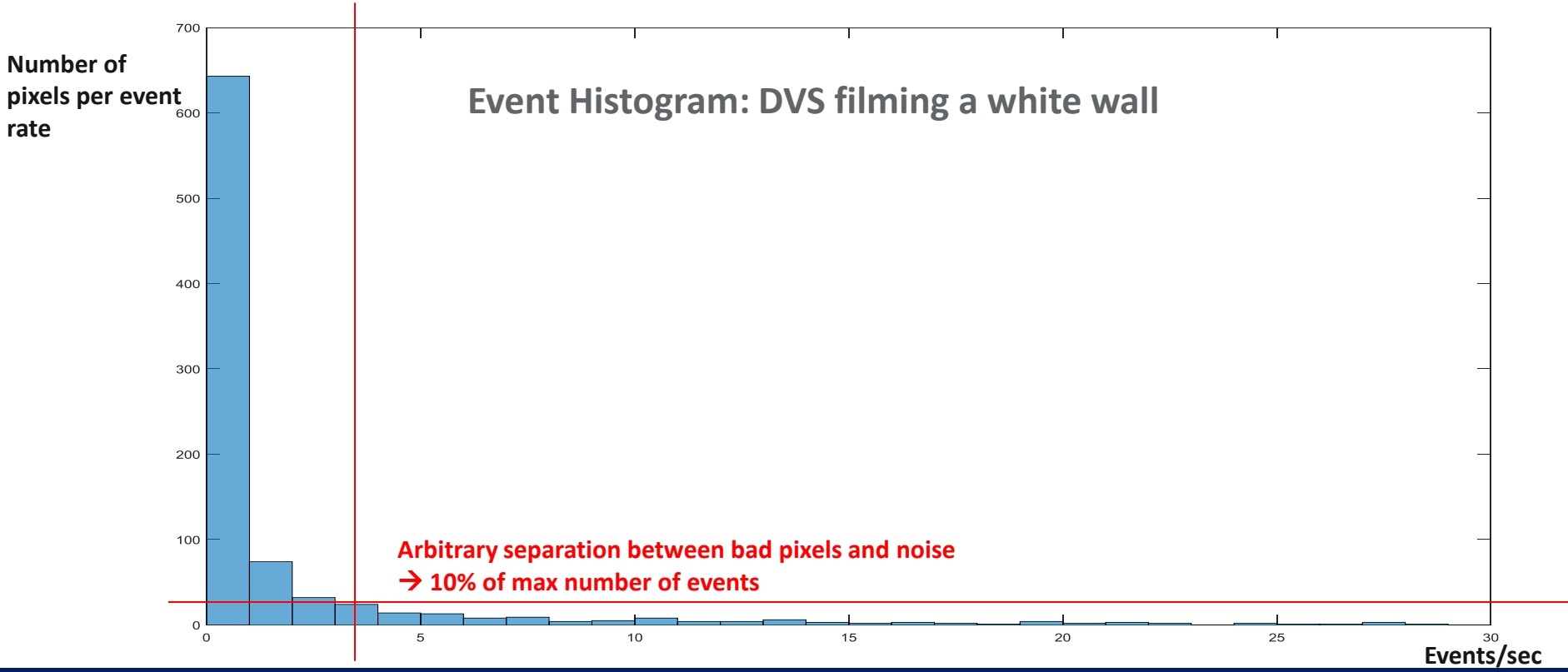
Examples:
- Overflow, Timestamp errors
- Filtered (mitigating redundant events)

Redundant Events: Bad Pixels; Noise

Basic model:

- **Bad Pixel: Specific** Pixels that generate events at **high** frequency (regardless of the scene)
- **Noise: Random** pixels that generate events at **low** frequency (regardless of the scene)

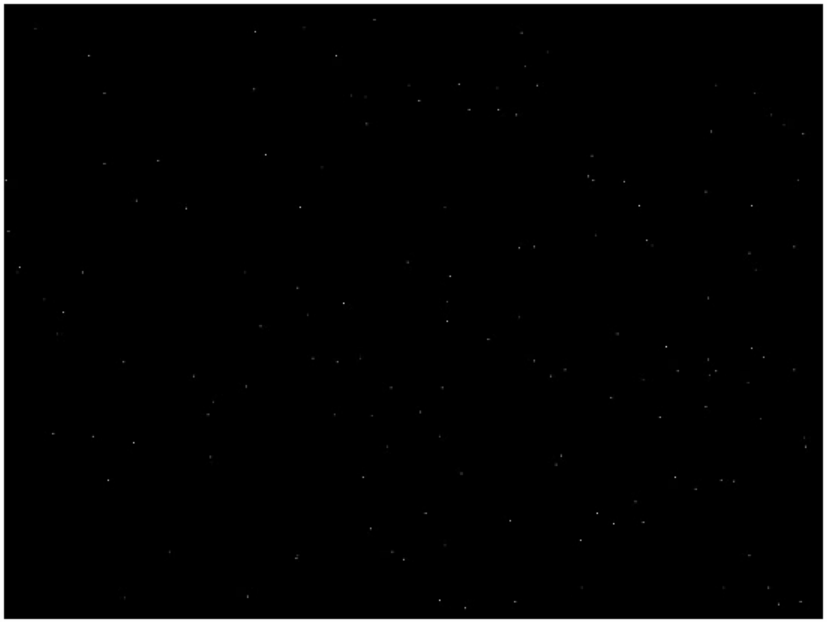
Actual behavior:



Redundant Events: Bad Pixels; Noise

Bad Pixel Suppression Example:

Before chain



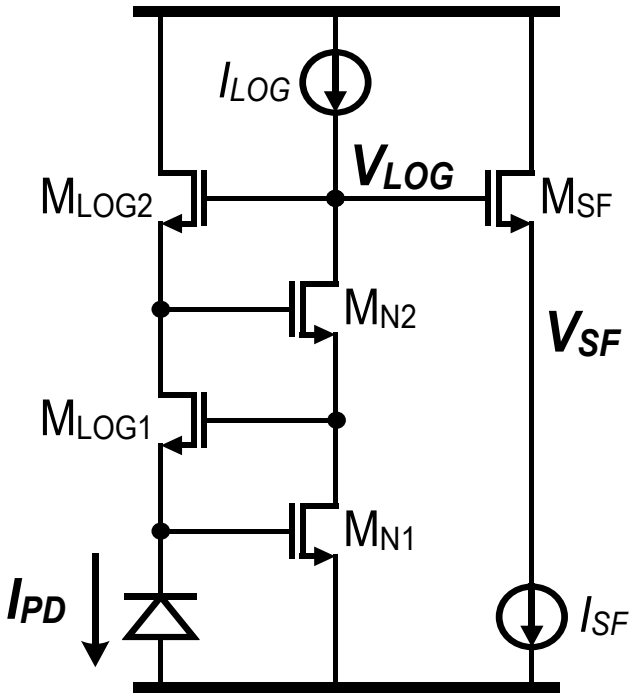
After chain

Redundant Events: Bad Pixels; Noise

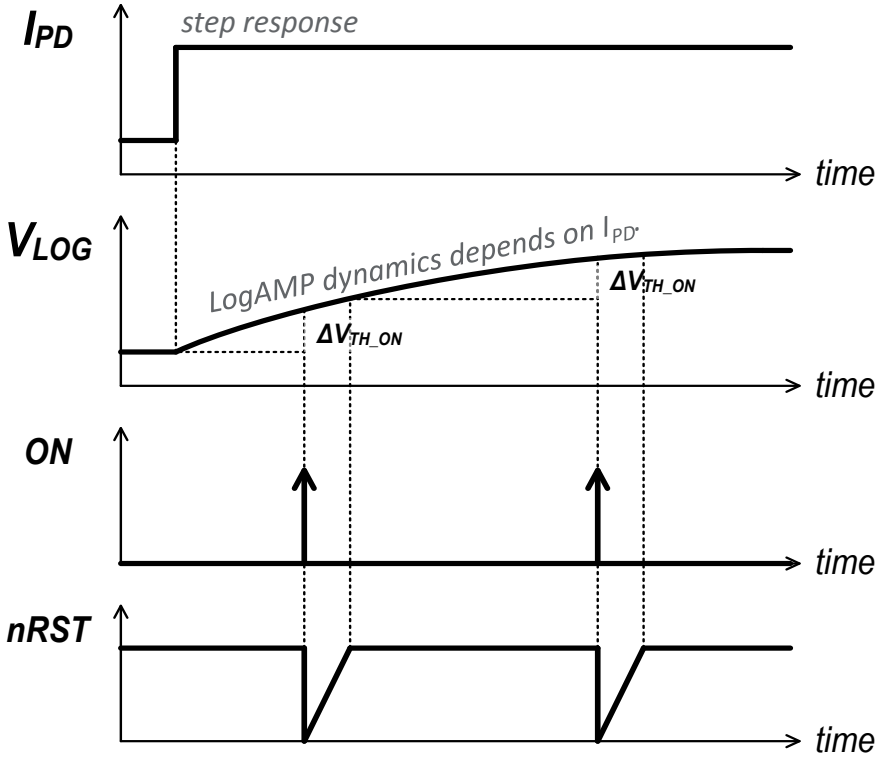
Noise Suppression Example:

Redundant Events: Multiple detections

- Overall pixel BW (speed) is limited by LogAMP BW
- LogAMP BW is mainly affected by 1) Its structure; 2) The amount of photocurrent
- **Redundant events generated at low light (I_{PD}) on @ large contrast change**



***LogAMP of DVS Gen2** (ISSCC'17)*



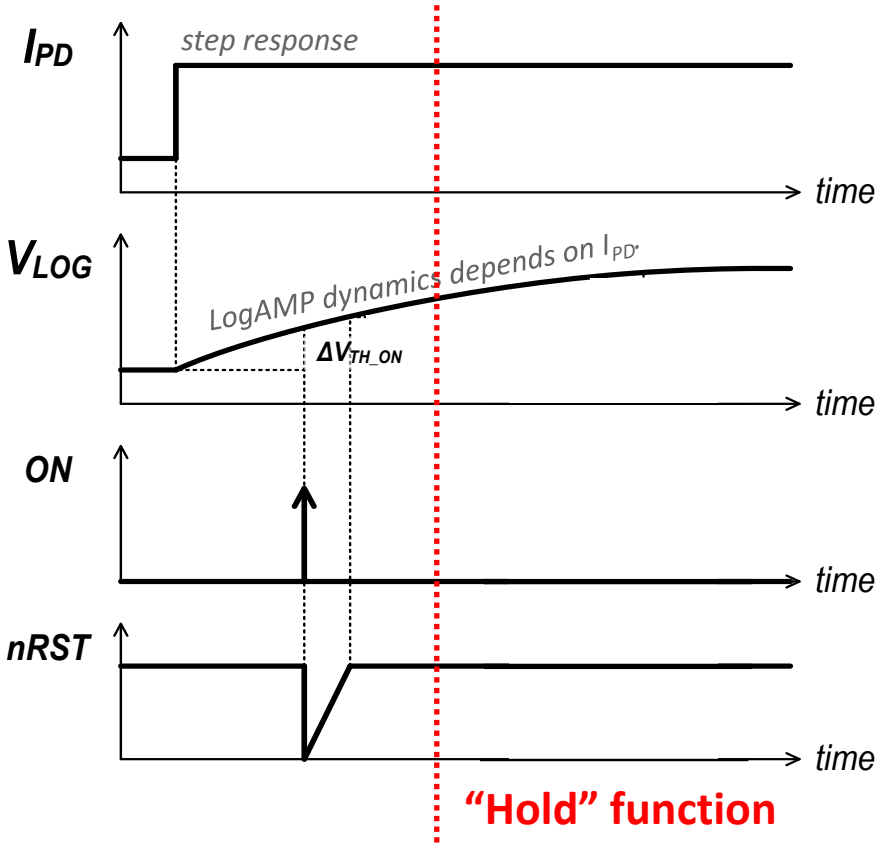
Redundant Events: Multiple detections

- Overall pixel BW (speed) is limited by LogAMP BW
- LogAMP BW is mainly affected by 1) Its structure; 2) The amount of photocurrent
- **Redundant events generated at low light (I_{PD}) on @ large contrast change**

Solutions:

- Structure:
  Increase ~x5-x10 LogAMP BW by new structure

- Limit photocurrent:
  Introduce “HOLD” signal to pixel

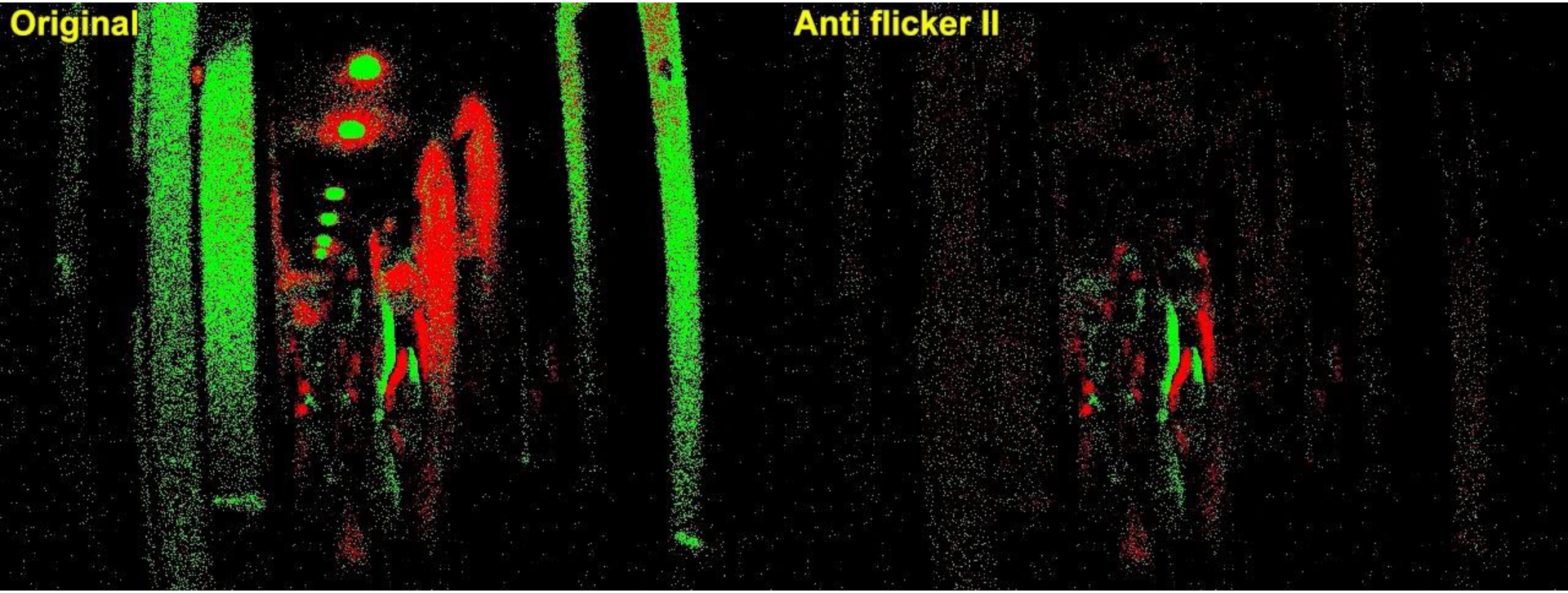


Redundant Events: Light Source Flicker

- **Light Source Flicker:**
  Events that are not generated by motion but by flickering lights (with different frequencies)
- **Note:**
  This has to be differentiated from periodic motion (for example a waving hand)

Redundant Events: Light Source Flicker

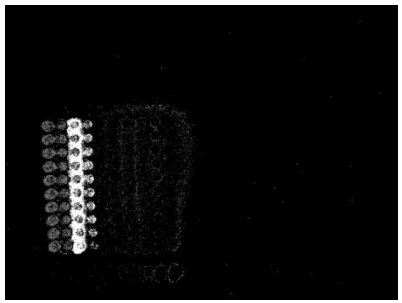
Flicker Suppression Example:



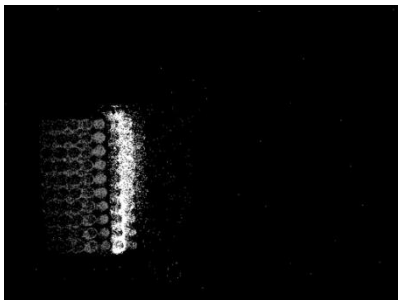
Missing Events: Overflow, Timestamp Errors

- **Overflow:**
  Number of events describing the scene change, exceeds the circuit maximal throughput.

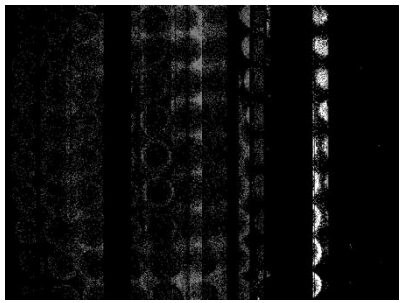
25% active area, 1ms refresh rate:

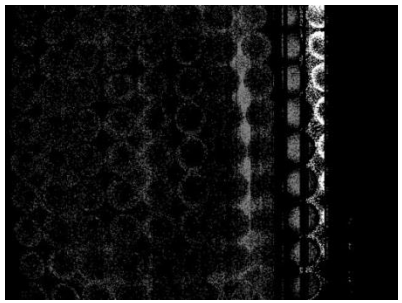

#1040

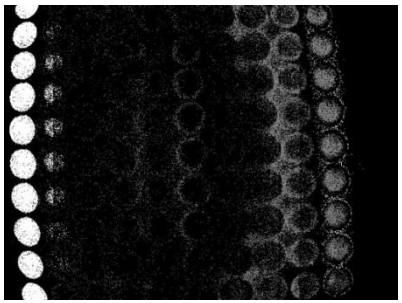

#1041


#1042

90% active area, 1ms refresh rate:

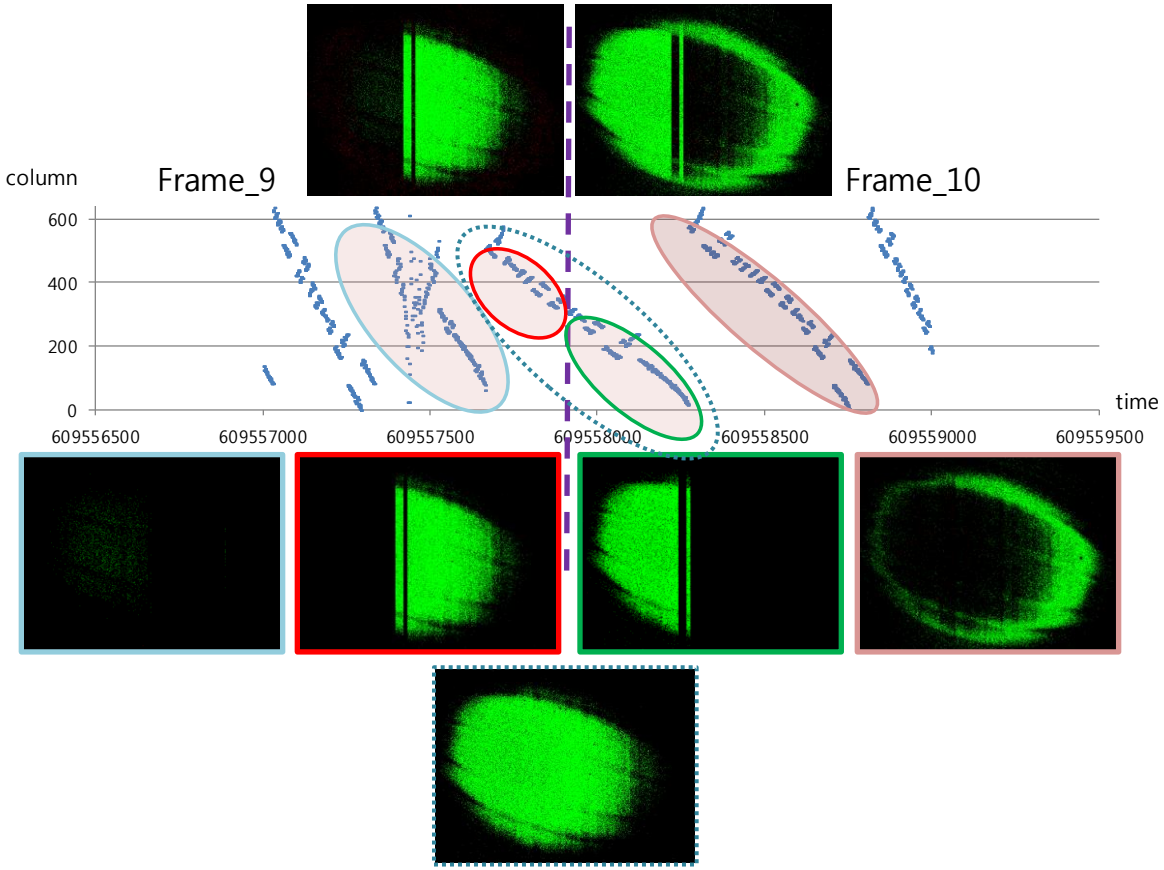

#109


#110


#110

Missing Events: Overflow, Timestamp Errors

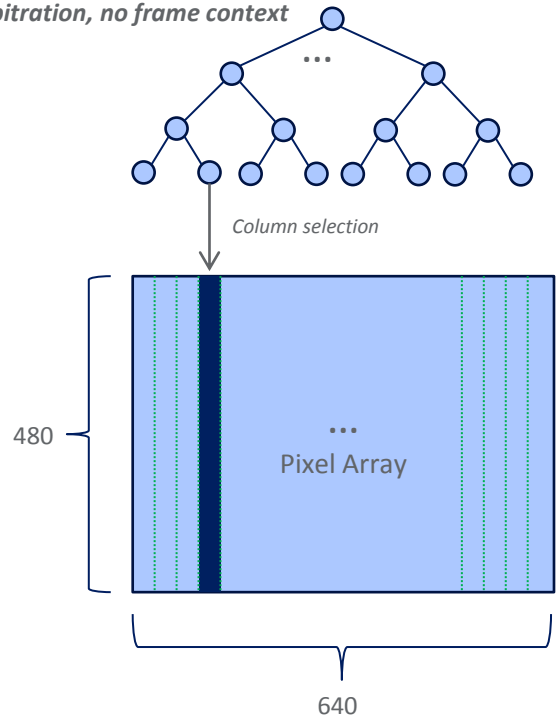
- **Timestamp Errors:**
  Events grouped according to incorrect timestamp



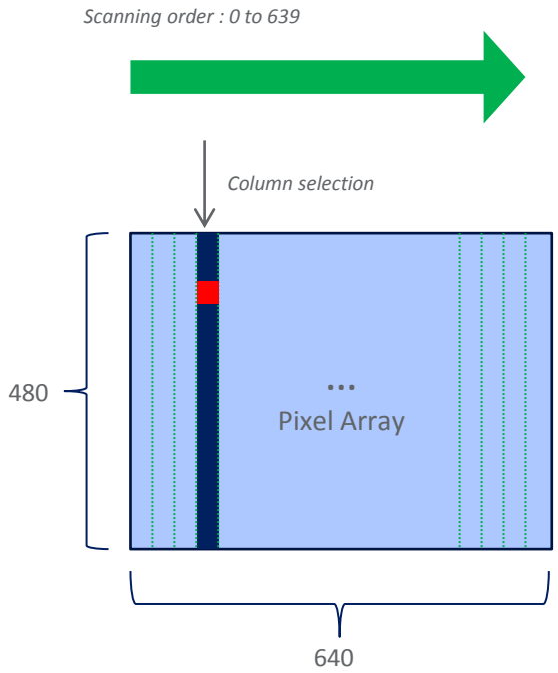
Solution: Pixel Readout Revision

Readout controls	Gen1 & Gen2	Gen3
Column selection	Unfair arbitration scheme	Scanning scheme
In-pixel memory	Free running (no event hold)	Free running (no event memory)
		Global hold (all pixels in pixel array)
Pixel reset	Pixels with an event in a column	Pixels with an event in a column
	-	All pixels in a column
	-	Global reset (all pixels in pixel array)

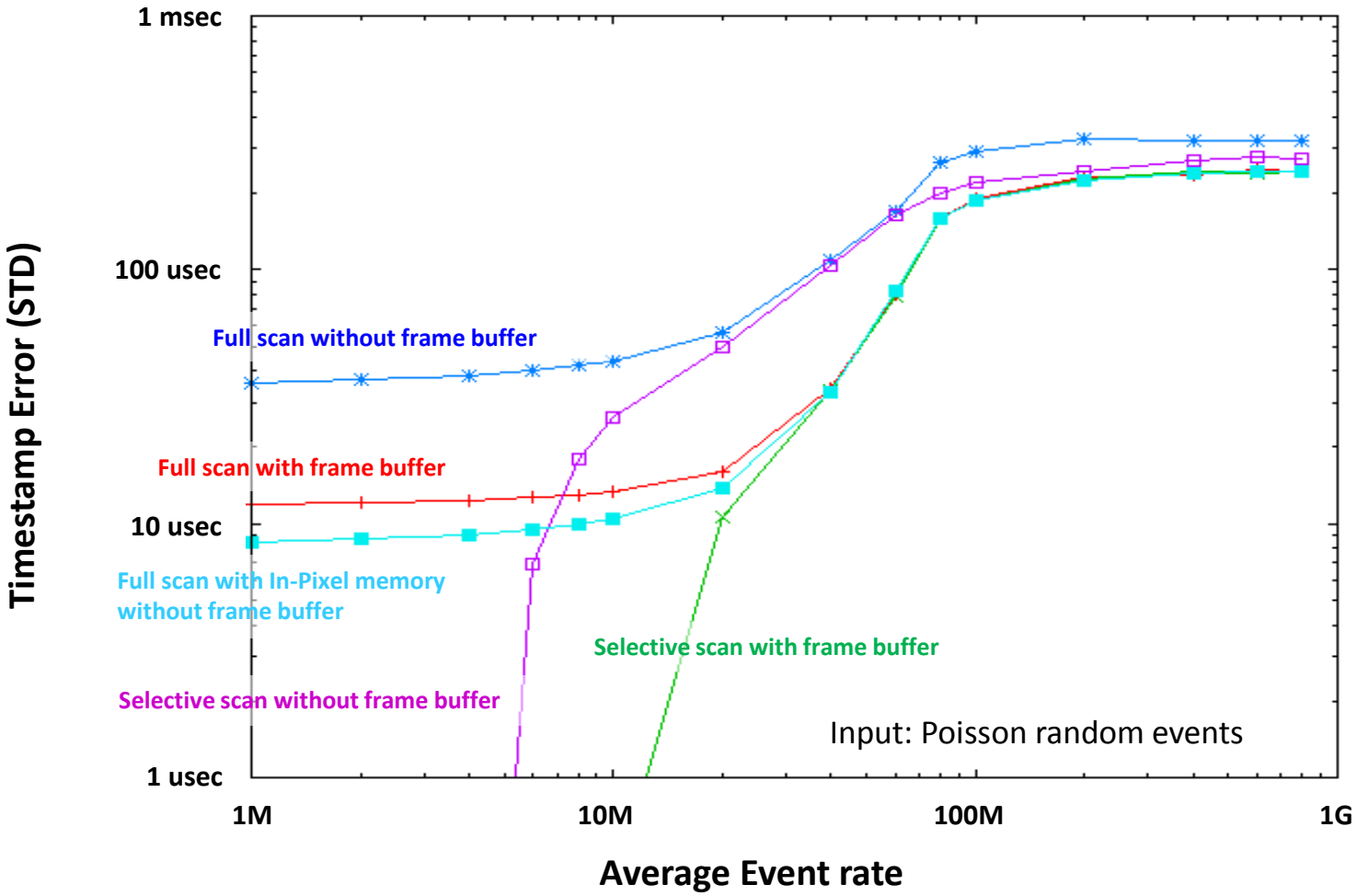
Prev.: Unfair arbitration, no frame context



New: Scanning scheme, frame context options



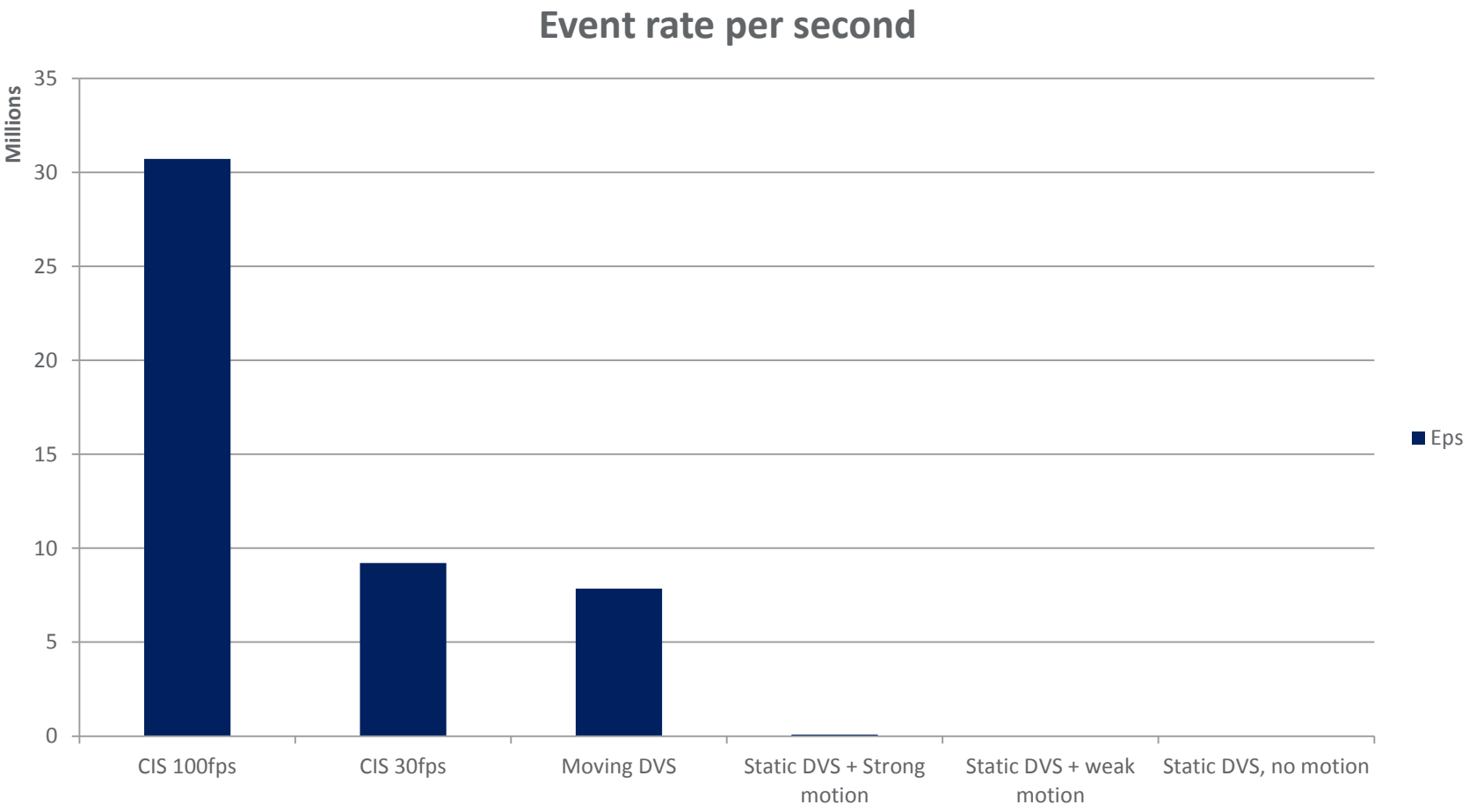
Pixel Readout: Timestamp Error Comparison



The Road to Market - Overview

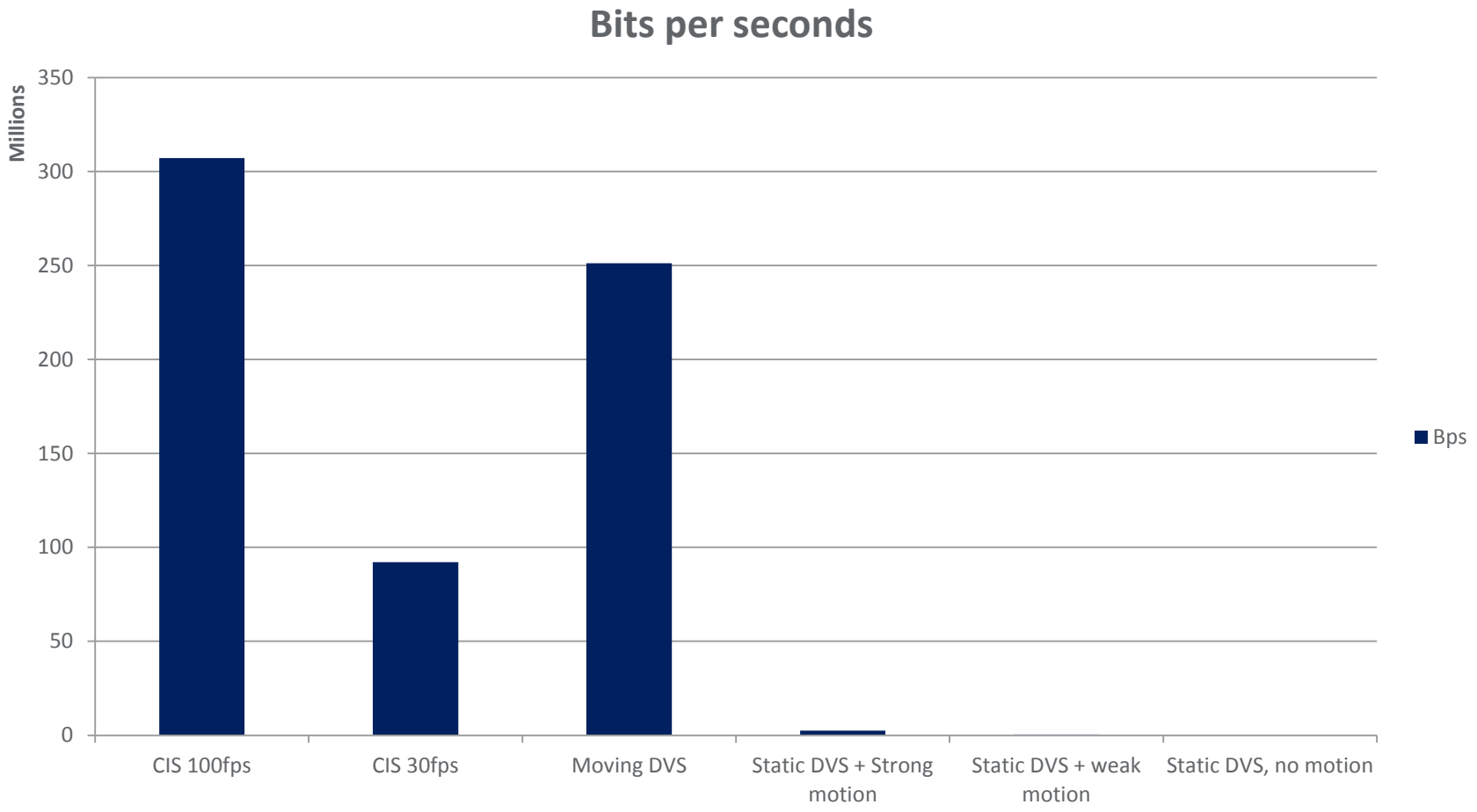
- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

Data Throughput: DVS Low Event Rate Example



Less Events → Easier to process

Data Bandwidth @ 32 bit/event Encoding



DVS AER 32bit per event

Group Address Event Representation (GAER)

AER (32 bits/event)

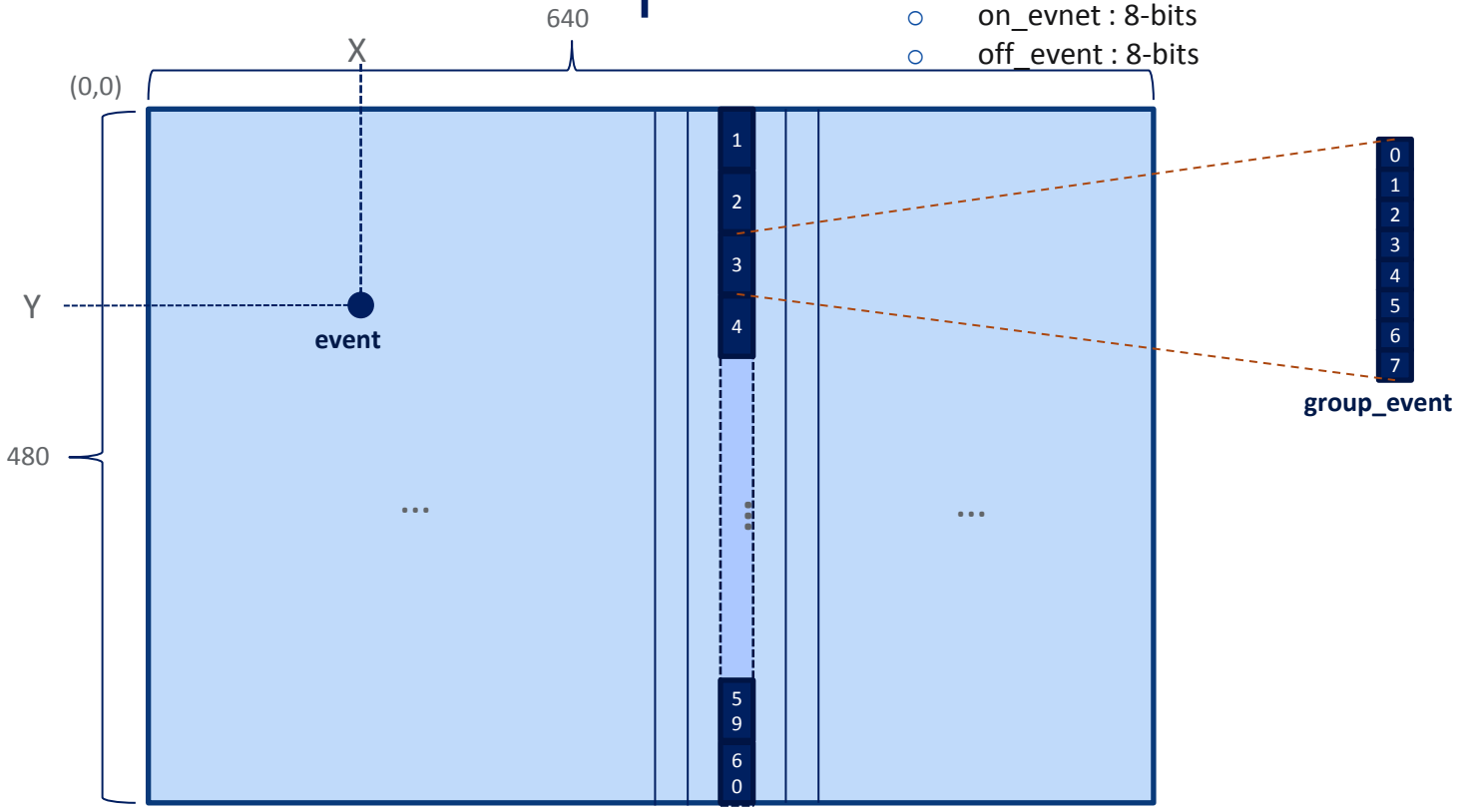
Address Event Representation (TS, X, Y, event)

- P: packet encoding (2bit)
- TS : timestamp (10bit)
- X : column address : 0 ~ 639 (10bit)
- Y : row address : 0 ~ 479 (9bit)
- event : 1-bit. on/off

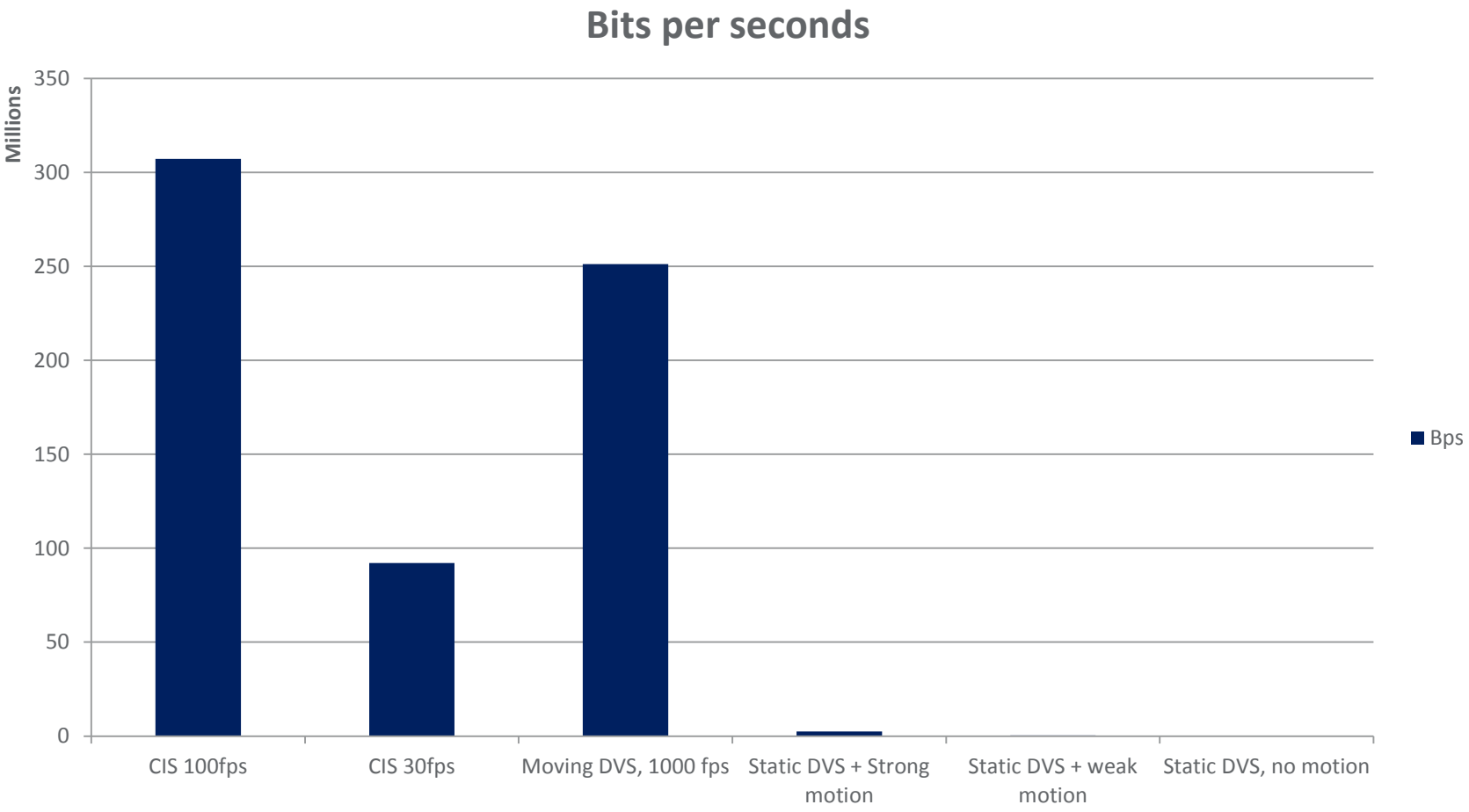
GAER (~64-4 bits/event – content dependent)

Group AER (TS, X, G, group_event)

- P: packet encoding (10bit group, 12bit per col)
- TS: timestamp (10bit)
- X : column address : 1 ~ 640 (10bit)
- G : group_address : 1 ~ 60 (6bit)
- group_event : a bundle of 8 events. 16-bits
  - on_evnet : 8-bits
  - off_event : 8-bits

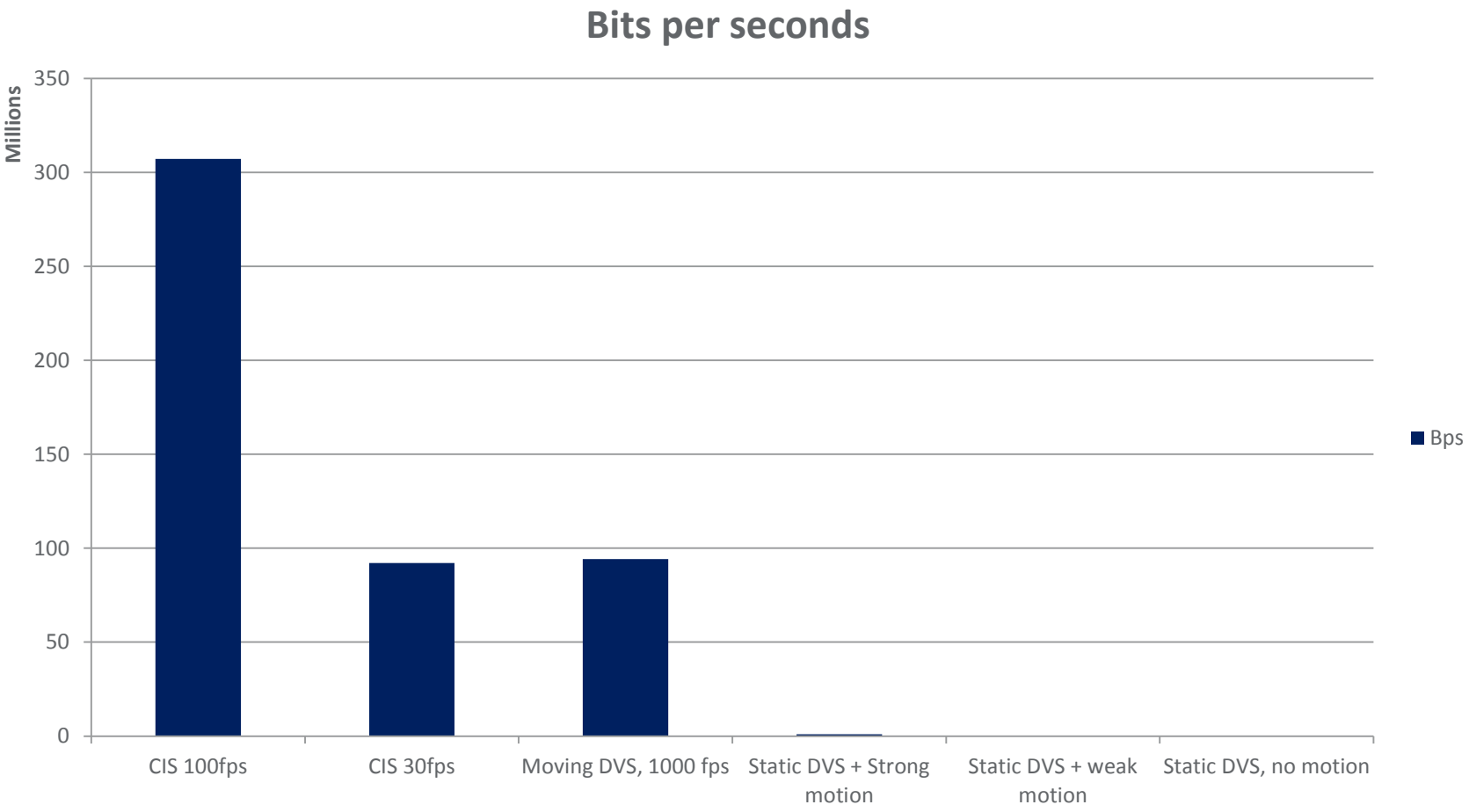


Data throughput: AER 32 bit/event



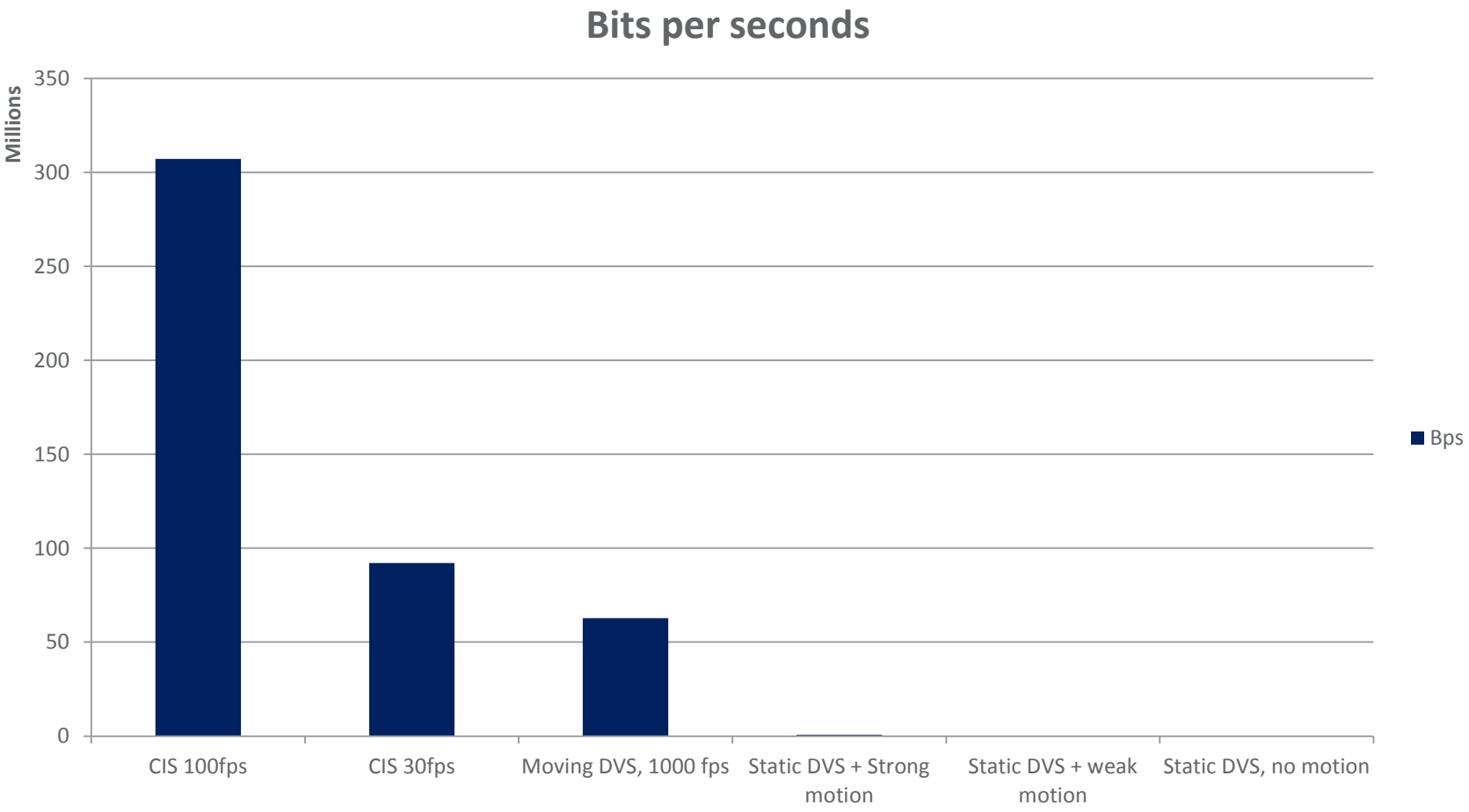
DVS AER 32bit per event

Data throughput: GAER ~12 bit / event



DVS GAER 12bit per event (average)

Data throughput: Improved GAER ~8 bit / event



DVS GAER with offset based column
8 bit per event (average)

The Road to Market - Overview

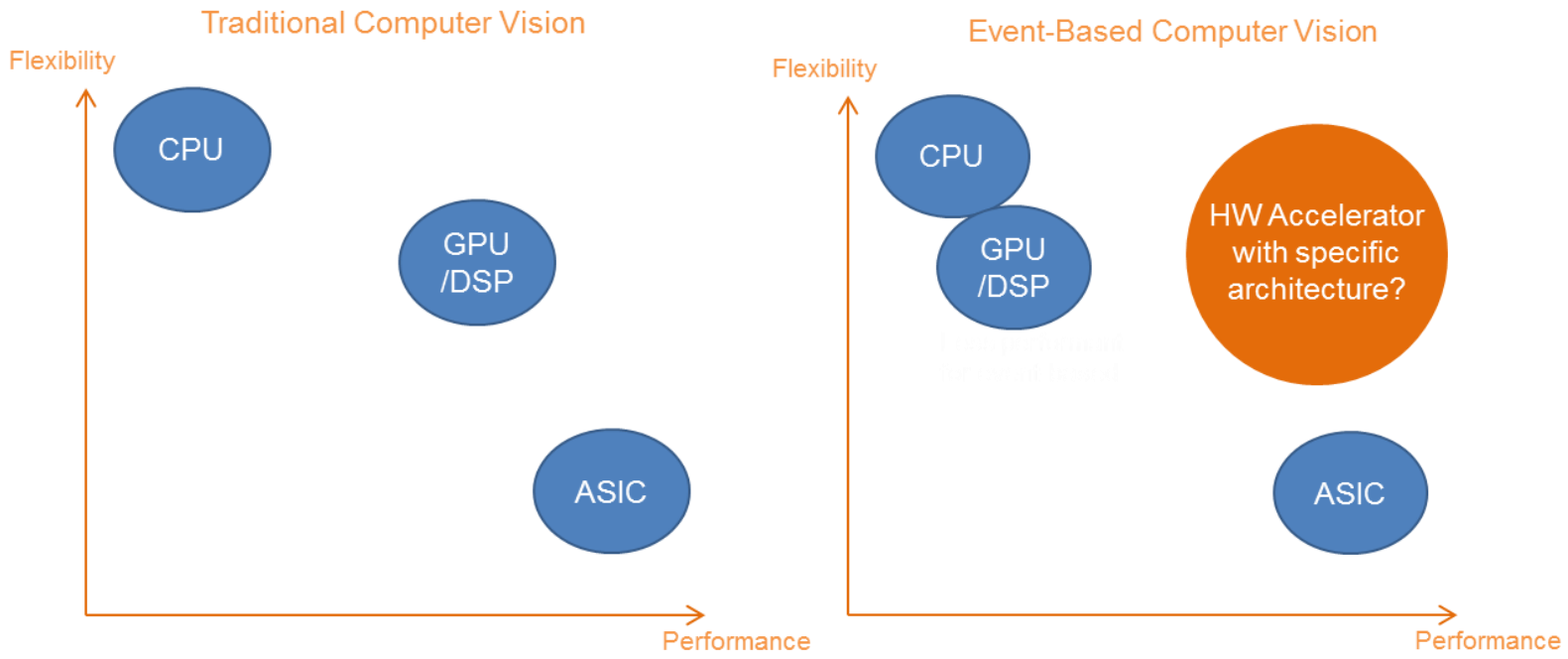
- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

Event Processing Acceleration

- GPU/DSP are not optimized for event-based computer vision:

	GPU/DSP	Requirements for Event Processing
Optimized for	High arithmetic density and high throughput	Low latency
Memory access	Privilege coalesced data	Optimized for random access

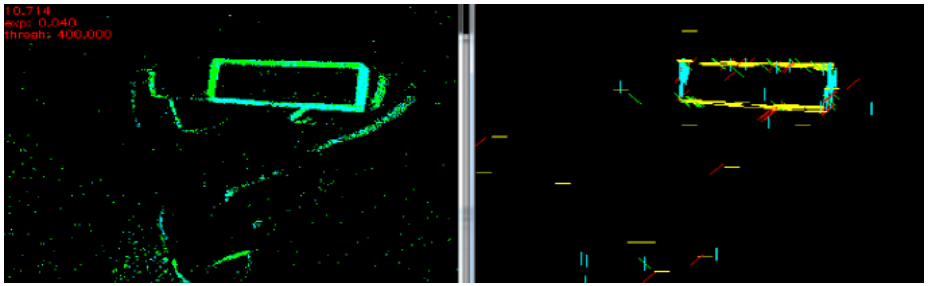
- Considering HW accelerator/processor with a DVS-specific architecture



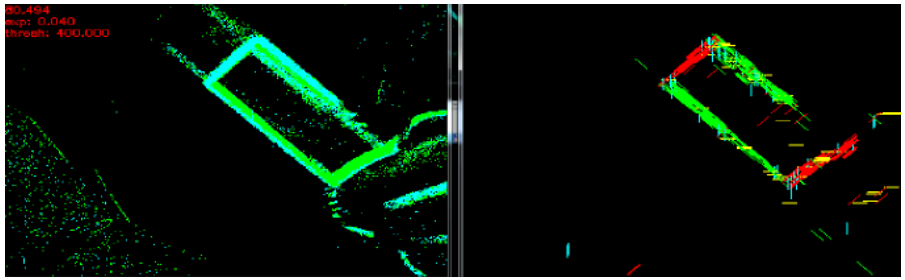
Concept: Less but More Informative Events

- Programmable feature
- A feature event is sent by the sensor when the feature appears strongly
- For example: direction detection (can be used to produce Histogram of Gradients in the Host for ***human detection***)

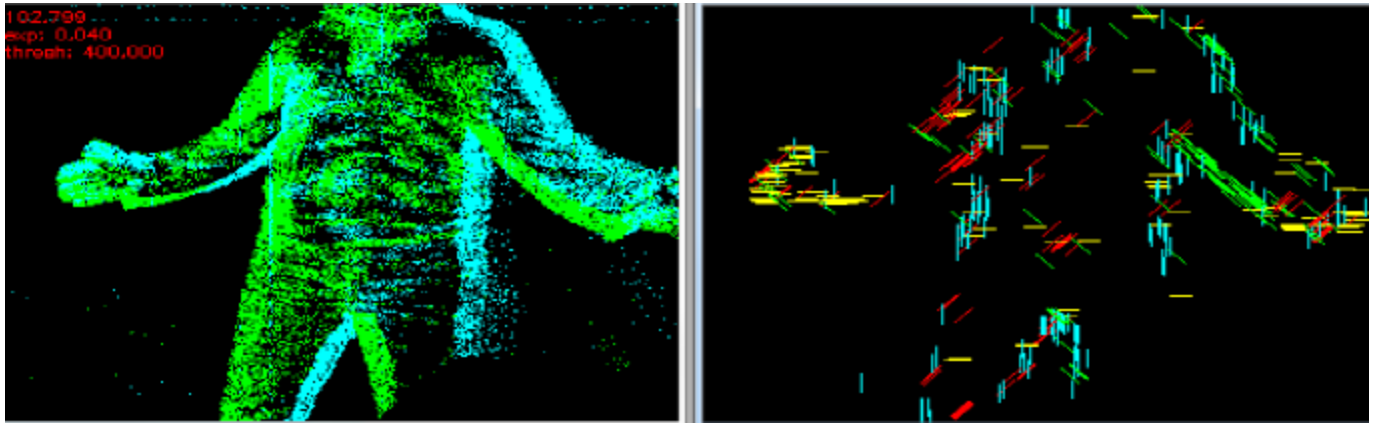
Horizontal (yellow) and vertical (cyan):



Slash (red) and Backslash (green):



Applied on Human captured is a DVS movie:



The Road to Market - Overview

- Reducing Cost and Module Size
- Reducing Power
- Delivering Good Event Quality
- Data Throughput Reduction
- Event Processing Acceleration
- Summary

Summary: Market Driven Requirements

Requirements	Samsung Early Prototypes	Samsung DVS Product Roadmap (2017,2018)
Low Cost	Large Pixel No standard for Quality / Die sorting	Similar size to 16-8MP mobile camera
Minimal Module Size	Large, due to large Pixel -> Large optical format	Smaller, Optical size reduced to 1/3” (VGA)
Ultra Low Power	Low power	Lower power
Minimal Optical Format	Large, due to large Pixel	Reduced to 1/3” (VGA)
Good Event Quality	Redundant events (especially at low light) Motion artifacts, Timing accuracy Noise, Flicker	Improved pixel circuit. On chip Bad Pixel, Noise and flicker suppression. External Triggered Scan, Global reset & hold capable
Low Data Rate	Scene dependent Might exceed CIS typical throughput	Compressed format in the bounds of slow CIS typical throughput
Edge device vision processing	Not available	Considering low level features

Thank You!