

Stream Processing – Performance Analysis

Filip Nálepa

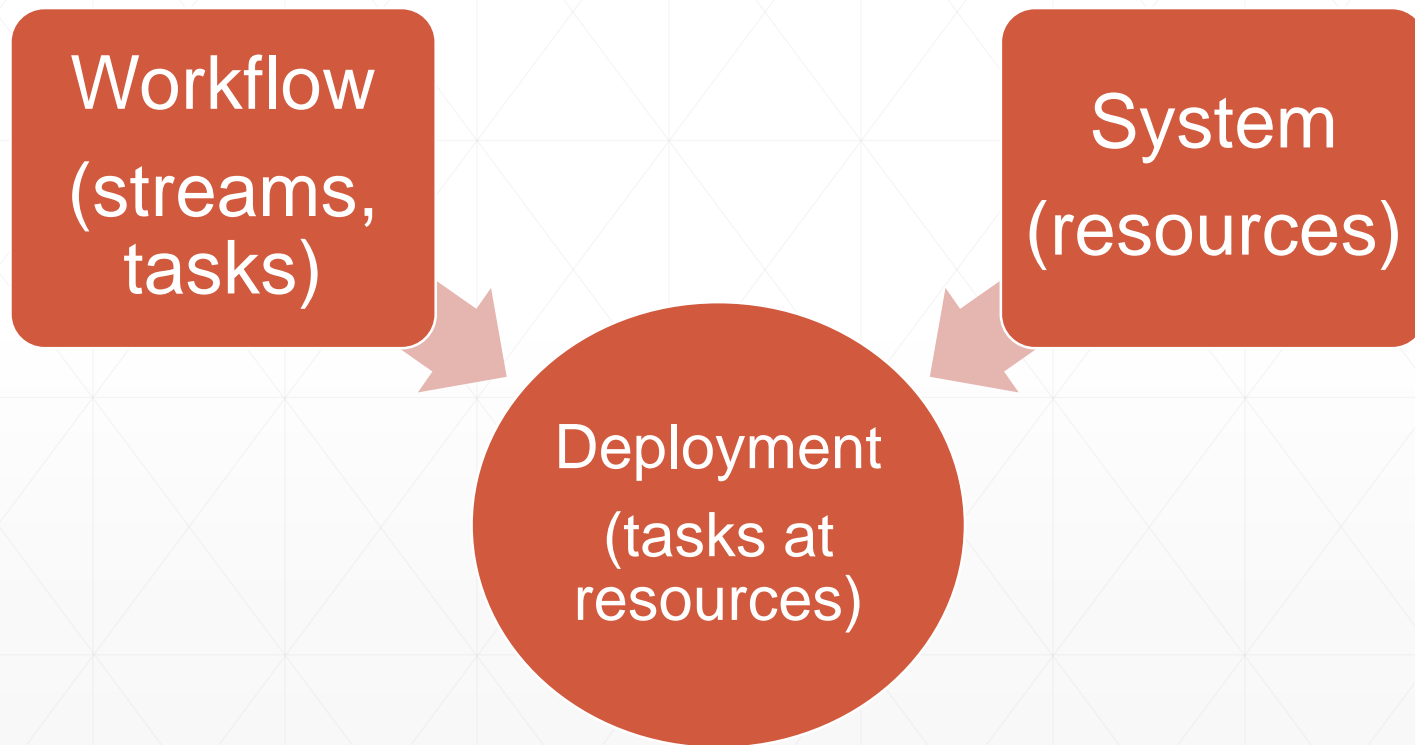
Outline

- Experiment description
 - Workflow model
 - System model
 - Deployment model
 - Results of performance analysis
-

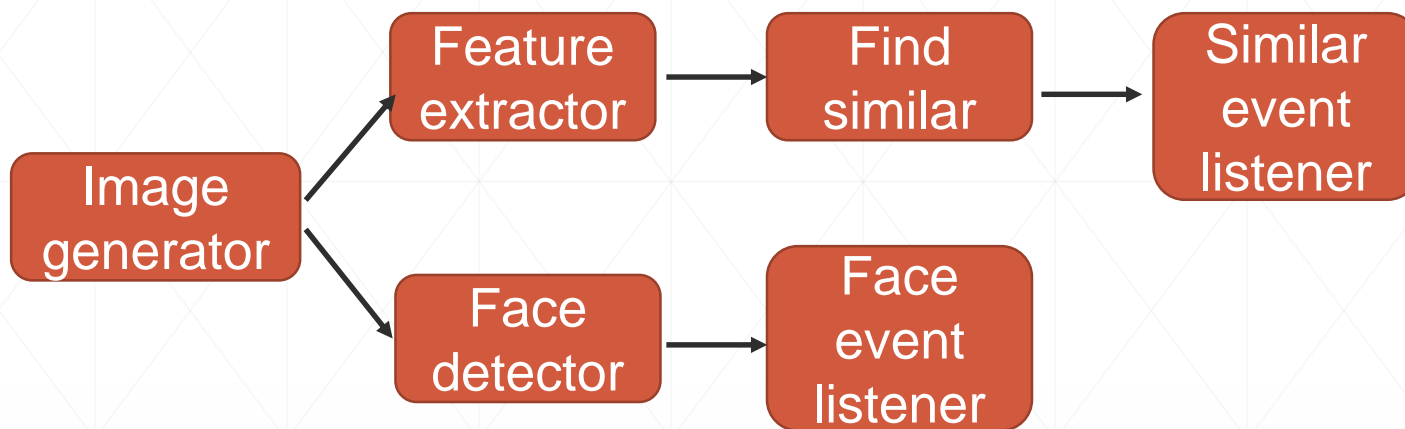
Experiment Description

- Image stream
 - Dataset ~1000 images
 - Images repeated in cycles → infinite stream
 - Tasks
 - Find similar images in DB
 - Detect faces
 - Deployment – Apache Storm
 - Measure delays
 - Model of the application
 - Convert to Colored Petri Net (CPN)
 - CPN simulation – derive minimal/maximal delays
-

Model Structure



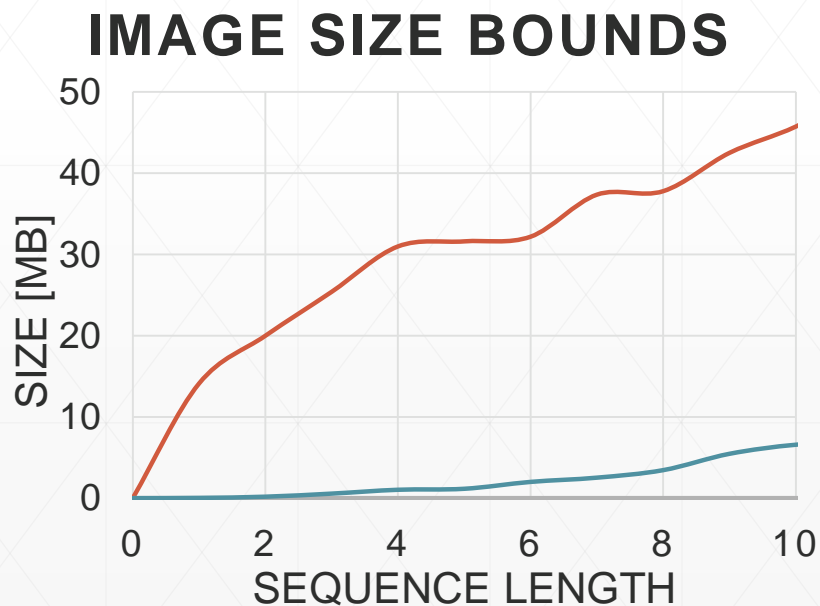
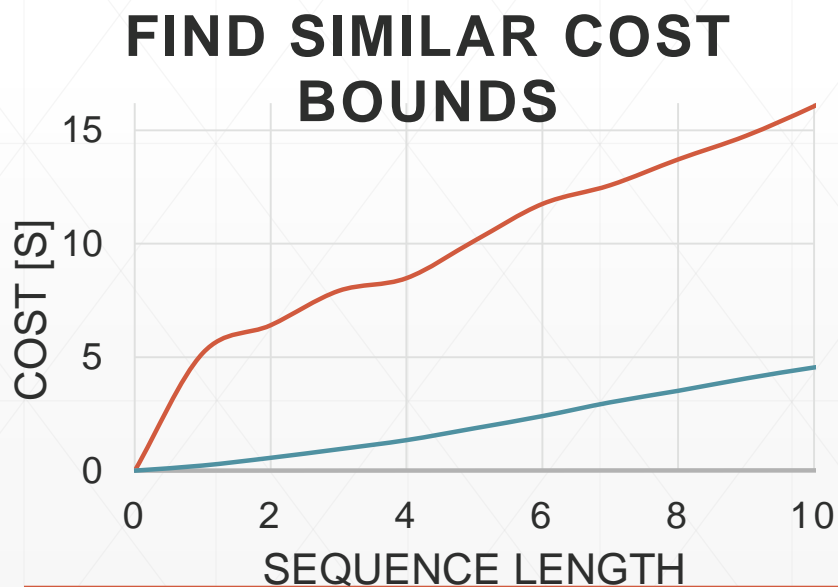
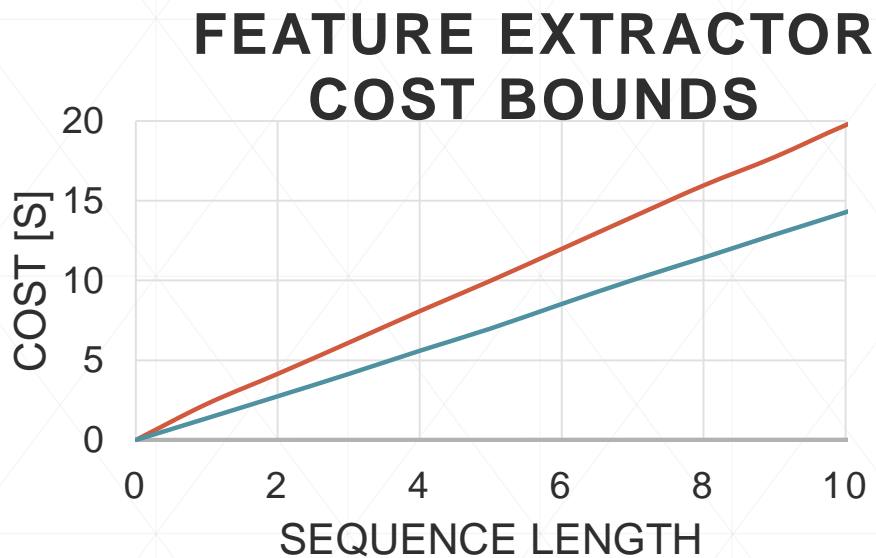
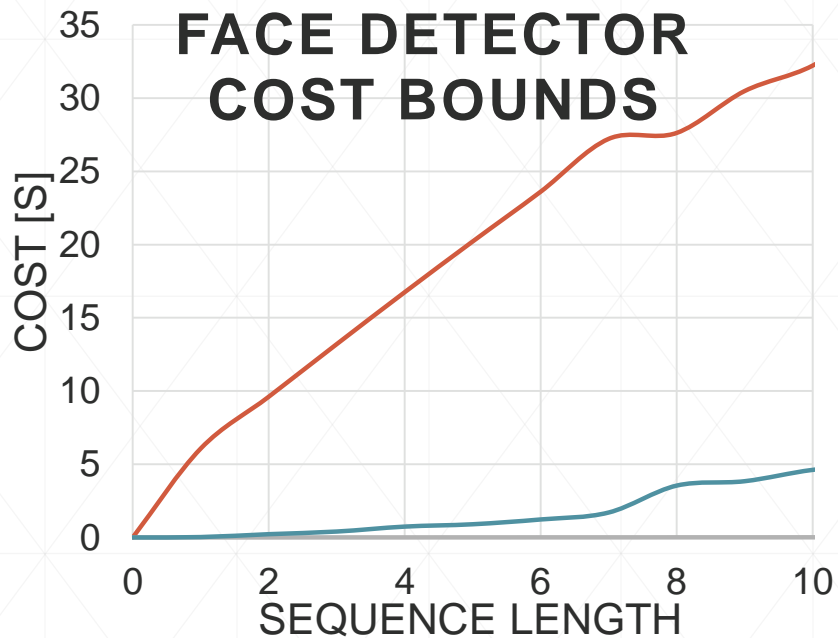
Workflow Model



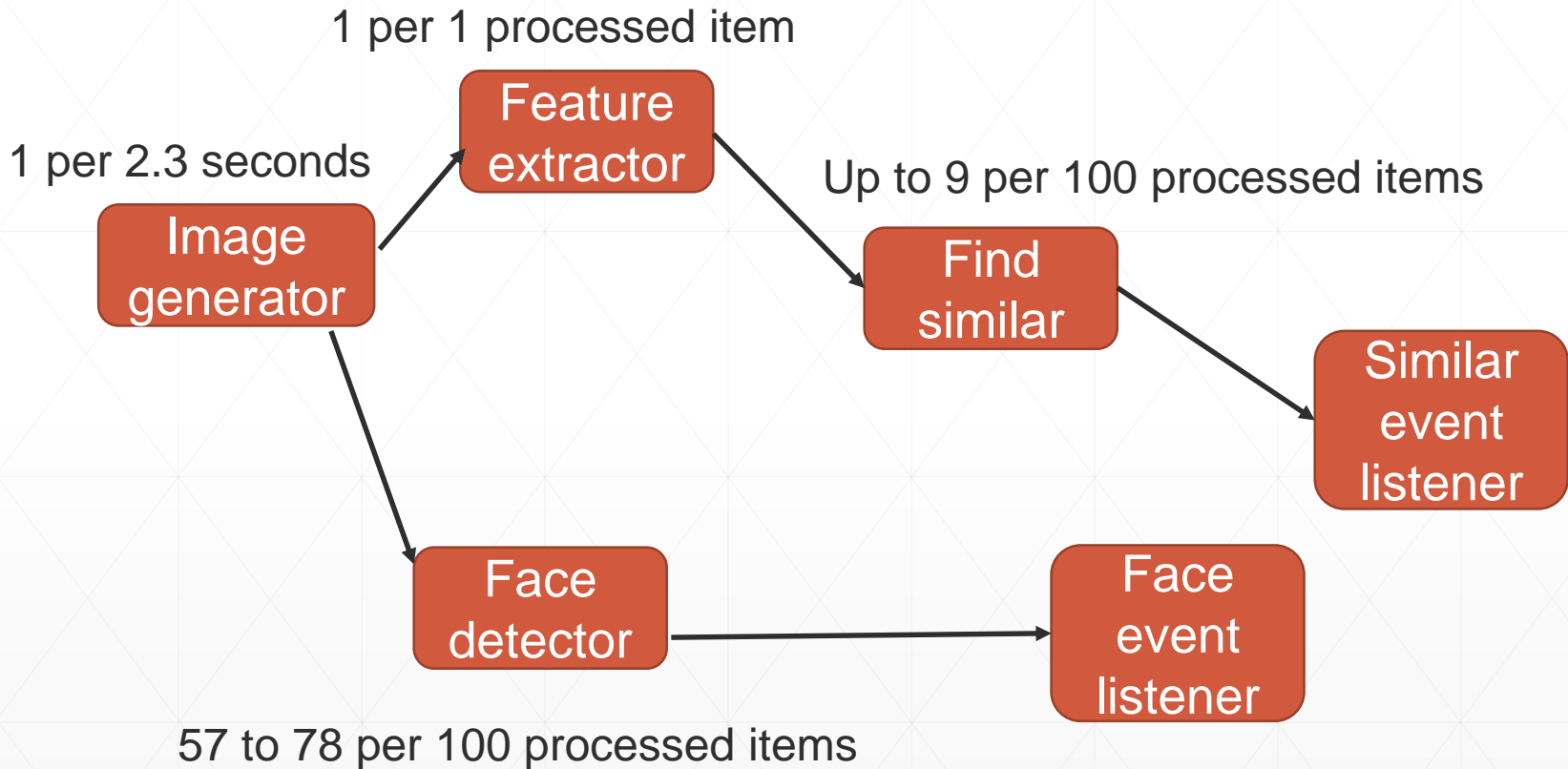
- Processing cost
 - Data size
 - Output frequency
-

Processing Cost

- $\text{cost}(\Delta) = x$, where x is the maximum number of processing units needed to process any sequence of data items of length Δ
 - Example:
 - Cost sequence: 4, 1, 1, 1, 1, 1, 1, 1, 1, 4, 4, 1, 1, 1, 1, 1
 - $\text{cost}(1) = 4$
 - $\text{cost}(2) = 8$
 - $\text{cost}(3) = 9$
 - Analogically minimal cost
-

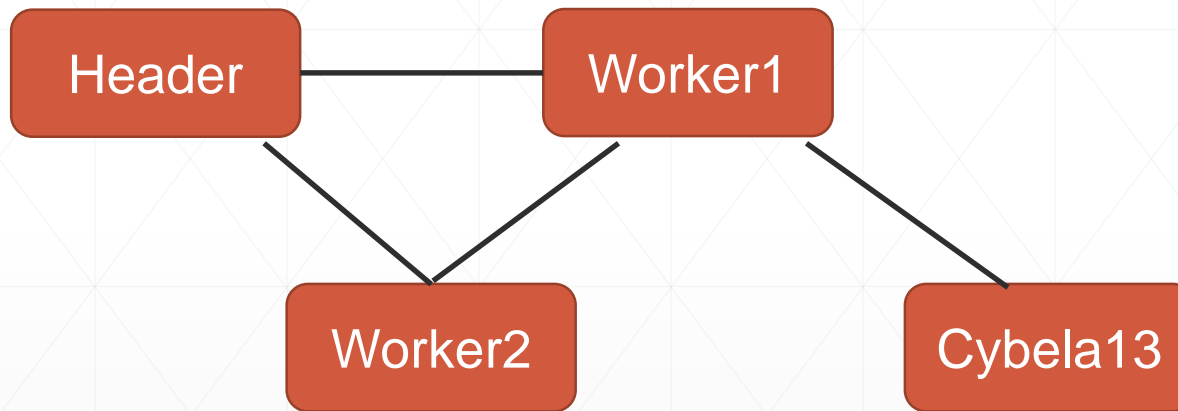


Output Frequency

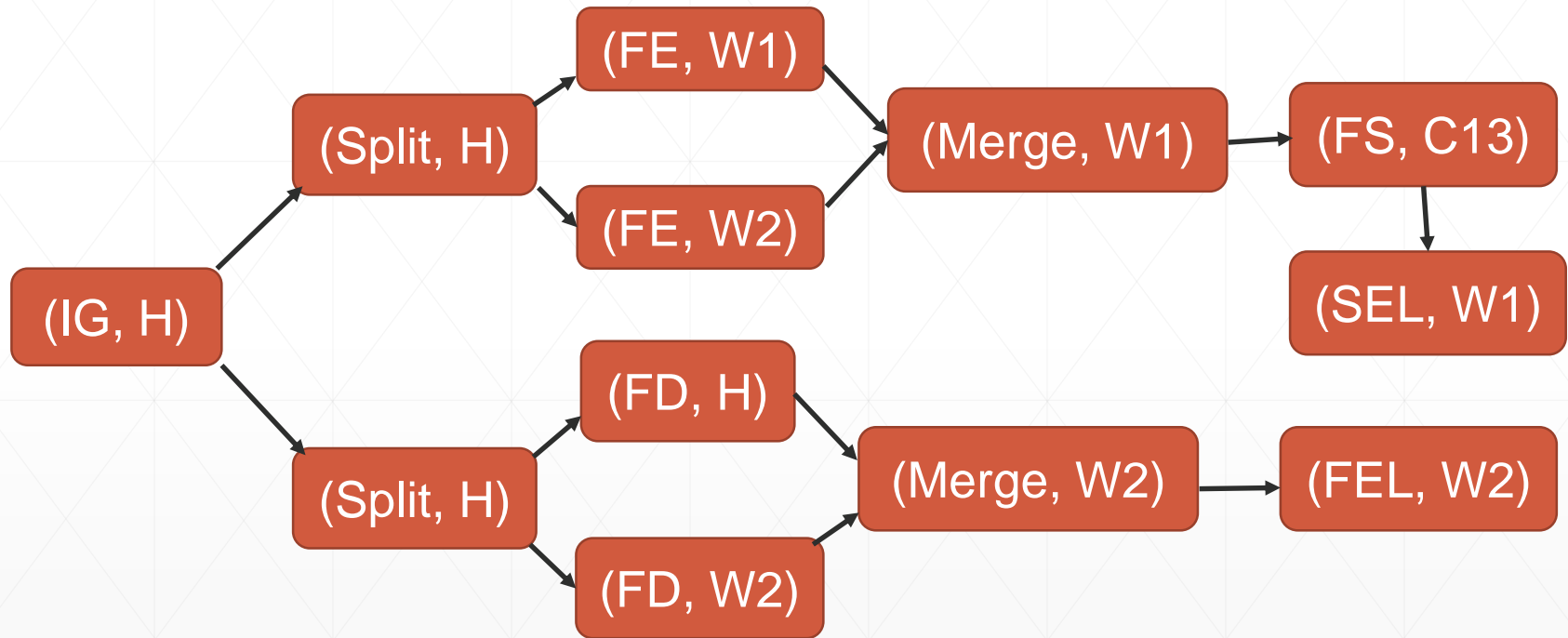


System Model

- Apache Storm cluster – Header, Worker1, Worker2



Deployment Model



- Split – equal distribution
-

Results

Storm: ~ 38,000 processed images

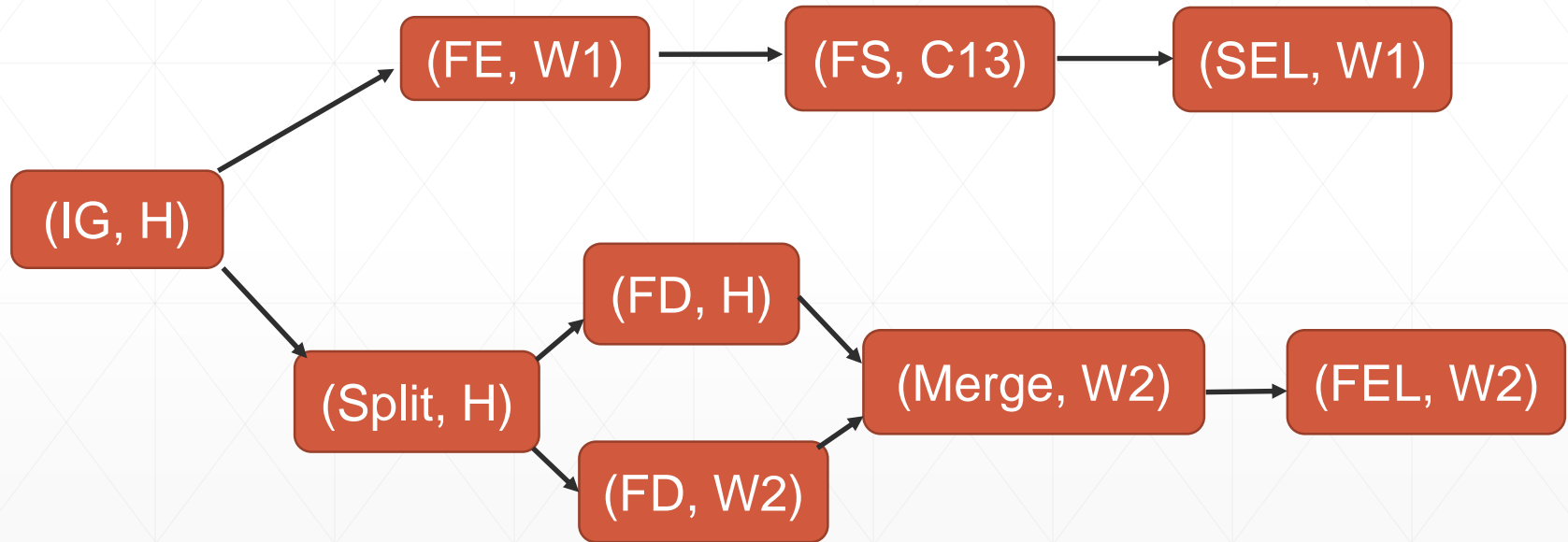
Delay – time since an image was output by the image generator until it was processed by a particular task

		FD H	FD W2	FE W1	FE W2	FS C13
CPN	max delay [ms]	7416	19529	2316	5611	12536
	min delay [ms]	27	51	1387	1362	1586
Storm	% of delays \leq CPN max	99.65	100*	99.70	96.31	99.99
	% of delays \geq CPN min	99.68	99.62	99.70	100	100

*Face detector at Worker2

- Storm max delay: 10153 ms

Deployment Model 2



- Split – equal distribution
-

Results 2

Storm: ~ 7400 processed images

	FD H	FD W2	FE W1	FS C13
CPN: max delay [ms]	7357	7500	2446	7624
Storm: % of delays \leq CPN max	99.84	99.54	99.72	99.92

Comparison

	Face event max delay estimate	Similar img event max delay estimate
Experiment #1 (two feature extractors)	19.5 s	12.5 s
Experiment #2 (one feature extractor)	7.5 s	7.6 s

Thank you for your attention

