

eFPGA for Neural Network based Image Recognition

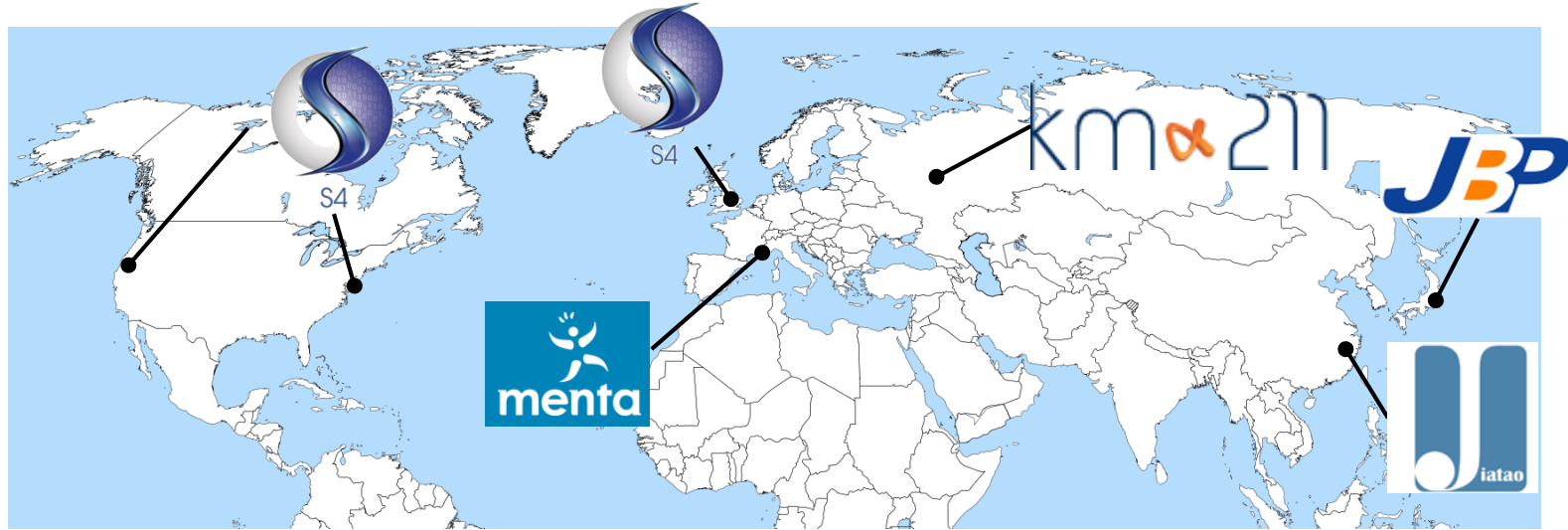
June 26, 2018

Yoan Dupret – Managing Director – Menta

yoan.dupret@menta-efpga.com



Menta Overview



17 employees



11 years of R&D



10 Tapeout on 4
nodes at 3 Foundries

Some partners

SECURE-IC  **surecore**



SYNOPSYS[®]



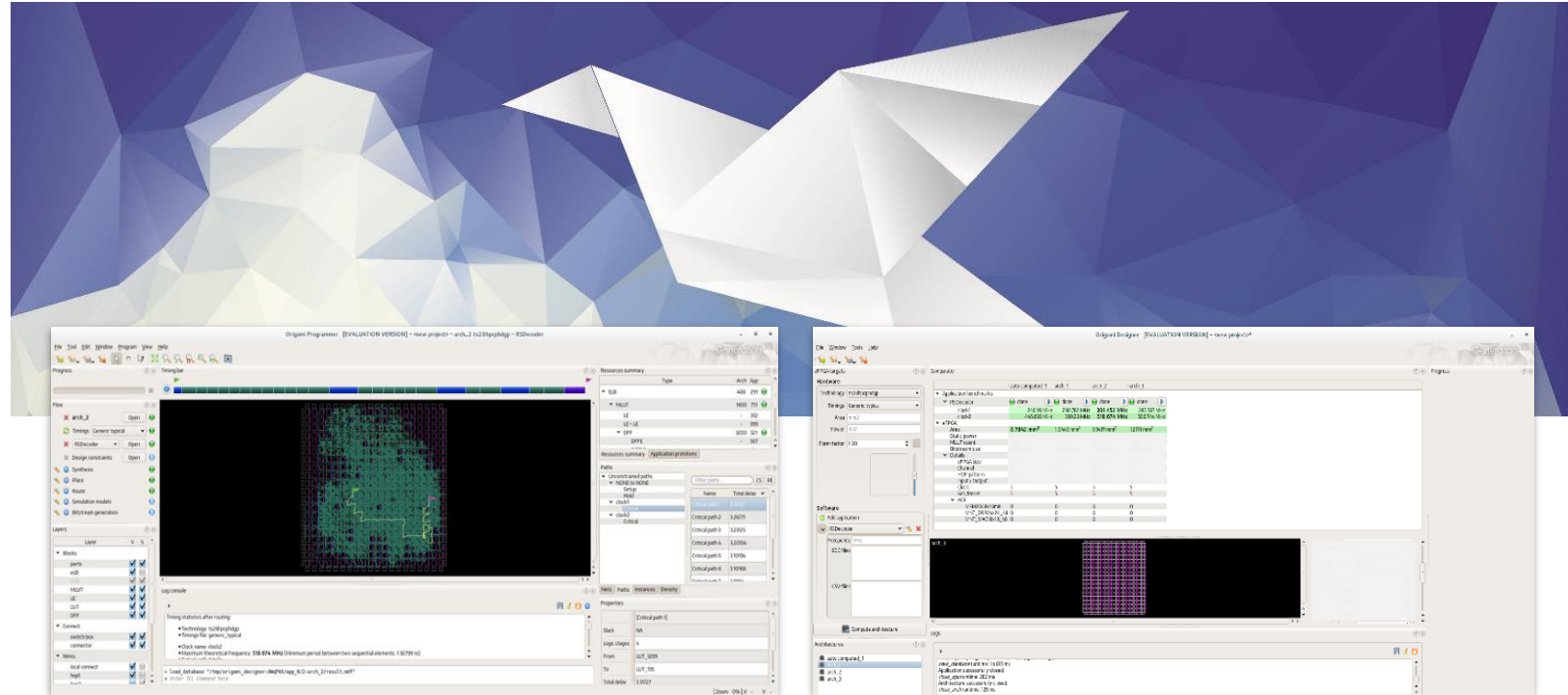
Some customers



Fully flexible standard cells eFPGA

Origami Tool suite

eFPGA IP



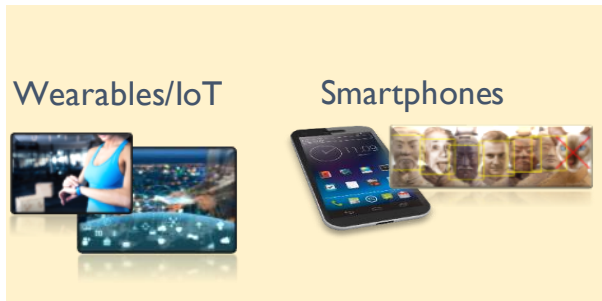
Origami Programmer

From RTL to bitstream

Origami Designer

eFPGA IP definition

Products offering



SMALL
<2K LC



MEDIUM
]2K to 6K[LC



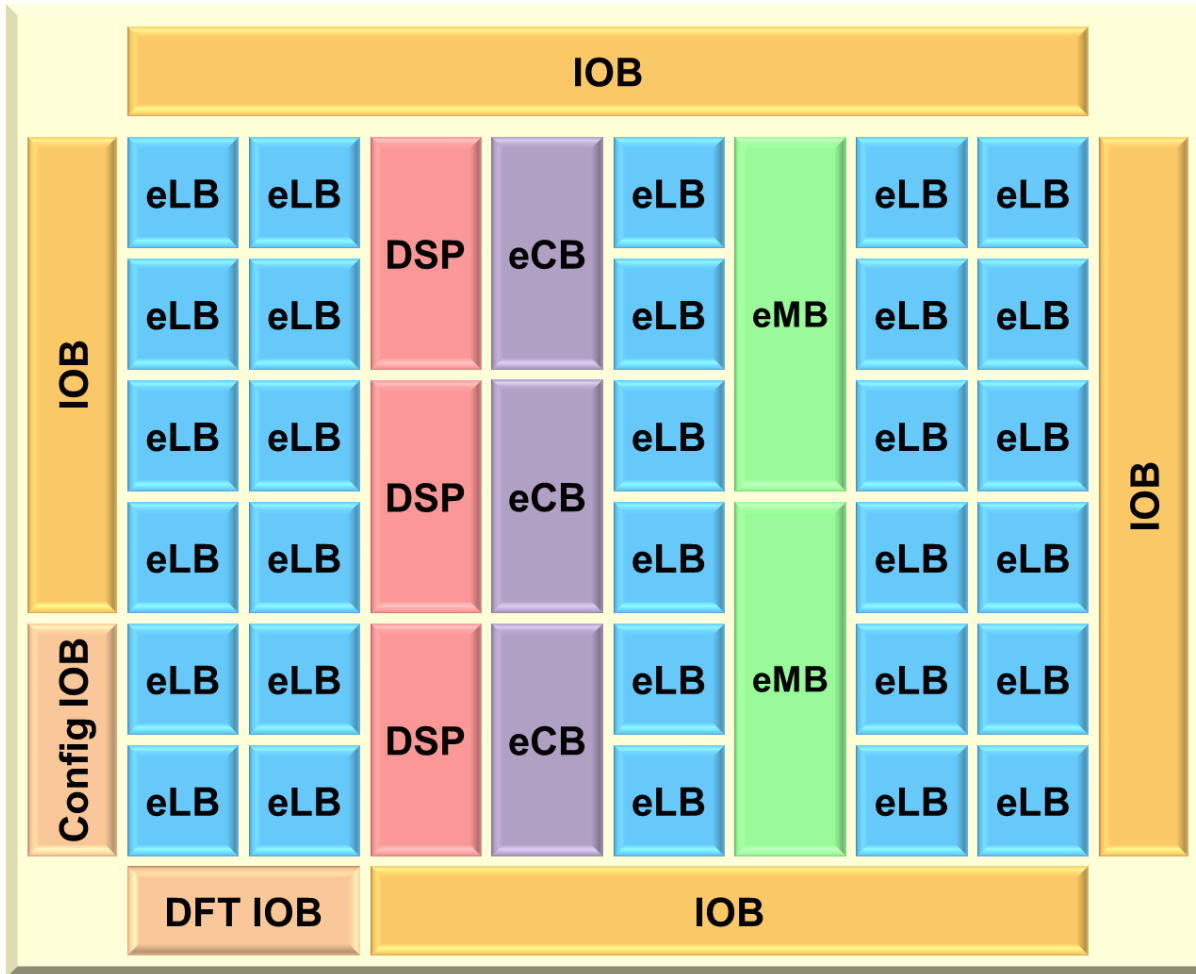
LARGE
]6K to 60K[LC

XLARGE
>60K LC

eFPGA IPs examples

	Small		Medium	Large		XLarge	
Name	M5S0.5	M5S2	M5M5	M5L15	M5L40	M5XL65	M5XL130
# LC	538	1 690	5 120	15 232	40 141	65 434	~130 000
# MAC	0	6	10	0	0	0	~1000
# CDSP	0	0	0	21	48	18	0
SRAM (kb)	0	0	1 311	0	2 097	2 359	5 252
# IOs	208	384	1 032	2 304	3 744	4 640	6 500
Interface / Sensor Hub	•	•	•	•			
Consumer Electronics	•	•	•	•			
Communication	•	•	•	•	•	•	
Signal Processing		•	•	•	•	•	
Industrial	•	•	•	•	•		
Aerospace & Defense		•	•	•	•	•	•
Security			•	•	•	•	
Automotive			•	•	•	•	
Data Center & HPC						•	•

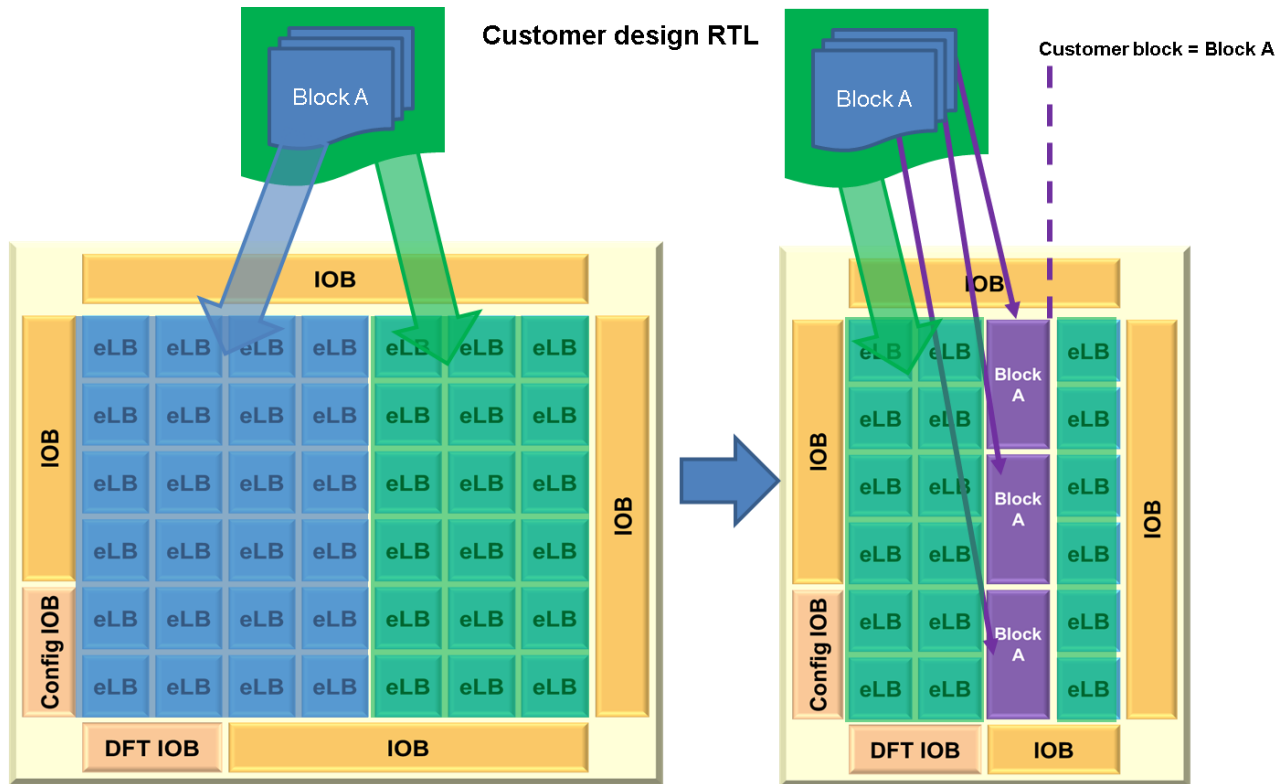
5th generation eFPGA IP



- High density & performances
 - Patented MLUTs - LUT6 based
- Fully flexible
 - Number of eLBs
 - ASIC like eMBs: type, quantity & size
 - eCBs:
 - Catalogue of DSP blocks (MAC, CDSP)
 - Custom blocks
 - Number of IOs
- No specific interface
 - Data: AXI, AHB, proprietary buses, direct connections, etc.
 - Configuration: SPI, AXI/AHB, JTAG, etc.
- Various ASIC like power management options
- Standard scan chain. TC > 99.7%
- 100% 3rd party standard cells based
- Robust verification flow

GDSII delivery in 1 to 6 months
Other models possible

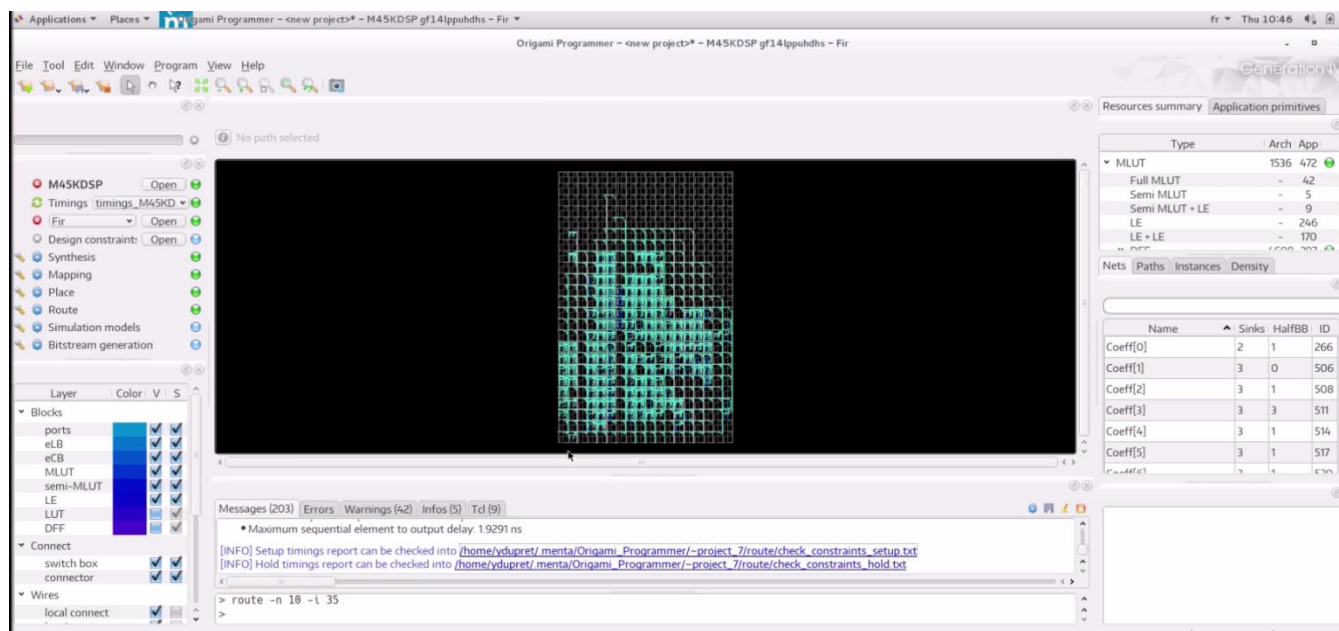
Custom blocks



- Embedded customers arithmetic blocks
 - Optimization of performances & area
 - Integration:
 - Complete integration possible (auto-inference, STA)
- or
- As a black box

Custom block to target AI applications under definition (Kortiq / Menta)

Origami Programmer – From RTL to bitstream



- Inputs:
 - Application RTL
 - IEEE Verilog / SystemVerilog / VHDL
 - Constraints: timing & IOs
- Outputs:
 - Bitstream
 - Simulation model
 - Timing reports
- GUI interface:
 - STA
 - Resources visualisation
 - Congestion maps
 - Move and reallocate resources
- Soft IPs generator

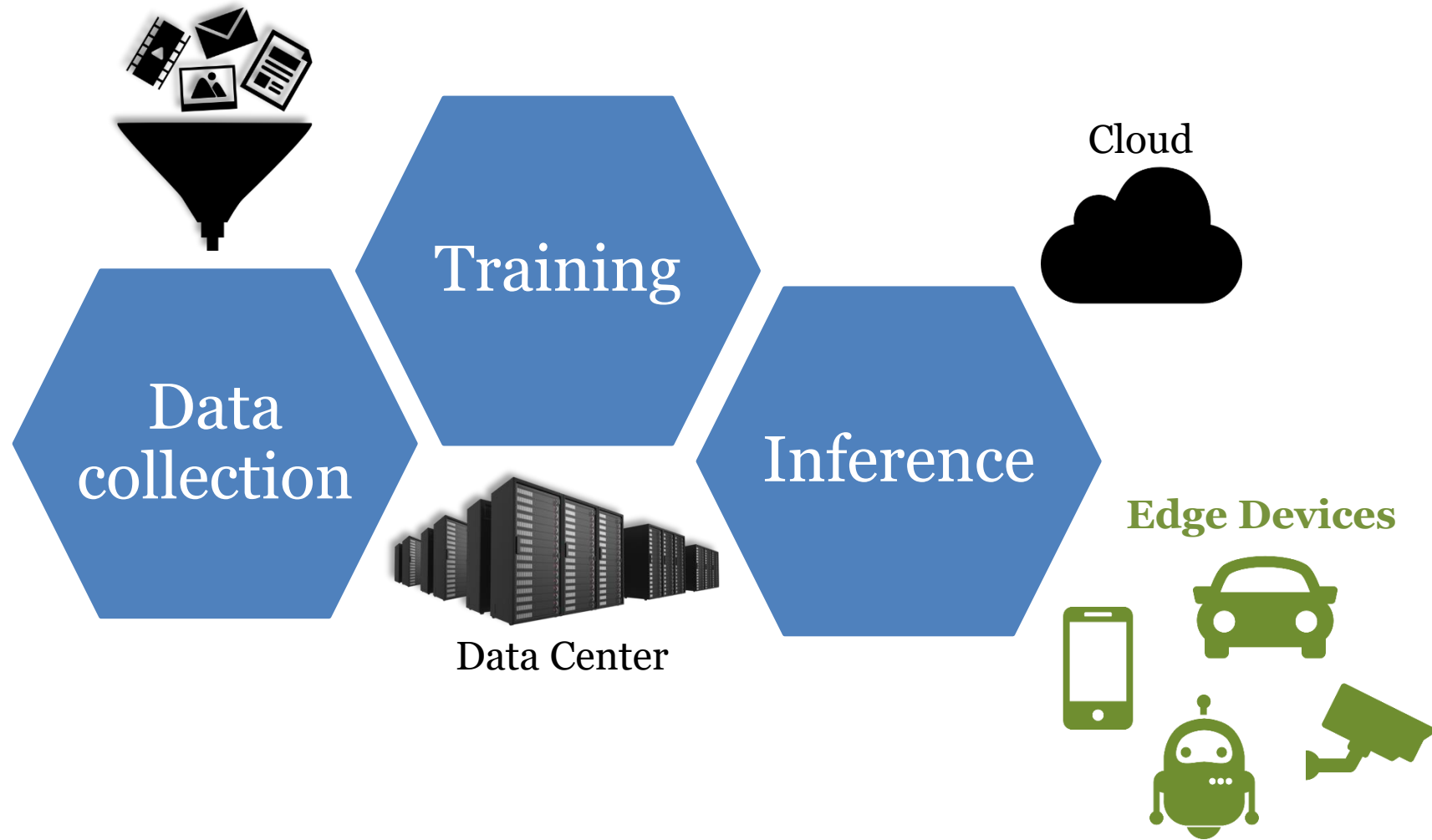
Various distribution models possible

eFPGA IP requirements & Menta answers

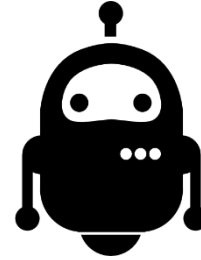
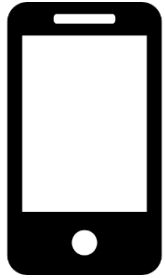
- Seamless SoC Integration
 - Menta 100% standard cells
 - No specific interface.
- High yield and reliability
 - Menta is using Flops instead of SRAM bitcells for bitstream storage
- Area sensitivity
 - Support of any kind of arithmetic blocks (DSPs) and SRAM (i.e. BRAM / LRAM)
 - IP is fully flexible. Specification software available: Origami Designer
 - PPA comparable to competitors
- Testability & test cost
 - Standard scan chain DfT with TC > 99.7% (patented)
- Fully verifiable in customer EDA environment
- Flexible Origami Programmer business model



AI inference at the edge



Inference on edge devices requirements



Cost for high volume production (mobile phones, security camera)

Flexibility over lifetime (cars, robots)

Power consumption (all)

One fit all solution

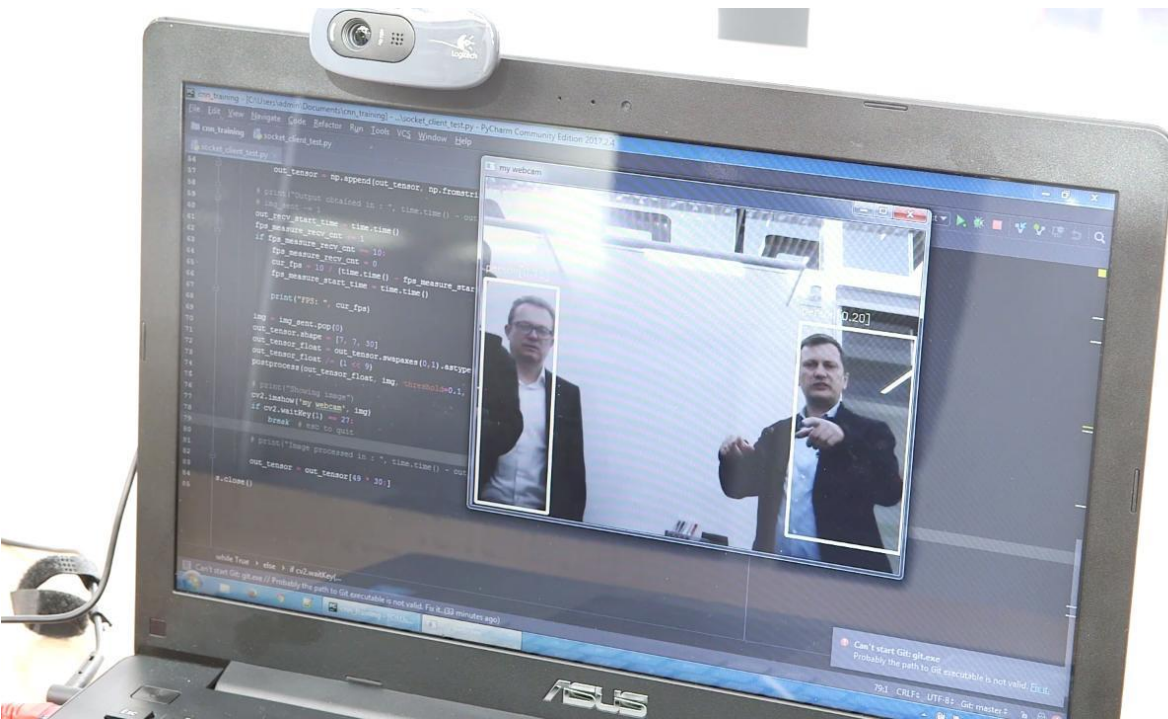
	ASIC	GPU	FPGA	DSP
Cost	+++	++	-	++
Flexibility	NA	++	+++	+
Power consumption	+++	-	+	+

eFPGA with custom Block Modules running highly efficient CNN algorithms

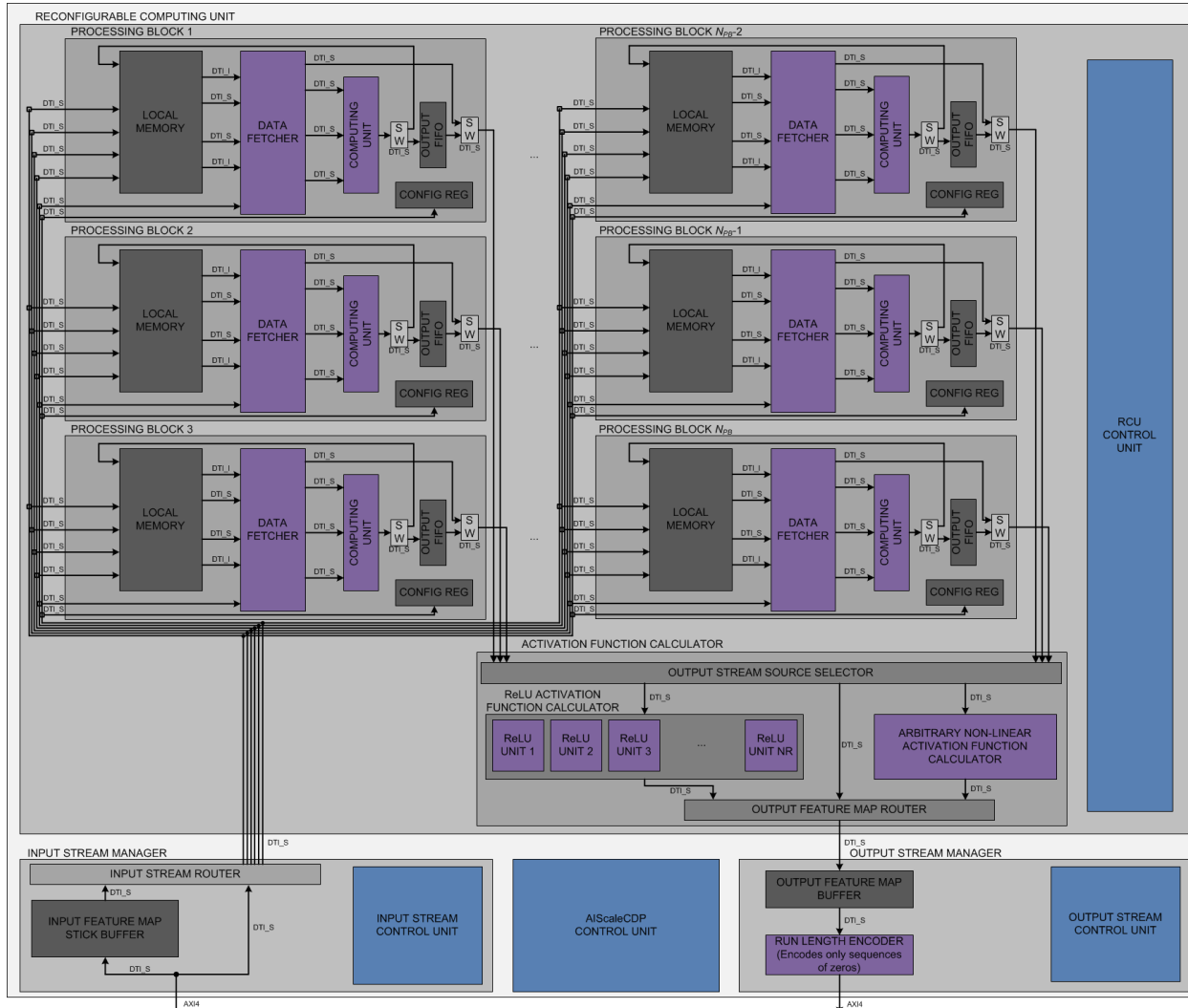
Menta + Kortiq	
Cost	++
Flexibility	+++
Power consumption	++

About Kortiq

- Munich based company specializing in IP cores for hardware acceleration of machine learning algorithms
- IP portfolio includes IP cores for accelerating decision trees, support vector machines, artificial neural networks and ensemble classifiers
- Latest addition to the IP portfolio is the AIScale, universal CNN accelerator IP core
- More information available at www.kortiq.com



Kortiq IP on FPGA



- Smallest & most efficient CNN accelerator
- Works with compressed CNNs and feature maps
- Employs novel “all zeros skipping” technology
- Capable of accelerating all standard CNN layers, including convolutional, pooling, adding, fully-connected
- Supports on-the-fly, dynamic changing of CNN architecture that needs to be accelerated
- Implementable using any Xilinx FPGA family

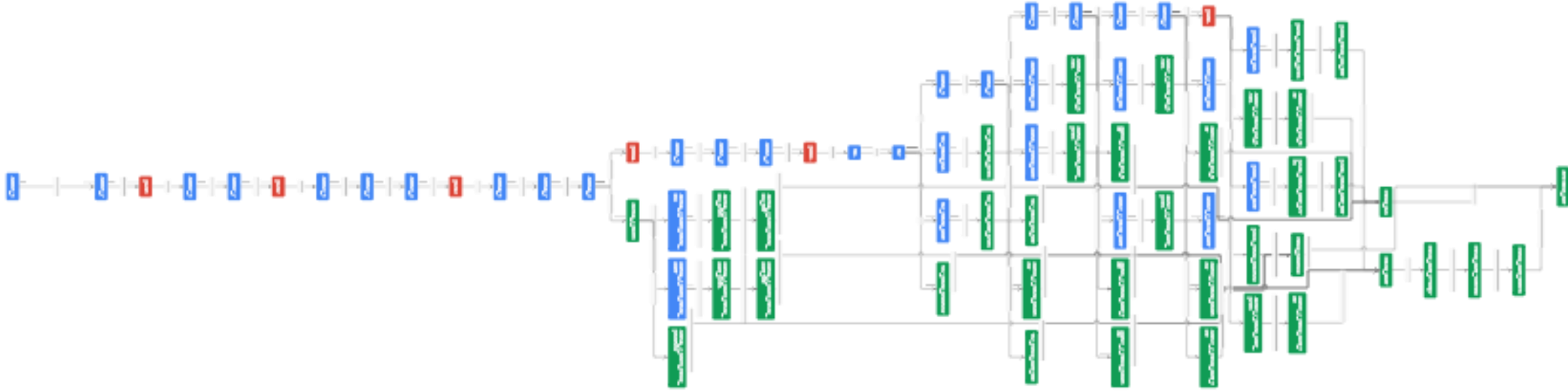
ADAS vision application



- Pilot project with a well-known car manufacturer
- Phase 1 – traffic sign detection and recognition
- Phase 2 – person detection
- Phase 3 – fully autonomous driving system

Phase 1 Details

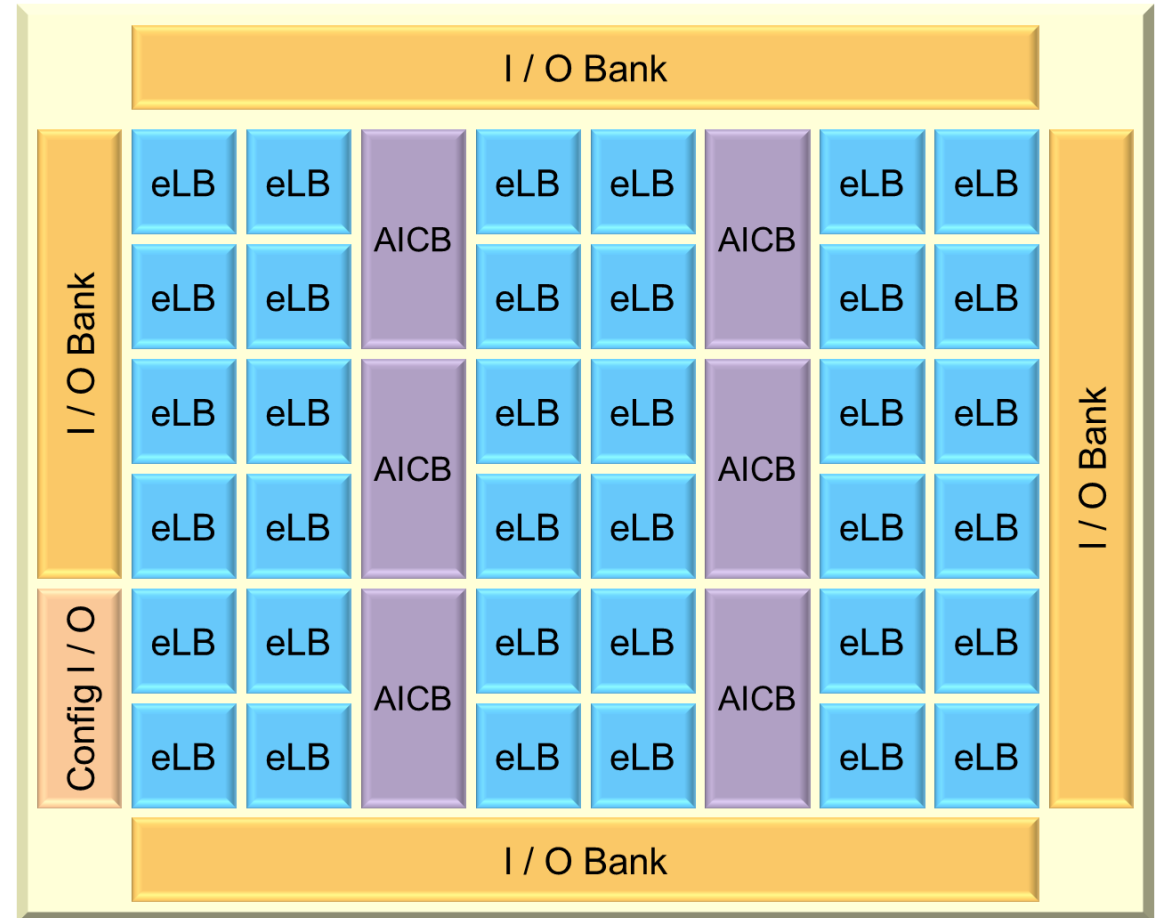
- Single Shot MultiBox Detector (SSD) method, using various CNN architectures as the base classifier
- Input image size 300x300
- Target frame rate > 25 fps



VGG16SSD CNN Visualization made by GenProto

Menta Kortiqli Cooperation

- KortiQ's ML IP cores on Menta eFPGA
- New Custom Block, AICB
- New AI Reconfigurable Architecture, AIRA
- ML SoC based on AIRA



A glowing blue brain is positioned on the right side of the image, resting on a complex network of blue circuit lines. The brain has a metallic, reflective texture and is surrounded by a dense web of circuitry that fills the entire background. The overall color scheme is a monochromatic blue with varying shades and highlights.

Thank you