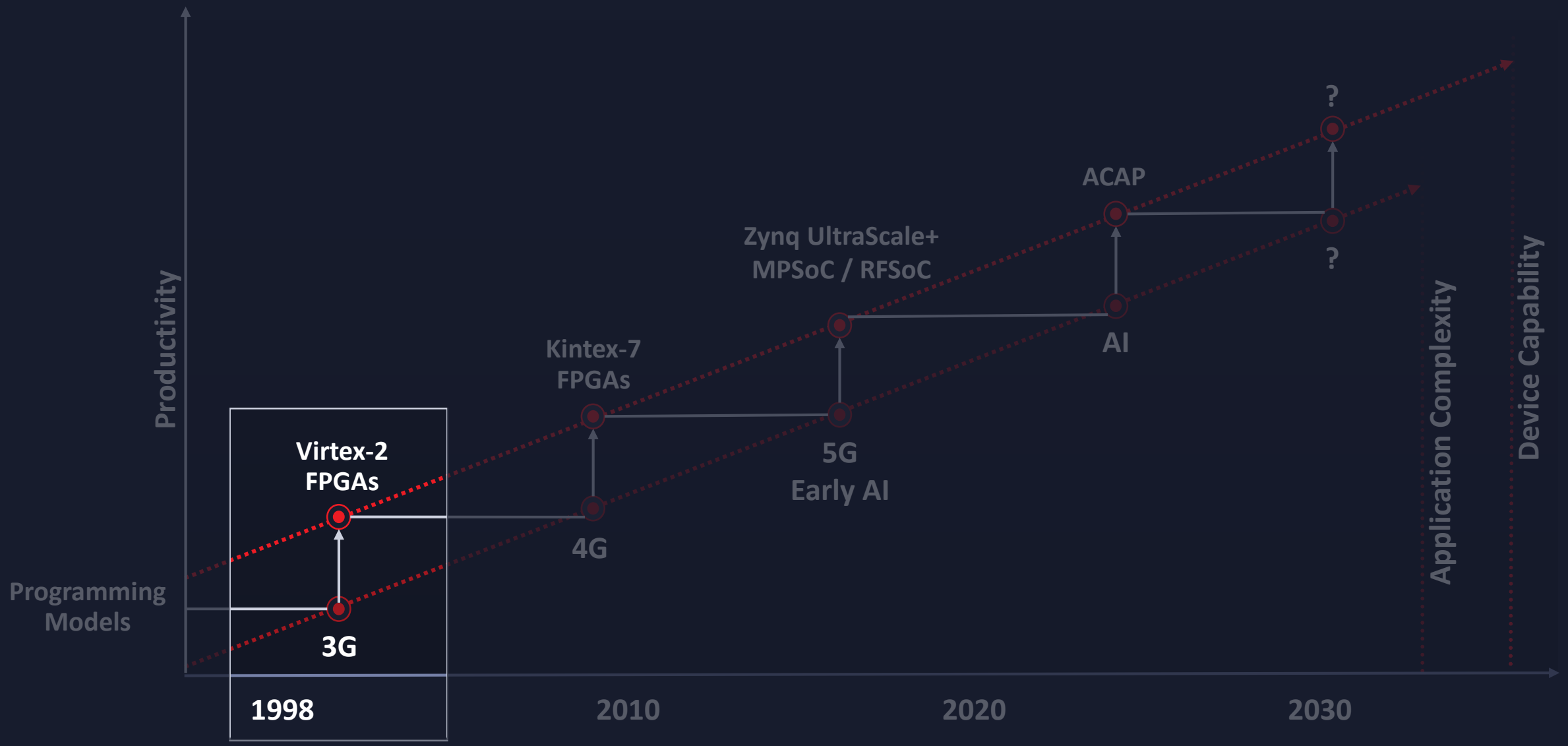
Unleashing the Power of FPGAs through Model-Based Design

Nabeel Shirazi
Senior Director
System Level Design Tools

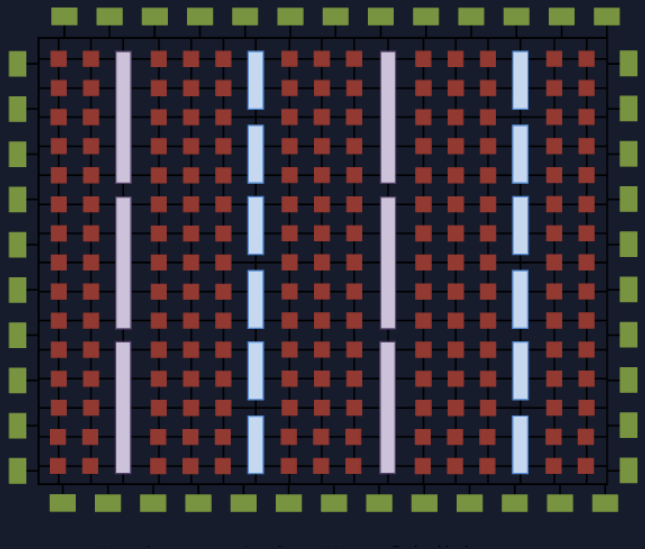
XILINX

Intersection of Applications, Devices, and Tools



In the Beginning: FPGAs for 3G Wireless Radios

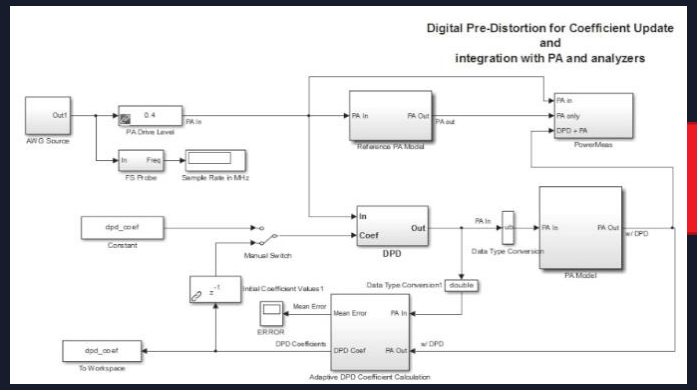
- Key application: Digital Baseband Pre-Distortion
  - Enables use of cost effective non-linear power amplifiers
  - Lowers spectral noise floor
- Why FPGAs:
  - Custom memory hierarchy and parallel processing
  - Enabled 3X cost reduction vs. analog implementations
- Programming Model:
  - Simulink for algorithm development
  - Traditional FPGA tools for implementation



FPGA

Gap between Simulink and FPGAs

System Architects & Algorithm Designers



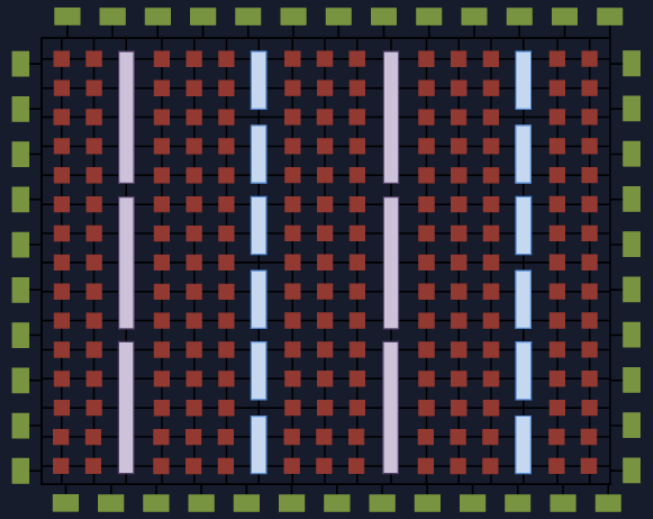
Simulink

Hardware Designer

Manual translation

RTL code

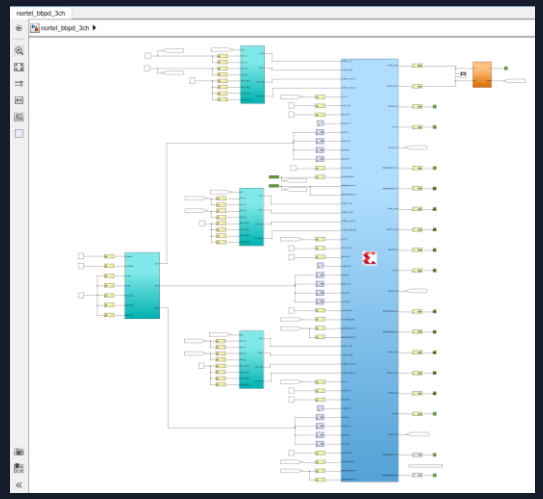
FPGA Tools



FPGA

Bridging the Gap between Simulink and FPGAs

System Architects & Algorithm Designers



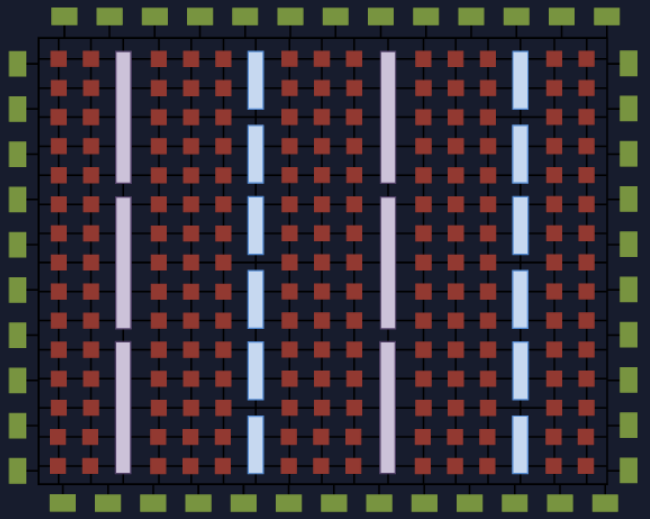
Simulink with
Xilinx System Generator
Blockset

Hardware Designer

Manual translation

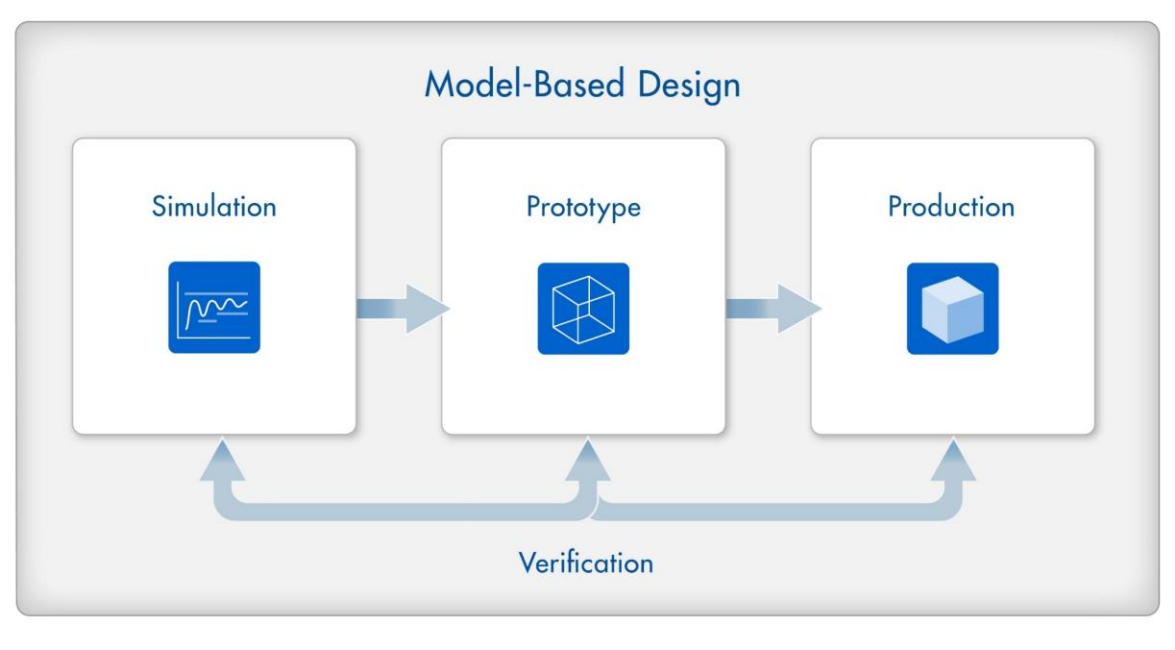
RTL code

FPGA Tools



FPGA

Why Model-Based Design



BAE Systems Achieves 80% Reduction in SW-defined Radio Development Time

645 hrs:

VHDL Expert using traditional design flow

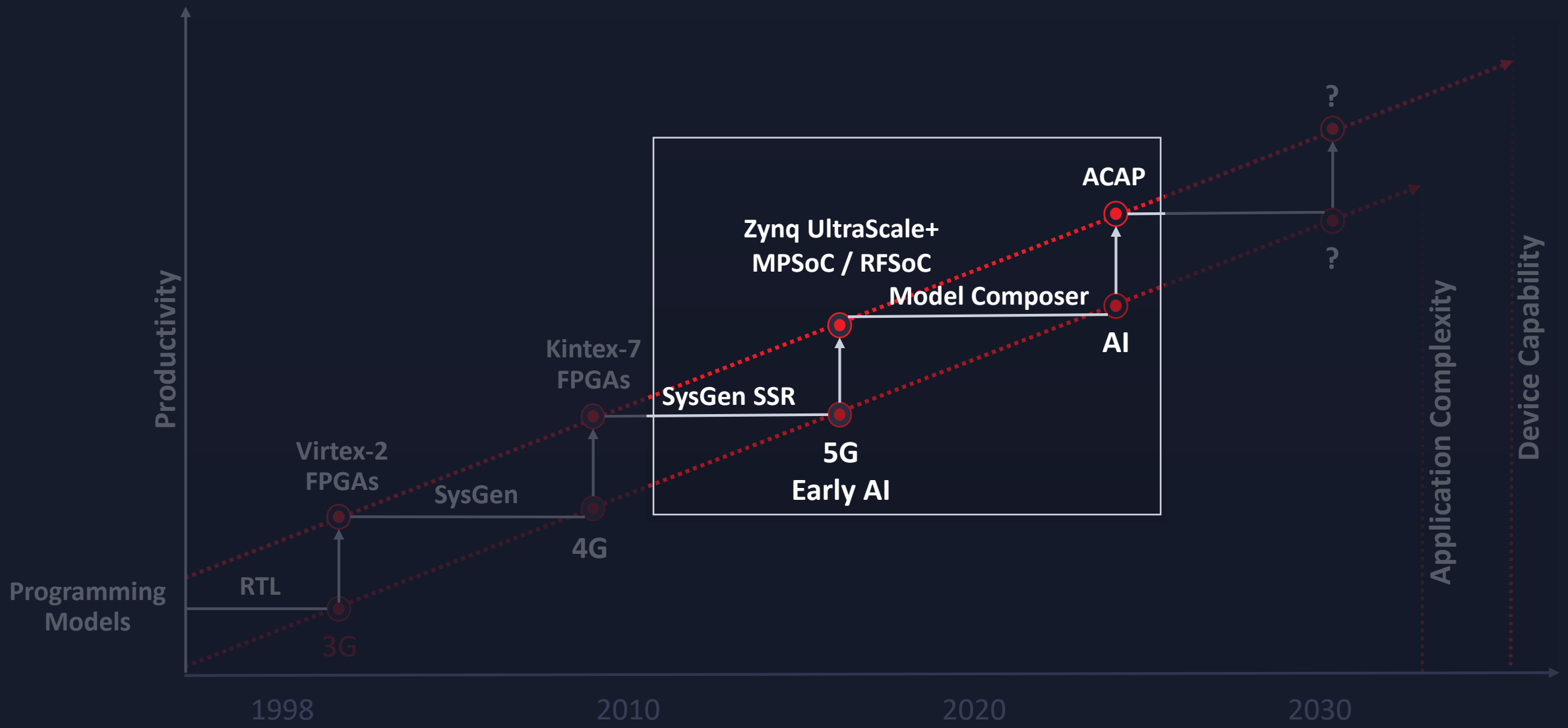
vs.

46 hrs:

Engineer with Simulink + System Generator flow

- Natural way to express parallelism
- Debug and test at the model level
- Reduce number hardware iterations
- Share models across different disciplines – FPGA, RF, Communications

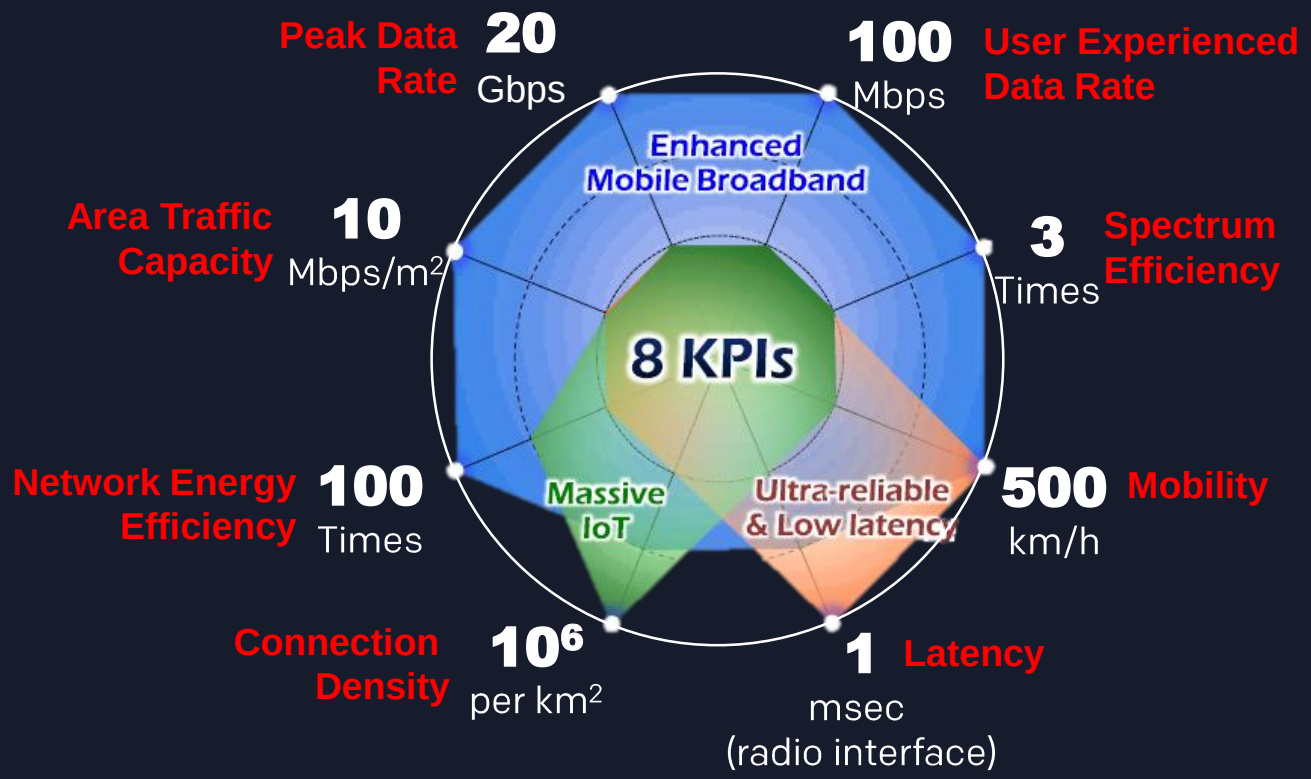
Meeting Today's Challenges



5G Wireless Radio Challenges

5G Complexity is 100X 4G

Still Evolving Standard



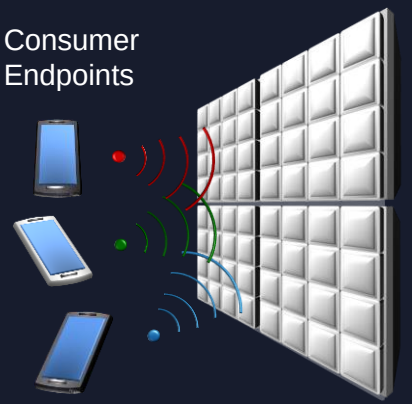
New Technologies in 5G

- Multi-user Massive MIMO
- New beamforming technology
- Millimeter wave transmission

ETRI RWS-150029,
5G Vision and Enabling Technologies: ETRI Perspective 3GPP RAN Workshop
Phoenix, Dec. 2015
http://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_70/Docs

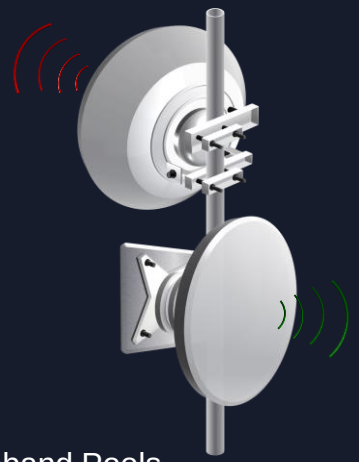
Zynq RFSoC Devices for 5G Applications

Remote Radio Head
& Fixed Wireless Access
Enabling Massive-MIMO
2D Antenna Arrays



5G

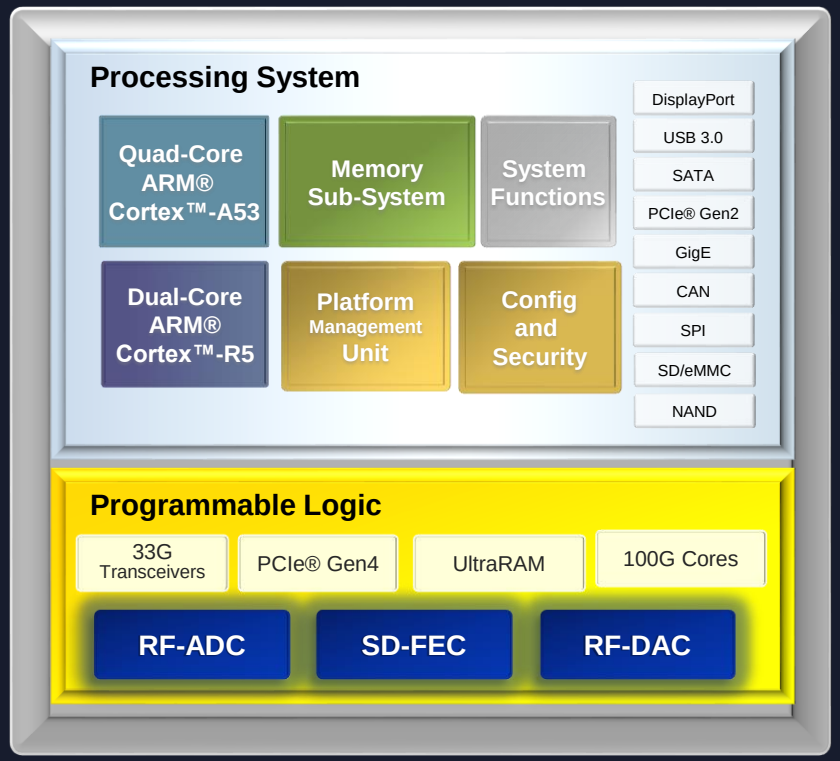
Wireless Backhaul
Enabling Throughput for
mmWave Transmission



Baseband
Maximizing Throughput in Baseband Pools

RFSoC

Heterogeneous Multi-processing



Analog-to-Digital Converters
Up to 4 GSPS

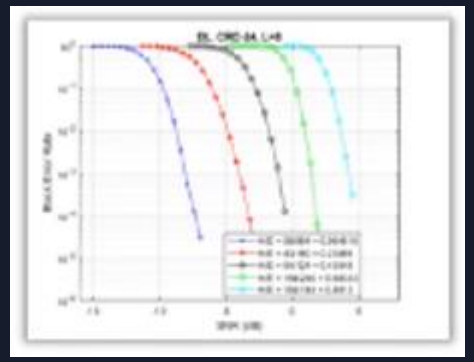
Soft Decision Forward Error Correction

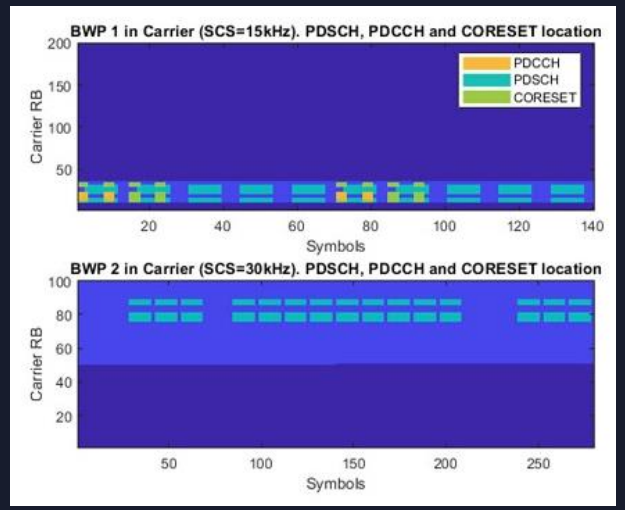
Digital-to-Analog Converters
Up to 6.4 GSPS

5G Design in MATLAB & Simulink

5G Toolbox

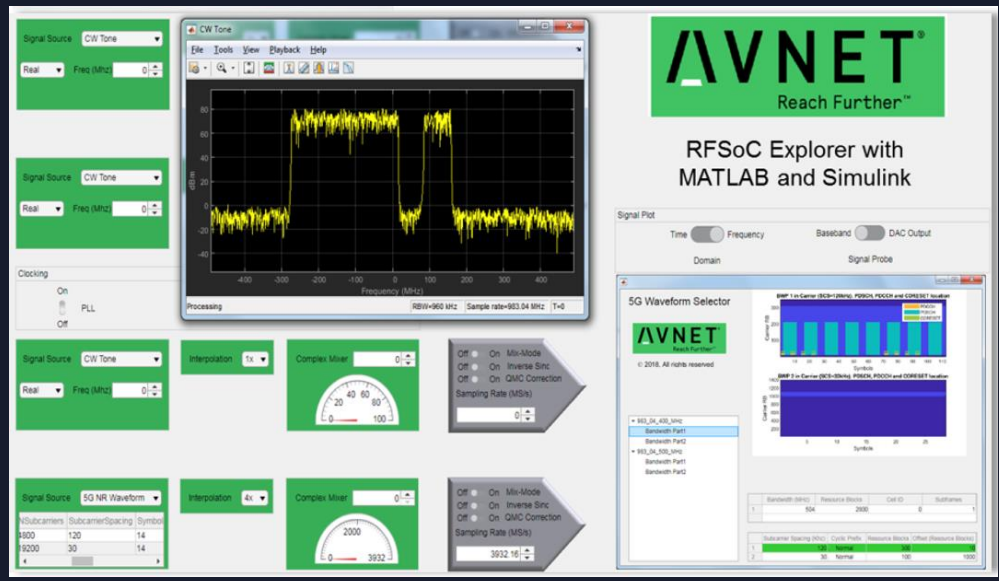
- End-to-End Link-Level Simulation
- Waveform Generation and Analysis
- Golden Reference Design Verification





Avnet RFSoC Explorer

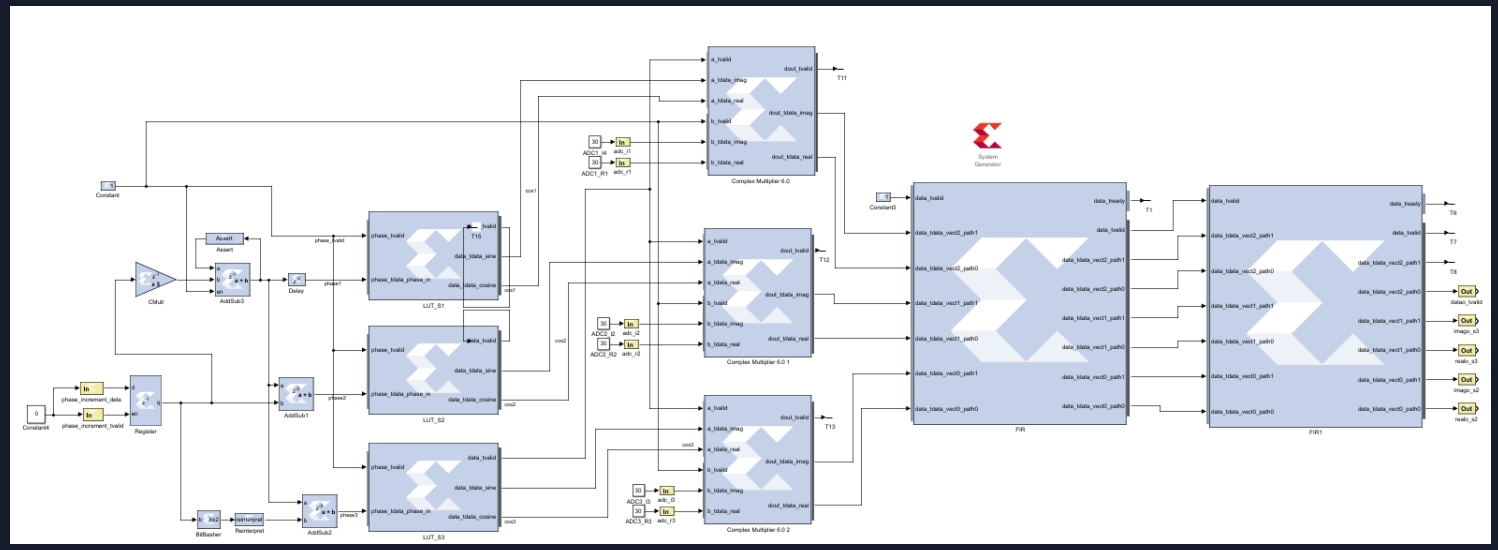
- Generate stimuli
- Setup of device
- Analyze signals from H/W



Super Sample Rate Processing in SysGen

Before:

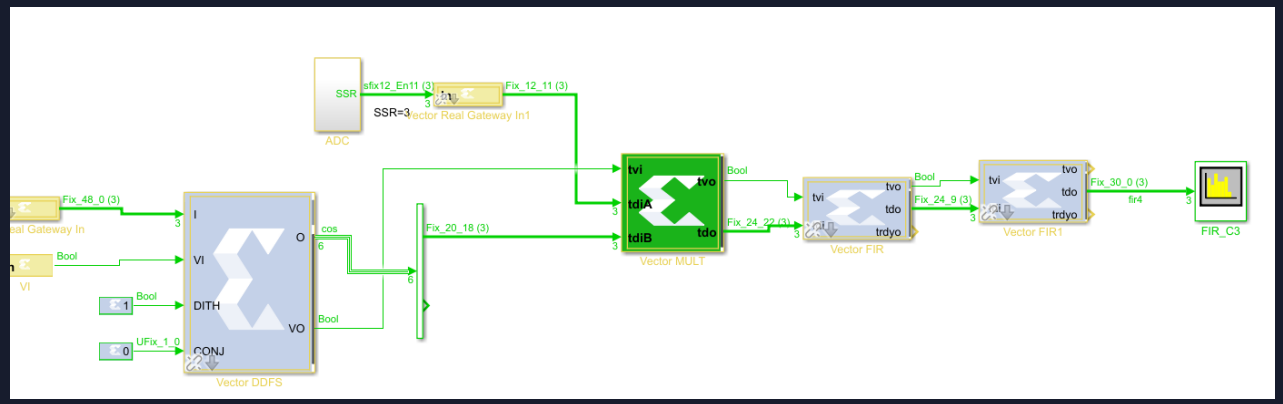
Input Serial Data Rate of 1.5 GHz



SSR = 3 Device Data Rate: 500MHz

After:

With SSR support of vectorized blocks

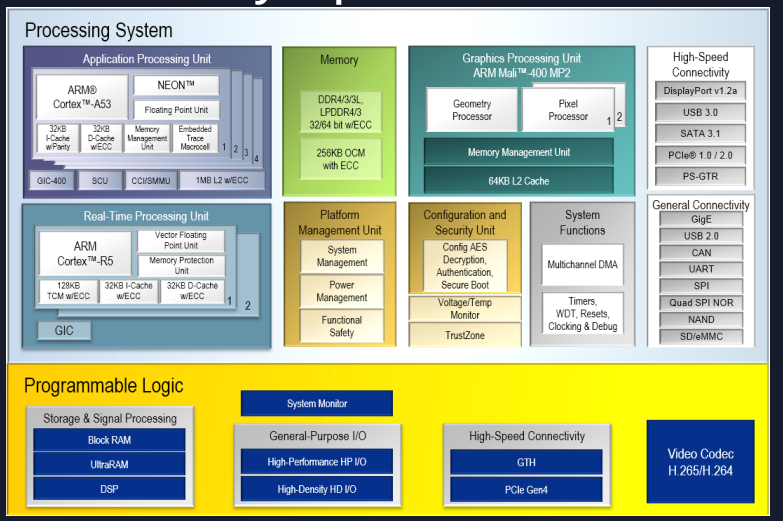


Zynq MPSoC for Embedded Vision

Power budget: 5W

Cost budget: $10-$40

Zynq MPSoC

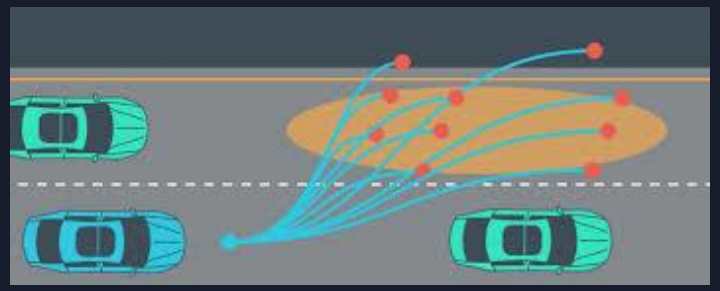


Driver Assistance Example:

Sensor input

Deep learning based perception

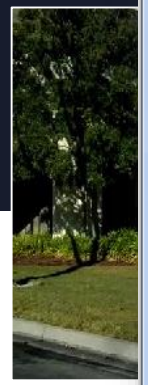
Path planning



Automatic braking

Latency < 30 ms

Moving Up in Abstraction: Model Composer Blockset with OpenCV



Simulink Library Browser

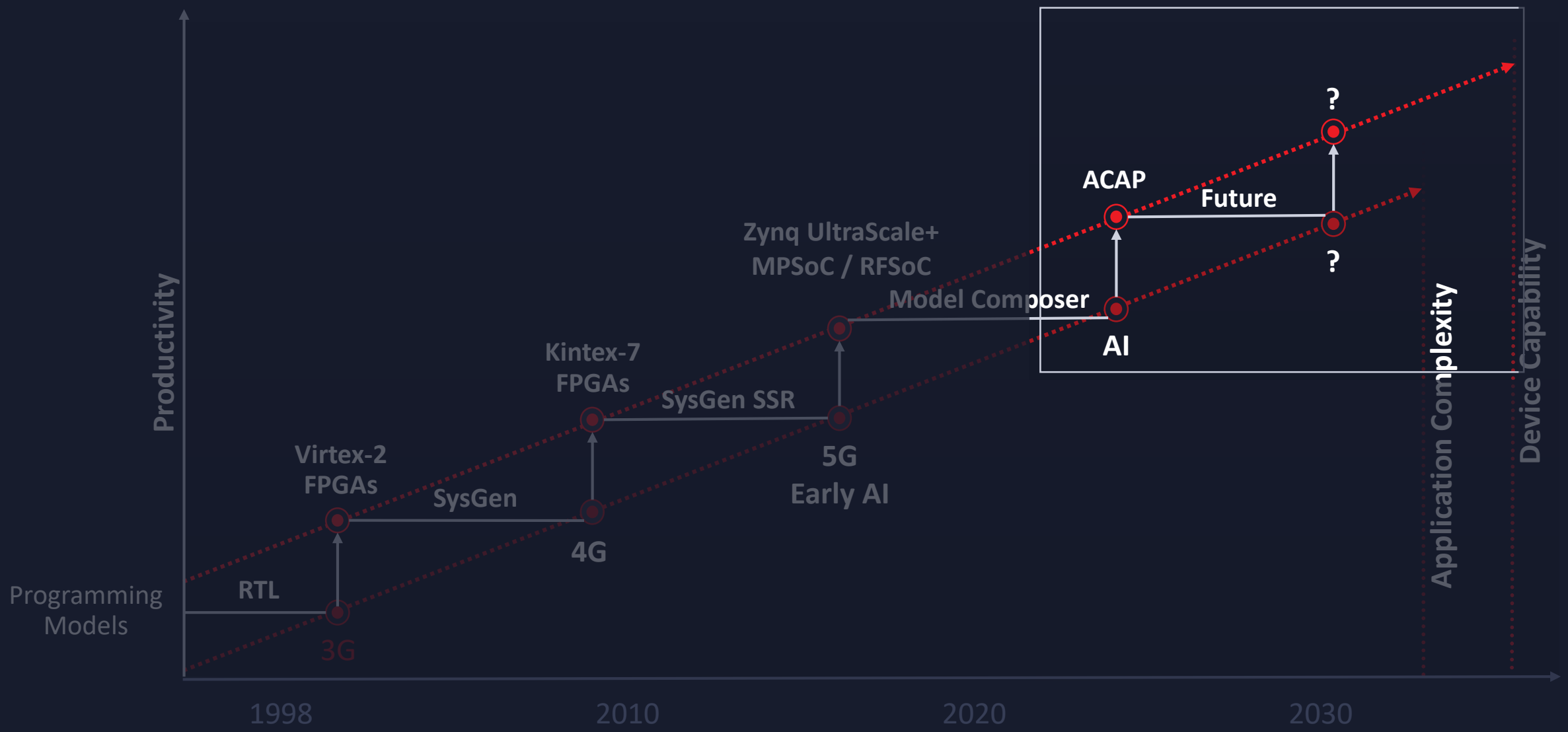
Enter search term

Xilinx Model Composer/Computer Vision

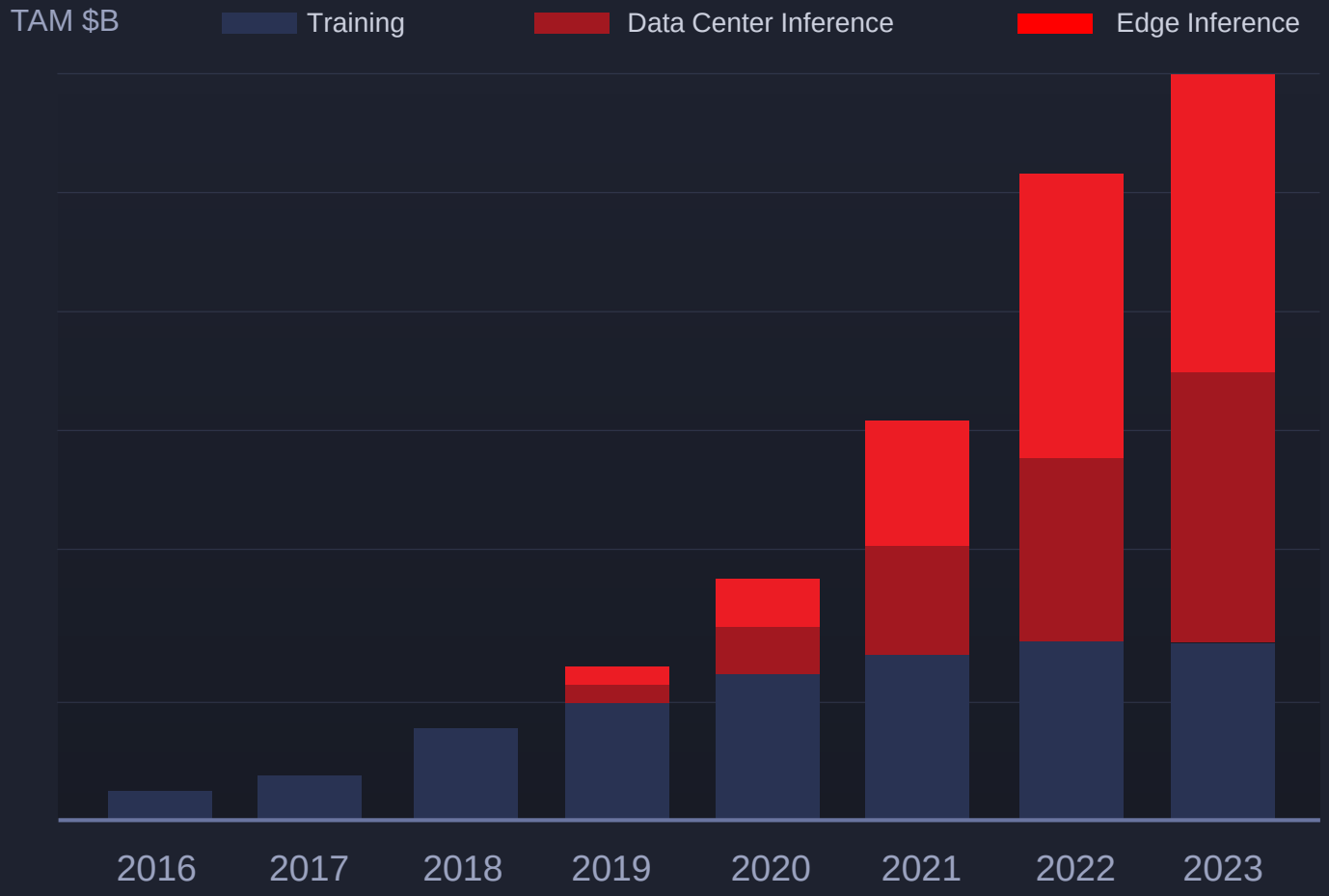
- Sinks
- Sources
- User-Defined Functions
- Additional Math & Discrete
- Communications System Toolbox
- Communications System Toolbox HDL
- Computer Vision System Toolbox
- Control System Toolbox
- DSP System Toolbox
- DSP System Toolbox HDL Support
- Embedded Coder
- HDL Coder
- HDL Verifier
- Image Acquisition Toolbox
- Phased Array System Toolbox
- Simulink 3D Animation
- Simulink Coder
- Simulink Extras
- Stateflow
- Xilinx Model Composer
  - Computer Vision
  - Logic and Bit Operations
  - Lookup Tables
  - Math Functions
  - Ports & Subsystems
  - Relational Operations
  - Signal Attributes
  - Signal Operations
  - Signal Routing
  - Sinks
  - Source
  - Tools
- Recently Used


The Road Ahead

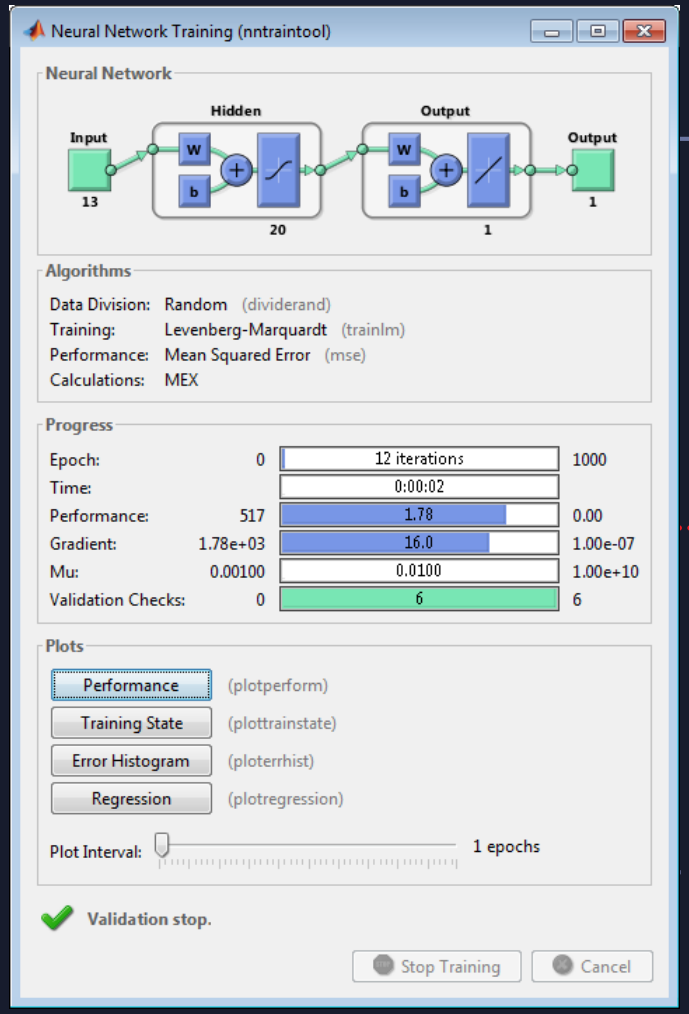
AI, AI, and More AI



AI/Machine Learning: Training vs. Inference



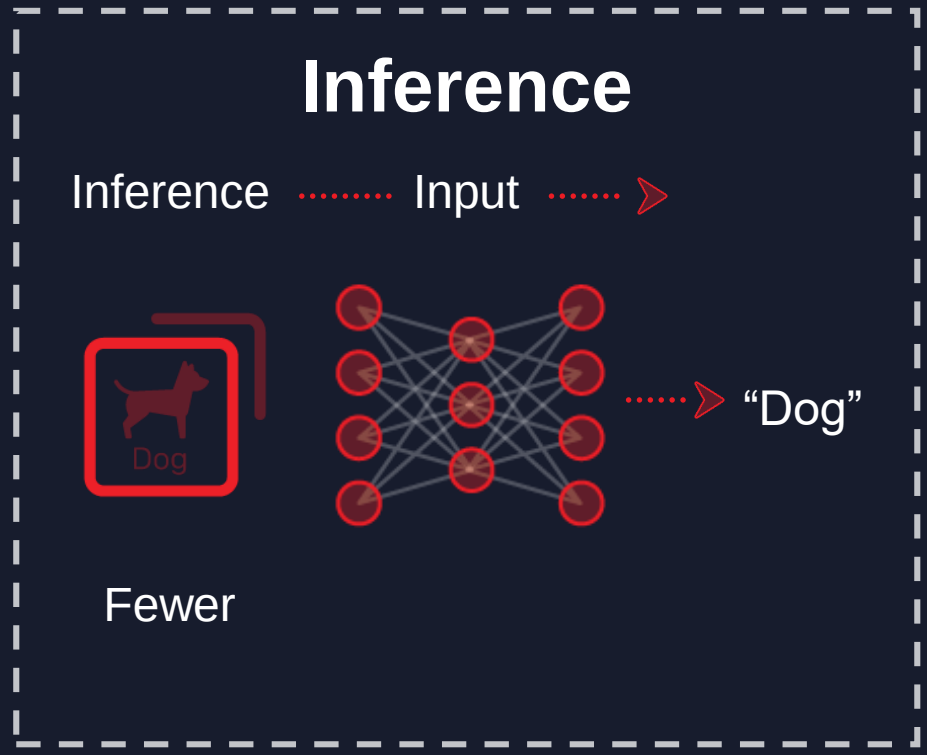
Barclays Research, Company Reports May 2018



"Dog"

Machine Learning Inference Challenges

- The rate of AI innovation
- Performance at low latency
- Low power consumption
- Whole app acceleration

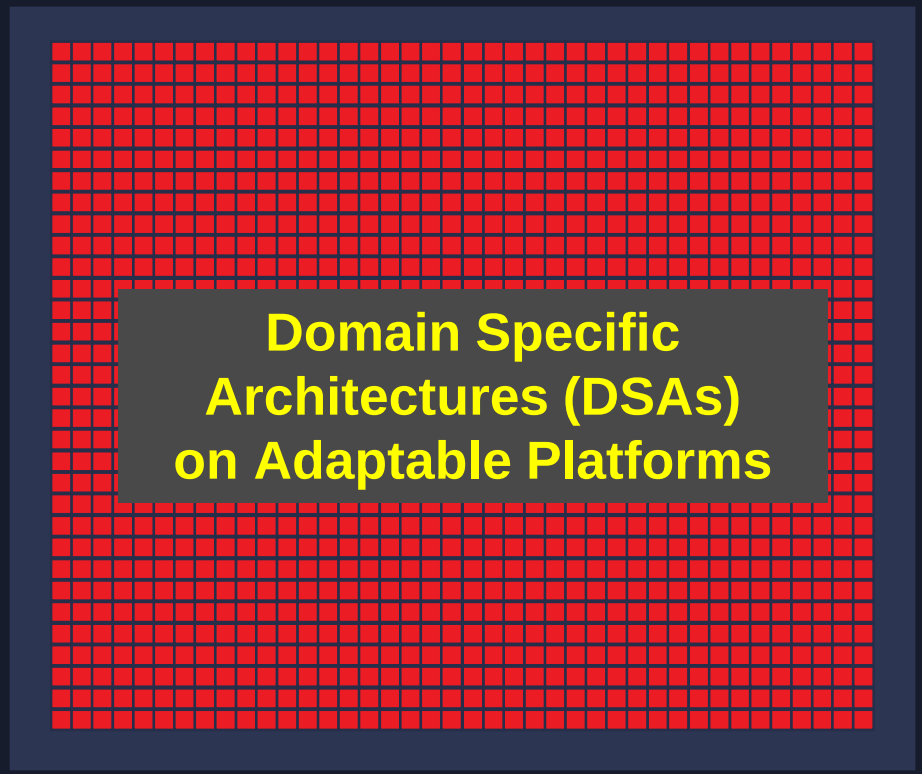


Only Adaptable Hardware Addresses Inference Challenges

Custom data flow

Custom memory hierarchy

Custom precision



Example Domain Specific Architecture: xDNN

Custom data flow

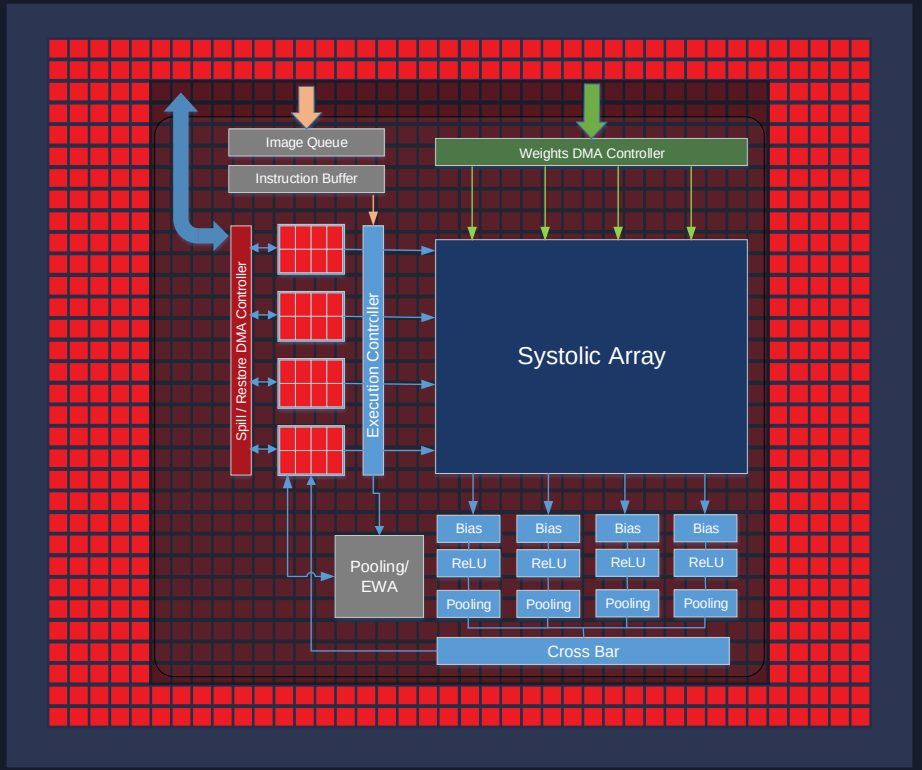
Optimized for latest CNN

Custom memory hierarchy

Optimized on-chip memory

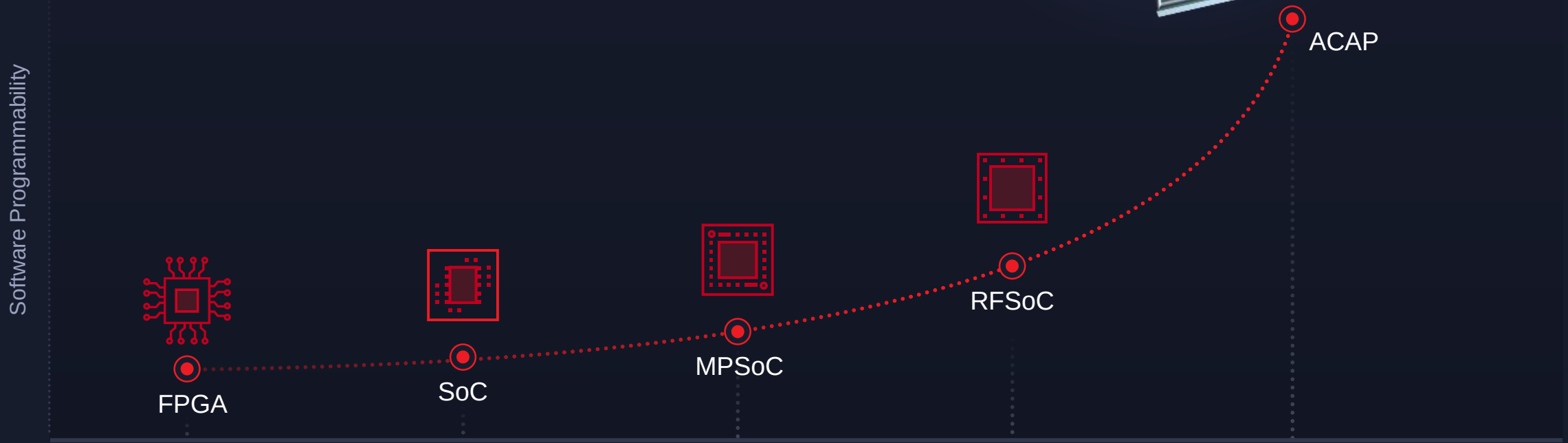
Custom precision

Int8

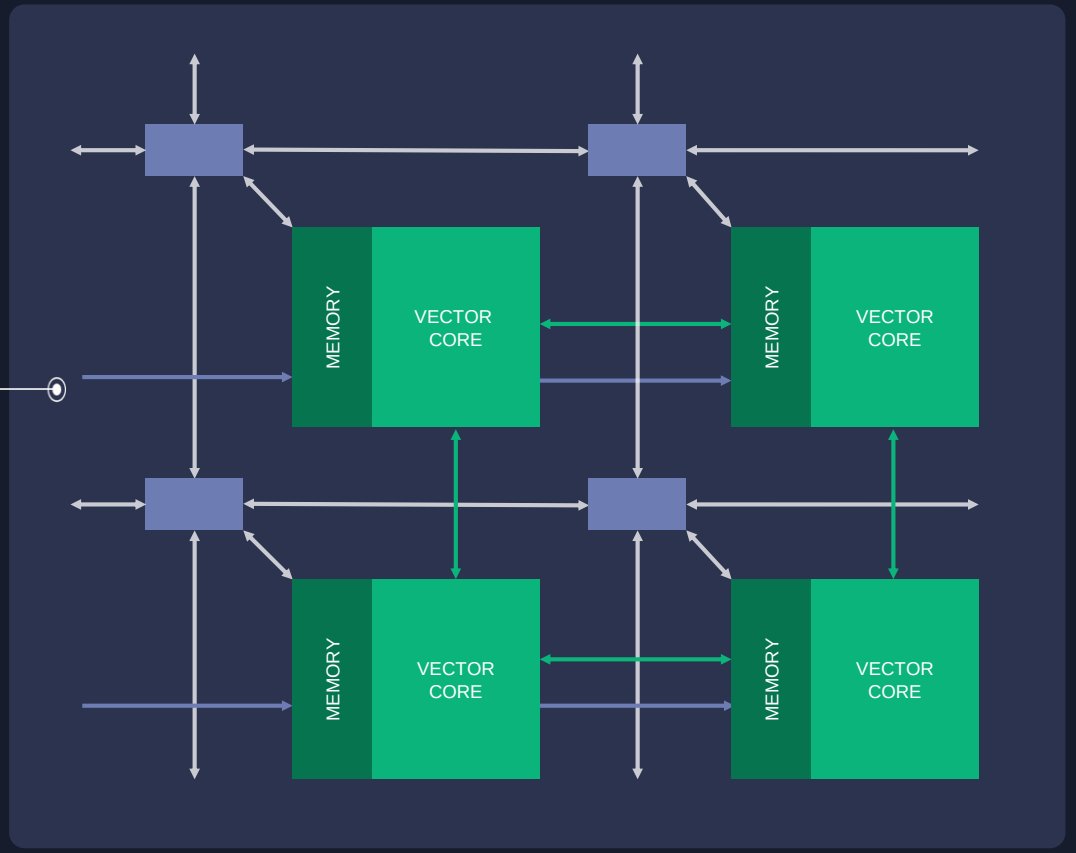


Adaptable Compute Acceleration Platform

A new class of devices for today and tomorrow's challenges



AI Engines



Optimized for AI Inference and
Advanced Signal Processing Workloads

- >1GHz VLIW/SIMD vector processor cores
- Massive array of interconnected cores with local memory
- Coupled to adaptable hardware enabling custom memory hierarchy
- Programmable with MATLAB/Simulink via Model Composer

Low-Latency CNN Inference Performance

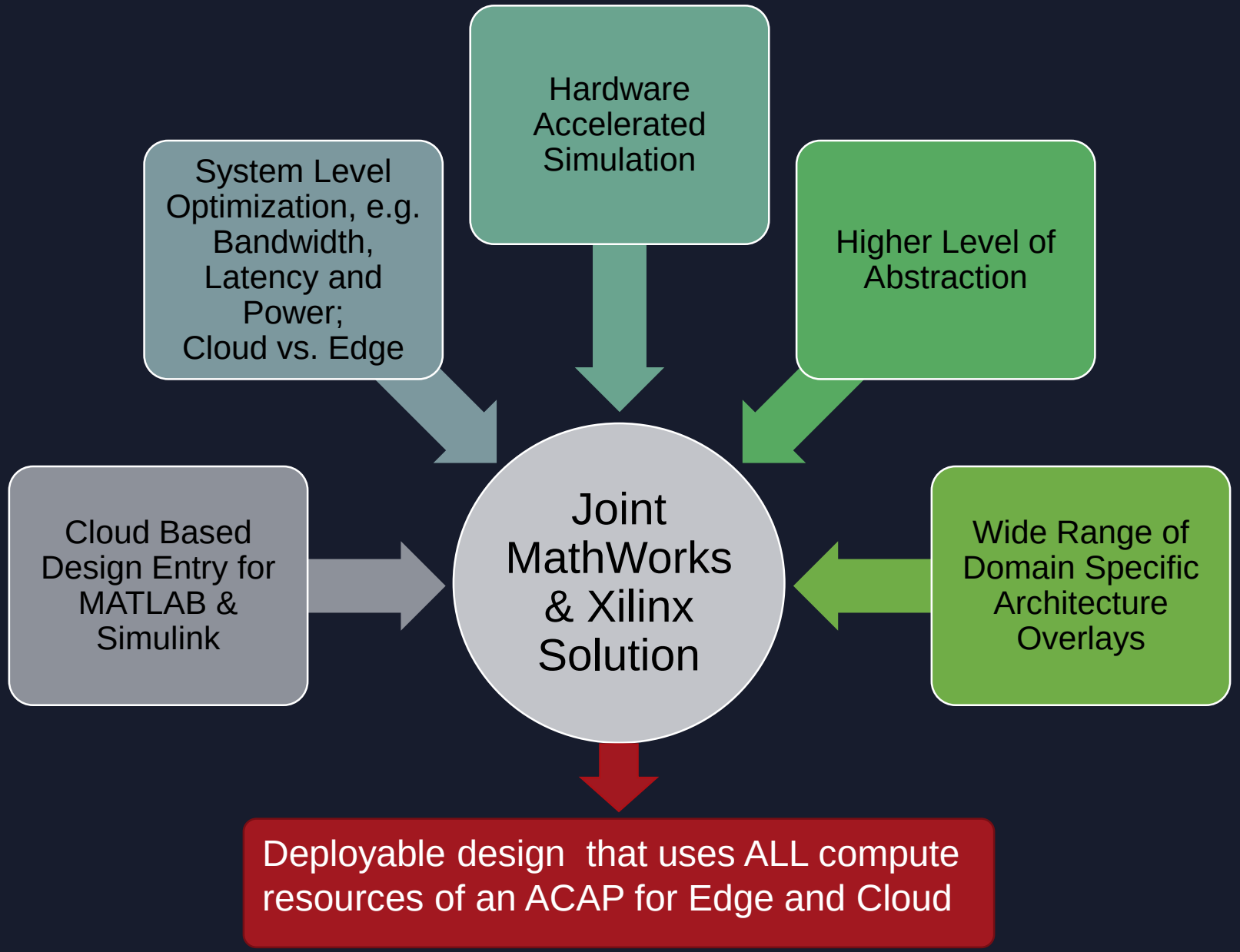
Pruning Technology

1.3x-8x

Performance improvement based on the network

Sources: Alveo - Published (INT8); Versal - Projected (INT8), 65% PL reserved for whole application; GPU 1 - P4 Published (INT8); GPU 2 - V100 Published (FP16/FP32); GPU 3 - T4 Projected

Future of Model-Based Design for ACAPs



Summary

Xilinx will continue to invest in Model-Based Design as a natural, productive on-ramp to our devices

Adaptable devices have a clear advantage in ML, ADAS and 5G to meet performance, latency and power requirements

The intersection of tools, silicon and platforms provides an inflection point for AI adoption