Big shifts in EDA, IP and HPC

Dr. Chris Rowen
CTO – IP Group
Cadence Design Systems, Inc.

June 7, 2015

cadence®

Agenda

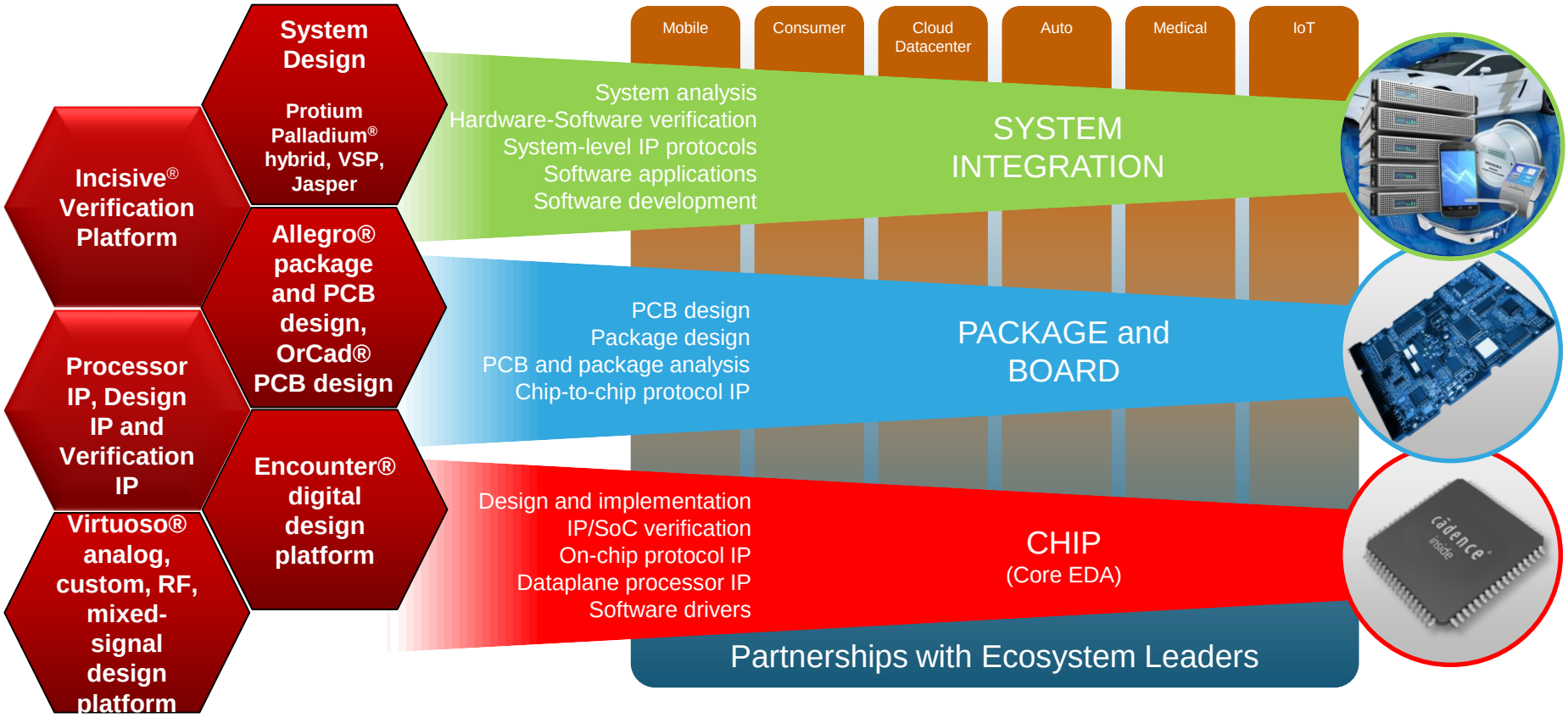
- A mainstream view of big compute and advanced EDA/IP
  - System Design Enablement
  - Design and Verification across digital and high-speed analog
  - Universal adoption of IP-rich development
  - New roles for processors
- But the world is changing…
- Example: vision super-computing

Big trends creating opportunities

cadence

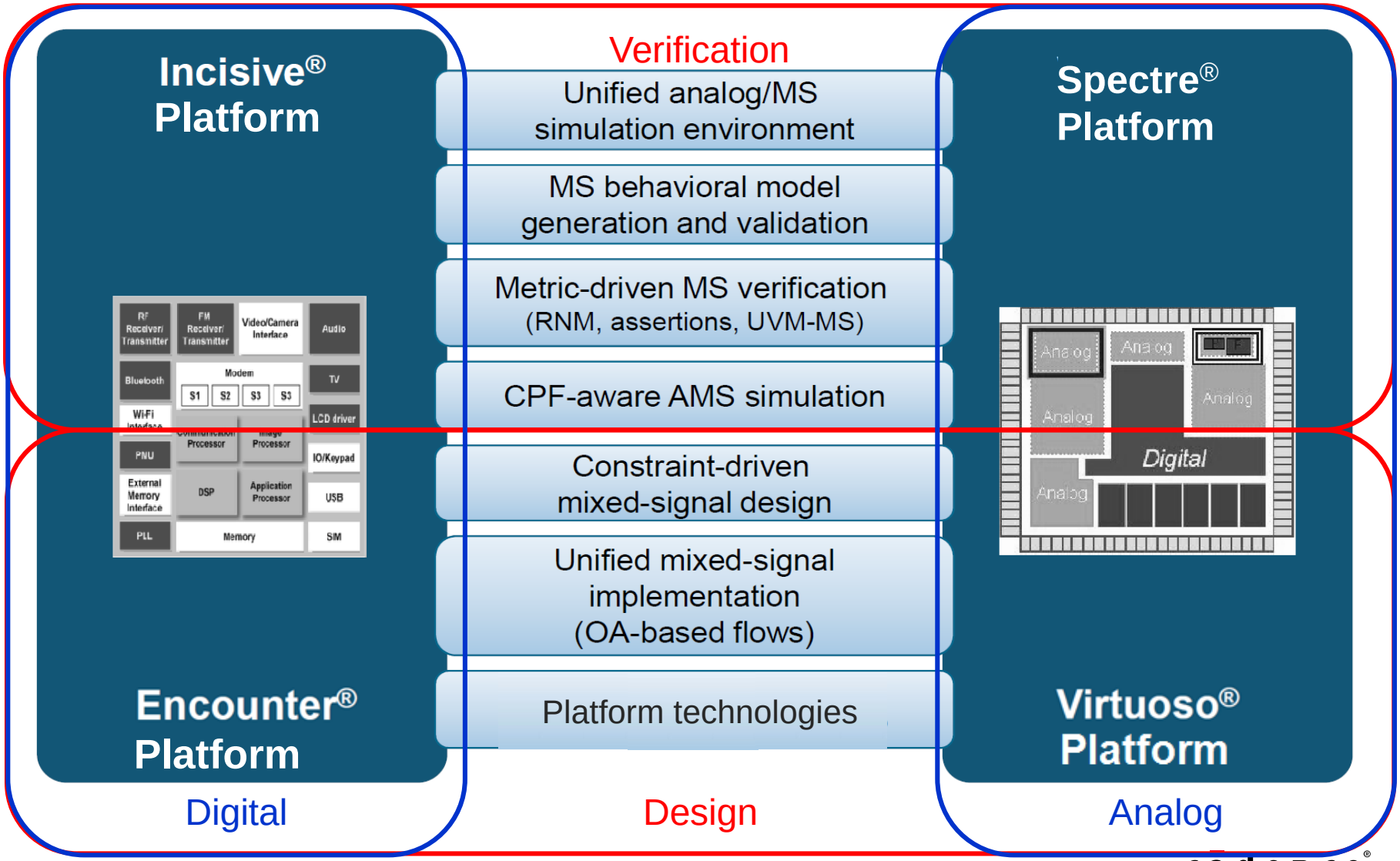
Cadence focus: System Design Enablement

From end product down to chip level



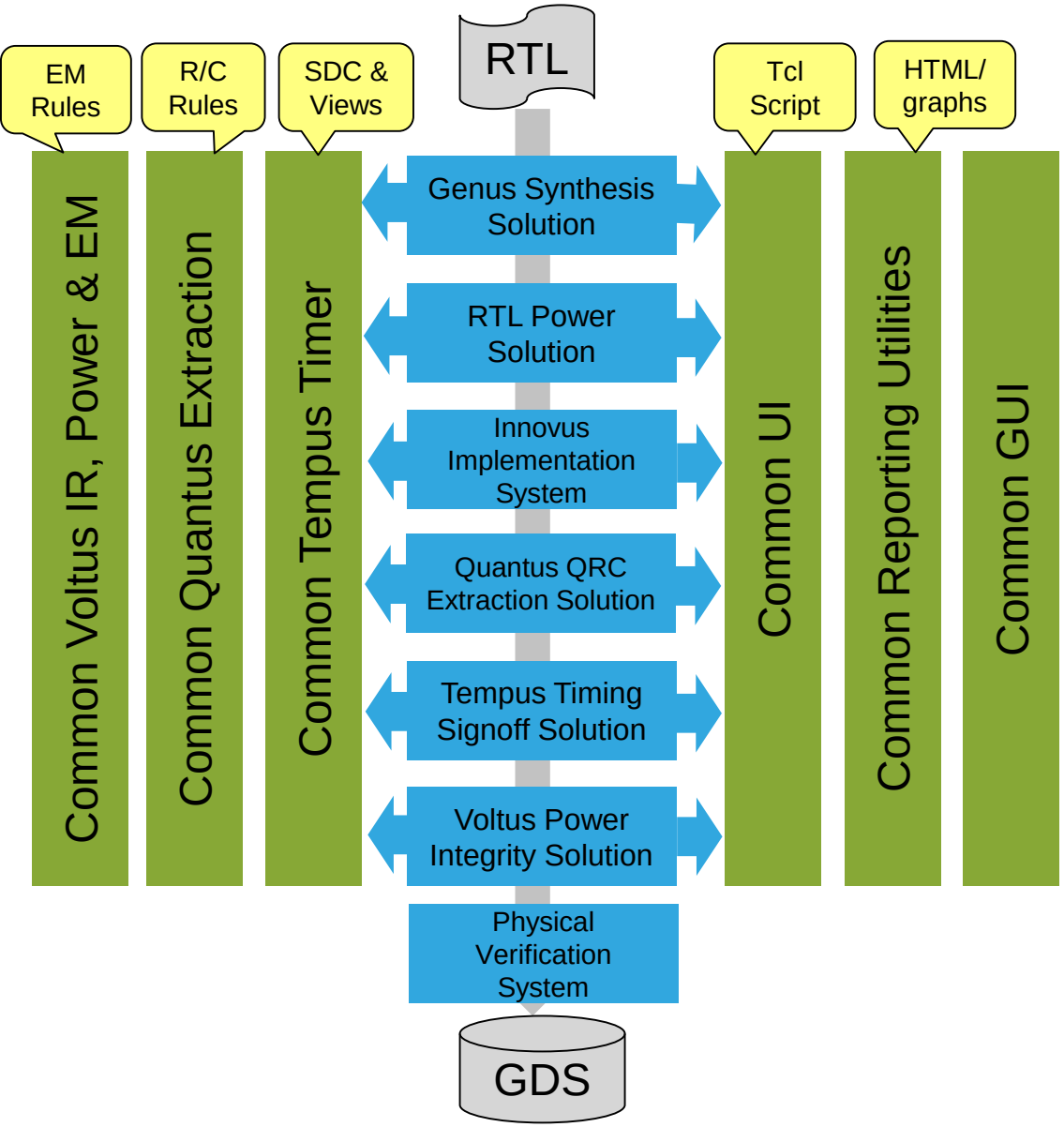
HPC: Compute + memory bandwidth + high-speed IO

Digitally-rich but high-speed analog also crucial



cadence®

New Integrated Digital Flow: Genus and Innovus

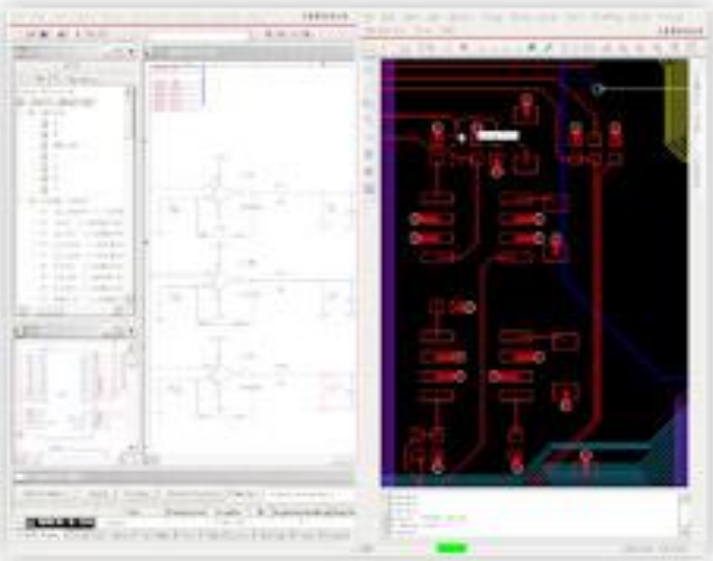


- Digital implementation tools share a common infrastructure
  - Analysis
    - Tempus Timer - Same SDC and View setup files for whole flow
    - Quantus Extractor - Same R/C rules for whole flow
    - Voltus IR, Power and EM - Same EM for whole flow
  - User Interface
    - Consistent tcl commands run whole flow
    - Easy to learn and reuse across different tools
    - Powerful reporting utilities
- 10-20% better PPA
- Up to 10x TAT & capacity

cadence

Integration is not just about the silicon design

- Allegro platform: system-in-package
  - Multi-radio gateway board & package design
  - Sigrity for gateway board & package signal integrity analysis
  - Mixed-signal design & verification across fabrics
  - OrCAD platform: system-on-board
- Virtuoso & Encounter platforms: system-on-chip
- Incisive & Virtual System platforms: system-on-chip HW/SW verification
- Palladium & Rapid Prototyping platforms: full system development



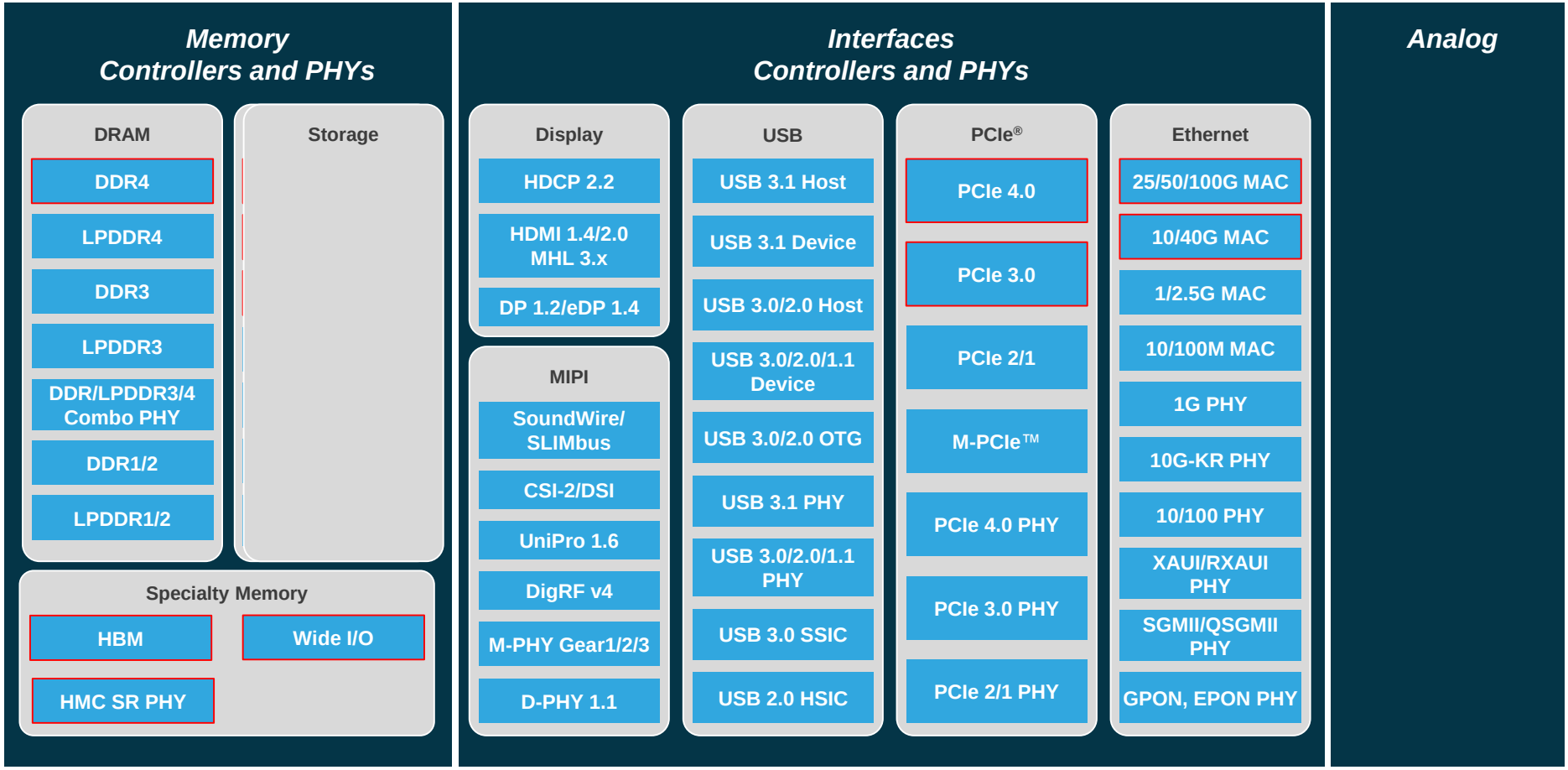
IP – A Path for Getting to Market Faster

Cadence is a trusted partner enabling customer innovation

Design IP	• Fastest Growing segment share • 1st to adv. nodes with DDR, PCIe® • Highest performance Analog IP
Tensilica® IP	• #1 in DSP IP licensing, #2 royalty bearing shipments • Shipping over 2 Billion cores/year • 225+ licensees world wide
Verification IP	• #1 in Verification IP segment share • 100% adoption by top 30 semiconductor companies • 500+ customers world wide

Cadence IP Portfolio

Broad and growing solution



Cadence VIP Catalog product line

Supports all major simulators and verification languages

Recent Addition

Simulation VIP

ARM AMBA 5 CHI	ARM AMBA 4 ACE	ARM AMBA AXI 3/4	ARM AMBA AHB	ARM AMBA 4 Stream
CAN	Display Port	Ethernet 10/100 1G/10G	Ethernet 25G/50G	Ethernet 40G/100G
HDMI 1.4	HDMI 2.0	I2C	JTAG cJTAG	LIN
MHL 3.0	MIPI CSI-2	MIPI CSI-3	MIPI C-PHY	MIPI DigRF
MIPI D-PHY	MIPI DSI incl. DBI, DPI	MIPI DSI2 incl. DBI, DPI	MIPI LLI 2.0	MIPI M-PHY
MIPI SLIMbus	MIPI Sound Wire	MIPI UniPro	NVM Express	OCP 2.2
OCP 3.0	PCI	PCIe© Gen2	PCIe Gen3	PCIe Gen4
PCIe SR-IOV	PCIe MR-IOV	M-PCIe	PLB 6	SAS 6G
SAS 12G	SATA 6G	SRIO 2.1	SRIO 3.0	UART
USB 2.0 w/ OTG	USB 3.0 w/ OTG	USB 3.1 w/ OTG	USB SSIC	Wireless 802.11 MAC

Memory Models

Cellular SRAM	Compact FLASH	DDR DIMM	DDR SDRAM	DDR Sync GFX RAM	DDR Sync RAM
DDR2	DDR3	DDR4 Incl. 3DS	DDR4 LRDIMM	DDR4 SDRAM	Delay line
DFI	Embed. SSRAM	Embed. SSRAM TI	eMMC 4.4	eMMC 4.5	eMMC 5.0
Enhanced SDRAM	FCRAM	FIFO	FLASH (basic)	FLASH ONFi	Flash ONFi 3/4
FLASH PPN DDR	FLASH Toggle NAND	FLASH Toggle NAND 2	GDDR2	GDDR3	GDDR4
HBM	HMC	LBA NAND	LL DRAM	LPDDR	LPDDR2
LPDDR3	LPDDR4	LR DIMM	Memory Stick	Memory Stick Pro	NAND FLASH
NOR FLASH Spansion	One NAND FLASH	PROM	Pseudo Burst SRAM	QDR SRAM	Rambus DRAM
Rambus Turbo Mode	Register File	RL DRAM	Scratch pad	SD Card	SD Card 3.0
SD Card 4.0	SDIO	Synch DRAM	Synch Mask ROM	Synch RAM NEC	UFS 1.0
UFS 2.0	Wide I/O	Wide I/O 2			

Accelerated VIP

ARM AMBA 5 CHI*	ARM AMBA 4 ACE	ARM AMBA AXI 3/4	ARM AMBA AHB
ARM AMBA APB	Ethernet 10/100 1G/10G	Ethernet 25G/50G*	Ethernet 40G/100G
HDMI 1.4	HDMI 2.0*	I2C	I2S
Keypad	MIPI CSI-2	MIPI DBI	MIPI DSI
MIPI UniPro*	NVM Express*	PCIe Gen2/3	PCIe SR-IOV*
SATA 3G/6G Device	SIM Card	USB 2.0 w/ OTG*	USB 3.0 Host*

*beta

Productivity Tools

PureView	Indago Protocol Debug App	
Interconnect Validator Basic	Interconnect Validator Coherent	Interconnect Workbench
TripleCheck Ethernet 40G/100G	TripleCheck MIPI UniPro	TripleCheck PCI Express

Assertion-Based VIP

ARM AMBA ACE	ARM AMBA AXI	ARM AMBA AHB	DFI
OCP			

cadence

Tensilica IP is #2 overall in shipments of royalty-bearing, licensable processors (CPU or DSP)

Used across most electronics markets

IoT, Mobile Phones, Storage/SSD, Networking, Video, Security, Cameras, Watches, Printers, others...

Control

Control
Embedded

- Advanced compiler and development tools support
- Energy & area efficient

Wide range of DSPs




HiFi
Audio / Voice / Speech

- Encode + Decode
- Voice Trigger
- Noise Reduction
- Post-Processing

ConnX
Communications

- Narrow to wide band wireless
- LTE/LTE-A, WiFi, SmartGrid
- Infrastructure & Terminals

Vision
Vision / Imaging

- Image Processing & Analytics
- Video Pre-Post Processing

Fusion
General Purpose

- Scalable DSP
- Always-alert
- Sensor processing
- Audio / Voice / Speech
- Comm's

Custom

Custom ISA
Application-specific

- Low Energy
- High Performance
- Application datatypes

Thousands of designs

Xtensa® Platform for Innovation

Common Architecture, Development Tools, 3rd Party Ecosystem

>2B cores per year

cadence

Tensilica automated hardware/software synthesis

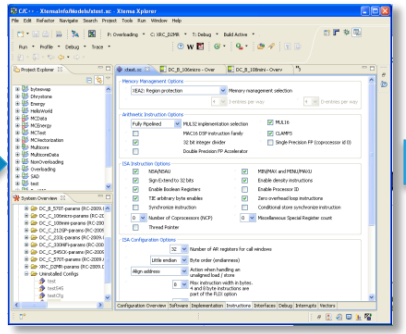
Architects' needs:

- Higher compute throughput
- More memory bandwidth
- Higher efficiency control structure update
- Continuous software upgrade of EVERYTHING

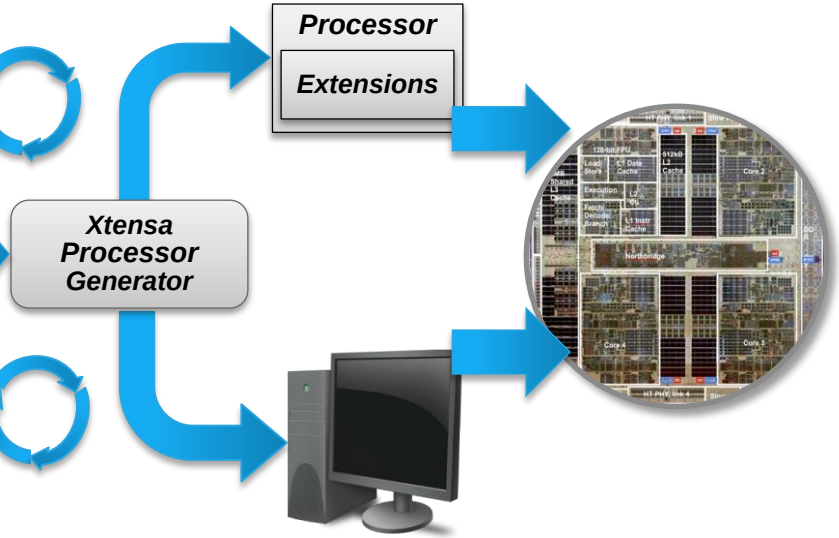
Full hardware design

Source pre-verified RTL, EDA scripts, test suite

Processor configuration



1. Select from menu
2. Explicit instruction description (TIE)



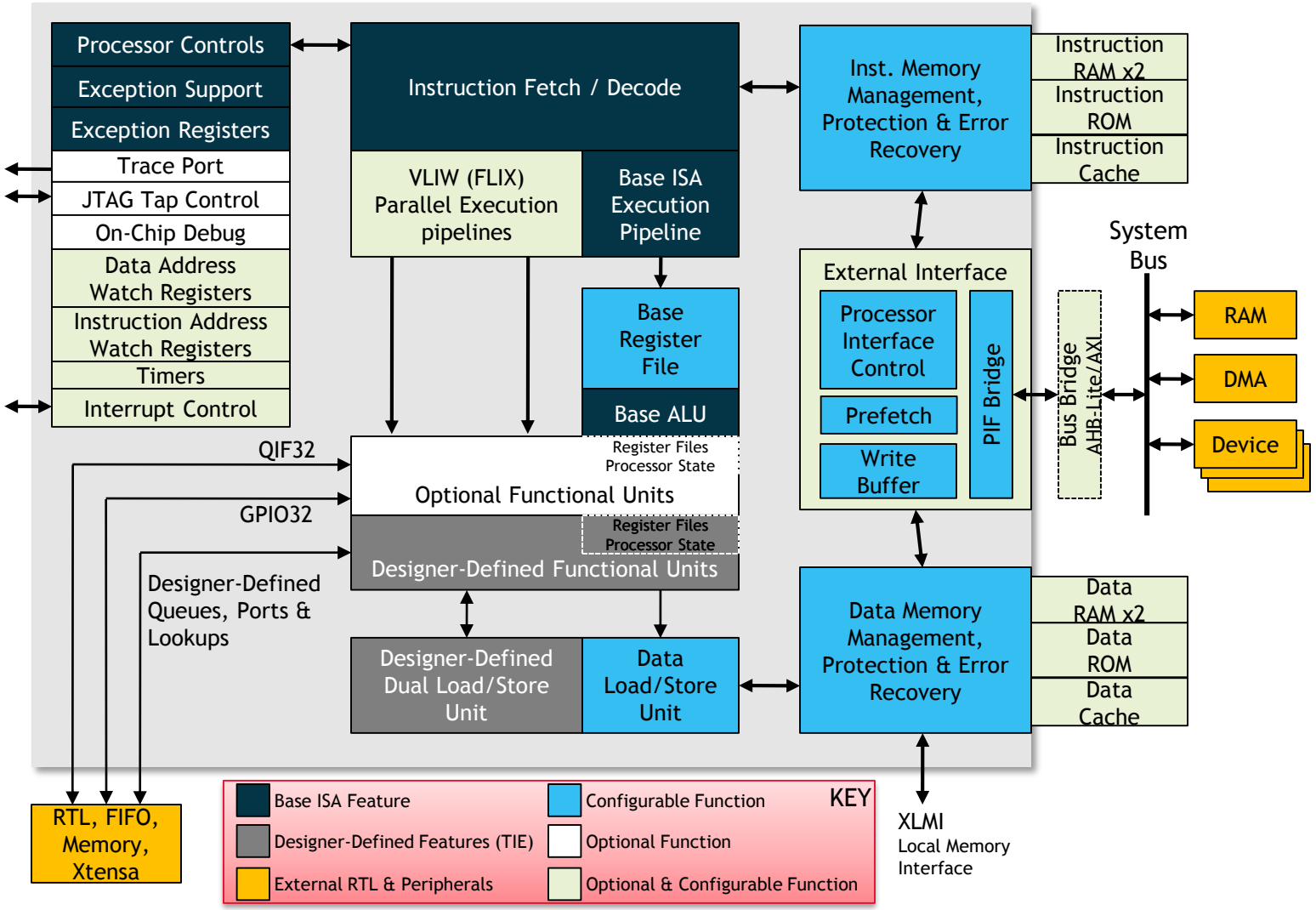
Customized software tools

C/C++ compiler Debuggers, Simulators, RTOSes

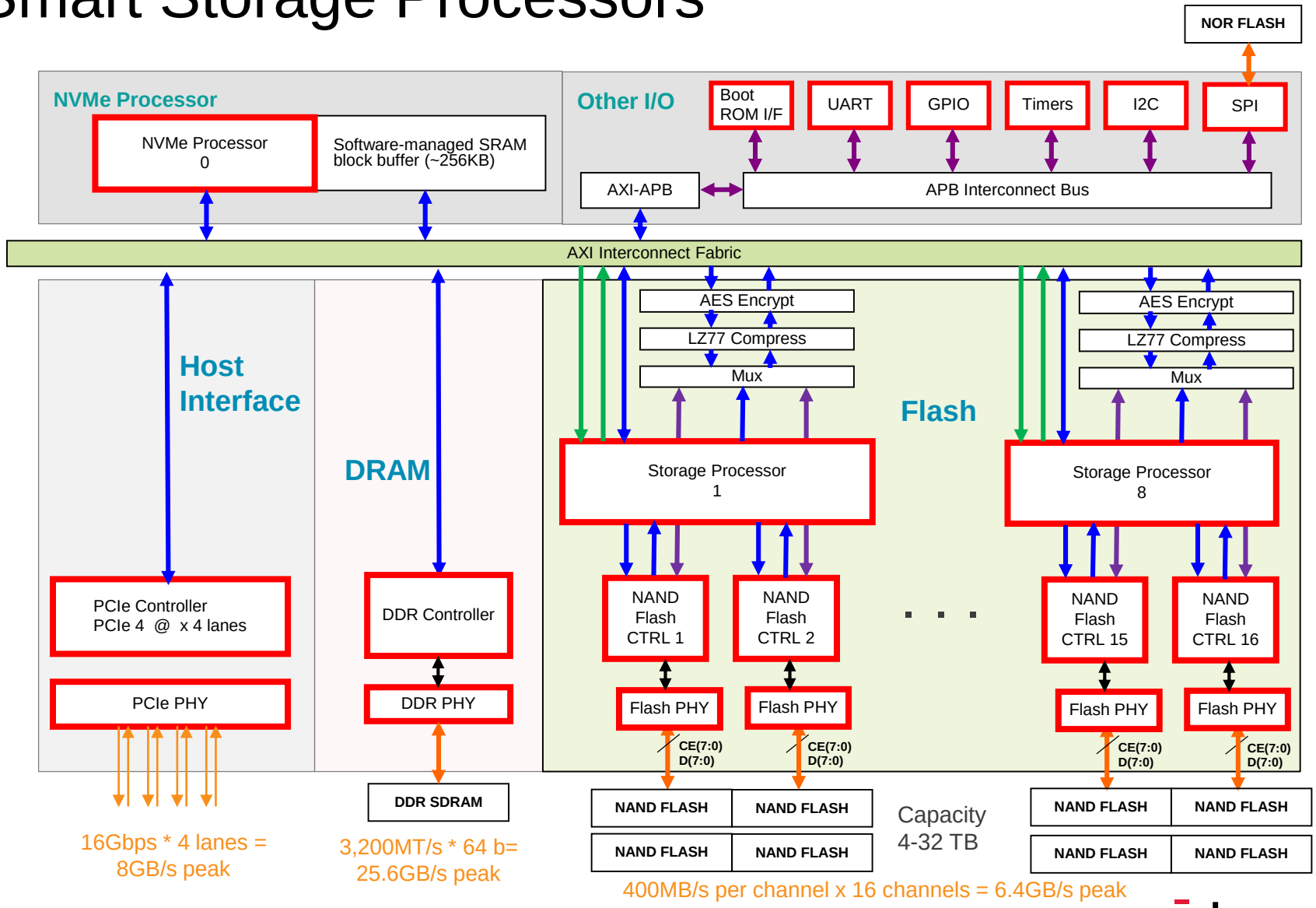
No manual intervention needed in hardware or software completeness

cadence

Xtensa LX6
Block Diagram - System



Example: Flash-Based Memory Systems with Smart Storage Processors



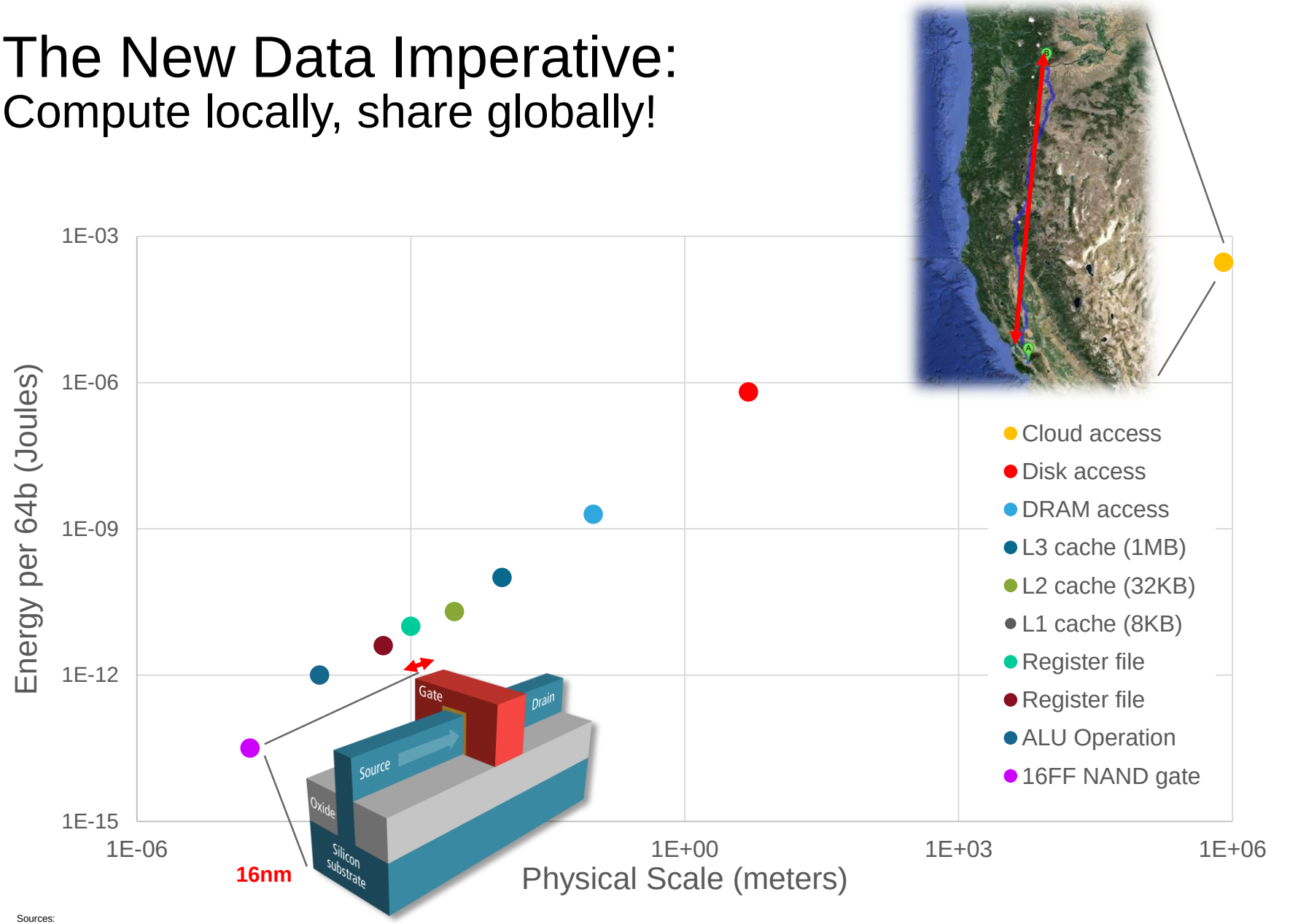
Agenda

- A mainstream view of big compute and advanced EDA/IP

- But the world is changing…
  - The new data imperative
  - Energy and layered cognition
  - Scaling to ultra-low power

- Example: vision super-computing

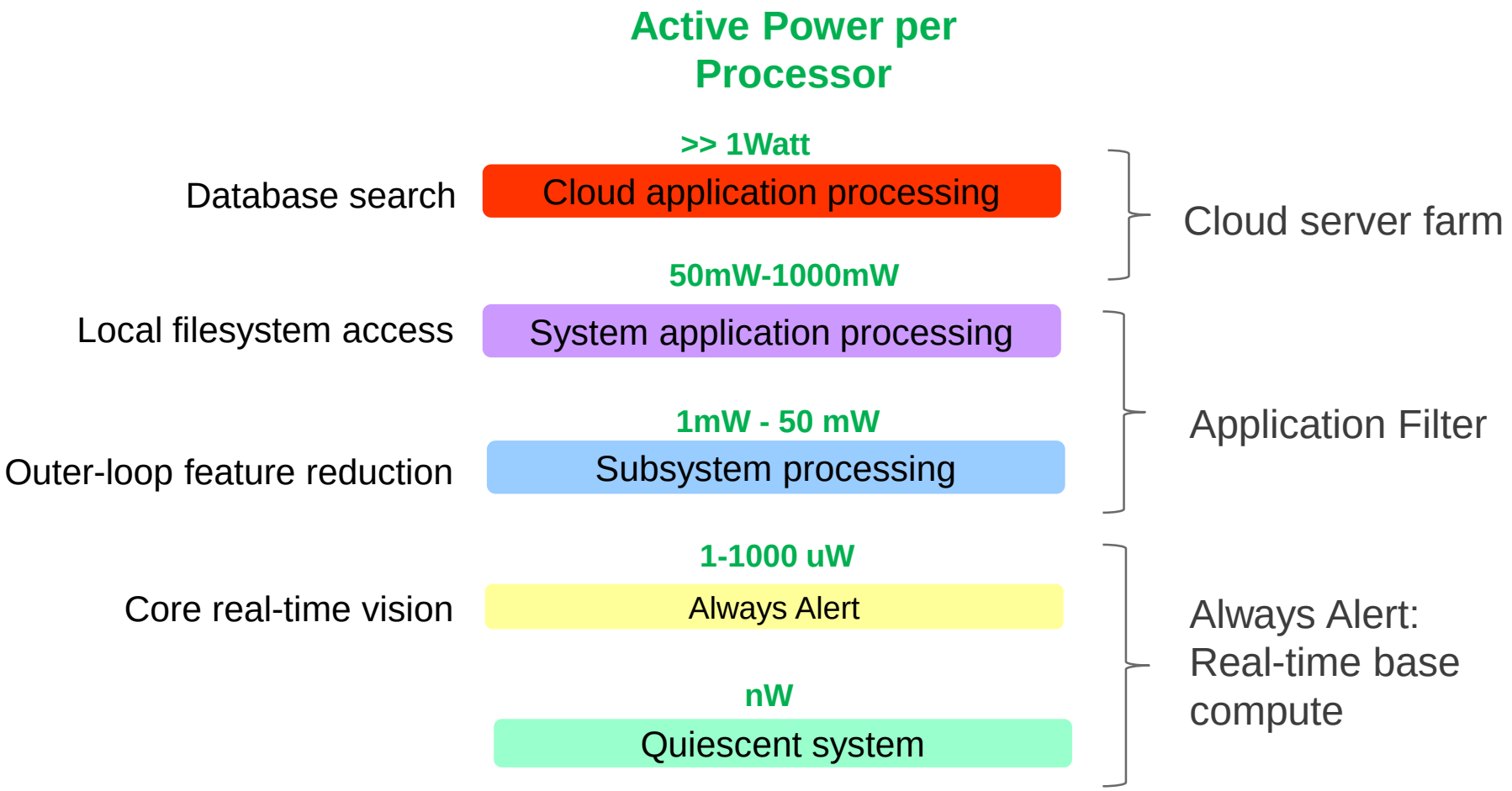
The New Data Imperative:
Compute locally, share globally!



Sources:
- Rick Zarr, TI, 2008- The True Cost of an Internet "Click" - estimate of transfer cost for 30KB page from server: http://energyzarr.typepad.com/energyzarrnationalcom/2008/08/the-true-cost-o.html
- J Kunkel et al, University of Hamburg 2010,Collecting Energy Consumption of Scientific Data
- Horowitz ISSCC 2014, 1300-2600 pJ per 64b access

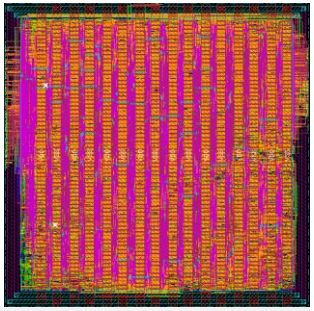
cadence

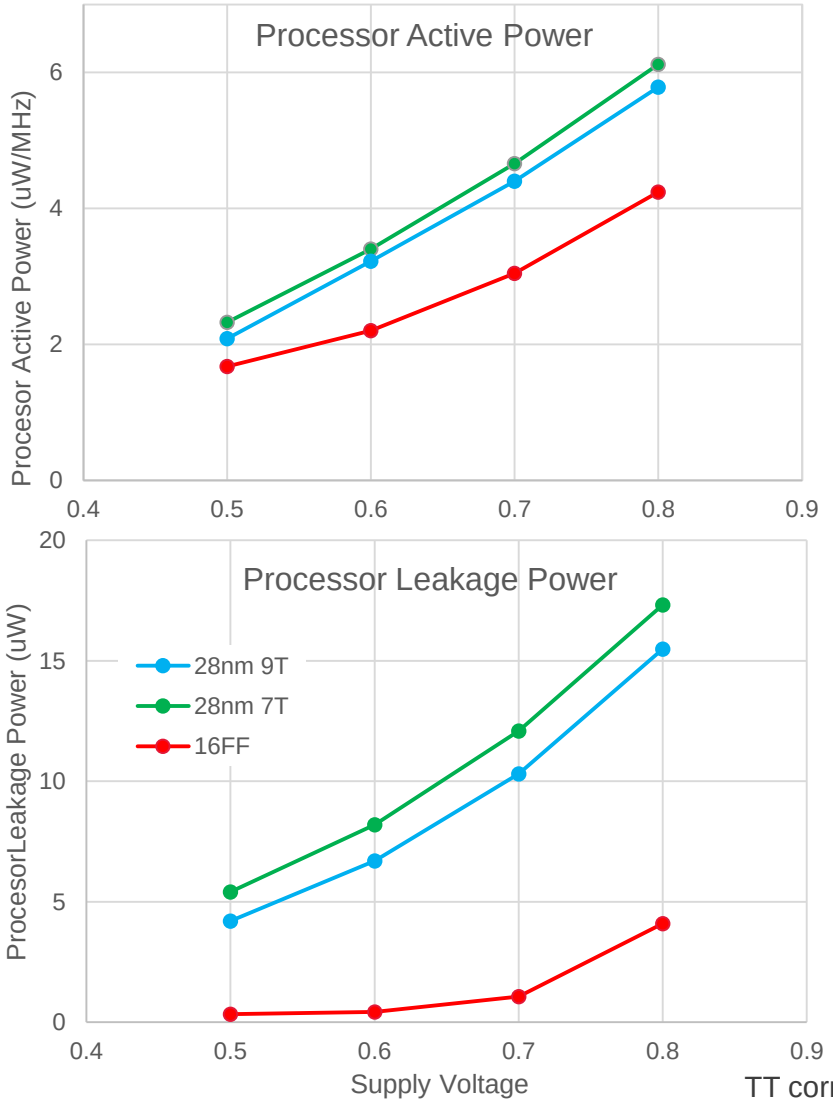
Energy and layered cognition



Scaling to ultra-low-energy

Xtensa core scales from ~1uW/MHz to 1.5GHz





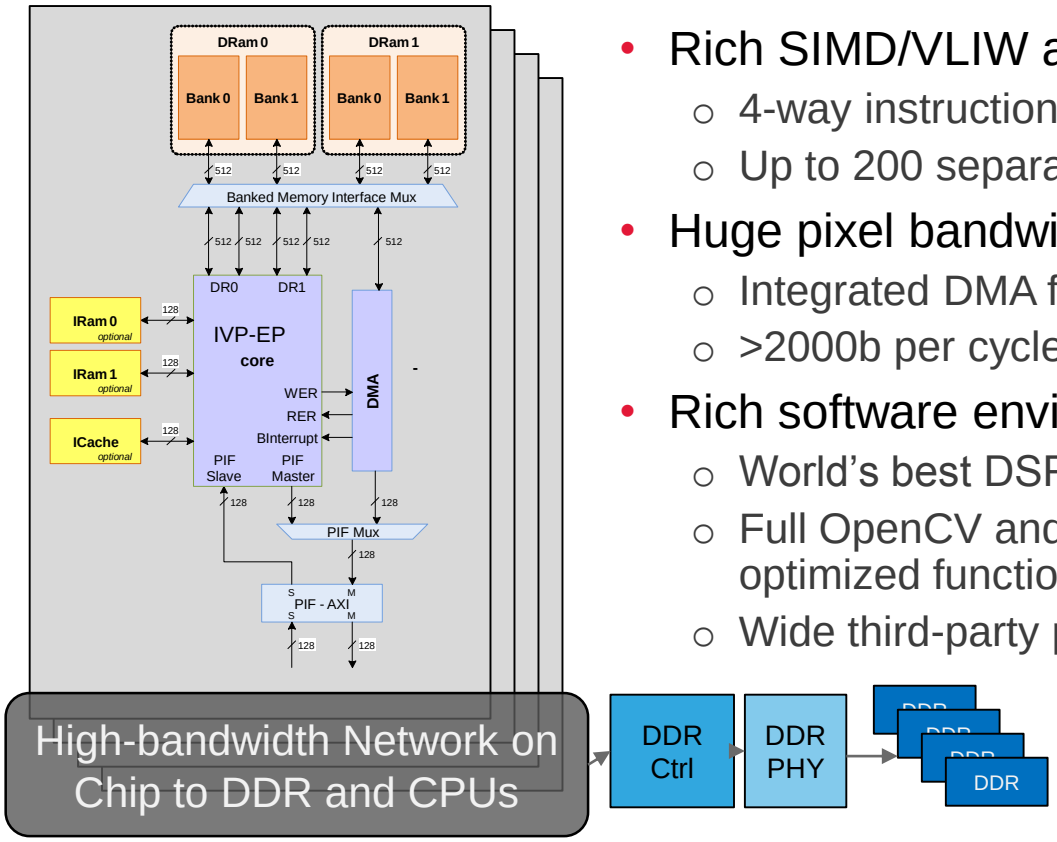
Near-threshold core development

- *SPICE (SPECTRE APS) simulation of full Tensilica base 32b processor core*
- *Verifies correctness, speed and power independent of library characterization.*
- *Full parasitics extracted circuit gives accurate power and functionality*
- *Running basic power and functionality diagnostics*
- *Optimally-pipelined core runs >1.5GHz in16FF, but same RTL also achieves micro-power level. Sustains high MHz at lower voltage*

Agenda

- A mainstream view of big compute and advanced EDA/IP
- But the world is changing…
- Example: vision super-computing
  - Vision DSPs
  - Convolutional Neural Network compute
  - Hybrid architectures for high-density computation

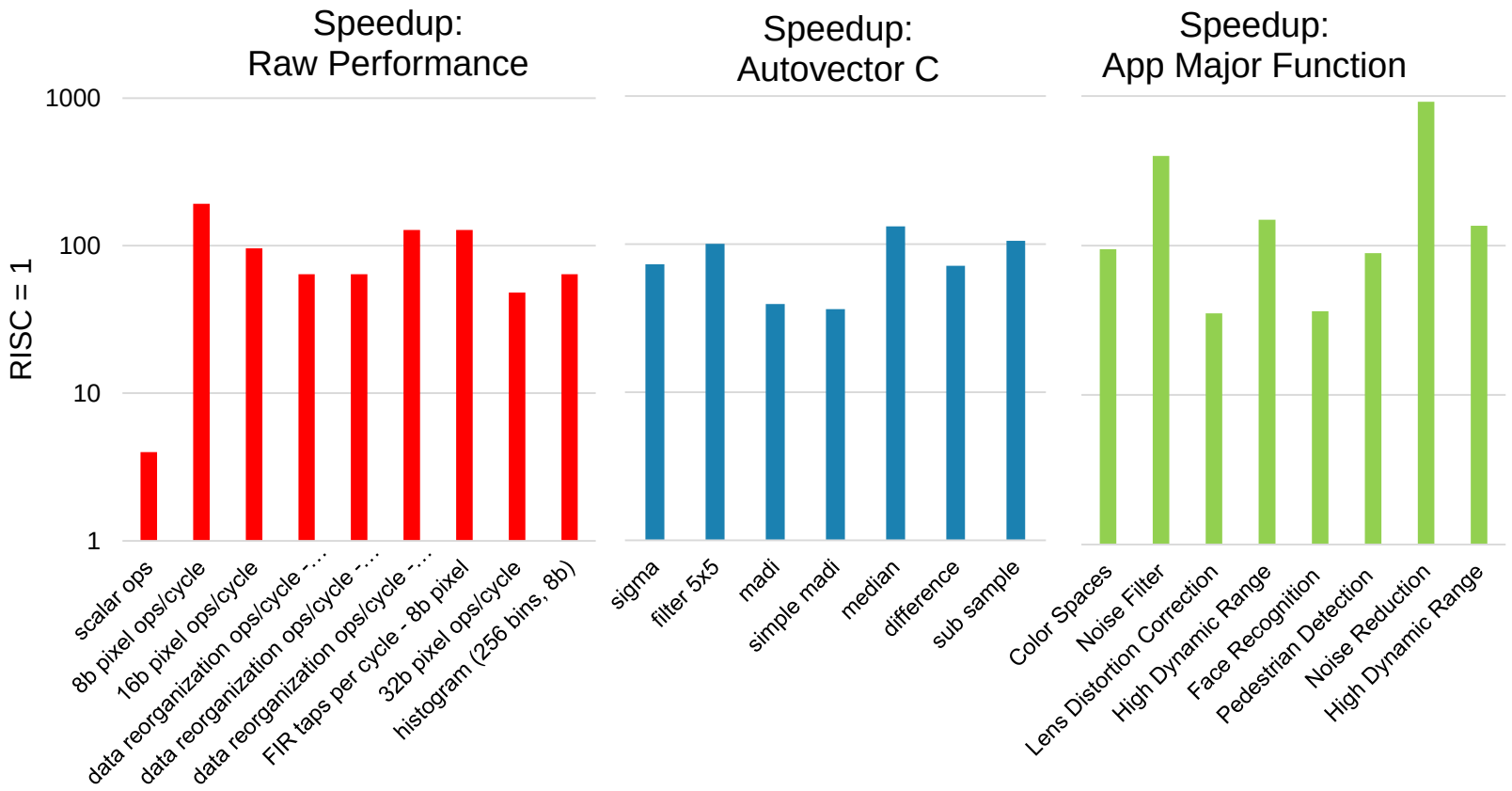
Tensilica IVP Vision DSPs



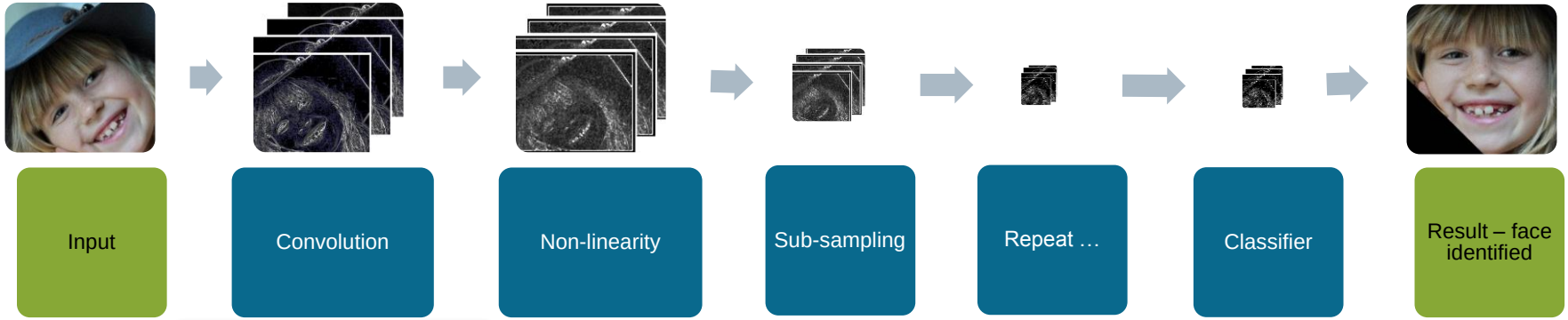
- A family of high-performance DSPs for imaging, video and vision
- Rich SIMD/VLIW architecture
  - 4-way instruction issue
  - Up to 200 separate ALU operations per cycle
- Huge pixel bandwidth
  - Integrated DMA for data streaming
  - >2000b per cycle data memory bandwidth
- Rich software environment
  - World's best DSP C compilers: 0 assembly code
  - Full OpenCV and OpenVX support with 800 optimized functions
  - Wide third-party program and support

cadence

IVP-EP
A new category of application performance



Example: Convolutional Neural Network (CNN)
Sorting through the hype



Convolutions model locally receptive visual cortex cells by sampling a small region and generating features

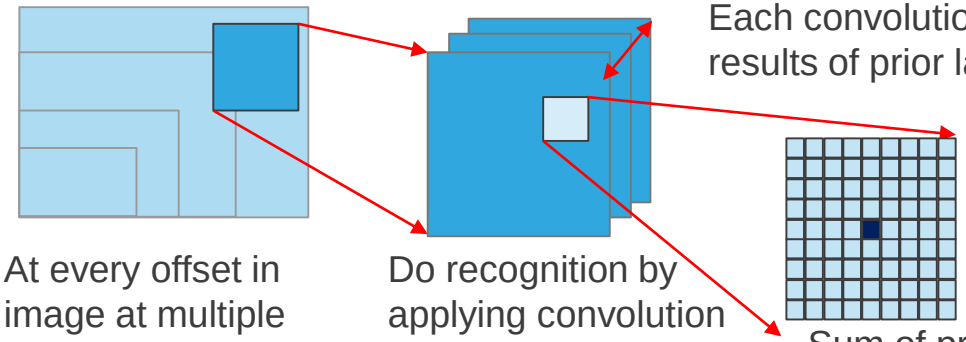
Non-linearity like tanh function models on-off behavior of neurons

Subsampling models cells with larger receptive fields (provide local invariance)

Repeat previous steps for neural network layers

Final classifier stage

one layer



Each convolution spans M results of prior layer

At every offset in image at multiple scales: K

Do recognition by applying convolution at every offset in patch: L

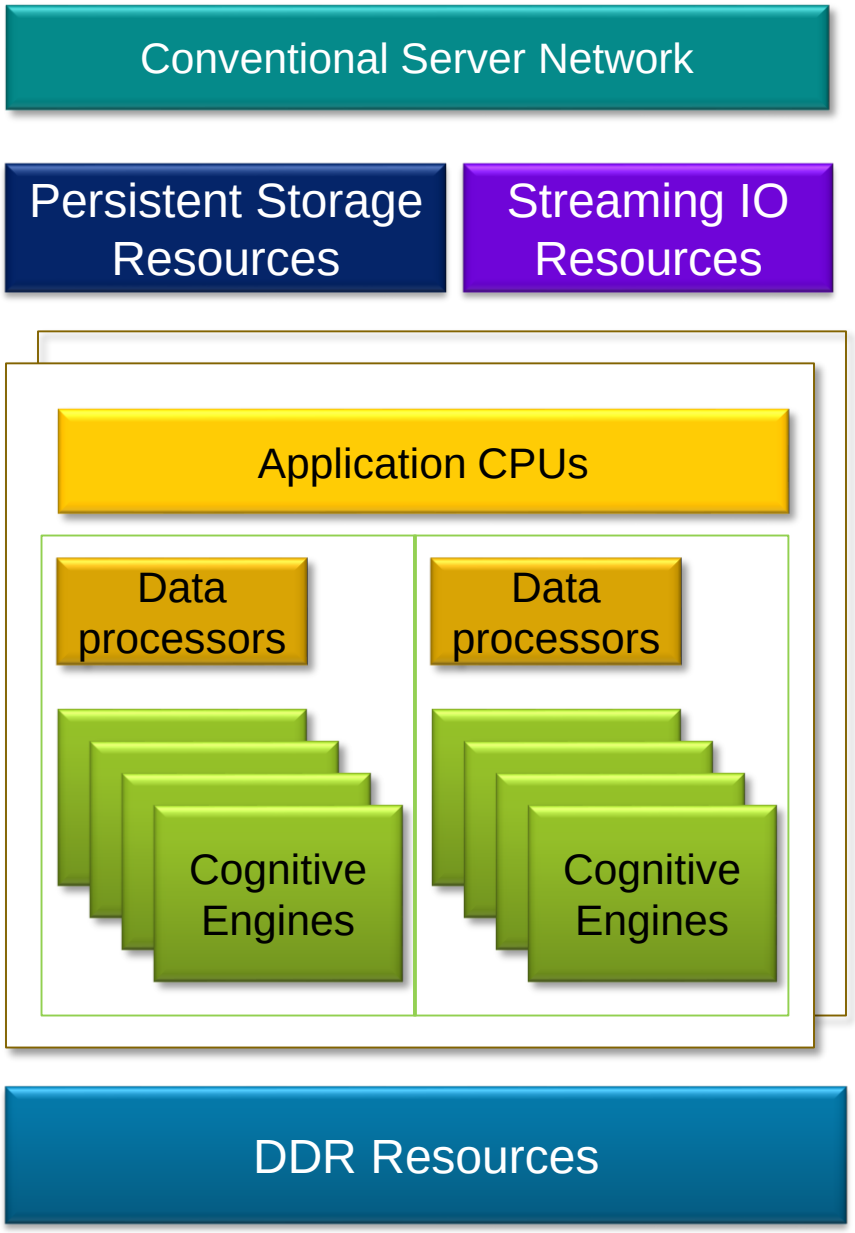
Sum of products of pixels x weights in NxN sub-patch

Compute for layer:
O (K * L * M * N * N)
Up to 30 x 10^{12} MACs

Kevin Jarrett , Koray Kavukcuoglu , Yann Lecun
What is the Best Multi-Stage Architecture for Object Recognition

Vision supercomputing

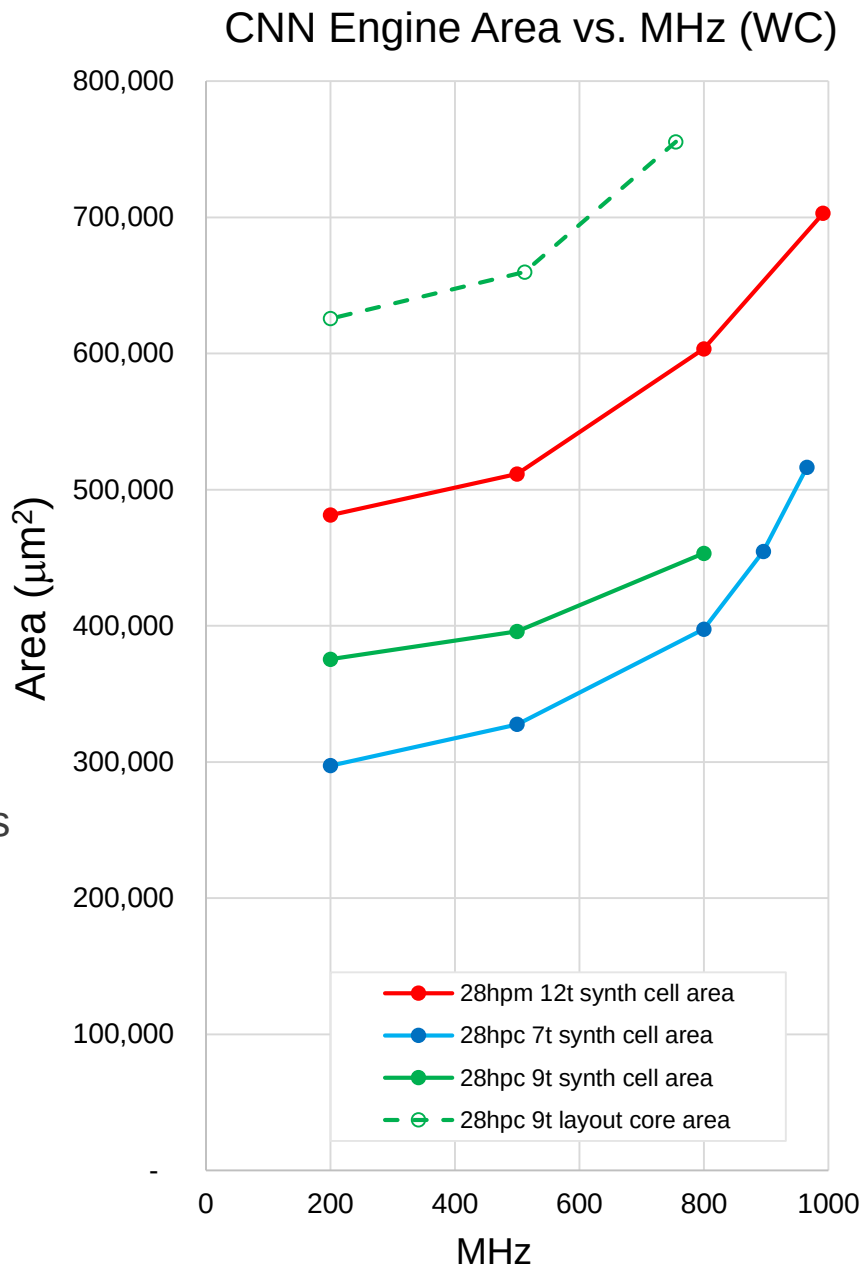
- 80-90% of internet traffic is video/image
- Expect image recognition, video analytics, augmented reality, visual AI to grow to similar % of compute
- By 2020, world's image sensors will theoretically capture 10^{28}B/year
- Key characteristics of visual HPC:
  1. Applications in flux
  2. High-bandwidth data * very large parameter sets – e.g. for CNN weights
  3. Need extreme compute efficiency, yet high flexibility in algorithms, data reference patterns and data types.
- Expect "cognitive layering" in computing architectures – 95% of ops in 5% of code



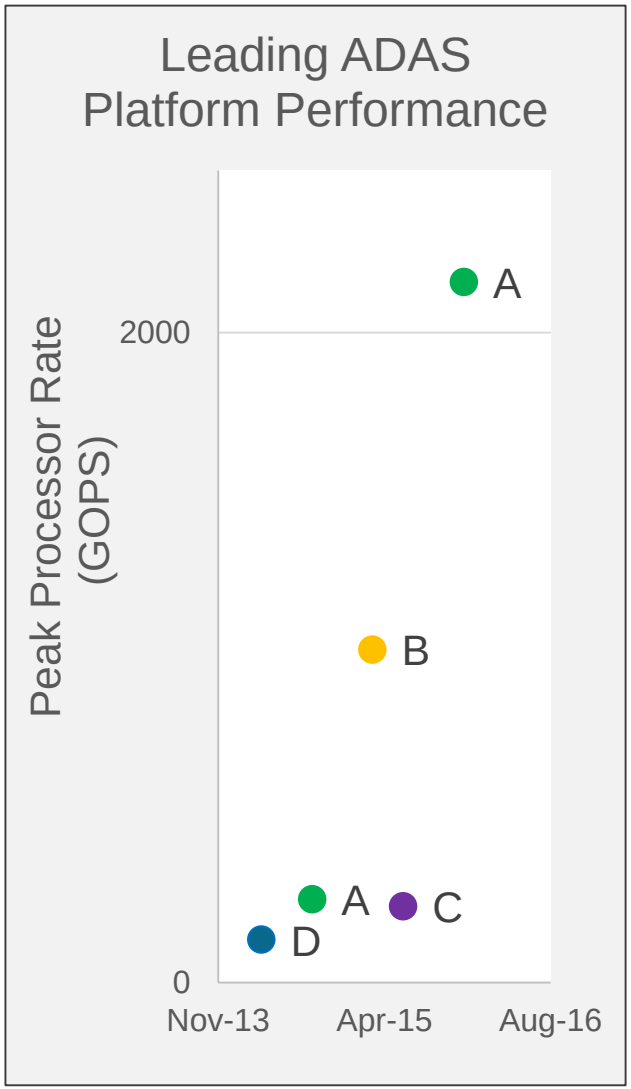
cadence

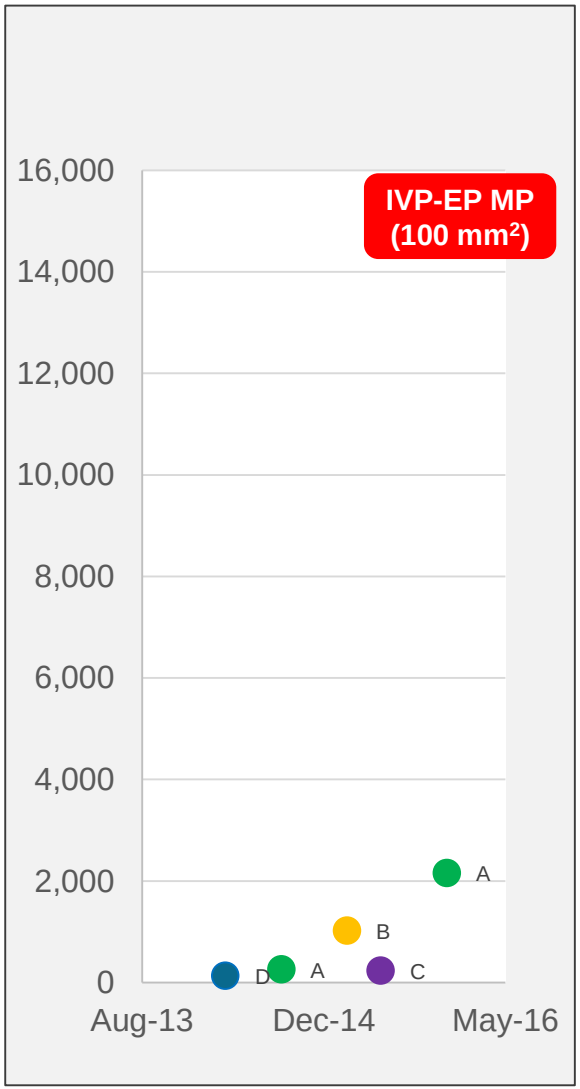
CNN Engine Example

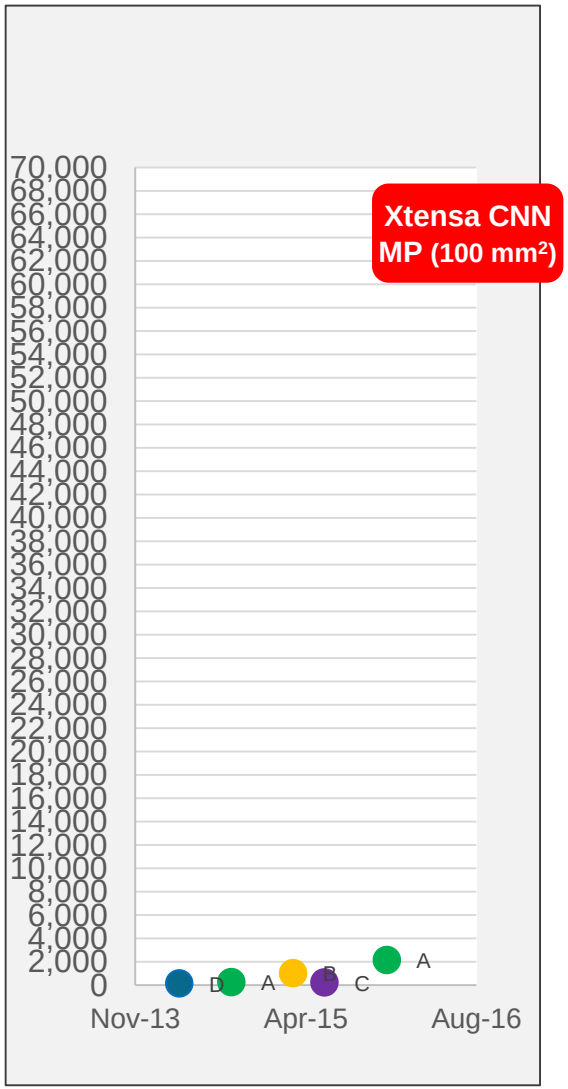
- Demonstration of a specialized DSP architecture using TIE on Xtensa – source available
- Very high throughput (256 16x8 MAC/cycle) on filter and convolution algorithms for vision and object recognition
- Architecture:
  - 512 x 8 data vectors (3r/1w)
  - 2048 x 2 accumulators (1r/1w)
  - 512 x 2 align regfile 1r/1w)
  - 2 slot, 40b instruction width, 47 added ops
  - 48GB/s local data BW/core
- High compute density:
  - 4x higher efficiency than IVP-EP
  - 9T full layout
  - With 32KB I/ 64 KB D
  - 0.95mm^2 @ 750MHz
  - 200 GMACS/mm^2



Huge upside in vision platform performance:
>60 Trillion ops per 100mm²







Core + memory 28nm

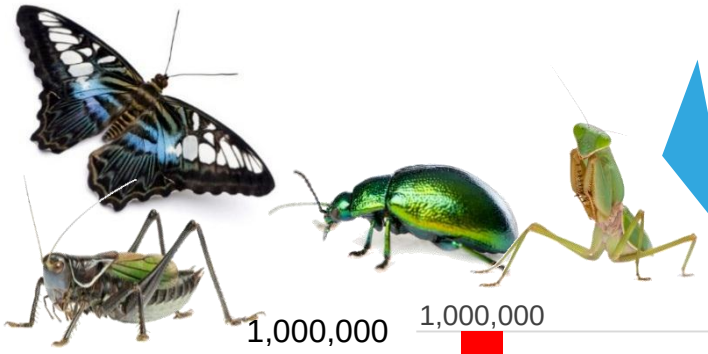
cadence

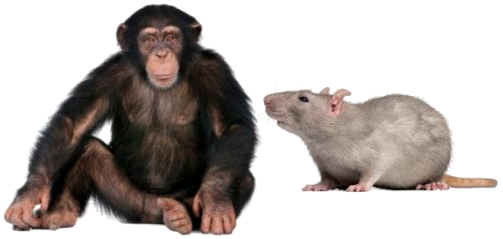
A closing thought
Specialization ➔ proliferation of designs

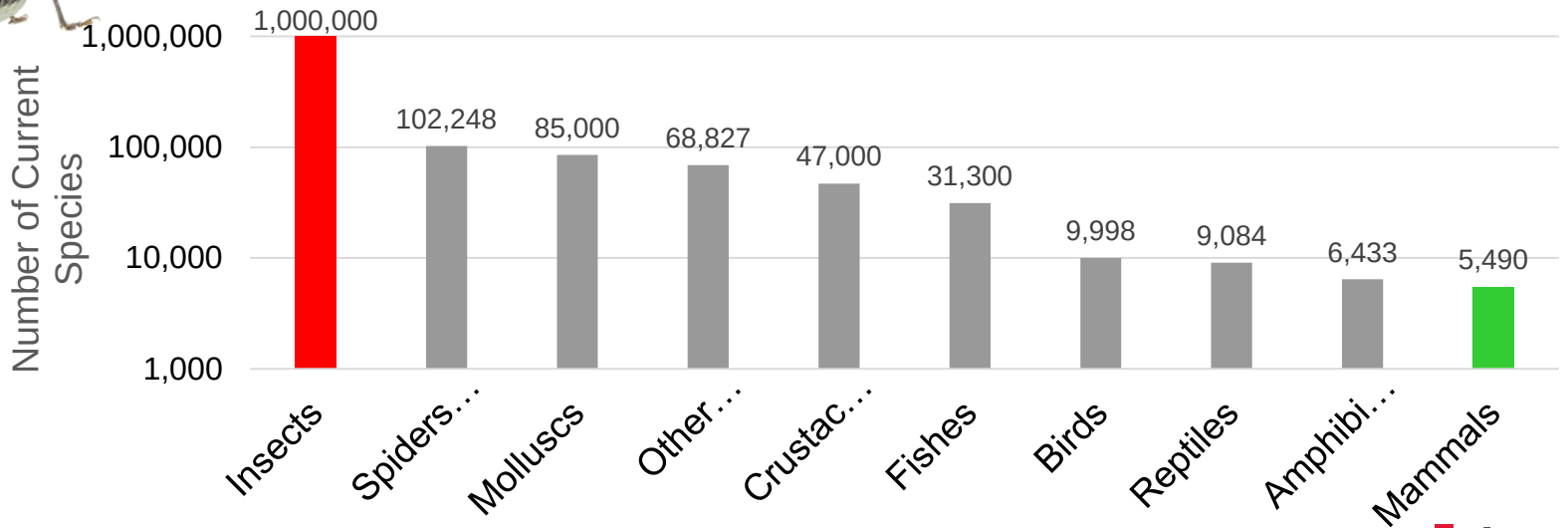
Specialized low-cost, low-energy platforms: wearables, smart-home, medical, industrial, mil-aero, sensor-rich **extreme fit** SoCs

All SOC Designs

General-purpose all-in-one system platforms: cellphone, PC, server, gateway – **extreme scale** SoCs







cadence

Agenda

- A mainstream view of big compute and advanced EDA/IP
  - System Design Enablement
  - Design and Verification across digital and high-speed analog
  - Universal adoption of IP-rich development
  - New roles for processors
- But the world is changing…
  - The new data imperative
  - Energy and layered cognition
  - Scaling to ultra-low power
- Example: vision super-computing
  - Vision DSPs
  - Convolutional Neural Network compute
  - Hybrid architectures for high-density computation


cadence®