

Lucian Codrescu  
Sr. Director, Technology  
Qualcomm Technologies, Inc.

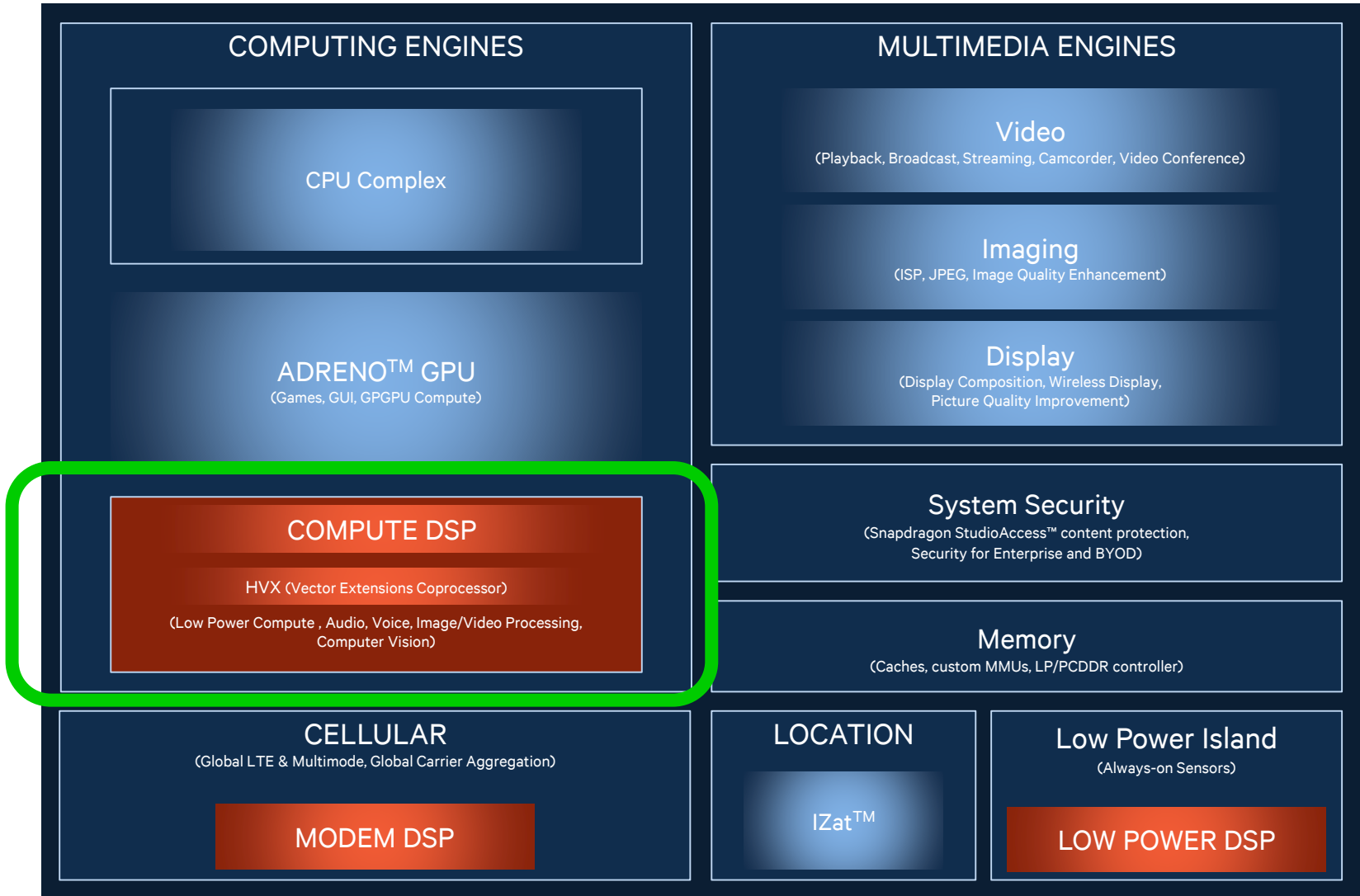
---

Architecture of the  
Hexagon™ 680 DSP for  
Mobile Imaging and  
Computer Vision

---



# Hexagon DSPs in Snapdragon™ 820



# Hexagon Vector eXtensions (HVX)

- DSP Extensions for Advanced Imaging and Computer Vision
- Achieve Performance / mW substantially better than CPU



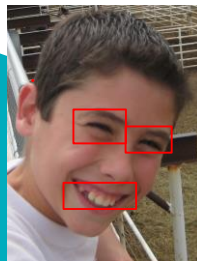
Camera



Augmented /  
Virtual Reality



Video



Computer Vision



HMD

# Low Light Video & Photos



## Algorithm

HVX accelerated  
local tone mapping and  
temporal noise reduction to  
brighten dark videos and photos

## Benefits

Adaptively brightening  
dark areas of video/images



# Content Adaptive Detail Enhancement



**Original**



**After HVX Processing**

---

# Hexagon Vector eXtensions (HVX)

**Domain Specific  
Architecture**

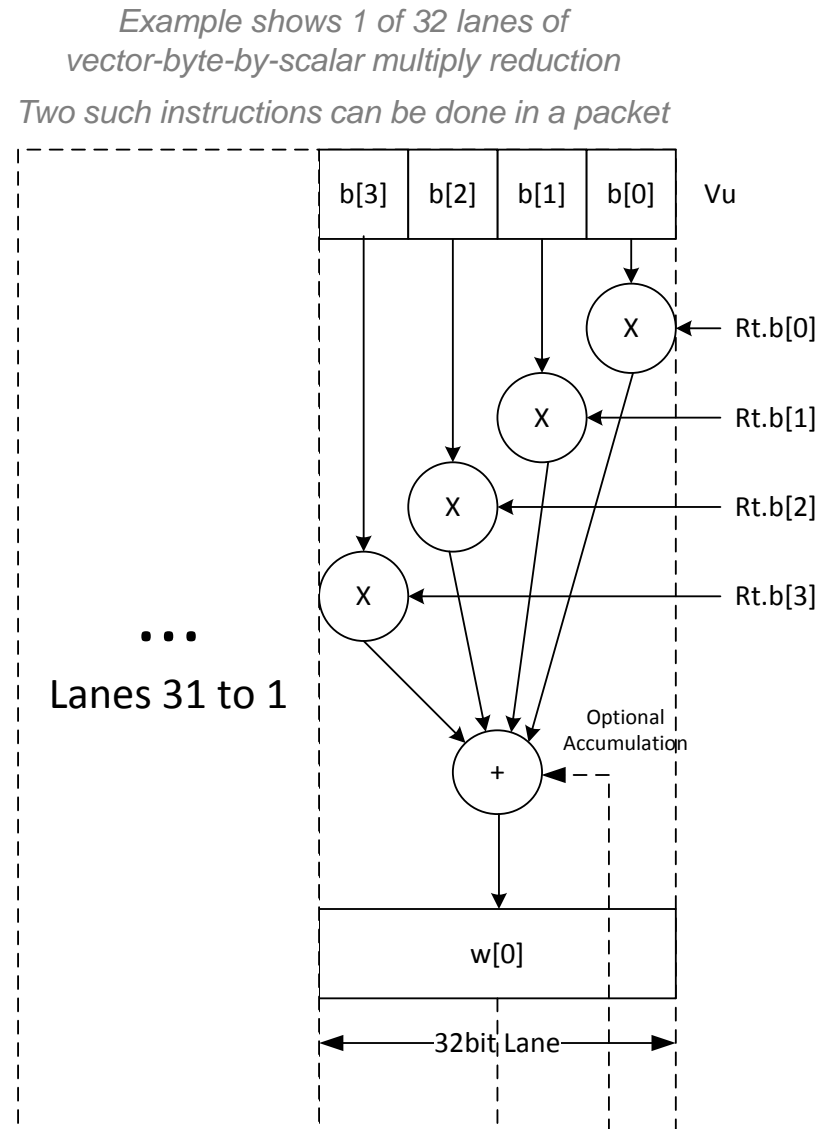
Familiar  
Programming  
Model



Tight System  
Integration

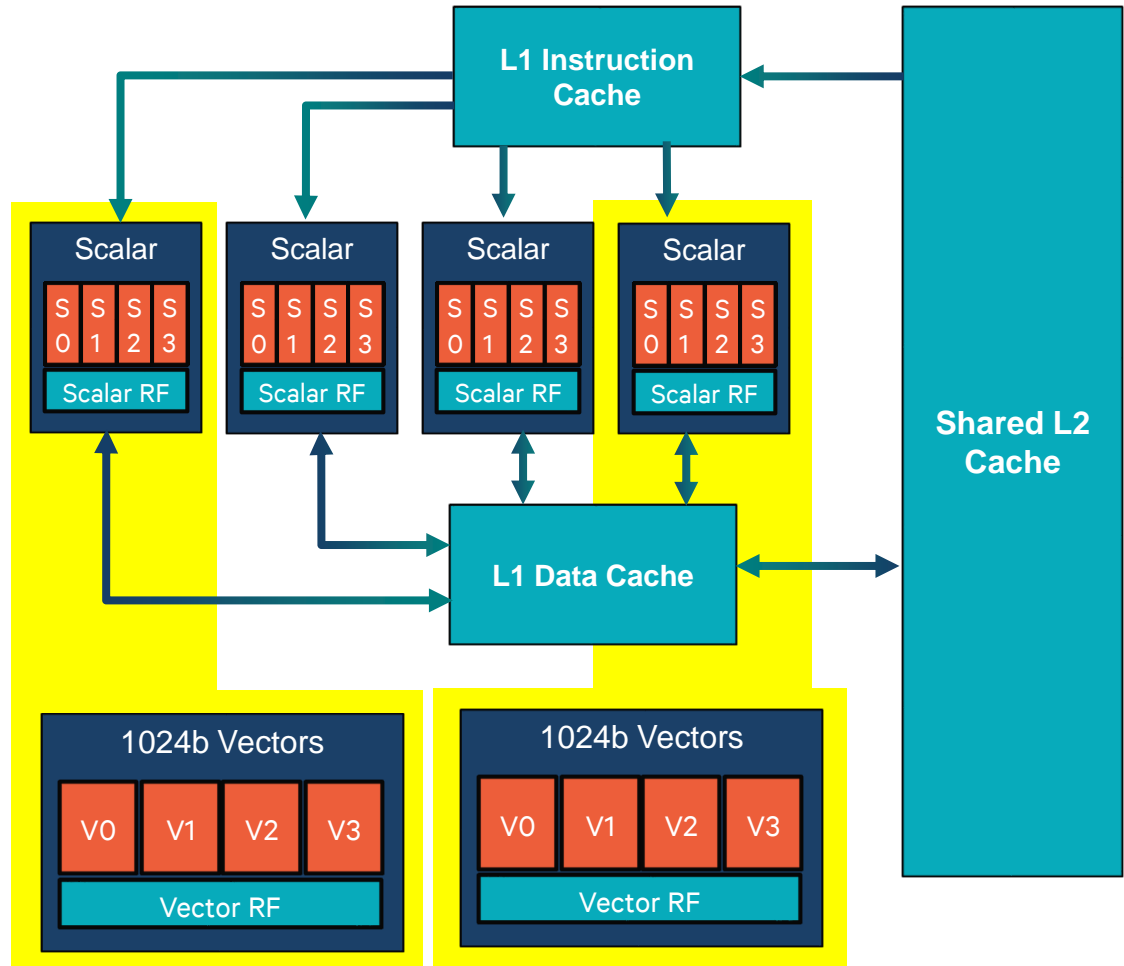
# HVX Architecture – SIMD Extensions

- Large SIMD Extensions
  - 1024b SIMD \* 4 vector-slot VLIW
  - 4096 result bits / cycle
- 256 8x8 mpy, 64 16x16 mpy
- 32 1024-bit vector registers
- 8/16/32 bit fixed point
- *NO floating-point*
  - *Smaller & Lower Energy Design*
  - *Algorithmically not needed for majority of CV/Imaging Apps*
- Special ISA: Sliding window filters, LUTs, Histograms
- Performance is sufficient for UHD video post-processing, 20Mpix camera burst mode processing ... and more



# HVX Architecture – Threading Model

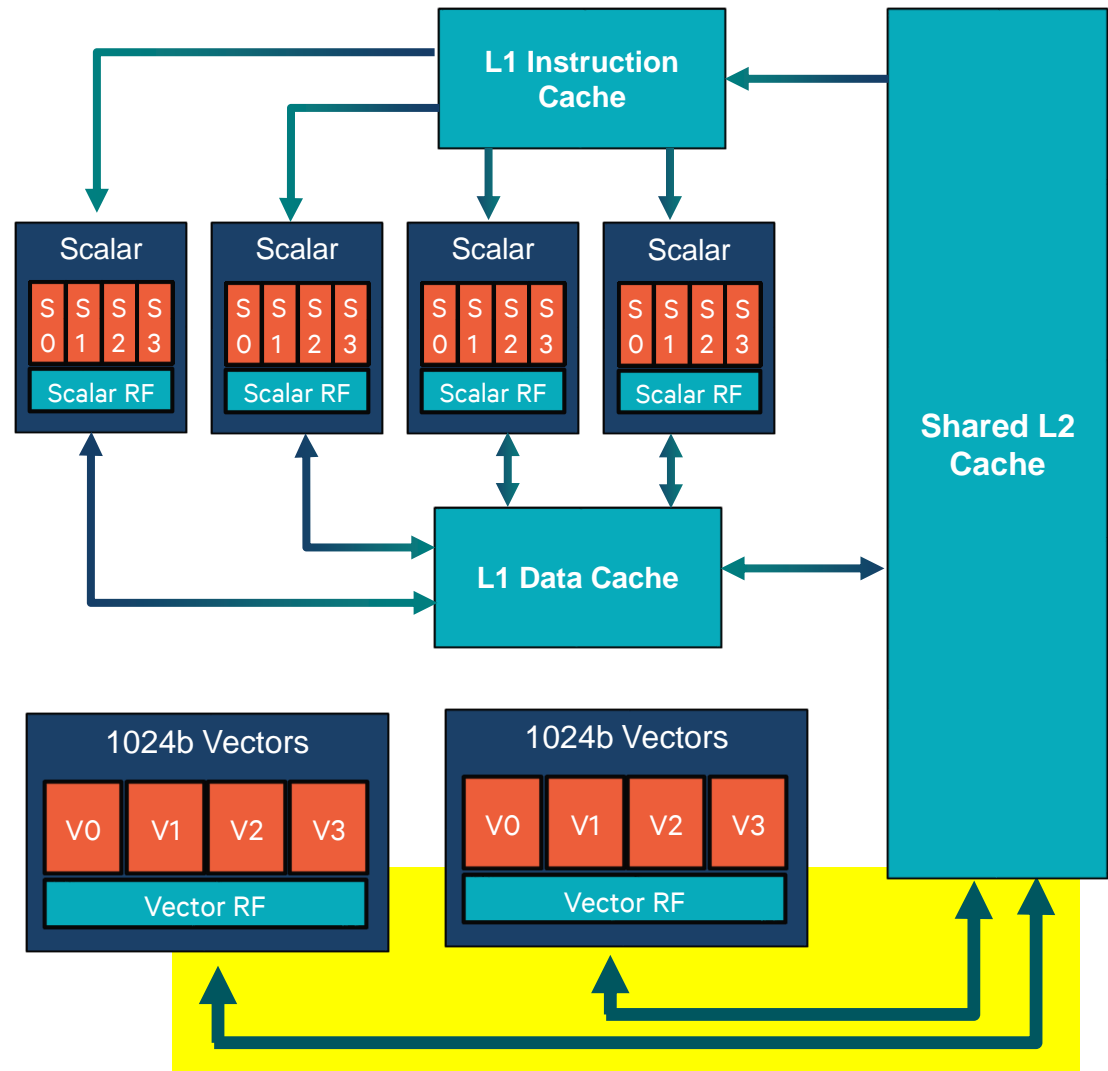
- 4 Parallel Scalar Threads each with 4-way VLIW and shared L1/L2
  - 500MHz per Thread
  - 2GHz total scalar performance
- 2 HVX Contexts, controllable by any two scalar threads
  - 500MHz per Thread
  - 1GHz total vector performance
  - Other 2 threads can do scalar work in parallel





# HVX Architecture – Memory

- L2 is the first level memory for the vector units
  - Large primary memory to hold image data reduces tiling overheads seen on small L1
  - Single cycle Load to Use
  - Supports full BW
  - Simplifies programming
- L1/L2 is kept HW coherent
- Streaming prefetch from DDR to L2
- Vector units support variety of Load/Store instructions:
  - Unaligned
  - Per-Byte Conditional



# Hexagon Vector eXtensions (HVX)

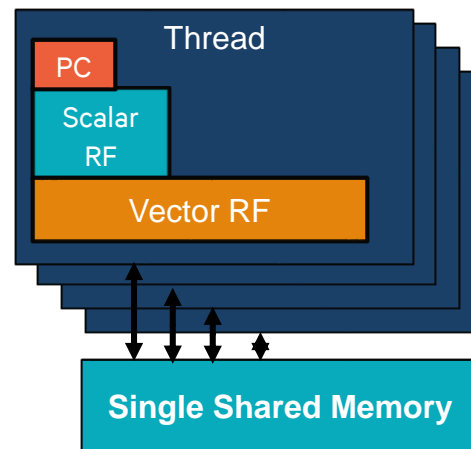
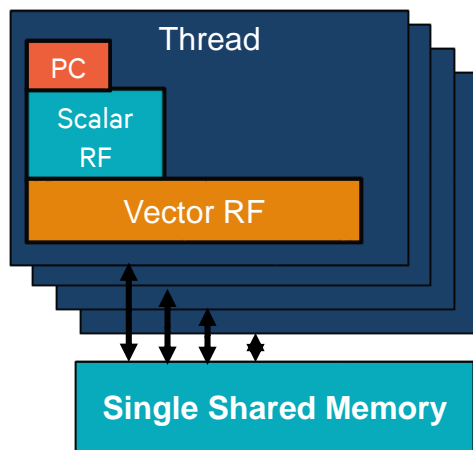
## Domain Specific Architecture

- Wide 1024-bit SIMD (for pixel data parallelism)
- Emphasis on low precision fixed-point + Special ISA
- Parallel and coordinated Scalar & Vector Threads
- Large primary cache for Imaging Working Sets



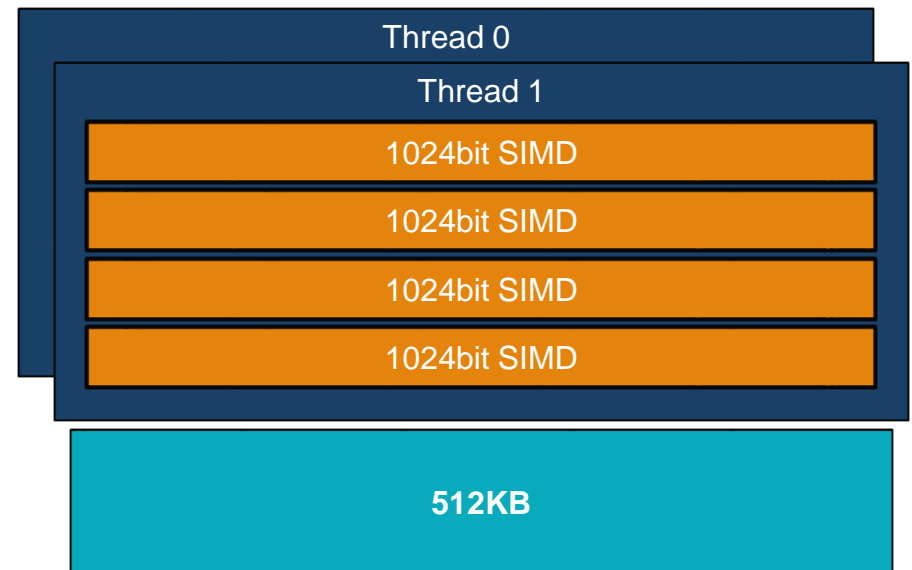
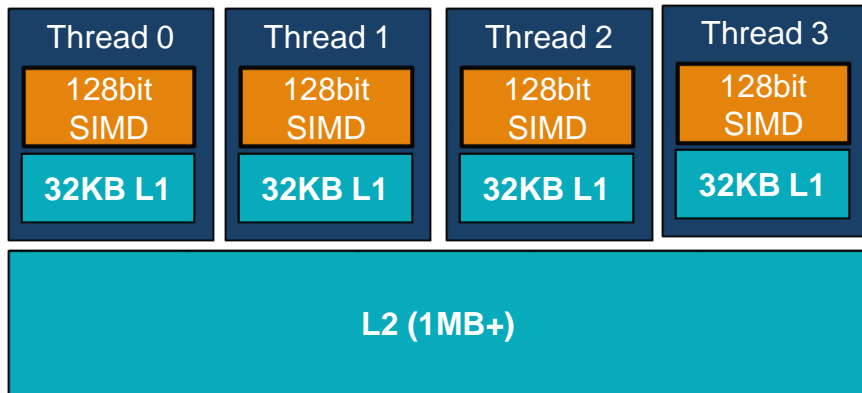
# DSP with HVX has a CPU-like Programming Model

CPU w/ SIMD	Hexagon DSP with HVX
Multi-Thread Programming w/ cache-based coherent shared memory	Multi-Thread Programming w/ cache-based coherent shared memory
SIMD Instructions on Vector RF	SIMD Instructions on Vector RF
Scalar Instructions on Scalar RF	Scalar Instructions on Scalar RF
Good Control Performance for mixed vector/scalar	Good Control Performance for mixed vector/scalar



# With Key Differences for Performance & Power

Quad CPU with Neon	Hexagon DSP with HVX	HVX Advantage
128 bit SIMD with 1 SIMD pipeline/CPU is common	1024 bit SIMD, 4 Pipelines	8x compute/cycle
SIMD thread on 32KB L1	SIMD threads share 512KB “L1”	8x more “L1” memory/thread Efficient Data Sharing
Floating-Point in SIMD	Only Fixed-Point in SIMD	Lower area & power



---

# Familiar Programming Model

- Conventional Tools & Techniques
  - Shared memory POSIX-like threads (on DSP RTOS)
  - LLVM compiler
  - Program with C/C++ and Intrinsics
  - Pre-Optimized Libraries for common filters & algorithms
- Easy DSP offload from Android
  - DSP code is dynamically loaded on invocation of synchronous Remote Procedure Call
  - Higher Level Android Frameworks Available for
    - Camera Pre-Processing in Real-Time
    - Video Post-Processing in Real-Time
    - Offline Camera Post-Proc and Computer Vision
- On-going Research: Halide Language on HVX



# Hexagon Vector eXtensions (HVX)

## Domain Specific Architecture

- Wide 1024-bit SIMD (for pixel data parallelism)
- Emphasis on low precision fixed-point + Special ISA
- Parallel and coordinated Scalar & Vector Threads
- Large primary cache for Imaging Working Sets

## Familiar Programming Model

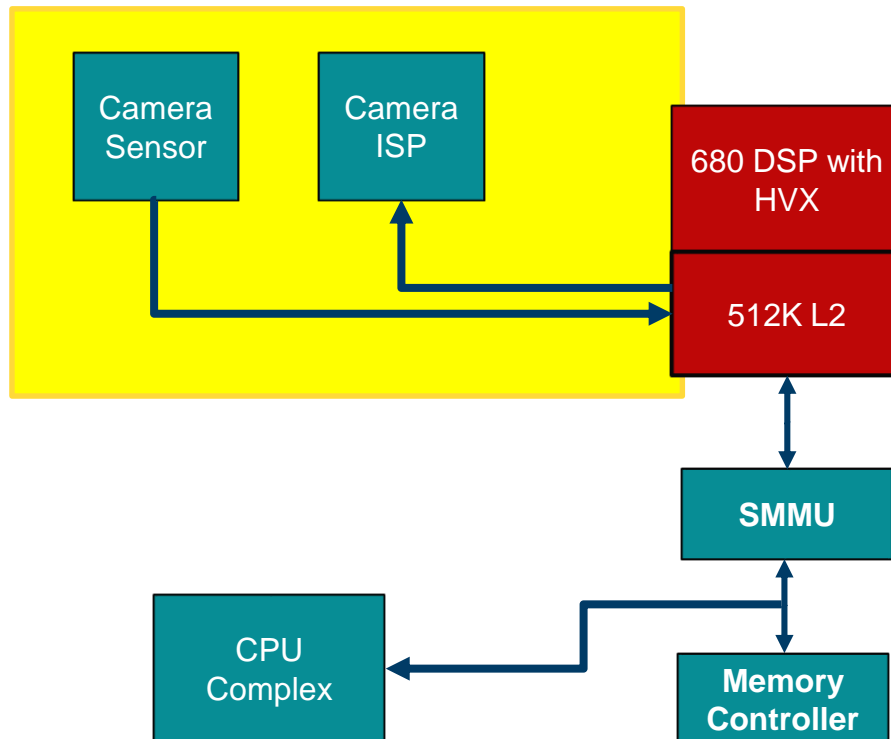
- SIMD + Multi-Thread
- Cache based coherent shared memory
- Programming in C/C++ with Intrinsics
- Pre-optimized libraries for common functions



## Tight System Integration

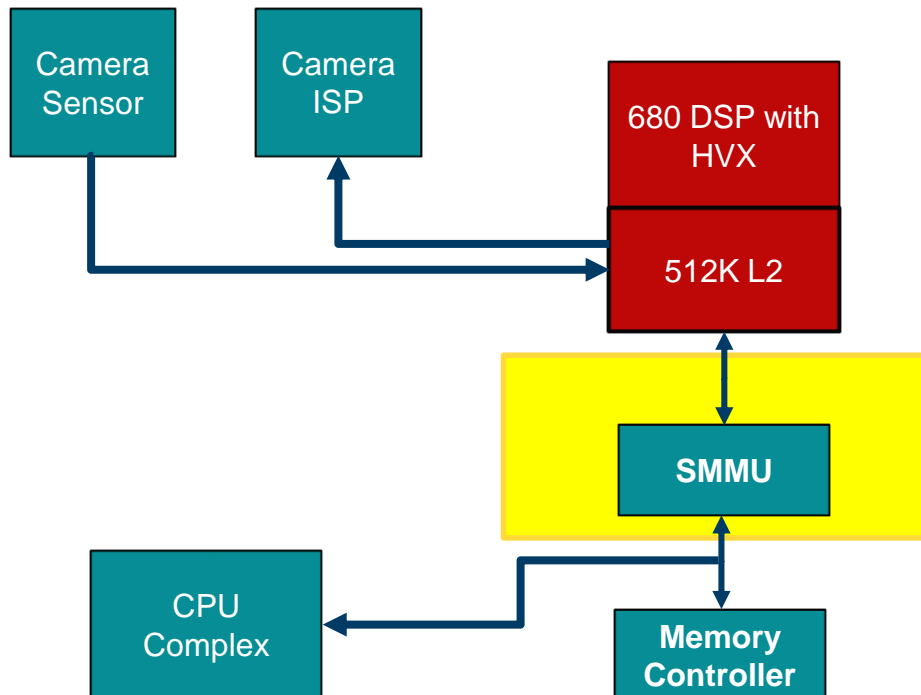
# System Features – Streaming Interface

- Custom Camera Pre-Processing without DDR access saves power
- Raw data from Image Sensor is streamed into L2 at up to 1.2Gpixels/sec
- Streamer formats / pads / aligns data for convenient vector processing
- Data held in input & output circular buffers
- HVX processed pixels can be streamed out to ISP HW
- Synchronization through memory-mapped control registers



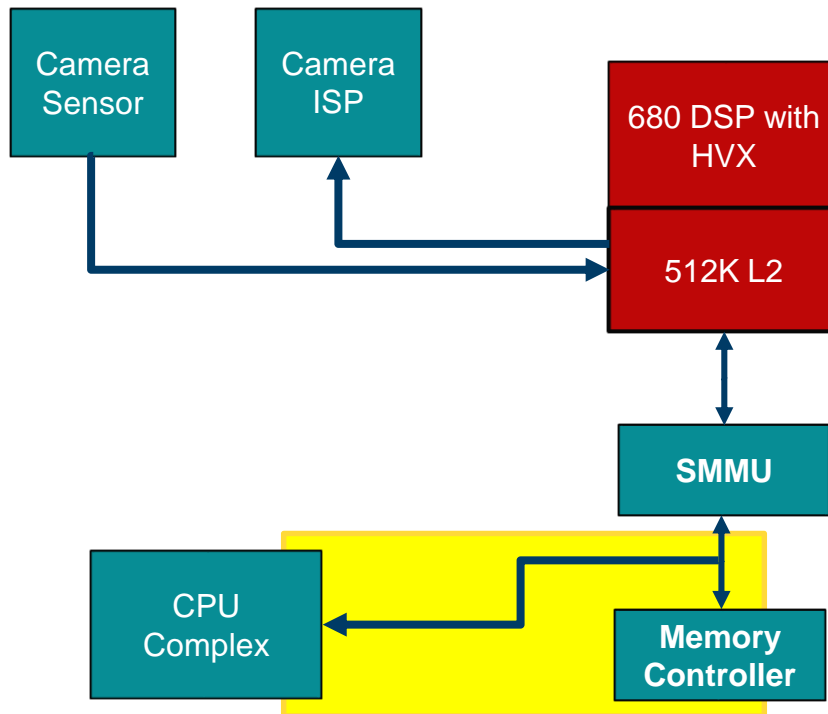
# System Features – SMMU

- ARM Compliant SMMU allows for Zero-Copy data sharing with CPU
- Multi-Threaded DSP can be servicing multiple offload sessions (concurrent apps for Audio, Camera, Computer Vision (CV), etc.)
- SMMU supports multiple Context Banks to allow sharing with multiple different address spaces on CPU
- SMMU can be used to support processing on Secure Content managed outside of HLOS



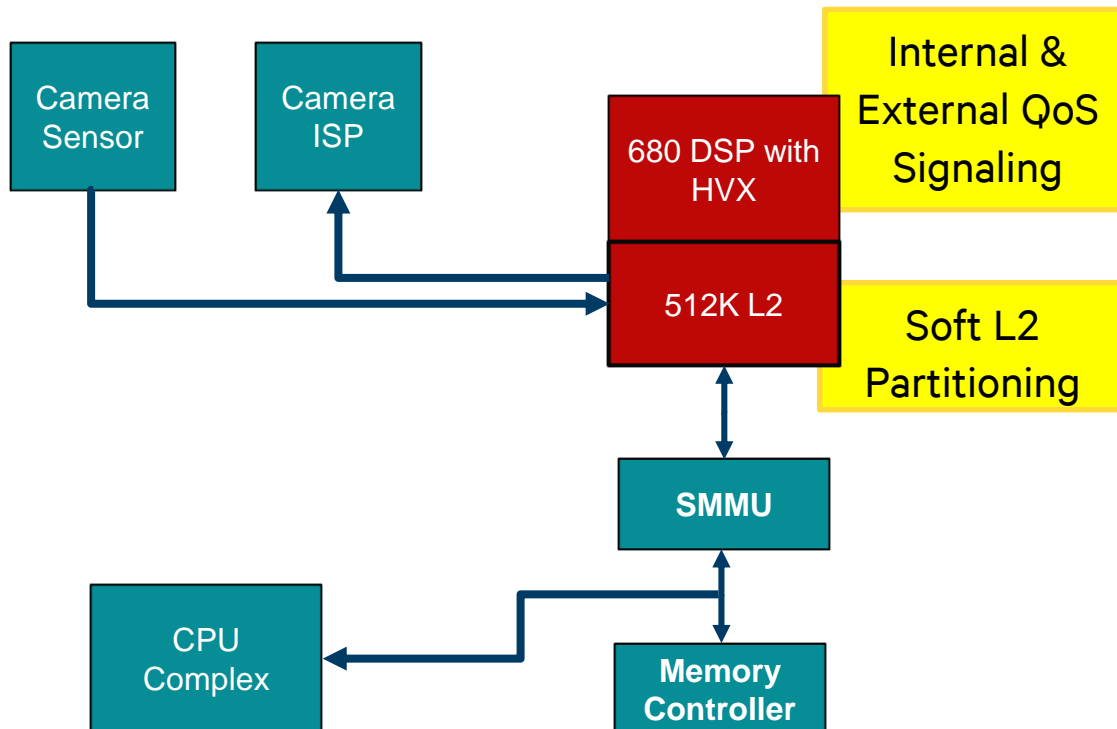
# System Features –One Way Coherency

- Bus writes from the DSP snoop-invalidate the CPU caches
  - Avoids expensive SW cache maintenance on CPU, saving time & energy
  - DSP side requires cache maintenance, but this is handled transparently in the RPC software interface layer



# System Features – Quality of Service

- DSP services multiple real-time clients (Audio, Camera, CV), each with their own timelines
  - Imaging Algorithms can consume large amounts of external bandwidth and cause congestion for other clients
- L2 partitioning: L2 can be soft partitioned into regions assigned to different threads
- Internal QoS: HW-based prioritization of memory requests; SW assigns priorities to threads
- External QoS: System Level HW algorithms can throttle DSP traffic to ensure system performance





# Hexagon Vector eXtensions (HVX)

## Domain Specific Architecture

- Wide 1024-bit SIMD (for pixel data parallelism)
- Emphasis on low precision fixed-point + Special ISA
- Parallel and coordinated Scalar & Vector Threads
- Large primary cache for Imaging Working Sets

## Familiar Programming Model

- SIMD + Multi-Thread
- Cache based coherent shared memory
- Programming in C/C++ with Intrinsics
- Pre-optimized libraries for common functions



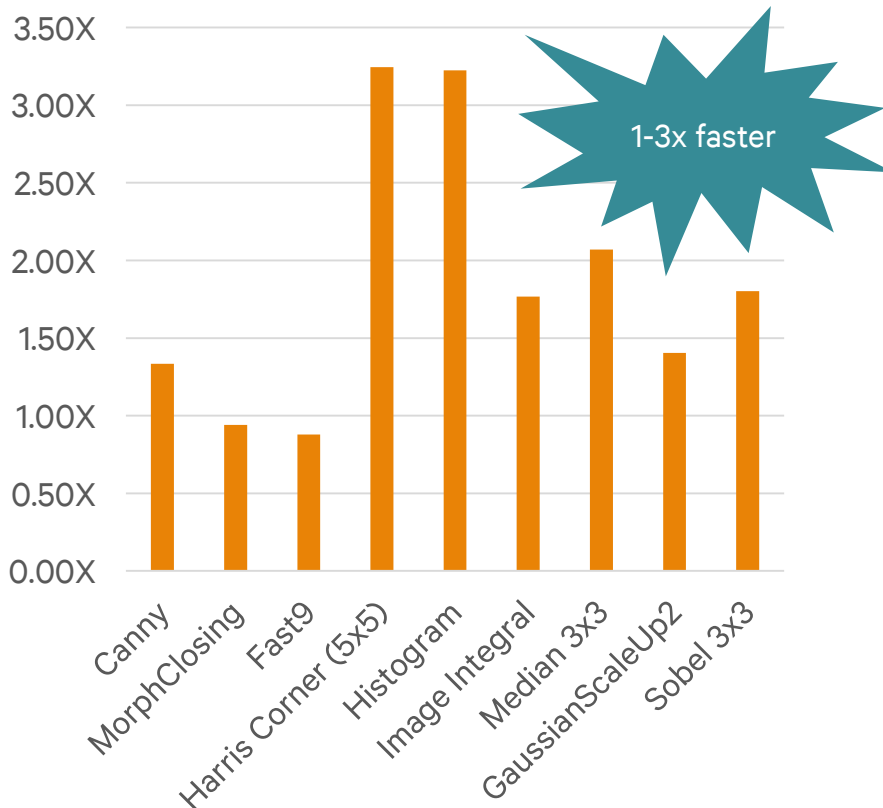
## Tight System Integration

- Camera Streaming Interface
- SMMU
- One way Coherency
- System & Internal QoS

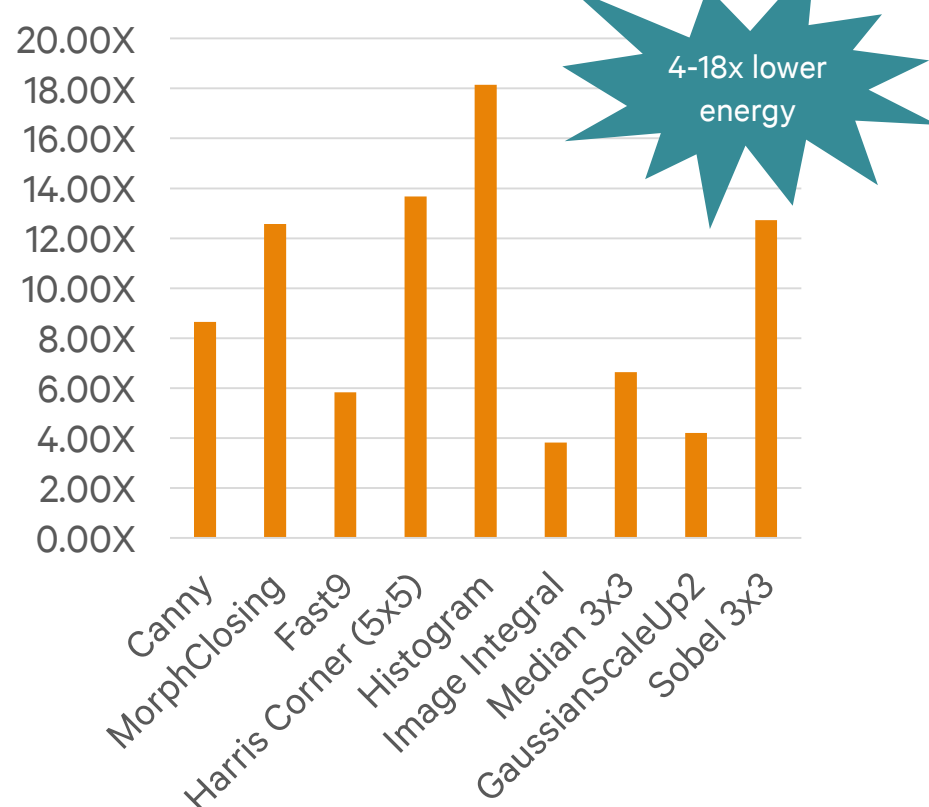
# Imaging & Vision Kernel Benchmarks

- DSP with HVX vs Quad Krait CPU *with full Neon-Optimization*
  - Quad Krait CPU clocked at 2.65GHz
  - Single DSP/HVX clocked at 725MHz
  - Core power only excluding SoC infrastructure, DDR, etc.

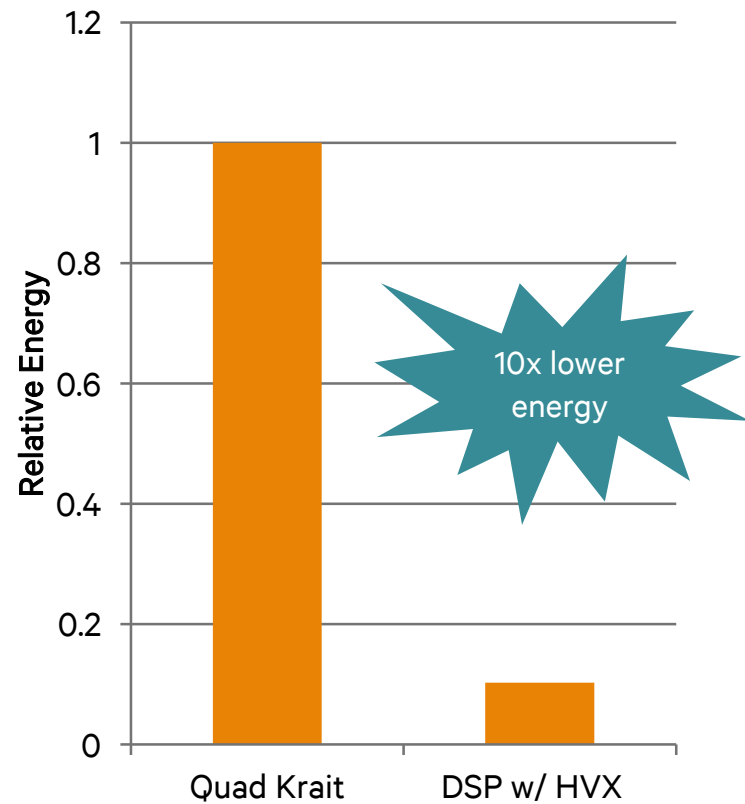
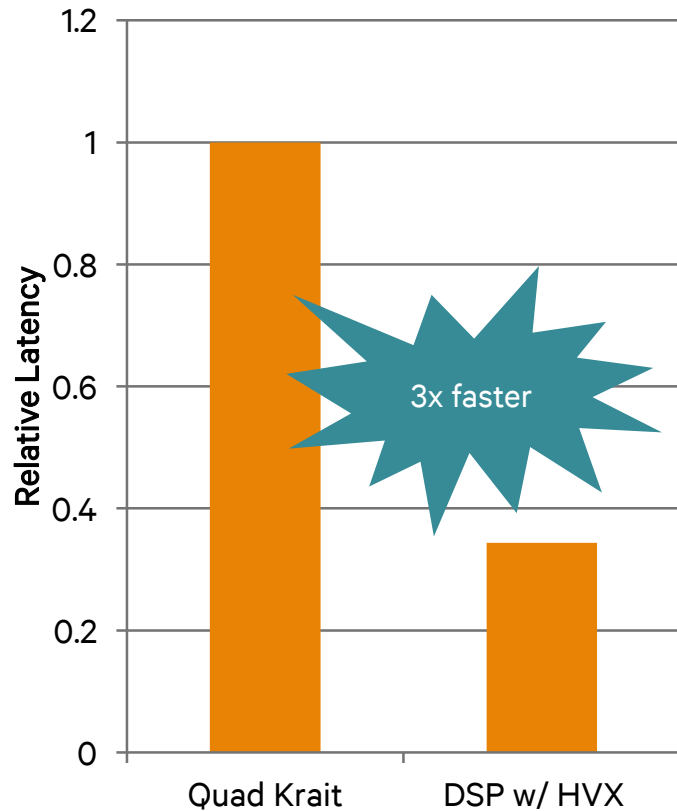
Speedup vs Quad CPU



Energy/Pixel Reduction vs Quad CPU



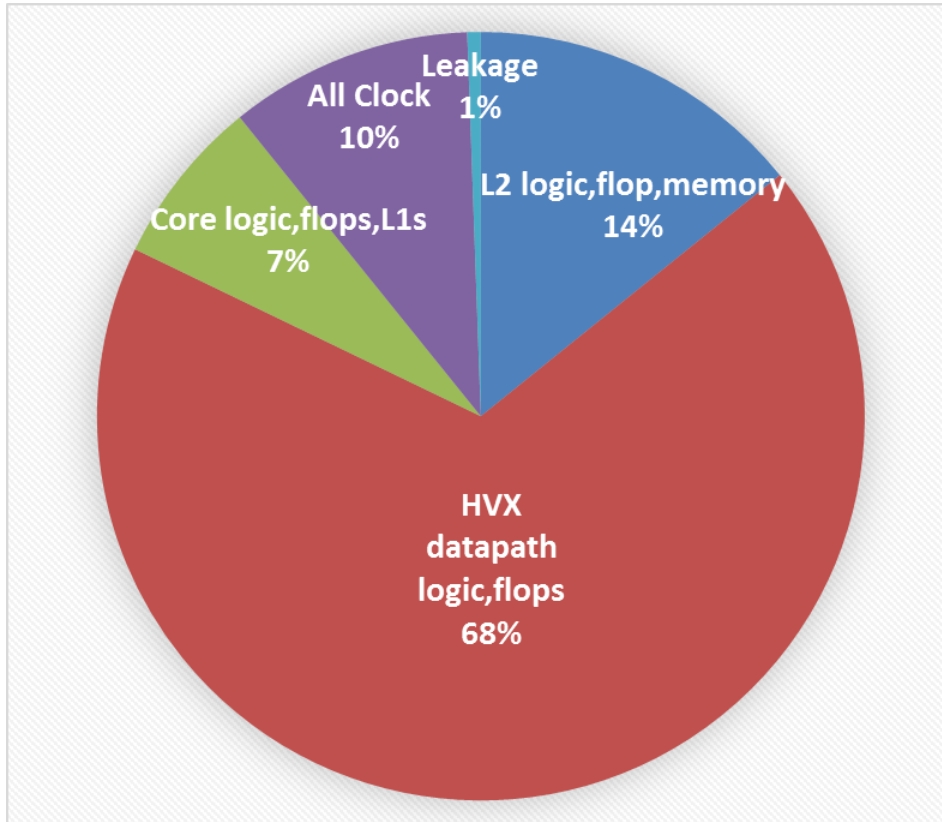
# Full Application Low Light Video Enhancement



# DSP vs CPU power

- Data for a typical imaging application
- Unlike CPUs, the bulk of the power is spent in compute datapath

## DSP w/ HVX

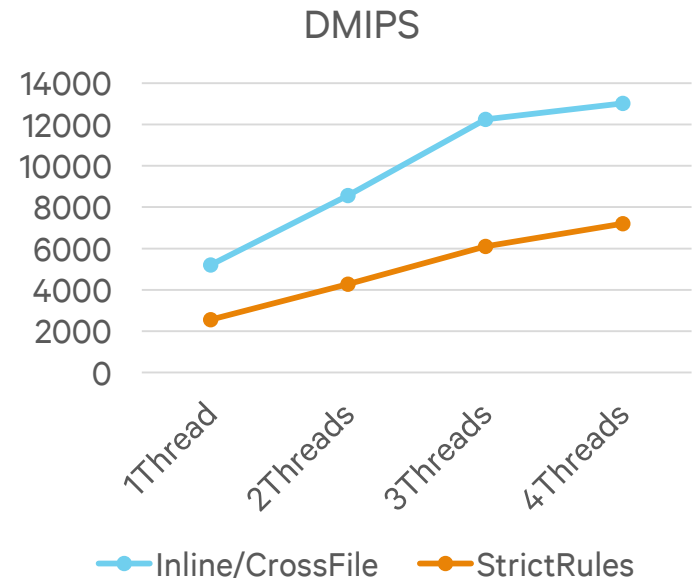
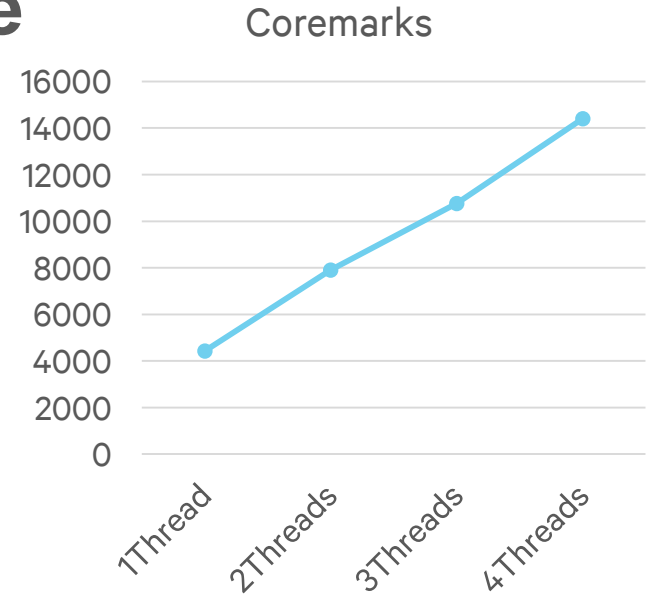


## Why power savings?

CPU	DSP
OOO/Superscalar	In-order/VLIW
Short Vector have high overhead	Long Vectors amortize overhead
L1 & L2 used for image data	Only L2 used for image data
Design Optimized for high MHz incurs high clock power overheads & high leakage	Design Optimized for low-power & lower MHz reduces overheads & leakage

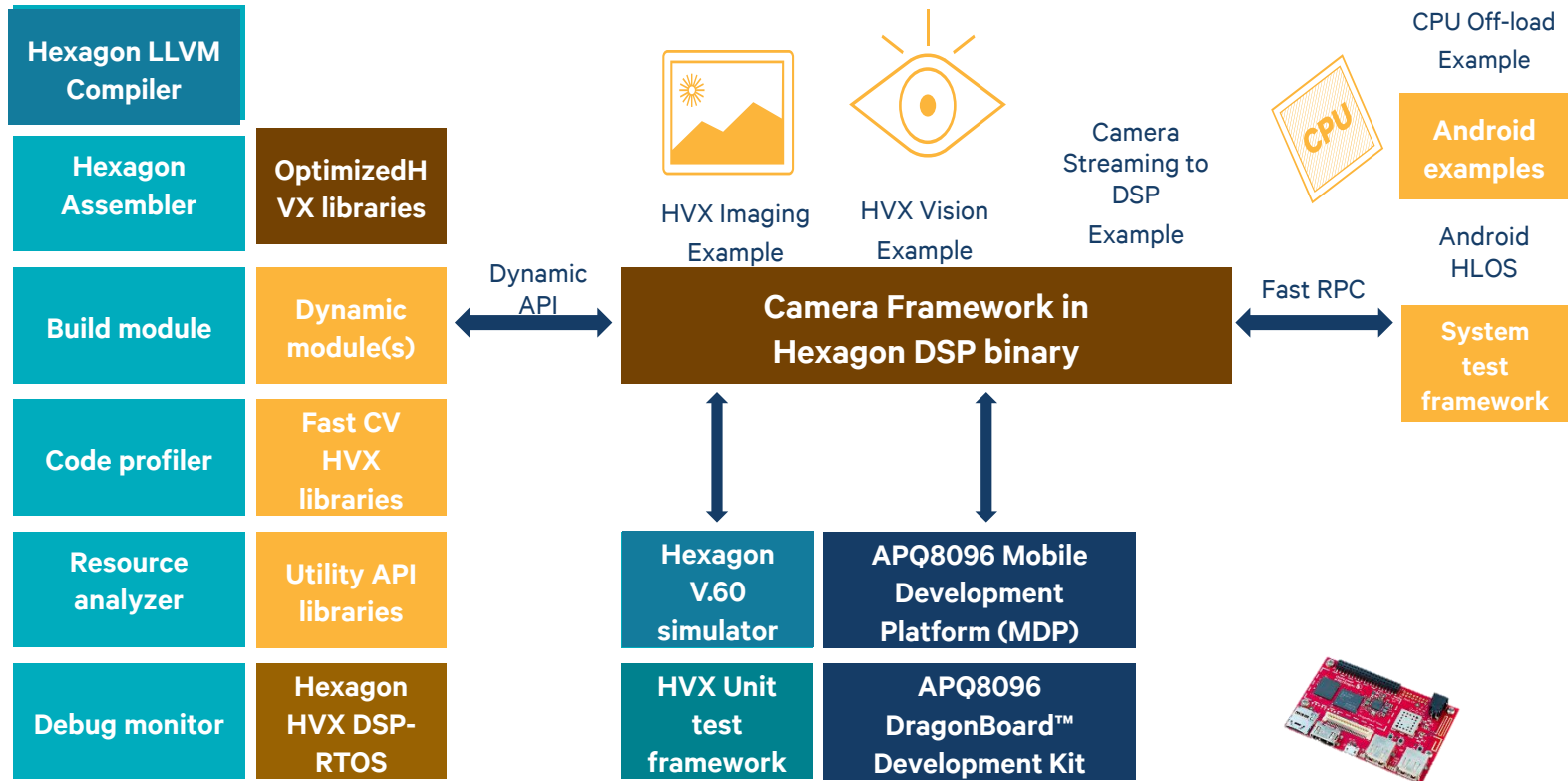
# HVX Architecture – Scalar core

- Many target applications have both scalar and vector components
- Good control performance means less need to move control parts of an application to the CPU:
  - Easier to offload a full algorithm rather than partition it
  - Support apps with vector→scalar→vector dependency loops
  - Keep data local in cache
  - Avoid CPU power





# Hexagon SDK 3.0 for Hexagon600™ Support



Visit [developer.qualcomm.com](http://developer.qualcomm.com) to request the latest Hexagon HVX SDK

# Hexagon Licensees Deploying HVX Solutions



*Almalence*



core|photonics



**Ittiam**



**OMRON**

**PATHPARTNER**



**SONY**



\*Not a complete listing, simple for OEM to enable their own visual partners

---

# Thank you

Follow us on:  

For more information on Qualcomm, visit us at:  
[www.qualcomm.com](http://www.qualcomm.com) & [www.qualcomm.com/blog](http://www.qualcomm.com/blog)

©2013 Qualcomm Technologies, Inc.

Qualcomm and Hexagon are trademarks of QUALCOMM Incorporated, registered in the United States and other countries. All QUALCOMM Incorporated trademarks are used with permission. Other product and brand names may be trademarks or registered trademarks of their respective owners. Hexagon is a product of Qualcomm Technologies, Inc.

