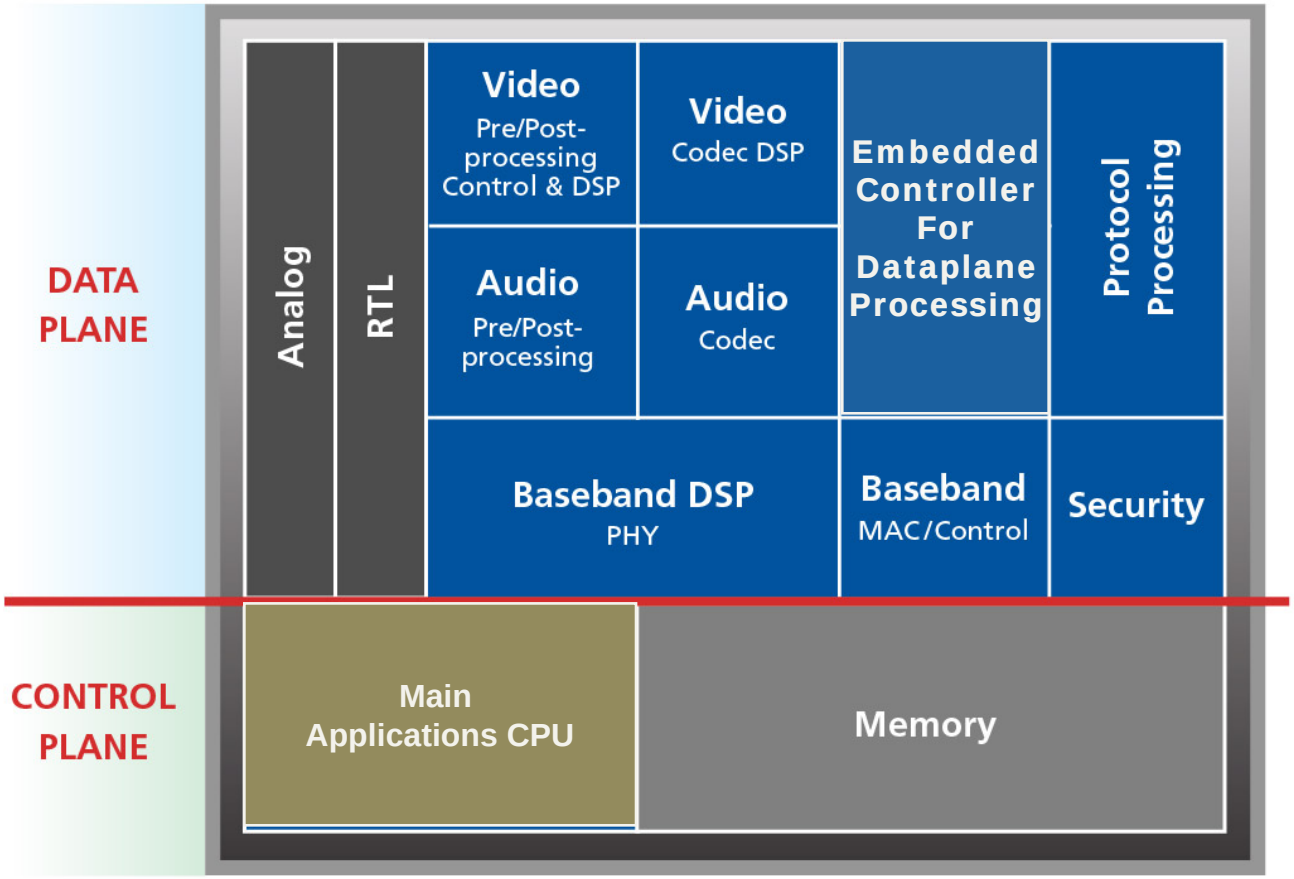
Energy-Efficient LTE Baseband with Extensible Dataplane Processor Units

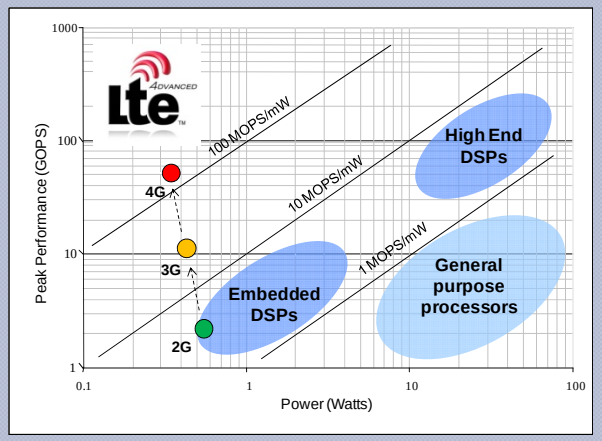
Chris Rowen
Founder and CTO
Tensilica Inc

Tensilica Focus:
Dataplane Processing Units (DPUs)

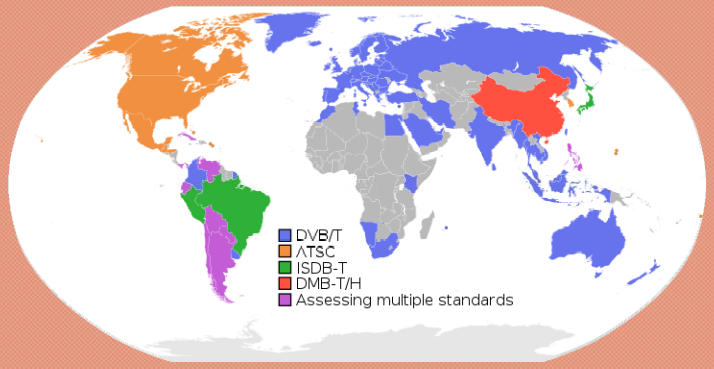
DPUs: Customizable CPU+DSP delivering >10x higher performance than CPU or DSP with better flexibility than RTL



Next-Generation Baseband Standards Drive Fundamental Change in Market



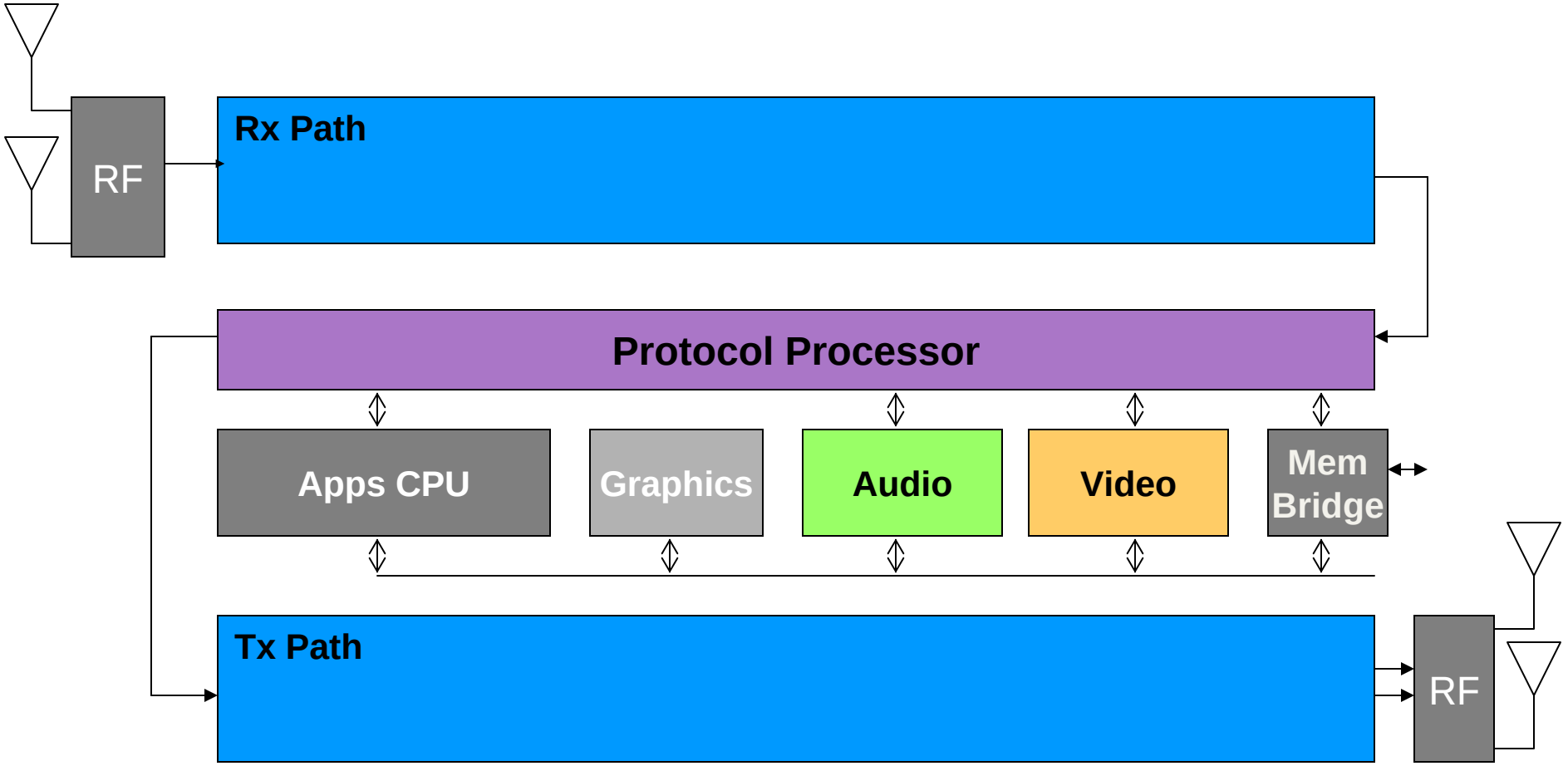
Emerging standards (LTE, WiMAX) require processing power exceeding the capabilities of today's DSPs



Drive towards multi-standard receivers requires programmable solutions

Push towards low-cost green infrastructure requires high performance at very low power

Dataplane Processors for Almost Every Wireless Systems Role



New Wireless Standards Drive Performance and Efficiency

Evolving from 2G to 4G:

- 100-1000x increase in op rate
- Baseband power budget reduced by 2-3x

Preferred implementation:

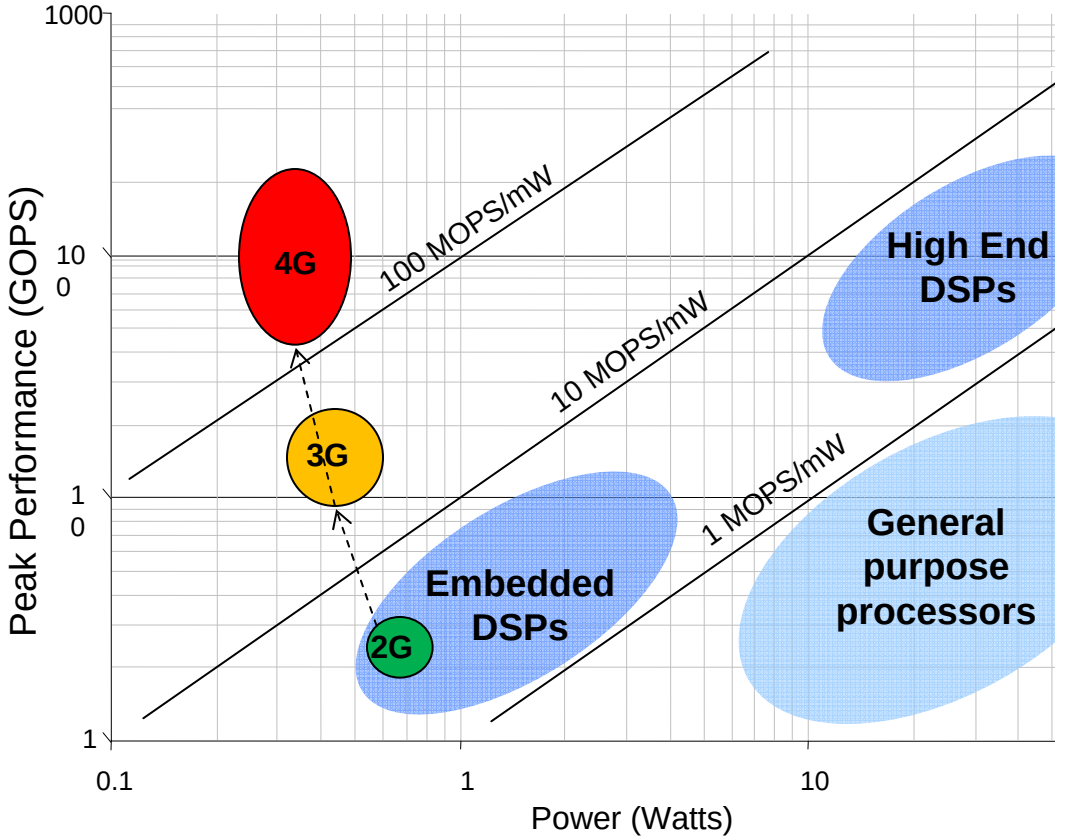
- 2G (GSM)➔DSP
- 3G (UMTS)➔DSP + function-specific coprocessors
- 3.9B/4G (LTE/LTE-A)➔DPU

Performance increase to 3.9G/4G

- 3G ➔ 3.9G (LTE @140Mbps)
  - 7x demodulation Mops/mW
  - 130x decoding Mops/mW
- 3G➔ 4G (LTE-Adv @ 1Gbps)
  - 85x demodulation Mops/mW
  - 870x decoding Mops/mW

DPU for baseband:

- Tight integration of DSP and special-purpose function units to increase efficiency and programmability



Source: From Multi-Core for Mobile Phones: C.H van Berkel, ST-NXP Wireless DATE09

Why DPUs?

Faster, More Flexible than DSPs and Coprocessors

- **Existing DSPs and Coprocessors Inadequate:**

1. Wireless computing requirement growing much more rapidly than GP DSP performance or Moore's Law silicon scaling
2. DSP + special-function coprocessor (RTL) too rigid
   - Multiple complex standards demand programmable acceleration
   - Excessive bus, cycle, power overhead in DSP to coprocessor communication

- **So Tensilica Invents DPU Technology**

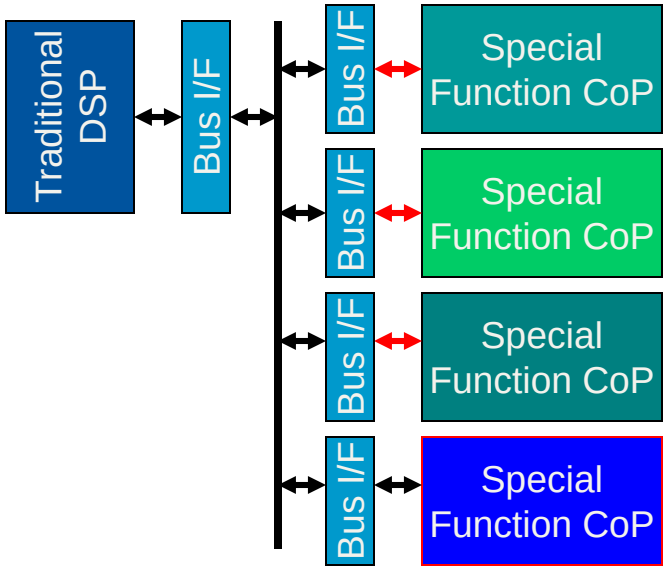
➔ Leading edge DSP foundations: ConnX family:

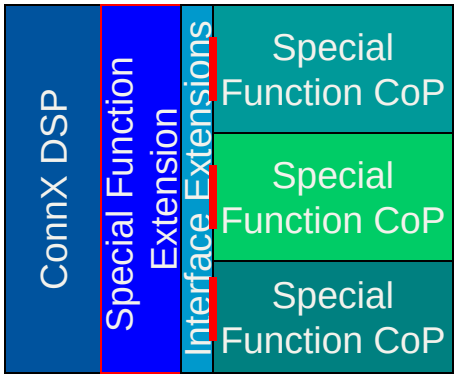
Introducing flagship Baseband Engine: up to 8 engines at 80ops/cycles per engine: 640 ops/cycle

➔ Direct integration of special-function coprocessors directly into DPU for programming and full debug

Processor instruction set extension

Direct RTL interface extension





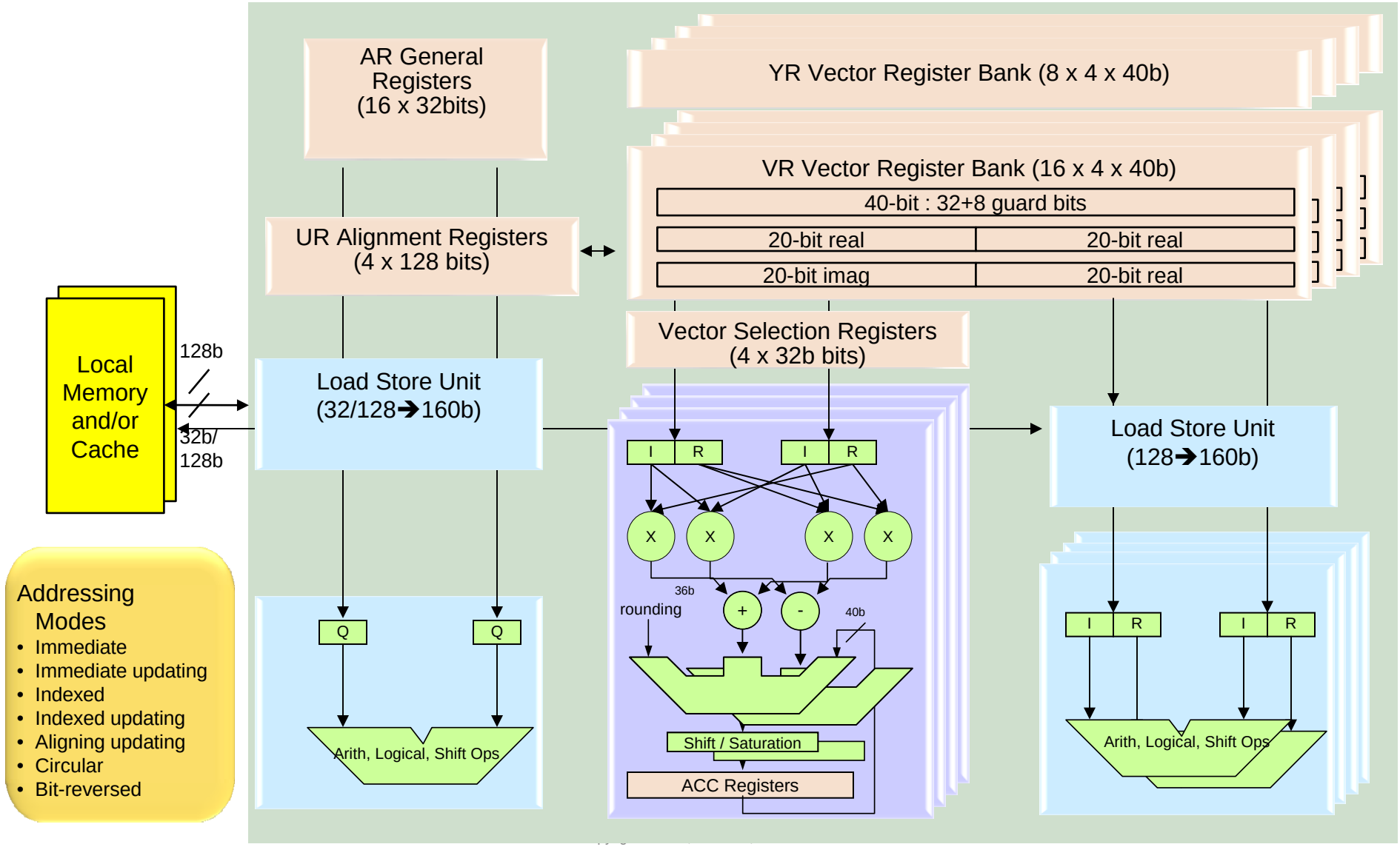
Three Key Ingredients

1. New Baseband DSP Options: ConnX Family
   - Flagship: ConnX Baseband Engine: 16-MAC throughput for OFDM-based wireless
2. Direct integration of RTL accelerators with control engines
3. Improved processor foundation for higher processor efficiency

The ConnX Baseband Engine

- High performance DSP for wireless communication
  - Wireless: LTE, LTE-Advanced, WiMax, WiFi
  - Broadcast: DVB-t, ATSC, ISDB-T
  - Mobile TV: ATSC-M/H, CMMB, DMB, MediaFLO, DVB-H, 1/3-seg
  - Radio: (RDS)FM, HD Radio, Satellite radio
- Industry leading computational throughput:
  - 8 way Real/4-way Complex SIMD per cycle + 3 way VLIW
  - 16 18-bit MACs/cycle
  - Radix-4 FFT butterfly per cycle
  - 4 complex FIR taps per cycle
- Configuration option for Tensilica's Xtensa LX customizable processor: memories, accelerator interface, and extra instructions can be added as required
- Scalable cluster architecture from 1-8 processors

ConnX Baseband Engine Architecture



ConnX Baseband Engine Instruction Set

- **Rich baseline instruction set: up to 153 operations**
- **DSP instruction set: 285 operations in 3 VLIW slots**

Load/Stores ops:
- Addressing Modes:
  - offset
  - offset-update
  - Index
  - index-update
  - circular
  - bit-reversed
- Load 16b/32b scalars and vectors
- Store 16b/32b scalar, vectors, transposed
- Load/store unaligned and masked delivers full bandwidth loads and stores with unaligned data

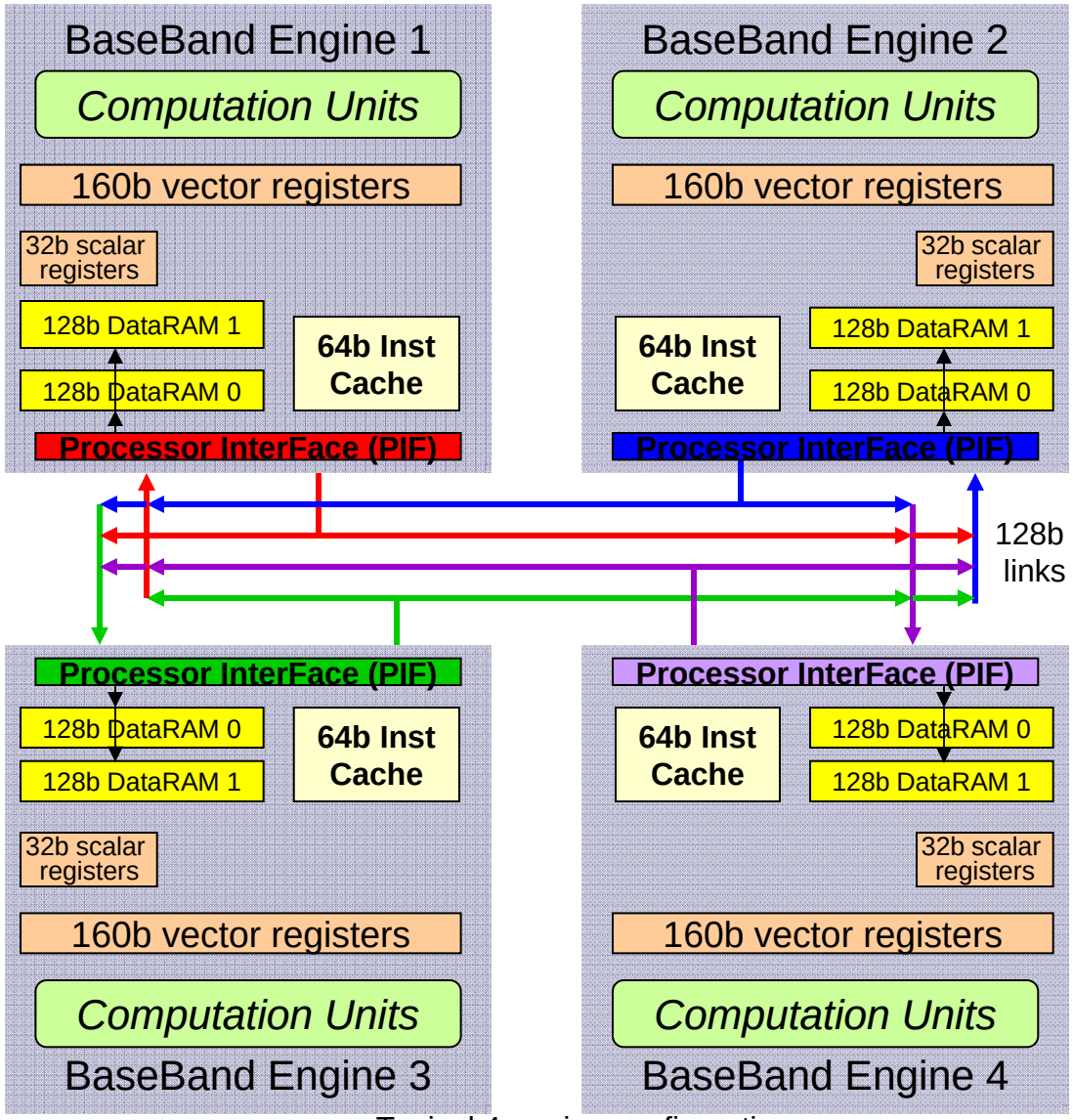
Multiply ops:
- Complex and scalar 18bx18b multiplies
- Multiply, multiply-round, multiply-add, multiply-subtract
- Multiply complex conjugate
- Magnitude-squared of complex
- Full precision and saturated/rounded outputs
- Up to 16 multiplies per operations
- FIR-optimized multiply-add

ALU ops:
- 20b/40b extended precision in 160b vectors
- Full arithmetic, logical and shift with saturation operations
- SIMD boolean setting for compares
- Ops: ABS, ADD, AND, ASUB, CLAMPS EQ, XOR, LE , MAX, MAXB, MAXU, MIN, MINB, NAND, NEG, NSA, NSAU, OR, PACK, SLL, SLLI, SLLV, SRA SRAI, RADD SUB

Other ops:
- Direct support for single-cycle radix-2 and radix-4 butterfly operations
- 8-way SIMD integer and fractional divide
- 4-way SIMD reciprocal square root
- Arbitrary permutation and selection from vector pairs
- Zero-overhead looping
- Conditional vector moves

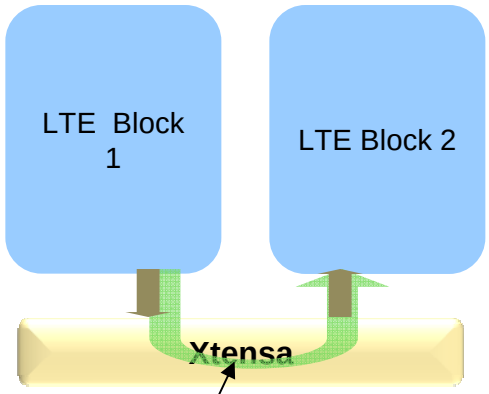
2-8 ConnX Baseband Engines Form Advanced Processor Cluster

- 2-8 Baseband Engines form powerful shared memory baseband processor platform
- 8 engine cluster:
  - 128 MACs/cycle
  - Up to 640 ops/cycle
  - 880K 2048pt complex FFTs per second
- Distributed DataRAM space visible to all engines accessed across 128b pipelined interconnect
- Write-buffered interface allows aggregate 120GB/s processor load/store data bandwidth and 60GB/s inter-engine data bandwidth (at 500MHz)
- Native SystemC modeling of multi-engine processors, including cycle-accurate and fast "Turbo" mode bit-accurate simulation



Typical 4-engine configuration

Optimized processing in LTE signal path

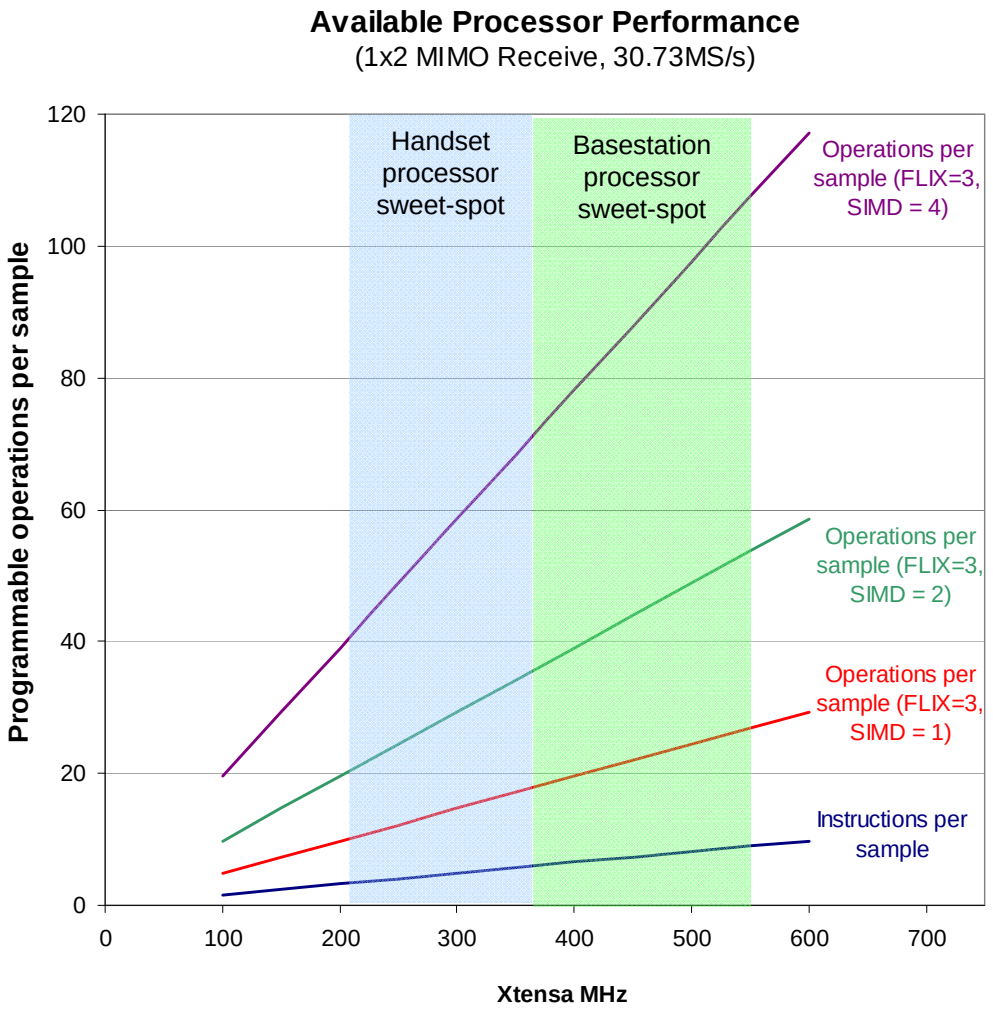


LTE Receive data stream:
- 1x2 MIMO
- 30.73M samples/sec
- Each sample is [I,Q] pair

- Typical Xtensa processors implement 1 to 3 VLIW (FLIX) slots per instruction
- Typical Xtensa processors implement 1 to 4-way SIMD for complex operands

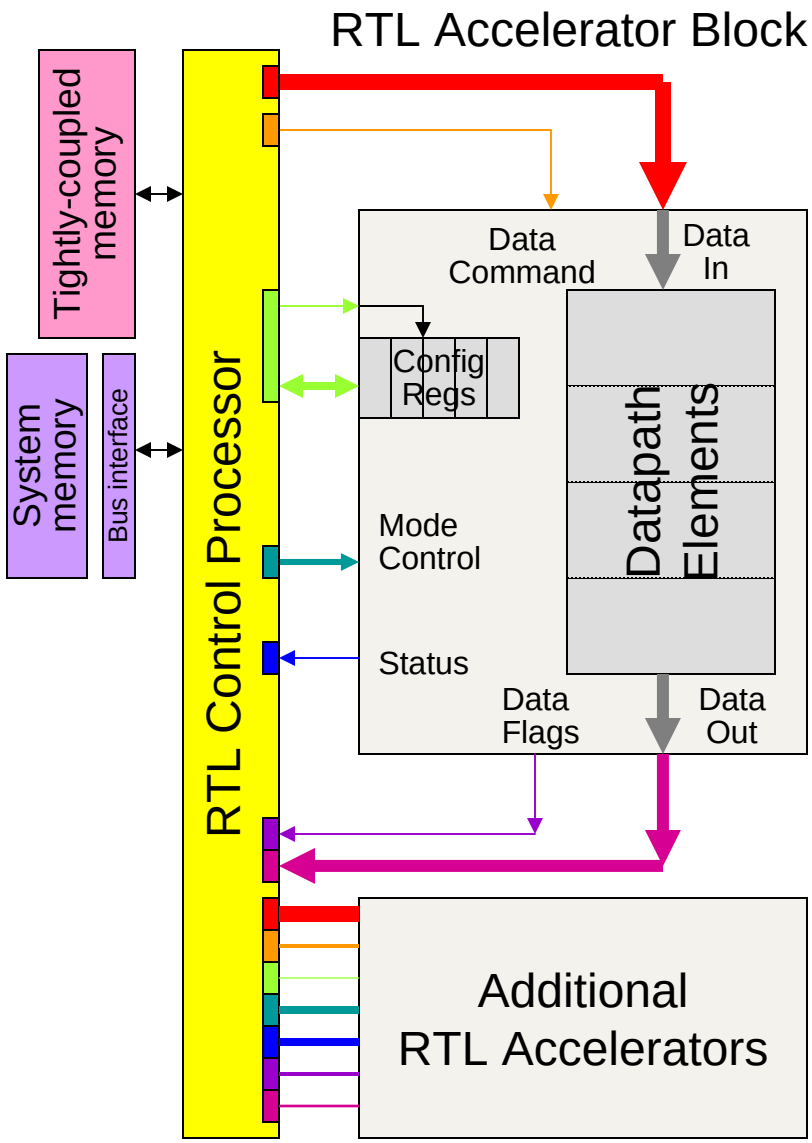
How many useful operations can you perform on the data stream (per Xtensa)?

20-50 ops per sample at modest MHz



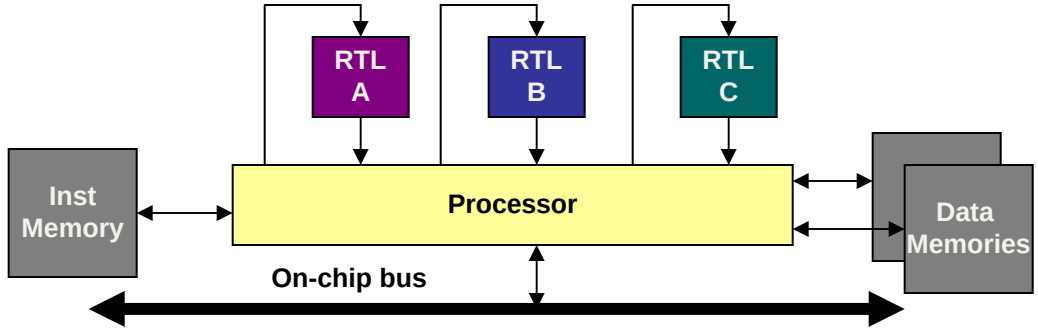
Enhanced RTL Integration

- RTL Accelerator blocks have a wide variety of interface types and widths
  - Data input stream
  - Data output stream
  - Data command inputs
  - Data output flags
  - Configuration registers
  - Mode control
  - Status outputs
- Extensible processor matches RTL interface type and width (to 1024b)
  - Output queues
  - Input queues
  - Read only lookups
  - Read/write lookups
  - Import wires
  - Export states
- Full software support for interfaces:
  - Mapped to instructions and compiler
  - Modeling in high-level and RTL tools
  - Visible to source debugger
- Processor performs “smart DMA” for RTL data transfers
- Multiple RTL blocks controlled by one processor



Direct Control of Multiple RTL Blocks Example

Input Data	Output Data	Control word	Cmd word
128b	128b	32b	3b



Interface Declaration:

```
regfile DR 128 16 d
lookup LUA {`128+32+8`, Mstage} {`128`, Mstage+3}
state ModeA 32 add_read_write
lookup LUB {`128+32+8`, Mstage} {`128`, Mstage +3}}
state ModeB 32 add_read_write
lookup LUC {`128+32+8`, Mstage} {`128`, Mstage +3}}
state ModeC 32 add_read_write
format f64 64 {l_slot,s_slot,a_slot,b_slot,c_slot}
table cmdA 8 8 {0, 1, 2, 3, 4, 5, 6, 7}
table cmdB 8 8 {0, 1, 2, 3, 4, 5, 6, 7}
table cmdC 8 8 {0, 1, 2, 3, 4, 5, 6, 7}
slot_opcodes l_slot {LDIU}
slot_opcodes s_slot {SDIU}
slot_opcodes a_slot {LUOpA}
slot_opcodes b_slot {LUOpB}
slot_opcodes c_slot {LUOpC}
operation LUOpA {out DR do, in DR di, in cmdA cmd}
   {in ModeA, out LUA_Out, in LUA_In} {
  assign LUA_Out = {cmd,ModeA,di};
  assign do = LUA_In;}
operation LUOpB {out DR do, in DR di, in cmdB cmd}
   {in ModeB, out LUB_Out, in LUB_In} {
  assign LUB_Out = {cmd,ModeB,di};
  assign do = LUB_In;}
operation LUOpC {out DR do, in DR di, in cmdC cmd}
   {in ModeC, out LUC_Out, in LUC_In} {
  assign LUC_Out = {cmd,ModeC,di};
  assign do = LUC_In;}
```

5-slot VLIW Instruction streams data from memory, through 3 RTL data-paths and back to memory:

Load → Reg[]	Reg[] → Store	Reg[] → RTL_A(Cmd) → Reg[]	Reg[] → RTL_B(Cmd) → Reg[]	Reg[] → RTL_C(Cmd) → Reg[]

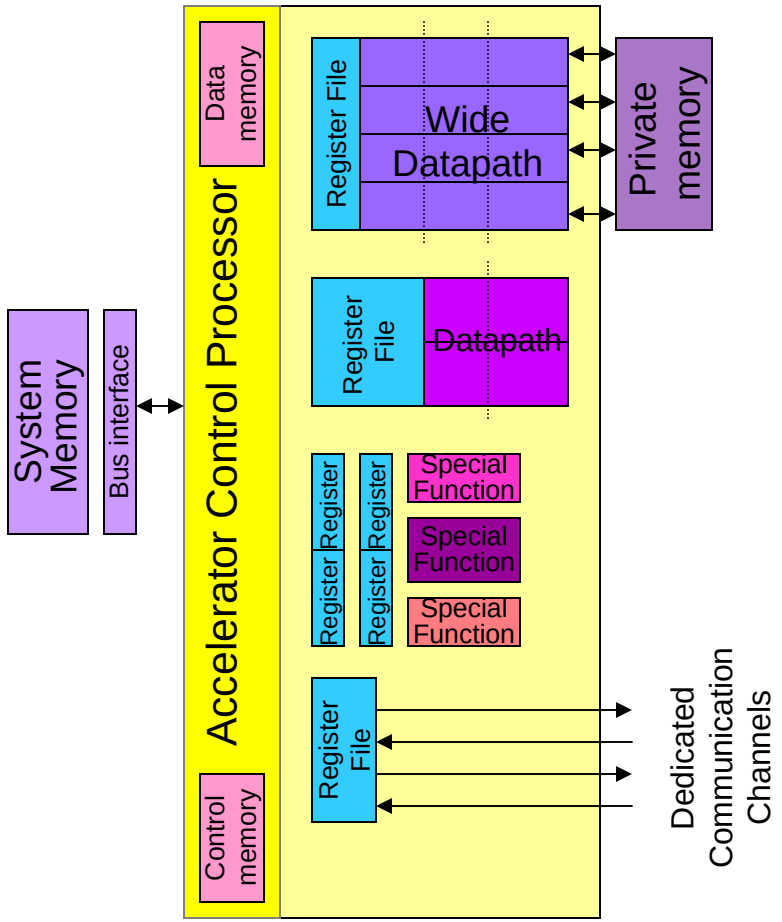
Typical operations per cycle:
1 128b read from memory
1 128b operation through RTL A
1 128b operation through RTL B
1 128b operation through RTL C
1 128b write to memory

Full Accelerator Integration

- For new functions, integrated acceleration is easy and efficient
- Your proprietary accelerators are fully integrated into instruction set and software tools for each processor
- Add any number of new data pipelines, registers, memories, inter-processor channels – up to 100s of ops per cycle
- Tensilica Instruction Extension (TIE) format typically 10x more concise than Verilog
- The cycle-by-cycle behavior of each accelerator written in standard C and modeled in fast cycle-accurate simulator
- Use multiple small processor for additional throughput on complex sets of tasks

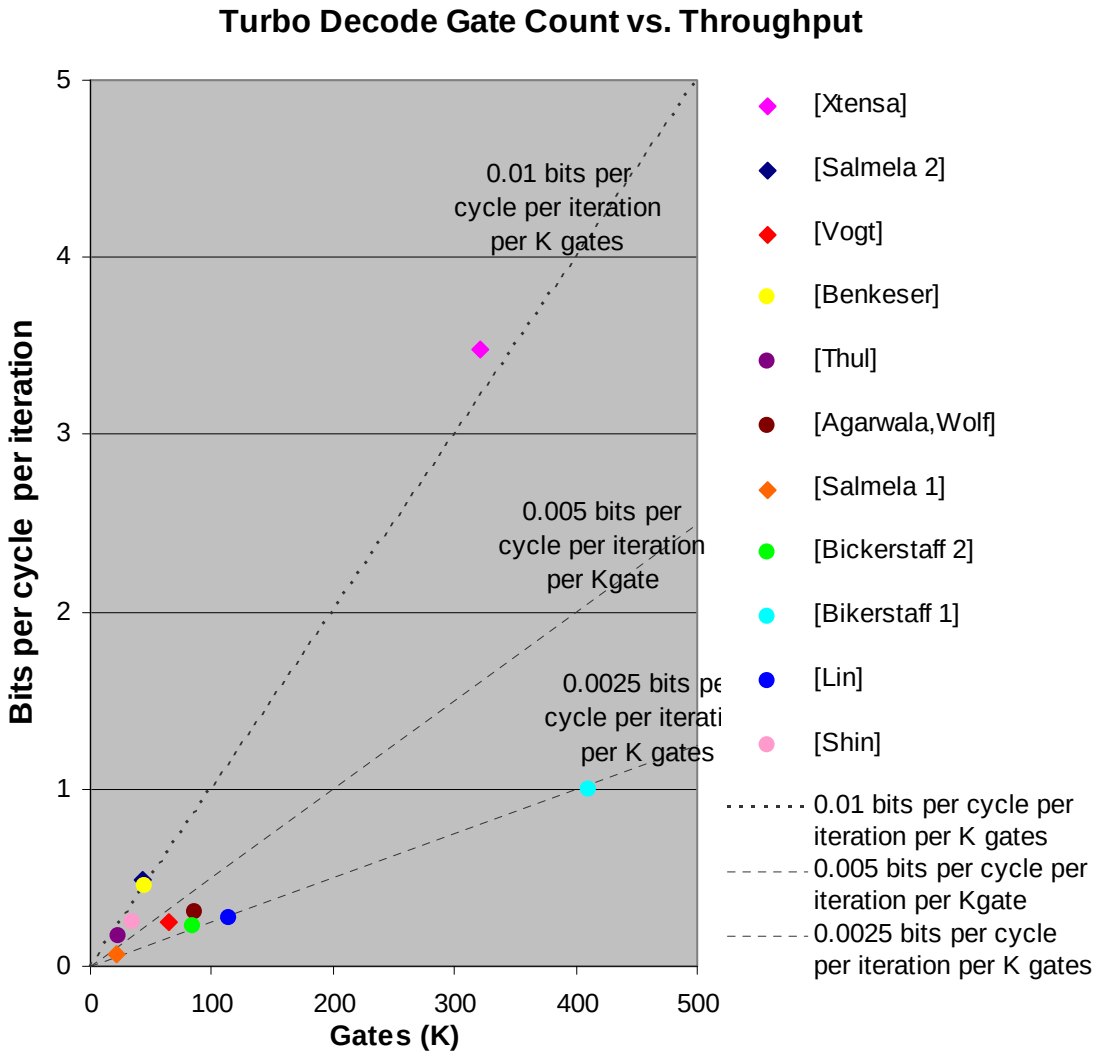
No other processor family offers this combination

1. **Performance**: special operations, VLIW, SIMD
2. **Efficiency**: low overhead in gates and power
3. **Automation**: Learn simple TIE format in hours
4. **Programmability**: all accelerators controlled in C



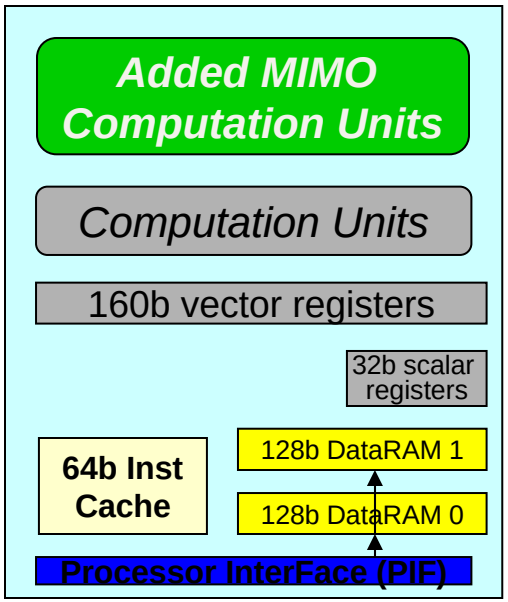
Specialized Processor as Efficient as RTL
Tensilica Turbo Engine

- LTE requires high-data-rate Turbo decoding: >5000 ops per bit
- Turbo decoding closely tied to HARQ error processing: favors programmability
- Xtensa's instruction extensions, wide data-paths and multiple wide memories enable efficient programmable Turbo Engine
- Method:
  - 2 blocks in parallel
  - 8 parallel windows per block
  - each window updates 8 states per cycle
  - 1 cycle for each forward and backward pass (4 cycles per iteration) per bit
- Implementation:
  - Dual 576b wide state memories
  - Single 128b interleave memory
  - Dual 128b load/store interface for main memory
  - 320K gates
  - 6 stage computation pipeline
  - 4 cycles per bit per iteration/16 bits in parallel = 0.25 cycles per bit per iteration.



MIMO Decoder for 3GPP LTE

- Worst-case system
  - 4x4 Spatial Multiplexing
  - 20MHz Channel with 64QAM Modulation
- Algorithm
  - SIC (Successive Interference Cancellation) LMMSE-SQRD (Sorted QR)
  - Sorted QR implemented with Givens Rotations
  - LLR Module
- Implementation Platform
  - Instruction extensions for ConnX Baseband Engine: (120K gates)
- For 10MHz uplink channel, channel estimation also fits in 350MHz budget



Advances in Processor Foundations

- Smaller: <15K gates for 32b RISC baseline (5-stage pipe)
- Faster: 10-15% performance gain across all configurations – to 1GHz (45GS)
- Lower power: Pipeline-specific clock-gating + power-optimized VLSI flow = ~3x active power reduction over non-clock-gated design
- More efficient register files: optimal port sharing across VLIW slots
- Huge step in memory system capability: Up to 32 ports into private memories to 1024b wide (25Tbps @ 750MHz)
- Tool advances:
  - Compiler operator overloading for user-defined types
  - Direct instantiation of fast processor model into standard Verilog simulation
  - Multi-processor application/configuration energy analysis
  - Expanded multi-processor library

Tensilica LTE PHY Development Architecture

ConnX *Baseband DSPs + Accelerators for Minimum Power /Area*