tensilica

The Engine of SOC Design

The Reinvention of the Microprocessor

June 2006

Chris Rowen
President and CEO
Tensilica, Inc.

Outline

- Basic changes in the electronics environment
- Moore Law gives us less
- The Old World: A short history of the microprocessor
- The New World:
  - A new processor class
  - Concurrency and multiple processors
  - Design flow for multiple processor embedded systems
- Wrap-up: Tensilica and the future of “system on chip”

Basic Changes in Industry Drive Innovation

Change #1: Products go mobile, connected, always-on, and media-rich

Change #2: Product complexity drives up development cost and risk

Change #3: Moore's Law drives lower silicon cost but not lower power

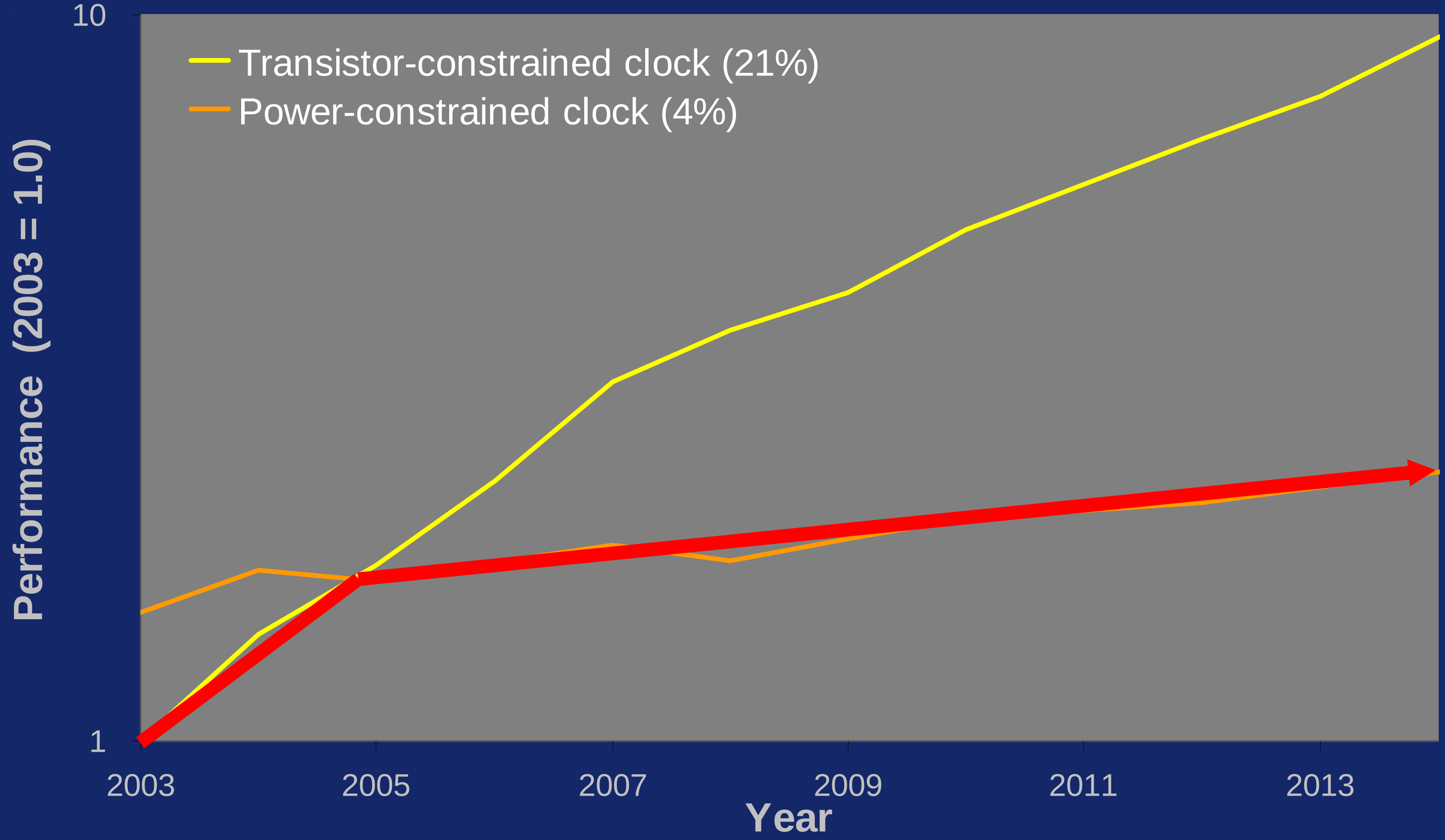




The Old World

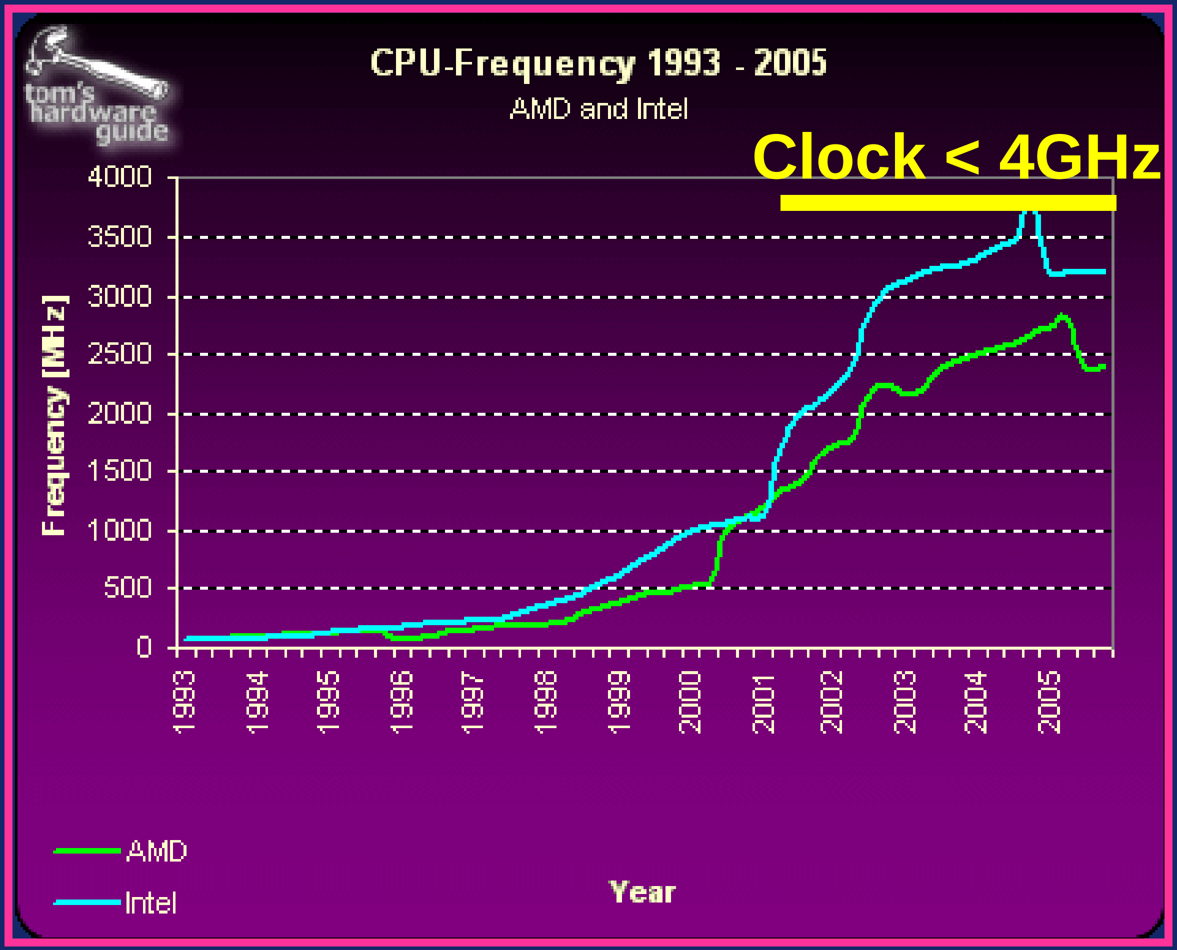
1. **Board-level using standard ICs, FPGA and ASIC**
   - Modest total hardware and software complexity
   - Pin-count-limited interfaces between sub-systems
   - Processor requires full IC
2. **Manual development of processor hardware and software tools forces one-size-fits-all processors**
   - Limitation of processor performance and energy efficiency
   - Long development cycles for new programmable platforms
3. **Deep sub-micron silicon design limited to advanced chip-makers**
   - High schedule and risk in hardwired functions

Inflexion Point in Clock Frequency



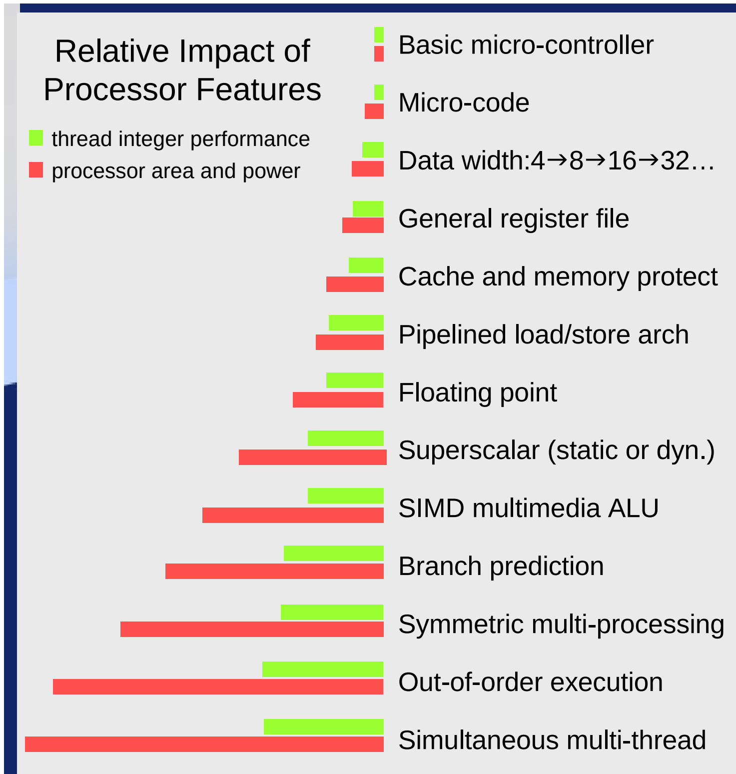
Based on International Technology Roadmap for Semiconductors 12/03

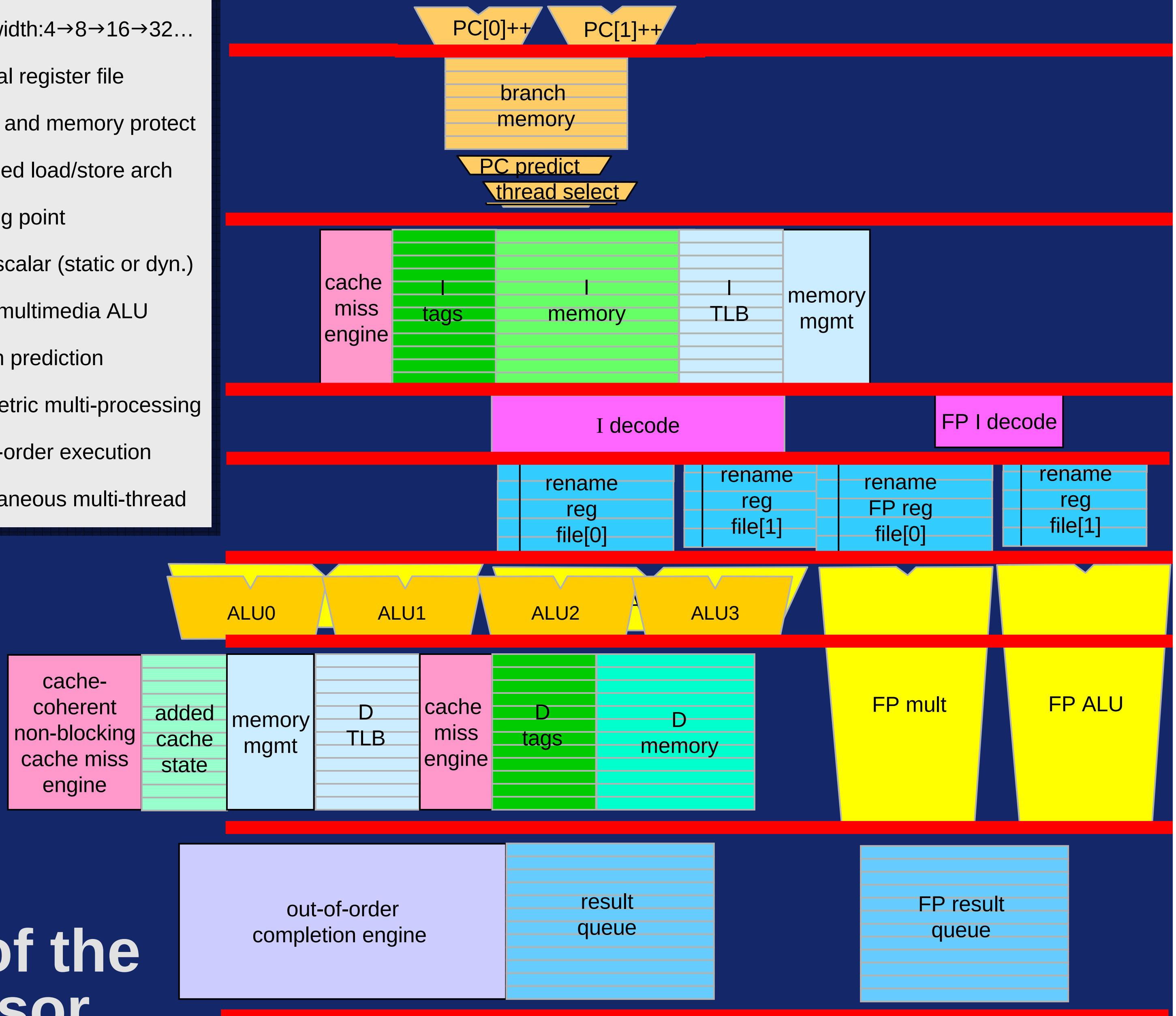
High-End Processors hit MHz ceiling



Basic Implications:

1. Get performance from better architecture instead of more MHz
2. Use multiple processors

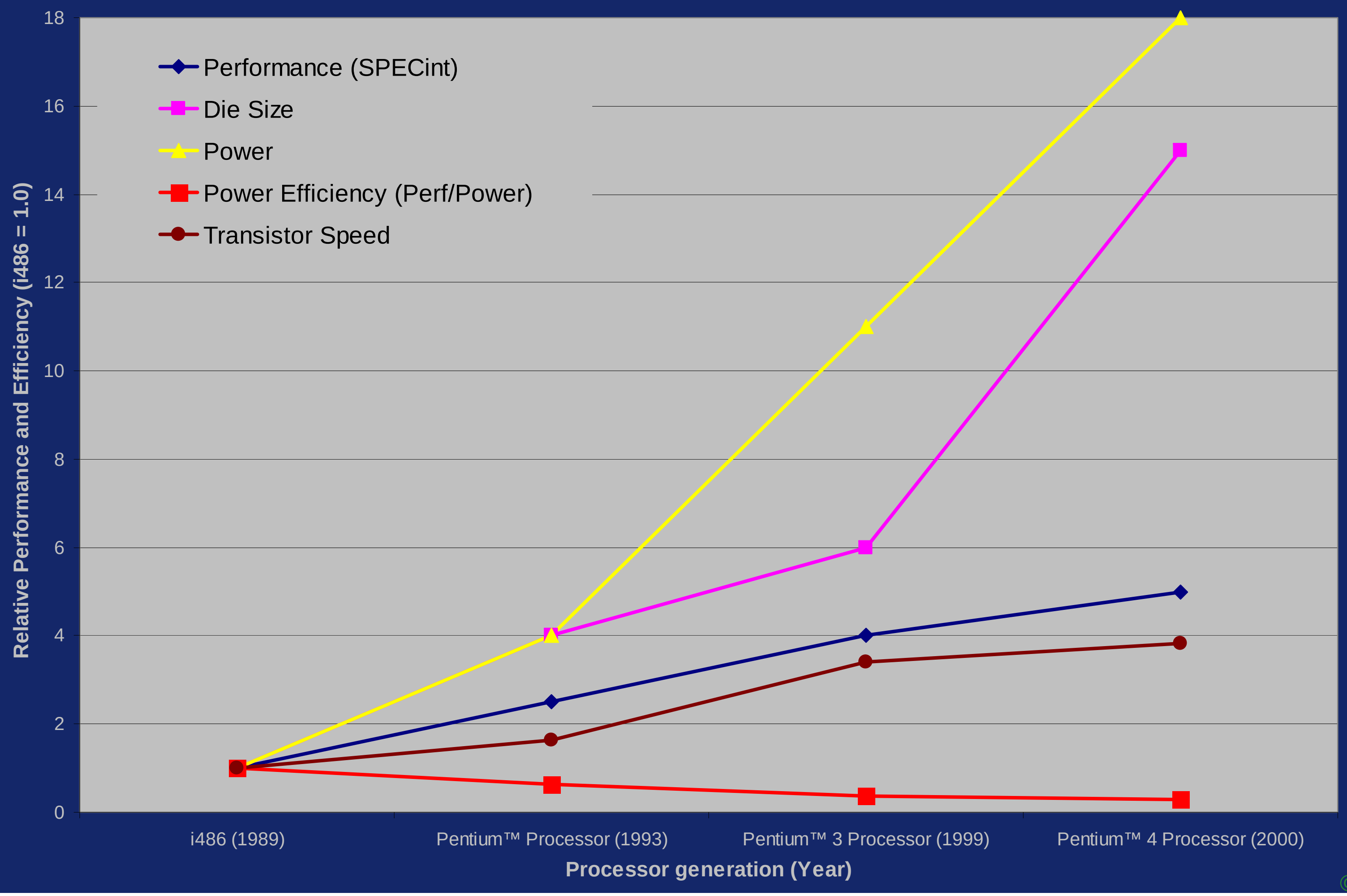




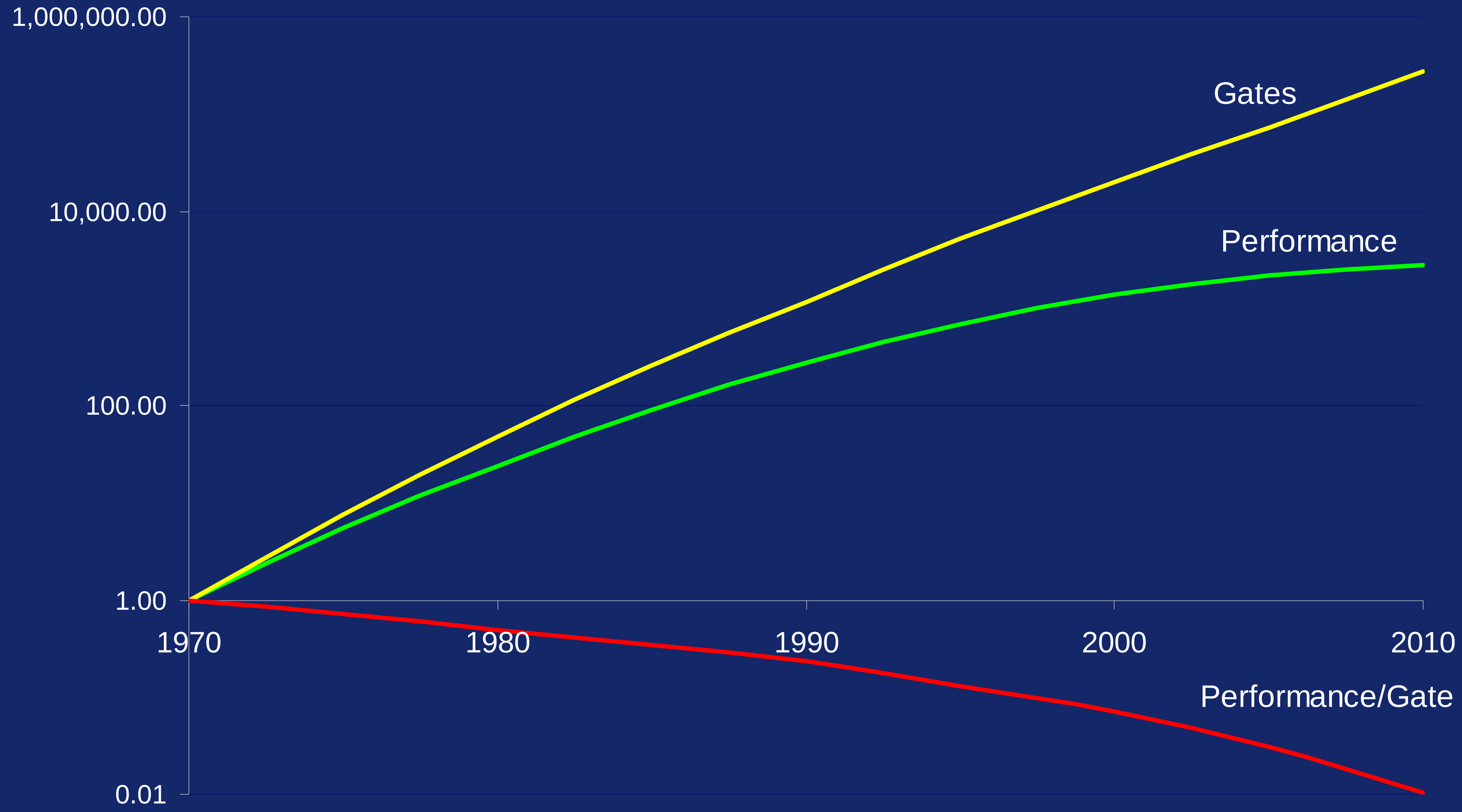
The history of the microprocessor

Intel's Own Assessment

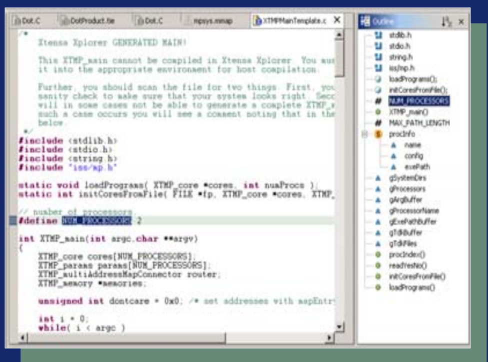
Power and Area increase more rapidly than Performance



Long-term uni-processor scaling trend



The New World
Automatic Processor Generation

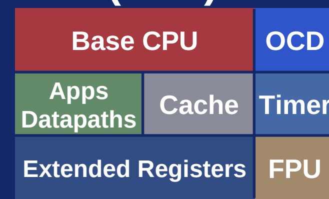


Processor configuration

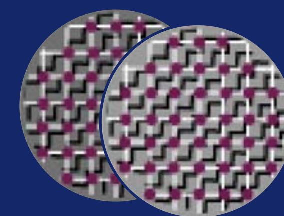
1. Select from menu
2. Automatic instruction discovery (XPRES Compiler)
3. Explicit instruction description (TIE)

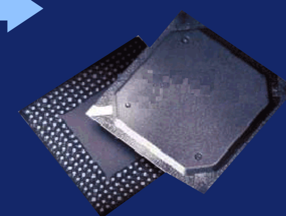
Tensilica Processor Generator

Application-optimized processor implementation (RTL)



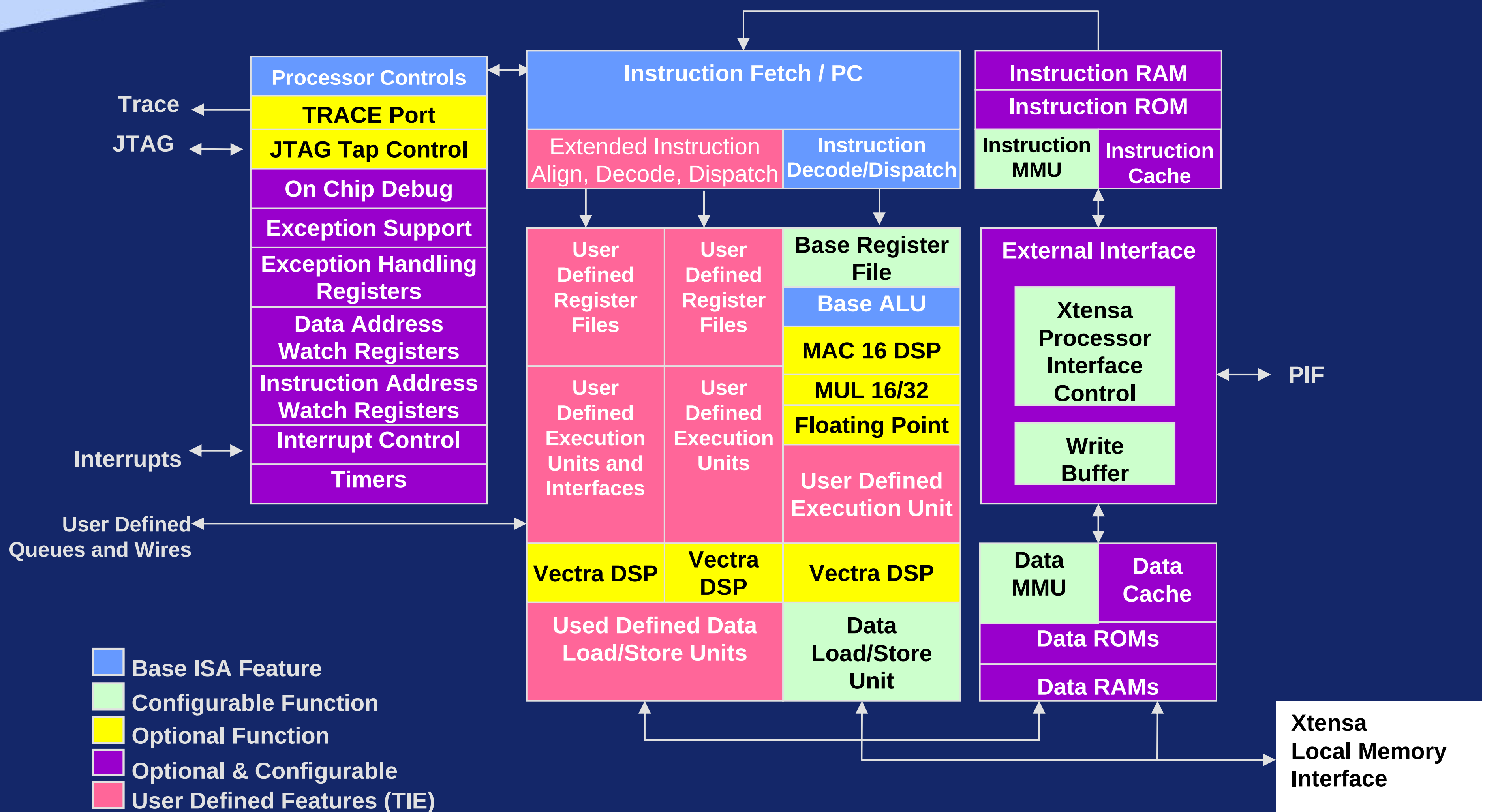
Tailored SW Tools: Compiler, debugger, simulators, OS ports





Build with any process in any fab

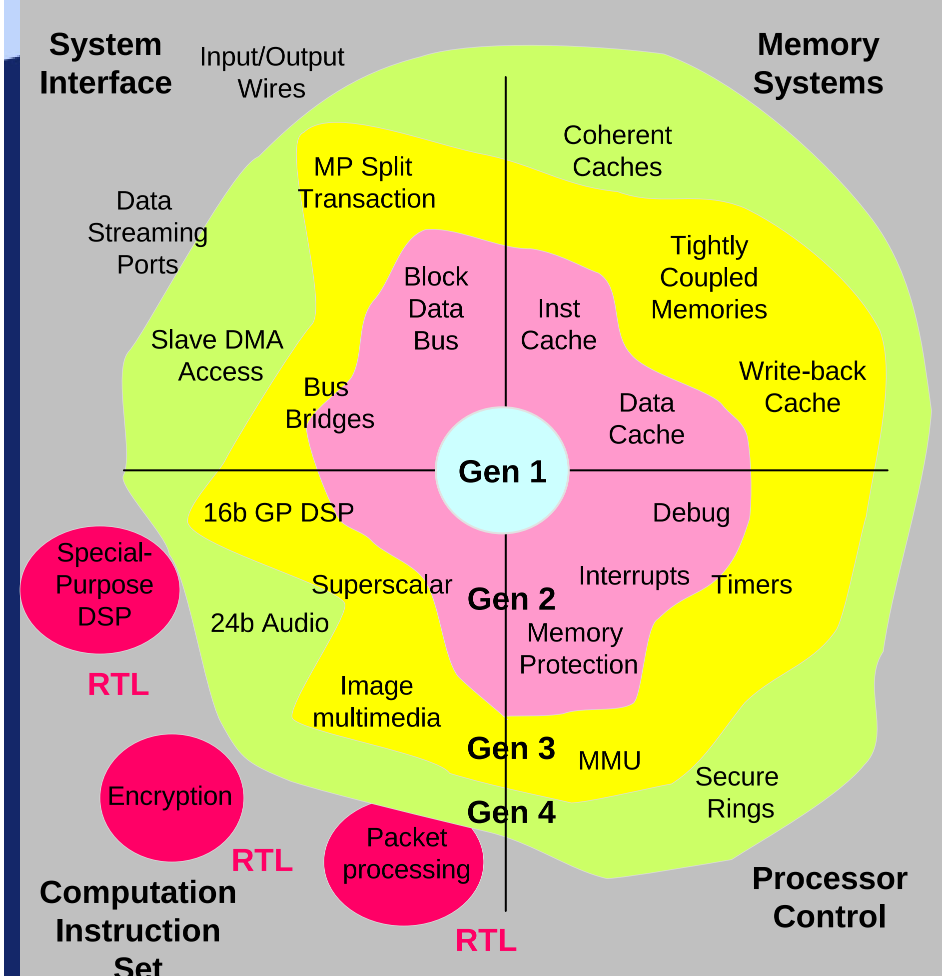
Build Almost Any Processor



Tensilica Confidential

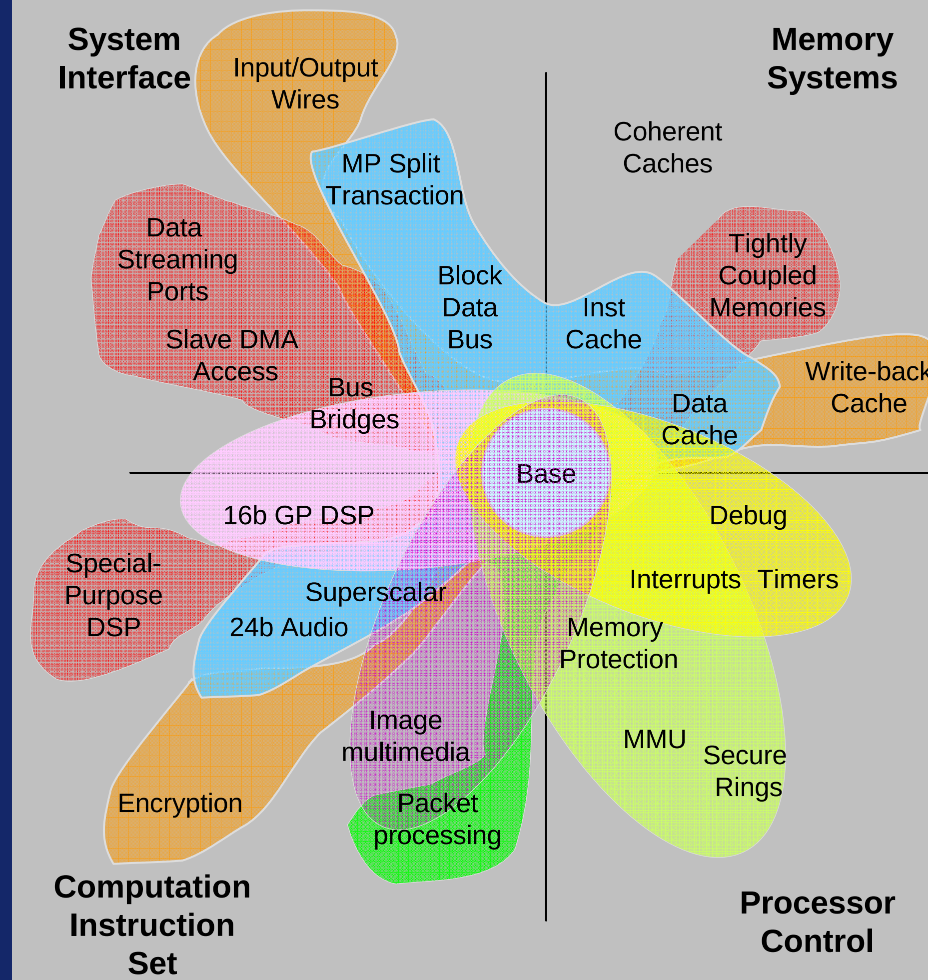
Two models of processor evolution

Traditional Processor Family



Time per variant: **years**

Configurable Processor Family



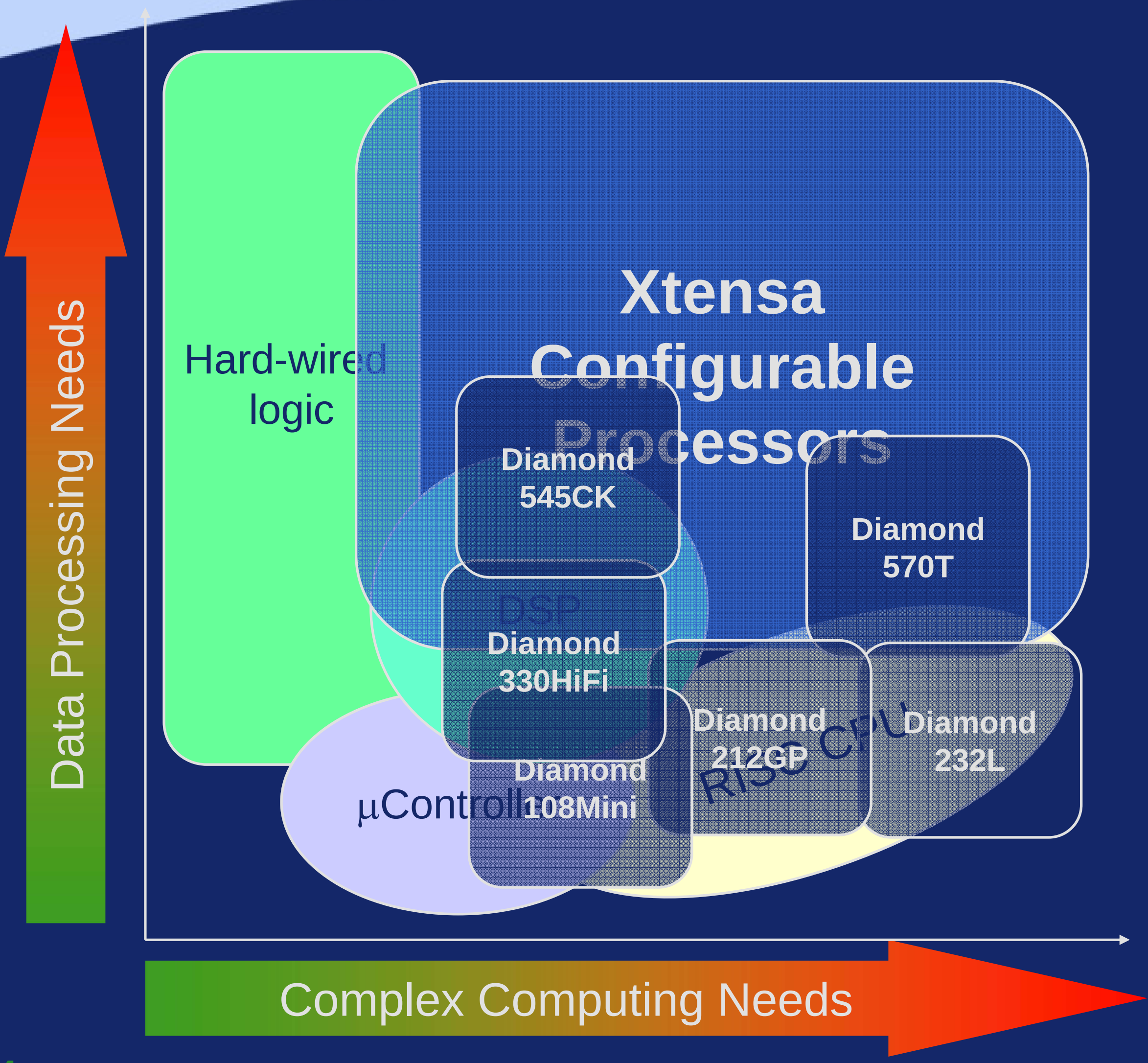
Time per variant: **days**

Area = silicon cost and power

Why Configurability Now?

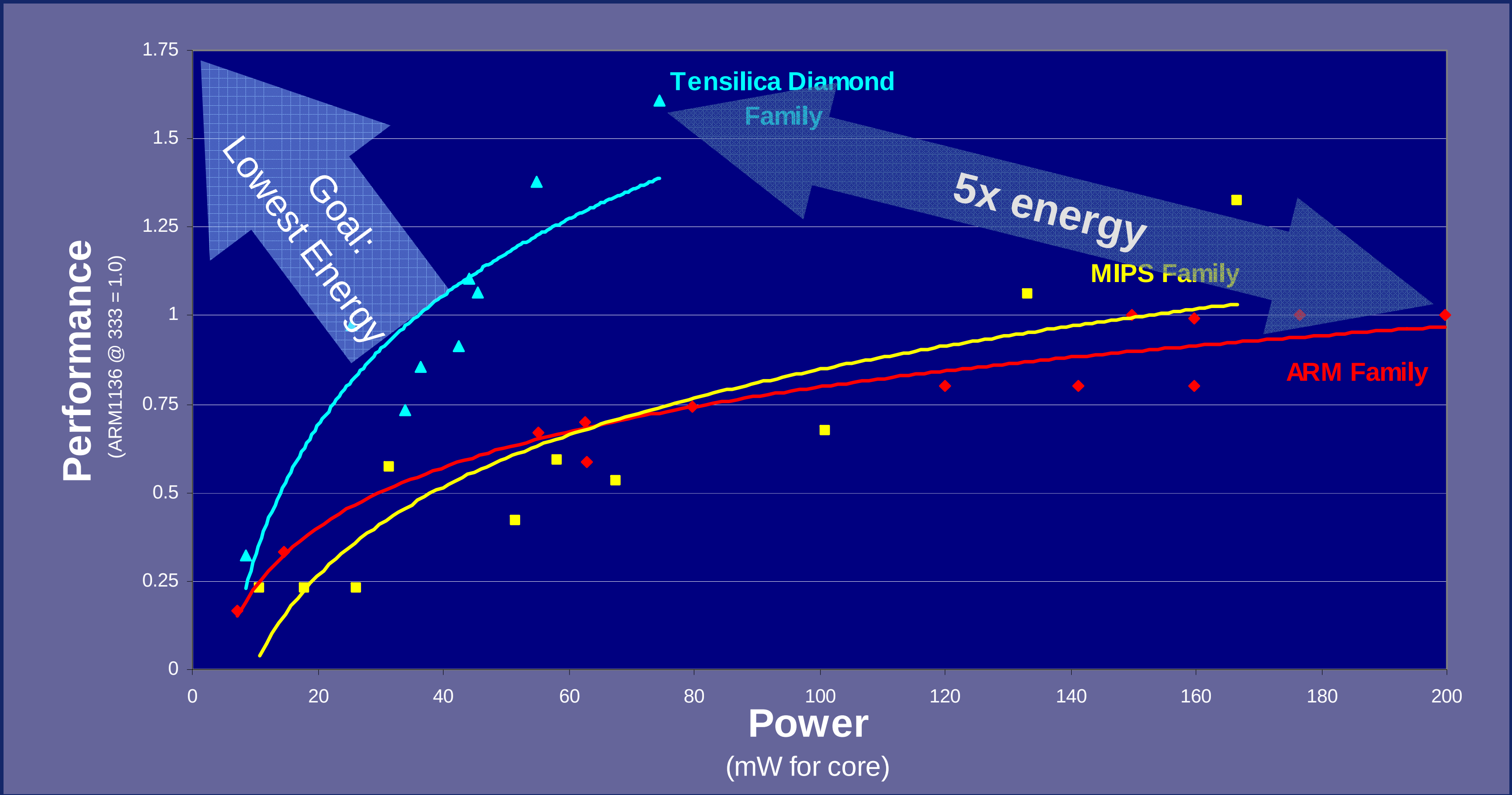
1. **Breakthrough on automatic generation of hardware + software**
   - Same quality and completeness of compilers, debuggers, development GUI, simulators, prototype emulation, OS ports as traditional one-size-fits all processor families
2. **Application-directed processor naturally fits into application-directed system-on-chip**
   - No extra mask cost in using unique processor variant
   - Smaller and lower power than generic processor
   - More flexible and programmable than hardwired accelerator blocks

Covering Breadth of SOC Processor Demands



- Tensilica is the largest supplier of configurable processors
- Diamond Standard cores introduced February 20:
  - Broadest line of controller, CPU and DSP cores in industry
  - Highest performance synthesizable general-purpose CPU
  - Highest performance DSP
  - Most complete low-power audio solution
- Tensilica spans general-purpose and application-specific processors and software

Processor Power and Performance
Standard Core Families



Performance on EEMBC benchmarks aggregate for Consumer, Telecom, Office, Network, based on ARM1136J-S (Freescale i.MX31), ARM1026EJ-S, Tensilica Diamond 570T, T1050 and T1030, MIPS 20K, NECVR5000). MIPS M4K, MIPS 4Ke, MIPS 4Ks, MIPS 24K, ARM 968E-S, ARM 966E-S, ARM926EJ-S, ARM7TDMI-S scaled by ratio of Dhrystone MIPS within architecture family. All power figures from vendor websites, 2/23/2006

Automatic Instruction Set Optimization

General-purpose processors all look alike (more or less)

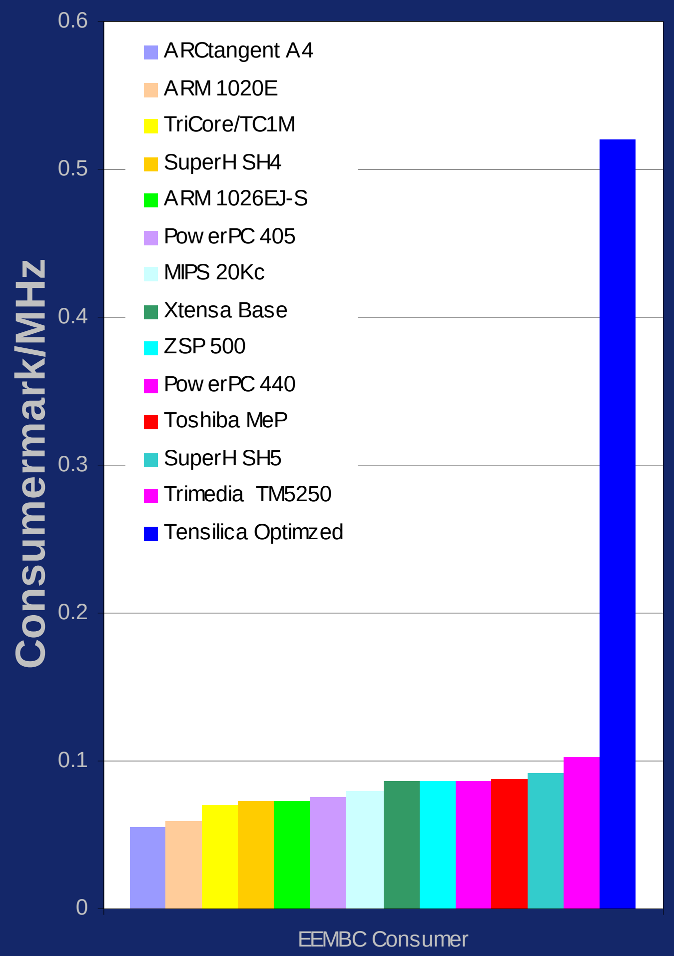
Automatic optimization from general-purpose C code with *XPRES Compiler*

- no intrinsic functions
- no assembly code
- no instruction set hardware definition coding

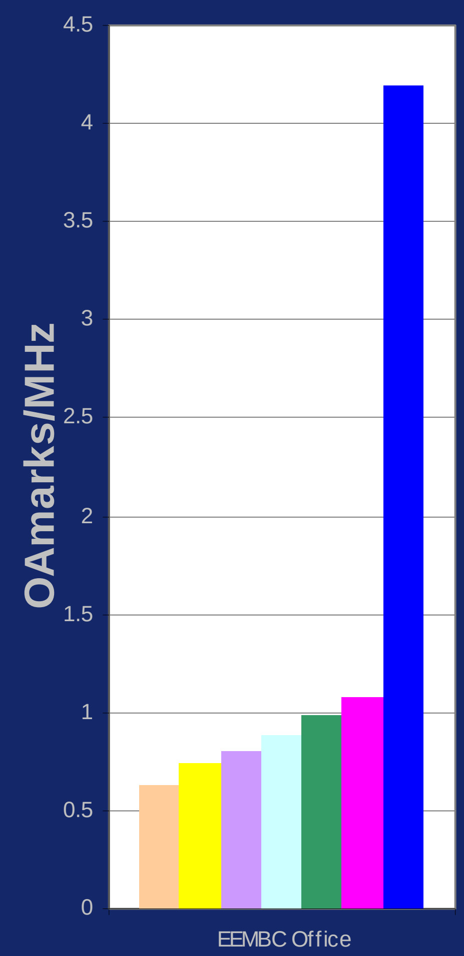
Results for all reported cores:

- Instruction set synthesis beats all other processors by 4-8x

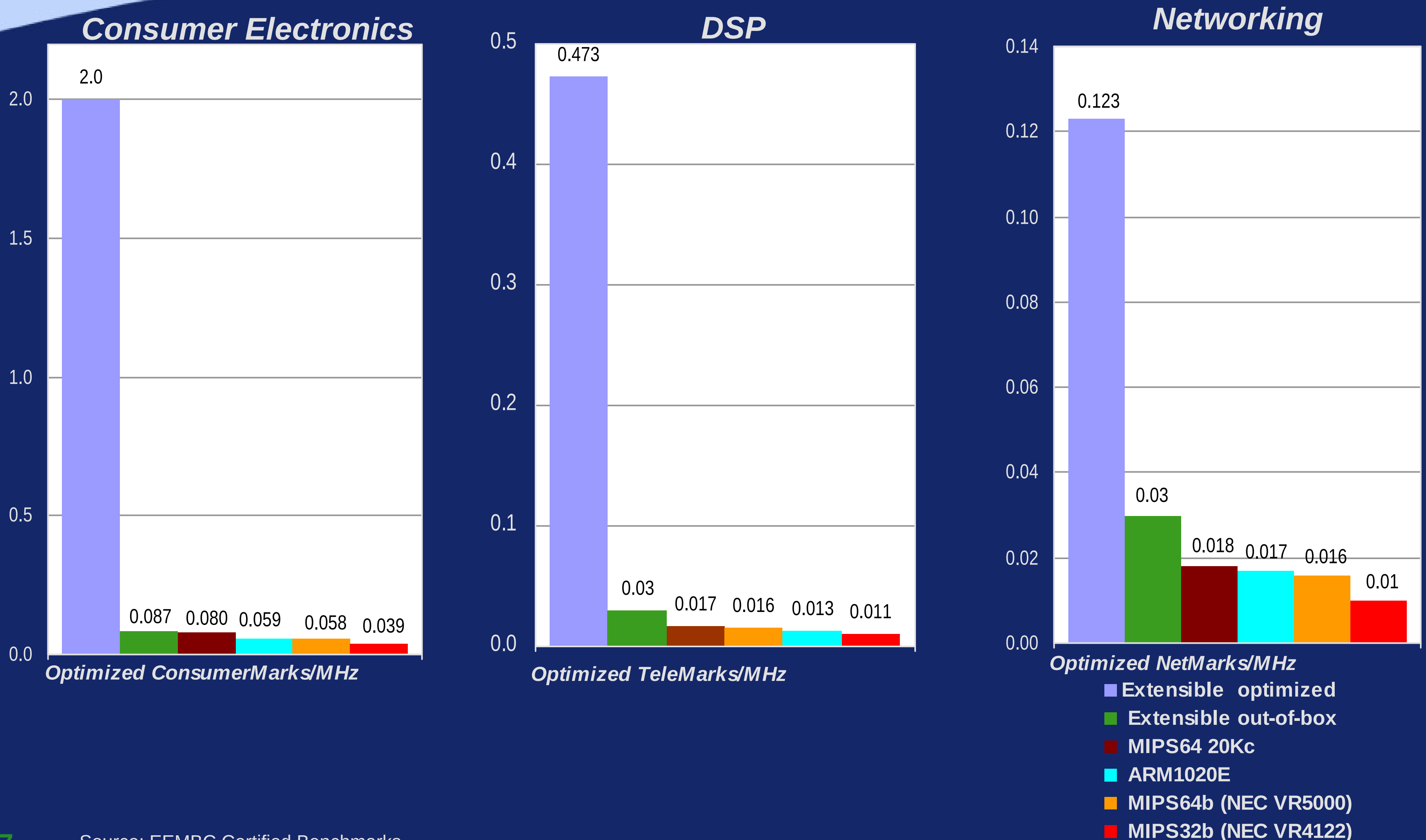
Consumer Electronics



Office Automation



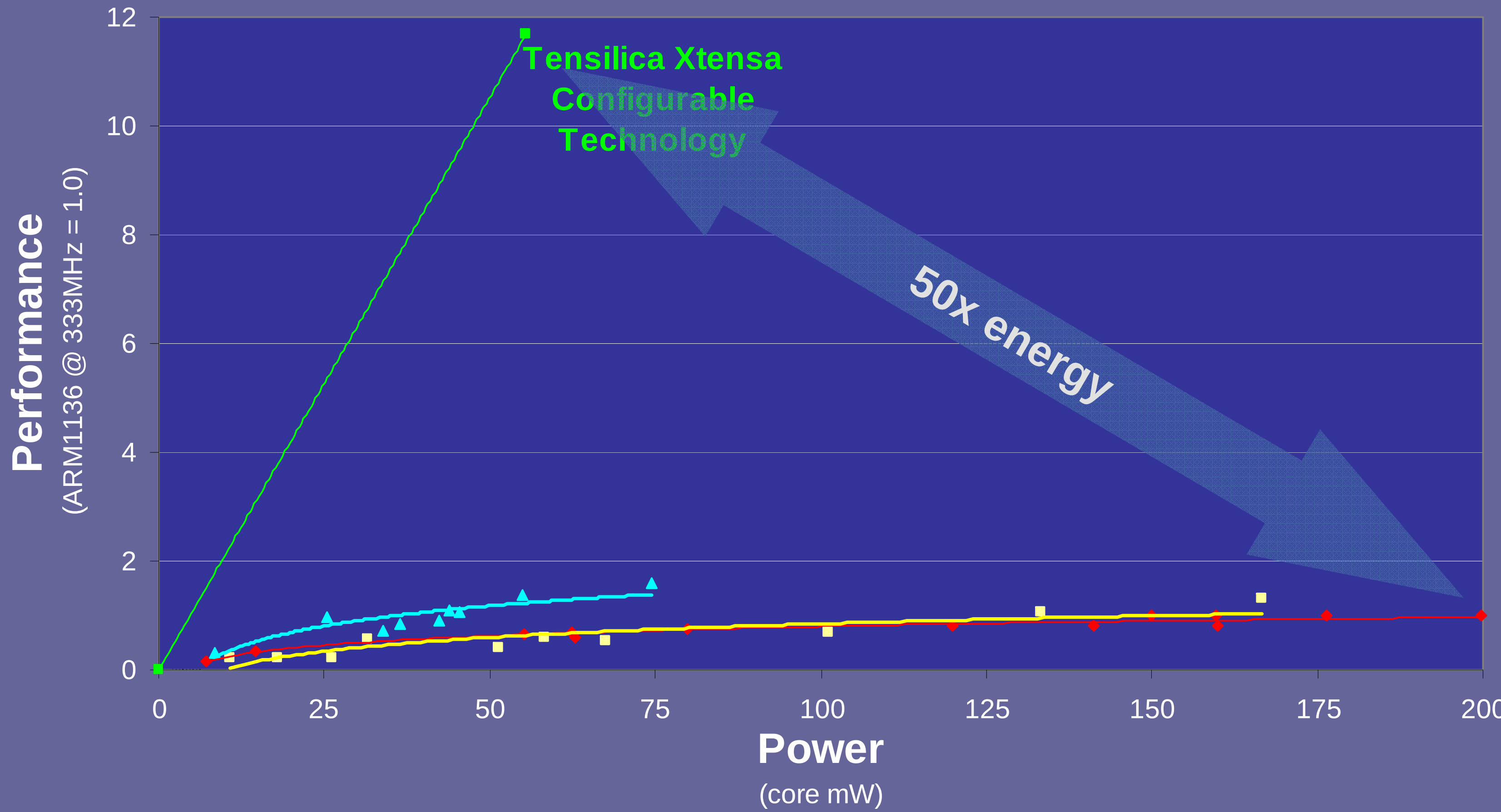
Explicit Instruction Set Optimization
Tensilica Instruction Extension



Source: EEMBC Certified Benchmarks

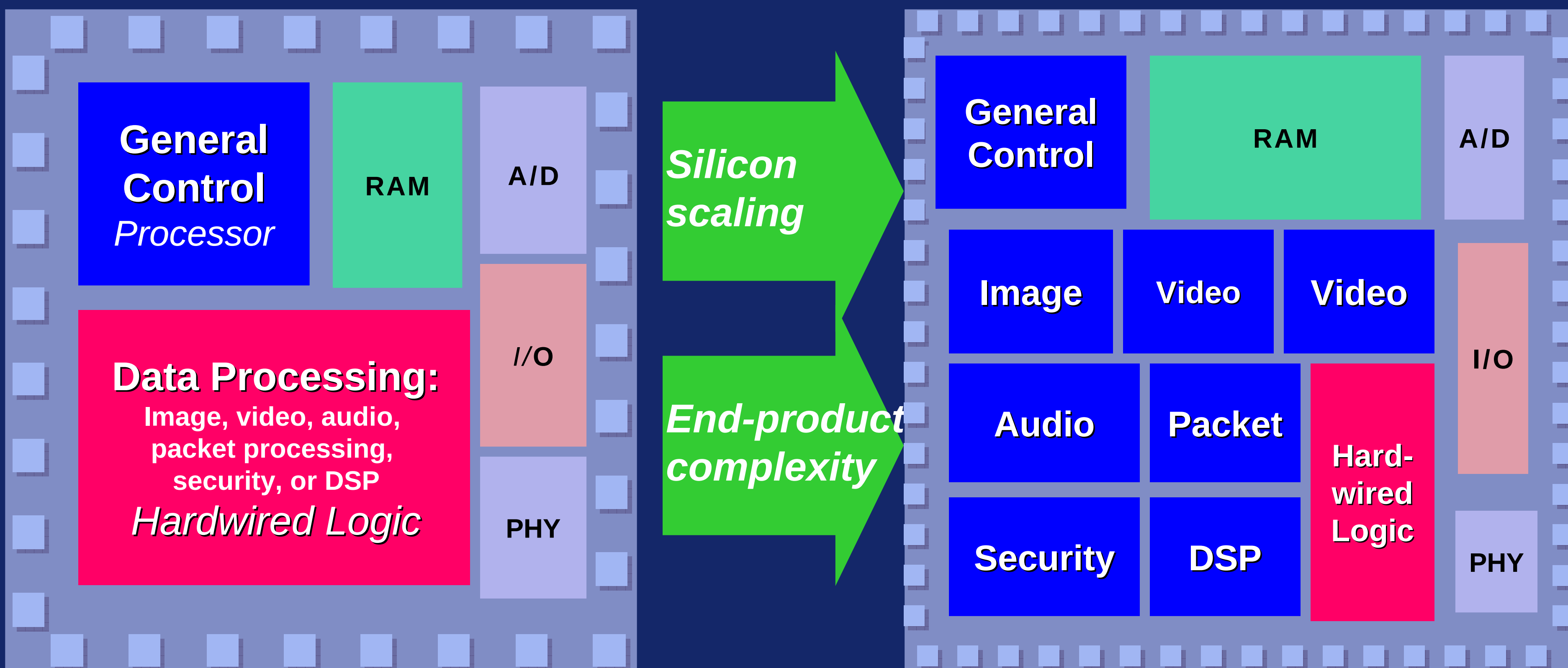
Processor Power and Performance

Xtensa Application-Specific Cores



Performance on EEMBC benchmarks aggregate for Consumer, Telecom, Office, Network, based on ARM1136J-S (Freescale i.MX31), ARM1026EJ-S, Tensilica Diamond 570T, T1050 and T1030, MIPS 20K, NECVR5000). MIPS M4K, MIPS 4Ke, MIPS 4Ks, MIPS 24K, ARM 968E-S, ARM 966E-S, ARM926EJ-S, ARM7TDMI-S scaled by ratio of Dhrystone MIPS within architecture family. All power figures from vendor websites, 2/23/2006

Basic Transition in Chip Design



Traditional View of "System on Chip"

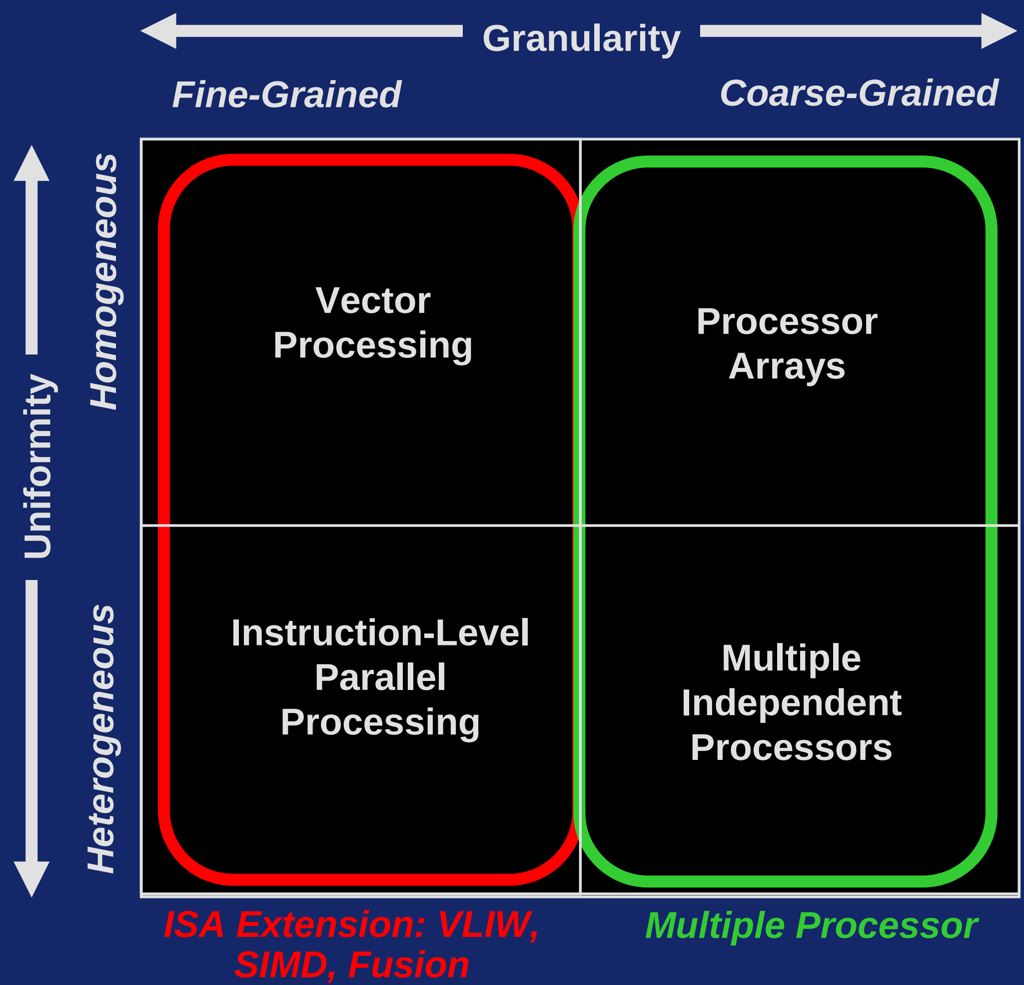
Multiple Processor SOC

The New World
Concurrency with Multiple Processors

- *Moore's Law scaling allows many more functions per chip*
- *Most applications have multiple forms of available concurrency*
- *Multiple concurrent processors much lower energy than a big processor*

Examples:

- *Dual core x86*
- *IBM/Sony Cell 1+8 Processor*
- *Cisco CRS-1 188 SPP*



Low Power and Multiple Processors

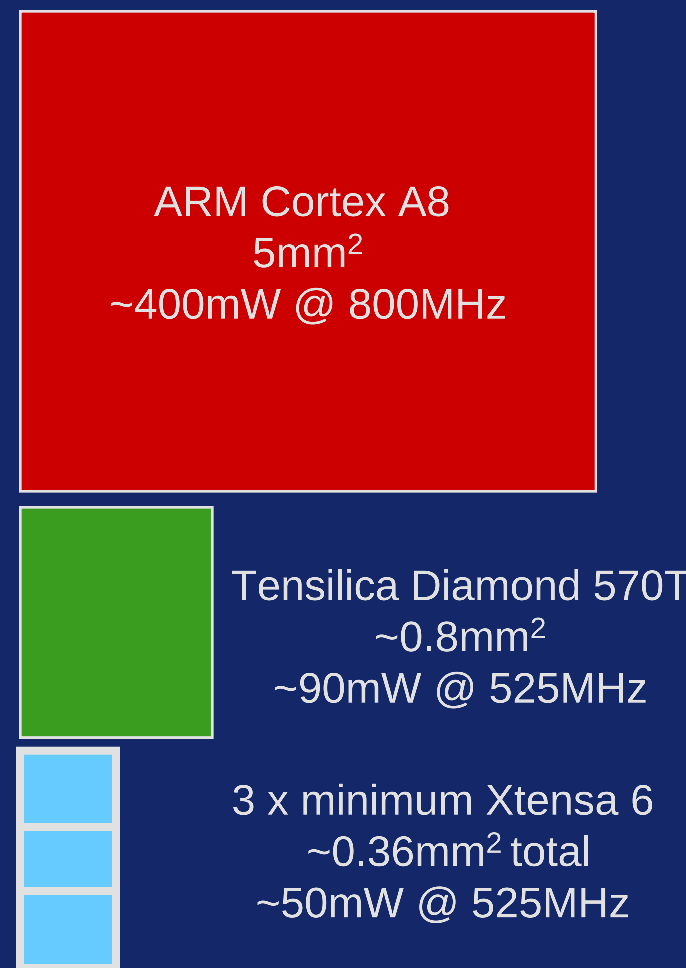
Why MP?

- Real systems implement multiple concurrent functions
- Demand for flexibility turns multiple hard-wired blocks into multiple processors

MP from the ground up

- MP architecture
- MP SW development tools
- High bandwidth/low latency
- Novel MP interconnects

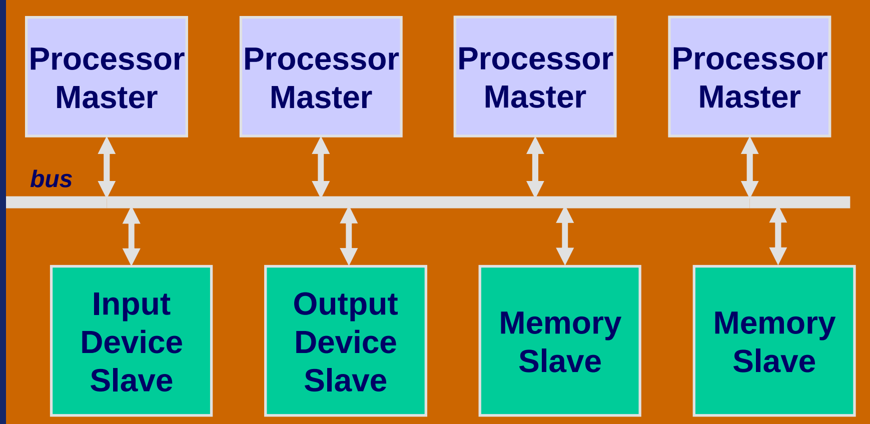
Low power and high performance



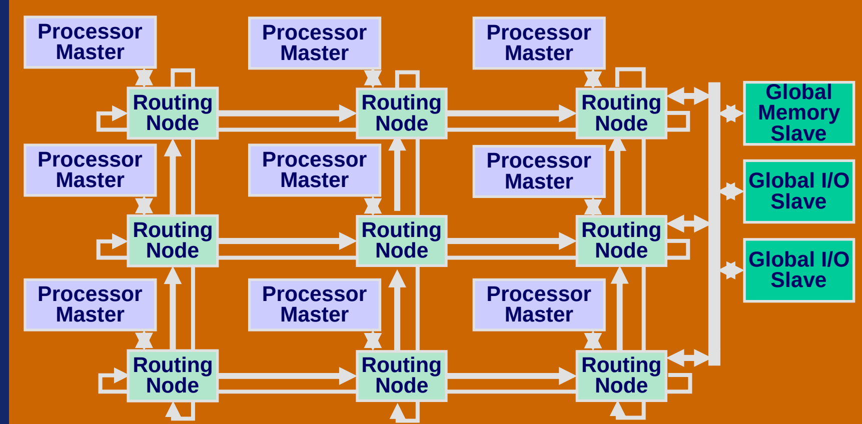
Keys to Efficient MP

Flexible range of topologies

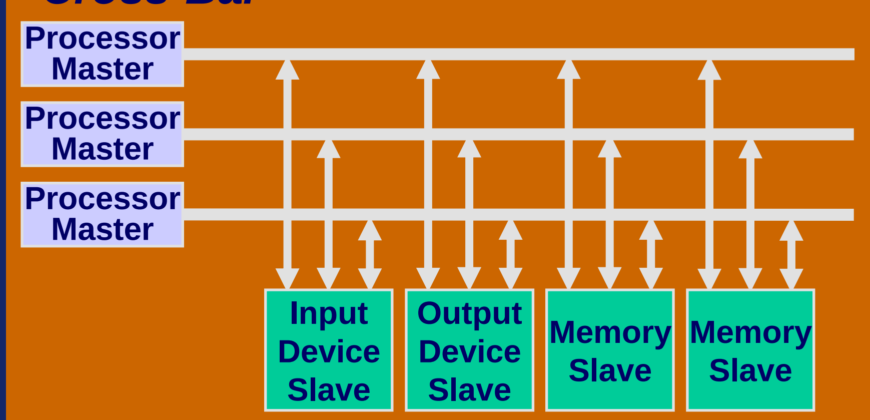
Shared Bus



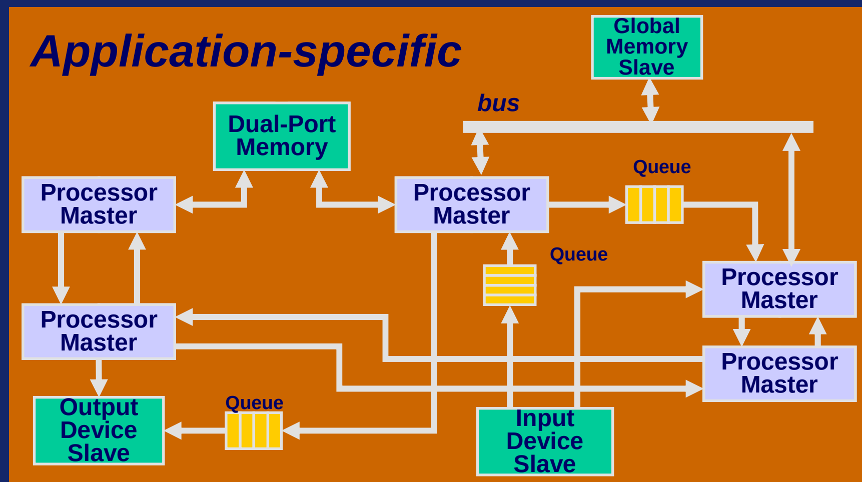
On-chip Routing Network



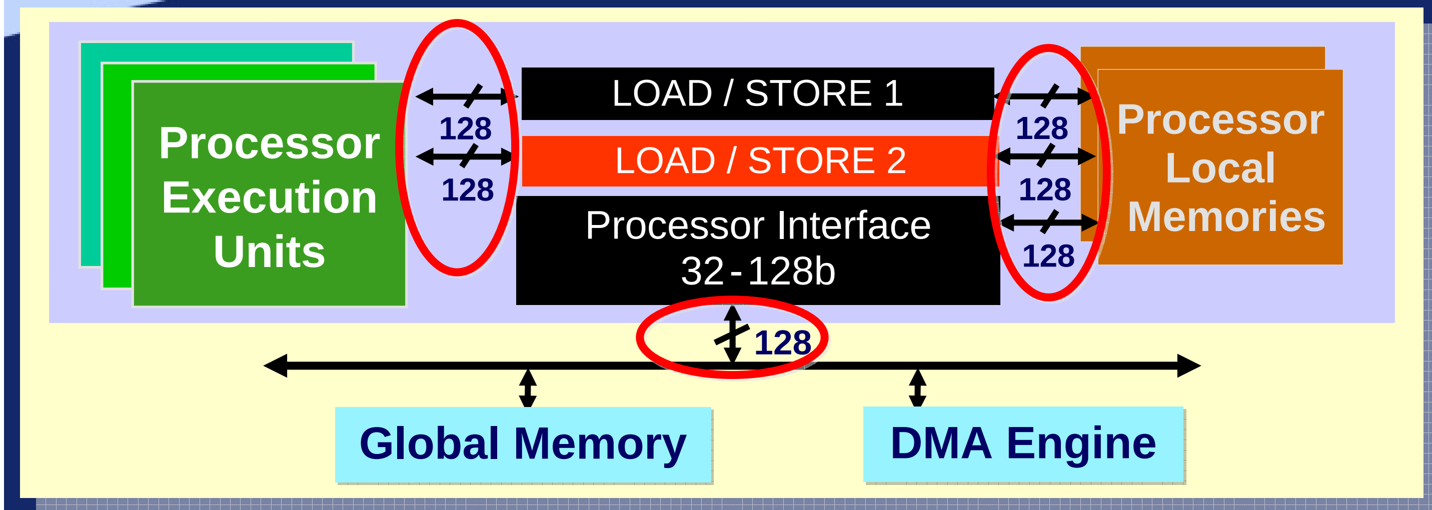
Cross-Bar



Application-specific

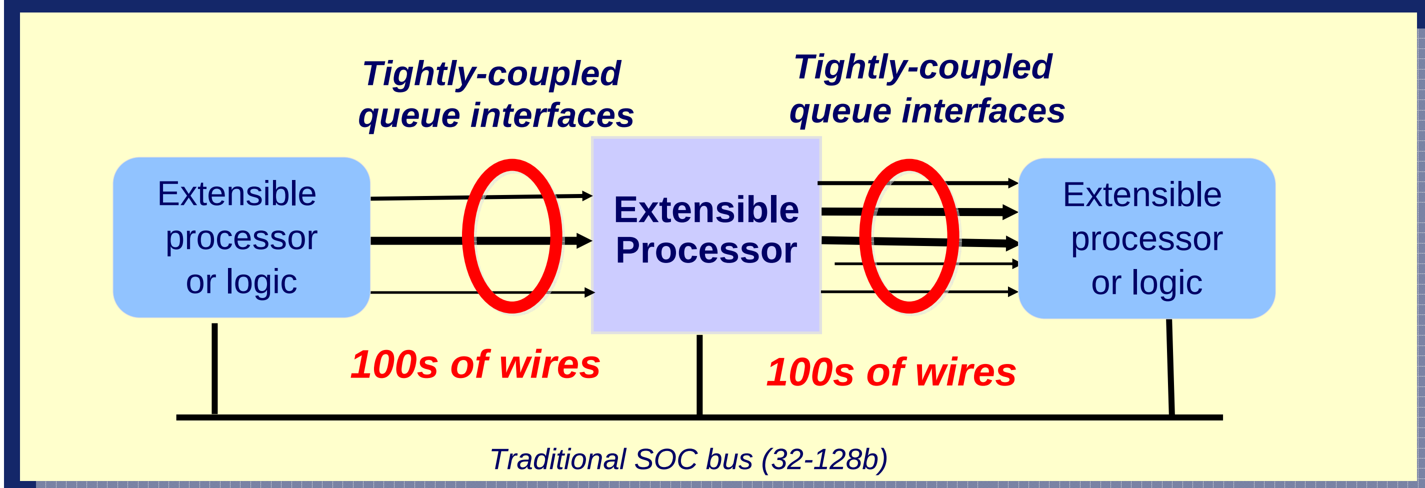


Keys to Efficient MP
High bandwidth, low energy



Remove RISC bottlenecks

- *> 23 GB/s memory bandwidth@500 MHz*
- *Energy-efficiency for memory-based tasks*

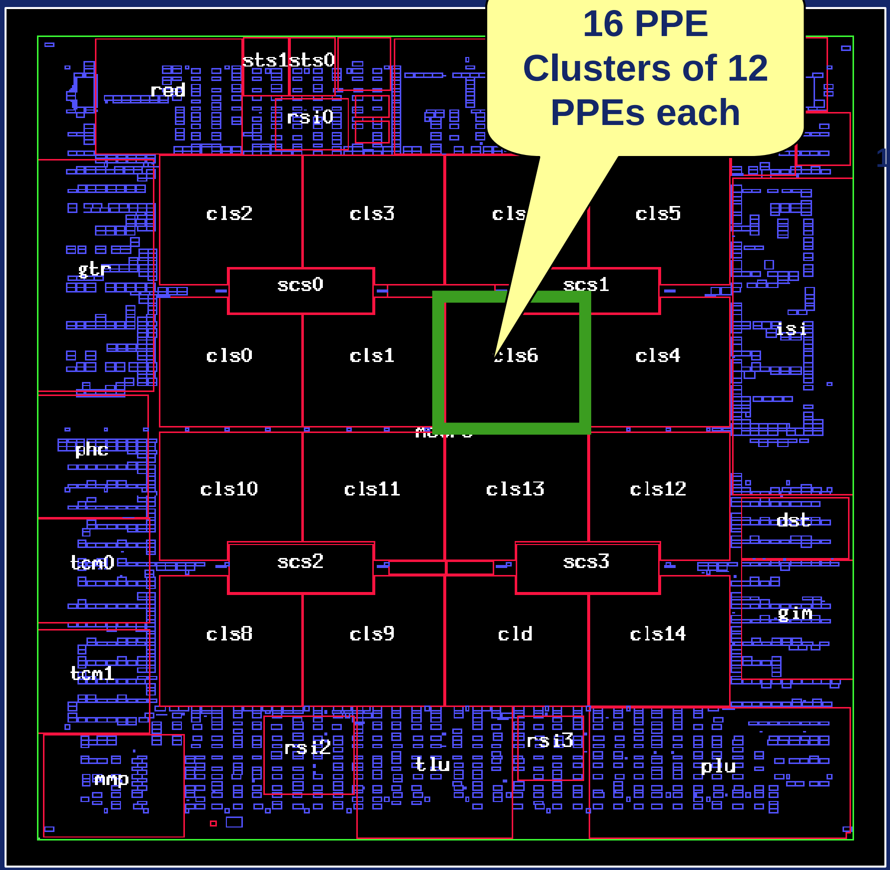


Remove bus bottlenecks

- *Arbitrary size queues*
- *Zero - overhead synchronization*
- *Lowest energy communication*

Enabling Parallel Architectures

Cisco CRS-1 Terabit Router



188 Xtensa network processing cores per Silicon Packet Processor

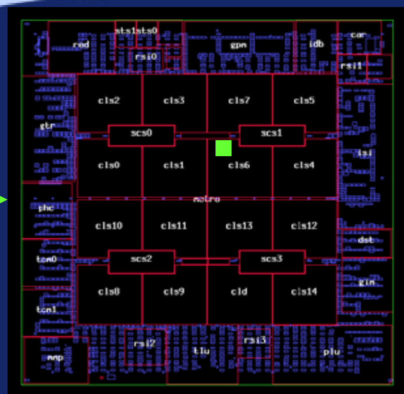
Up to 400,000 processors per system

CISCO SYSTEMS

tensilica

Massive general-purpose throughput

96Gb/s



96Gb/s

18mm x 18mm IBM 0.13μm
18M gates
8Mbit SRAMs

50,000 general purpose MIPS
175 Gb/s memory bandwidth

Programmability also means

- Ability to juggle feature ordering
- Support for heterogeneous mixes of feature chains
- Rapid introduction of new features

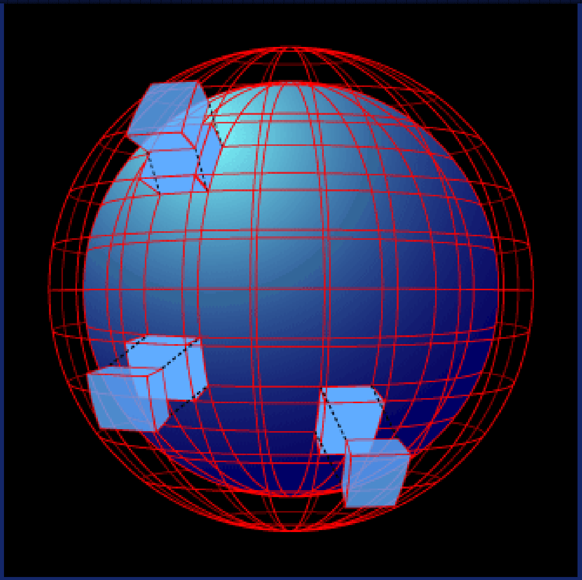
Routing task runs to completion on each processor:

- IPv4 Unicast
- MPLS–3 Labels
- Link Bundling (v4)
- Load Balancing L3 (v4)
- 1 Policier Check
- Marking
- TE/FRR
- Sampled Netflow
- WRED
- ACL
- IPv4 Multicast
- IPv6 Unicast
- Per prefix accounting
- GRE/L2TPv3 Tunneling
- RPF check (loose/strict) v4
- Load Balancing V3 (v6)
- Link Bundling (v6)
- Congestion Control

High Performance Computation

Technical/ Economic Challenges

1. Variance in data-reference/ communication patterns for codes
2. Potential for extreme scalability via large-scale processing arrays
3. Size, cost and maintenance strongly correlated to system power dissipation
4. General-purpose CPUs optimized for integer applications – unimpressive performance per $, per watt



Parallel Climate Modeling

Lenny Oliker and Michael Wehner of Lawrence Berkeley Lab speculation: a much more parallel climate model

- 1.5km grid for Earth
- 20,000,000 domains
- 500 MFLOPS/domain
- 500 MB/s per domain
- Complex algorithms require general-purpose programmability in double precision floating point
- 2D communications mesh @~20MB/s per domain

System Architecture Approach

- Highly suitable for distributed array computation
- Two design challenges:
  - Total memory bandwidth: 5-10 peta-bytes/s – many parallel local DRAM channels
  - Power: GFLOPS/W best predictor of system cost, size
- Best off-the-shelf processor (IBM Cell) is about 1.5 DP GFLOPS/W
- Domain-specific processor approach offers significant potential advantage (>10 DP GFLOPS/W)

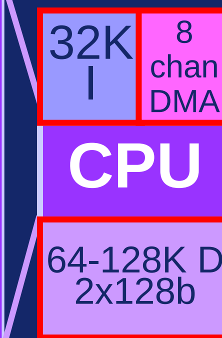
tensilica

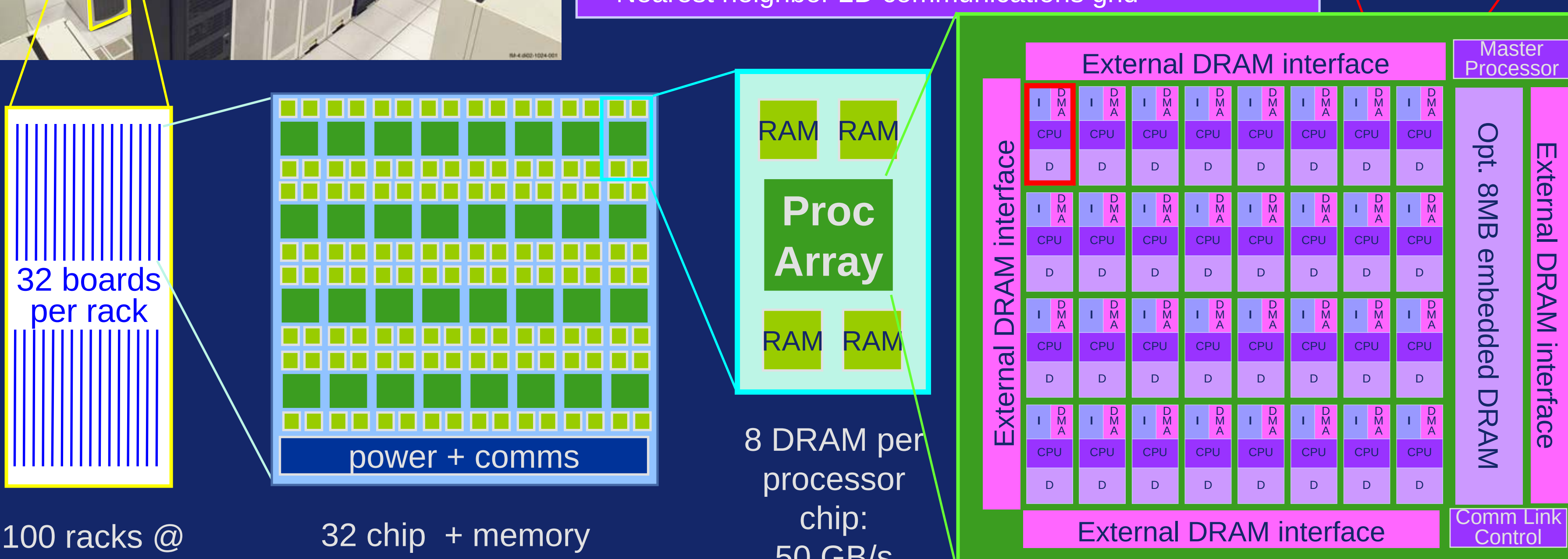
10 PetaFLOPS System Concept

3.8M processor array

VLIW CPU:

- 128b load-store + 2 DP MUL/ADD + integer op/ DMA per cycle:
- Synthesizable at 650MHz in commodity 65nm
- 1mm^2 core, 1.8-2.8mm^2 with inst cache, data cache data RAM, DMA interface, 0.25mW/MHz
- Double precision SIMD FP : 4 ops/cycle (2.7GFLOPs)
- Vectorizing compiler, cycle-accurate simulator, debugger GUI
- 8 channel DMA for streaming from on/off chip DRAM
- Nearest neighbor 2D communications grid





100 racks @ ~25KW

32 chip + memory clusters per board (2.7 TFLOPS @ 700W

8 DRAM per processor chip: 50 GB/s

32 processors per 65nm chip 83 GFLOPS @ 7W

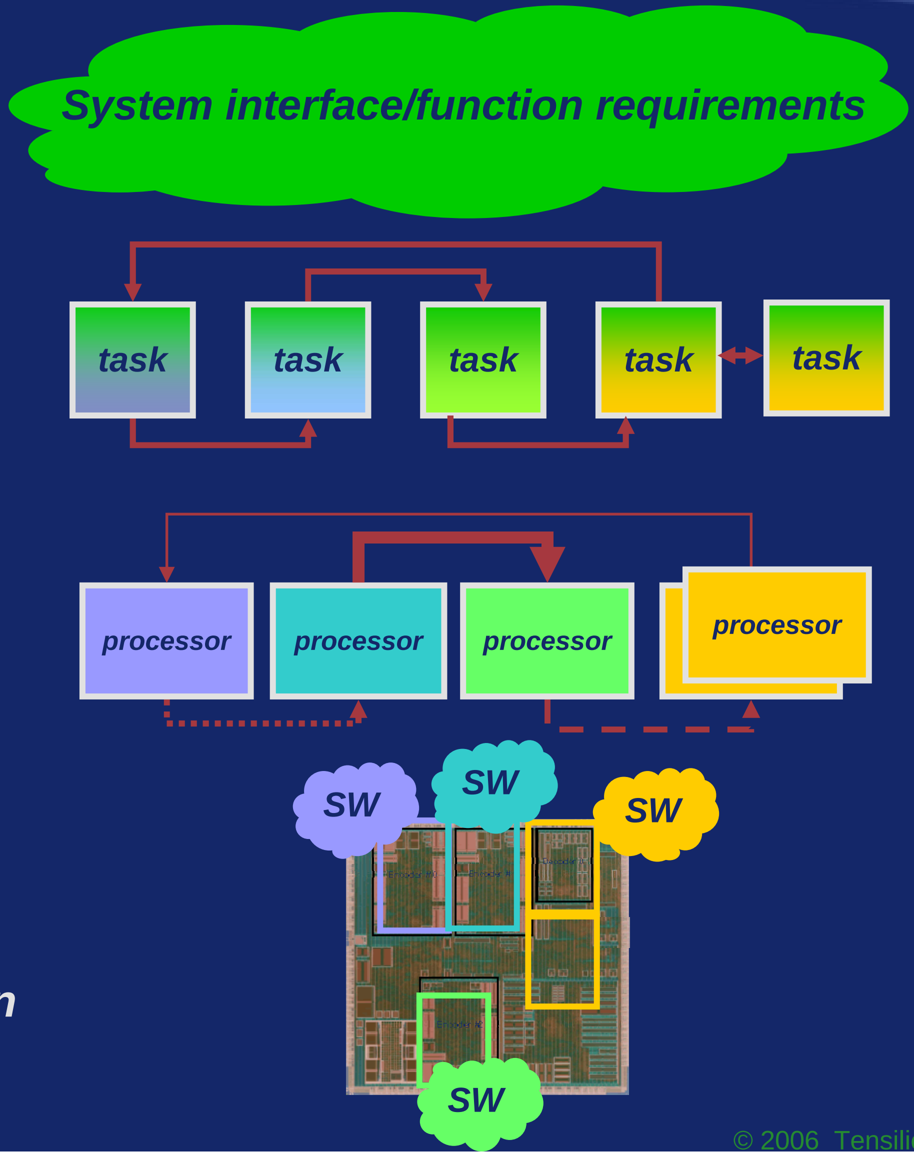
The New World
Automated MP SOC Design Flow

Complex system definition

Rapid partitioning into a set of communicating tasks

Optimized HW/SW communication on tuned processors and interconnect

Physical synthesis of hardware + high-accuracy system software integration



Shift towards embedded electronics changes the processor ideals

Ideal for general computing processor:

- Target tethered systems: integration at board level
- Maximum generic performance at package power limit
- Small scale symmetric MP-ready
- One size fits all instruction set
- Consumes large faction of chip area

Ideal for system-on-chip processor:

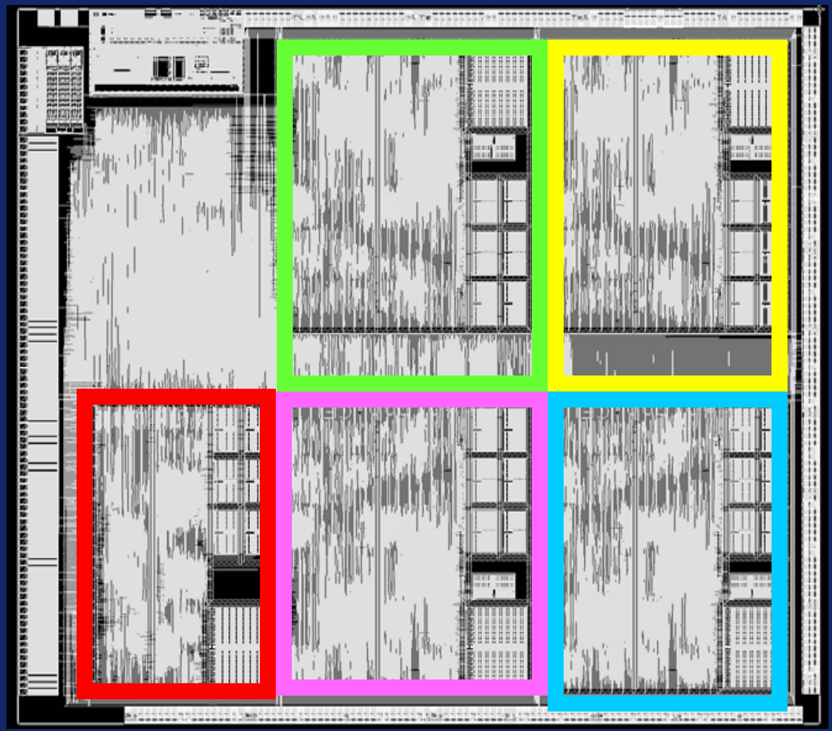
- Target mobile systems: integration at chip level
- Lowest energy at application-required performance
- Large scale heterogeneous MP-ready
- Extensible without compromise on generic performance
- Multiple dimensions of extensibility
  1. Instruction set
  2. System control functions
  3. Memory systems and interconnect
  4. Non-memory-mapped interfaces and interconnect

Wrap-Up

System-on-chip platform trends:

- More functions per chip
- More SW content per function
- Intense pressure on power, cost, performance and feature evolution pace

Result: Multiple configurable processor design now widely adopted



Impact of unified MP architecture

1. Essential base code compatibility across all processors
2. Common tools, models across all processors
3. Seamless mix-n-match of features and performance optimization