The Cavium 32 Core OCTEON II 68xx

R. E. Kessler

Cavium, Inc.

Hot Chips 23

August, 2011

Agenda

CAVIUM

- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - Efficient CPU cores
  - Cache coherence, interconnect & memory bandwidth, chip floorplan
  - Power scalability
  - Hardware work queueing, scheduling, synchronization, and ordering assist
  - Coprocessor acceleration
- Chip Performance Results

About Cavium, Inc.

- Founded 2001
- NASDAQ IPO (CAVM) 2007
- 825 Employees, 625 in Engineering
- $250+M annual revenue run-rate, among fastest-growing public technology companies
- Profitable with Strong Financials: ~$70M cash & cash equivalents, no debt, strong cash flow

- Global Footprint: US, India, Taiwan, China, Canada
- MIPS and ARM based Processor SOCs
- Addressing Multi-billion dollar Networking, Communications and Digital Home markets
- Voted #1 Multi-core processor vendor by Heavy Reading 2010 survey of 50+ worldwide networking OEMs

Cavium SoC's for Range of Target Markets

CAVIUM

Networking			Consumer	Wireless		Storage
Enterprise & Edge Routers	Enterprise, Metro Switches & L4-L7 Equipment	Security & DPI Equipment	Home, Video, High Bandwidth Broadband	3G, 4G Infra-structure	LAN: Controllers & Enterprise AP's	Storage Networking, Arrays & Adapters

Networking : Range of Router, Appliance, Gateway SoCs

OCTEON

Security, Compression, DPI Offload & Virtualization SoCs

NITROX

Video SoCs

PureVu

Media & Set Top SoCs

Celestial

SMB, Home & NAS SoCs

ECONA

3G/ 4G Data-card SoCs

ODYSSEY

Base-station, RAN, Core SoCs

OCTEON

WLAN Controller & AP SoCs

OCTEON

Intelligent Adapter SoCs

OCTEON

Highly Integrated SOCs enable Lower Real-Estate, Cost & Power

OCTEON Design Philosophy

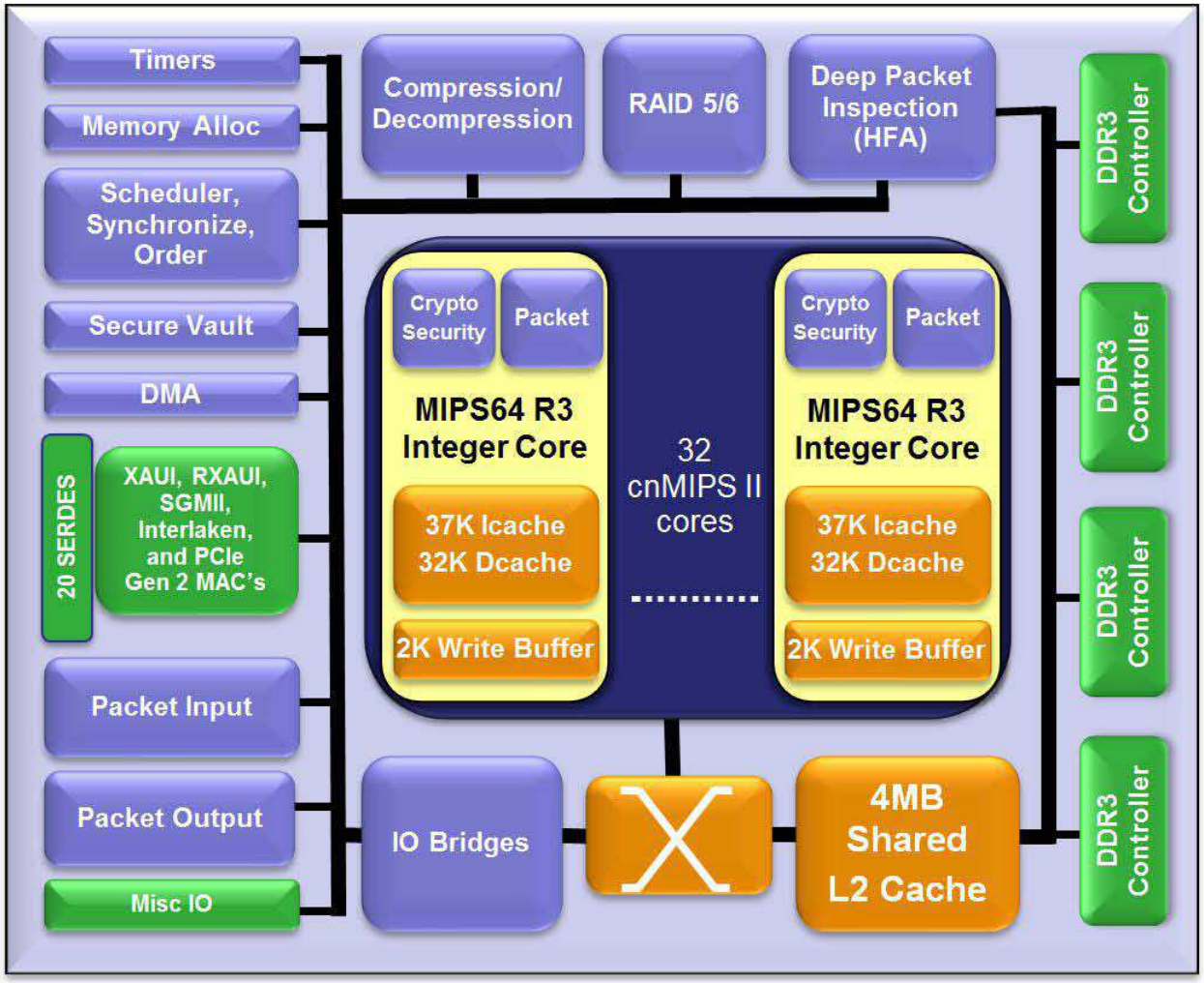
CAVIUM

High Application Performance at Low Power and Cost	• Many power and area efficient MIPS64 CPU cores • Hardware acceleration for high packet throughput, and content processing, e.g. compression, RegEx pattern match, RAID5/6 • Integrated networking and memory controllers
Scalable Performance	• Take advantage of packet and flow-level parallelism • Linear performance scaling with increasing number of cores enabled by proven hardware features
Optimized ISA	• MIPS64 version 3 instruction set with OCTEON enhancements • More than 80 instructions added on top of MIPS ISA • Full C programming and OS support
Flexible Hardware Security Acceleration	• Hardware accelerators in each core for a comprehensive set of asymmetric and symmetric algorithms: RSA, DH, ECC, IPSec, SSL/TLS, KASUMI, SNOW3G, others • Adapt to new algorithms through software updates
Software Compatible Roadmap	• Single SDK to develop software for all OCTEONs • Software compatible from 1-32 cores

OCTEON II CN68XX Block Diagram

CAVIUM

- 32 custom designed MIPS64 cores
- Up to 1.5 GHz
- Up to 96G inst/sec, 40+Gbps
- 4 72-bit DDR3 interfaces up to 1600 MHz data rate
- Optimized for service-rich networking , security, wireless, and storage apps
- HW Acceleration:
  - ✓ DPI acceleration with integrated HFA (RegEx Engine)
  - ✓ Comprehensive crypto algorithms and RNG
  - ✓ TCP, Packet Processing
  - ✓ Compression
  - ✓ RAID5/6, De-dup
  - ✓ Multi-core scaling



Agenda

- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - **Efficient CPU cores**
  - Cache coherence, interconnect & memory bandwidth, chip floorplan
  - Power scalability
  - Hardware work queueing, scheduling, synchronization, and ordering assist
  - Coprocessor acceleration
- Chip Performance Results

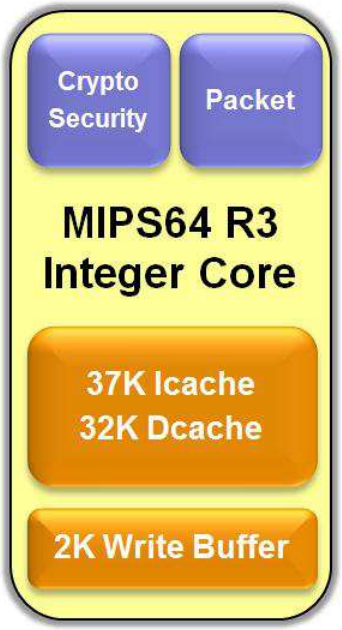
Small CPU Core or Big CPU Core?

CAVIUM

- Many potential Big Core features:
  - Huge caches
  - Very high frequency, deep pipeline
  - Many-way issue
  - Out-of-order issue
  - Floating-point
  - …
- Important questions:
  - Does the feature add more performance than area/power?
  - Is the feature difficult or expensive to implement, take to production, and support?

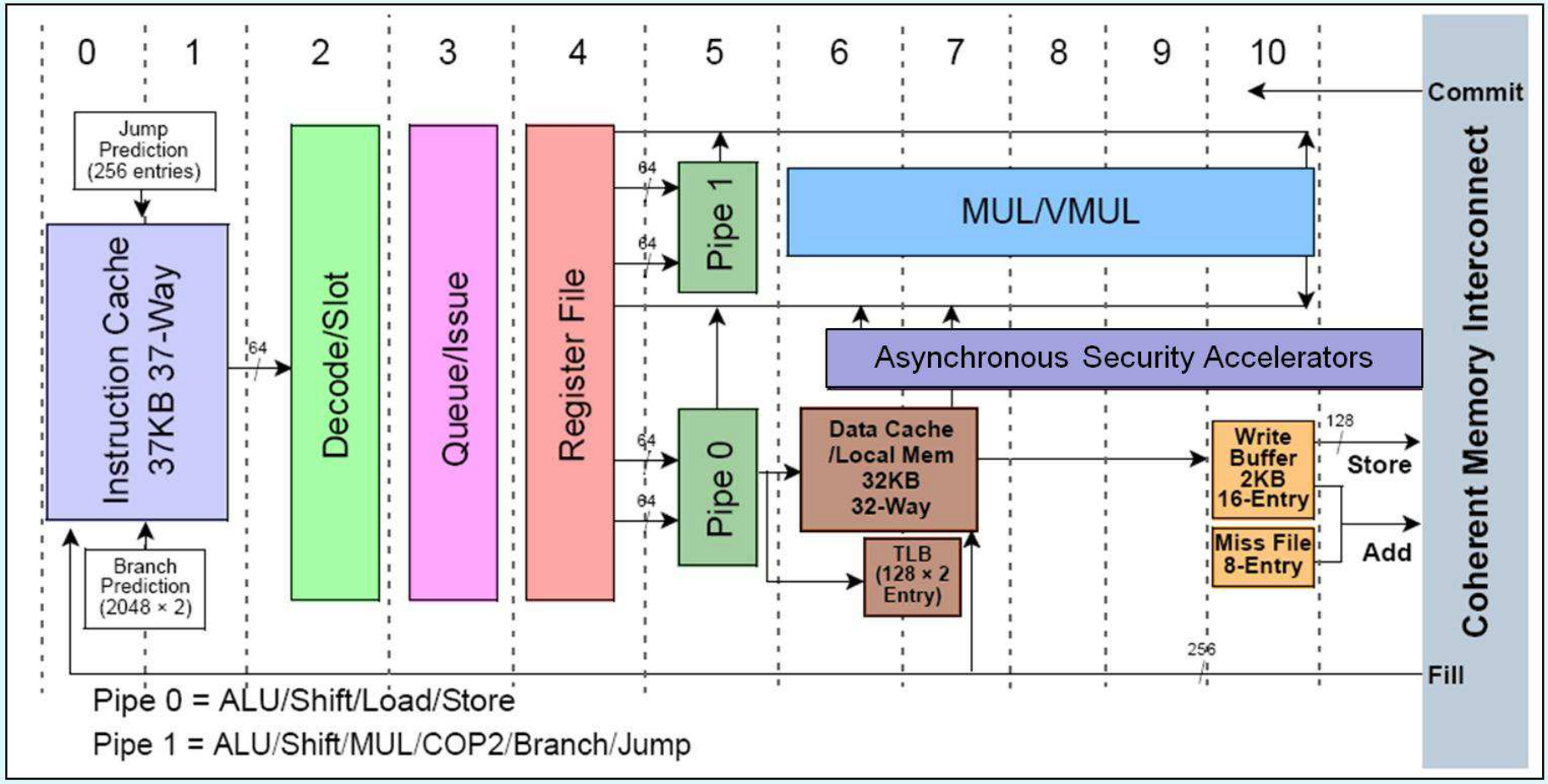
OCTEON II CPU (i.e. cnMIPS II) Design

CAVIUM



- cnMIPS II Core Goals:
  - General-purpose, industry-standard 64-bit ISA
  - Great fit for networking, security, wireless
  - Excellent MIPS/area & MIPS/watt
  - Use multi-core to scale product line up and down
  - Low latency, deterministic performance
- cnMIPS II Core Non-goals:
  - Highest power
  - Highest cost
  - Greatest complexity (longer implementation time)
  - Largest customer support cost
- Not directly mentioned:
  - Frequency
  - Single-thread performance
  - …

cnMIPS II Core 8+ Stage Pipeline



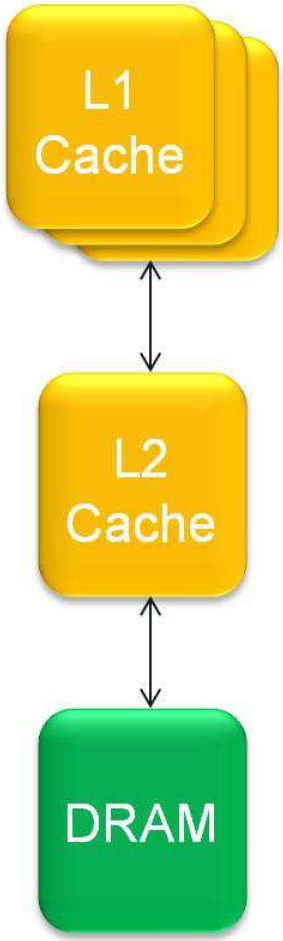
- Shipping at up to 1.8 GHz in 65nm
- Thread-dedicated resources = very deterministic CPU performance
- Highly-associative L1 caches = equivalent miss rate to much larger caches

Agenda

CAVIUM

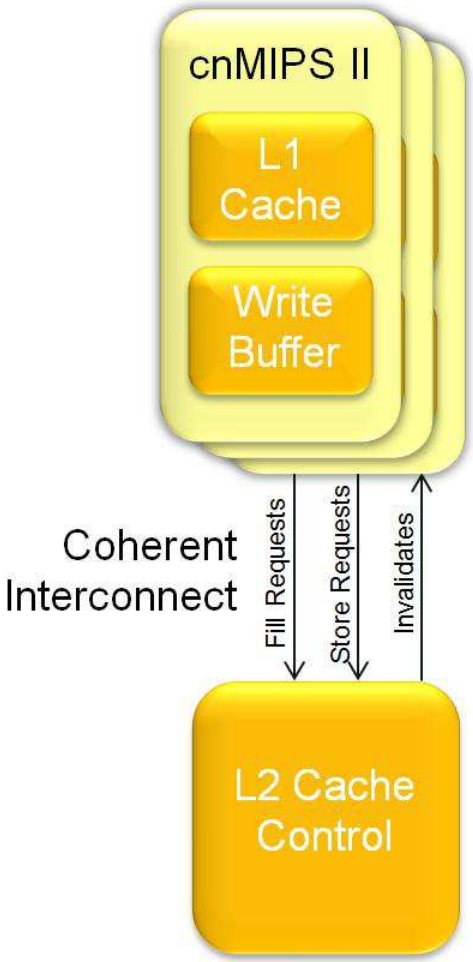
- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - Efficient CPU cores
  - **Cache coherence, interconnect & memory bandwidth, chip floorplan**
  - Power scalability
  - Hardware work queueing, scheduling, synchronization, and ordering assist
  - Coprocessor acceleration
- Chip Performance Results

OCTEON Cache Policies



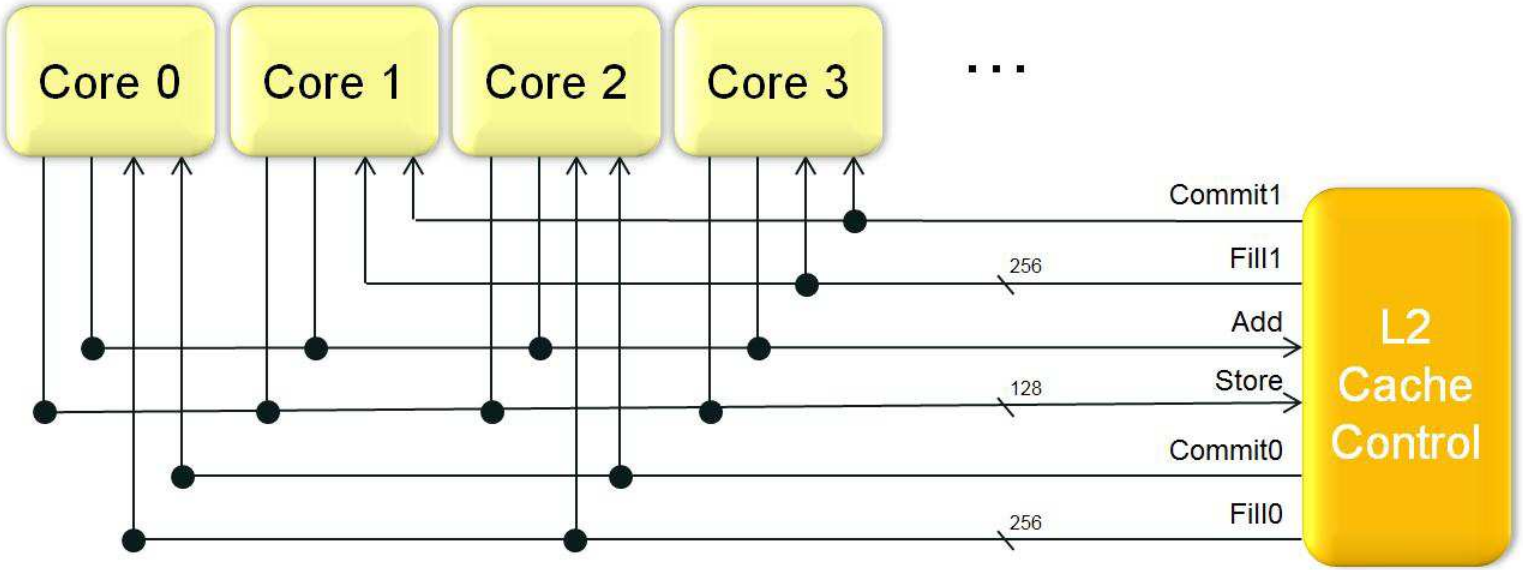
- L1 <-> L2 Cache: Write-through
  - Excellent performance for networking and mobile applications
  - Minimal per-CPU-core cost
  - Simple and highest performance
  - Lowest possible read latencies
  - Allows many outstanding stores, optimizations
  - Automatic L1 error correction
- L2 Cache <-> DRAM: Write-back
  - Standard DDR3 DRAM DIMM's are highest performance with block transfers
  - Minimizes required DRAM bandwidth
  - Don't-write-back feature (e.g. for most of packet data) plus additional cache hints

OCTEON L1<->L2 Coherence and Memory Model

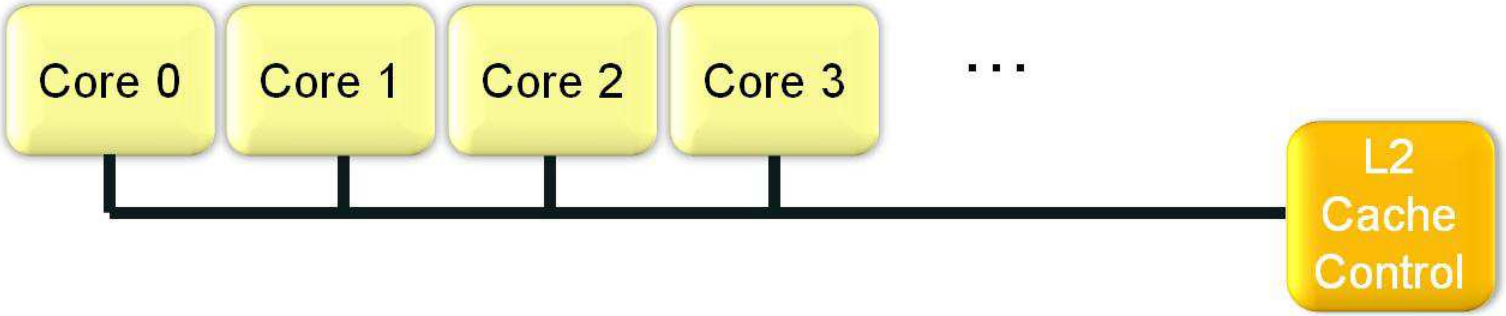


- Write-through, write-invalidate coherence protocol
- L2 Cache Controller is the coherence point
  - L2 controller tracks L1 cache contents
  - Invalidates to maintain L1 coherence
- Aggressive write-buffering in cnMIPS II cores eliminates writes
  - 2 KB merging write buffer
  - Fully-coherent, loosely-consistent memory model
  - Page-wise hints to eliminate write-buffer flushes of private data

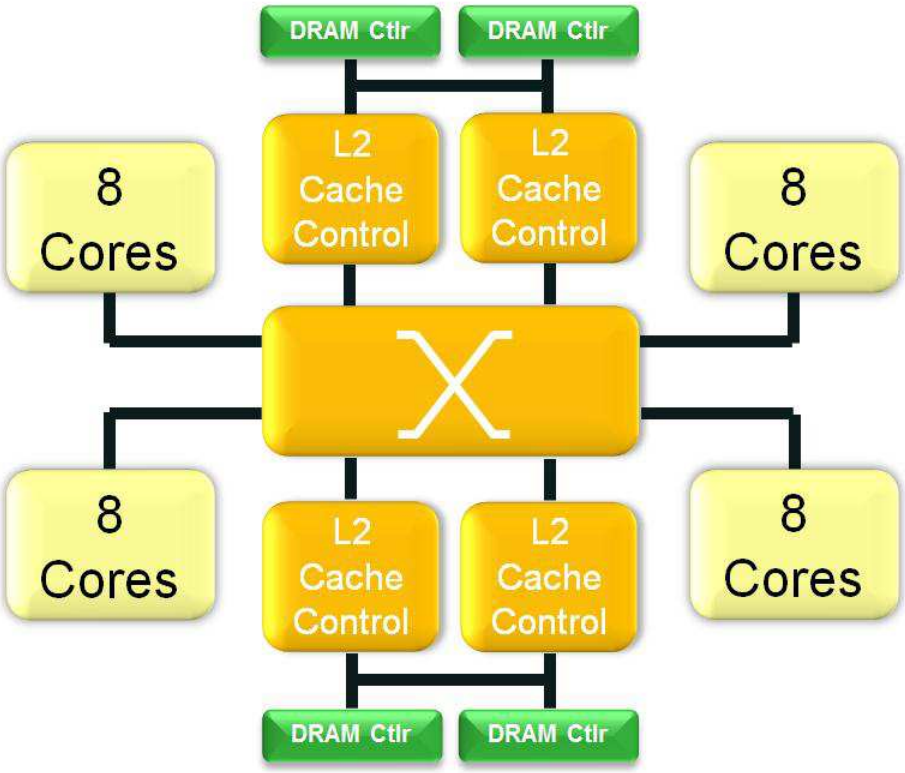
Lower Core Count OCTEON Coherent Interconnect



Redrawn:

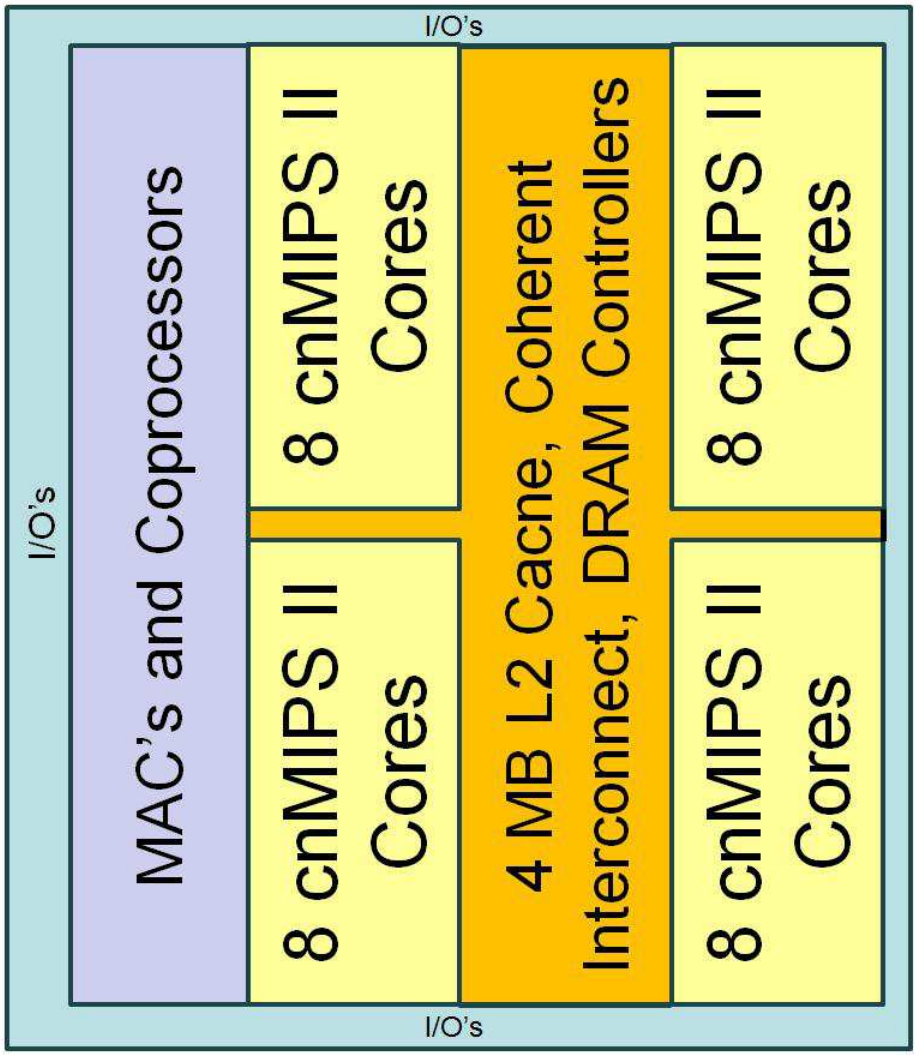


32 Core OCTEON Coherent Interconnect



- Crossbar interconnect easily scales to 32 cores
- Optimized for both low latency and high bandwidth
- Flat, deterministic latency profile
- Interconnect provides best combination of scalability and low-power

32 Core OCTEON Chip Floorplan



Approximately 800 Million Transistors in 65nm

Agenda

CAVIUM

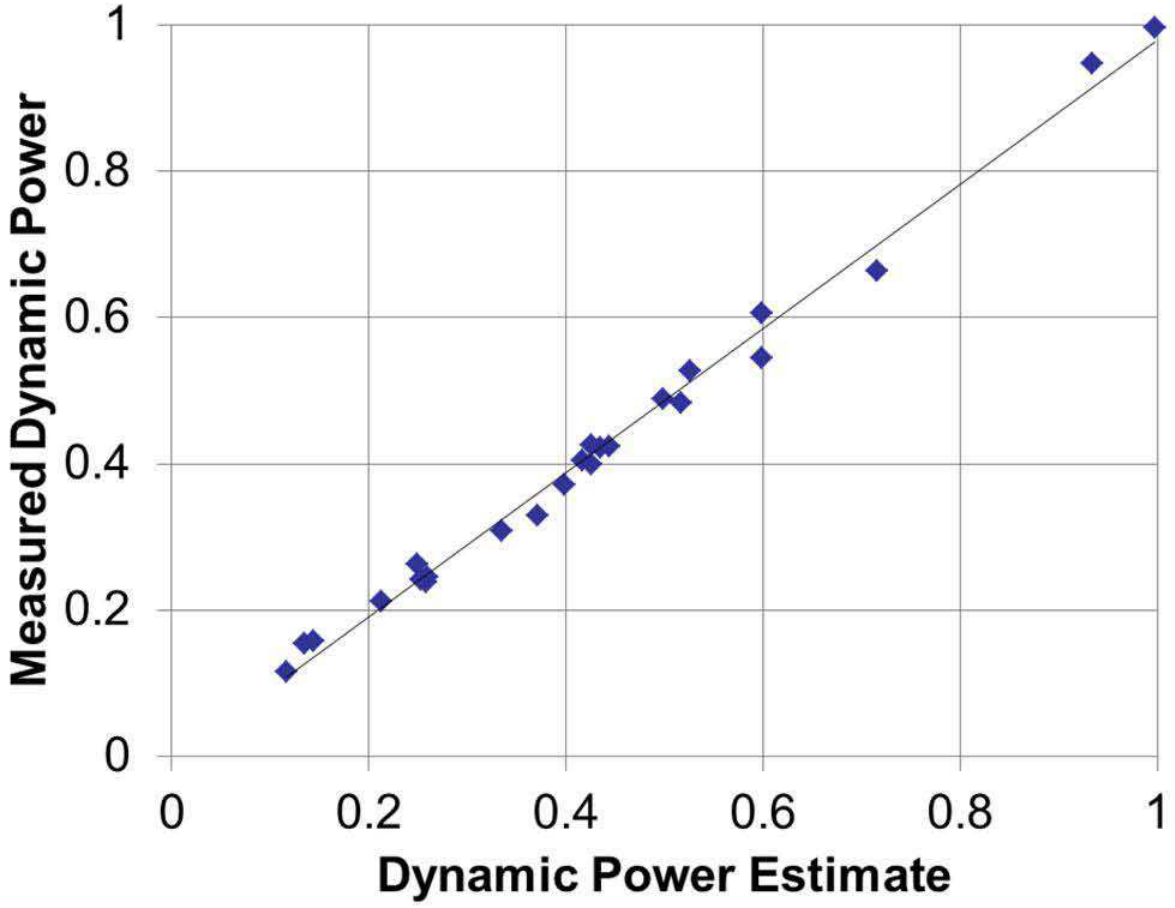
- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - Efficient CPU cores
  - Cache coherence, interconnect & memory bandwidth, chip floorplan
  - **Power scalability**
  - Hardware work queueing, scheduling, synchronization, and ordering assist
  - Coprocessor acceleration
- Chip Performance Results

OCTEON 68xx Power Optimizer Technology

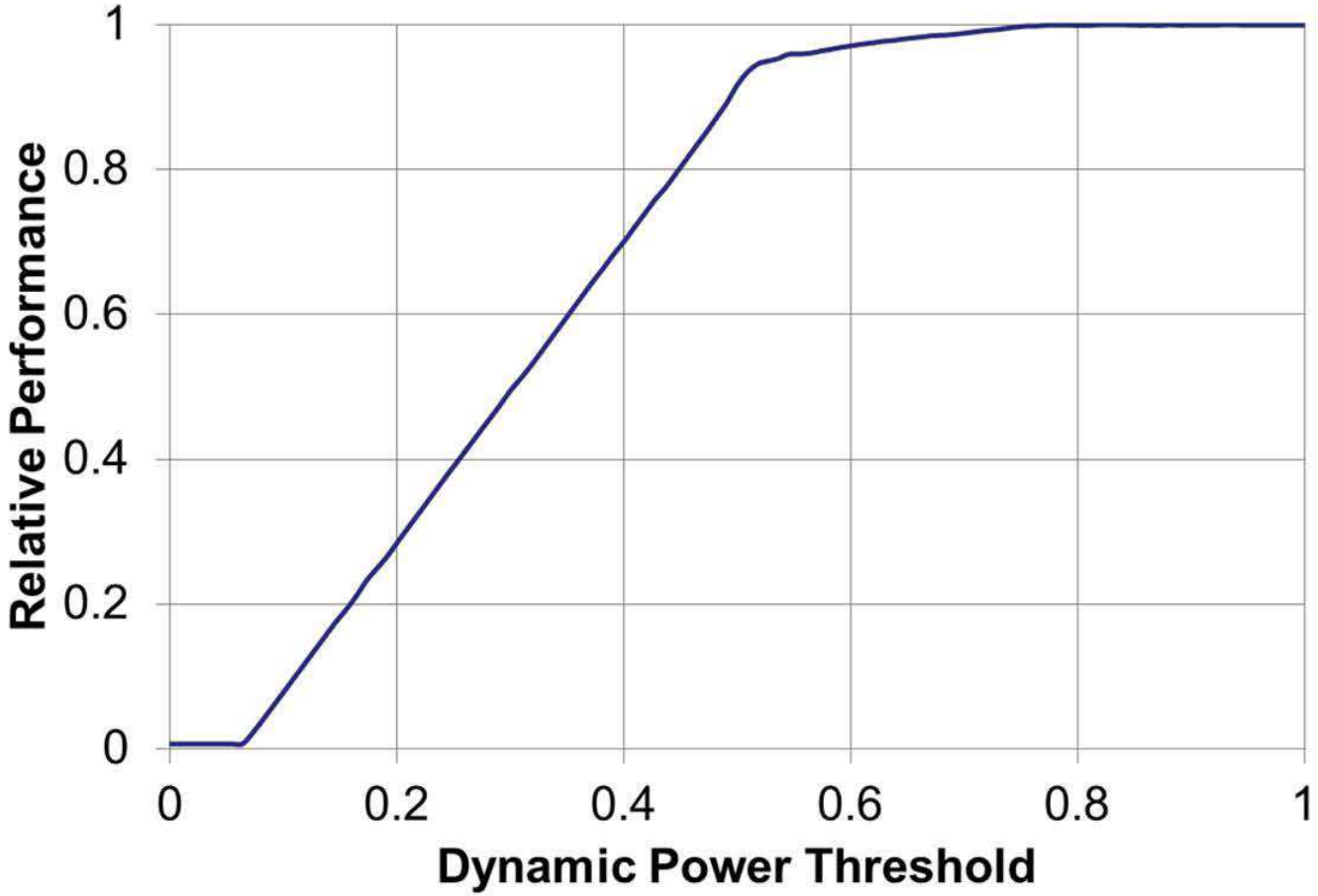
CAVIUM

- Per-CPU dynamic power consumption estimates
  - Digital logic that monitors CPU behavior
  - Pipe, functional unit, bus and clock activity monitored
- Per-CPU dynamic power threshold
  - CPU forced to idle when estimate exceeds threshold
  - Power controlled over intervals of 256-1024 cycles
- Per-CPU threshold in a register
  - Software can quickly and easily change it
- Suitable for thermal design or average power reduction:
  - Closed-loop (e.g. thermal sensor) or open-loop thermal solution

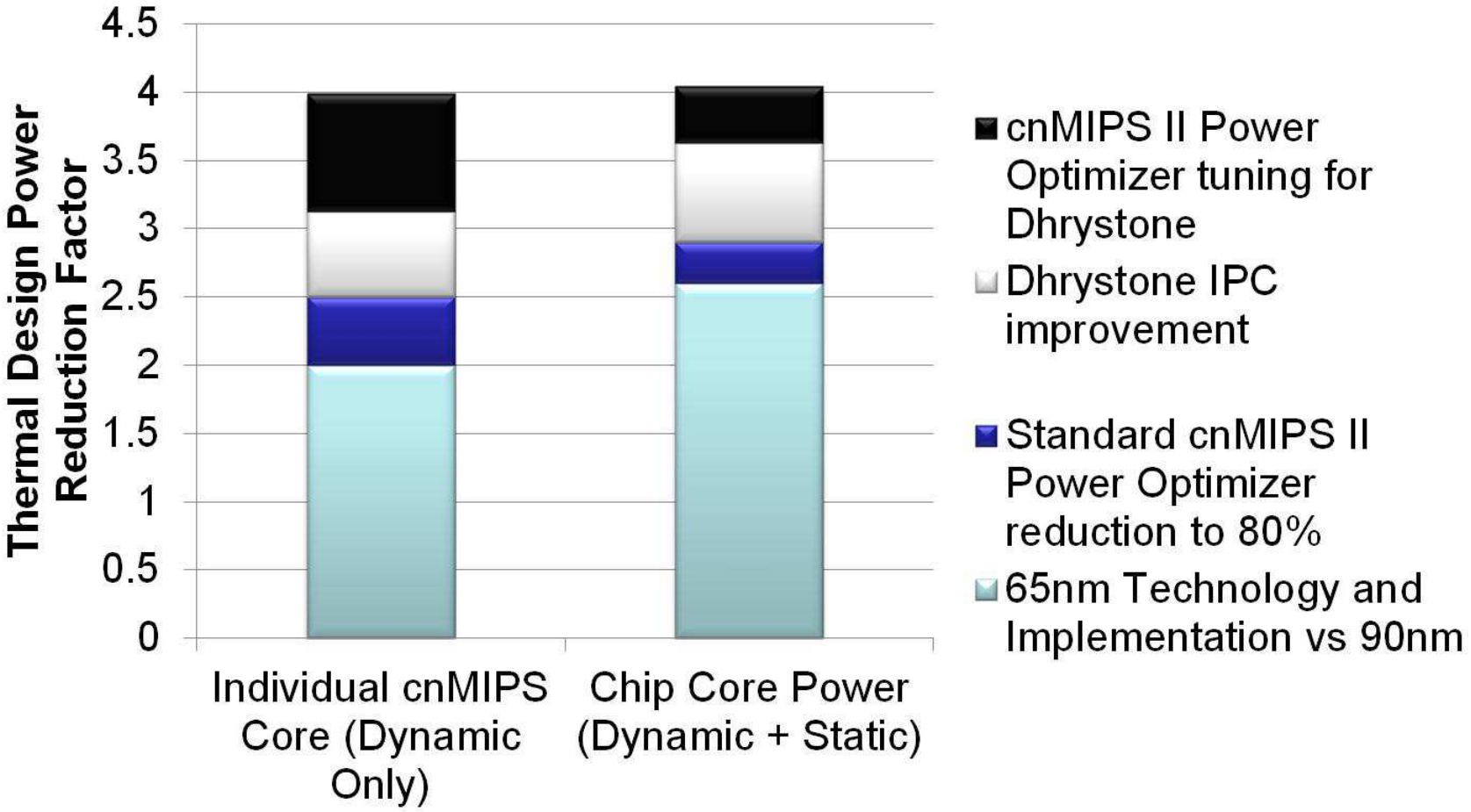
Accuracy of Dynamic Power Estimate for Various Applications



Performance vs. Dynamic Power Threshold for Zlib



Thermal Improvements from 16 Core 58xx to 32 Core 68xx OCTEON (per Dhrystone instruction)



OCTEON Power Optimizer Comparison

CAVIUM

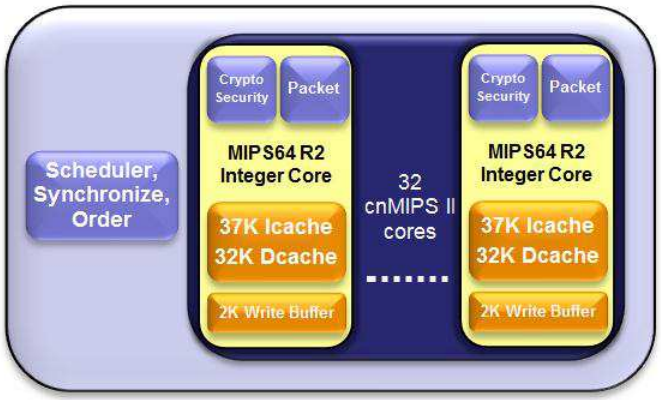
- The OCTEON solution is unique
- Advantages compared to dynamic voltage and frequency scaling (DVFS):
  - Very fine-grained core-by-core power control
  - A low power application is not penalized
    - Frequency reduction affects all applications
    - 95+% of applications don't even achieve 80% of max spec power
  - Power optimizer settings can change instantly with minimal software interruption
  - Simpler chip and system design
    - Voltage and frequency do not need to change

Agenda

CAVIUM

- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - Efficient CPU cores
  - Cache coherence, interconnect & memory bandwidth, chip floorplan
  - Power scalability
  - **Hardware work queueing, scheduling, synchronization, and ordering assist**
  - Coprocessor acceleration
- Chip Performance Results

OCTEON Schedule/Synch/Order Hardware



- Work queueing
  - Unlimited-size queues for work
  - Work can be created by software
  - Work can be created by hardware
    - e.g. packet arrival
- Work/Packet Ordering
- Automatic synchronization and lock-removal
- Dynamic work scheduling
  - Hardware selects from amongst input queues
    - Quality of service
  - Different cores can receive different work
  - Integrated with ordering and synchronization
    - Work proceeds only when ordering and synchronization allows

Agenda

CAVIUM

- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - Efficient CPU cores
  - Cache coherence, interconnect & memory bandwidth, chip floorplan
  - Power scalability
  - Hardware work queueing, scheduling, synchronization, and ordering assist
  - **Coprocessor acceleration**
- Chip Performance Results

68xx Coprocessors

CAVIUM

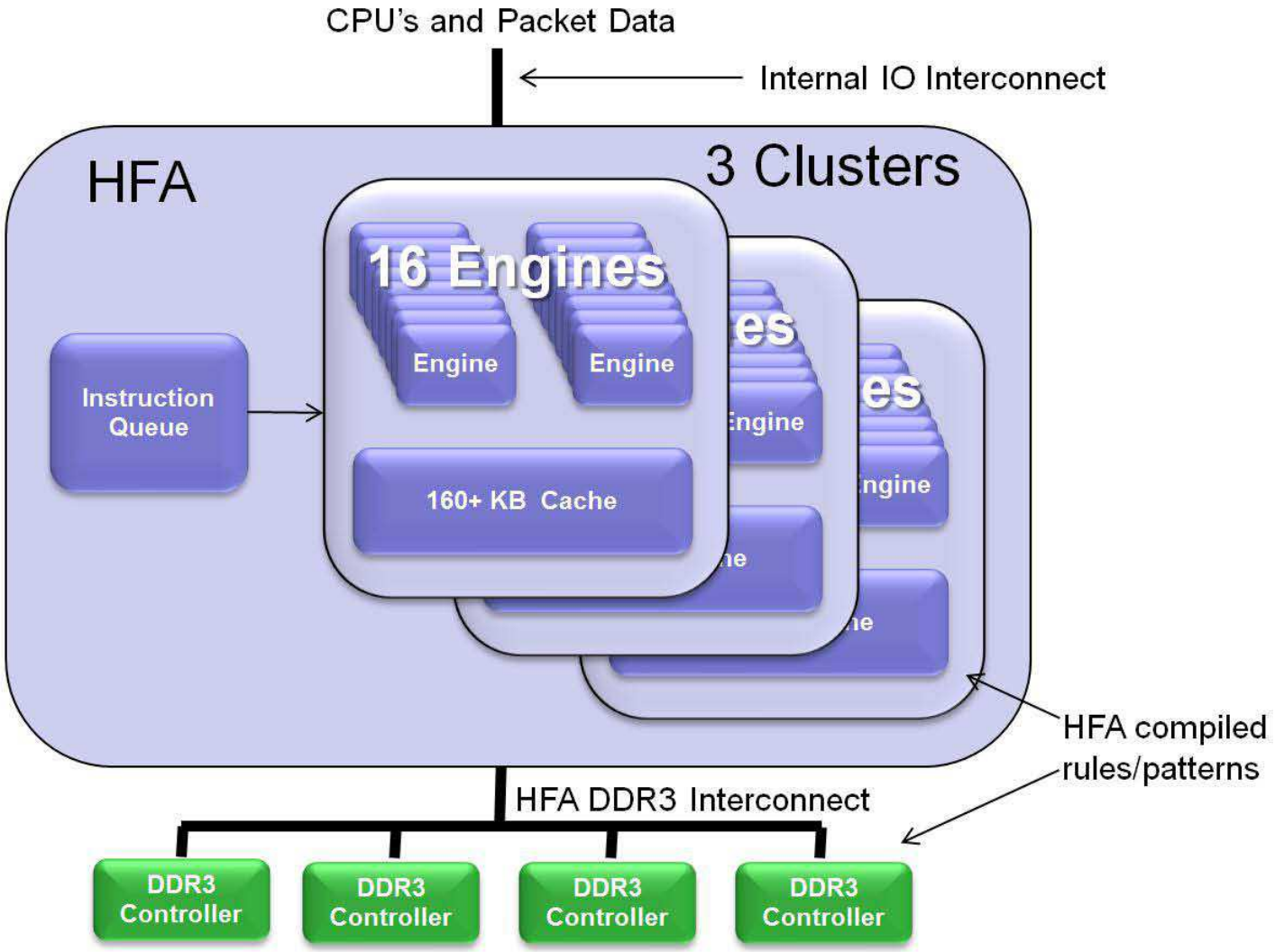
- 68xx has many coprocessors suitable for many tasks:
  - Packet processing acceleration
  - TCP processing acceleration
  - Security acceleration
  - Compression/Decompression acceleration
  - …
- But the next few slides focus only on regular expression matching, which is required by Deep-Packet Inspection applications

OCTEON 68xx Deep Packet Inspection

CAVIUM

- Many applications require Deep-Packet Inspection (DPI):
  - Intrusion detection/prevention, Packet classification, …
- We focus on pattern matching here
  - DPI may also require packet, TCP, and other processing that can be accelerated by other OCTEON coprocessors
  - The percentage of data scanned for matches varies for different applications
    - a few percent (e.g. Application Recognition) to most packet bytes (e.g. Anti-Virus, IPS)
- Patterns/rules are often regular expressions
  - Pre-compiled into hardware state machines
- Cavium OCTEON 68xx HFA processing technology:
  - Searches for regular expressions via both:
    - Deterministic Finite Automata (DFA), and
    - Non-deterministic Finite Automata (NFA)
  - Includes graph compression and caching to maximize coverage and performance
  - Compatible with stand-alone NITROX DPI processors from Cavium

68xx Deep Packet Inspection HFA Hardware

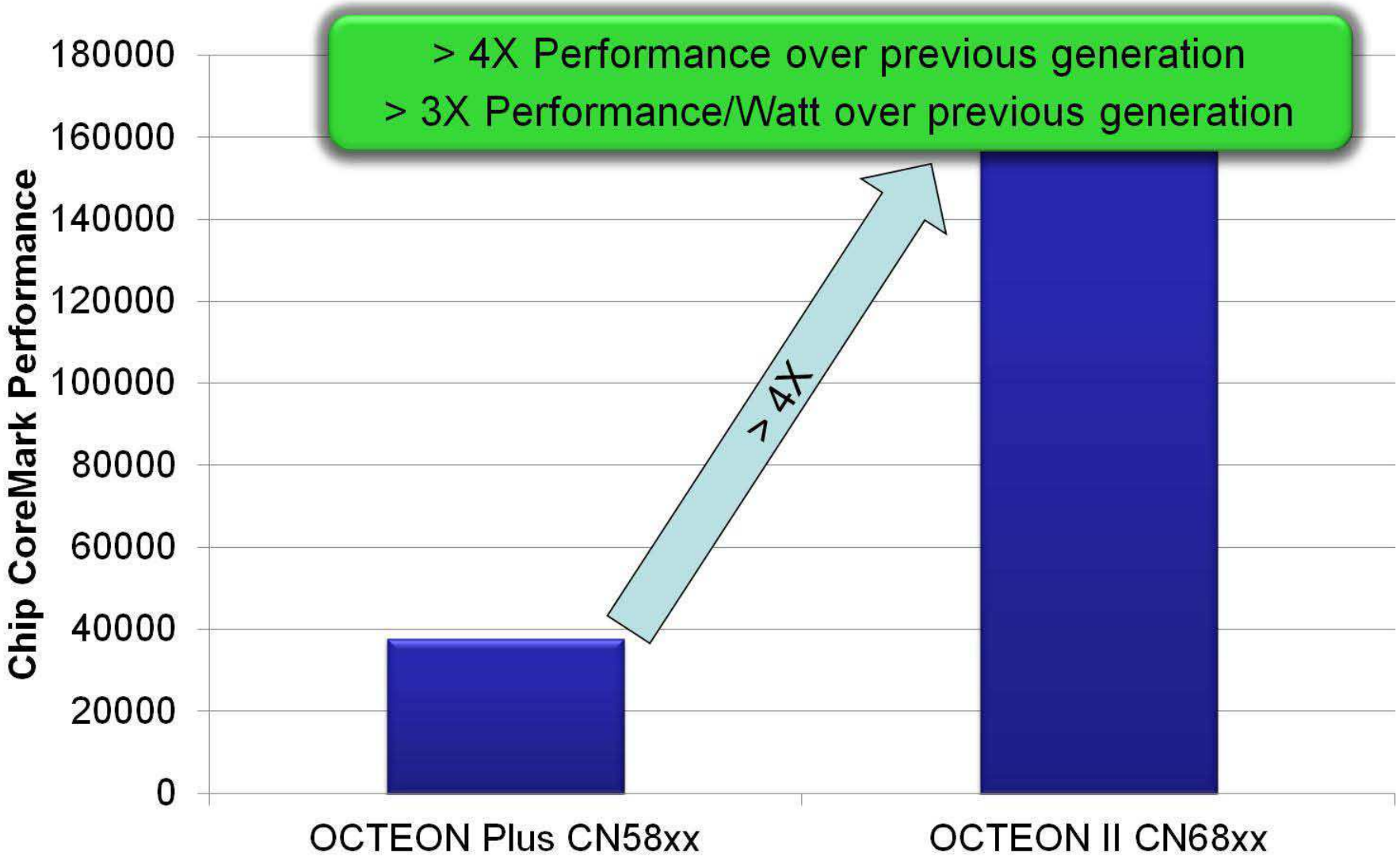


Agenda

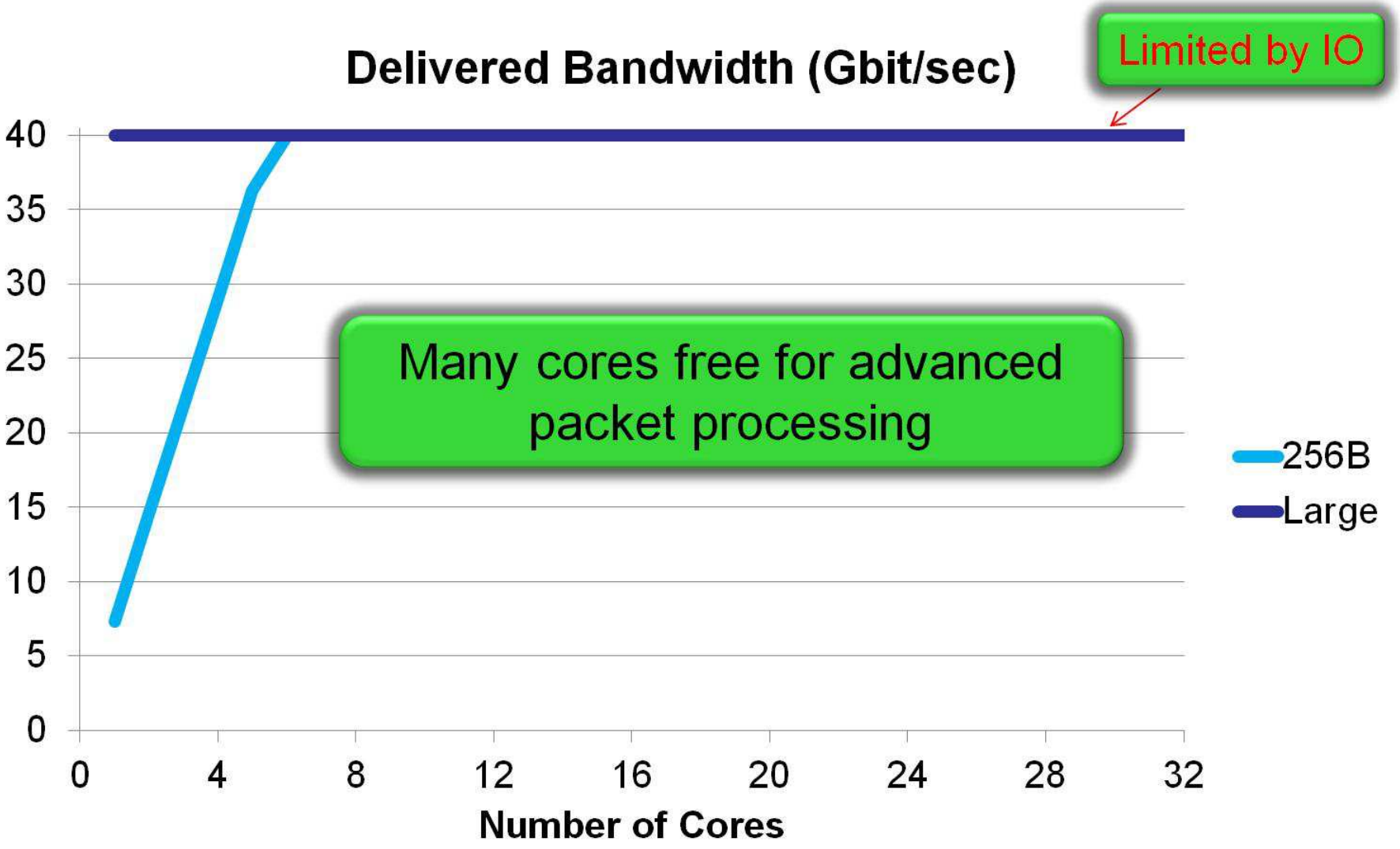
CAVIUM

- Introduction to OCTEON and 68xx
- OCTEON II 68xx Scalability Techniques
  - Efficient CPU cores
  - Cache coherence, interconnect & memory bandwidth, chip floorplan
  - Power scalability
  - Hardware work queueing, scheduling, synchronization, and ordering assist
  - Coprocessor acceleration
- **Chip Performance Results**

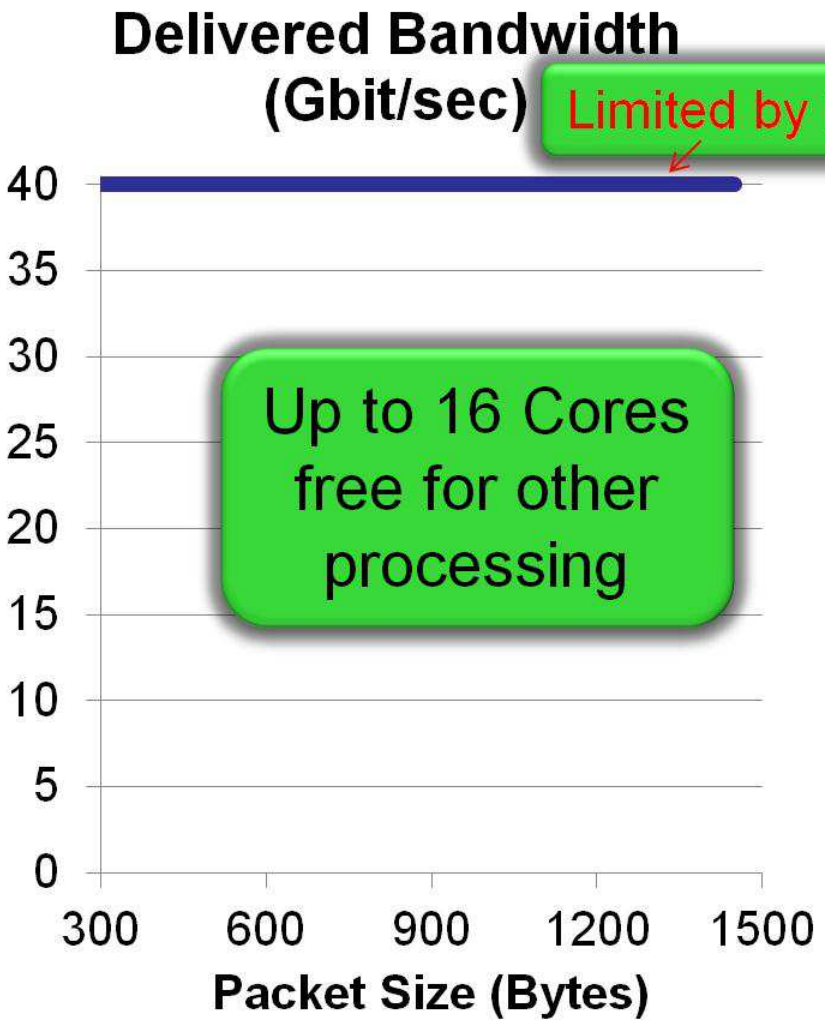
EEMBC CoreMark

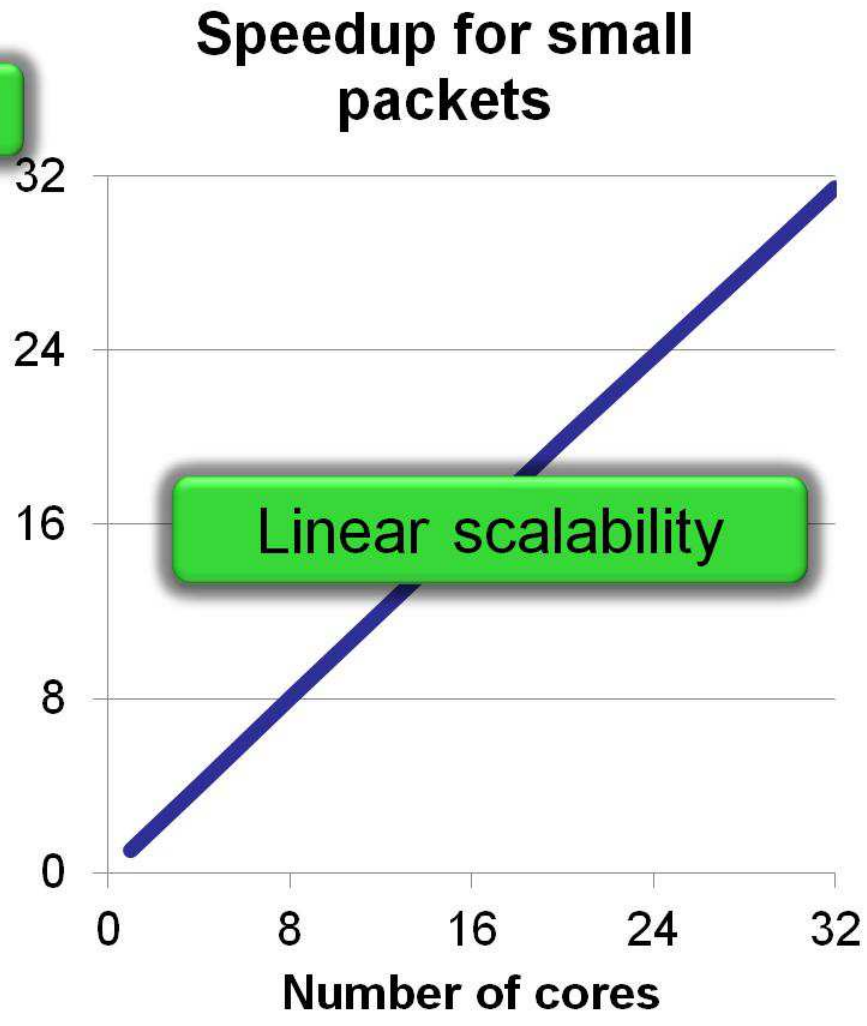


IPv4 Packet Forwarding



Full IPSEC Application





Conclusions

CAVIUM

- The 32 Core Cavium OCTEON II 68xx:
  - Up to 96 billion industry-standard 64-bit instructions per second
  - Including a coherent crossbar interconnect and other features delivering scalable general-purpose processing power
  - Including Power Optimizer technology that maximizes compute capability within a thermal envelope
  - Including integrated coprocessors tuned for networking, security, wireless, and storage