



ThunderX3

Next-Generation Arm-Based Server

Rabin Sugumar

August 2020



MARVELL™

ThunderX2®

32 core die
128 threads
Arm®v8.1

First Arm-based
Top 500 System

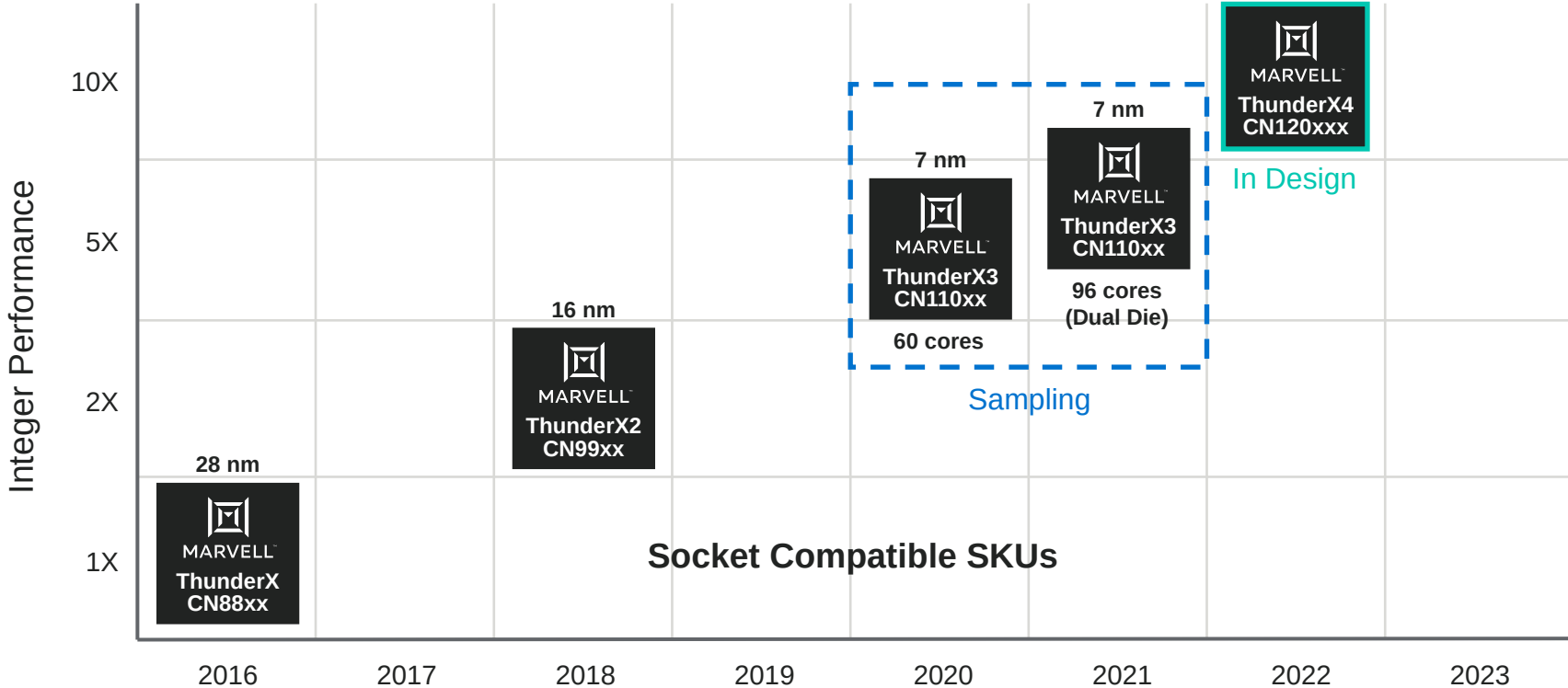
First non-x86
CPU in
Microsoft Azure

Most widely
deployed
Arm-based
server processor

Industry-leading
performance
at time of release

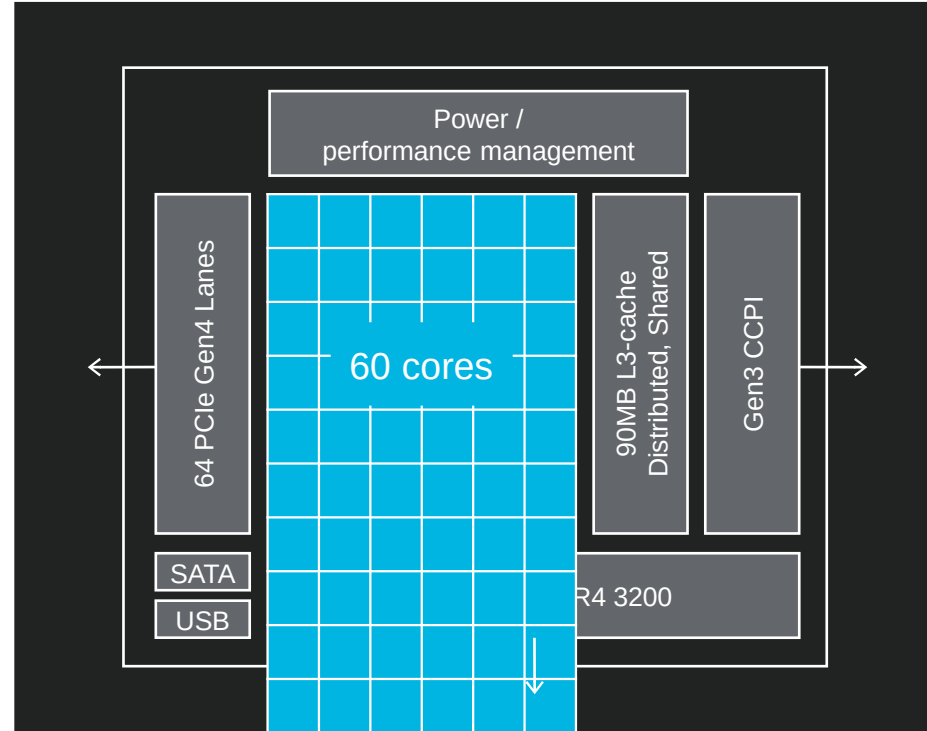
Proven
production
quality solution
at scale

Marvell server processor roadmap

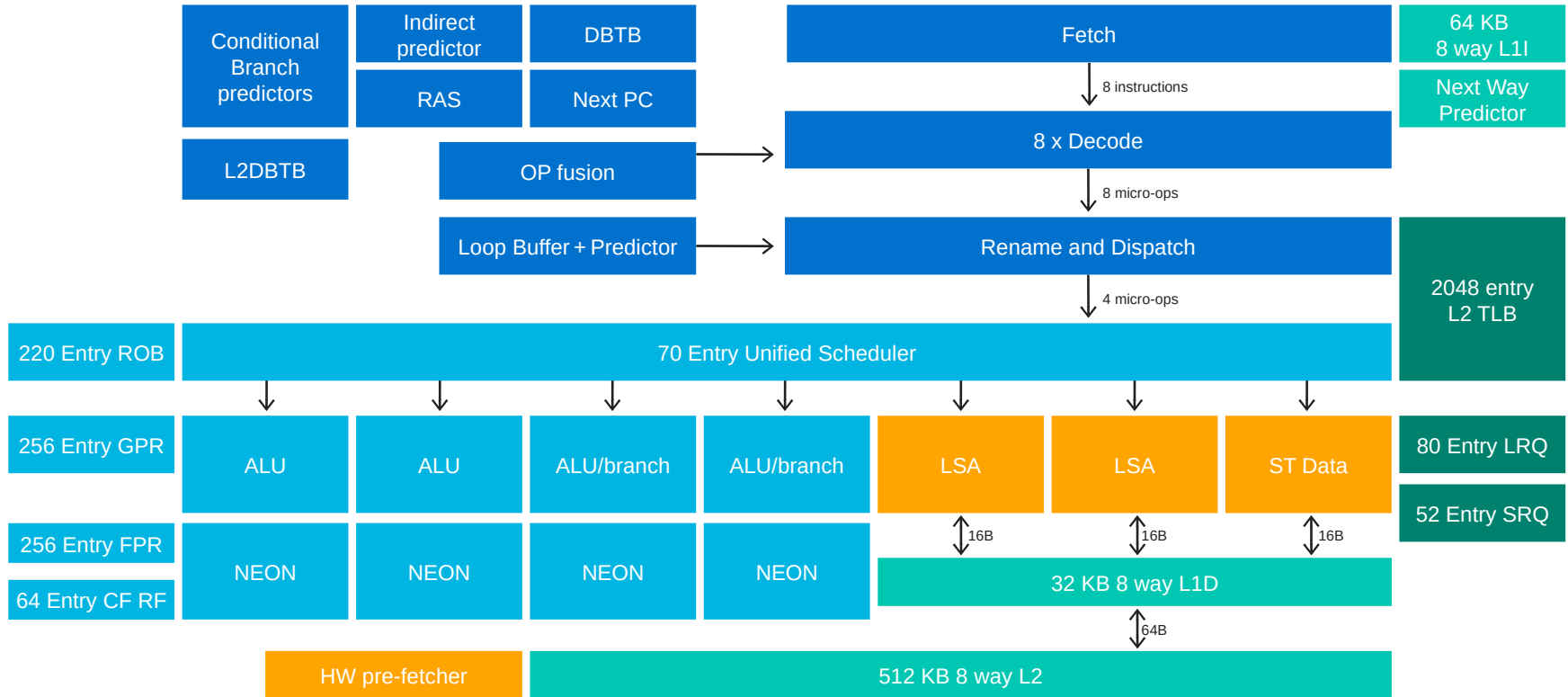


ThunderX3™ overview

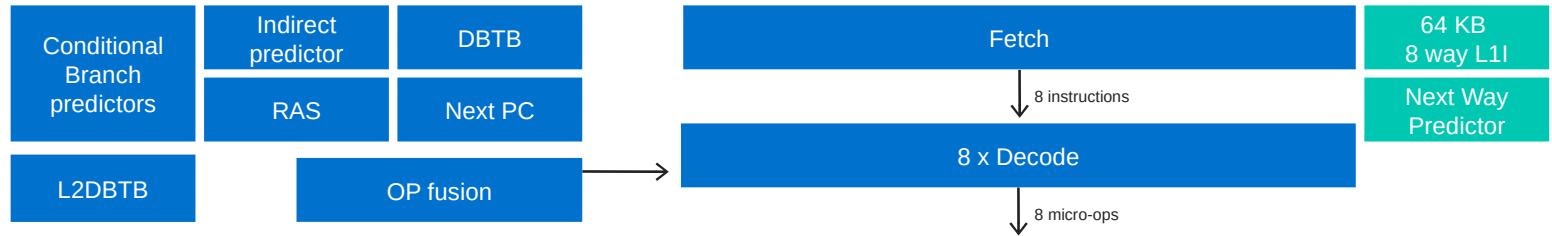
- Single die: Up to 60 cores
- Dual die: Up to 96 cores
- Arm v8.3 with select v8.4/v8.5 features
- 30% single thread gain at equal frequency over ThunderX2
- Up to four threads per core
- High bandwidth switched ring interconnect
- Up to 8 DDR4-3200 channels
- Single die: 2X-3X perf over ThunderX2 at equal power
 - Further gains from dual die
- Up to 64 PCIe Gen4, 16 PCIe controllers
- Fine grain power monitoring/management
- TSMC 7nm



ThunderX3 core block diagram



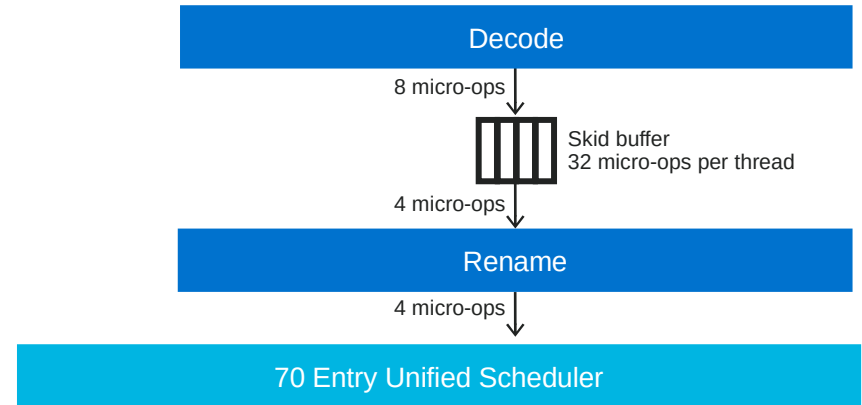
Core microarchitecture – Fetch



- 64KB Icache, 8 way set associative, 64B line size, next line pre-fetch, way prediction
- Decoupled fetch for large instruction footprint codes
- 8-wide instruction fetch
- Fetch breaks on 64B line boundary, or on a taken branch
- Large condition branch predictors, indirect and return address predictors
- Fetched bundle is decoded 8 instructions at a time
- Decode breaks a few instruction types into multiple micro-ops

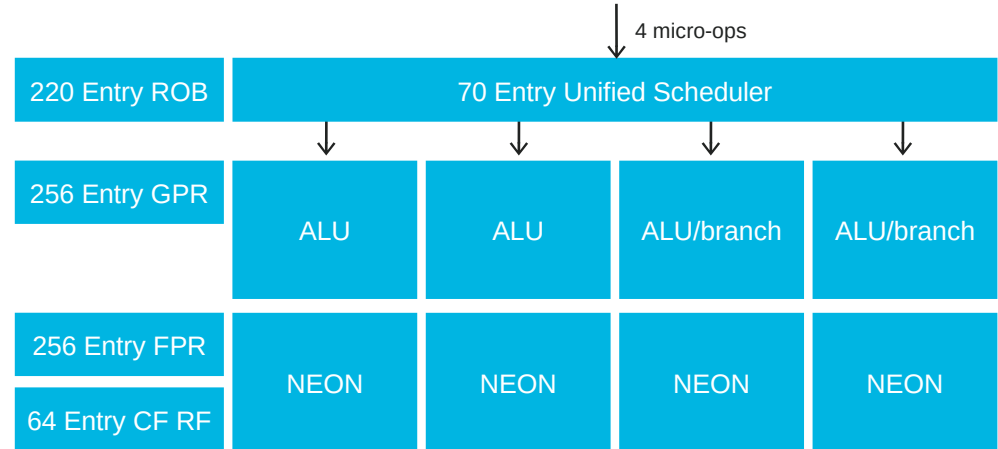
Core microarchitecture – Decode/dispatch

- Decoded micro-ops enter skid buffer –
 - Up to 8 micro-ops per cycle
- Each thread has a 32 micro-ops skid buffer – 8 four micro-ops bundles
- 4 micro-ops dispatched per cycle to scheduler
- NOP not dispatched to scheduler – Go to ROB and retire
- Some merging between bundles in skid buffer



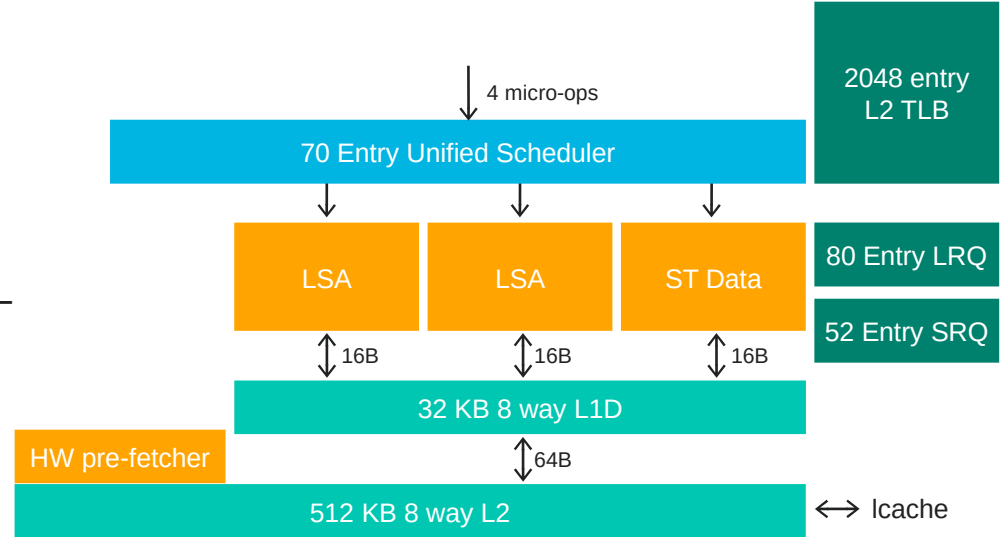
Core microarchitecture – Scheduler

- Out of order issue from unified issue queue
 - 70 entries
- Seven issue ports:
 - Port 0: ALU, FP/SIMD
 - Port 1: ALU, FP/SIMD, integer mul/div
 - Port 2: ALU, Branch, FP/SIMD
 - Port 3: ALU, Branch, FP/SIMD
 - Port 4: Ld/St
 - Port 5: Ld/St
 - Port 6: Store data



Core microarchitecture – D-cache / DTLB / L2-cache

- 32KB D-cache, 8-way associative, 64B line size, write back
- Small L1TLBs for zero impact translation in the common case
- 2K entry L2 TLB, 8-way associative
- 512KB L2-cache, 8-way associative – private to core
 - Larger L2-cache increases area and latency with minor incremental performance benefit
- Hardware prefetcher into L2-cache
 - Next line
 - Strides
 - Region

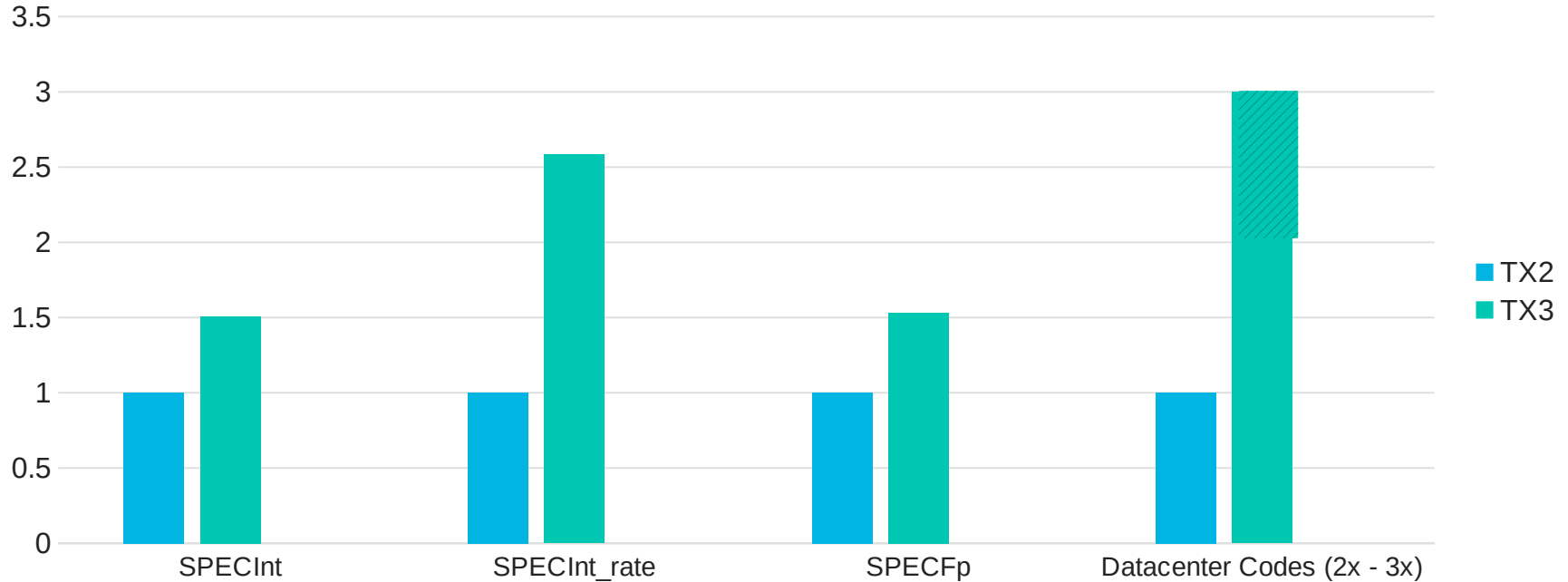


ThunderX3 core performance enhancements over ThunderX2

	Feature	Approx. pct gain over ThunderX2 (SPECInt)
Size	Icache Size	0.5%
	512KB L2-cache	2.5%
	Larger out-of-order structures	5%
Width	Wider decode	2%
	Additional ALU port	1.5%
	Two branches per cycle	1.5%
Algorithm	Branch prediction enhancements	3%
	Front end resteer enhancements	1%
	Reduce micro-op expansion	6%
	D-cache bank conflict reduction	0.5%
	Reduce FP structural hazards	1%
	Prefetch enhancements	1.5%
Latency	FP latency reduction	0.5%

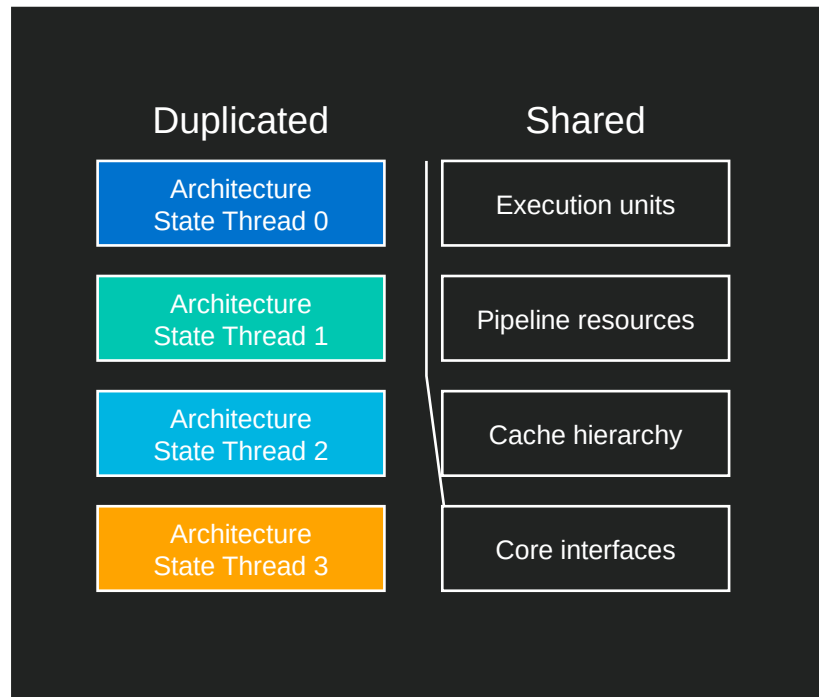
ThunderX3 performance – Single die

Substantial performance gains



Multithread execution

- Four hardware threads per core
- Each thread includes full copy of Arm architecture state
- Threads share core pipeline resources
- To OS each thread appears as a regular Arm CPU
 - So four CPUs per core
- Area impact of 4-way SMT relative to no SMT: ~5%
- ThunderX3 has 60 cores / 240 threads per die



Thread arbitration

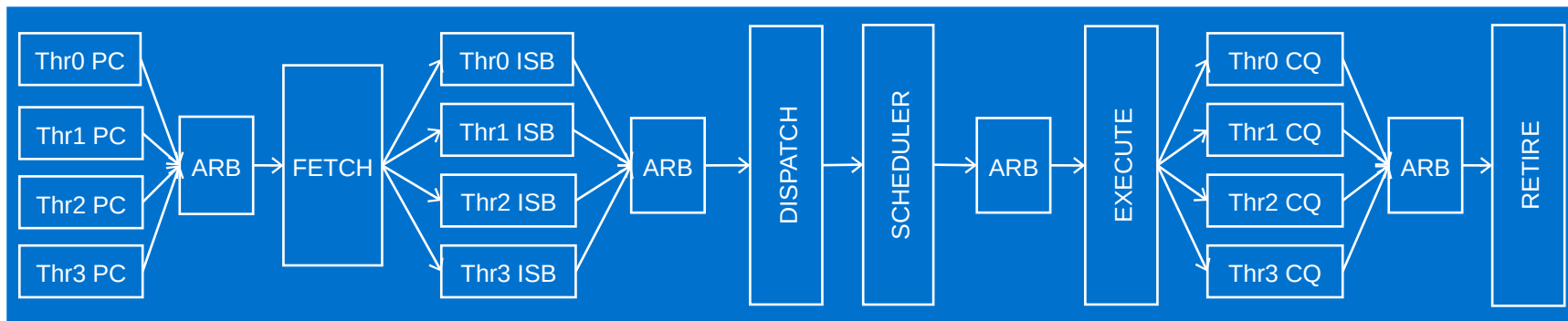
Goals

- Fair sharing of pipeline resources among threads
- Maximize pipeline utilization

Four points of arbitration:

- **Fetch:** Prioritize threads with fewer instructions in pipeline over threads with more instructions
- **Dispatch:** Similar to Fetch but just considering stages after Dispatch
- **Scheduler (issue):** Age based priority
- **Retire:** Favor threads with more instructions to retire

Dynamic sharing of caches, branch predictor structures



Multithread scaling performance – Single core

Low IPC (~0.5)

MySQL	1 thread	2 threads	4 threads
Relative performance	1.00	1.79	2.21

Medium IPC (~1.25)

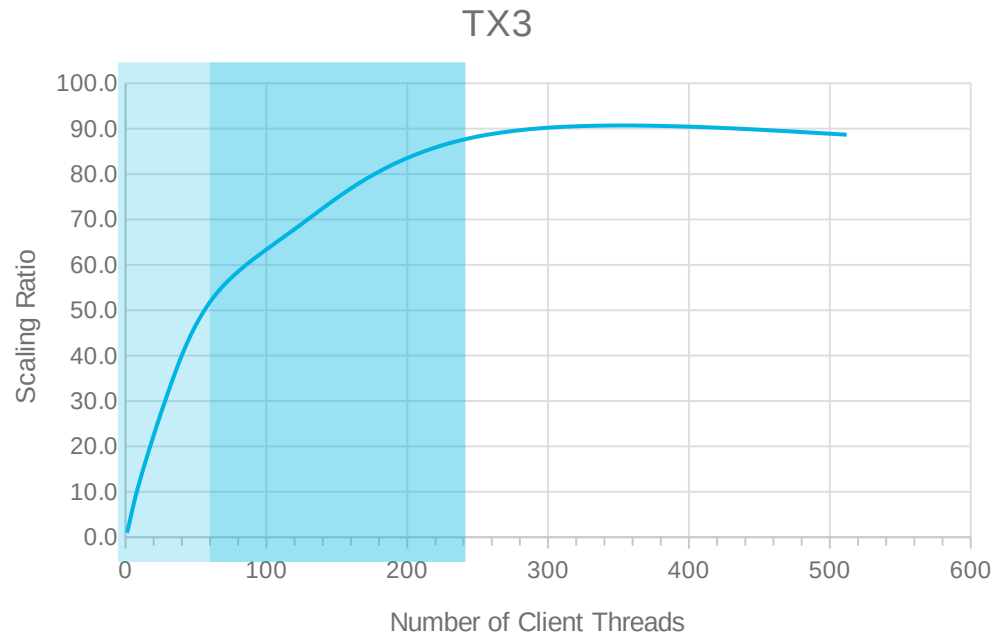
Leela	1 thread	2 threads	4 threads
Relative performance	1.00	1.38	1.73

High IPC (>2)

X264	1 thread	2 threads	4 threads
Relative performance	1.00	1.18	1.28

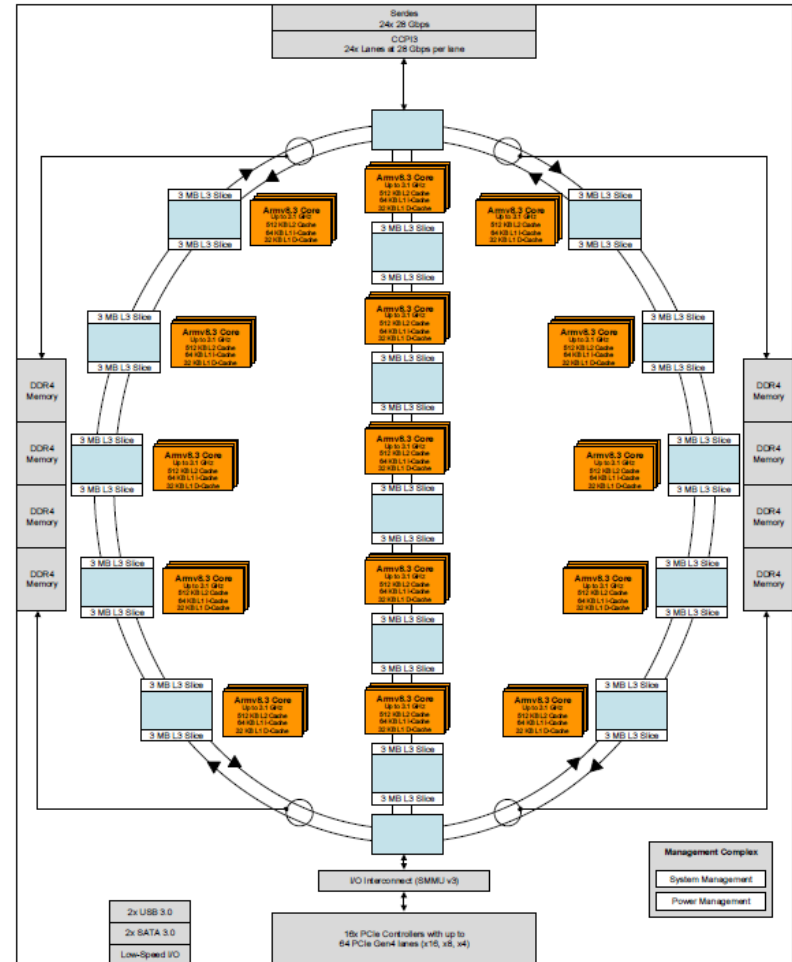
Socket level performance – MySQL

- Roughly linear scaling in Core region
- Scaling flattens out in threaded region, but still good gains
- Net 89x over single thread



L3-Cache and interconnect

- Cores / L3-caches organized as switched rings
- DDR channels, I/O tap into rings
- L3-cache organized as tiles that are cache line striped
 - 1 ½ MB per core
 - No notion of L3 cache affinity to cores
 - Good for shared text and shared data
- Exclusive L3-cache – filled on evict from L2-cache
- Snoop based coherence with snoop filters
 - Single socket and two socket



ThunderX3 Arm-based server processor summary

1

Up to 3x performance over industry-leading ThunderX2 within same power envelope

2

Evolutionary design approach – leverage ThunderX2 platform and proven production quality solution at scale

3

Four-way threading provides ~50% performance advantage on data center codes over competitor systems

4

Marvell offers a competitive solution with technology roadmap built on legacy of processor expertise



Thank You



Essential technology, done right™