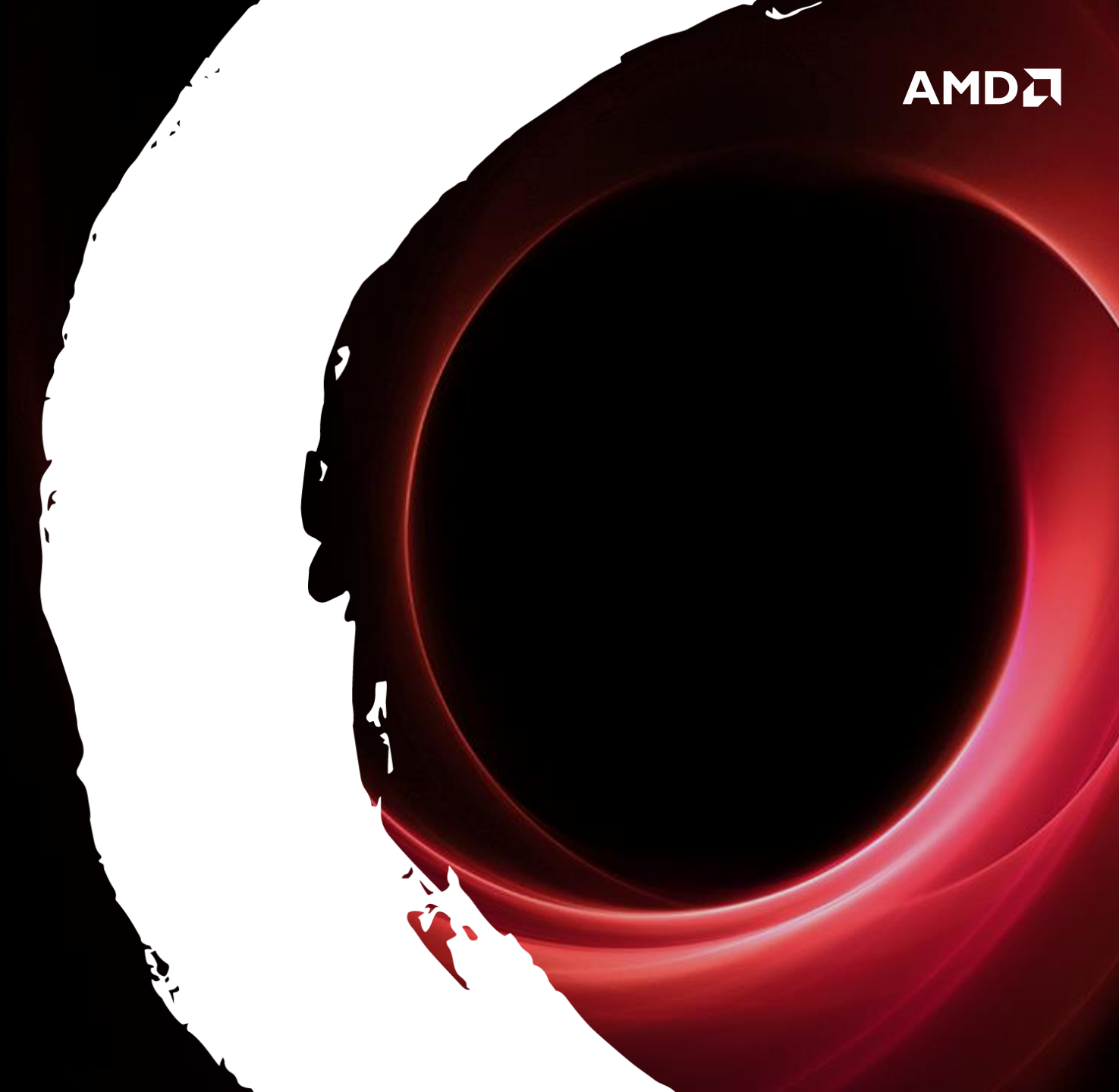
A NEW X86 CORE ARCHITECTURE FOR THE NEXT GENERATION OF COMPUTING

MIKE CLARK
SENIOR FELLOW



AGENDA

THE ROAD TO ZEN

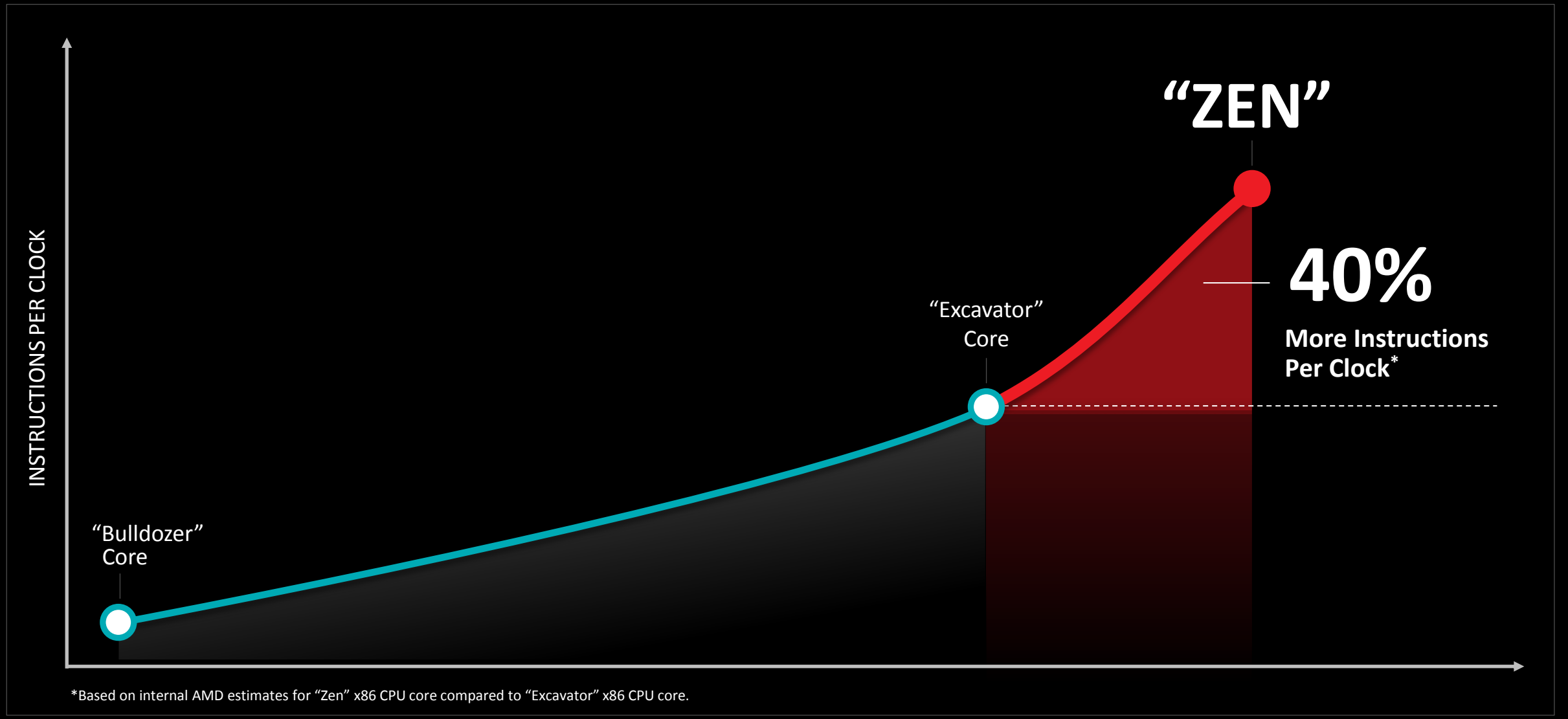
HIGH LEVEL ARCHITECTURE

- IMPROVEMENTS IN CORE ENGINE
- FLOATING POINT
- IMPROVEMENTS IN CACHE SYSTEM
- SMT DESIGN TO MAXIMIZE THROUGHPUT
- NEW ISA EXTENSIONS

SUMMARY

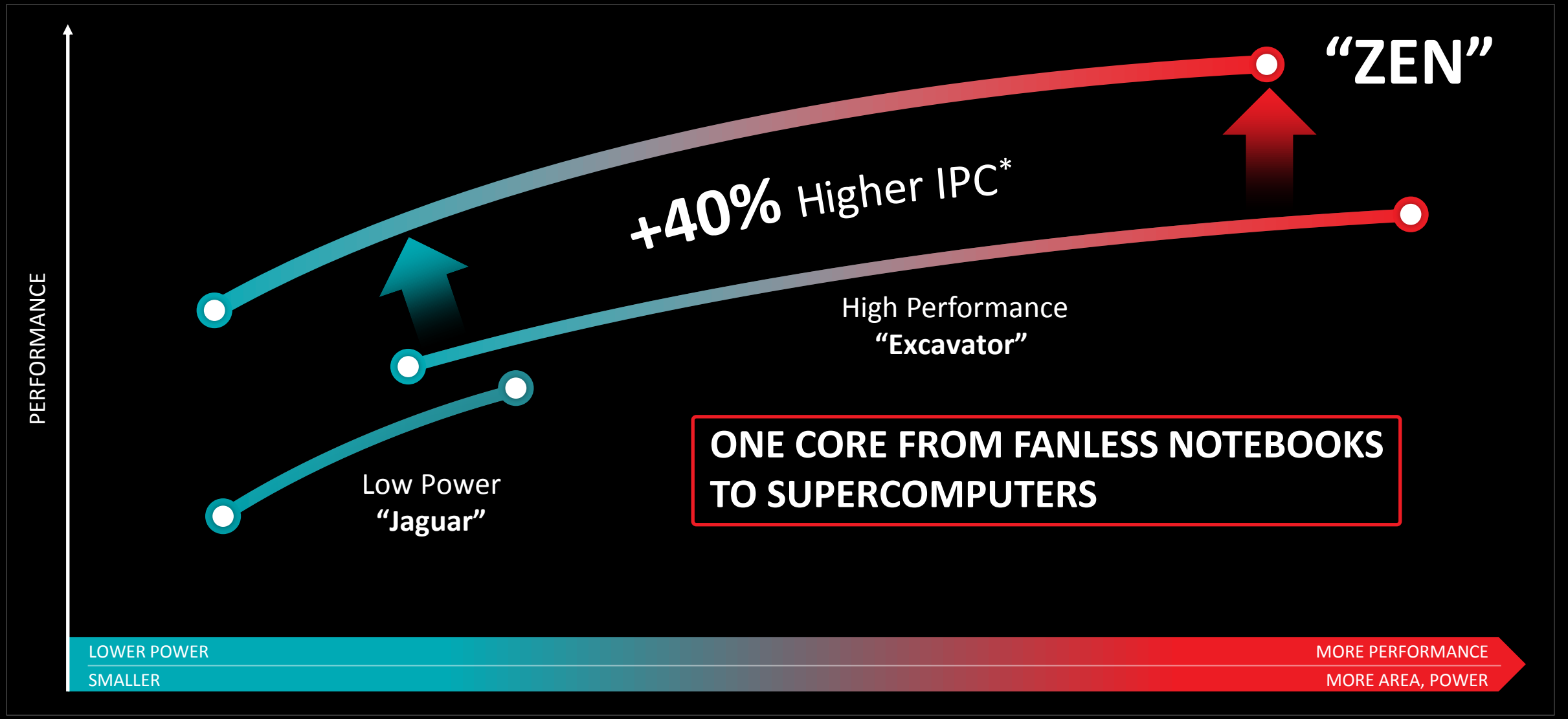
NEXT STEP UP

AMD X86 CORES: DRIVING COMPETITIVE PERFORMANCE



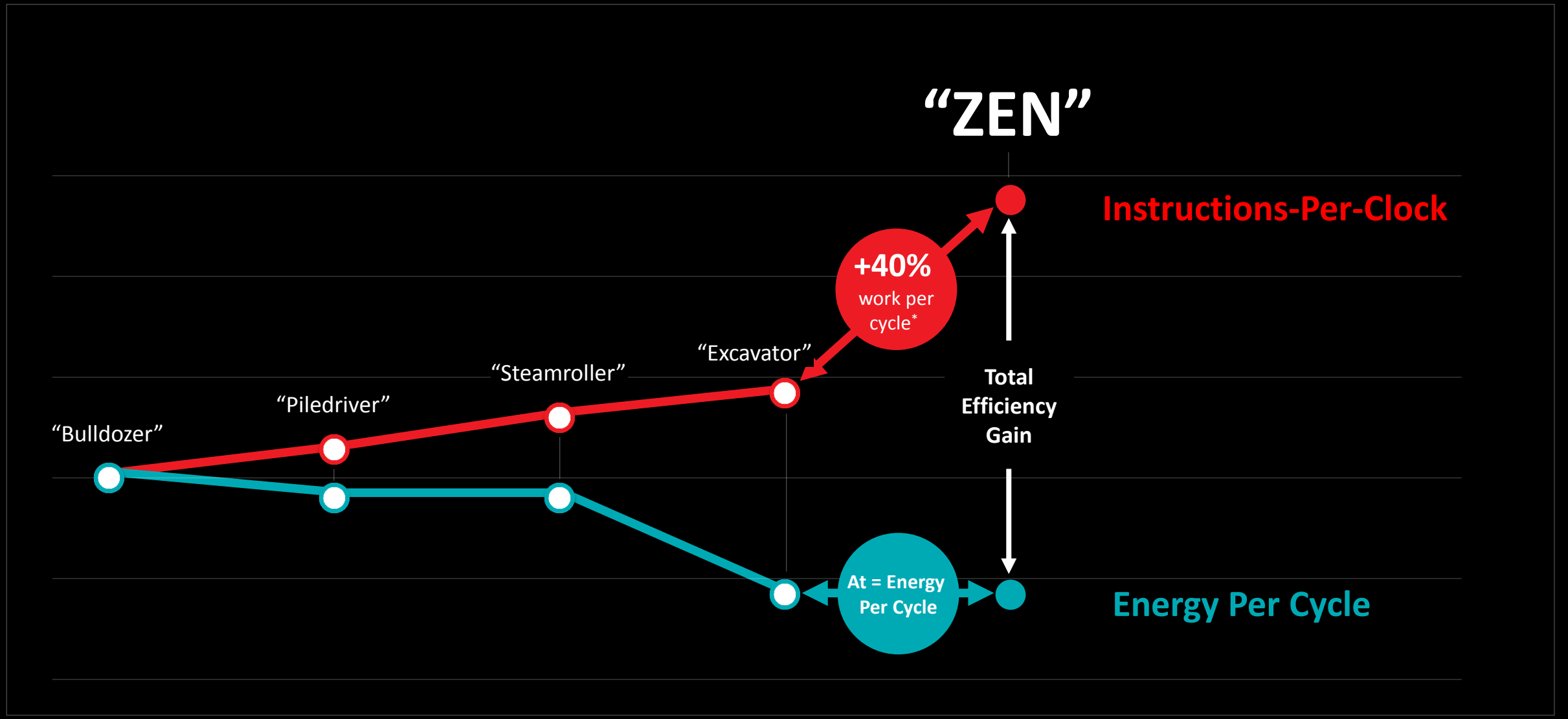
*Based on internal AMD estimates for "Zen" x86 CPU core compared to "Excavator" x86 CPU core.

AMD CPU DESIGN OPTIMIZATION POINTS



*Based on internal AMD estimates for "Zen" x86 CPU core compared to "Excavator" x86 CPU core.

DEFYING CONVENTION: A WIDE, HIGH PERFORMANCE, EFFICIENT CORE



*Based on internal AMD estimates for "Zen" x86 CPU core compared to "Excavator" x86 CPU core.

ZEN PERFORMANCE & POWER IMPROVEMENTS

AMD

BETTER CORE ENGINE

- Two threads per core
- Branch mispredict improved
- Better branch prediction with 2 branches per BTB entry
- Large Op Cache
- Wider micro-op dispatch 6 vs. 4
- Larger Instruction Schedulers Integer: 84 vs. 48 | FP: 96 vs. 60
- Larger retire 8 ops vs. 4 ops
- Quad issue FPU
- Larger Retire Queue 192 vs. 128
- Larger Load Queue 72 vs. 44
- Larger Store Queue 44 vs. 32

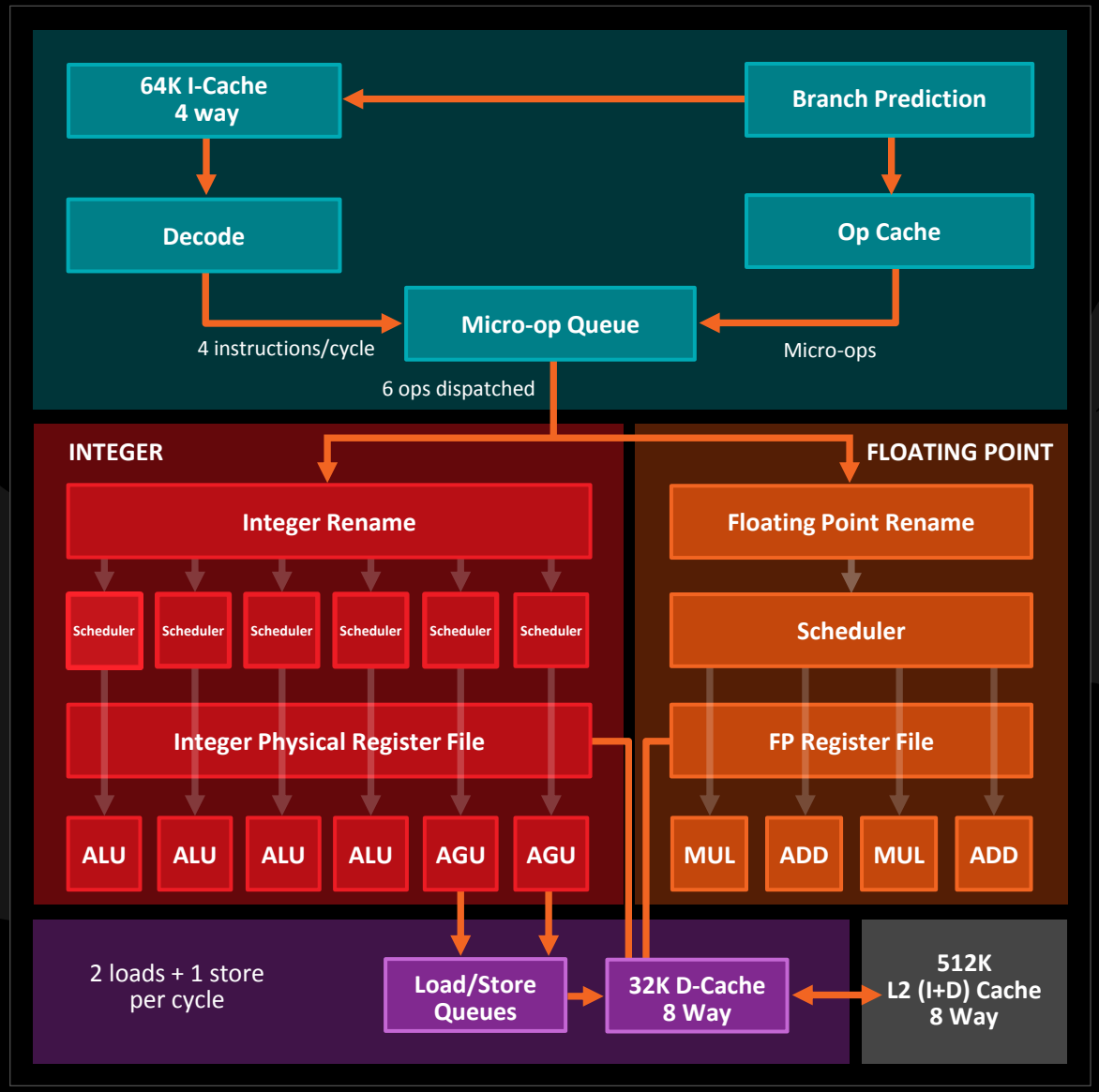
BETTER CACHE SYSTEM

- Write back L1 cache
- Faster L2 cache
- Faster L3 cache
- Faster Load to FPU: 7 vs. 9 cycles
- Better L1 and L2 data prefetcher
- Close to 2x the L1 and L2 bandwidth
- Total L3 bandwidth up 5x

LOWER POWER

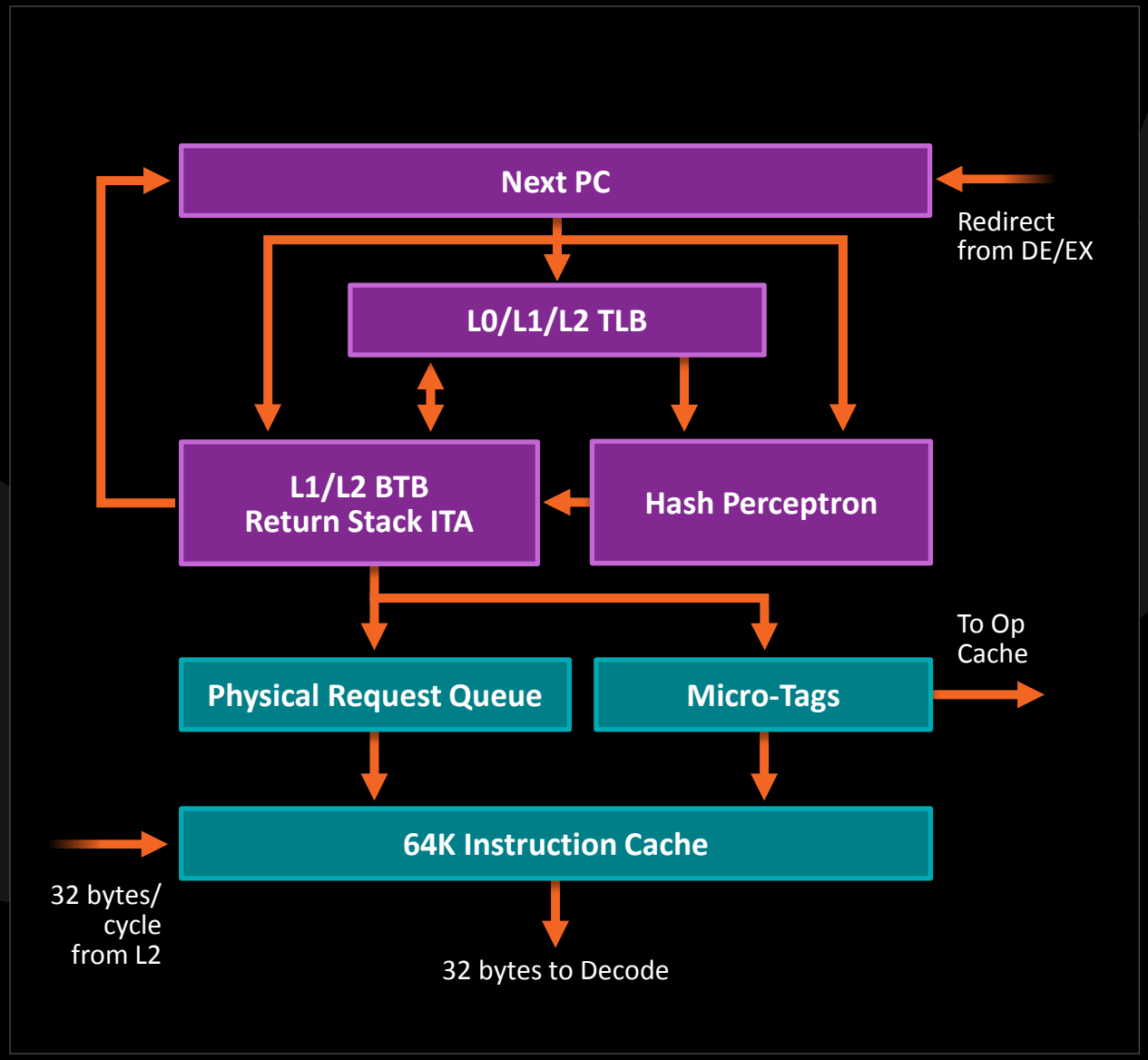
- Aggressive clock gating with multi-level regions
- Write back L1 cache
- Large Op Cache
- Stack Engine
- Move elimination
- Power focus from project inception
- Low Power Design Methodologies

40% IPC PERFORMANCE UPLIFT



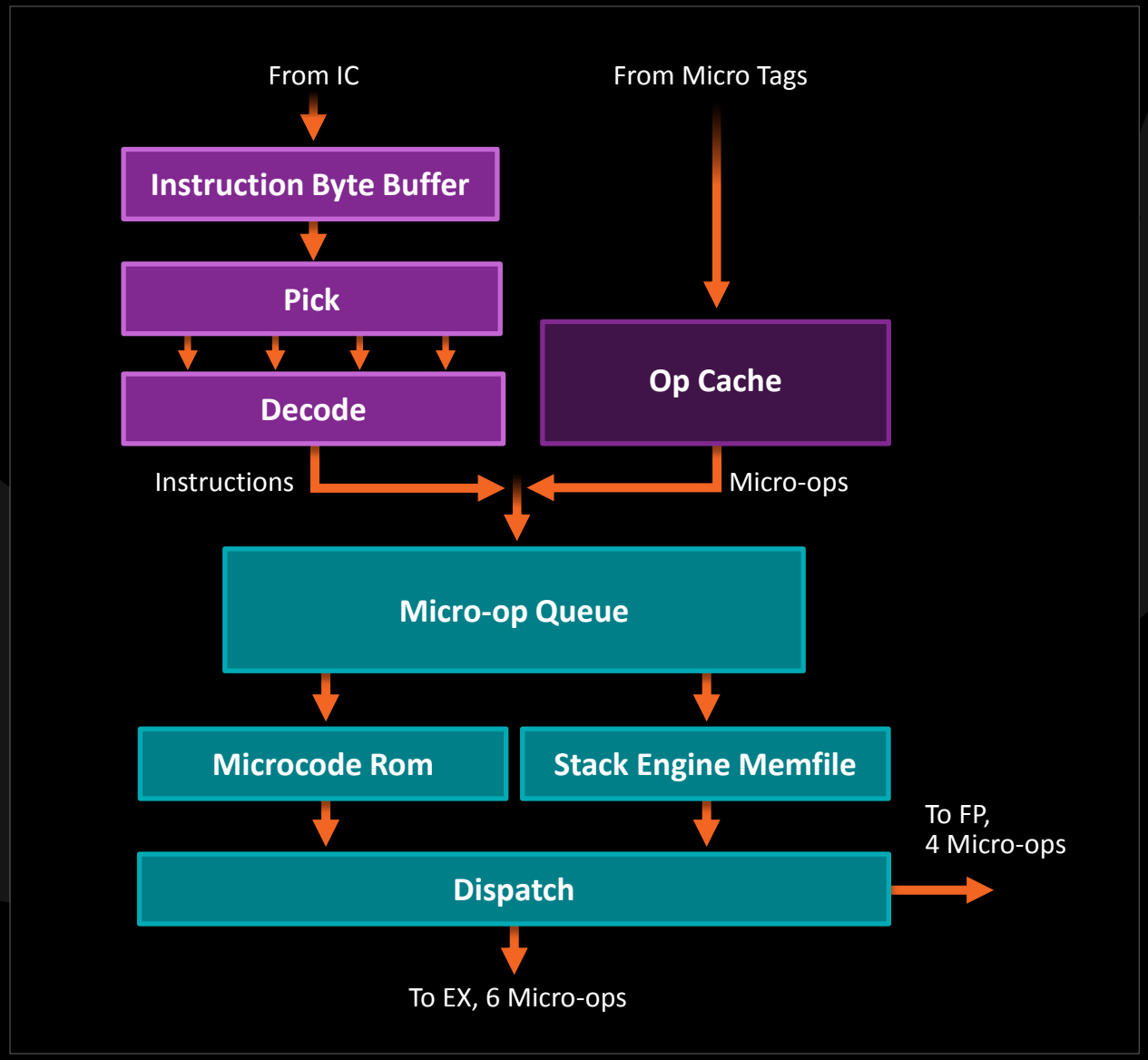
ZEN MICROARCHITECTURE

- Fetch Four x86 instructions
- Op Cache instructions
- 4 Integer units
  - Large rename space – 168 Registers
  - 192 instructions in flight/8 wide retire
- 2 Load/Store units
  - 72 Out-of-Order Loads supported
- 2 Floating Point units x 128 FMACs
  - built as 4 pipes, 2 Fadd, 2 Fmul
- I-Cache 64K, 4-way
- D-Cache 32K, 8-way
- L2 Cache 512K, 8-way
- Large shared L3 cache
- 2 threads per core



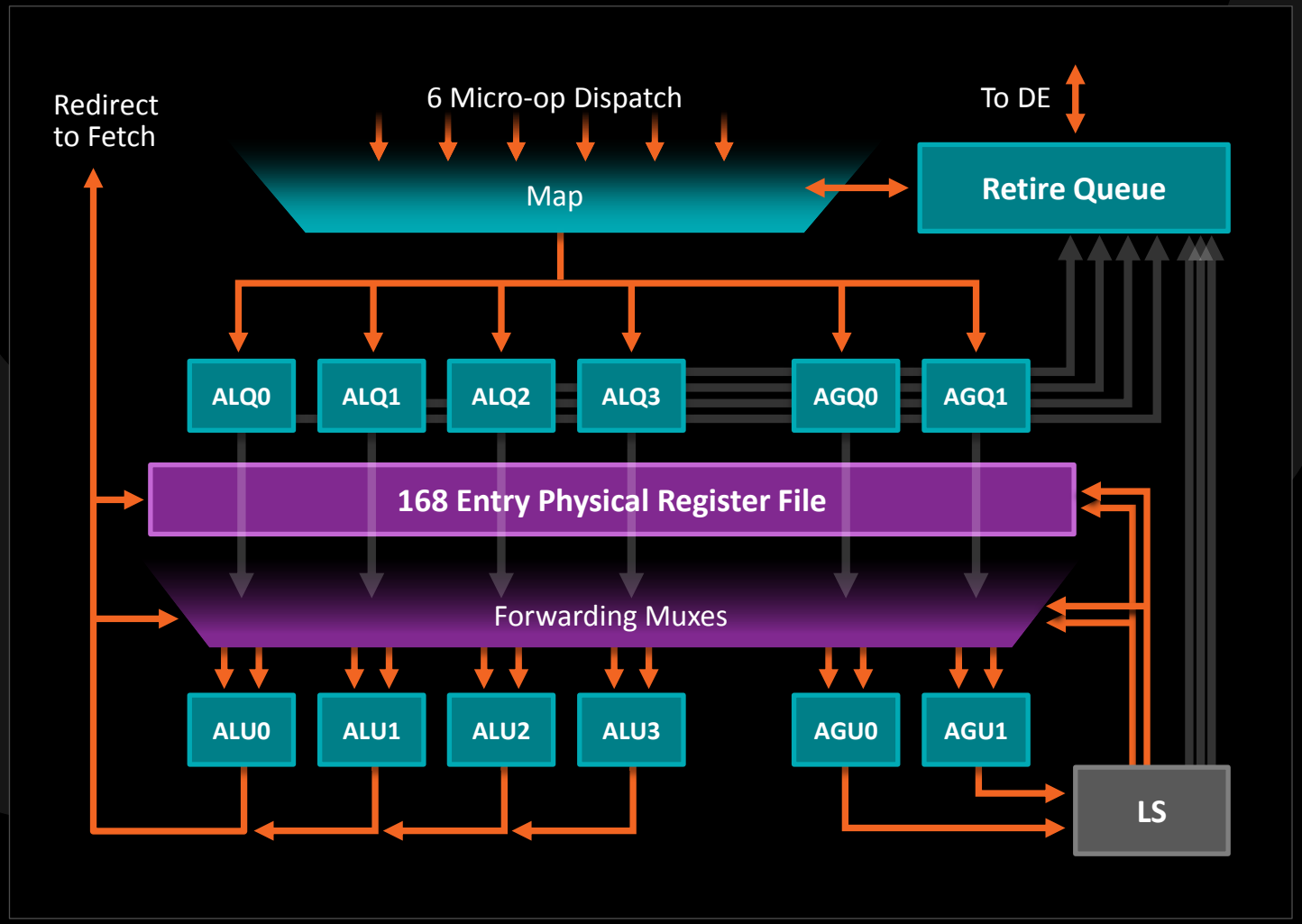
FETCH

- Decoupled Branch Prediction
- TLB in the BP pipe
  - 8 entry L0 TLB, all page sizes
  - 64 entry L1 TLB, all page sizes
  - 512 entry L2 TLB, no 1G pages
- 2 branches per BTB entry
- Large L1 / L2 BTB
- 32 entry return stack
- Indirect Target Array (ITA)
- 64K, 4-way Instruction cache
- Micro-tags for IC & Op cache
- 32 byte fetch



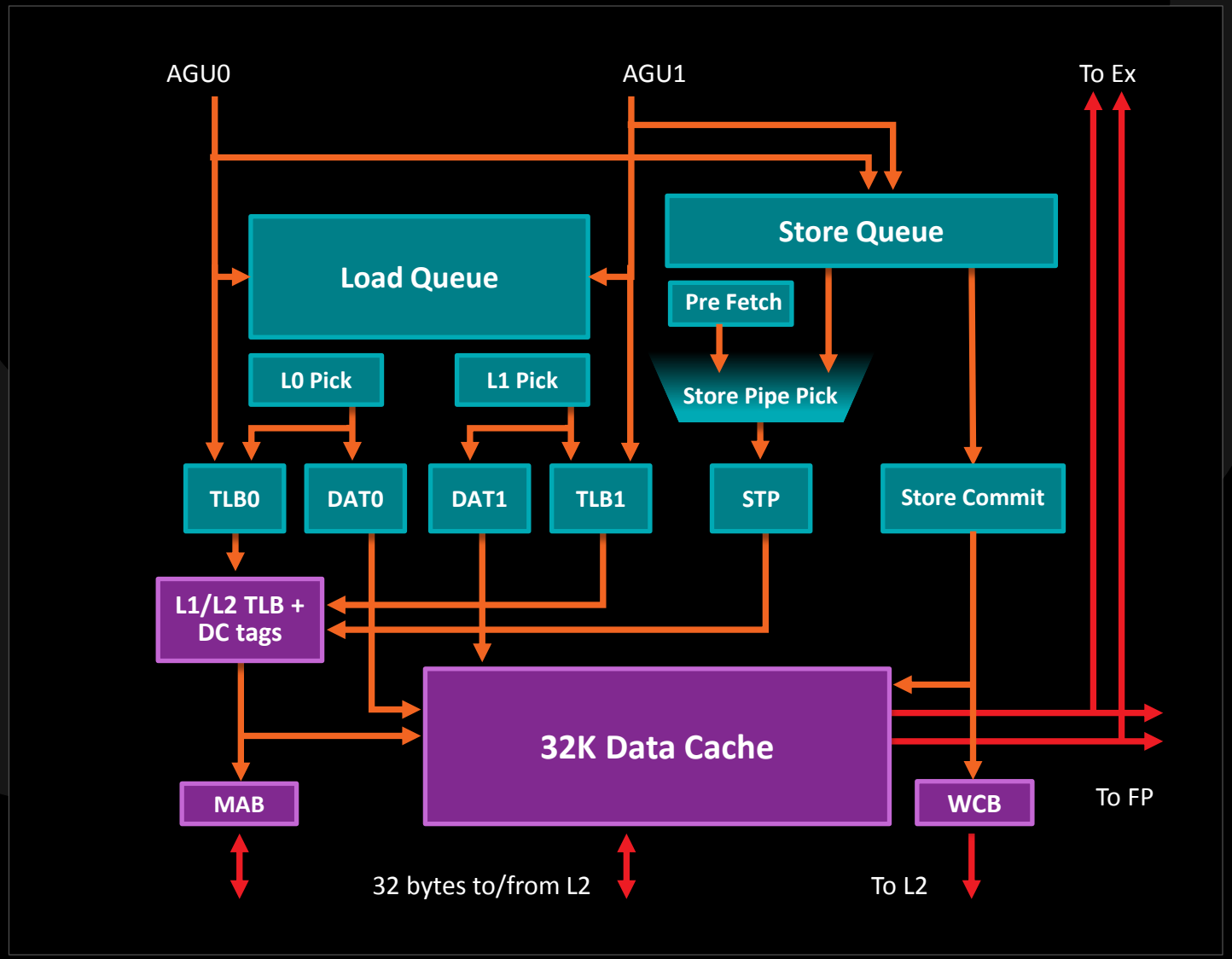
DECODE

- Inline Instruction-length Decoder
- Decode 4 x86 instructions
- Op cache
- Micro-op Queue
- Stack Engine
- Branch Fusion
- Memory File for Store to Load Forwarding



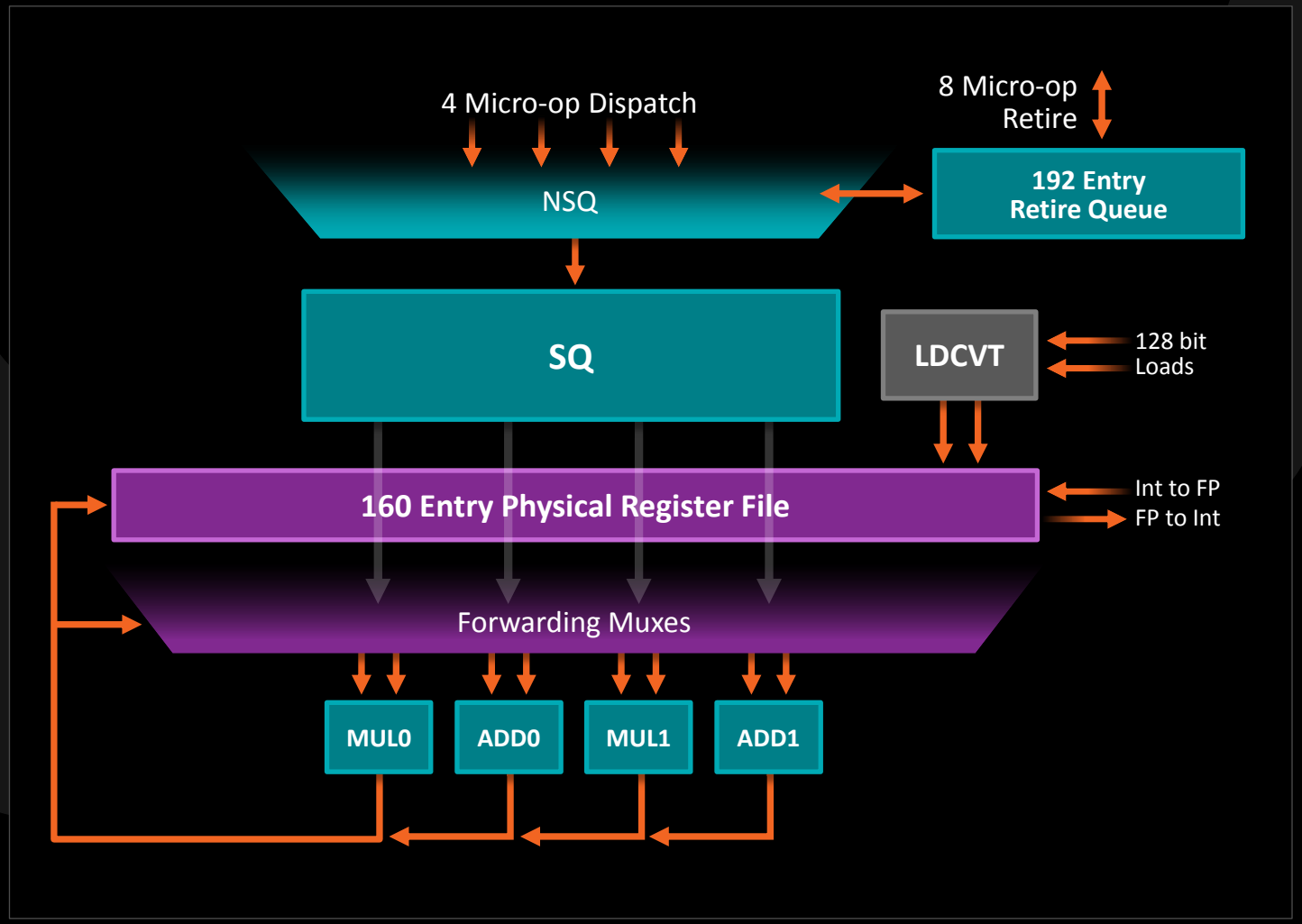
EXECUTE

- 6x14 entry Scheduling Queues
- 168 entry Physical Register File
- 6 issue per cycle
  - 4 ALU's, 2 AGU's
- 192 entry Retire Queue
- Differential Checkpoints
- 2 Branches per cycle
- Move Elimination
- 8-Wide Retire



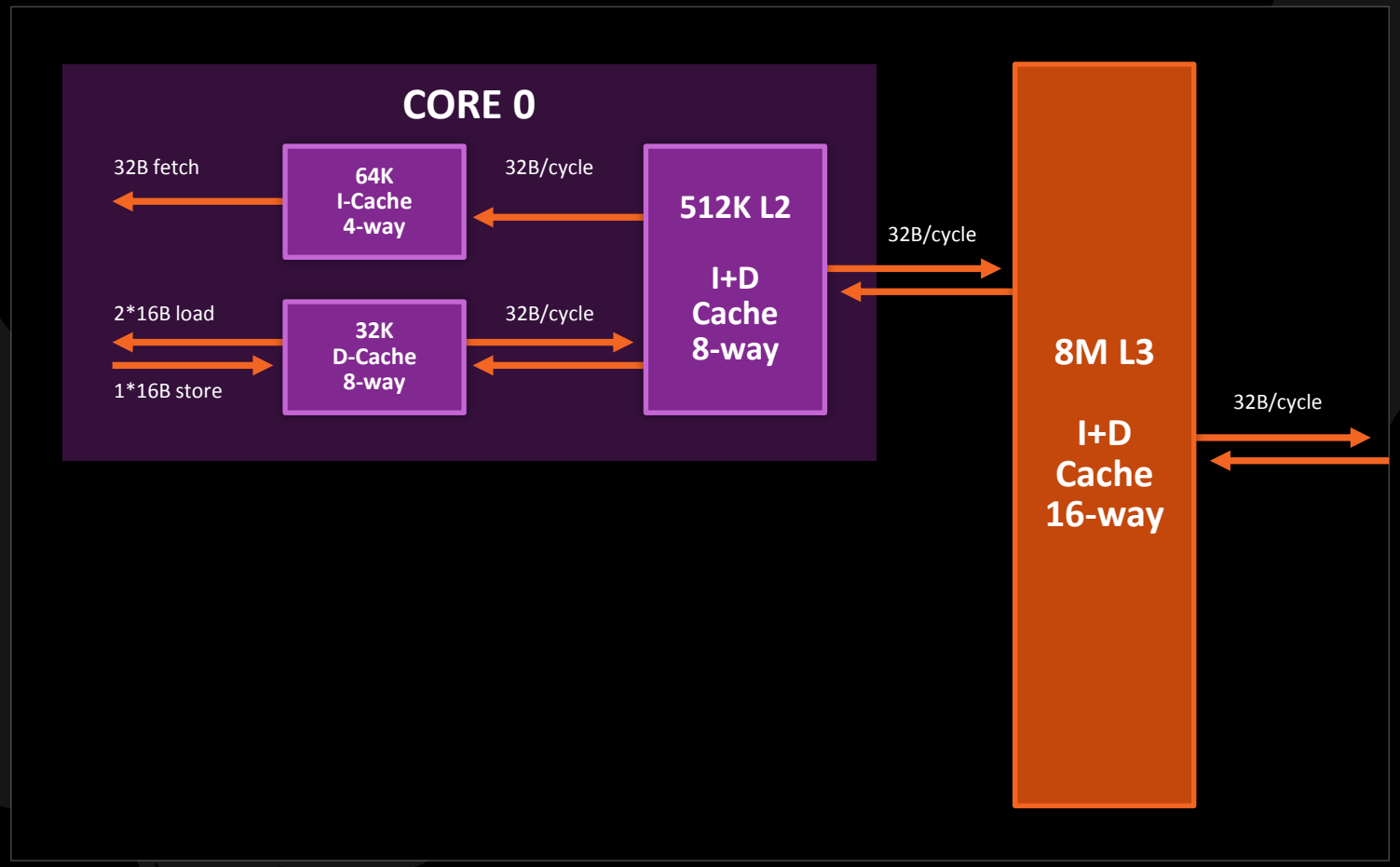
LOAD/STORE AND L2

- 72 Out of Order Loads
- 44 entry Store Queue
- Split TLB/Data Pipe, store pipe
- 64 entry L1 TLB, all page sizes
- 1.5K entry L2 TLB, no 1G pages
- 32K, 8 way Data Cache
  - Supports two 128-bit accesses
- Optimized L1 and L2 Prefetchers
- 512K, private (2 threads), inclusive L2



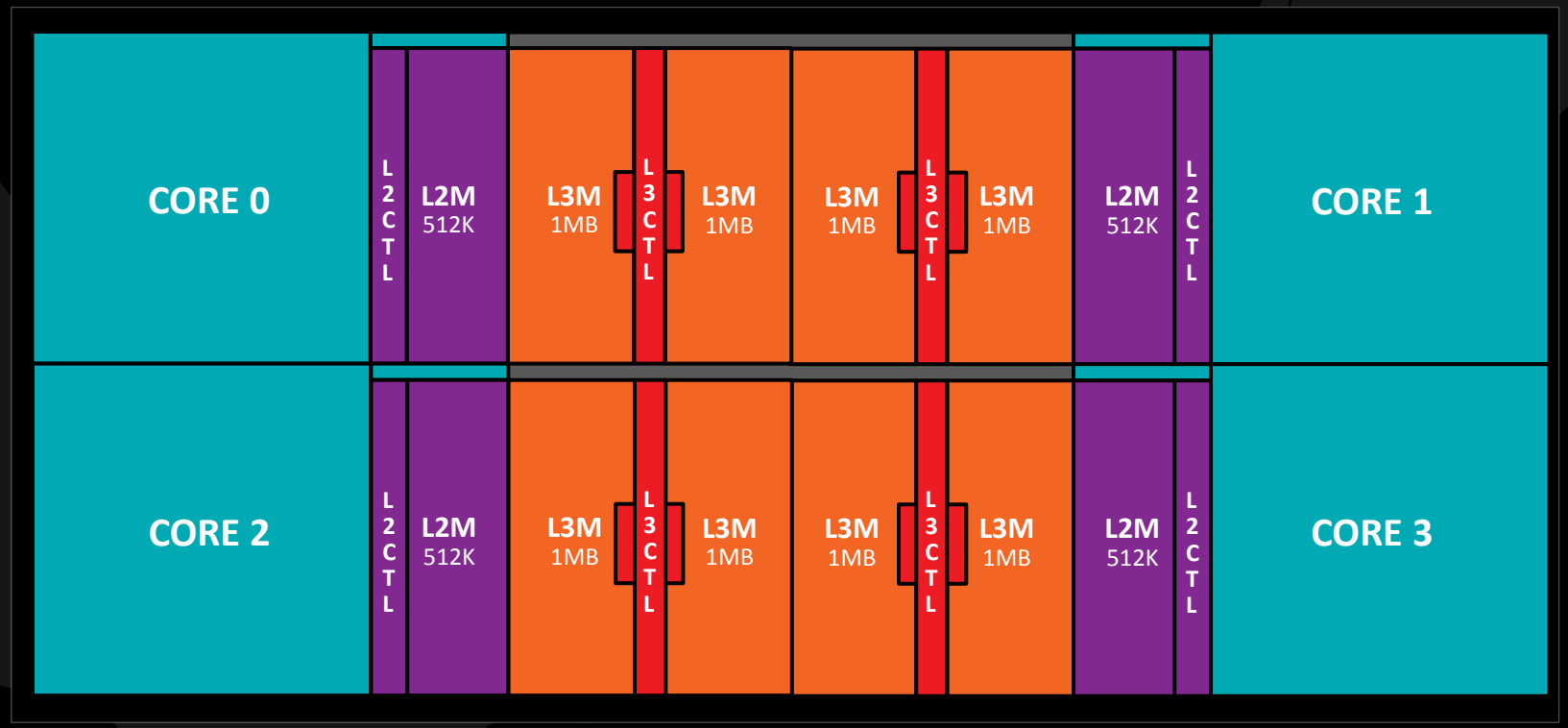
FLOATING POINT

- 2 Level Scheduling Queue
- 160 entry Physical Register File
- 8 Wide Retire
- 1 pipe for 1x128b store
- Accelerated Recovery on Flushes
- SSE, AVX1, AVX2, AES, SHA, and legacy mmx/x87 compliant
- 2 AES units



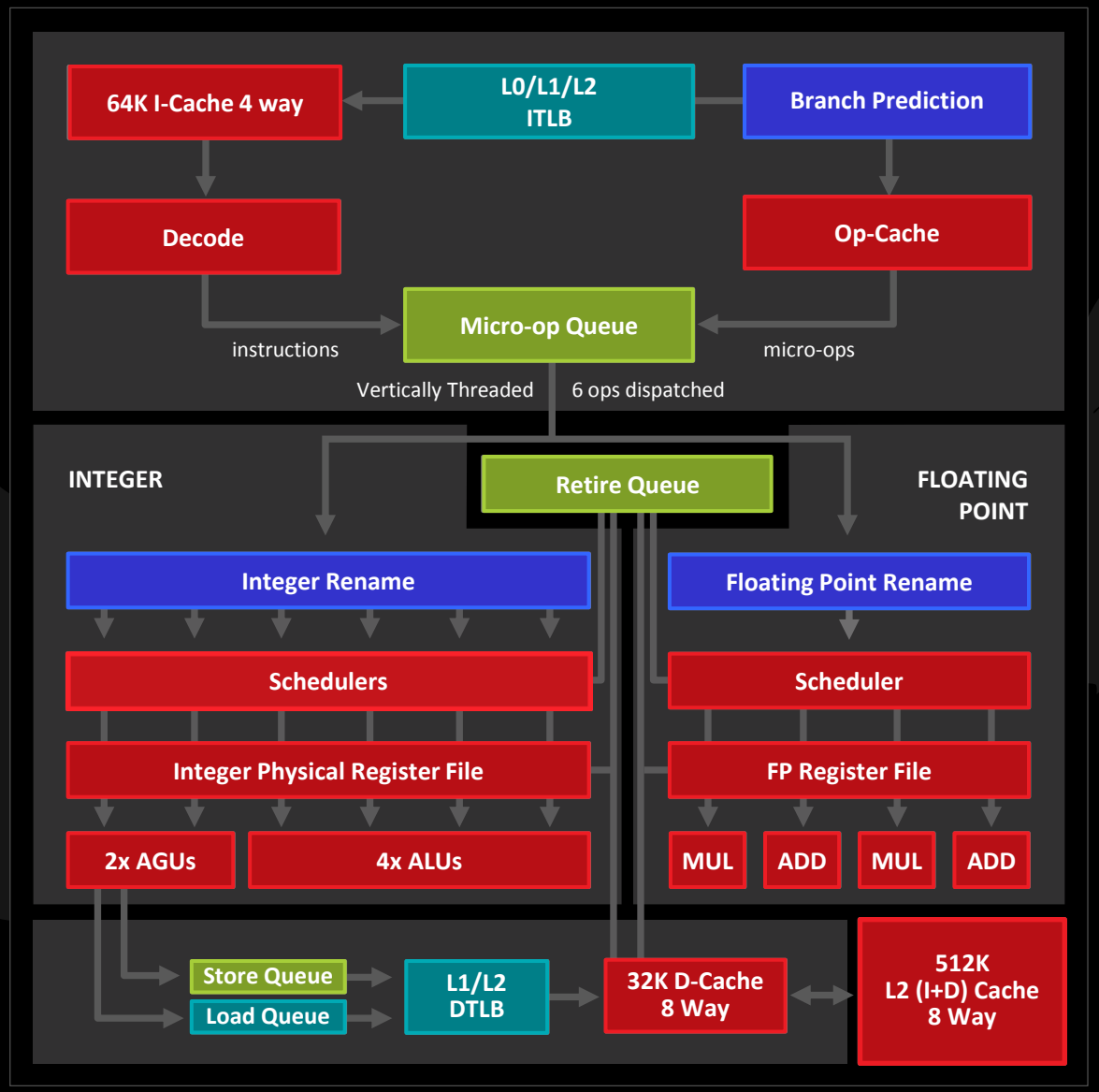
ZEN CACHE HIERARCHY

- Fast private 512K L2 cache
- Fast shared L3 cache
- High bandwidth enables prefetch improvements
- L3 is filled from L2 victims
- Fast cache-to-cache transfers
- Large Queues for Handling L1 and L2 misses



CPU COMPLEX

- A CPU complex (CCX) is four cores connected to an L3 Cache.
- The L3 Cache is 16-way associative, 8MB, mostly exclusive of L2.
- The L3 Cache is made of 4 slices, by low-order address interleave.
- Every core can access every cache with same average latency



SMT OVERVIEW

- All structures fully available in 1T mode
- Front End Queues are round robin with priority overrides
- Increased throughput from SMT

- Competitively shared structures
- Competitively shared and SMT Tagged
- Competitively shared with Algorithmic Priority
- Statically Partitioned

NEW INSTRUCTIONS

AMD

Feature	Notes	Excavator	Zen
ADX	Extending multi-precision arithmetic support		✓
RDSEED	Complement to RDRAND random number generation		✓
SMAP	Supervisor Mode Access Prevention		✓
SHA1/SHA256	Secure Hash Implementation Instructions		✓
CLFLUSHOPT	CLFLUSH ordered by SFENCE		✓
XSAVEC/XSAVES/XRSTORS	New Compact and Supervisor Save/Restore		✓
CLZERO	Clear Cache Line		✓
PTE Coalescing	Combines 4K page tables into 32K page size		✓

AMD Exclusive

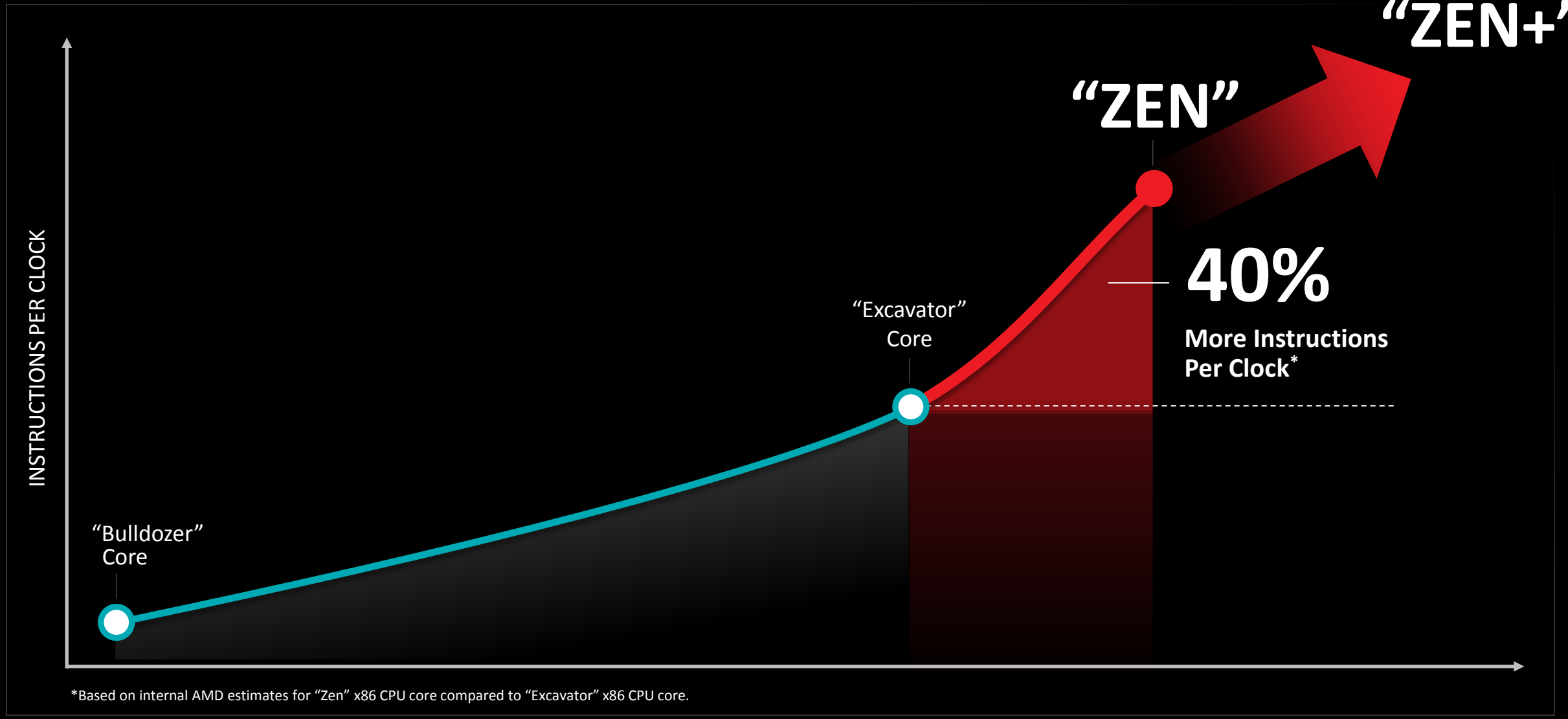
We support all the standard ISA including AVX &AVX-2, BMI1 & BMI2, AES, RDRAND, SMEP

"ZEN"

DESIGNED FROM THE GROUND UP FOR OPTIMAL BALANCE OF PERFORMANCE AND POWER

- Totally New High-performance Core Design
- New High-Bandwidth, Low Latency Cache System
- Simultaneous Multithreading (SMT) for High Throughput
- Energy-efficient FinFET Design Scales from Enterprise to Client Products

A COMMITTED ROADMAP TO X86 PERFORMANCE



*Based on internal AMD estimates for "Zen" x86 CPU core compared to "Excavator" x86 CPU core.

AMD

DISCLAIMER & ATTRIBUTION

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ATTRIBUTION