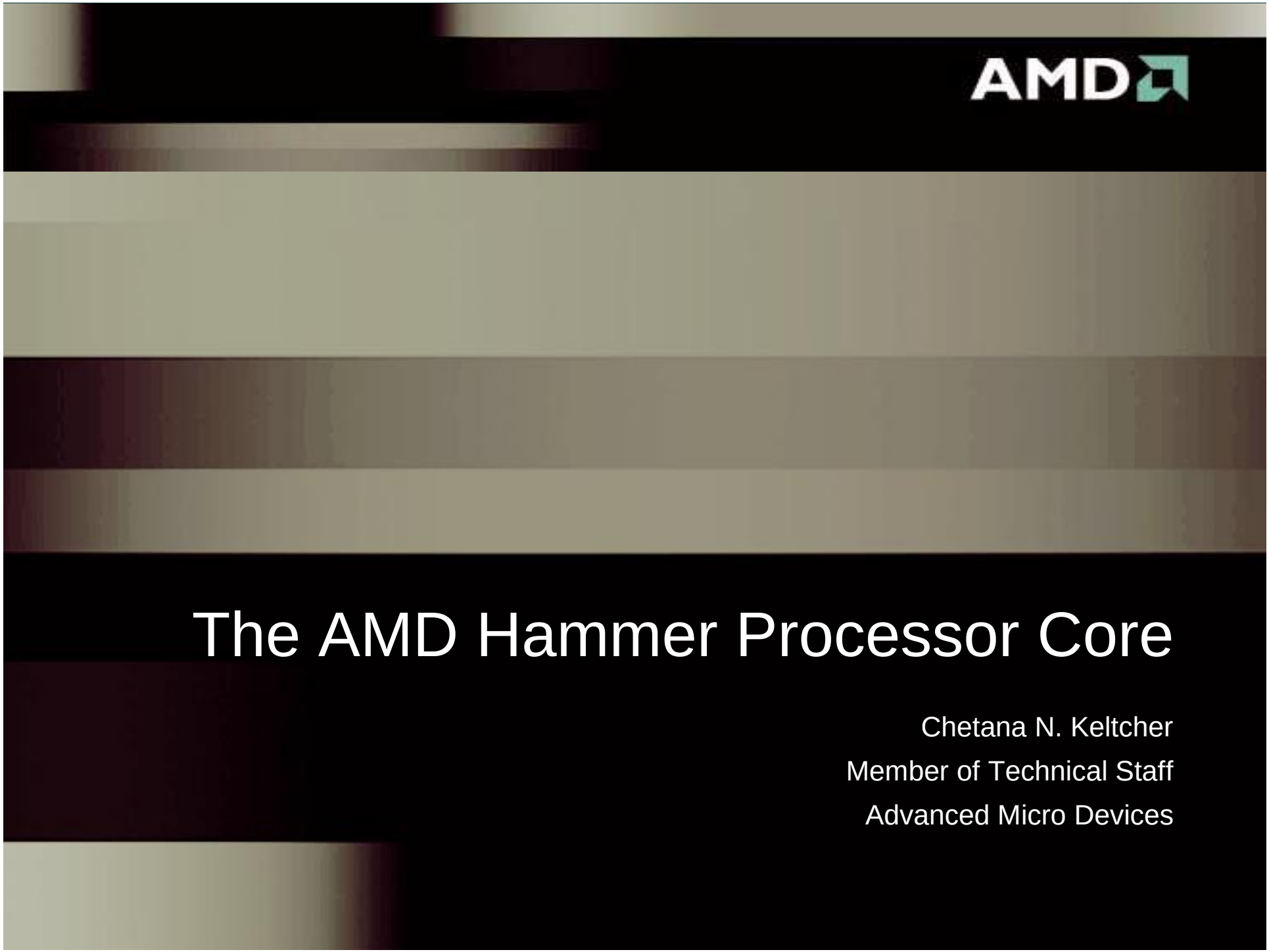
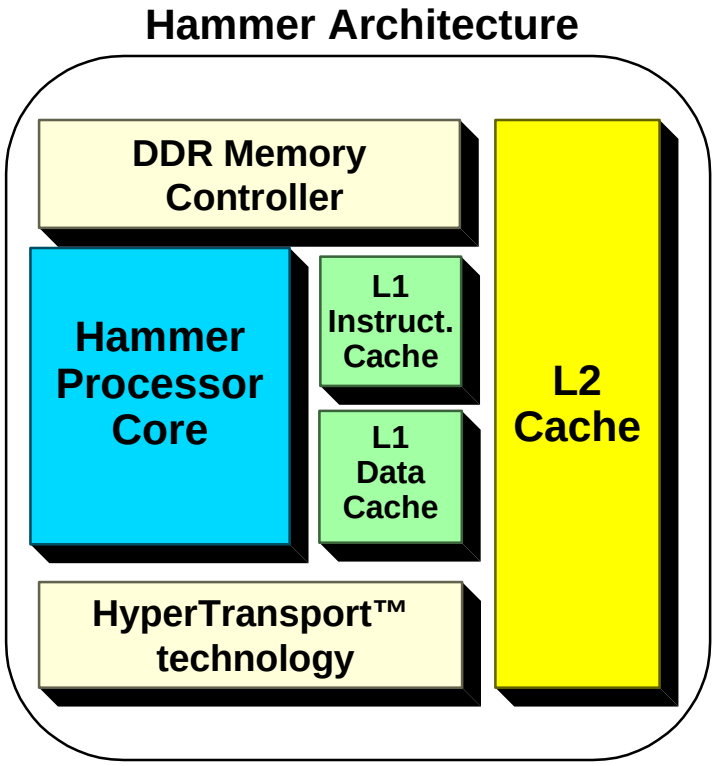


The AMD Hammer Processor Core

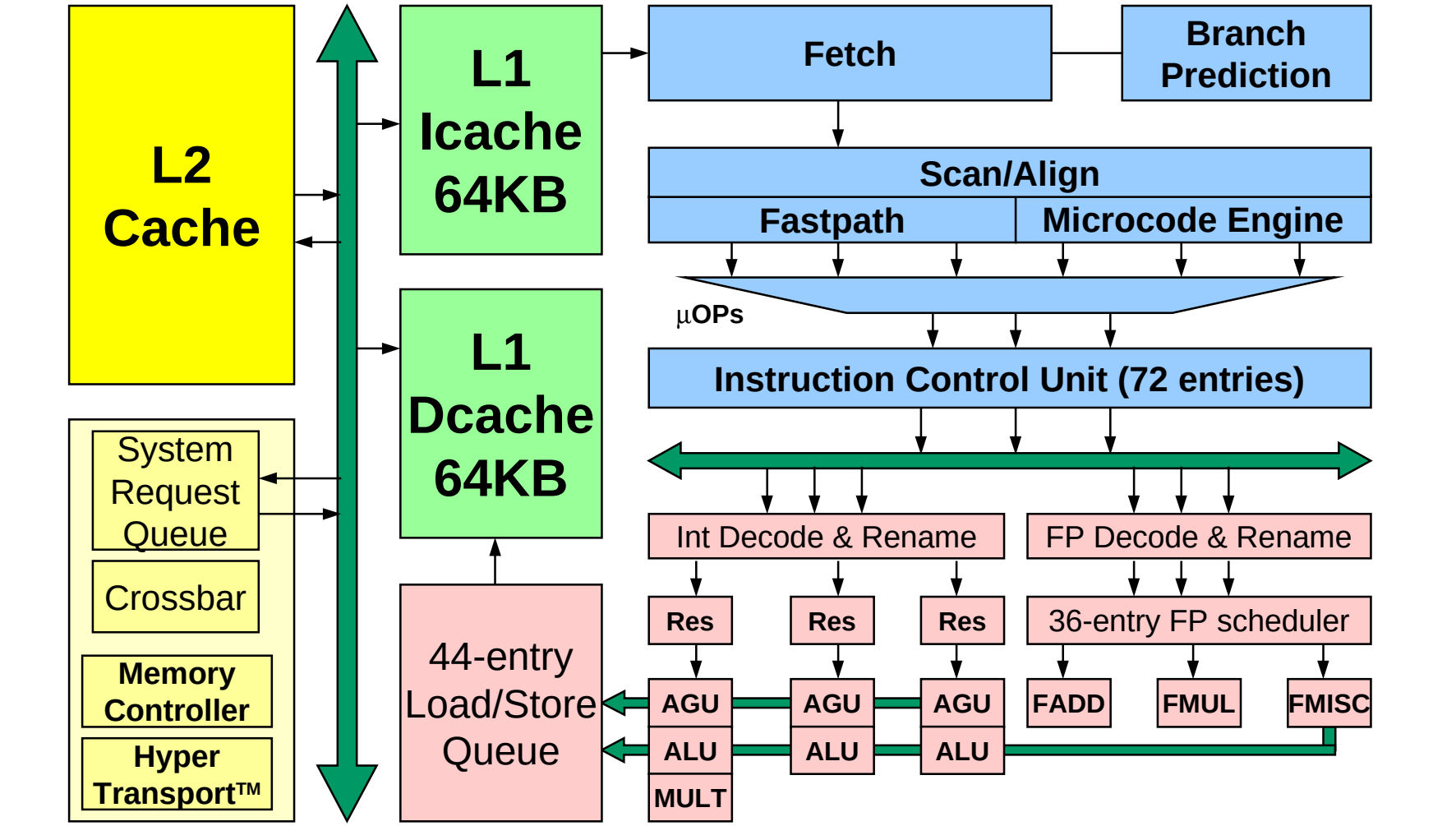
Chetana N. Keltcher
Member of Technical Staff
Advanced Micro Devices

Hammer Architecture Overview

- First x86-64 based processor
- Aggressive out-of-order, 9-issue superscalar processor
- Integrated DDR memory controller
- Leading performance in integer, floating point and multimedia
  - x86-64, x87, MMX™, 3DNow!™, SSE, SSE2

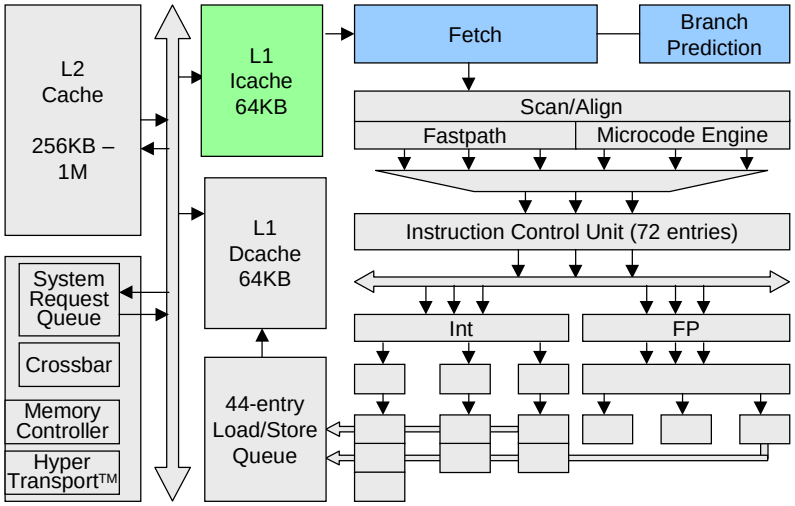


Hammer Core Overview



Instruction Fetch

- Supply 16 instruction bytes to the decoder per cycle
- 64KB instruction cache, 2-way set associative
  - Linearly-indexed, physically-tagged, 64-byte block size
  - Prefetch next sequential block on a miss



- 2 sets of instruction cache tags (fetch port, snoop)
- Predecode instruction
  - 1 end bit per-byte
  - Decode some branch types
- Branch prediction

Branch Prediction

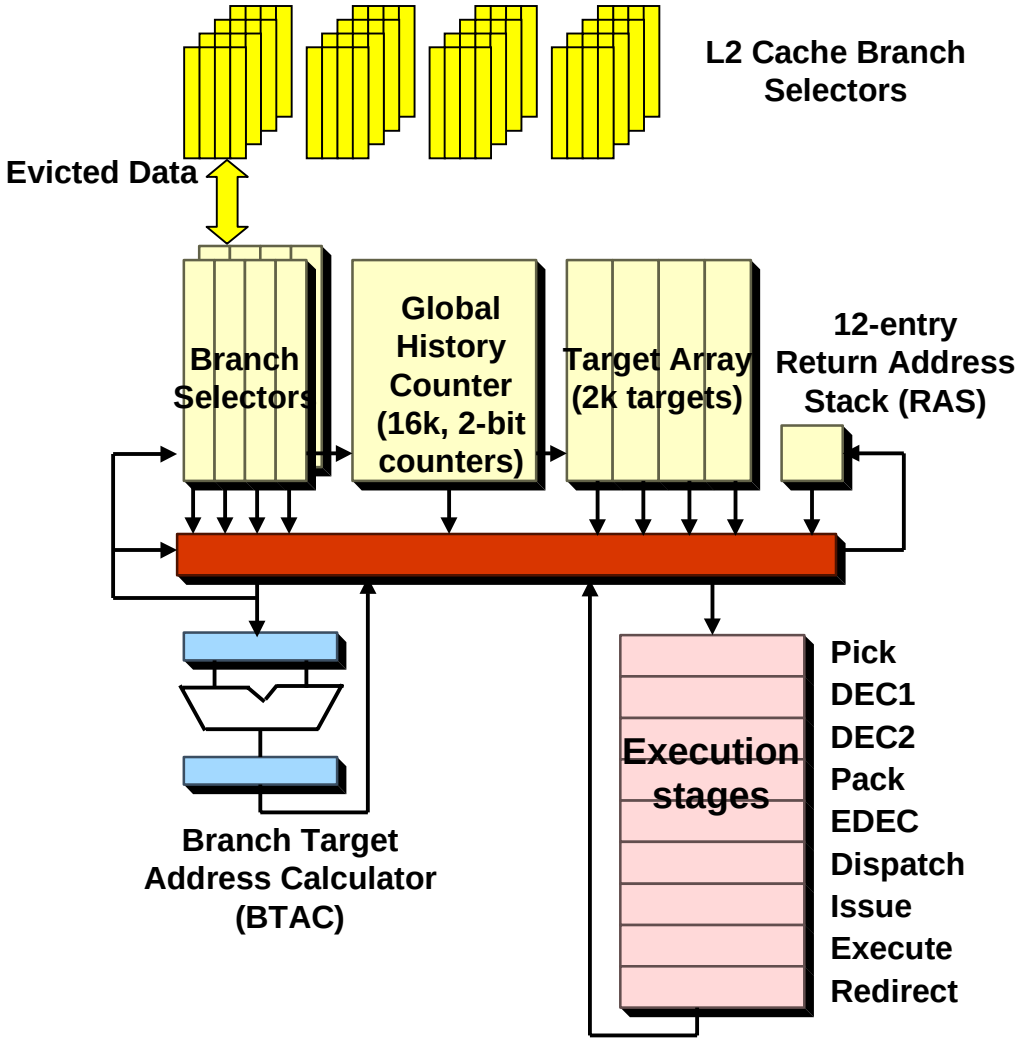
- Sequential Fetch

- Predicted Fetch

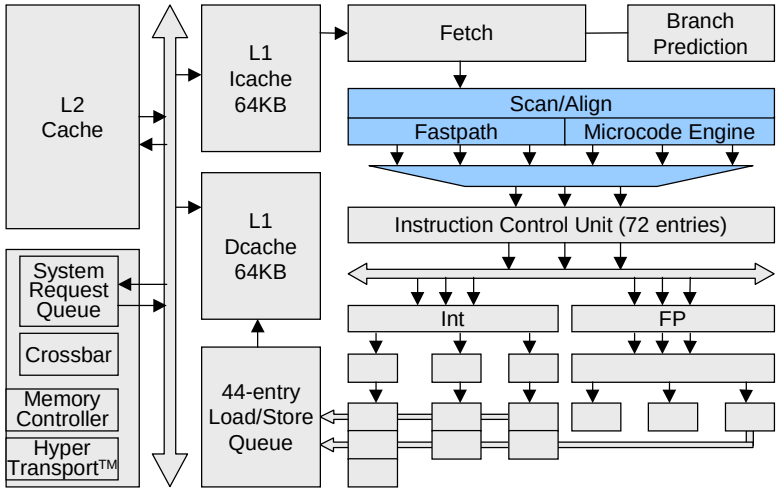
- Branch Target Address Calculator Fetch

- Mispredicted Fetch

- 5-10% improvement in prediction accuracy vs. AMD Athlon™



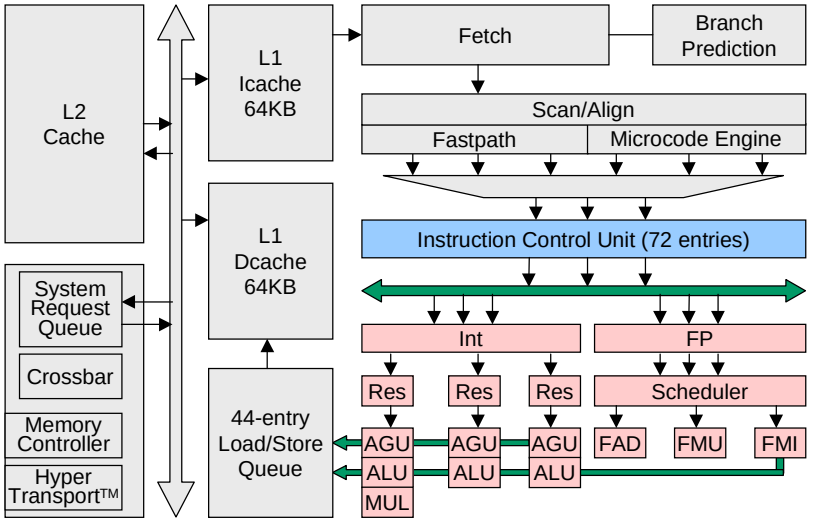
Scan / Align



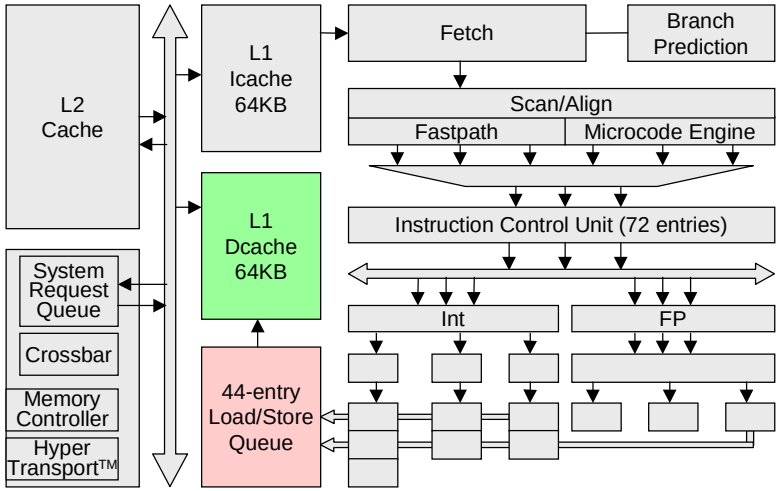
- Convert x86 instructions to fixed length µOPs
- Dispatch 3 µOPs per cycle to integer/FP schedulers
- Instructions use one of two decoding pipelines
  - Fastpath: instructions decoding to two or fewer µOPs are decoded by hardware, packed into 3 dispatch positions
  - Microcode: x86 instructions decoding to more than two µOPs, calculate ROM entry point, fetch sequence from ROM
- Compared to AMD Athlon™, more instructions use the fastpath
  - Eg: Packed SSE is microcoded in AMD Athlon and fastpath in Hammer
  - Hammer has 8% fewer microcoded instructions for Specint2000
  - Hammer has 28% fewer microcoded instructions for Specfp2000

Execution Units

- 3 integer units
- 3 address generation units
- 3 superscalar floating point units
- Integer
  - Full 64-bit data path
  - 3 x 8-entry reservation stations
  - Single cycle 32 and 64-bit add, sub, rotate, shift, logical, etc.
  - 32-bit multiply: 3 cycle latency
  - 64-bit multiply: 5 cycle latency
- Floating point
  - Handles x87, MMX™, 3DNow!™, SSE and SSE2
  - 36-entry scheduler
  - Out-of-order, fully pipelined design



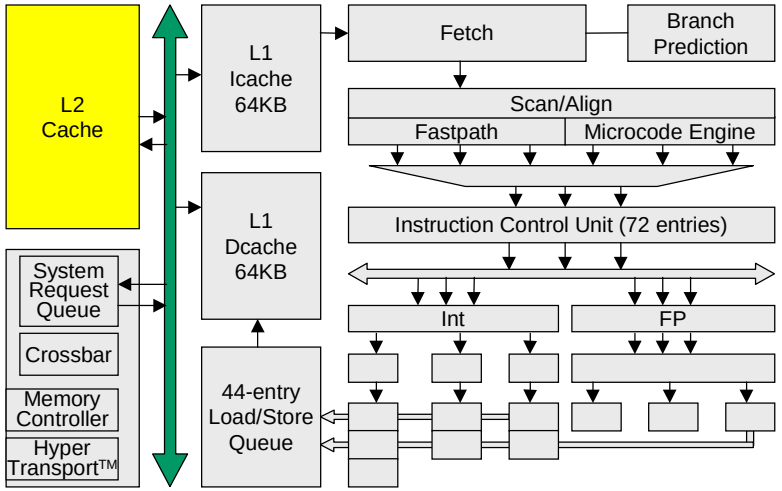
Load/Store and Data Cache



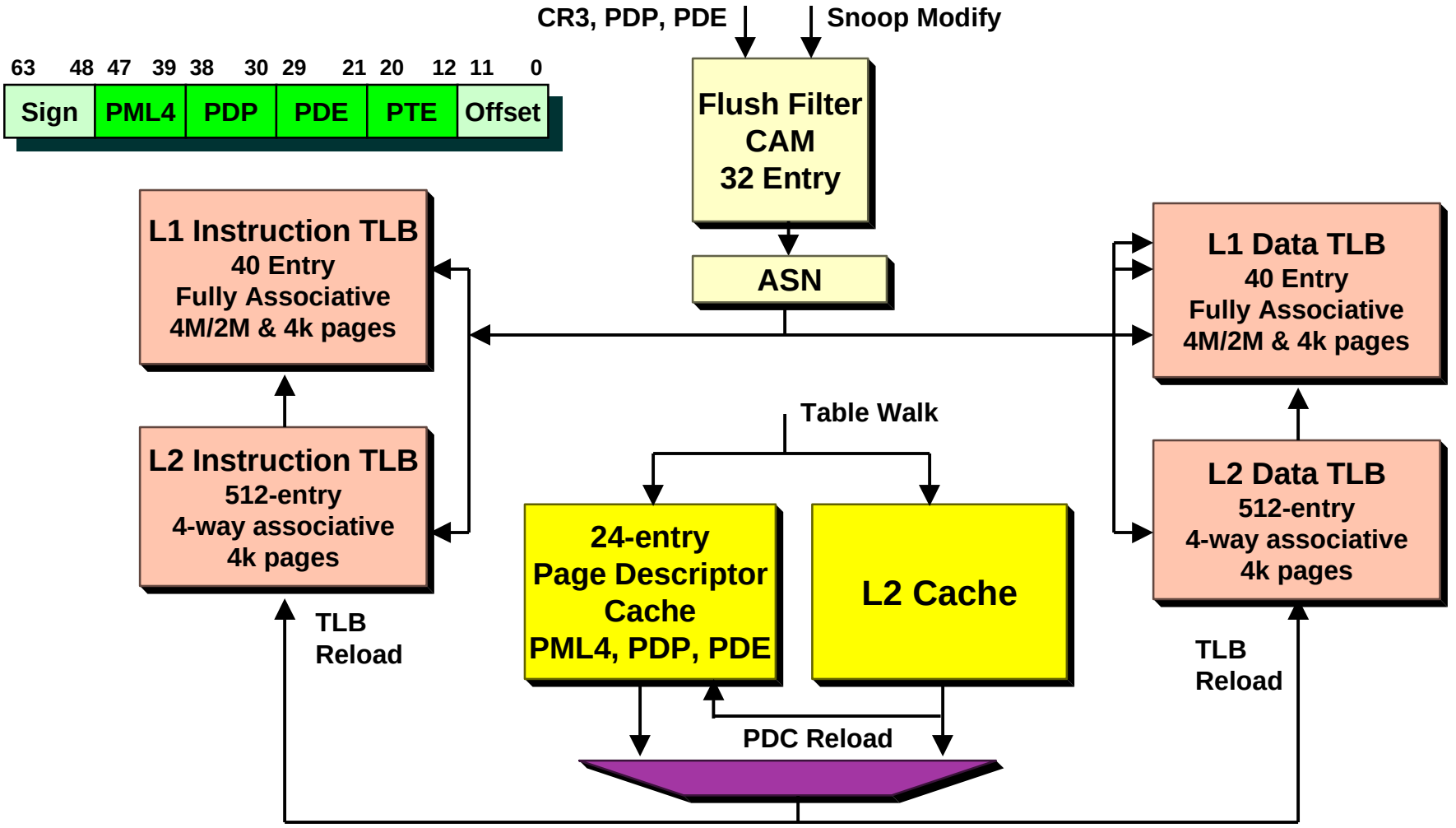
- 64KB data cache
  - 2-way set associative
  - Linearly-indexed, physically-tagged
  - 40-bit physical address
  - 48-bit linear address
  - MOESI coherency
  - 64-byte block size
- Banked and dual ported
  - 2 64-bit reads/writes each cycle to different banks
- 3 sets of data cache tags (port A, port B, snoop)
- Load->use latency is 3 cycles (zero segment base)
  - 1 extra cycle to handle misaligned (quadword boundary) loads
- Data forwarding from stores to dependent loads
- Hardware prefetch

L2 Cache

- Configurable sizes up to 1MB
- 16-way set associative
- L1 and L2 storage is mutually exclusive
- Pseudo-LRU scheme to reduce the number of LRU bits by half
- Stores IC predecode and branch prediction bits
- 10 outstanding miss requests
  - 8 DC
  - 2 IC
- System interface
  - Victim Buffer (8-entry)
  - Snoop Buffer (8-entry)
  - Write Buffer (4-entry)

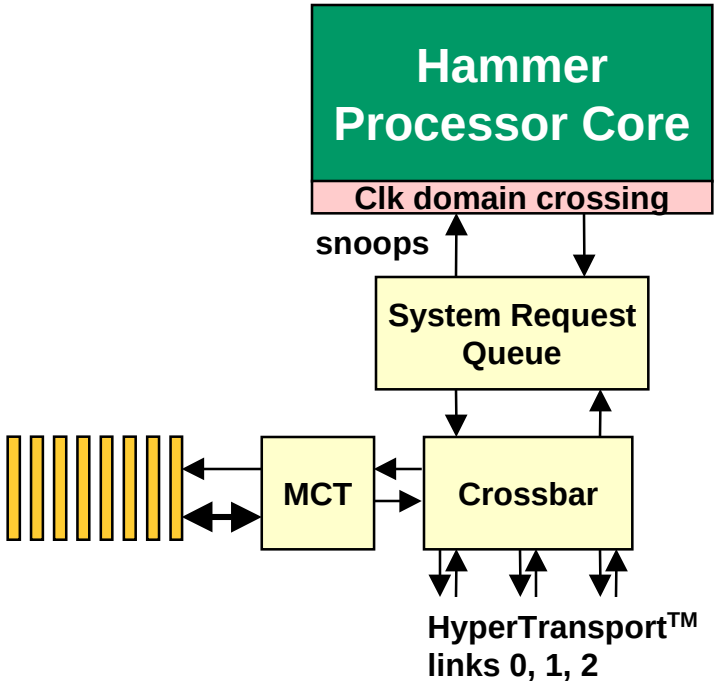


TLB for Large Workloads

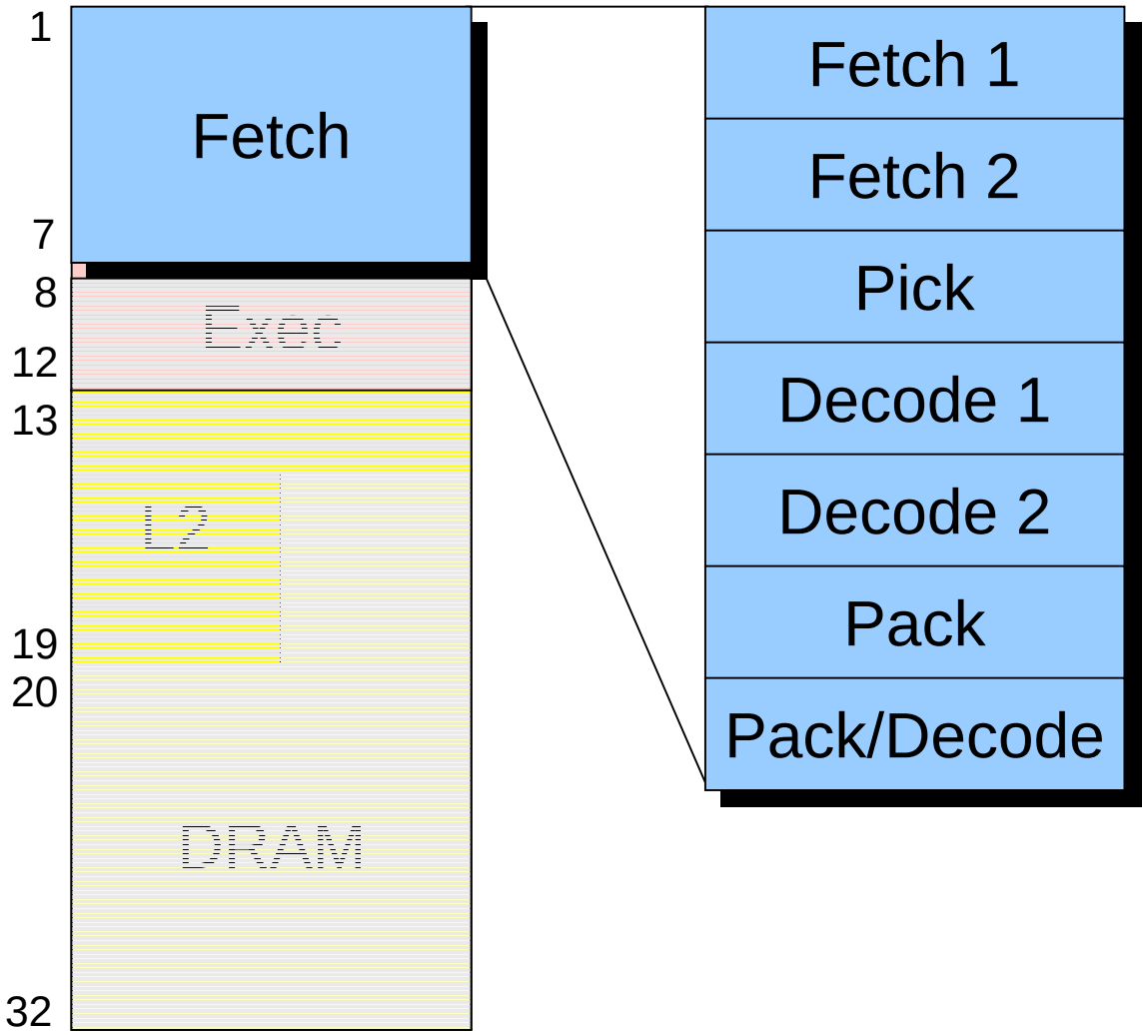


Integrated Memory Controller

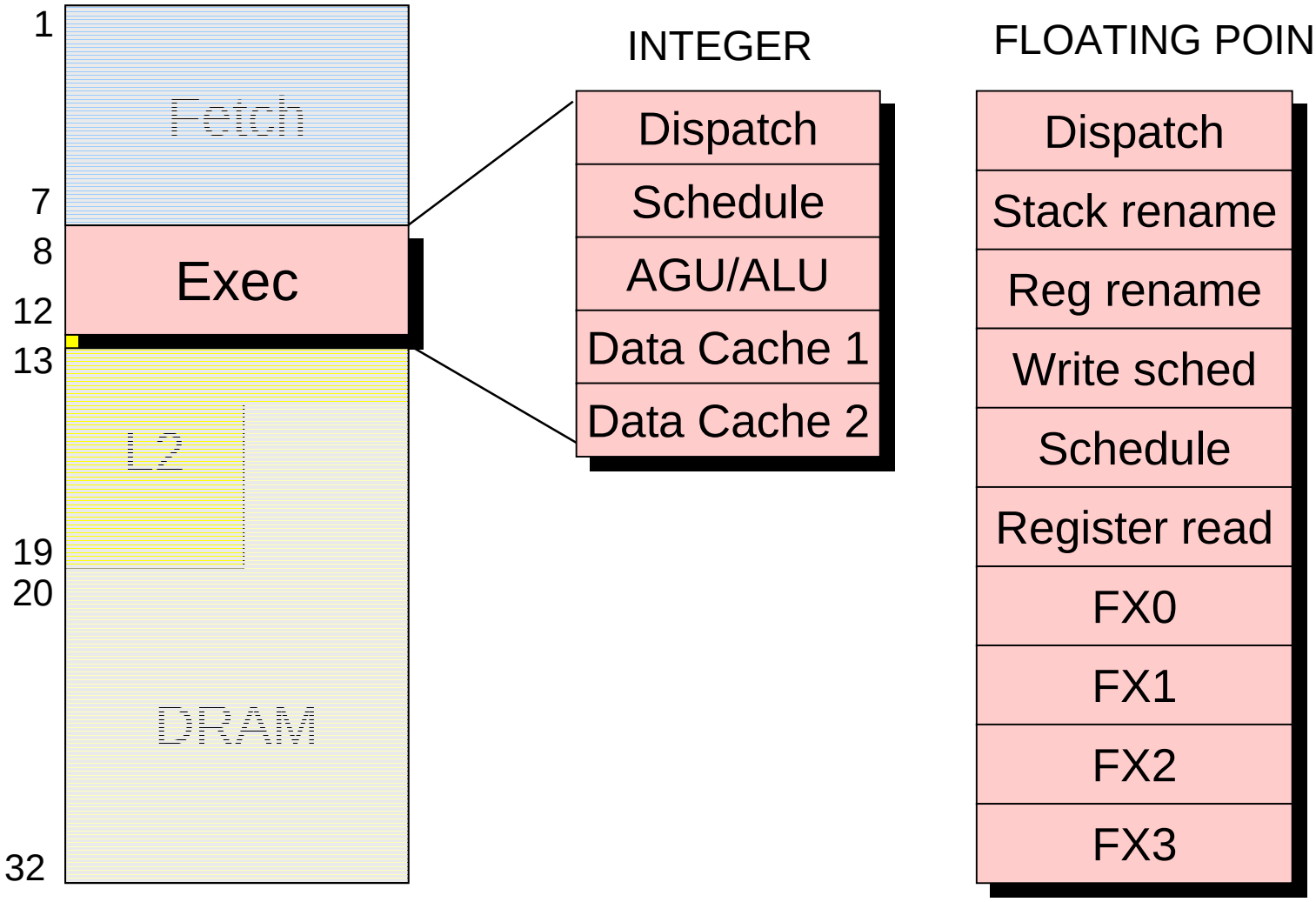
- Integrated DDR memory controller
  - 8-byte or 16-byte interface
  - Unbuffered or Registered DIMMs
  - 16-byte interface supports direct connection to 8 registered DIMMs and chipkill ECC
  - Significantly reduces memory latency
  - Memory latency improves as CPU and HyperTransport™ link speed improves
  - Performance improves by approximately 20% compared to AMD Athlon™ topology
  - Snoop throughput scales with CPU frequency
- Integrated Northbridge Functionality
  - Processes requests from CPU/IO to DRAM/IO
  - HyperTransport™ routing
    - peak bandwidth = 6.4GB/s
  - Handles transaction ordering and cache coherence
  - Runs at the same frequency as CPU core



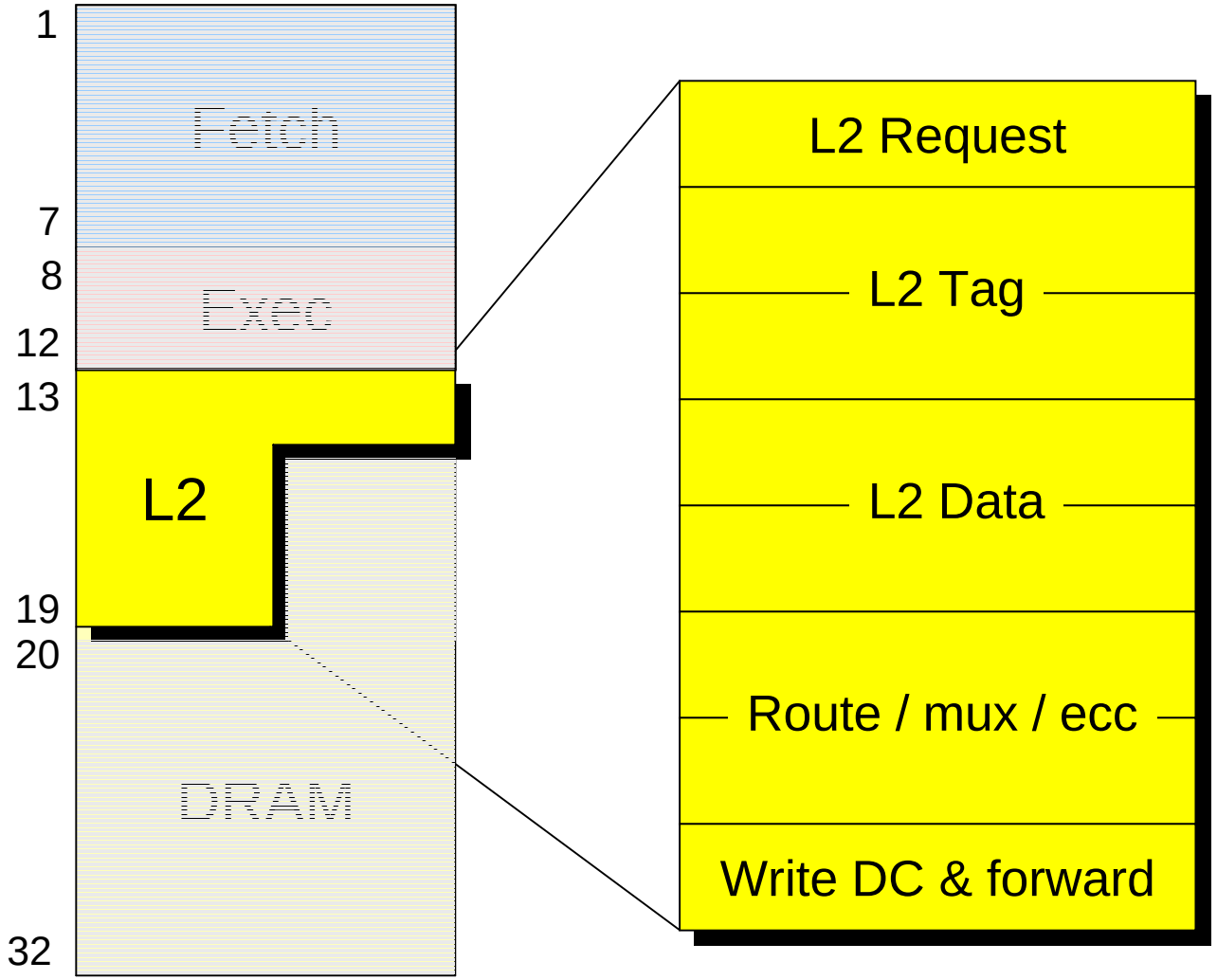
Fetch/Decode Pipeline



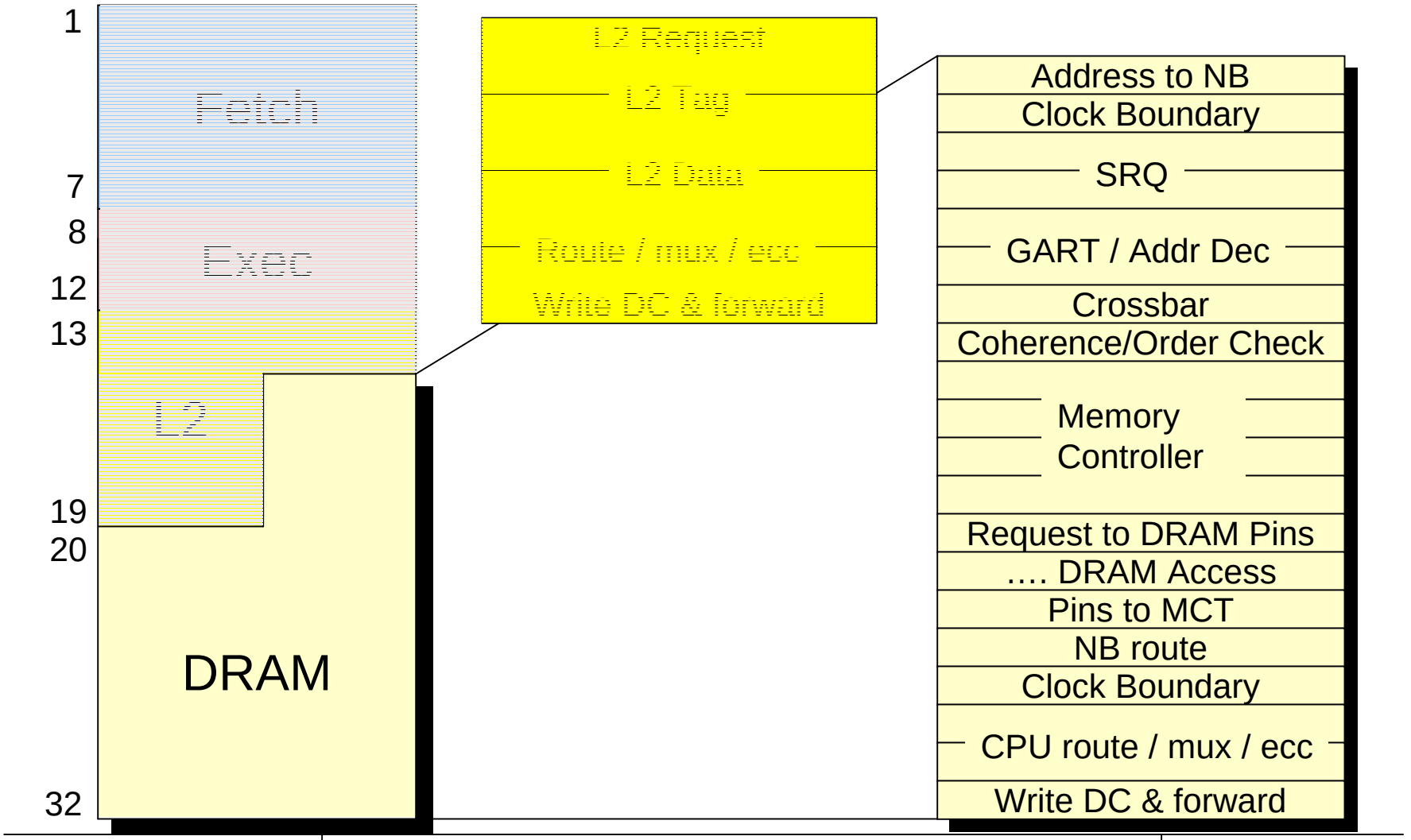
Execute Pipeline



L2 Pipeline


1
Fetch
7
8
Exec
12
13
L2
19
20
DRAM
32
L2 Request
L2 Tag
L2 Data
Route / mux / ecc
Write DC & forward

DRAM Pipeline



Reliability features

- L1 cache data is ECC protected
- L2 cache and tags are ECC protected
- DRAM is ECC protected with chipkill ECC support
- On-chip and off-chip ECC protected arrays include autonomous, background hardware scrubbers
- Remaining arrays are parity protected
  - Instruction cache, tags and TLBs
  - Data tags and TLBs
  - Generally read only data which can be recovered
- Machine Check Architecture
  - Report failures and predictive failure results

Hammer Family of Processors

- Process Technology
  - 130nm SOI
  - 9-layer
  - copper interconnect
  - Dresden, Germany
- AMD Opteron™ with Hammer technology
  - 940 pin mPGA package
  - PC1600, PC2100, or PC2700 DDR memory
  - dual-channel DDR memory interface
  - Up to 3 HyperTransport™ links
    - 6.4GB/s each
    - 19.2GB/s aggregate
- AMD Athlon™ with Hammer technology
  - 754 pin mPGA package
  - PC1600, PC2100, or PC2700 DDR memory
  - Single-channel DDR memory interface
  - 1 HyperTransport™ link
    - 6.4GB/s

Summary

- AMD's Hammer architecture provides a foundation for market-specific solutions:
  - Desktop, mobile, workstation (1-2 way), server (1-8 way)
- AMD's next-generation microprocessor core
  - Leading edge performance for 16-, 32- and 64-bit applications
- Cache subsystem
  - Enhanced TLB structures
  - Improved branch prediction
  - Reliability features
- Integrated DDR memory controller
  - Reduced memory latency
  - 1:1 scaling
- HyperTransport™ technology
  - Fast I/O for chip-to-chip communication
  - Enables glueless MP
- Innovation with compatibility

Authors

Chetana Keltcher

Ashraf Ahmed

Michael Clark

Hongwen Gao

Kelvin Goveas

Ramsey Haddad

Bruce Holloway

Bill Hughes

Richard Klass

Dave Kroesche

Phil Madrid

Kevin McGrath

Teik Tan

Scott White

Tim Wood

Gerald Zuraski, Jr.

Trademark Attribution