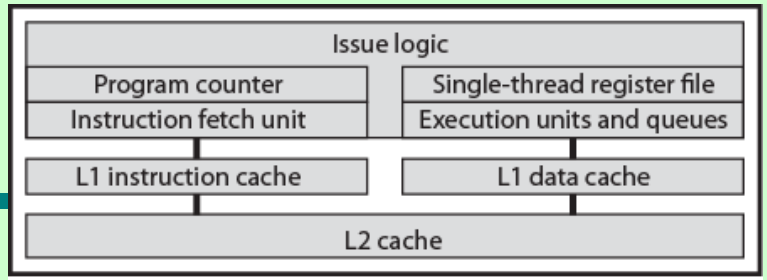
William Stallings
Computer Organization and Architecture
8th Edition

Chapter 18
Multicore Computers

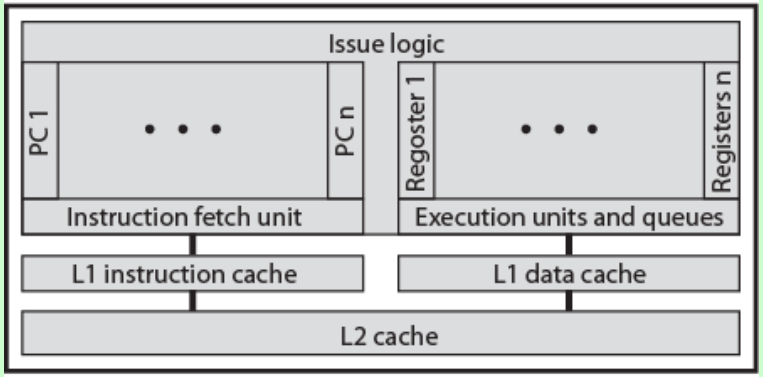
Hardware Performance Issues

- Microprocessors have seen an exponential increase in performance
  - Improved organization
  - Increased clock frequency
- Increase in Parallelism
  - Pipelining
  - Superscalar
  - Simultaneous multithreading (SMT)
- Diminishing returns
  - More complexity requires more logic
  - Increasing chip area for coordinating and signal transfer logic
    - Harder to design, make and debug

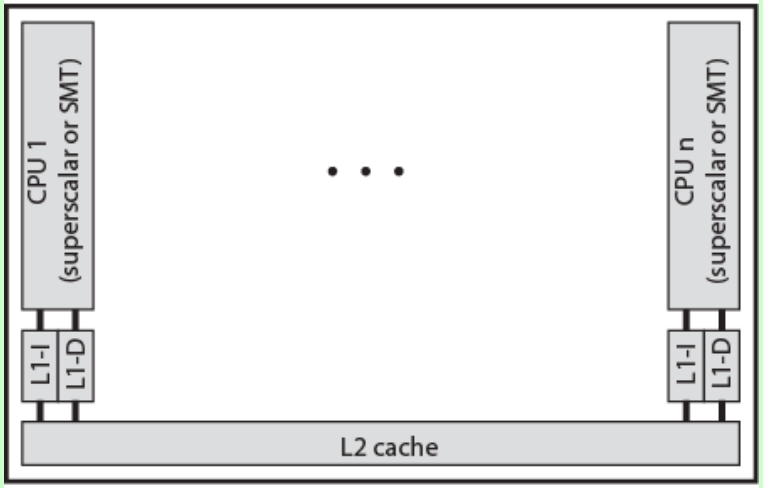
Alternative Chip Organizations



(a) Superscalar

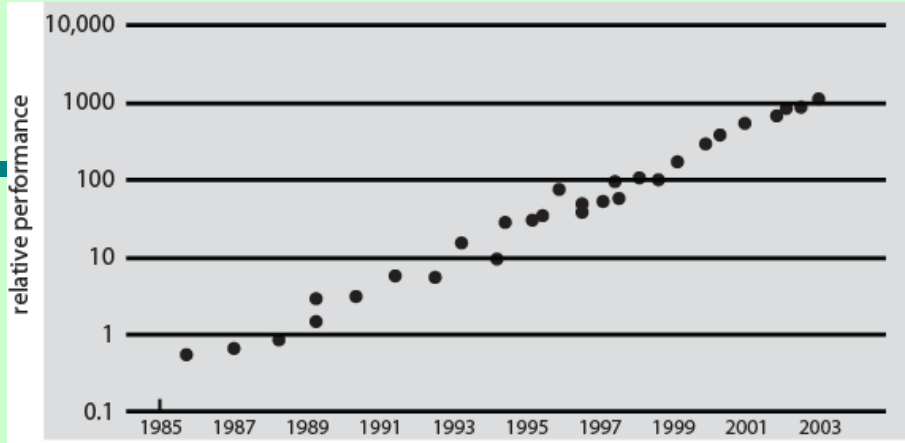


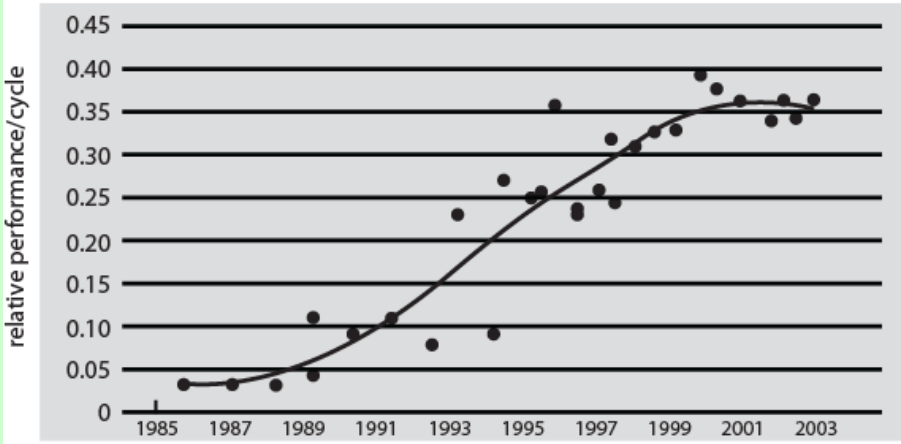
(b) Simultaneous multithreading

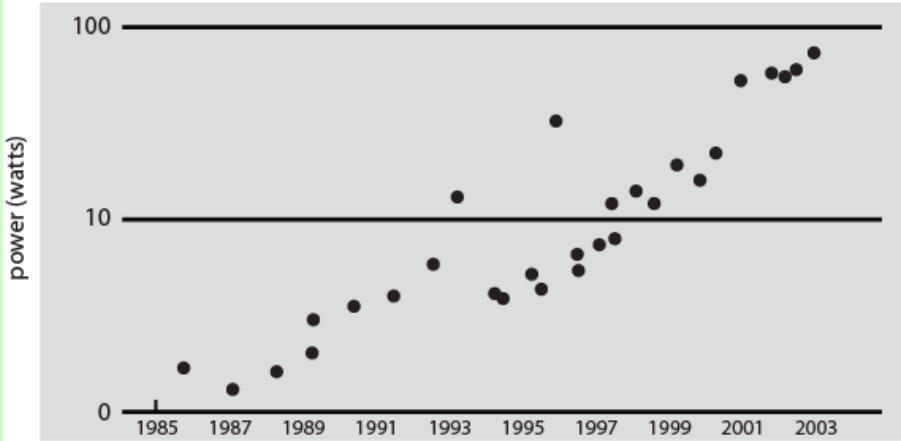


(c) Multicore

Intel Hardware Trends



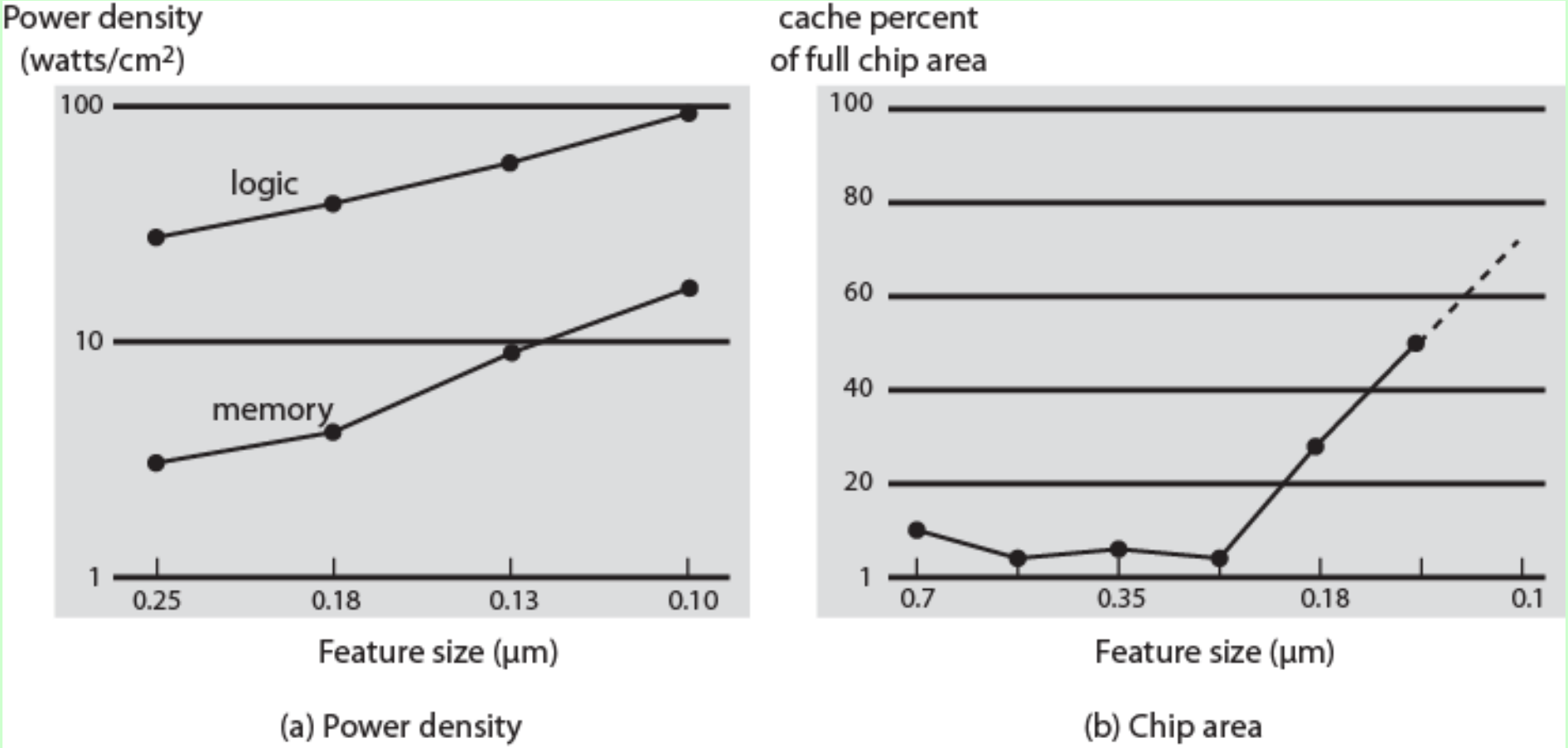




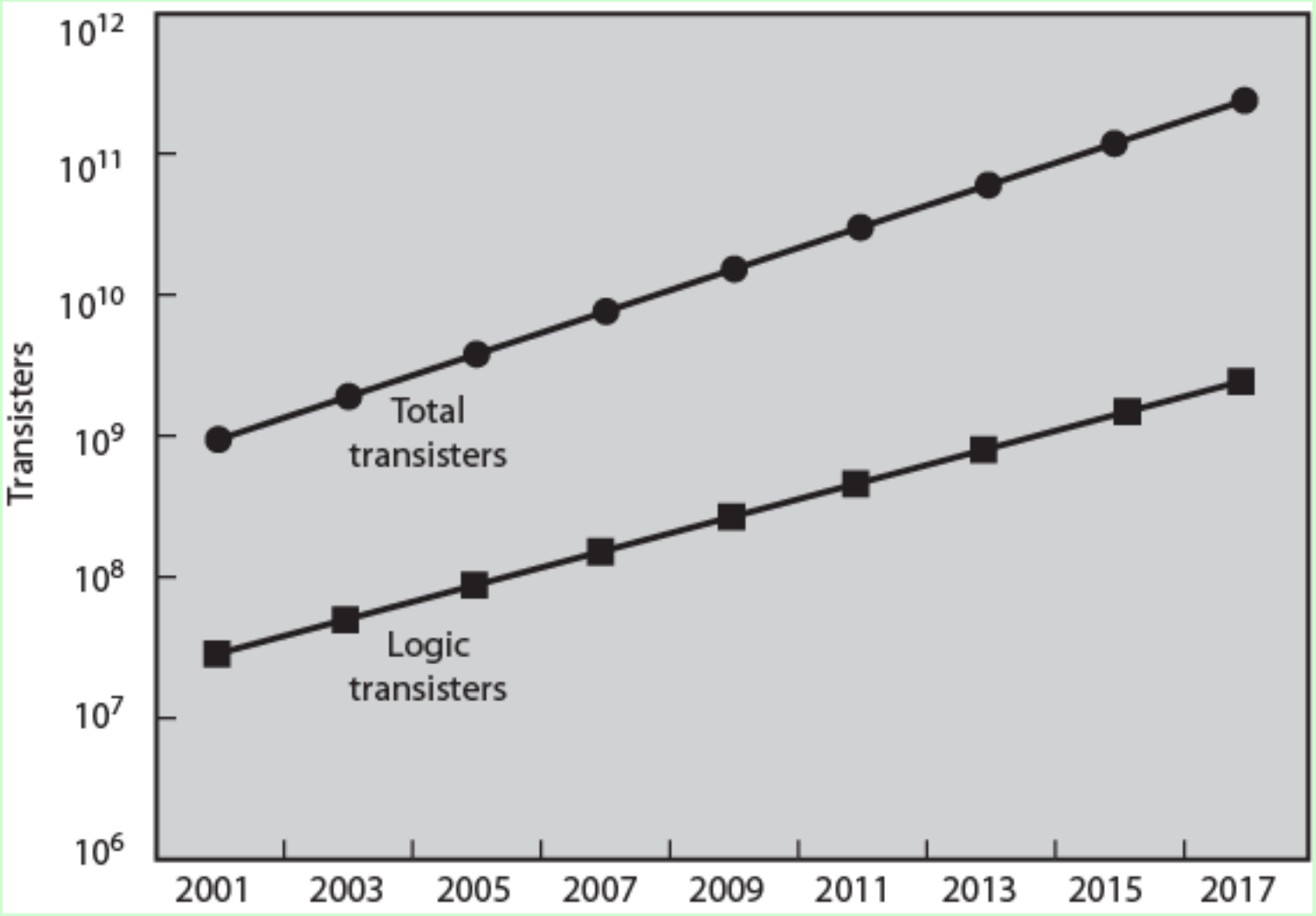
Increased Complexity

- Power requirements grow exponentially with chip density and clock frequency
  - Can use more chip area for cache
    - Smaller
    - Order of magnitude lower power requirements
- By 2015
  - 100 billion transistors on 300mm^2 die
    - Cache of 100MB
    - 1 billion transistors for logic
- Pollack's rule:
  - Performance is roughly proportional to square root of increase in complexity
    - Double complexity gives 40% more performance
- Multicore has potential for near-linear improvement
- Unlikely that one core can use all cache effectively

Power and Memory Considerations



Chip Utilization of Transistors



Software Performance Issues

- Performance benefits dependent on effective exploitation of parallel resources
- Even small amounts of serial code impact performance
  - 10% inherently serial on 8 processor system gives only 4.7 times performance
- Communication, distribution of work and cache coherence overheads
- Some applications effectively exploit multicore processors

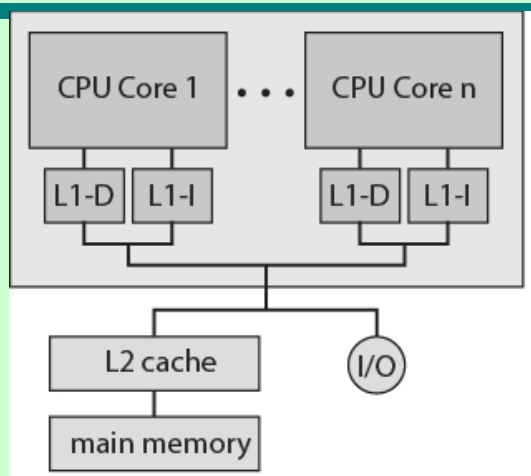
Effective Applications for Multicore Processors

- Database
- Servers handling independent transactions
- Multi-threaded native applications
  - Lotus Domino, Siebel CRM
- Multi-process applications
  - Oracle, SAP, PeopleSoft
- Java applications
  - Java VM is multi-thread with scheduling and memory management
  - Sun's Java Application Server, BEA's Weblogic, IBM Websphere, Tomcat
- Multi-instance applications
  - One application running multiple times
- E.g. Value Game Software

Multicore Organization

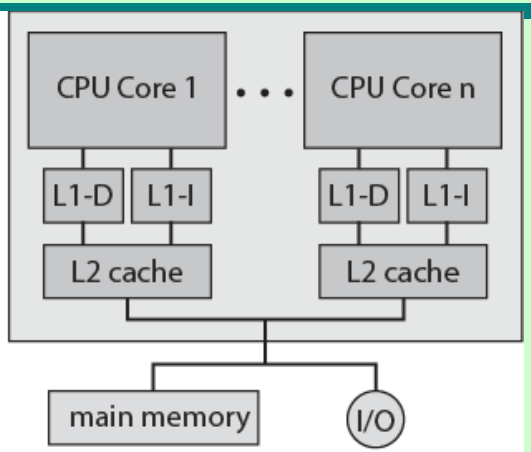
- Number of core processors on chip
- Number of levels of cache on chip
- Amount of shared cache
- Next slide examples of each organization:
- (a) ARM11 MPCore
- (b) AMD Opteron
- (c) Intel Core Duo
- (d) Intel Core i7

Multicore Organization Alternatives



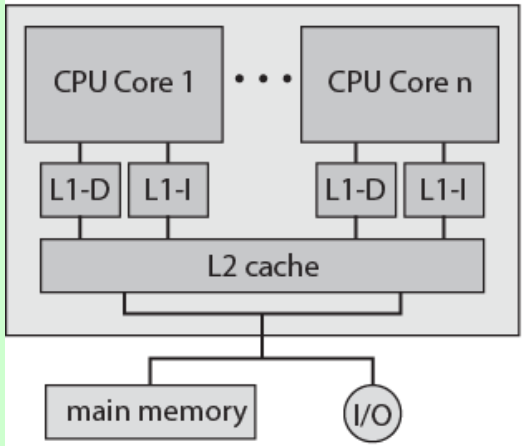
(a) Dedicated L1 cache

ARM11 MPCore



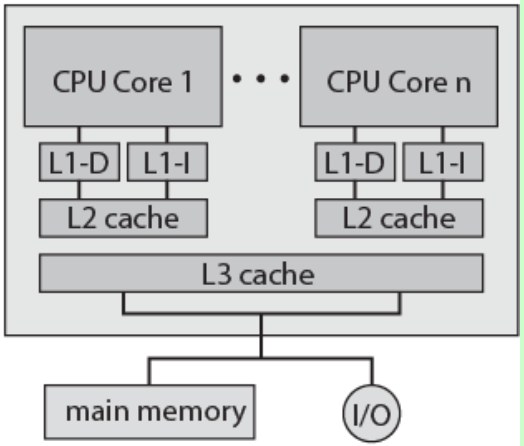
(b) Dedicated L2 cache

AMD Opteron



(c) Shared L2 cache

Core Duo



(d) Shared L3 cache

Core I7

Advantages of shared L2 Cache

- Constructive interference reduces overall miss rate
- Data shared by multiple cores not replicated at cache level
- With proper frame replacement algorithms mean amount of shared cache dedicated to each core is dynamic
  - Threads with less locality can have more cache
- Easy inter-process communication through shared memory
- Cache coherency confined to L1
- Dedicated L2 cache gives each core more rapid access
  - Good for threads with strong locality
- Shared L3 cache may also improve performance

Individual Core Architecture

- Intel Core Duo uses superscalar cores
- Intel Core i7 uses simultaneous multi-threading (SMT)
  - Scales up number of threads supported
    - 4 SMT cores, each supporting 4 threads appears as 16 core

Intel x86 Multicore Organization - Core Duo (1)

- 2006
- Two x86 superscalar, shared L2 cache
- Dedicated L1 cache per core
  - 32KB instruction and 32KB data
- Thermal control unit per core
  - Manages chip heat dissipation
  - Maximize performance within constraints
  - Improved ergonomics
- Advanced Programmable Interrupt Controlled (APIC)
  - Inter-process interrupts between cores
  - Routes interrupts to appropriate core
  - Includes timer so OS can interrupt core

Intel x86 Multicore Organization - Core Duo (2)

- Power Management Logic
  - Monitors thermal conditions and CPU activity
  - Adjusts voltage and power consumption
  - Can switch individual logic subsystems
- 2MB shared L2 cache
  - Dynamic allocation
  - MESI support for L1 caches
  - Extended to support multiple Core Duo in SMP
    - L2 data shared between local cores or external
- Bus interface

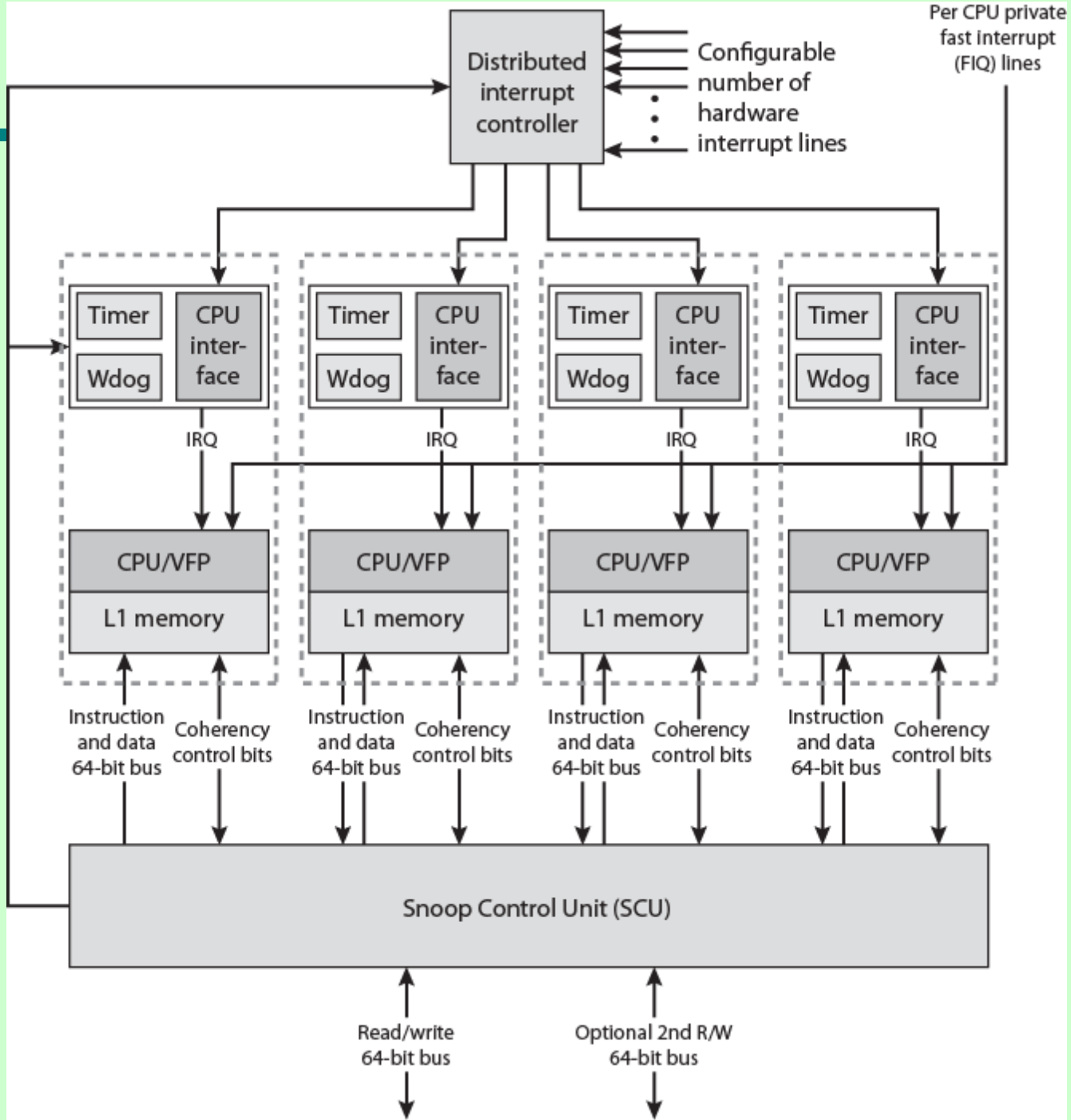
Intel x86 Multicore Organization - Core i7

- November 2008
- Four x86 SMT processors
- Dedicated L2, shared L3 cache
- Speculative pre-fetch for caches
- On chip DDR3 memory controller
  - Three 8 byte channels (192 bits) giving 32GB/s
  - No front side bus
- QuickPath Interconnection
  - Cache coherent point-to-point link
  - High speed communications between processor chips
  - 6.4G transfers per second, 16 bits per transfer
  - Dedicated bi-directional pairs
  - Total bandwidth 25.6GB/s

ARM11 MPCore

- Up to 4 processors each with own L1 instruction and data cache
- Distributed interrupt controller
- Timer per CPU
- Watchdog
  - Warning alerts for software failures
  - Counts down from predetermined values
  - Issues warning at zero
- CPU interface
  - Interrupt acknowledgement, masking and completion acknowledgement
- CPU
  - Single ARM11 called MP11
- Vector floating-point unit
  - FP co-processor
- L1 cache
- Snoop control unit
  - L1 cache coherency

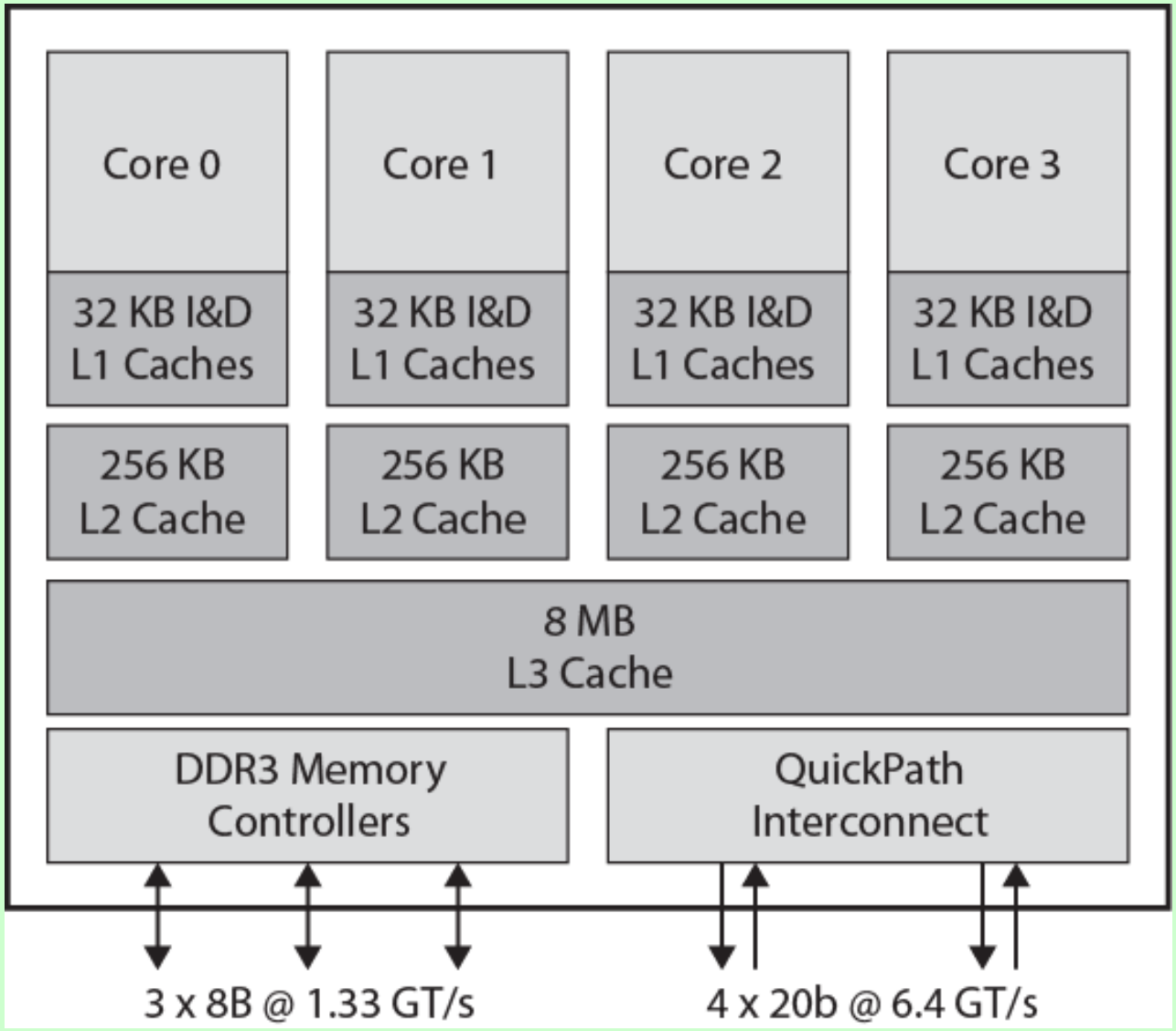
ARM11 MPCore Block Diagram



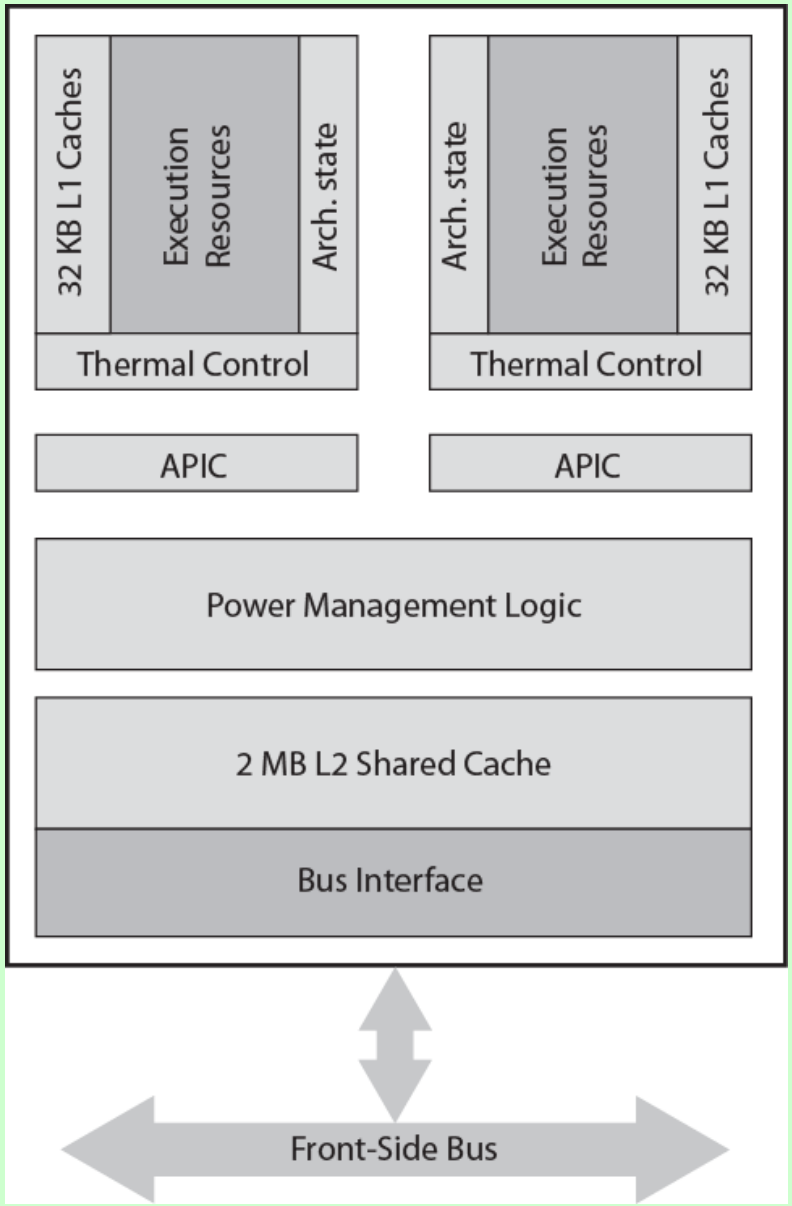
Recommended Reading

- Stallings chapter 18
- ARM web site

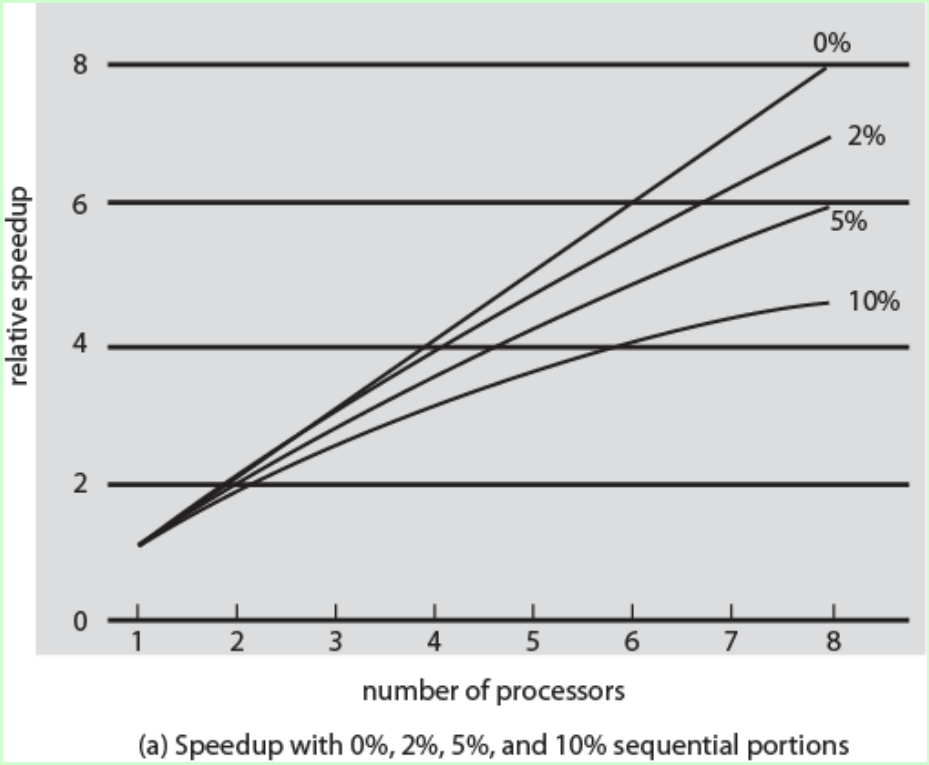
Intel Core i& Block Diagram



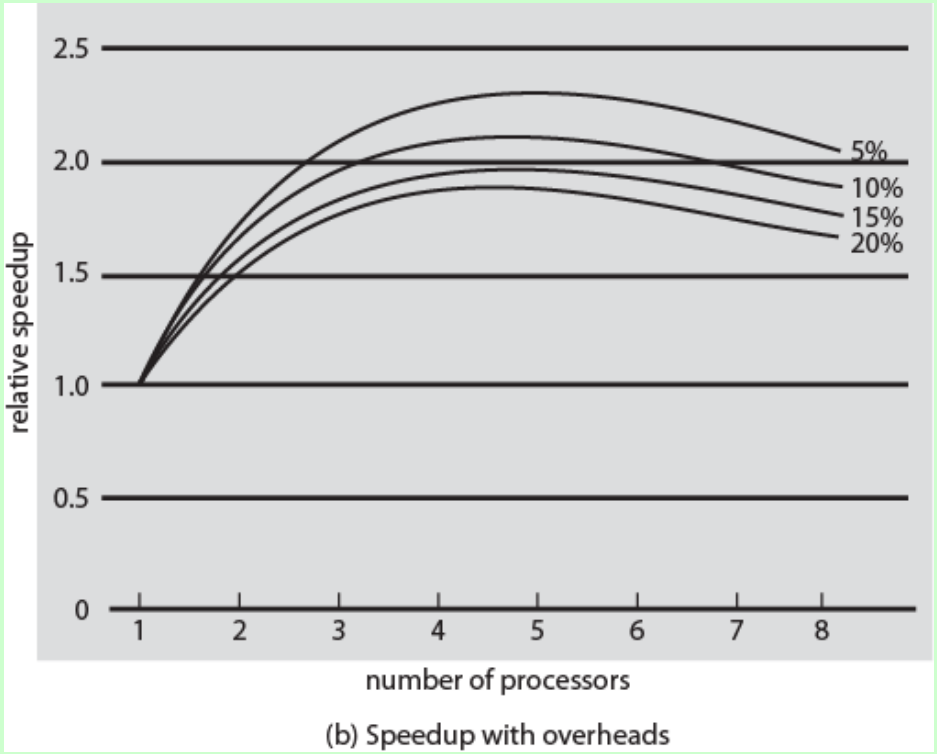
Intel Core Duo Block Diagram



Performance Effect of Multiple Cores



(a) Speedup with 0%, 2%, 5%, and 10% sequential portions



(b) Speedup with overheads

Recommended Reading

- Multicore Association web site
- ARM web site