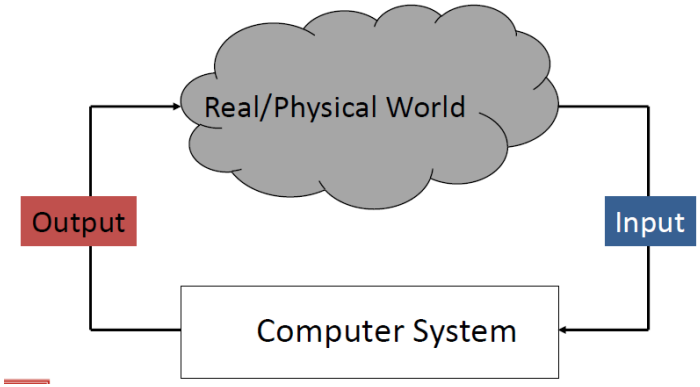
Chapter 5
Large and Fast : Exploiting the Memory Hierarchy

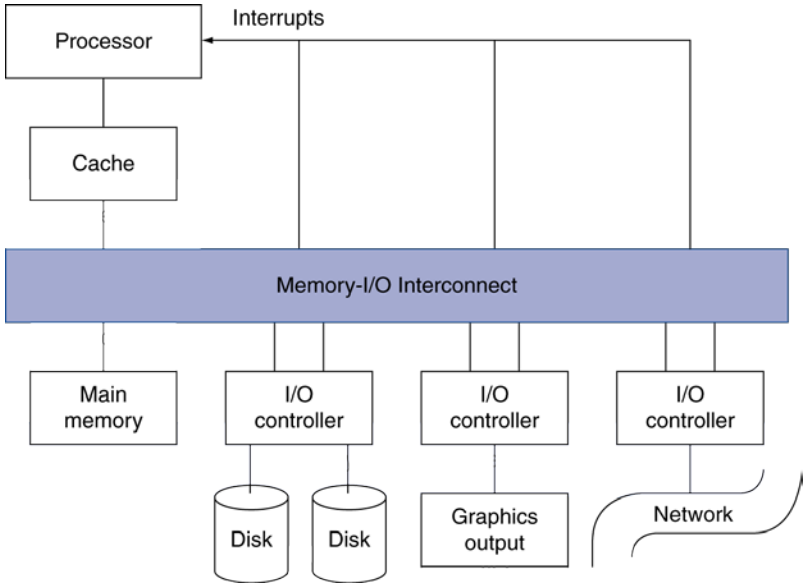
I/O Topics and Storage (Online Section 5.11)

Introduction

- Input/Output Devices



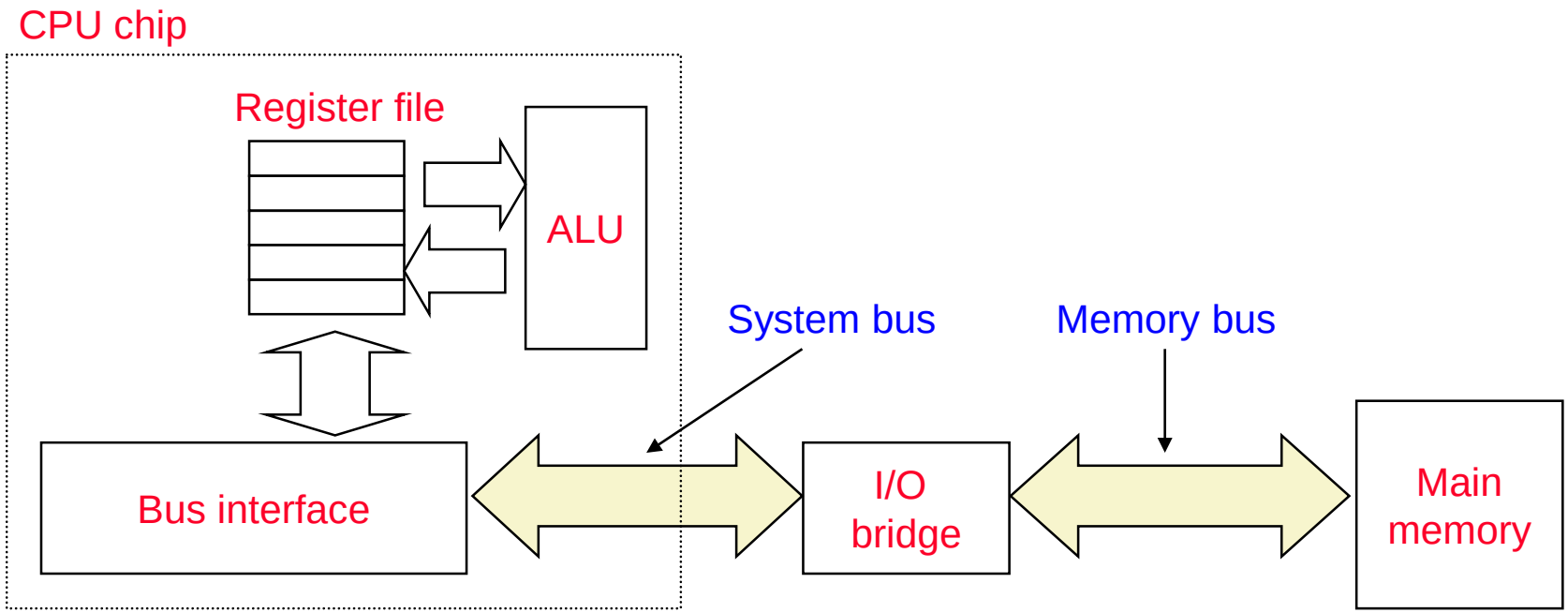
- I/O Devices Connections



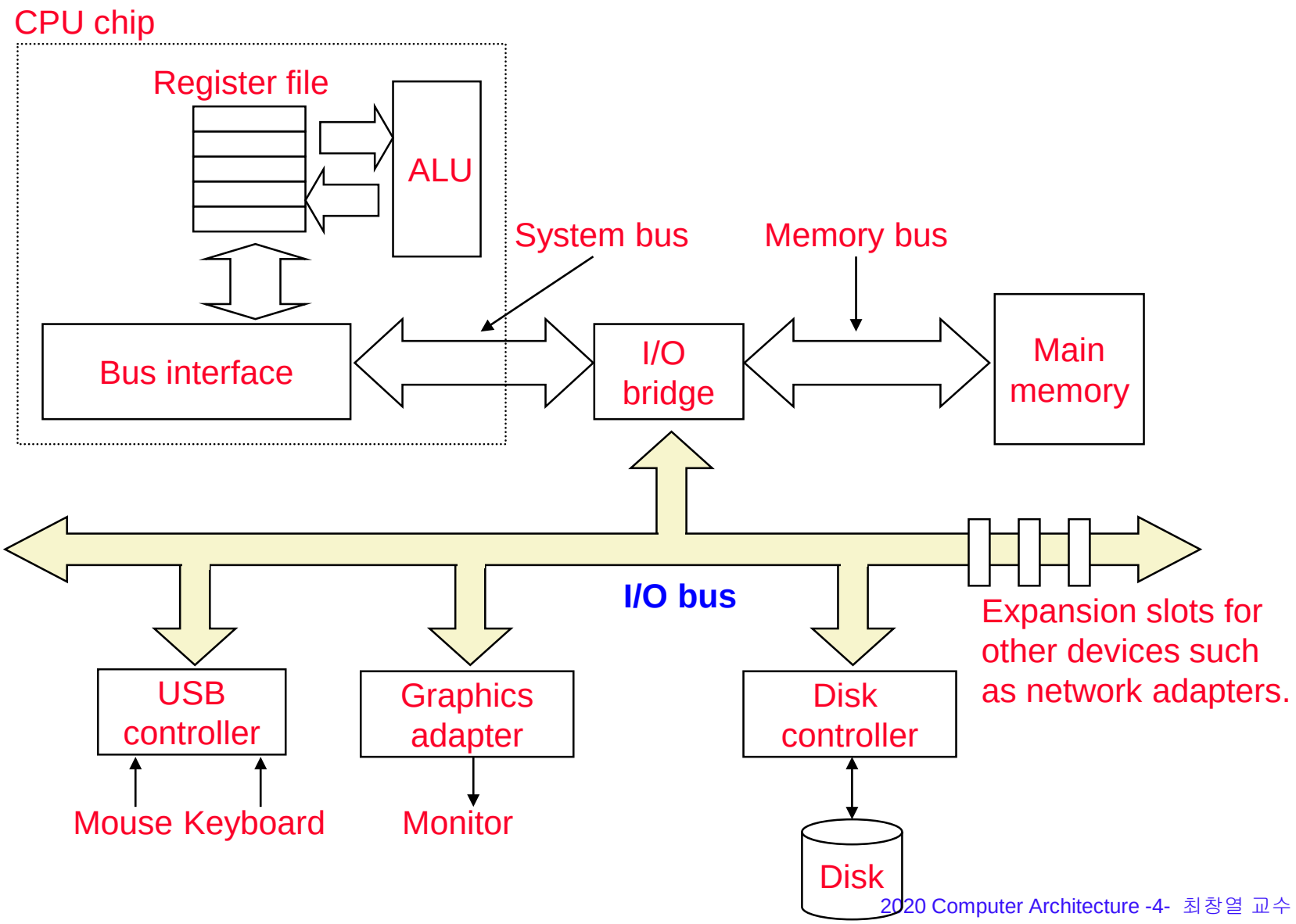
Traditional Bus Structures

- A Bus is a shared communication link.

- Memory Read / Write transaction

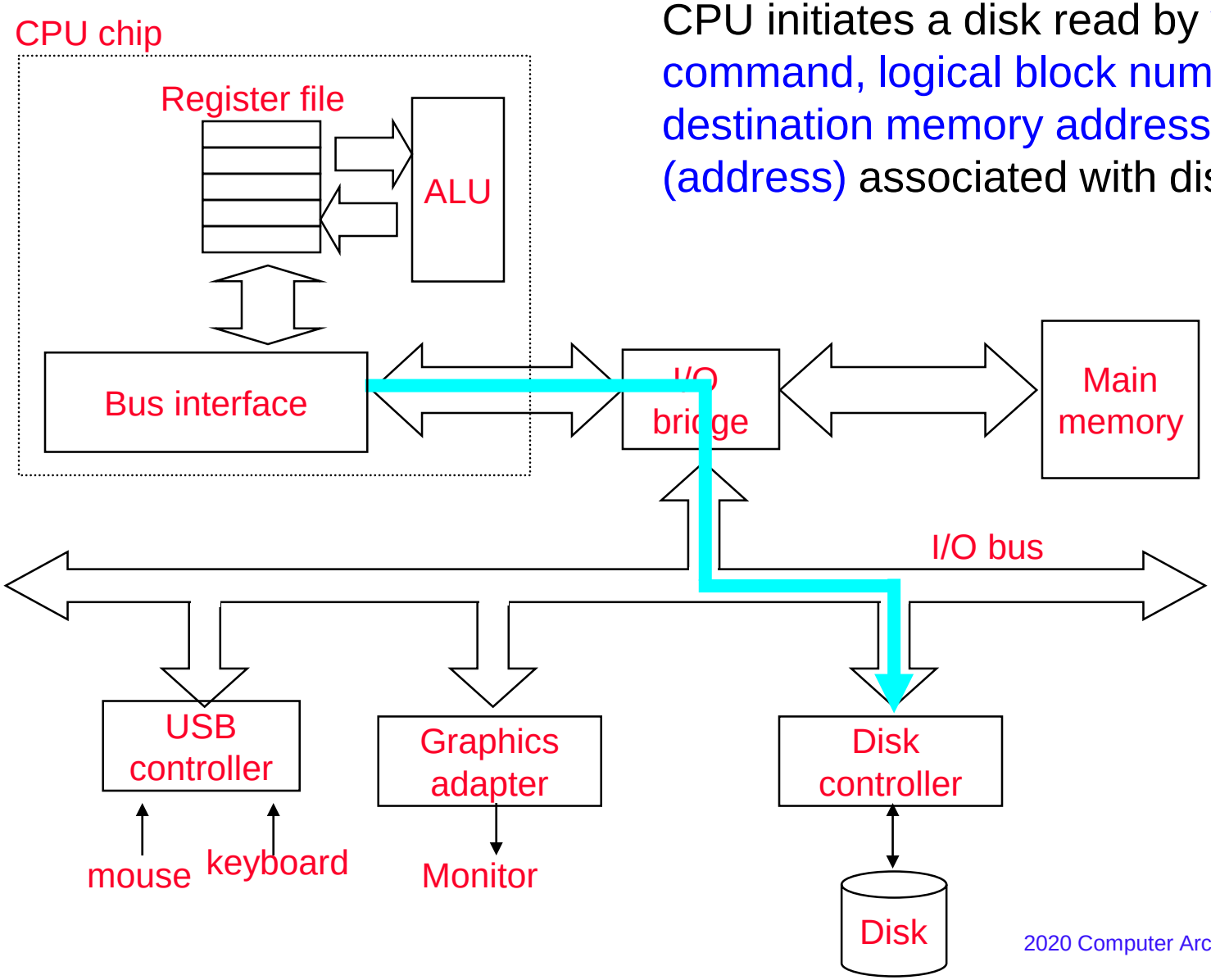


I/O Bus


CPU chip
Register file
ALU
System bus
Memory bus
Bus interface
I/O bridge
Main memory
I/O bus
Expansion slots for other devices such as network adapters.
USB controller
Graphics adapter
Disk controller
Mouse
Keyboard
Monitor
Disk

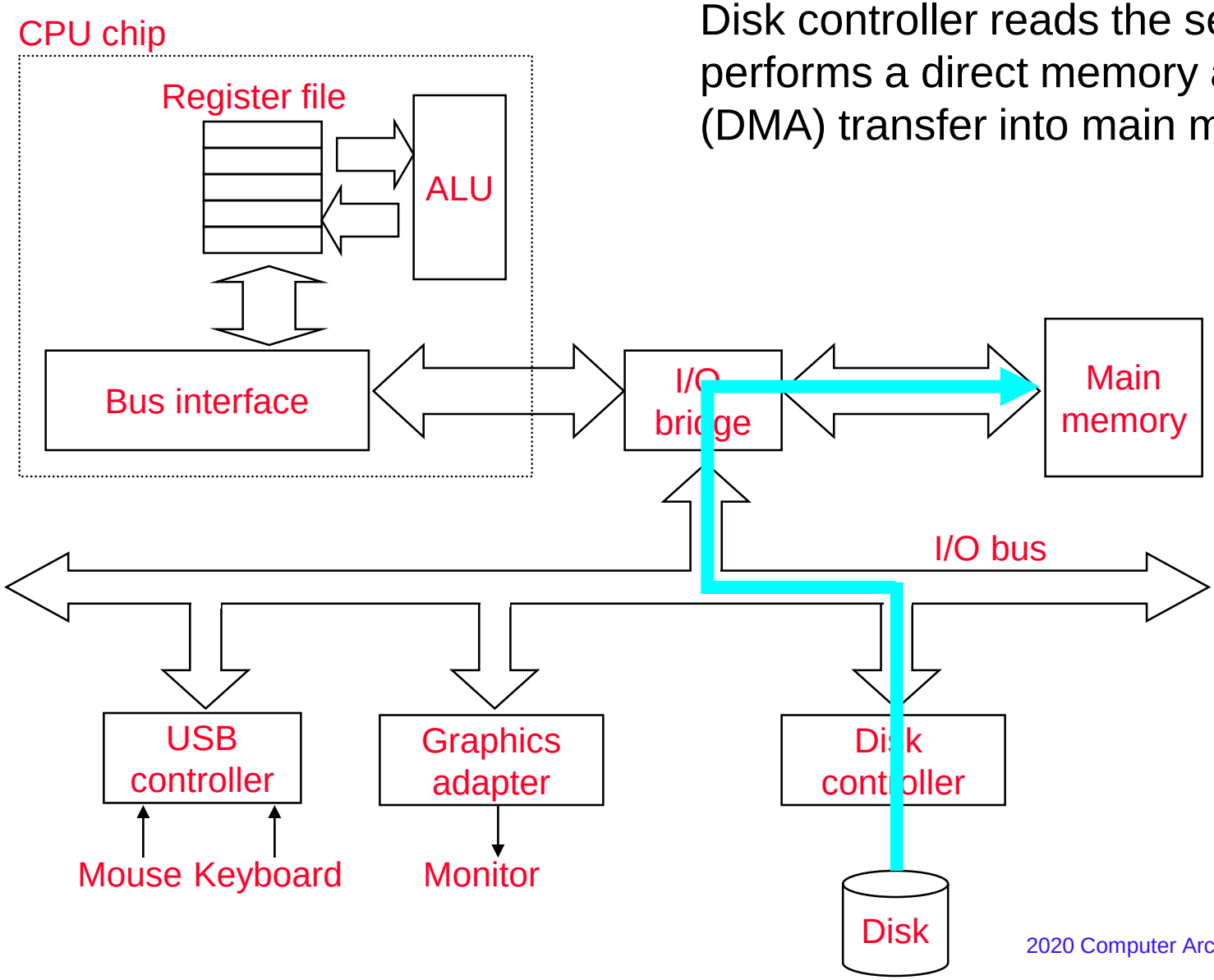
Reading a Disk Sector (1)

CPU initiates a disk read by writing a command, logical block number, and destination memory address to a port (address) associated with disk controller.



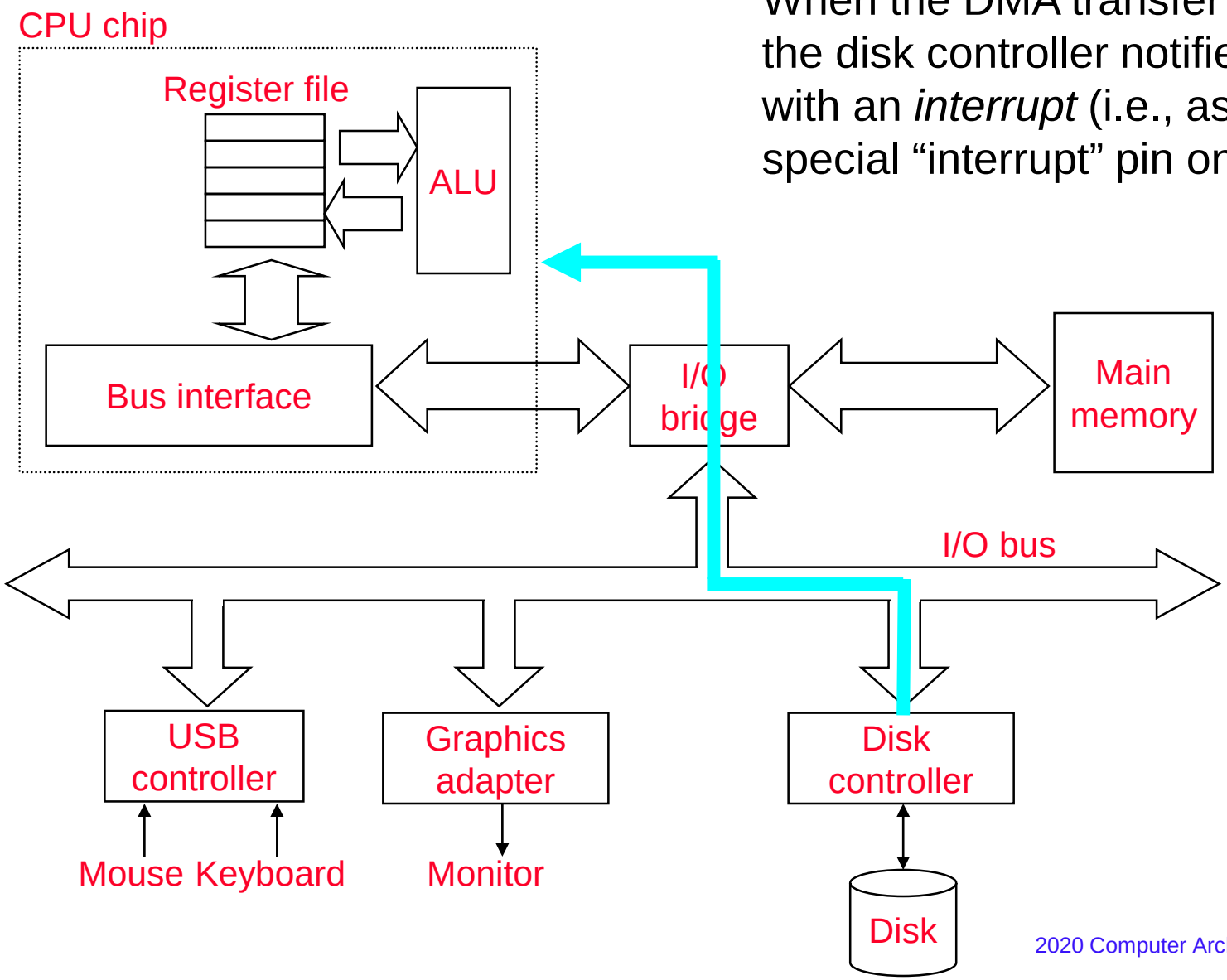
Reading a Disk Sector (2)

Disk controller reads the sector and performs a direct memory access (DMA) transfer into main memory.



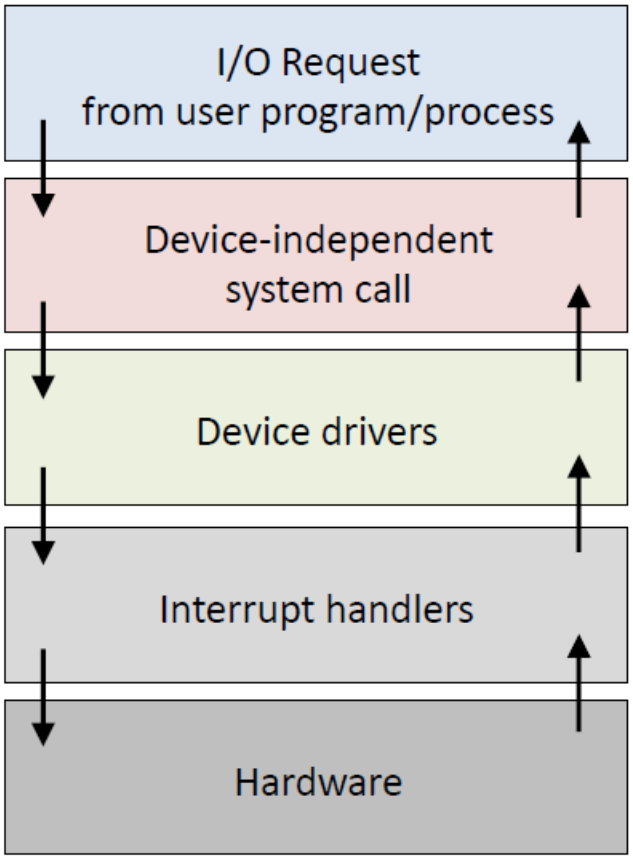
Reading a Disk Sector (3)

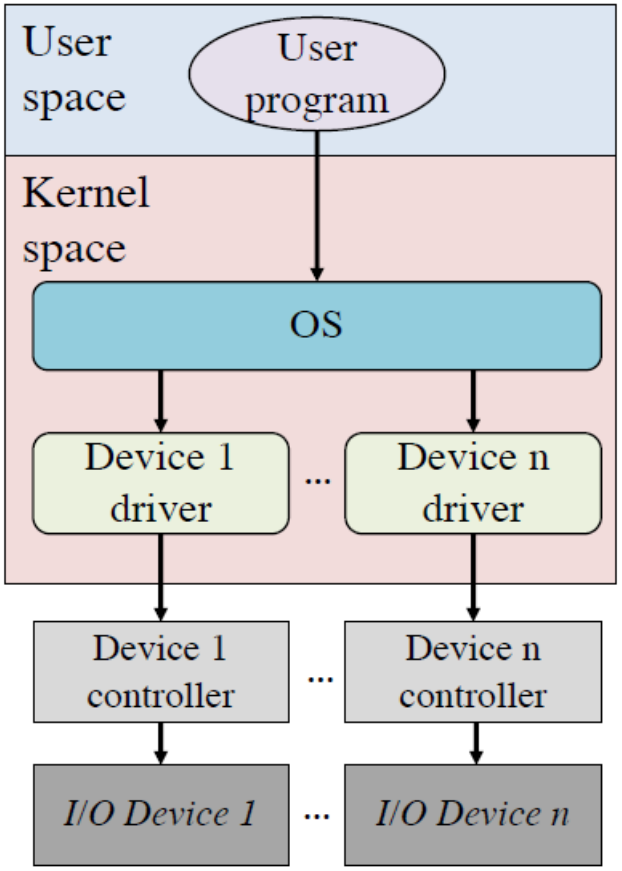
When the DMA transfer completes, the disk controller notifies the CPU with an *interrupt* (i.e., asserts a special “interrupt” pin on the CPU)



I/O Management

- I/O devices are managed by I/O controller hardware
  - Transfers data to/from device
  - Synchronizes operations with software
- I/O Execution Process





I/O Commands and I/O Register Mapping

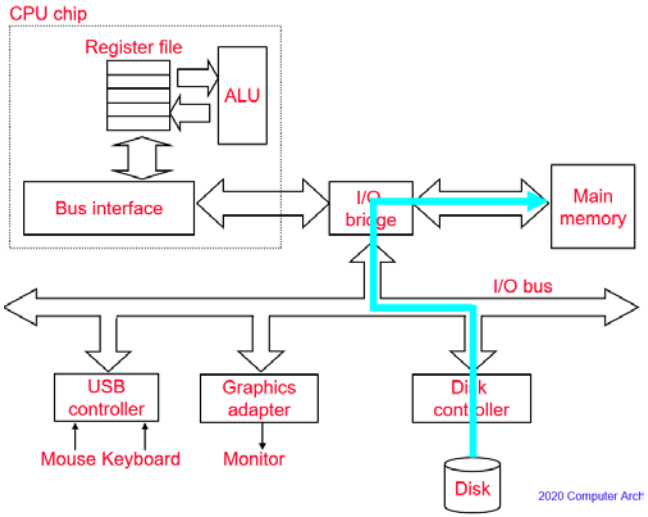
- Command registers
  - Cause device to do something
- Status registers
  - Indicate what the device is doing and occurrence of errors
- Data registers
  - Write: transfer data to a device
  - Read: transfer data from a device
- Memory mapped I/O
  - I/O Registers are addressed in same space as memory
  - Address decoder distinguishes between them
  - OS uses address translation mechanism to make them only accessible to kernel
- I/O instructions : x86
  - Separate instructions to access I/O registers
  - Can only be executed in kernel mode

Polling vs. Interrupts

- Periodically check I/O status register
  - If device ready, do operation; otherwise wait
  - If error, take action
- Common in small or low-performance real-time embedded systems
  - Predictable timing, Low hardware cost
- In other systems, wastes CPU time
- When a device is ready or error occurs
  - Controller interrupts CPU
- Interrupt is like an exception
  - But not synchronized to instruction execution
  - Can invoke handler between instructions
  - Cause information often identifies the interrupting device
- Priority interrupts
  - Devices needing more urgent attention get higher priority
  - Can interrupt handler for a lower priority interrupt

I/O Data Transfer

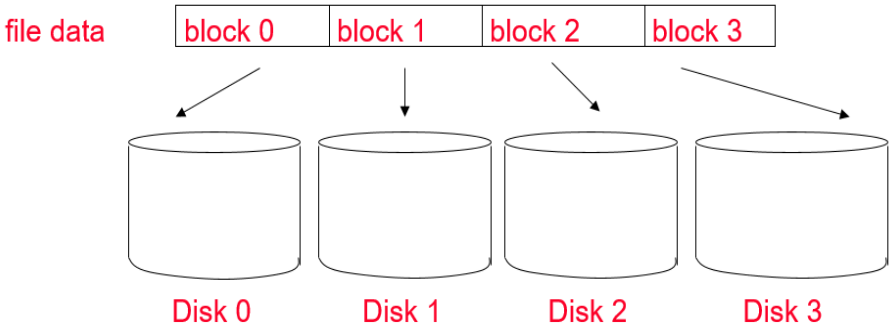
- Polling and interrupt-driven I/O
  - CPU transfers data between memory and I/O data registers
  - Time consuming for high-speed devices



- Direct memory access (DMA)
  - OS provides starting address in memory
  - Special I/O controller (DMA controller) transfers to/from memory autonomously
  - DMA controller interrupts on completion or error
- If DMA writes to a memory block that is cached
  - Cached copy becomes stale
- Need to ensure cache coherence
  - Invalidate blocks from cache on DMA writes to memory blocks

Mass Storage

- Many systems today need to store many terabytes of data
- Don't want to use single, large disk
  - too expensive
  - failures could be catastrophic
- Would prefer to connect many smaller disks to provide
  - large storage capacity
  - faster access to reading data
  - redundant data

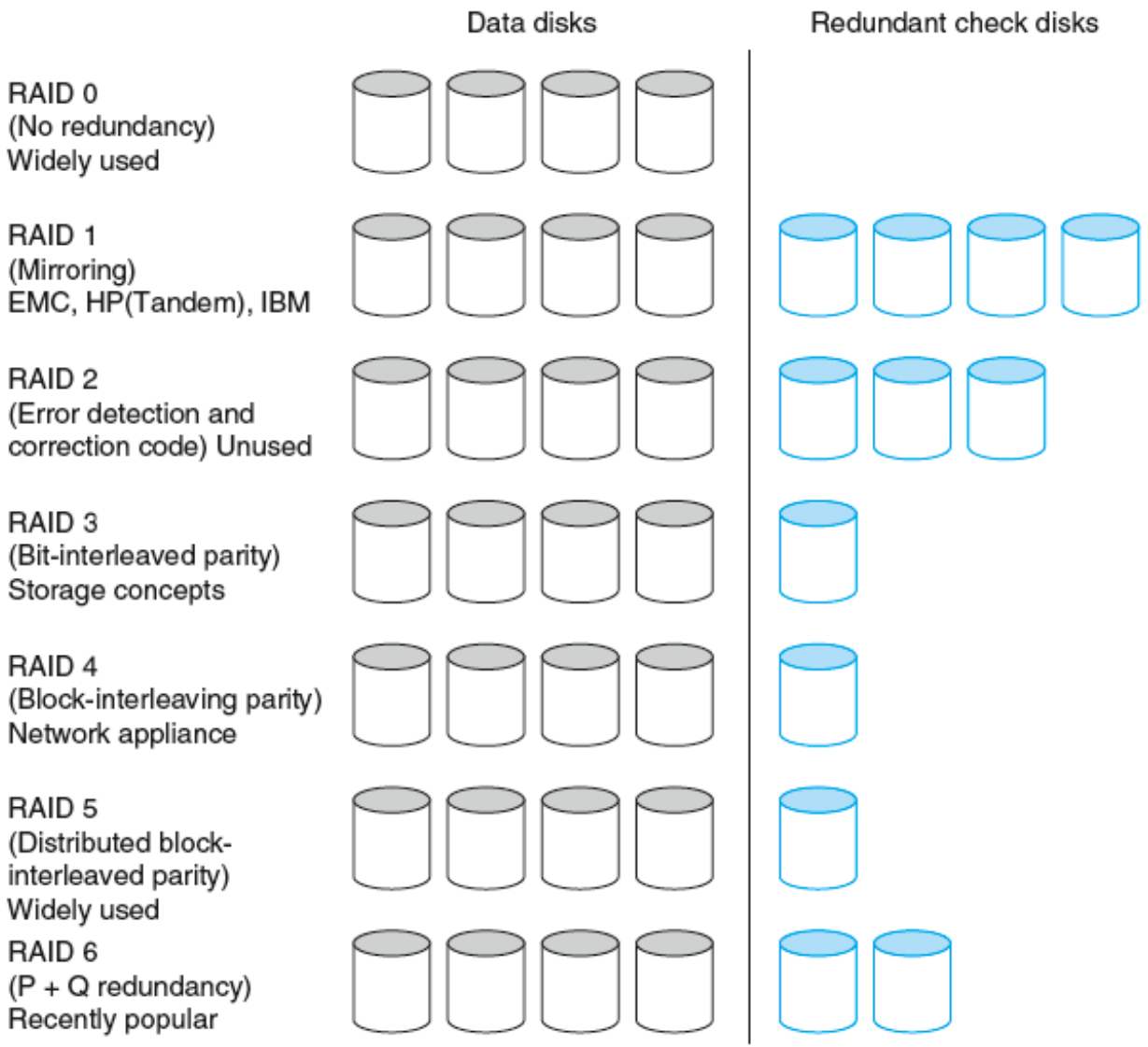


- Dependability Measures
  - Reliability
  - Availability

RAID

- Redundant Array of Inexpensive/Independent Disks
- Technology of managing multiple storage devices
  - Typically in a single machine/array, due to limitations of fault-tolerance
- Differing levels of redundancy, error checking, capacity, and cost
- RAID can improve performance and availability
  - High availability requires hot swapping

RAID Levels


Data disks
Redundant check disks
RAID 0
(No redundancy)
Widely used
RAID 1
(Mirroring)
EMC, HP(Tandem), IBM
RAID 2
(Error detection and
correction code) Unused
RAID 3
(Bit-interleaved parity)
Storage concepts
RAID 4
(Block-interleaving parity)
Network appliance
RAID 5
(Distributed block-
interleaved parity)
Widely used
RAID 6
(P + Q redundancy)
Recently popular

Concluding Remarks

- I/O performance measures
  - Throughput, response time
  - Dependability and cost also important
- I/O Management
- Buses used to connect CPU, memory, I/O controllers
  - Polling, interrupts, DMA
- RAID
  - Improves performance and dependability