Lecture 14: Input/Output & Storage Management

Adam Hawley

March 12, 2019

Contents

1 I/O Management

1.1 Introduction

I/O subsystem is responsible for controlling devices connected to a computer. It must provide processes with a sufficiently simple interface and also take device characteristics into account to maximise performance and efficiency.

There is a large variety of I/O devices:

- Storage (e.g disk drives, non-volatile memory)
- Communications (e.g. Ethernet, WiFi, Bluetooth, USB)

- User Interface (e.g. mouse, touch, keyboard, display, sound)

1.2 Device Drivers

Device drivers are low-level sofware that interacts directly with device hardware, they hide the hardware details to the higher levels of the OS and user applications and are often developed by the hardware vendor. They track the status of the device and enforce access/allocation policies. Types of drivers:

Dedicated Each device is allocated to a single process

Shared Each device is shared between multiple processes

Virtual Hides sharing from processes

1.3 Devices

Devices usually have registers where the device driver places commands, addresses and data to write or read data from registers after command execution. A minimum setup usually consists of the following:

- Data-In Register
- Data-Out Register
- Status Register
- Control Register

Where each register is typically 1-4 bytes and they may be contained in a FIFO buffer.

Devices themselves also have addresses used by direct I/O instructions or memory-mapped I/O.

1.4 I/O Management

The I/O subsystem provides interfaces to access devices via device drivers (or access to specific devices in a family of devices hidden by the device driver). There are three main device communication mechanisms:

- Polling & Interrupts
- Direct Memory Access (DMA)
- Buffering

1.5 Polling

Polling is about checking if a device is ready for communication so for each byte of I/O:

1. Read busy bit from status register until 0
2. Host sets read or write but and if write copies data into data-out register.
3. Host sets command-ready bit
4. Controller sets busy bit, executes transfer
5. Controller clears busy bit, error bit, command-ready bit when transfer done.

Step 1 is a busy-wait cycle to wait for I/O from device. This is reasonable if the device is fast but inefficient if it is slow. The CPU could switch to other tasks, but if miss a cycle data could be overwritten/lost.

1.6 Interrrupts

Polling can happen in 3 instruction cycles:

1. Read status
2. Extract status bit
3. Branch if not zero

How to be more efficient if devices are seldom ready?

CPU Interrupt-Request line triggered by I/O device (checked by processor after each instruction). The interrupt handler receives interrupts (maskable to ignore or delay some interrupts). Interrupt vector to dispatch interrupt to correct handler. This has a context switch at the start and at the end. We get **interrrupt chaining** if more than one device at same interrrupt number.

1.7 Direct Memory Access (DMA)

This is used to avoid programmed I/O (one byte at a time) for large data movement. This requires a DMA controller and bypasses CPU to transfer data directly between I/O device and memory.

The OS writes DMA command block into memory.

- Source and destination addresses
- Reade or write mode
- Count of bytes
- Writes location of command block to DMA controller.

2 Storage Devices

2.1 Introduction

The hierarchy of storage devices is driven by performance and volatility of data. **Data access time** includes:

Ready time Time to prepare set up storage media to read/write data at the appropriate location (e.g. wind/rewind tape, rotate disk, charge memory row)

Transfer time Time to read/write data from media

Different devices may impose access latencies at different orders of magnituse and hence, the OS should manage each of them appropriatly and mediate transfers (e.g. buffering).

2.2 Tertiary Storage

Tertiary storage is usually used for backups, storage of infrequently used data and transfer between systems. The two main forms of tertiary storage are:

- Magnetic tapes:
  - GB to TB capacity
  - Very slow access time (must wind and rewind to position tape under read-write head but once in place, reasonable transfer rates >140 MB/s)
- Optical discs:
  - MB to GB capacity
  - Read-only or read-write using high intensity laser beams

2.3 Secondary Storage

Secondary storage is mainly used for non-volatile storage, high-capacity storage supporting swapping/paging.

2.3.1 Magnetic Disks (HDDs)

- Made of n disks (2/n/ sides), each side is divided into tracks (circular), and each track into sectors.
- **Cyclinder**: Set of tracks at the same position on all sides
- **Access Time**: Seek time (disk head movement) + Search time (rotational delay) + Transfer time
- Typical Avg Values:
  - Seek = 25ms
  - Search = 4ms
  - Transfer = 0.00094ms/MB
  - Rotation speed = 7200rpm (120rps)

2.3.2 Non-volatile memory (NVMs, SSDs)

- Made of no mechanical components
- Redundant Arrays of Independent Disks (RAIDs)