IB CS Course Notes

Uzen

Oct 29, 2020

Contents

Chapter 1

System Fundamentals

1.1 Systems in Organizations

1.1.1 System life cycle

System life cycle discusses the process of developing a new system.

- Existing System Analysis
- Requirement Specification
- Software Design
- Software Implementation
- Testing & Debugging
- New System Installation
- Maintenance

Often times, we can simplify this to the `Software Life Cycle`

- Planning and Analysis
- Design Overview
- Development
- Evaluation

1.1.2 Planning a New System

We can use the acronym `TELOS` to plan whether or not to use a new system.

Technical Feasibility: Can we actually implement this system?
Economic Feasibility: Is the system cost effective?
Legal Feasibility: Is there any legal issues?
Operational Feasibility: Is the current practices sufficient to support the new system
Schedule Feasibility: How long do we need to wait?

1.1.3 Change Management

`Change Management` refer to the process of shifting individuals, teams, departments, and organizations. The purpose of Change Management is move a company into a more efficient state, while minimizing the negative effects of the process.

1.1.4 Compatibility Issue

There may be compatibility issues when encountering a `Legacy System`. A Legacy System is a general term referring to old technology. For example, an old printer may be called a legacy printer. Keeping around legacy systems is troublesome, since companies may no longer support the technology (replacement parts are no longer produced, software is incompatible). There may also be compatibility issues when interacting with legacy systems, as it might no support new protocol, or have a outdated proprietary software.

Four Strategies of Integration

1. Keep both information system, and develop them to have the same functionality (high maintenance cost)
2. Replace both information systems with a new one (increased inital cost)
3. Select the best information system from each company and combine them (it is very difficult for the employees to work information systems from another company)
4. Select one company's information systems and drop the other companies' (policy problems)

1.1.5 SaaS

Nowadays, there is many different methods of providing business software. Some clients install their software on their own premises, while other have the software installed on dedicated servers that provide these services.

SaaS is one of the methods of providing services. Software and data is hosted on a remote data center, where clients pay fees to access the services. SaaS relies heavily on the Internet. Some examples of SaaS are Dropbox, Gmail, and Salesforce.

Nowadays, SaaS is less expensive and have a lower initial cost. It is easier to install, maintain, and upgrade. It is easily scalable, allowing for rapid growth. It requires less IT staff to manage.

However, SaaS requires a constant Internet connection. It is based on a monthly subscription. Since the information is stored on a remote server, there is no guarantee of security.

1.1.6 Alternative Installation Process

Installing a new system is a necessary evil in many enterprises. Many users do not like to adjust to a new system, and changeover can cause many compatibility issue. Yet, installing new software is an essential part of the everyday operation of the IT Department.

There are four common types of software changeovers:

Parallel

In a parallel changeover, both software work simultaneously for a period of time. If the new system fails, the old system is usable. Only when the new system have been running smoothly can the old system be retired.

Parallel changeover is the safest method, but the most time consuming. As both system is required to be running, additional resources is needed to support both systems. However, this method is not effective if the old system and new system does different tasks.

Big Bang

Big Bang, or direct changeover is a risk operation where the old system is retired immediately, and the new system is immediately put in use. This system is risky, and should be only used for non critical systems. In this case, all users need to be trained before-hand on the new system.

Pilot

A pilot changeover is used in a company with a lot of sites. In a pilot changeover, a new system is introduced in sites one at a time. This method allows for a enterprises to test out the new system before implementing it fully. And the feedback received from each system can be used to improve the changeover at other sites.

Phased

A phased conversion is used a system with many modules. A company may change one module at a time, so that the system is converted at different times. This approach allows for an extended training an adoption of the system, but takes more time.

1.1.7 Data Migration

Data migration occurs when data in transfered between different storage devices or different formats. There are many problems that will occur when data is transfered between two systems. Errors such as incompatibility, corruption, virus, etc, can cause irreparable lost to the data. To guarantee the safety of the data, following three steps when migrating data.

1. Plan
2. Migrate
3. Validate

Chapter 2

Computer Organization

2.1 Computer Architecture

The Central Processing Unit (CPU) is an integral part of a computer. The CPU follows the `input, process, output model`, and is responsible for doing most of the calculations our computers need to do in order to run. The CPU is the brain of a computer.

A CPU contains:

- Control Unit
- Arithmetic Logic Unit
- Memory Address Register
- Memory Data Register

CU is responsible for operation of the CPU. It retrieves instructions from memory, then execute.

ALU performs math/logic/IO operations. The CU provides ALU data and instructions.

MAR is responsible for the memory address of data needed, while MDR is responsible for retrieving the data from the memory address.

2.2 Primary Memory

- 1 bit = one or zero
- 8 bits = 1 byte
- 64 bit system = 64 bits in each Memory Bus data transfer
- MB = Mega Byte, Mb = Mega bit

Memory is separated into `Random Access Memory` and `Read Only Memory`.

2.2.1 RAM

RAM stores program instructions and relevant data. Data in RAM is stored in unique places: Memory Addresses.

Note, RAM allows for frequent read-write, and is volatile. If power is lost, all information is lost.

RAM is separated into two different types:

- Dynamic RAM
- Static RAM

DRAM

DRAM is slower than SRAM, but cheaper. This is the RAM sticks we use.

SRAM

SRAM is faster but more expensive (and smaller). SRAM is placed between the CPU and RAM. It is also called a cache.

Cache holds useful information CPUs frequently uses. This is so that the CPU does not have to use the slower DRAM when fetching instructions. The CPU always read and write through the cache memory.

There are 3 levels of cache memory. L1 is on the processor itself. L2 lies between primary memory and the CPU. Newer L3 cache is fulfilling similar roles to a L2 cache.

2.2.2 ROM

ROM holds instructions. But ROM is read only, and stores permanent information. Such as instructions for boot or the BIOS (Basic Input Output System).

Often, ROM stores embedded data that is non-volatile, and is used to store permanent programs. ROM is also commonly smaller than RAM.

2.3 Registers

Registers are a small storage location that holds data. We use the word `word` to indicate the register size (in bytes).

Both MAR and MDR are registers. MAR stores memory address, MDR stores data (which is then saved to RAM). A Memory (Address) Bus helps communicate between the MAR and the primary memory. A Data Bus helps communicate between the MDR and the ALU.

2.4 Machine Instruction Cycle

This is the steps in which instructions are carried out in machine code.

1. Fetch instructions from Primary Memory to Control Unit
2. Decode instructions in Control Unit (loads additional data)
3. Execute instructions
4. Stores results and check for next instruction

2.5 Secondary Memory

Slower memory than is non-volatile. Used to store data permanently from RAM.

During startup, RAM is empty. And thus Secondary Memory is needed to load data onto RAM.

When primary memory is not enough for one application, Virtual Memory enables program to store their data on Secondary Memory. Virtual Memory moves the memory of unused application to secondary memory, freeing up space for newly opened application. The memory is loaded into primary memory back on when the unused application is accessed again.

Some examples of secondary memory are:

- Hard Drive
- DVD/CD
- USB
- SD Card
- Floppy Disk
- Magnetic Tape

2.6 Operating and application Systems

An `Operating System` is a set of software that controls the computer's hardware resources and provides services for computer programs. It interactive between the user and the hardware of the computer system.

The main services of an OS are:

- Peripheral Communication
- Memory Management
- Resource Monitoring and Multitasking
- Networking
- Disk Access and Data Management
- Security

2.7 Software Application

Computer system uses software application to complete some task. The main types of Software Application includes:

- Word Processors
- Spreadsheets
- Database Management Systems
- Web Browsers
- Email
- Computer Aided Design
- Graphic Processing Software

2.7.1 Word Processor

A software used for production of any sort of document. Used for composing, editing, formatting, and printing of a document.

2.7.2 Spreadsheets

Spreadsheets are programs used for the analysis of data. Data is represented on cells, organized in rows and columns. Spreadsheets are used for operation of arithmetic and mathematical functions.

2.7.3 Database Management System

A DBMS is an application application that manages a database. Common features of a DBMS includes:

- Create Queries
- Update Stores
- Modifies Queries
- Extract Information

This program provides an interface between a database between an user. A database can often be represented as table, with rows and columns of data.

2.7.4 Web Browser

A web browser is a software that is used to access, retrieve, and present content on the Internet.

2.7.5 Email

Electronic mail is an application that sends digital message as mail. Email use the `Simple Mail Transfer Protocol` (SMTP).

2.7.6 Computer Aided Design

CAD is a type of software that assists engineers is creating a design. CAD engineers to inspect a design from many positions or angles, allowing them to create designs better.

2.7.7 Graphics Processing Software

A graphic processing software allows designers to edit, create, crop, erase, an image.

2.8 Common Features of an Application

The two typical applications in a computer system is a GUI and a CLI.

A Command Line Interface (CLI) is an application than runs in a terminal. A CLI is faster, use less memory, and more powerful. Yet, it is harder to learn.

A Graphical User Interface (GUI) is an application that utilize icons and pictures to make the application more accessible. It is easy to use compared to a CLI. But requires more memory, a graphical display and a mouse, and is more difficult to implement.

2.8.1 GUI Implementation

A common GUI includes several elements:

- Toolbar
- Menu
- Dialogue box

2.9 Binary Representation

Computer systems use bits and bytes to store information. This is a binary system.

There is 8 bits in a byte. A bit is a boolean, either yes or a no. Or more accurately, a bit is a digit in binary.

The number system we commonly use is base-10. This means for every digit incremented, the number will add by 10^n. Binary is base-2, where every digit incremented, the number will add by 2^n.

2.9.1 Binary

Here's a table showing the multipliers of an eight bit binary:

Digit	8	7	6	5	4	3	2	
Power of 2	2^7	2^6	2^5	2^4	2^3	2^2	2^1	
Multiplier	128	64	32	16	8	4	2	

2.9.2 Negative Numbers In Binary

There is two methods of indicating a negative number: `two's complement` and

Chapter 3

Networks

Vocabulary

- Bus Topology: Computer network in which a "bus" connects all the devices together with a cable
- Cable: a wire that is used for networking
- Check Digit: Extra digit added to check data integrity after input, transmission, storage, or processing
- Data Integrity: Accuracy of data after input, transmission, storage, or processing
- Data Packet: Portion of a message that is transmitted through a network. Contains data such as check digits and destination address.
- Gateway: Link that resides between computer networks that convert data so that it is understood by the receiving network.
- Handshaking: exchange of predetermined signals to signify that a connection has been established between two systems.
- Hub: Network connection point for devices. Data arriving at a hub is copied over and sent to all the devices on the network.
- ISDN (Integrated Services Digital Network): An international communication standard that allows for the transmission of media over digital telephone lines.
- LAN (Local Area Network): A computer network where the computers are connected locally.
- Microwave Transmission: Wireless electronic communication

- Modem (Modulator/Demodulator): A device that converts electrical device to audio signals that travels on telephone lines.
- Network: Computer systems that are interconnected.
- Packet: Group of bits that form some sort of data.
- Packet Switching: method of network communication that transmit packets through a network.
- Networking: Using a network.
- Parity bit: Error detecting bit that is appended on data in order to check for data integrity
- Protocol: International rules that ensure the transfer of data between system.
- TCP/IP (Transmission Control Protocol/Internet Protocol): Communication protocol used for connection on the Internet.
- WAN (Wide Area Networks): A connected network of LANs.

3.1 Network Fundamentals

3.1.1 Different Types of Computer Networks

In a computer network, there is often a *Server* and a *Client*. A *Server* provides services to the *client*.

In a network, there is three important device. A `hub` is a connection point for devices on a network. Network devices are connected to the hub through a cable. When a client request some data from a server, the hub copies the data and send it to all devices connected to it. This is sometimes problematic, as it causes extra traffic in the network.

Similar to a hub, a `switch` is a connection point for devices on a network. However, a switch is able to identify the device on each port, and transmit data accordingly to the receiver. Nowadays, hubs are obsolete.

A `router` is device that join multiple networks. One potential job of a router is joining the LAN with the Internet. A simple switch cannot accomplish this job. However, nowadays, routers and switch are mostly integrated into one device, allowing the creation of wireless networks connected to wired ones.

Below is a list of the common network types:

- LAN
- WLAN (Wireless LAN)
- VLAN (Virtual LAN)
- WAN

- SAN (Storage Area Network)
- Internet
- Extranet
- VPN (Virtual Private Network)
- PAN (Personal Area Network)
- Peer to Peer (P2P)

In fact, the Internet is a really big WAN that is connected to thousands of LAN around the world.

Local Area Network

In a LAN, peripheral devices and data can be shared. There is a high data-transfer rate between computers on the network.

WLAN

A WLAN is just a wireless LAN. WLANs generally are not as fast as an wired network, but allows the wireless devices to move freely. However, the use of WiFi also reduce the safety of the data transfered. To work around this, secure protocals (WEP, WPA, WPA2) should be used.

VLAN

VLAN stands for Virtual Local Area Network. A VLAN artificially divides a network into separate compartments. Logically separated computer networks are similar to physically separate networks. Unless separate rules have been set, the computer cannot see across a VLAN. VLANs allow for easy management of the resources on a network, and improves security.

WAN

A WAN connects many LANs across a wide geographic area. For example, a company may have a WAN connecting its many headquarters around the world.

SAN

A SAN network is a network specifically designed for the sharing of large storage devices.

Intranet and Extranet

An intranet is a collection of private computer networks within an organization that is not connected to the Internet. Intranets are secure and fast, but hard to access outside of the LAN.

Similarly, an Extranet is a collection of private computer network, but opened to the Internet. This is less secure, but allows for more flexibility.

Internet

The Internet is just a very large WAN connecting billions of LANs around the world together.

IoT

Internet of Things are small devices that can connect to the Internet. These small devices may send and receive data.

VPN

A Virtual Private Network allows for an encrypted tunnel to be established between two computers. This allows for remote workers to access the resources on a company Intranet while ensuring safety of data.

PAN

A Personal Area Network is composed of devices that are centered around an individual's workspace. These devices have very short ranges.

P2P

A Peer to Peer network do not uses the client/server network, but instead a distributed network architecture where all computer systems are decentralized. The same device can be both a server and a client.

3.1.2 Importance of Standards

There are two main organization setting standards for computer communication. Institute of Electrical and Electronics Engineers (IEEE) and Internet Engineering Task Force (IETF). The standards they create help create compatibility between hardware and software, computers of different brands and origins. Without standards, communication between computers are nearly impossible.

3.1.3 Networks, Communication, and Layers

In order to simplify the many layers of inter computer communication, a network is split into a different layers, where each layer represents a module. The `OSI Model` is created for this reason. The OSI Model stands for Open System Interconnection Model.

The OSI Model is split into 7 parts.

1. Application: Various services used by the end user
2. Presentation: Provides data format, compression, encryption information.
3. Session: Manages a connection between two users
4. Transport: Definition of data segments, assignment of numbers, data transfer, re assemblage of data
5. Networking: Handling packet switching
6. Data Link: Error handling for transitions, transmission rate, flow control
7. Physical: Transmission of 1 and 0 through cables and microwave. Definition of media and specification.

The `TCP/IP` protocal describes functions that take place in each layer.

1. Application: Perform services for software.
2. Transport: The transformation of data.
3. Internet: Packet Switching
4. Network Access: Media and devices

3.1.4 Technology required for a VPN

Hardware requirements

- Internet Access
- VPN Software
- VPN Routers
- VPN Appliances
- VPN Concentrators (handles large number of incoming VPN tunnels)
- VPN Servers

For *Secure* VPN (encrypted, tunneled VPN)

1. IPSec (Internet Protocal, Security Protocal)

- AES (Encryption)

2. SSL (Secure Socket Layer) or TLS (Transport Layer Security)

For *Trusted* VPN (Secure on provider side)

1. ATM (Asyncronous Transfer Mode) circuits
2. Frame delay circuits
3. Transport of layer 2 frames over MultiProtocol Label Switching

`Hybrid` VPNs are a mix of Trusted and Secure VPN.

3.1.5 Common VPN Types

- Site to site VPN (Connecting network or facilites together)
- Remote access VPN (remote into an intranet)

3.1.6 Uses of a VPN

VPN Benefits

- Easy Communication
- Secure connection
- Decrease operational cost when compared to WAN
- Extended connection
- Allows for remote work
- Allows global networking
- Simplify network topology

3.2 Data Transmission

3.2.1 Necessity of Protocols

Requirements for communication to take place:

- Presence of an identified sender
- Presence of an identified receiver
- Presence of an agreed-upon method of communicating
- Presence of a common language
- Presence of a common grammar

- Presence of an agreed-upon speed and timing of delivery
- Presence of confirmation or acknowledgment requirements

Computer Networks Protocols also provides:

- Rules about the message format
- Rules about the way intermediary devices should facilitate communication
- Rules about initiation and termination of a communication session
- Rules about the type of error checking to be used
- Rules about an error detection and correction mechanism
- Rules about recovery and resending data

Computer Network Protocol guarantee:

- Data Integrity
- Source Integrity
- Flow Control
- Congestion Management
- Deadlock Prevention
- Error Checking
- Error Correction

3.2.2 Speed of Data Transmission across a network

We measure the ability of a medium to transfer data as bandwidth. Bandwidth is measured in $Mbps$, or Mega-bits per second. This is different from $MBps$, which means Mega-Byte per second.

The theoretical speed of data is called bandwidth, the actual transfer rate is called throughput. While the slowest part of a network is called the bottleneck. Goodput measures the transfer rate of usable data.

There are several factors that affect the speed of data transmission in a network:

- Bandwidth
- Data transfer rate of storage devices
- Interference
- Malicious software

- Number of connected deviecs
- Packet loss
- Bottleneck
- Client/Server specification

3.2.3 Compression of Data

In order to reduce bandwidth, compression is sometimes used to reduce the file size. There is two types of compression, Lossy and Lossless. Text have to be compressed losslessy, but some medias can be compressed with lossy compression.

3.2.4 Characteristics of different transmission media

There are three most common wires. Fiber Optic cable, coaxial cables, and Unshielded Twisted Pair Cable (UTP). There are many wireless communication methods.

- Microwave Radio
- Satellites
- Infrared
- RFID
- Bluetooth
- Free Space Optics

In general, fiber optics is the fastest, most secure, and most reliable. Wireless is not as secure, and not as fast or reliable. UTP lies between the two.

3.2.5 Packet Switching

Packet switching is the communication method in which data packets are transfered through a network. Packets are sent individually, and may take different paths to the same destination. There is two method types of packet switching.

Datagram packet switching, where each packet is sent with an address. The route can be different per packet.

Virtual Circuit packet switching, where the route and the destination is predetermined before hand.

3.2.6 Network Topology

There are eight basic topologies:

- Point to point
- bus
- star
- ring
- mesh
- tree
- fully connected
- hybrid

3.3 Wireless Networking

3.3.1 Advantages and Disadvantages

Advantages:

- Ease to setup
- Cheap
- Wireless/Convenient
- Easier to plan

Disadvantage:

- Slow
- High error rates
- Weather dependent
- Weakest security
- Unreliable

3.3.2 Hardware and Software Components

Hardware needed

- Modem
- Wireless Router/Access Point
- Wireless Network Adapter (NIC)
- Device
- Wireless antennas
- Wireless repeater
- Ethernet repeater
- Ethernet over power line

Software needed

- DHCP (IP Address assigning)
- Software Firewall
- Name/SSID (Service Set Identification)
- NIC Drivers (Network Interface Card)
- Operating System
- Security Software
- WAP (Wireless Application Protocol)
- Web Browser