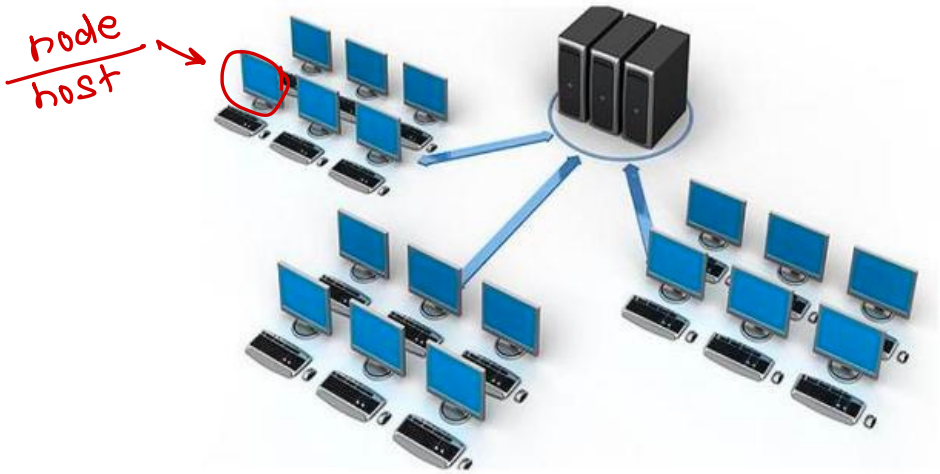
Cloud Computing

Networking

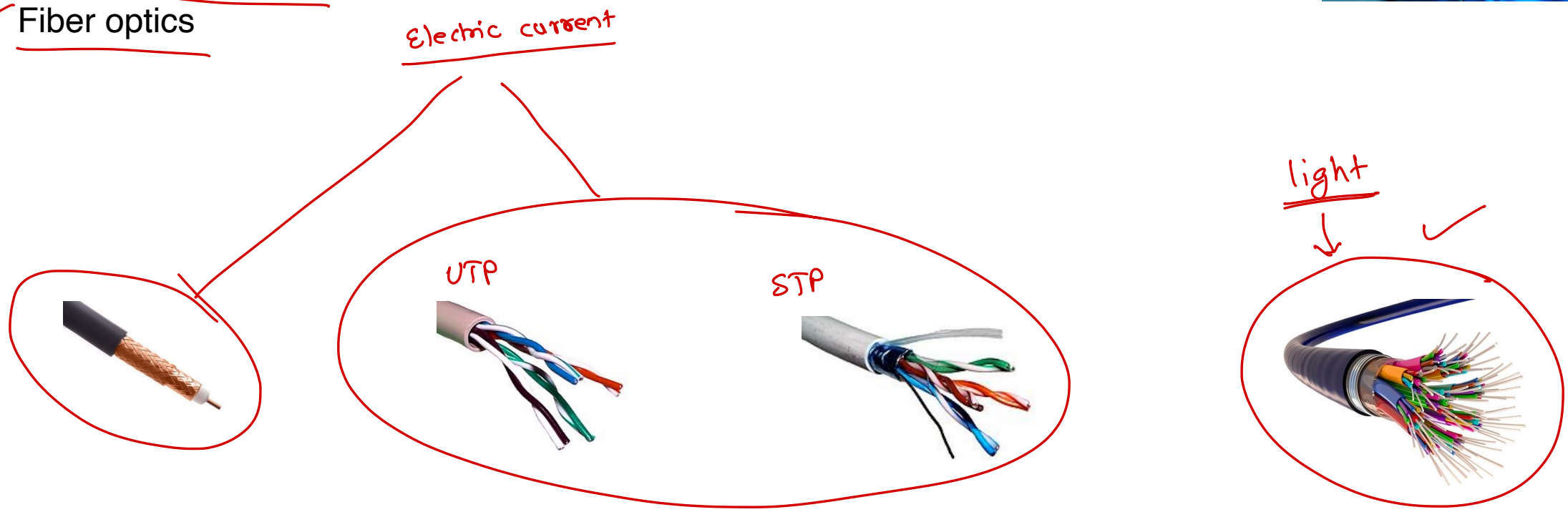
What is network ?

- It is the interconnection of multiple devices, generally termed as Hosts connected using multiple paths
- The purpose of network is: sending/receiving data or media
- It involves various devices like hubs, switches, routers etc.



Wired network

- The network build by connecting devices together using wires/cables as a medium to transfer the data
- Cables
  - Coaxial cable
  - Twisted pairs cables
  - Fiber optics



Wireless network

- The network build by connecting devices together using air as a medium to transfer the data
- EM Waves are used to transfer data from sender to receiver

EM Waves

Network Types

- Personal Area Network (PAN)
  - Smallest network which is very personal to the user
  - E.g. BlueTooth
- Local Area Network (LAN)
  - Spans across building(s) and operated under single administrative system
  - E.g. company, school network
  - Technologies: TokenRing or Ethernet
- Metropolitan Area Network (MAN)
  - Spans across cities
  - E.g. cable network
  - Technologies: high speed fiber optics
- Wide Area Network (WAN)
  - Spans across countries
  - Technologies: ATM, Frame Relay

school, college, organization

What is a network topology ?

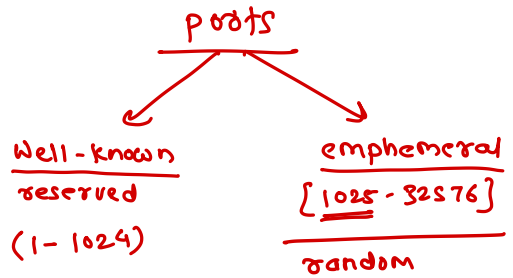
- Physical arrangement of computers is known as topology
- Famous topologies
  - Bus
  - Ring
  - ~~Token Ring~~ connected Ring
  - Star
  - Mesh

ISO OSI model [Open System Interconnect]

- Conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard to its underlying internal structure and technology
- Goal is the interoperability of diverse communication systems with standard communication protocols
- Layered architecture having 7 layers
  - Application
  - Presentation
  - Session
  - Transport
  - Network
  - Data Link
  - Physical

Application Layer

- Specifies interface methods used by hosts in a communications network
- Contains communication protocols
  - **HTTP [80]**: Hyper Text Transfer Protocol
  - **HTTPs [443]**: Secure Hyper Text Transfer Protocol
  - **FTP [20, 21]**: File Transfer Protocol
  - **SFTP [115]**: Simple FTP
  - **DNS [53]**: Domain Name Service
  - **NFS [1023]**: Network File System
  - **POP3 [110]**: Post Office Protocol
  - **SMTP [25]**: Simple Mail Transfer Protocol
  - **SSH [22]**: Secure Shell
  - **LDAP [389]**: Lightweight Directory Access Protocol



Presentation Layer

- Serves as the data translator for the network
- Also known as syntax layer
- Responsible for
  - Translation
  - Compression/Decompression
  - Encoding/Decoding
  - Encryption/Decryption

Sender: ASCII → EBCDIC | Receiver: EBCDIC → ASCII

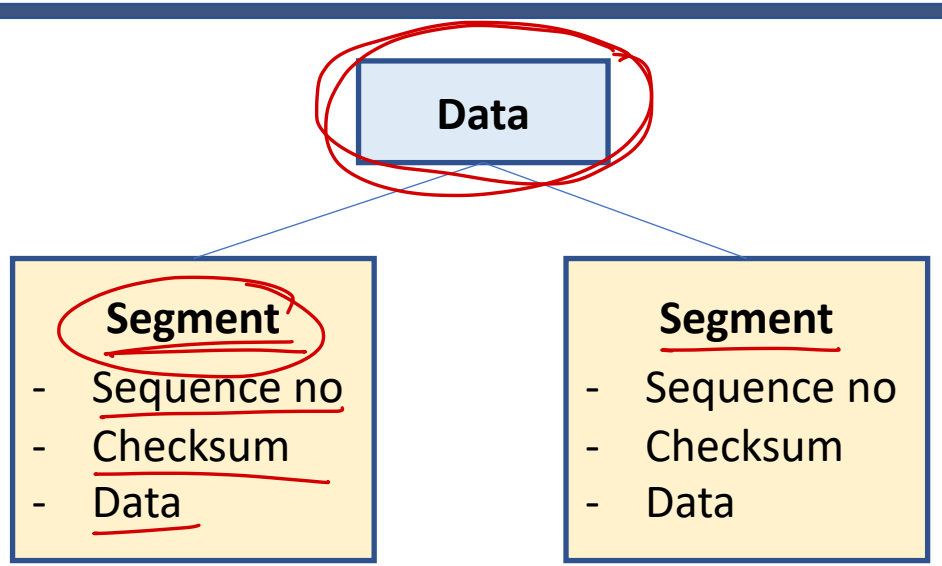
: saves bandwidth & time

Session Layer

- Provides mechanism for opening, closing and managing session between processes
- Communication sessions consist of requests and responses that occur between applications
- Protocols
  - **ASP**: AppleTalk Session Protocol
  - **ADSP**: AppleTalk Data Stream Protocol
  - **NetBIOS**: Network BIOS
  - **PAP**: Password Authentication Protocol
  - **PPTP**: Point to Point Tunnelling Protocol (VPN)
  - **RPC**: Remote Procedure Call
  - **SCP**: Session Control Protocol
  - **SDP**: Socket Direct Protocol

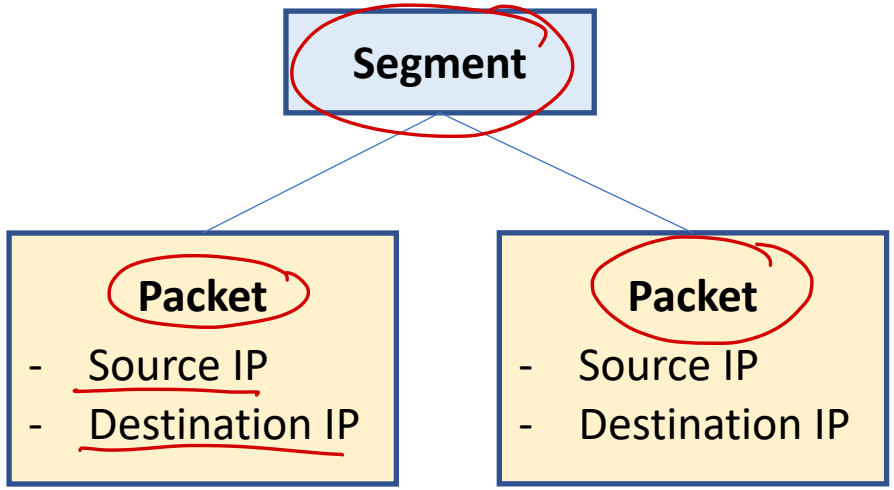
Transport Layer

- Provide host-to-host communication services for applications
- Creates Segment (data unit) containing
  - Sequence number
  - Checksum : used to check if the receiver has received data intact
  - Port number
- Protocols
  - TCP
    - Connection oriented protocol
    - Provides: Flow Control, Error checking
    - Guarantees data delivery
    - Slower than UDP
    - E.g. WWW, HTTP
  - UDP
    - Connectionless protocol
    - Does not provide flow control
    - Does not guarantee data delivery
    - Faster than TCP
    - E.g. streaming, online games



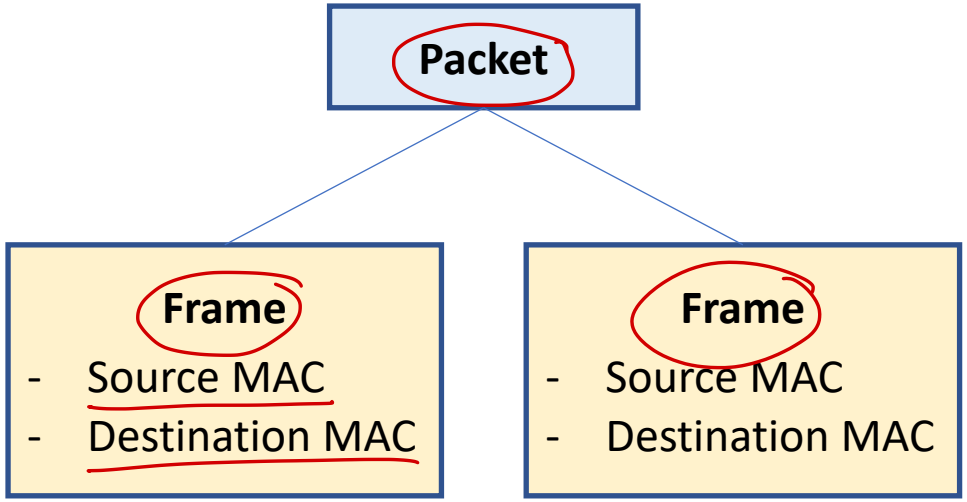
Network Layer

- Responsible for packet forwarding including routing through intermediate routers
- Responsible for splitting segment into packets containing
  - Source IP address
  - Destination IP address
- Protocols
  - **IP**: Internet Protocol
  - **IPX**: Internetwork Packet Exchange
  - **IPSec**: Internet Protocol Security
  - **EGP**: Exterior Gateway Protocol



Data Link Layer

- Transfers data between
  - adjacent network nodes in a wide area network (WAN) or
  - between nodes on the same local area network (LAN) segment
- Encapsulates packet into Frames containing
  - Source MAC Address
  - Destination MAC Address
- Sublayers
  1. Logical Link Layer
  2. Media Access Control Layer



Data Link Layer: Logical Link Layer

- The uppermost sublayer multiplexes protocols running at the top of data link layer, and optionally provides flow control, acknowledgment, and error notification
- Provides addressing and control of the data link
- Services
  - Error control (automatic repeat request, ARQ)
  - Flow control [Data-link-layer flow control is not used in LAN protocols such as Ethernet, but in modems and wireless networks]

Data Link Layer: Media Access Control Layer

- Refers to the sublayer that determines who is allowed to access the media at any one time (CSMA/CD)
- Determines where one frame of data ends and the next one starts (frame synchronization)
- Frame synchronization uses: time based, character counting, byte stuffing and bit stuffing.
- Services
  - Multiple access protocols for channel-access control,
    - CSMA/CD protocols for collision detection and re-transmission in Ethernet networks
    - CSMA/CA protocol for collision avoidance in wireless networks
  - Physical addressing (MAC addressing)
  - LAN switching (packet switching), including MAC filtering, Spanning Tree Protocol (STP) and Shortest Path Bridging (SPB)
  - Data packet queuing or scheduling

Physical Layer

- Consists of the electronic circuit transmission technologies of a network
- Fundamental layer underlying the higher level functions in a network which provides means of transmitting raw bits rather than logical packets or segments
- The bitstream may be grouped into code words or symbols and converted to a physical signal that is transmitted over a transmission medium
- Translates logical communications requests from the data link layer into hardware-specific operations to cause transmission or reception of electronic signals
- Services
  - Modulation/Demodulation
  - Multiplexing
- Consists of
  - Cables/wires
  - Devices like hub, repeaters etc.

Addressing Modes: MAC Address

- Used to identify NIC uniquely
- Consists of 6 bytes [48 bits]
- First 3 bytes represents manufacturer
- Next 3 bytes represents NIC's unique address

Windows : ipconfig /all

Linux/macos : ifconfig

Addressing Modes: IP Address

- Used to identify every device uniquely
- Set by operating system running on the device
- Can be written in
  - Decimal: 192.168.100.10
  - Binary: 11000000.10101000.01100100.00001010
- Types
  - Private: used to communicate with other devices in local network
  - Public: used to communicate with other devices over internet
- Versions
  - IPv4
    - 32 bit [4 bytes] address
    - Classful and Classless addressing
  - IPv6
    - 128 bit address

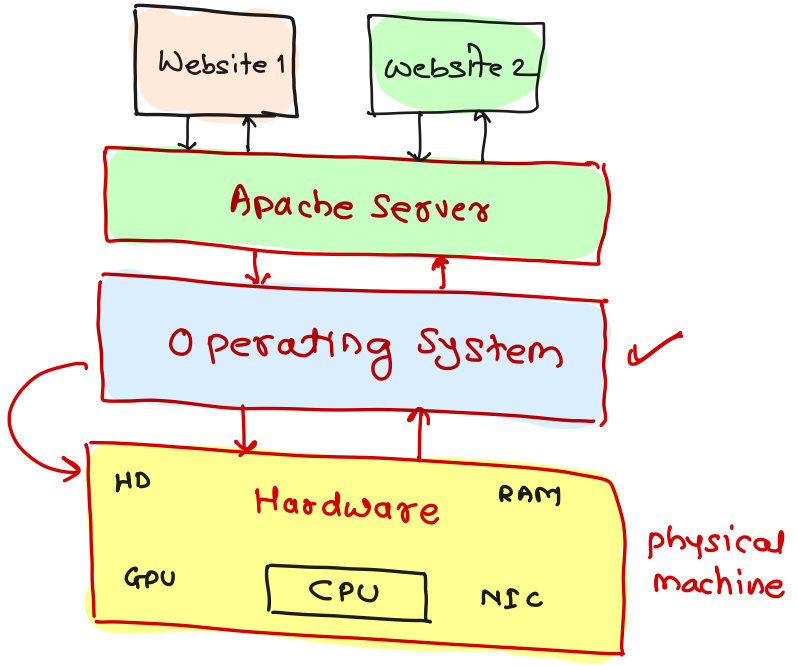
Virtualization

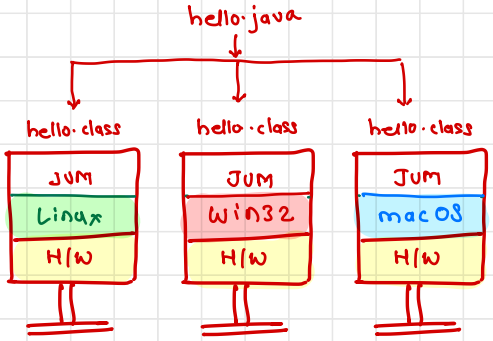
Software Deployment

- Software deployment defines a process of making the software available for the users
- E.g. a web site can be available for the end users, when it is hosted on a machine which can be accessed from anywhere in the world

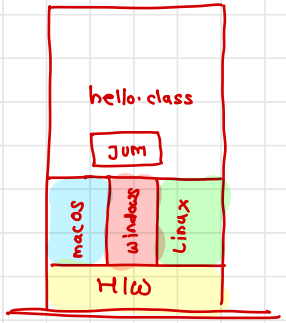
Traditional Deployment

- Early on, organizations ran applications on physical servers
- There was no way to define resource boundaries for applications in a physical server, and this caused resource allocation issues
- For example, if multiple applications run on a physical server, there can be instances where one application would take up most of the resources, and as a result, the other applications would underperform
- A solution for this would be to run each application on a different physical server
- But this did not scale as resources were underutilized, and it was expensive for organizations to maintain many physical servers





physical machines



multi-booting

What is virtualization

- Virtualization is the creation of a virtual -- rather than actual -- version of something, such as an operating system (OS), a server, a storage device or network resources
- Virtualization uses software that simulates hardware functionality in order to create a virtual system
- This practice allows IT organizations to operate multiple operating systems, more than one virtual system and various applications on a single server
- Types
  - Network virtualization
  - Storage virtualization
  - Data virtualization
  - Desktop virtualization
  - Application virtualization
  - Hardware virtualization

Network Virtualization

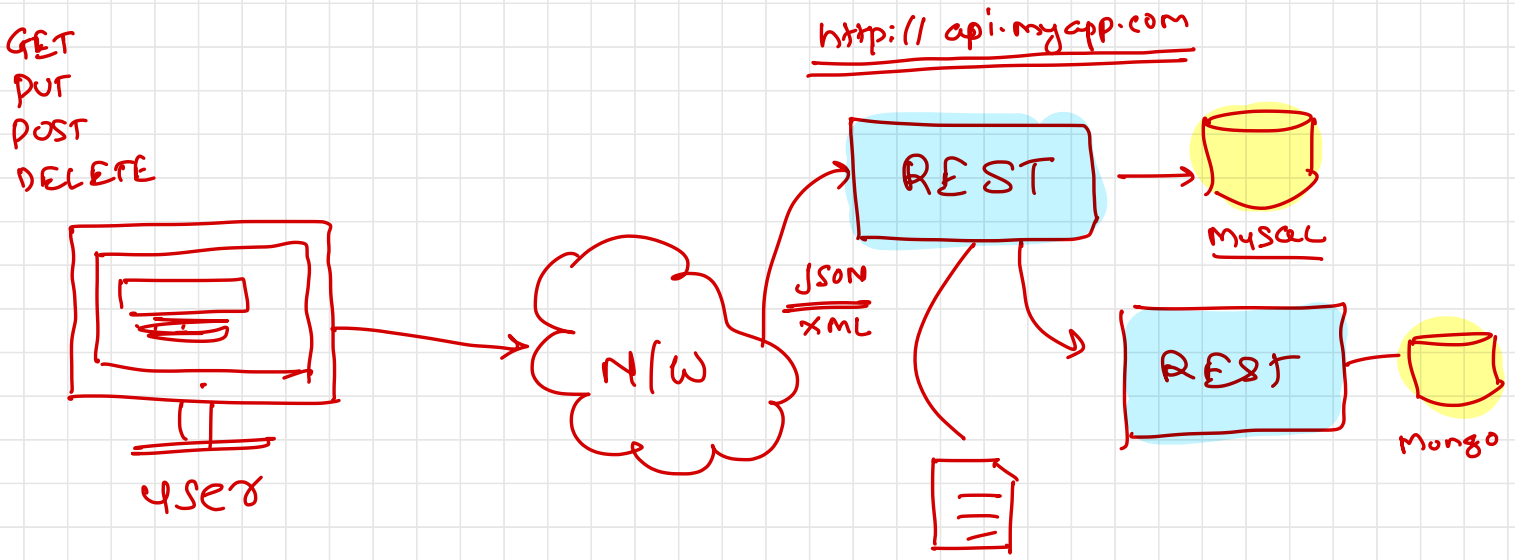
- Network virtualization takes the available resources on a network and breaks the bandwidth into discrete channels
- Admins can secure each channel separately, and they can assign and reassign channels to specific devices in real time
- The promise of network virtualization is to improve networks' speed, availability and security, and it's particularly useful for networks that must support unpredictable usage bursts

Storage Virtualization

- Storage virtualization is the pooling of physical storage from multiple network storage devices into what appears to be a single storage device that is managed from a central console
- Storage virtualization is commonly used in storage area networks
- Applications can use storage without having any concern for where it resides, what technical interface it provides, how it has been implemented, which platform it uses and how much of it is available
- Benefits
  - Makes the remote storage devices appear local
  - Multiple smaller volumes appear as a single large volume
  - Data is spread over multiple physical disks to improve reliability and performance (RAID)
  - All operating systems use the same storage device
  - Provided high availability, disaster recovery, improved performance and sharing

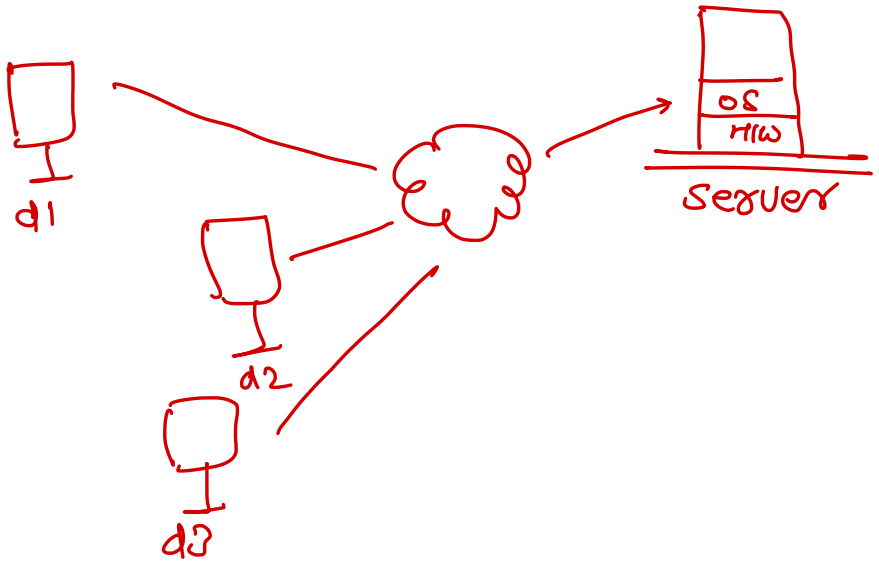
Data virtualization

- Data virtualization is the process of aggregating data from different sources of information to develop a single, logical and virtual view of information so that it can be accessed by front-end solutions such as applications, dashboards and portals without having to know the data's exact storage location
- The process of data virtualization involves abstracting, transforming, federating and delivering data from disparate sources
- The main goal of data virtualization technology is to provide a single point of access to the data by aggregating it from a wide range of data sources
- Benefits
  - Abstraction of technical aspects of stored data like APIs, Language, Location, Storage structure
  - Provides an ability to connect multiple data sources from a single location
  - Provides an ability to combine the data result sets across multiple sources (also known as data federation)
  - Provides an ability to deliver the data as requested by users


GET
PUT
POST
DELETE
http:// api.myapp.com
REST
MySQL
JSON
XML
N/W
REST
Mongo
User

Desktop virtualization

- With desktop virtualization, the goal is to isolate a desktop OS from the endpoint that employees use to access it
- It provides an ability to connect to the desktop from remote site
- When multiple users connect to a shared desktop, as is the case with Microsoft Remote Desktop Services, it's known as shared hosted desktop virtualization



Application Virtualization

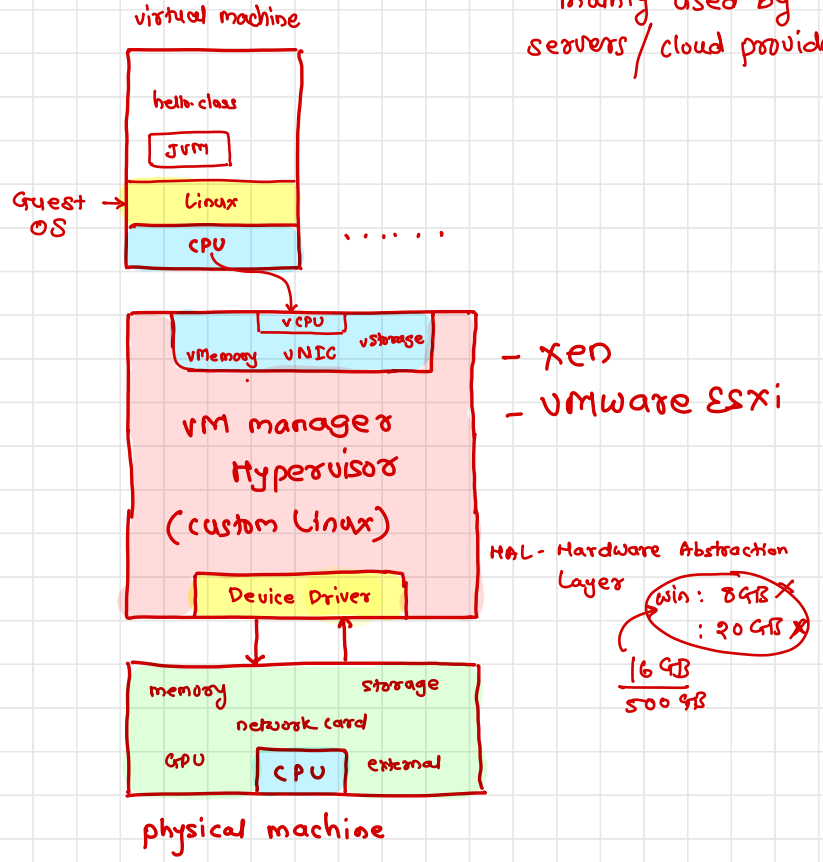
- With application virtualization, an app runs separately from the device that accesses it
- Application virtualization makes it possible for IT admins to install, patch and update only one version of an app rather than performing the same management tasks multiple times

Hardware Virtualization [simulating Hardware]

- Hardware virtualization or platform virtualization refers to the creation of a virtual machine that acts like a real computer with an operating system
- The process of masking the hardware resources like
  - CPU
  - Storage
  - Memory
- For example, a computer that is running Microsoft Windows may host a virtual machine that looks like a computer with the Ubuntu Linux operating system; Ubuntu-based software can be run on the virtual machine
- The process of creating Machines

Type I

mainly used by servers / cloud providers



- xen
- VMware ESXi

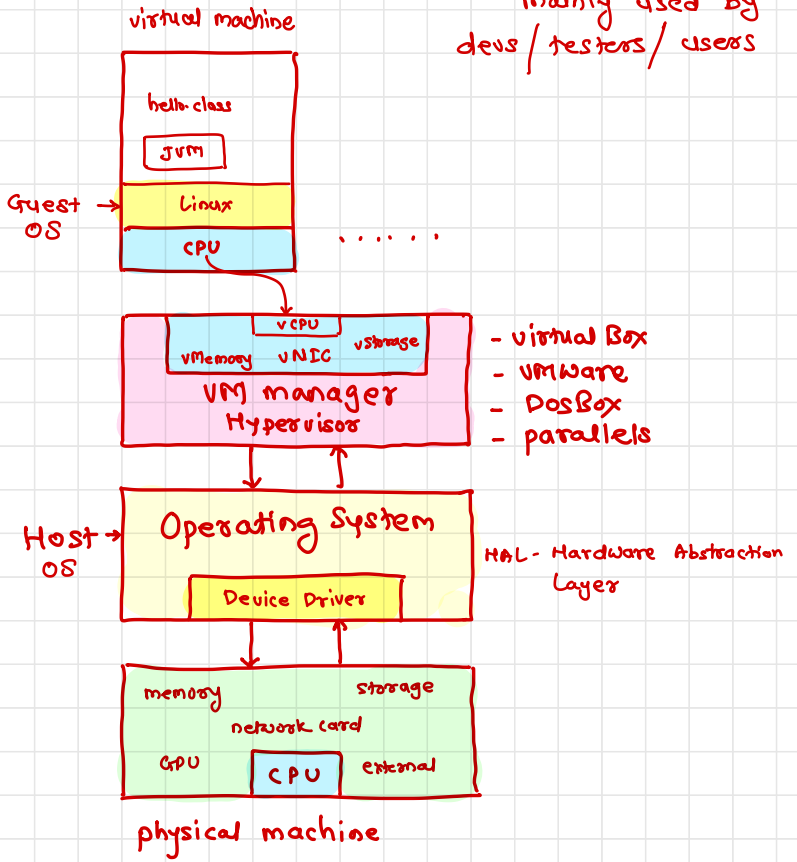
HAL - Hardware Abstraction Layer

win : 8GB ✗
: 20GB ✗

16 GB
500 GB

Type II

mainly used by devs / testers / users



- virtual Box
- VMware
- DosBox
- parallels

HAL - Hardware Abstraction Layer

Virtual Machine

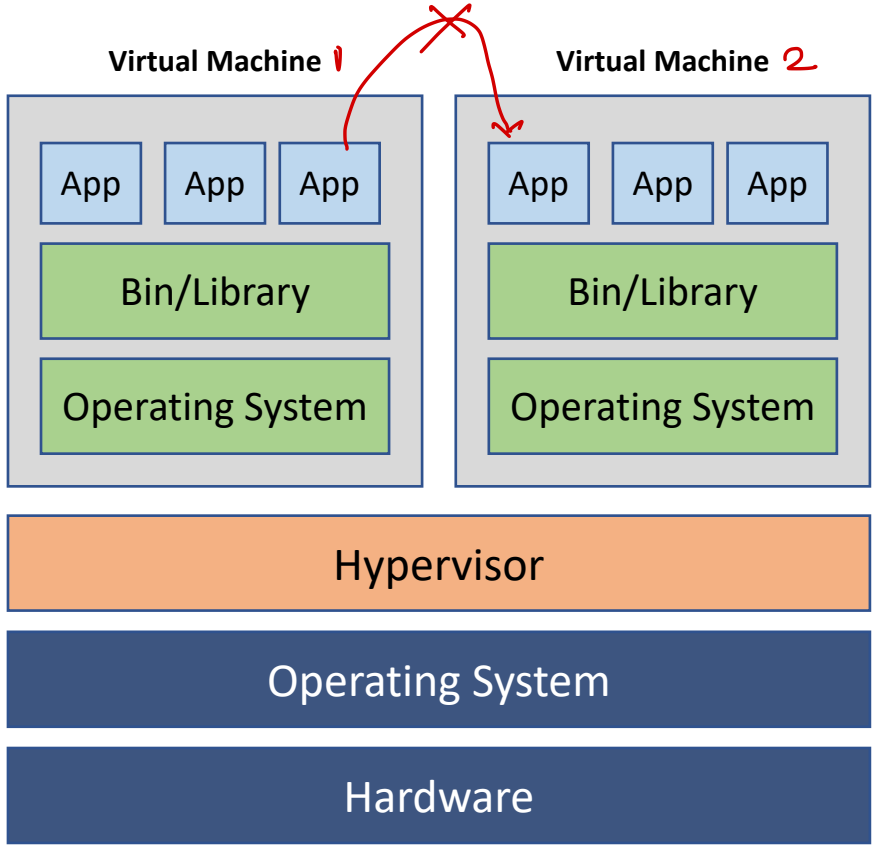
- A virtual machine is the emulated equivalent of a computer system that runs on top of another system
- Virtual machines may have access to any number of resources
  - Computing power - through hardware-assisted but limited access to the host machine's CPU
  - Memory - one or more physical or virtual disk devices for storage
  - A virtual or real network interfaces
  - Any devices such as
    - video cards,
    - USB devices,
    - other hardware that are shared with the virtual machine
- If the virtual machine is stored on a virtual disk, this is often referred to as a disk image

Types of hardware virtualization

- Type I
  - A Type 1 hypervisor runs directly on the host machine's physical hardware, and it's referred to as a bare-metal hypervisor
  - It doesn't have to load an underlying OS first
  - With direct access to the underlying hardware and no other software, it is more efficient and provides better performance
  - It is best suited for enterprise computing or data centers
  - E.g. VMware ESXi, Microsoft Hyper-V server and open source KVM
- Type II
  - A Type 2 hypervisor is typically installed on top of an existing OS, and it's called a hosted hypervisor
  - It relies on the host machine's pre-existing OS to manage calls to CPU, memory, storage and network resources
  - E.g. VMware Fusion, Oracle VM VirtualBox, Oracle VM Server for x86, Oracle Solaris Zones, Parallels and VMware Workstation

Virtualized Deployment

- It allows you to run multiple Virtual Machines (VMs) on a single physical server's CPU
- Virtualization allows applications to be isolated between VMs and provides a level of security as the information of one application cannot be freely accessed by another application
- Virtualization allows better utilization of resources in a physical server and allows better scalability because
  - an application can be added or updated easily
  - reduces hardware costs
- With virtualization you can present a set of physical resources as a cluster of disposable virtual machines
- Each VM is a full machine running all the components, including its own operating system, on top of the virtualized hardware



Advantages of virtualization

- **Lower costs**
  - Virtualization reduces the amount of hardware servers necessary within a company and data center
  - This lowers the overall cost of buying and maintaining large amounts of hardware
- **Easier disaster recovery**
  - Disaster recovery is very simple in a virtualized environment
  - Regular snapshots provide up-to-date data, allowing virtual machines to be feasibly backed up and recovered
  - Even in an emergency, a virtual machine can be migrated to a new location within minutes
- **Easier testing**
  - Testing is less complicated in a virtual environment
  - Even if a large mistake is made, the test does not need to stop and go back to the beginning
  - It can simply return to the previous snapshot and proceed with the test.

Advantages of virtualization

- **Quicker backups**
  - Backups can be taken of both the virtual server and the virtual machine
  - Automatic snapshots are taken throughout the day to guarantee that all data is up-to-date
  - Furthermore, the virtual machines can be easily migrated between each other and efficiently redeployed
- **Improved productivity**
  - Fewer physical resources results in less time spent managing and maintaining the servers
  - Tasks that can take days or weeks in a physical environment can be done in minutes
  - This allows staff members to spend the majority of their time on more productive tasks, such as raising revenue and fostering business initiatives