Underlying Technologies for Cloud Computing

Underlying Technologies for Cloud Computing

- A set of primary technology components that collectively enable key features and characteristics associated with contemporary cloud computing:
  - Broadband Networks and Internet Architecture
  - Data Center Technology
  - Virtualization Technology
  - Web Technology
  - Multitenant Technology
  - Service Technology

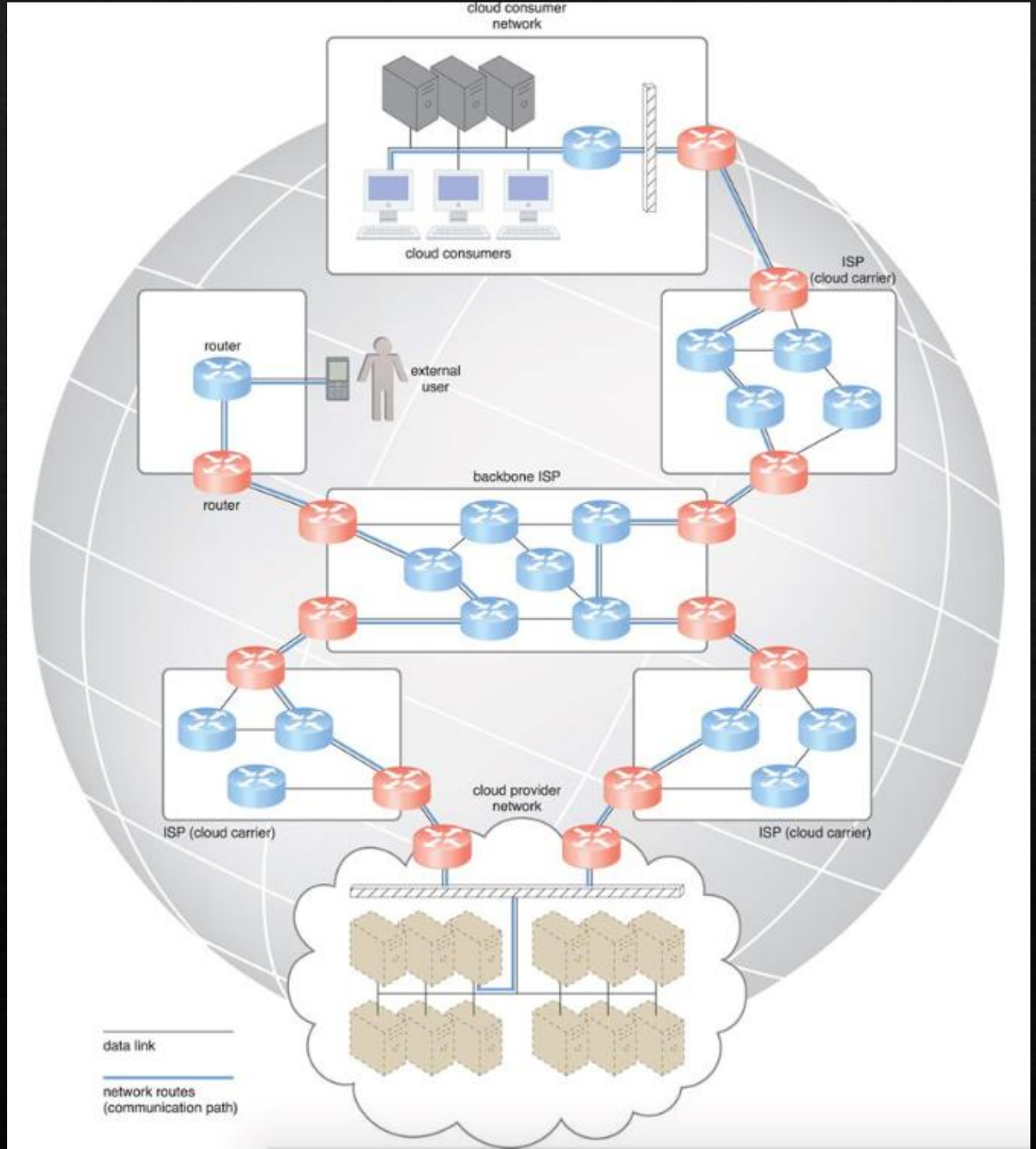
Why Broadband Networks and Internet?

- All clouds must be connected together somehow.
- Internetworks (or Internet) allow for the remote provisioning of IT resources.
- Internet support ubiquitous access.
- Architecture components?

Messages travel over dynamic network routes in this ISP internetworking configuration.

Internet Principles

- decentralized provisioning and management model.
- dynamic and aggregation of ISPs.
- No one govern the Internet.

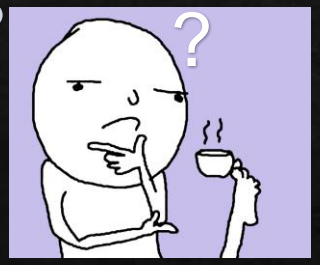


Connectionless Packet Switching (Datagram Networks)

- End-to-end (sender-receiver pair) data flows are divided into packets of a limited size.
- Received and processed through network switches and routers, then queued and forwarded from one intermediary node to the next.
- Important addresses for switching?

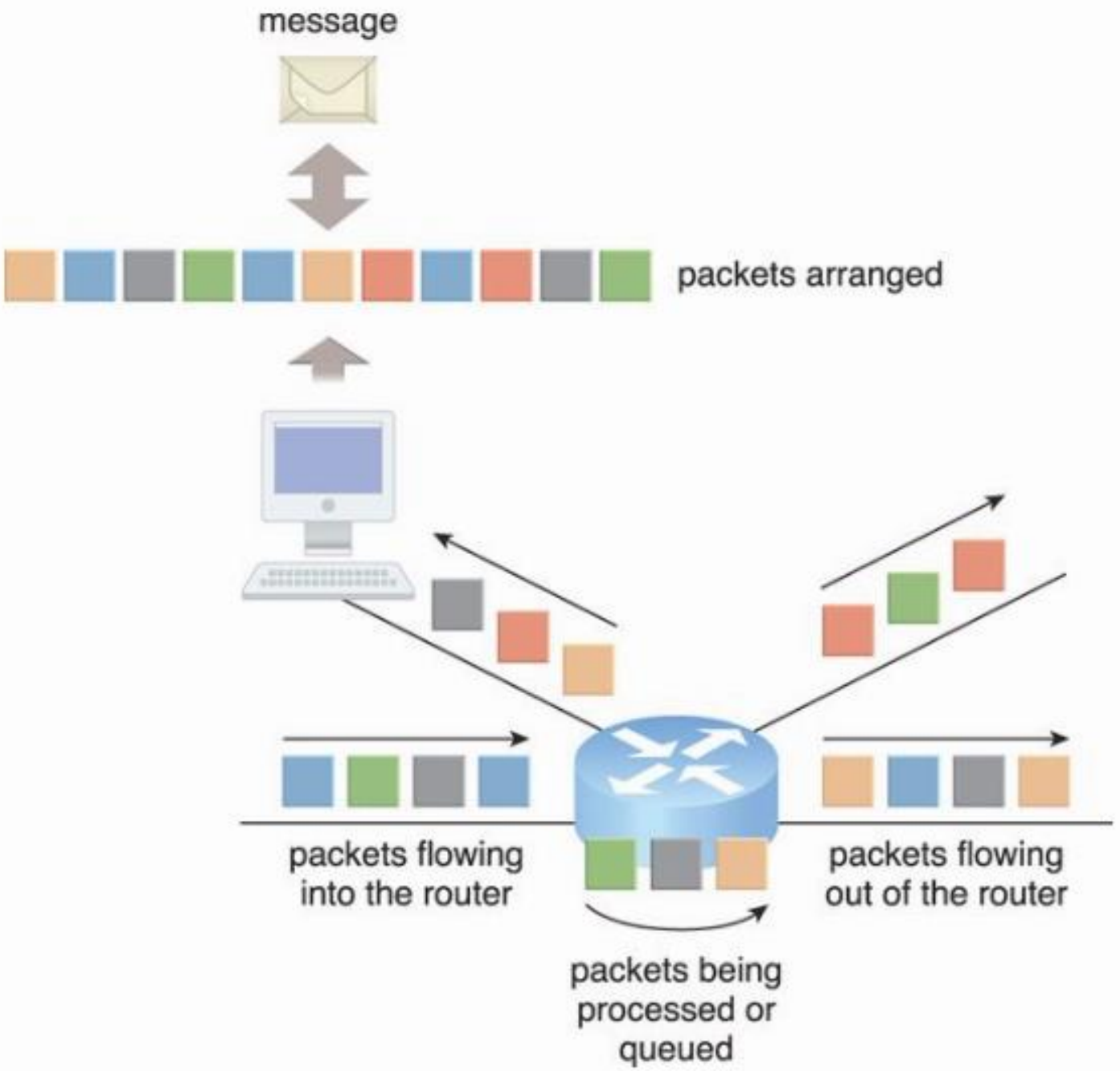
Router-based Interconnectivity

- Routers process and forward each packet individually while maintaining the network topology information.
- What is network topology information?



- Routers manage network traffic and gauge the most efficient hop (in terms of number of hops or link cost) for packet delivery.

The communication path that connects a cloud consumer with its cloud provider may involve multiple ISP networks.

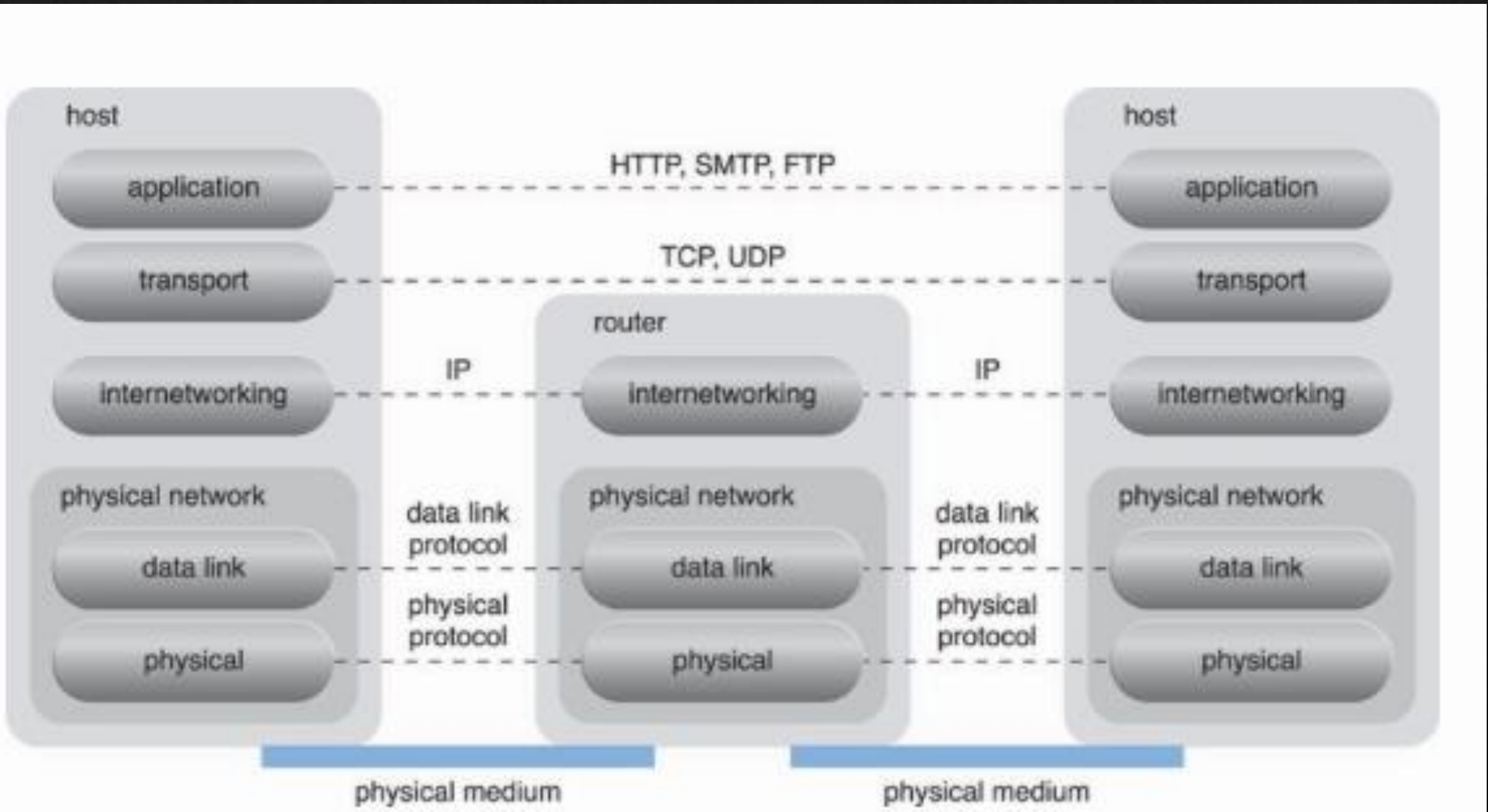


ISPs Internet's internetworking Layer

- Physical Layer
  - IP packets are transmitted through underlying physical networks that connect adjacent nodes, such as Ethernet, ATM network, and the 3G mobile HSDPA.
  - Data link layer that controls data transfer between neighboring nodes, and a physical layer that transmits data bits through both wired and wireless media.
- Transport Layer
  - Transport Control Protocol (TCP)
  - User Datagram Protocol (UDP)

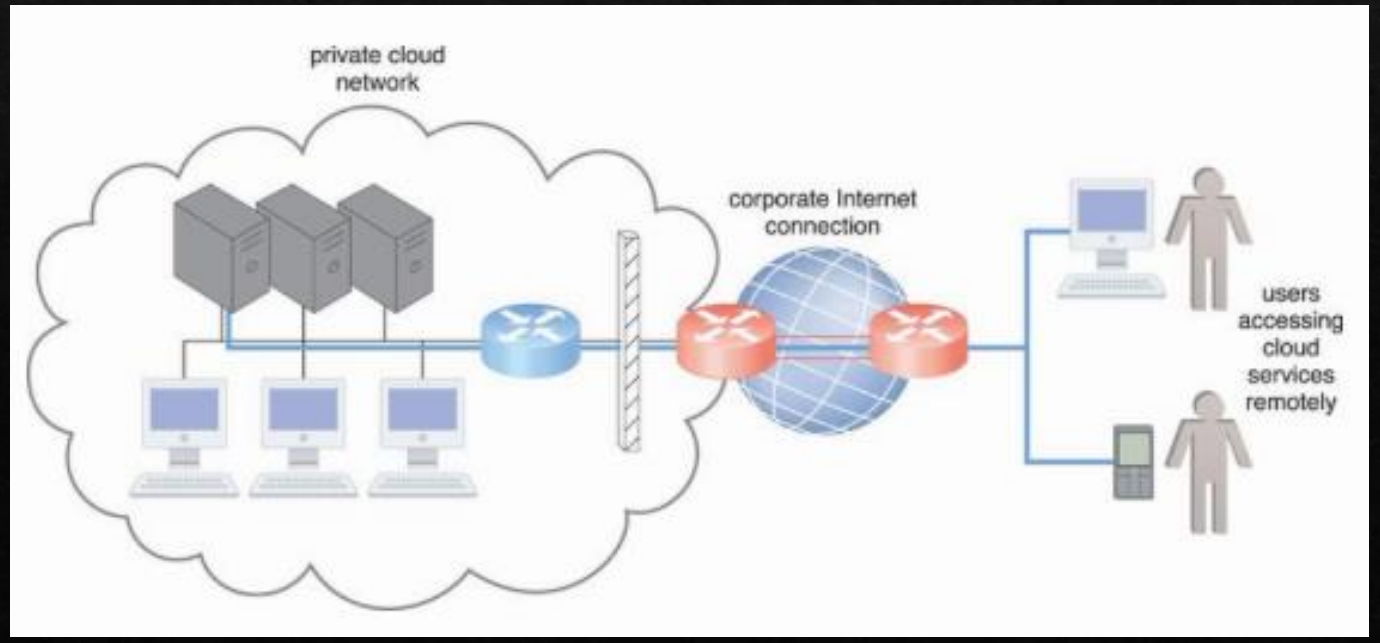
ISPs Internet's internetworking Layer (2)

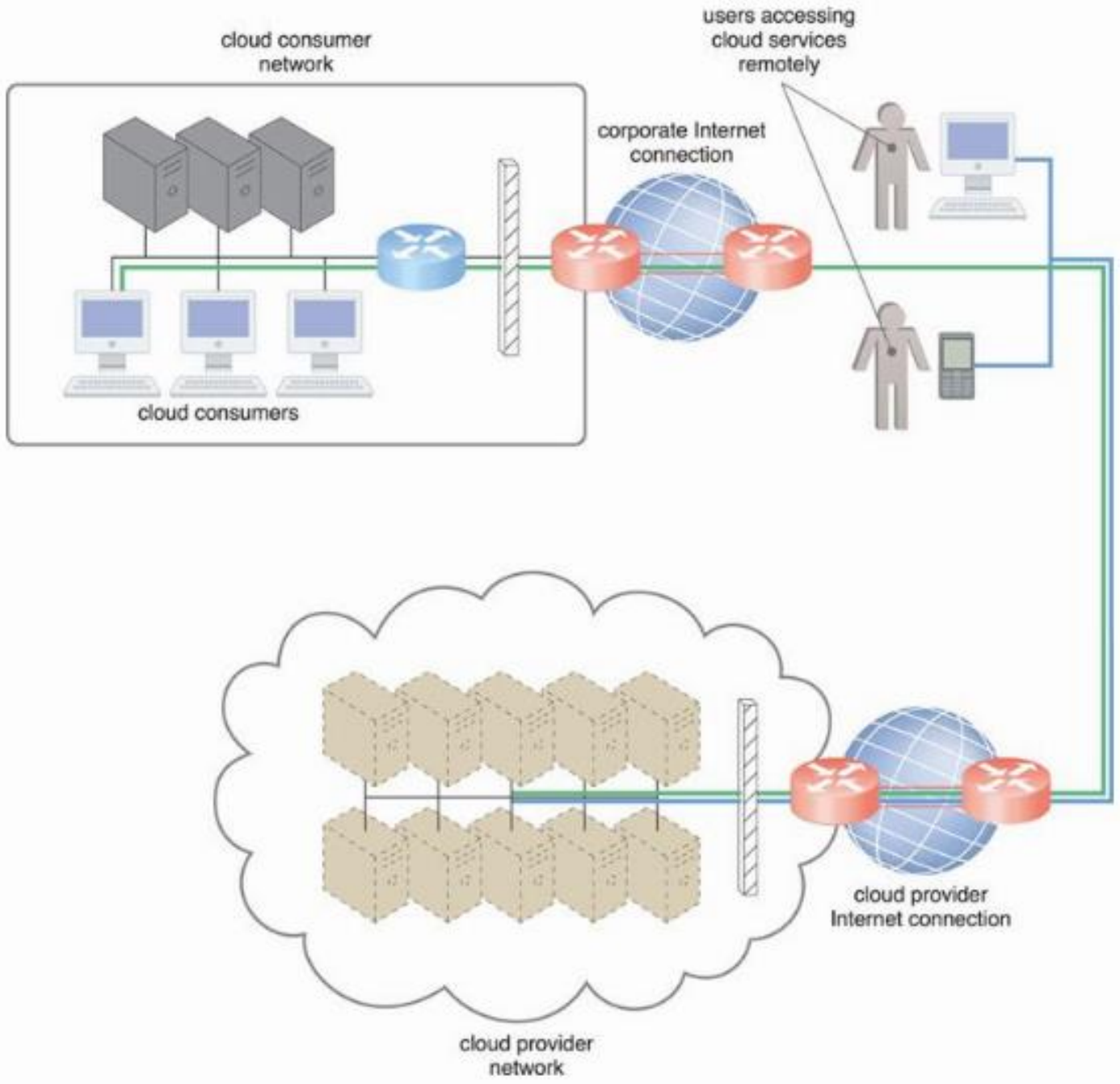
- Application Layer
  - Protocols such as HTTP, SMTP for e-mail, BitTorrent for P2P, and SIP for IP telephony use transport layer protocols to standardize and enable specific data packet transferring methods over the Internet.
  - Many other protocols also fulfill application-centric requirements and use either TCP/IP or UDP as their primary method of data transferring across the Internet and LANs.


host
application
transport
internetworking
physical network
data link
physical
HTTP, SMTP, FTP
TCP, UDP
IP
data link protocol
physical protocol
router
internetworking
physical network
data link
physical
IP
data link protocol
physical protocol
host
application
transport
internetworking
physical network
data link
physical
physical medium
physical medium

Technical Business Consideration

- Connectivity issues
  - On-premise services has no connectivity issues if accessed from corporate Internet networks.
  - Users can directly access the network traffic to and from the Internet.
  - Have complete control over and can safeguard their corporate networks using firewalls and monitoring software.

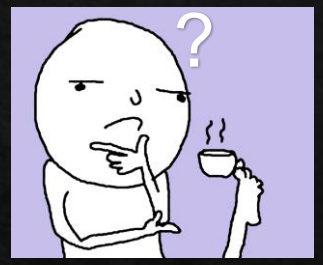




On-Premise IT Resources	Cloud-Based IT Resources
internal end-user devices access corporate IT services through the corporate network	internal end-user devices access corporate IT services through an Internet connection
internal users access corporate IT services through the corporate Internet connection while roaming in external networks	internal users access corporate IT services while roaming in external networks through the cloud provider's Internet connection
external users access corporate IT services through the corporate Internet connection	external users access corporate IT services through the cloud provider's Internet connection

Technical Business Consideration (2)

- Network bandwidth and latency issues
  - End-to-end bandwidth is determined by the transmission capacity of the shared data links.
  - Latency is the amount of time it takes a packet to travel from one data node to another.
  - Factors that increase latency?



  - Bandwidth is critical for applications that require substantial amounts of data to be transferred to and from the cloud.
  - Latency is critical for applications with a business requirement of swift response times.

Technical Business Consideration (3)

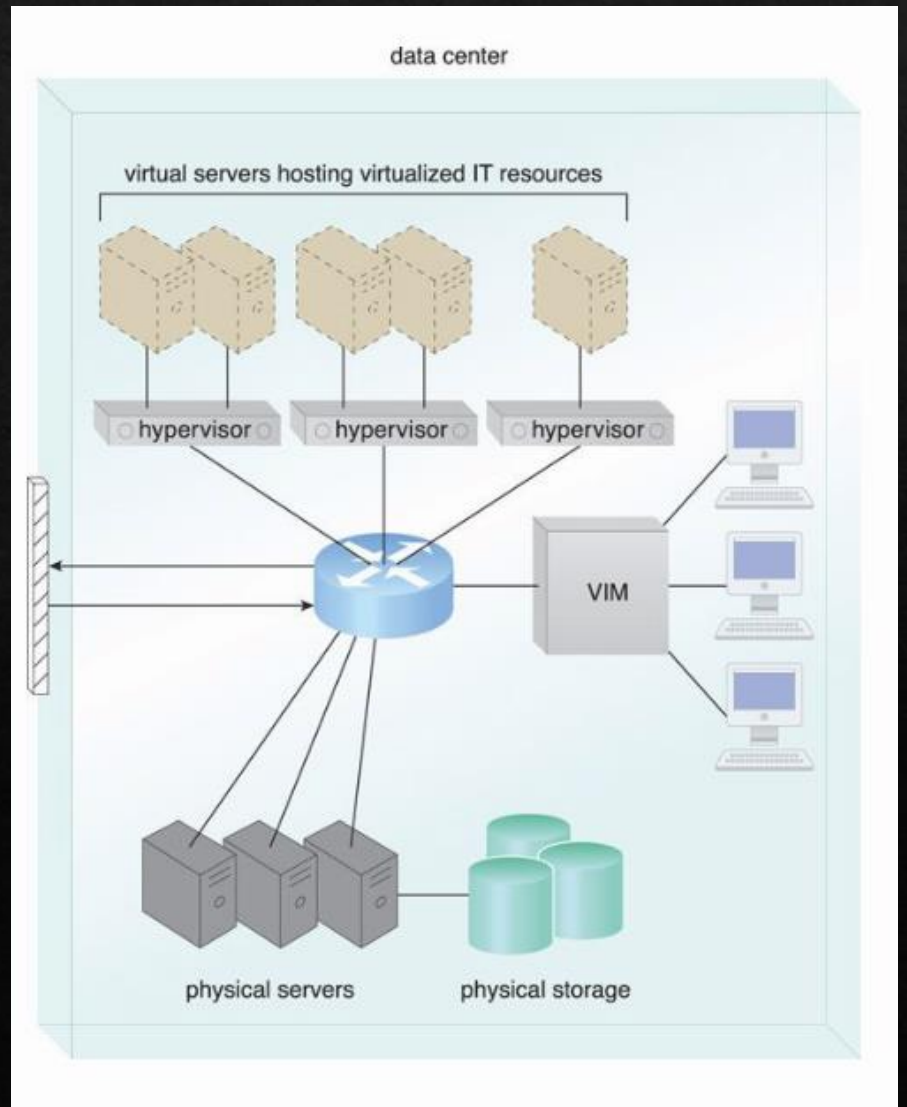
- Cloud carrier and cloud provider selection
  - The service levels of Internet connections between cloud consumers and cloud providers are determined by their ISPs (multiple ISPs involved)
  - QoS management across multiple ISPs is difficult to achieve in practice.
  - Multiple cloud carriers may be needed in order to achieve the necessary level of connectivity and reliability for their cloud applications (between cloud consumers and cloud providers), resulting in additional costs.
  - Cloud adoption can therefore be easier for applications with more relaxed latency and bandwidth requirements.

Data Center Technology

- A group of IT resourced located in close proximity offering:
  - power sharing,
  - higher efficiency in shared IT resource usage, and
  - improved accessibility for IT personnel.
- Modern data centers exist as specialized IT infrastructure used to house centralized IT resources, such as servers, databases, networking and telecommunication devices, and software systems.

Data Center Components

- Virtualization
  - A platform that abstracts the physical computing and networking IT resources as virtualized components that are easier to allocate, operate, release, monitor, and control.
- Standardized and modularity
  - Aggregation of multiple identical building blocks of facility infrastructure and equipment to support scalability, growth, and speedy hardware replacements.
- Automation
  - To automate tasks like provisioning, configuration, patching, and monitoring without supervision.
- Remote operation and Management
  - Remote control



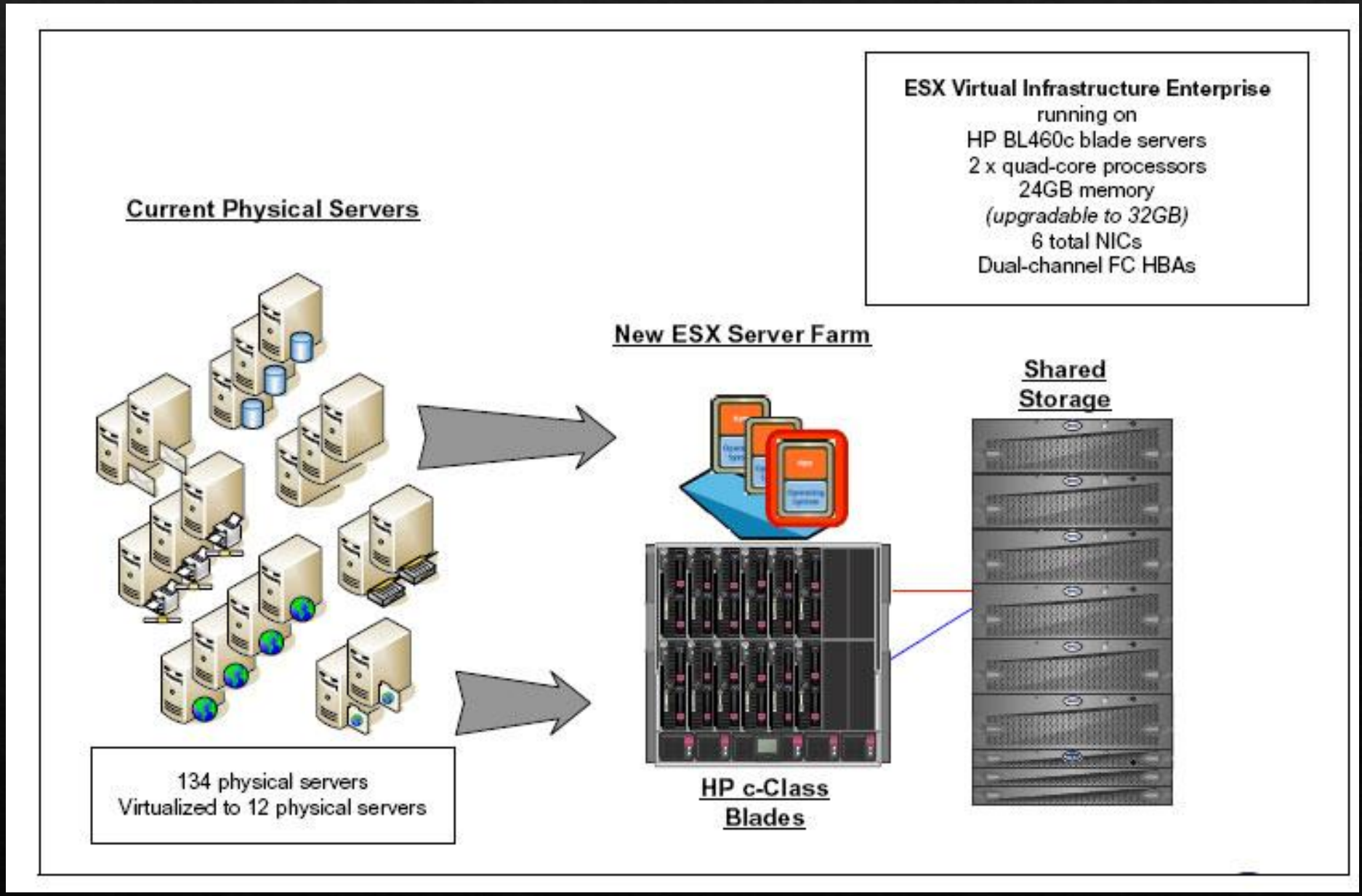
Data Center Components

- High availability
  - Designed to operate with increasingly higher levels of redundancy to sustain availability through redundant power supplies, UPSs, hardware, etc.
- Security-aware design, operation and management
  - Thorough and comprehensive physical and logical accesses for data centers
- Facility
  - Custom-designed locations that are outfitted with specialized computing, storage, and network equipment.

Data Center Components

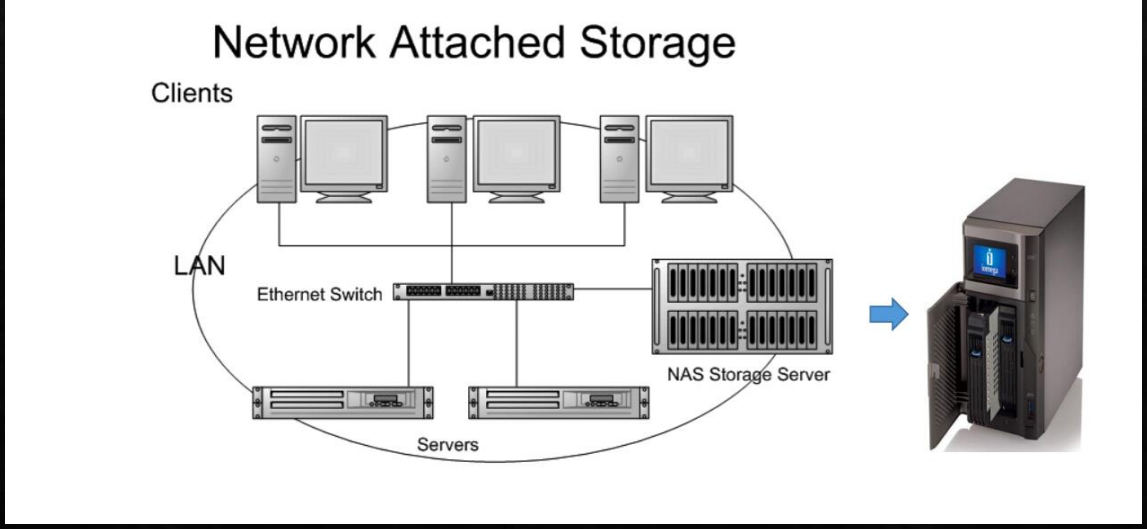
- Computing hardware
  - Rack-mounted, blade server, etc.
  - x86-32bits, x86-64bits, and RISC architecture (multiple cores).
  - Redundant and hot-swappable components, such as hard disks, power supplies, network interfaces, and storage controller cards.

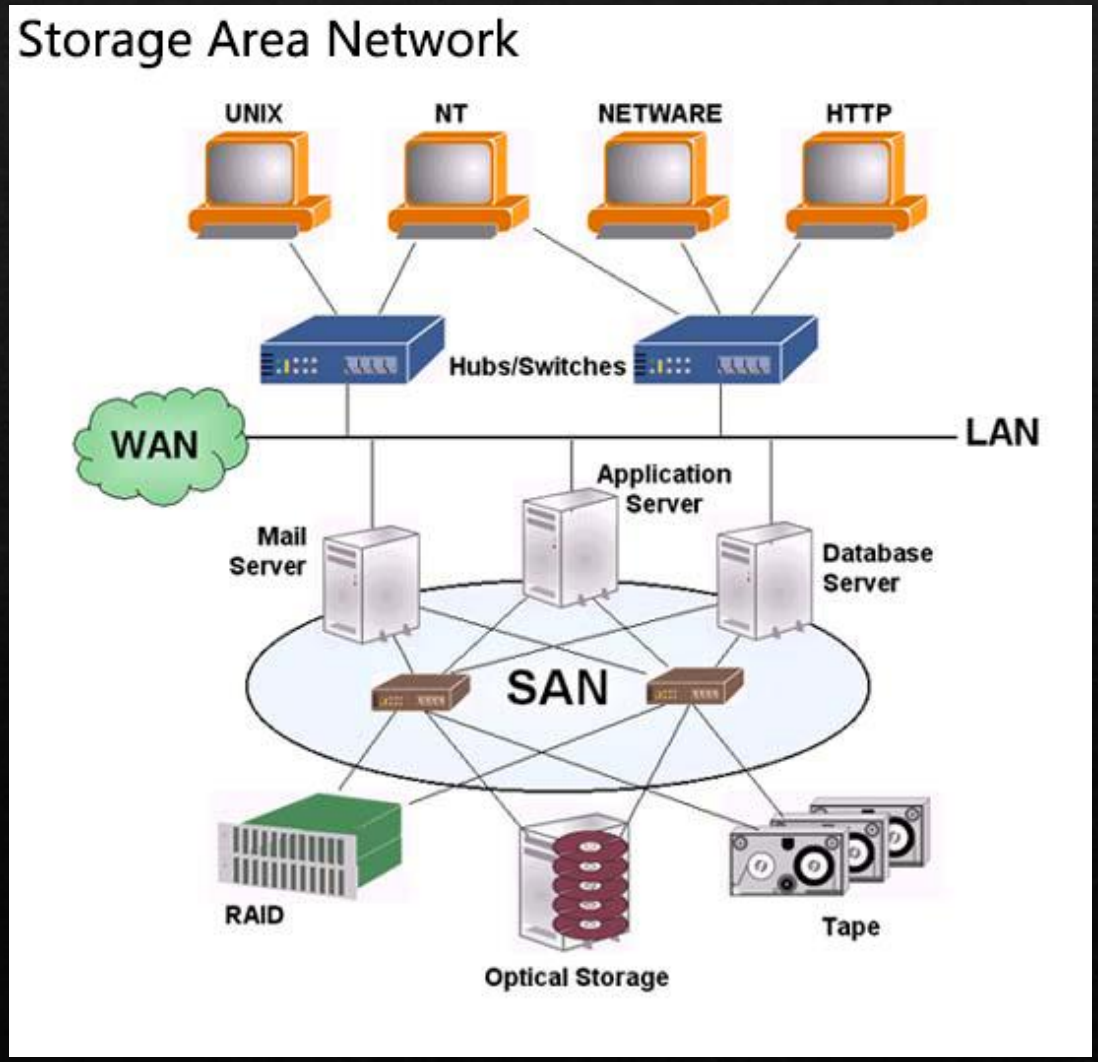
An example from HPE solution.



Data Center Components (4)

- Storage
  - NAS, SAN, Cloud-based storage, etc.
  - Underlying technologies
    - Disk arrays, I/O caching, hot-swappable HDD, storage virtualization, fast data replication technology





Data Center Components (5)

- Network hardware
  - Carrier and external networks interconnection
  - Web-tier load balancing and acceleration
    - Examples include XML pre-processors, encryption/decryption appliances, and layer 7 switching devices that perform content-aware routing.
  - LAN fabric (LAN networks)
    - Gigabit Ethernet.
  - SAN fabric
    - Fiber optic connection. Fibre Channel (FC), Fibre Channel over Ethernet (FCoE), etc.

Virtualization Technology

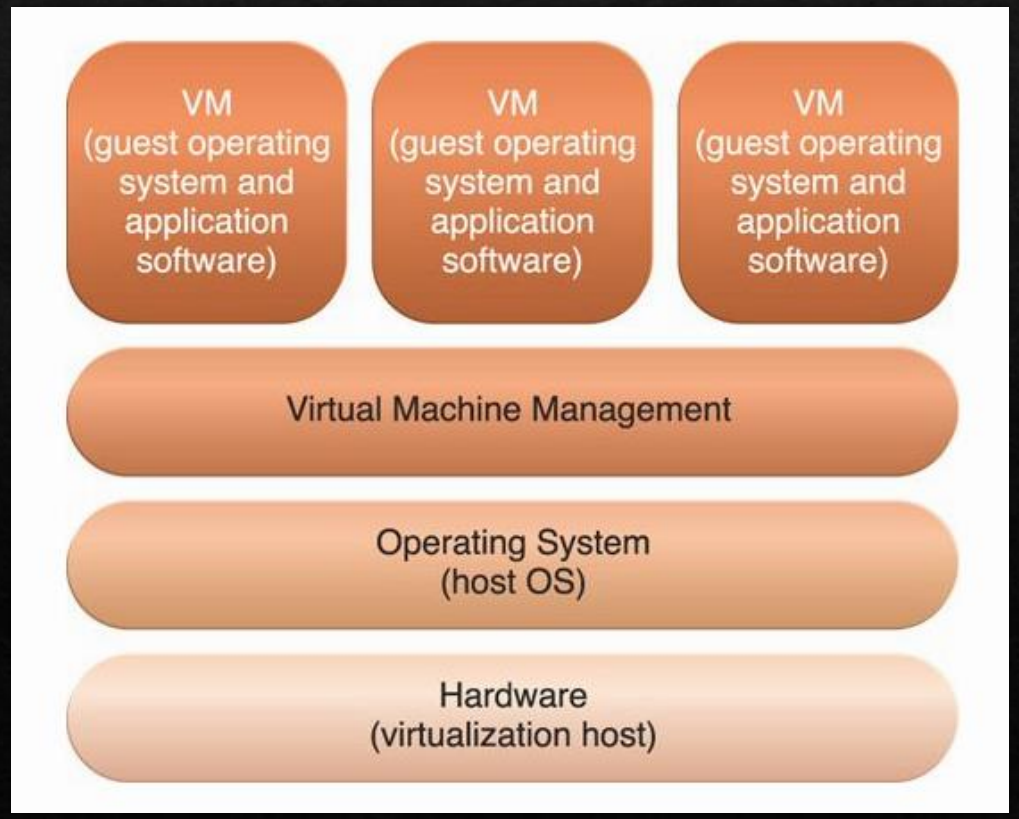
- A process of converting a physical IT resource into a virtual IT resource. Most types of IT resources can be virtualized, including:
  - Servers – A physical server can be abstracted into a virtual server.
  - Storage – A physical storage device can be abstracted into a virtual storage device or a virtual disk.
  - Network – Physical routers and switches can be abstracted into logical network fabrics, such as VLANs.
  - Power – A physical UPS and power distribution units can be abstracted into what are commonly referred to as virtual UPSs.
- Steps?
  - Allocation of physical IT resources,
  - Installation of host OS,
  - Installation of guest OS.
- Virtualization software is sometimes referred to as a virtual machine manager or a virtual machine monitor (VMM), but most commonly known as a hypervisor.

Its Characteristics

- Hardware Independency
  - A conversion process that translates unique IT hardware into emulated and standardized software-based copies.
- Server consolidation
  - Multiple virtual servers to be simultaneously created in the same virtualization host.
- Resource replication
  - Virtual server as virtual disk image.
  - Simple file operations, such as copy, move, and paste, can be used to replicate, migrate, and back up the virtual server.
  - Agility, rapidly scale out and up.
  - Ability to roll back.
  - Support business continuity with efficient backup (e.g., disk snapshot)

Operating System-based Virtualization

Operating system-based virtualization can rectify hardware compatibility issues even if the hardware driver is not available to the virtualization software.



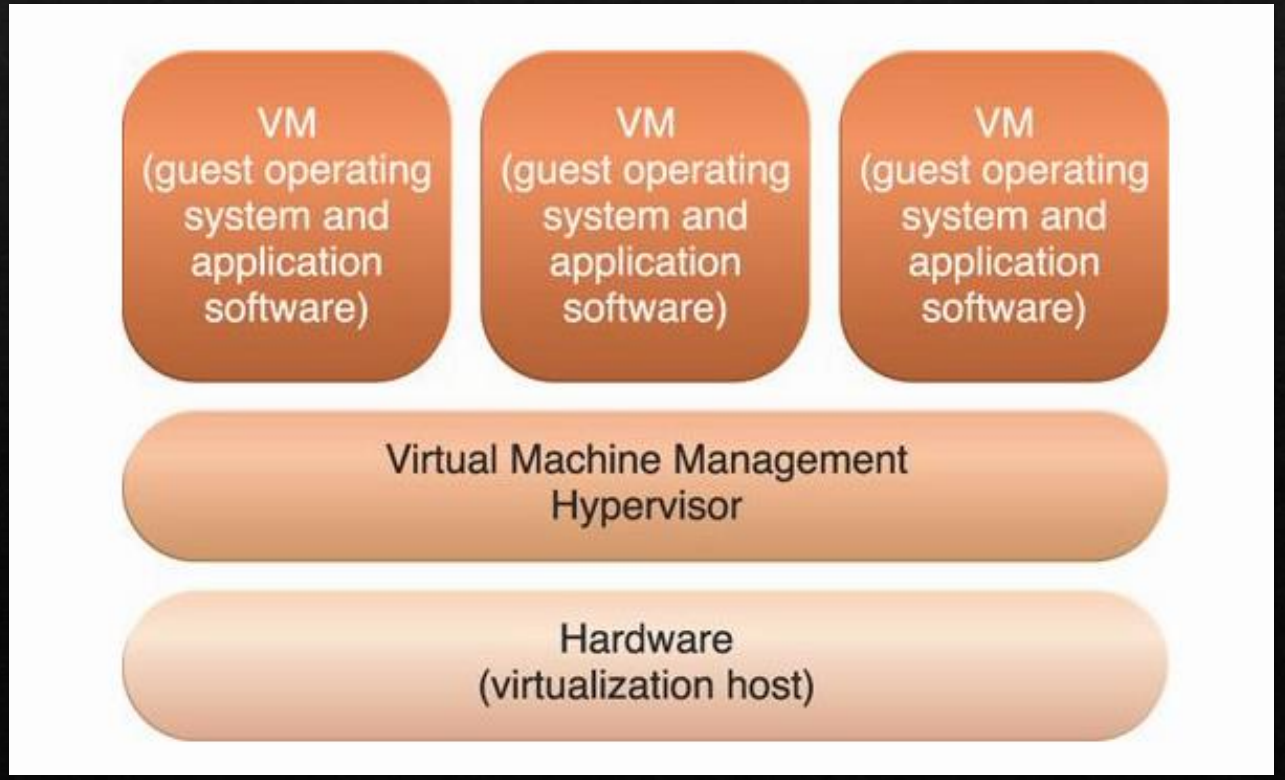
Operating System-based Virtualization (2)

- Operating system-based virtualization can introduce demands and issues related to performance overhead such as:
  - The host operating system consumes CPU, memory, and other hardware IT resources.
  - Hardware-related calls from guest operating systems need to traverse several layers to and from the hardware.
  - Licenses are usually required for host operating systems, in addition to individual licenses for each of their guest operating systems.

Hardware-based Virtualization

- An installation of virtualization software directly on the physical host hardware.

- Virtual servers to interact with hardware without requiring intermediary action from the host operating system, resulting in better efficiency.
- A hypervisor has a simple user-interface that requires a negligible amount of storage space, existing as a thin layer of software that handles hardware management functions to establish a virtualization management layer.



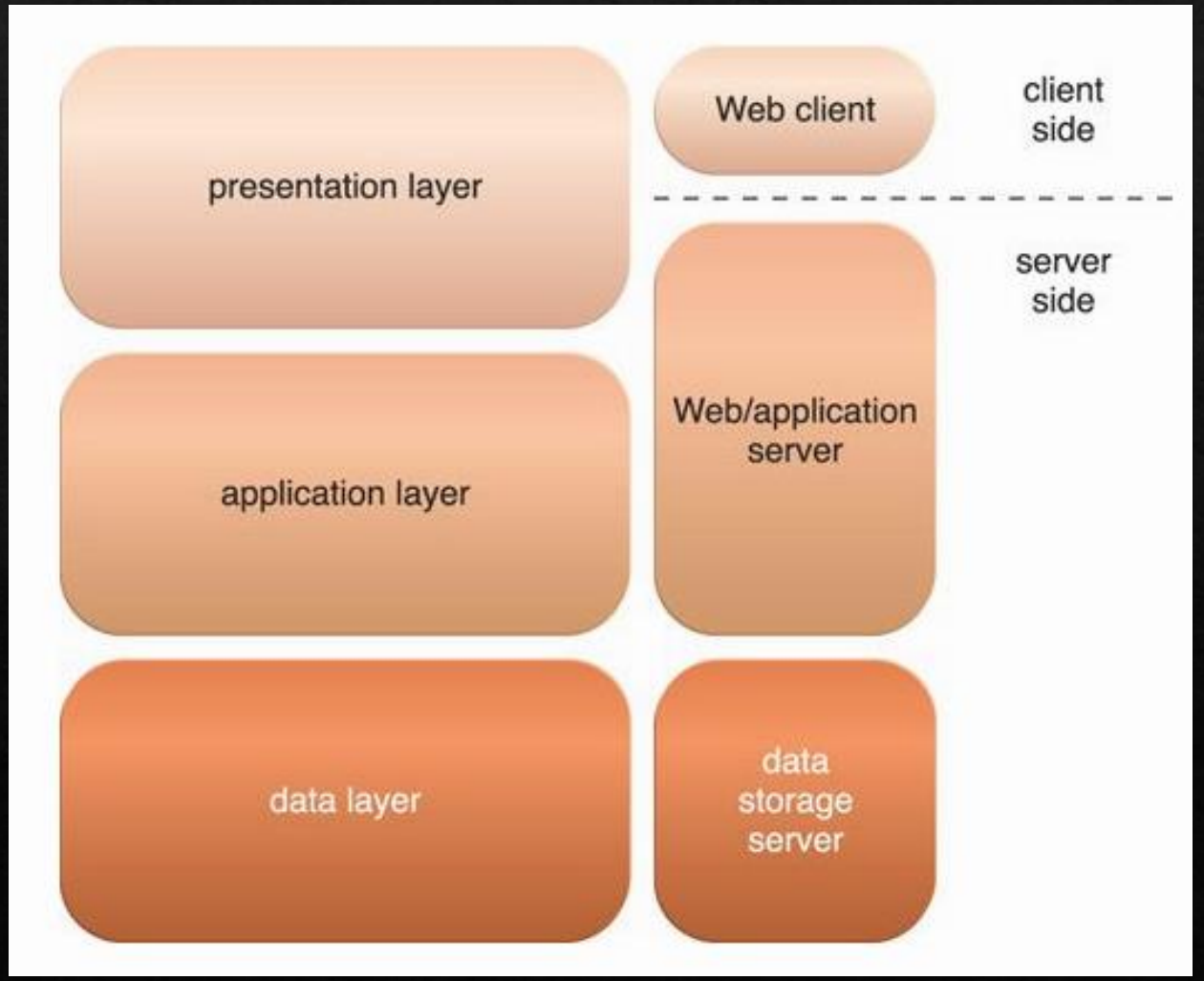
Considerations in Virtualization Technology

- **Performance Overhead** – Virtualization may not be ideal for complex systems that have high workloads with little use for resource sharing and replication.
- **Special Hardware Compatibility** – Many hardware vendors that distribute specialized hardware may not have device driver versions that are compatible with virtualization software.
- **Portability** – The programmatic and management interfaces that establish administration environments for a virtualization program to operate with various virtualization solutions can introduce portability gaps due to incompatibilities.

Web Technology Terminology

- Web resources vs. IT resources
  - Web resources – artifacts accessible via WWW.
  - IT resources - a physical or virtual IT-related artifact that can be software or hardware-based.
- Basic web technologies
  - URL, HTTP, XML
- Web application
  - Distributed applications that use web-based technologies.

Common Architectural Abstraction of Web Application



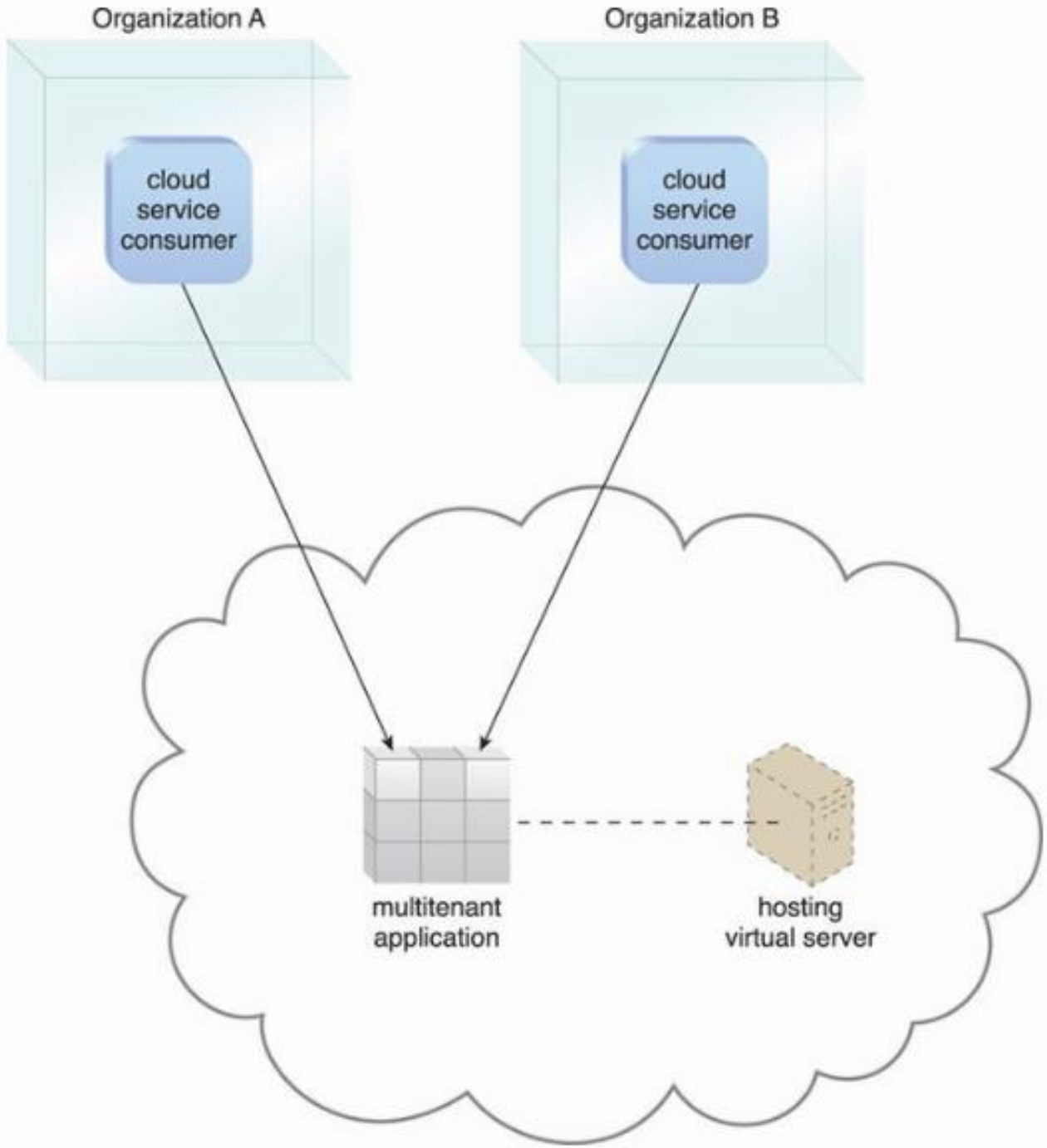
Multitenant Technology

- A technology that enables multiple users (tenants) to access the same application logic simultaneously (a dedicated instance of software).
- Multitenant applications ensure that tenants do not have access to data and configuration information that is not their own.
- Tenants can individually customize features of the application, such as:
  - Interfaces (look and feel)
  - Business process – customize rules, logics and workflows.
  - Data model – freely include, exclude, rename fields.
  - Access control

Common Characteristics of Multitenant Applications

- **Usage Isolation** – The usage behavior of one tenant does not affect the application availability and performance of other tenants.
- **Data Security** – Tenants cannot access data that belongs to other tenants.
- **Recovery** – Backup and restore procedures are separately executed for the data of each tenant.
- **Application Upgrades** – Tenants are not negatively affected by the synchronous upgrading of shared software artifacts.
- **Scalability** – The application can scale to accommodate increases in usage by existing tenants and/or increases in the number of tenants.
- **Metered Usage** – Tenants are charged only for the application processing and features that are actually consumed.
- **Data Tier Isolation** – Tenants can have individual databases, tables, and/or schemas isolated from other tenants.

A multitenant application that is being concurrently used by two different tenants with a typical SaaS implementation.



Multitenancy vs. Virtualization

- Multitenancy is sometimes mistaken for virtualization because the concept of multiple tenants is similar to the concept of virtualized instances.
- The differences lie in what is multiplied within a physical server acting as a host:
  - **With virtualization**: Multiple virtual copies of the server environment can be hosted by a single physical server. Each copy can be provided to different users, can be configured independently, and can contain its own operating systems and applications.
  - **With multitenancy**: A physical or virtual server hosting an application is designed to allow usage by multiple different users. Each user feels as though they have exclusive usage of the application.