Networking Support

CHAPTER 06

Networking support

- A cloud is built around a high-performance interconnected networks.
- The servers in a cloud communicate through high-bandwidth and low-latency specialized networks. It is thus obvious why networking will continue to play a crucial role in the evolution of cloud computing and a content-centric.
- Cloud computing and delivery of content stored on a cloud are feasible only due Internet (packet-switched network) and World Wide Web (Remote access).

Packet-switched networks

- A packet-switched network transports data units called packets through a maze of switches where packets are queued and routed towards their destination.
- A packet-switched network consists of:
  - Network core made up from routers and control systems interconnected by very high bandwidth communication channels.
  - Network edge where the end-user systems/hosts reside.
- Packet → consists of a header which contains control information necessary for its transport through the network and a payload or data.

Packet-switched networks

- Packets are subject to a variable delay, errors, and loss.
- A network architecture describes the protocol stack.
- Protocol → a discipline for communication, it specifies the actions taken by the sender and the receiver of a data unit.
- Host → a system located at the network edge capable to initiate and to receive communication, e.g., computer, mobile device, sensor.

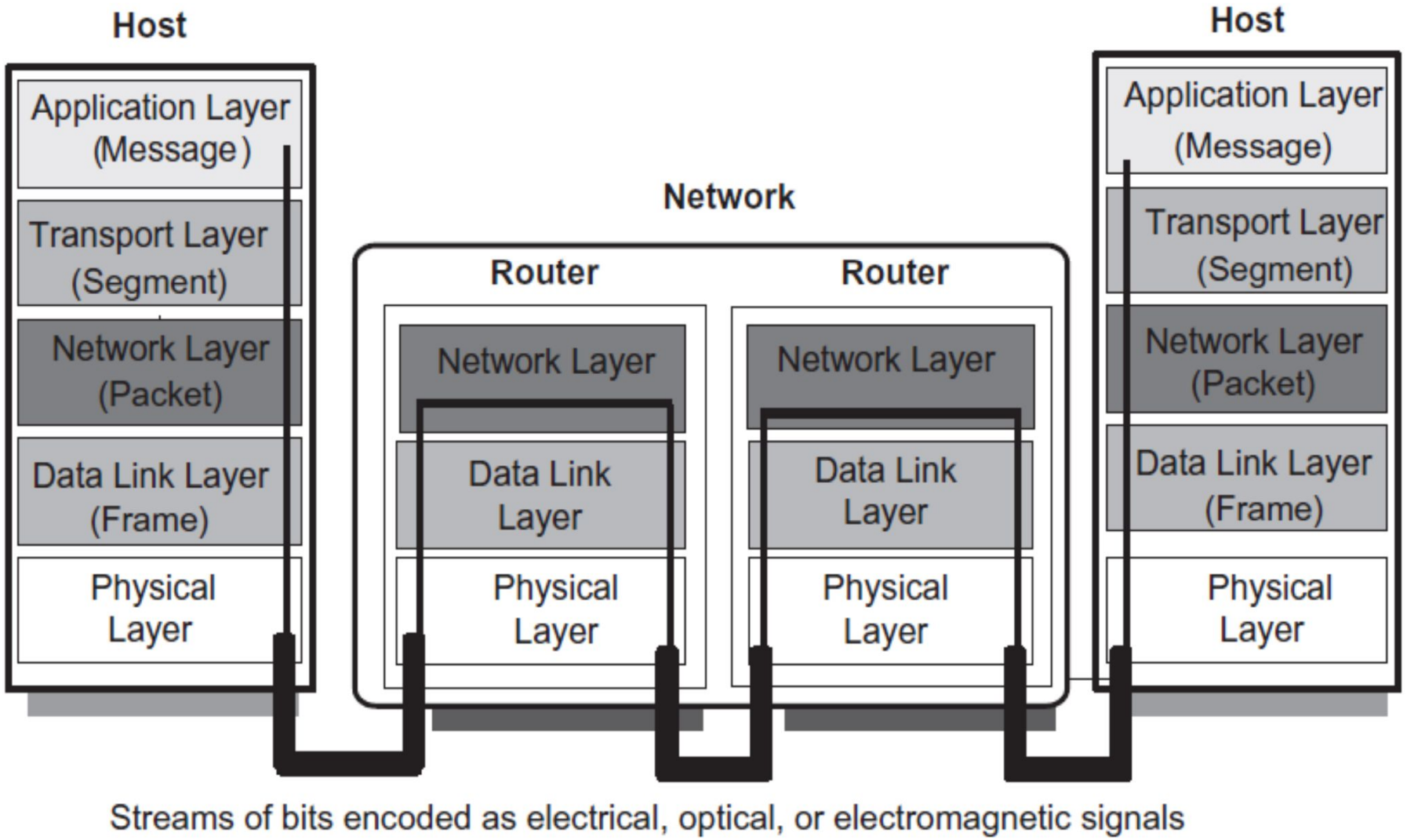
The Internet

- Collection of separate and distinct networks.
- All networks operate under a common framework consisting of:
  - globally unique IP addressing.
  - IP routing.
  - global Border Gateway Routing (BGP) protocols.
- IP only provides best effort delivery - any router on the path from the source to the destination may drop a packet if it is overloaded.

The Internet

- The Internet uses two transport protocols
  - UDP (User Datagram Protocol) - a connectionless datagram protocol. The UDP transport protocol assumes that error checking and error correction are either not necessary or performed by the application. Datagrams may arrive out of order, duplicated, or may not arrive at all.
  - TCP (Transport Control Protocol) - a connection-oriented protocol. TCP provides reliable, ordered delivery of a stream of bytes from an application on one system to its peer on the destination system.

Protocol stack



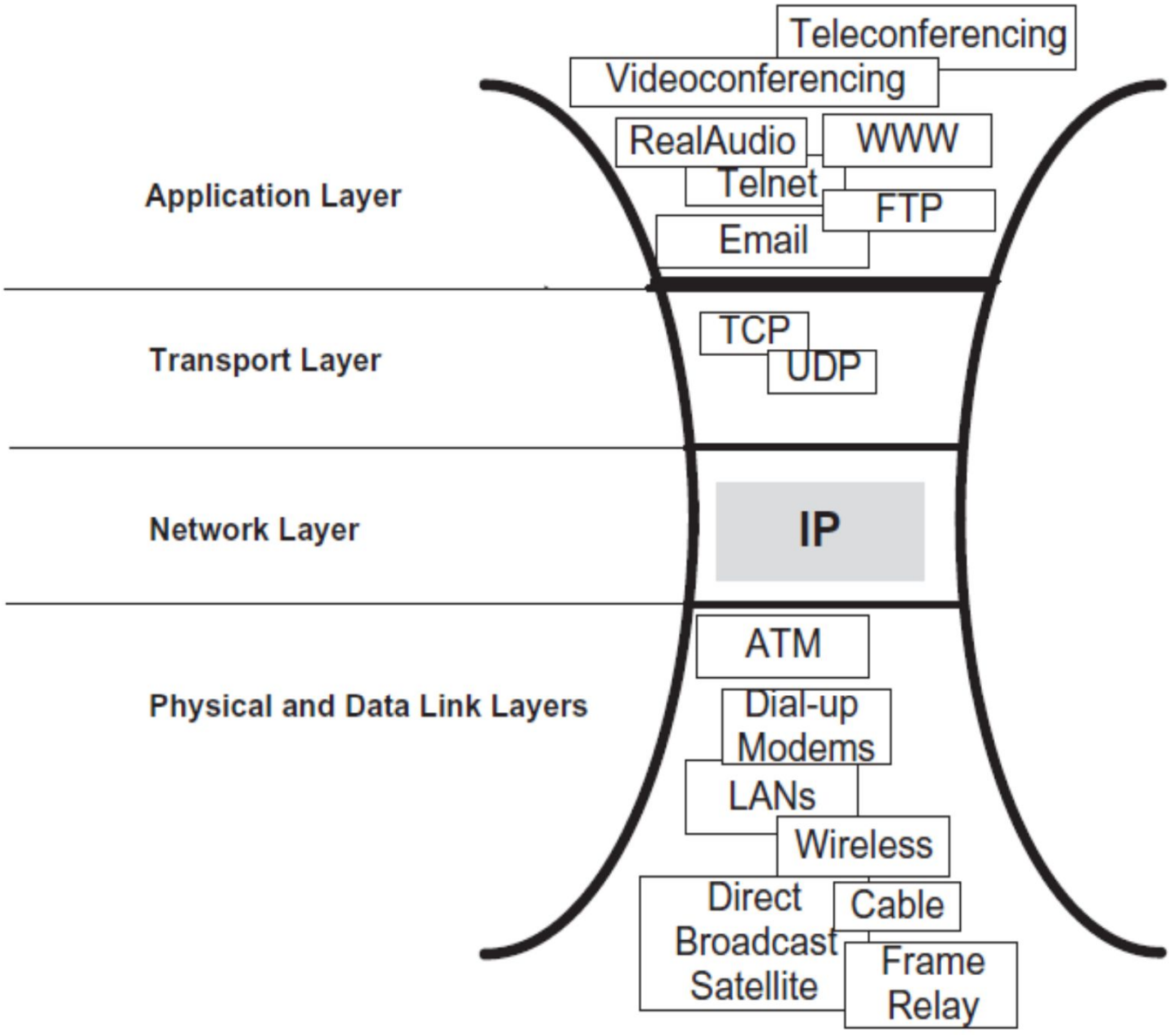
Protocol stack

- Applications running on hosts at the edge of the network communicate using application layer protocols.
- The transport layer deals with end-to-end delivery.
- The network layer is responsible for routing a packet through the network.

Protocol stack

- The data link layer ensures reliable communication between adjacent nodes of the network, and the physical layer transports streams of bits encoded as electrical, optical, or electromagnetic signals (the thick lines represent such bit pipes).
- The corresponding data units for the five-layer architecture are messages, segments, packets, frames, and encoded bits, respectively.

The hourglass architecture of the Internet



Note

- Regardless of the application, the transport protocol, and the physical network, all packets are routed through the Internet from the source to the destination by the IP protocol, and routing is based on the destination IP address.

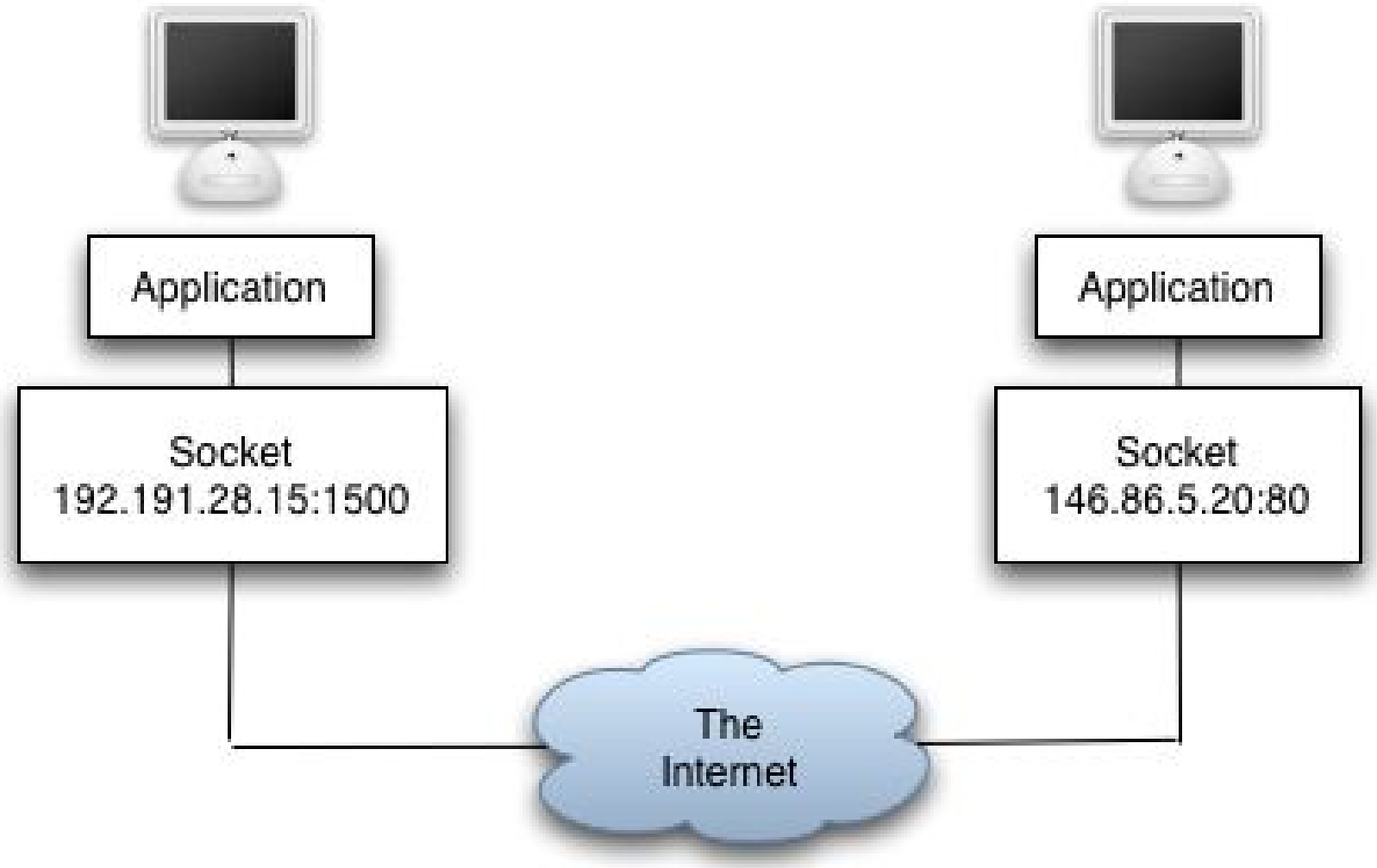
IP and MAC addresses, ports and sockets

- IP address → logical address assigned dynamically by a DHCP server. A host may have multiple IP addresses as it may be connected to more than one network.
- MAC address → unique physical address of each network interface.
- Network interface → hardware connecting a host with a network.

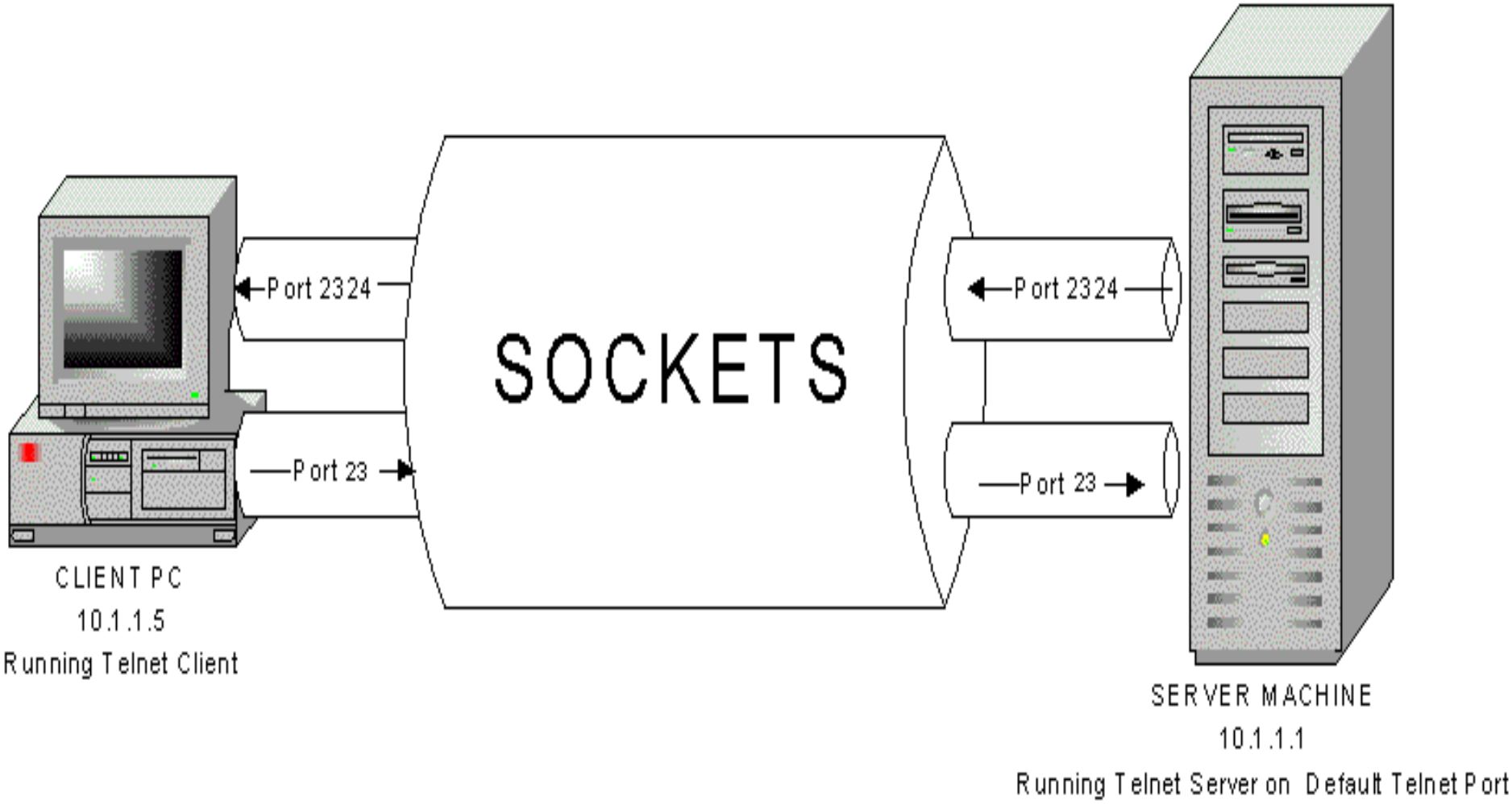
IP and MAC addresses, ports and sockets

- Port → software abstraction for message delivery to an application.
- Sockets → software abstraction allowing an application to send and receive messages at a given port; implemented as two queues, one for incoming and the other for outgoing messages.
- A **socket** is a software object that acts as an end point establishing a bidirectional network communication link between a server-side and a client-side program.

Sockets


Application
Application
Socket
192.191.28.15:1500
Socket
146.86.5.20:80
The
Internet

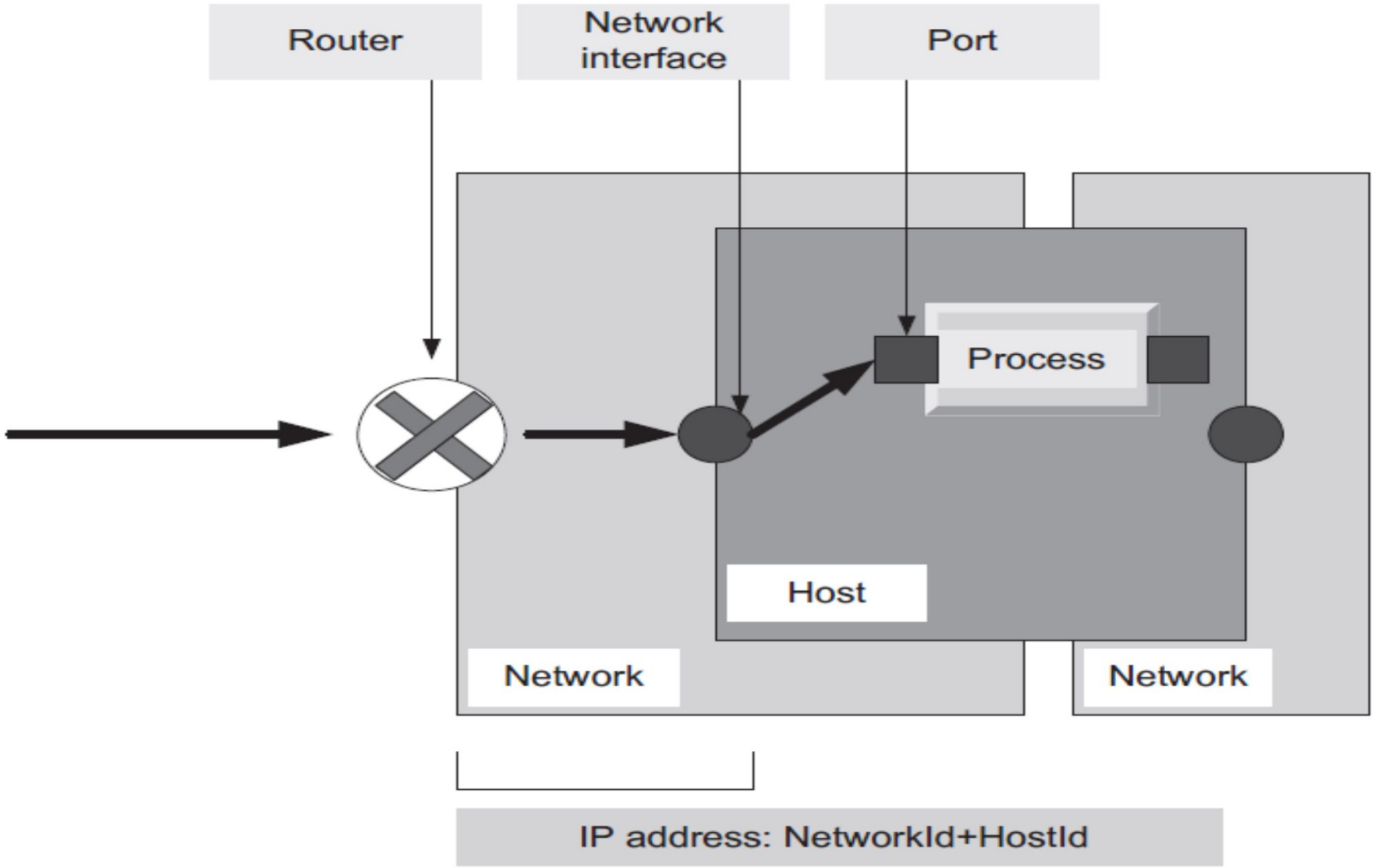
Sockets


Port 2324
Port 23
SOCKETS
Port 2324
Port 23
CLIENT PC
10.1.1.5
Running Telnet Client
SERVER MACHINE
10.1.1.1
Running Telnet Server on Default Telnet Port

Packet delivery to processes and threads

- The packet is first routed by the IP protocol to the destination network and then to the host specified by the IP address.
- Each application listens to an abstraction of the endpoint of a logical communication channel called a port.
- The processes or threads running an application use an abstraction called *socket* to send and receive data through the network

Packet delivery to processes and threads



Internet migration to IPv6

- The Internet addressing capabilities and the migration to IPv6 are important for cloud computing.
- IPv4 has an addressing capability of 2^{32}, or approximately 4.3 billion addresses, a number that proved to be insufficient.
- IPv6 has an addressing capability of 2^{128}, or $3.4x10^{38}$ addresses.
- Unfortunately, migration to IPv6 is a very challenging and costly proposal.

Other major differences between IPv4 and IPv6

- IPv6 supports new multicast solutions and but not traditional IP broadcast.
- IPv6 hosts can configure themselves automatically when connected to a routed IPv6 network using the Internet Control Message Protocol version 6 (ICMPv6).
- Mandatory support for network security. Internet Network Security(IPsec) is an integral part of the base protocol suite in IPv6.

The transformation of the Internet

- The Internet is continually evolving under the pressure of its own success and the need to accommodate new applications and a larger number of users.
- Initially conceived as a data network, a network supporting only the transport of data files, it has morphed into today's network for data with real-time constraints for multimedia applications.

The transformation of the Internet

- New technologies such as Web applications, cloud computing, and content-delivery networks are reshaping the definition of a network.
- The Web, gaming, and entertainment are merging, and more computer applications are moving to the cloud.
- Data streaming consumes an increasingly large fraction of the available bandwidth as high-definition TV (HDTV) sets become less expensive and content providers such as Netflix and Hulu offer customers services that require a significant increase in network bandwidth.

The transformation of the Internet

- The “last mile” - the link connecting the home to the Internet Service Provider (ISP) network is the bottleneck.
- Google has initiated the Google Fiber Project which aims to provide 1Gb/s access speed to individual households through FTTH.
- Fiber-to-the-home (FTTH) is a broadband network architecture that uses optical fiber to replace the copper-based local loop used for the last mile of network access to the home.

The relations between Internet networks

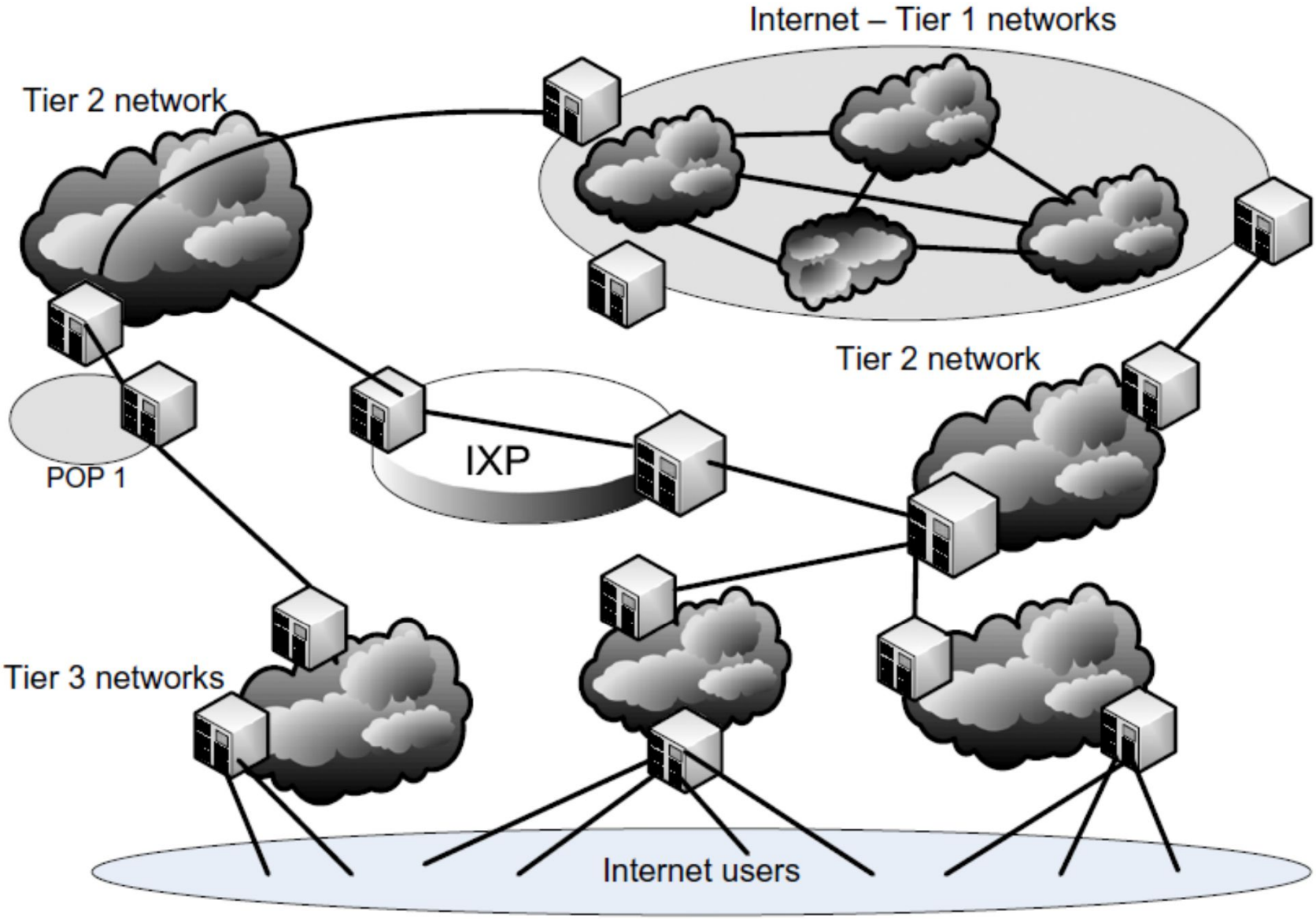
- Three type of relations:
  - Peering - two networks exchange traffic between each other's customers freely.
  - Transit - a network pays to another one to access the Internet.
  - Customer - a network is paid to allow Internet access.

Classification of networks

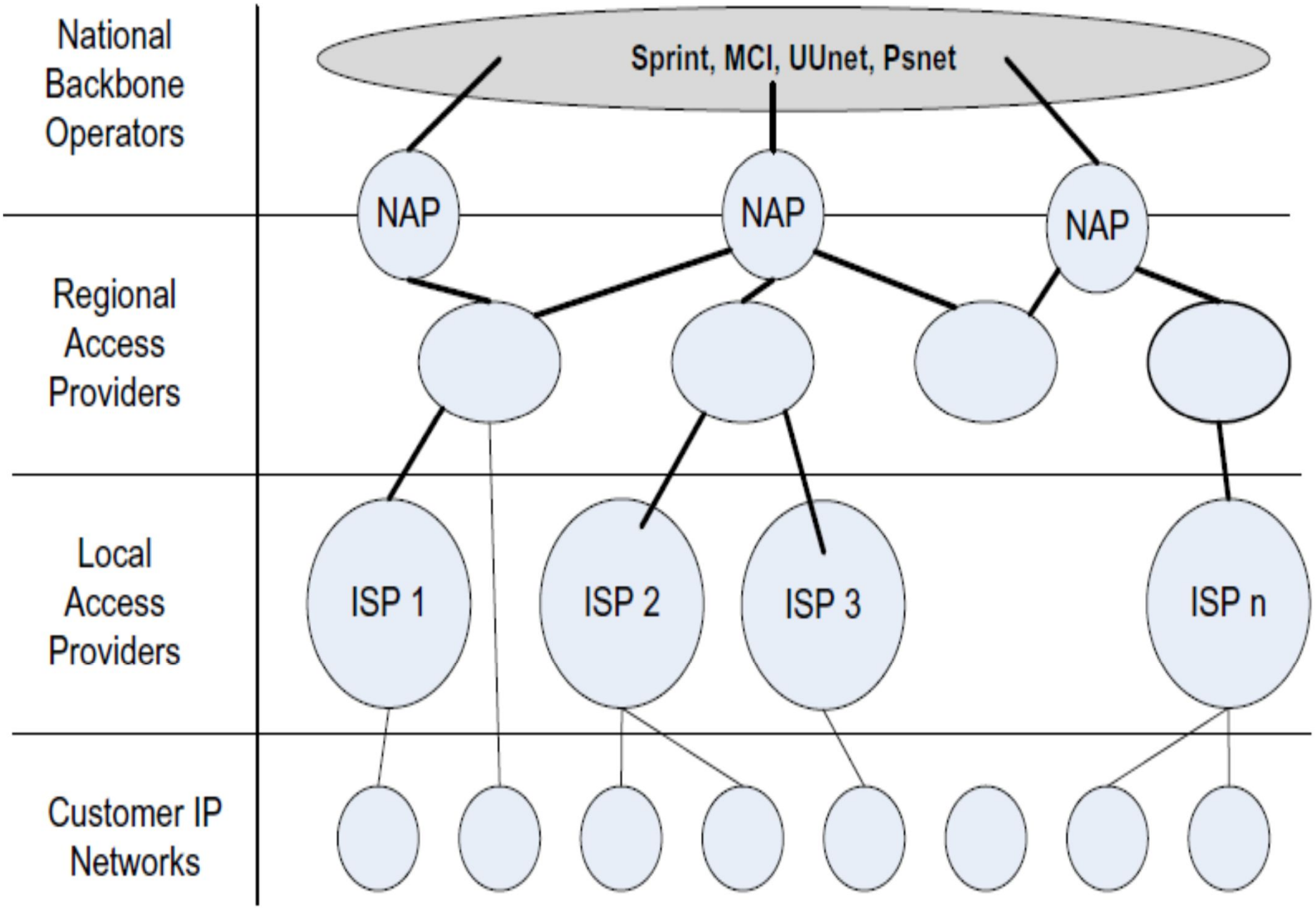
- Tier 1 - can reach every other network on the Internet without purchasing IP transit or paying settlements.
- Tier 2 - an Internet service provider who engages in the practice of peering with other networks, but who still purchases IP transit to reach some portion of the Internet; the common providers on the Internet.
- Tier 3 - purchases transit rights from other networks (typically Tier 2 networks) to reach the Internet.

Classification of networks

- Tier 1 - can reach every other network on the Internet without purchasing IP transit or paying settlements.
- Tier 2 - an Internet service provider who engages in the practice of peering with other networks, but who still purchases IP transit to reach some portion of the Internet; the common providers on the Internet.
- Tier 3 - purchases transit rights from other networks (typically Tier 2 networks) to reach the Internet.



A point-of-presence (POP) is an access point from one place to the rest of the Internet. An *Internet exchange point (IXP)* is a physical infrastructure allowing ISPs to exchange Internet traffic.


National Backbone Operators
Regional Access Providers
Local Access Providers
Customer IP Networks
Sprint, MCI, UUnet, Psnet
NAP
NAP
NAP
ISP 1
ISP 2
ISP 3
ISP n

Web access and the TCP congestion control

- The Web supports access to content stored on a cloud; virtually all cloud computing infrastructures allow users to interact with their computations on the cloud using Web-based systems.
- HTTP, the application protocol for Web browsers, uses TCP for web access.
- TCP supports mechanisms to avoid congestion and limit the amount of data transported over the Internet.

Web access and the TCP congestion control window

- Thus, it should be clear that the metrics related to Web access are important for designing and tuning networks.
- Web access requires the transfer of large amounts of data as we can see in measurements reported by Google. Metrics, such as the average size of a page or the number of GET operations, are useful to explain the results of performance measurements carried out on existing systems and to propose changes to optimize the performance.

Table 7.1 Web statistics.

Metric	Value
Number of sample pages analyzed	4.2×10^9
Average number of resources per page	44
Average number of GETs per page	44.5
Average number of unique host names encountered per page	7
Average size transferred over the network per page, including HTTP headers	320 KB
Average number of unique images per page	29
Average size of the images per page	206 KB
Average number of external scripts per page	7
Number of sample SSL (HTTPS) pages analyzed	17×10^6

Congestion

- Congestion is an important issue that can arise in packet switched network.
- Congestion is a situation in Communication Networks in which too many packets are present in a part of the subnet, performance degrades.
- Congestion in a network may occur when the load on the network *(i.e.* the number of packets sent to the network) is greater than the capacity of the network *(i.e.* the number of packets a network can handle.)

Congestion Control

- congestion control refers to the network mechanism and techniques used to control congestion and keep the load below the networks capacity.

TCP congestion control

- TCP seeks to achieve high channel utilization, avoid congestion, and, at the same time, ensure a fair sharing of the network bandwidth.
- TCP uses a sliding window flow control protocol. If W is the window size, then the data rate S of the sender is:

$$S = \frac{W \times \text{MSS}}{\text{RTT}} \text{bps} = \frac{W}{\text{RTT}} \text{ packets/second}$$

TCP congestion control

- TCP seeks to achieve high channel utilization, avoid congestion, and, at the same time, ensure a fair sharing of the network bandwidth.
- TCP uses a sliding window flow control protocol. If W is the window size, then the data rate S of the sender is:

$$S = \frac{W \times \text{MSS}}{\text{RTT}} \text{bps} = \frac{W}{\text{RTT}} \text{ packets/second}$$

TCP congestion control

- where MSS and RTT denote the maximum segment size and the round-trip time, respectively.
- If *S* is too, small, the transmission rate is smaller than the channel capacity, whereas a large *S* leads to congestion.
- The channel capacity in the case of communication over the Internet is not a fixed quantity, but, different physical channels are shared among many flows and it depends on the load of the network.

TCP congestion control

- Tahoe is an popular algorithm used for controlling congestion and is based on: (1) slow start, (2) congestion avoidance, and (3) fast retransmit. It allows faster recovery after losses through fast retransmission..
- **Slow-start** is used to avoid sending more data than the network is capable of transmitting, that is, to avoid causing network congestion.

Network resource management

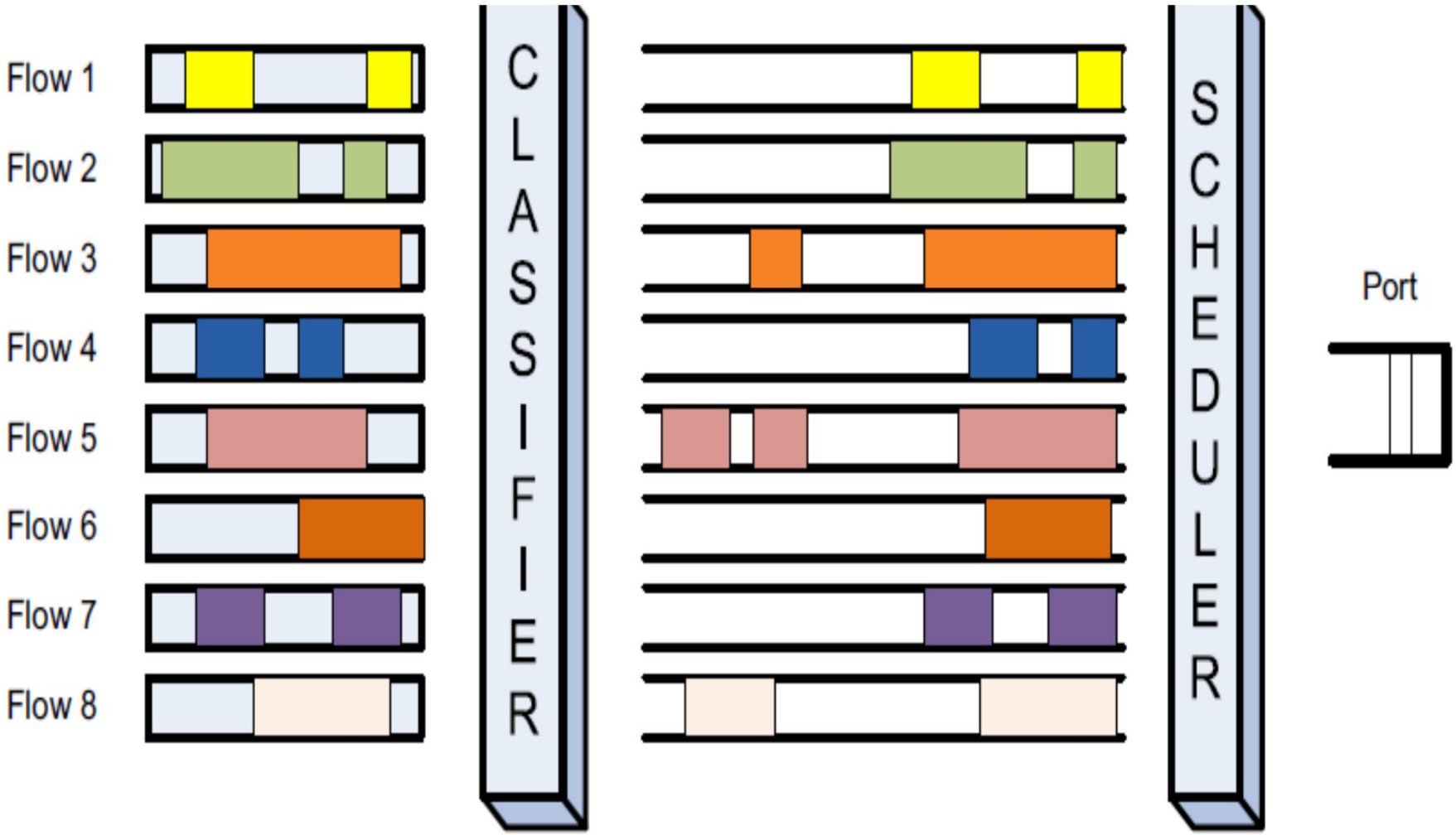
- *Cloud computing is dependent on communication; thus, network resource management is a very important aspect of the management of computer clouds.*
- *A critical aspect of resource management in cloud computing is to guarantee the communication bandwidth required by an application as specified by a service-level agreement (SLA).*
- .

Techniques

- stochastic fairness queuing (SFQ) algorithm
- class-based queuing (CBQ) method

stochastic fairness queuing (SFQ) algorithm

- Variant of **fairness queuing.** The goal is to ensure fairness so that each flow is able to send data in turn, thus preventing any single flow from drowning out the rest.
- which takes into account data packet sizes to ensure that each flow has the opportunity to transmit an equal amount of data.
- Packets are first classified into flows by the system and then assigned to a queue dedicated to the flow; queues are serviced one packet at a time in round-robin order.

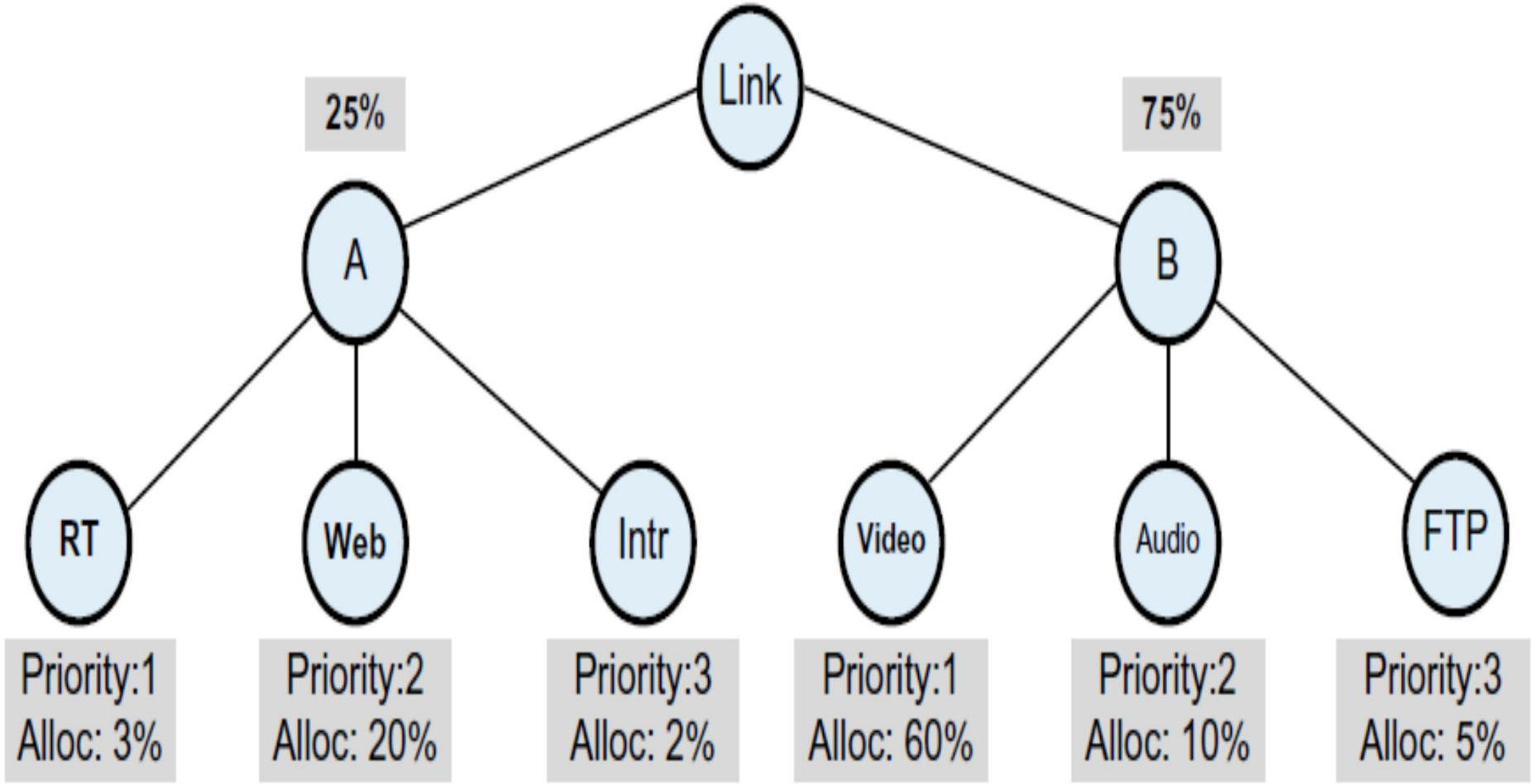

Flow 1
Flow 2
Flow 3
Flow 4
Flow 5
Flow 6
Flow 7
Flow 8
CLASSIFIER
SCHEDULER
Port

Class-based queuing (CBQ) method

- **Class-based queuing** (CBQ) is a **queuing** discipline for the network scheduler that allows traffic to share bandwidth equally, after being grouped by **classes**.
- The **classes** can be **based** upon a variety of parameters, such as priority, interface, or originating program.
- The objective of CBQ is to support flexible link sharing for applications that require bandwidth guarantees such as VoIP, video streaming, and audio streaming.
- CBQ supports some balance between short-lived network flows, such as Web searches, and long-lived ones, such as video streaming or file transfers.

Class-based queuing (CBQ) method

- Class-Based Queueing feature provides the ability to define multiple groups (classes) of users. You can prioritized groups and guarantee minimum bandwidth on a per-group basis.



CBQ link sharing for two groups: A, of short-lived traffic, and B, of long-lived traffic, allocated 25% and 75% of the link capacity. Intr (interactive applications)

Interconnection networks for computer clouds

- The speed of the Ethernet has increased from 1 Gbps in 1997 to the current speed 100 Gbps.
- The networking infrastructure of a cloud must satisfy several requirements:
  - Scalability.
  - Low cost.
  - Low-latency.
  - High bandwidth.
  - Provide location transparent communication between servers.

Location transparent communication

- Every server should be able to communicate with every other server with similar speed and latency.
- This requirement ensures that *applications need not be location aware* and, at the same time, it reduces the complexity of the system management.

Elements of the interconnection fabric

- Important elements of the interconnection fabric are routers and switches.
- Routers are switches with a very specific function: joining multiple networks, LANs, and WANs. They receive IP packets, look inside each packet to identify the source and target IP addresses, then forward these packets as needed to ensure that data reaches its final destination.

Networking infrastructure

- Organized hierarchically. The servers are packed into racks and interconnected by a top-of-the-rack router; then rack routers are connected to cluster routers, which in turn are interconnected by a local communication fabric.

Interconnection networks - InfiniBand

- Interconnection network used by supercomputers and computer clouds.
  - Has a switched fabric topology designed to be scalable.
  - Supports several signaling rates.
  - The energy consumption depends on the throughput.
  - Links can be bonded together for additional throughput.
- The *InifiniBand* fabric is used to connect compute nodes, compute nodes with storage servers.

Interconnection networks - InfiniBand

- The data rates.
  - single data rate (SDR) - 2.5 Gbps in each direction per connection.
  - double data rate (DDR) - 5 Gbps.
  - quad data rate (QDR) – 10 Gbps.
  - fourteen data rate (FDR) – 14.0625 Gbps.
  - enhanced data rated (EDR) – 25.78125 Gbps.

Interconnection networks - InfiniBand

- Advantages.
  - high throughput, low latency.
  - supports quality of service guarantees and failover - the capability to switch to a redundant or standby system

Routers and switches

- The cost of routers and the number of cables interconnecting the routers are major components of the cost of interconnection network.
- Better performance and lower costs can only be achieved with innovative router architecture.

Routers and switches

- Router – switch interconnecting several networks.
  - low-radix routers – have a small number of ports; divide the bandwidth into a smaller number of wide ports.
  - high-radix routers - have a large number of ports; divide the bandwidth into larger number of narrow ports

Network characterization

- The diameter of a network is the average distance between all pairs of nodes; if a network is fully-connected its diameter is equal one.
- When a network is partitioned into two networks of the same size, the bisection bandwidth measures the communication bandwidth between the two.
- The cost.
- The power consumption.

Content-delivery networks

- Computer clouds support not only network-centric computing, but also network-centric content. For example, Internet video is expected to generate over 18 Exabytes of data per month.
- The vast amount of data stored on the cloud has to be delivered efficiently to a large user population.

Content-delivery networks

- Content-delivery networks (CDNs) offer fast and reliable content delivery and reduce the communication bandwidth by caching and replication. A CDN receives the content from an *origin* server, then replicates it to its *edge* cache servers.
- The content is delivered to an end user from the “closest” edge server.

Content-delivery networks

- CDNs are designed to support scalability, increase reliability and performance, and provide better security.
- The resources provided by a CDN are replicated, and when one of the replicas fails, the content is available from another one. The replicas are “close” to the consumers of the content, and this placement reduces the startup time and the communication bandwidth.

Content-delivery networks

- A CDN uses two types of server: the *origin* server updated by the content provider and *replica* servers that cache the content and serve as an authoritative reference for client requests.
- A CDN can deliver static content and/or live or on-demand streaming media.
  - Static content - media that can be maintained using traditional caching technologies as changes are infrequent. Examples: HTML pages, images, documents, software patches, audio and video files.
  - Live media - live events when the content is delivered in real time from the encoder to the media server.

Live events

- On-demand delivery of audio and/or video streams, movie files, and music clips.

Akamai

- The first CDN was set up by *Akamai*, a company evolved from a Massachusetts Institute of Technology (MIT) project to optimize network traffic.
- *Akamai* has placed some 20,000 servers in 1,000 networks in 71 countries since its inception. In 2009 it controlled some 85% of the market.

Akamai

- *Akamai* mirrors the contents of clients on multiple systems placed strategically through the Internet.
- Though the domain name is the same, the IP address of the resource requested by a user points to an *Akamai* server rather than the customer's server.
- Then the *Akamai* server is automatically picked, depending on the type of content and the network location of the end user.

Other CDNs

- *EdgeStream*,
- *Limelight Networks*,
- *Coral* (academic)
- *Globule* is an open-source collaborative CDN

Protocols used by CDNs

- *Network Element Control Protocol* (NECP)
- *Web Cache Coordination Protocol* (WCCP),
- *SOCKS*
- *Cache Array Routing Protocol* (CARP)
- *Internet Cache Protocol* (ICP)
- *Hypertext Caching Protocol* (HTCP)
- *Cache Digest*,

Important design and policy decisions for a CDN

- The placement of the edge servers.
- The content selection and delivery.
- The content management.
- The request routing policies.

The critical metrics of CDN performance

- Cache hit ratio, which is the ratio of the number of cached objects versus total number of objects requested.
- Reserved bandwidth for the origin server.
- Latency, which is based on the perceived response time by the end users.
- Edge server utilization.
- Reliability, which is based on packet-loss measurements.

Note

- CDNs will face considerable challenges in the future due to increased appeal of data streaming and to the proliferation of mobile devices such as smartphones and tablets.
- On-demand video streaming requires enormous bandwidth and storage space as well as powerful servers.
- CDNs for mobile networks must be able to dynamically reconfigure the system in response to spatial and temporal demand variations.

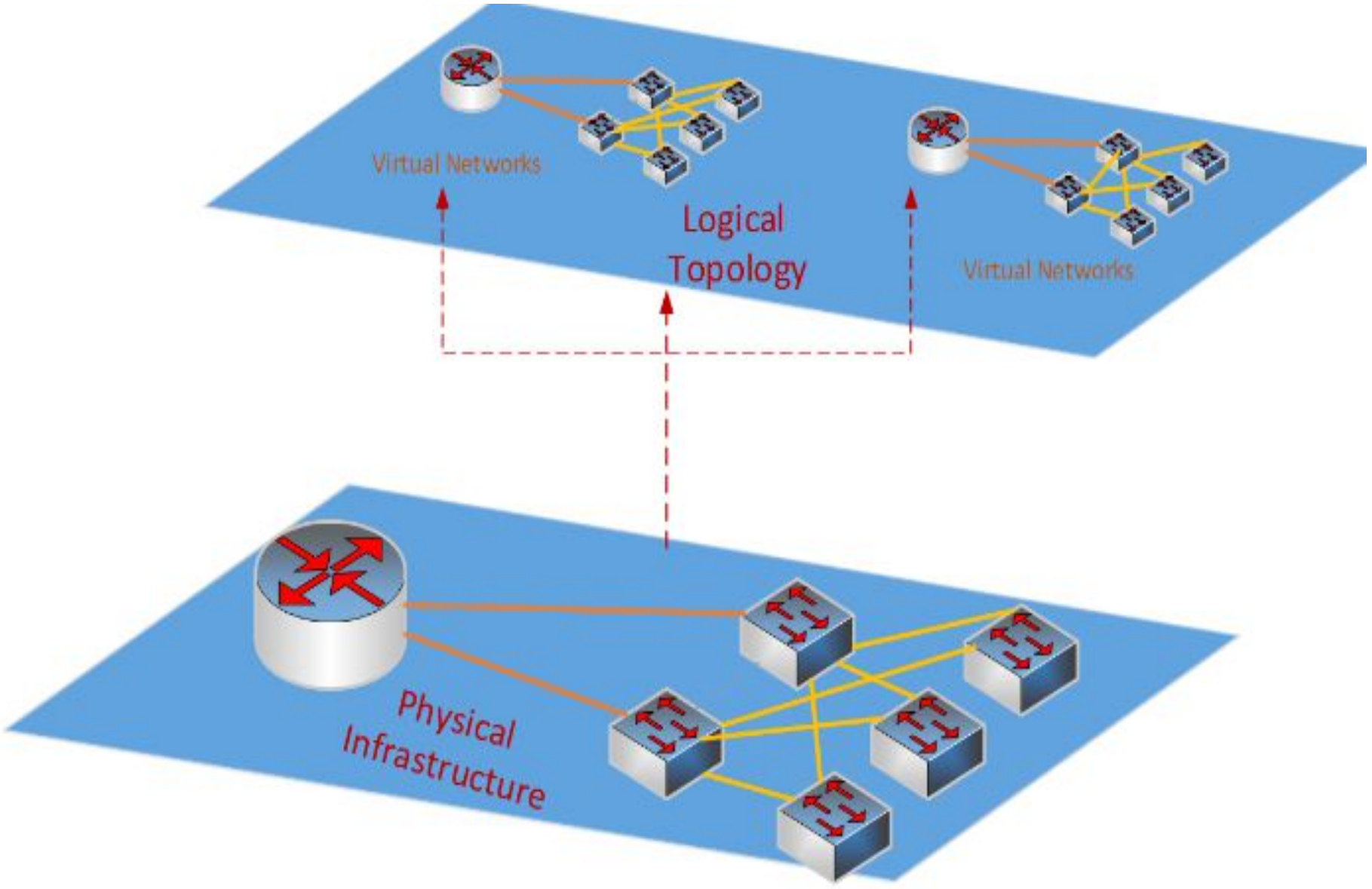
Overlay networks

- An *overlay network*, or virtual network, is a network built on top of a physical network.
- The nodes of an overlay network are connected by virtual links and each of these links correspond to a path in the underlying network.
- Overlay networks are widely used in many distributed systems such as peer-to-peer systems, content-delivery systems, and client-server systems; in all these cases the distributed systems communicate through the Internet.

Overlay networks

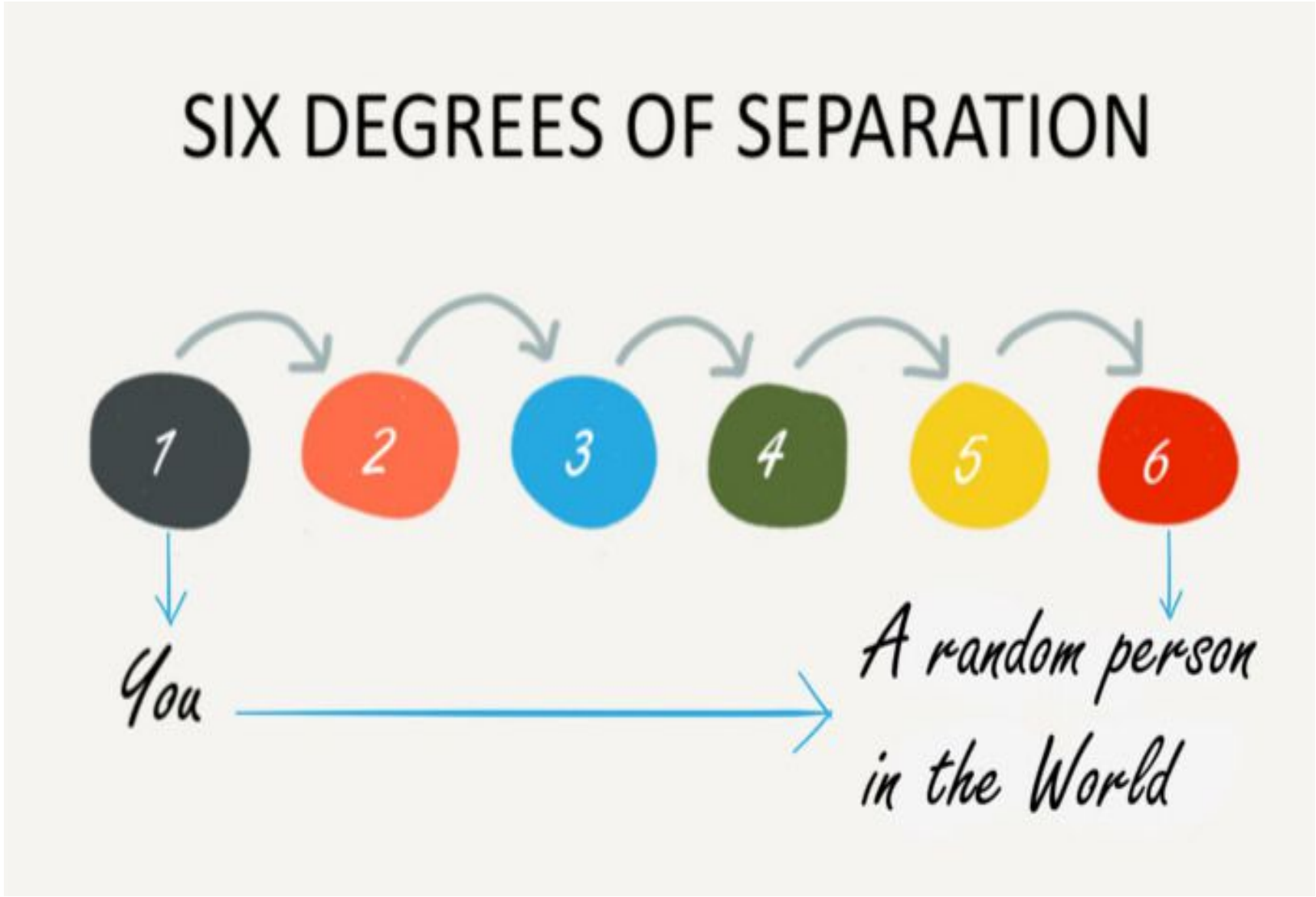
- Overlay networking is a method of using software to create layers of network abstraction that can be used to run multiple separate, discrete virtualized network layers on top of the physical network, often providing new applications or security benefits.
- An overlay network can support QoS guarantees for data-streaming applications through improved routing over the Internet.

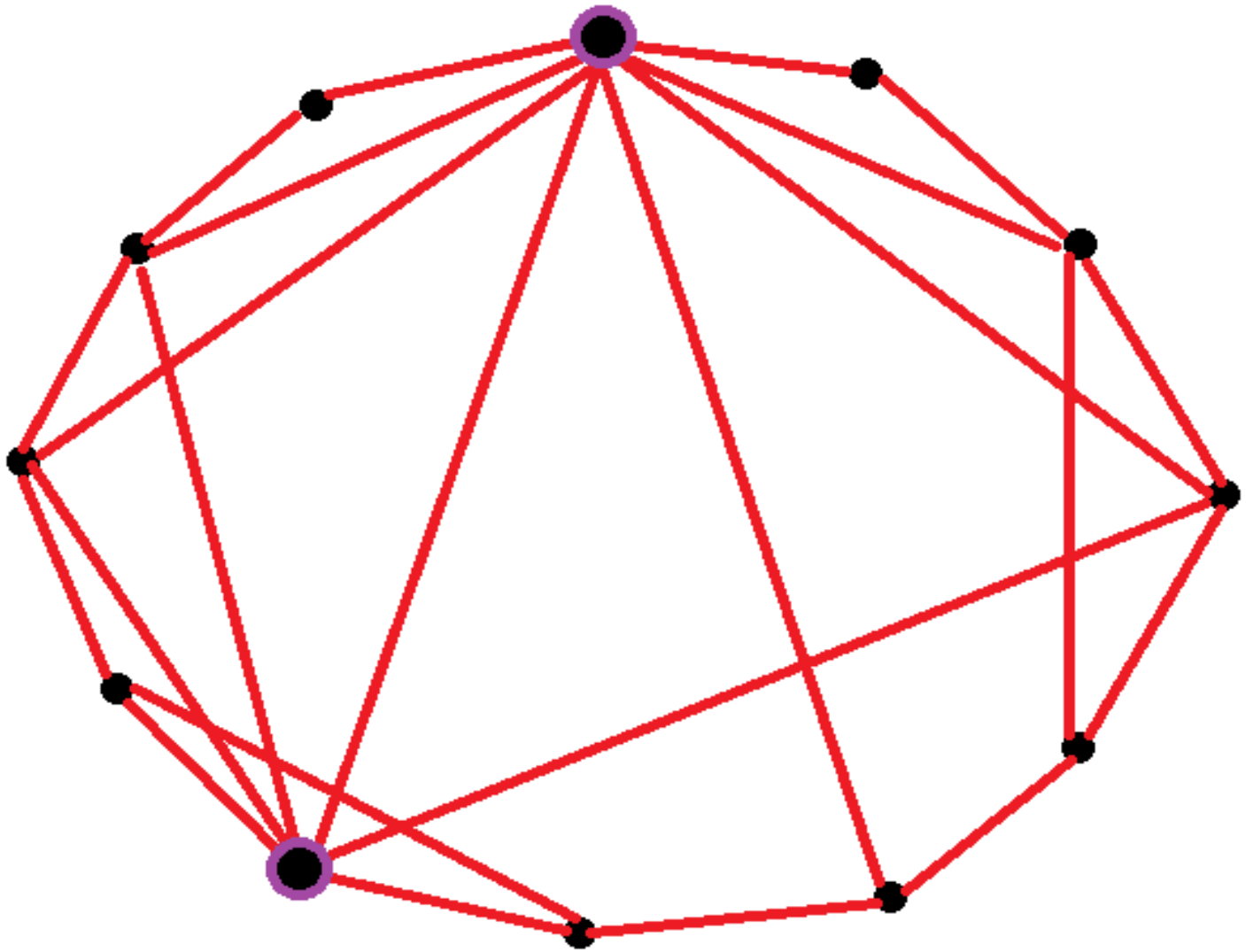
Overlay networks

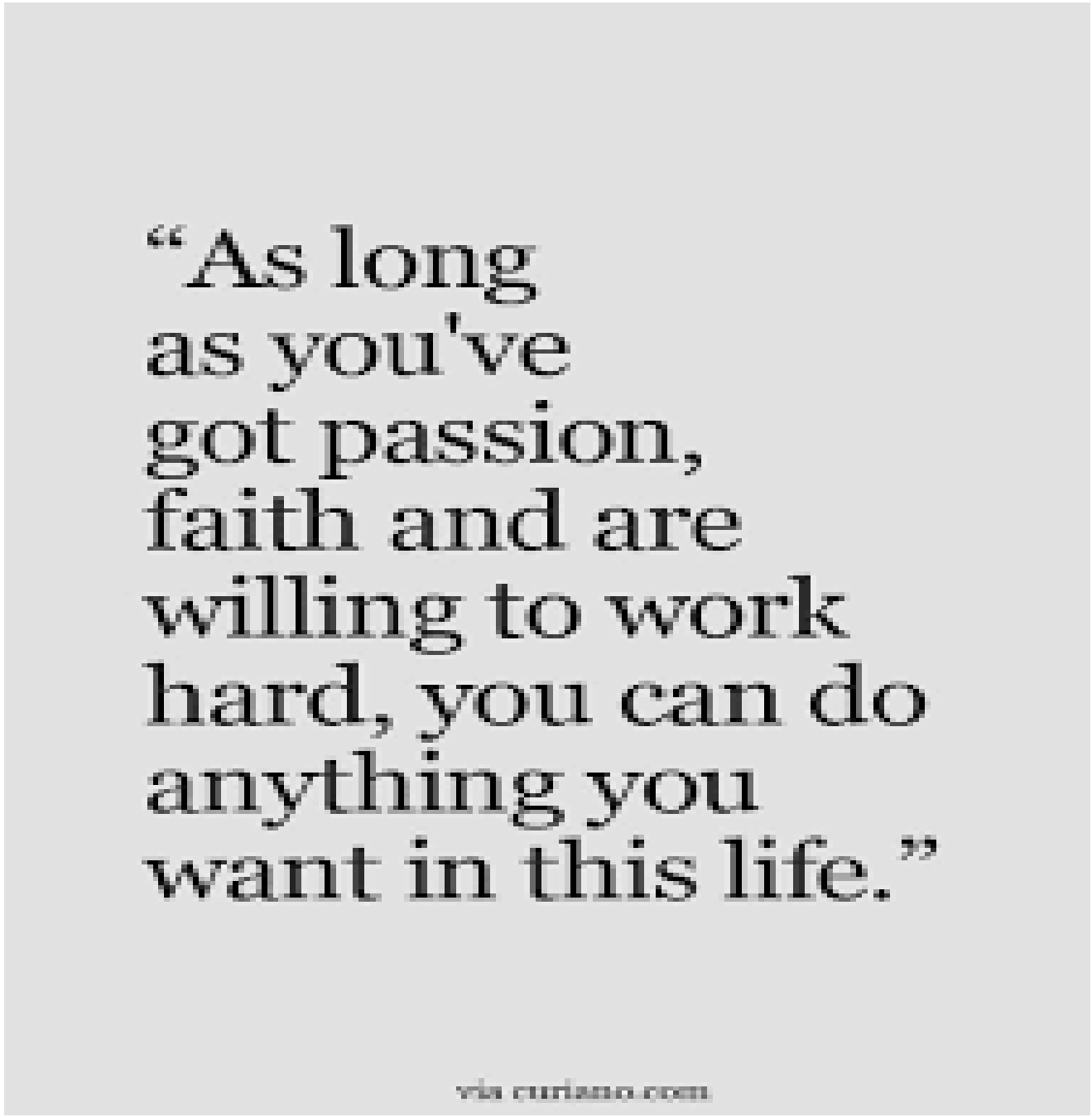


Small-World Networks

- A **small-world network** is a type of mathematical graph in which most nodes are not neighbors of one another, but the neighbors of any given node are likely to be neighbors of each other and most nodes can be reached from every other node by a **small** number of hops or steps.
- It is sometimes also known as "six degrees of separation" since, in the social network of the world, any person turns out to be linked to any other person by roughly six connections.


SIX DEGREES OF SEPARATION
1
2
3
4
5
6
You
A random person
in the World




“As long
as you've
got passion,
faith and are
willing to work
hard, you can do
anything you
want in this life.”
via curiano.com