Introduction to TCP/IP

Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

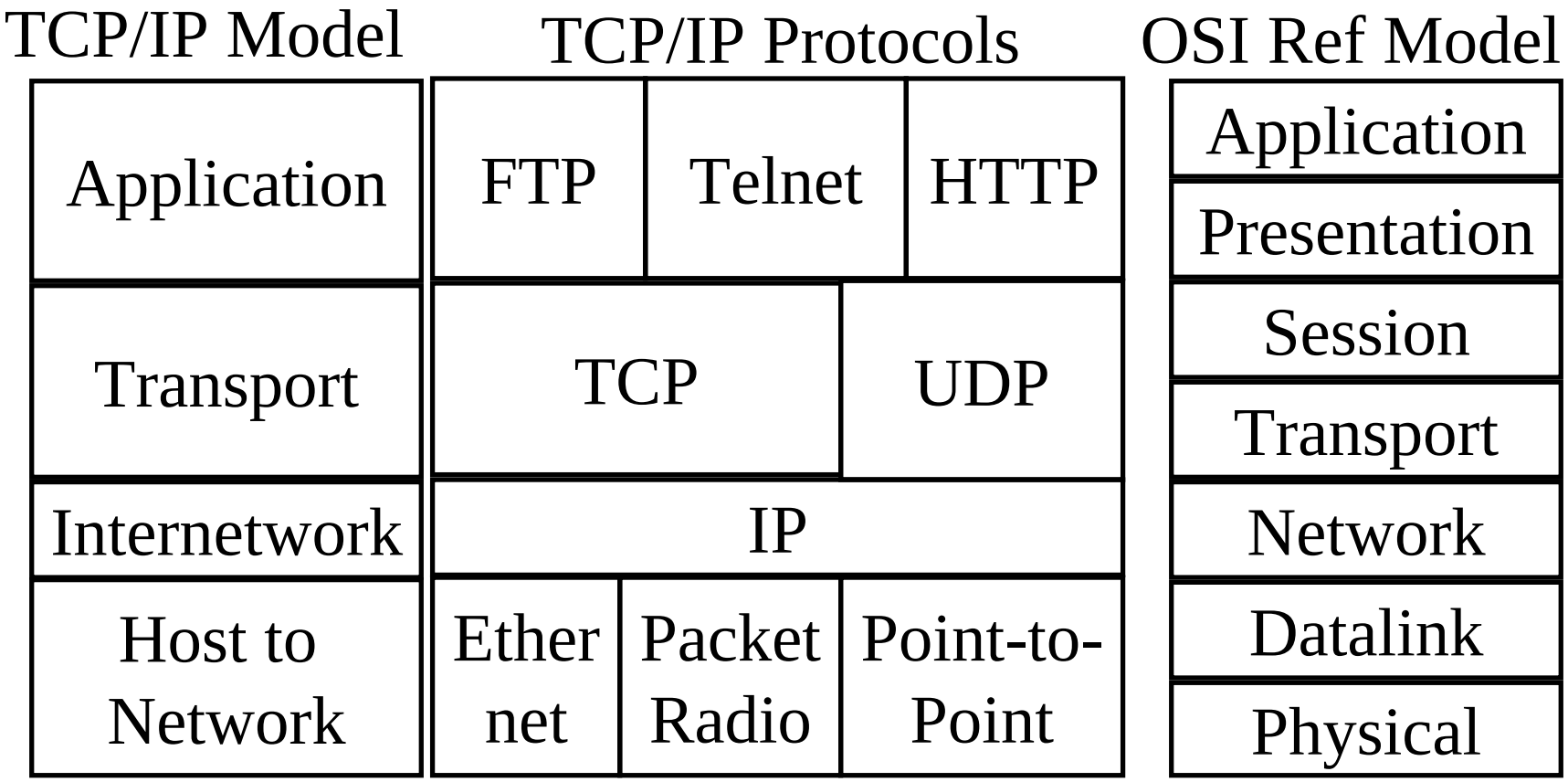
The
(
s
35

Overview

- Internetworking Protocol (IP)
- IP Addressing
- Domain Name System
- Routing Protocols: RIP, OSPF, BGP
- Transport Protocols: TCP, UDP

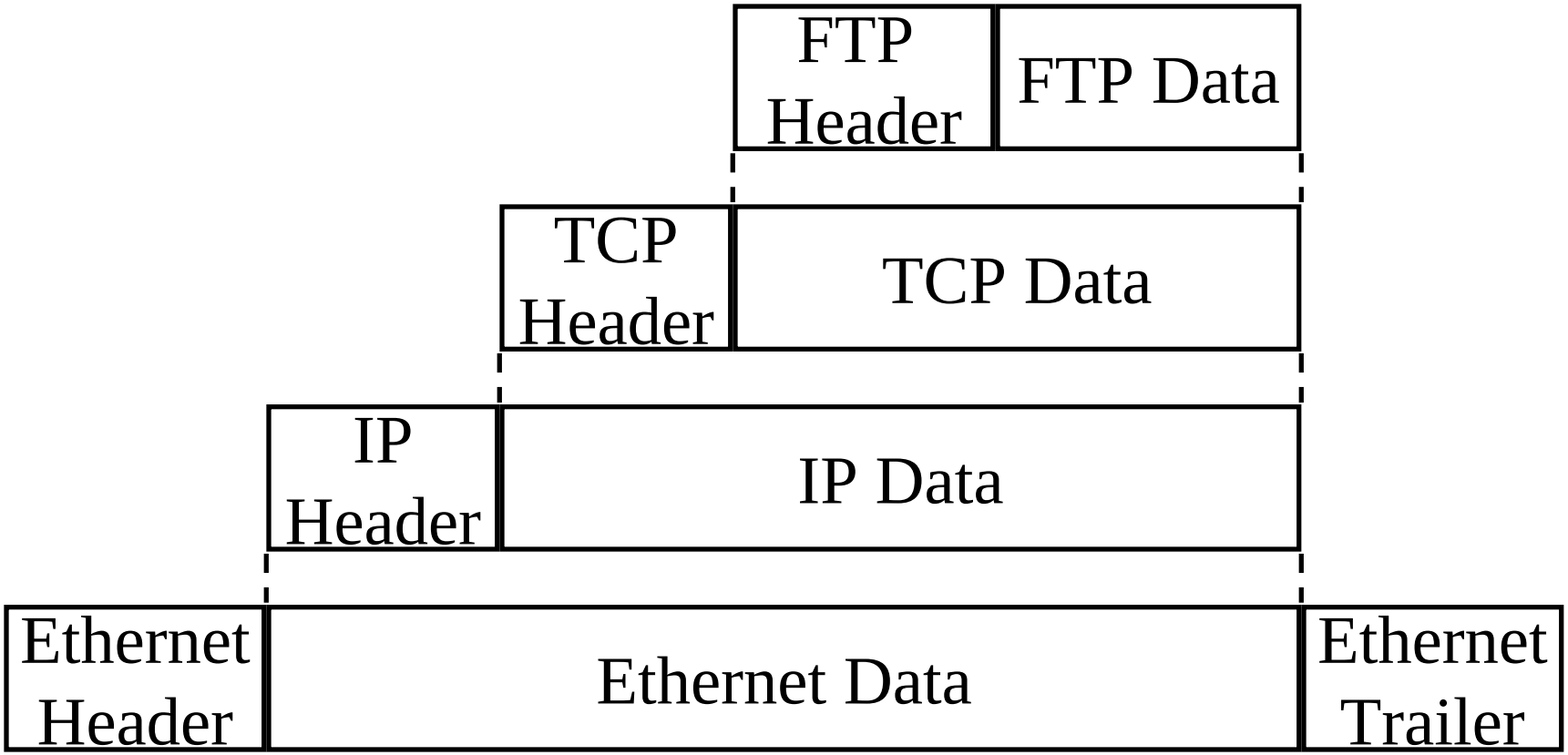
TCP/IP Reference Model

- TCP = Transport Control Protocol
- IP = Internet Protocol (Routing)



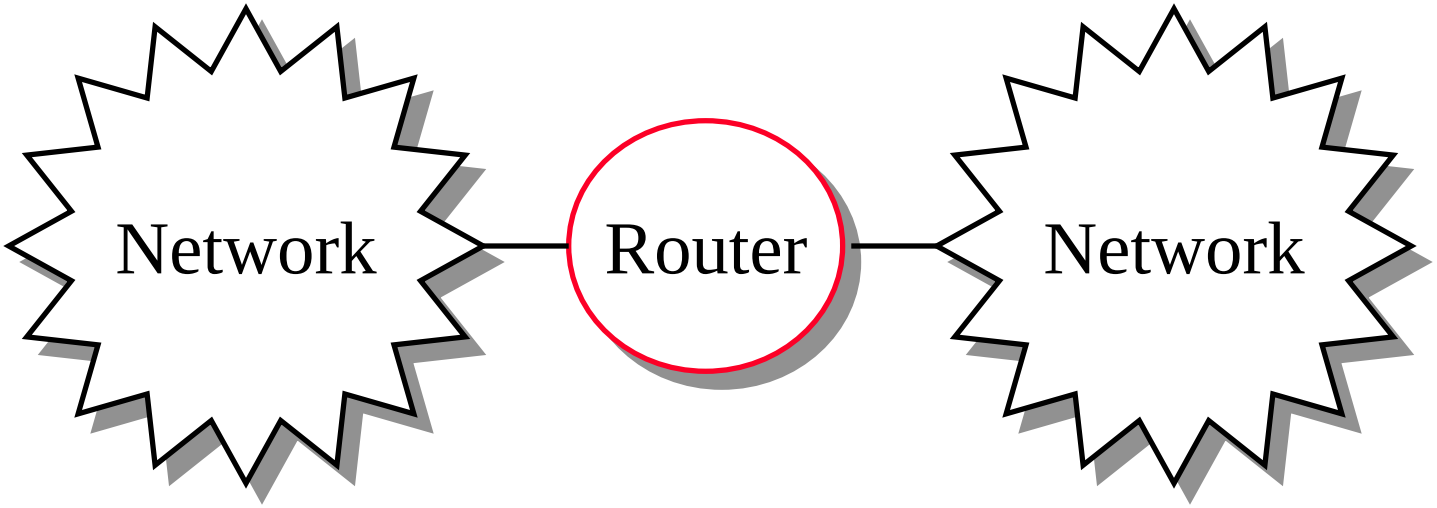
Layered Packet Format

- Nth layer control info is passed as N-1th layer data.
  Example: File Transfer Protocol (FTP)



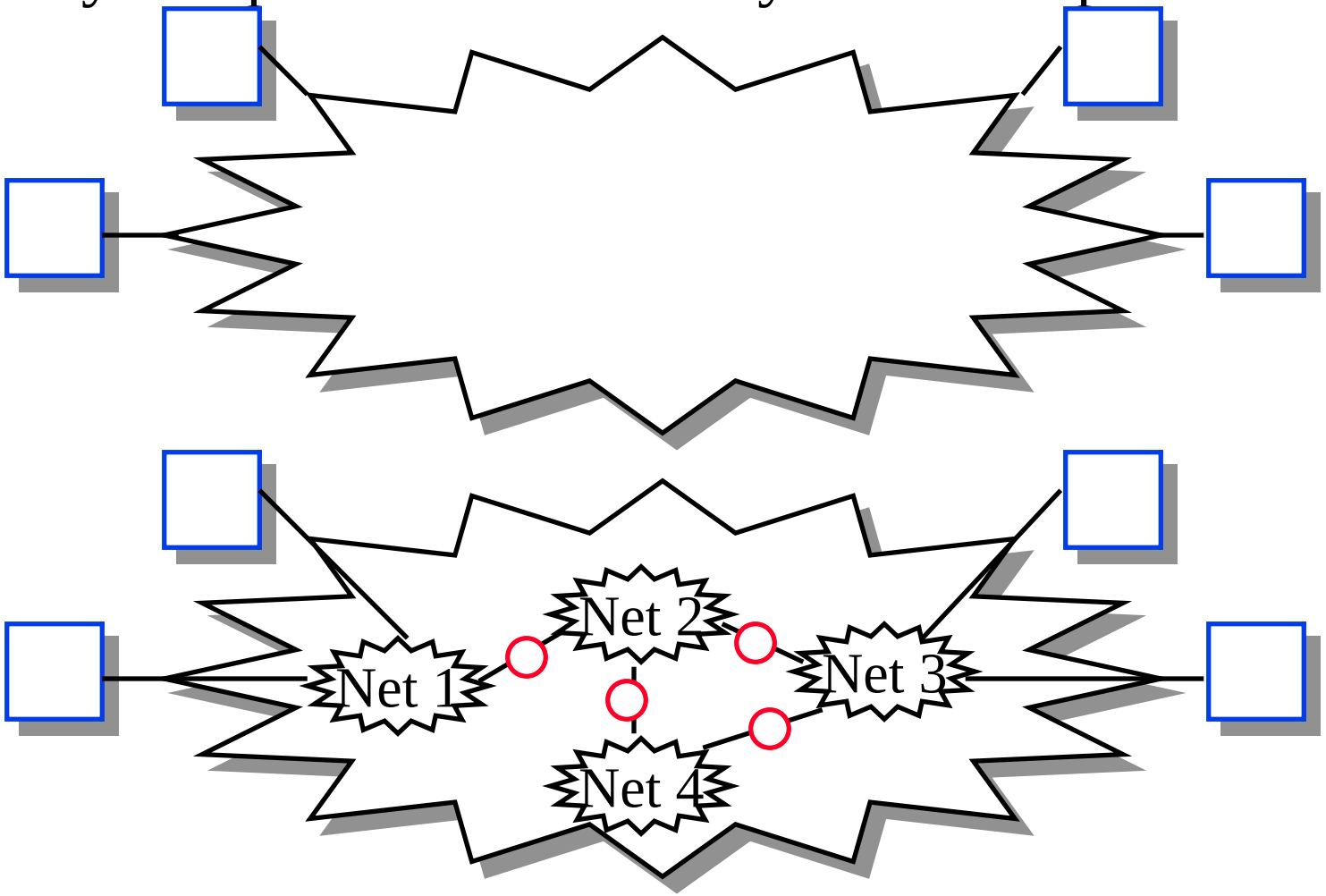
Internetworking

- Inter-network = Collection of networks
  Connected via routers



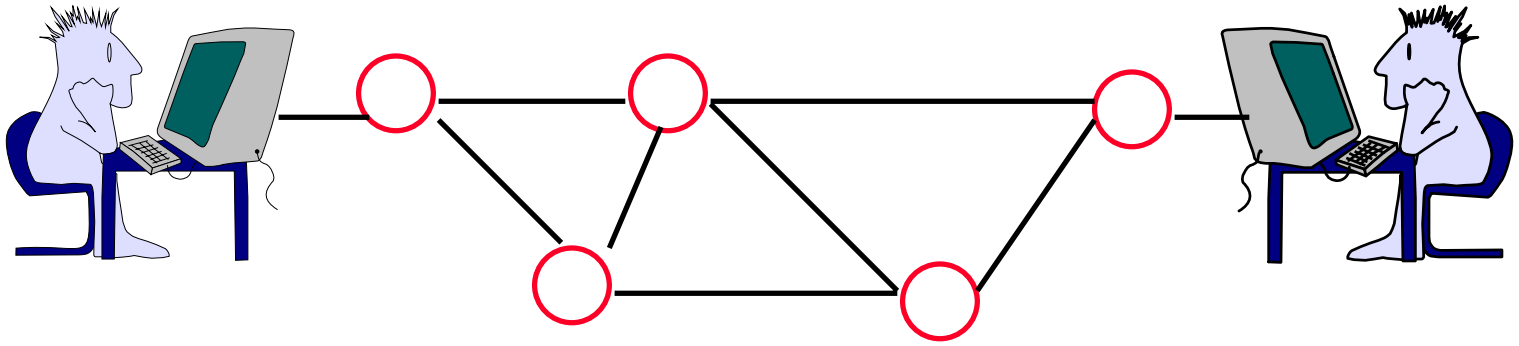
Internet = Collection of Networks

- Any computer can talk to any other computer

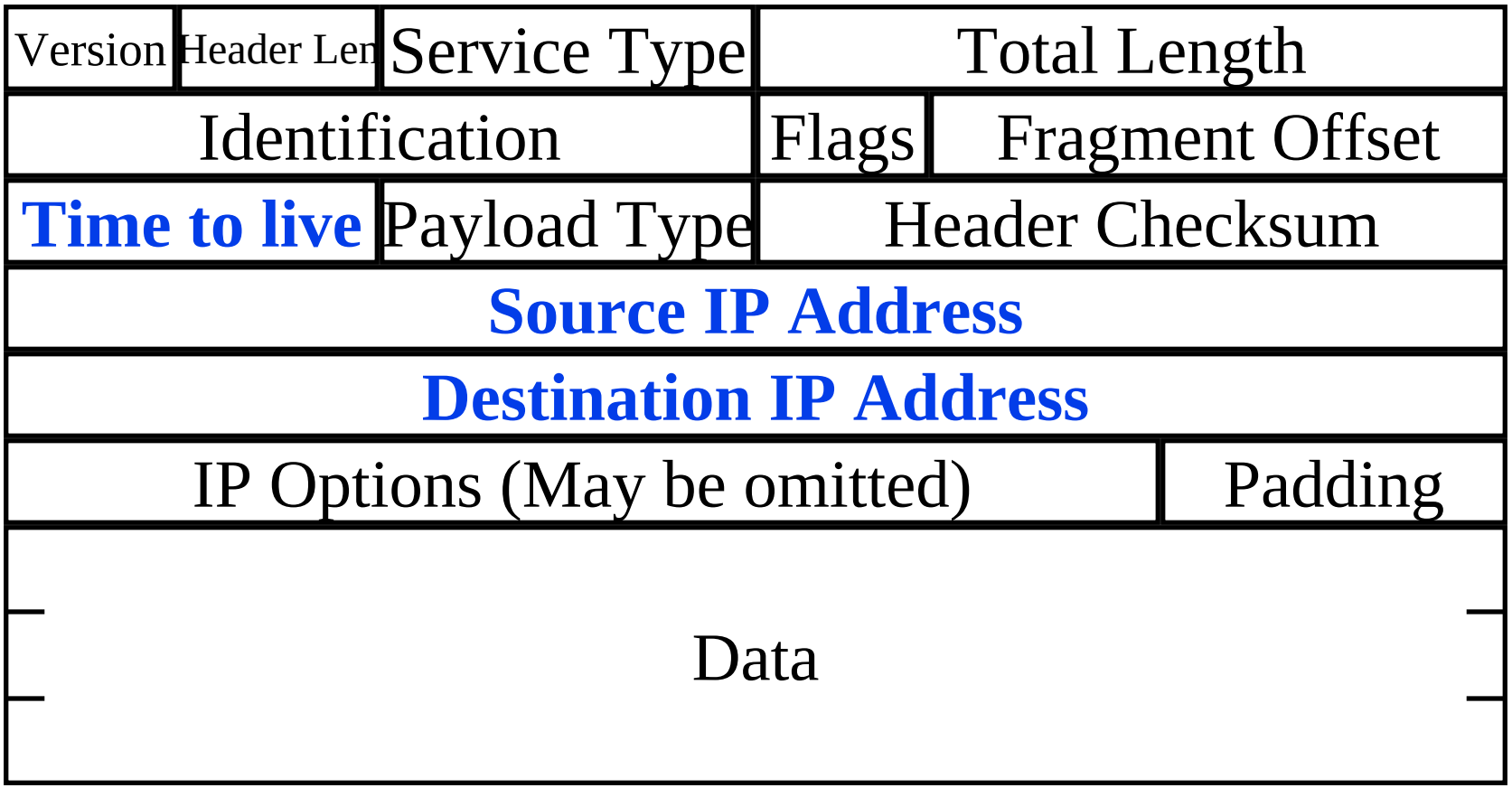


Internet Protocol (IP)

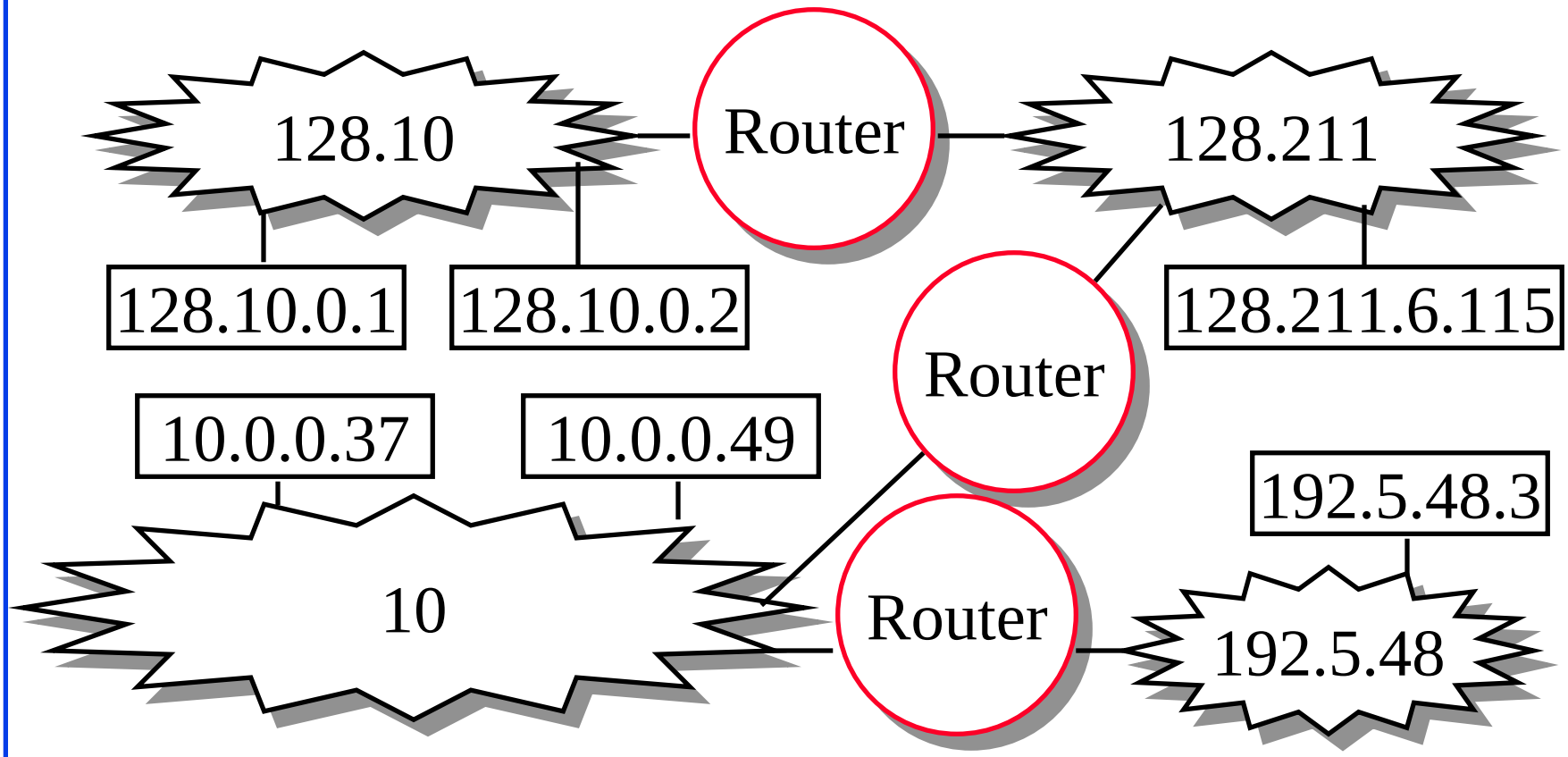
- Layer 3 protocol that *forwards* datagrams across internet
- Uses routing tables prepared by routing protocols, e.g., Open Shortest Path First (OSPF), Routing Information Protocol (RIP)
- Connectionless service
  vs connection-oriented (circuits)



IP Datagram Format

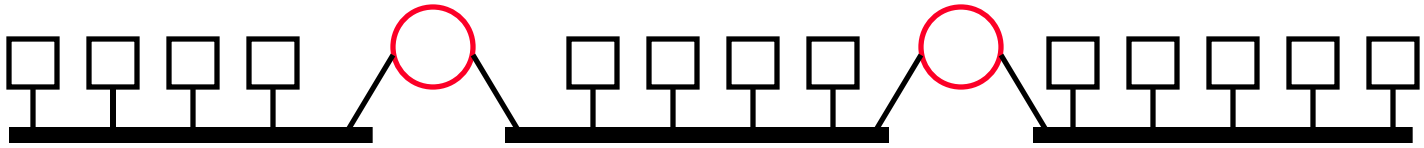


IP Addressing



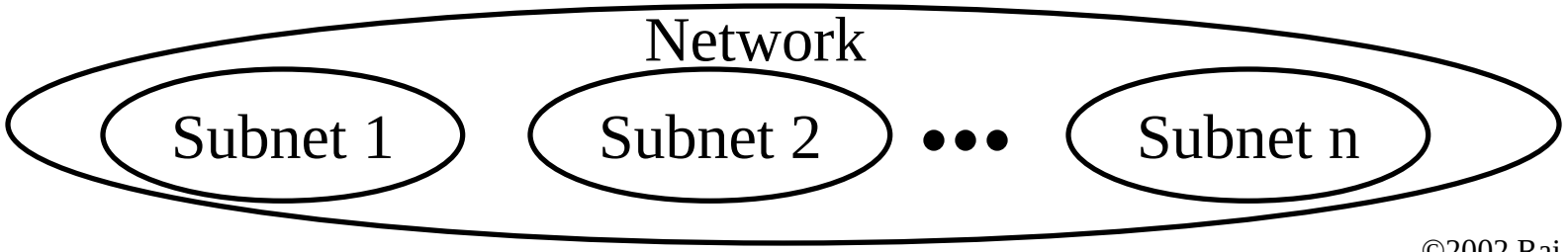
- All IP hosts have a 32-bit address.128.10.0.1 =1000 0000 0000 1010 0000 0000 0000 0001
- All hosts on a network have the same network prefix

Subnetting



- All hosts on a subnetwork have the same prefix. Position of the prefix is indicated by a "subnet mask"
- Example: First 23 bits = subnet

  Address: 10010100 10101000 00010000 11110001
  Mask: 11111111 11111111 11111110 00000000
  .AND. 10010100 10101000 00010000 00000000



Forwarding an IP Datagram

- Delivers **datagram**s to destination network (subnet)
- Routers maintain a "routing table" of "next hops"
- Next Hop field does not appear in the datagram

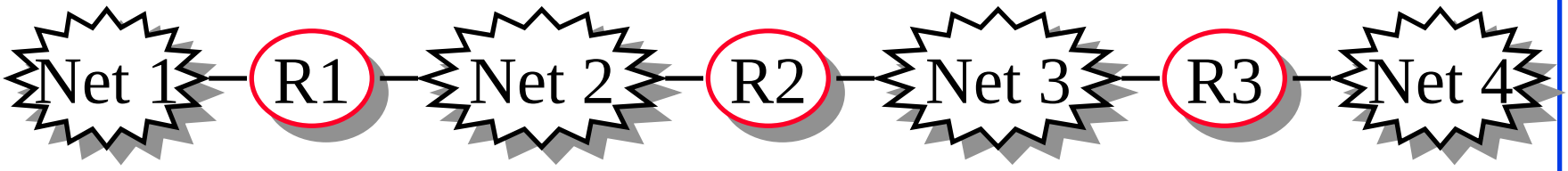
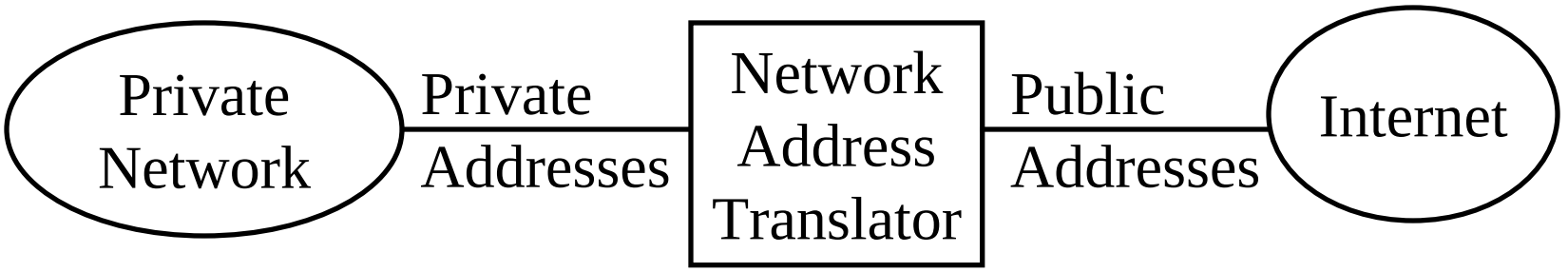


Table at R2:

Destination	Next Hop
Net 1	Forward to R1
Net 2	Deliver Direct
Net 3	Deliver Direct
Net 4	Forward to R3

Private Addresses

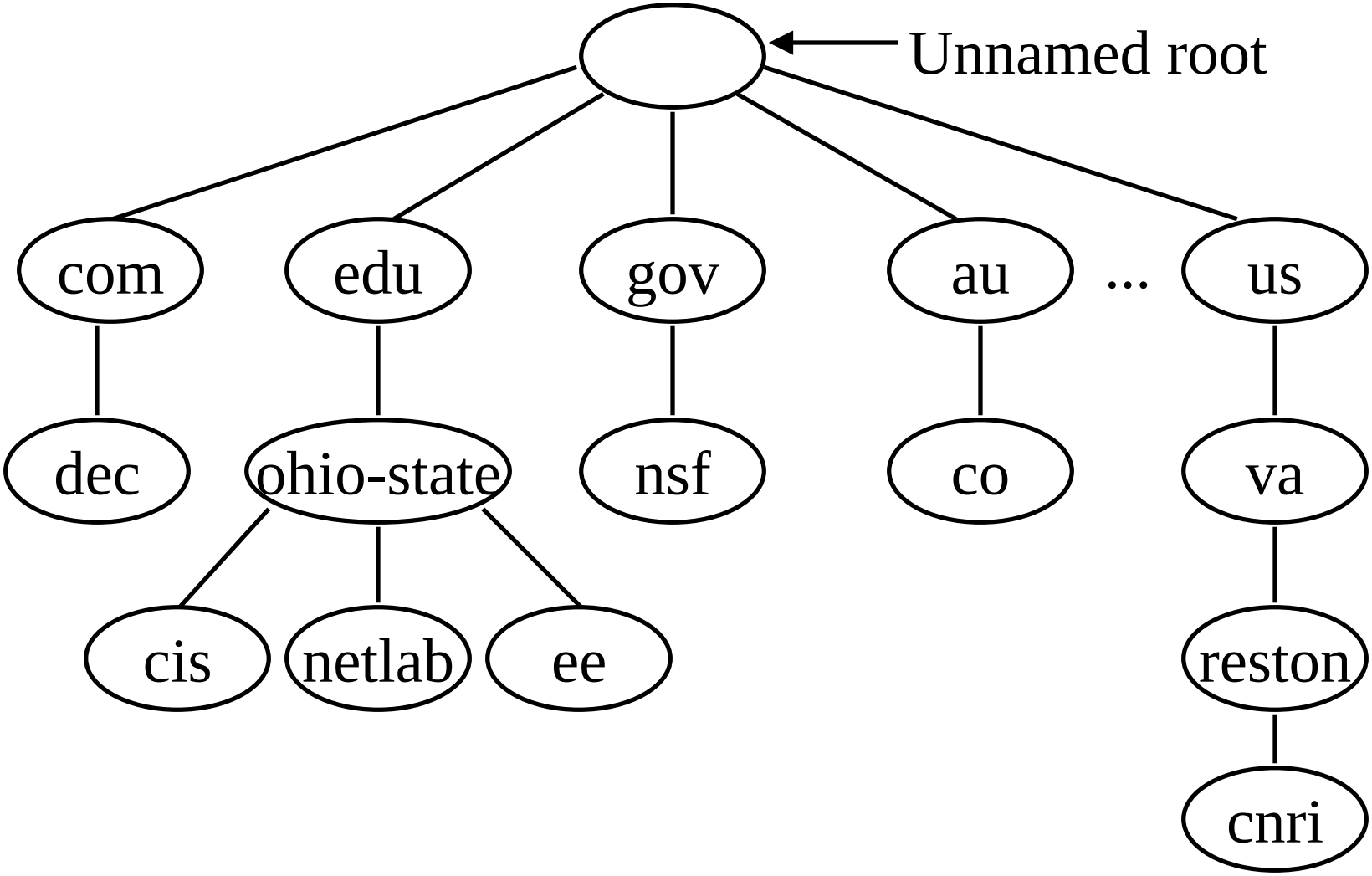
- Any organization can use these inside their network Can't go on the internet. [RFC 1918]
- 10.0.0.0 - 10.255.255.255 (10/8 prefix)
- 172.16.0.0 - 172.31.255.255 (172.16/12 prefix)
- 192.168.0.0 - 192.168.255.255 (192.168/16 prefix)



Domain Name Service

- Computers use addresses
- Humans cannot remember IP addresses
  $\Rightarrow$ Need names
  Example, Liberia for 164.107.51.28
- Simplest Solution: Each computer has a unique name and has a built in table of name to address translation
- Problem: Not scalable
- Solution: DNS (Adopted in 1983)
- Hierarchical Names: Liberia.cis.ohio-state.edu

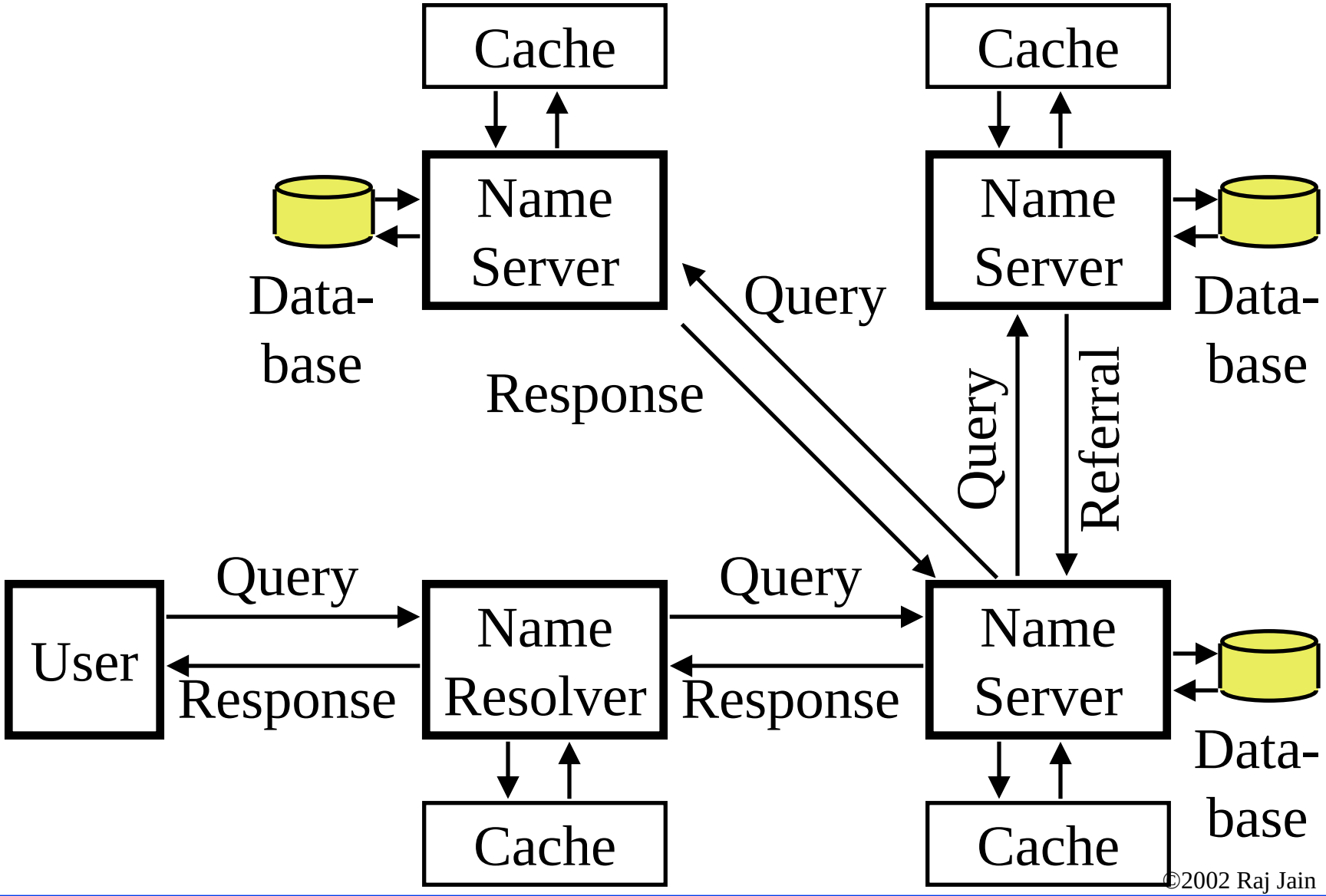
Name Hierarchy



Name Hierarchy

- Unique domain suffix is assigned by Internet Assigned Number Authority (IANA)
- The domain administrator has complete control over the domain
- No limit on number of sub-domains or number of levels
- computer.site.division.company.com
  computer.site.subdivision.division.company.com
- Name space is not related to physical interconnection, e.g., math.ohio-state and cis.ohio-state could be on the same floor or in different cities

Name Resolution

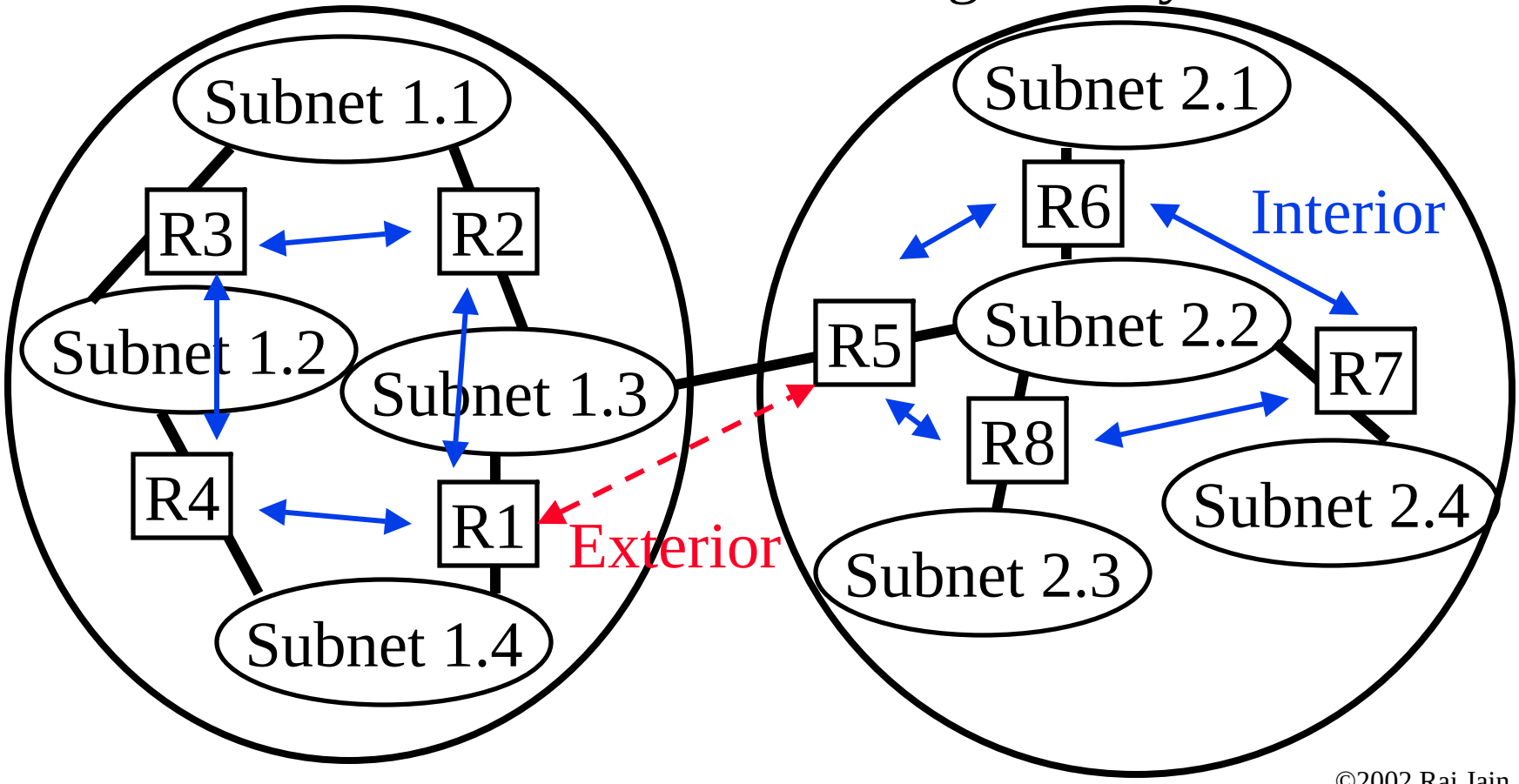


Name Resolution (Cont)

- Each computer has a name resolver routine, e.g., gethostbyname in UNIX
- Each resolver knows the name of a local DNS server
- Resolver sends a DNS request to the server
- DNS server either gives the answer, forwards the request to another server, or gives a referral
- Referral = Next server to whom request should be sent
- Servers respond to a full name only
  However, humans may specify only a partial name
  Resolvers may fill in the rest of the suffix, e.g.,
  Liberia.cis = Liberia.cis.ohio-state.edu

Autonomous Systems

- An internet connected by homogeneous routers under the administrative control of a single entity



Routing Protocols

- Interior Router Protocol (IRP): Among routers internal to an autonomous system.
  Also known as IGP.
  - Examples: Routing Information Protocol (RIP), Open Shortest Path First (OSPF)
- Exterior Router Protocol (ERP): Among routers between autonomous systems. Also known as EGP.
  - Examples: Exterior Gateway Protocol (EGP), Border Gateway Protocol (BGP), Inter-Domain Routing Protocol (IDRP)
    Note: EGP is a class as well as an instance in that class.

Routing Information Protocol

- RIP uses distance vector $\Rightarrow$ A vector of distances to all nodes is sent to neighbors every 30 seconds
- Each router computes new distances and replaces entries with new lower hop counts

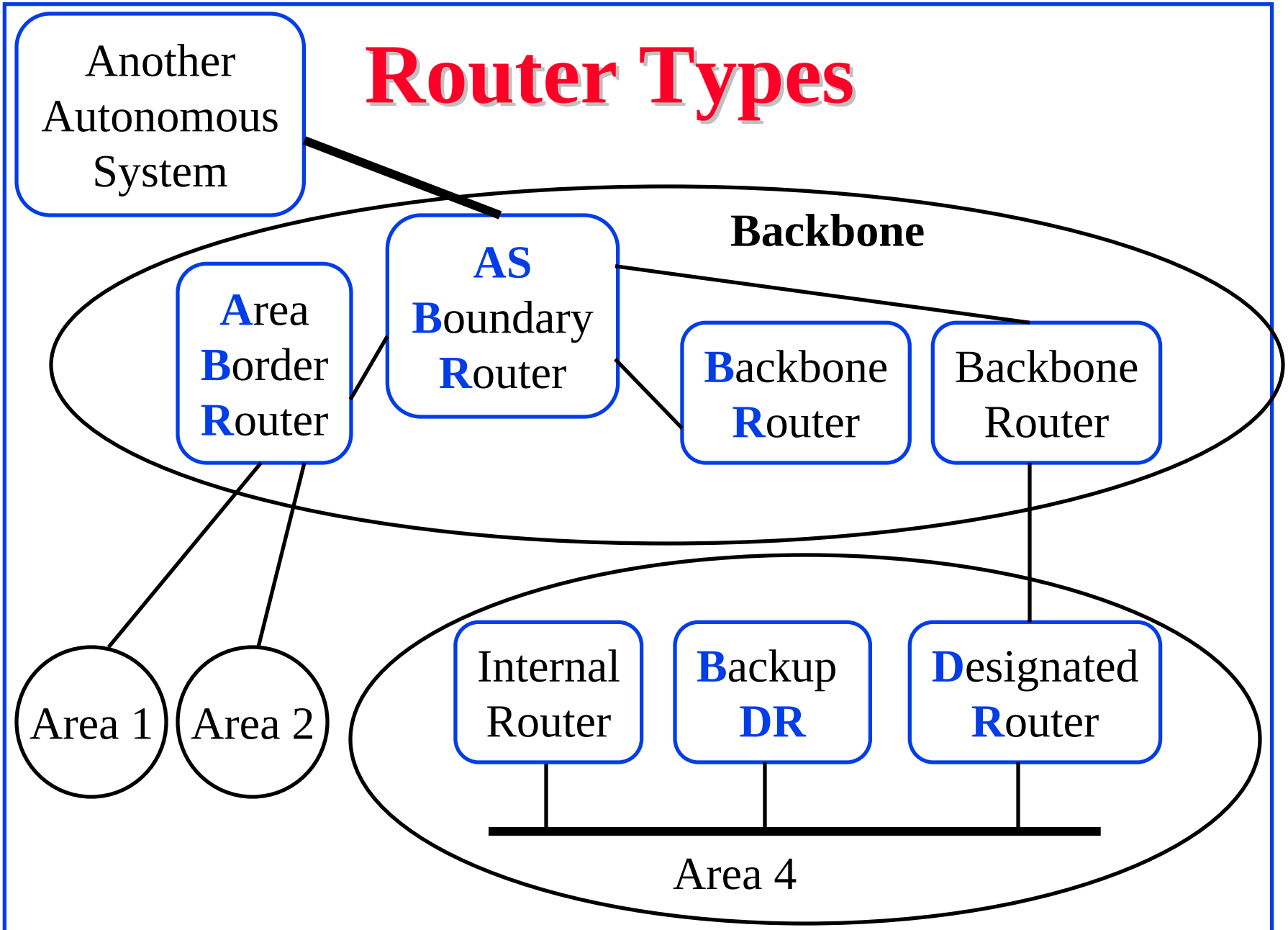
Shortcomings of RIP V1

- Maximum network diameter = 15 hops
- Cost is measured in hops
  Only shortest routes. May not be the fastest route.
- Entire tables are broadcast every 30 seconds.
  Bandwidth intensive.
- Uses UDP with 576-byte datagrams.
  Need multiple datagrams.
  300-entry table needs 12 datagrams.
- An error in one routing table is propagated to all routers
- Slow convergence

Open Shortest Path First (OSPF)

- Uses true metrics (not just hop count)
- Allows load balancing across equal-cost paths
- Authenticates route exchanges
- Quick convergence
- Large networks are subdivided into a backbone network and areas
- Each area has multiple subnets. Each subnet has a designated router.
- Link state routing $\Rightarrow$ Each router broadcasts its connectivity with neighbors to entire area

Router Types



Router Types (Cont)

- **Internal Router (IR)**: All interfaces belong to the same area
- **Area Border Router (ABR):** Interfaces to multiple areas
- **Backbone Router (BR):** Interfaces to the backbone
- **Autonomous System Boundary Router (ASBR):** Exchanges routing info with other autonomous systems
- **Designated Router (DR):** Generates link-state info about the subnet
- **Backup Designated Router (BDR):** Becomes DR if DR fails.

OSPF Operation

- Periodic "Hello" packets are multicast on the subnet to find other routers and elect "designated router" and "backup designated router"
- Designated routers and routers on point-to-point links form "adjacency." Exchange "Link State Advertisements (LSAs)." New info flooded to all other adjacent routers in the area.
- Area border routers (ABRs) send "summary LSAs" to other ABRs
- Autonomous system border routers (ASBRs) use exterior routing protocol to exchange routing information

Border Gateway Protocol

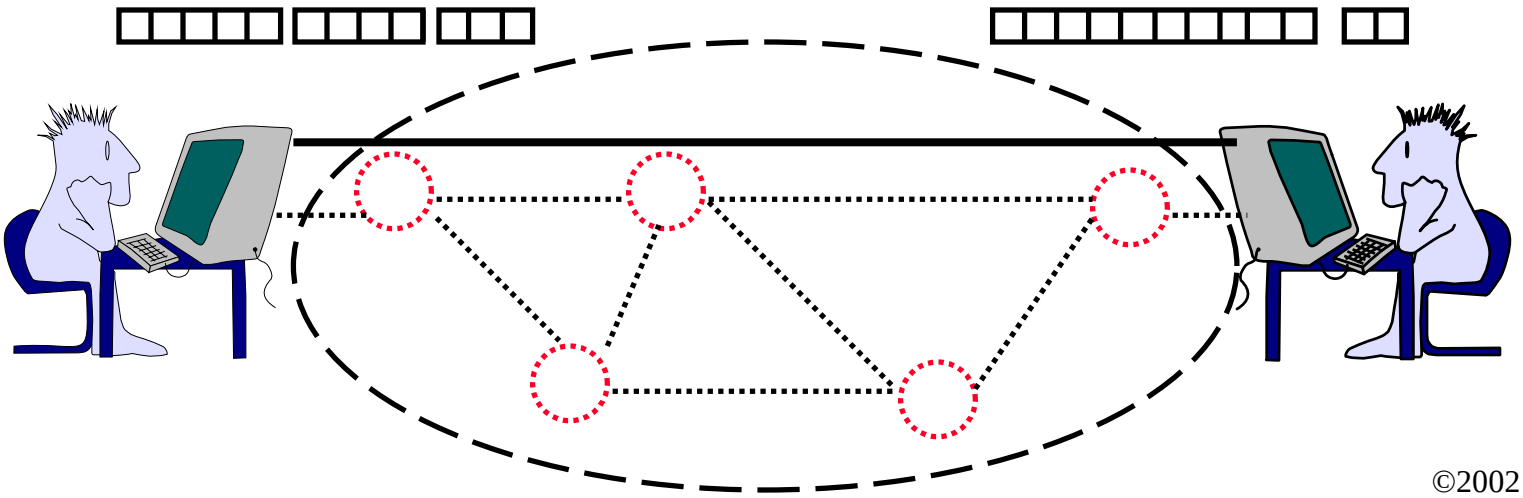
- Inter-autonomous system protocol [RFC 1267]
- Used since 1989 but not extensively until recently
- Advertises all transit ASs on the path to a destination address
- A router may receive multiple paths to a destination
  $\Rightarrow$ Can choose the best path

BGP Operations

- BGP systems initially exchange entire routing tables. Afterwards, only updates are exchanged.
- BGP messages have the following information:
  - Origin of path information: RIP, OSPF, …
  - AS_Path: List of ASs on the path to reach the dest
  - Next_Hop: IP address of the border router to be used as the next hop to reach the dest
  - Unreachable: If a previously advertised route has become unreachable
- BGP speakers generate update messages to all peers when it selects a new route or some route becomes unreachable.

TCP: Key Features

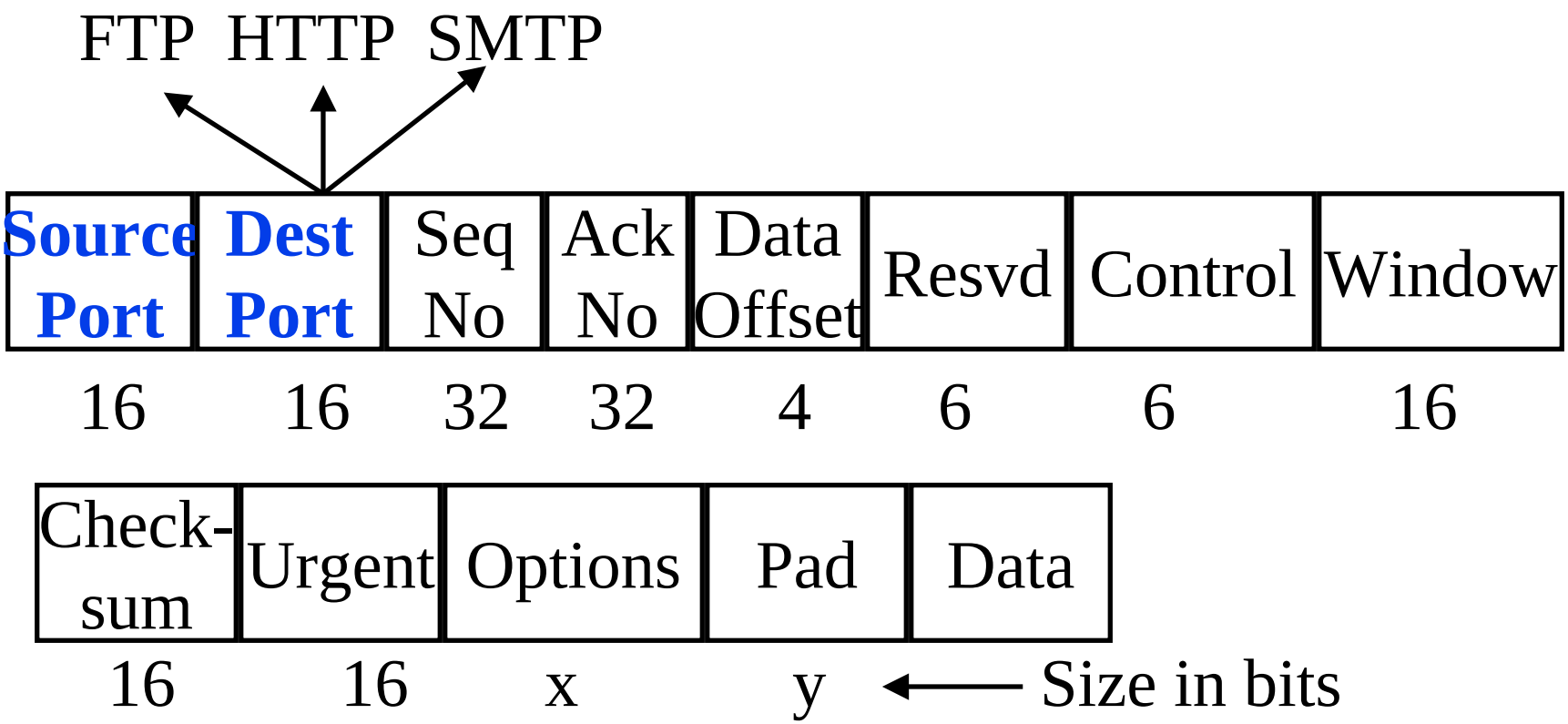
- Point-to-point communication: **Two** end-points
- **Connection** oriented. Full duplex communication.
- **Reliable** transfer: Data is delivered in order
  Lost packets are retransmitted.
- **Stream** interface: Continuous sequence of octets



Transport Control Protocol (TCP)

- Key Services:
  - Send: Please send when convenient
  - Data stream push: Please send it all now, if possible.
  - Urgent data signaling: Destination TCP! please give this urgent data to the user
    (Urgent data is delivered in sequence. Push at the should be explicit if needed.)
  - Note: Push has no effect on delivery.
    Urgent requests quick delivery

TCP Header Format

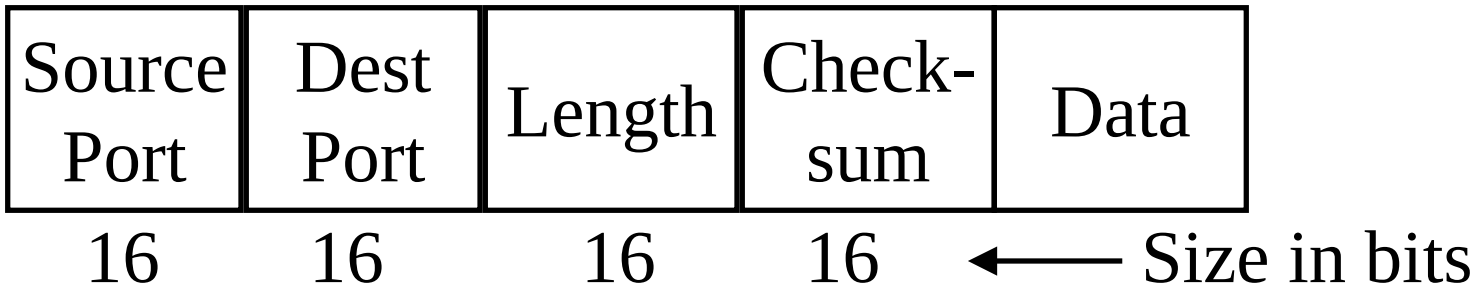


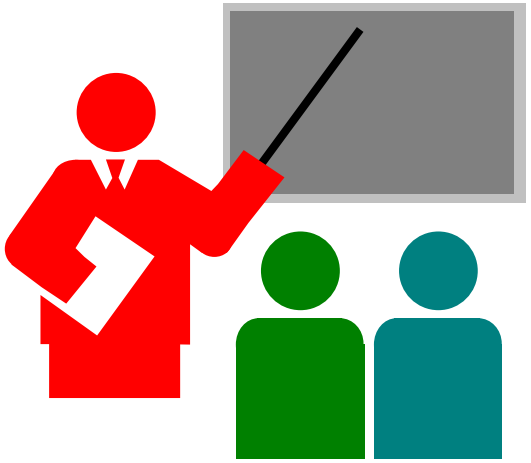
TCP Header

- Source Port (16 bits): Identifies source user process 20 = FTP, 23 = Telnet, 53 = DNS, 80 = HTTP, ...
- Destination Port (16 bits)
- Sequence Number (32 bits): Sequence number of the first byte in the segment.
- Ack number (32 bits): Next byte expected
- Data offset (4 bits): # of 32-bit words in the header
- Reserved (6 bits)
- Window (16 bits): Will accept [Ack] to [Ack]+[window]

User Datagram Protocol (UDP)

- **Connectionless** end-to-end service
- **Unreliable**: No flow control.
  No error recovery (No acks. No retransmissions.)
- Used by network management and Audio/Video.
- Provides port addressing
- Error detection (Checksum) optional.





Summary

- IP is the forwarding protocol between networks
- IPv4 uses 32-bit addresses
- DNS: Maps names to addresses
- OSPF uses link-state advertisements
- BGP is used between autonomous systems
- TCP provides reliable full-duplex connections.
- UDP is connectionless and simple. No flow/error control.

Homework 2

True or False?

T F

❑ ❑ A sample IP address is 10.0.110.357

❑ ❑ Two computers cannot have the same IP address

❑ ❑ Two computers cannot have the same complete DNS name

❑ ❑ IANA assigns all names used inside a company.

❑ ❑ Each DNS server database stores all computer names in the world.

❑ ❑ Routing tables used by IP are prepared using routing protocols like OSPF, BGP.

❑ ❑ RIP is used in small networks

❑ ❑ OSPF area border routers connect to other autonomous systems.

❑ ❑ OSPF Hellos are flooded through out the area.

❑ ❑ BGP is used between autonomous systems

❑ ❑ TCP delivers all packets to the destination exactly as received at the source.

❑ ❑ TCP port numbers are related to applications using them.

❑ ❑ UDP is unreliable transport protocol.

Marks = Correct Answers _____ - Incorrect Answers_____ = _____