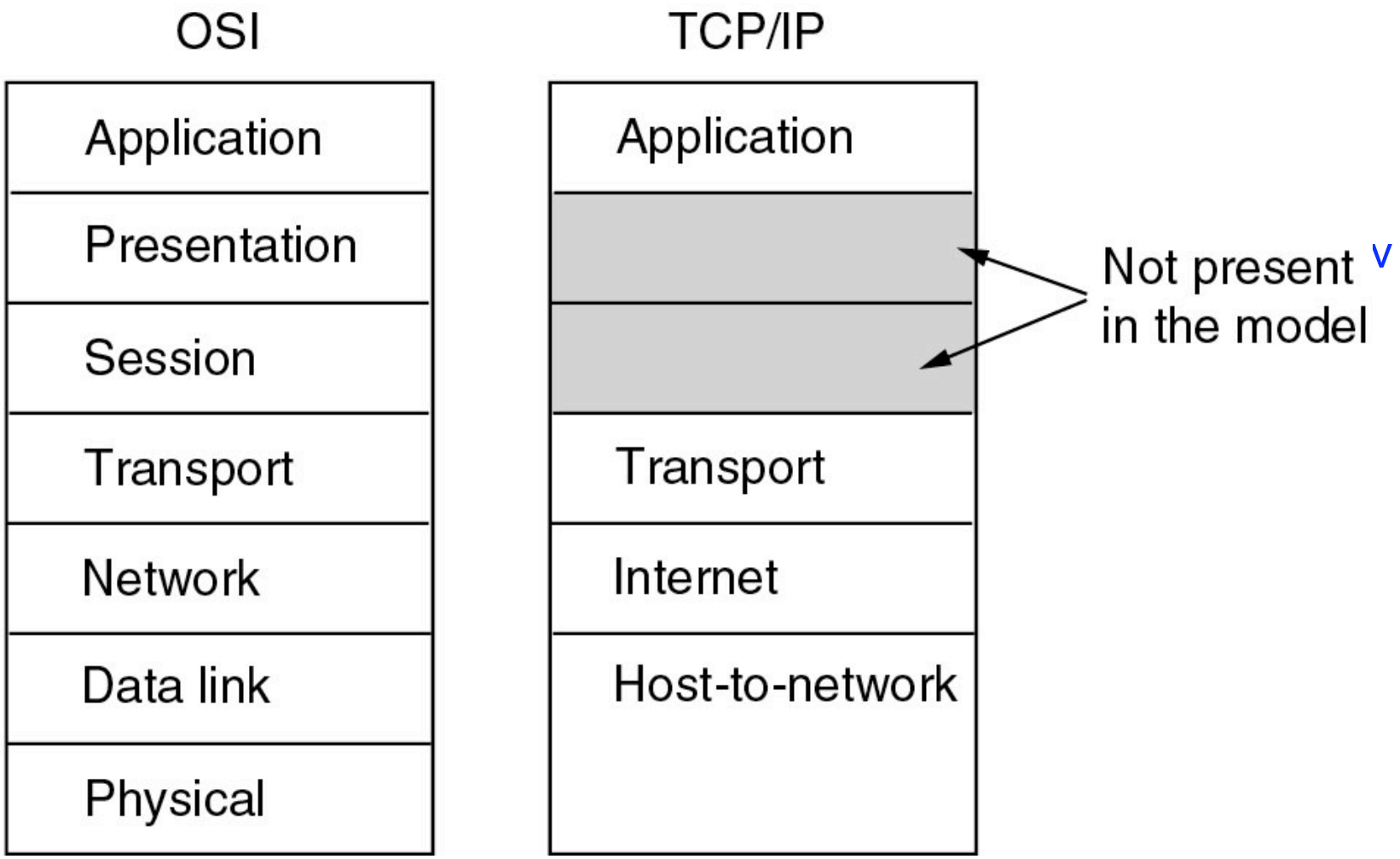
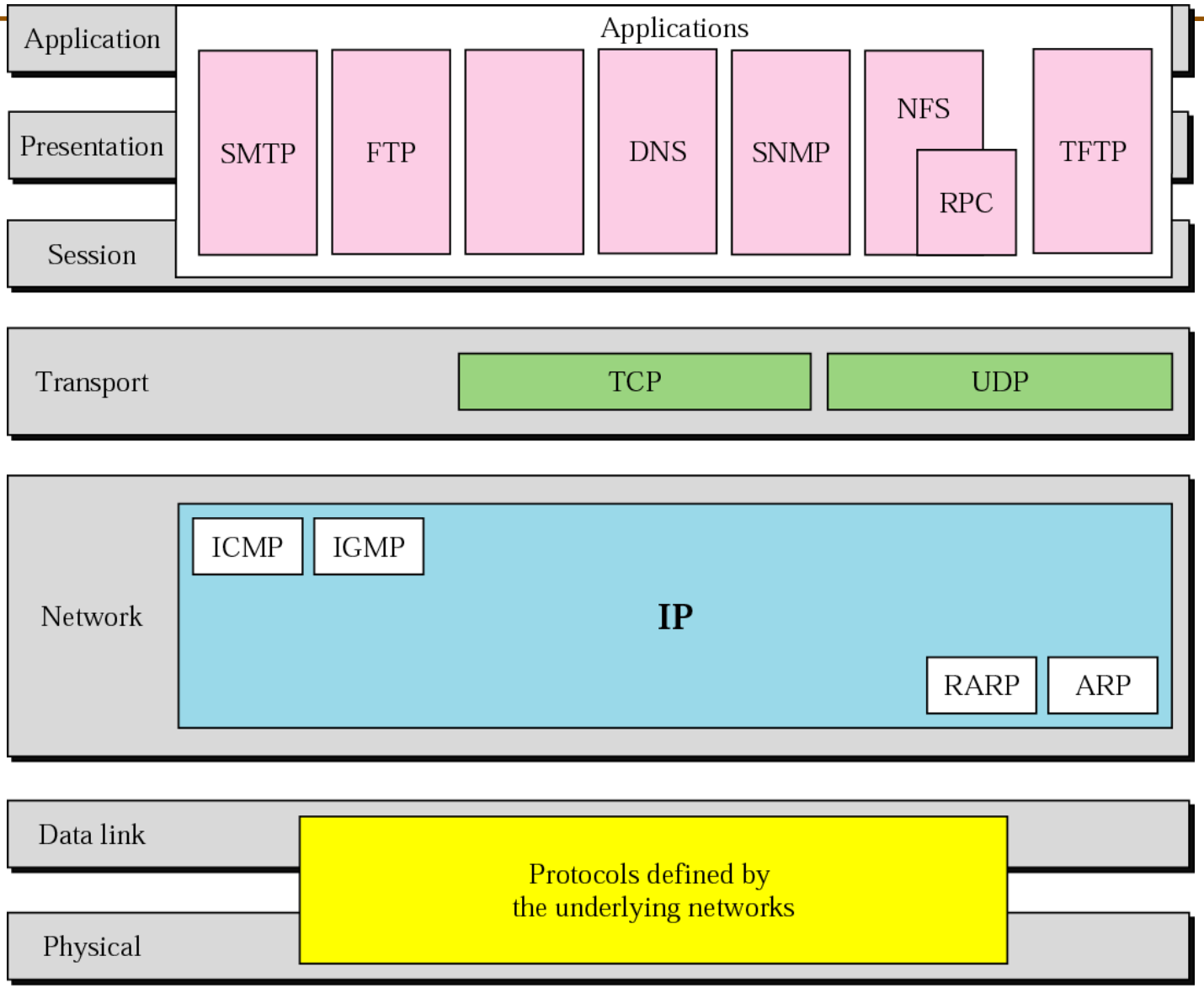
Overview of Internet Protocols

Διαδικτυακά Πρωτόκολλα

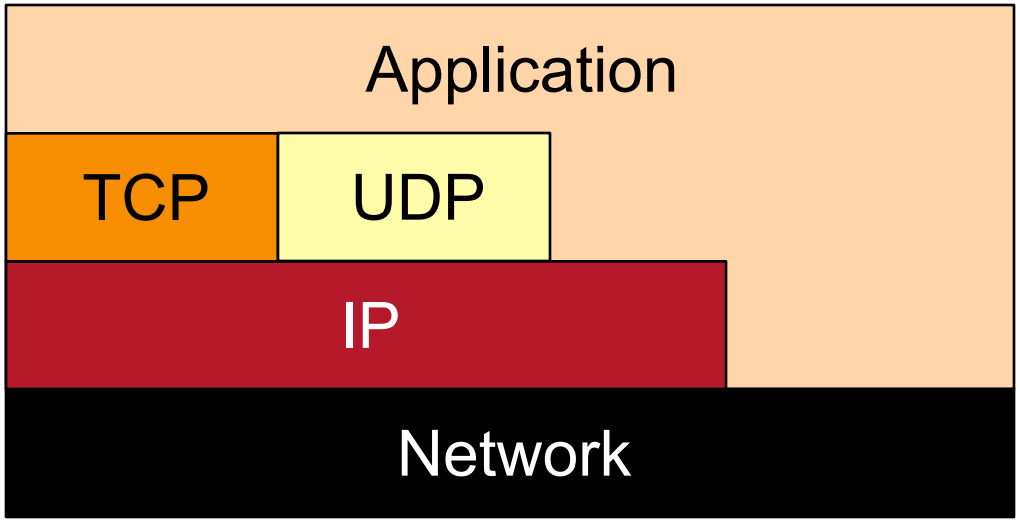
Διαστρωματώσεις Διαδικτύου



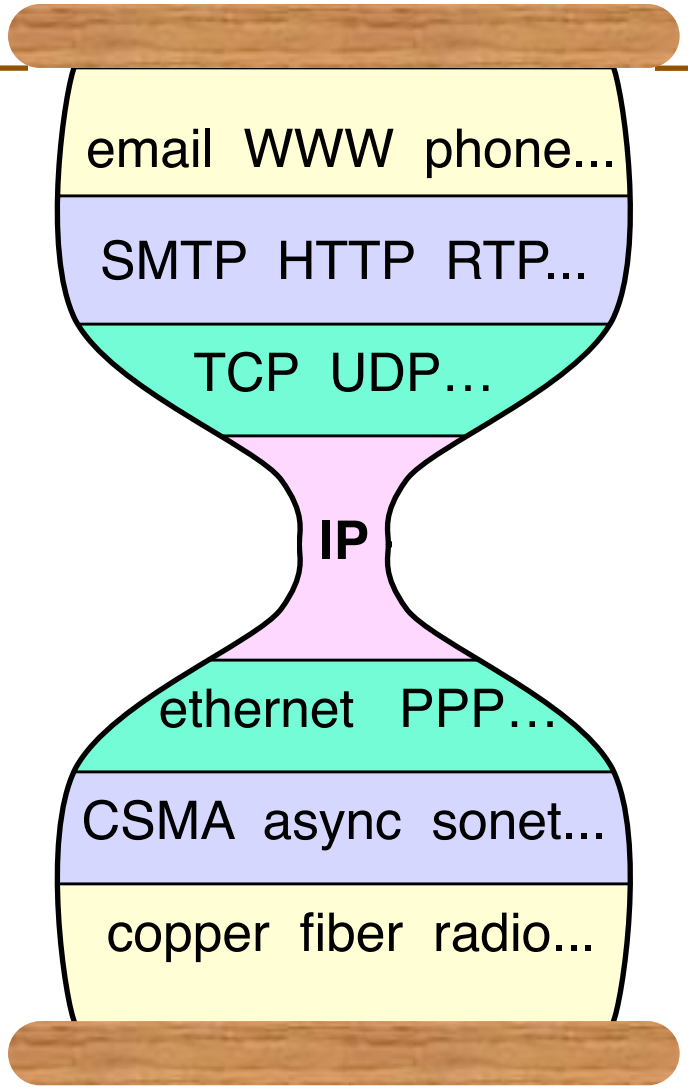
OSI και Διαδίκτυο



Διαδικτυακή στοίβα πρωτοκόλλων



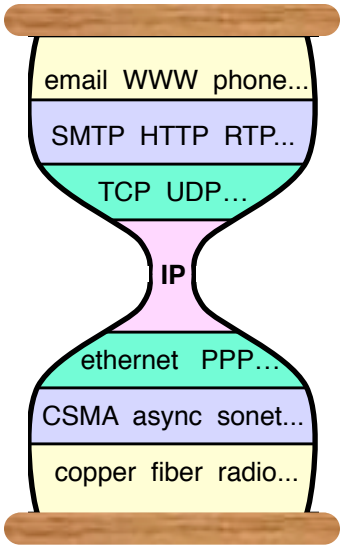
Αρχιτεκτονική Κλεψύδρας



Source: Hoffman & Beaumont

Why the Hourglass Architecture?

- Why an internet layer?
  - make a bigger network
  - global addressing
  - virtualize network to isolate end-to-end protocols from network details/changes
- Why a *single* internet protocol?
  - maximize interoperability
  - minimize number of service interfaces
- Why a *narrow* internet protocol?
  - assumes least common network functionality to maximize number of usable networks



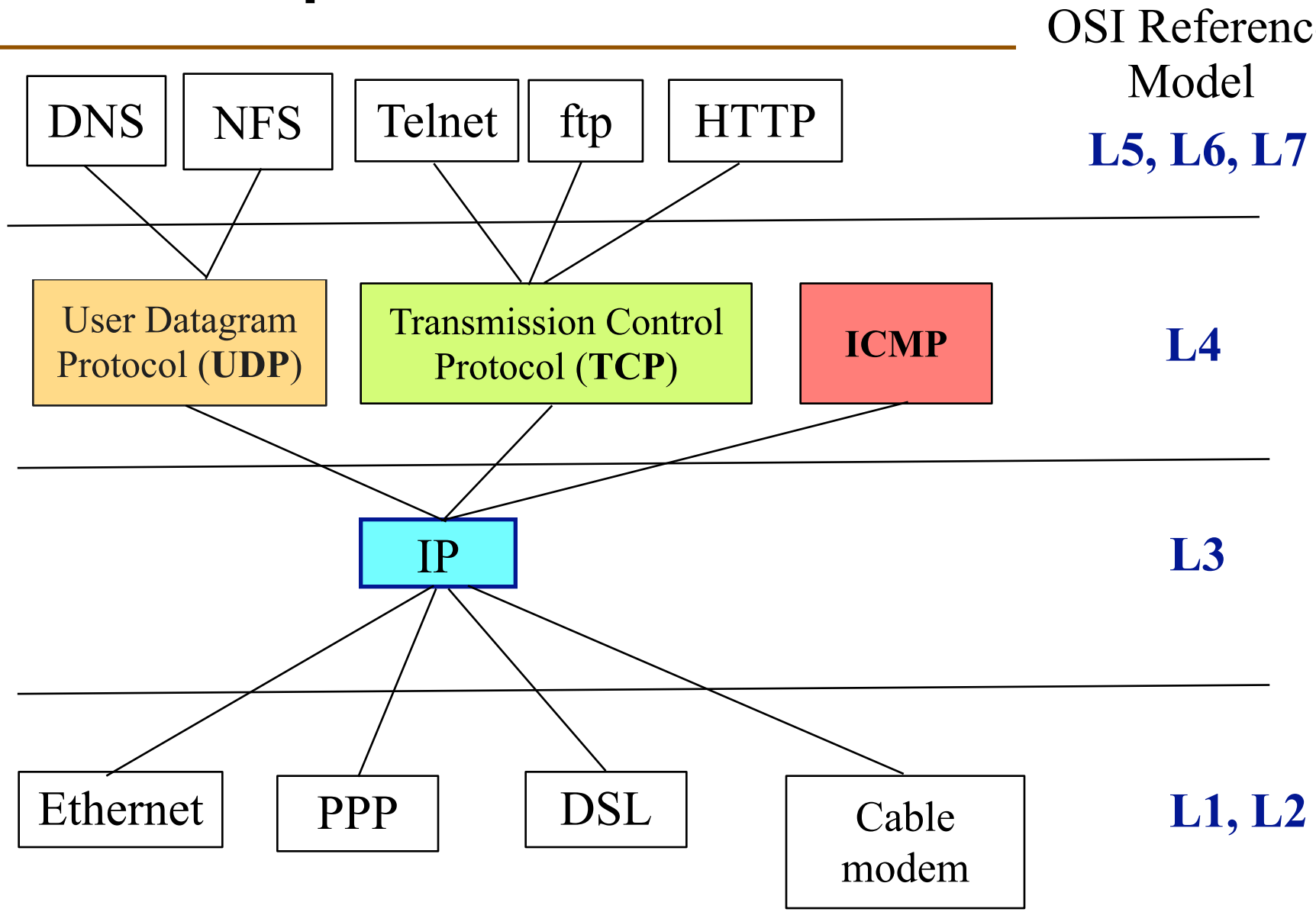
Source: Hoffman & Beaumont

Πρωτόκολλα Διαδικτύου

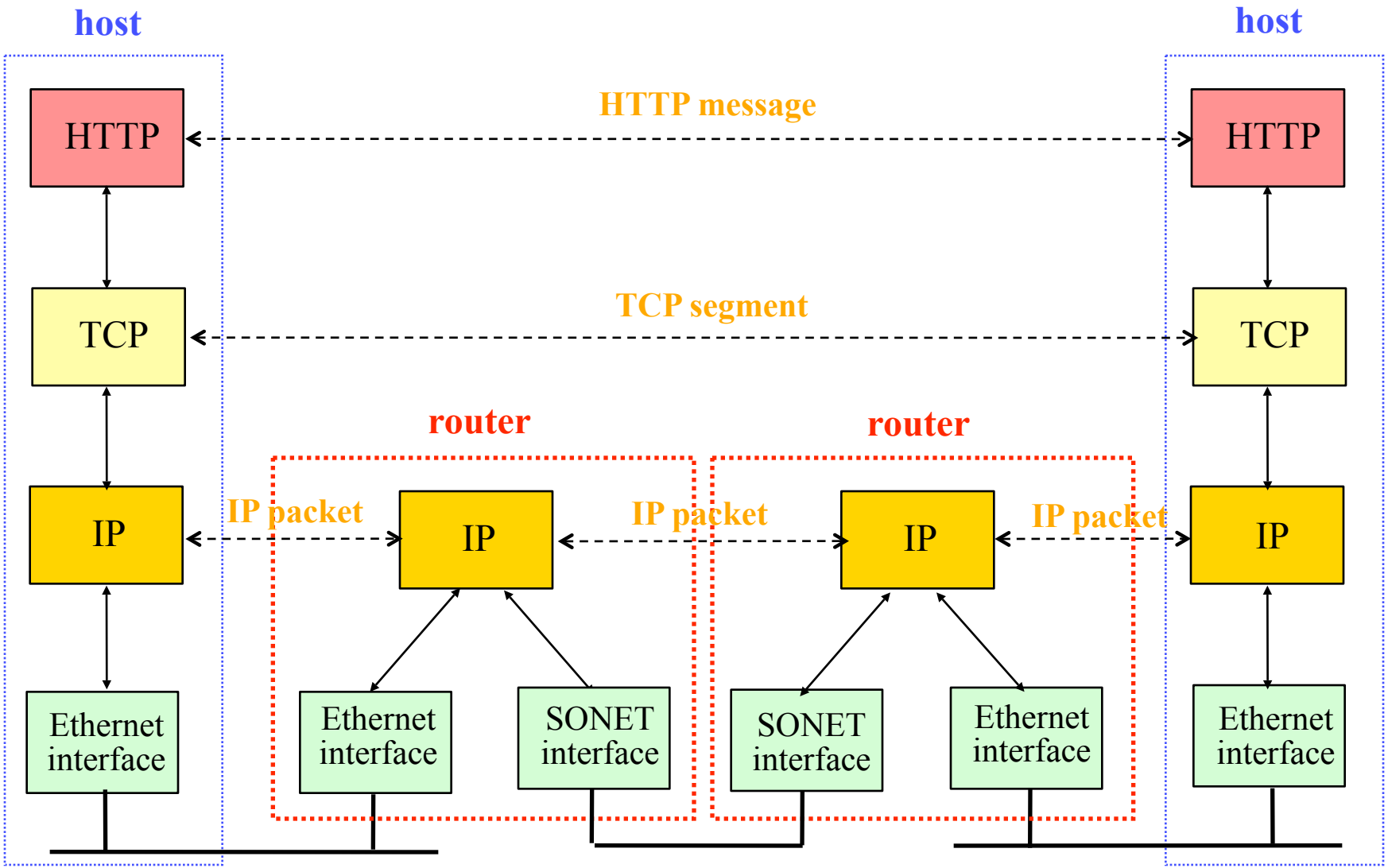
Το Διαδίκτυο βασίζεται:

- Σε δύο πρωτόκολλα μεταφοράς:
  - TCP (Transmission Control Protocol) – αξιόπιστο, προσανατολισμένο σε διατήρηση συνόδων (connection-oriented).
  - UDP (User Datagram Protocol) – πρωτόκολλο πακέτων (datagram protocol) το οποίο δεν εγγυάται αξιόπιστη μετάδοση.
- Στο πρωτόκολλο IP, το οποίο αποτελεί το δικτυακό «υπόστρωμα» - τα datagrams του IP είναι ο βασικός μηχανισμός μετάδοσης για το Διαδίκτυο και άλλα δίκτυα TCP/IP.
- Η επιτυχία του TCP/IP οφείλεται σε μεγάλο βαθμό στην ανεξαρτησία του από τις τεχνολογίες μετάδοσης.

Common protocols in the Internet



IP Suite: End Hosts vs. Routers



Βασικά Χαρακτηριστικά TCP/IP

- The popularity of TCP/IP is due to a number of important features:
  - Open protocol standards: freely available and developed independently from any computer hardware or OS.
  - Independence from specific physical network hardware.
  - A common addressing scheme.
  - Standardized high-level protocols.
- TCP/IP standards and protocols are published publicly as *Requests for Comments* (RFCs).
  - Where??

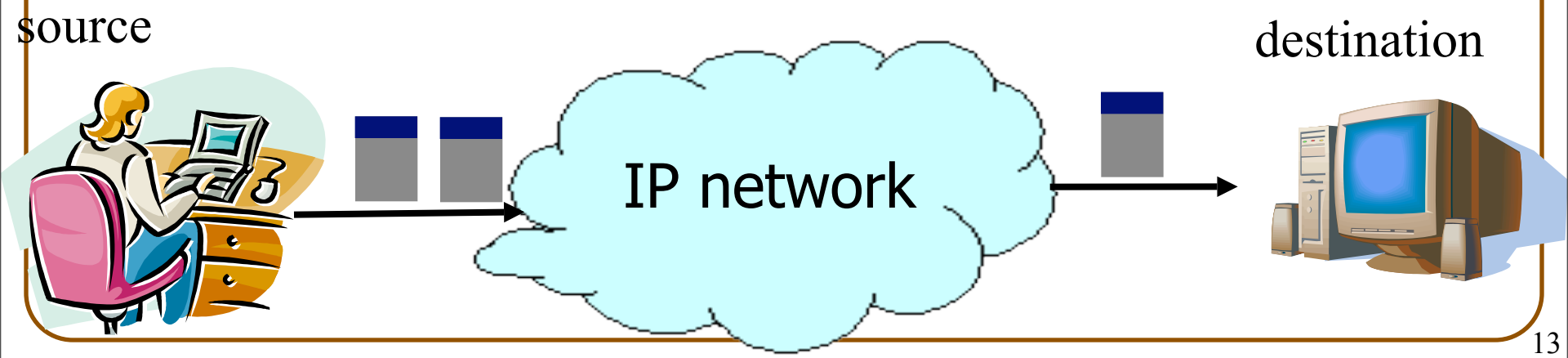
Ανασκόπηση του IP

Διαδικτύωση

- Για την δημιουργία ενός διαδικτύου πρέπει να συναρμολογήσουμε πολλά υποδίκτυα (subnets) τα οποία πιθανότατα βασίζονται σε διαφορετικές τεχνολογίες (Ethernet, ATM, sat, ISDN, DSL). Για το σκοπό αυτό χρειαζόμαστε:
  - Ένα ενοποιημένο σχήμα διευθυνσιοδότησης, το οποίο να επιτρέπει σε πακέτα να στέλνονται σε οποιονδήποτε κόμβο συνδεδεμένο σε οποιοδήποτε υποδίκτυο: IP addresses.
  - Ένα πρωτόκολλο που να προσδιορίζει τον μορφότυπο των πακέτων διαδικτύωσης και τους κανόνες διαχείρισής τους: IP Protocol.
  - Συστατικά διαδικτύωσης τα οποία να δρομολογούν πακέτα προς τις διευθύνσεις των παραληπτών τους, εκπέμποντας πακέτα με βάση την τεχνολογία των διασυνδεδεμένων υποδικτύων: Internet routers.

IP Service: Best-Effort Packet Delivery

- Packet switching
  - Divide messages into a sequence of packets
  - Headers with source and destination address
- Best-effort delivery
  - Packets may be lost
  - Packets may be corrupted
  - Packets may be delivered out of order



IP Service Model: Why Packets?

- Data traffic is bursty
  - Logging in to remote machines
  - Exchanging e-mail messages
- Don't want to waste reserved bandwidth
  - No traffic exchanged during idle periods
- Better to allow multiplexing
  - Different transfers share access to same links
- Packets can be delivered by almost anything
  - RFC 2549: IP over Avian Carriers (aka birds)
- … still, packet switching can be inefficient
  - Extra header bits on every packet

IP Service Model: Why Best-Effort?

- IP means never having to say you're sorry…
  - Don't need to reserve bandwidth and memory
  - Don't need to do error detection & correction
  - Don't need to remember from one packet to next
- Easier to survive failures
  - Transient disruptions are okay during fail-over
- … but, applications *do* want efficient, accurate transfer of data in order, in a timely fashion

IP Service: Best-Effort is Enough?

- No error detection or correction
  - Higher-level protocol can provide error checking
- Successive packets may not follow the same path
  - Not a problem as long as packets reach the destination
- Packets can be delivered out-of-order
  - Receiver can put packets back in order (if necessary)
- Packets may be lost or arbitrarily delayed
  - Sender can send the packets again (if desired)
- No network congestion control (beyond “drop”)
  - Sender can slow down in response to loss or delay

Other Main Driving Goals (In Order)

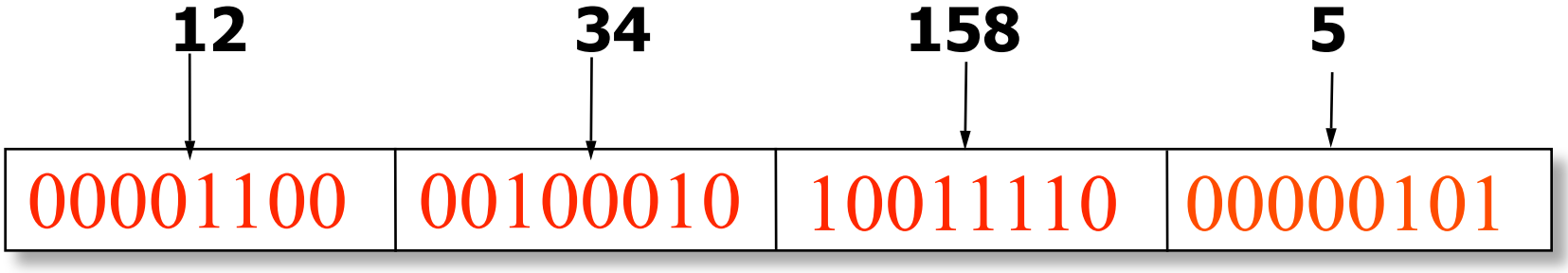
- Communication should continue despite failures
  - Survive equipment failure or physical attack
  - Traffic between two hosts continue on another path
- Support multiple types of communication services
  - Differing requirements for speed, latency, & reliability
  - Bidirectional reliable delivery vs. message service
- Accommodate a variety of networks
  - Both military and commercial facilities
  - Minimize assumptions about the underlying network

Other Driving Goals, Somewhat Met

- Permit distributed management of resources
  - Nodes managed by different institutions
  - … though this is still rather challenging
- Cost-effectiveness
  - Statistical multiplexing through packet switching
  - … though packet headers and retransmissions wasteful
- Ease of attaching new hosts
  - Standard implementations of end-host protocols
  - … though still need a fair amount of end-host software
- Accountability for use of resources
  - Monitoring functions in the nodes
  - … though this is still fairly limited and immature

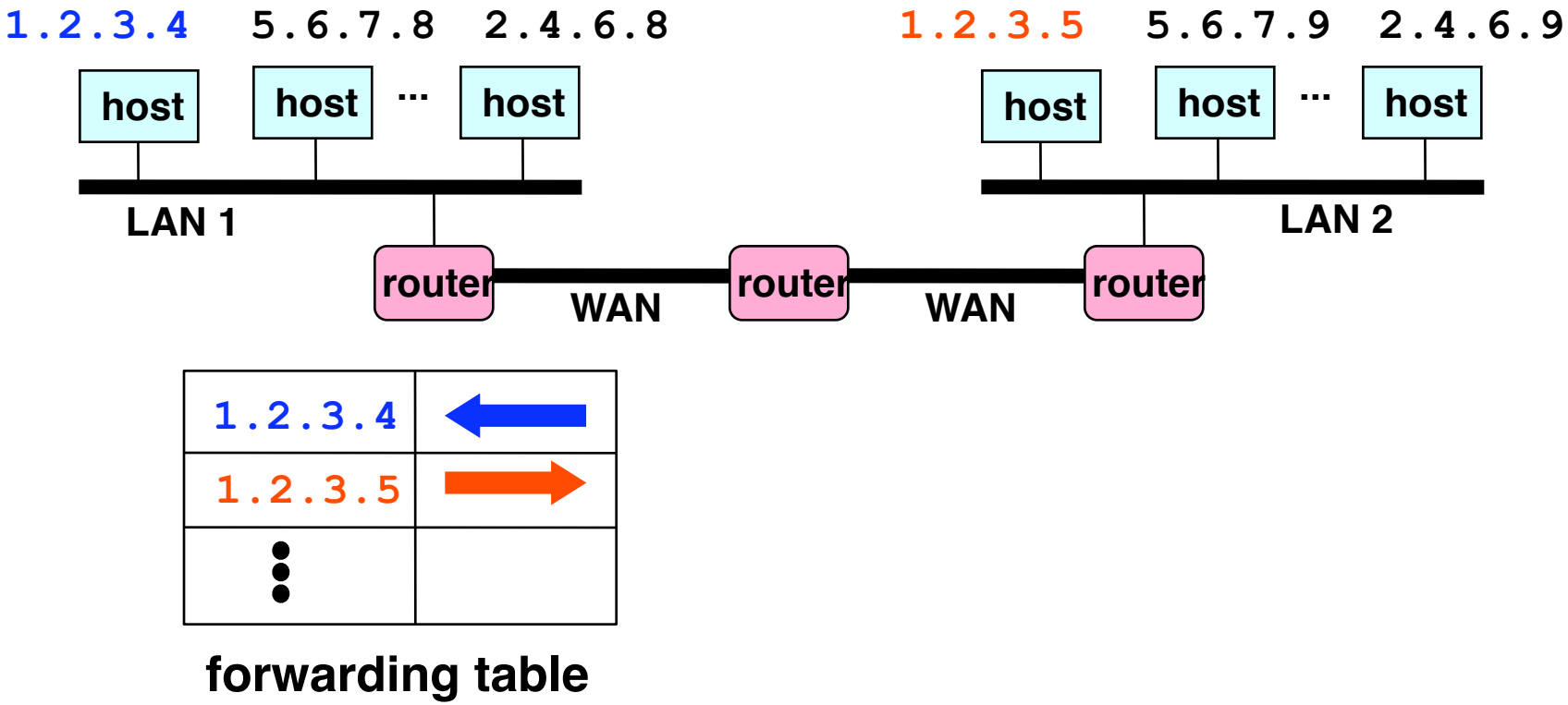
Διευθυνσιοδότηση σε Δίκτυα IP

- Για τη δημιουργία ενός παγκόσμιου δικτύου, είναι αναγκαία η ταυτοποίηση κάθε κόμβου του δικτύου.
- Το IP επιτυγχάνει την ταυτοποίηση με την ανάθεση μιας μοναδικής IP διεύθυνσης σε κάθε κόμβο:
  - A unique 32-bit number
  - Identifies an interface (on a host, on a router, …)
  - Represented in dotted-quad notation
- Η επικεφαλίδα ενός πακέτου IP περιλαμβάνει όλη την πληροφορία που χρειάζονται οι δρομολογητές για να παραδώσουν το πακέτο στον κατάλληλο παραλήπτη, υπό μορφή της IP διεύθυνσης



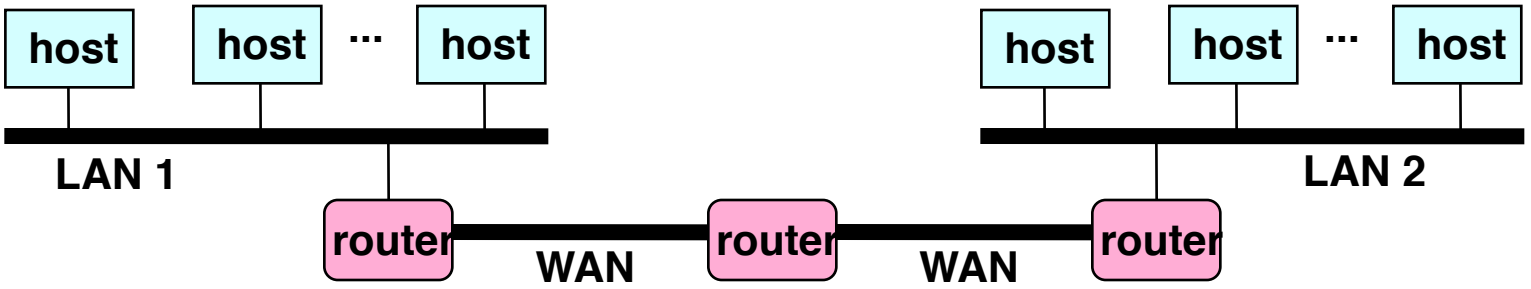
Scalability Challenge

- Suppose hosts had arbitrary addresses
  - Then every router would need a lot of information
  - …to know how to direct packets toward the host



Grouping Related Hosts

- The Internet is an "inter-network"
  - Used to connect *networks* together, not *hosts*
  - Needs a way to address a network (i.e., group of hosts)



LAN = Local Area Network
WAN = Wide Area Network

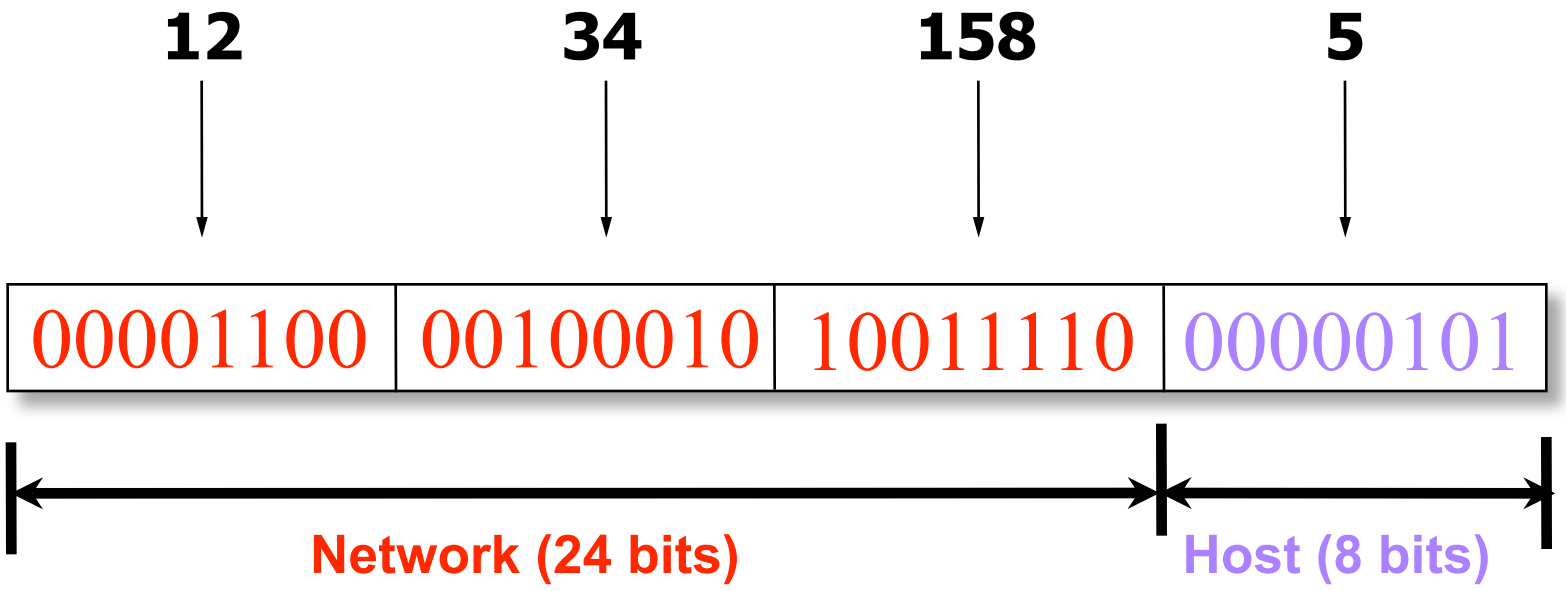
Hierarchical Addressing in U.S. Mail

- Addressing in the U.S. mail
  - Zip code: 08540
  - Street: Olden Street
  - Building on street: 35
  - Room in building: 306
  - Name of occupant: Jennifer Rexford

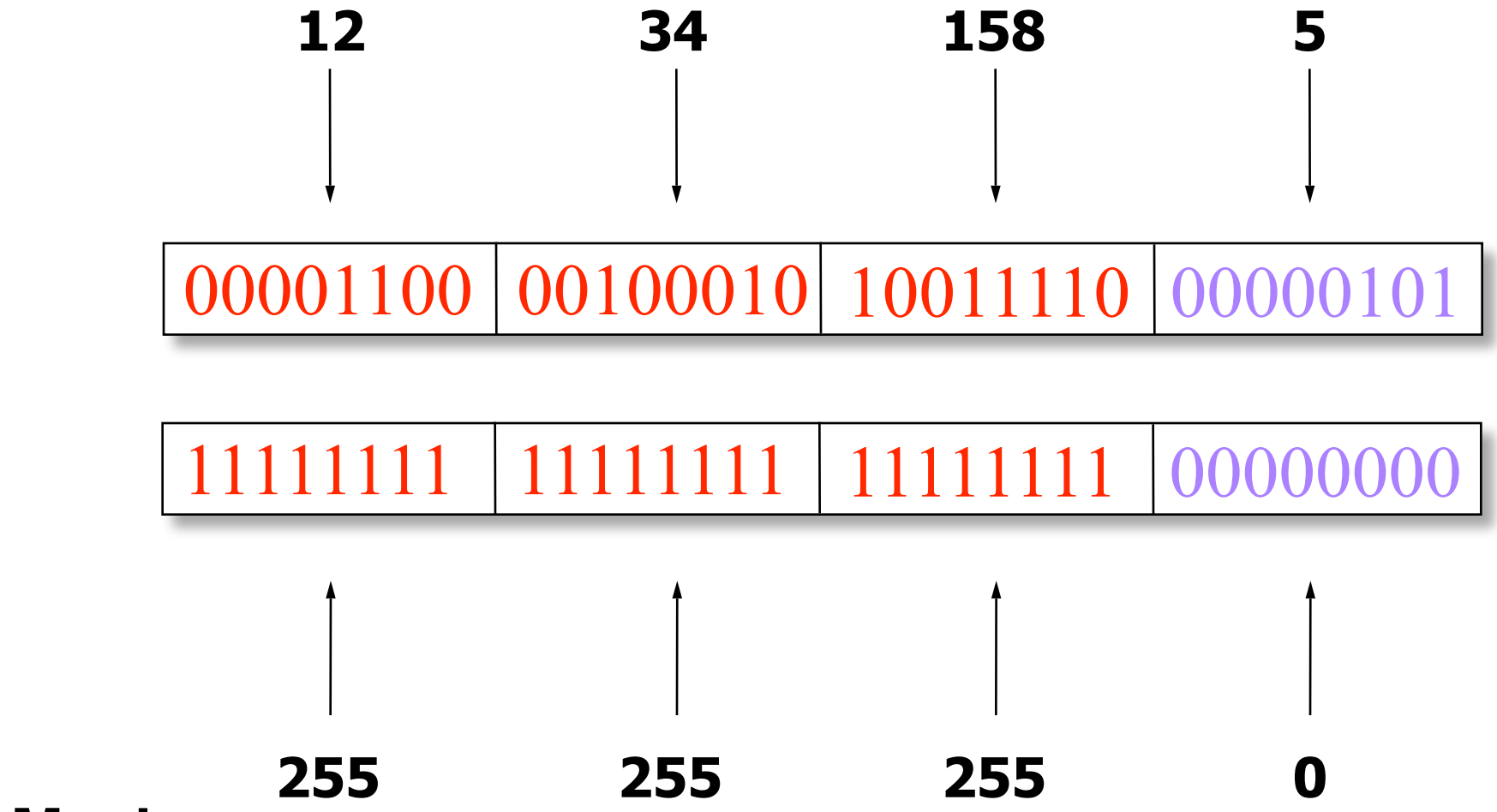
- Forwarding the U.S. mail
  - Deliver letter to the post office in the zip code
  - Assign letter to mailman covering the street
  - Drop letter into mailbox for the building/room
  - Give letter to the appropriate person

Hierarchical Addressing: IP Prefixes

- Divided into network & host portions (left and right)
- 12.34.158.0/24 is a 24-bit prefix with 2^8 addresses

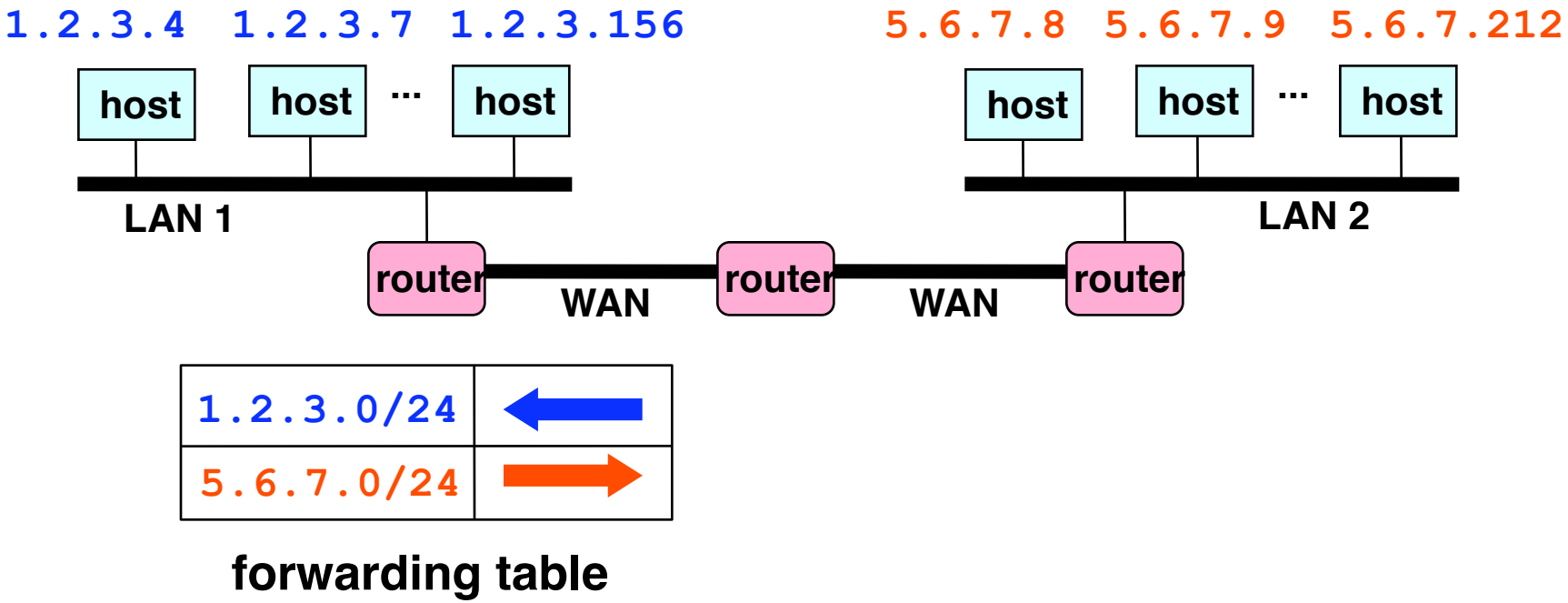


IP Address and a 24-bit Subnet Mask



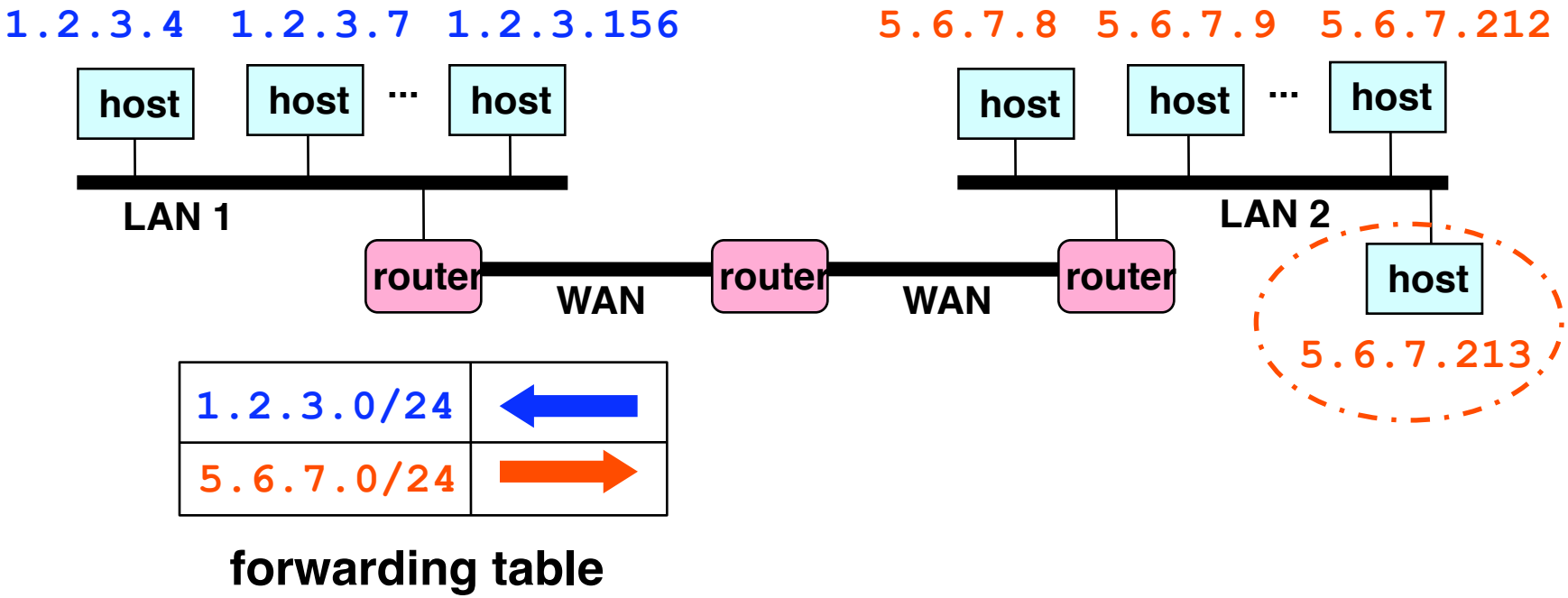
Scalability Improved

- Number related hosts from a common subnet
  - 1.2.3.0/24 on the left LAN
  - 5.6.7.0/24 on the right LAN



Easy to Add New Hosts

- No need to update the routers
  - E.g., adding a new host 5.6.7.213 on the right
  - Doesn't require adding a new forwarding entry



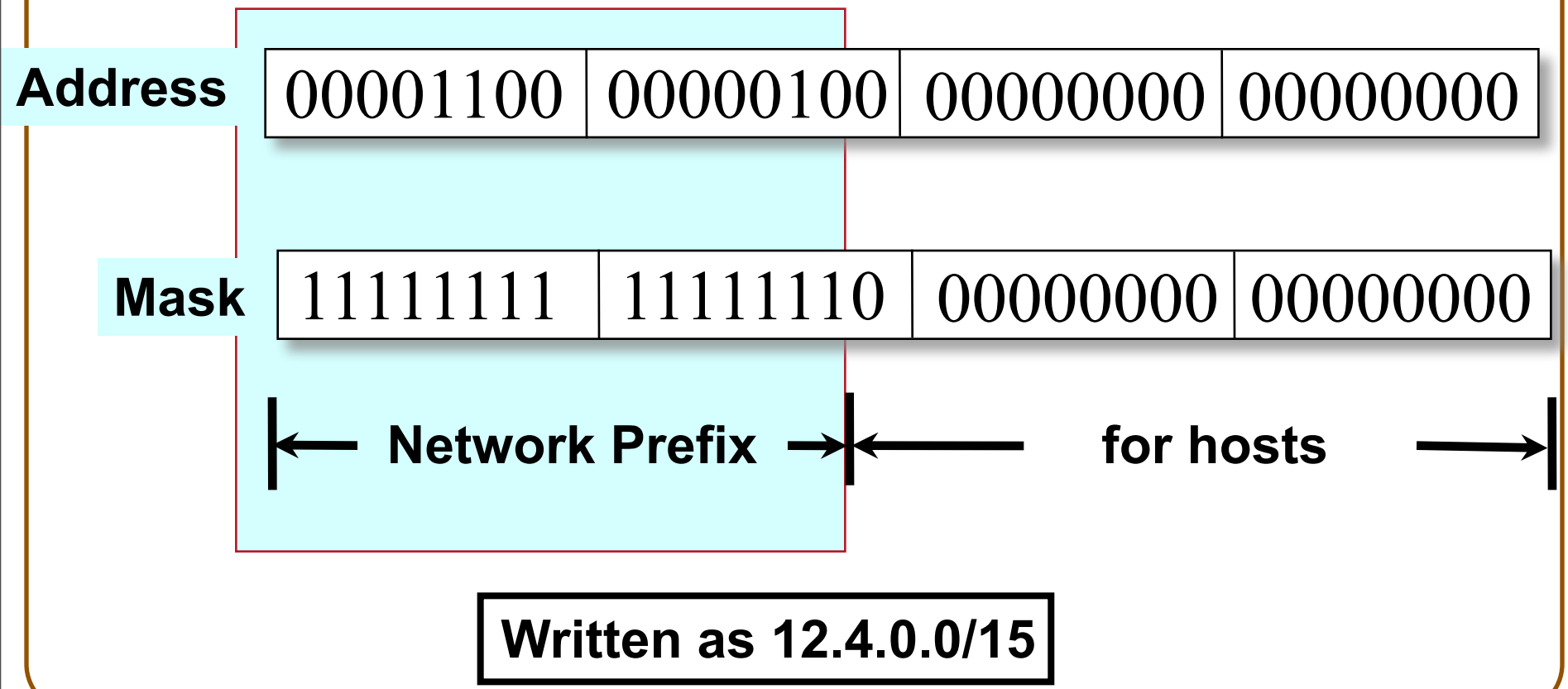
Classful Addressing

- In the older days, only fixed allocation sizes
  - Class A: 0*
    - Very large /8 blocks (e.g., MIT has 18.0.0.0/8)
  - Class B: 10*
    - Large /16 blocks (e.g,. Princeton has 128.112.0.0/16)
  - Class C: 110*
    - Small /24 blocks (e.g., AT&T Labs has 192.20.225.0/24)
  - Class D: 1110*
    - Multicast groups
  - Class E: 11110*
    - Reserved for future use
- This is why folks use dotted-quad notation!

Classless Inter-Domain Routing (CIDR)

Use two 32-bit numbers to represent a network.
Network number = IP address + Mask

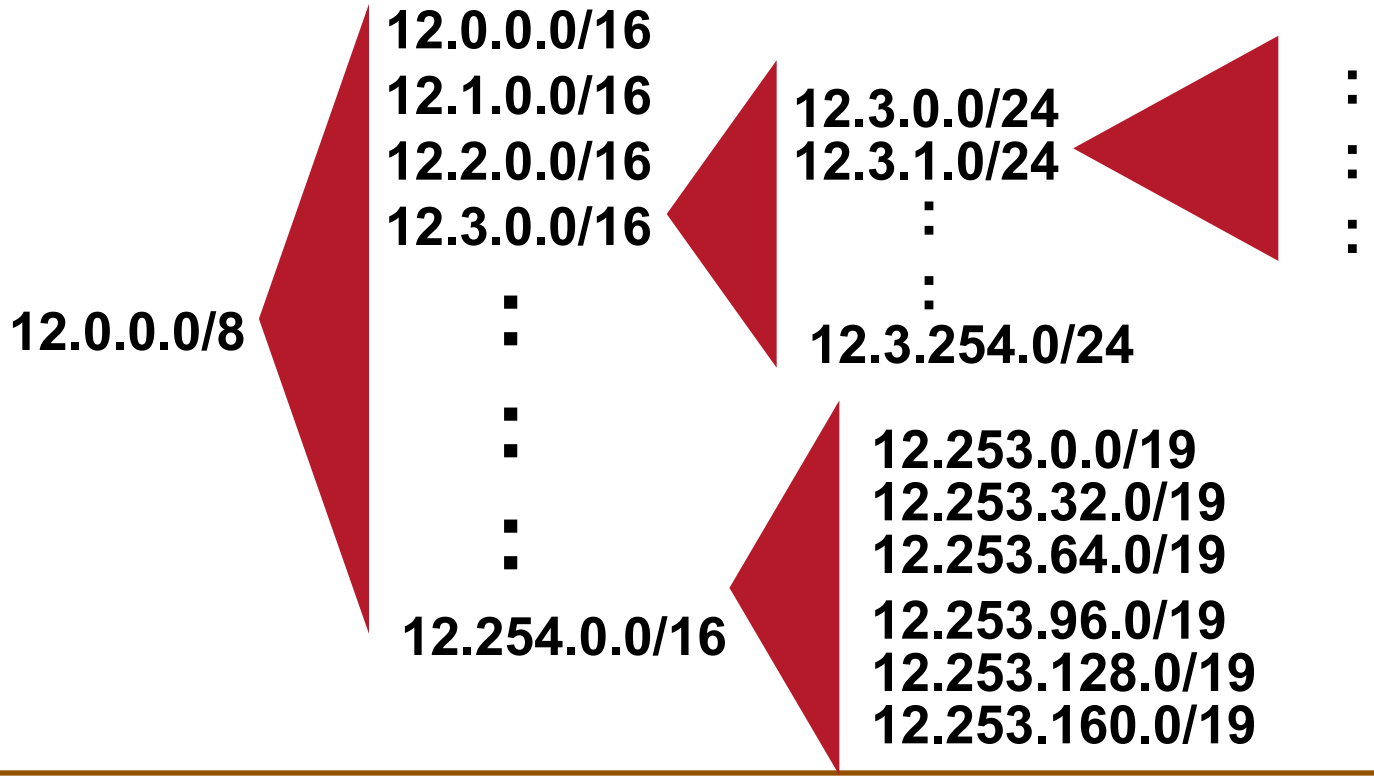
IP Address : 12.4.0.0 **IP Mask: 255.254.0.0**



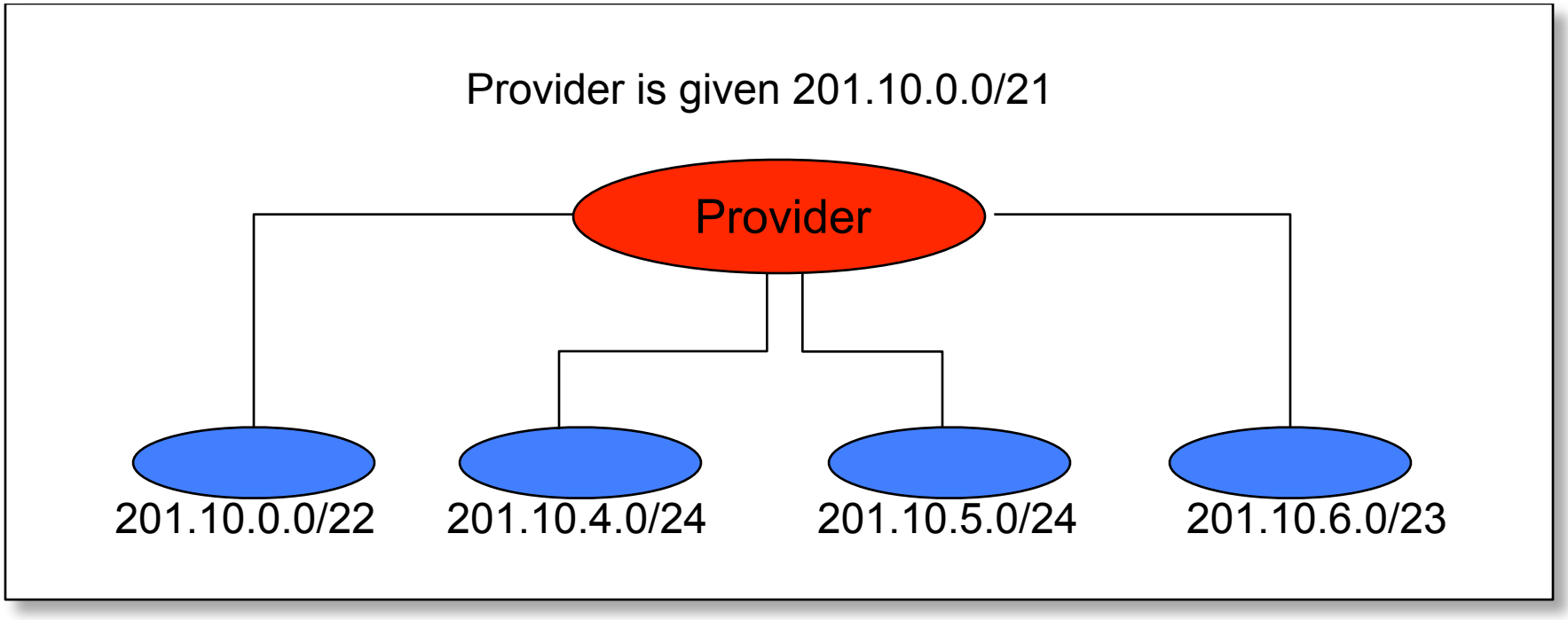
Written as 12.4.0.0/15

CIDR: Hierarchal Address Allocation

- Prefixes are key to Internet scalability
  - Address allocated in contiguous chunks (prefixes)
  - Routing protocols and packet forwarding based on prefixes
  - Today, routing tables contain ~150,000-200,000 prefixes

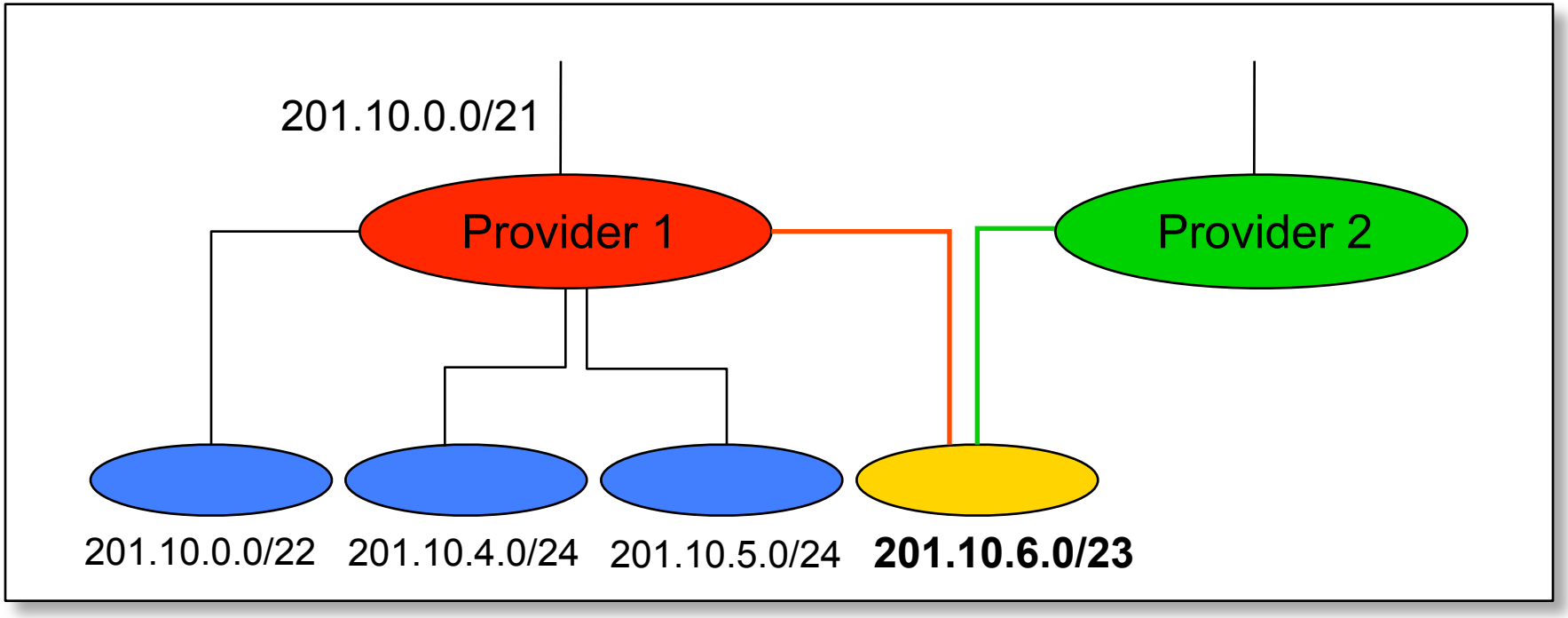


Scalability: Address Aggregation



Routers in the rest of the Internet just need to know how to reach 201.10.0.0/21. The provider can direct the IP packets to the appropriate customer.

But, Aggregation Not Always Possible

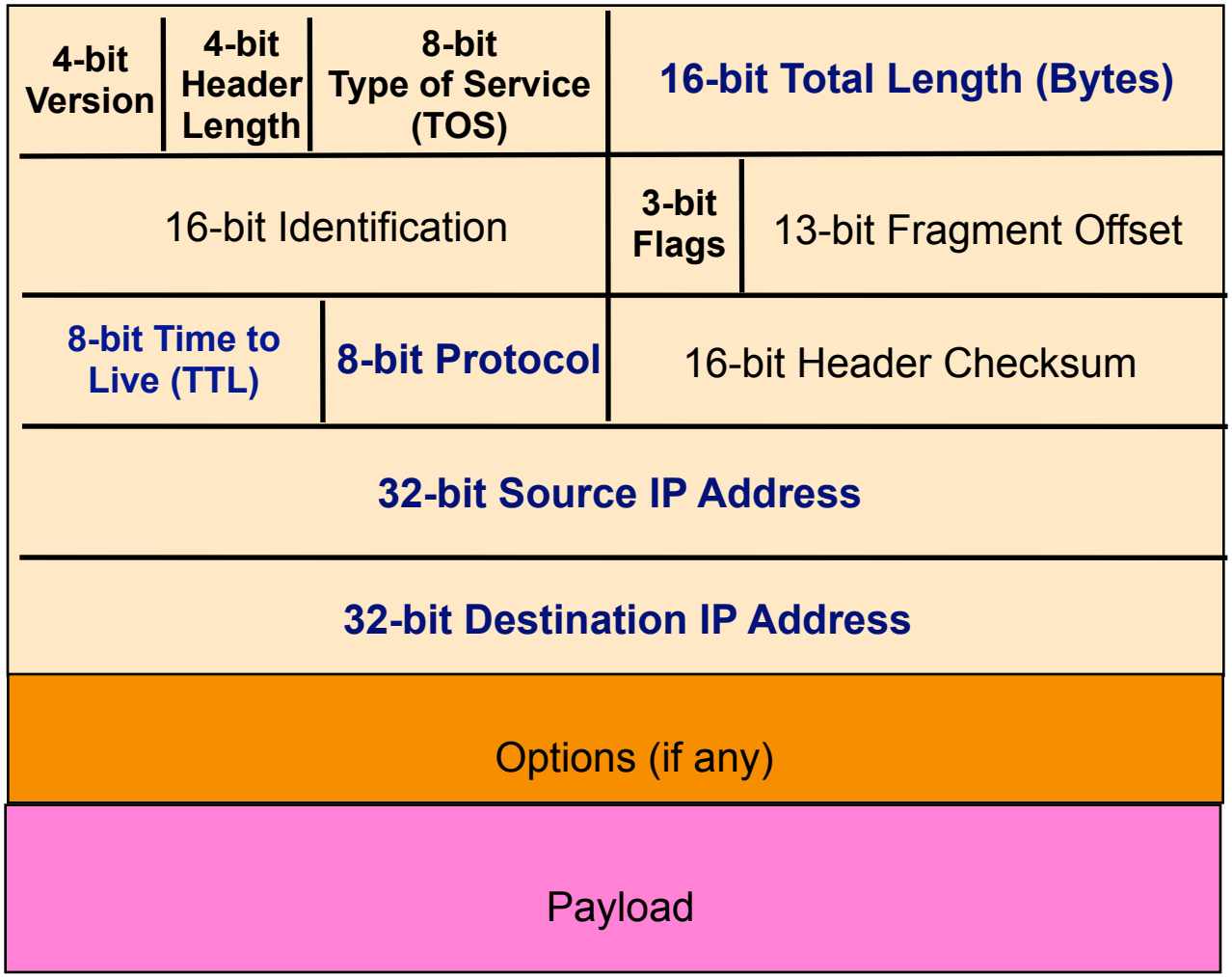


Multi-homed **customer with 201.10.6.0/23 has two providers. Other parts of the Internet need to know how to reach these destinations through *both* providers.**

Obtaining a Block of Addresses

- Separation of control
  - Prefix: assigned to an institution
  - Addresses: assigned to nodes by the institution
- Who assigns prefixes?
  - Internet Corp. for Assigned Names and Numbers
    - Allocates large blocks to Regional Internet Registries
  - Regional Internet Registries (RIRs)
    - E.g., ARIN (American Registry for Internet Numbers)
    - Allocated to ISPs and large institutions in a region
  - Internet Service Providers (ISPs)
    - Allocate address blocks to their customers
    - Who may, in turn, allocate to their customers…

IP Packet Structure



IP Packet Header Fields

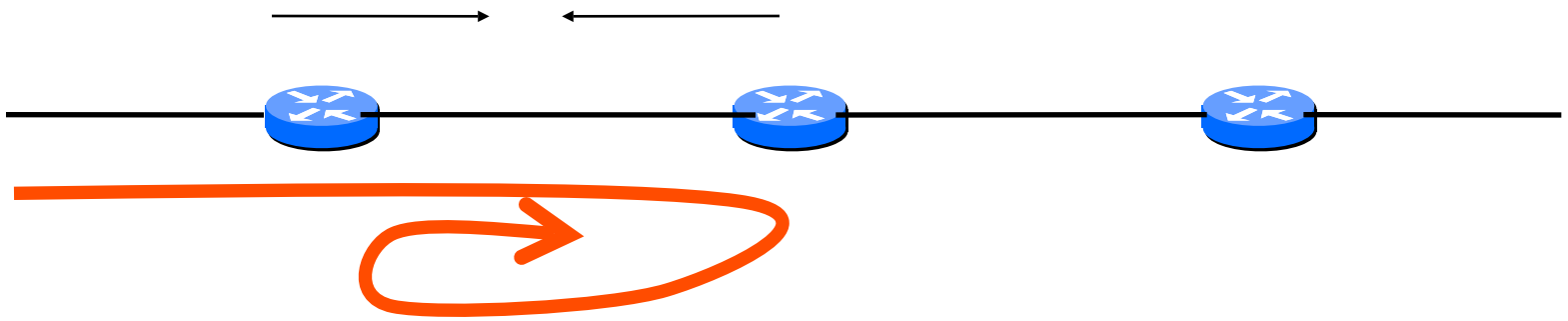
- Version number (4 bits)
  - Indicates the version of the IP protocol
  - Necessary to know what other fields to expect
  - Typically “4” (for IPv4), and sometimes “6” (for IPv6)
- Header length (4 bits)
  - Number of 32-bit words in the header
  - Typically “5” (for a 20-byte IPv4 header)
  - Can be more when “IP options” are used
- Type-of-Service (8 bits)
  - Allow packets to be treated differently based on needs
  - E.g., low delay for audio, high bandwidth for bulk transfer

IP Packet Header Fields (Continued)

- Total length (16 bits)
  - Number of bytes in the packet
  - Maximum size is 63,535 bytes (2^{16} -1)
  - … though underlying links may impose harder limits
- Fragmentation information (32 bits)
  - Packet identifier, flags, and fragment offset
  - Supports dividing a large IP packet into fragments
  - … in case a link cannot handle a large IP packet
- Time-To-Live (8 bits)
  - Used to identify packets stuck in forwarding loops
  - … and eventually discard them from the network

Time-to-Live (TTL) Field

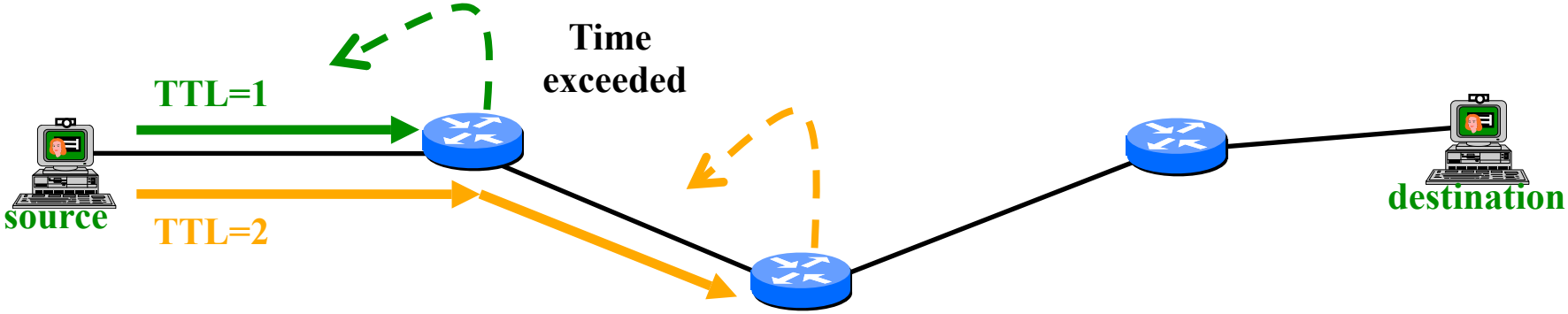
- Potential robustness problem
  - Forwarding loops can cause packets to cycle forever
  - Confusing if the packet arrives much later



- Time-to-live field in packet header
  - TTL field decremented by each router on the path
  - Packet is discarded when TTL field reaches 0…
  - …and "time exceeded" message is sent to the source

Application of TTL in Traceroute

- Time-To-Live field in IP packet header
  - Source sends a packet with a TTL of *n*
  - Each router along the path decrements the TTL
  - "TTL exceeded" sent when TTL reaches *0*
- Traceroute tool exploits this TTL behavior



Send packets with TTL=1, 2, … and record source of "time exceeded" message

Example Traceroute: Berkeley to CNN

Hop number, IP address, DNS name

1	169.229.62.1	inr-daedalus-0.CS.Berkeley.EDU
2	169.229.59.225	soda-cr-1-1-soda-br-6-2
3	128.32.255.169	vlan242.inr-202-doecev.Berkeley.EDU
4	128.32.0.249	gigE6-0-0.inr-666-doecev.Berkeley.EDU
5	128.32.0.66	qsv-juniper--ucb-gw.calren2.net
6	209.247.159.109	POS1-0.hsipaccess1.SanJose1.Level3.ne
7	*	?
8	64.159.1.46	?
9	209.247.9.170	pos8-0.hsa2.Atlanta2.Level3.net
10	66.185.138.33	pop2-atm-P0-2.atdn.net
11	*	?
12	66.185.136.17	pop1-atl-P4-0.atdn.net
13	64.236.16.52	www4.cnn.com

No response from router

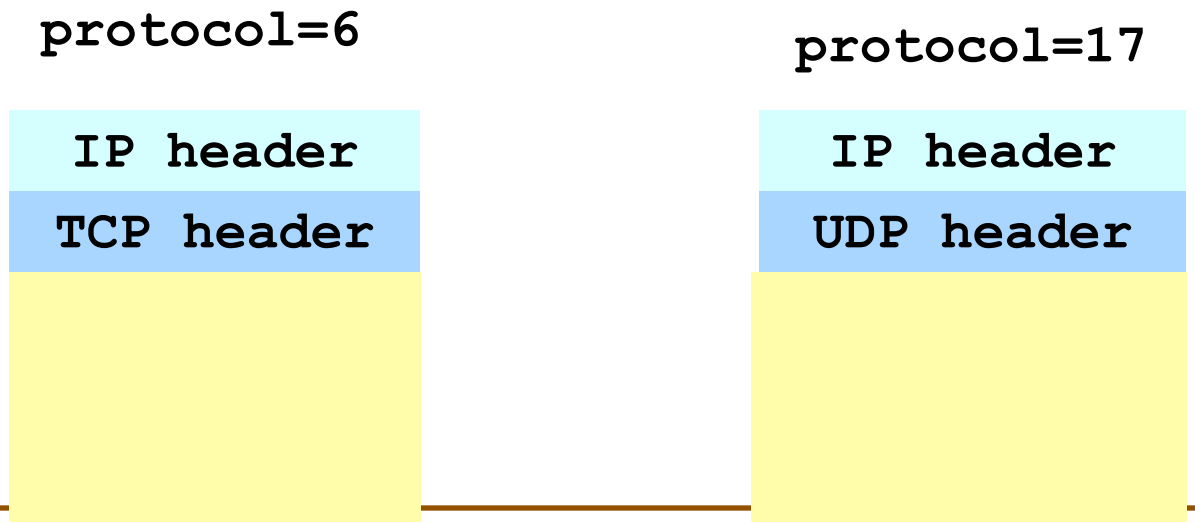
No name resolution

Try Running Traceroute Yourself

- On UNIX machine
  - Traceroute
  - E.g., “traceroute www.cnn.com” or “traceroute 12.1.1.1”
- On Windows machine
  - Tracert
  - E.g., “tracert www.cnn.com” or “tracert 12.1.1.1”
- Common uses of traceroute
  - Discover the topology of the Internet
  - Debug performance and reachability problems

IP Packet Header Fields (Continued)

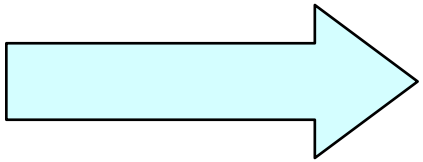
- Protocol (8 bits)
  - Identifies the higher-level protocol
    - E.g., “6” for the Transmission Control Protocol (TCP)
    - E.g., “17” for the User Datagram Protocol (UDP)
  - Important for demultiplexing at receiving host
    - Indicates what kind of header to expect next



IP Packet Header Fields (Continued)

- Checksum (16 bits)
  - Sum of all 16-bit words in the IP packet header
  - If any bits of the header are corrupted in transit
  - … the checksum won't match at receiving host
  - Receiving host discards corrupted packets
    - Sending host will retransmit the packet, if needed

```
   134
 + 212
-----
 = 346
```



```
   134
 + 216
-----
 = 350
```

Mismatch!

IP Packet Header (Continued)

- Two IP addresses
  - Source IP address (32 bits)
  - Destination IP address (32 bits)
- Destination address
  - Unique identifier for the receiving host
  - Allows each node to make forwarding decisions
- Source address
  - Unique identifier for the sending host
  - Recipient can decide whether to accept packet
  - Enables recipient to send a reply back to source

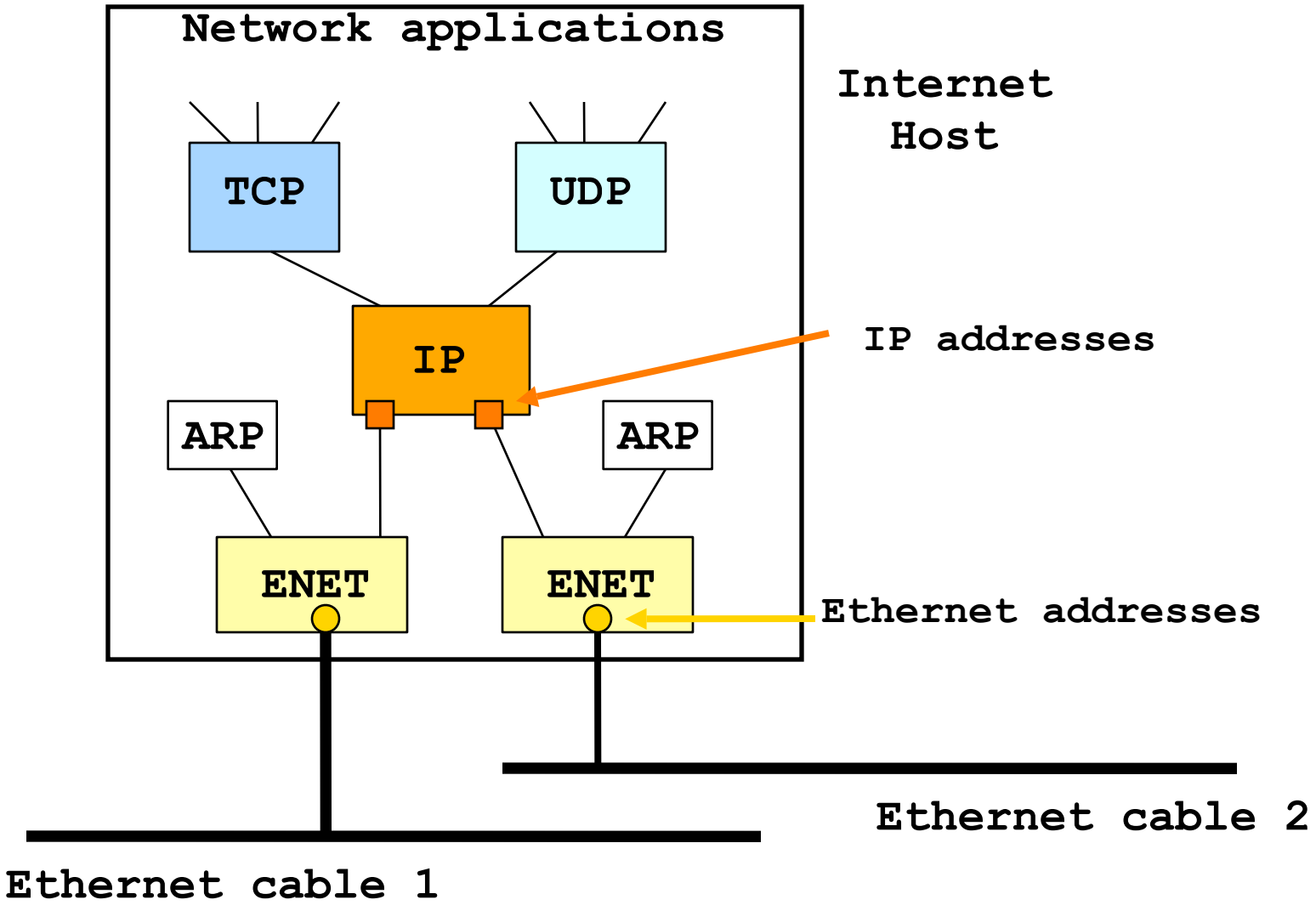
What if the Source Lies?

- Source address should be the sending host
  - But, who's checking, anyway?
  - You could send packets with any source you want
- Why would someone want to do this?
  - Launch a denial-of-service attack
    - Send excessive packets to the destination
    - … to overload the node, or the links leading to the node
  - Evade detection by "spoofing"
    - But, the victim could identify you by the source address
    - So, you can put someone else's source address in the packets
  - Also, an attack against the spoofed host
    - Spoofed host is wrongly blamed
    - Spoofed host may receive return traffic from the receiver

Summary: Packet Switching Review

- Efficient
  - Can send from any input that is ready
- General
  - Multiple types of applications
- Accommodates bursty traffic
  - Addition of queues
- Store and forward
  - Packets are self contained units
  - Can use alternate paths – reordering
- Contention (i.e., no isolation)
  - Congestion
  - Delay

IP και Ethernet



Λίγα λόγια για το Ethernet

- Ethernet frame contains:
  - Destination address
  - Source address
  - Type field
  - Data (payload)
- Ethernet address:
  - Size: 6 bytes
  - Format: 08-00-10-5A-54-C3 (hex)
- Applies CSMA/CD:
  - Carrier Sense and Multiple Access with Collision Detection

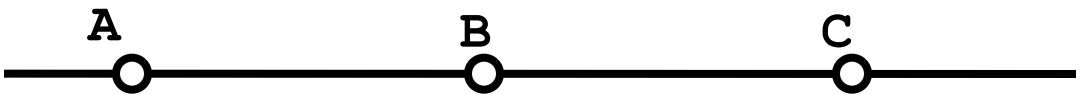
Address Resolution Protocol (ARP)

- Used to translate IP addresses to Ethernet addresses.
  - Translation done *only for outgoing IP packets*: this is when the IP header and the Ethernet header are created.
  - Translation is performed with a table look-up in an ARP Table; this is stored in memory and contains a row for each computer:

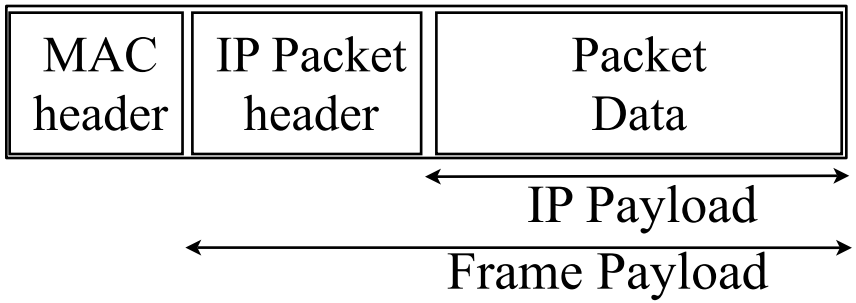
IP address	Ethernet address
223.1.2.1	08-00-39-00-2F-C3
223.1.2.3	08-00-5A-21-A7-22
223.1.2.4	08-00-10-99-AC-54

- The ARP table is necessary because the IP address and Ethernet address are selected independently:
  - the IP address is selected by the network manager based on the location of the computer on the internet.
  - the Ethernet address is selected by the manufacturer based on the Ethernet address space licensed by the manufacturer.
- Each host has separate ARP tables for each of its Ethernet interfaces.

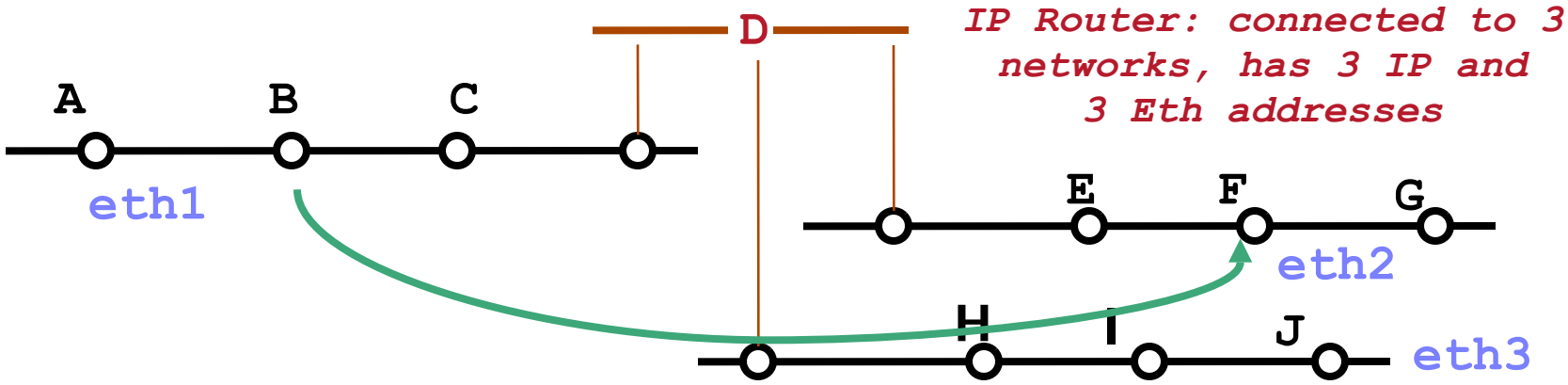
Άμεση δρομολόγηση στο IP



- Direct routing
  - A sends packet to C
  - The IP header contains A's IP address (source) and B's IP address (destination)
  - The Ethernet frame header contains A's ethernet address (source) and B's ethernet address (destination)



Έμμεση δρομολόγηση IP

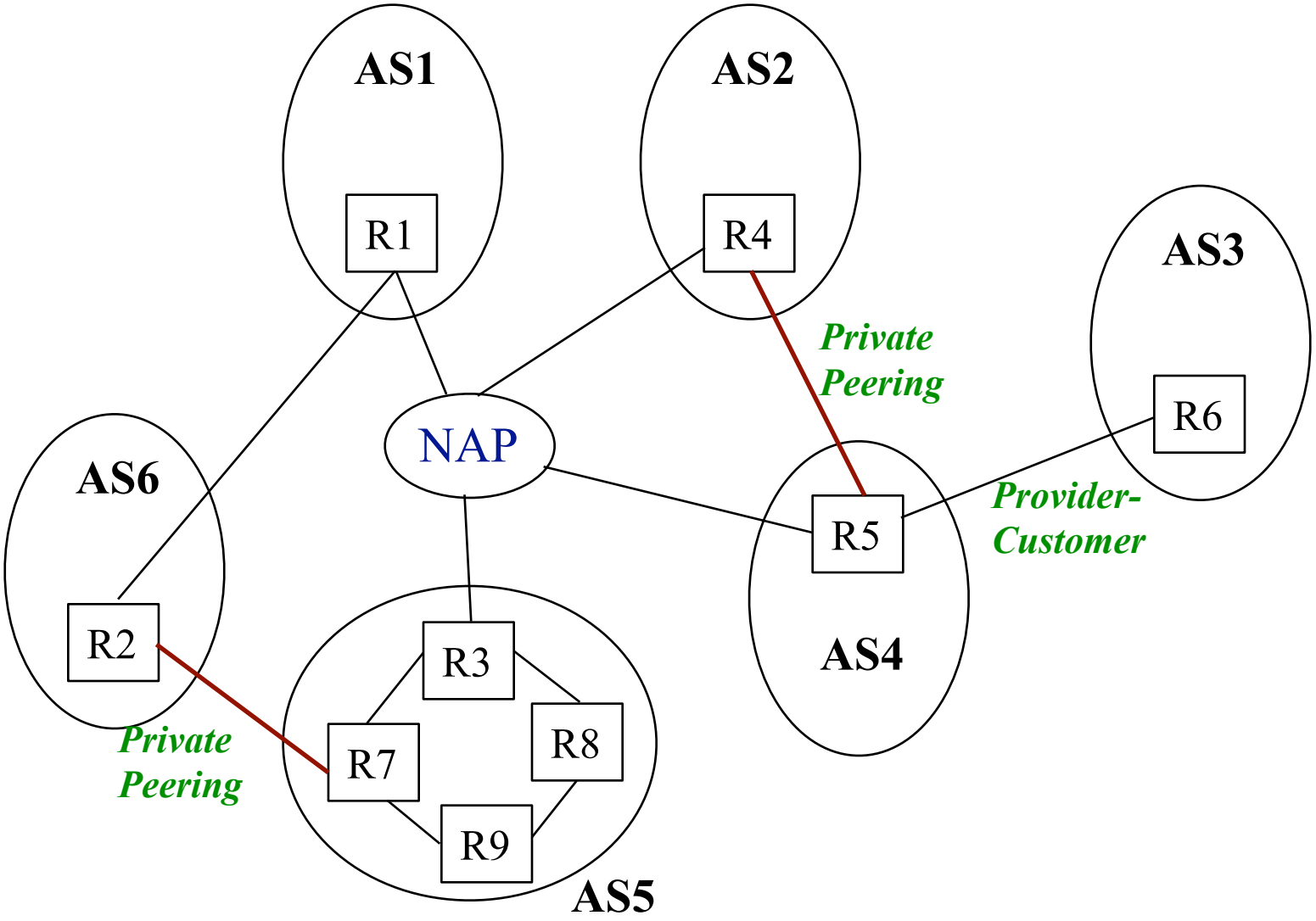


- Network managers assign a unique number (the IP network number) to each of the ethernets.
- B sends msg to D: direct routing
- B sends msg to F: indirect routing, done by IP modules, transparently to TCP, UDP and apps.
  - IP header contains A's IP and eth address, as source addresses
  - Ethernet header carries F's IP and D's eth. address as destination.

Δρομολόγηση στο IP

- The Internet is divided into regions under a single administrative control, each of which is called an Autonomous System and contains a group of network IDs.
- Routing inside an AS is completely hidden from the rest of the Internet.
- Routes between ASs are computed in terms of AS hops: lists of intermediary ASs from the source AS to the destination AS.
  - All outside networks that belong to the same AS share the same route, expressed as the list of intermediary ASs.
  - To route to arbitrary ASs on the Internet, a router need only know the next-hop router to every AS, rather than to individual destinations.

Autonomous Systems



Routing within ASs

- Under complete control of AS owner
- Small ASs have only one router, which does the internal and external routing (e.g., AS1, AS2, AS3, AS4, AS6)
- Larger ASs may have more than one routers:
  - some of them are *border routers* and deal with outgoing or incoming traffic to the AS (e.g., AS5)
  - for internal routing the Open Shortest Path First (OSPF) protocol is used

AS categories

- Transit AS:
  - an AS with connections to more than one AS that is also willing to carry datagrams that neither originate nor terminate on its own hosts
- Multihomed AS:
  - an AS that is connected to more than one AS and that does not accept datagrams not destined to itself
- Stud AS:
  - an AS that is connected to only one other AS

AS Connections

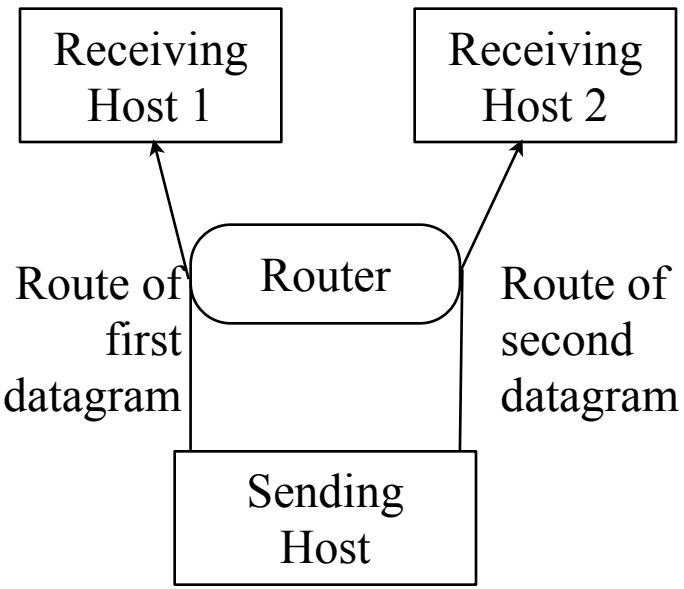
- Network Access Point (NAP): a physical network that connects routers from different ASs
- Private Peering Links: private physical networks connecting only the routers from two ASs
  - Peering agreements usually limit traffic over peering link to non-transit traffic
- Provider-customer relationship: similar to peering linkage but usually complies to different business model:
  - the provider delivers transit traffic to the customer

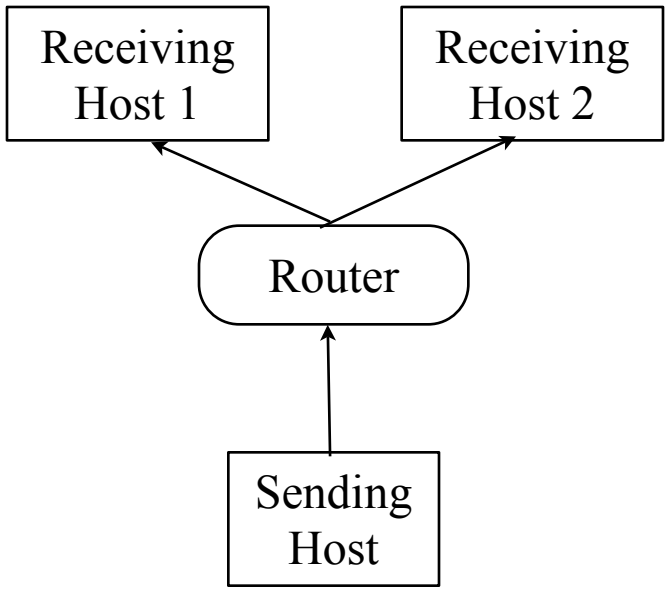
Routing between ASs

- Done according to the Border Gateway Protocol (BGP):
  - Each router advertises reachability information to neighbor routers for:
    - all networks within its AS and
    - for outside networks reachable via its AS (for transit ASs)
  - Reachability information includes a list of reachable networks and performance cost expressed as the number of hops to the destination
  - An AS can set routing policies that determines which reachability information is advertised to which routers

IP Multicast

- Allows the efficient delivery of a datagram to multiple hosts on the Internet that are members of the same multicast group and are interested in the same content.
- IP multicast seeks to optimize the transmissions of IP datagrams in a way that an identical datagram traverses the link between two routers only once.





IP Addressing of Multicast groups

- Multicast groups are addressed using Class D IP addresses:
  - A host can send an IP packet to all hosts in a group, using the group's multicast IP address as the packet destination address
  - To receive multicast packets, a host must join the multicast group by notifying its next-hop router that it would be interested in a particular multicast address
  - The router then cooperates with other Internet routers to ensure the delivery to this host of packets with this destination address
- Scalability challenge:
  - every router on a path from any host to any member of a multicast group must know about this group
  - amount of routing state substantially greater
  - in some multicast protocols routing of mcast packets depends on the sender: a router must maintain an entry in its table for each mcast group and potential sending host!