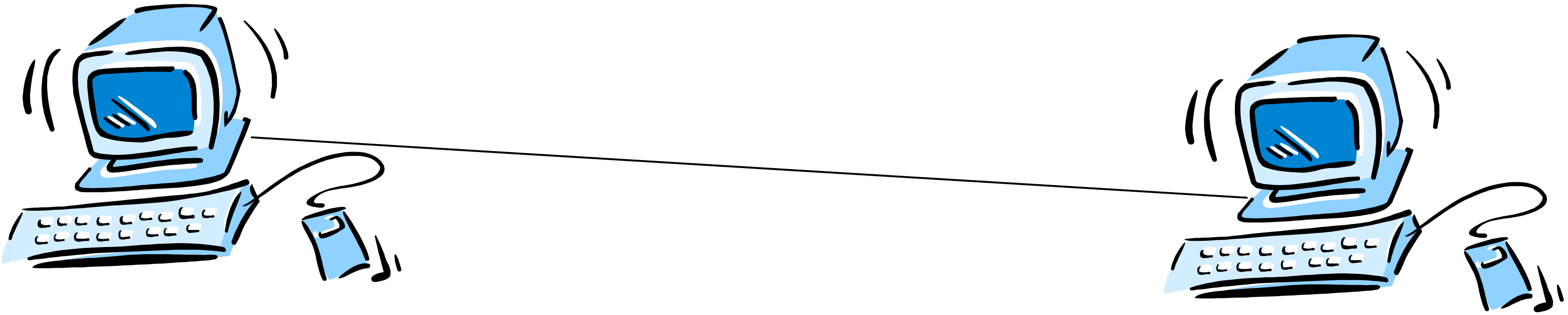


95-702 Distributed Systems

Lecture 7: Internetworking

See Chapter 3 of Colouris

Learning Goals Preamble:

MISM and MSIT grads may need to plan, develop, and manage distributed systems.

These distributed systems run on networks and internetworks. Therefore they need to understand their basic operation, the most prevalent of which is the Internet.

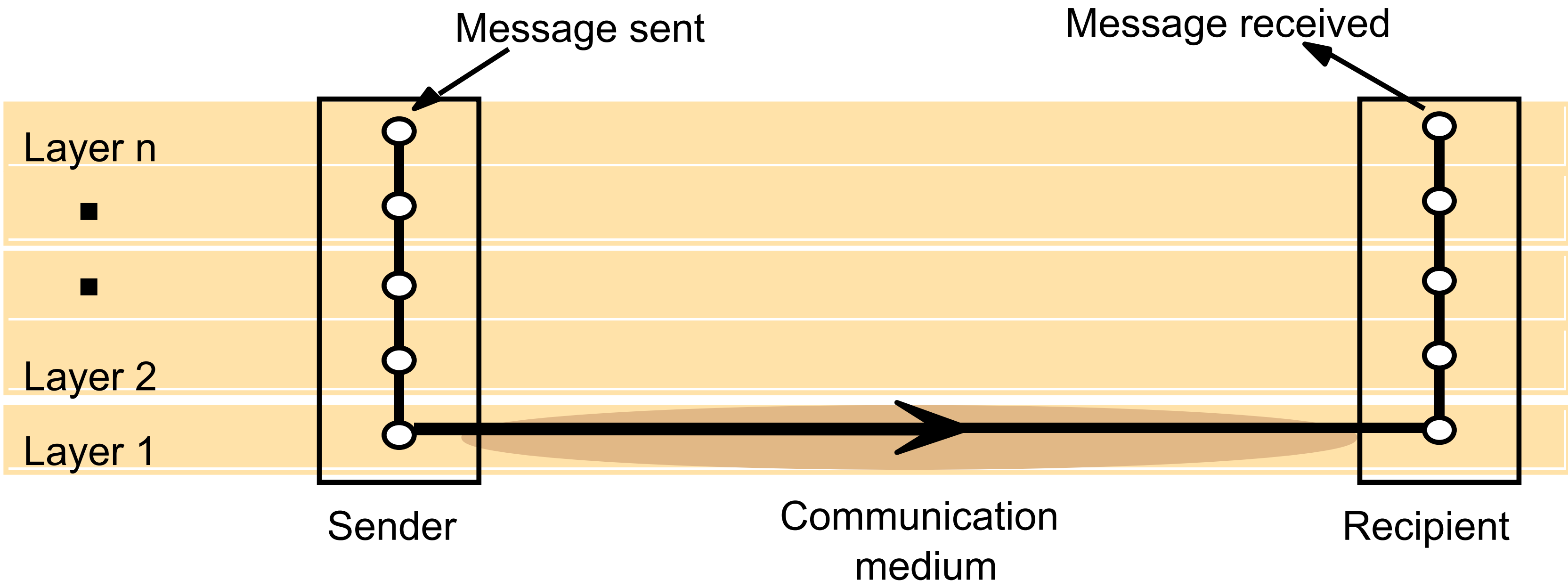
Therefore, today's learning goals are to:

1. Be comfortable with terminology used concerning the Internet
2. Understand the role of protocols, and the layering of protocols, in the architecture of the Internet. And how this layering provides levels of abstraction below which a developer need not be (too) concerned.
3. Understand the basic functionality of how packets of information travel between one system and another. This will inform design and configuration choices in building and maintaining systems.
4. Understand IP addressing.

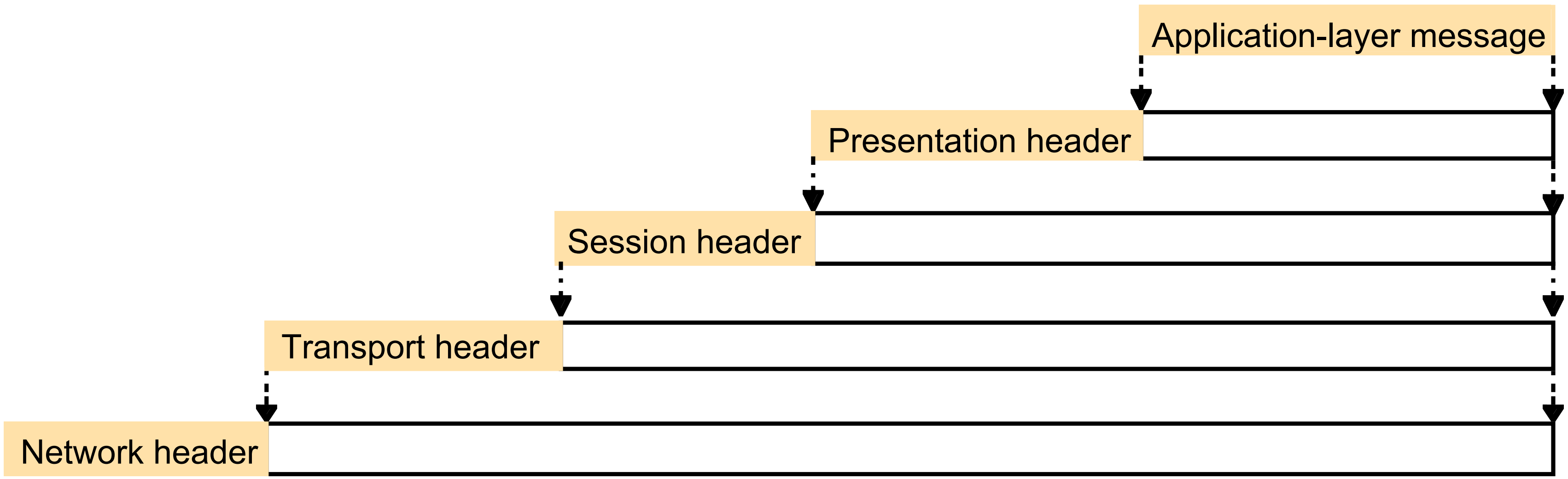
Basics

- When we speak of a network we will be speaking about a single technology network (Ethernet, Token Ring, ATM, Point to Point, WaveLan, etc.)
- An internetwork is an interconnected collection of such networks.
- The Internet Protocol (IP) is the key toll used today to build scalable, heterogeneous internetworks

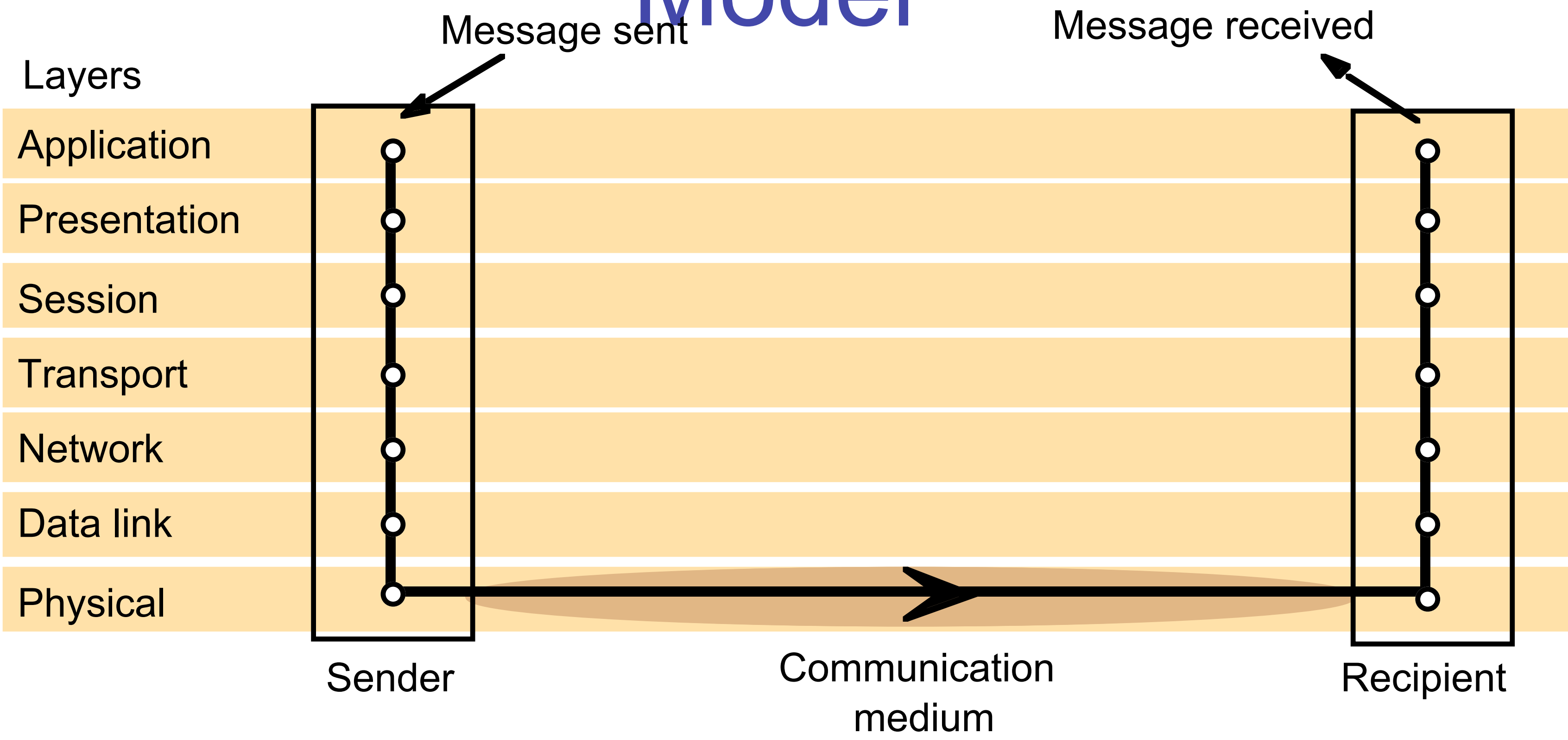
Conceptual Layering of Protocol Software



Encapsulation as it is Applied in Layered Protocols



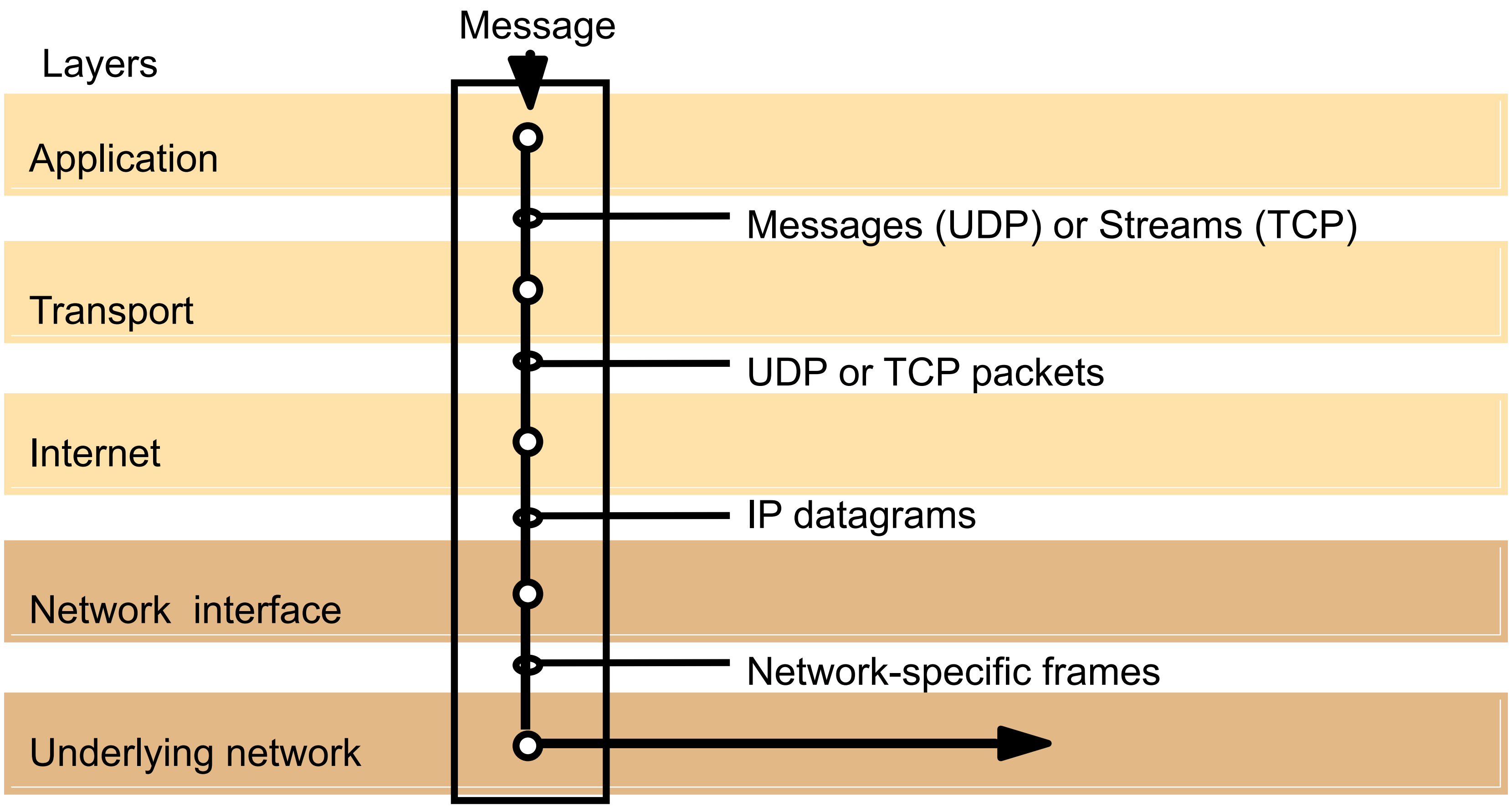
Protocol Layers in the ISO Open Systems Interconnection (OSI) Model



OSI Protocol Summary

Layer	*Description*	*Examples*
Application	Protocols that are designed to meet the communication requirements of specific applications, often defining the interface to a service.	HTTP,FTP, SMTP, CORBA IIOP
Presentation	Protocols at this level transmit data in a network representation that is independent of the representations used in individual computers, which may differ. Encryption is also performed in this layer, if required.	Secure Sockets (SSL),CORBA Data Rep.
Session	At this level reliability and adaptation are performed, such as detection of failures and automatic recovery.	SIP
Transport	This is the lowest level at which messages (rather than packets) are handled. Messages are addressed to communication ports attached to processes, Protocols in this layer may be connection-oriented or connectionless.	TCP, UDP
Network	Transfers data packets between computers in a specific network. In a WAN or an internetwork this involves the generation of a route passing through routers. In a single LAN no routing is required.	IP, ATM virtual circuits
Data link	Responsible for transmission of packets between nodes that are directly connected by a physical link. In a WAN transmission is between pairs of routers or between routers and hosts. In a LAN it is between any pair of hosts.	Ethernet MAC, ATM cell transfer, PPP
Physical	The circuits and hardware that drive the network. It transmits sequences of binary data by analogue signalling, using amplitude or frequency modulation of electrical signals (on cable circuits), light signals (on fibre optic circuits) or other electromagnetic signals (on radio and microwave circuits).	Ethernet base- band signalling, ISDN

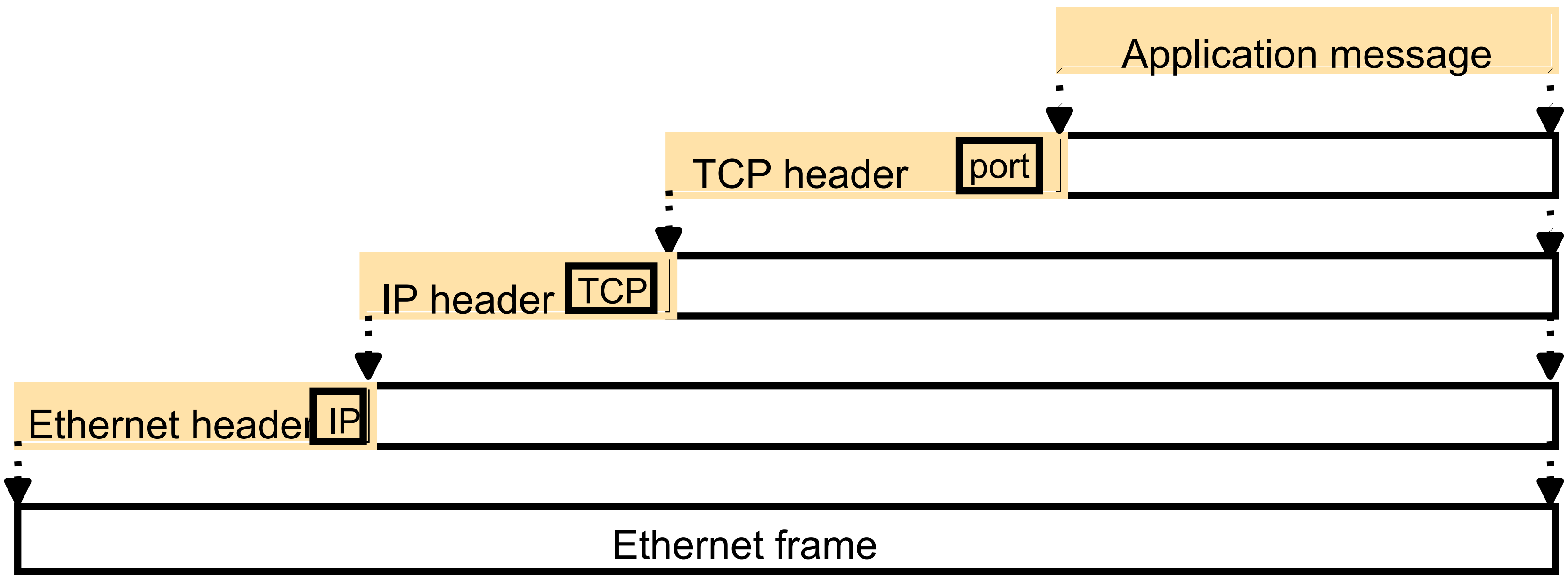
TCP or UDP Over IP



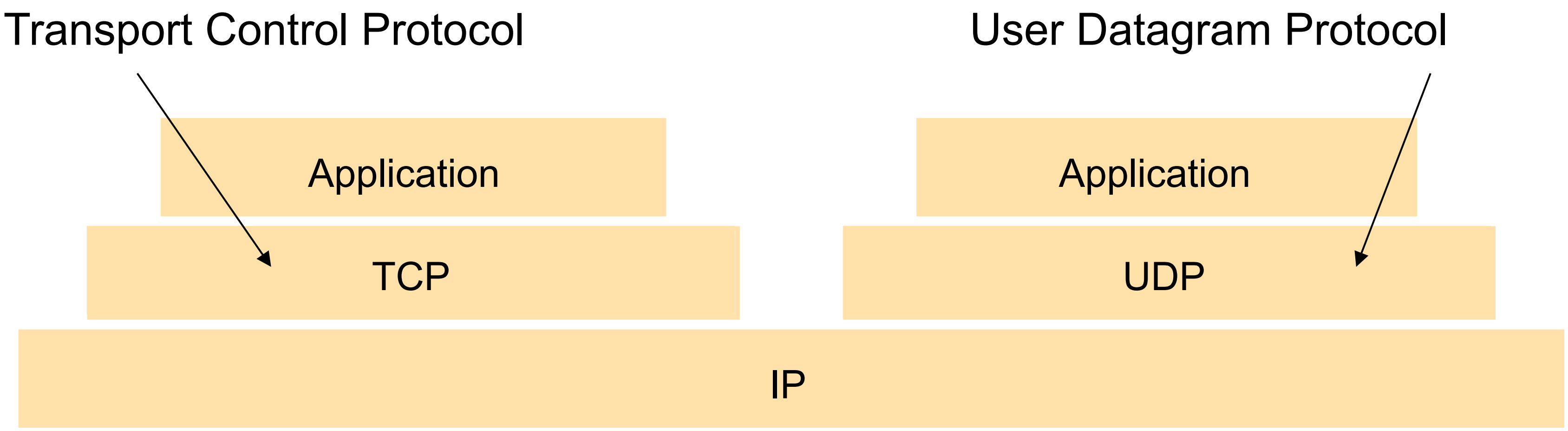
TCP and UDP Quick Notes

- TCP is stream based, connection oriented and stateful.
- The TCP message sender gets acknowledgements.
- This makes it a "reliable" protocol.
- TCP "plays nice" with others. If problems are detected it backs off by ½. If no problems it ramps up by 1.
- UDP uses datagrams and does not establish a connection.
- UDP fires and forgets.
- UDP does not necessarily "play nice". If problems occur UDP is not even aware.
- UDP can be made reliable by the application. Require acknowledgements and do retries when acknowledgements do not arrive in time.
- UDP also allows for broadcasting messages to many hosts.
- If you are willing to occasionally lose some bits and need high performance, UDP is a strong candidate.

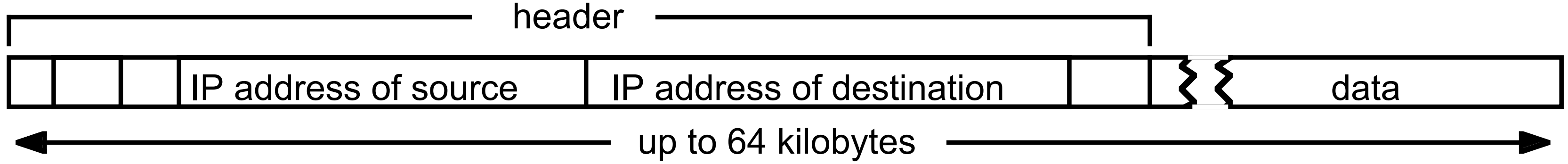
Encapsulation in a Message Transmitted via TCP over an Ethernet



The Programmer's Conceptual View of a TCP/IP Internet



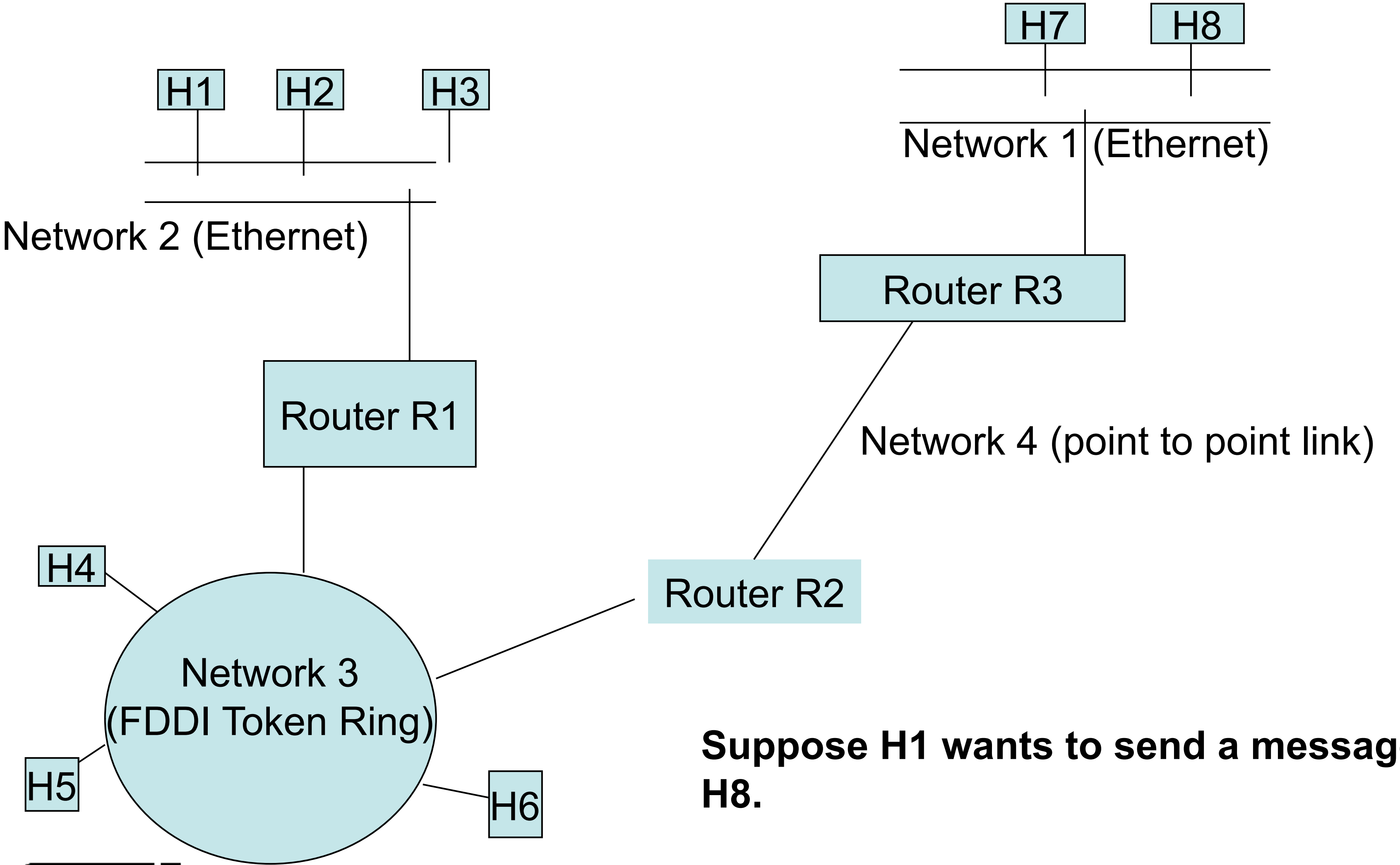
IP Packet Layout



IEEE 802 Network Standards

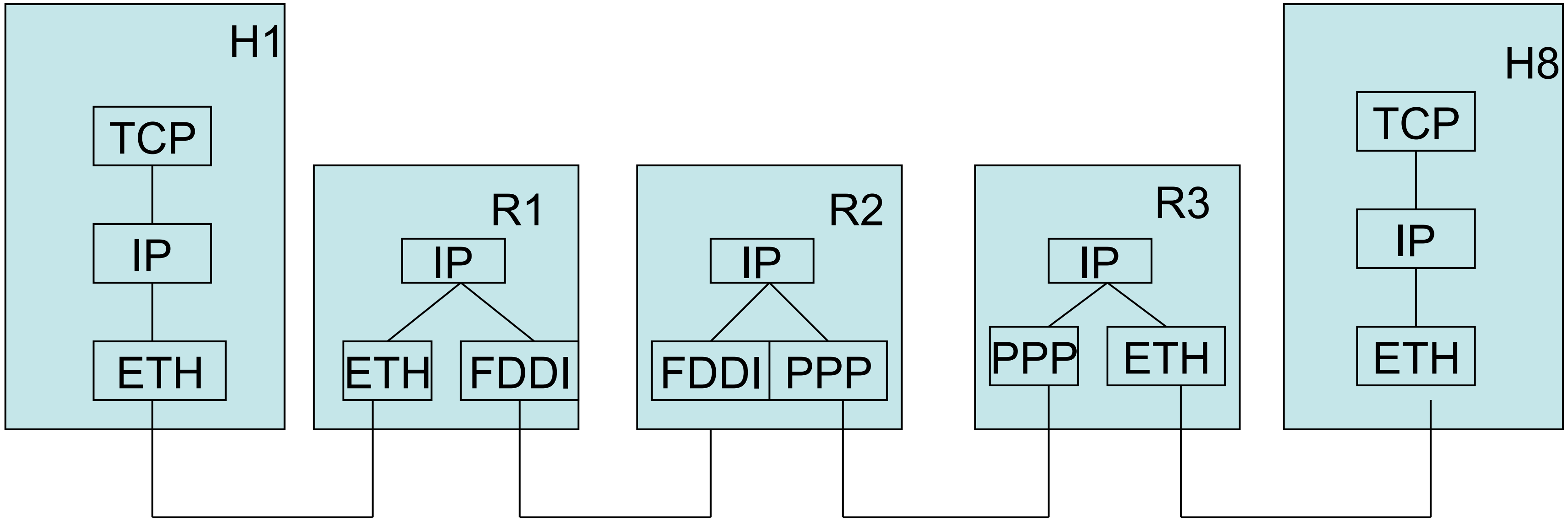
IEEE No.	*Title*	*Reference*
802.3	CSMA/CD Networks (Ethernet)	[IEEE 1985a]
802.4	Token Bus Networks	[IEEE 1985b]
802.5	Token Ring Networks	[IEEE 1985c]
802.6	Metropolitan Area Networks	[IEEE 1994]
802.11	Wireless Local Area Networks	[IEEE 1999]

Example Internetwork



Suppose H1 wants to send a message to H8.

H1 To H8



Protocol Layering

IP

- Requires that lower level protocols provide services…
- And therefore was designed to be undemanding…
- In this way, IP can make use of a wide variety of underlying networks

IP

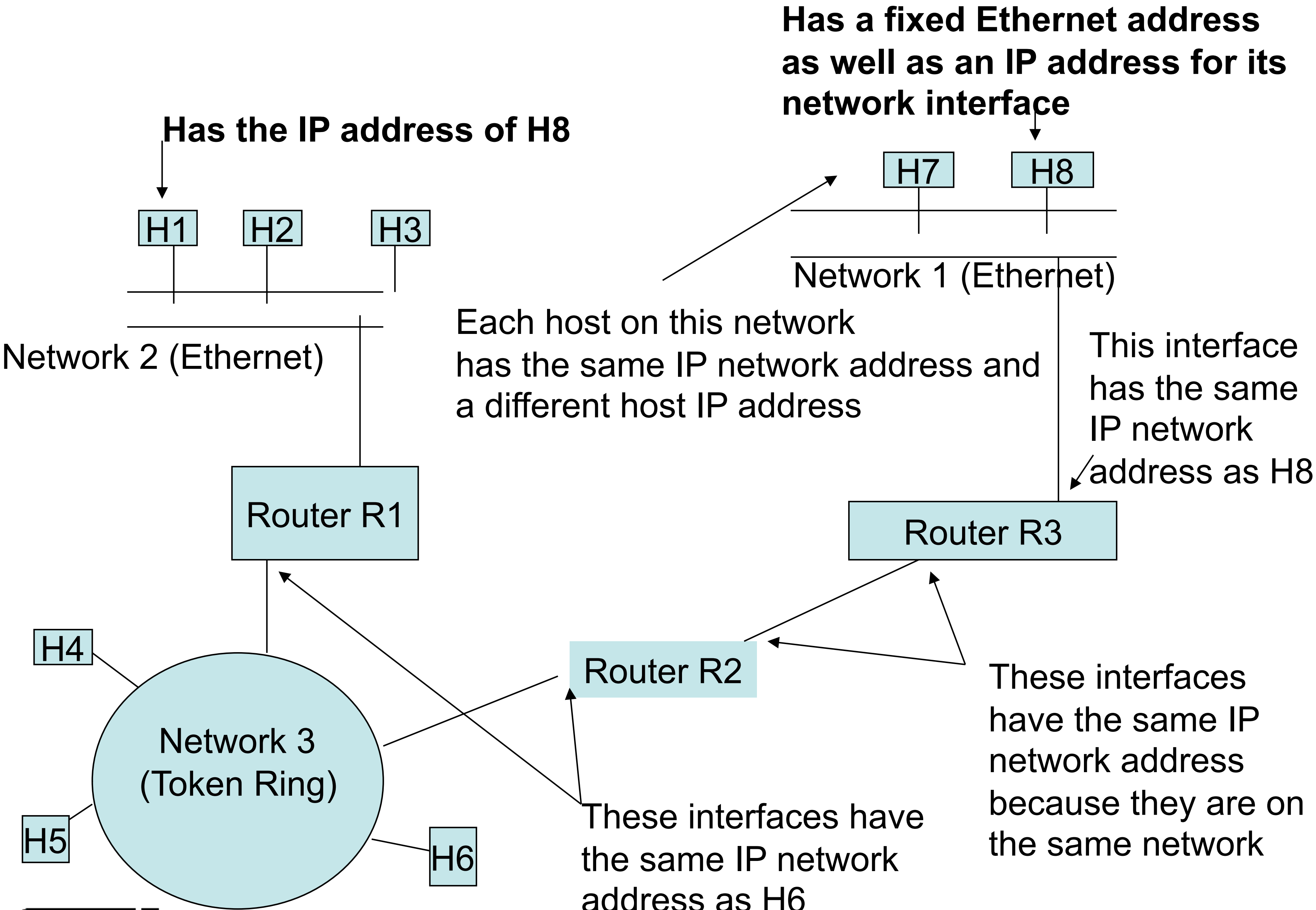
- Has an addressing scheme which identifies each host on the internetwork
- Has a best effort datagram delivery model
- Could be run over carrier pigeons
- Many of the technologies that IP runs on were invented well after IP was defined.

Addressing

Every Ethernet device has a network adapter with a 48-bit globally unique ID. Each manufacturer is assigned 24 bits. The other 24 bits are assigned by the manufacturer. These addresses have little structure and provide very few clues as to their location.

IP addresses have a network part and a host part.

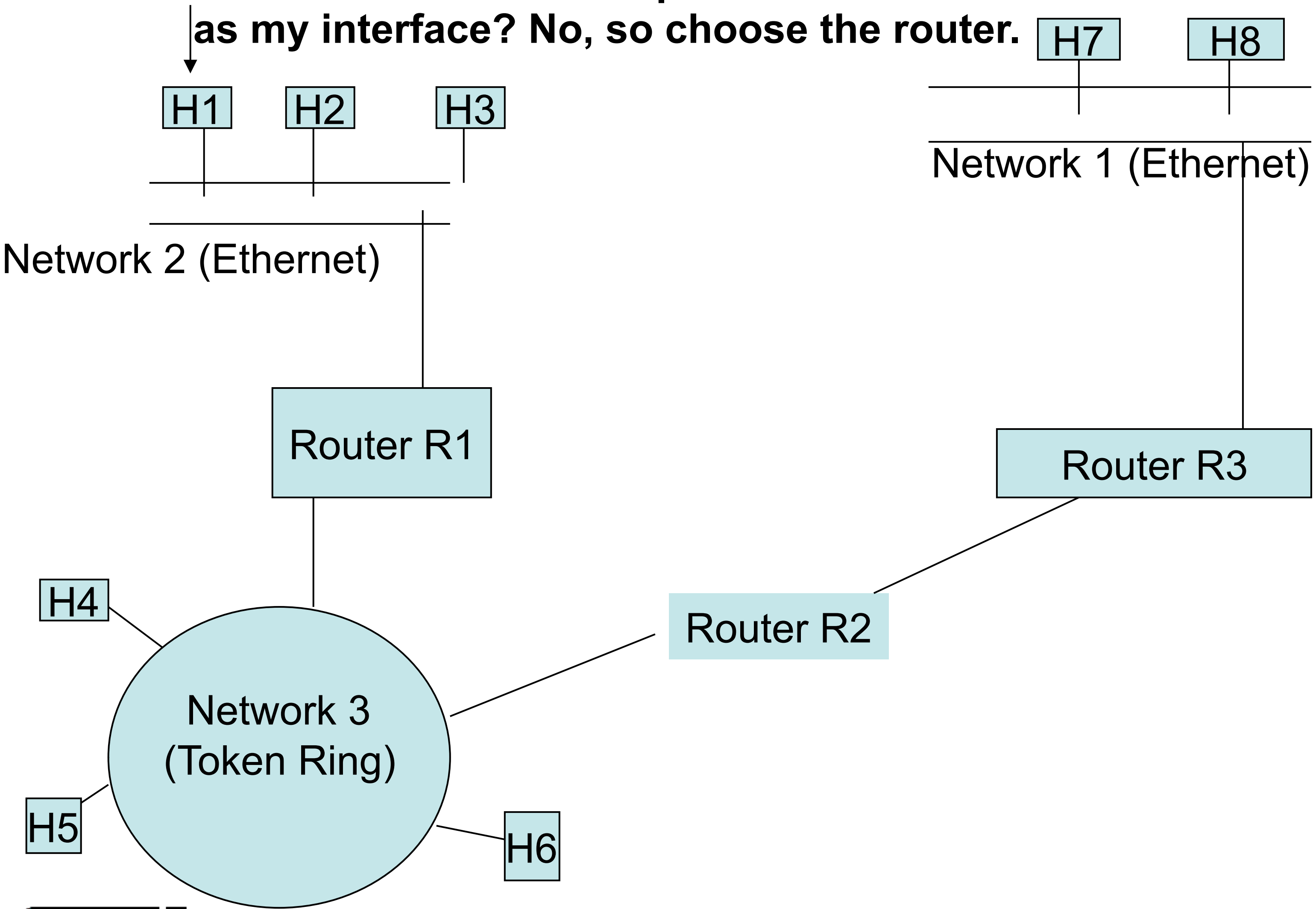
Suppose H1 has the IP address of H8…


Has a fixed Ethernet address as well as an IP address for its network interface
Has the IP address of H8
H7
H8
H1
H2
H3
Network 1 (Ethernet)
Network 2 (Ethernet)
Each host on this network has the same IP network address and a different host IP address
This interface has the same IP network address as H8
Router R1
Router R3
Router R2
H4
Network 3 (Token Ring)
H5
H6
These interfaces have the same IP network address because they are on the same network
These interfaces have the same IP network address as H6

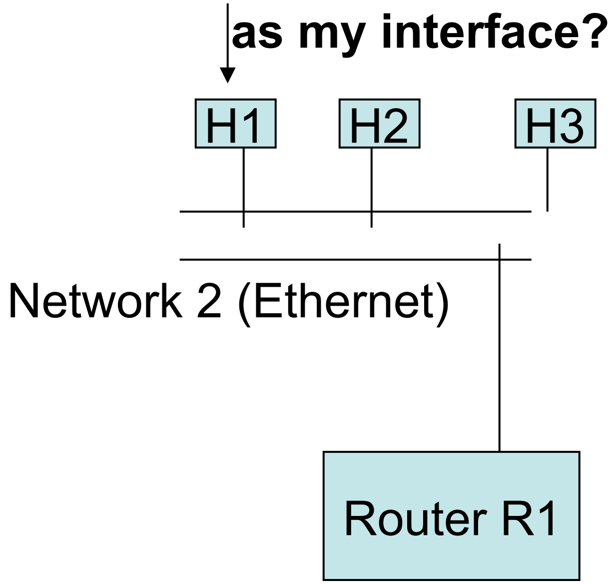
IP Addressing

- Every IP datagram contains the IP address of the destination host.
- The “network part” of an IP address uniquely identifies a single physical network that is part of the larger Internet.
- All hosts and routers that share the same network part of their address are connected to the same physical network and can thus communicate with each other by sending frames over the network.
- Every physical network that is part of the Internet has at least one router that, by definition, is also connected to at least one other physical network; this router can exchange packets with hosts or routers on either network.

H1 has the IP address of H8. Does H8 have the same network part address as my interface? No, so choose the router.



H1 has the IP address of H8. Does H8 have the same network part address as my interface? No, so choose the router.



But, how is this decision made?

Suppose this is a /24 network.

The leftmost 24 bits represent the network identifier. The remaining 8 bits represent the 2^8 hosts.

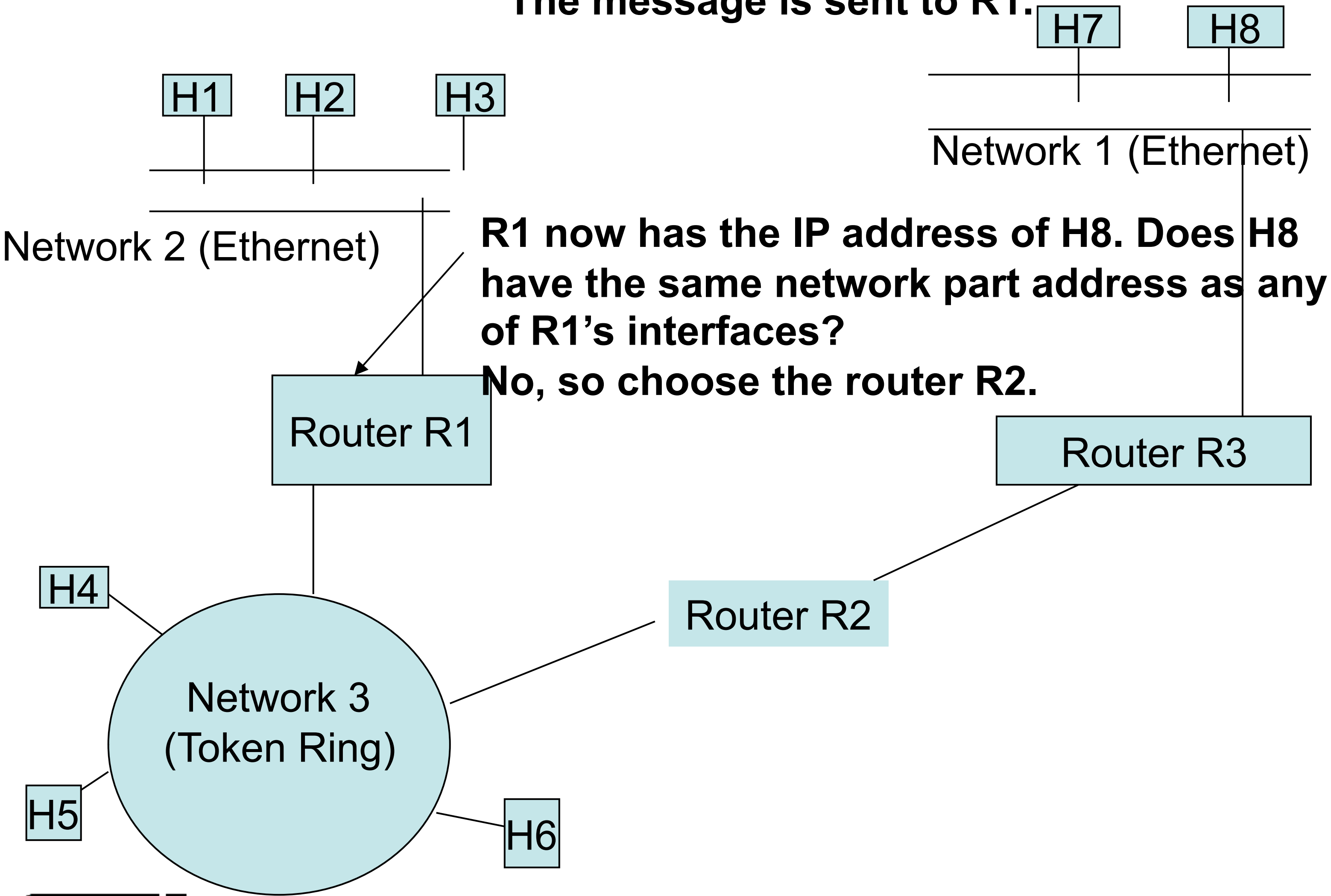
Therefore, H1 has a subnet mask of 255.255.255.0.

H1 performs a bitwise and of the subnet mask with H8's 32-bit IP address.

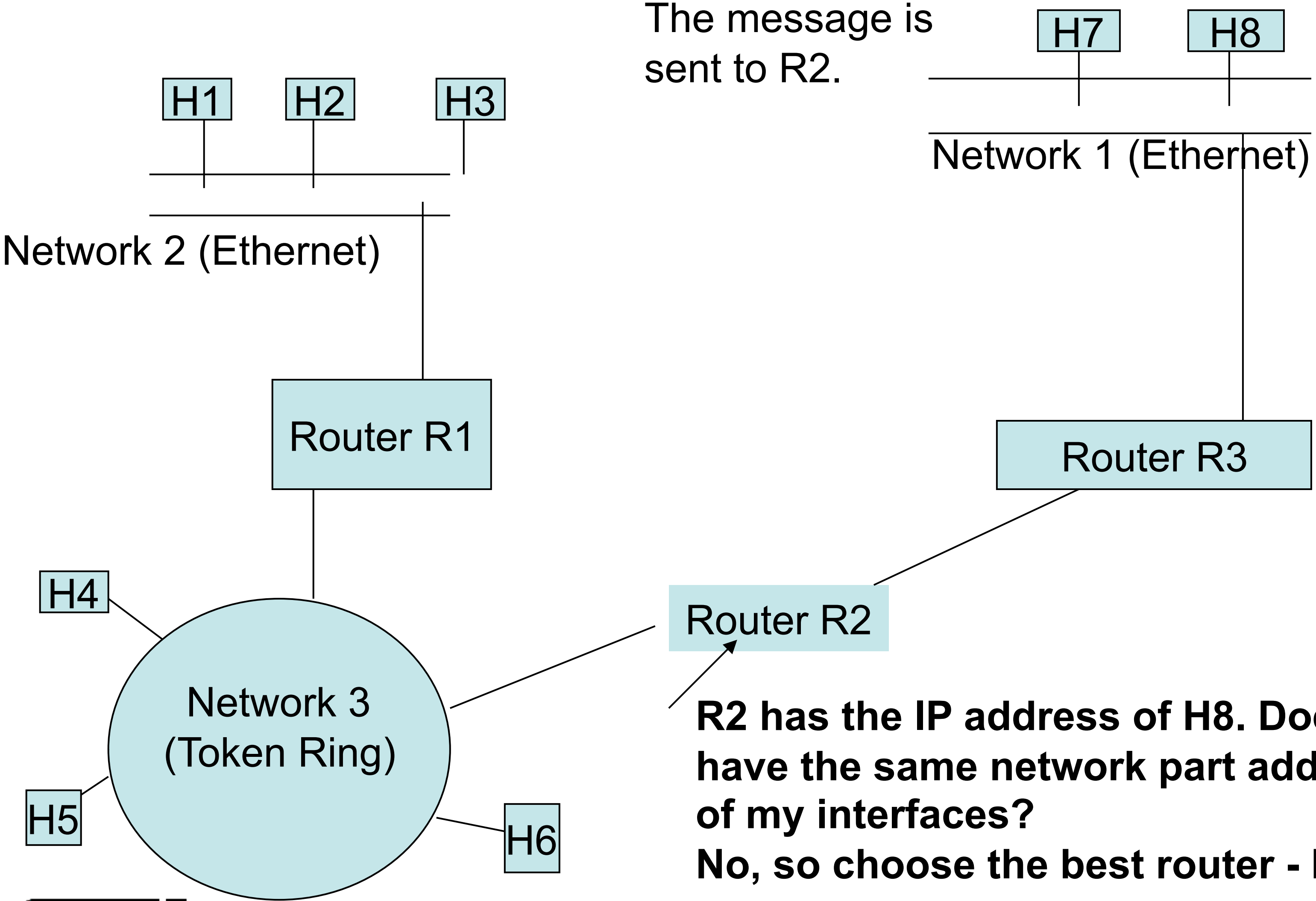
If the result does not match H1's network Identifier then H8 is a foreign machine.

The message is sent to R1.

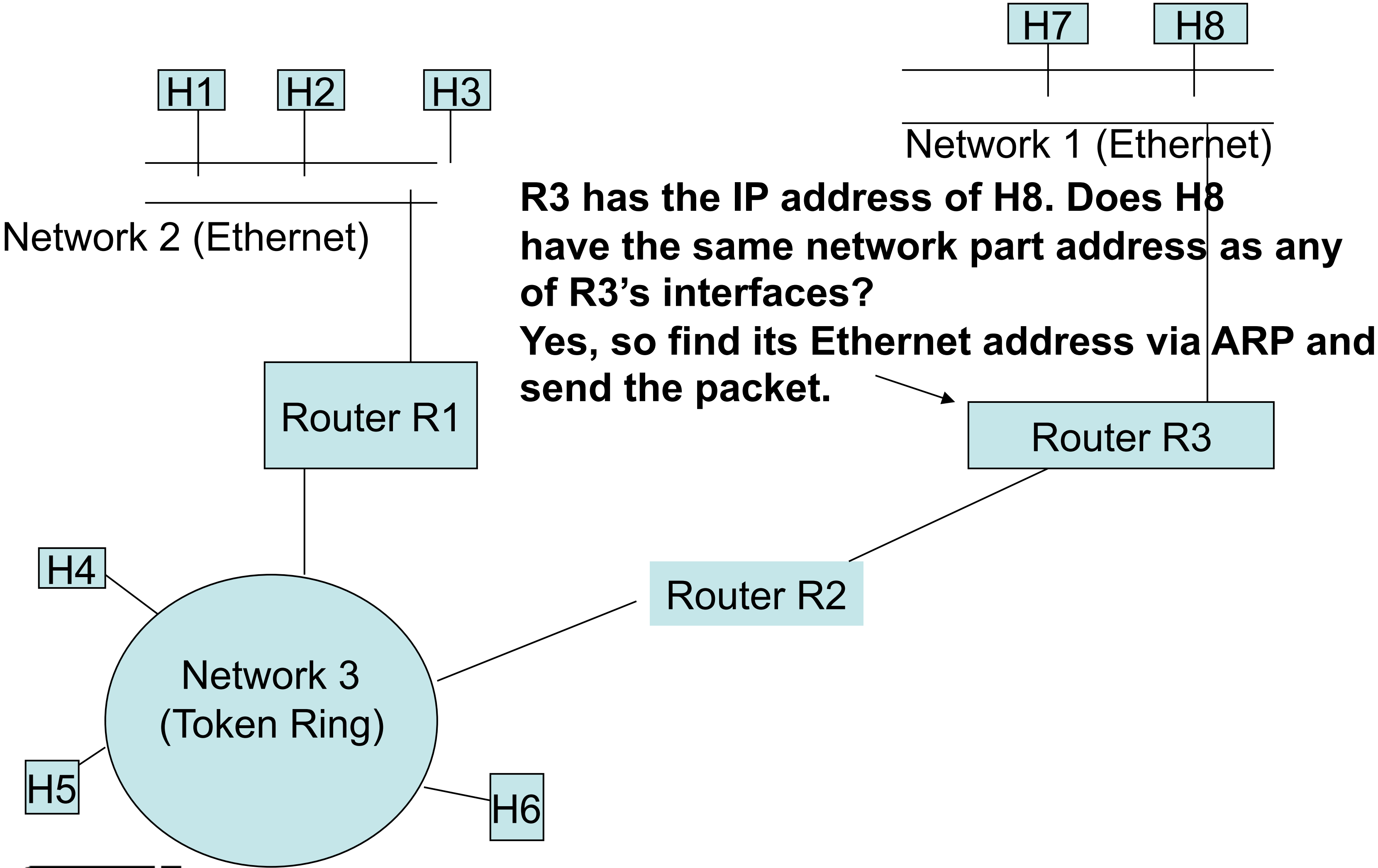
R1 now has the IP address of H8. Does H8 have the same network part address as any of R1's interfaces?
No, so choose the router R2.



The message is sent to R2.



R2 has the IP address of H8. Does H8 have the same network part address as any of my interfaces?
No, so choose the best router - R3.



R3 has the IP address of H8. Does H8 have the same network part address as any of R3's interfaces?

Yes, so find its Ethernet address via ARP and send the packet.

ARP

- Address Resolution Protocol
- The IP address needs to be translated to a link level address that is specific to the particular type of network.
- For example, Ethernet addresses are 48 bits. We must find the 48 bits associated with an IP address.
- Suppose a letter arrives at camp addressed to Billy. How does Billy get the letter?

Without ARP

- Without ARP, each host might hold a table of pairs:

  (IP address, Particular network address)

  (Billy, Bunk #4)
- If a host or router needs to reach a particular IP in its network it simply looks up the physical address in the table.
- This letter is for Billy and we do a lookup to find his bunk number.

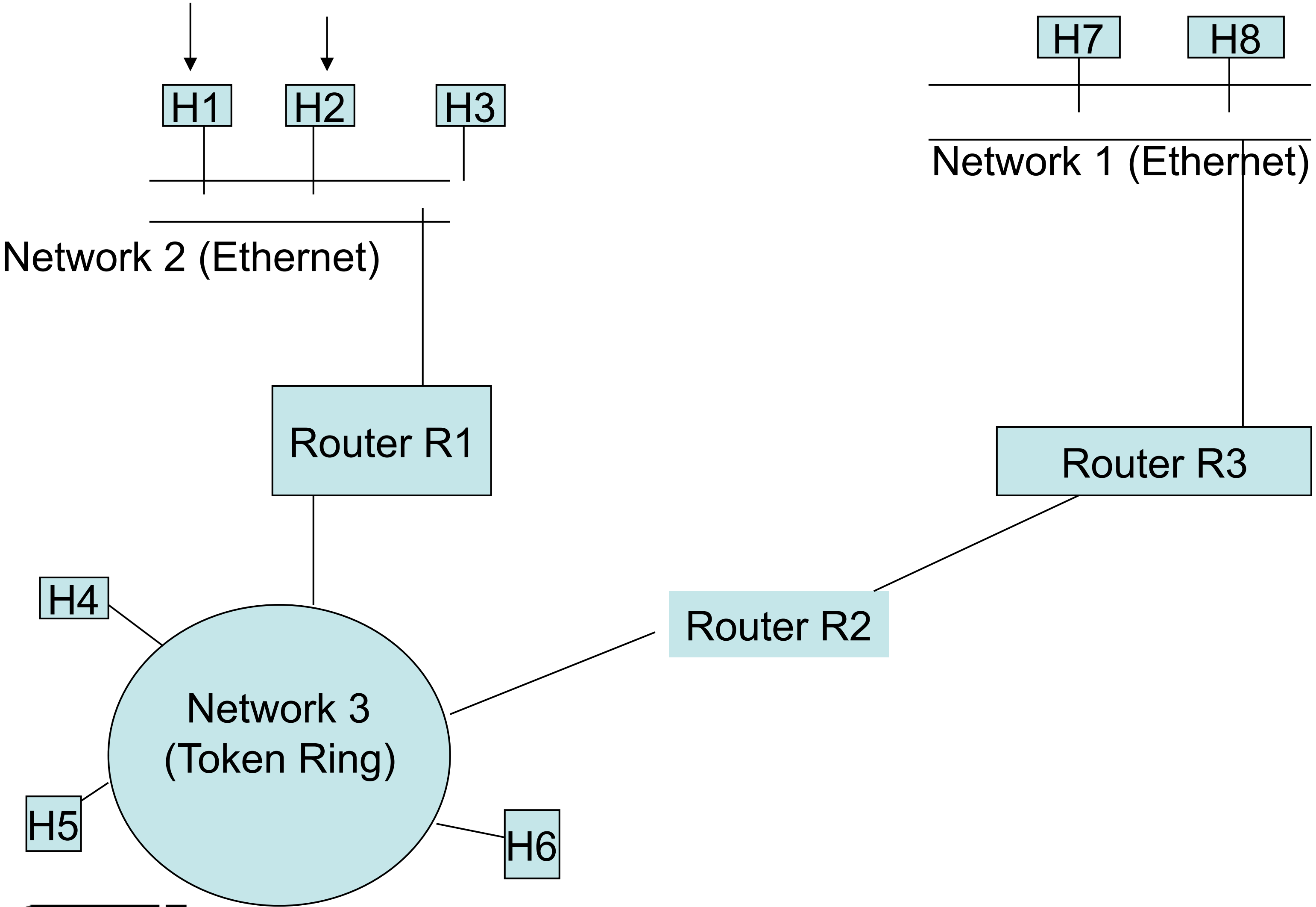
ARP

- But hosts might come and go. Billy might change bunks often.
- Each host dynamically builds up a table of mappings between IP addresses and link level addresses.
- The ARP cache times out every 15 minutes or so and construction begins anew.

ARP

- Host A wants to contact host B on the same network.
- First, A checks its cache to see if it already contains the IP address, physical address pair. If it does then use the physical address.
- If it does not then broadcast the IP address to all hosts on this network. The matching host sends back its physical address. A then adds this mapping to its cache.
- Other hosts on the network will see this interaction and build tables of their own.

H1 has H2's IP address. It finds H2's physical address with ARP.



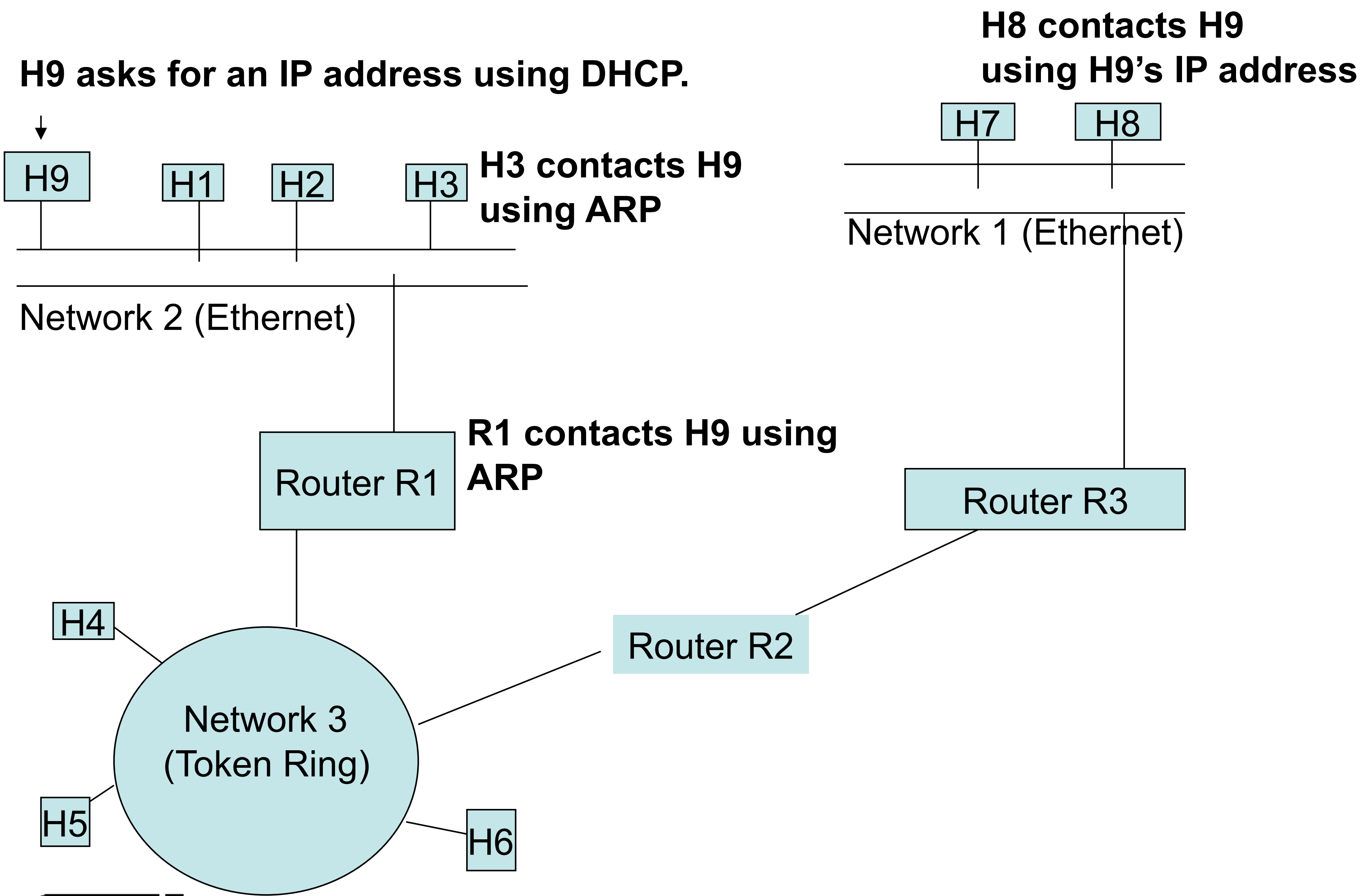
DHCP

- Dynamic Host Configuration Protocol
- Ethernet addresses are globally unique and fixed during the manufacture of Ethernet devices.
- IP addresses cannot be configured once into a host. The IP address has a network part and a host part. (You could never move the host to a different network!)
- Devices need IP addresses and the address of the default router.

DHCP

- A DHCP server provides configuration information to hosts.
- But how does the host find a DHCP server?
- Service discovery:

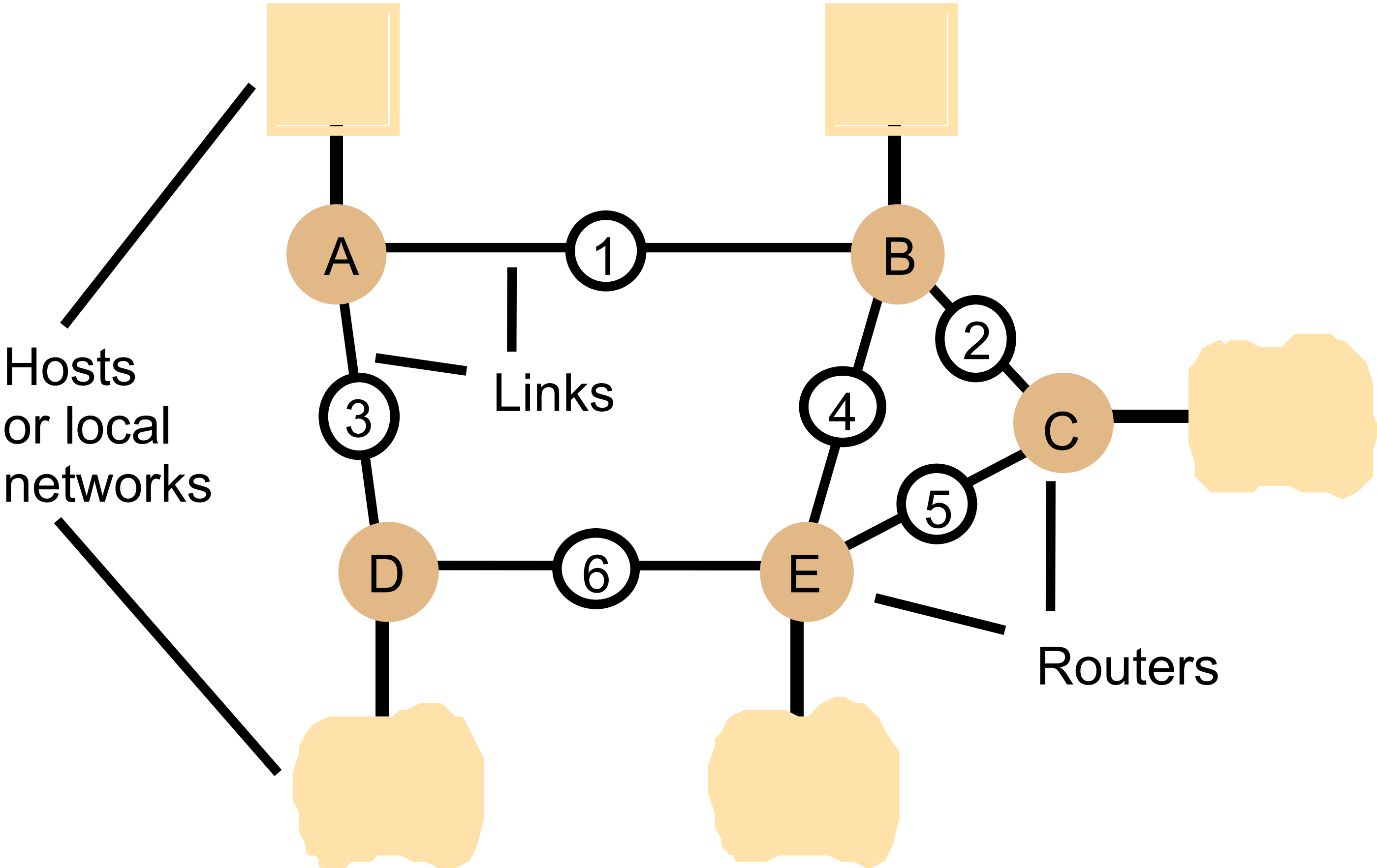
  The host broadcasts a DHCPDISCOVER over UDP/IP and the DHCP server sends back a leased IP address


H8 contacts H9
using H9's IP address
H9 asks for an IP address using DHCP.
H7
H8
H9
H1
H2
H3
H3 contacts H9
using ARP
Network 1 (Ethernet)
Network 2 (Ethernet)
R1 contacts H9 using
ARP
Router R1
Router R3
H4
Router R2
Network 3
(Token Ring)
H5
H6

Routers

- Keep messages flowing **between** networks rather than **within** networks
- Come in different sizes
- The largest have more in common with supercomputers than office servers - MIPS processors

Routing in a Wide Area Network



Initial Routing Tables for the Network

Routings from A

To	*Link*	*Cost*
A	local	0
B	1	1
C	-	inf
D	3	1
E	-	inf

Routings from B

To	*Link*	*Cost*
A	1	1
B	local	0
C	2	1
D	-	inf
E	4	1

Routings from C

To	*Link*	*Cost*
A	-	inf
B	2	1
C	local	0
D	-	inf
E	5	1

Routings from D

To	*Link*	*Cost*
A	3	1
B	-	inf
C	-	inf
D	local	0
E	6	1

Routings from E

To	*Link*	*Cost*
A	-	inf
B	4	1
C	5	1
D	6	1
E	local	0

RIP Routing Algorithm

Fault on n discovered: set cost to inf for each destination using that link and execute a send

Send: Each *t* seconds or when *Tl* changes, send *Tl* on each non-faulty outgoing link.

Receive: Whenever a routing table *Tr* is received on link *n*:

```
for all rows Rr in Tr {
    if (Rr.link <> n) {        // if the plan is not to come through here
        Rr.cost = Rr.cost + 1;  // Then I too could get there with a higher cost
        Rr.link = n;             // and I would travel through n
        if (Rr.destination is not in Tl) add Rr to Tl;  //add new destination toTl
        else for all rows Rl in Tl {
            if (Rr.destination = Rl.destination and
                    (Rr.cost < Rl.cost or Rl.link = n)) Rl = Rr;
            // Rr.cost < Rl.cost : remote node has better route
            // Rl.link = n : remote node is more authoritative
        }
    }
}
```

Suppose the Routers Transfer Tables as Follows:

A -> B

B -> A

B -> C

E -> C

A -> D

B -> E

Updated Routing tables

Routings from A		
To	*Link*	*Cost*
A	local	0
B	1	1
C	1	2
D	3	1
E	1	2

Routings from B		
To	*Link*	*Cost*
A	1	1
B	local	0
C	2	1
D	1	2
E	4	1

Routings from C		
To	*Link*	*Cost*
A	2	2
B	2	1
C	local	0
D	5	2
E	5	1

Routings from D		
To	*Link*	*Cost*
A	3	1
B	3	2
C	6	2
D	local	0
E	6	1

Routings from E		
To	*Link*	*Cost*
A	4	2
B	4	1
C	5	1
D	6	1
E	local	0

Simplified View of the QMW Computer Science Network(1)

240=11110000

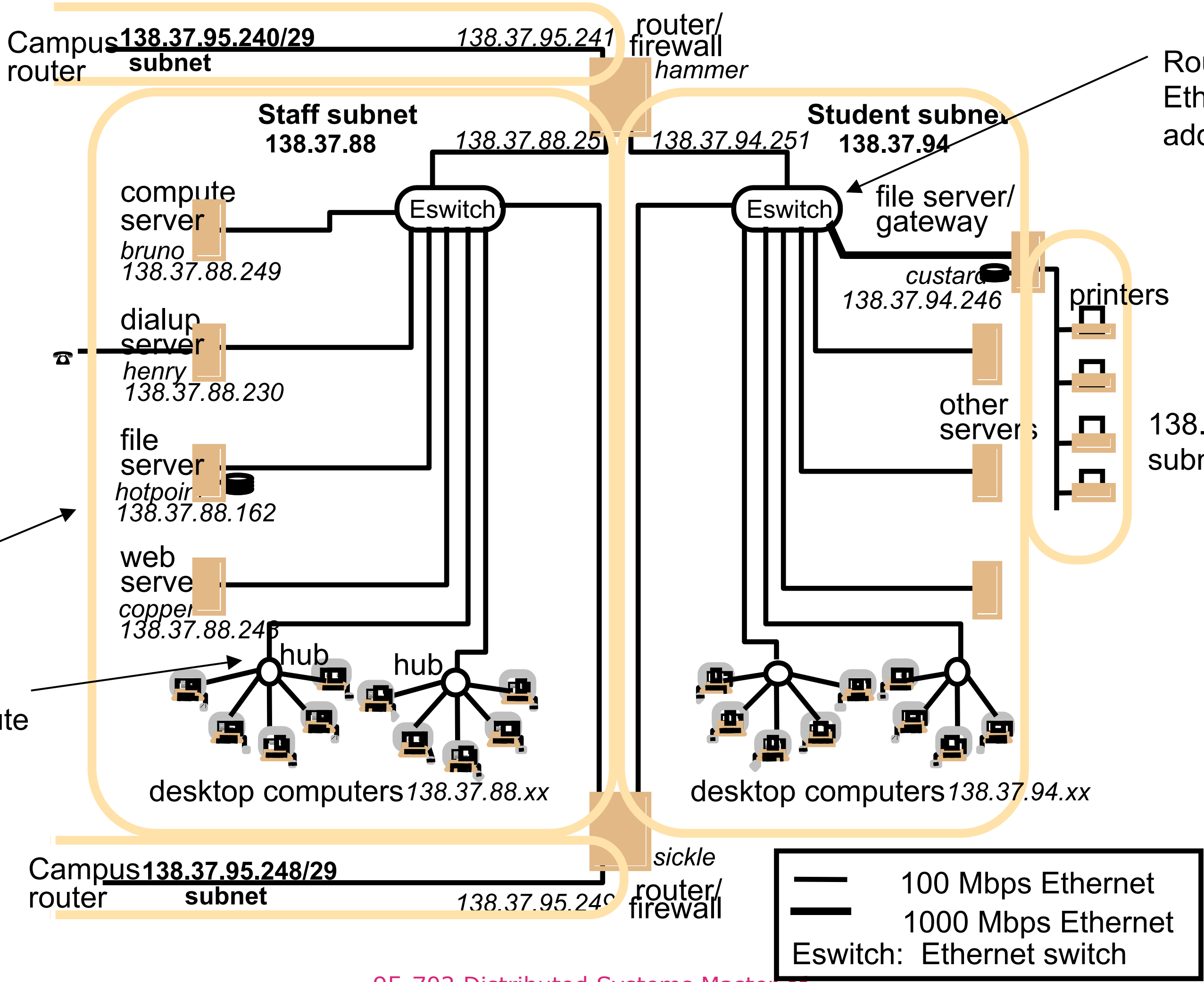
248=11111000

232=11101000

Class C
or /24

Hubs don't route

Routes at the Ethernet address level



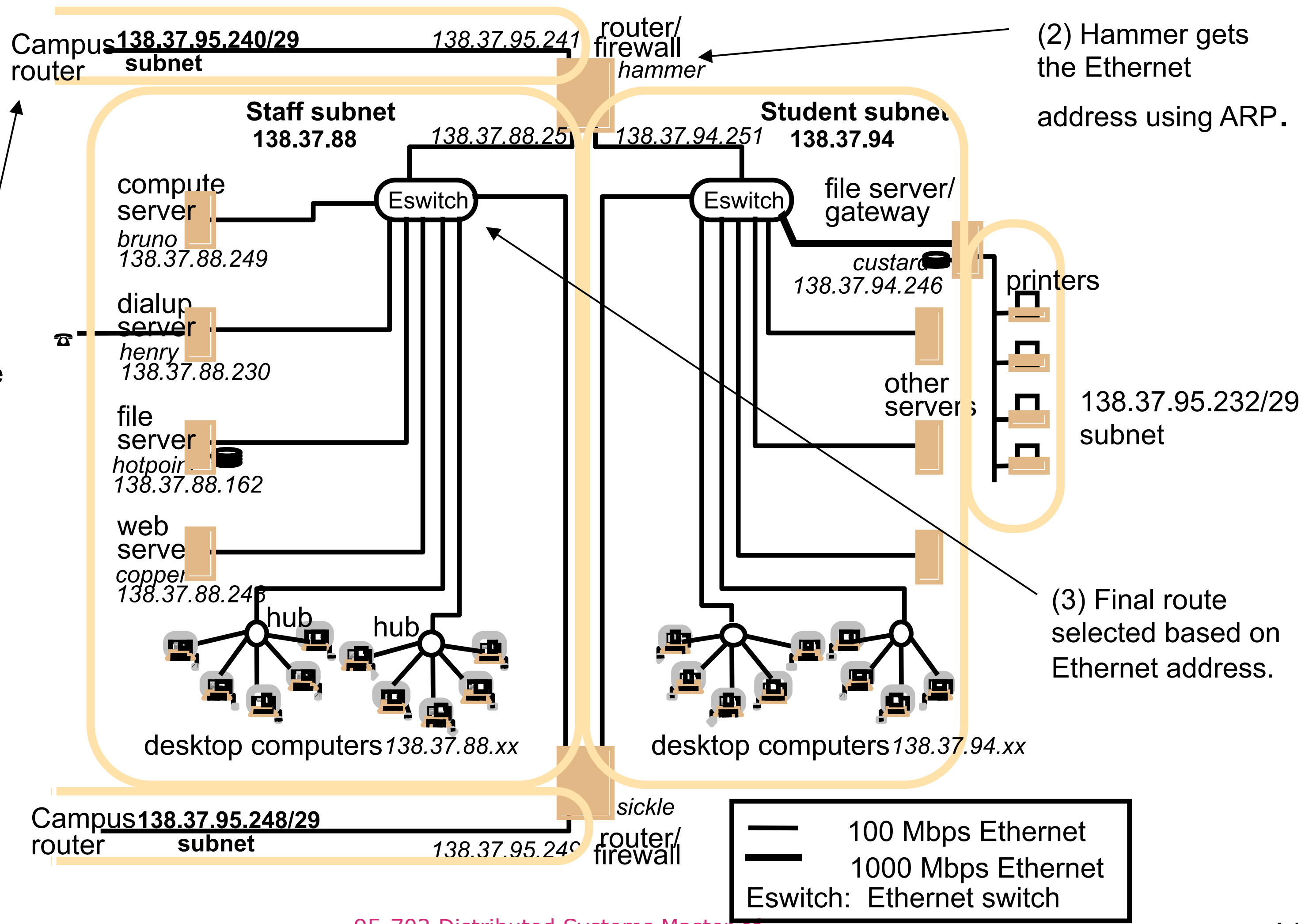
Simplified View of the QMW Computer Science Network(2)

240=11110000

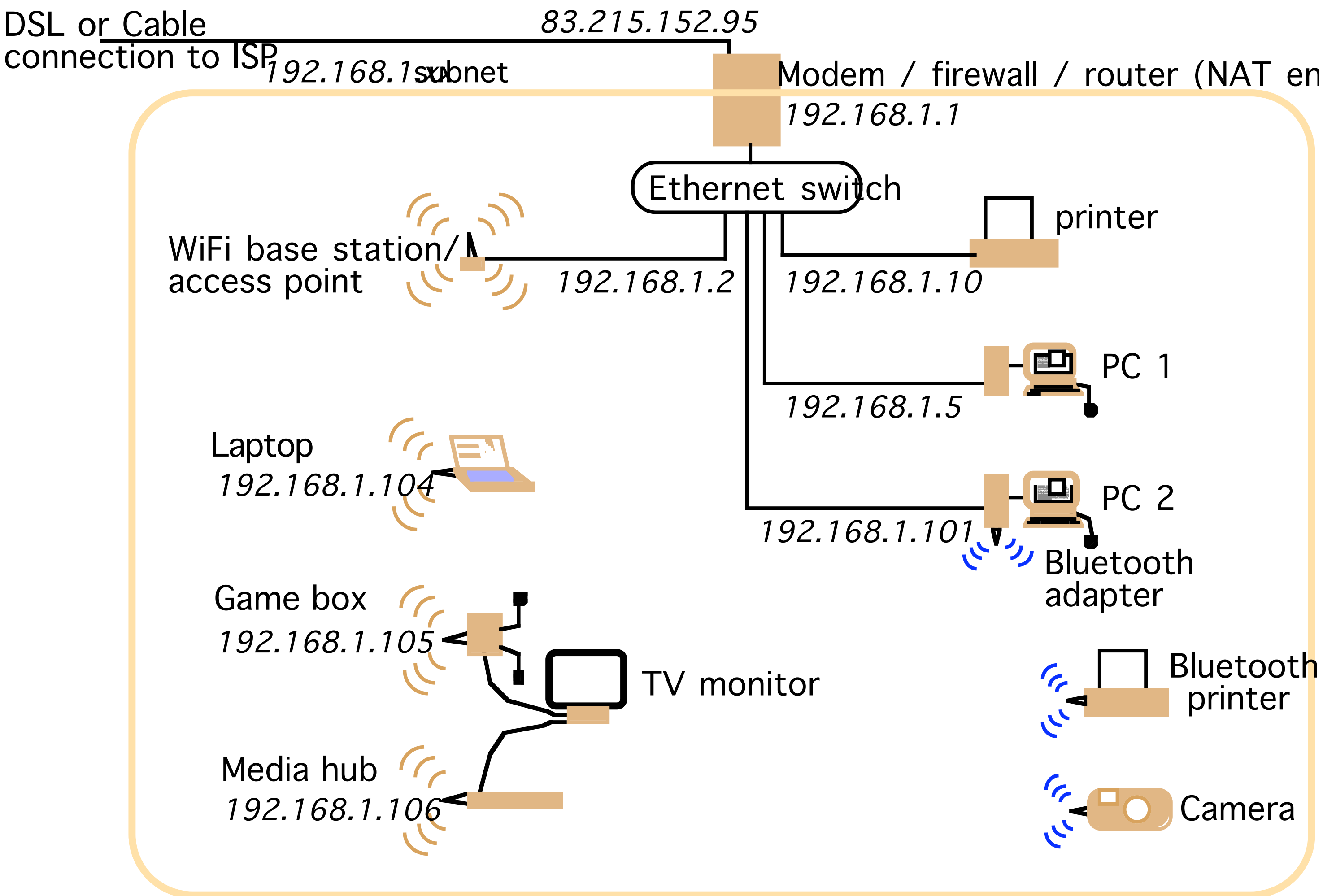
248=11111000

232=11101000

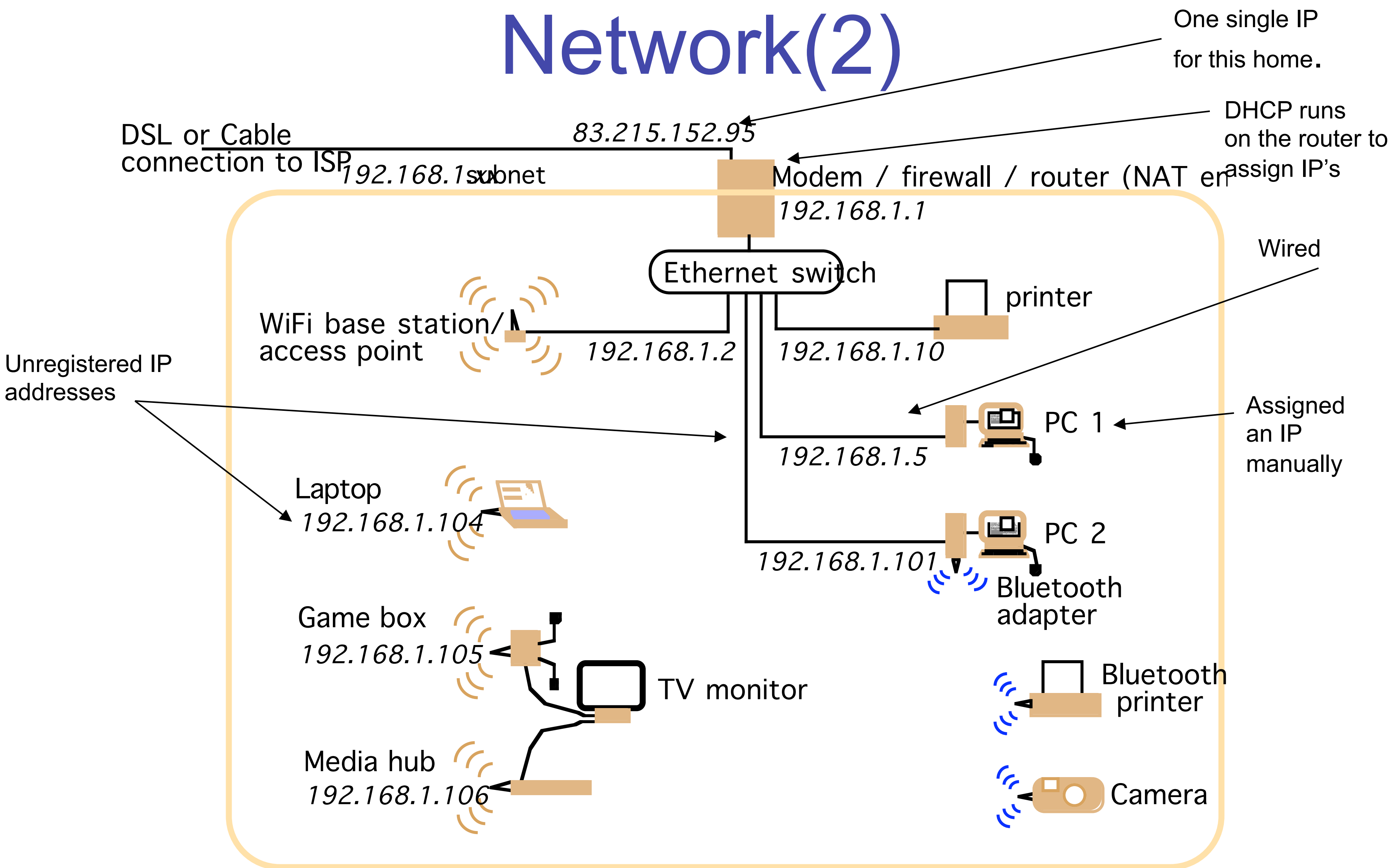
(1) Suppose we have An IP packet for Cooper 138.37.88.248



A Typical NAT-based Home Network(1)

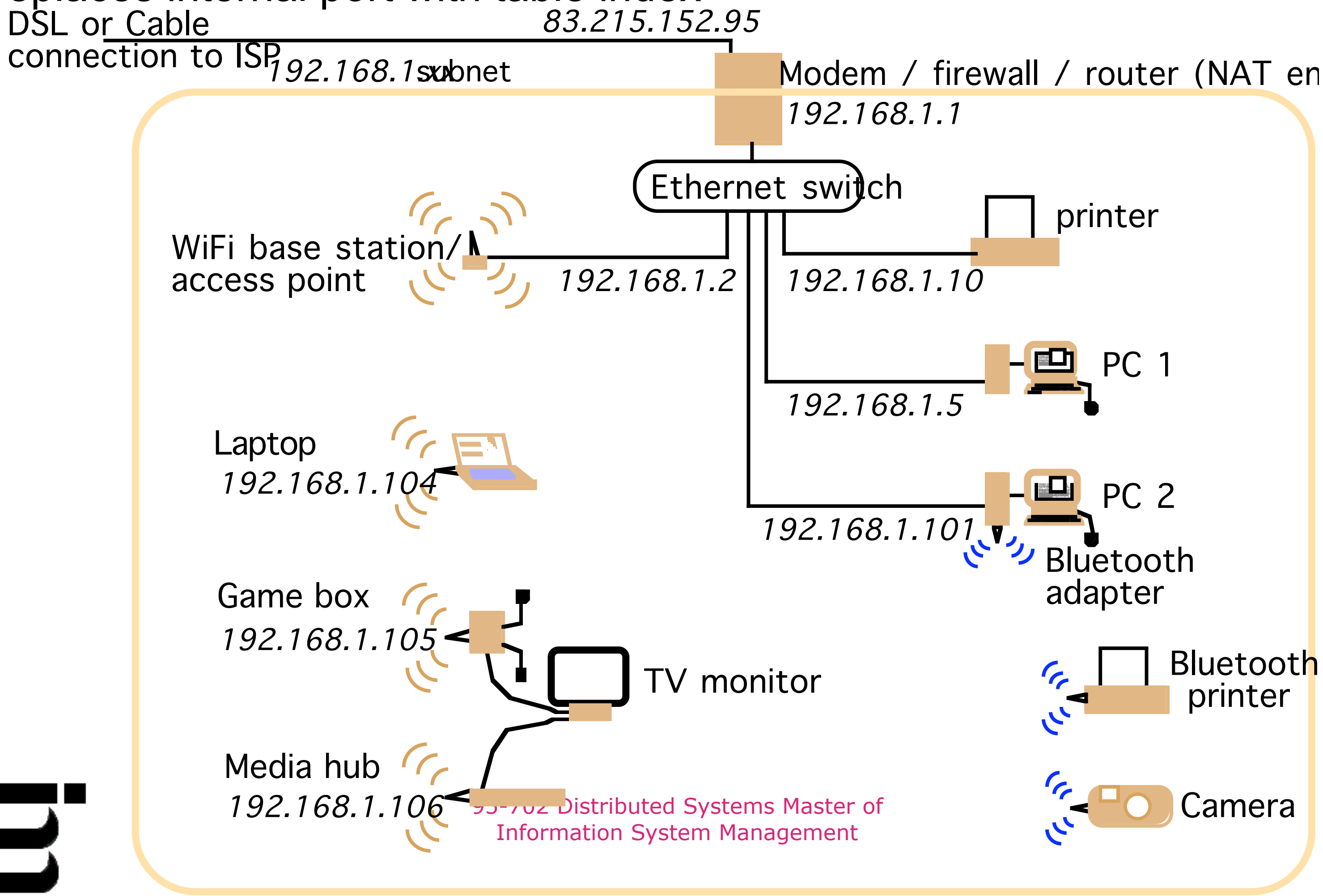


A Typical NAT-based Home Network(2)



The NAT router maintains an address translation table.
For outgoing TCP or UDP messages, modify the source IP address and port.

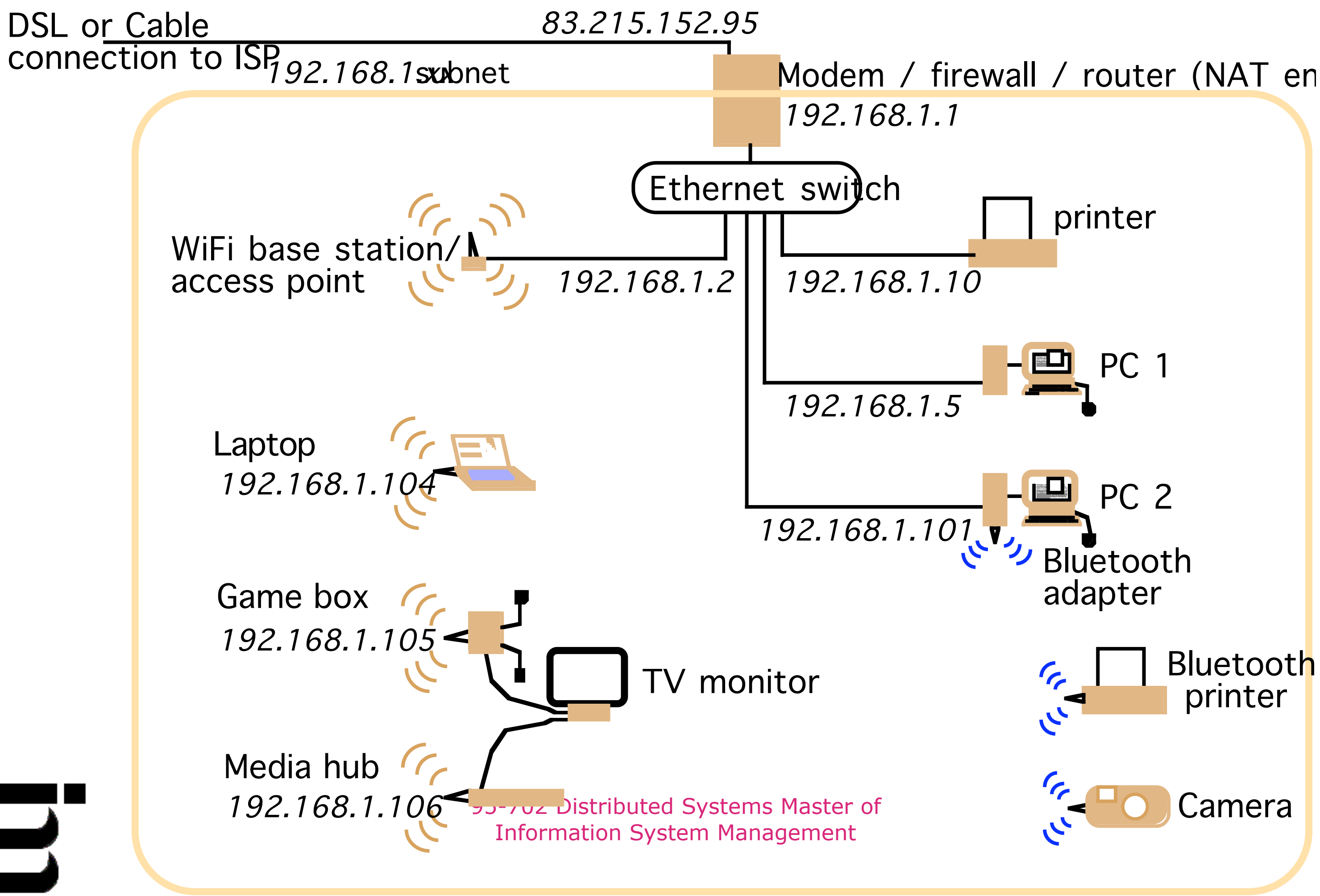
- save internal IP and Port in table
- replaces internal IP with external IP
- replaces internal port with table index



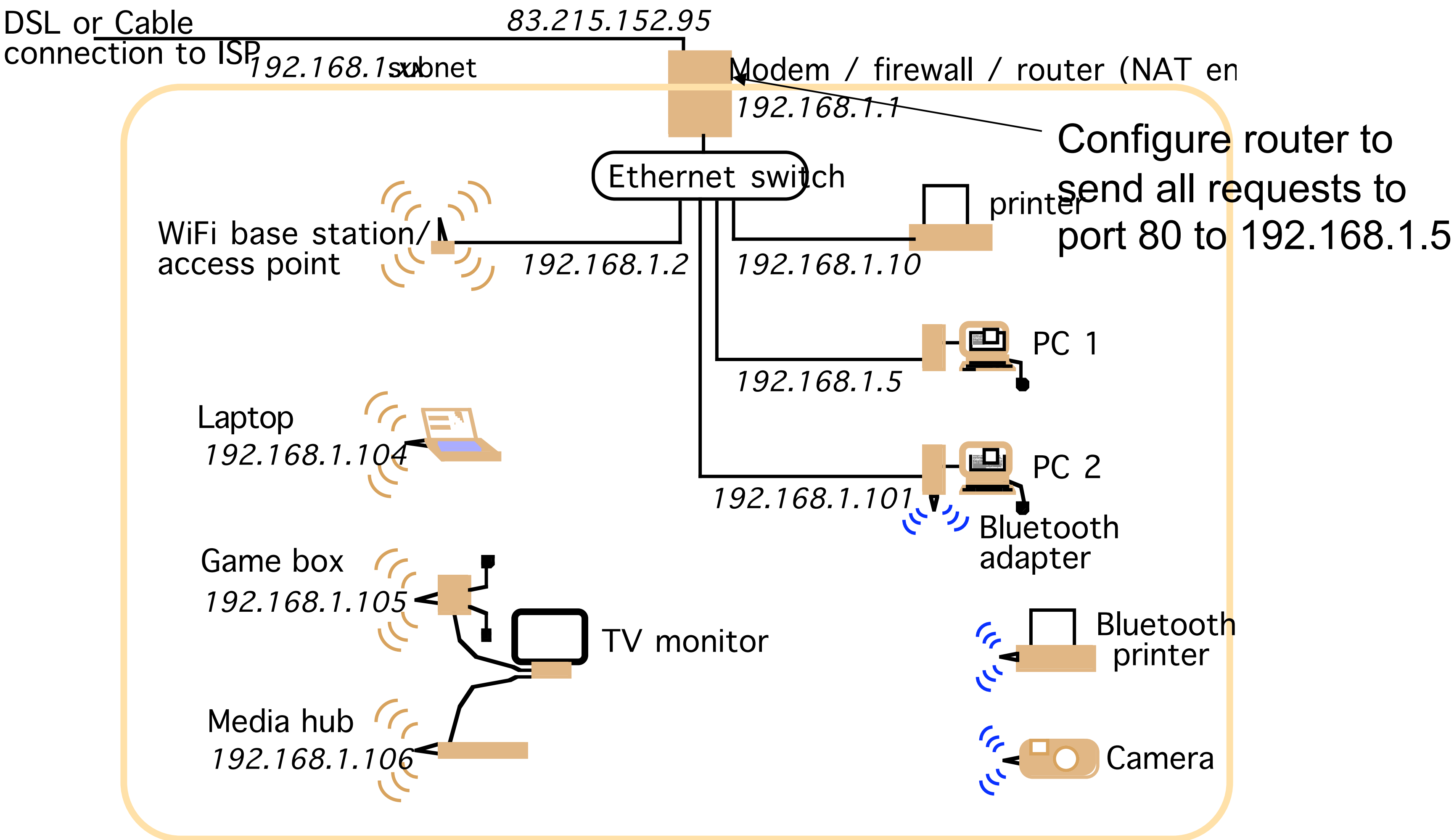
NAT router maintains an address translation table.

For incomming TCP or UDP messages:

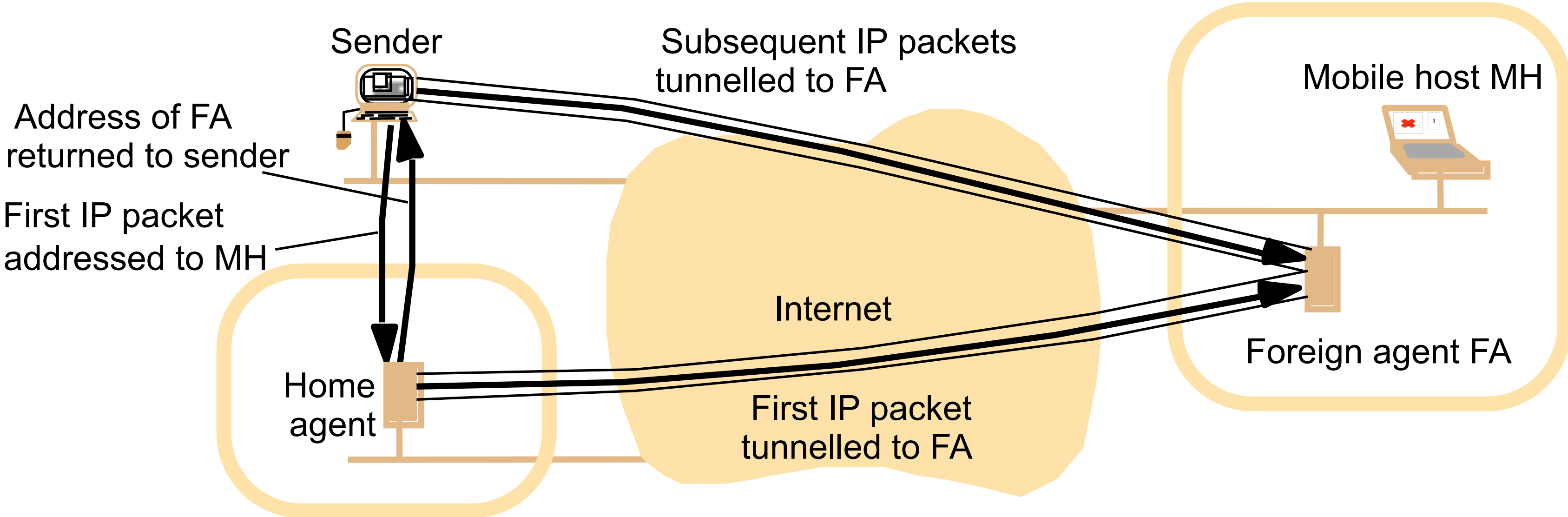
- Use the port number to look up internal address in table



But How Do We Serve?

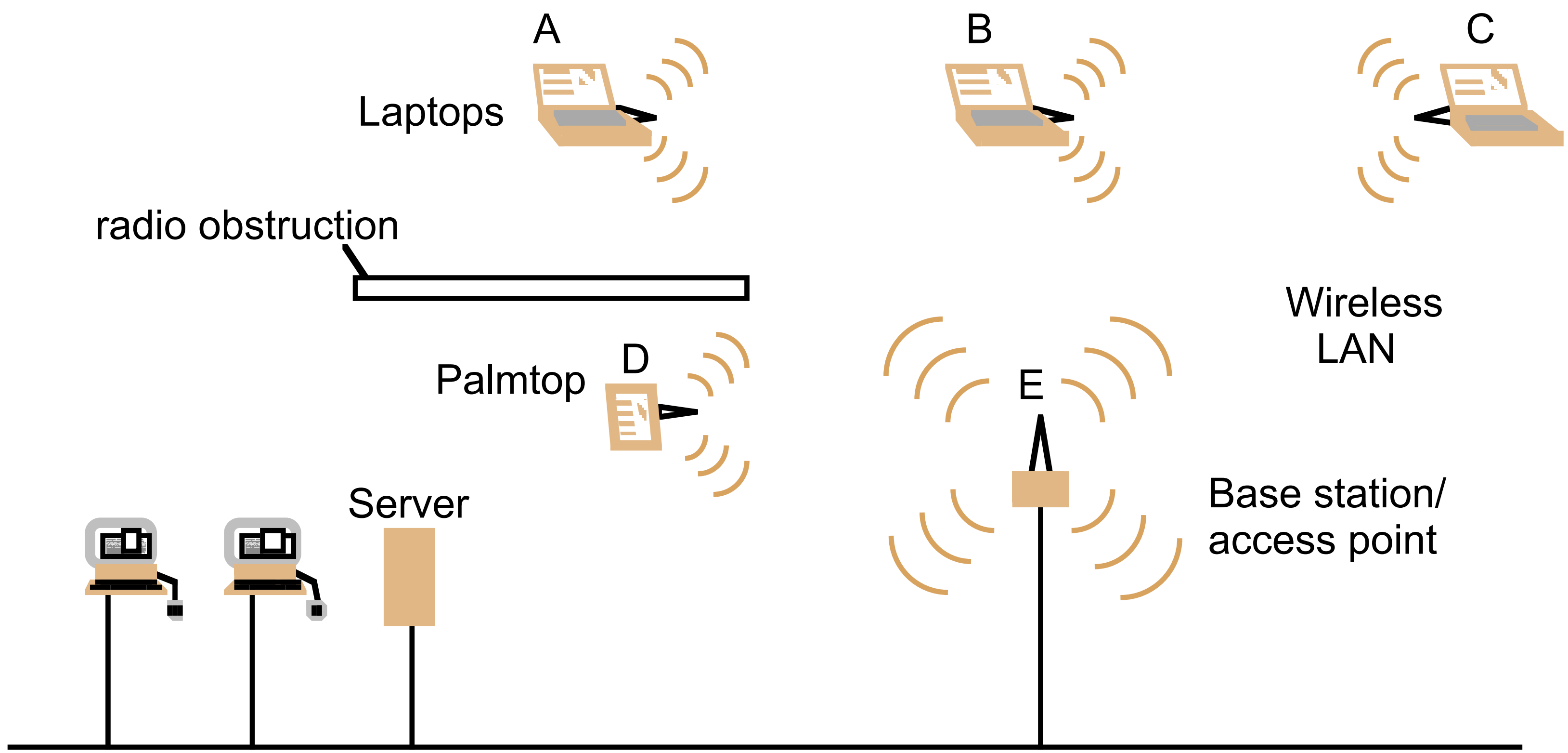


The MobileIP Routing Mechanism



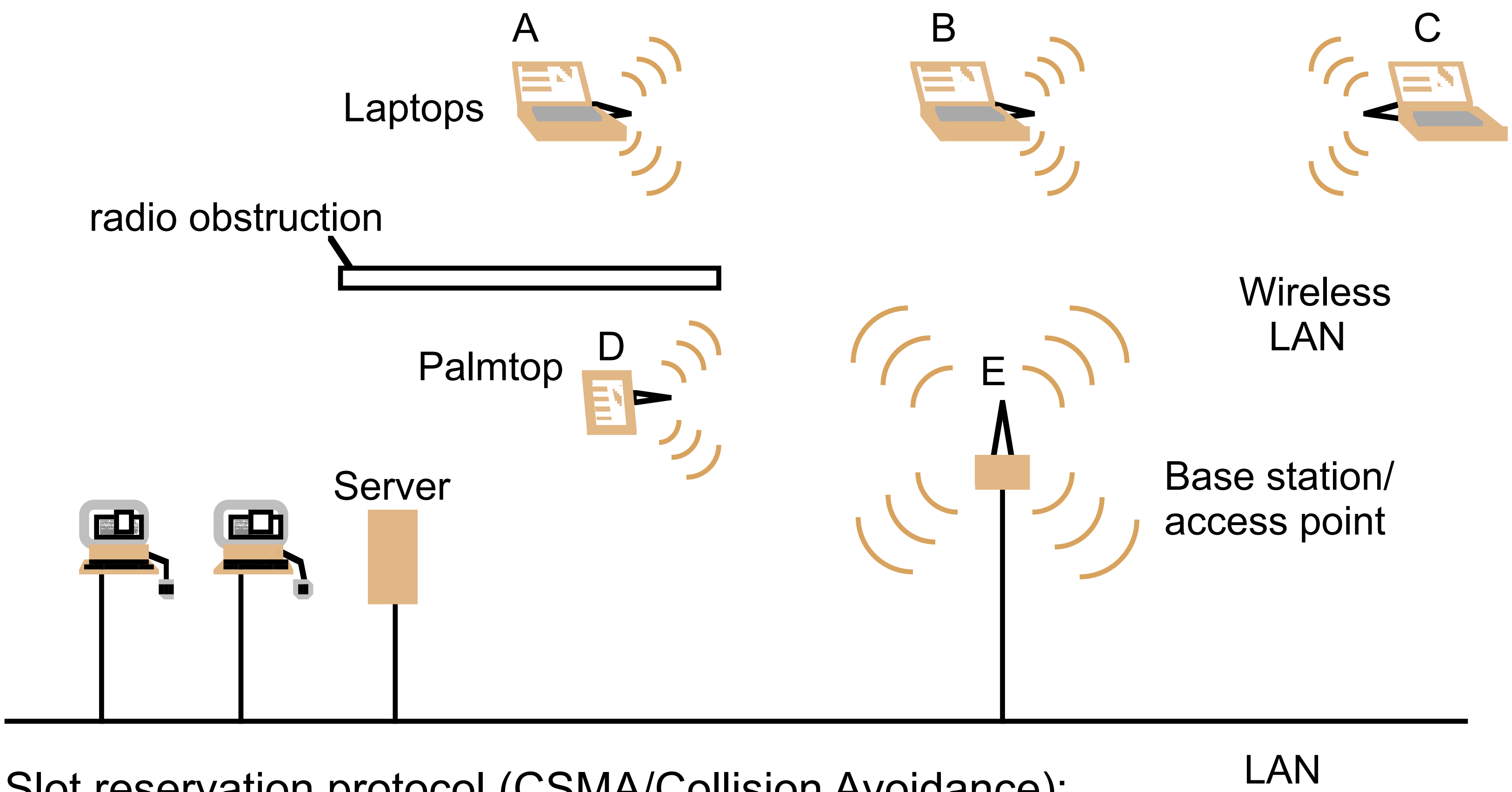
The case of a Mobile host making a request is easy – it has a new IP on the new network. No problem.
The case of the Mobile host acting as a server is described in the picture. Messages to it must be re-routed to its new home.

Wireless LAN Configuration



Challenges to the CSMA/CD approach:
  Hidden stations: A may not be able to sense D's signal to E.
  Fading: A may not be able to detect a transmission by C.
  Collision Masking: Locally generated signals are stronger than distant signals.

Wireless LAN Configuration



Slot reservation protocol (CSMA/Collision Avoidance):

A sends a request to send (RTS) message carrying a duration to E.

E responds with a clear to send (CTS) message repeating the duration.

All those near A or E back off for that period.

Multimedia Applications

- Typically divided into two types: conferencing applications and streaming applications.
- See the vat tool for audio conferencing.
- See the vic tool for video conferencing.
- Streaming applications deliver an audio or video stream.
- See Real Audio for a commercial stream application.
- Real-Time Transport Protocol (RTP) commonly runs over UDP.