COMP2001J Computer Networks

Lecture 7 – Network Layer
(IP address)

Dr. Shen WANG (王燊)

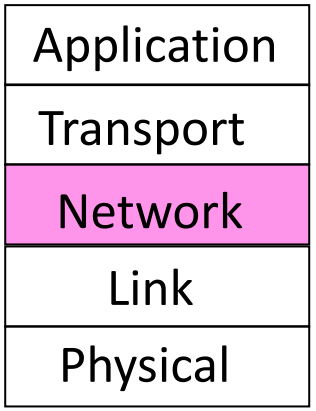
shen.wang@ucd.ie

Outline

- Introduction
- MAC/IP/Port address
- IP address
  - Prefixes
  - Subnets
  - CIDR
    - route aggregation
    - longest matching prefix
  - Class
- ARP
- ICMP

Network Layer

- Responsible for delivering packets between endpoints over multiple links.
  - It clearly contrasts with that of the data link layer, which has the more modest goal of just moving frames from one end of a wire to the other (point-to-point).
  - Thus, the network layer is the lowest layer that deals with **end-to-end** transmission.



Network Layer

- To achieve its goals, the network layer must know about the topology of the network (i.e., the set of all routers and links) and choose appropriate paths through it, even for large networks.
- It must also take care when choosing routes to avoid overloading some of the communication lines and routers while leaving others idle.
- Finally, when the source and destination are in different networks, new problems occur. It is up to the network layer to deal with them.

Network Layer Services

- The network layer provides services to the transport layer.
- The services need to be carefully designed with the following goals in mind:
  - The services should be independent of the router technology.
  - The transport layer should be shielded from the number, type, and topology of the routers present.
  - The network addresses made available to the transport layer should use a uniform numbering plan, even across LANs and WANs.

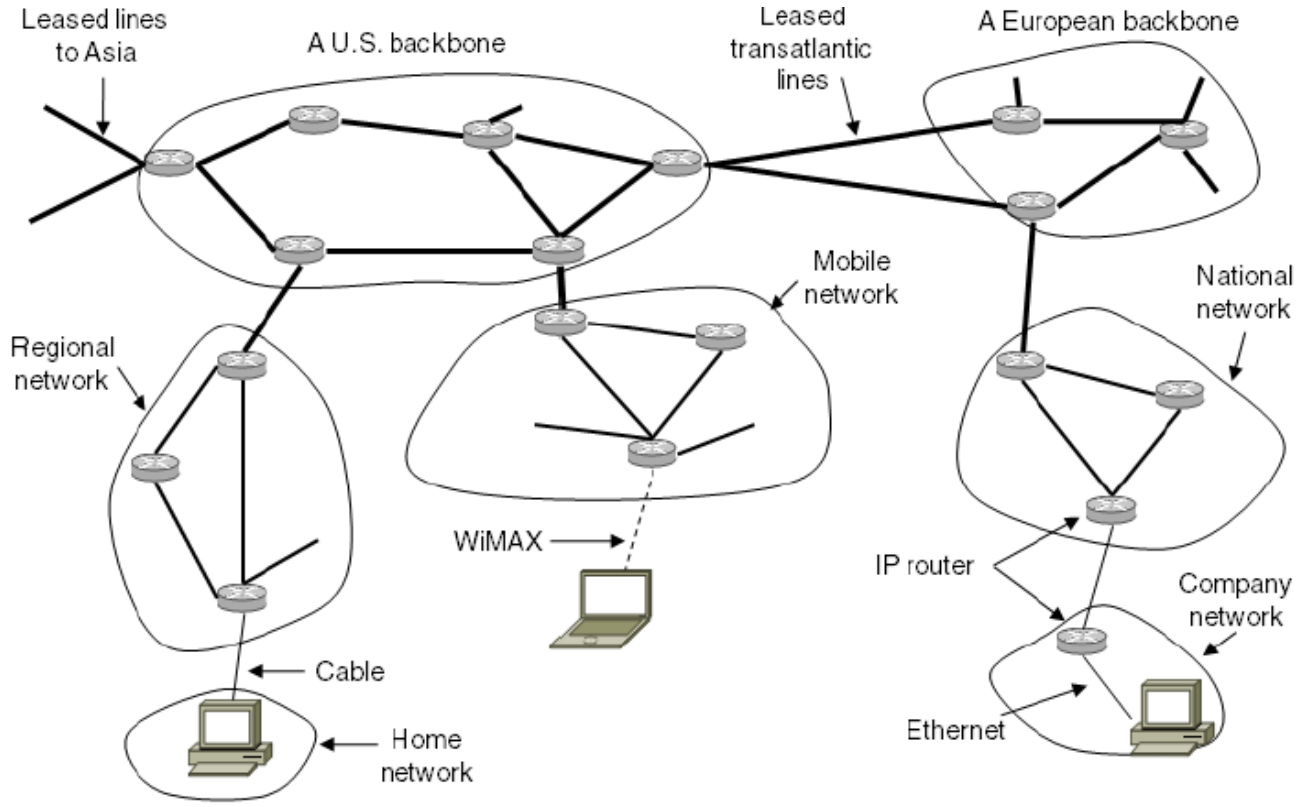
How Networks Differ

- Differences can be large; complicates internetworking

Item	Some Possibilities
Service offered	Connectionless versus connection oriented
Addressing	Different sizes, flat or hierarchical
Broadcasting	Present or absent (also multicast)
Packet size	Every network has its own maximum
Ordering	Ordered and unordered delivery
Quality of service	Present or absent; many different kinds
Reliability	Different levels of loss
Security	Privacy rules, encryption, etc.
Parameters	Different timeouts, flow specifications, etc.
Accounting	By connect time, packet, byte, or not at all

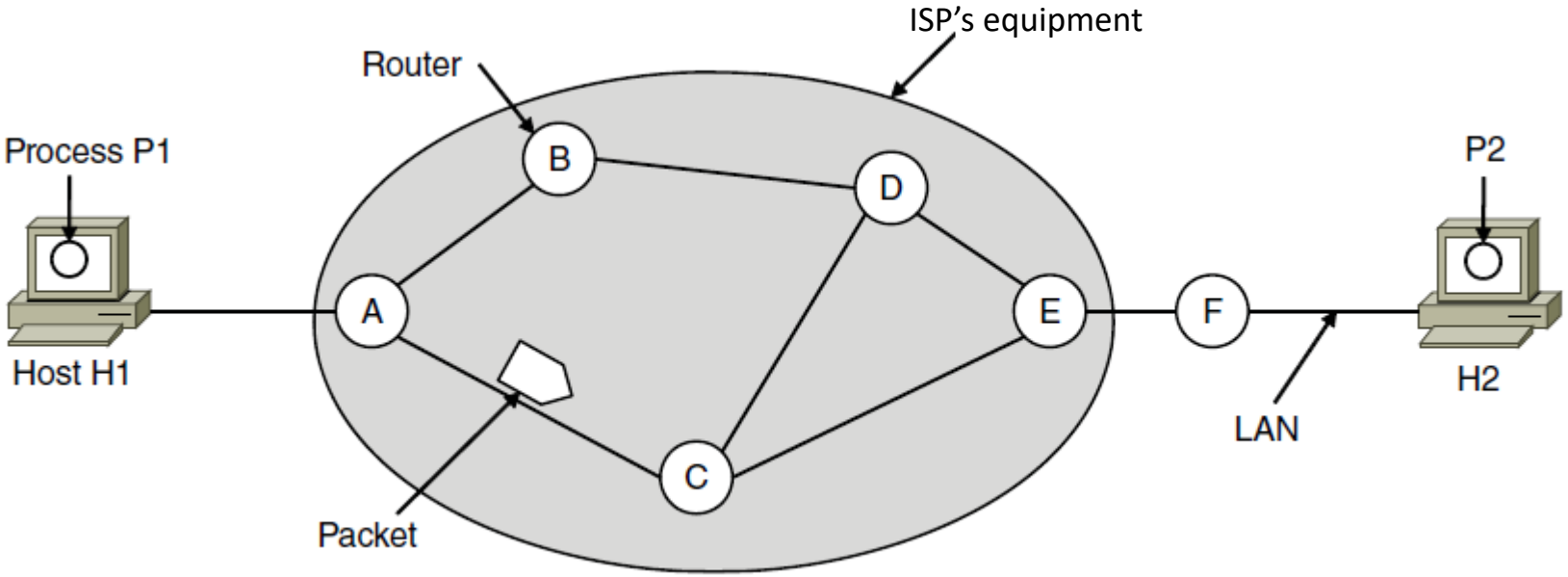
Network Layer in the Internet

- Internet is an interconnected collection of <u>many networks</u> that is held together by the IP protocol



Store-and-Forward Packet Switching

- Hosts send packets into the network; packets are forwarded by routers
  - Routers treat packets as messages, receiving (storing) them and then forwarding them based on how the message is addressed.
  - For completeness, it is a process running on the host that sends the packet into the network and receives packets at the destination.

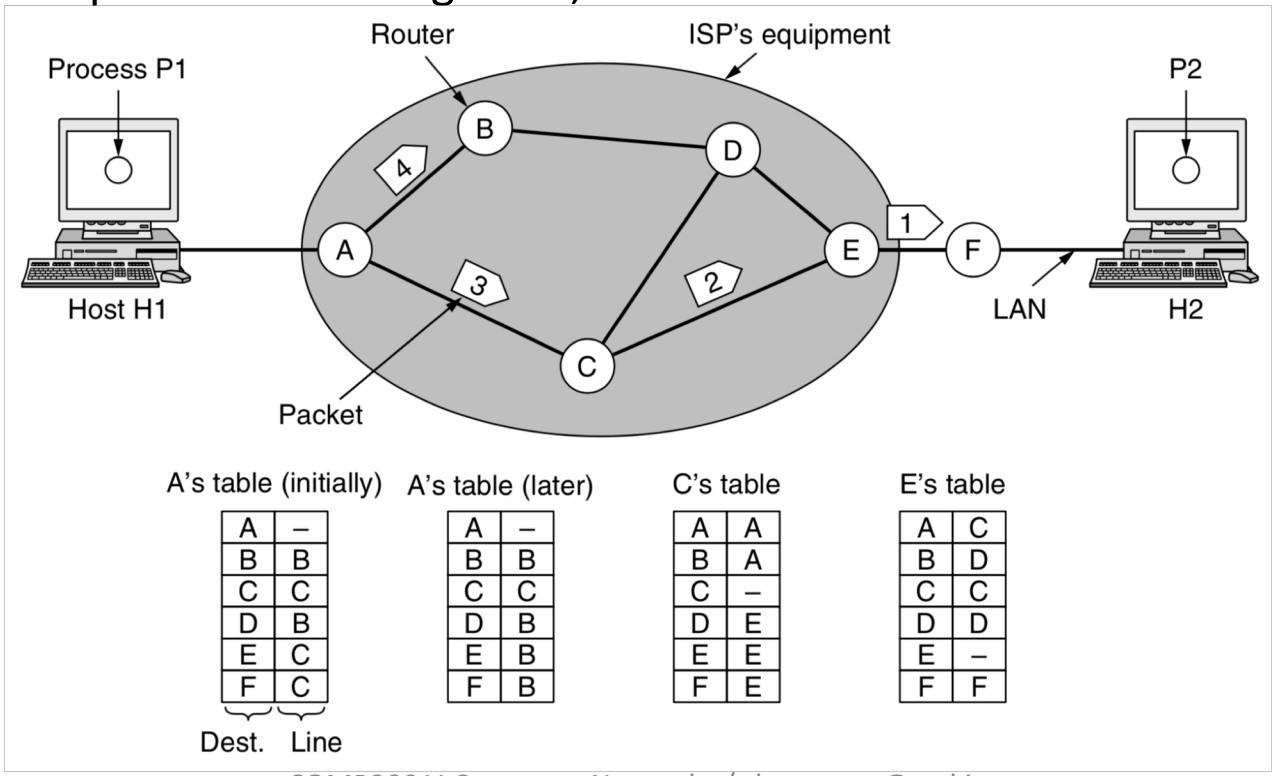


Connectionless Service – Datagrams

- Packet is forwarded using destination address inside it
  - Different packets may take different paths
  - This model is like the postal service – each letter is sent through the network independently.
  - Datagram is a packet that contains an absolute destination address
  - Routers need only look up the destination address in a table to find the outgoing line to send the packet on its way.

Connectionless Service – Datagrams

- E.g: P1 on H1 sends to P2 on H2
  - A's forwarding table changes for the entries that target on E and F
  - Perhaps it has learned of a traffic jam somewhere along the *ACE* path and updated its routing table, as shown under the label ''later.''



Dest.	Line
A	–
B	B
C	C
D	B
E	C
F	C

Dest.	Line
A	–
B	B
C	C
D	B
E	B
F	B

Dest.	Line
A	A
B	A
C	–
D	E
E	E
F	E

Dest.	Line
A	C
B	D
C	C
D	D
E	–
F	F

Addressing

- MAC address
- IP address
- Port address

MAC (Ethernet/Physical) address

- Data link layer (point-to-point)
- A 48-bit binary number.
  - 2^{48} -> 281,474,976,710,656
  - If all bits are “1”, it is a broadcast MAC address that allows a device to address every device in its LAN.
- This is a world-wide unique identifier that bounds to the NIC (network interface card) when manufactured.
  - It is fixed for that NIC and will stay the same for its lifetime
- Use the Hex pair notation for human readable purposes.

```
100100011100111100100110111010000011010110100001
                              ↓
                    91:CF:26:E8:35:A1
```

IP address

- Network layer (end-to-end)
- A 32-bit binary number.
  - 2^{32} -> 4,294,967,296
- It is a hierarchical address, composed of network id and host id. Top-tier IP addresses are assigned by ICANN (Internet Corporation for Assigned Names and Numbers). Others can be decided by network designers so that:
  - All devices in the same network (e.g. LAN) should share the same network id of their IP addresses.
  - Each port or interface of network devices should have a unique address in the Internet.
- Usually represented in decimal dot notation:

```
11000000.10101000.00110111.000000101
                    ↓
            192.168.55.5
```

Port address

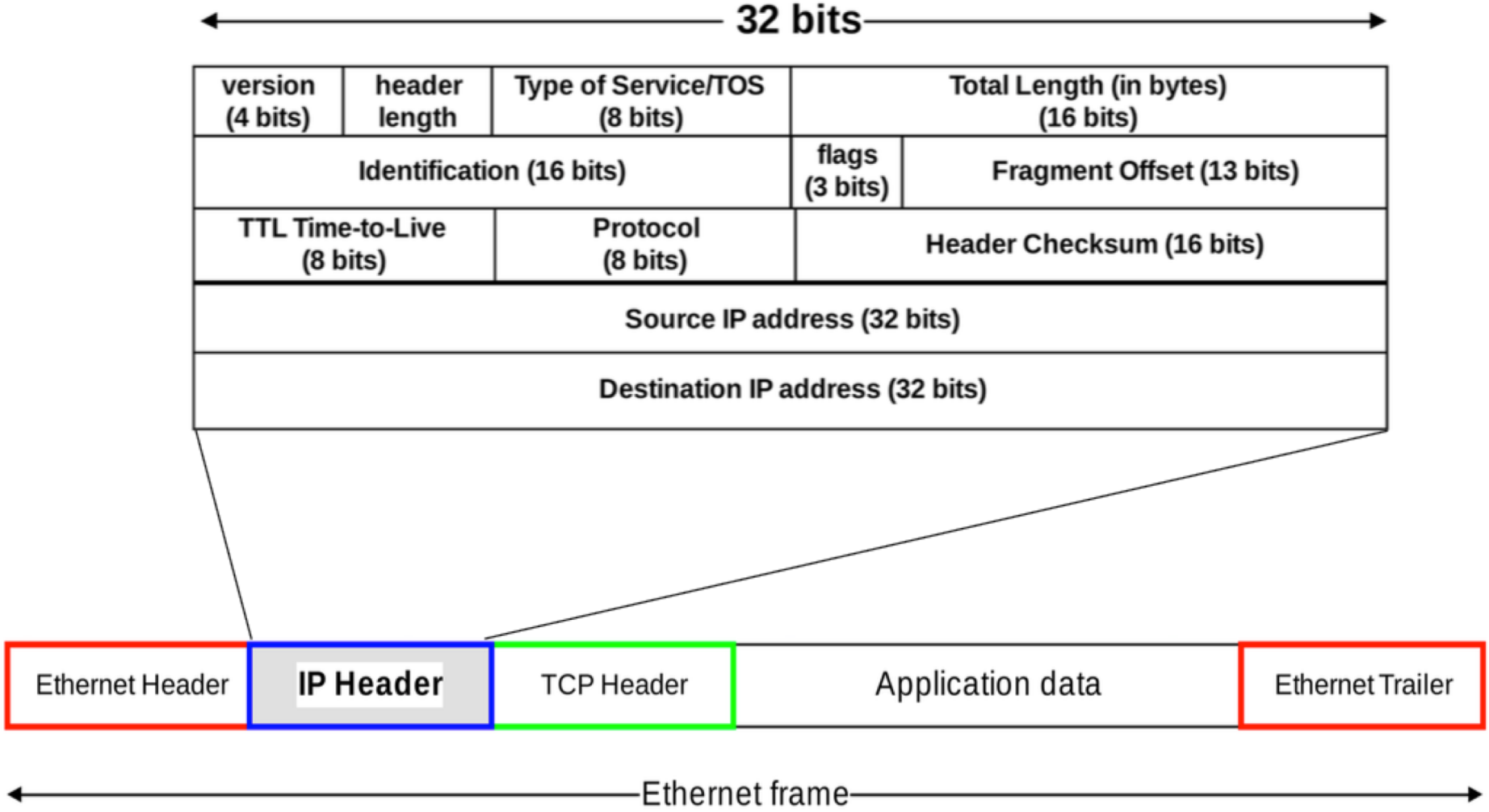
- Transport layer (end-to-end)
- A 16-bit binary number.
  - 2^{16} -> 65,536
  - Often presented in an integer number.
- The address of a process (e.g. application) that is unique locally on each network device.
- Ports 0-1023 well-known ports by The Internet Assigned Number Authority (IANA)

Some well-known ports

- Over 700 assigned already.
- To avoid overlaps, other ports from 1024 through 49151 can be registered with IANA for use by unprivileged users.
- The rest (49152-65535) are used for client to choose flexibly for temporary use.

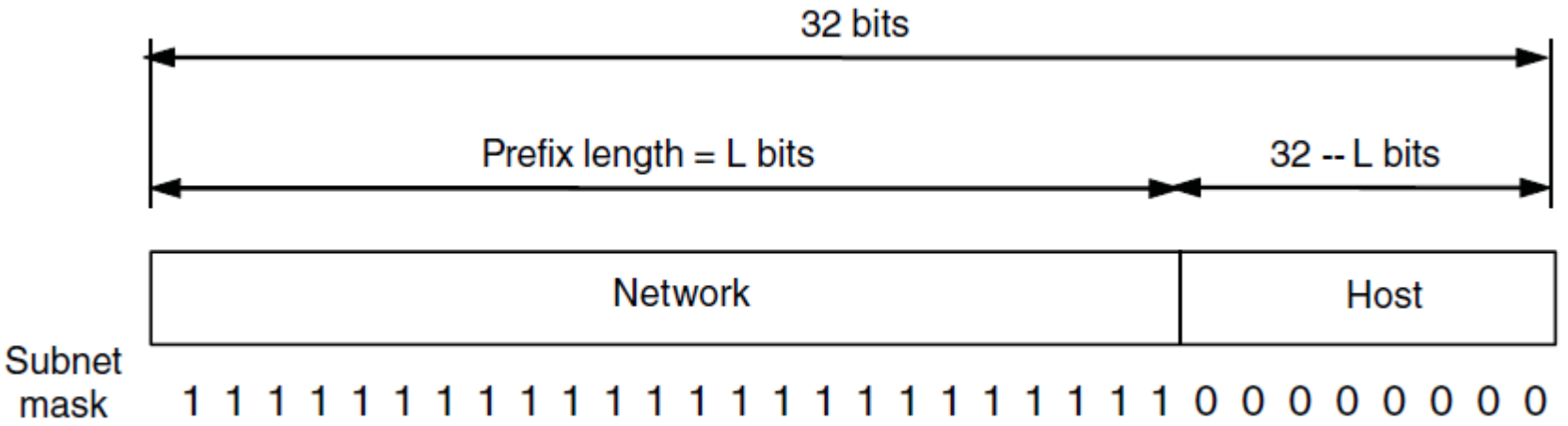
Port	Protocol	Use
20, 21	FTP	File transfer
22	SSH	Remote login, replacement for Telnet
25	SMTP	Email
80	HTTP	World Wide Web
110	POP-3	Remote email access
143	IMAP	Remote email access
443	HTTPS	Secure Web (HTTP over SSL/TLS)
543	RTSP	Media player control
631	IPP	Printer sharing

Where are these addresses?



Prefixes

- Addresses are allocated in blocks called prefixes
  - Prefix is determined by the network portion
  - Has 2^L addresses aligned on 2^L boundary
  - Written address/length, e.g., 18.0.31.0/24
  - Can also be written using a **subnet mask:** the length of the prefix corresponds to a binary mask of 1s in the network portion. It can be ANDed with the IP address to extract only the network portion.



Hierarchical Address - Pros

- Routers can forward packets based on only the network portion of the address, as long as each of the networks has a unique address block.
  - The host portion does not matter to the routers because all hosts on the same network will be sent in the same direction. It is only when the packets reach the network for which they are destined that they are forwarded to the correct host.

- This makes the routing tables much smaller than they would otherwise be.
  - Consider that the number of hosts on the Internet is approaching one billion. That would be a very large table for every router to keep. However, by using a hierarchy, routers need to keep routes for only around 300,000 prefixes.

Hierarchical Address - Cons

- Unlike MAC address, the IP address of a host depends on where it is located in the network.
  - Once a host moves to another network, it's IP address should be reconfigured. E.g. a special design is needed to support "mobile IP" that even a host moves between networks it can use the same IP address.
- The hierarchy is wasteful of addresses unless it is carefully managed.
  - If addresses are assigned to networks in (too) large blocks, there will be (many) addresses that are allocated but not in use.

IP address allocation

- ICANN delegates parts of the address space to various regional authorities, which dole out IP addresses to ISPs and other companies.
- Routing by prefix requires all the hosts in a network to have the same network number. This property can cause problems as networks grow.

IP address allocation

- Consider a university that started out with our example **/16** prefix for use by the Computer Science Dept. for the computers on its Ethernet. A year later, the Electrical Engineering Dept. wants to get on the Internet. The Art Dept. soon follows suit. What IP addresses should these departments use?
  - Once /16 prefix is allocated, it is difficult to change. Getting further blocks requires going outside the university and may be expensive or inconvenient.
  - Moreover, the /16 already allocated has enough addresses for over 60,000 hosts. It might be intended to allow for significant growth, but until that happens, it is wasteful to allocate further blocks of IP addresses to the same university.

Subnets

- **Subnetting** allows the block of addresses to be split into several parts for internal use as multiple networks(**subnets**), while still acting like a single network to the outside world.

- The single /16 can be split into pieces.
  - This split does not need to be even
  - but each piece must be aligned so that any bits can be used in the lower host portion.

Subnets

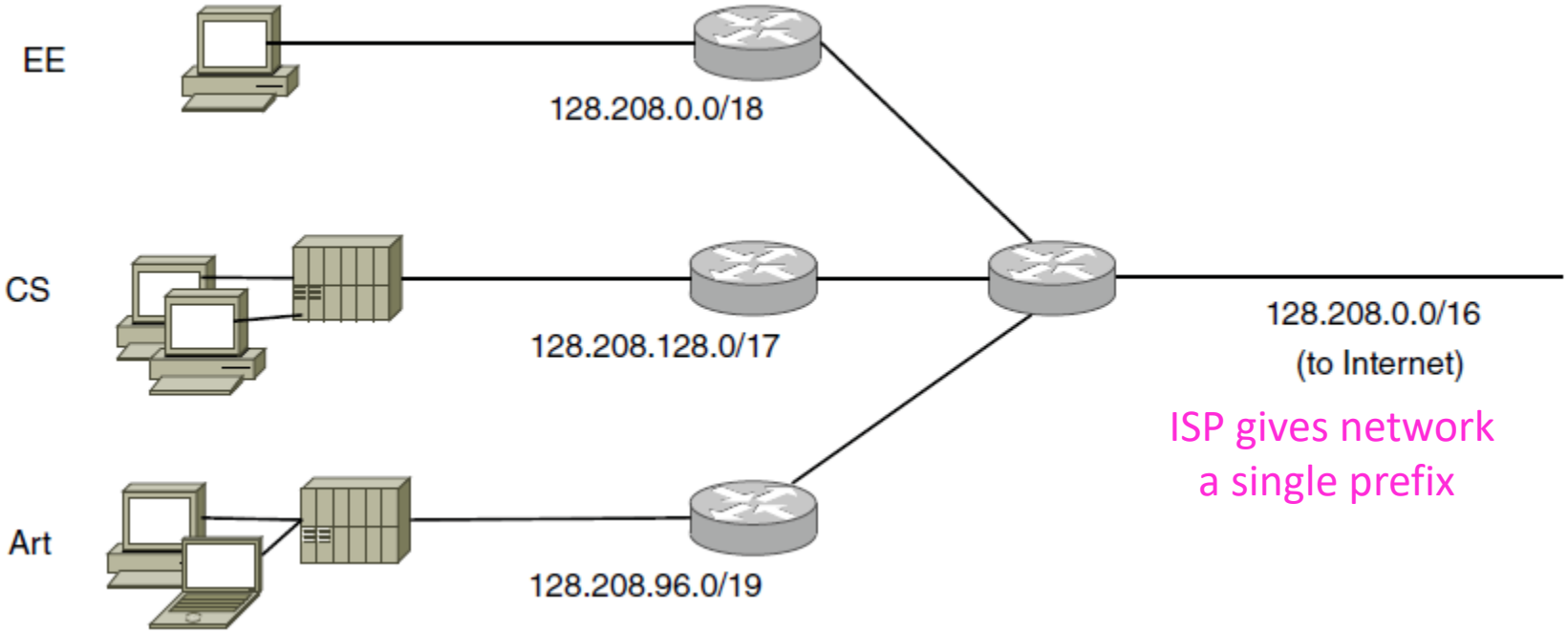
- Half of the block (a /17) is allocated to the Computer Science Dept
- A quarter is allocated to the Electrical Engineering Dept. (a /18)
- One eighth (a /19) to the Art Dept.
- The remaining eighth is unallocated.
- A different way to see how the block was divided is to look at the resulting prefixes when written in binary notation.

	128	208		
Computer Science:	10000000	11010000	1|xxxxxxx	xxxxxxxx
Electrical Eng.:	10000000	11010000	00|xxxxxx	xxxxxxxx
Art:	10000000	11010000	011|xxxxx	xxxxxxxx

- Here, the vertical bar (|) shows the boundary between the subnet number and the host portion.

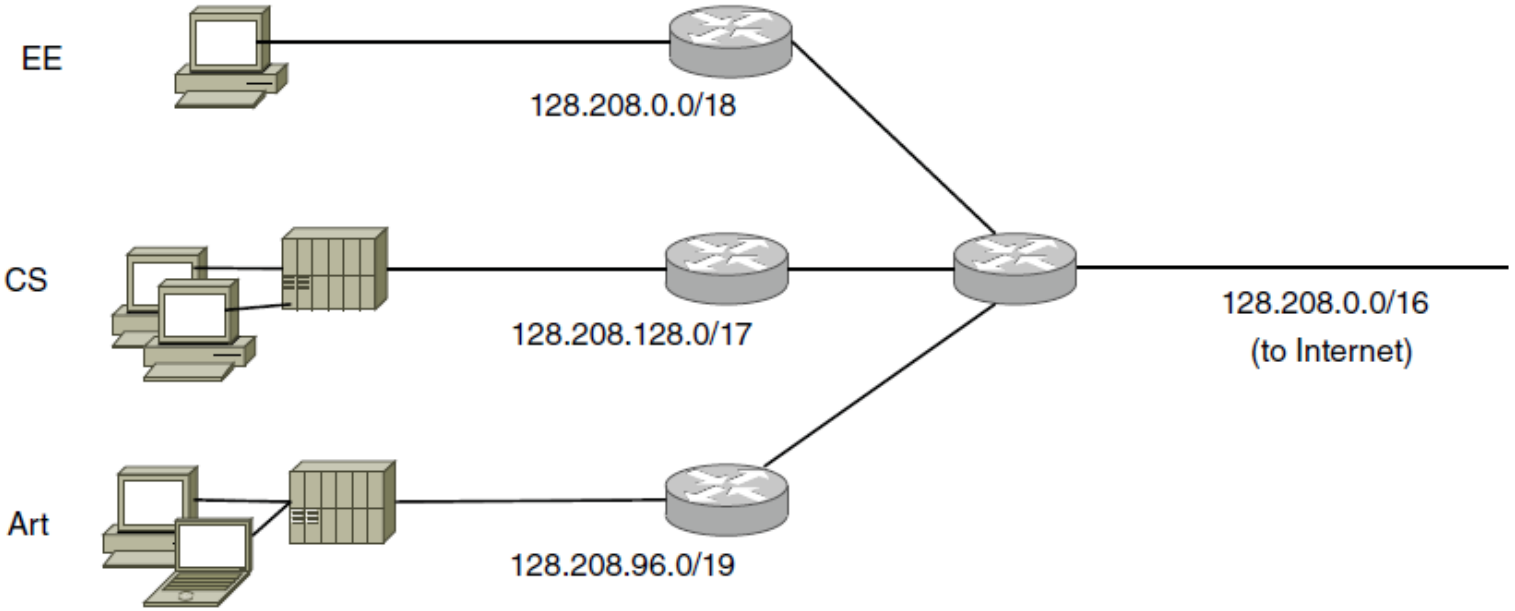
Subnets

- Subnetting splits up IP prefix to help with management
  - Looks like a single prefix outside the network



Subnets

- When a packet arrives, the router looks at the destination address of the packet and checks which subnet it belongs to.
  - The router can do this by ANDing the destination address with the mask for each subnet and checking to see if the result is the corresponding prefix.
- E.g. A packet to 128.208.2.151:
  - AND 255.255.128.0 (/17, CS) -> 128.208.0.0: not match
  - AND 255.255.192.0 (/18, EE) -> 128.208.0.0; match



Subnets

- The subnet divisions can be changed later if necessary, by updating all subnet masks at routers inside the university.

- Outside the network, the subnetting is not visible, so allocating a new subnet does not require contacting ICANN or changing any external databases.

Routing table explosion

- Routers in ISPs and backbones in the middle of the Internet (called **default-free zone**) must know which way to go to get to every network (probably at least a million). This can make for a very large table.
- Routers must perform a lookup in this table to forward every packet, and routers at large ISPs may forward up to millions of packets per second.
- Routing algorithms require each router to exchange information about the addresses it can reach with other routers. The larger the tables, the more information needs to be communicated and processed. This increases the likelihood that some parts will get lost, at least temporarily, possibly leading to routing instabilities.

Route aggregation

- Instead of splitting an address block into subnets, we combine multiple small prefixes into a single larger prefix.
  - This process is called **route aggregation**.
  - The resulting larger prefix is sometimes called a **supernet.**
- With aggregation, IP addresses are contained in prefixes of varying sizes. The same IP address that
  - one router treats as part of a /22 (a block containing 2^{10} addresses)
  - may be treated by another router as part of a larger /20 (which contains 2^{12} addresses).
  - It is up to each router to have the corresponding prefix information.
- This design works with subnetting and is called **CIDR** (**Classless Inter-Domain Routing**)

CIDR—Example

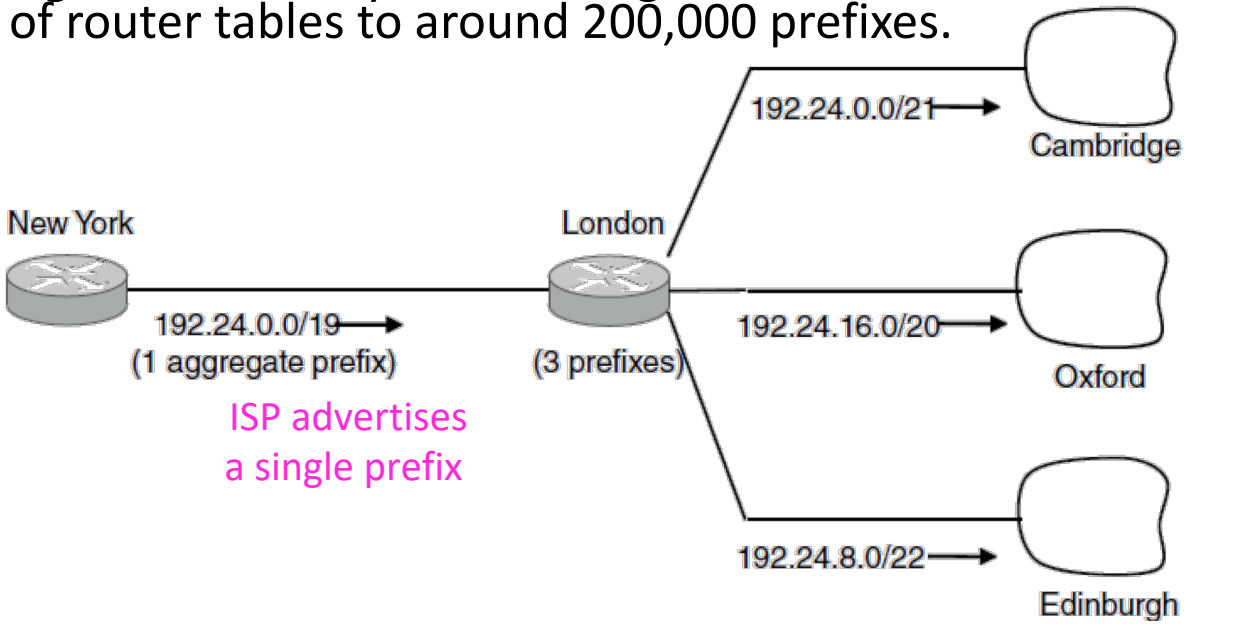
WARNING: Convert necessary parts of IP addresses into binary representations!

- A block of 8192 IP addresses is available starting at 194.24.0.0.
- Cambridge University needs 2048 addresses.
  - It is assigned the addresses 194.24.0.0 through 194.24.7.255, along with mask 255.255.248.0. This is a /21 prefix.
- Oxford University asks for 4096 addresses.
  - Since a block of 4096 addresses must lie on a 4096-byte boundary, Oxford cannot be given addresses starting at 194.24.8.0. Instead, it gets 194.24.16.0 through 194.24.31.255, along with subnet mask 255.255.240.0.
- Finally, the University of Edinburgh asks for 1024 addresses.
  - It is assigned addresses 194.24.8.0 through 194.24.11.255 and mask 255.255.252.0.

University	First address	Last address	How many	Prefix
Cambridge	194.24.0.0	194.24.7.255	2048	194.24.0.0/21
Edinburgh	194.24.8.0	194.24.11.255	1024	194.24.8.0/22
(Available)	194.24.12.0	194.24.15.255	1024	194.24.12.0/22
Oxford	194.24.16.0	194.24.31.255	4096	194.24.16.0/20

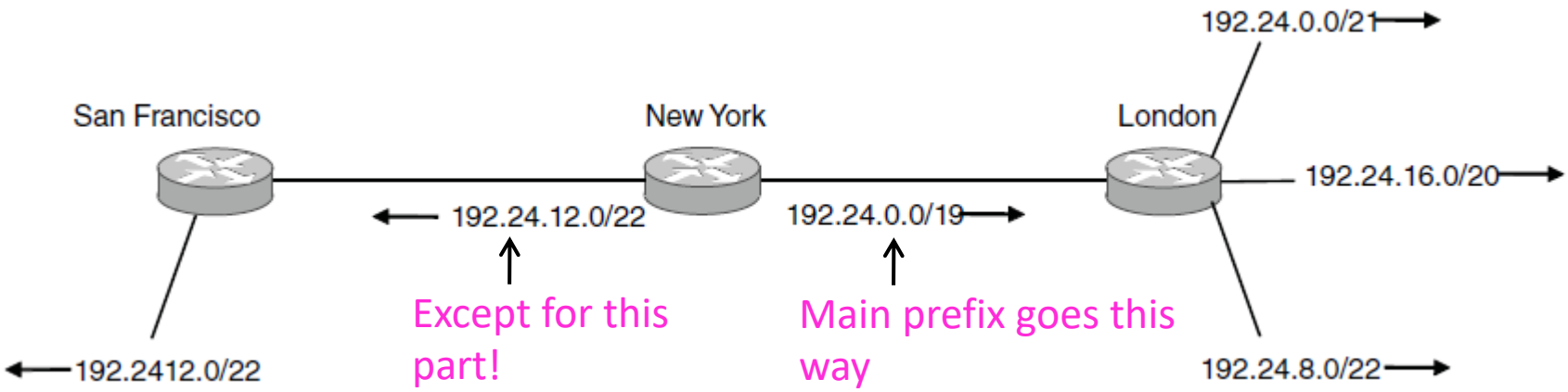
Aggregation

- Aggregation joins multiple IP prefixes into a single larger prefix to reduce routing table size.
  - Before: “New York” has to have 3 routing entries for 3 different networks although they have the same “next hop”
  - After: “New York” only need 1 aggregated routing entry.
- Aggregation is heavily used throughout the Internet and can reduce the size of router tables to around 200,000 prefixes.



Longest Matching Prefix

- The previously available block of addresses within this prefix has now been allocated to a network in San Francisco.
- Prefixes are allowed to overlap. The rule is that packets are sent in the direction of the most specific route, or the **longest matching prefix** that has the fewest IP addresses.
  - Complicates forwarding but adds flexibility



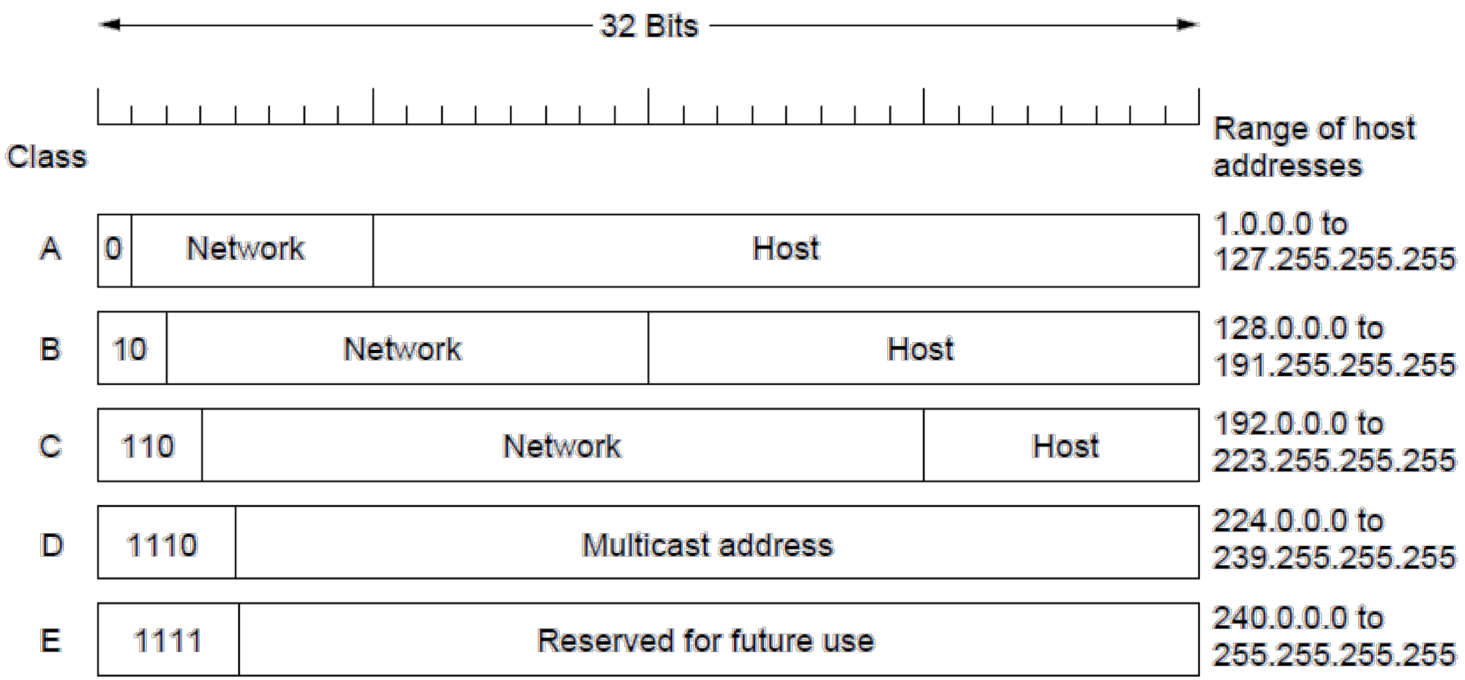
The first 19bits of 192.24.12.0 and 192.24.0.0 are the same!

CIDR – Summary

- When a packet comes in, the routing table is scanned to determine if the destination lies within the prefix.
- It is possible that multiple entries with different prefix lengths will match, in which case the entry with the longest prefix is used.
- Thus, if there is a match for a /20 mask and a /24 mask, the /24 entry is used to look up the outgoing line for the packet.

Classful Addressing

- Old addresses came in blocks of fixed size (A, B, C)
  - Carries size as part of address, but lacks flexibility
  - Called classful (vs. classless) addressing



Classful Addressing

- Class A: 126 available networks with 16 million hosts each
- Class B: 16,383 available networks with up to 65,534 hosts
- Class C: 2 million networks with up to 254 hosts each.

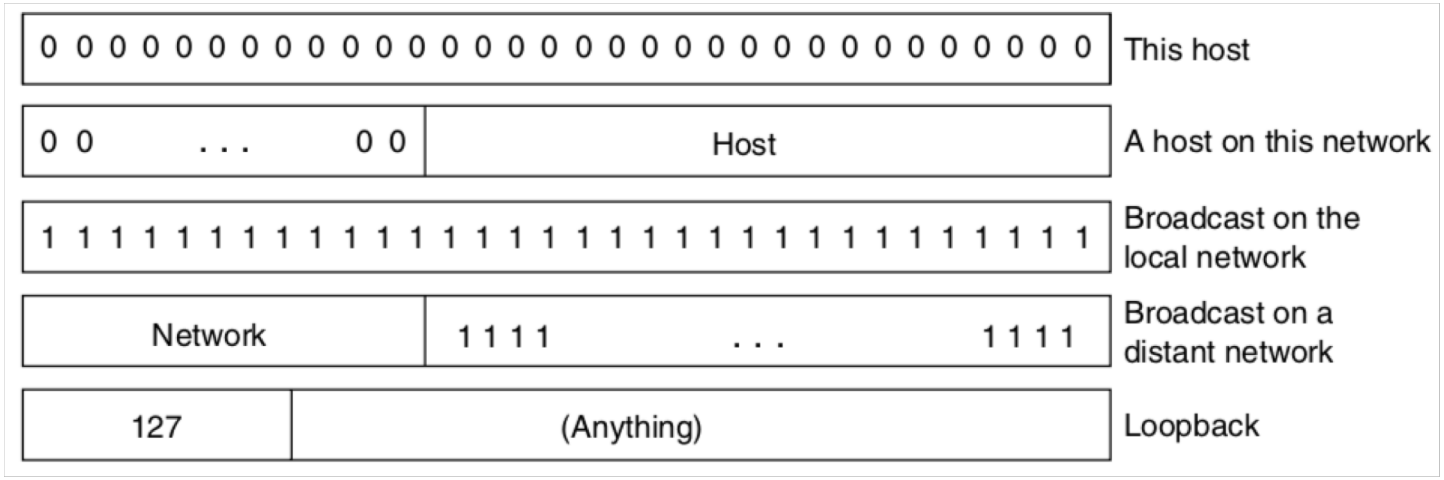
Class	Max. # Available Networks	First available network id	Last available network id	Max. # Available Hosts per network
A	2^7-**2**	1	126	2^{24}-2
B	2^{14}-1	128.1	191.255	2^{16}-2
C	2^{21}-1	192.0.1	223.255.255	2^8-2

Classful Addressing

- Unlike CIDR, the sizes of the address blocks are fixed.
- Studies show that more than half of all class B networks have actually fewer than 50 hosts.
- A class C network would have done the job, but no doubt every organization that asked for a class B address thought that one day it would outgrow the 8-bit host field.

Special IP addresses

- **0.0.0.0**: used by hosts when they are being booted. It means ‘‘this network’’ or ‘‘this host.’’
- **IP addresses with 0 as the network number** refer to the current network. These addresses allow machines to refer to their own network without knowing its number .
- **255.255.255.255**: is used to mean all hosts on the indicated network. It allows broadcasting on the local network, typically a LAN.
- **The addresses with a proper network number and all 1s in the host field** allow machines to send broadcast packets to distant LANs anywhere in the Internet. However, many network administrators disable this feature as it is mostly a security hazard.
- **All addresses of the form 127.*xx.yy.zz*** are reserved for loopback testing. Packets sent to that address are processed locally and treated as incoming packets without putting out onto the wire.



Internet Control Protocols

- IP works with the help of several control protocols:
  - ICMP is a companion to IP that returns error info
    - Required, and used in many ways, e.g., for traceroute
  - ARP finds Ethernet address of a local IP address
    - Glue that is needed to send any IP packets
    - Host queries an address and the owner replies
  - DHCP assigns a local IP address to a host
    - Gets host started by automatically configuring it
    - Host sends request to server, which grants a lease
    - Will be introduced in the next few weeks

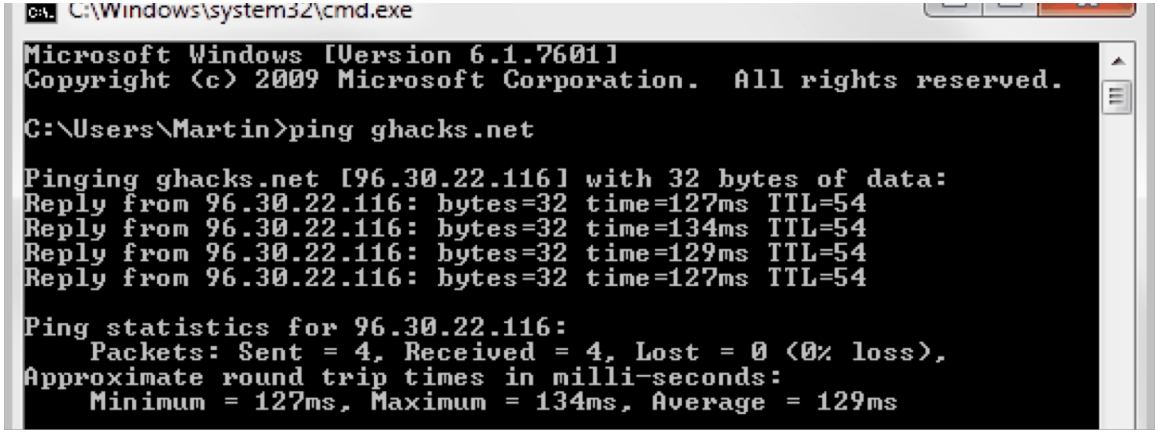
ICMP

- Main ICMP (Internet Control Message Protocol) types:

Message type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	Time to live field hit 0
Parameter problem	Invalid header field
Source quench	Choke packet
Redirect	Teach a router about geography
Echo and Echo reply	Check if a machine is alive
Timestamp request/reply	Same as Echo, but with timestamp
Router advertisement/solicitation	Find a nearby router

Echo and Echo reply

- They are sent by hosts to see if a given destination is reachable and currently alive.
- Upon receiving the ECHO message, the destination is expected to send back an ECHO REPLY message.
- These messages are used in the **ping** utility that checks if a host is up and on the Internet.



- Equivalent to “dragging and dropping a message” in your Packet Tracer.
- ”ping” does not belong to the ICMP message type: “Destination unreachable”.

Time exceeded

- The TIME EXCEEDED message is sent when a packet is dropped because its *TtL (Time to live)* counter has reached zero.
  - This event is a symptom that packets are looping, or that the counter values are being set too low.
- The *TtL (Time to live)* field is a counter used to limit packet lifetimes.
  - In practice, it just counts hops and must be decremented on each hop.
  - This feature prevents packets from wandering around forever.

Traceroute

- Traceroute finds the routers along the path from the host to a destination IP address.
- The method is simply to send a sequence of packets to the destination, first with a TtL of 1, then a TtL of 2, 3, and so on.
- The counters on these packets will reach zero at successive routers along the path.

```
SG350X#traceroute ip software.cisco.com ttl 20
Tracing the route to software.cisco.com (184.26.111.212) from , 20 hops
max, 18 byte packets
Type Esc to abort.
 1  192.168.100.1 (192.168.100.1)  <10 ms  <10 ms  <10 ms
 2  124.6.177.113 (124.6.177.113)  <20 ms  <10 ms  <20 ms
 3  124.6.149.117 (124.6.149.117)  <20 ms  <30 ms  <30 ms
 4  120.28.0.61 (120.28.0.61)  <20 ms  <20 ms  <30 ms
 5  120.28.10.101 (120.28.10.101)  <40 ms  <30 ms  <30 ms
 6  120.28.9.158 (120.28.9.158)  <40 ms  <40 ms  <40 ms
 7   *  *  *
 8   *  *  *
 9  63.218.2.189 (63.218.2.189)  <50 ms  <50 ms  <50 ms
10  63.223.17.162 (63.223.17.162)  <60 ms  <50 ms  <50 ms
11  63.223.17.162 (63.223.17.162)  <50 ms  <50 ms  <50 ms
12  213.254.227.77 (213.254.227.77)  <50 ms  <60 ms  <50 ms
13   *  *  *
14  184.26.111.212 (184.26.111.212)  <190 ms  <200 ms  <200 ms

Trace complete.
```

Traceroute

- These routers will each obediently send a TIME EXCEEDED message back to the host.
- From those messages, the host can determine the IP addresses of the routers along the path, as well as keep statistics and timings on parts of the path.
- It is not what the TIME EXCEEDED message was intended for, but it is perhaps the most useful network debugging tool of all time.

```
SG350X#traceroute ip software.cisco.com ttl 20
Tracing the route to software.cisco.com (184.26.111.212) from , 20 hops
max, 18 byte packets
Type Esc to abort.
 1  192.168.100.1 (192.168.100.1)  <10 ms  <10 ms  <10 ms
 2  124.6.177.113 (124.6.177.113)  <20 ms  <10 ms  <20 ms
 3  124.6.149.117 (124.6.149.117)  <20 ms  <30 ms  <30 ms
 4  120.28.0.61 (120.28.0.61)  <20 ms  <20 ms  <30 ms
 5  120.28.10.101 (120.28.10.101)  <40 ms  <30 ms  <30 ms
 6  120.28.9.158 (120.28.9.158)  <40 ms  <40 ms  <40 ms
 7   *  *  *
 8   *  *  *
 9  63.218.2.189 (63.218.2.189)  <50 ms  <50 ms  <50 ms
10  63.223.17.162 (63.223.17.162)  <60 ms  <50 ms  <50 ms
11  63.223.17.162 (63.223.17.162)  <50 ms  <50 ms  <50 ms
12  213.254.227.77 (213.254.227.77)  <50 ms  <60 ms  <50 ms
13   *  *  *
14  184.26.111.212 (184.26.111.212)  <190 ms  <200 ms  <200 ms

Trace complete.
```

ARP

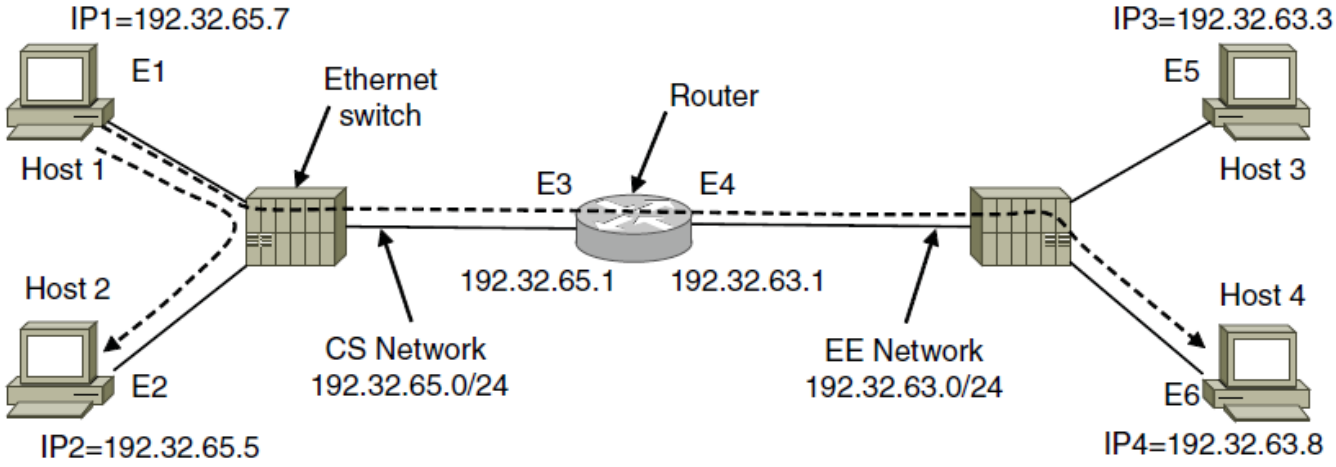
- Although every machine on the Internet has one or more IP addresses, these addresses are not sufficient for sending packets.
- The data link layer NICs send and receive frames based on 48-bit Ethernet addresses. They do not understand 32-bit IP addresses.
- How do IP addresses get mapped onto data link layer addresses, such as Ethernet?
  - Using ARP (Address Resolution Protocol)!

ARP - Steps

- When A intends to send IP messages to B
  - A has its own MAC address and IP address, it also knows B's IP address
- If B is in the same network as A:
  - Checks if B's MAC address exists in A's ARP cache.
    - If yes, return B's MAC address directly.
  - If not, A sends a broadcast frame to everyone in the same network saying "I'm IP-A, MAC-A, who is IP-B?"
  - B receives it then replies to A directly with its own MAC address.
  - At the same time, when others but not B receive this broadcast frame, they update A's entry in their ARP cache, then drop the frame.
- If B is NOT in the same network as A:
  - A should find the MAC address of its default gateway using the similar process described above.

ARP - Examples

- ARP lets nodes find target Ethernet addresses (in pink) from their IP addresses



Frame	Source IP	Source Eth.	Destination IP	Destination Eth.
Host 1 to 2, on CS net	IP1	E1	IP2	E2
Host 1 to 4, on CS net	IP1	E1	IP4	E3
Host 1 to 4, on EE net	IP1	E4	IP4	E6