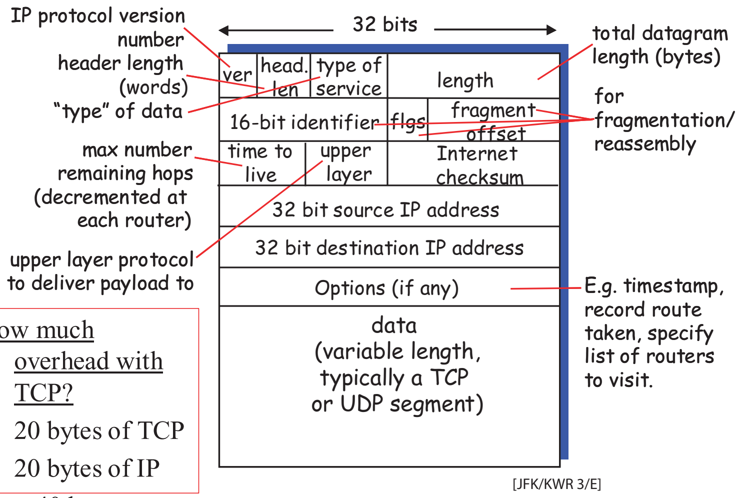
Slide 1

4. IPv4: The Internet Protocol Version 4 (Section 4.4)

Datagram Format

Variable length header (due to the options field):



how much overhead with TCP?

- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

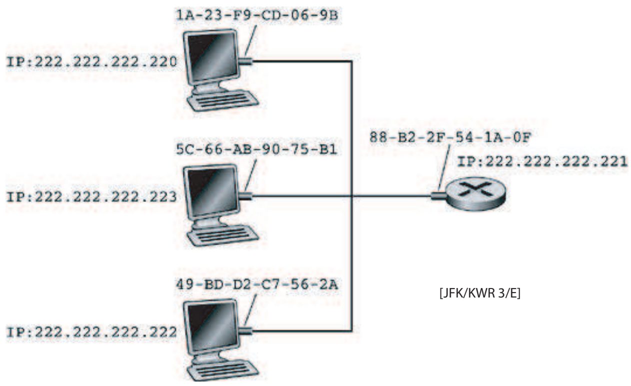
Slide 2

- *Version number.* 4 for IPv4
- *Header length.* 4 bits (counting 32-bit words) for the header only part. Min. length is 5w (20 bytes), Max. length is 15w (60 bytes).
- *Datagram length (bytes).* Header + payload.
- *Type of service (TOS).* 6 bits intended to identify the datagram required service (e.g., delay, throughput, reliability). Not standardized.
- *Identifier, flags, fragmentation offset.* For fragmentation and reassembly.
- *Time-to-live (TTL).* decremented by one each hop.
- *Header checksum.* Internet checksum on the header part only. Must be recomputed at each hop (because of the TTL field).
  Note. In contrast, TCP computes checksum over the entire segment.
- *Options.* Originally, five options were defined: *security*, *strict source routing*, *loose source routing*, *record route*, and *timestamp*.

Slide 3

IP Addresses

- Each network interface (e.g., an Ethernet card) has two addresses:
  1. a *link layer* address: a unique built-in (manufacturer assigned) number (6-byte numbers for Ethernet cards): e.g.,



     Note. A link layer address is also know as *medium access control* (MAC) address, *physical* addr, *LAN* address.

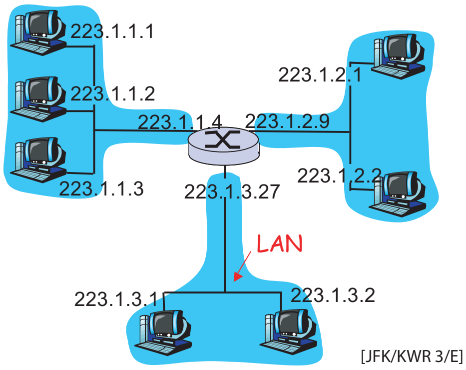
  2. a *network layer* IP address: typically provided by the ISP.

Slide 4

- Decimal-dotted notation for IP addresses: e.g., 223.1.1.1
- Q. Why are MAC addresses **not** used for network layer routing?

  Q. Are IP addresses more like SINs, or postal addresses?
- Each 32-bit IP address carries two parts:
  - a *network address* (high order bits), and
  - a *host (interface) address* (low order bits).
- IP jargon: an *IP network (or subnetwork)* is a group of hosts that share the same network address. e.g.,

Slide 5



network consisting of 3 subnets

- How can a router determine the network part of an IP address? Two schemes:
  - the classful addressing scheme, and
  - the *Classless InterDomain Routing (CIDR)* scheme.

Slide 6

Classful Addressing (old)

- Classful addressing simplifies:
  - allocation of IP addresses in *fixed* chunks to organizations/purposes
  - fast extraction of the "network part" of an IP address (without extra information)
- How? The scheme defines four classes:

32 Bits

Class				Range of host addresses
A	0	Network	Host	1.0.0.0 to 127.255.255.255
B	10	Network	Host	128.0.0.0 to 191.255.255.255
C	110	Network	Host	192.0.0.0 to 223.255.255.255
D	1110	Multicast address		224.0.0.0 to 239.255.255.255
E	1111	Reserved for future use		240.0.0.0 to 255.255.255.255

- Some reserved numbers:

00000000000000000000000000000000		This host
00 . . . 00	Host	A host on this network
11111111111111111111111111111111		Broadcast on the local network
Network	1111 . . . 1111	Broadcast on a distant network
127	(Anything)	Loopback

[Tan 4/E]

Slide 7

- Drawbacks:
  - A class C network can accommodate only up to $2^8 - 2$= 254 hosts (too small for many organizations).
  - A class B network can accommodate about $2^{16} = 65,634$ hosts (too large for many organizations).
  - This led to rapid depletion of the class B address space, and poor utilization of the assigned address space

Classless InterDomain Routing (CIDR) Addresses (nowadays)

Slide 8

- Allow IP addresses to be allocated in *variable-sized* chunks.
- Address format: $a.b.c.d/x$, where x is the # of bits in the network portion of the address, e.g.

← subnet part → ← host part →

11001000 00010111 00010000 00000000

200.23.16.0/23 [JFK/KWR 3/E]

- x can also be defined by the pattern: 255.255.254.0 (the high-order 23 bits are set to 1s), called a *network (or subnet) mask* .
- That is, the network part is the AND of the IP# and the network mask.
- Note. To build the routing tables, routers must exchange both the IP address and the network mask of each IP network (e.g., as part of OSPF link state advertisements).

Slide 9

■ CIDRized Addresses: Aggregation and the Longest Prefix Matching Rule

- Consider the following scenario: millions of addresses are available starting at 194.24.0.0 .
- Now, three universities are allocated IP addresses as follows:

University	First address	Last address	How many	Written as
Cambridge	194.24.0.0	194.24.7.	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	194.24.12/22
Oxford	194.24.16.0	194.24.31.255	4096	194.24.16.0/20

- Note that all of the above addresses agree on the high-order 19 bits.
- So, for a router in Edmonton, one aggregate entry that directs to UK any incoming datagram whose destination network # matches 194.24.0.0/19 should work fine.
- Now, suppose that some of the 194.24.12.0/22 addresses are allocated to a company in California. How can routing be done?

Slide 10

- Answer: routing should be done using the entry that gives the *longest prefix match* between addresses.

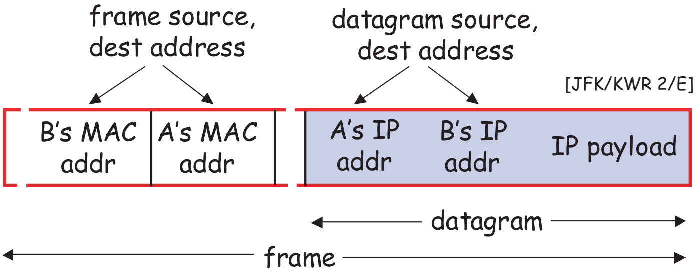
5. More Internet Control Protocols

ARP: Address Resolution Protocol [RFC 826] (Chapter 5)

■ Routing to the same LAN.
For host A to send a datagram to same-LAN host B, A should know both B's IP address and MAC address.

■ Why? Because the MAC address is required by the link-level protocol:

Slide 11



- Q: How to determine B's MAC address knowing B's IP address (in a plug-and-play manner)?
- The ARP Protocol:
  - Runs on each host and router.
  - Maintains an *ARP table* (initially empty) for each network interface:

IP Address	LAN Address	TTL
222.222.222.221	88-B2-2F-54-1A-0F	13:45:00
222.222.222.223	5C-66-AB-90-75-B1	13:52:00

  $*$ Typically, TTL (Time To Live) is initialized to 20 minutes.

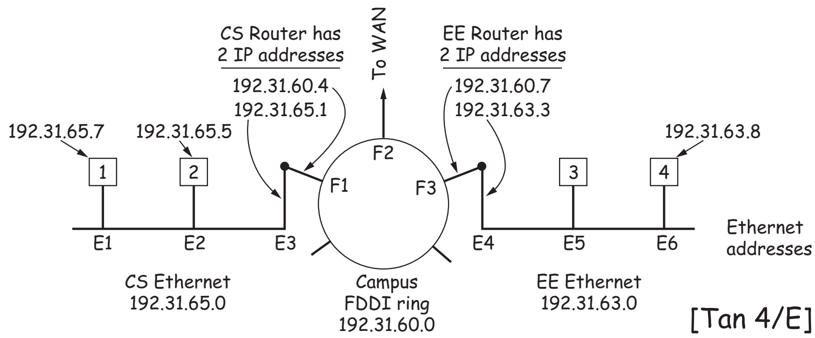
Slide 12

  $*$ Q. Why does ARP need the TTL field?

  - If B's MAC address is not cached then A *broadcasts* an ARP query pkt $<$IP#= 255.255.255.255, MAC= FF-FF-FF-FF-FF-FF$>$, containing B's IP address.
  - B replies to A with its (B's) MAC address (this transaction is *unicast*).
  - A caches the IP-to-MAC address pair in its ARP table until the TTL field times out.
- Exercise (routing to another LAN). Suppose that initially host E1 knows only:
  - its $<$IP addr, MAC addr$>$,
  - the IP addr of the CS router, and
  - host E6's IP addr.

  How many ARP tables are kept in the CS router? How many ARP msgs are used to send a pkt from E1 to E6?

Slide 13



[Tan 4/E]