

EEC-684/584 Computer Networks

Lecture 14

Wenbing Zhao

wenbing@ieee.org

(Lecture notes are based on materials supplied by
Dr. Louise Moser at UCSB and Prentice-Hall)

Outline

- Review of last lecture
 - Internetworking
 - Network layer in Internet (part 1)
- Network layer in Internet (part 2)
 - NAT
 - Internet control protocols: ICMP, ARP, RARP, BOOTP, DHCP
 - OSPF, BGP, IP Multicast, Mobile IP
 - IPv6

25 October 2005

EEC484/584

Wenbing Zhao

Review

- Tunneling
- Fragmentation
- IPv4 header
- Subnetting
- CIDR

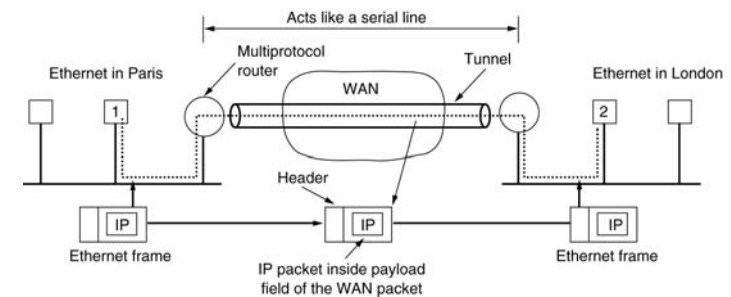
25 October 2005

EEC484/584

Wenbing Zhao

Tunneling

- Tunneling - common solution for interconnecting multiple heterogeneous networks **when the source and destination hosts are on the same type of network**



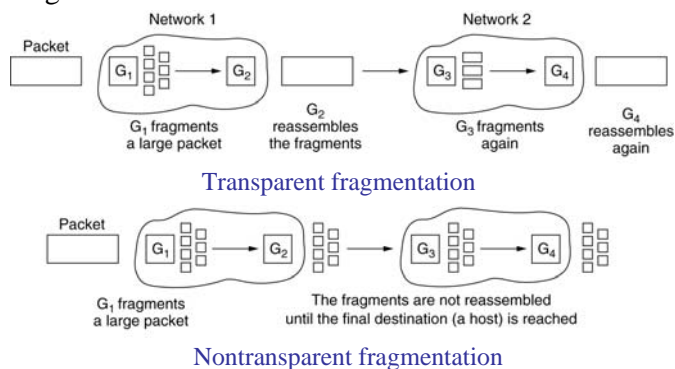
25 October 2005

EEC484/584

Wenbing Zhao

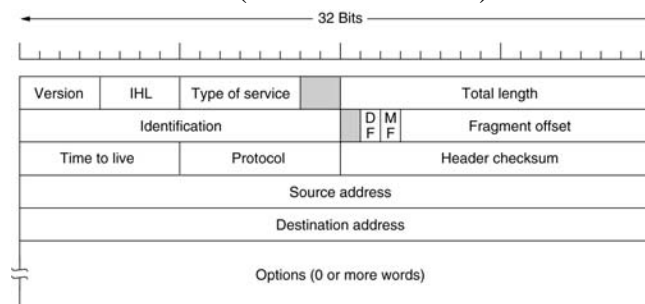
Fragmentation

- **Fragmentation** - if a packet has to be transmitted over a network that imposes a smaller max size, it has to be fragmented



The IP Protocol

The IPv4 (Internet Protocol) header



25 October 2005

EEC484/584

Wenbing Zhao

The IPv4 Header

- IHL - length of header in 32-bit words (5-15)
- Total length - header and data $\leq 65,535$ ($2^{16}-1$) bytes
- Identification - allows destination to determine which datagram a fragment belongs to
- DF - Don't Fragment; MF - More Fragments
- Fragment offset - where in datagram this fragment belongs
- Time to live - counter to limit packet lifetimes
- Protocol - which transport layer protocols being used
- Header checksum - verifies header

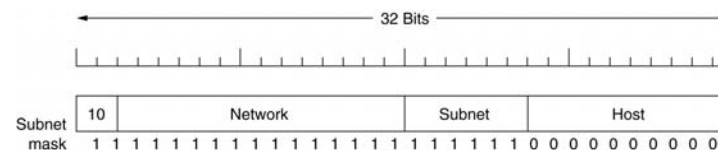
25 October 2005

EEC484/584

Wenbing Zhao

Subnets

- Allow network to be split into several parts for internal use, but to act as **single network** to outside world
 - Take some bits away from host numbers
 - **Subnet mask** - needed by the main router. Indicates split between network + subnet number and host
 - Write the address and the mask as a binary number
 - If mask bit is 1, then corresponding bit of address matters
- Example: a class B network can be subnetted into 64 subnets
 - Originally 16 bits for host info. Now, 6 bits used for subnet and 10 bits for host numbers
 - Subnet mask can be written as 255.255.252.0 or /22



CIDR – Classless InterDomain Routing

- Shortage of IP addresses caused by the classful addressing
- Solution - for the remaining IP addresses, CIDR is used
 - Allocate remaining IP addresses in variable-sized blocks, without regard to the classes
 - **The starting address must fall on the boundary of the block size**
 - E.g., if a site needs, say, 2000 addresses, it is given a block of 2048 addresses on a 2048-byte boundary

NAT – Network Address Translation

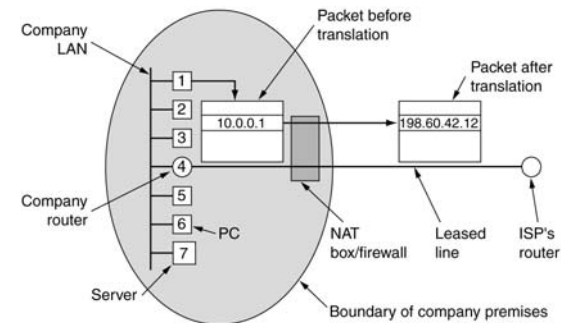
- Problem - IP addresses are scarce
 - For an ISP that provides dial-up service, it can support many times more subscribers than its capacity because only some users are active at a time
 - For ADSL/cable/business users, they need an IP address for each of their machines 24/7

NAT – Network Address Translation

- Temporary solution - network address translation
 - Each company or family is assigned a single IP address (or a small number of them)
 - Within the company, every computer gets a unique private IP address, which is used for routing intramural traffic
 - When a packet exits the company and goes to the ISP, an address translation takes place
 - Three ranges of IP addresses have been declared as private:
 - 10.0.0.0 - 10.255.255.255 (16,777,216 hosts)
 - 172.16.0.0 - 172.31.255.255/12 (1,048,576 hosts)
 - 192.168.0.0 - 192.168.255.255/16 (65,536 hosts)

NAT – Network Address Translation

Placement and operation of a NAT box.



NAT – What about the Incoming Traffic?

- **Solution is based on the assumption all traffic is TCP/UDP**
- TCP/UDP has two port fields, one for source port, the other for destination port, each 16 bits wide
- The source port is used as an index to an internal table maintained by the NAT box
- The internal sender's private IP and original port info are stored in the table
- When the reply comes back, it will carry the index as the destination port, the NAT box then translates the address back
- For both outgoing and incoming address translations, the TCP/UDP and IP header checksums are recomputed

Drawback of NAT

- NAT violates the architectural model of IP, which states that every IP address uniquely identifies a single machine worldwide
- NAT box must maintain mapping info for each connection passing through it. This changes the Internet from a connectionless network to a kind of connection-oriented network
- NAT violates the most fundamental rule of protocol layering: layer k may not make any assumptions about what layer k+1 has put into the payload field
- NAT only support UDP/TCP traffic
- NAT has problem supporting apps that include local IPs in payload, such as FTP and H.323
- Each NAT box can support at most 65,536 (2^{16}) hosts

Internet Control Protocols

- ICMP
- ARP
- RARP
- BOOTP
- DHCP

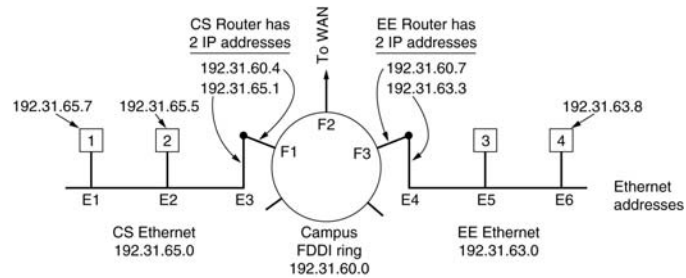
ICMP – Internet Control Message Protocol

- When something unexpected occurs in Internet, the event is reported by routers using ICMP
- It is also used to test Internet
- Principal ICMP message 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 request	Ask a machine if it is alive
Echo reply	Yes, I am alive
Timestamp request	Same as Echo request, but with timestamp
Timestamp reply	Same as Echo reply, but with timestamp

ARP–Address Resolution Protocol

- How do IP addresses get mapped onto data link layer addresses, such as Ethernet?



ARP–Address Resolution Protocol

- Sender A wants to send something to B
- A can find out B's IP address through DNS
- A broadcasts "who owns IP xxx.xxx.xxx.xxx?"
- The broadcast arrives at every machine on the same LAN
- Each one checks its IP address
- Only the machine that carries the IP address (B) responds with its Ethernet address
- A then builds a Ethernet frame with the right Ethernet address (B's Ethernet address)

ARP–Address Resolution Protocol

- Strength of ARP
 - Simple, no manual configuration except IP assignment to each host

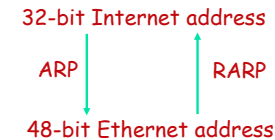
ARP Optimization

- Once a machine has run ARP, it caches the result in case it needs to contact the same machine shortly
- When A wants to communicate with B, A includes its IP-to-Ethernet mapping in the ARP packet so that B knows the mapping right away
 - In fact, all machines on the Ethernet can enter this mapping into their ARP caches
- Have every machine broadcast its mapping when it boots, so that everyone else knows the mapping
- To accommodate changes, entries in the ARP cache should time out after a few minutes

ARP—Address Resolution Protocol

- Proxy ARP - A router is configured to answer ARP requests on one of its networks for a host on another of its networks
 - The router is acting as a proxy agent for the destination host, relaying packets to it from other hosts

RARP - Reverse Address Resolution Protocol



- RARP - This protocol allows a newly-booted diskless-workstation (e.g., X terminal) to broadcast its Ethernet address and ask for its IP address
 - **RARP server** responds to a RARP request with the assigned IP address

Limitations of RARP

- RARP uses a link-layer broadcast, RARP requests are not forwarded by routers, therefore, an RARP server must be present on every network.
- The only thing returned by the RARP server is the IP address

BOOTP - Bootstrap Protocol

- BOOTP - uses UDP
 - When a client is bootstrapping using BOOTP, the request is normally a link-layer broadcast and the destination IP address in the IP header is normally 255.255.255.255
 - The source IP address is set to 0.0.0.0 if client does not know its own IP address yet
 - Recall that 0.0.0.0 is a valid source IP address when a system is bootstrapping itself
 - Port number: 67 for server, 68 for client
- BOOTP drawbacks
 - Requires manual configuration of tables mapping IP address to Ethernet address at the BOOTP server

DHCP –

26

Dynamic Host Configuration Protocol

- Allows both manual IP address assignment and automatic assignment. DHCP has largely replaced RARP and BOOTP
- A **DHCP relay agent** is needed on each LAN. The only piece of information the relay agent needs is the IP address of the DHCP server.
- To find its IP address, a newly-booted machine broadcasts a DHCP DISCOVER packet. The DHCP relay agent on its LAN receives all DHCP broadcasts
- When it finds a DHCP DISCOVER packet, it sends the packet as a unicast packet to the DHCP server, possibly on a distant network
- IP address assignment is lease-based (to cope with client failure)

25 October 2005

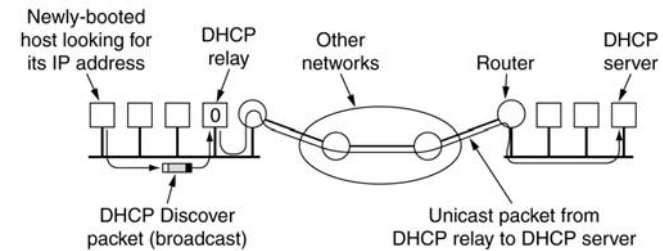
EEC484/584

Wenbing Zhao

Dynamic Host Configuration Protocol

27

Operation of DHCP



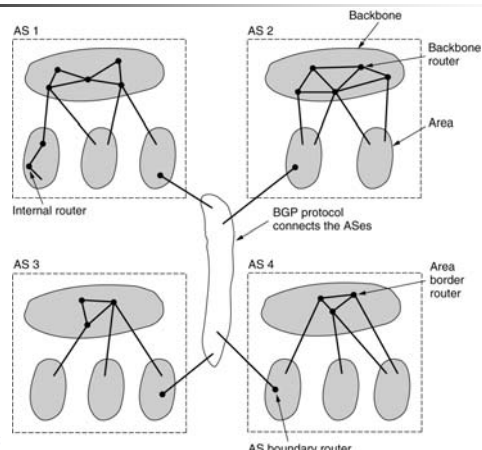
25 October 2005

EEC484/584

Wenbing Zhao

Internet Routing Protocols

28



25 October 2005

AS boundary router

Wenbing Zhao

Interior Gateway Routing Protocol

29

- Uses Open Shortest Path First (OSPF)
 - Open, dynamic, and support multiple distance metrics
 - Routing based on type of service
 - Load balancing
 - Hierarchical
 - Security
 - Tunneling
- Supports 3 types of connections
 - Point-to-point between 2 routers
 - Multiaccess (multiple routers that communicate with each other) - with broadcasting and without broadcasting

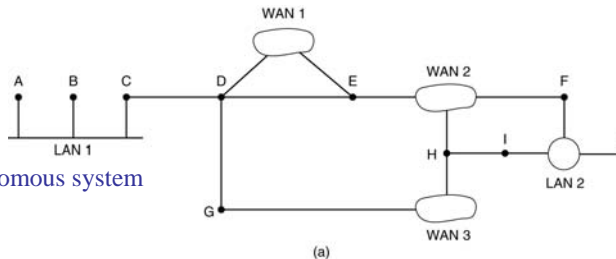
25 October 2005

EEC484/584

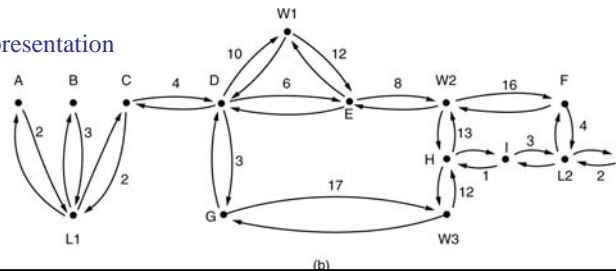
Wenbing Zhao

OSPF

An autonomous system



A graph representation



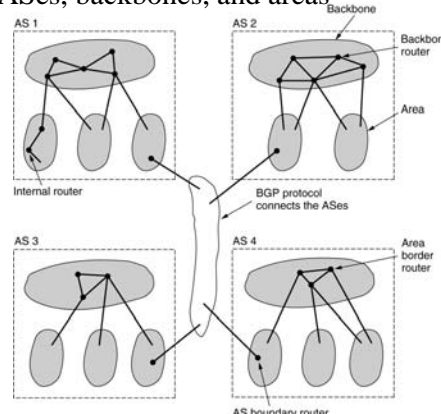
OSPF

- OSPF operates by abstracting the collection of actual networks, routers, and lines into a directed graph
 - Each arc is assigned a cost (distance, delay, etc.)
 - It then computes the shortest path from every router to every other router, based on the weights on the arcs.
 - A serial connection between two routers is represented by a pair of arcs, one in each direction. Their weights may be different.
 - A multiaccess network is represented by a node for the network itself plus a node for each router

OSPF

- The relation between ASes, backbones, and areas

- Four types of routers
 - Internal
 - Backbone
 - Area border
 - AS boundary



OSPF

- Within area, each router has same link state database
- Each router periodically floods LINK STATE UPDATE packets to each of its **adjacent** routers
- Each router constructs the graph for its area and computes shortest path
- Backbone router also does:
 - Accepts info from area border routers
 - Computes best route from each area border router to every other area border router
 - This info propagated back to area border routers which advertise it within their areas
- Using this info, router about to send interarea packet selects best exit router to backbone

OSPF

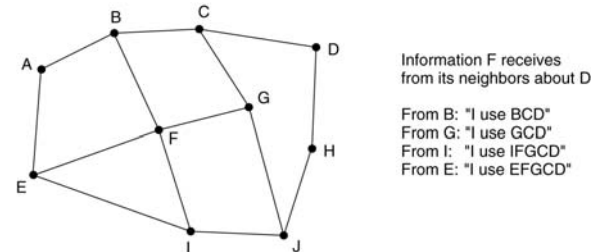
- The five types of OSPF messages

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

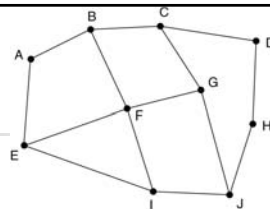
Exterior Gateway Routing Protocol

- Border Gateway Protocol (BGP)

- Used between autonomous systems
- Main concerns: politics, security, economic
- Uses distance vector routing except keeps track of exact path instead of cost to destination and periodically tells its neighbors that path



BGP



- Does not suffer from count-to-infinity problem
- Example: suppose *G* crashes or the line *FG* goes down. *F* then receives routes from its three remaining neighbors. These routes are *BCD*, *IFGCD*, and *EFGCD*
 - It can immediately see that the two latter routes are pointless, since they pass through *F* itself, so it chooses *FBCD* as its new route

Internet Multicasting

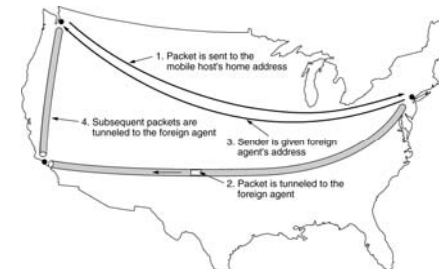
- Uses class D addresses
 - 28 bits to identify groups of hosts
- 2 kinds of group addresses
 - Permanent, e.g., 224.0.0.1 all systems on LAN
 - Temporary
- Implemented by special multicast routers
 - Once a minute, each multicast router sends hardware multicast to hosts on its LAN asking them to report back on groups to which their processes belong
 - Each host sends back response

Internet Multicasting

- Queries and responses use
 - Internet Group Management Protocol (IGMP) query and response packets
- Uses spanning tree per group
 - Constructed by exchanging info with neighbors using distance vector protocol

Mobile IP

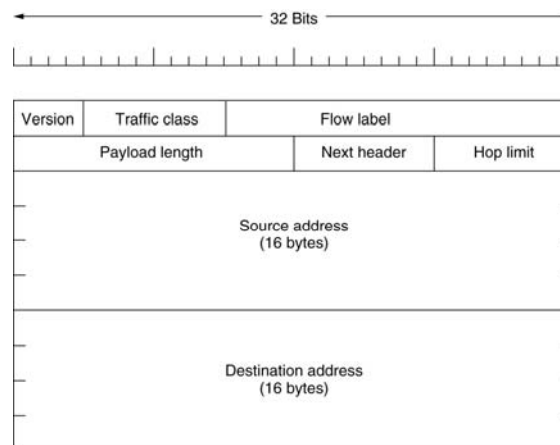
- Basic problem: routers use class and network number to do their routing, but mobiles move around and are not tied to one fixed network
- Solution: every mobile has home agent. Every site that allows a mobile visitor creates a foreign agent
 - When mobile reaches foreign site, it contacts foreign host, which contacts mobile's home agent and provides its own IP address as the mobile's "in-care-of" address
 - Tunneling and mappings between Ethernet addresses and IP addresses are used, as previously described



Internet Protocol Version 6 (IPv6)

- IPv4 current version
- IPv5 experimental real-time stream protocol
- IPv6
 - Longer addresses than IPv4 - 16 bytes instead of 4 bytes
 - Simplified header - 7 fields instead of 13 fields
 - Speed packet processing and improves throughput
 - Better support for options
 - Security
 - Different types of service (data, video, audio, etc.)

The Main IPv6 Header



The Main IPv6 Header

- Version 6
- Priority
 - 0-7 slow down in event of congestion
 - 8-15 real-time traffic
- Flow label - allows source and destination to set up pseudo-connection with particular properties and requirements
- Payload length (as opposed to total length in IPv4)
- Next header - additional optional extension header
- Hop limit (time to live in IPv4)
- Source/destination address - 128 bits (32 bits in IPv4)

IPv6 Extension Headers

43

Extension header	Description
Hop-by-hop options	Miscellaneous information for routers
Destination options	Additional information for the destination
Routing	Loose list of routers to visit
Fragmentation	Management of datagram fragments
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents

25 October 2005

EEC484/584

Wenbing Zhao

IPv6 Extension Headers

45

- Two types of extension header formats
 - Fixed format
 - Variable number of variable-length fields
 - Each item is encoded as a (type, length, value) tuple
 - Type: 1-byte field telling which option this is
 - Length: 1-byte field telling how long the value is (0-255 bytes)
 - Value: any info required

25 October 2005

EEC484/584

Wenbing Zhao

IPv6 Extension Headers

46

- **Hop-by-hop options:** used for info that all routers along the path must examine
 - One option has been defined to support datagrams exceeding 64K
 - Next header: 1-byte field telling type of header
 - Length field: 1-byte field telling how long the hop-by-hop header is in bytes, excluding the first 8 bytes, which are mandatory
 - 1-byte field indicating that this option defines the datagram size
 - 1-byte field telling the size is a 4-byte number
 - 4-byte field: size of datagram

Next header	0	194	4
Jumbo payload length			

25 October 2005

EEC484/584

Wenbing Zhao

IPv6 Extension Headers

- **Routing option header:** lists one or more routers that must be visited on the way to destination
 - Routing type field: 1-byte giving format of the rest of the header
 - Segments left field: 1-byte keeping track of how many of the addresses in the list have not yet been visited

