

EEC-484/584 Computer Networks

Lecture 13

Wenbing Zhao

wenbing@ieee.org

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

Outline

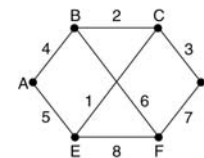
- Review of lecture 12
 - Routing
 - Congestion control
 - Quality of service
- Today's topics
 - Internetworking
 - The network layer in the Internet
 - IP protocol
 - IP address (classful addressing, subnet, CIDR)

Link State Routing

- Discover its neighbors, learn their network address
- Measure the delay or cost to each of its neighbors
- Construct a packet telling all it has just learned
- Send this packet to all other routers
- Compute the shortest path to every other router

Building Link State Packets

- Packets contain identity of sender, sequence number, age, list of **neighbors** and delay to that neighbor
- When are link state packets constructed?
 - Periodically at regular intervals
 - When link or nodes goes down or comes back



A subnet

Link		State		Packets	
A	B	C	D	E	F
Seq.	Seq.	Seq.	Seq.	Seq.	Seq.
Age	Age	Age	Age	Age	Age
B 4	A 4	B 2	C 3	A 5	B 6
E 5	C 2	D 3	F 7	C 1	D 7
	F 6	E 1		F 8	E 8

The link state packets for this subnet

Possible Strategies for Broadcast Routing

- Source sends separate copy of packet to all destinations
- Flooding
- Multi-destination routing
- **Spanning tree** - the sink tree (no loops) rooted at source that includes all routers. A router copies an incoming broadcast packet onto all the spanning tree lines except the one it arrived on
- **Reverse path forwarding**: if router gets packet on optimal (reverse) path to root, then it forwards the packet

Warning Bit

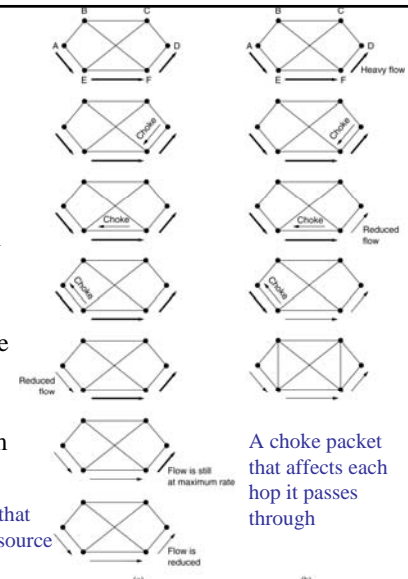
- Signal the warning state by setting a special bit in the packet's header
 - When packet arrived at its destination, transport entity copied the bit into the next ack sent back to the source
 - The source then cut back on traffic
 - Every router along the path could set the warning bit, traffic increases only when no router is in trouble

Choke Packets

- Tell the source directly to reduce traffic
 - Each router monitors utilization of each of its output lines
 - When utilization becomes greater than threshold, output line enters warning state
 - For each newly arriving packet, check if output line in warning state
 - If so, router sends choke packet back to source giving it the destination found in the packet
 - Packet is tagged so does not generate any more choke packets and forwarded as usual
 - When source receives choke packet, it must reduce traffic to specific destination

Hop-by-Hop Choke Packets

- The router that receives a choke packet must reduce the flow to its downstream router
 - This is achieved by allocating more buffer to the incoming flow
 - The router also passes the choke packet to its upstream router



A choke packet that affects only the source

A choke packet that affects each hop it passes through

Load Shedding

■ RED - Random Early Detection

- Having routers drop packets before situation becomes hopeless
- Routers maintain a running average of their queue lengths
- When average queue length on some line exceeds a threshold, the line is said to be congested and action is taken
- Just discard selected packet and not report it.
- Sender respond to lost packets by slowing down transmission rate
- Appropriate for wired networks

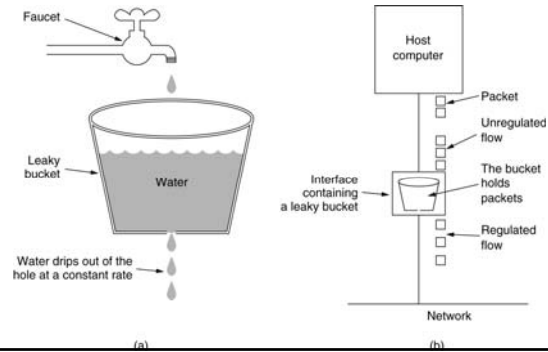
25 October 2005

EEC484/584

Wenbing Zhao

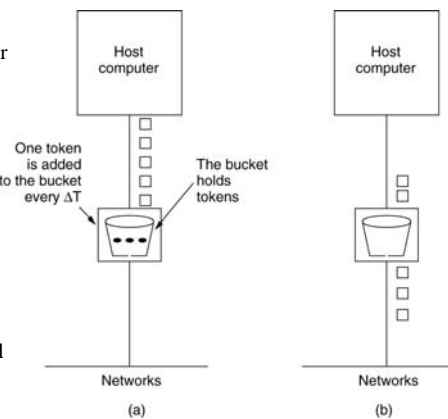
The Leaky Bucket Algorithm

- No matter what rate water enters the bucket, water flows out of the bucket at constant rate ρ (provided bucket is not empty)
 - When bucket is full, water slops over sides and is lost
 - Analogous to finite queue in network switch

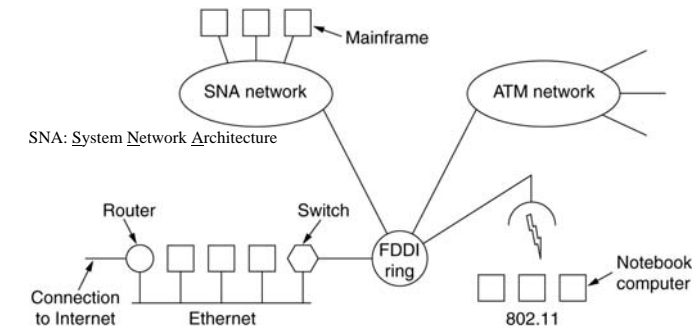


The Token Bucket Algorithm

- The bucket hosts tokens in stead of data packets
- A token represents a packet or k bytes
- Tokens are added into the bucket with a constant rate
- Bucket has certain capacity. If bucket is full, new tokens are thrown away
- A data packet can be transmitted only if enough tokens present in bucket
- **Token bucket algorithm allows some burstiness because a full bucket of tokens saved can be used all at once**



Internetworking



25 October 2005

EEC484/584

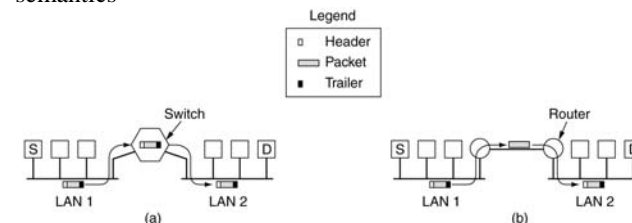
Wenbing Zhao

How Networks Differ

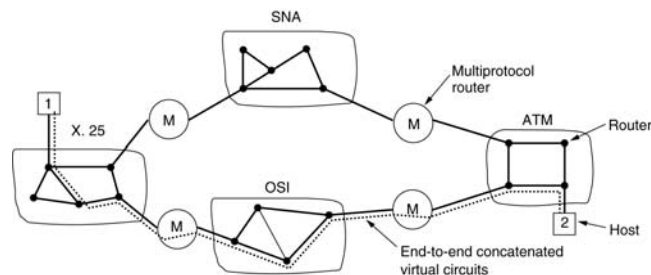
Item	Some Possibilities
Service offered	Connection oriented versus connectionless
Protocols	IP, IPX, SNA, ATM, MPLS, AppleTalk, etc.
Addressing	Flat (802) versus hierarchical (IP)
Multicasting	Present or absent (also broadcasting)
Packet size	Every network has its own maximum
Quality of service	Present or absent; many different kinds
Error handling	Reliable, ordered, and unordered delivery
Flow control	Sliding window, rate control, other, or none
Congestion control	Leaky bucket, token bucket, RED, choke packets, etc.
Security	Privacy rules, encryption, etc.
Parameters	Different timeouts, flow specifications, etc.
Accounting	By connect time, by packet, by byte, or not at all

How Networks Can Be Connected

- Physical layer – repeaters, hubs
- Link layer – bridges, switches
- Network layer – routers
 - Multiprotocol router: a router that can handle multiple protocols
- Transport layer – transport gateways
 - E.g., allow packets to flow between a TCP network and an SNA network
- Application layer – application gateways to translate message semantics



Concatenated Virtual Circuits

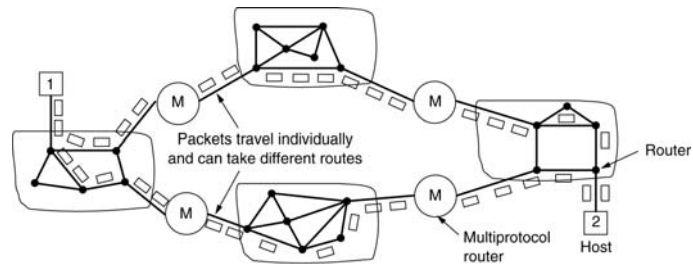


Concatenated Virtual Circuits

- Connection setup
 - Subnet sees that destination is remote and builds a virtual circuit to router nearest the destination network
 - Then it constructs a VC from that router to an external gateway (multiprotocol router)
 - This gateway records the existence of the VC in its tables and proceeds to build another VC to a router in the next subnet
 - This process continues until destination host has been reached
- After connection setup
 - Each gateway relays incoming packets, converting between packet formats and VC numbers as needed
 - All data packets traverse same sequence of gateways. Packets are not reordered

Connectionless Internetworking

17



25 October 2005

EEC484/584

Wenbing Zhao

Challenges in Connectionless Internetworking

18

- Different network layer protocols. How to transmit a packet over a foreign protocol subnet
 - Format translation. But what if the formats of two protocols are not compatible? (format translation rarely used)
- Addressing
 - IP - each interface card has an address
 - SNA - each hardware device has an address
 - At least a mapping database is needed

25 October 2005

EEC484/584

Wenbing Zhao

Tunneling

19

- Tunneling - common solution for interconnecting multiple heterogeneous networks **when the source and destination hosts are on the same type of network**
 - Multiprotocol router extracts the IP packet, inserts it in the payload field of the WAN network layer packet, and addresses the latter to the WAN address of the destination multiprotocol router
 - Destination multiprotocol router removes the IP packet and sends it the destination host

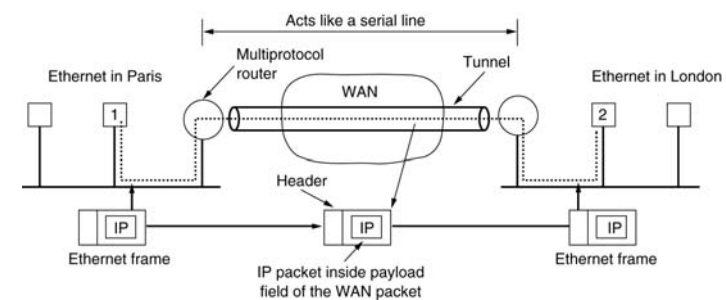
25 October 2005

EEC484/584

Wenbing Zhao

Tunneling

20



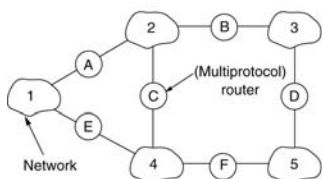
25 October 2005

EEC484/584

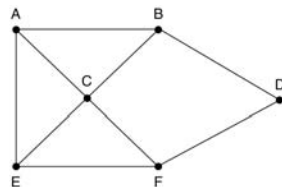
Wenbing Zhao

Internetwork Routing

- Two levels of routing algorithms
 - Interior gateway protocol – used within each network
 - Exterior gateway protocol – used between networks



(a) An internetwork



(b) A graph of the internetwork

Internetwork Routing

- A typical internet packet starts out on its LAN addressed to local multiprotocol router
- After it gets there, the network layer code decides which multiprotocol router to forward the packet to, using its own routing tables
- If that router can be reached using the packet's native network protocol, the packet is forwarded there directly
- Otherwise, it is **tunneled** there, encapsulated in the protocol required by the intervening network
- This process is repeated until the packet reaches the destination network

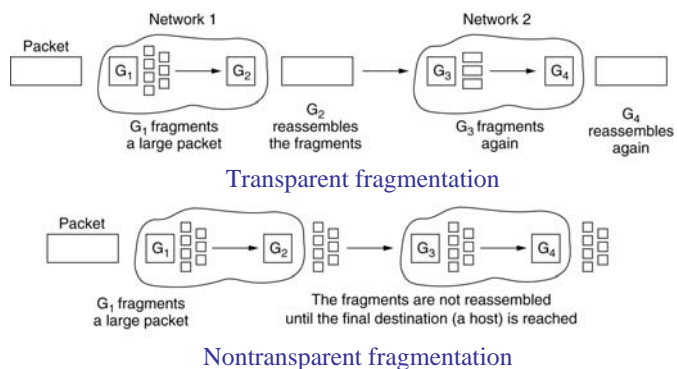
Fragmentation

- Each network imposes some maximum size on its packets. What if a packet has to be transmitted over multiple networks?
- **Fragmentation** - if a packet has to be transmitted over a network that imposes a smaller max size, it has to be fragmented

Two Fragmentation Strategies

- **Transparent** - fragmentation caused by a "small-packet" network transparent to any subsequent networks through which the packet must pass on its way
 - How does the exit gateway know if it has received all the fragments - use a count field, or an "end of packet" bit
 - All fragments must travel through same exit gateway
 - Overhead - fragmentation/reassembly might happen many times
- **Nontransparent** - each fragment is treated as though it were an original packet. Recombination occurs only at the destination host
 - **IP works this way**

Fragmentation



The Network Layer in the Internet

- The IP Protocol
- IP Addresses
- Internet Control Protocols
- OSPF – The Interior Gateway Routing Protocol
- BGP – The Exterior Gateway Routing Protocol
- Internet Multicasting
- Mobile IP
- IPv6

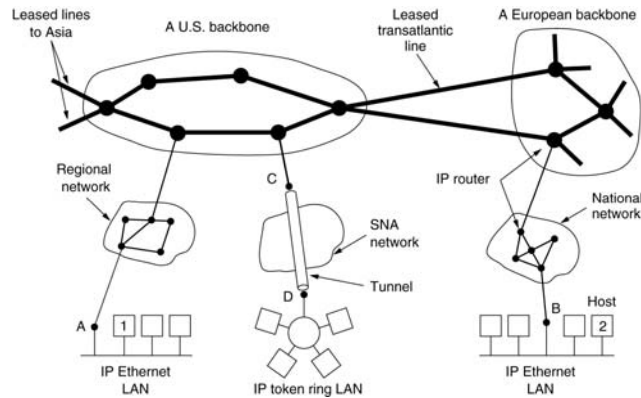
Design Principles for Internet

- Make sure it works
 - Do not finalize the design or standard until multiple prototypes have successfully communicated with each other
- Keep it simple
 - When in doubt, use the simplest solution
- Make clear choices
 - If there are several ways of doing the same thing, choose one
- Exploit modularity
 - This principle leads directly to the idea of having protocol stacks, each of whose layers is independent of all the other ones

Design Principles for Internet

- Expect heterogeneity
 - Different types of hardware, transmission facilities, and applications will occur on any large network
- Avoid static options and parameters
 - If parameters are unavoidable, it is best to have the sender and receiver negotiate a value than defining fixed choices
- Look for a good design; it need not be perfect
- Be strict when sending and tolerant when receiving
- Think about scalability
- Consider performance and cost

Collection of Subnetworks



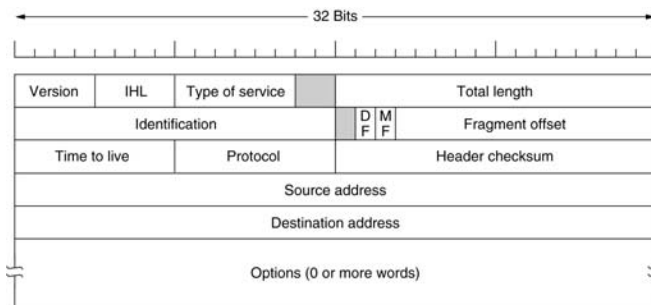
The Internet is an interconnected collection of many networks, or **Autonomous Systems (ASes)**

Communication in the Internet

- The glue that holds the whole Internet together is the network layer protocol, Internet Protocol (IP)
- Communication in the Internet
 - Transport layer takes data stream and breaks them up into datagrams, typically no more than 1500 bytes
 - Each datagram is transmitted through the Internet, possibly being fragmented into smaller units as it goes
 - When all pieces finally get to the destination machine, they are reassembled by the network layer into the original datagram
 - This datagram is then handed to the transport layer, which inserts it into the receiving process' input stream

The IP Protocol

The IPv4 (Internet Protocol) header



The IPv4 Header

- Version - 4
- IHL - length of header in 32-bit words
 - Min 5, max 15 - i.e., 60 bytes
- Type of service - to distinguish different classes of service
 - Original: 3 bits Precedence fields, and 3 flags D, T, R
 - Delay, Throughput, Reliability
 - Current use: to accommodate differentiated services (which class this packet belongs to)
- Total length - header and data $\leq 65,535 (2^{16}-1)$ bytes
- Identification - allows destination to determine which datagram a fragment belongs to

The IPv4 Header

- DF - tells routers "Don't Fragment" because destination can't reassemble
- MF - More Fragments.
 - All fragments except last have this set. Used as check against total length
- Fragment offset - where in datagram this fragment belongs.
 - All fragments (payload in the IP packet) except last must be multiples of 8 bytes
 - The number of 8 byte blocks is called **Number of Fragment Blocks (NFB)**
 - The unit of the offset is NFB
- Time to live - counter to limit packet lifetimes
 - Max lifetime 255sec
 - Packet is destroyed when counter becomes 0

The IPv4 Header

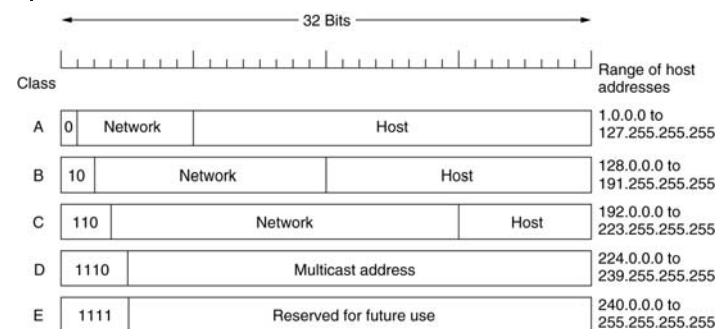
- Protocol - which transport layer protocols being used
- Header checksum - verifies header
- Options - security, error reporting, etc.
 - Some of the IP options

Option	Description
Security	Specifies how secret the datagram is
Strict source routing	Gives the complete path to be followed
Loose source routing	Gives a list of routers not to be missed
Record route	Makes each router append its IP address
Timestamp	Makes each router append its address and timestamp

IP Addresses

- Classful addressing - every host and router has unique IP address consisting of network number and host number (2 level hierarchy)
 - E.g., Class A: up to $2^7 = 128$ networks with up to $2^{24} = 16,777,216$ hosts **each**
 - No longer used, but references to it are still common
 - Network numbers are managed by ICANN (Internet Corporation for Assigned Names and numbers) to avoid conflicts

IP Addresses



How IP Packets Are Processed at a Router?

- Each router has a table listing some number of (network, 0) IP addresses and some number of (this-network, host) IP addresses
 - (network, 0) tells how to get to distant networks
 - (this-network, host) tells how to get to a local host
- When an IP packet arrives, its destination address is looked up in the routing table
 - If the packet is for a distant network, it is forwarded to the next router on the interface given in the table
 - If it is a local host, it is sent directly to the destination
 - If the network is not present, the packet is forwarded to a default router with more extensive tables
- Each router only has to keep track of other networks and local hosts, not (network, host) pairs, greatly reducing the size of routing table

How Subnets Work

- Entries in a routing table take the form (this-network, subnet, 0) and (this-network, this-subnet, host)
- A router on subnet k knows how to get to all the other subnets and also how to get to all the hosts on subnet k
- Do a Boolean AND of the destination address with the network's subnet mask to get rid of the host number and look up the resulting address in the routing table
- For example, a packet address to 130.50.15.6, the subnet mask is 255.255.252.0/22, AND them, we get 130.50.12.0 and this address is looked up in the routing table to find out which output line to use

CIDR – Classless InterDomain Routing

- Shortage of IP addresses caused by the classful addressing
 - A class is obviously too large for any organization
 - C class is too small (only 256 addresses available)
 - B class is requested and allocated, but it is still too large for most organizations => many IP addresses are wasted
- 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

Routing with CIDR

- Each routing table is extended by giving it a 32-bit mask
- A single routing table for all networks consisting of an array of (IP address, subnet mask, outgoing line) triples
- When a packet comes in, its destination IP address is first extracted
- Then, the routing table is scanned entry by entry, masking the destination address and comparing it to the table entry looking for a match
- It is possible that multiple entries (with different subnet mask lengths) match, in which case the longest mask is used
 - E.g., if there is a match for a /20 mask and a /24 mask, the /24 mask is used

CIDR – Classless InterDomain Routing

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

- Routing tables are now updated with the three assigned entries. Each entry contains a base address and a subnet mask

C: 11000010 00011000 00000000 00000000 11111111 11111111 11111000 00000000
 E: 11000010 00011000 00001000 00000000 11111111 11111111 11111100 00000000
 O: 11000010 00011000 00010000 00000000 11111111 11111111 11110000 00000000

Base address

Subnet mask

CIDR – Classless InterDomain Routing

- Now consider what happens when a packet comes in addressed to 194.24.17.4, in binary
11000010 00011000 00010001 00000100
- First it is Boolean ANDed with the Cambridge mask to get
11000010 00011000 00010000 00000000
- This value does not match the Cambridge base address, so next try Edinburgh mask, to get
11000010 00011000 00010000 00000000
- This value still does not match, so Oxford is tried, yielding
11000010 00011000 00010000 00000000
- This value matches the Oxford base. If no longer matches are found, the Oxford entry is used and the packet is sent along the line named in it

CIDR – Classless InterDomain Routing

- Aggregate entry - all three new entries can be combined into a single aggregate entry
194.24.0.0/19 with a binary address and submask as follows:
11000010 00000000 00000000 00000000 11111111 11111111 11100000 00000000
- By aggregating the three entries, a router has reduced its table size by two entries. Aggregation is heavily used throughout the Internet