Internetworking/Internetteknik, Examination 2G1305
Date: August 18th 2004 at 9:00 – 13:00

SOLUTIONS

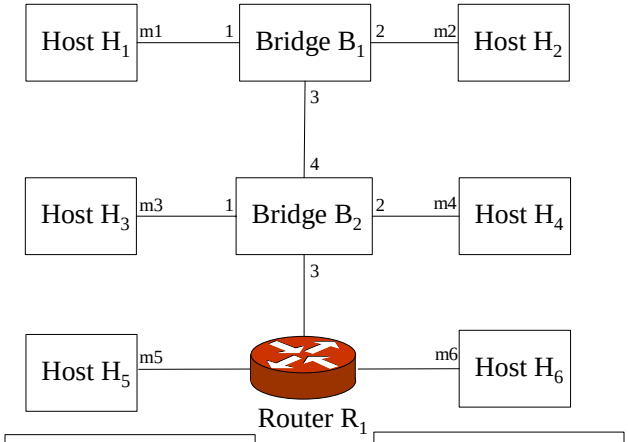
1. General (5p)

a) The so-called “hourglass” model (sometimes referred to as a “wine-glass”) has been used to illustrate the power of the Internet protocol design. Briefly describe the hourglass model. How does it illustrate the power of the Internet design in terms of support for applications and different link-layer technologies? (3p)
IP is the thin layer providing the least common denominator of different applications and varying link-level technologies. In this way, IP can be used as a “glue” to provide interconnection between many different applications and networks: IP over anything, anything over IP.

b) Typically, every layer in the TCP/IP protocol stack adds a header (or trailer) containing protocol-specific data. What is this mechanism called? (1p)
Encapsulation.

c) The Internet protocols are mainly standardized by the IETF (the Internet Engineering Task Force). The IETF produces and maintains a set of standard documents according to a standards track (from Internet drafts to Internet standards). What are these documents called? (1p)
RFC – Request For Comments.

2. Link Layer: bridging (5p)



Bridge B_1, Learning Table	
MAC	Port

Bridge B_2, Learning Table	
MAC	Port

The figure above illustrates six hosts H_1- H_6 running IPv4 over a bridged and routed network. The MAC addresses of the hosts and routers (m1-m6) are shown in the figure. The bridges B_1 and B_2 are learning bridges which switch packets(frames) between its ports. The port numbers are shown in the figure. The empty learning tables (station caches) of B_1 and B_2 are also shown in the figure. The router R1 performs IP routing between its interfaces.

a) Assume the network is in an initial state (no traffic has been sent). Now, host H_1 sends a unicast message to host H_4. H_4 then sends a reply unicast message to host H_1. Complete the bridging tables of B1 and B2 after these two packets have been

sent. Assume no other traffic has occurred.(2p)

Bridge B_1, Learning Table	
MAC	Port
m1	1
m4	3

Bridge B_2, Learning Table	
MAC	Port
m4	2
m1	4

b) Assume the network is in the state after the traffic in the previous exercise has been sent. If H_3 now sends a unicast packet to H_2, to which nodes (hosts and routers) will the packet arrive? (1p)
The bridges have not learned H2's address. It is therefore flooded to H1,H2, H4,R1

c) Assume now H1 sends a broadcast message on the link it is attached to. To which nodes (hosts and routers) will the broadcast message arrive? (1p)
H2, H3, H4, R1

d) Assume H5 sends a broadcast message on the link it is attached to. To which nodes will the broadcast message arrive? (1p)
Only R1

3. ARP (5p)

ARP – the Address Resolution Protocol – is primarily used to resolve IPv4 addresses to link-layer addresses. ARP typically works with a cache and a number of timeouts.

a) What is the purpose of the ARP cache? If no cache were used in the ARP protocol, what would happen? (2p)
The role of the ARP cache is to store resolved addresses in order to limit the amount of ARP traffic on a subnet. If resolved addresses were not stored, an ARP request/reply would be needed for every IP packet transmitted.

b) Under what circumstance would the ARP cache be useless? (1p)
The cache works because of the correlation in the use of IP addresses. A sender typically sends many IP datagrams in a row to the same destination. The cache would therefore be useless if there was no such correlation (e.g. a sender sends packets to random hosts on the link) and there is a large number of connected hosts on the link.

c) The entries in the ARP cache are controlled by timers. There are two variants of ARP cache timeouts that are used in two different situations. Which are the two variants and what is the purpose of each? What happens when the timers expire? (2p)
First, a timer is set for an incomplete entry, when an ARP request has been sent. If no ARP reply is received, the timer expires, a new ARP request is sent, and eventually the entry is cleared. The purpose of this timer is to give up if there is no host on the link with the given IP address.
Second, a timer is sent for a completed entry. When the timer expires, the entry is cleared. The purpose of this timer is to reinitiate address resolution in case the remote host crashes, or changes addresses.

4. IPv4 Addressing (5p)

Assume a network N with address/prefix 143.12.34.64/26. Two nodes are attached to the network: Router R with address 143.12.34.65 and host H with address 143.12.34.66.

a) H sends a datagram to the net-directed broadcast address of N. Which are the source and destination addresses of the IP datagram? (1p)
src=143.12.34.66, dst=143.12.34.127

b) H sends a datagram to the limited broadcast address of N. Which are the source and destination addresses of the IP datagram (1p)
src=143.12.34.66, dst=255.255.255.255

c) H sends a datagram to the loopback interface (to itself). Which are the source and destination addresses of the IP datagram (1p)
src=127.0.0.1, dst=127.0.0.1

d) Assume H has not yet retrieved its IP address and H sends an initial DHCP request. Which are the source and destination addresses of the IP datagram (1p)
src=0.0.0., dst=255.255.255.255

e) What is the subnet mask of N? (1p)
255.255.255.192

5. IPv4 and ICMP (5p)

a) Traceroute is a tool to explore the path to a given destination. Traceroute uses two methods where ICMP messages are involved to detect each hop on the way to the destination. Describe these two methods and name the ICMP messages involved. (3p)
Traceroute starts with TTL=1 and sends IP datagrams with increasing TTL levels to the destination. The intermediate routers send ICMP time exceeded back to the source, if the TTL is decremented to zero. When the destination is reached, traceroute sends a UDP datagram with an unlikely port. The destination returns an ICMP Port unreachable.

b) Using IP options, an alternative method to traceroute can be used to find the path to a given destination. Describe this method, and name at least one reason why it is of limited use. (2p)
The IPv4 record route option can be used as an alternative. Limitations of this approach include
- The limited (20-byte) option field that can only be used for a limited (9) set of hops
- Not all routers implement this functionality.

6. TCP/UDP (5p)

a) A UDP datagram is sent and the optional checksum is used. Upon reception of this packet, the receiving UDP calculates the checksum over the UDP datagram (including the checksum field itself). The result of this check is 0. Is this packet considered corrupted? Why/why not? (1p)
No, the datagram is not considered corrupted. The checksum is designed to yield the result 0 when checksum calculation is computed including the checksum field.

b) You are writing an application to send interactive unicast real-time audio over the Internet, such as an IP telephony service. Should you use TCP or UDP as your transport protocol? Briefly motivate you answer. (2p)
The preferred choice is UDP. TCP provides retransmissions and a reliable service. This may give extra delay that is undesirable for such an application.

c) You are writing an application to send video files over IP multicast. Is it possible to use TCP as your transport protocol? Briefly motivate your answer. (2p)
No. TCP is not defined for IP multicast transmissions.

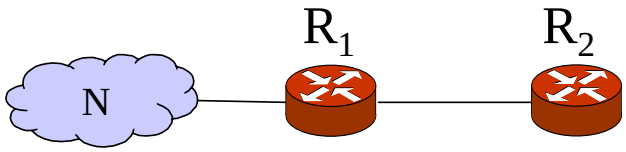
7. TCP (5p)

a) A TCP connection is using a window size of 16,000 bytes. The sending side receives a segment with acknowledgement number 48,001. Give the window at the sending side, as a range of byte numbers after the reception of this ACK. (1p)
The range is 48,001 – 54,000.

b) TCP sends a segment at 4:30:20. It does not receive an acknowledgement. At 4:30:28, it retransmits the previous segment. It receives an acknowledgement at 4:30:30. Give the values of both the RTT (Round Trip Time) and the RTO (Retransmission Time-Out) after reception of the ACK according to Karn's algorithm. When the original TCP segment was sent, the RTT was 4 seconds. (2p)
*According to Karn, RTT estimation of a retransmitted segment should not be considered. RTT should not be updated until an ACK is received without the need for retransmission. In addition, exponential back off should be used meaning that the RTO will be doubled for each retransmission. The value of RTO was 8 seconds when the original TCP segment was sent, given by 4:30:28 – 4:30:20 (also given by the formula RTO = 2*RTT = 2*4 = 8 seconds). Thus, the new values of RTT and RTO are as follows:*
RTT_{new} is 4 seconds
RTO_{new} is $2*RTO_{prev} = 2*8 = 16$ seconds

c) Briefly describe the slow start phase of TCP's congestion control mechanism. (2p)
Slow start is the first phase of TCP congestion control. The congestion window is initially set to 1 MSS (Maximum Segment Size) and then increased exponentially as the ACKs are received.

8. Dynamic routing (5p)



RIP uses the distance-vector routing algorithm. In the figure above, network N is reachable via routers R_1 and R_2. Suppose R_1 and R_2 runs RIP. If the link between R_1 and N is broken, the count-to-infinity problem may occur.

a) What is the count-to-infinity problem and why does it occur? (2p)
When the link is broken, R1 will set the cost to N to infinity (=16 for RIP). The count-to-infinity problem occurs if R2 sends an update to R1 before R1 sends it to R2. Then, R1 will believe that R2 has a valid route to N and will update its route to N to go via R2, and a loop will occur. The problem will resolve eventually after R1 and R2 have sent periodical updates, each time incrementing the cost to N, until it reaches infinity.

b) Suppose you implement the *split horizon* algorithm in the RIP implementation to solve the problem. How does *split horizon* work? (1p)
If a router receives route update information from an interface, then it may not send back updated information on that interface. This means that R2 will not send an update to R1 and R1 will not be fooled into updating its entry with a path to N via R2.

c) BGP uses path-vector instead of distance-vector. Describe how path-vector enhances distance-vector and how the count-to-infinity problem is avoided in BGP. (2p)
Path-vector adds the sequence of autonomous systems to pass (a path) in order to reach a destination network. With this information, loops that can occur in distance vector (e.g. count-to-infinity) can be avoided.

9. DHCP (5p)

a) What is the major improvement of DHCP over BOOTP? (2p)
DHCP can provide dynamic configuration of IP addresses, while BOOTP uses a predetermined mapping between physical addresses and IP addresses. The dynamics means that DHCP can give a client a temporary IP address from a pool of available IP addresses, by issuing a lease for a specified period of time. When the lease expires, the client must either stop using the IP address or renew the lease. DHCP can also provide static IP addresses, like in BOOTP. DHCP is backward compatible with BOOTP.

b) A diskless client is powered on. Since the machine has no disk, there is no configuration available during start-up. Describe how DHCP can be used to assist the client during the start-up phase in addition to providing the client's IP address. Mention at least three other pieces of information that the client can get through DHCP. (3p)
The client can get e.g., netmask, router addresses, and name server addresses. In addtion, the client can ask DHCP for a pointer to a server where the operating system software image is located. Finally, the diskless client can fetch the operating system software image from the server, e.g., through the use of TFPT.

10. MPLS (5p)

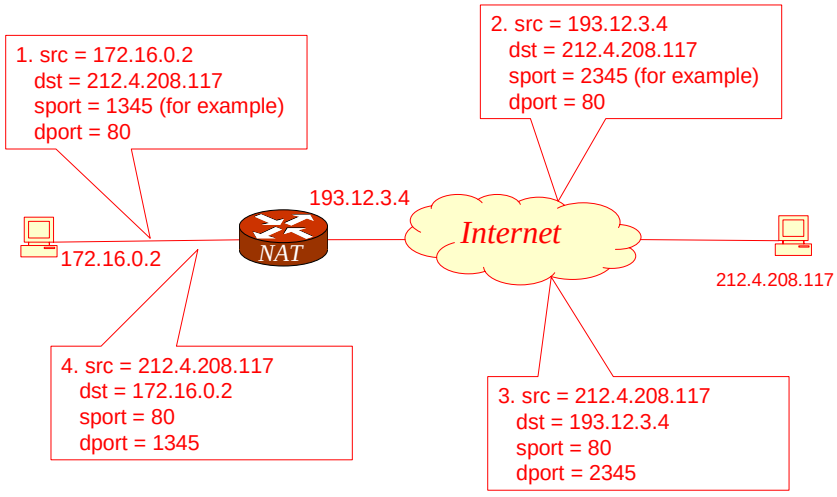
a) Briefly explain how packet forwarding in MPLS differs from regular IPv4 packet forwarding. (2p)
In an MPLS cloud, packet forwarding is based on a fixed length label instead of longest prefix match on an IP destination address. The label is applied at the ingress of an MPLS cloud and removed at the egress of the cloud.

b) There are two basic mechanisms to do path selection in MPLS. Which are the two and how do they differ? (3p)
Independent control or Hop-by-hop, where a regular routing protocol is used to determine the path.
Ordered control or Explicit routing, where the path is completely specified by the edge LSR (Label Switching Router).

11. NAT/NAPT (5p)

Assume you are administrating a private network 172.16.0.0/16, and you have just been assigned the public IP address 193.12.3.4 by your ISP. On your private network, you have many hosts and servers, all who you wish be able to access public hosts on the Internet. You therefore setup a NAT/NAPT box, which you attach to the Internet, with the global IP address on the "WAN" port and the private network on the "LAN" port. You also turn on the dynamic native address port translation (NAPT).
Assume now that host 172.16.0.2 on your private network wants to access the public web server 212.4.208.117, for example. The host issues an http request and the server replies with

a corresponding http update message. Which source and destination IP address and which TCP source and destination port numbers will the datagram have in the following locations: (1) when the http request is on the private network; (2) when the http request is on the Internet; (3) when the http reply from the server is on the Internet; and (4) when the http reply is on the private network. Assume that the web service uses the standard port 80, and that the source port used by 172.16.0.2 is already in use on the NAT/NAPT box WAN-side. (5p)



12. IPv6 (5p)

a) Show the shortest form of the following IPv6 address: 2340:0000:0000:000F:7000:119A:A001:0000 (1p)
2340::F:7000:119A:A001:0

b) What is the difference between fragmentation in IPv6 versus IPv4? (2p)
In IPv4, packets can be fragmented by the host and by routers along the path between sender and receiver (hop-by-hop fragmentation).
In IPv6, only the sending host is allowed to perform fragmentation. The sending host should learn the path MTU through path MTU discovery, or transmit packets that are small enough to fit any MTU limit.

c) Name two main problems in IPv4 that were addressed by the original design IPv6. (2p)
1) The address space in IPv4 was not considered large enough
2) The routing tables in IPv4 were getting too large
3) Security improvements needed
4) Better support for autoconfiguration (plug-and-play)