Internetworking/Internetteknik, Examination 2G1305
Date: October 19th 2004 at 14:00 – 18:00

SOLUTIONS

1. General (5p)

The OSI reference model tries to abstract communication into separate layers. The TCP/IP protocol stack is more practical and is based on the original implementation of the protocol stack. The two models are usually used interchangeably, and often map nicely to each other.

a) How do the OSI and TCP/IP layers map to each other? (2p)
OSI layers 1and 2 -> TCP/IP link layer (sometimes also physical layer)
OSI layer 3 -> TCP/IP network layer
OSI layer 4 -> TCP/IP transport layer
OSI layers 5,6,7 -> TCP/IP application layer

b) Following the OSI reference model (and the terminology used in the course), where would you place each of the following communication devices: Bridge, Router, Repeater, Gateway? (2p)
Bridge: L2, Router: L3, Repeater: L1, Gateway: L4-L7 (Gateway at L3v also accepted since "gateway" is a common name for a router)

c) Which TCP/IP layer handles naming of nodes? (1p)
The application layer in the form of DNS.

2. IP routing versus bridging (5p)

Layer 2 bridging is a popular way to build networks. But using IP routing is seen as a more scalable networking technology, in the sense that it can handle a larger number of nodes and handle more aggregated traffic. Why is this the case? Specifically, you should cover the difference in addressing and forwarding, and how this leads to better scaling in terms of nodes and traffic. (5p)
Large bridged networks are unpractical for the following reasons: (1) MAC addresses can not be aggregated, so the size of the learning tables increases linearly, (2) broadcast (eg ARP requests or unlearned destinations) are flooded over the whole broadcast domain and leads to a linear increase in the background traffic. IP routing limits the broadcast domains and aggregates the addresses and is therefore more scalable in this respect.

3. IP Addressing (5p)

a) A diskless client requesting an IP address sends an IP packet to the limited broadcast address. What are the source and destination IP addresses of this packet? (1p)
source: 0.0.0.0 destination: 255.255.255.255

b) A router is connected to the 202.33.44.128/25 subnet, using the address 202.33.44.129. It sends a multicast packets destined to all hosts on this network. What are the source and destination IP addresses of the packet? (1p)
source address: 202.33.44.129, destination address: 224.0.0.1

c) What is the subnet mask of 202.33.44.128/25? (1p)
255.255.255.128

d) A site creates an IP subnet of a block of classless addresses: 140.34.23.0/27. What is the network address of the subnet? What is the directed broadcast address of the subnet? (1p)
network address=140.34.23.0, directed broadcast address = 140.34.23.31

e) Which interval of unicast addresses can be used by hosts and routers on this subnet? (1p)
140.34.23.1-140.34.23.30

4. IPv4 fragmentation (5p)

An application sends a message which is 2400 bytes long. It sends the data as one UDP datagram. The message is then transferred from a sender to a receiver over three hops. The first link has an MTU of 1500, the second has an MTU of 1000 bytes, and the third has an MTU of 1200 bytes.

Assume that no PATH MTU discovery and that the IP header is 20 bytes (without options), and a UDP header is 8 bytes. MTU on a link means Max Transmission Unit and is the max *payload* of the link-level frame. The four fragmentation fields in the IPv4 header are: identification, don't-fragment flag (DF), more fragment flag (MF) and fragmentation offset

a) How many fragments will arrive to the IP module at the receiving host and how large will each IP fragment be (headers included)? (3p)
Two datagrams on the first hop, three on the second and third. There will be two fragments on the first hop; three fragments on the second and third hop.
At the first hop the sizes of the two fragments are:
1472(payload) +20(IPhdr) +8(UDPhdr) = 1500 bytes.
928(payload) +20(IPhdr) = 948 bytes
At the second hop (and later), the three fragments are:
968(payload) +20(IPhdr) +8(UDPhdr) = 996 bytes,
504(payload) + 20 (IPhdr) = 524 bytes
928(payload)+20(IPhdr) bytes = 948 bytes.
Note 1: UDP header only in the first fragment
Note 2: Payload length of all fragments except the first must be on 8-byte boundaries.

b) What will be the fragmentation field values of each fragment when they have arrived at the receiver? (2p)
Fragment 1: id is a 16-bit number A, DF=0, MF=1, frag off: 0
Fragment 2: id = A, DF=0, MF=1, frag off= 976/8 = 122
Fragment 3: id = A, DF=0, MF=0, frag off = (976+504)/8 = 185
Note: fragmentation offset in 8-byte units.

5. TCP (5p)

a) There is something called the "silly window" syndrome in TCP. Describe what this is and why it is a problem. (1p)
The silly window syndrome appears if TCP sends very small datagrams (called tinygrams). Sending tinygrams over a network uses the bandwidth of that network in an inefficient way and may cause (unnecessary) congestion.

b) There are two variants of the silly window syndrome. What are the two causes for the silly window syndrome? (1p)
The sender produces data slowly, or the receiver consumes data slowly

c) Name and explain two ways to solve the silly window syndromes (one for each variant). (3p)
Sender initiated: With Nagle's algorithm, the sender only has one outstanding tinygram on the network at any point in time. In this way, it accumulates one-byte

messages into larger segments until an ack is received for the previous transmission. Receiver initiated: Delayed acks: delay the sending of an acknowledgement. This potentially slows the sender, and also allows the receiver to accumulate more data in the reply message, such as piggybacking the ack on return traffic. An alternative solution is Clarks's solution which temporarily announces a window size of zero.

6. Dynamic Routing (5p)

OSPF and RIP are dynamic routing protocols for routing within an autonomous system.

a) How does OSPF handle network topology: How does OSPF partition the network, and what are the limitations of this partitioning (with respect to how partitions are connected)? (2p)
OSPF partitions network topologies by *areas*. The backbone area (area 0) may have sub-areas (but sub-areas may not be further partitioned). All traffic must pass through the backbone area. Routers connecting areas are called Area Border Routers.

b) OSPF and RIP use fundamentally different algorithms, but OSPF is said to *converge* faster than RIP. What does this mean, and why is this so? In your answer, you should compare the two protocols with respect to convergence. (3p)
OSPF is based on link-state routing, whereas RIP uses distance-vector. Link-state routing distributes original link information by flooding to every other node. Thus, every node has complete link information fast. The system can reach a correct routing state in a short time. In contrast, distance-vector uses periodic updates between neighbours to distribute information. In addition, distance-vector re-computes routes – nodes do not have access to original data. Therefore, RIP takes longer time to reach a correct state. Thus, the system may be inconsistent causing routing loops during a relatively long time.

7. DNS (5p)

a) DNS stores its data in general mapping entries called Resource Records (RRs). The following are four examples of such entries: PTR, MX, NS, SOA. What is the purpose of each entry: what mapping does each entry define (for each entry, state "from" and "to" data-type)? (2p)
PTR: IP address → name
MX: domain name → mail server name
NS: domain name → name-server name
SOA: name → zone information

b) What is a DNS domain? (1p)
A DNS domain has to do with naming. DNS names form a tree of nodes. A domain name is a part of such a tree: a sub-tree. It includes the node with the domain-name and all nodes under it.

c) What is a DNS zone? In particular, how does it differ from a domain? (1p)
DNS zones have to do with authority: A zone is a part of a domain that is managed by a primary name-server and a set of secondary name-servers. A zone is a subset of a domain: a zone may delegate parts of the domain to other zones.

d) What is a root server? (1p)
A root server is a name-server holding complete information about the top-level domains.

8. Applications: Sockets (5p)

Sketch the socket system calls of a *connection-oriented concurrent server*. You do not need to have correct data-types or error handling, just try to describe which sequence of socket system calls a connection-oriented concurrent server typically calls. You can use pseudo-code or a flow-chart, for example. An example list of system calls are the following: socket, bind, connect, accept, read, write, recvfrom, sendto, select, close. (5p)

Main-process: 1. Socket → 2. bind → 3. listen → 4. accept → 5. fork sub-process→ go to 4.
Sub-process: 1. read → 2. (process) → 3. write → 4. go to 1.

9. Applications (5p)

a) TELNET uses a specific TCP feature to transfer high priority control data, such as ^C (control-C). Which is this feature, and why does TELNET need this feature? (2p)
TELNET uses out-of-band signalling by using the TCP URG pointer and the OOB-socket feature to transfer critical control data (eg ^C) to the application. In this way, the important control data does not have to wait in line on a potential large amount of data. One example is the abortion of the listing of a large file.

b) FTP does not need to use this feature for its control messages. How does FTP solve this problem? (1p)
FTP uses separate TCP control and data channels. In this way, FTP control traffic can be transferred without having to wait for data traffic, if the application processes control before data.

c) Most Internet applications encode its protocol data in clear text (ASCII/NVT), rather than in binary encoded format (such as, for example, TCP and IP). Name one advantage and one disadvantage with this solution, in comparison with binary encoding. (2p)
Advantages: ASCII data is easier to debug, no need for byte-swapping.
Disadvantages: Less efficient in terms of bandwidth. More difficult to parse.

10. IPv6 (5p)

The IPv6 header has changed drastically from the IPv4 header. For each of the IPv4 fields below, describe what change has been made in IPv6, and why the change was made. Changes may include renaming, removal, or movement to some other header. You should be able to describe how the field has changed, with primary focus on the reason for the change.

a) Header length (1p)
The header length field has been removed, since the IPv6 header is fixed-size

b) Time-to-live (1p)
The TTL has been renamed to hop-limit, since this is the actual semantics of TTL – not time spent in the network.

c) Protocol (1p)
The field has been renamed to "next header" to better reflect the generic use of headers in IPv6.

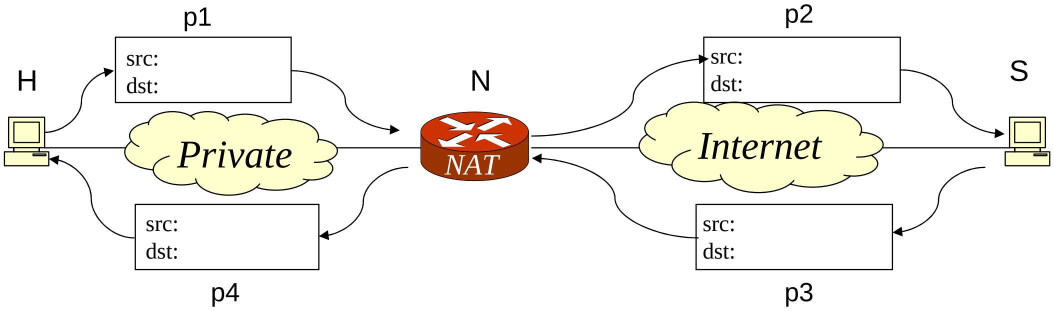
d) Header checksum (1p)
The header checksum has been removed since IP requires lower layers to provide a strong CRC protection.

e) Options (1p)
The options have been removed – most of the fields have moved to next-headers.

11. Internet security (5p)

Briefly describe a typical use of digital signatures. You should describe how the sender can produce a digital signature, and how the receiver can verify it, and why the method works. You should also name which aspects of security (according to Forouzan) that a digital signature addresses. (5p)
Public-key encryption is used. The sender makes a digest using a one-way hash function of the message. This digest is encrypted with (the sender's) private key. The signed digest is appended to the message and sent to the receiver. The receiver does two things: it decrypts the signed digest using the sender's public key, it also makes a digest using the same hash function of the message, and compares the two results. If it matches the message is accepted. The method works because it can only be the sender who knows the private key that can encode the digest properly. Digital signatures may address authentication, integrity and non-repudiation.

12. NAT (5p)



A NAT/NAPT box, N, is placed between a private and public network as shown in the figure. A host, H, on the private network with address 10.0.0.23 intends to communicate with a web server, S, on the public Internet at 123.23.4.90 on port 80. The NAT/NAPT box has one public address 23.1.2.3.

a) Fill in the source and destination address and port numbers in the four positions (p1, p2, p3 and p4) as a packet is sent from the host to the web server, and the reply is sent back. In this exercise, the NAT/NAPT box should change port numbers as well – assume that the port number used by the host was already used by some other flow. (2p)
p1: src 10.0.0.23:13450 (13450 is an example ephemeral port)
dst 123.23.4.90:80
p2: src 23.1.2.3:10390 (10390 is an example port assigned by N)
dst 123.23.4.90:80
p3: src 123.23.4.90:80
dst 23.1.2.3:10390
p4: src 123.23.4.90:80
dst 10.0.0.23:13450

b) What is the address/port mapping that the NAT/NAPT box has after the transmission in the example? (1p)
10.0.0.23:13450 ←→ 23.1.2.3:10390

c) What does it mean that a NAT/NAPT box is a “full cone NAT”? (1p)
A full-cone NAT-box makes no filtering of the specific destination host: the “hole” opened to the internal address/port can be re-used by another destination. (The mapping to the internal address/port is also deterministic, it does not change in time).

d) What does it mean if the NAT/NAPT box is a “symmetric NAT”? (1p)
A symmetric NAT opens the “hole” in the NAT only for that specific destination address/port. That is, it associates a filter with the destination, so that no other external host can re-use the mapping.