Session 2003-2004 Exam 1

EG/ES 3567 Worked Solutions. (revised)

Please note that both exams have identical solutions, however the level of detail expected in ES is less, and the questions are phrased to provide more guidance on how to provide the solution.

Dr Gorry Fairhurst

G.Fairhurst@eng.abdn.ac.uk

Marks

1.

(a) Compare the operation of the protocol layers above and below the Network Layer.
[6 marks]

The two lowest layers operate between adjacent systems connected via the physical link and are said to work "hop by hop". The protocol control information is removed after each "hop" across a link (i.e. by each System) and a suitable new header added each time the information is sent on a subsequent hop. The network layer (layer 3) operates network-wide and is present in all systems and responsible for overall co-ordination of all systems along the communications path.

Physical layer and Link layer: Provide links between network nodes (Intermediate Systems, IS).
Network layer: Provides independence from link technology. Includes routing ; transfer data between end users.
The layers above layer 3 operate end-to-end and are only used in the End Systems (ES) which are communicating. The Layer 4 - 7 protocol control information is therefore unchanged by the IS in the network and is delivered to the corresponding ES in its original form. Layers 4-7 (if present) in Intermediate Systems (IS) play no part in the end-to-end communication.

6

(b) An End System uses the "ping" program to determine if a destination End System is operational. If each sent message is of size 150B, what is the total size of the Ethernet frame sent?
[4 marks]

Determine packet headers:
Ethernet Frame Header (14B); IP Header (20B); ICMP Mesage (150 B); Ethernet Trailer (4B)
Size = 188B.
4 N.B. This calculation ignores the Inter-Frame Gap (IFG) introduced between Ethernet Frames.

(c) By comparing the operation of the "ping" program and the "traceroute" programs describe the key differences between these two programs. [8 marks]

The "ping" program contains a client interface to ICMP. It may be used by a user to verify an end-to-end Internet Path is operational. The ping program also collects performance statistics (i.e. the measured round trip time and the number of times the remote server fails to reply. Each time an ICMP echo reply message is received, the ping program displays a single line of text. The text printed by ping shows the received sequence number, and the measured round trip time (in milliseconds). Each ICMP Echo message contains a sequence number (starting at 0) that is incremented after each transmission, and a timestamp value indicating the transmission time.

The "traceroute" program also contains a client interface to ICMP. Like the "ping" program, it may be used by a user to verify an end-to-end Internet Path is operational, but also provides information on each of the Intermediate Systems (i.e. IP routers) to be found along the IP Path from the sender to the receiver. Traceroute uses ICMP echo messages. These are addressed to the target IP address. The sender manipulates the TTL (hop count) value at the IP layer to force each hop in turn to return an error message.

The program starts by sending an ICMP Echo request message with an IP destination address of the system to be tested and with a Time To Live (TTL) value set to 1. The first system that receives this packet decrements the TTL and discards the message, since this now has a value of zero. Before it deletes the message, the system constructs an ICMP error message (with an ICMP message type of "TTL exceeded") and returns this back to the sender. Receipt of this message allows the sender to identify

Marks

which system is one link away along the path to the specified destination.
The sender repeats this two more times, each time reporting the system that received the packet. If all packets travel along the same path, each ICMP error message will be received from the same system. Where two or more alternate paths are being used, the results may vary.
If the system that responded was not the intended destination, the sender repeats the process by sending a set of three identical messages, but using a TTL value that is one larger than the previous attempt. The first system forwards the packet (decrementing the TTL value in the IP header), but a subsequent system that reduces the TTL value to zero, generates an ICMP error message with its own source address. In this way, the sender learns the identity of another system along the IP path to the destination.
This process repeats until the sender receives a response from the intended destination (or the maximum TTL value is reached).

Some Routers are configured to discard ICMP messages, while others process them but do not return ICMP Error Messages. Such routers hide the “topology” of the network, but also can impact correct operation of protocols. Some routers will process the ICMP Messages, providing that they do not impose a significant load on the routers, such routers do not always respond to ICMP messages. When “traceroute” encounters a router that does not respond, it prints a “*” character.

8

(d) Explain what is meant by the term Service Access Point. [2 marks]

The protocol for each layer is concerned with providing a peer-to-peer service with the corresponding layer at the other end of the path (a hop for the lower three layers, end-to-end for the upper four). Each layer uses the services of the layers below it, by communicating via a Service Access Point (SAP).

During peer-to-peer communication, information at the sender (i.e. a Protocol Data Unit, PDU) flows down through each of the lower layers in the same node. At the lowest (physical layer) the information passes over the communications cable to the corresponding physical layer entity. When information is received, the information (a Service Data Unit, SDU) is passed up to the next higher layer.

The boundaries between adjacent layers in the same system are called Interfaces. Service Primitives are used to pass the information, and the protocol entity to which the information is delivered is called a Service Access Point (SAP). Examples of SAPs are the type field in the Medium Access Control (MAC) protocol, the address field in HDLC, the protocol field in the IP network header, and the port identifier in UDP and TCP. 2

Marks

2.

(a) Explain the algorithm used by a Network Interface Card (NIC) when transmitting frames over a shared Ethernet cable. [10 marks]

The transmitter initialises the number of transmissions of the current frame (n) to zero, and starts listening to the cable (using the carrier sense logic (CS) - e.g., by observing the Rx signal at transceiver to see if any bits are being sent). If the cable is not idle, it waits (defers) until the cable is idle. It then waits for a small Inter-Frame Gap (IFG) (e.g., 9.6 microseconds) to allow to time for all receiving nodes to return to prepare themselves for the next transmission.

Transmission then starts with the preamble, followed by the frame data and finally the CRC-32. After this time, the transceiver Tx logic is turned off and the transceiver returns to passively monitoring the cable for other transmissions.During this process, a transmitter must also continuoulsy monitor the collision detection logic (CD) in the transceiver to detect if a collision ocurs. If it does, the transmitter aborts the transmission (stops sending bits) within a few bit periods, and starts the collision procedure, by sending a Jam Signal to the transceiver Tx logic. It then calculates a retransmission time.

If all nodes attempted to retransmit immediately following a collision, then this would certainly result in another collision. Therefore a procedure is required to ensure that there is only a low probability of simultaneous retransmission. The scheme adopted by Ethernet uses a random back-off period, where each node selects a random number, multiplies this by the slot time (minimum frame period, 51.2 µS) and waits for this random period before attempting retransmission. The small Inter-Frame Gap (IFG) (e.g., 9.6 microseconds) is also added.

On a busy network, a retransmission may still collide with another retransmission (or possibly new data being sent for the first time by another node). The protocol therefore counts the number of retransmission attempts (using a variable N in the above figure) and attempts to retransmit the same frame up to 15 times. For each retransmission, the transmitter constructs a set of numbers:
{0, 1, 2, 3, 4, 5, ... L} where L is ([2 to the power (K)]-1) and where K=N; K<= 10;
A random value R is picked from this set, and the transmitter waits (defers) for a period
R x (slot time) i.e. R x 51.2 Micro Seconds

The scaling is performed by multiplication and is known as exponential back-off. This is what lets CSMA/CD scale to large numbers of nodes - even when collisions may occur. The first ten times, the back-off waiting time for the transmitter suffering collision is scaled to a larger value. The algorithm includes a threshold of 1024. The reasoning is that the more attempts that are required, the more greater the number of computers which are trying to send at the same time, and therefore the longer the period which needs to be deferred. Since a set of numbers {0,1,...,1023} is a large set of numbers, there is very little advantage from further increasing the set size.

Each transmitter also limits the maximum number of retransmissions of a single frame to 16 attempts (N=15). After this number of attempts, the transmitter gives up transmission and discards the frame, logging an error. In practice, a network that is not overloaded should never discard frames in this way.

(b) How is the algorithm modified when Network Interface Card operates in the full duplex mode? [2 marks]

2 The CSMA/CD algorithm is disabled, since the media is not shared.

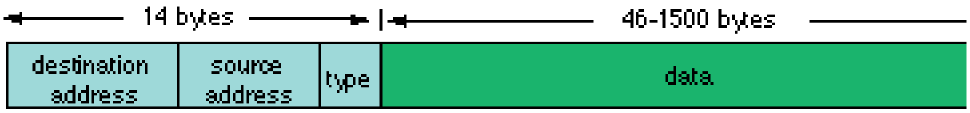
Marks

(c) Is it possible to use the full duplex mode with (i) a Hub (ii) a switch? [2 marks]

(i) No - half duplex is required for a shared media.
(ii) Yes.

2

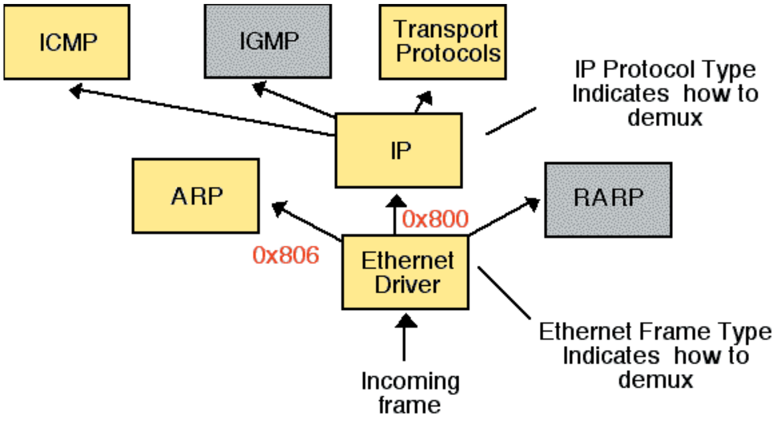
(d) Using suitable diagrams, explain the purpose of the Ethernet Frame Type Field. [4 marks]



A 2-byte type field, which provides a Service Access Point (SAP) to identify the type of protocol being carried. (Answer may also note that in the case of IEEE 802.3 LLC, this is used to indicate the length of the data part, although this is not expected.)

Answer must describe or provide diagrams to show that a SAP is both an INTERFACE between layers in the OSI reference model and a TYPE field used in demultiplexing.

4


ICMP
IGMP
Transport Protocols
IP Protocol Type Indicates how to demux
IP
ARP
RARP
0x800
0x806
Ethernet Driver
Ethernet Frame Type Indicates how to demux
Incoming frame

(e) Provide 2 examples of protocols whose operation relies on the presence of the Ethernet Frame Type Field. [2 marks]

2

two example values are:

0x0800 which is used to identify the IP network protocol

0x806 correponds to the Address Resolution Protocol, ARP

other values are used to indicate other network layer protocols.

Note the answer does NOT have to give the numeric value of the type fields.

Marks

3.

(a) An End System sends 5 packets per second using the User Datagram Protocol (UDP) over a full duplex 100 Mbps Ethernet LAN connection.
The UDP message is 1000 bytes in size (including the UDP Protocol Control Information).

(i) What is the throughput, when measured at the transport layer? [4 marks]

(ii) What is the utilisation of the link? [4 marks]

(i) Throughput

Throughput is the number of bytes transferred per second by a protocol layert using the services of thelayer below. It does not include protocol header information added by the layer itself of thelayers below. It is usally measured in bits per second.

UDP message = 1000 B = UDP Header (PCI) + Payload
Payload = 1000-8 bytes
Throughput (at UDP Layer)= (992x8)x5 bps **= *39.68 kbps.***

4

(ii) Utilisation

Utilsiation is a meaure of the cpacity used in the physical layer. It includes all protocol header information added by the layer itself andthelayers below. It is usally measured as a percentage.

Ethernet Frame Size = 1000+IP+MAC Header+CRC
= 1000 + 20 + 14 + 4 B
Utilisation = (1038x8x5) x100 /10E8 %
= 4%
(N.B. Ignoring Interframe gap).

4

(b) What may limit the maximum utilisation of a shared Ethernet network? [2 marks]

A drawback of sharing a medium using CSMA/CD, is that the sharing is not necessarily fair. When each node connected to the LAN has little data to send, the network exhibits almost equal access time for each node. The access time may increase, and utlisation decrease from “collisions” as the network becomes overloaded.
If one node starts sending an excessive number of packets, it may dominate the network. Such conditions may occur, for instance, when one node in a LAN acts as a source of high quality packetised video. The effect is known as “Ethernet Capture”.

2

(c) What is the smallest size of frame that is permitted in an Ethernet network? [2 marks]

No frame may have less than 46 bytes of payload - i.e. 64 B in total.

2

(d) Why does Ethernet define a minimum frame size, and what would be the implication of sending a frame smaller than this size? [4 marks]

To ensure that no node may completely receive a frame before the transmitting node has finished sending it, Ethernet defines a minimum frame size. The minimum frame size is related to the distance

Marks

which the network spans, the type of media being used and the number of repeaters which the signal may have to pass through to reach the furthest part of the LAN. Together these define a value known as the Ethernet Slot Time, corresponding to 512 bit times at 10 Mbps.

4

(e) Given that the Ethernet CRC-32 protects the integrity of frames sent across a Local Area Network, why does a transport protocol (e.g., the User Datagram Protocol, UDP) also include a checksum? [4 marks]

The link layer CRC protects the frame from corruption while being transmitted over the physical mediuym (cable). The CRC is removed by routers - as partr of the processing. A new CRC is added if the packet is forwarded by the router on another Ethernet link. While the packet is being processed by the router the packet data is protected by the CRC. Router processing errors may otherwise pass undetected.

The transport layer CRC therefore provides an end-to-end integrity check to ensure correctness of the data transferred.

The main purpose of the UDP Checksum is to detect problems that may arise in Intermediate Systems (where there is no CRC on the data).

4

Marks

4. Consider the network shown below in figure 1:

Figure 1: An Ethernet LAN

(a) Provide a diagram of this network clearly labelling each Collision Domain [4 marks]

Domain 1: A - Switch Port for A
Domain 2: Switch Port for Hub I; B;Hub I;C; Router Port to C
Domain 3: Router Port for Hub II; B;Hub II ;D; E
Domain 4: Router Port for Internet Feed 4

(b) Given there are two IP networks, list the End Systems present in each IP network.
[2 marks]

IP 1: A,B,C
IP 2: D,E

2 **(c) Sketch the contents of the Address Resolution Protocol (ARP) cache after the computer C has communicated with the computers A,B, and D, E, explaining the set of MAC addresses used.**
[4 marks]

System	MAC Address
A	MAC Address of A
B	MAC Address of A
?	MAC Address of Router

4 N.B. There are no entries for D & E, since these computers are in a different IP network.

(d) If computer B is reconnected directly to the switch, does the ARP cache change? [2 marks]

No. The address table in the switch will change, as the switch learns the new address, but the MAC address is associated with the interface and is a flat address, not changed by topology of the L2 network. 2

(e) If computer C wishes to communicate with a remote server in the Internet. Explain the process by which the C uses the name of the server to identify where to send the packets. [8 marks]

Find the network ID of the sender (i.e. the End System s own network ID).

Convert the source interface IP address to hex (or binary)
Convert netmask to hex (or binary)
Perform logical AND between the two

Repeat the process with the destination address to identify the destination network ID.

Compare source and destination network IDs.

If the two match, use ARP to find a MAC address, and send directly.
If they do not match, send the packet to a router (this may require a route-lookup - if routing is used, otherwise the default route is used). If the router sMAC address is not in the ARP cache, then ARP is first used. 8

Marks

(a) The End System A (in figure 2) uses the Transmission Control Protocol (TCP) to send a packet to the End System B with a payload of 10 bytes.

Sketch the Ethernet frame that is sent, showing each of the protocol headers, and the packet payload. Ensure that your sketch also shows the addresses at both the MAC and IP layers. [6 marks]

A	**MAC 0x00:11:22:33:44:55**	*****	**IP 192.7.1.1**
J	**MAC 0x22:33:44:55:66:77**	*****	**IP 192.7.1.2 (towards A)**
J	**MAC 0x44:55:66:77:88:99**	*****	**IP 192.8.8.8 (towards K)**
B	**MAC 0x66:77:88:99:00:11**	*****	**IP 192.8.8.2 ****

* One version of the exam paper had an extra byte in the MAC address field, students who copied the values in the question were not disadvantaged.
** One version of the exam paper had an invalid IP address byte for ES B, students who copied the values in the question were not disadvantaged.

Figure 2: An Internet Path between two End Systems, A and B

```
  +-----------------------------+-------------------------------------------+--------------+-----+---------+
6 | MAC J:MAC A:0x800:  | IP-SRC=A:IP-DST=B:IP-TYPE=TCP |TCP Header| Data|CRC-32 |
  +-----------------------------+-------------------------------------------+--------------+-----+---------+
```

Other appropriate fields may be supplied, based on the information in the PDU Header sheet.
Students must recognise that the ROUTER MAC address is used.

(b) Explain how the Switch I (in figure 2) learns the correct place to forward the frames it receives. [4 marks]

A bridge works within the data link layer (layer 2) of the OSI reference model. The format of PDUs at this layer in a LAN is defined by the Ethernet frame format (also known as MAC - Medium Access Control) consists of two 6 byte addresses and a one byte protocol ID / length field. The address field allows a frame to be sent to single and groups of stations. The MAC protocol is responsible for access to the medium and for the diagnosis of failure in either the hardware or the cabling.

4 The bridge learns which MAC addresses belong to the computers on each conected subnetwork by observing the source address values which originate on each side of the bridge. This is called “learning”. The learned addresses are stored in the corresponding interface address table. Once this table has been setup, the bridge examines the destination address of all packet, and forwards them only if the address does not correspond to the source address of a computer on the local subnetwork. A system administrator may overide the normal forwarding by inserting entries in a filter table to inihibit forwarding between different workgroups (for example to provide security).

Summary:
MAC Sources address observed for learning
Associated with a port in the address table
MAC Destination address observed for forwarding
Learned addreses -> forward only to specified port
Discard frames to own address

Marks

Flood frames with unkonwn addresses to all ports
Aging required and re-learning when computers change the port they are connected to

(c) Explain how the Switch I recognises multicast and broadcast frames sent by A and whether each of these are forwarded by the switch. [4 marks]

* Ethernet supports broadcast, unicast, and multicast addresses. The appearance of a multicast address on the cable (in this case an IP multicast address, with group set to the bit pattern 0xxx xxxx xxxx xxxx xxxx xxxx) is therefore as shown below (bits transmitted from left to right):

```
0                    23 IP Multicast Address Group  47
|                    | <--------------------------->|
1000 0000 0000 0000 0111 1010 xxxx xxx0 xxxx xxxx xxxx xxxx
|                           |
Multicast Bit               0 = Internet Multicast
                            1 = Assigned for other uses
```

4

* The all 1 s multicast address is used for Broadcast, i.e., 0xFFFFFF::FFFFFF

* Switches normally forward all multicast and broadcast frames.

(d) Explain the term Maximum Transmission Unit (MTU), and the procedure by which computer A may determine the smallest MTU available on the path between A and B in Figure 2. [6 marks]

The MTU is the largest size of IP datagram which may be transferred using a specific data link connection The MTU value is a design parameter of a LAN and is a mutually agreed value (i.e. both ends of a link agree to use the same specific value) for most WAN links. The size of MTU may vary greatly between different links

Path MTU Discovery
This is now the normal way of operation. The way in which the end system finds out this packet size, is to send a large packet (up to the MTU of the link to which it is connected). The packet is sent with the Don t Fragment (DF) flag set in the IP protocol header. If a router finds that the MTU of the next link exceeds the packet size, the DF flag tells the router not to segment the packet, but instead to discard the packet. An ICMP message is returned by the router (R1 in the example below) to the sender (H0), with a code saying the packet has been discarded, but IMPORTANTLY, also saying the reason and indicating the maximum MTU allowed (in this case the MTU of the link between R1 and R2).

If the end system receives an ICMP message saying a packet is too large, it sets a variable called the PATH-MTU (P-MTU) to the appropriate maximum size and then itself fragments the packet to make sure it will not be discarded next time. The end system keeps a set of P-MTU values for each IP address in use. When there are a series of links between routers along the path, each with smaller MTU s, the above process may take place a number of times, before the sender finally determines the minimum value of the P-MTU. Once the P-MTU has been found, all packets are sent segmented to this new value. Routers do not therefore have to do any additional processing for these packets.
6 Occasionally the end system will generate a large packet, just to see if a new Internet path has been found (i.e. a different route). The new path may allow a larger P-MTU.