Session 2004-2005 Exam 1

EG/ES 3567 Worked Solutions.

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) Use the Open Systems Interconnection Reference Model to explain the operation of the Transport 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.

The layers above layer 3 operate end-to-end and are only used in the End Systems (ES) which are communicating. The Transport Layer exists at Layer 4 is and is the first true end-to-end layer. The transport layer is the fourth layer of the OSI reference model. It provides transparent transfer of data between end systems using the services of the network layer (e.g. IP) below to move PDUs of data between the two communicating systems.

The transport service is said to perform “peer to peer” communication, with the remote (peer) transport entity. The data communicated by the transport layer is encapsulated in a transport layer PDU and sent in a network layer SDU. The network layer nodes (i.e. Intermediate Systems (IS)) transfer the transport PDU intact, without decoding or modifying the content of the PDU. In this way, only the peer transport entities actually communicate using the PDUs of the transport protocol.

The transport layer relieves the upper layers from any concern with providing reliable and cost effective data transfer. It provides end-to-end control and information transfer with the quality of service needed by the application program. It is the first true end-to-end layer, implemented in all End Systems (ES).

6

b) Provide two examples of protocols that operate at the Transport Layer. [2 marks]

The Internet Protocol (IP) provides two transport layer protocols:

* The Universal Datagram Protocol (UDP) (Best Effort Service)
* The Transmission Control Protocol (TCP) (Reliable Service)
- Other valid transport protocols (such as XTP, SCPS, SCTP, RMT will also be accepted).

2

(c) The “ping” program sends a message of 1000B, what is the total size of the Ethernet frame? [4 marks]

1000B message size (includes PCI for ICMP)

20B IP header = 20 B (assume no network loayer options are present)
14B Ethernet MAC Header (assume LLC/SNAP is not used)
8B Ethernet 10 MHz Preamble
4B CRC-32 Integrity Check
= 1046B (fortunately this is less than the 1500B MTU)

Ignore Idle period 9.5 microsecs (10.5 microsecs also valid IFG)

4

(d) 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 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 << diagrams may be used to illustrate these operations and will be credited>>

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]

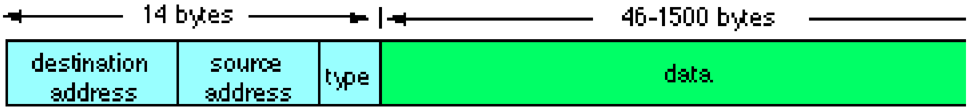
2 The CSMA/CD algorithm is disabled, (CD is ignored) since the media is not shared.

Marks

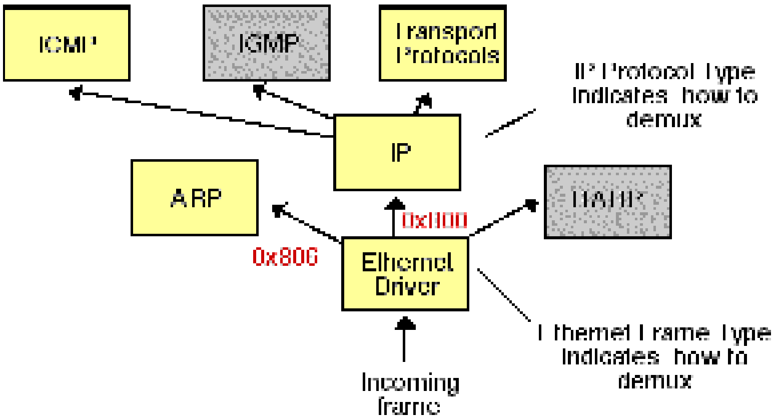
(c) Is it possible to use the full duplex mode with (i) a Hub (ii) a switch? [2 marks]

(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. In the case of IEEE 802.3 LLC, this is used to indicate the length of the data part.


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

4

(e) What types of cable are supported in the Gigabit Ethernet specification? [2 marks]

UTP CAT-5
Fibre Optic Cable.

The co-axial cable types of 10 MHz Ethernet are not supported in either 100,1000,or 100000 Etherent. 2

3.

(a) An End System sends 10 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)x10 bps **=** ***80 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 = (1038x8x10) x100 /10E8 %
= 8%
(N.B. Ignoring Interframe gap).

4 **(b) 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. This minimum frame size is determined by the Ethernet Slot Time, a design decision in DIX Ethernet v2 Spec. The minimum frame size is related to the distance 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.

2

(c) 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

(d) Figure 1 shows a part of an Ethernet Preamble. Describe Manchester Encoding and use this to explain which part of the waveform indicates the start of the MAC header.

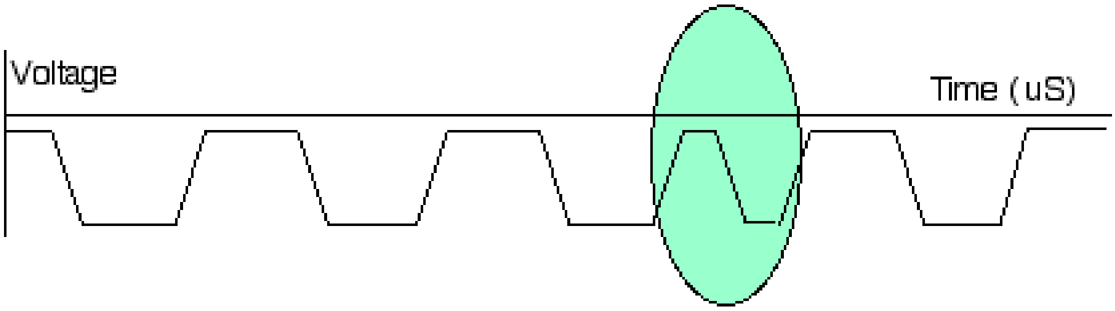
This question calls for an understandin of the preamble sequence.

A node starts transmission by sending an 8 byte (64 bit) preamble sequence. This consists of 62 alternating 1's and 0's followed by the pattern 11. Strictly speaking the last byte which finished with the '11' is known as the "Start of Frame Delimiter". When encoded using Manchester encoding, at 10 Mbps, the 62 alternating bits produce a 5 MHz square wave. n the Manchester encoding shown, a logic 0 is indicated by a 0 to 1 transition at the centre of the bit and a logic 1 is indicated by a 1 to 0 transition at the centre of the bit. Note that signal transitions do not always occur at the 'bit boundaries' (the division between one bit and another), but that there is always a transition at the centre of each bit. (N.B. since most line driver electronics actually inverts the bits prior to transmission, you may observe the opposite coding on an oscilloscopeconnected to a cable).

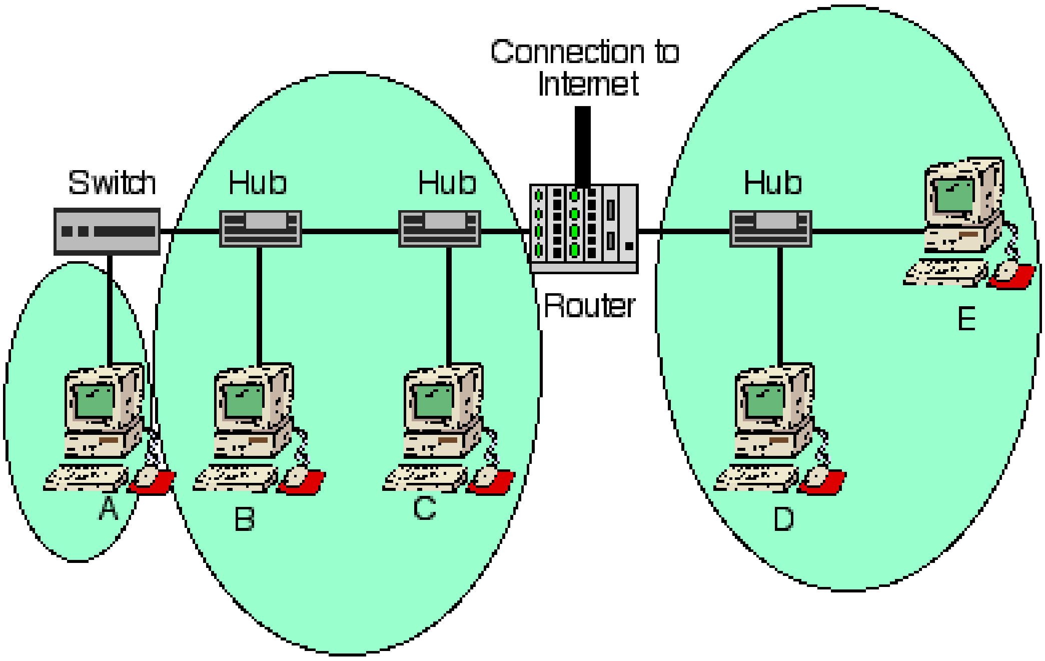
The encoding may be alternatively viewed as a phase encoding where each bit is encoded by a postive 90 degree phase transition, or a negative 90 degree phase transition. The Manchester code is therefore sometimes known as a Biphase Code.

The purpose of the preamble is to allow time for the receiver in each node to achieve lock of the receiver Digital Phase Lock Loop which is used to synchronise the receive data clock to the transmit data clock. At the point when the first bit of the preamble is received, each receiver may be in

6



Marks



(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) Which End Systems are in the same Broadcast Domain as system B?
[2 marks]

IP 1: A,C

2

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

System	MAC Address
C	MAC Address of C
A	MAC Address of A
?	MAC Address of Router

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 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.

(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

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

Marks

5 (a)

```
 0x0000:  0100 5e00 000d 00e0 f726 3ff1 0800
                                          45c0
 0x0010:  0036 5a3f 0000 0167 226b 8b85 d064 e000
 0x0020:  000d
               2300 ad3c 0100 8b85 d0d2 0001 00d2
```

Source IP address (in hex) 8b85 d064 (in standard form: 139.133.208.100

Not required:
IP 139.133.208.100 > 224.0.0.13: pim v2 Join/Prune upstream-neighbor=139.133.208.210

4 **(b) Explain how the Switch I (in figure 4) may dynamically build an Address Table that indicates the correct place to forward the frames that 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.

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

6 **(c) What is meant by the term Mulitcast? How does the Switch I recognise a multicast frame sent by A?**
* Ethernet supports broadcast, unicast, and multicast addresses. The appearance of a multicast address on the cable is therefore as shown below (bits transmitted from left to right):

```
1000 0000 0000 0000 0111 1010 xxxx xxx0 xxxx xxxx xxxx xxxx
|
Multicast Bit : 0 = Unicast   1 = Multicast (or broadcast)
```

Marks

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

* Switches normally forward all multicast and broadcast frames.

Multicast is the term used to describe communication where a piece of information is sent from one or more points to a set of other points. In this case there is may be one or more senders, and the information is distributed to a set of receivers (theer may be no receivers, or any other number of receivers).

One example of an application which may use multicast is a video server sending out networked TV channels. Simultaneous delivery of high quality video to each of a large number of delivery platforms will exhaust the capability of even a high bandwidth network with a powerful video clip server. This poses a major salability issue for applications which required sustained high bandwidth. One way to significantly ease scaling to larger groups of clients is to employ multicast networking.

Multicasting is the networking technique of delivering the same packet simultaneously to a group of clients. IP multicast provides dynamic many-to-many connectivity between a set of senders (at least 1) and a group of receivers. The format of IP multicast packets is identical to that of unicast packets and is distinguished only by the use of a special class of destination address (class D IP address) which denotes a specific multicast group. Since TCP supports only the unicast mode, multicast applications must use the UDP transport protocol The majority of installed LANs (e.g. Ethernet) are able to support the multicast transmission mode. 4

(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

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 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. 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. 6