THE UNIVERSITY *of* ADELAIDE

Primary Examination, Semester 1, 2016

Computer Networks & Applications
COMPSCI 3001, 7039

Official Reading Time:	10 mins
Writing Time:	120 mins
Total Duration:	130 mins

Questions	Time	Marks
Answer all 5 questions	120 mins	120 marks
		120 Total

Instructions

- Begin each answer on a new page in the answer book.
- Examination material must not be removed from the examination room.

Materials

- Simple calculators without alpha-numeric memory or communication capabilities allowed.
- Foreign language paper dictionaries permitted.

DO NOT COMMENCE WRITING UNTIL INSTRUCTED TO DO SO

Application Layer

Question 1

(a) What information is used by a process running on one host to identify a process running on another host?

[2 marks]

(b) Why does HTTP run on top of TCP rather than on UDP?

[2 marks]

(c) Consider an HTTP client that wants to retrieve a Web document at a given URL. The IP address of the HTTP server is initially unknown. What transport and application-layer protocols besides HTTP are needed in this scenario (list all that will be used)?

[3 marks]

(d) Consider a short, 10-meter link, over which a sender can transmit at a rate of 150 bits/sec in both directions. Suppose that packets containing data are 100,000 bits long, and packets containing only control (e.g., ACK or handshaking) are 200 bits long.
Assume that N parallel connections each get $1/N$ of the link bandwidth.
Now consider the HTTP protocol, and suppose that each downloaded object is 100 Kbits long, and that the initial downloaded object contains 10 referenced objects from the same sender.
Firstly, consider parallel downloads via parallel instances of non-persistent HTTP. Secondly, consider persistent HTTP. Do you expect significant performance gains over the non-persistent case? Justify and explain your answer (please clearly show your reasoning and summary calculations).

[12 marks]

(e) i. The equation for determining the download time for a peer-to-peer file sharing is given below.

$$D_{p2p} \geq \max(F/u_s, F/d_{min}, \frac{NF}{u_s + \sum_{i=1}^{N} u_i})$$

Explain what each of the following terms represents (for example: "The time for one peer to download a single copy of the file")

1. F/u_s
2. F/d_{min}
3. $u_s + \sum_{i=1}^{N} u_i$

[3 marks]

ii. Using the same variables (F, u, d) give an equation for the minimum download time for a client/server file download.

[3 marks]

1. a) port number , ip address

b) TCP provides reliable transport , as opposed to UDP best effort. Nessecary for HTTP to have loss control.

c). DNS - translate the domain url into an ip address.
TCP - reliable transport needed for HTTP

d) Yes, there is a bit of overhead in non persistent connections - needing to establish a connection for each download,

Non persistent.

10 objects at 100,000 bits each , 10 packets

assume we use a connection for each packet

each packet has 3 connection control packets and 4 termination control packets

$= 10 \times (4+3) \times 200$ bits

$= 70 \times 200$

= extra 14,000 bits to send

persistent

assume a single connection sending everything.

$= 1 \times (4+3) \times 200$ bits

= 1400 extra bits to send

Assume that the full bandwidth is avaible in both scenarios.

Non Persistent

$= 100{,}000 \times 10 + 14{,}000$

$= 1{,}014{,}000$ bits , 150 b/s

= 6760 seconds

= ~113 minutes

Persistent

$= 100{,}000 \times 10 + 1400$

$= 1{,}001{,}400$ bits, 150 b/s

= 6676 seconds

= ~111 minutes

∴ persistent = better.

e) i) 1. The time that the server (original seed) can upload the entire file elsewhere such that availability becomes ≥ 2 (the file exists in the swarm twice)

2. The time the peer with smallest download speed can download the file.

3. The total upload speed for the entire swarm of peers.

ii). $D_{cs} \geq \max\left(\frac{F}{u_s}, \frac{F}{d_{min}}\right)$

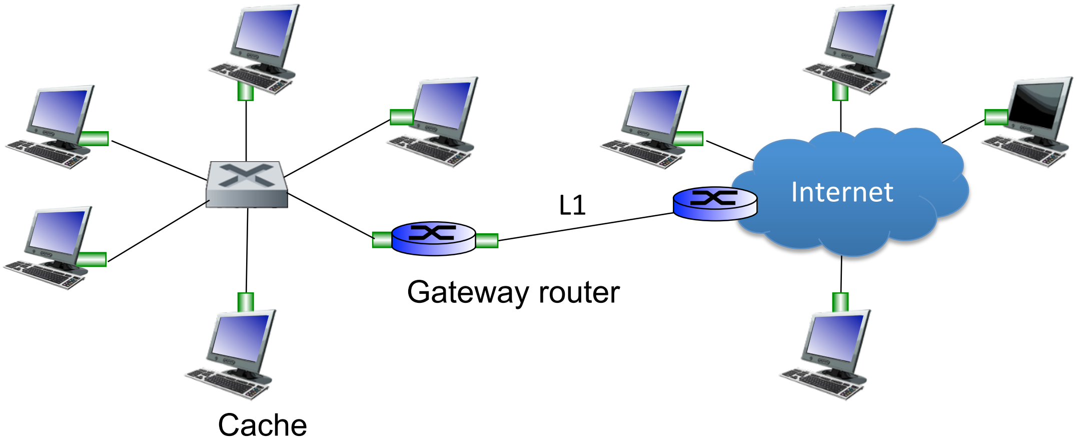


Figure 1: Network with HTTP cache

(f) Consider the network shown in figure 1. The hosts in the left network are requesting resources using HTTP from the origin servers on the Internet. 30% of web requests result in a cache hit (the request has been previously downloaded and can be served from the cache). Each host (not including the cache) generates a request every 20 seconds and each request requires the download of 1Mbit. What download bandwidth is needed on the access link, L1, in order for this HTTP download traffic to utilise a no more than 80% of the link download capacity?

[4 marks]

[Total for Question 1: 29 marks]

f) 4 requests every 20 seconds

30% are cache hits

$4 \times \frac{7}{10}$ = 2.8 requests every 20 seconds

= 2.8 mbit/20s

= 0.14 mbit/s

$\frac{0.14}{x} = 0.8$

$x = \frac{0.14}{0.8}$

= 0.175 mbit/s for 80% link utilisation

Transport Layer

Question 2

(a) Answer True (T) or False (F). Please note that **each wrong answer will incur a penalty of 1 mark** so guess at your own risk. Please only write the number of the statement and T or F in your answer books.

1. Each TCP socket is identified with a 2-tuple: source port number; and destination port number.
2. Host A is sending Host B a large file over a TCP connection. Assume Host B has no data to send Host A. Host B will not send acknowledgements to Host A because Host B cannot piggyback the acknowledgements on data.
3. Suppose Host A is sending Host B a large file over a TCP connection. The number of unacknowledged bytes that A sends cannot exceed the size of the receive buffer.
4. Suppose Host A sends one segment with sequence number 38 and 4 bytes of data over a TCP connection to Host B. In this same segment the acknowledgement number must be 42.
5. With the SR protocol, the receiver will only acknowledge packets within it's current window.
6. With GBN, it is possible for the sender to receive an ACK for a packet that falls outside of its current window.

[12 marks]

(b) Why does TCP include a deviation in its timeout calculation? Why doesn't it just use the Estimated RTT?

[4 marks]

(c) Consider a reliable data transfer protocol that uses only negative acknowledgements (NAK).

i. Suppose the sender sends data only infrequently. Would a NAK-only protocol be preferable to a ACK-only protocol? Why (think about the receiver's view and recovery from a failed transmission)?

[3 marks]

ii. Now suppose the sender has a lot of data to send and the end-to-end connection experiences few losses. In this second case, would a NAK-only protocol be preferable to a protocol that uses ACKs? Why ?

[2 marks]

iii. Given a choice of the ACK-only scheme in TCP and the NAK-only scheme just described which scheme will you use if you protocol needed to work under high data rate and low BER conditions?

[2 marks]

2a).

1. False - 4 tuple (source IP, port . Destination IP, port)
2. False - they can be sent standalone
3. True
4. False. we don't know what B sent before - the ACK could be anything
5. True.
6. False. An ACK can only be sent

2 b) deviation is used to account for variability in RTTs. If estimated RTT was used standalone, you might run into issues with early timeouts / waiting too long to timeout during volatile network conditions.

c) i. NO. The primary goal with removing NaKs is to reduce control overhead. If messages are infrequent, NAKs can be beneficial in identifying garbled data faster / more reliably than a timeout can. Timeouts can be innacurate and lead to sending data again unnessecarily. Naks fix this.
Overall communication overhead is also reduced. Only send a NAK on a bad packet, rather an ACK for every good one. However, with loss, NAK only can lead to packets being lost forever. A Nak is lost in transit and the sender thinks it went ok.

ii. Still no, communication overhead reduced significantly reduced, but even with some loss, this protocol is still unreliable. Perhaps if there was a baseline for reliability (eg. 1/100000 loss), then NAK only could be preffered.

iii. ACK only. Reliabity is the most important thing here, over efficiency.

(d) Slow-start and Congestion avoidance are two phases of TCP's congestion avoidance mechanism.

  i. At what point does TCP change from slow-start to congestion avoidance and how is this value determined?

  [3 marks]

  ii. Why should three duplicate ACKs trigger fast retransmit/fast recovery instead of dropping back to slow start?

  [2 marks]

[Total for Question 2: 28 marks]

d). i). Once congwin ≥ ss thresh. ssthresh is initially some large value, and halved every time there is a timeout.

ii). - Reduces time spent in recovery
- better utilisation of bandwidth to halve and go straight to congestion avoidance
- faster convergence to optimal congestion window.

Routing and Internet Protocol

Question 3

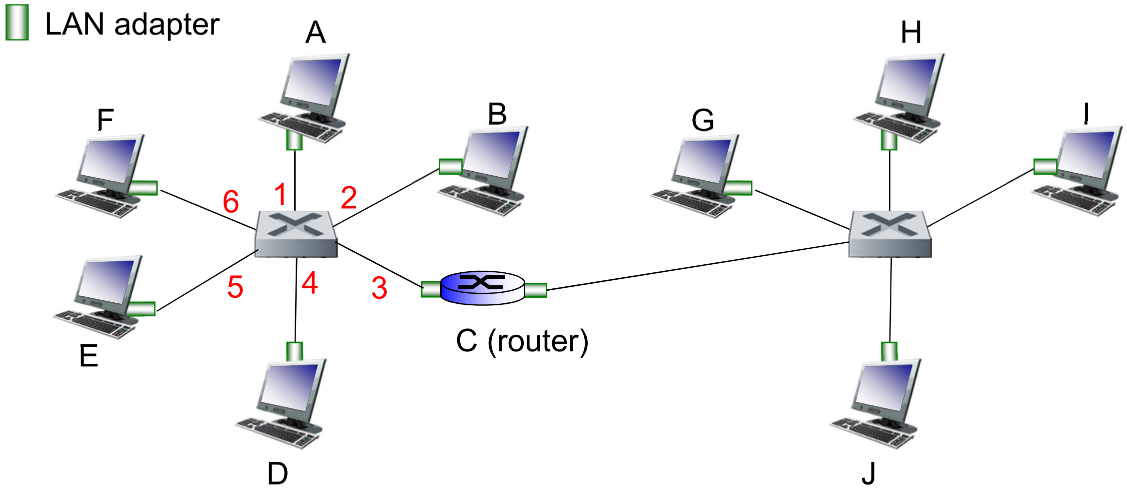


Figure 2: LAN network diagram

(a) You are asked to assign IP addresses to all of the interfaces found in the network identified by A, B, C, D, E, F shown in Figure 2. In your assignment you must make sure that you do not waste any IP addresses given that you only have **five** hosts and a router. Please clearly show your work to justify the IP address assignment and write down three IP addresses for A, B, and C.

[4 marks]

(b) Consider a datagram network using 8-bit host addresses. Suppose a router uses longest prefix matching (like the example in lecture where the organisation changed providers) and has the following forwarding table:

Prefix Match	Interface
1	0
10	1
111	2
otherwise	3

For each of the four interfaces, give the associated range of destination host addresses and the number of addresses in the range.

[6 marks]

3 a) Using CIDR, we can waste as little IP address space as possible.

6 addresses + broadcast = 7 total

111 (32 bits total)

= 29 bits for network part

$$\underbrace{10000000 \quad 11000000 \quad 10000000 \quad 00000}_{\text{network part}}\,\underbrace{111}_{\text{host part}}$$

converting to decimal

1×2^7 $\quad 1 \times 2^7 + 1 \times 2^6$ $\quad 1 \times 2^7$ $\quad 2^0 + 2^1 + 2^2$

128.192.128.1 /29 -> A ip

128.192.128.2 /29 -> B ip

128.192.128.3 /29 -> C ip

b). interface 0

= 1000 0000 -> 1111 1111

total : 2^7 = 128 total addresses

interface 1

= 10000000 -> 1011 1111

total = 2^6 = 64 total addresses

interface 2

= 1110 0000 -> 1111 1111

total = 2^5 = 32 total addresses

interfacl 3 (otherwise is 0 prefix)

= 0 000 000 -> 0111 1111

total = 2^7 = 128 total addresses

(c) The Figure 3 shows a three-node network that uses a distance-vector protocol for routing.

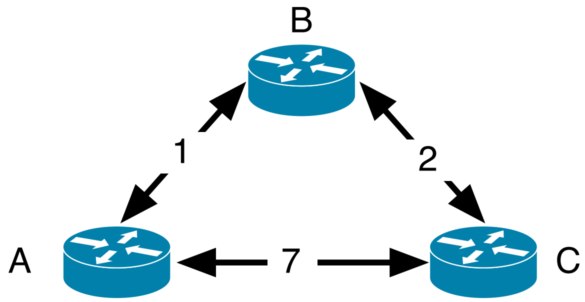


Figure 3: A three-node network showing bi-directional link costs

i. Assuming that 999 represents the value 'infinity', what are the initial routing tables for each node?

[3 marks]

ii. Using the distance-vector protocol, calculate and show the final, converged, routing tables for all nodes. You should clearly show all working – marks will be deducted if you fail to show your working.

[12 marks]

(d) What network characteristic(s) would make distance vector a better choice than link state route determination?

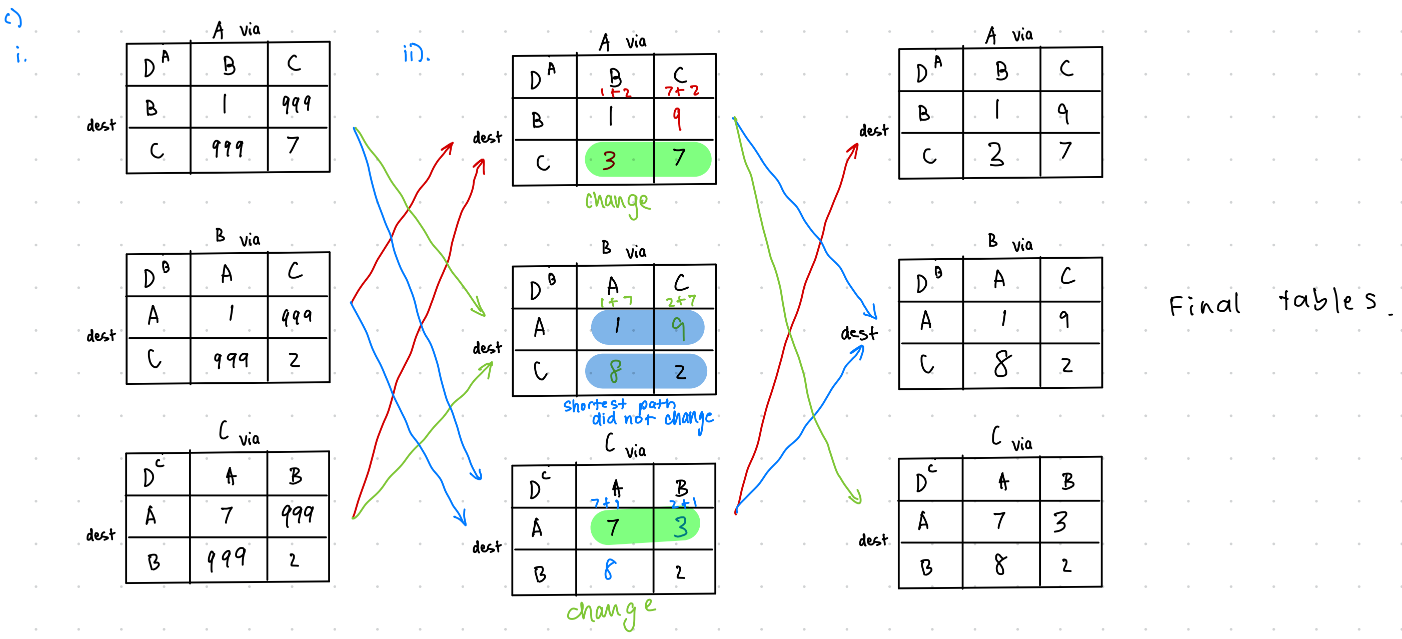
[2 marks]

(e) When IPv4 fragments a packet, how does it know which fragments belong to the original packet when it has to reassemble them?

[2 marks]

[Total for Question 3: 29 marks]

c)



i.

A via

dest

D^A	B	C
B	1	999
C	999	7

B via

dest

D^B	A	C
A	1	999
C	999	2

C via

dest

D^C	A	B
A	7	999
B	999	2

ii).

A via

dest

D^A	B 1+2	C 7+2
B	1	9
C	3	7

change

B via

dest

D^B	A 1+7	C 2+7
A	1	9
C	8	2

shortest path did not change

C via

dest

D^C	A 7+1	B 2+1
A	7	3
B	8	2

change

A via

dest

D^A	B	C
B	1	9
C	3	7

B via

dest

D^B	A	C
A	1	9
C	8	2

C via

dest

D^C	A	B
A	7	3
B	8	2

Final tables.

d). simpler network topology, low bandwidth / CPU & memory constraints

e). Headers, Identification says which fragments belong with each other, offset says their order.

Switching and Link Layer Protocols

Question 4

(a) Consider the LAN in Figure 2

i. How many subnets can you find? Write down the labels (letters) of the hosts in each subnet.

[2 marks]

ii. What is the physical topology of the subnet(s)?

[1 mark]

iii. How does the switch know A is reachable via interface 1, B reachable via interface 2 and so on (Use a simple example to briefly explain your answer)?

[3 marks]

(b) Explain how a packet enters a *Multi-Protocol Label Switching* (MPLS) network, how it is routed through the network and how it is ultimately delivered. Consider using a simple diagram to help you explain your answer.

[6 marks]

(c) Suppose two nodes start to transmit at the same time a packet of length L over a broadcast channel of rate R. Now, denote the propagation delay between the two node as d_{prop}. Will there be a collision if $d_{prop} < L/R$? Explain why or why not.

[2 marks]

(d) Suppose the information portion of a packet contains 10 bytes consisting of the integers 1 through to 10 where each integer is represented using a single 8-bit unsigned binary number representation. Compute the Internet checksum for this data.

[4 marks]

(e) You have received the following data and crc code: 100110001. The generator being used is: 1101 Is there an error in this data or has it been received correctly? Show your work.

[5 marks]

[Total for Question 4: 23 marks]

4 a). i. 2 subnets

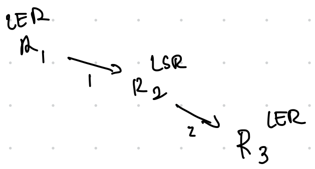
- A,B, D,E,F
- G,H,I,J

ii. Star

iii. MAC address tables employ self learning to store which hosts are reachable via which interface.

eg. A sends packet to B via switch. MAC table at switch learns where A is. It does not know where B is, so it floods all ports. B now sends packet to A, switch learns where B is and forwards directly to A.

b). ingress LER assigns label to packet, forwards to outgoing interface. LSRs unpack label and add a new one, routing it until reaches an egress LER, where the packet is extracted.



Router	incoming	outgoing	interface
R_1	–	100	1
R_2	100	200	2
R_3	200	–	–

c). Yes, this equation doesn't matter, because both packets have the same transmission delay, and are propagating at the same time. = collision

d). 1 and 2 => 0000000100000010

3 and 4 => 000000110000 0100 +

000001000000 0110

5 and 6 => 0000 0101 0000 0110 +

0000 1001 0000 1100

7 and 8 0000 0111 0000 1000 +

0001 0000 0001 0100

9 and 10 0000 1001 0000 1010 +

sum = 0001 1001 0001 1110

one's complement = 1110 0110 1110 0001 = checksum

e). convert to polynomial

8 7 6 5 4 3 2 1 0

$$100110001 = 1 + x^4 + x^5 + x^8$$

3 2 1 0

$$1101 = 1 + x^2 + x^3$$

divide

$$
\begin{array}{r}
x^5 + x^4 + x^3 + x^2 + x \\
x^3 + x^2 + 1 \overline{\big) x^8 + x^5 + x^4 + 1} \\
-(x^8 + x^7 + x^5) \\
\hline
x^7 + x^4 + 1 \\
-(x^7 + x^6 + x^4) \\
\hline
x^6 + 1 \\
(-x^6 + x^5 + x^3) \\
\hline
x^5 + x^3 + 1 \\
-(x^5 + x^4 + x^2) \\
\hline
x^4 + x^3 + x^2 + 1 \\
-(x^4 + x^3 + x) \\
\hline
\end{array}
$$

remainder should be 0

remainder = $x^2 + x + 1$

∴ error has occurred.

```
                1 1 1 1 1
          _____________________
1 1 0 1 ) 1 0 0 1 1 0 0 0 1
          1 1 0 1 ↓ ↓ ↓ ↓ ↓
        ---------------------
          Ø 1 0 0 1 │ │ │ │
            1 1 0 1 ↓ │ │ │
        ---------------------
            Ø 1 0 0 0 │ │ │
              1 1 0 1 ↓ │ │
        ---------------------
              Ø 1 0 1 0 │ │
                1 1 0 1 ↓ │
        ---------------------
                Ø 1 1 1 0 │
                  1 1 0 1 ↓
        ---------------------
                  Ø Ø 1 1 1
                      └───┘
```

remainder $= x^2+x+1$

ICMP and Security

Question 5

(a) Explain how traceroute uses ICMP in order to trace the routers along the path to a destination

[5 marks]

(b) What are the differences between message confidentiality and message integrity?

[2 marks]

(c) Suppose N people want to communicate with each of N-1 other people using symmetric key encryption. All communication between any two people, i and j, is visible to all other people in this group of N, and no other person in this group should be able to decode their communication. How many keys are required in the system as a whole? Now suppose that public key encryption is used. How many keys are required in this case?

[4 marks]

[Total for Question 5: 11 marks]

End of exam

5a). Traceroute sends packets to the destination with increasing TTL from 1. Each router in the path will decrement it by 1. If the packet reaches TTL = 0, the router will send an ICMP time exceed message. Since we are sending incrementing TTLs from 1, each router in the path will send a time exceeded message in order, allowing the trace of the route.

b) confidentiality: things hidden can only be read by the intended reciever

integrity: messages are kept intact and not altered.

c) no public key: a key for every pair: $\frac{N(n-1)}{2}$

publick key: a key for every person: 2N