Network Systems (201600146/201600197), Test 1 resit

September 7, 2020, 18:15–20:15

(Taken online on campus)

1. Applications

1 pt **Q 1.1** Which of the following protocols can be used to read emails in a mailbox? Mark all that apply.

A. IMAP
B. POP3
C. SMTP
D. none of the above

→ A,B

1 pt **Q 1.2** When delivering messages to users in a certain domain, how can a mail server ensure that it delivers the message to the right host, where the mail server of that domain is running?

A. By requesting the MX record for that domain from the DNS system
B. The network address translator of the domain will forward the TCP packets to the right server
C. By probing port 25 at the IP address of the destination domain
D. By delivering the message to a gmail server, which will deliver the message to the mail server for the destination domain.
E. By making use of the so-called anycast service.
F. By hashing the ZIP-code and P.O.Box number of the destination domain

→ A

Many chose answer B. However, servers would typically not be behind a NAT. Furthermore, even if it were, still the MX record from the DNS (answer A) would be used to find which IP address to connect to, in this case, the address of the public side of the NAT.

2 pt **Q 1.3** HTTP requests contain a Referer header, which indicates the URL of the page on which the user found the URL he/she is currently fetching. Which statement is true? Mark all that apply.

A. This compromises the user's privacy.
B. This improves security.
C. This is an alternative to sending a cookie.
D. This speeds the protocol up by allowing caching.
E. This is needed for the web server to deliver the correct file.
F. This is needed in case the user clicks the back button of the browser.
G. This is of interest to the owner of the website.
H. None of the above is true.

→ A,G

1 pt **Q 1.4** Compare HTTP/1.0 without parallel connections, to HTTP/1.1 with persistent and parallel connections, for loading an HTML page with 3 images.

A. 1.1 is faster than 1.0 regardless of where the images are hosted.
B. 1.1 is faster than 1.0 only if the images are hosted on a different server than the HTML.
C. 1.1 is faster than 1.0 only if the images are hosted on the same server as the HTML.
D. 1.1 is faster than 1.0 only if the images are hosted on different servers.
E. In this case, 1.1 and 1.0 are equally fast.
F. In this case, 1.0 is faster than 1.1.

→ A

1 pt **Q 1.5** Compare HTTP/1.1 with persistent and parallel connections to HTTP/2, for loading an HTML page with 3 images.

A. the ”push” function of HTTP/2 makes it faster than 1.1 regardless of where the images are hosted.
B. the ”push” function of HTTP/2 makes it faster than 1.1 only if the images are hosted on a different server than the HTML.
C. the ”push” function of HTTP/2 makes it faster than 1.1 only if the images are hosted on the same server as the HTML.
D. the ”push” function of HTTP/2 makes it faster than 1.1 only if the images are hosted on different servers.
E. the ”push” function of HTTP/2 does not affect the download speed compared to 1.1 for this particular case
F. the ”push” function of HTTP/2 makes it slower than 1.1 for this particular case.

→ C

Many chose answer A. However, the push function won't help in all cases. The push function means that when the client requests an HTML file, the server can already also start sending (”pushing”) the image files needed for that web page. But the server serving the HTML cannot serve image files that are hosted elsewhere, so there's only something to be gained if the images are on the same server as the HTML.

Donald, Boris, and Jair want to exchange the new computer game Autocracy. Boris' and Jair's homes have a fast internet connection (1 Gbit/s up and down). Unfortunately, Donald's great white house only has a slow Internet connection (10 Mbit/s up and 100 Mbit/s downlink). They agree to use a peer-to-peer file distribution system to distribute the computer game, which consists of a file of 1 GByte. We assume one of them has the original file, and are interested in the time it takes to distribute the file to the two others.

2 pt **Q 1.6** How much time will file distribution take if Donald has the original file?

→ 800 s

Q 1.7 Please explain how you got to that answer:

→ In general, there are three possible limiting factors for the file distribution time in a peer-to-peer system, as expressed by the max(...) in the formula. In this case, the time based on the server's uplink (in this case, that's Donald) is the largest of the three, so it takes F/u_s = 1 GByte / 10 Mbit/s = 8 Gbit / 10 Mbit/s = 800 s.

Note that there are no points marked in the margin for this question. That means the points marked at the previous question are actually for both of these questions together.

2 pt **Q 1.8** How much time will file distribution take if Boris has the original file?

→ 80 s

Q 1.9 Please explain how you got to that answer:

→ In this case, Donald's downlink is the limiting factor, so it takes F/d_{min} = 1 GByte / 100 Mbit/s = 8 Gbit / 100 Mbit/s = 80 s.

Continued on next page...

2. Performance

In order to have a hotline whenever they are playing their favourite game Autocracy, Donald and Boris have installed a dedicated fiber-optic link between the Great White House and the house at 10 Drowningstreet. This fiber optic link is 6 250 km long, and signals travel at 200 000 km/s through it. The datarate is 10 gigabit/s and packets are 10 000 bits long.

2 pt **Q 2.1** Give the propagation time of a packet.

→ 31.25 ms

Q 2.2 Please explain how you got to that answer:

→ distance / speed = 6 250 / 200 000

2 pt **Q 2.3** Give the transmission time of a packet.

→ 1 μs

Q 2.4 Please explain how you got to that answer:

→ number of bits / bit rate = 10 000 / 10 000 000 000

2 pt **Q 2.5** Suppose the stop and wait algorithm is used for reliable transmission. What is the maximum possible throughput?

→ 16 packet/s or 160 kbit/s.

Q 2.6 Please explain how you got to that answer:

→ A packet transmission takes 1 μs. After the one-way propagation time of 31.25 ms, the packet arrives at the destination. Transmitting an ACK packet takes negligible time. After another 31.25 ms propagation time this ACK will arrive at the original transmitter, which can then send the next packet. So, the max throughput is 1 / 0.062501 = 16 packet / s. This equals 16 $\times$ 10 000 = 160 kbit/s.

Disappointed with the performance of their fiber optic link, Donald asks one of his genius engineers to improve the performance. The engineer replaces the stop and wait algorithm with a sliding window algorithm with send window size 200 and receive window size 100.

2 pt **Q 2.7** What is the maximum possible throughput for this improved system?

→ 3200 packets/s or 32 Mbit/s

Q 2.8 Please explain how you got to that answer.

→ 200 / 0.062501 = 3200 packets /s = 32 Mbit/s

2 pt **Q 2.9** How many different sequence numbers should be available in the header of the packet for a sliding window protocol with the mentioned window sizes?

→ 300

This is SWS+RWS. Note that 300 is enough, 301 are not strictly needed. We still gave half a point for 301 though.

Q 2.10 Explain (e.g., by giving an example) what would go wrong if the number of available sequence numbers would be one less than the number mentioned above

→ If 299 sequence numbers would be available, they would (for instance) range from 0 – 298. Suppose 0 – 199 correctly received but acks lost. Sender will retransmit 0. The receiver cannot distinguish this packet from a packet that would be received if the acks are not lost, and the sender will transmit the next 100 packets (with sequence numbers 200 until 298, and after wrap around 0) and the first 99 of these are lost.

Continued on next page...

3. Information theory

Boris is setting up a monitoring system for the monitoring of swim water quality along the coast. Each sensor periodically generates a message, which can be one of four possibilities:

message	probability
Clean	50%
Dirty	40%
Unsafe	9%
Error	1%

2 pt **Q 3.1** Calculate the amount of information in each message. Outcome:

→ 1.40786 bits

Q 3.2 Explain your calculation:

→ $.5 \cdot \log_2(1/.5) + .4 \cdot \log_2(1/.4) + .09 \cdot \log_2(1/.09) + 0.01 \cdot \log_2(1/.01)$

1 pt **Q 3.3** Calculate the amount of information in each message. If the outcome of such a calculation is between 0 and 1, what does that mean?

A. This is not possible, there must be a calculation error.
B. Then the formula is not applicable; since less than 1 bit makes no sense, the correct answer then is 1 bit.
C. The data is so redundant that occasionally, the message can be sent without sending any bit.
D. By combining consecutive messages, the data can be compressed to less than 1 bit per message.
E. There were too many errors (the last message) in the data.

→ D

Propose a suitable code (bit pattern for each of the four messages) that achieves an average message length of not more than 1.66 bits;

2 pt **Q 3.4** Code for "Clean":

→ 0

Q 3.5 Code for "Dirty":

→ 10

Q 3.6 Code for "Unsafe":

→ 110

Q 3.7 Code for "Error":

→ 111

Of course, there are many variations possible here; but watch out that the code must be uniquely decodable, i.e., there must be only one possible sequence of messages for a given received string of 0s and 1s.

The messages need to be sent over a binary symmetric channel, with a raw speed of 10 bits/second and an error probability p of 5

2 pt **Q 3.8** Outcome:

→ 5.068

Q 3.9 Explain your calculation:

→ First calculate the channel capacity for this binary symmetric channel, using the formula from the reader: 7.136 bits/s; then divide this by the entropy of the messages calculated in question 1.

Many forgot to divide by the entropy (that's the amount of information per message); with that error, the answer is still worth 1 point. Some divided not by the entropy but by the average length of their own code (from the previous question). This is not right, as the question is about an optimal encoding; but you could still earn 1.5 points if you *made that error.*

1 pt **Q 3.10** "Channel capacity" is: how many bits can be sent per second through a channel ...

A. ... swimming from England to France.
B. ... without using an error-correcting code.
C. ... and achieve a given error probability.
D. ... and achieve an extremely low error probability.
E. ... while guaranteeing no errors will occur.

→ D

The actual statement is that one can achieve an "arbitrarily low" error probability: as low as one wants. Hence the answer D.
We've also accepted C, since "given error probability" can be read as "any given arbitrarily low error probability". (But when writing the question I thought of C as "given in advance, not necessarily very low", and then it's not quite correct.)

1 pt **Q 3.11** Suppose we need to transport 64 databits, and have a choice between using (i) four 5 x 5 parity matrices, or (ii) one 9 x 9 parity matrix. What's the difference?

A. Option (i) can correct more bit errors than option (ii)
B. Option (i) can correct fewer bit errors than option (ii)
C. Both options can correct equally many bit errors
D. Parity matrices are not suitable for error correction, only for detection.
E. Option (i) is not possible, all bits must go into a single matrix.

→ A

Many chose option E, but that is incorrect. We can simply split the 64 databits into 4 groups of 16, and apply a parity matrix to each group of 16, if we want to do that.
We know that each parity matrix can correct one bit error. So if we put all 64 databits into a single matrix, we can correct only 1 bit error in total. O.t.o.h., putting them into 4 matrices, we can correct up to 4 bit errors, if *the bit errors happen to all be in different matrices. (But of course this advantage comes at the cost of needing more parity bits.)*

1 pt **Q 3.12** Of course, the water quality doesn't change very quickly: in 99

A. Nothing, as the probabilities of each of the messages are the same in both cases.
B. It is higher than in the case of independent messages, as each message also informs us about the previous one.
C. It is lower than in the case of independent messages, as the previous message already informs us about the next one.
D. That depends on whether $0.99 \log_2(1/0.99)$ is lower or higher than the entropy of the messages themselves.

→ C

Continued on next page...

4. Physical layer

1 pt **Q 4.1** What's the role of total internal reflection in a glass fiber?

A. It causes a speed limit.
B. It ensures most light reaches the end of the fiber.
C. It ensures light from the outside doesn't get in.
D. It must be avoided by choosing the right index of reflection.
E. It must be avoided by choosing the right index of refraction.

→ B

1 pt **Q 4.2** Suppose we would make a glass fiber where the core is air, rather than glass as in a normal step-index fiber. Would this still work properly?

A. Only if the glass does not contain any impurities.
B. Only if there's no ambient light.
C. Yes, because the air's index of refraction is lower than the glass's.
D. Yes, because the air's index of refraction is higher than the glass's.
E. No, because the air's index of refraction is lower than the glass's.
F. No, because the air's index of refraction is higher than the glass's.

→ E

2 pt **Q 4.3** Which of the following factors contribute(s) to the limited range of radio signals?

A. Diffraction
B. Doppler shift
C. Inverse-square law
D. Curvature of the earth
E. Shadowing
F. None of the above

→ C,D,E

We were asking which factor limit how far radio signals can come.
Answer A, diffraction, is something that actually helps signals to get farther then they would without diffraction, so it doesn't contribute to limiting the range.
However, we recognize that the question could be read more generally as asking for factors that determine rather than reduce the range of the signals. Therefore, we've decided to simply ignore this answer in when giving points.
Answer B, Doppler shift, is a change of frequency when receiver and/or transmitter are moving, but it does not influence how far the signals can travel.
Answer C, Inverse-square law, says that received signal power decreases with the square of the distance: clearly, this makes the signal weaker at larger distance, and beyond some distance it's too weak to be received.
Answer D, Curvature of the earth: radio signals in principle travel in straight lines, so they can't go beyond the horizon. The curvature of the earth determines how far away that is.
Answer E, Shadowing, means that e.g. a building can block radio signals, just it blocks light giving a shadow. Clearly, this limits the range of the radio signals.

1 pt **Q 4.4** What is the purpose of 4B5B encoding?

A. Improving efficiency by squeezing 5 bits into 4.
B. Preventing clock recovery problems
C. Adding redundancy for error correction.
D. Adding a parity bit for error detection.
E. Nowadays only compatibility with old systems using it.

→ B

4B5B encoding ensures that it can't happen that there are lots of consecutive 0s or lots of consecutive 1s; having lots of consecutive 0s or 1s causes difficulty for the receiver to "recover the clock", i.e., to find out at which moment the next bit begins, and thus how many consecutive 0s/1s there are.

2 pt **Q 4.5** Suppose we use flags with bit stuffing as our framing method, with as flag the bit pattern 01111110. Suppose our frame to be sent is 10011111001111111111110010. Give the result of applying the flags and the bit stuffing. (Feel free to insert spaces in your answer for legibility; we'll only look at the 0s and 1s.)

→ 01111110 10011111 0 0011111 0 11111 0 10010 01111110

Spaces in the above answer are only for legibility, marking the flag at the beginning and end, and the stuffed 0 after every 5 consecutive 1s.

---

5. Medium Access Control

Boris, Jair and Donald set up their own network for playing Autocracy. Their network consists of a shared medium using a satellite. The satellite makes sure they can all hear each other and detect collisions, but with a (propagation) delay of 0.2 seconds (due to the distance to the satellite and back). Each packet's transmission duration is only 0.01 seconds.

2 pt **Q 5.1** Which of the following medium-access mechanisms would be the best choice in this situation?

A. pure CSMA
B. CSMA/CD
C. CSMA/CA
D. ALOHA

→ D

Q 5.2 Explain your choice:

→ CSMA variants can't work properly here, because the carrier detect would be far too slow. The propagation time is 0.2 s, very much longer than the transmission time of 0.01 s. So if say node A starts to transmit, it takes 0.2 s before node B can detect that. By that time, node A's transmission has already stopped 0.19 s ago, so it's totally useless for node B to use this information to decide about when it can safely transmit.
That leaves ALOHA as the best option.

1 pt **Q 5.3** Consider normal ALOHA and slotted ALOHA. Which of these require(s) time synchronization between the stations?

A. Neither.
B. Only normal ALOHA
C. Only Slotted ALOHA
D. Both normal and Slotted ALOHA

→ C

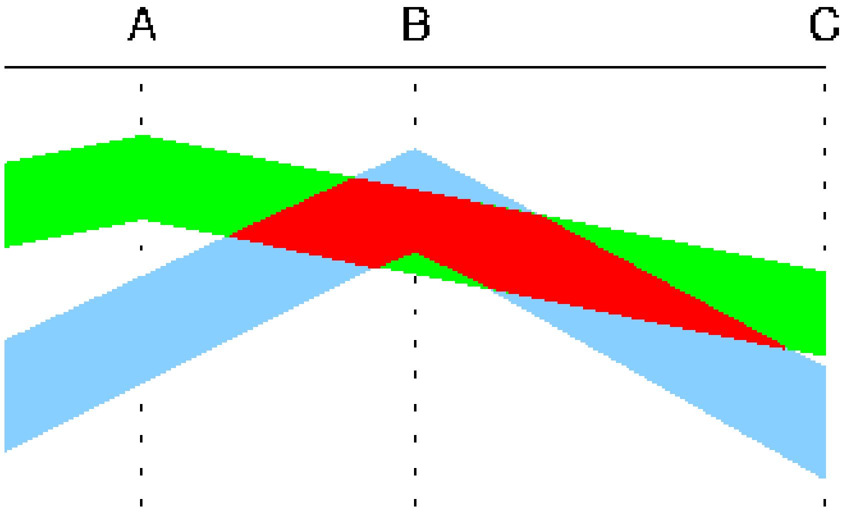
2 pt **Q 5.4** Compare TDMA and CSMA/CA; which of the following statements is/are true? (choose all that apply)

A. If all nodes need equally much bandwidth, TDMA is more efficient than CSMA/CA.
B. For bursty traffic, CSMA/CA is more efficient than TDMA.
C. TDMA needs collision detect.
D. CSMA/CA needs collision detect.
E. None of the above is true.

→ A,B

Quite many students also selected option D. This is not correct; CSMA/CA (Collision Avoidance) does not need collision detect. It is used in wireless systems, where collision detect is essentially impossible, because your own transmitter overpowers any colliding signal that you may receive from someone else.

Boris, Jair and Donald are disappointed by the performance of their satellite network. Defying the global pandemic, they decide to travel to one place, so they can use a shared cable network. The following picture shows a space-time diagram with three nodes on their shared cable, running the CSMA algorithm. What is happening here?



1 pt **Q 5.5**

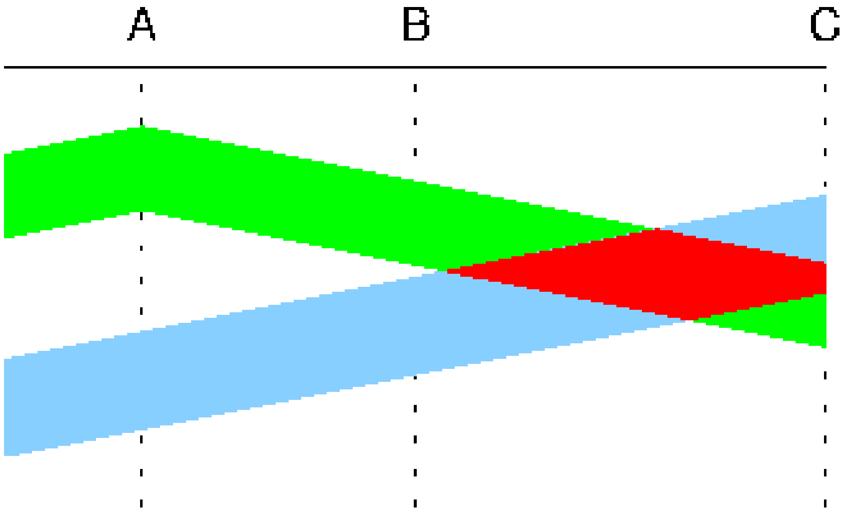
A. This is physically impossible.
B. This physically possible, but a CSMA node would never do this. (*)
C. This is correct behaviour, but there is a collision. (**)
D. This is correct behaviour, and no collisions occurs.

→ A

Recall that this diagram shows how a signal propagates along the cable, with distance on the horizontal axis and time on the vertical axis. The slopes of the "green" transmission from node A and of the "blue" transmission from node B are different. This would imply that the signal from node A needs less time (vertical axis in the diagram) than the signal from node B for propagating the same distance along the cable (horizontal axis in the diagram). That's impossible: all signals must propagate at the same speed (typically about 200 000 km/s in a copper cable).

Q 5.6 If you chose (*), indicate here which node(s) is/are not behaving correctly:
If you chose (**), indicate here which node(s) detect(s) the collision:
→ Not applicable

2 pt **Q 5.7** Same questions about the following picture:

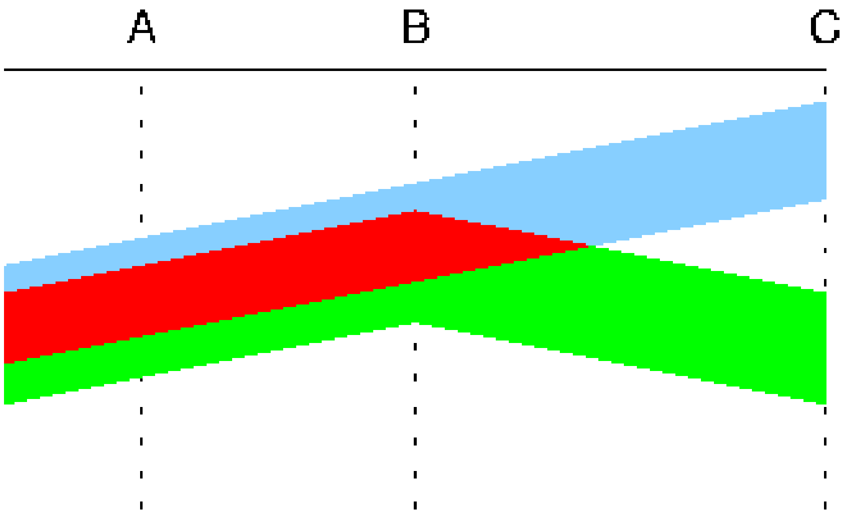


A. This is physically impossible.
B. This physically possible, but a CSMA node would never do this. (*)
C. This is correct behaviour, but there is a collision. (**)
D. This is correct behaviour, and no collisions occurs.

→ C

Q 5.8 If you chose (*), indicate here which node(s) is/are not behaving correctly:
If you chose (**), indicate here which node(s) detect(s) the collision:
→ collision only detected at node C

2 pt **Q 5.9** Same questions about the following picture:



A. This is physically impossible.
B. This physically possible, but a CSMA node would never do this. (*)
C. This is correct behaviour, but there is a collision. (**)
D. This is correct behaviour, and no collisions occurs.

→ B

Q 5.10 If you chose (*), indicate here which node(s) is/are not behaving correctly:
If you chose (**), indicate here which node(s) detect(s) the collision:
→ node B misbehaves

Node B starts it own transmission while it is already receiving the transmission from node C (blue in the figure). Using carrier sense, it shouldn't do that (since it will immediately cause a collision).

Grade calculation

The grade was calculated using the following formula:

$$\text{grade} = \frac{\text{points} - 4.6}{47 - 4.6} \times 9 + 1$$

47 is the maximum number of points for this test.
4.6 is the "guessing factor": it's the number of points one would get on average from giving totally random answers at the multiple-choice questions.