ETH zürich

Institut für Technische Informatik und Kommunikationsnetze

Department ITET
August 2022

Prof. Dr. Laurent Vanbever
C. Busse-Grawitz, T. Bühler, A. Dietmüller, T. Schneider

Exam: Communication Networks

27 August 2022, 09:30–12:00, Room HIL G 41

General remarks:

- ▷ Write your **name** and your **ETH student number** below on this front page and **sign it**.
- ▷ Put your **legitimation card** on the top right corner of your desk. Make sure that the side containing your name and **student number** is visible.
- ▷ Check that you have received **all task sheets** (Pages **1 − 35**).
- ▷ Do **not separate** the **task sheets** as we collect the exams **only after you left** the room.

- ▷ Write your answers directly on the task sheets.
- ▷ **All answers fit within the allocated space and often in much less.**
- ▷ If you need more space, use the three extra sheets at the **end of the exam**. Indicate the **task** in the corresponding field.

- ▷ **Read each task completely before you start solving it**.
- ▷ **For the best mark, it is not required to score all points.**

- ▷ Please answer either in **English or German**.
- ▷ **Write clearly** in blue or black ink (not red) using a **pen**, not a pencil.
- ▷ **Cancel** invalid parts of your solutions **clearly**.
- ▷ At the end of the exam, **place the exam face up on the top left corner** of your desk. Then collect all your belongings and **exit the room** according to the given instructions.

Special aids:

- ▷ All written materials (vocabulary books, lecture and lab scripts, exercises, etc.) are allowed.
- ▷ Using a calculator is allowed, but the use of electronic communication tools (mobile phone, computer, etc.) is strictly forbidden.

Family name:

Student legi nr.:

First name:

Signature:

Do not write in the table below (used by correctors only):

Task	Points
Ethernet & IP	/31
Intra-domain routing	/25
Inter-domain routing	/39
Reliable transport	/36
Applications	/19
Total	/150

leave blank

leave blank

Task 1: Ethernet & IP 31 Points

a) Warm-Up (6 Points)

(i) Name one scenario where a single packet may have two different IP headers. (1 Point)

____

(ii) Explain why a router (a layer 3 device) also needs to be able to parse and modify MAC addresses (layer 2 header fields) while forwarding IP packets. (1 Point)

____

____

(iii) Can you use the `ping` command to ping a layer-2 switch, a layer-3 router, or both? (1 Point)

____

(iv) Which protocol do layer-2 networks commonly use to avoid forwarding loops? (1 Point)

____

(v) What are the network address, the broadcast address, and the usable range of host IPs of the IP address `56.32.122.3/18`? (1 Point)

Network address: ______________________________

Broadcast address: ______________________________

Host address range: ______________ – ______________

(vi) What does best-effort delivery mean? (1 Point)

____

____

b) From the Bottom Up **(9 Points)**

leave blank

Consider the ETH and SWITCH networks in Figure 1, connected via two routers. All circles represent layer-2 switches.

The DHCP server connected to the HG switch in the ETH network assigns unused IPs in the private ETH prefix `10.132.0.0/16` and also returns the default gateway: `10.132.0.1`.

The ETH router uses Network Address Translation (NAT) to translate IPs in the private ETH prefix to the public IP (`1.2.3.4`), and randomly assigns a port in the range `1000`–`2000`.

Throughout this question, you may abbreviate MAC addresses, e.g. write `11` **instead of** `11:11:11:11:11:11`**.**

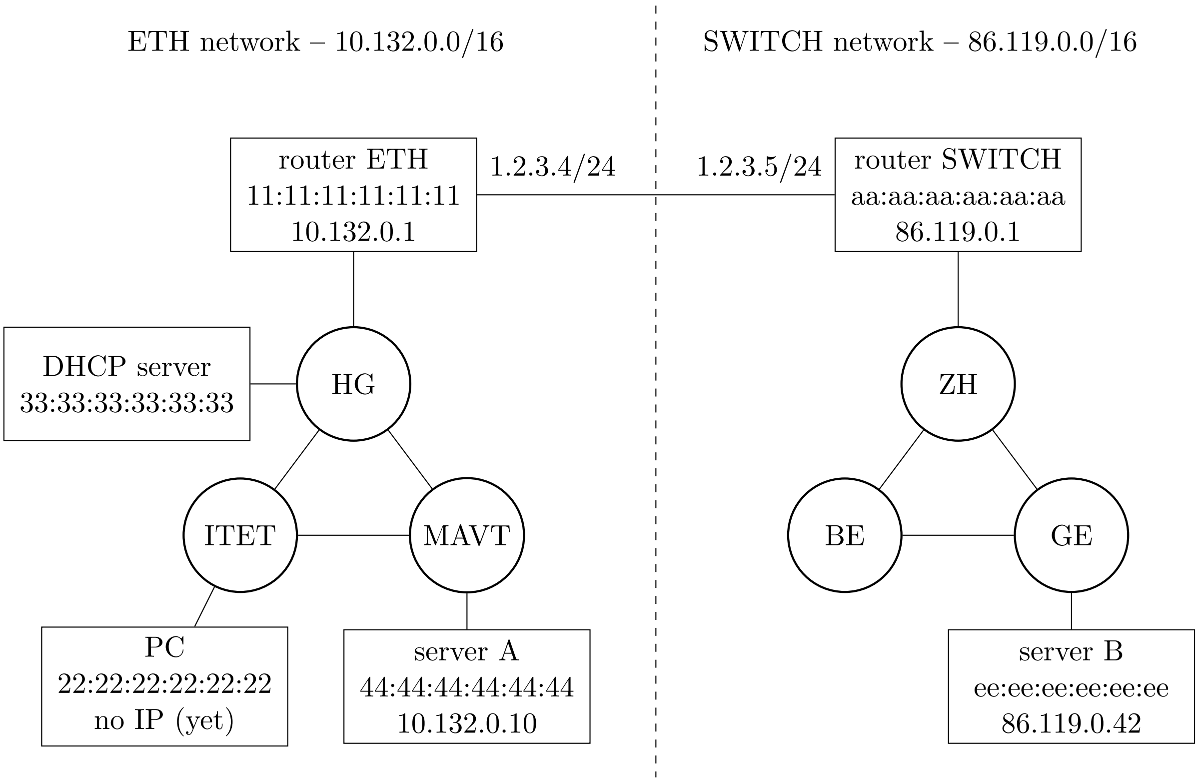


Figure 1: The ETH and SWITCH networks.

(i) You enable DHCP on your PC that you just connected in the ITET building. Answer the following questions: (2 Points)

What does your PC's first DHCP request look like?

Src MAC	
Dst MAC	
Request (summarize)	

Table 1: DHCP request.

Which DHCP response will you receive (include all relevant information as precisely as possible)?

Src MAC	
Dst MAC	
Response	

Table 2: DHCP response.

(ii) Next, you want to send a packet to server A (`10.132.0.10`) and server B (`86.119.0.42`) using their (known) IP addresses. For each of these servers, your host first needs to send an ARP request. Fill in the tables for the two requests and explain the differences between them: (3 Points)

Src MAC	
Dst MAC	
Request (summarize)	

Table 3: ARP request for server A

Src MAC	
Dst MAC	
Request (summarize)	

Table 4: ARP request for server B

Difference between the two requests: ______________________

____

____

(iii) Finally, your host sends a packet to server B. The table below shows all the devices on the path the packet takes. Complete the table by adding the missing MAC and IP addresses as well as the ports (the initial ports are already given to you). For each table row, write down the header values **just after the packet exits the indicated device**. Your PC still knows the MAC addresses from the previous question. (4 Points)

	MAC		**IP**		**Port**	
Device	**src**	**dst**	**src**	**dst**	**src**	**dst**
ITET					5678	4321
HG					5678	4321
Router ETH						
Router SWITCH						
ZH						
GE						

Table 5: For each step, fill in the missing IP, MAC, and port values.

c) Spanning Tree Protocol **(8 Points)**

leave blank

Consider the layer-2 network with 8 switches (circles) in Figure 2. Each link in the network has a capacity of 1 Gbps. Attached to the switches are three senders and three receivers:

- Sender 1 (S_1) tries to send 1 Gbps of traffic to Receiver 1 (R_1).
- S_2 tries to send 1 Gbps of traffic to R_2.
- S_3 tries to send 1 Gbps of traffic to R_3.

The switches in the network run the Spanning Tree Protocol (STP). If there exist multiple shortest paths to the root node, a switch picks the next hop which has the lower switch ID. Your goal is to maximize the aggregated throughput between the three senders and receivers.

(i) Apply the STP to the network in Figure 2. Cross out all links which the STP disables. What is the aggregated throughput between the three senders and receivers? (3 Points)

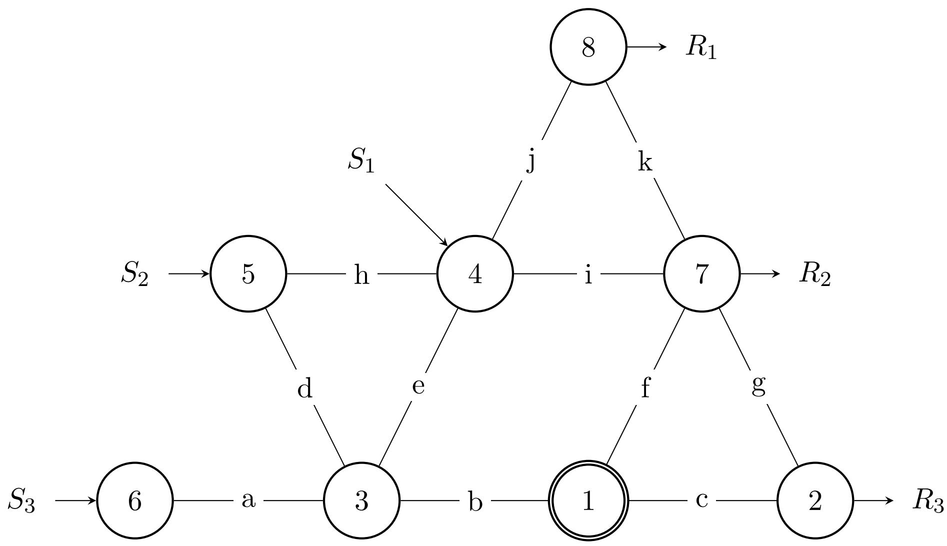


Figure 2: Cross out disabled links in the final spanning tree.

Aggregated throughput: ________________

(ii) You can now exchange IDs of two switches to increase throughput, however, the root of the spanning tree should not change (i.e., the switch with ID 1 cannot move). Which two IDs do you exchange and what is the new aggregated throughput? (2 Points)

Exchange switch ID: ________________ with switch ID: ________________

Aggregated throughput: ________________

(iii) You are still not happy with the throughput. You can now also physically remove *two* links ($a, b, \ldots, k$) from the network in Figure 2. Which two links do you remove and what is the aggregated throughput once the spanning tree has adapted? Continue with the spanning tree from task **(ii)**. (3 Points)

1st removed link: ________________ 2nd removed link: ________________

Aggregated throughput: ________________

d) Forwarding Tables **(8 Points)**

leave blank

You receive the forwarding table in Table 6. Each /16 prefix has a next hop from 1 to 4. Leverage longest-prefix matching to reduce the forwarding table to as few entries as possible. All existing addresses in the table should still be forwarded to the same next hop. Additionally, you can add a single default route (`0.0.0.0/0`). Pick the next hop of the default route in such a way that you achieve the highest forwarding table compression. Write your solution directly into Table 7.

Prefix	Next hop
142.112.0.0/16	2
142.113.0.0/16	2
142.114.0.0/16	2
142.115.0.0/16	2
142.116.0.0/16	1
142.117.0.0/16	2
142.118.0.0/16	3
142.119.0.0/16	3
142.120.0.0/16	4
142.121.0.0/16	4
142.122.0.0/16	3
142.123.0.0/16	4
142.124.0.0/16	2
142.125.0.0/16	1
142.126.0.0/16	4
142.127.0.0/16	2

Table 6: Original IP forwarding table.

Prefix	Next hop
0.0.0.0/0	

Table 7: Optimized forwarding table with a default route. Note, you might not need all the rows.

leave blank

leave blank

Task 2: Intra-domain routing **25 Points**

a) Warm-Up **(6 Points)**

(i) In one sentence, explain which problem "poisoned reverse" solves in a distance-vector protocol. (1 Point)

(ii) How would you set the link weights in a network such that the computed shortest paths have the highest bandwidths? (1 Point)

(iii) Explain why a global forwarding state is not valid if it only prevents dead ends. (1 Point)

(iv) The Round-Trip Time (RTT) of traffic forwarded over a shortest path computed by Dijkstra is higher than the RTT of traffic which takes a different (non-shortest) path towards the same destination. How is that possible? (1 Point)

(v) Explain one possible problem if you apply the Dijkstra algorithm to a network which contains links with negative weights. (1 Point)

(vi) Explain why routers normally load balance traffic between two shortest paths in such a way that all packets belonging to one flow follow the same path. (1 Point)

b) Reverse Distance Vector **(10 Points)**

leave blank

Table 8 shows the evolution of shortest paths from node A towards all other nodes in the network (Figure 3) while running a distance vector algorithm. It takes one time step (i.e., one row in the table) for a distance vector message to be sent from one node to another on a link. A node can send distance vector messages to all direct neighbors at the same time.

step	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
0	0	∞	∞	∞	∞	∞	∞	∞	∞
1	0	20	3	∞	∞	∞	∞	∞	∞
2	0	**18**	3	25	26	13	∞	∞	∞
3	0	18	3	**23**	**24**	13	27	16	∞
4	0	18	3	23	24	13	**24**	16	36
5	0	18	3	23	24	13	24	16	36

Table 8: Node A's evolution of shortest paths towards all other nodes while running a distance vector algorithm. ∞ indicates that A does not yet know a shortest path.

(i) Given the information in Table 8, use Figure 3 and fill in all identified links with their corresponding link weights. If the given information could lead to multiple solutions how the nodes are connected, only indicate one of them. (7 Points)

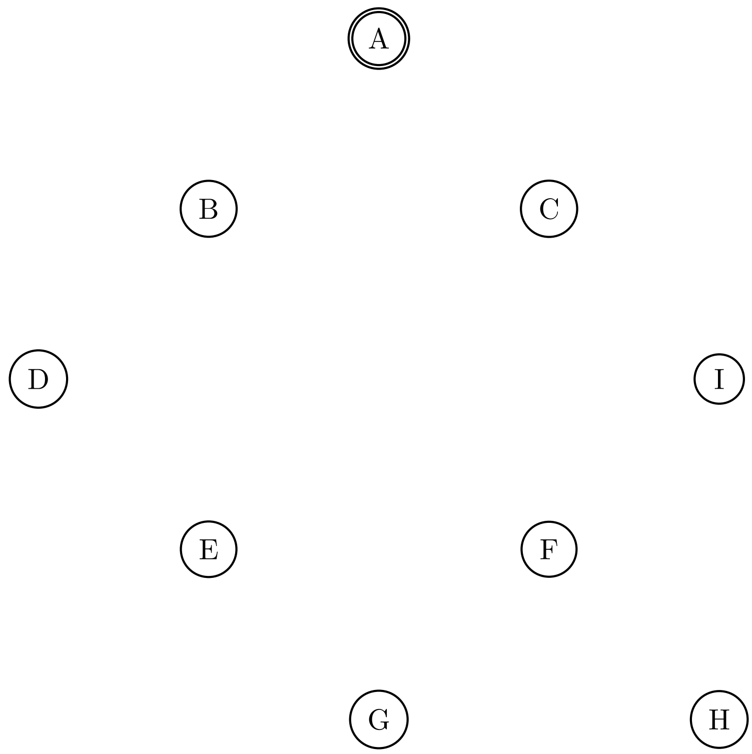


Figure 3: Use this figure to fill in all the detected links together with the corresponding weights according to the output of the distance vector algorithm in Table 8

(ii) As the previous question mentions, there might be multiple possible solutions given the information in Table 8. Assume now that the nodes instead run the Dijkstra algorithm to find all shortest paths and you receive the corresponding output for node A. Could the Dijkstra output of node A still lead to multiple solutions or would you only end up with one possible link and link weight assignment? Explain your answer.
Note: you do *not* have to figure out how the Dijkstra output or the link assignment would look like for the given network. (3 Points)

c) Label-based Forwarding **(9 Points)**

leave blank

This question compares destination-based forwarding introduced in the lecture with so-called "label-based" forwarding.

Figure 4 shows a simple example. The network contains four routers (A - D) and three hosts (H_1 - H_3) which are the destinations we want to reach.

The four tables at the **top** of the figure represent normal destination-based forwarding. Each router contains one forwarding entry for each destination ("dst" column) and knows how to forward packets such that they eventually reach the destination ("out" column).

The four tables at the **bottom** show label-based forwarding. The two edge routers (A and D), which are connected to the hosts, have a full forwarding table. For each packet, they add labels ("added label" column) to an extra header field. The two core routers (B and C) only forward based on these labels ("label" column).

dst	**out**
H_1	B
H_2	H_2
H_3	B

dst	**out**
H_1	C
H_2	A
H_3	C

dst	**out**
H_1	D
H_2	B
H_3	D

dst	**out**
H_1	H_1
H_2	C
H_3	H_3

dst	**added label**	**out**
H_1	L_1	B
H_2	–	H_2
H_3	L_1	B

label	**out**
L_1	C
L_2	A

label	**out**
L_1	D
L_2	B

dst	**added label**	**out**
H_1	–	H_1
H_2	L_2	C
H_3	–	H_3

Figure 4: Dst-based (top) and label-based (bottom) forwarding.

(i) Find *two* advantages and *one* disadvantage of label-based forwarding compared to the destination-based forwarding we saw in the lecture. (3 Points)

1st advantage: ______________________________

____

2nd advantage: ______________________________

____

Disadvantage: ______________________________

____

(ii) We want to forward traffic for D destinations in an arbitrary network with N routers out of which M are edge routers such as router A in Figure 4. How many different labels do you need *at most* in order to achieve any forwarding state which is possible with the destination-based forwarding we saw in the lecture? Explain your answer. (2 Points)

Number of labels: ____________________

Explanation: ____________________

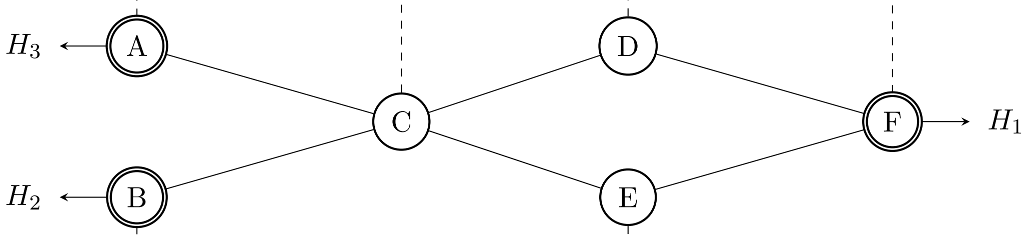
(iii) Figure 5 contains a network with three edge routers (A, B, F) which are connected to one host each (H_1-H_3). Can you find labels which result in a forwarding behavior which is *not* achievable with destination-based forwarding? If you think such a forwarding behavior is possible, show one example by filling in corresponding entries in the "added label"/"label" (L_1, L_2, ...) and "out" (A, B, ...) columns. Make sure that packets entering an edge router are correctly forwarded to the three hosts.
If you think it is impossible, explain your reasons below. (4 Points)

dst	**added label**	**out**
H_1		
H_2		
H_3	–	H_3

label	**out**

label	**out**

dst	**added label**	**out**
H_1	–	H_1
H_2		
H_3		



dst	**added label**	**out**
H_1		
H_2	–	H_2
H_3		

label	**out**

Figure 5: Show a forwarding behavior which is *not* possible with destination-based forwarding.

Explanation if impossible: ____________________

leave blank

leave blank

Task 3: Inter-domain routing **39 Points**

a) Warm-Up (8 Points)

You are an operator of an AS and you have decided to peer with two new networks, namely AS 195 and AS 466, at router `ZURI`. Before establishing the eBGP sessions, you receive the BGP routes that both of them would advertise to you. `ZURI` performs regular BGP path selection as presented in the lecture, breaking ties according to the **smaller next hop IP address**. Further, assume that `ZURI` does not know any additional routing information.

For each subtask, you are given a table with three routes, one from each AS (195 and 466), and the currently known route. Clearly indicate the selected route for the given destination **directly in the table**.

(i) Which route will be picked for the destination `5.21.2.35`? (1 Point)

	from	prefix	next hop	local pref.	MED	AS path	IGP Cost
☐	Current	`5.21.2.0/24`	`12.16.9.1`	50	200	260 590	4600
☐	AS 195	`5.21.2.0/24`	`20.79.3.2`	100	50	195 439 590	0
☐	AS 466	`5.21.2.0/24`	`13.8.47.25`	100	120	466 338 10 590	0

(ii) Which route will be picked for the destination `15.8.2.250`? (1 Point)

	from	prefix	next hop	local pref.	MED	AS path	IGP Cost
☐	Current	`15.8.2.0/24`	`12.16.9.2`	100	130	195 13 13 290	1250
☐	AS 195	`15.8.2.0/24`	`20.79.3.2`	100	140	195 13 13 290	0
☐	AS 466	`15.8.2.0/24`	`13.8.47.25`	100	80	466 20 6 13 290	0

(iii) Which route will be picked for the destination `68.7.5.6`? (1 Point)

	from	prefix	next hop	local pref.	MED	AS path	IGP Cost
☐	Current	`68.7.5.0/24`	`12.16.9.1`	100	120	42 91 20 590	4600
☐	AS 195	`68.7.5.0/24`	`20.79.3.2`	100	150	195 13 20 590	0
☐	AS 466	`68.7.5.0/24`	`13.8.47.25`	100	180	466 92 20 590	0

(iv) Which route will be picked for the destination `2.7.8.22`? (1 Point)

	from	prefix	next hop	local pref.	MED	AS path	IGP Cost
☐	Current	`2.7.8.0/23`	`56.22.219.29`	150	50	30 89 59 20	0
☐	AS 195	`2.7.8.0/24`	`20.79.3.2`	100	100	195 338 89 59 20	0
☐	AS 466	`2.7.8.0/23`	`13.8.47.25`	100	80	466 439 20	0

Consider the following table. There exists no single, most preferred route for `9.19.2.0/20`.

from	prefix	next hop	local pref.	MED	AS path	IGP Cost
Current	`9.19.2.0/20`	`12.16.9.1`	100	130	466 20 120	4600
AS 195	`9.19.2.0/20`	`20.79.3.2`	100	120	195 338 120	0
AS 466	`9.19.2.0/20`	`13.8.47.25`	100	150	466 20 120	0

(v) Compare the three routes for `9.19.2.0/20` individually with each other. Assume that only two routes are present at the same time. For each pair, indicate which route is preferred, and write down the deciding attribute. (2 Points)

Current vs. AS 195: ______________________________

Current vs. AS 466: ______________________________

AS 195 vs. AS 466: ______________________________

(vi) Why is there no single, most preferred route for `9.19.2.0/24`, and what can you do as a network operator to resolve this issue? (2 Points)

____

____

____

b) Traffic Engineering and Load Balancing **(8 Points)**

leave blank

You are a network operator of AS 1, as depicted in Figure 6.

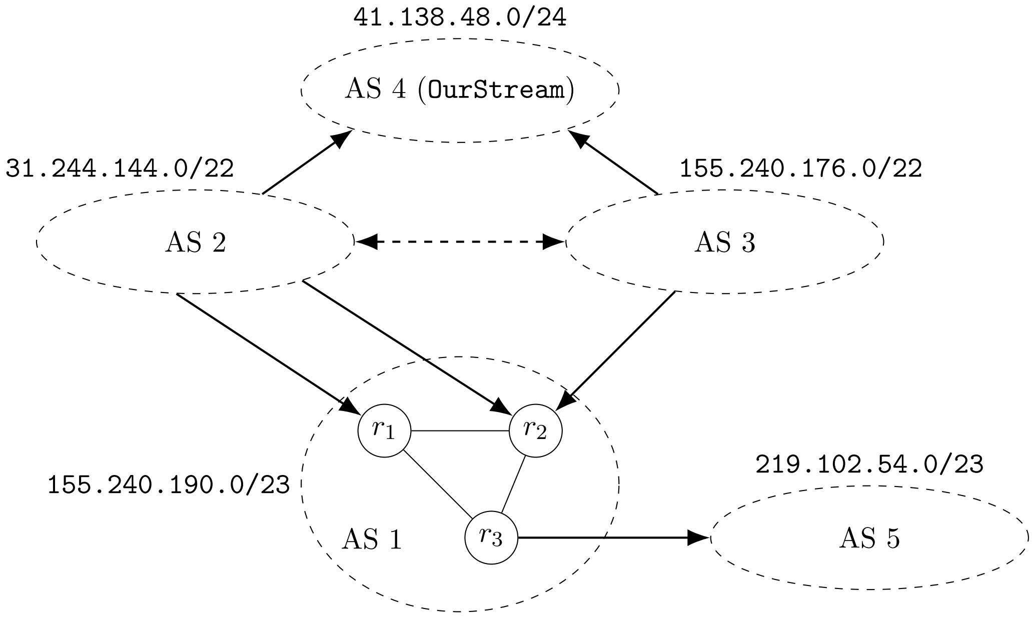


Figure 6: Network Topology

Single-headed plain arrows point from providers to their customers (AS 2 is the provider of AS 1), while double-headed dashed arrows connect peers (AS 2 and AS 3 are peers). Each AS applies the default selection and exportation BGP policies based on their customers, peers and providers.

(i) Consider the current topology. Which packets are routed over the direct peering session between AS 2 and AS 3? Give the prefixes of those packets. (2 Points)

Explain why those prefixes will be routed directly, and why all others will not.

(ii) Will router r_3 of AS 1 perform inter-domain load balancing for traffic towards destination `41.138.48.0/22` (AS 4)? What about r_2 of AS 1? Justify your answer. (2 Points)

In the following, you will perform inter-domain load balancing for traffic from the very popular streaming platform `OurStream` hosted in AS 4 to your customer AS 5. You notice that traffic from `OurStream` **traverses AS 3** since AS 4 sets a *higher local preference* to routes received from AS 3 than AS 2. Consequently, the link from AS 3 to r_2 is congested.

(iii) You want to load balance the traffic from `OurStream` to AS 5 between AS 2 and AS 3. How do you change the routing announcements to load balance traffic regardless of the local preference configured by AS 4? Your solution should still provide reachability, even if any link from AS 1 to either AS 2 or AS 3 fails. (2 Points)

(iv) Your load balancing works, and you now see traffic coming from both AS 2 and AS 3. However, AS 2 chooses to send the traffic towards r_2, and now, the link between r_2 and r_3 is congested. You now want to make AS 2 to send you traffic for AS 5 towards r_1. Mention two techniques (different from task **(iii)**) for influencing AS 2 to forward traffic with destination `219.120.54.0/23` towards r_1 instead of r_2. Both should not reduce the robustness against link failures. (2 Points)

c) BGP Hijack **(11 Points)**

leave blank

Consider the small Internet topology depicted in Figure 7. A user, located in AS 4, communicates with a server in AS 5. You are owning and maintaining AS 6. AS 4 is advertising the prefix `244.41.174.0/23` towards AS 2, and AS 5 announces `78.35.58.0/23` to AS 3. We assume that no AS validates any BGP announcement (e.g., using RPKI).

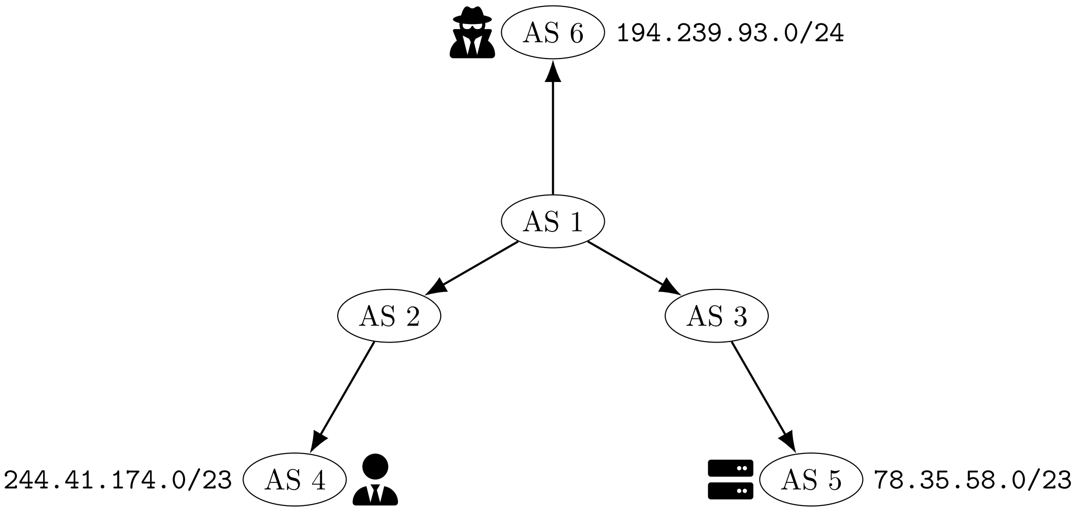


Figure 7: Network Topology

Single-headed plain arrows point from providers to their customers (AS 1 is the provider of AS 2). Unless stated otherwise, each AS applies the default selection and exportation BGP policies based on their customers, peers and providers.

In the following questions we consider different attacks that you can perform from AS 6. Your first goal is to **impersonate the server** (AS 5). You wish to attract traffic from AS 4 towards AS 5 and answer back to AS 4, without AS 5 ever receiving any traffic from AS 4.

(i) What BGP route do you advertise to AS 1, such that you can impersonate the server? We assume that no AS is manipulating the prefixes they receive. However, your attack should work no matter how AS 1, AS 2, and AS 3 are configured. (2 Points)

(ii) Could AS 5 detect the BGP hijack? Justify your answer. (1 Point)

You now wish to **eavesdrop** on the traffic **from the user to the server** (only in this direction). This means that you must attract traffic from AS 4 and forward it to AS 5.

(iii) What is the problem with the current topology? Why is it impossible to eavesdrop on traffic from the user to the server as an attacker located at AS 6? (2 Points)

(iv) To make it possible to eavesdrop on traffic from the user to the server, you decide to add a single new link (with an eBGP session) either to AS 2 or to AS 3. (6 Points)

With which AS would you establish a session, and which business relationship do you choose (i.e., are you a customer, a provider, or a peer)?

Explain which destinations you would advertise to AS 1 and over the new session, such that you can eavesdrop on traffic from the user to the server.

Does your attack work even if AS 1, AS 2, or AS 3 prepend their AS number multiple times to the AS path? Explain why or why not.

Explain how AS 6 handles incoming packets to perform the eavesdrop attack.

d) Inferring AS Relationships **(12 Points)**

leave blank

Consider the following Internet with 11 different ASes in Figure 8. AS 15, AS 17, and AS 18 are Tier-1 ISPs without any provider, whereas all other ASes have at least one provider. Some BGP sessions between ASes are already shown in the figure, but not all of them.

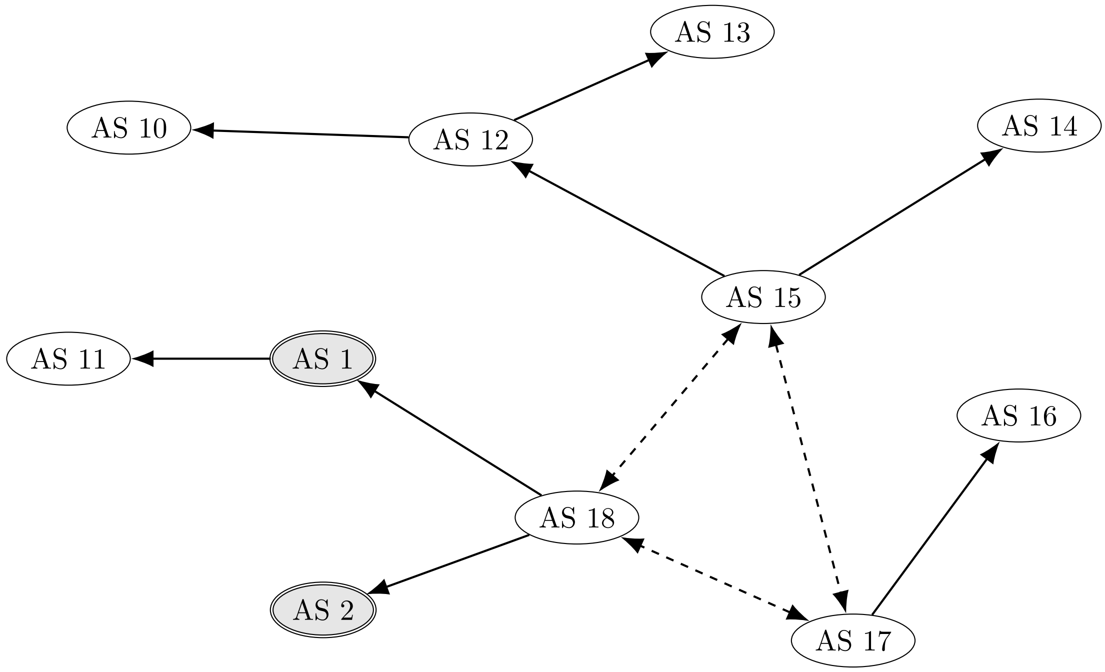


Figure 8: Network topology with a **subset** of existing BGP sessions.

Single-headed plain arrows point from providers to their customers, while double-headed dashed arrows connect peers. Each AS applies the default selection and exportation BGP policies based on their customers, peers and providers. ASes break ties by preferring the neighbor with the **lowest AS number**.

Each AS `XX` advertises its prefix `XX.0.0.0/24` to all of its neighbors. For instance, AS 12 advertises `12.0.0.0/24`. The Tier-1 ISPs (AS 15, 17, and 18) do not advertise any prefix. You are given a complete and sorted list of **all** incoming BGP messages of AS 1 and AS 2.

AS 1			
#	kind	prefix	AS path
1	U	`2.0.0.0/24`	1 18 2
2	U	`10.0.0.0/24`	1 11 10
3	U	`10.0.0.0/24`	1 12 10
4	W	`10.0.0.0/24`	1 11 10
5	U	`11.0.0.0/24`	1 11
6	U	`12.0.0.0/24`	1 12
7	U	`13.0.0.0/24`	1 12 13
8	U	`13.0.0.0/24`	1 18 15 12 13
9	U	`14.0.0.0/24`	1 18 15 14
10	U	`16.0.0.0/24`	1 18 17 16

AS 2			
#	kind	prefix	AS path
11	U	`1.0.0.0/24`	2 18 1
12	U	`10.0.0.0/24`	2 18 1 11 10
13	U	`10.0.0.0/24`	2 18 15 12 10
14	U	`11.0.0.0/24`	2 18 1 11
15	U	`12.0.0.0/24`	2 18 15 12
16	U	`13.0.0.0/24`	2 18 15 12 13
17	U	`14.0.0.0/24`	2 18 15 14
18	U	`16.0.0.0/24`	2 18 17 16

Figure 9: Stream of BGP messages received by AS 1 and AS 2. “U” abbreviates a BGP Update message, and “W” abbreviates a BGP Withdraw message.

(i) Consider only the BGP messages of AS 1 (the first table of Figure 9). What might have caused message number 4? (1 Point)

Recall that only some BGP sessions are given, but not all of them. As an example, consider message 6 from Figure 9. AS 12 advertises its own prefix directly towards AS 1 with AS path $[1, 12]$. This indicates that there must exist a BGP session between AS 1 and AS 12.

(ii) Which kind of BGP session is possible between **AS 15** and **AS 16**? For each kind of business relationship, justify your answer. (3 Points)

Is a peer to peer session possible? ☐ Yes ☐ No

Justify:

Is a customer (AS 15) to provider (AS 16) session possible? ☐ Yes ☐ No

Justify:

Is provider (AS 15) to customer (AS 16) session possible? ☐ Yes ☐ No

Justify:

(iii) Which kind of BGP session is possible between **AS 13** and **AS 14**? For each kind of business relationship, justify your answer. (4 Points)

Is a peer to peer session possible? ☐ Yes ☐ No

Justify: ______________________________

Is a customer (AS 13) to provider (AS 14) session possible? ☐ Yes ☐ No

Justify: ______________________________

Is a provider (AS 13) to customer (AS 14) session possible? ☐ Yes ☐ No

Justify: ______________________________

(iv) Which kind of BGP session is possible between **AS 1** and **AS 12**? For each kind of business relationship, justify your answer. (4 Points)

Is a peer to peer session possible? ☐ Yes ☐ No

Justify: ______________________________

Is a customer (AS 1) to provider (AS 12) session possible? ☐ Yes ☐ No

Justify: ______________________________

Is a provider (AS 1) to customer (AS 12) session possible? ☐ Yes ☐ No

Justify: ______________________________

leave blank

leave blank

Task 4: Reliable transport **36 Points**

a) Warm-Up **(7 Points)**

(i) Explain what *the congestion collapse* refers to and why it almost happened. (1 Point)

What it is: ______

______

Why it almost happened: ______

______

(ii) Why is the TCP sequence number no longer based on a timestamp? (1 Point)

Answer: ______

______

(iii) Explain with an example how cumulative ACKs can cause unnecessary retransmission. (1 Point)

Answer: ______

______

______

(iv) Compute the max-min fair bandwidth allocation for the flows A, B, C, D in Figure 10. The bandwidth on each edge is shared for both directions. (4 Points)

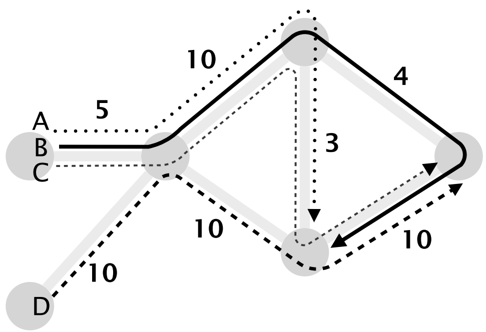


Figure 10: Graph with four concurrent flows. The edges are labeled with their bandwidth.

A: ______

B: ______

C: ______

D: ______

b) GBN protocol **(13 Points)**

leave blank

In this task, you complete partial time-sequence diagrams of the Go-Back-N (GBN) protocol. *Attention: Each subtask presents a new situation, they do not relate to each other.*

Here is a **non-exhaustive** list of implementation choices made for the GBN sender and receiver. **Read them carefully.**

- The sender and receiver window have a size of 4 packets;
- The receiver's out-of-order buffer stays empty in this task and is therefore not shown;
- The receiver uses cumulative ACKs which acknowledge all previous segments and point to the next expected data segment;
- The sender uses Fast Retransmit after three duplicate ACKs. For instance, if the sender gets ACKs [A9, A10, A10, A10], it will immediately retransmit the data segment D10;
- For each tick in the diagrams below, the sender can send one data segment and the receiver can send one ACK. Sender and receiver will first analyze the incoming packet and then send a data segment/ACK;
- The sender uses a retransmission timer (timeout) of 6 ticks. Each time it sends a data segment or receives an ACK, the timer is reset. After a timeout, the sender retransmits all current segments in its sender buffer (in order, one segment per tick);
- A data segment needs *two* ticks, and an ACK *one* tick to travel to the other end of the connection. See the given start in the diagrams.

Important notes for filling in the solutions:

- For all provided boxes and brackets, mark it clearly if they are empty (e.g. with a dash "—"). This can happen frequently.
- Indicate timeouts with an arrow. Draw the arrow start where the timeout counter starts, and draw the arrow end where the timeout stops (if it fits into the diagram).
- You do not need to add ticks to the time arrows or draw arrows that point into empty space (except for timeouts where the end does not fit into the diagram).

(i) Fill in the given boxes and brackets in Figure 11 and continue the arrows. Here, the sender intends to send D1 – D5. It has already sent D1 – D3 once, and just sent D4 for the first time (depicted). (2 Points)

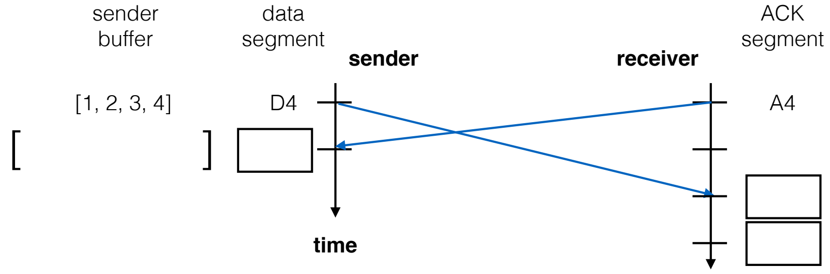


Figure 11

(ii) Fill in the given boxes and brackets in Figure 12 and continue the arrows. Here, the sender already sent D1 – D4. Currently, it fast retransmits D1 (depicted), but it is lost (marked with X). (4 Points)

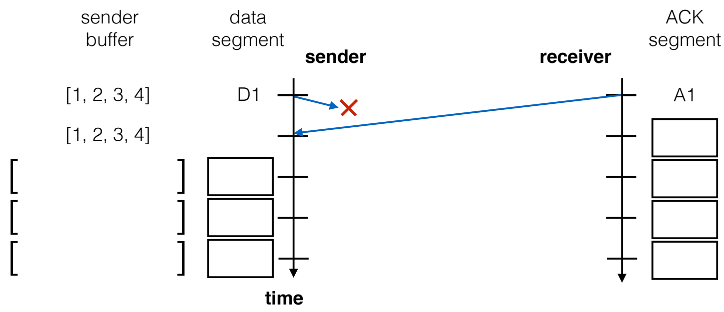


Figure 12

(iii) Only fill in the ACKs in Figure 13. (All arrows are already given.) (3 Points)

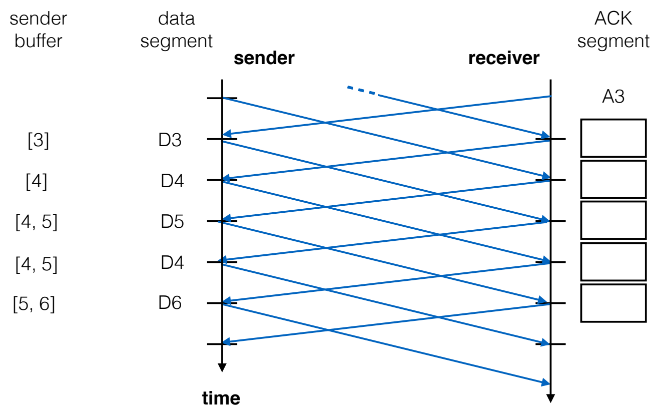


Figure 13

(iv) Assume that we additionally have a congestion window. It does not change with the regular TCP rules, however. Instead, it is simply given that it starts with size 2 and shifts to size 1 at the dotted line in Figure 14. Fill in the given boxes and brackets in Figure 14 and continue the arrows. The sender is in the middle of sending D1 – D20. (4 Points)

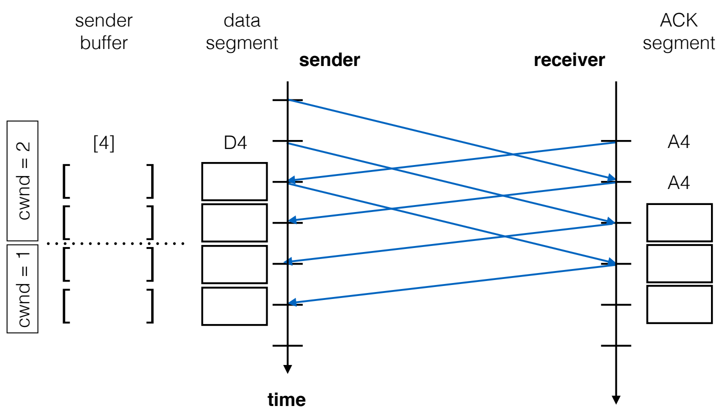


Figure 14

c) Congestion Window Evolution (12 Points)

Figure 15 depicts an evolution of the TCP congestion window. In this task, you will complete the further evolution of the congestion window and compute how the RTT estimate and timeout value evolve alongside.

Assume that:

- RTT = 1 s for all packets, except for the lost ones
- the entire window is sent *simultaneously*, meaning that the `cwnd` directly jumps from e.g. 2 (at $t = 2\,\text{s}$) to 4 kB (at $t = 3\,\text{s}$)
- all sent packets have a size of 1 kB
- RTT_{est} means the current RTT estimate, RTO means the current timeout value
- in this task, you should only consider the propagation delay, all other delay types are negligible
- the slow-start phase does not continue in the *next* iteration if in the *next* iteration, the congestion window is bigger than the slow-start threshold `ssthresh`.

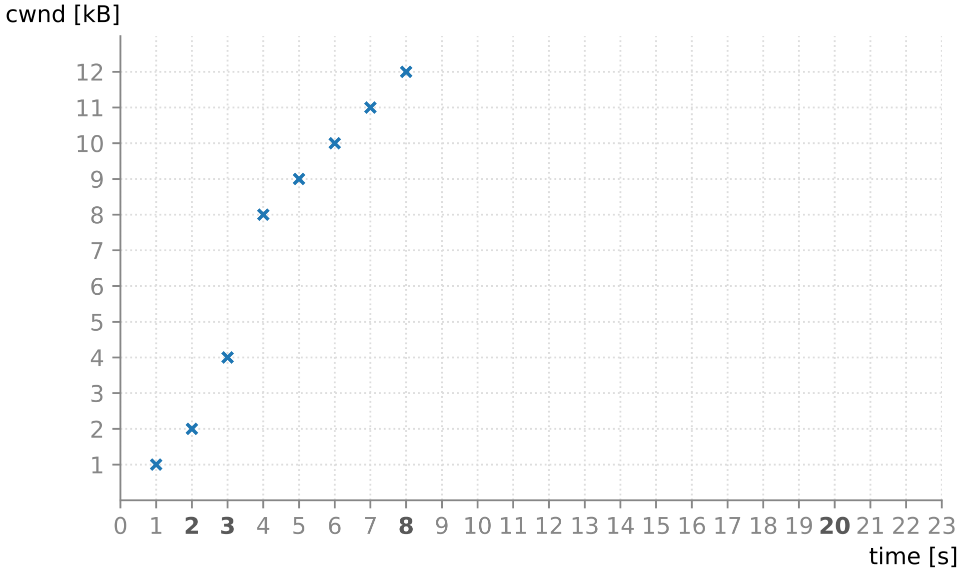


Figure 15: Congestion window evolution; y-axis: congestion window size, x-axis: time.

(i) Assume that at $t = 2\,\text{s}$ the RTT estimate is $\text{RTT}_{\text{est}} = 3\,\text{s}$. If $\alpha = 0.5$, what value does RTT_{est} have at $t = 3\,\text{s}$? Document your computation below. (3 Points)

Computation: ______________________________

____

____

(ii) The packets sent at $t = 8\,\text{s}$ are all lost, resulting in a timeout. Assume that $\text{RTT}_{\text{est}} = 1\,\text{s}$ at $t = 8\,\text{s}$. If that is the case, how long will the timeout be? Draw (in Figure 15) the correct next `cwnd` at the point where the timeout expires. Document your computation below. (2 Points)

Computation: ______________________________

____

(iii) After the first timeout expired, the timeout expires another time. Draw (in Figure 15) the next `cwnd` value at the point where the second timeout expires. Document your computation below. (1 Point)

Computation: ______________________________

____

(iv) As soon as the second timeout expires, the sender sends the packets in the congestion window. From now, all packets are sent successfully. Draw (in Figure 15) the resulting `cwnd` values up to and including $t = 20\,\text{s}$. *Assume that the second timeout does not affect the slow-start threshold `ssthresh`; also assume that there are no duplicate ACKs for this timespan.* Document your computation below. (3 Points)

Computation: ______________________________

____

(v) At $t = 20\,\text{s}$, the sender receives 3 duplicate ACKs. Explain why this can happen even though the data packets did arrive at the receiver. Furthermore, draw (in Figure 15) the resulting `cwnd` for the following two timesteps, assuming there are no more duplicate ACKs afterwards. (3 Points)

Explanation: ______________________________

____

d) AIMD and Fairness **(4 Points)**

leave blank

In this task, we call AIMD's coefficient for the additive increase $\alpha \in \mathbb{R}$, and the coefficient for the multiplicative decrease $\beta \in \mathbb{R}$. Concretely, we define AIMD as:

$$\texttt{cwnd}_{i+1} = \begin{cases} \texttt{cwnd}_i + \alpha & \text{no congestion signal} \\ \texttt{cwnd}_i \,/\, \beta & \text{congestion signal} \end{cases}$$

where $\texttt{cwnd}_{i+1}$ is the next iteration of $\texttt{cwnd}_i$, $i \in \mathbb{N}$. For this task, you can work without rounding, therefore $\texttt{cwnd}_i \in \mathbb{R} \;\; \forall\, i \in \mathbb{N}$.

Assume that two senders A and B run AIMD. Both use the same coefficients $\alpha, \beta \in \mathbb{R}$. What are the minimal conditions for α and β such that A and B converge to a fair bandwidth usage? If both oscillate around some values, a fair bandwidth usage means that they get the same amount of bandwidth *on average*.
Can it happen that A and B converge to a single point of bandwidth allocation (if so, under what circumstances?), or will they always "converge" to an oscillation? Explain.

Condition for α and β: ________________________________

Oscillation reasoning: ________________________________

leave blank

leave blank

Task 5: Applications **19 Points**

a) Warm-Up (4 Points)

(i) Briefly explain why we need $n + 1$ RTTs to retrieve n small objects using persistent connections across multiple HTTP requests. (1 Point)

Answer: ______________________________

____

(ii) Briefly explain why web caching can reduce the delay in rendering a web page. (1 Point)

Answer: ______________________________

____

(iii) Your friend has hired a service to serve DNS records for her domain. She adds an A record, looks it up, and does not share it with anyone. A few days later, she modifies the A record but realizes that DNS lookups are still going to the old address, even when she tries on your computer. She suspects this is due to DNS caching but you have *never* visited her website before. Briefly explain her how that could be. (2 Points)

Answer: ______________________________

____

____

____

b) Putting it All Together, Again **(7 Points)**

leave blank

You just arrived at ETH Zürich. You boot your laptop, open your browser, and type in the following URL:

`https://isnt.routing.fun`

Assume that:

1. Your laptop has already received a public IP address from ETH's DHCP server.
2. Your laptop has already issued all the ARP requests needed. (Its ARP table is full).
3. Your laptop uses ETH's DNS resolver (which it also learned via DHCP).
4. No DNS entries are cached. (Everywhere).
5. The website is working correctly and fits entirely within a single segment/packet.
6. No packet is lost, ever.

Describe each packet sent by your host, the ETH's DNS resolver, and the webserver, in the order in which they are sent.

A short description of the packets suffices, meaning you do not need to write down all the packet headers. That said, for each packet exchange, make sure to specify: the source, the destination, the transport protocol used, together with a description of the content of the packet (DNS request for X, TCP SYN, HTTP GET, etc.).

c) Load, Balanced **(8 Points)**

leave blank

The ETHZ networking team needs your help: the webserver hosting `www.ethz.ch` is frequently overloaded, leading to unhappy users.

They are currently considering acquiring three identical webservers for `www.ethz.ch` and would like to load balance the incoming user requests on them. The webservers would be hosted on `129.132.19.216`, `129.132.19.217`, and `129.132.19.218`.

In order to load-balance, the first technique they consider is a DNS-based round-robin, but they are not sure how to do it. This is where you come into action.

(i) Explain how DNS can be used to load-balance incoming requests. (1 Point)

(ii) What exact DNS resource records would you add to ETHZ's nameservers? Use the format (name, value, type, TTL). (2 Points)

(iii) Explain how the TTL affects the quality of the load-balancing. (1 Point)

Another solution would involve relying on a load-balancing device. Here, the DNS record for `www.ethz.ch` would point to the IP address of the load balancer (`129.132.19.215`) which would then randomly load balance the incoming traffic onto the replicas by rewriting the destination IP of the incoming IP packets to the one of the chosen replica.

(iv) One issue is that they do not know how to guarantee that the load balancer ends up sending the IP packets belonging to the same TCP connection to the same replica. Describe a load-balancing technique the load balancer could use that does *not* require any state. (That is, your solution should not require the load balancer to remember where each TCP connection is going.) (3 Points)

(v) Besides consistent load-balancing decisions, you realize that their solution is missing another key element to work properly. Briefly explain. (1 Point)

Extra Sheet 1

In case you need more space, use the following pages. Make sure to always indicate the task to which the answer belongs (e.g., *3 d) (ii)*).

Task: ______________

Task: ______________

Extra Sheet 2

Task: ______________

Task: ______________

Extra Sheet 3

Task: ________________

Task: ________________