SCIPER: _____

First name: _____ Family name: _____

EXAM
TCP/IP NETWORKING
Duration: 3 hours

January 2024

INSTRUCTIONS

1. Verify that you have 4 problems + one figure sheet.
2. Write your solution into this document and return it to us (you do not need to return the figure sheet). You may use additional sheets if needed. Do not forget to write your name on **each of the four problem sheets** and **all** additional sheets of your solution.
3. All problems have the same weight.
4. Briefly justify your answer. For grading, the justification is as important as the solution itself.
5. If you find that you need to make additional assumptions in order to solve some of the questions, please describe such assumptions explicitly.
6. Figures are on a separate sheet, for your convenience.
7. You can bring and use 4x A4 sheets = 8x AA4 pages of hand-written or type-written notes or the exam booklet that we offer in Moodle (in printed form). You can also use your pocket calculator (i.e. a simple calculator without extra storage or graph plotter).

PROBLEM 1

Consider the network for Problem 1 in the figure sheet. H1, H2, H3, and H4 are hosts. H4 runs a web server and H2 runs a DHCP server, as will be explained in the relevant questions. R1, R2, and R3 are routers. S1, S2, S3, S4 and S5 are switches. N1 is an IPv4 NAT. O1, O2, O3, O4, and O5 are observation points. All machines are dual-stack. All necessary IPv4 and IPv6 addresses are shown in the figure, as well as necessary MAC addresses (denoted with e.g., H1, H2, S1e, R2n, ...)

All links are full duplex Ethernet. We assume that all machines are correctly configured (unless otherwise specified), proxy ARP is not used and there is no VLAN. **The hosts, switches and routers have been running for some time, their different protocols have converged and the forwarding tables of all routers and switches are in their final state.** There is no other system or interface than those shown on the figure.

Question 1:

1. Write the two IPv6 address of H2 in uncompressed format.

2. Give the possible values of x, y, z (in the IPv4 addresses of H2, H1, and H3, respectively) and the possible network masks at H1, H2, H3.

Proposed values for x, y, z	valid	invalid
$x = 1, y = 1, z = 1$		
$x = 1, y = 2, z = 1$		
$x = 2, y = 2, z = 1$		
$x = 2, y = 1, z = 1$		
$x = 1, y = 1, z = 2$		
$x = 1, y = 2, z = 2$		
$x = 2, y = 2, z = 2$		
$x = 2, y = 1, z = 2$		

Proposed v4 subnet masks at H1, H2, and H3	valid	invalid
255.0.0.0		
255.255.0.0		
255.255.254.0		
255.255.255.0		

3. The IPv6 prefix lengths are /60 at H1, H2, and H3. Give the possible values of p, q (in the IPv6 addresses of H1 and H3).

Proposed values for p, q	valid	invalid
$p = 1, q = 1$		
$p = 2, q = 2$		
$p = 11, q = 1$		
$p = 1, q = 11$		
$p = 11, q = 11$		
$p = 13, q = 13$		

Question 2: H1 sends a UDP packet to H4's IPv6 address fe80::3. How many packets will be observed at O1, O3, and O5? Explain why.

Question 3: H1 sends a ping message to each of H2, H3, and H4 IPv4 addresses. Assume the TTL is equal to 64 in all IP packets generated by the hosts in the topology. What are the TTLs in the packets observed at O1, O3, O5?

Question 4:

1. H1 downloads a huge file from a web server `www.h4ipv4.com` running at H4 using HTTP through IPv4. H2 also downloads the same file using HTTP through IPv4. H1 and H2 happen to use the same local port number, 4567. What are the packets observed at O2 and O4? Give possible values in the table below. (You are allowed to use *x, y, z, p, q*)

Direction from H1 to H4							
	MAC addresses		IPv4 addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O2							
O4							

Direction from H2 to H4							
	MAC addresses		IPv4 addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O2							
O4							

2. H1 downloads a huge file from a web server `www.h4ipv6.com` running at H4 using TLS through IPv6. H2 also downloads the same file using TLS through IPv6. H1 and H2 happen to use the same local port number, 4567. What are the packets observed at O2 and O4? Give possible values in the table below.(You are allowed to use *x, y, z, p, q*)

Direction from H1 to H4							
	MAC addresses		IPv6 addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O2							
O4							

Direction from H2 to H4							
	MAC addresses		IPv6 addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O2							
O4							

Question 5: Suppose that suddenly H1 reboots and its caches become empty.

1. Suppose that H1 correctly configures its IPv4 address with the help of a DHCP server running at H2. After H1 is configured, it directly sends an HTTP request to `www.h4ipv4.com` using IPv4. You observe all packets resulting from this activity and up to receiving the first HTTP response from the web server, at observations points O1 and O2. Write the values of the fields (if present) in the table below. In each row, use as many lines as needed. The "type" field is the one contained in the MAC header (Ethertype).

At observation point O1						
MAC addresses		IPv4 addresses			port numbers	
srce	dest	srce	dest	type	srce	dest

At observation point O2						
MAC addresses		IPv4 addresses			port numbers	
srce	dest	srce	dest	type	srce	dest

2. Further suppose that H1 uses SLAAC in order to configure its IPv6 interface. List the messages that it needs to send and to which device in order to be able to send an HTTP request to `www.h4ipv6.com`. *[Hint: Just explain the steps that H1 needs to take, you do not need to write the values of any packet header.]*

Sender (hostname, NOT address)	Receiver (hostname, NOT address)	Type of message	Reason

SCIPER: _____

First name: _____ Family name: _____

PROBLEM 2

In the following questions, we assume that the BGP decision process uses the following criteria in decreasing order of priority.

1. Highest LOCAL-PREF.
2. Shortest AS-PATH.
3. E-BGP is preferred over I-BGP.
4. Shortest path to NEXT-HOP, according to IGP.
5. Lowest BGP identifier of sender of route is preferred; the comparison is lexicographic, with A<B<C<D and $1 < 2$; for example A1 is preferred over A2, A2 is preferred over B1, etc.

Furthermore, unless otherwise specified:

- When receiving an E-BGP announcement, every BGP routers tags it with LOCAL-PREF = 0. No other optional BGP attribute (such as MED, etc.) is used in BGP messages.
- No aggregation of route prefixes is performed by BGP.
- The policy in all ASs is that all available routes are accepted and propagated to neighbouring ASs, as long as the rules of BGP allow it.
- Every router redistributes internal OSPF destinations into BGP.
- Every router performs recursive forwarding-table lookup.
- No confederation or route reflector is used.
- Besides what is shown on each figure, there are no other stub networks.

Question 1: Consider the network for Problem 2, Question 1 in the figure sheet. There are four ASs (AS1 to AS4), with border routers (A1, A2, B1, C1, etc), and internal routers (IA1, IA2, etc). Only border routers use BGP. **Justify each answer**.

1. Consider the situation at t_1 after both BGP and OSPF have converged. List all BGP routes received by Router A_1.

At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
Justification:				

2. At time $t_2 > t_1$, Router D_2 crashes. List all the announcements received by router A_1.

At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
Justification:				

3. At time $t_3 > t_2 > t_1$ Router D_2 is still down, and AS2 decides to change its policy, forbidding all types of transit traffic. After this policy is in effect, list all routes received by Router A_1. *[Hint: look at the routes advertised by Router B_2].*

At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
Justification:				

Question 2: In the topology for Problem 2, Question 2 in the figure sheet, there are three ASs (AS1 to AS3) with routers A1, B1, C1, etc. **Justify each answer**.

1. Initially, only border routers use BGP with injection. Within each AS, RIP is used as the intra-domain protocol. What is the AS-level path taken by a packet sent from AS1 to the following destinations?
   (a) 3001:a:b::10
   (b) 2001:a:b::10

2. The network reboots and the following change is applied: now redistribution of BGP routes to RIP is used instead of injection.
   (a) How many new routes are added to Router C_4's table?
   (b) Reason about convergence time of routing protocols: in our setting, would you prefer injection or redistribution? Why?
   (c) Propose **two** different solutions that address the challenges you have encountered until this point.

3. The network reboots and the following change is applied: OSPF is used instead of RIP; all routers (including internal routers) use BGP with injection, but BGP routes learnt via E-BGP are also redistributed to OSPF.
   (a) How many routes to 1001:a:b::/48 does Router C_4 learn?
   (b) Which route is chosen? Is the route chosen optimal? Why?

Question 3: Consider the network for Problem 2, Question 3 in the figure sheet. There are seven ASs (AS1 to AS7) with routers A1, B1, C1, C2, C3, etc. In each domain, there is an I-BGP mesh. **For the purposes of this question, aggregation is used by BGP. Justify each answer.**

1. List all the BGP routes received by Router A_2.

At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
Justification:				

2. What is the AS-level path taken by a packet sent from AS7 with the following destinations?
   (a) 10.21.8.56
   (b) 10.44.12.36
   (c) 10.48.8.10

3. AS4 starts (maliciously) sending bogus announcements in addition to its normal announcements for prefix 10.112.0.0/12.
   (a) First, it announces 10.0.0.0/11. What is the route taken by a packet sent from AS7 to 10.8.8.56?
   (b) Then, AS4 stops announcing 10.0.0.0/11, and it starts announcing prefix 10.8.0.0/13. What is the AS-level path taken by a packet sent from AS7 to 10.8.8.56?

Question 4: Consider the network for Problem 2, Question 4 in the figure sheet. There are five ASs (AS1 to AS5) with routers A1, A2, A3, B1, C1, etc. In each domain, there is an I-BGP mesh. **For the purposes of this question, LOCAL-PREF is set-up by the different ASs as shown on the figure.**

1. Will BGP converge?
   (a) If yes, list all routes received by Router A_1.
   (b) If no, give **one** BGP route announcement that would make BGP converge.

At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
Justification:				

SCIPER: _____

First name: _____ Family name: _____

PROBLEM 3

PART A

Host A uses TCP to send a file of size 17 bytes to Host B, by using segments of size $MSS = 1$ byte each. A uses TCP Reno for congestion control and B's receive buffer is infinite (so that the offered window is always larger than A's congestion window). The first segment that A transmits has sequence number 1 and B acknowledges each segment that it receives (sends one ACK for each segment).

The transfer has been going on for some time, and at time T_0:

- the size of the congestion window of A, $cwnd_A$, is 8 bytes;
- the slow start threshold of A, $ssthresh_A$, is 10 bytes;
- all unacknowledged segments have been dropped because of a link error;
- A experiences a timeout for the segment with sequence number 10.

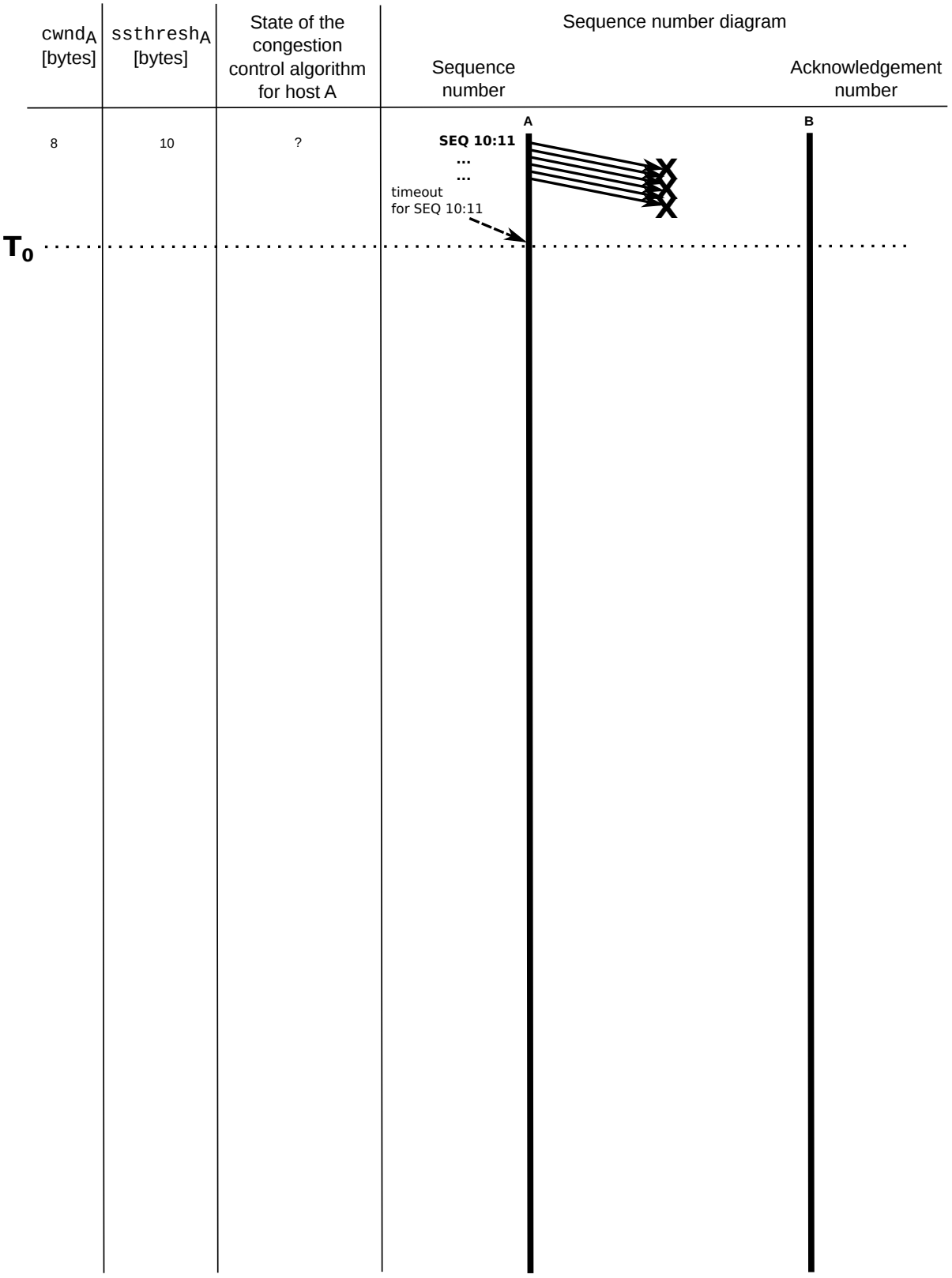
After time T_0 the network does **not** drop any more segments or acknowledgments (ACKs); no other timeout events or packet reordering events ever happen.

Question 1:

Describe the file transfer process between A and B from time T_0 until the file transfer completes by completing the provided sequence diagram that shows:

- all packets exchanged between A and B after T_0;
- the sequence numbers (SEQ) sent by A and the acknowledgment numbers (ACK) sent by B;
- the phase the congestion control algorithm is in: "slow start" (exponential increase) or "congestion avoidance" (linear increase);
- the size of the congestion window of A, $cwnd_A$;
- the slow start threshold of A, $ssthresh_A$.

[Hint: You may consult the FSM in the figure sheet.]


cwnd_A [bytes]
ssthresh_A [bytes]
State of the congestion control algorithm for host A
Sequence number diagram
Sequence number
Acknowledgement number
8
10
?
A
B
SEQ 10:11
...
...
timeout
for SEQ 10:11
T0

PART B

Consider the network for Problem 3 (Part B) on the figure sheet.

- R1, R2, R3, R4, R5, and R6 are routers, **not** switches, connected via point to point links.
- A, B, C, and D are unidirectional flows (as indicated by the arrows). There is no other system and no other flow than those shown on the figure.
- The capacity of each link is 1 Mb/s. The links are full duplex with the same rate in both directions. There is no other capacity constraint.
- We neglect the impact of the acknowledgement flows in the reverse direction.
- We also neglect all overheads and assume that the link capacities can be fully utilized at bottlenecks.

Question 1: Assume the rates are allocated by some central bandwidth manager according to max-min fairness. What are all the possible rate allocations? Justify.

Question 2: Assume the rates are allocated by some central bandwidth manager. Answer the following questions and justify:

1. What is the value of the maximum aggregate throughput that can be attained?

2. Find a Pareto-efficient allocation that maximizes the aggregate throughput of the flows.

Question 3: Assume the rates for the flows A, B, C, and D are allocated as $\frac{2}{3}$, $\frac{1}{3}$, $\frac{1}{3}$, and $\frac{1}{3}$ Mb/s respectively. Answer the following questions:

1. Is this allocation Pareto-efficient ? Justify.
2. Show that this allocation is *not* proportionally fair.

Question 4: In this question, flow A is turned off and all other flows use TCP Reno. The round trip times (RTTs) are 1000 ms for flow B and 10 ms for flows C and D; these RTTs include all processing times. All flows use the same MSS, the offered window is very large, and the application layer has always data to send (i.e. B, C and D are 3 long-running TCP flows).
Answer the following questions and **justify** your answers:

1. Assume that all routers use RED (Random Early Detection) queuing. What are the rates attained by each flow in the long run?

2. Further assume that not only routers use RED, but they also support ECN and the flows make use of it. Would the flows benefit from this change?

3. In the same scenario as above (sub-question *2.*), assume now that flows use TCP Cubic instead of TCP Reno, which flow(s) may experience a rate increase and why?

4. Now, suppose that routers use FIFO tail-drop queuing instead of RED, and therefore flows stop using ECN (because the routers cannot be ECN-enabled). How are the 3 flows affected by this?

Question 5: Consider the same settings as in Question 4.1 I.e., flows B, C, and D use TCP Reno and all routers use RED queuing. Suppose that the TCP flows have been running for a long time, and then flow A is turned on as a UDP flow.
Answer the following questions and justify your answers:

1. What is the maximum rate w that flow A can attain without affecting the rate of the other flows? *[Hint: If you have not calculated the flow rates above, in sub-question 4.1, then here, you can assume them to be known. I.e., the flow rates of B, C and D are r_b, r_c and r_d, respectively.]*
2. Assume that A starts sending at a higher rate v, where $w < v < 1$Mb/s, with a very dense packet sending pattern (i.e. UDP packets are send out with 0 inter-arrival time). How would this affect the rate of all other flows? Explain qualitatively how each flow rate will change.
3. Describe two mechanisms we can use, so that flow A cannot affect the rate.

SCIPER: _____

First name: _____ Family name: _____

PROBLEM 4

Question 1: Consider the network for Problem 4, Question 1 in the figure sheet. A1, B1, B2 and R1-R6 are routers that belong to the same AS and run OSPF with Equal Cost Multipath. The network is split into three OSPF areas. Area 0 is backbone, whereas areas 1 and 2 are not. Routers R3 and R4 are border routers between areas 0 and 1, whereas routers R1 and R2 are border routers between areas 0 and 2. All lines represent physical links and the numbers are their OSPF costs. Networks n1 and n2 are stub.

1. What is the best path, cost and next-hop from routers R1 and R2 to network n2? Explain how these two routers find this information. Specifically, what types of messages need to be exchanged in Area 2, and what is the algorithm used to compute the best path (no need to show the algorithm steps)?

At	Destination network	Cost	Next hop
R1	n2		
R2	n2		
Justification:			

2. What are the best paths from R3 and R4 to network n2? Which messages do R1 and R2 need to send, and how do R3 and R4 compute their best path?

At	Destination network	Cost	Next hop
R3	n2		
R4	n2		
Justification:			

3. Provide the routing table information at A1 with destination n2. Write down all the entries. How does A1 compute this information?

At A1 :		
Destination Network	Cost	Next hop
n2 ...		
Justification:		

4. If a host in n1 sends a large stream of packets to a host in n2, which path will the packets follow?

5. Now assume that the network is a software-defined network (SDN) and the routers are centrally managed by a controller (while they keep using OSPF). The network operator needs to ensure that all traffic *going from n1 to n2* passes through router R6 for security reasons. What exactly should they do at A1 and R3 to achieve this? Please state any assumptions you make.

6. Is there any other way that we could achieve the result of sub-question *1.5* without SDN?

Question 2: Consider the network for Problem 4, Question 2 in the figure sheet. Y1-Y_M, R1-R4 and X1-X_N are routers and n1-n_M are stub networks. S1-S_N are multicast sources.

1. Assume that all routers run PIM and only S1 streams traffic to multicast address m. Further assume there are hosts **only in networks n1, n2, n4 and n5** that are subscribed to group (S1,m). How many copies of the same message does S1 need to send to X1? Justify your answer.

2. For this question, S1 is still the only multicast source and the same hosts as sub-question *2.1* are subscribed to group (S1,m), but all routers run BIER. Specifically, Y1-Y_M are BIER egress routers (BFERs), X1-X_N are ingress routers, R1-R4 are BIER backbone routers, and there exists a centralized BIER Multicast flow overlay (not shown in the figure). Fill in the values of the forwarding bit masks in the BIER Index Forwarding table at R4. You can use either binary or set notation, and state your assumptions.

At R4:		
Destination BFER	Forwarding Bit Mask	Next-Hop

3. Consider the above BIER scenario of sub-question *2.2*, and suppose that S2 is a bogus streamer that wants to stream its own traffic to multicast group m (which is normally associated with source S1). Describe a mechanism that prevents BIER routers from forwarding S2's traffic towards the subscribers of (S1,m).

4. Now suppose that S2 is a legitimate multicast source. And assume that all multicast sources S1-S_N (with N very large) stream to various source-specific multicast groups (S_i,m_i). Also, assume that there are multiple hosts in all edge networks n1,...,n_M (with M large) that have subscribed to various such multicast groups. Which solution between PIM and BIER is the best choice for the routers R1-R4? Justify your answer.

Question 4: The company RomandeTech, located in the Canton de Vaud, needs to find a solution to connect their various departments in their building in Vevey (see Problem 4, Question 4 in the figure sheet). On the ground floor are the engineering department and security departments A and B, whereas on the first floor are the finance and sales departments. The company has purchased two Ethernet switches, S1 and S2, both with many ports, and wants to use one for each floor, as shown in the figure. Switch S1 is also connected to a router, which provides access to the public internet.

1. Assume that the finance, sales, and engineering departments need to be in three separate LANs, whereas the two security departments need to be connected to the same separate LAN. How can RomandeTech achieve this without purchasing new hardware?
2. Assume now that security B moves to the first floor and is connected to switch S2, instead of switch S1, but still needs to be in the same LAN as security A. Do we need to change anything from the previous solution?
3. Now assume that security B moves to the second building of RomandeTech, which is located in Lausanne. However, it still needs to be connected to the same LAN as security A. Is it possible to achieve this and with which mechanism?