EXAM
TCP/IP NETWORKING
Duration: 3 hours
With Solutions

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. Briefly justify your answer. For grading, the justification is as important as the solution itself.
4. If you find that you need to make additional assumptions in order to solve some of the questions, please describe such assumptions explicitly.
5. Figures are on a separate sheet, for your convenience.
6. 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 (25PTS)

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 (8 pts):

1. **(2pt)** Write the two IPv6 address of H2 in uncompressed format.

   Solution.

   (a) fe80:0000:0000:0000:0000:0000:0000:0002

   (b) 2001:000a:000a:0012:0000:0000:0000:0002

2. **(3pts)** 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	
$x = 1, y = 2, z = 1$	X	
$x = 2, y = 2, z = 1$		X
$x = 2, y = 1, z = 1$		X
$x = 1, y = 1, z = 2$	X	
$x = 1, y = 2, z = 2$	X	
$x = 2, y = 2, z = 2$		X
$x = 2, y = 1, z = 2$		X

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

3. **(3pts)** 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
$p = 1, q = 1$		X
$p = 2, q = 2$		X
$p = 11, q = 1$	X	
$p = 1, q = 11$		X
$p = 11, q = 11$		X
$p = 13, q = 13$		X

Question 2 (2 pts): 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.

At O1: 1 packet will be observed as it leaves H1. This will be the NS packet of the NDP protocol, with which H1 wants to learn the MAC address of fe80::3. If someone assumes that the link-local IPv6 address

of N1w or R1s is fe80::3, then there will be 2 additional packets (the NA and the UDP packet) that will be observed at O1.

At O3: 0 packets will be observed because link-local traffic is not routed (the packet is dropped at R1).

At O5: 0 packets will be observed because link-local traffic is not routed (the packet is dropped at N1).

Question 3 (3 pts): 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 TTL values in the packets observed at O1, O3, O5?

Only layer-3 devices, such as NATs and routers, decrement the TTL value (by 1 at each layer-3 hop). So:

At O1: The packet have not yet passed through any routers, so their TTLs are all 64.

At O3: The packet traverses R1. Thus, TTL = 63.

At O5: The packet traverses N1, R3, R2 and any other layer-3 devices inside the ISP. Thus, TTL $\leq$ 61.

Question 4 (6 pts):

1. **(3pts)** 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	H1	N1w	10.1.y.1	10.4.1.4	TCP	4567	80
O4	N1e	R3w	10.3.3.21	10.4.1.4	TCP	a number > 1024 (e.g. 20000)	80

Direction from H2 to H4							
	MAC addresses		IPv4 addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O2	H2	N1w	10.x.1.2	10.4.1.4	TCP	4567	80
O4	N1e	R3w	10.3.3.21	10.4.1.4	TCP	a number > 1024 (e.g. 20001)	80

2. **(3pts)** 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	H1	N1w	2001:a:a:p::1	2001:b:b:1::1	TCP	4567	443
O4	N1e	R3w	2001:a:a:p::1	2001:b:b:1::1	TCP	4567	443

Direction from H2 to H4							
	MAC addresses		IPv6 addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O2	H2	N1w	2001:a:a:12::2	2001:b:b:1::1	TCP	4567	443
O4	N1e	R3w	2001:a:a:12::2	2001:b:b:1::1	TCP	4567	443

Question 5 (6 pts): Suppose that suddenly H1 reboots and its caches become empty.

1. **(3pts)** 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).

   By assumption, we do not include the DHCP discovery/offer, as this has been done already. Instead, we focus only on H1's communication with the DNS, the gateway router, and the web server. I.e. possible ARP requests/replies, DNS query/reply, TCP SYN/SYNACK, ACK and HTTP request/response. There are multiple possible answers in this question, depending on where the DNS is located. In the following, we assume that the DNS is hosted at H2. Hence, there is no need for ARP resolution of the DNS (the MAC address of H2 is known to H1 after receiving the DHCP offer):

At observation point O1						
MAC addresses		IPv4 addresses			port numbers	
srce	dest	srce	dest	type	srce	dest
H1	H2	10.1.y.1	10.x.1.2	IP	45000	53
H2	H1	10.x.1.2	10.1.y.1	IP	53	45000
H1	ff:ff:ff:ff:ff:ff	-	-	ARP (for MAC of gateway)	-	-
N1w	H1	-	-	ARP	-	-
H1	N1w	10.1.y.1	10.4.1.4	IP (for TCP SYN)	45800	80
N1w	H1	10.4.1.4	10.1.y.1	IP (TCP SYNACK)	80	45800
H1	N1w	10.1.y.1	10.4.1.4	IP (TCP ACK, HTTP get request)	45800	80
N1w	H1	10.4.1.4	10.1.y.1	IP (HTTP response)	45800	80

   where 45000 and 45800 are just random numbers selected from [1025, 65536].

   The table for O2 has to be completed only if the link S1-S2 is activated by the Spanning Tree protocol, and in this case, it is the same as the table above.

2. **(3pts)** 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 (host)	Receiver (host)	Type of message	Reason
H1	multicast address derived algorithmically	Neighbor solicitation	Duplicate address test
H1	multicast address of all routers on-link	Router solicitation	Learn the network prefix from gateway router (N1) and obtain a globally valid IPv6 address; also learn gateway's IP address
H1	multicast address of all routers on-link	Request for a Router Advertisement	Learn the IP of the DNS that will serve H1
H1	(*) DNS server	DNS query	Learn the IP address of `www.h4ipv6.com`
H1	H4	TCP SYN	Open connection to `www.h4ipv6.com`
H1	H4	HTTP request	Get homepage of `www.h4ipv6.com`

Explanation (not needed for getting full points):

Step1. H1 creates a link local address via SLAAC and performs an address duplication test by sending a multicast packet to the solicited node multicast address with the same 24 last bits as its own IP address.

Step2. H1 sends a Router Solicitation message to ask any router that is on link for the common network prefix, so that it can automatically configure a globally valid IPv6 address.

Step3. H1 sends a request for a router advertisement, so that it learns the IP address of the DNS server. Alternatively, it uses the stateless DHCP server to get the same—the latter was also considered a valid answer.

Step 4. H1 contacts the DNS server with a query message for learning the IP address of `www.h4ipv6.com`.

(*) Depending on where the DNS is located, H1 may need to run NDP to learn the MAC address of the DNS server. E.g. if the DNS is at H2, H1 needs to resolve the MAC address, but if the DNS is at router N1 that advertized the network prefix, then NDP is not necessary. H1 will already know its MAC from the route advertisement. No points were subtracted for missing the NDP.

Step5. H1 sends out the HTTP request after opening a TCP connection to `www.h4ipv6.com`.

Before sending out the HTTP request, H1 does *not* need to run NDP in order to learn the MAC address of N1w (gateway router), because the N1 will have responded with a router advertisement earlier and its MAC address will be known to H1.

PROBLEM 2 (25PTS)

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 (6 pts): 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. **(2pts)** 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 ?
B1	2001:a:b::/48	B1	AS2	yes
D2	2001:a:b::/48	D2	AS4 AS2	no
A2	2001:a:b::/48	D1	AS4 AS2	no
D2	4001:a:b::/48	D2	AS4	yes
B1	4001:a:b::/48	B1	AS2 AS4	no
A2	4001:a:b::/48	D1	AS4	no
B1	3001:a:b::/48	B1	AS2 AS3	yes
D2	3001:a:b::/48	D2	AS4 AS3	no
A2	3001:a:b::/48	D1	AS4 AS3	no
Justification: Both 1001:a:b::/48 and 1002:a:b::/48 are not BGP routes, so they are not included. A1 will receive routes from its three BGP peers (B1 and D2 through E-BGP, and A2 through I-BGP). The next hop of the routes learnt from I-BGP has to be the first hop outside the AS. For 2001:a:b::/48 and 4001:a:b::/48, best routes are selected by distance.				

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

At At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
B1	4001:a:b::/48	-	-	withdrawal
B1	4001:a:b::/48	B1	AS2 AS3	update
Justification:				

Explanation:

(a) (Optional) To remove the old paths:

i. A1 losses connection to D1, removes path to 4001:a:b::/48 through D2.
ii. B2 losses connection to D2, removes path to 4001:a:b::/48 through D1, sends withdrawal to B1.
iii. B1 removes path to 4001:a:b::/48, sends withdrawal to A1.

(b) To keep connectivity:

i. B2 selects new best route to 4001:a:b::/48 through D3. B2 sends update to B1.
ii. B1 receives update, selects new route to 4001:a:b::/48 through B2 (I-BGP). B1 sends update to A1.

3. **(2pts)** 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].*
Note: If the above answer does not include the BGP withdraw message of the route to 4001:..., then this withdraw must be included in this answer.

At A_1 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
B1	2001:a:b::/48	B1	AS2	yes
A2	2001:a:b::/48	D1	AS4 AS2	no
A2	4001:a:b::/48	D1	AS4	yes
A2	3001:a:b::/48	D1	AS4 AS3	yes
Justification: AS2 only sends routes for its prefix (2001:a:b/48), the other 2 prefixes that need to be learned by AS1 all arrive through $D1$.				

Question 2 (7 pts): 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. **(1pt)** Initially, *only* border routers use BGP with injection and there are no other routers or stub networks. AS3 uses RIP for its 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

   (a) 3001:a:b::10: AS3 (AS1, the origin, is not included in the AS-level path).
   (b) 2001:a:b::10: unreachable. The packet is sent from AS1: as the destination is in AS2, it first goes to Router C1. At C1, due to injection there is a prefix with next hop the IP address of the external interface of B1, but therefore C1 applies recursive forwarding table lookup in order to forward the packet to the correct interface—which is its interface on the link C1-C4; so, the packet will be forwarded to C4. C4 receives a packet with destination address 2001:..., but has no entry in its forwarding table to forward it properly. The reason why C1 knows about how to reach 2001:..., whereas C4 does not, is that C1 runs iBGP and therefore learns a path to that prefix from C2, while C4 does not run iBGP and hence has no way to know about how to reach it (no redistribution from BGP to RIP causes this issue). Therefore, the packet is dropped.

2. **(1pt)** Some time later, AS1 and AS2 become popular providers, and each one signs 50 new customer ASs (not shown in the figure); each customer advertises 1000 stub networks to its provider AS1 or AS2, and AS3 learns about them only from AS1 and AS2. Given the large number of entries, AS3 now decides to redistribute the routes learnt via E-BGP into its IGP (i.e. RIP). How many new routes will be added to Router C_4's table?
   C_4 will receive all routes from AS1 and AS2, so 100,000 new routes ($2 \times 50 \times 1{,}000$).

3. **(2pts)** Further later in time, AS3 also increases in size and hundreds of internal routers are added to its intra-domain topology. Reason about the convergence time of the routing protocols used by AS3, in the unfortunate case that some of the routes that are announced by AS1 and/or AS2 to AS3 are unstable (i.e. some of the routers on the route often fail and need to be rebooted). In this setting, was redistribution to RIP a good choice of IGP? Why?

   No, it was not. Redistribution has the downside that all routes learnt by BGP are passed to all routers, so, in our case, every single router will have 100,000 new entries. With so many routes constantly changing, the time it will take for RIP to converge will be large enough that the network will be unable to forward packets for a considerable portion of time.
   It is important to note that the convergence time of BGP does not change significantly because AS3 is only connected to AS1 and AS2 and therefore accepts any announcement about their (100,000) stub networks as sent by routers A1 and B1.

4. **(1pt)** Propose **two** alternative solutions that address the challenges AS3 has encountered up to this point (both in sub-question *2.1* and *2.3*).
   Choose 2 out of the following:
   (a) Use BGP with MPLS (as explained in the relevant lecture slide).
   (b) Use BGP with Source routing (as explained in the relevant lecture slide).
   (c) Use convinient combinations: redistribution only with IGPs that scale (e.g. OSPF).
   (d) Injection + use BGP in all routers with route reflectors.

5. **(2pts)** Due to a change of strategy, AS3 network reboots and deploys the following change: it now uses OSPF instead of RIP with the cost of every physical link being equal to 1; all routers (including internal routers) use BGP with injection, and BGP routes learnt via E-BGP are also redistributed to OSPF with a cost equal to 100.
   (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?

   (a) C_4 learns two different routes to 1001:a:b::/48: *route1* comes from I-BGP (sent from C_1) and *route2* from OSPF.
   (b) This is a case of an injection conflict: two routes to the same destination are learnt from different protocols. In these cases, the conflict is solved by using the administrative distance: the route with smallest administrative distance is chosen.
      i. *route1*: from I-BGP, distance = 200.
      ii. *route2*: from OSPF, distance = 110.

   Then, the route learnt from OSPF is chosen. This route is **not** optimal as it is longer that its alternative.

Question 3 (7 pts): 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** that is not shown in the figure. **For the purposes of this question, aggregation is used by BGP. Justify each answer.**

1. **(2pts)** List all the BGP routes received by Router A_2.

At A_2 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
A1	10.48.0.0/12	B2	AS2	yes
C2	10.32.0.0/11	C2	AS3	yes
A3	10.0.0.0/11	G1	AS7 AS6	yes
C2	10.112.0.0/12	C2	AS3 AS4	yes
A1	10.112.0.0/12	B2	AS2 AS3 AS4	no
Justification: The prefixes 10.32.0.0/12 and 10.48.0.0/12 are aggregated at AS3, and arrive at $A2$ as one single entry. An entry for 10.48.0.0/12 still exists as it is received from $A1$ through I-BGP.				

2. **(3pts)** 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

   (a) 10.21.8.56: AS6
   (b) 10.44.12.36: AS1 AS3 AS5
   (c) 10.48.8.10: AS1 AS2

3. **(2pts)** 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?

   (a) The prefix 10.0.0.0/11 advertised by AS4 can not be aggregated by AS3: the entire prefix is added to $A2$'s table. Then, AS7 has two possible paths to 10.0.0.0/11: through AS6 and through AS1. As the former is shorter (in terms of the number of AS-level hops), packets sent from AS7 go to AS6.
   (b) The router $G2$ will then learn the following routes:

At A_2 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
G1	10.8.0.0/13	A3	AS1 AS3 AS4	yes
F2	10.0.0.0/11	F2	AS6	yes
Justification: The prefix 10.8.0.0/13 advertised by AS4 can not be aggregated by AS3: the entire prefix is added to $A2$'s table. Then, that route is propagated by $C2$ to $A2$ and then to $G1$.				

   $G2$ has two routes that match destination 10.8.8.56, and will prefer the one with the longest prefix match (through $G1$), although it is longer in terms of the number of AS-level hops.

Question 4 (5 pts): 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** that is not shown in the figure. **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:				

1. No, BGP will not converge because there are circular dependencies in the preferences: the routes will constantly oscillate. Alternatively, one can try to follow the route propagation process and check that the steps start to repeat.
2. To make BGP converge, we need to break the circular dependencies: for example, sending a withdrawal from AS1-AS4 removing any of the routes between them.

PROBLEM 3 (30PTS)

PART A (8 PTS)

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.

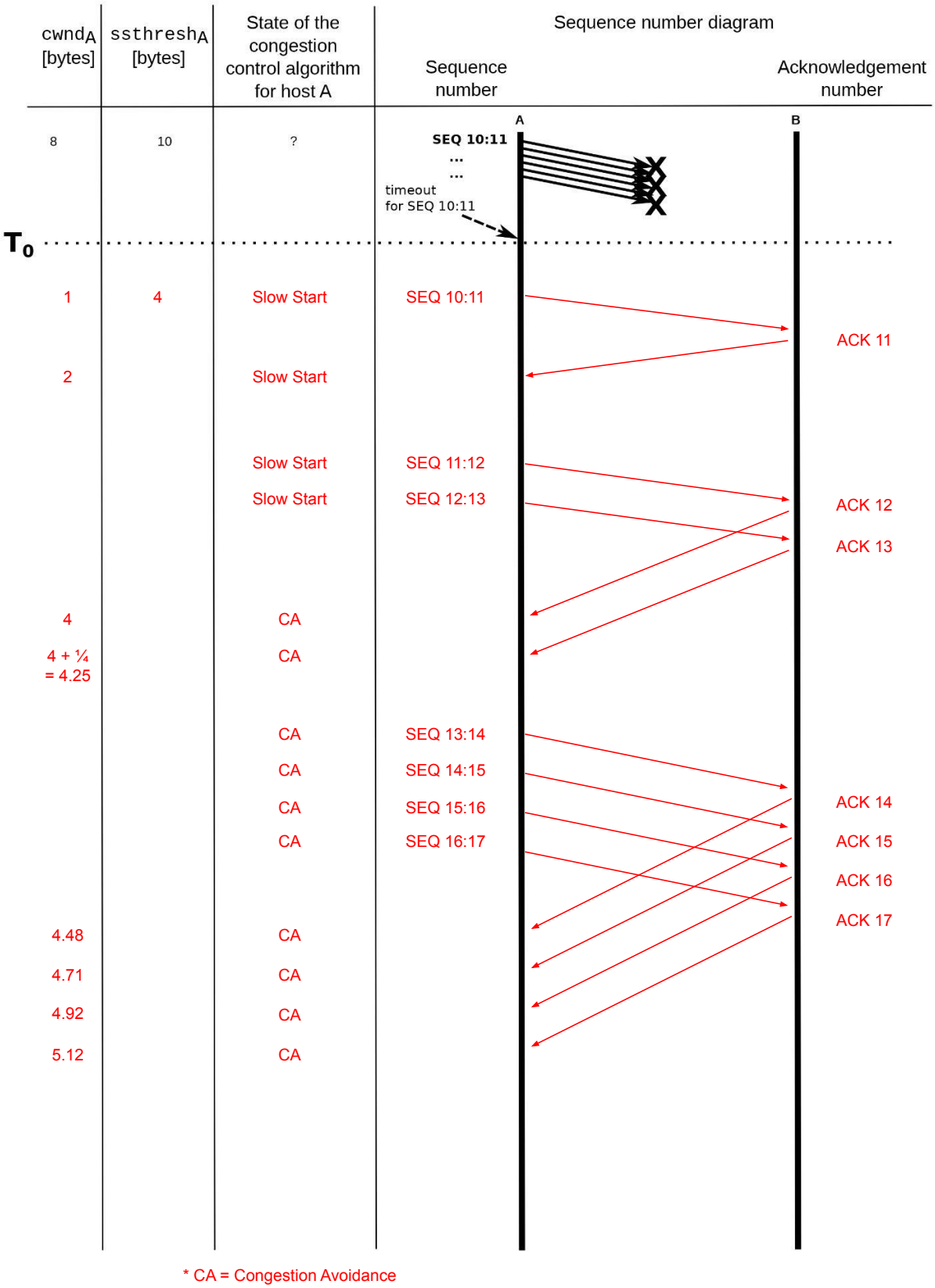
After time T_0 the network does **not** drop any more segments or acknowledgments (ACKs); no other timeout events or packet reordering events ever happen.

Question 1 (8pts):

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.]



* CA = Congestion Avoidance

PART B (22 PTS)

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

Using water-filling algorithm: $a + b + c + d = t = \frac{1}{3}$. $link_{R4-R5}$ gets saturated, so we freeze b, c, and d. Now, we can increase a till $\frac{2}{3}$. $link_{R2-R3}$ gets saturated and so we freeze a. Thus, $a = \frac{2}{3}$, $b = c = d = \frac{1}{3}$. This is a unique soliution

Question 2 (4pts): Assume the rates are allocated by some central bandwidth manager. Answer the following questions and justify:

1. **(1pt)** What is the value of the maximum aggregate throughput that can be attained?
   It is 2 Mb/s because of bottleneck links $link_{R2-R3}$ and $link_{R4-R5}$.
2. **(3pts)** Find a Pareto-efficient allocation that maximizes the aggregate throughput of the flows.
   $a + b + c + d \leq 2$ (maximum)
   From bottleneck links: $b + c + d \leq 1$, $a + b \leq 1$
   Max aggregate throughput can be achieved by turning off b. So, $c + d \leq 1$, $a = 1$
   Possible solution: $c = d = 0.5$, $a = 1$. It is Pareto-efficient since every flow has a saturated bottleneck link.

Question 3 (3 pts): 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. **(1pt)** Is this allocation Pareto-efficient ? Justify.
   Yes, because every flow has a bottleneck link (i.e., for every flow there exists a link, used by that flow, which is saturated). The bottleneck link for flows b, c and d is $link_{R4-R5}$, while the bottleneck for flow a is $link_{R2-R3}$. So, none of the flows can be increased without reducing some other.
2. **(2pts)** Show that this allocation is *not* proportionally fair.
   If we reduce b by $\delta > 0$, then a can increase by δ while either c or d can be increased by δ.
   Total relative change: $\sum_i \frac{x_i' - x_i}{x_i} = -\frac{\delta}{b} + \frac{\delta}{a} + \frac{\delta}{c} = -3\delta + \frac{3\delta}{2} + 3\delta = \frac{3\delta}{2} > 0$
   Thus, this allocation is not proportionally fair.

Question 4 (9 pts): 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. **(3pts)** Assume that all routers use RED (Random Early Detection) queuing. What are the rates attained by each flow in the long run?

TCP Reno loss throughput:

$$r_{reno} = \frac{MSS \times 1.22}{RTT \times \sqrt{q}}$$

Let, flow rates for flows B, C, and D be b, c, and d respectively.

For B:

$$b = \frac{MSS \times 1.22}{1000 \times \sqrt{q}} = \frac{k}{1000}$$

For C and D:

$$c = d = \frac{MSS \times 1.22}{10 \times \sqrt{q}} = \frac{k}{10}$$

Therefore, $c = d = 100 \times b$.

Now,

$$b + c + d \leq 1$$

$$b + 100b + 100b \leq 1$$

$$b = \frac{1}{201}$$

So,

$$c = \frac{100}{201} \quad , \quad d = \frac{100}{201}$$

2. **(2pts)** 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?
   There would be no packet drops but the flow rates would remain the same.

3. **(2pts)** 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?
   Flow b may experience rate increase because a TCP-Cubic connection gets more throughput than TCP-Reno when RTT is large.

4. **(2pts)** 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?
   The flows will experience higher latencies due to buffer bloat and irregular drop patterns.

Question 5 (4 pts): 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. **(2pts)** 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.]
   $r_a = 1 - r_b = \frac{200}{201}$. Rates for other flows remain the same.

2. **(1pt)** Assume that A starts sending at a higher rate v (with $w < v <$ 1Mb/s) and its packet sending pattern is very dense (i.e. UDP packets are send out with 0 inter-arrival time). How would this affect the rate of the other flows? Explain qualitatively how each flow rate will change (i.e. increase, decrease, stay the same).
   r_b will decrease, while r_c and r_d will increase.

3. **(1pt)** Describe a mechanism we can use, so that UDP-flow A cannot affect the rate of the other flows; i.e. attain at most a rate equal to w, even if its application writes data in the UDP socket at a higher rate.
   Possible mechanisms:

   - Per-class Queuing
   - UDP-based congestion control. E.g, QUIC
   - Application-based rate control. E.g, TFRC (TCP-Friendly Rate Control)

PROBLEM 4 (20PTS)

Question 1 (10 pts): 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. **(2pts)** 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	1	B1
R2	n2	4	B2
Justification: All the routers in Area 2 build their link state database by flooding their LSAs to the area. Then, each one computes the shortest path to n2 with Dijkstra's algorithm. The shortest paths from R1 and R2 can be found with observation.			

2. **(2pts)** 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	8	R5
R4	n2	5	R2
Justification: Routers R1 and R2 flood a summary LSA in Area 0 indicating their distances to n2. Routers R3 and R4 compute their distances to n2 using the Bellman-Ford formula. The computations are for R3: $d(R3, n2) = \min\{7+1, 8+4\} = 8$ via R5, and for R4: $d(R4, n2) = \min\{5+1, 1+4\} = 5$ via R2. The distances from R3 and R4 to R1 and R2 are computed again using Dijkstra's algorithm in Area 0.			

3. **(2pts)** 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	9	R3
n2	9	R4
Justification: The border routers R3 and R4 advertise a summary LSA to Area 1 with their distances towards n2. The two possible paths to n2 are via R3 with cost $d(A1, n2) = 1 + d(R3, n2) = 9$ and via R4 with cost $d(A1, n2) = 4 + d(R4, n2) = 9$.		

4. **(1pt)** If a host in n1 sends a large stream of packets to a host in n2, which path will the packets follow?
   There are two paths from A1 with equal costs, namely A1-R3-R5-R1-B1 and A1-R4-R2-B2-B1. Assuming A1 uses equal cost multipath, roughly half of the packets will follow one path and half will follow the other.
5. **(2pts)** 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.
   We need to setup a forwarding rule to the flow table of A1 so that packets with source IP addresses in n1 and destination IP addresses in n2 will be forwarded to R3 and another rule in the flow table of R3 so that the same packets will be forwarded to R6.
6. **(1pt)** Is there any other way that we could achieve the result of sub-question *1.5* without SDN?
   We can use source routing, either strict or loose. In strict source routing the source A1 puts the whole path into the extension header (i.e. R3-R6-R2-B2-B1), while in loose source routing A1 needs to indicate some intermediate router in the header by putting the path segment: R6-R2-B2-B1. Either type of source routing was considered correct.

Question 2 (6 pts): 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. **(1pt)** 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.
   S1 will only send one copy. The other routers will duplicate copies according to the number and the location of the subscribers.
2. **(3pts)** 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.

   The BFERs that can be reached via R2 are Y1 and Y2; the BFERs that can be reached via R3 are Y4 and Y5. If we are using set notation, the forwarding table and masks will look like this:

Destination BFER	Forwarding Bit Mask	Next-Hop
Y1	$\{1, 2\}$	R2
Y2	$\{1, 2\}$	R2
Y4	$\{4, 5\}$	R3
Y5	$\{4, 5\}$	R3

   If we are using a bit string of 5 bits (assuming that all routers $Y_N, N > 5$ are not active), the bit masks would look like this.

Destination BFER	Forwarding Bit Mask	Next-Hop
Y1	00011	R2
Y2	00011	R2
Y4	11000	R3
Y5	11000	R3

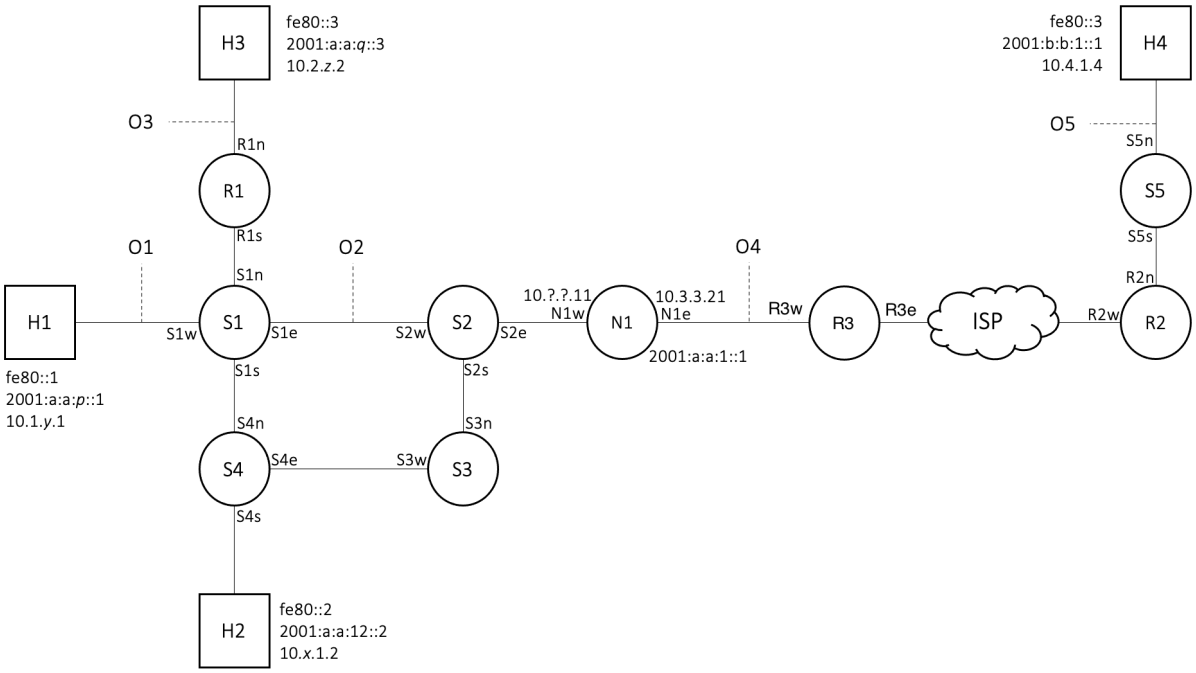
3. **(1pt)** Consider the above BIER scenario of sub-question *2.2*, and suppose that S2 is a bogus streamer that wants to stream its own traffic to multicast group m (which is normally associated with source S1). Describe a mechanism that prevents BIER routers from forwarding S2's traffic towards the subscribers of (S1,m).
   In BIER, the group membership information is stored in a centralized database, a.k.a. the *multicast flow overlay*. Therefore, the ingress BIER router X2 will communicate with the multicast overlay in order to create the BIER header (i.e. the set of BFERs that are connected to subscribers), but the multicast overlay will return an empty set, because there are no subscribers to multicast group (S2,m). So, X2 will simply drop the bogus traffic.
4. **(1pt)** 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. In such a scenario, with which architecture, PIM or BIER, routers R1-R4 will be "stressed" the least (i.e. they will need fewer resources)? Justify your answer.
   With PIM, backbone routers need to keep per-flow state information, whereas in BIER the state is kept in a centralized entity. If we have many flows, then BIER causes less stress to the backbone routers.

Question 3 (4 pts): 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 3 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.

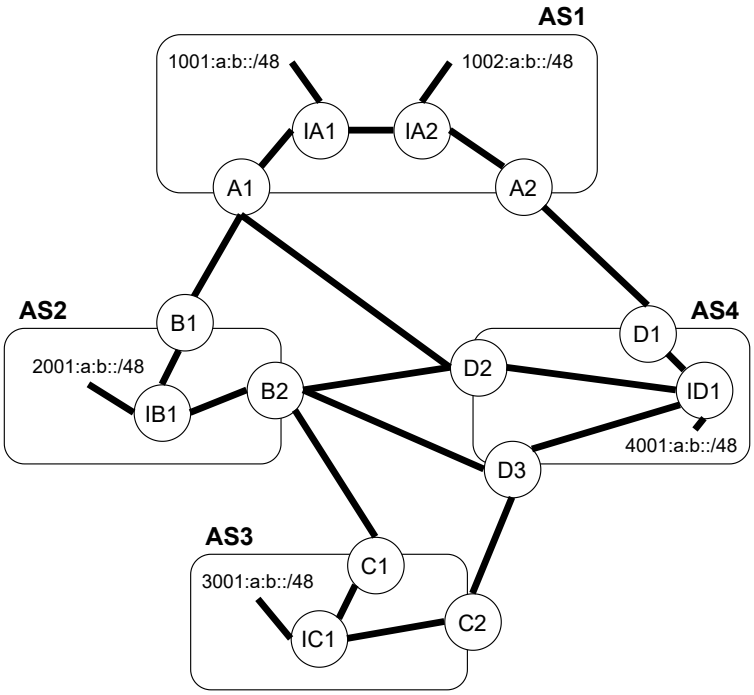
1. **(1pt)** Assume that the finance, sales, and engineering departments need to be in three separate LANs, whereas the two security departments need to be connected to the same separate LAN. How can RomandeTech achieve this without purchasing new hardware?
   We can configure the ports of switches 1 and 2 to construct four VLANs. Two in switch 1 for Engineering and one for Security A and B and two in switch 2 for Finance and Sales. The two switches have to be connected with a *trunk* link (i.e. an ethernet cable interconnecting their *trunk ports*.)
2. **(1pt)** 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?
   We can move all hosts belonging to Security B from switch 1 to switch 2 (assuming there are enough available ports). Then, we should reconfigure the ports of switch 2 so that they belong to the same VLAN as before.
3. **(2pts)** 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?
   We can create a VLAN tunnel ("MAC in IP") between the router in Vevey and the one in Lausanne. The mechanism is called LS bridging. Anything related to the "MAC-in-IP" idea was graded as correct despite the terminology that was used.

TCP IP EXAM - FIGURES

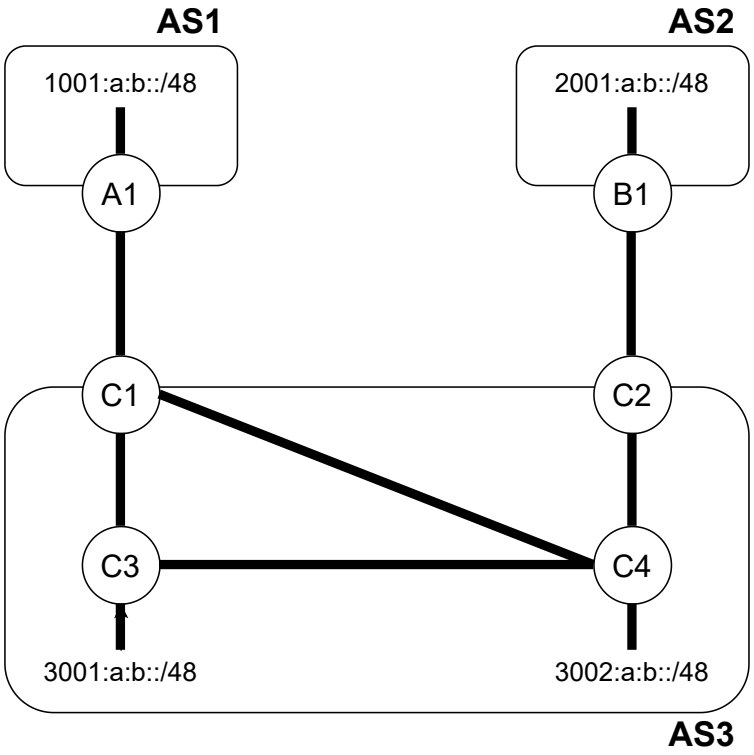
For your convenience, you can separate this sheet from the main document. Do not write your solution on this sheet, use only the main document. You do not need to return this sheet.



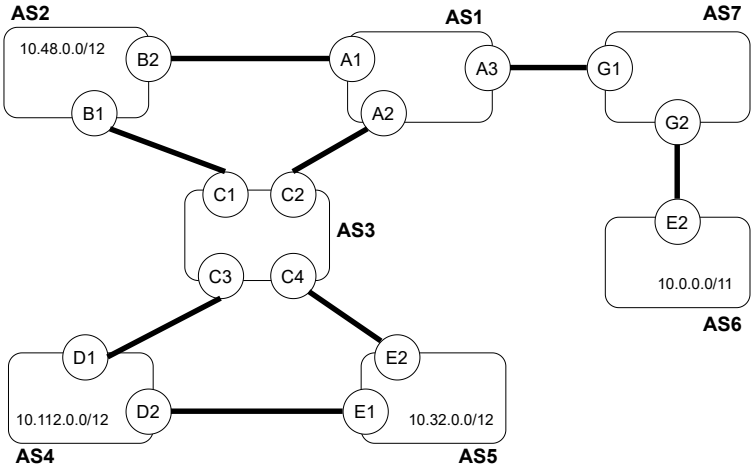
Problem 1.



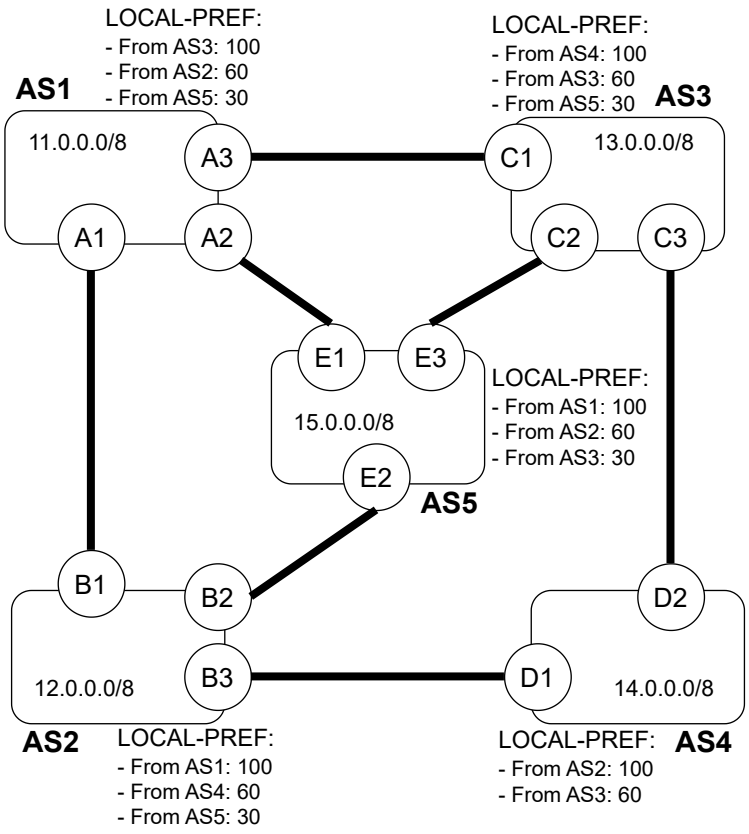
Problem 2, Question 1.



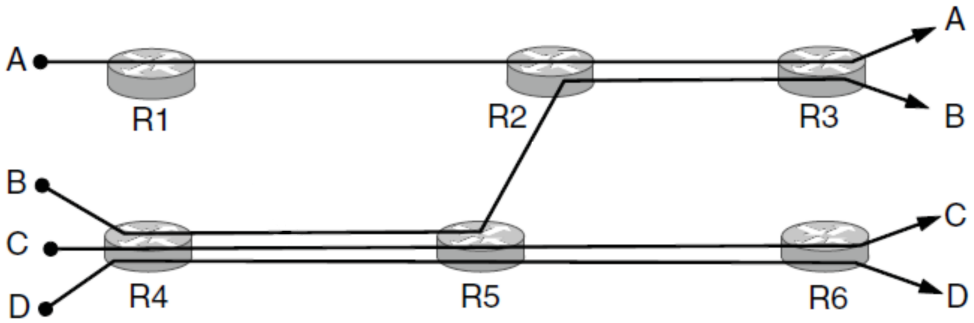
Problem 2, Question 2.



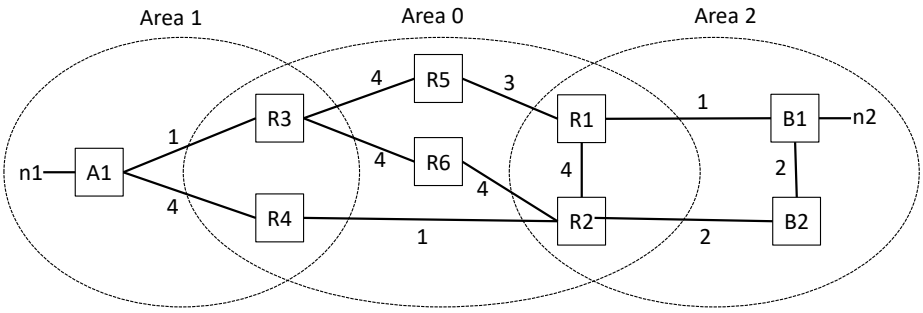
Problem 2. Question 3.



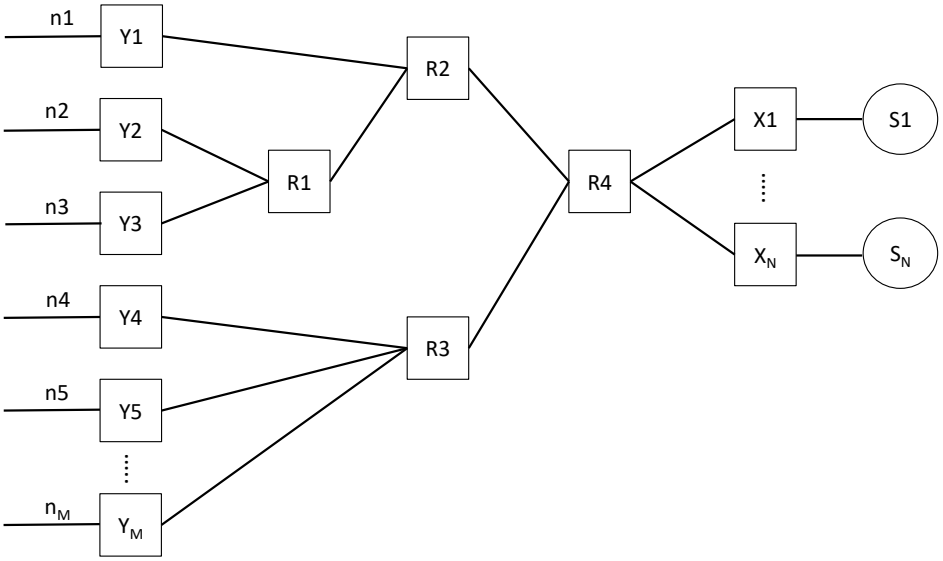
Problem 2. Question 4.



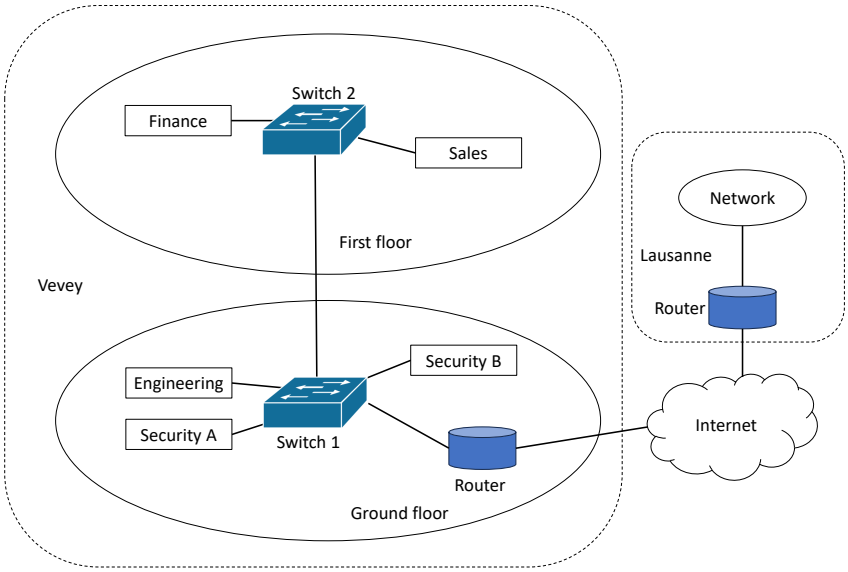
Problem 3 (Part B)



Problem 4, Question 1.



Problem 4, Question 2.



Problem 4, Question 3. ***All links are Ethernet links.***

TCP finite state machine

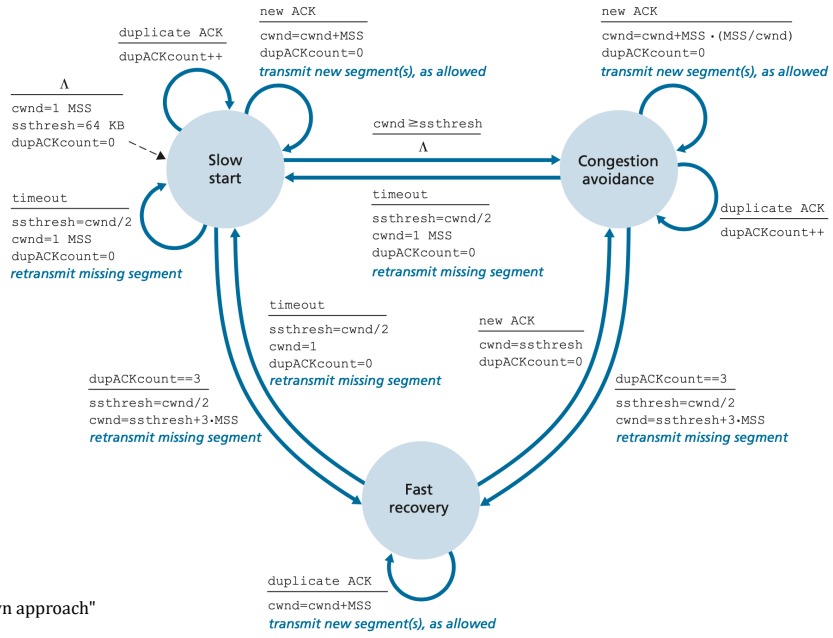


Figure from our textbook:
"Computer Networking: A top-down approach"
by J. Kurose and K. Ross