EXAM
TCP/IP NETWORKING
Duration: 3 hours
With Solutions

Jean-Yves Le Boudec

January 2023

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. No documents, no electronic equipments are allowed.

PROBLEM 1

Consider the network in the figure sheet. A, B, S1 and S2 are hosts. R1 and R2 are routers. BR1 to BR6 are bridges. The cloud represents the public internet. O1, O2, , ... , O11 are observation points. Machines where both IPv4 and IPv6 addresses appear are dual-stack, others are IPv6 only.

- The network box N1 operates as a NAT for IPv4 and IPv6. The internal port of the NAT is N1n.
- The network boxes N2 operates as a NAT for IPv4 (with internal port N2n) and as a router for IPv6.

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. A's default gateway is N1 and B's default gateway is N2. The hosts, bridges and routers have been running for some time and their different protocols are in steady-state. There is no other system or interface than shown on the figure.

The figure shows explicit IP addresses. Note that the IPv4 addresses of A and B are identical, as are N1n and N2n's. The symbols A and B also represent the MAC addresses of the hosts A and B. Similarly, the notation N2s, N2n, etc. denotes the MAC addresses of the corresponding interfaces.

1. Explain why it is possible that A and B have same IPv4 addresses.

   N1 and N2 are NATs for IPv4 hence they map the IP addresses of A and B, which are on the internal side, to the address of the external port, i.e. 10.0.3.2 (for N1) and 10.0.2.2 (for N2). The internal address of A [resp. B] is not visible to systems that are not onlink with A [resp. B], so the same address can be re-used.

2. Write S1 and B's IPv6 addresses in non-compressed forms (use hexadecimal digits).

Host	Compressed address	Non-compressed address
S1	3000:a:b:c::20	3000:000a:000b:000c:0000:0000:0000:0020
B	3000:a:b:c:2::b1	3000:000a:000b:000c:0002:0000:0000:00b1
B	fe80::1	fe80:0000:0000:0000:0000:0000:0000:0001

3. S1, N1s, N2s and R2n all have the same IPv4 subnet mask. Say which of the following values are valid for this common subnet mask.

   For this table and the next, the usual rules of quizzes apply: (i) if you cross the correct box in one row you obtain the full score for this row; (ii) if you cross zero or two boxes in one row, your score is 0 for this row; (iii) if you cross exactly one box in one row and it is incorrect your score for this row is negative, namely a penalty equal to half the score for this row.

Proposed v4 subnet masks at S1, N1s, N2s and R2n	valid	invalid
255.0.0.0	☒	☐
255.255.0.0	☒	☐
255.255.255.0	☐	☒

   The constraints are that S1, N1s, N2s and R2n all have the same subnet prefix.

4. B and S1 have the same IPv6 subnet mask length. Say which of the following values are valid for this common subnet mask length.

Proposed v6 subnet mask lengths at B and S1	valid	invalid
64	☐	☒
78	☐	☒
80	☒	☐
82	☒	☐

   The constraints are that

- B and N2n are in the same subnet
- N1s, N2s, R2n and S1 are in the same subnet, and this subnet is different than the subnet of B and N2n

5. The routing table at S1 is as follows.

dest	next-hop
3000:a:b:c:2::/80	3000:a:b:c::12
::	3000:a:b:c::1

There is no other entry, in particular, there is no IPv4 destination in S1's routing table. Now assume that S1 has an IPv6 packet to send to B. Will this work ? If so, to which destination MAC address does S1 send the packet to ? Same question for an IPv4 packet.

S1 does longest prefix match for the destination address of B, which is 3000:a:b:c:2::b1. The longest prefix match is the first row, and the packet should be sent to the next-hop 3000:a:b:c::12. This next-hop is in the same subnet, therefore S1 can send the packet directly to the MAC address of the next-hop, i.e. N2s. If needed, S1 learns this MAC address with NDP.

For IPv4, both answers are possible, since it depends on the system you consider. Either all lines in the routing table have been deleted. There is no entry and it is not possible to forward a packet. Or it remains the onlink lines which are put by default inside the routing table. Since B is behind a NAT, the IPv4 address that S1 knows for B is N2's external address, namely 10.0.2.2. This address is onlink with S1 (on the same bridged LAN). When S1 has a packet to sent to B, it sends it directly to the destination MAC address equal to N2s. If needed, S1 learns this MAC address with ARP.

6. A sends one UDP packet to S1, S1 sends one to S2, S2 sends one to B and B sends one to A. All are sent using IPv6. We observe the IPv6 packets containing these messages at observation points O3, O5, O6, and O9 (in both directions at every observation point). For each observation point, say whether such packets are observed and if so, give the IPv6 destination address. Give your answer in the table below. Put one line per packet (leave blank if no packet is observed). Justify your answer in the box below the table.

At	IPv6 destination address(es)
O3	
O5	(from A to S1) 3000:a:b:c::20 (from B to A) 3000:a:b:c::11
O6	(from A to S1) 3000:a:b:c::20 (from S1 to S2) 4100:b:b:b:b:2
O9	(from A to S1) 3000:a:b:c::20 (from S2 to B) 3000:a:b:c:2::b1

Justification:

- The spanning tree disables one of the links in the topology of bridges BR3-BR4-BR5-BR6. I assume the link B3-B4 is disabled.
- N1 is a NAT and the destination address for A is the IPv6 address of N1s.

7. B downloads a large web page from S2 using QUIC and IPv6. The local port used by B is 8765. At about the same time, A also downloads a large web page from S2 using QUIC over IPv6. By coincidence, the local port used by A is also 8765. The port used by S2 for HTTP over QUIC is 443. With Wireshark, we observe the packets sent by S2 to B and A at O11, O10, O8, O2, O7 and O1. Write the values of the fields in the tables below.

From S2 to B

At	MAC addresses		IP addresses		protocol	port numbers	
	srce	dest	srce	dest		srce	dest
O11	S2	R1n	4100:b:b:b:b::2	3000:a:b:c:2::b1	udp	443	8765
O10	R2n	N2s	4100:b:b:b:b::2	3000:a:b:c:2::b1	udp	443	8765
O8	R2n	N2s	4100:b:b:b:b::2	3000:a:b:c:2::b1	udp	443	8765
O7	N2n	B	4100:b:b:b:b::2	3000:a:b:c:2::b1	udp	443	8765

From S2 to A

At	MAC addresses		IP addresses		protocol	port numbers	
	srce	dest	srce	dest		srce	dest
O11	S2	R1n	4100:b:b:b:b::2	3000:a:b:c::11	udp	443	some value q
O10	R2n	N1s	4100:b:b:b:b::2	3000:a:b:c::11	udp	443	q
O2	R2n	N1s	4100:b:b:b:b::2	3000:a:b:c::11	udp	443	q
O1	N1n	A	4100:b:b:b:b::2	fc12:a:b:c::a1	udp	443	8765

Comment: N1 acts as a NAT for IPv6 and needs to map A's source address to its own south address. The source address seen by N1 is the one chosen by A when sending the first QUIC packet to S2. Since the destination (S2) is not onlink with A, the source address used by A is, in principle, its global private address fc12:a:b:c::a1 and not its link-local address fe80::a1.

8. B now downloads a large web page from S1 using TLS and IPv4. The local port used by B is 4321. At about the same time, A also downloads a large web page from S1 using TLS and IPv4. Again, by coincidence, the local port used by A is also 4321. The port used by S1 for HTTP over TLS is 443.

With Wireshark, we observe the packets sent by S1 to B and A at O4, O1 and O7. Write the values of the fields in the tables below.

From S1 to B							
	MAC addresses		IP addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O4	S1	N2s	10.0.4.1	10.0.2.2	tcp	443	some value p
O7	N2n	B	10.0.4.1	192.168.1.23	tcp	443	4321

From S1 to A							
	MAC addresses		IP addresses			port numbers	
At	srce	dest	srce	dest	protocol	srce	dest
O4	S1	N1s	10.0.4.1	10.0.3.2	tcp	443	some value $p' \neq p$
O1	N1n	A	10.0.4.1	192.168.1.23	tcp	443	4321

9. S1 now does a traceroute to B's second IPv6 address, namely to `fe80::1`. We observe the traffic resulting from this activity at observation point O7, in both directions. Which packets do we observe?

   S1 should not use the destination address `fe80::1` because it is link-local and B is not onlink with S1. In the likely case that there is no system in S1's subnet that has the same link-local address `fe80::1`, the traceroute activity fails, because S1 will try to send the traceroute packets to a MAC address that corresponds to `fe80::1`, and there is none. In the unlikely event that there is a system in S1's subnet, say X, that has the same link-local address `fe80::1`, the traceroute packets will be sent to X (and not to B), and nothing will be visible at O7.

   Thus in any case, we will not observe any packet resulting from this activity at O7.

PROBLEM 2

Consider the network for problem 2 in the figure sheet. There are four ASs, A, B, C and D with routers A1, A2, B1, C1, C2, C3, R1 and D1. The physical links are shown with plain lines. Each AS uses OSPF with Equal Cost Multipath as IGP, and every router inside each AS uses OSPF. The cost of every link and every directly attached network is 1, except when otherwise specified. C1 injects into OSPF the route `::` (the route with empty prefix), i.e., in C, OSPF will compute shortest path to `::` as if the empty prefix were directly attached to C1.

The figure shows one stub network at router A1 and one at router D1, each shown with its IPv6 network prefix. The lower case symbols such as a1e, b1n, etc. represent IPv6 addresses. H1 is a host, its IPv6 address is also indicated on the figure.

Routers A1, A2, B1, C1, C2 , C3, and D1 use BGP with their external neighbours and as required with their internal neighbours. The router R1 may or may not use BGP, depending on the question. No confederation or route reflector is used.

We assume that the BGP decision process use the following criteria in decreasing order of priority. BGP identifiers are router names such as A1, A2...

1. Highest LOCAL-PREF
2. Shortest AS-PATH
3. Lowest MED, if taken seriously by this network
4. E-BGP is preferred over I-BGP
5. Shortest path to NEXT-HOP, according to IGP
6. Lowest BGP identifier of sender of route is preferred. Here 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, we assume the following.

- 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 A, B, C and D is such that all available routes are accepted and propagated to neighbouring ASs, as long as the rules of BGP allow.
- Every router redistributes internal OSPF destinations into BGP. Unless otherwise specified, there is no other redistribution.
- Every router performs recursive forwarding-table lookup.
- Equal Cost Multi-Path routing is supported by all routers.
- When writing the AS-PATH attribute received by a router X, do not write the AS that this router is in. For example, if router X in AS A receives a route from router Y in AS B, who received it from AS C, the AS-PATH stored by X in its RIB is `AS-PATH = B C`. However, if X propagates this route over E-BGP to a router in AS D, the message sent by X will have `AS-PATH = A B C`.

1. At time t_1, BGP and OSPF have converged in all ASs.

   (a) At time t_1, what is the best BGP route to `3000:a:b::/48` selected by C1 ? From which BGP peer was it received ? What are its BGP NEXT-HOP and AS-PATH attributes ? Same question for C2 and C3. Give your answers in the table below, with a short justification.

At	From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH
C1	A1	3000:a:b::/48	a1s	A
C2	C1	3000:a:b::/48	a1s	A
C3	A2	3000:a:b::/48	a2e	A

Justification: Because OSPF-internal is redistributed into BGP, A1 announces `3000:a:b::/48, next-hop = a1s, AS-PATH=A` to C1 and A2 announces `3000:a:b::/48, next-hop = a2e, AS-PATH=A` to C3. C1 also possibly receives `3000:a:b::/48, next-hop = a1s, AS-PATH=A` from C3 but the route received from A1 is preferred because received from an external peer. The same holds when exchanging the roles of C1 and C3.
C2 receives one route to `3000:a:b::/48` from C1 and one from C3. The route obtained from C1 is preferred because it has the shortest distance to next-hop.
Eventually, all routers in AS C may also receive routes to `3000:a:b::/48` from D or from B. These routes have larger AS-PATH and are not preferred.

   (b) Still at time t_1, what is the best BGP route to `3000:a:b::/48` selected by B1 ? From which BGP peer was it received ? What are its BGP NEXT-HOP and AS-PATH attributes ? Give your answer in the table below, with a short justification.

At	From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH
B1	A2	3000:a:b::/48	a2s	A

Justification: B1 may receive routes from A2 (by E-BGP, with AS-PATH = A) or from C3 (by E-BGP, with AS-PATH = C A). B1 prefers the former because it has shorter AS path length.

   (c) Still at time t_1, what is the list of all BGP routes received and stored by C3 in its RIB, with destination = 3000:a:b::/48 ? For every route, indicate from which BGP peer it was received from, the BGP NEXT-HOP, the AS-PATH attributes, and whether it is selected as best. Give your answer in the table below, with a short justification (put as many rows as necessary).

At C3 :				
From BGP Peer	Destination Network	BGP NEXT-HOP	AS-PATH	Best route ?
C1	3000:a:b::/48	a1s	A	no
A2	3000:a:b::/48	a2e	A	yes
B1	3000:a:b::/48	b1e	B A	no
Justification: The route received from C1 is not selected because is has same AS-PATH length but comes from I-BGP. The route received from B1 is not selected because it has longer AS-PATH length. There is no router received from C2 because C2's best route is learnt by I-BGP (from C1).				

(d) Still at time t_1, H1 sends a packet to `3000:a:b:c:d:e:f::1`. Which path will this packet follow? At every router except the final router A1, say by which protocol (BGP or OSPF, or both) the next-hop is obtained. Justify your answer.
R1 does not run BGP and does not have a route to `3000:a:b::/48`, but, by OSPF, has obtained routing information to `::`, in the form of shortest paths to C1. The packet destination address matches `::`, and will be sent to the next-hop towards `::` (computed by OSPF). There are two equal cost paths to C1, one via C3, one via C2. Any of these will be selected with probability 50%. If the former path is chosen the packet reaches C3; here, C3 has a route to `3000:a:b::/48`, obtained from BGP. Therefore C3 sends the packet to the corresponding BGP NEXT-HOP, which results in the packet reaching A2. A2 has routing information obtained from OSPF and the packet reaches A1.
If the latter path is chosen, the packet reaches C2; here, too, C2 has a route to `3000:a:b::/48`, obtained from BGP, and forwards the packet to the corresponding BGP NEXT-HOP, i.e., a1s. By recursive table look-up, the packet is sent to the next-hop towards a1s, as indicated by OSPF. Thus the packet goes next to C1. At C1 the next hop is A1 (obtained by BGP).
In summary, the packet takes one of the following paths with equal probabilities; the string in parentheses indicates by which protocol the next hop decision was obtained:

- R1 (OSPF) - C3 (BGP) - A2 (OSPF) - A1
- R1 (OSPF) - C2 (BGP and OSPF) - C1 (BGP) - A1

2. In this question we assume that the network is restarted, with the following change in the configuration:

- R1 also runs BGP.

(a) At time t_2, BGP and OSPF have converged in all ASs and H1 sends a packet to `3000:a:b:c:d:e:f::1`. Which path will this packet follow? At every router except the final router A1, say by which protocol (BGP or OSPF, or both) the next-hop is obtained. Justify your answer.
Since R1 does not have external BGP peers, any BGP decision taken by R1 does not influence the other BGP routers, hence there is no change to the BGP information in all other routers. R1 obtains a BGP route to `3000:a:b::/48` from C1 and from C3, and prefers the one from C3 as it has shorter distance to next-hop. The BGP NEXT-HOP at R1 is thus a2e.
The packet is sent at R1 towards a2e and then, using recursive table lookup and the routing information obtained from OSPF, to C3. From there it is as in the previous question. The solution is thus

- R1 (BGP and OSPF) - C3 (BGP) - A2 (OSPF) - A1

(b) At time $t_3 > t_2$, C1 is re-started. Unfortunately, there is a configuration error in the new configuration file at C1: the OSPF interface database of C1 is configured with the prefix `3000:a:b:c::/64` (i.e the OPSF software believes that the network `3000:a:b:c::/64` is directly attached at C1). The default prefix `::` continues to be injected into OSPF by C1. At time $t_4 > t_3$, OSPF and BGP have converged again and R1 has a packet to send to `3000:a:b:c:d:e:f::1`. Which path will this packet follow? At every router except the final router A1, say by which protocol (BGP or OSPF, or both) the next-hop is obtained. Justify your answer.

C1 injects the bogus prefix `3000:a:b:c::/64` into OSPF. Since routers do not aggregate, both prefixes `3000:a:b::/48` and `3000:a:b:c::/64` are propagated into OSPF and BGP. However, the former is shorter and longest prefix match is used in forwarding tables. Thus, at R1, by OSPF, has obtained routing information to `3000:a:b:c::/64`, in the form of shortest paths to C1. Similar to the solution of 1(d), there are two equal cost paths to C1, one via C3, one via C2. Any of these will be selected with probability 50%. Again, at C2 or C3, longest prefix match is used in forwarding tables thus, by OSPF, the packet is sent to C1. At C1, there is no interface that matches the bogus prefix `3000:a:b:c::/64` but by BGP, C1 has obtained a routing information to the prefix `3000:a:b:c::/48` via A1.

In summary, the packet takes one of the following paths with equal probabilities; the string in parentheses indicates by which protocol the next hop decision was obtained:

- R1 (OSPF) - C2 (OSPF) - C1 (BGP) - A1
- R1 (OSPF) - C3 (OSPF) - C1 (BGP) - A1

3. In this question we assume that the network is restarted, with the following change in the configuration:

   - Inside AS A: routers tag all incoming E-BGP announcements received from AS C with LOCAL-PREF = 75 and from AS-D with LOCAL-PREF = 50. Other incoming E-BGP announcements are tagged with LOCAL-PREF = 0, same as before.
   - Inside AS C: routers tag all incoming E-BGP announcements received from AS A with LOCAL-PREF = 100 and from AS-D with LOCAL-PREF = 50. Other incoming E-BGP announcements are tagged with LOCAL-PREF = 0, same as before.

   The rest of the configuration is as in the previous question.
   At time t_3, BGP and OSPF have converged in all ASs. What is the best route to `4000:c:d::/48` selected by B1 ? Justify your answer.

   There are two possible stable equilibriums that might happen, depending on the order the BGP announcements were sent by the routers.

   - Equilibrium 1: PATH = C D.
   - Equilibrium 2: PATH = A D.

   Justification for equilibrium 1:
   Assume that a router in AS C first announces a route to `4000:c:d::/48` to a router in AS A, i.e., so far routers in AS C do not know a route to `4000:c:d::/48` through AS A, but a router in AS A receives a route `4000:c:d::/48` through AS C. Specifically, assume that C2 receives a route to `4000:c:d::/48` from D1 with PATH=D, and as it is the only known route to this destination, it is promoted as the best route; thus C2 announces it to its BGP peers, i.e., other routers in AS C. C1 and C3 receive this route, and with the same reasoning, they promote this route as the best one, thus announce it to their BGP peers. Hence, A1 and A2 receives a route to `4000:c:d::/48` with PATH= C D, and they will promote this one to best one. This is because there are two cases, depending whether A1 has already received a route to `4000:c:d::/48` by D1 or not: For the former, if a route is already received by D1, due to the highest LOCAL-PREF, the route with PATH= C D will be promoted as the best one; note that A1 and A2 will not announce the best route to their peers in AS C, because its best route includes AS C in the path, so that would create a loop. Else, for the later, if the route with PATH= C D is the only known route so far to `4000:c:d::/48` , then it will be promoted as the best one; also, if later A1 receives a route to `4000:c:d::/48` from D1, again due to the highest LOCAL-PREF, it will not be promoted as the best path, hence A1 announces nothing, and the best route to `4000:c:d::/48` remains the one with PATH = C D. Eventually, B1 receives two routes to `4000:c:d::/48`: 1) From A2 with PATH = A C D, and 2) From C3 with PATH = C D. The sec-

ond one will be promoted as the best route as it is shorter in AS path. Therefore, the best route to `4000:c:d::/48` at B1 is PATH = C D.

Justification for equilibrium 2:
Replace the role of AS A and AS C, and observe that B1 receives two routes to `4000:c:d::/48`: 1) From A2 with PATH = A D, and 2) From C3 with PATH = C A D. The first one will be promoted as the best route as it is shorter in AS path. Therefore, the best route to `4000:c:d::/48` at B1 is PATH = A D.

Note that there is also a unstable scenario: Assume we restart the network, and C2 receives a route with PATH = D from D1, and A1 also receives a route with PATH = D from D1, which they both promote it as the best. Then, A1 and C3 will announce it internally, and at some time t, the best route for routers in AS A is the one with PATH = D, and the best route for routers in AS C is the one with PATH = D. Then, in parallel, 1) A1 announces a route to C1 with PATH = A D and 2) C3 announces to A2 a route with PATH = C D; because of the highest LOCAL-PREF, A1 and C3 promote these new routes as best one, and will announce it internally. Then, it is possible that after a while, the best route for routers in AS A is the one with PATH = C D, and the best route for routers in AS C is the one with PATH = A D. Then, in parallel, 1) C1 sends a withdrawal for path with PATH = C D to A1 and 2) A2 sends a withdrawal for path with PATH = A D to C3; thus, A1 and C3 will announce this withdrawal it internally as well. Hence, at some time t' the best route for routers in AS A is the one with PATH = D, and the best route for routers in AS C is the one with PATH = D. This is exactly the state that we observed at time t, hence we might infinitely repeat this, and never get stabilized.

This unstable scenario was not considered when grading. But, note that LOCAL-PREF attribution might cause instability

PROBLEM 3

Consider the network for Problem 3 on the figure sheet.

- R1, R2 and R3 are routers. H1, H2, H3, S1, S2 and S3 are hosts. The links are full duplex with same rate in both directions.
- S1 [resp. S2, S3] sends one large flow of data to H1 [resp. H2, H3] (there are thus 3 unidirectional flows in total).
- There is no other system and no other flow than shown on the figure. There is no other capacity constraint than the link capacities shown on the figure. We also neglect the impact of the acknowledgement flows in the reverse direction.
- We neglect all overheads and assume that the link capacities can be fully utilized at bottlenecks.
- For $i = 1, 2, 3$ we call x_i the rate of the flow from Si to Hi .

1. Which of the following allocations, in Mb/s, are Pareto-efficient ? Justify your answer.

   (a) $x_1 = 1,\ x_2 = 2,\ x_3 = 6$
   (b) $x_1 = 0,\ x_2 = x_3 = 4.5$
   (c) $x_1 = x_2 = x_3 = 3$
   (d) $x_1 = 7,\ x_2 = x_3 = 1$

   Solution.
   All allocations are Pareto-efficient: it is not possible to unilaterally increase any rate because the capacity of the link R1-R2 is attained.

2. Assume the rates x_1, x_2, x_3 of the three flows are allocated according to max-min fairness. What are all the possible allocations ? Which ones are Pareto-efficient ?
Solution. Every link gives one constraint, i.e., the constraints are

$$x_1 + x_2 + x_3 \leq 9 \quad (1)$$
$$x_1 \leq 10'000 \quad (2)$$
$$x_2 \leq 10'000 \quad (3)$$
$$x_3 \leq 10'000 \quad (4)$$

If (**??**) holds then so do all other inequalities, hence the set of constraints is (**??**).
There is a unique max-min fair allocation and it is Pareto efficient. It can be computed by water-filling. At step 1 we assign $x_1 = x_2 = x_3 = 3$ Mb/s. The link between R1 and R2 saturates therefore all sources are frozen. Thus, the max-min fair allocation is $x_1 = x_2 = x_3 = 3$ Mb/s.

3. Assume the rates x_1, x_2, x_3 of the three flows are allocated according to proportional fairness. What are all the possible allocations ? Which ones are Pareto-efficient ? There is a unique proportionally fair allocation and it is Pareto efficient. The allocation is obtained by solving the following optimization problem:

$$\max \log x_1 + \log x_2 + \log x_3$$

over $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$ subject to:

$$x_1 + x_2 + x_3 \leq 9$$

The objective function and the constraints are symmetric and there is a unique solution, hence the solution must be such that $x_1 = x_2 = x_3$ (otherwise, if for example $x_1 \neq x_2$, then x_2, x_2, x_3 would be a second, distinct solution).
Therefore, $x_1 = x_2 = x_3 = 9/3 = 3$.

4. In this question we assume that the three flows use TCP Reno with ECN. Queuing at all routers is FIFO with RED enabled. The round trip times are:

   - 40 ms for flow 1,
   - 10 ms for flows 2 and 3.

   These numbers include all processing times. The MSS is the same for all flows and is equal to 1250 Bytes = 10^4 bits. We assume that the offered window is very large. Compute the rates of flows 1, 2 and 3.
   Since TCP allocations are Pareto efficient, the constraint in (**??**) holds with equality, i.e.

$$x_1 + x_2 + x_3 = 9$$

By using the loss-throughput formula for TCP Reno, we obtain

$$x_1 = \frac{MSS \cdot 1.22}{0.04\sqrt{q_1}} \quad (5)$$
$$x_2 = \frac{MSS \cdot 1.22}{0.01\sqrt{q_2}} \quad (6)$$
$$x_3 = \frac{MSS \cdot 1.22}{0.01\sqrt{q_3}} \quad (7)$$
$$(8)$$

where q_i is the proportion of marked ECN packets for flow i. Because the link capacities are all very large except for R1-R2, congestion occurs only at the output of R1 to R2 and hence we can assume

$$q_1 = q_1 = q_3$$

Combining with (**??**) we obtain

$$x_1 = 1, \; x_2 = x_3 = 4$$

5. In this question we continue with the same assumptions as in the previous question, except for the following modifications:
   - R3 now acts as an application layer gateway for flow 1.
   - The connection S1 - R3 has the same RTT as flows 2 and 3, namely 10 ms.
   - The RTT of the connection R3 - H1 is 122 ms.

   What is the rate achieved by flow 1 (from S1 to H1)? What are the rates achieved by flows 2 and 3?
   Flow 1 now consists of two piped connections: let us call them flow 1a (S1 - R3) and flow 1b (R3-H1).
   Flows 1a, 2 and 3 compete for 9 Mb/s; they have the same RTT and MSS, thus if there is no other constraint they obtain the same rate, namely 3 Mb/s.
   Flow 1b is alone and can obtain 10 Gb/s.
   However flow 1b is fed by the output of flow 1a hence it cannot have a larger rate- Thus the end-to-end rate of flow 1 is that of flow 1a, namely, 3 Mb/s.
   In summary, $x_1 = x_2 = x_3 = 3$.
6. In this question we continue with the same assumptions as in the previous question, except for the following modifications:
   - The link R3-H1 is wireless and has a loss ratio equal to $\ell = 0.0001$

   What is the rate achieved by flow 1 (from S1 to H1)? What are the rates achieved by flows 2 and 3?
   As before, flow 1 now consists of two piped connections: flow 1a (S1 - R3) and flow 1b (R3- H1).
   Flows 1a, 2 and 3 compete for 9 Mb/s; they have the same RTT and MSS, thus if there is no other constraint they obtain the same rate, namely 3 Mb/s.
   Flow 1b is alone and is constrained by its loss rate. Using the loss-throughput formula, we obtain the maximum value for its rate:

   $$x_{1b} = \frac{MSS \cdot 1.22}{0.122\sqrt{\ell}} = 10$$

   This value is larger than the rate available for flow 1a hence the end-to-end rate of flow 1 is that of flow 1a, namely, 3 Mb/s.
   In summary, $x_1 = x_2 = x_3 = 3$.
7. In this question we continue with the same assumptions as in the previous question, except for the following modifications:
   - the loss ratio of the link R3-H1 is now equal to $\ell = 0.01$

   What is the rate achieved by flow 1 (from S1 to H1)? What are the rates achieved by flows 2 and 3?
   As before, flow 1 now consists of two piped connections: flow 1a (S1 - R3) and flow 1b (R3- H1).
   Flows 1a, 2 and 3 compete for 9 Mb/s; they have the same RTT and MSS, thus if there is no other constraint they obtain the same rate, namely 3 Mb/s.
   Flow 1b is alone and is constrained by its loss rate. Using the loss-throughput formula, we obtain the maximum value for its rate:

   $$x_{1b} = \frac{MSS \cdot 1.22}{0.122\sqrt{\ell}} = 1$$

   This value is smaller than the rate available for flow 1a, therefore the flow 1a is also limited by this rate (because of flow control, the long-term rate of flows 1a and 1b must be equal).
   Hence flow 1ab is also limited by the constraint

   $$x_{1a} \leq 1$$

Since TCP is Pareto-efficient, flows 2 and 3 compete for 9-1=8 Mb/s; they have the same RTT and MSS, thus they obtain the same rate, namely 4 Mb/s.
In summary, $x_1 = 1,\ x_2 = x_3 = 4$.

PROBLEM 4

1. Lisa uses a media server to stream music in her house (see on the figure sheet). The audio server uses IP multicast, with source specific multicast. It uses the multicast address `232.1.2.3`. The music stream is unidirectional, from the media server to whomever is listening.

   (a) Lisa and Bart receive the music stream, each on their own separate device. Homer also would like to receive the audio stream. Say what happens at the IP layer when Homer decides to receive the audio stream. In particular, among the following machines: the media server, Lisa, Bart and Homer's machines, say which have to send specific IP layer control messages for Homer to be be able to receive the audio stream. (Application layer messages are not considered here).
   Solution. With IP multicast, it is the receiver that needs to join the group. Here, only Homer's device will send IGMP / MLD messages. The media server does nothing, it is not aware at the IP layer of the fact that Homer joins.

   (b) We assume in this question that Lisa's home network is a single bridged LAN. Lisa does a packet capture at the media server and at her device; she observes only the packets that contain the audio stream. Which addresses does she see in the packets sent by the media server ? Put the answer in the table below (see on the figure sheet for device addresses).

At media server:			
MAC source	MAC dest	IP source	IP dest
S	01:00:5e:01:02:03	10.0.0.9	232.1.2.3
At Lisa's device:			
MAC source	MAC dest	IP source	IP dest
S	01:00:5e:01:02:03	10.0.0.9	232.1.2.3

   (c) Assume that Lisa's home network contains several routers and Lisa is more than one hop away from the media server. Is there any change to the answer of the previous question ?
   Solution. At media server there is no change. At Lisa's device, the only change is the MAC source address, which is now the MAC address of the last router on the path between the media server and Lisa.

2. Consider the network for Problem 4, Question 2 on the figure sheet. All routers are BIER routers; R1 to R8 are BFERs (i.e., BIER Egress routers). The IP forwarding table of R10 is shown (only the lines that contain destination IP address = R1 to R8 are shown).
   The BIER network uses a bit-string of 8 bits, where position 1 is the least significant bit and position 8 is the most significant bit. Fill in the values of the forwarding bit masks in the Bier Index Forwarding table at R10. Use hexadecimal notation.

Destination BFER	Forwarding Bit Mask	Next-Hop
1	03	R9
2	03	R9
3	cc	R12
4	cc	R12
5	10	R5
6	20	R6
7	cc	R12
8	cc	R12

- For the destination BFER 1, the next-hop in the forwarding table of R10 is R9. According to the forwardig table of R10, the set of destination BFERs that are reached by the same next hop (R9) is $\{1, 2\}$. Hence the forwarding bit mask is $\{1, 2\}$ (see the lecture mcastsols.pdf, page 9). In binary, the mask is 00000011 (the bits corresponding to BFER IDs 1 and 2 are set to 1). In hexadecimal, this gives 03.
- For the destination BFER 2, the same reasoning applies and its forwarding bit mask is also 03 in hexadecimal motation.
- BFERs 3,4,7 and 8 share the same next hop in the forwarding table of R10 (next-hop is R12). Hence, for each of them the forwarding bit mask is $\{3, 4, 7, 8\}$, which in binary gives 11001100 (the bits for BFER IDs 3,4,7,8 are set to 1). In hexadecimal, this gives cc.
- BFER 5 is the only BFER destination with next-hop R5 in the forwarding table of R10. Its masks is hence 00010000 in binary, or 10 in hexadecimal.
- BFER 6 is the only BFER destination with next-hop R6 in the forwarding table of R10. Its masks is hence 00100000 in binary, or 20 in hexadecimal.

3. Consider the network for Problem 4, Question 3 in the figure sheet. Borducom is an Internet Service Provider (ISP) that has an IPv4 backbone and provides IPv6 connectivity by means of 6rd. Borducom has a box R which is a 6rd relay router; it connects to IPV4 on one side, and to IPv6 on the other side. The IPv6 side of R6 is connected to Tornado Networks, another ISP, which provides only IPv6 connectivity and has an IPv6-only backbone. Tornado networks has a box N that acts as a NAT64. The block of addresses 100.64.0.0/16 is reserved by Tornado Networks for the IPv4 side of the NAT64. The prefix 3000:a:b:ffff::/96 is reserved by Tornado Networks for the IPv6 side of the NAT64.
   You are a customer of Borducom and have a box M that connects your home network to Borducom. The box M acts as a 6rd router, and receives from Borducom the address shown on the figure.
   The prefix 2bbb:bbbb::/32 is reserved by Borducom for 6rd. The IPv6 prefix allocated to a customer network (such as our home network) is the concatenation of this prefix and the customer's IPv4 address provided by Borducom. This makes a 64-bit prefix, which can be used freely by the home network. The address 3.88.88.88 is reserved by Borducom to represent the IPv6 internet and is allocated to Borducom's 6rd relay router R.
   The hosts and routers have been running for some time and their different protocols are in steady-state.
   (a) Your home network is IPv6 only. Give a possible value for A's IPv6 address.
   The address prefix allocated by Borducom is obtained by concatenating the 32 bits of the reserved prefix with the 32 bits of our IPv4 address. In hexa, our IPv4 address is 0303:0404 therefore the allocated prefix is `2bbb:bbbb:0303:0404::/64` which can also be written as `2bbb:bbbb:303:404::/64`. Depending on how we allocate subnet prefixes in our home network, we obtain for A an address of the form
   `2bbb:bbbb:303:404:xxxx:xxxx:xxxx:xxxx`, where x stands for arbitrary hexadecimal digits.
   (b) The host A downloads a file from S6, using HTTP over TLS. We observe the IP headers of the packets resulting from this transfer, in the direction from S6 to A, at observation points O1, O2, O4 and O5. For each observation point, say whether there is a packet, and if so, give the IPv4

or/and IPv6 source and destination addresses. Write your answer in the table below.

Packets from S6 to A	
At	IPv4 or/and IPv6 source and destination addresses
O5	source IPv6 address : 3000:a:b:c::12, destination IPv6 address 2bbb:bbbb:303:404:h where h is the host part chosen for A.
O4	Nothing.
O2	Outer header: IPv4. Source : 3.88.88.88; destination: 3.3.4.4 Inner header: IPv6. Source : 3000:a:b:c::12; destination: 2bbb:bbbb:303:404:h
O1	IPv6 packet, source addr = 3000:a:b:c::12, dest addr = 2bbb:bbbb:303:404:h
Justification: The packets transit via R and not via N since it is v6 to v6 communication.	

(c) S6 downloads a file from S4, using HTTP over TLS. We observe the IP headers of the packets resulting from this transfer, in the direction from S4 to S6, at observation points O2, O3, O4 and O5. For each observation point, say whether there is a packet, and if so, give the IPv4 or/and IPv6 source and destination addresses. Write your answer in the table below.

Packets from S4 to S6	
At	IPv4 or/and IPv6 source and destination addresses
O2	Nothing.
O3	It is an IPv4 packet. Source : 3.4.5.6; destination: 100.64.x.y, where x, y are allocated on the fly by the NAT64.
O4	It is an IPv6 packet. Source: 3000:a:b:ffff::304:506 Destination: 3000:a:b:c::12. The source address is an IPv4-embedded-IPv6 address, using the prefix 3000:a:b:ffff::/96 reserved by Tornado Networks for the IPv6 side of the NAT64.
O5	Same as at O4.
Justification: The packets transit via N and not via R since it is v4 to v6 communication and N translates IPv4 packets into IPv6 packets.	

4. Consider the network for Problem 4, Question 4 in the figure sheet. A,B are LERs; the ports a,b of A and B are MPLS ports; the ports c,d,e,f of A and B are non-MPLS Ethernet ports. M,N,P and Q are LSRs. R is a non-MPLS router.
The host H sends a ping packet to the destination address 10.1.1.201; the destination replies and the ping reply is received by H. We observe the ping request and reply packets at observation points O1 to O6 (if they appear at such points). Give the MPLS labels and IP destination addresses (if any) that we see in such packets. Leave the space blank if there is no such header or no such packet.

Ping packet from H			
At	Outermost MPLS label	Next MPLS label	IP destination address
O1			
O2			10.1.1.201
O3			
O4			
O5	34	24	10.1.1.201
O6	99	24	10.1.1.201

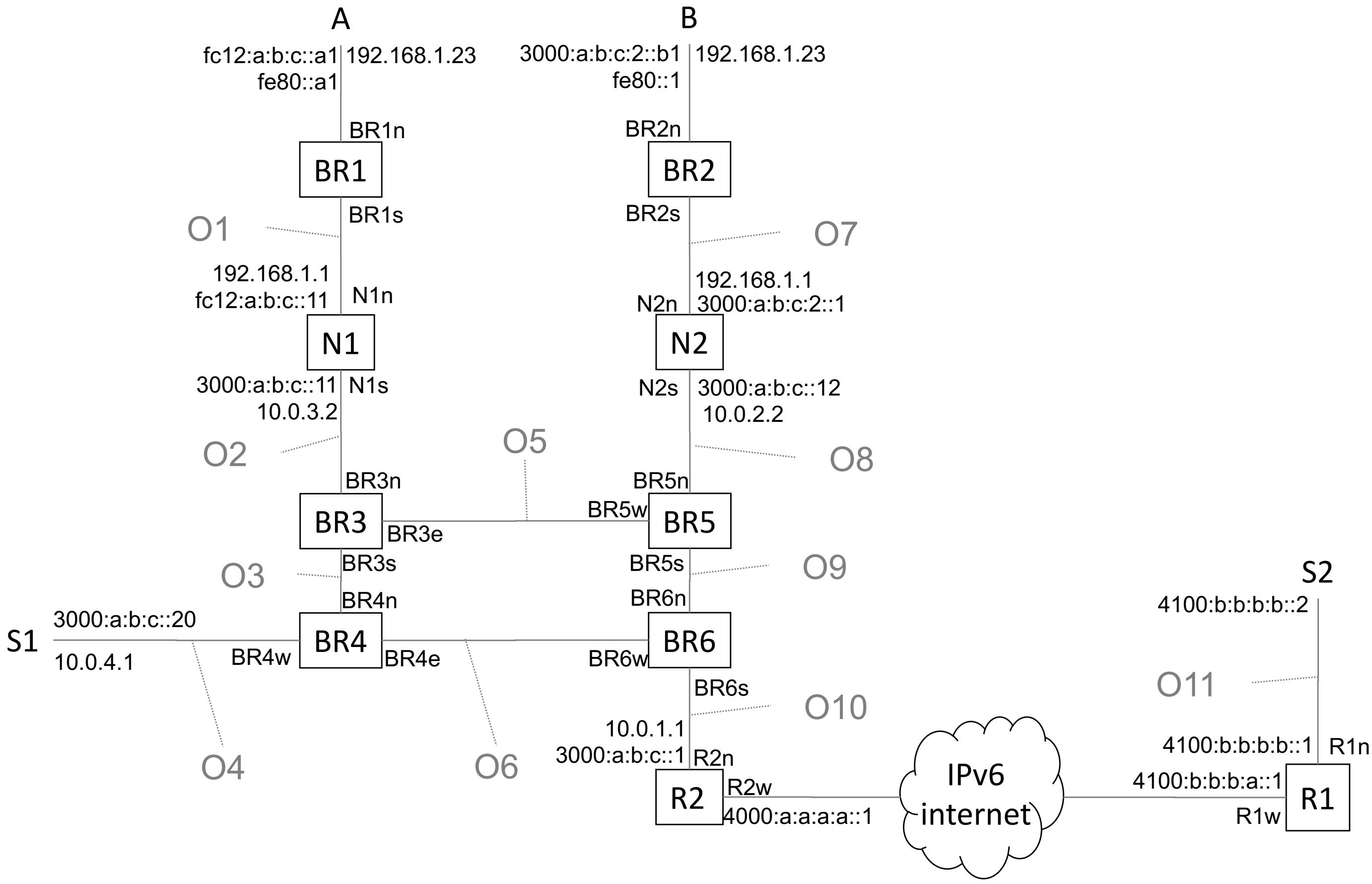
Ping packet to H			
At	Outermost MPLS label	Next MPLS label	IP destination address
O1			
O2			10.2.1.192
O3	50	24	10.2.1.192
O4	53	24	10.2.1.192
O5			
O6			

Comment:

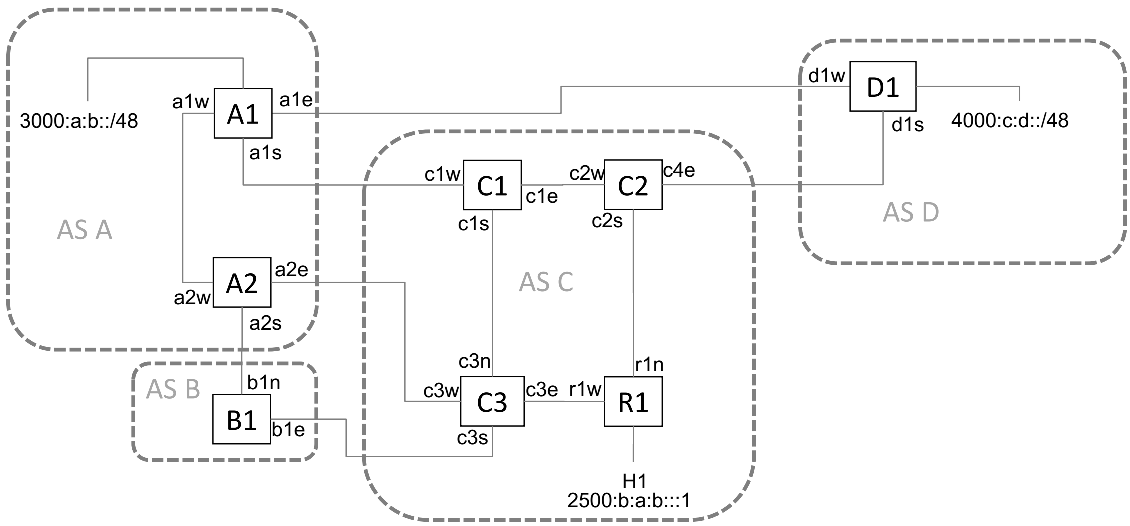
The ping packet is associated with VRF 1 and receives inner label 24 and outer label 99. The label switching tables at the LSRs indicate that the packet travels via B - N - M - A. At A the outer label is removed and the remaining label, namely 24, indicates VRF 1. The virtual routing table indicates that the outgoing port is e, therefore the ping packet goes to port e. I assume that the destination of the ping can reach A only via one port, hence the ping reply arrives on port e. The ping packet is associated with VRF 1 and receives inner label 24 and outer label 50. The label switching tables at the LSRs indicate that the packet travels via A - P - Q - B.

TCP IP EXAM - FIGURES

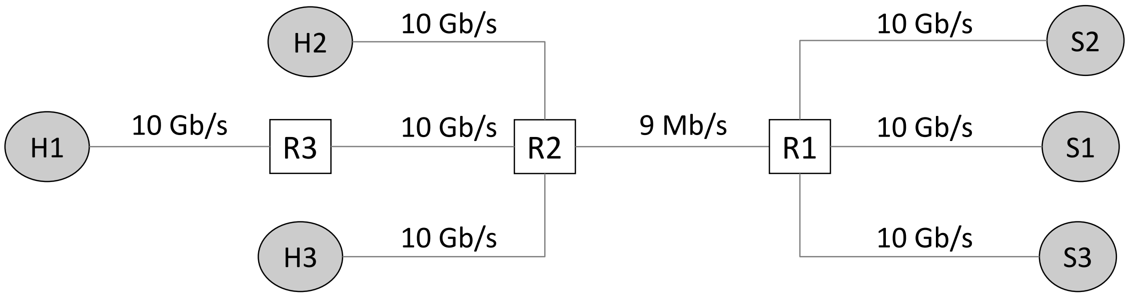
For your convenience, you can separate this sheet from the main document. Do not write your solution on this sheet, use only the main document. You do not need to return this sheet.



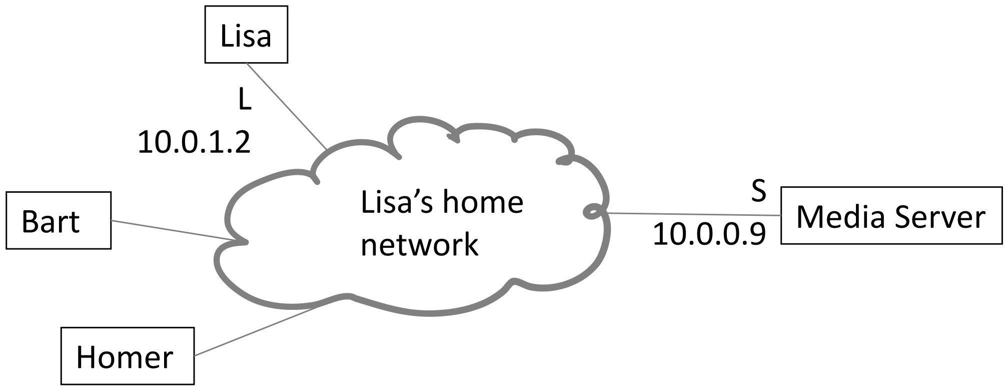
Problem 1.



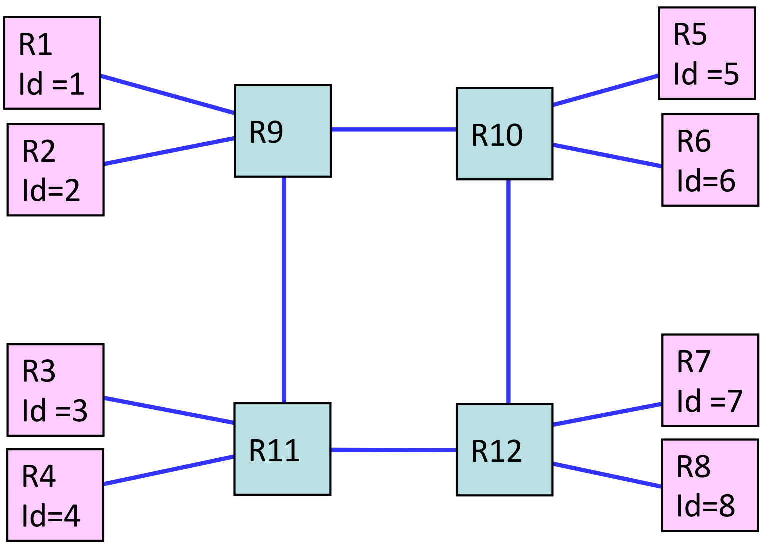
Problem 2.



Problem 3.



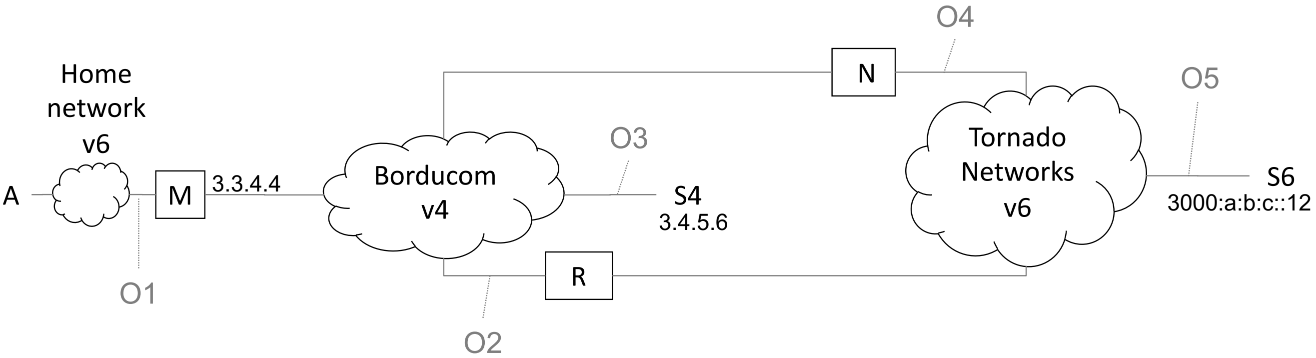
Problem 4, Question 1. L and S are MAC addresses.



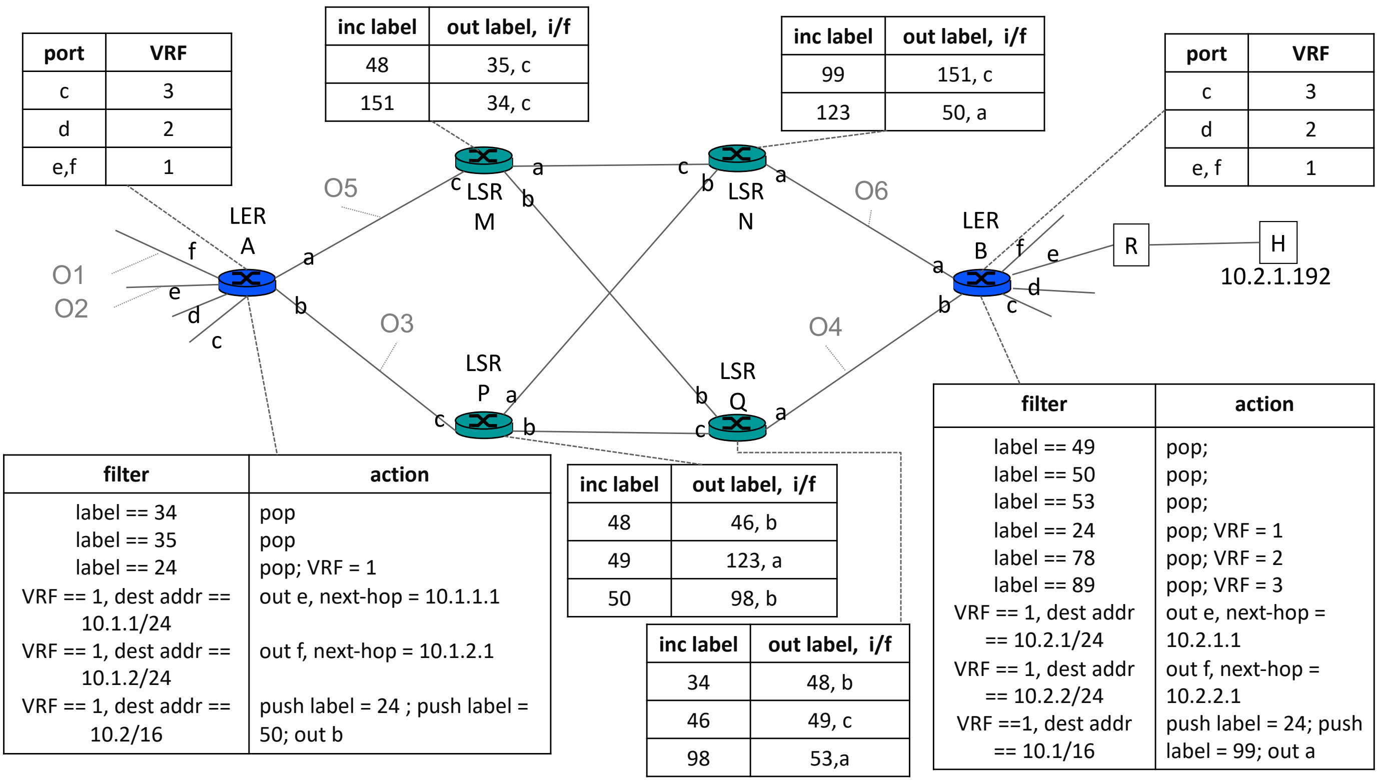
IP Forwarding Table at R10

Dest. IP addr	Next-Hop
R1	R9
R2	R9
R3	R12
R4	R12
R5	R5 (onlink)
R6	R6 (onlink)
R7	R12
R8	R12

Problem 4, Question 2.



Problem 4, Question 3.



LER A:

port	VRF
c	3
d	2
e,f	1

filter	action
label == 34	pop
label == 35	pop
label == 24	pop; VRF = 1
VRF == 1, dest addr == 10.1.1/24	out e, next-hop = 10.1.1.1
VRF == 1, dest addr == 10.1.2/24	out f, next-hop = 10.1.2.1
VRF == 1, dest addr == 10.2/16	push label = 24 ; push label = 50; out b

LSR M:

inc label	out label, i/f
48	35, c
151	34, c

LSR N:

inc label	out label, i/f
99	151, c
123	50, a

LSR P:

inc label	out label, i/f
48	46, b
49	123, a
50	98, b

LSR Q:

inc label	out label, i/f
34	48, b
46	49, c
98	53,a

LER B:

port	VRF
c	3
d	2
e, f	1

filter	action
label == 49	pop;
label == 50	pop;
label == 53	pop;
label == 24	pop; VRF = 1
label == 78	pop; VRF = 2
label == 89	pop; VRF = 3
VRF == 1, dest addr == 10.2.1/24	out e, next-hop = 10.2.1.1
VRF == 1, dest addr == 10.2.2/24	out f, next-hop = 10.2.2.1
VRF ==1, dest addr == 10.1/16	push label = 24; push label = 99; out a

Problem 4, Question 4.