SCIPER: ____________________

First name: ________________________ Family name: ________________________

EXAM
TCP/IP NETWORKING
Duration: 3 hours
With Solutions

Jean-Yves Le Boudec

January 2018

INSTRUCTIONS

1. 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.
2. All problems have the same weight.
3. You may need to make additional assumptions in order to solve some of the questions. If this happens, please describe such assumptions explicitly.
4. Figures are on a separate sheet, for your convenience.
5. No documents, no electronic equipments are allowed.
6. Justify every answer with a short explanation.

PROBLEM 1

Consider the network for problem 1 in the figure sheet. A, B and C are hosts; $BR1$, $BR2$ and $BR3$ are bridges; $R1$ and $R2$ are routers. X is a network box that can be configured in different ways, as explained next. $O1$ to $O6$ are observation points where we observe traffic in both directions of the link. Some selected IP addresses are shown, as well as some selected MAC addresses (denoted with e.g. A, B, $BR1w$, ..., $R1e$, ...). You may need to specify other IP or MAC addresses. 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 network masks are `ffff:ffff:ffff:ffff::` at all interfaces shown on the figure, unless otherwise specified. Addresses such as `fd24:a:b:1::2` are private IPv6 addresses.

1. Give the uncompressed version (all hexadecimal digits) of the address `fd24:a:b:1::2`.

   Solution. `fd24:000a:000b:0001:0000:0000:0000:0002`.

2. In this question X is a NAT.

   (a) Give possible values for the complete IP addresses of the interfaces $R1e$ and $R2e$ (i.e. give possible values for the unknown parts of the IP addresses indicated as ?). Shortly justify your answer.

   Solution. The value of the unknown part will be 2 as host B is in the same LAN and the IPv6 address of host B has 2 in its network part of the address.

   (b) A sends one ping message to B. We observe the ping request packets resulting from this activity at observation point $O1$. What are the MAC and IP source and destination addresses in such packets ? What is the Hop Count field, knowing that the HC value is equal to 64 in all IP packets generated by all hosts in this problem ? Put your answers in the table below.

At observation point $O1$:				
MAC source	MAC dest	IP source	IP dest	HC value
`R1e`	`MAC address of B`	`fd24:a:b:1::2`	`fd24:a:b:2::2`	`63`

   (c) A downloads a huge file from a web server at C using HTTP. A uses the local port 4567. At the same time, B also downloads a file from C, also using HTTP. By coincidence, B uses the same local port number, namely 4567. We observe the IP headers in the packets resulting from this transfer at $O5$ and $O6$, in the direction from A to C and from B to C. Give possible values of the protocol, the source and destination port numbers and the source and destination IP addresses. Give the answers in the tables below.

At observation point $O5$, from A and B to C:				
IP source	IP dest	protocol	source port	dest. port
`fd24:a:b:1::2`	`2001:bebe:baba:1::1`	`http`	`4567`	`80`
`fd24:a:b:2::2`	`2001:bebe:baba:1::1`	`http`	`4567`	`80`

At observation point $O6$, from A and B to C:				
IP source	IP dest	protocol	source port	dest. port
`2001:a:b:1::1`	`2001:bebe:baba:1::1`	`http`	`e.g., 45678`	`80`
`2001:a:b:1::1`	`2001:bebe:baba:1::1`	`http`	`e.g., 45679`	`80`

3. In this question X is an application layer gateway, acting as web proxy.

   (a) A downloads a huge file from a web server at C using HTTP. A uses the local port 4567. At the same time, B also downloads a file from C, also using HTTP. By coincidence, B uses the same local port number, namely 4567. We observe the IP headers in the packets resulting from this transfer at $O5$ and $O6$, in the direction from A to C and from B to C. Give possible values of the protocol, the source and destination port numbers and the source and destination IP addresses. Give the answers in the tables below.

At observation point $O5$, from A and B to C:				
IP source	IP dest	protocol	source port	dest. port
`fd24:a:b:1::2`	`fd24:a:b:4::1`	`http`	`4567`	`e.g., 8080`
`fd24:a:b:2::2`	`fd24:a:b:4::1`	`http`	`4567`	`e.g., 8080`

At observation point $O6$, from A and B to C:				
IP source	IP dest	protocol	source port	dest. port
`2001:a:b:1::1`	`2001:bebe:baba:1::1`	`http`	`e.g., 8932`	`80`
`2001:a:b:1::1`	`2001:bebe:baba:1::1`	`http`	`e.g., 8933`	`80`

   (b) B sends one ping message to X. Where, among observation points $O2$, $O3$ and $O4$, is the ping request packets resulting from this activity visible ? Justify your answer.

   Solution. The bridges build a tree and must therefore disable one of the links. Assume it is the link where $O2$ is. Once the bridges have had enough time to learn all the MAC addresses, BR1 knows that X is towards BR3 and the ping request will be visible only at $O3$.

   (If the link that is disabled by the spanning tree protocol is $O3$, then the ping request is visible at $O2$ and $O4$. If the link that is disabled by the spanning tree protocol is $O4$, then the ping request is visible at $O3$ only).

4. In this question we would like that (1) X is configured as a router, (2) the IP address of Xw is unchanged, and (3) the ISP that offers internet service at point $O6$ to N has to inject only one network prefix into its routing tables for the whole of N. IP addresses and masks of interfaces in N (other than Xw) can be modified; no other change is allowed to the configuration of N. Give a possible way to achieve this goal. Give all details of all addresses and masks that may need to be changed by your solution.

   Solution. New network prefix at Xe, R2e, and R1w may contain 66 first bits where last two bits can be 01, 10, and 11. For example, the network prefix for Xe and R2w can be `ffff:ffff:ffff:ffff:4:0:0:0`. X3 and R2w can have any ip addresses in the host addresses, e.g., `2001:a:b:1:4:0:0:1` and `2001:a:b:1:4:0:0:2`. In the same way, other ip addresses and network masks can be computed.

   Alternatively, address prefix delegation through DHCPv6 can be used at X.

SCIPER: ____________________

First name: ________________________ Family name: ________________________

PROBLEM 2

Consider the network for problem 2 in the figure sheet. Unless otherwise specified, there are three ASs, A, B and C. Each AS uses OSPF with Equal Cost Multipath as IGP, and all routers inside each AS uses OSPF. The cost of every link and every directly attached network is 1, except for two links shown on the figure, which have a cost of 10.

The figure shows two stub networks, at routers $A3$ and $A4$, with their IPv6 address prefixes. The lower case symbols such as $b1e$, $b1w$ also represent IPv6 addresses.

Routers $A1, A2, B1, B2, C1$ and $C2$ use BGP with their external neighbours and as required with their internal neighbours. Unless otherwise specified, the other routers do not use BGP. We assume that the BGP decision process use the following criteria in decreasing order of priority.

1. Shortest AS-PATH
2. Lowest MED, if taken seriously by this network
3. E-BGP $>$ I-BGP
4. Shortest path to NEXT-HOP, according to IGP
5. Lowest BGP identifier

Furthermore, we assume that no optional BGP attribute (such as MED, LOCAL-PREF etc.) is used in any BGP message and that no aggregation is performed.

1. Each AS is configured to redistribute internal OSPF routes into BGP and to redistribute E-BGP routes into OSPF.

   (a) When BGP and OSPF have converged in all ASs, which routes are selected by $B1$, $B2$, $A1$ and $A2$? Give the answers in the tables below. Justify your answers.
   Solution.

At B1	DESTINATION NETWORK	NEXT-HOP	AS-PATH
	2001:baba:bebe/48	c1w	C
	2001:baba:bebf/48	c1w	C
At B2	DESTINATION NETWORK	NEXT-HOP	AS-PATH
	2001:baba:bebe/48	c2w	C
	2001:baba:bebf/48	c2w	C

Justification:
Both B1 and B2 will receive announcements for the prefixes from their peers in AS C over E-BGP and from the other router in AS B over I-BGP. Since E-BGP is preferred over I-BGP, the next hop in AS C will be selected.

At A1	DESTINATION NETWORK	NEXT-HOP	AS-PATH
	2001:baba:bebe/48	b1w	B C
	2001:baba:bebf/48	b1w	B C
At A2	DESTINATION NETWORK	NEXT-HOP	AS-PATH
	2001:baba:bebe/48	b2w	B C
	2001:baba:bebf/48	b2w	B C

Justification:
Both A1 and A2 will receive announcements from their peers in AS B over E-BGP and from their peer in AS A over I-BGP. Because E-BGP is preferred over I-BGP, the next hop in AS B will be selected.

(b) Recall that each AS is configured to redistribute E-BGP routes into OSPF. When redistributing BGP into OSPF, the OSPF cost of the redistributed route is set to the OSPF cost to the BGP next-hop. When OSPF has converged, what are the routing table entries for the destinations `2001:baba:bebe/48` and `2001:baba:bebf/48` at routers $A3$ and $A4$? Justify your answer.

Solution.

At A3	DESTINATION NETWORK	NEXT-HOP	DISTANCE
	2001:baba:bebe/48	a2w	2
	2001:baba:bebf/48	a2w	2
At A4	DESTINATION NETWORK	NEXT-HOP	DISTANCE
	2001:baba:bebe/48	a1w	2
	2001:baba:bebf/48	a1w	2

Justification: In OSPF, the two subnets (redistributed from BGP) are treated as on-link at A1 and A2. Then, according to the shortest path calculation, we obtain the above results.

(c) Which path does a packet follow from $A4$ to `2001:baba:bebe::1` and `2001:baba:bebf::1` ? Give the answer in the form of a sequence of routers. Give the answer in the table below.
Solution.

From $A4$ to	Path
`2001:baba:bebe::1`	A1 B1 C1 C2 C3 C4
`2001:baba:bebf::1`	A1 B1 C1 C2 C3
Justification: Based on the previous questions, the packets will first pass by A1, B1, C1. Then, they will respectively follow the shortest path inside AS C.	

(d) Assume, in this question only, that $B2$ is compromised and now sends to $A2$ the (bogus) BGP announcement

```
DESTINATION = 2001:baba:bebe::/49, AS-PATH = B C, NEXT-HOP=c1w
```

Also assume that $A2$ accepts this BGP announcement. When BGP and OSPF have stabilized again in AS A, which path inside AS A does a packet follow from $A4$ to `2001:baba:bebe::1` ? Give the answer below in the form of a sequence of routers. Justify your answer.
Solution.

From $A4$ to	Path
`2001:baba:bebe::1`	A1 (or A3) A2
Justification: By longest prefix match, A4 will send the packet to A2 via A1 (or A3). Note that there are two possibilities due to Equal Cost Multipath.	

2. The configuration of AS A is now modified: BGP is not redistributed into OSPF at A (but continues to be redistributed at B and C). Explain what can be done in AS A to maintain full connectivity, in particular, we would like all routers in AS A to be able to forward packets to `2001:baba:bebe/48` and `2001:baba:bebf/48` (propose only one solution). Note that, in this question, you may make changes to the assumptions that we put at the beginning of Problem 2.
Solution. We could configure A1 or A2 as the default gateway of A3, A4. Alternatively, we could run BGP on all routers in AS A.
3. Following a re-organization, the three domains A, B and C are now merged into one single large AS. We assume in this question that all routers in the figure use OSPF with the assumptions given at the beginning of Problem 2 (note that all routers in the figure are in the same OSPF domain for the moment). BGP is not used in the rest of this problem.

(a) Assume that there is a single area, i.e. all routers in the figure are in the same OSPF area. When OSPF has converged, what are the routing table entries for the destinations `2001:baba:bebe/48` and `2001:baba:bebf/48` at routers $A3$ and $A4$? Justify your answer.
Solution.

At A3	DESTINATION NETWORK	NEXT-HOP	DISTANCE
	2001:baba:bebe/48	a2w	6
	2001:baba:bebf/48	a2w	5
At A4	DESTINATION NETWORK	NEXT-HOP	DISTANCE
	2001:baba:bebe/48	a1w (or a3n)	7
	2001:baba:bebf/48	a1w (or a3n)	6
Justification: Routers compute the shortest paths to the destinations. Paths with the same cost are all kept.			

(b) We would now like to configure this OSPF network with 3 areas, such that:

i. $B1$ and $B2$ are in the backbone area
ii. $A1, A2, A3, A4$ are in a common area
iii. $C1, C2, C3, C4$ are in a common area
iv. no router is a single point of failure

Give a possible allocation of routers to areas that satisfy these constraints. Justify your answer.

Solution. For example,

- Area 0: A1, A2, B1, B2, C1, C2;
- Area 1: A1, A2, A3, A4;
- Area 2: C1, C2, C3, C4.

Justification: Area 0 must interconnect areas 1 and 2 and Border Area routers must belong to both area 0 and one other area.

SCIPER: ____________________

First name: ________________________ Family name: ________________________

PROBLEM 3

Consider the network for problem 3 on the figure sheet.

- A, B and C are routers. The capacity of the links between them is 480 Mb/s. The links are full duplex with same rate in both directions.
- There are 4 unidirectional flows, as shown on the figure. There is no other system and no other flow than shown on the figure. There is no other capacity constraint than the two 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.

1. Assume the rates x_1, x_2, x_3, x_4 of the four flows are allocated according to max-min fairness. Compute the values of x_1, x_2, x_3, x_4.
   Solution. The rates can be computed via water-filling. At step 1 we assign $x_1 = x_2 = x_3 = x_4 = 160$ Mb/s. The link between B and C saturates and thus x_1, x_3, x_4 freeze at 160 Mb/s. At step 2 x_2 increases to 320 Mb/s and then the link between A and B saturates. Thus, for the max-min fair allocation $x_1 = x_3 = x_4 = 160$ Mb/s and $x_2 = 320$ Mb/s.

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

   (a) $x_1 = 160,\ x_2 = 320,\ x_3 = x_4 = 160$
   (b) $x_1 = 240,\ x_2 = 240,\ x_3 = x_4 = 120$
   (c) $x_1 = 240,\ x_2 = 240,\ x_3 = x_4 = 100$
   (d) $x_1 = 240,\ x_2 = 240,\ x_3 = 140,\ x_4 = 100$

   Solution. The Pareto-efficient allocations are (a), (b) and (d). (a) is the max-min fair allocation which is Pareto-efficient. These three allocations fully utilize the resources of the network (i.e., the link capacities) and thus, in each of them for increasing one source another source must be decreased; that is the definition of Pareto-efficiency. The allocation (c) is not Pareto-efficient. Specifically, x_3, x_4 can be increased without the need of decreasing any of x_1, x_2, which contradicts the definition of Pareto-efficiency.

3. Assume the rates x_1, x_2, x_3, x_4 of the four flows are allocated according to proportional fairness. Compute the values of x_1, x_2, x_3, x_4. Is this a Pareto-efficient allocation ?

Solution. The rates corresponding to the proportional fair allocation should solve the following optimization problem:

$$\max \log x_1 + \log x_2 + \log x_3 + \log x_4$$

subject to:

$$x_1 + x_2 \leq 480$$

$$x_1 + x_3 + x_4 \leq 480$$

$$x_i \geq 0, \; i = 1, 2, 3, 4$$

The first two constraints should be satisfied with equality (otherwise the objective is not maximized over all feasible allocations), therefore,

$$x_2 = 480 - x_1$$

$$x_3 = x_4 = \frac{480 - x_1}{2}.$$

Note that $\log x_3 + \log x_4$ is maximized for $x_3 = x_4$ due to the concavity of the log function and the considered capacity constraints.

Thus, the optimization problem can be transformed as

$$\max \log x_1 + \log(480 - x_1) + 2 \log \frac{480 - x_1}{2}$$

$$x_1 \geq 0$$

$$x_1 \leq 480$$

which is solved by $x_1 = 120$ Mb/s. Thus, $x_2 = 360$ Mb/s and $x_3 = x_4 = 180$ Mb/s.

This is a Pareto-efficient allocation since Proportionally Fair allocations are Pareto-efficient.

4. In this question flow 1 is using UDP and sends at a constant rate equal to 80 Mb/s. Flows 2, 3, 4 use TCP Reno with ECN. Queuing at all routers if FIFO with RED enabled. The round trip times are:

   - 300 ms for flows 2 and 3,
   - 100 ms for flow 4.

   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 2, 3 and 4.

   Solution. UDP source is a greedy source that will not adapt its rate. Therefore the leftover capacity at each link for the TCP flows is equal to 400 Mb/s. The flow 2 does not compete with another TCP source and since TCP allocations are Pareto efficient,the flow 2 must fully utilize the first link. Thus, $x_2 = 400$ Mb/s. In the link between B and C, the TCP sources x_3 and x_4 compete each other and TCP Reno allocations depend on the RTT values. Specifically, for symmetric flows (such as x_3 and x_4) higher RTT implies a smaller rate. By using the loss-throughput formula for TCP Reno, we obtain

$$x_3 = \frac{MSS \cdot 1.22}{300\sqrt{q}}$$

   and

$$x_4 = \frac{MSS \cdot 1.22}{100\sqrt{q}}$$

   where q is the percentage of the ECN-marked packets, the same for both x_3 and x_4. Therefore, $x_4 = 3x_3$. Since TCP Reno allocations are Pareto-efficient, $x_3 + x_4 = 400$ and as a result, $x_4 = 300$ Mb/s and $x_3 = 100$ Mb/s. To conclude, $x_2 = 400$ Mb/s, $x_4 = 300$Mb/s and $x_3 = 100$ Mb/s.

5. Assume now that the flows 2, 3, 4 use TCP Cubic with ECN instead of TCP Reno with ECN; the rest is as in the previous question. Can you guess how the allocation of rates would differ ? We don't ask you to compute the rates; simply put an X in the correct boxes in the table below, with a short justification.

Solution:

the rate achieved with TCP Cubic and ECN is ...	flow 2	flow 3	flow 4
higher		X	
same	X		
lower			X

Justification: For flow 2 there is no difference because it is a single flow and all TCP allocations are Pareto efficient, so flow 2 must fully utilize the first link in all cases.

For flows 3 and 4, the rates depend on the bias of TCP against large RTTs. At large RTT and large throughputs, Cubic is less sensitive to RTT than Reno. We need to see if we are in the region of large RTT and large throughput, where Reno does differ from Cubic. We see this on the slide given in the exam booklet; here the RTTs of interest are 100 and 300 msec. We see on the slide that for such RTTs and for throughputs larger than ≈ 50 Mb/s we are in the region where Cubic differs. Therefore we expect the rates of flows 3 and 4 to be more similar than with TCP Reno, namely higher for flow 3 and lower for flow 4.

SCIPER: ____________________

First name: ______________________ Family name: ______________________

PROBLEM 4

1. $R1$ is a router and uses RIP, a distance-vector routing protocol that implements the distributed Bellman-Ford algorithm. All link costs are equal to 1. At time t_0, the routing table at $R1$ contains

Destination Network	Next-Hop	Distance
10.0.0.0 / 8	10.1.1.1	5
10.1.0.0 / 16	10.1.2.1	2

The router $R2$ is a neighbour of $R1$ and has IP address `10.1.2.1`. At $t_1 > t_0$, $R1$ receives from $R2$ the distance-vector message

```
Destination = 10.0.0.0/8, distance = 3
Destination = 10.1.0.0/16, distance = 3
```

No other message is received between t_0 and t_1. Just after processing this message, what is the state of the routing table at $R1$? Give your solution by filling the empty cells in the table below.

Destination Network	Next-Hop	Distance
10.0.0.0 / 8	10.1.2.1	4
10.1.0.0 / 16	10.1.2.1	4

Solution. The next-hop to 10.0.0.0 / 8 is updated to 10.1.2.1, since it is path with smaller distance, which is equal to 4. And next-hop to 10.1.0.0/16 is increased to 4, since there is no other paths and here only distance is updated.

2. 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.

3. The web servers that serve the url `https://sovkom.coolstuff` are distributed across several sites. The managers of sovkom would like to spread the load between the different sites. How can that be done without modifying anything on user devices ?

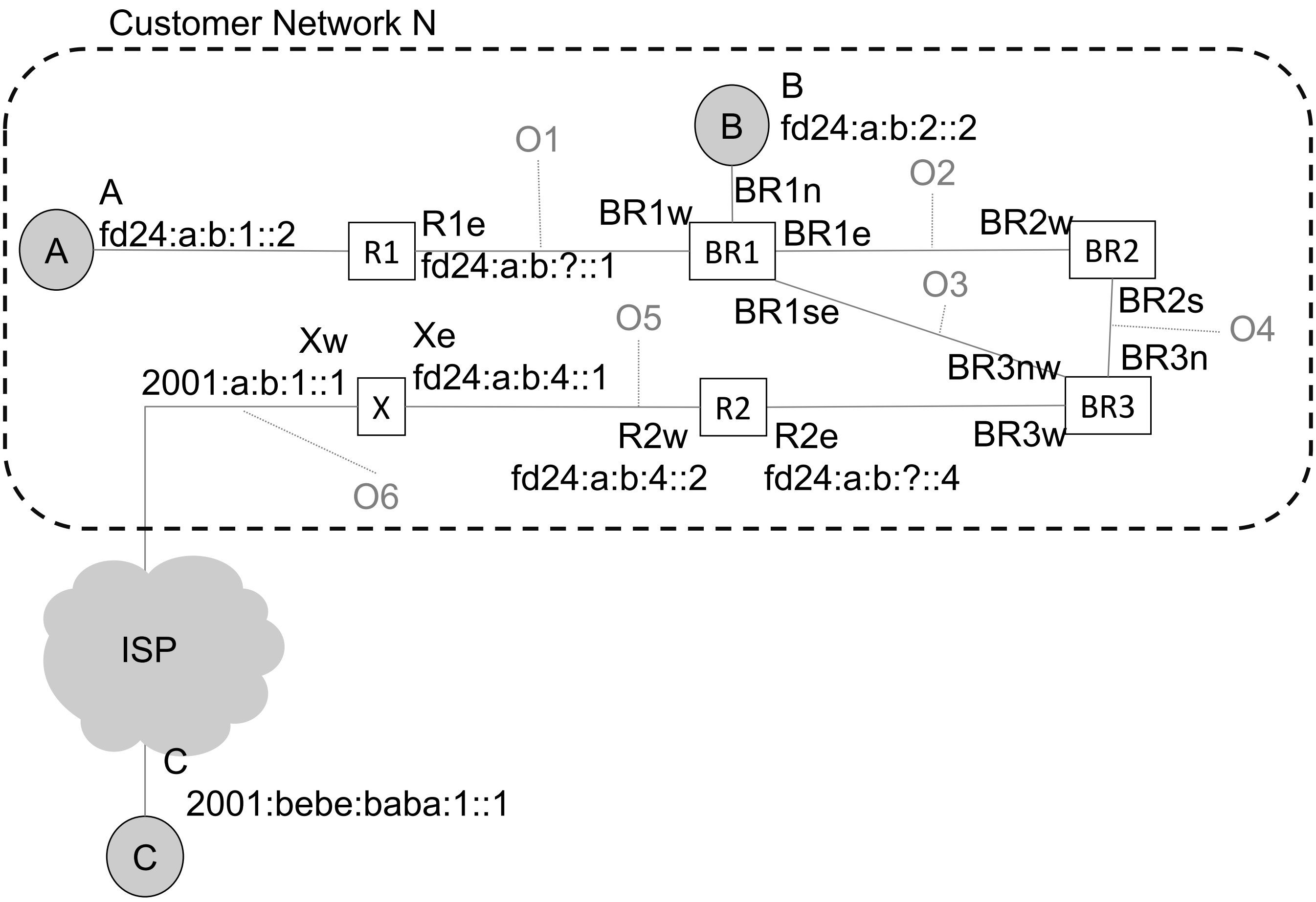
   Solution. Allocate several IP addresses to the same name `https://sovkom.coolstuff`, each address corresponding to one site. Put these IP addresses into DNS table. Put a short time-to-live in the DNS for these records, to ensure that they are not cached for too long (in case one of the sites crashes).

4. Joe has a dual stack IPv4 / IPv6 machine and has both IPv4 and IPv6 access. Joe connects to `https://lca.epfl.ch` and `https://infoscience.epfl.ch`. The web server `lca.epfl.ch` uses IPv6 only whereas the web server `infoscience.epfl.ch` uses IPv4 only. By which mechanism does Joe's machine know whether IP4 or IPv6 should be used ?

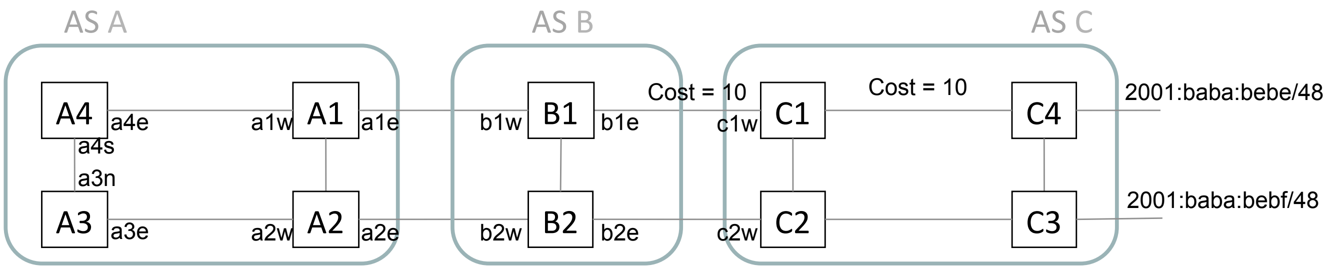
   Solution. This is done via DNS server, when name `lca.epfl.ch` and `infoscience.epfl.ch` mapped to IP addresses. Joe machine will receives an A record it uses IPv4. If it receives an AAAA record, it uses IPv6.

TCP IP EXAM - FIGURES

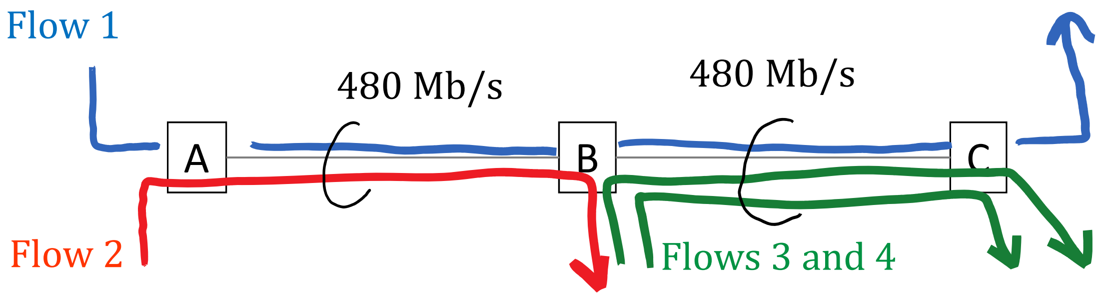
For your convenience, you can separate this sheet from the main document. Do not write your solution on this sheet, use only the main document. You do not need to return this sheet.



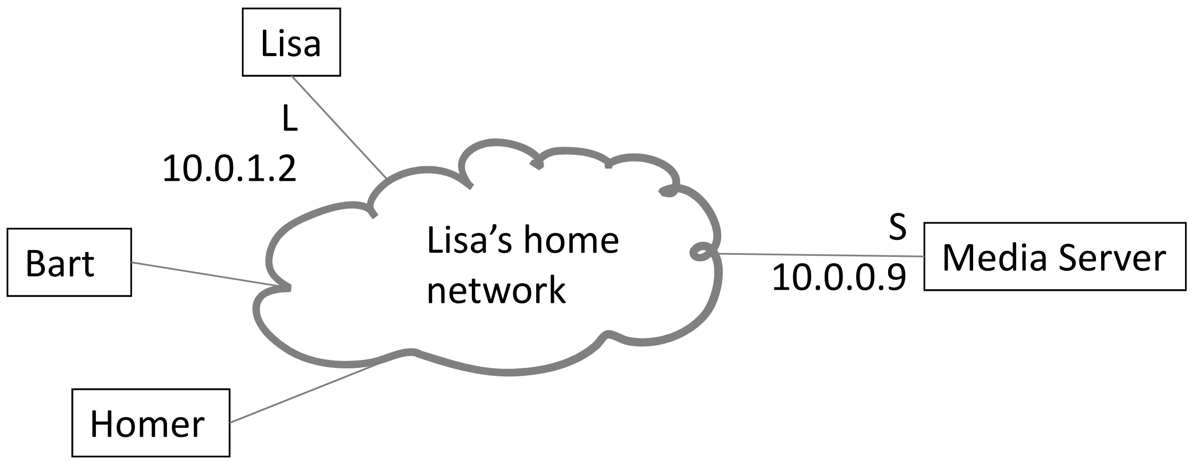
Problem 1



Problem 2



Problem 3



Problem 4, Question 2. L and S are MAC addresses.