ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

TIK Institut für Technische Informatik und Kommunikationsnetze

Department ITET
August 2017

Prof. Dr. Laurent Vanbever
Tobias Bühler, Rüdiger Birkner, Thomas Holterbach, Roland Meier

Exam: Communication Networks

9 August 2017, 09:00–11:30, Room HPH G 1

General Remarks:

- ▷ Write your **name** and your **ETH student number** below on this front page.
- ▷ Put your **legitimation card** on your desk.
- ▷ Check if you have received **all task sheets** (Pages **1 - 21**).
- ▷ Do **not separate** the **task sheets**.

- ▷ Write your answers directly on the task sheets.
- ▷ **All answers fit within the allocated space but often in much less.**
- ▷ If you need more space, please use your own extra sheets, in which case use a **new sheet of paper** for **each task** and write your name and the exam task number in the **upper right corner**.

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

- ▷ Please answer either in **English or German**.
- ▷ **Write clearly** in blue or black ink (not red) using a **pen**, not a pencil.
- ▷ **Cancel** invalid parts of your solutions **clearly**.
- ▷ At the end of the exam, hand your **solutions in together with all extra sheets**.

Special aids:

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

Family name: Student legi nr.:

First name: Signature:

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

Task	Points	Sig.
Ethernet & Switching	/20	
Intra-domain routing	/24	
Inter-domain routing	/42	
Reliable transport	/34	
Security & Applications	/30	
Total	/150	

Task 1: Ethernet & IP forwarding **20 Points**

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

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true	false	
☐	☐	The spanning tree protocol computes the shortest paths between any two switches in an Ethernet network.
☐	☐	Consider a switch willing to transmit some data on a link on which Carrier Sense Multiple Access/Collision Detection (CSMA/CD) is enabled. If the switch senses the link is busy, it will send a jamming signal and wait for the link to become available.
☐	☐	There can be only one router acting as gateway for the same IP subnet.
☐	☐	Consider hosts located in two different IP subnets connected by a router. Hosts located in one subnet would see the ARP requests sent by the hosts located in the other subnet (and vice-versa).
☐	☐	The IP address 8.0.1.0/255.0.0.0 identifies a network and as such cannot be assigned to an actual host.
☐	☐	Let $S1$ and $S2$ be the sets of IP addresses contained in two *distinct* subnets. If an IP address i is both in $S1$ and $S2$, then one of these two statements is *necessarily* true: $S1$ is a subset of $S2$ or $S2$ is a subset of $S1$.

b) Can your hear me now? **(4 Points)**

Consider two hosts (A and B) possessing a single network interface card connected to the same Ethernet switch. A's network interface is configured with 11.0.15.3/19 as IP address, while B's network interface is configured with 11.0.33.2/255.255.224.0 as IP address. Can a client (TCP-based) application running on A communicate with a server application running on B through the switch? Briefly explain why or why not.

c) Moving target **(10 Points)**

Consider the switched network depicted in Figure 1. It is composed of 5 Ethernet switches, two hosts (connected to switch 3 and 4, respectively) and one IP router acting as default gateway for the hosts. For redundancy reasons, the network exhibits cycles and each switch therefore runs the Spanning Tree Protocol (STP). All links have a unary cost. When equal-cost paths to the root are encountered, switches break the tie based on the sender ID (lower is better).

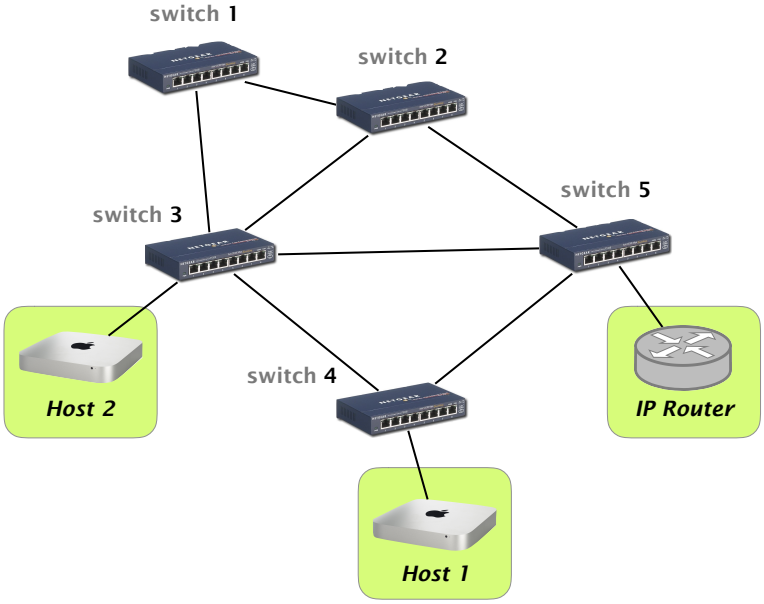


Figure 1: An Ethernet network running the spanning tree protocol.

(i) In the Figure 2 below, cross all the links that end up **deactivated** in the final state, once all the switches have converged on the final spanning tree. (2 Points)

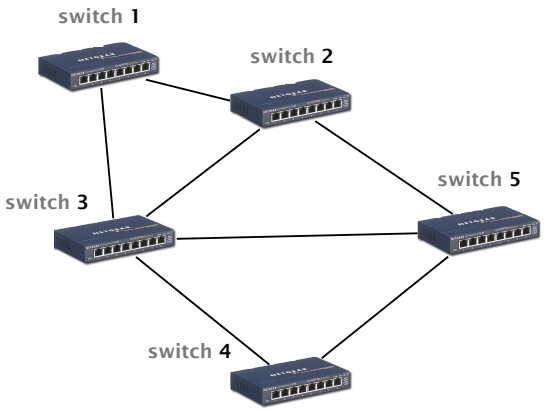


Figure 2: Cross the deactivated links.

(ii) Perhaps unsurprisingly, a *lot* of traffic is exchanged between Host 1 (resp. Host 2) and Internet destinations. Briefly explain **two distinct reasons** why this configuration is not optimal in terms of network utilization/throughput. (3 Points)

Reason 1:

Reason 2:

(iii) Realizing that there is a problem with their configuration, the network operators ask you (a fresh network engineer!) to help them improve their network performance. Briefly explain how you would adapt the configuration of the spanning tree protocol (i.e., the switches identifier and/or the link costs) so as to maximize the throughput between Host 1 (resp. Host 2) and Internet destinations. (2 Points)

(iv) The network operators are happy with your changes. But they now realize that Host 1 and Host 2, in addition to exchanging a lot of Internet traffic, also exchange a lot of traffic between themselves. The network operators ask for your help again! They ask you to find a spanning tree configuration such that: *(i)* the number of hops between any of these three hosts (Host 1 and 2, and the router) is equivalent; and *(ii)* the number of hops is minimum.

Briefly explain how you would configure the spanning tree protocol to achieve these requirements, or why these requirements are impossible to achieve. (3 Points)

Task 2: Intra-domain routing **24 Points**

a) Warm-up **(5 Points)**

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true	false	
☐	☐	To enable connectivity inside a network, an operator must use a link-state protocol or a distance-vector protocol.
☐	☐	Link-state routing protocols such as OSPF work by having routers flood their routing tables network-wide.
☐	☐	Link-state routing protocols such as OSPF do not suffer from the "count-to-infinity" problem.
☐	☐	OSPF routers verify that the forwarding entries they have computed are loop-free before updating their forwarding tables.
☐	☐	Any subpath of a shortest path is also necessarily shortest.

b) Weighing in **(3 Points)**

Consider the (currently unweighted) OSPF network depicted in Figure 3. The operators ask you to find weights such that each router has two paths of minimum length to reach each destination (so as to enable maximum load-balancing). As an illustration, the weights should be such that 1 knows two paths to reach 2, 3 and 4. Give the weight of each link, or briefly explain why it is not possible.

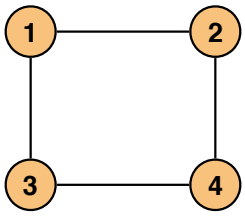
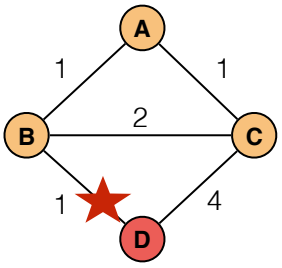


Figure 3: OSPF network with 4 routers.

c) Dying hard **(16 Points)**

Consider the topology depicted in Figure 4a where routers A, B, C and D are running RIP (initially, without poisoned reverse). The cost of a link is indicated next to it.



(a)

Step #0: Before the failure

Router A

NH	Cost
B	2*
C	4

Router B

NH	Cost
A	3
C	5
D	1*

Router C

NH	Cost
A	3*
B	3*
D	4

(b)

Figure 4: A simple RIP network undergoing an unfortunate failure.

In this question, we focus our attention on the routes towards router D while the routers converge upon the failure of (B, D). Figure 4b shows the routing tables of each router (towards D) before the failure. The best paths (used for forwarding) are indicated with a star. Observe that router C knows two minimum cost paths to reach D (via A and B, with a cost of 3) and uses them both for forwarding.

We will use Figure 5 to depict the evolution of the routing tables of each router during the convergence upon the failure of (B, D). We divide time into discrete steps. Each step corresponds to one router sending its current distance vector to its neighbors which then update their own vector accordingly. As an illustration, we pre-completed the first step corresponding to B sending its vector to A and C, and the resulting state reached on them. As in the course, we indicate with ∞ that a router cannot reach a destination via a specific next-hop.

(i) Compute the routing table of router A, B and C for destination D after each step of the convergence until the network has converged or up to the step #6. Consider that poisoned reverse is not used. Answer directly on Figure 5. (8 Points)

(ii) To speed-up the convergence time, the network operators decide to use RIP with poisoned reverse. Compute the routing table of router A, B and C for destination D **before the failure**. Answer directly on Figure 6. (2 Points)

(iii) Now consider the failure of link (B, D). Compute the routing table of router A, B and C for destination D after each step of the convergence until the network has converged or up to the step #6. Recall that poisoned reverse is now used. Answer directly on Figure 7. (6 Points)

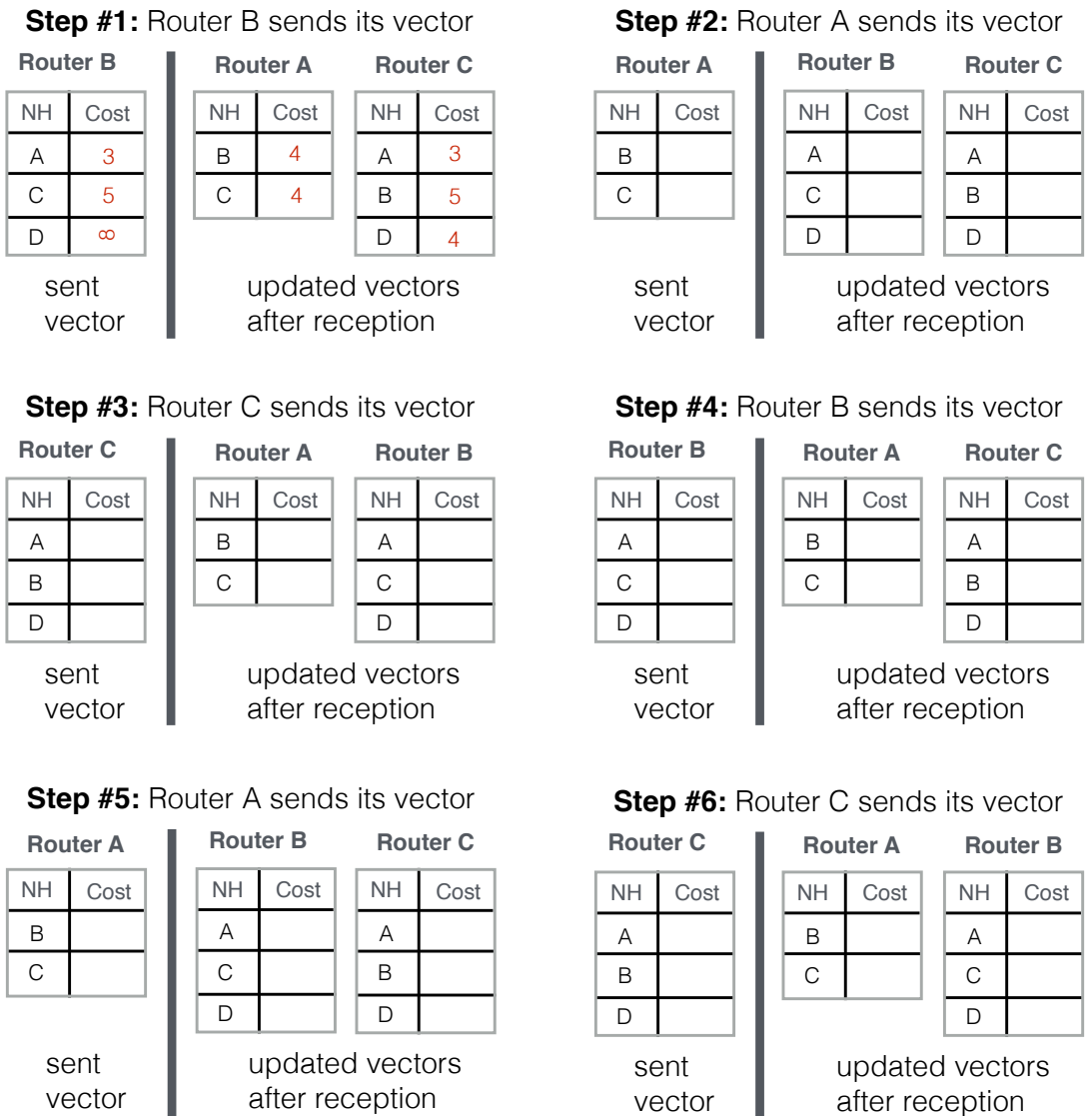
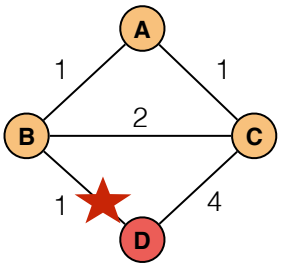


Figure 5: Indicate the routing table of router A, B and C for destination D at the end of each step of the convergence, **without poisoned reverse**.



(a)

Step #0: Before the failure

Router A		Router B		Router C	
NH	Cost	NH	Cost	NH	Cost
B		A		A	
C		C		B	
		D		D	

(b)

Figure 6: Indicate the routing table of router A, B and C for destination D **before the failure *and* considering poisoned reverse is used**.

Step #1: Router B sends its vector

Router B		Router A		Router C	
NH	Cost	NH	Cost	NH	Cost
A		B		A	
C		C		B	
D				D	
sent vector		updated vectors after reception			

Step #2: Router A sends its vector

Router A		Router B		Router C	
NH	Cost	NH	Cost	NH	Cost
B		A		A	
C		C		B	
		D		D	
sent vector		updated vectors after reception			

Step #3: Router C sends its vector

Router C		Router A		Router B	
NH	Cost	NH	Cost	NH	Cost
A		B		A	
B		C		C	
D				D	
sent vector		updated vectors after reception			

Step #4: Router B sends its vector

Router B		Router A		Router C	
NH	Cost	NH	Cost	NH	Cost
A		B		A	
C		C		B	
D				D	
sent vector		updated vectors after reception			

Step #5: Router A sends its vector

Router A		Router B		Router C	
NH	Cost	NH	Cost	NH	Cost
B		A		A	
C		C		B	
		D		D	
sent vector		updated vectors after reception			

Step #6: Router C sends its vector

Router C		Router A		Router B	
NH	Cost	NH	Cost	NH	Cost
A		B		A	
B		C		C	
D				D	
sent vector		updated vectors after reception			

Figure 7: Indicate the routing table of router A, B and C for destination D at the end of each step of the convergence, **with poisoned reverse**.

Task 3: Inter-domain routing **42 Points**

a) Warm-up **(9 Points)**

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true ☐ false ☐ In the classical BGP selection and exportation policies (with providers, peers, customers), an Autonomous System (AS) will never announce a route received from a provider to another provider.

true ☐ false ☐ An AS has full control over its outgoing traffic.

true ☐ false ☐ The forwarding table of a BGP router contains all routes received from its BGP peers whereas the routing table only contains the BGP best path.

true ☐ false ☐ Tier-1s only have Tier-2s as customers.

Consider the simple BGP network in Figure 8. Single-headed plain arrows point from providers to their customers (AS A is the provider of AS D), while double-headed dashed arrows connect peers (AS D and AS E are peers). Each AS in the network originates a unique prefix that it advertises to all its BGP neighbors. Each AS also applies the default selection and exportation BGP policies based on their customers, peers and providers.

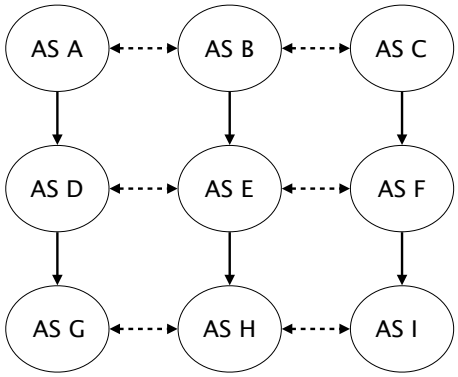


Figure 8: A simple BGP network.

true ☐ false ☐ The path [G, D, A, B, E, H] from AS G to AS H is valid.

true ☐ false ☐ AS A receives at least one route traversing the link between AS C and AS F.

true ☐ false ☐ AS A's best route to reach AS I has an AS-PATH length of 4.

true ☐ false ☐ AS D uses the path [D, E, H] to reach AS H.

true ☐ false ☐ AS H uses the path [H, I, F] to reach AS F.

b) Primary *vs* backup **(15 Points)**

Consider the two neighboring ASes in Figure 9. Swisscom is a customer of Deutsche Telekom and they interconnect in two locations: Zürich and Geneva. Swisscom announces the same route for the IP prefix 185.105.144.0/23 to Deutsche Telekom at both locations. Deutsche Telekom is using the standard BGP decision process and has no special policies in place.

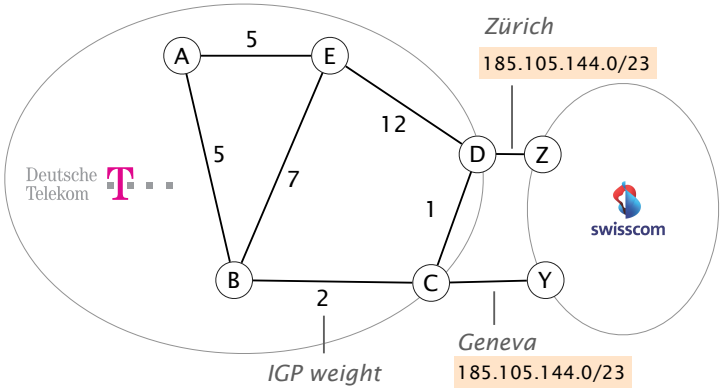


Figure 9: Swisscom and Deutsche Telekom.

(i) Given the network in Figure 9, indicate through which location (Zürich or Geneva) each router in Deutsche Telekom sends its traffic to Swisscom. (2 Points)

A: _____ B: _____

C: _____ D: _____

E: _____

(ii) It turns out that the border router Swisscom is using in Zürich is much more powerful than the one in Geneva. Hence, Swisscom would like to modify its configuration such that it receives the traffic originating from Deutsche Telekom in Zürich, unless there is a failure. One of Swisscom's network operators asks you for advice on how to achieve this by tweaking the MED values in Swisscom's announcements. Explain the meaning and use of the MED value. (3 Points)

(iii) Swisscom engineers configure their routers to advertise different MED values: 10 in Zürich and 20 in Geneva. Which locations (Zürich or Geneva) do the routers use now?
(2 Points)

A: _____ B: _____

C: _____ D: _____

E: _____

(iv) Deutsche Telekom noticed the change and is not happy with it. It turns out that for Deutsche Telekom, router C is much more powerful than D. As such, it wants to shift back all the traffic to Geneva. How could Deutsche Telekom achieve this? Explain briefly.
(2 Points)

(v) Deutsche Telekom successfully shifts all traffic to Geneva. Explain two **different approaches** for Swisscom to prevent that and mention any limitations/drawbacks of them.
(6 Points)

Approach 1: _____

Approach 2: _____

c) Left? Right? Both? (18 Points)

Consider the BGP network composed of 4 routers depicted in Figure 10. Two of these routers, $R1$ and $R4$ are egress routers and maintain eBGP sessions with external neighbors. $R1$ is configured to associate a local-preference of 100 to externally-learned routes, while $R4$ is configured to associate a local-preference of 200 to externally-learned routes. $R2$ and $R3$ are internal routers. All four routers are connected in an iBGP full-mesh. OSPF is used as intra-domain routing protocol. The link weights are indicated in the figure, e.g. the $(R1, R2)$ link is configured with a weight of 20. Figure 10 also indicates the propagation delay for each link (e.g., it takes 5ms for a packet to propagate between $R1$ and $R2$).

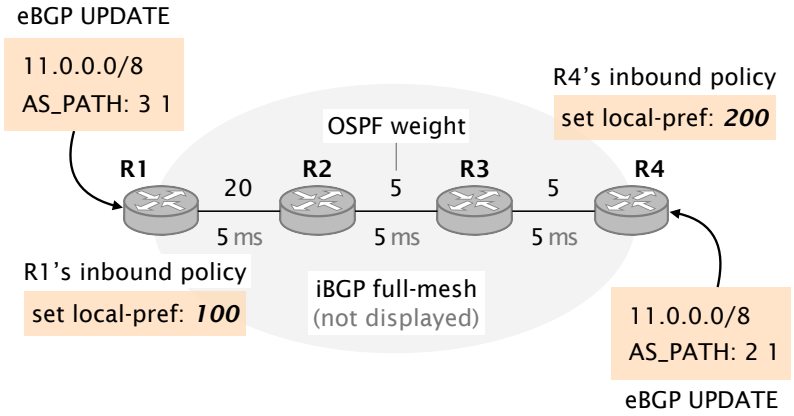


Figure 10: A simple BGP network learning external routes via eBGP on $R1$ and $R4$.

(i) Considering the above configuration, indicate the next-hop used by each router in the steady state, *i.e.,* once the network has fully converged. Use the keyword "external" to indicate that an edge router is forwarding outside of the domain. (2 Points)

R1: _____ R2: _____

R3: _____ R4: _____

(ii) One of the network operator decides to lower the local-preference associated by $R4$ to externally-learned routes to 50 (instead of the original 200). Indicate the sequence of BGP messages sent which is triggered following that change along with the timestamps at which they are generated. You can consider that the BGP process on each router is infinitely fast meaning only propagation delay matters. Only indicate when messages are sent, not when messages are received. (4 Points)

Use this template to answer (replace the content within the square brackets):
Timestamp [YY ms] [RX] sends the message [$msg\_content$] to [RA, RB, and RC]

(iii) Was a forwarding loop induced due to the configuration change? Briefly explain why or why not. If a loop was created, also indicate its duration (in ms). (3 Points)

(iv) It turns out that the network operator changed her mind. This time, she configures $R4$ to associate a local-preference of 100 to externally-learned routes (i.e. the same local-preference value as on $R1$). Indicate the next-hop used by each router in the steady state (once the network has fully converged). Again use the keyword "external" to indicate that an egress router is forwarding outside of the domain. (2 Points)

R1: _____ R2: _____

R3: _____ R4: _____

(v) Soon after the network has fully converged due to the configuration change of $R4$, a failure happens disconnecting $R4$ from all its external neighbors. The connection between $R4$ and $R3$ is still working fine though. Indicate the sequence of BGP messages sent following that failure along with the timestamps at which they are generated. Only indicate when messages are sent, not when messages are received. (4 Points)

Use this template to answer (replace the content within the square brackets):
Timestamp [YY ms] [RX] sends the message [$msg\_content$] to [RA, RB, and RC]

(vi) Was a forwarding loop induced due to the failure? Briefly explain why or why not. If a loop was created, also indicate its duration (in ms). (3 Points)

Task 4: Reliable Transport **34 Points**

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

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true	false	
☐	☐	The content of the TCP header is independent of the network layer technology used below (i.e. IPv4 or IPv6).
☐	☐	It is not possible to reliably transmit data over the Internet using UDP as a transport protocol.
☐	☐	In a loss-free and congestion-free network, sending small amounts of data with UDP is faster than with TCP.
☐	☐	Consider a TCP connection. Doubling the sender and receiver windows will double the observed bandwidth.
☐	☐	In a sliding window protocol with cumulative ACKs, a new ACK (observed for the first time) will always move the sender window.
☐	☐	In a GBN protocol with SACK (Selective ACKnowledgments), the segments indicated by the receiver in the SACK header (blocks of correctly received out-or-order segments) cannot be directly removed from the sender buffer.

b) Go Back **(15 Points)**

Figure 11 shows the beginning of a Go-Back-N (GBN) time-sequence diagram. Here is a **non-exhaustive** list of implementation choices made for the GBN sender and receiver implementation. **Read them carefully.**

- The sender and receiver window have a size 4;
- The receiver saves out-of-order segments in an infinite buffer and removes them as soon as the missing segment(s) arrive;
- The receiver uses cumulative ACKs which acknowledge all previous segments and point to the next expected data segment;
- The sender uses Fast Retransmit. After three duplicate ACKs, the sender immediately retransmits the corresponding data segment. For instance, if the sender gets the following ACKs [A1, A1, A1], it will immediately retransmit the data segment D1;
- For each tick in the diagram below, the sender can send one data segment and the receiver can send one ACK. Sender and receiver will first analyze the incoming packet and then send a data segment/ACK;
- The sender uses a retransmission timer of 5 ticks. Each time it sends a data segment or receives an ACK, the timer is reset. After a timeout, the sender retransmits all current segments in its sender buffer (in order, one segment per tick);
- A data segment or ACK needs two ticks to travel to the other end of the connection. See the given start in the diagram.

Your task: Draw the successful transmission of 6 data segments (D0 to D5) if the **first data segment** (D1, already indicated) **is lost** as well as **ACK A5 is lost the first time it is sent**. For each tick, indicate which data segment or ACK is transmitted (if any) as well as the content of the sender and out-of-order buffer.

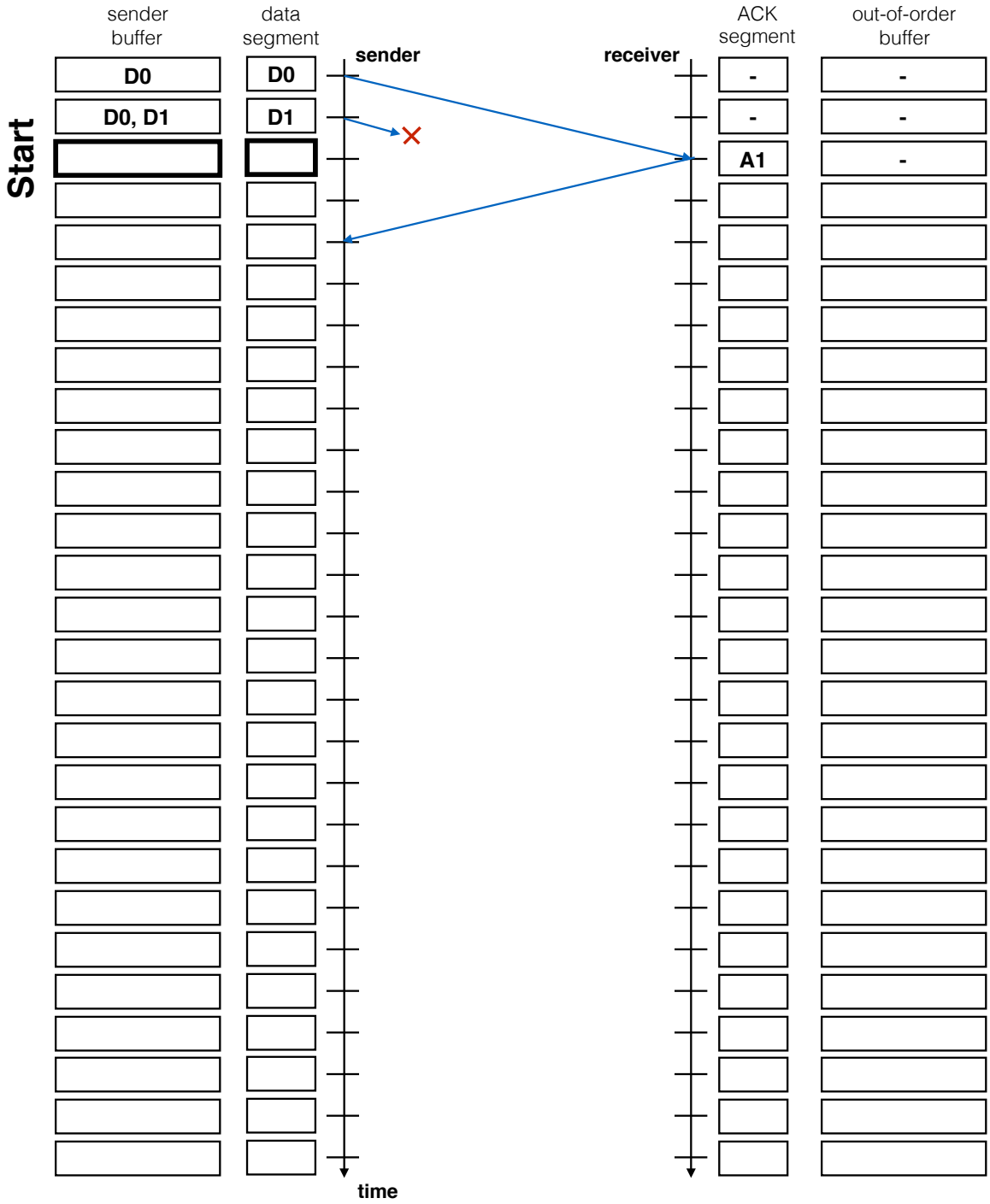


Figure 11: Time-sequence diagram of a GBN protocol with Fast Retransmit and cumulative ACKs.

c) Keeping counts **(4 Points)**

Say that a TCP connection with a Maximum Segment Size (MSS) of 1000 bytes sees a congestion window of 8000 bytes. What is the size of the congestion window (in bytes) after the connection has sent out 4 extra packets and received acknowledgments for two of them?

If the connection is in slow-start: _____

If the connection is in congestion avoidance: _____

d) Dissecting TCP connections **(9 Points)**

Consider Figure 12 which depicts the evolution of the size of the TCP congestion window of the sender after each transmission round.

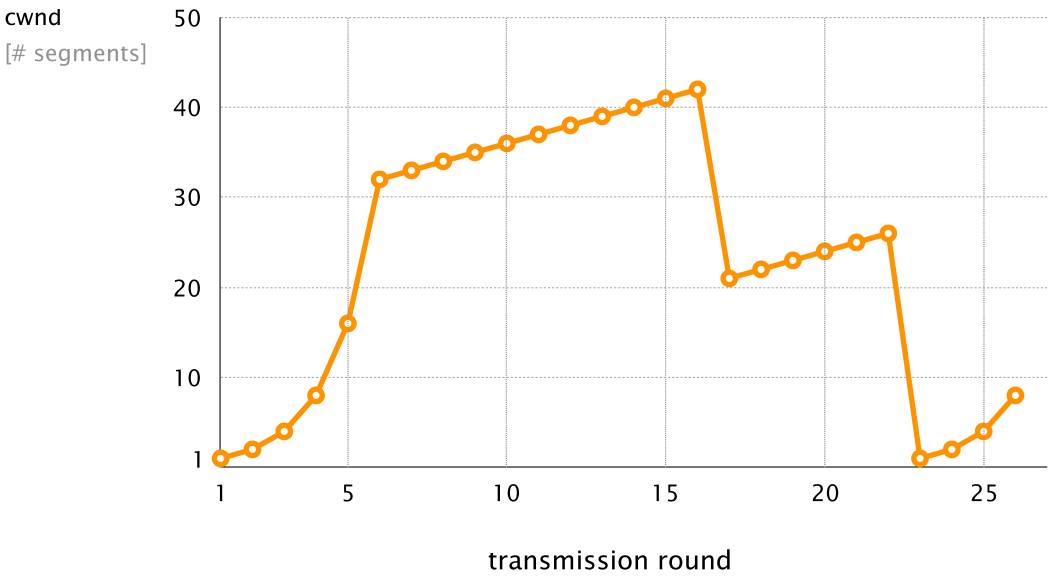


Figure 12: Evolution of the size of the congestion window.

(i) List all the intervals during which slow-start is operating. (2 Points)

(ii) List all the intervals during which congestion avoidance is operating. (2 Points)

(iii) Explain what happens after the 16th transmission round. (2 Points)

(iv) During what transmission round is the 100th segment sent? Briefly explain. (3 Points)

Task 5: Security & Applications **30 Points**

a) Warm-up **(8 Points)**

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true	false	
☐	☐	Resolving the IP address for the website `abc.def.ghi.jkl.ch` involves queries to at least 5 DNS servers (assuming all caches are empty).
☐	☐	ICMP can be used to determine properties of the network such as round trip times (RTTs), maximum transmission units (MTUs) and the size of the network.
☐	☐	Network address translation (NAT) is required for enabling one HTTP server (with a single public IP address) to host multiple websites.
☐	☐	NAT uses the transport protocol source and destination ports to identify which connections belong to which host. This means that if one host originates multiple connections with the same (source port, destination port) to different IP addresses, NAT does not work correctly. However, this rarely occurs because the source port is chosen randomly.
☐	☐	Google search would not list a website that is hosted on a server with IP address 10.20.30.40 unless this server is also reachable by another IP address.
☐	☐	If an attacker manages to turn off all DNS root servers, both `http://ethz.ch` and `http://comm-net.ethz.ch` will eventually become unreachable outside of ETH.
☐	☐	Thanks to the checksum in the IPv4 header, the receiver can make sure that the packet wasn't modified by a Man-In-The-Middle attacker.
☐	☐	Consider two neighboring ASes, AS1 and AS2, that physically peer with each other in one location and that AS2 is using OSPF internally. A malicious operator in AS1 could attract traffic away from AS2 by establishing an OSPF adjacency between the two border routers.

b) BGP, this time, with a security twist **(10 Points)**

Consider the Internet topology composed of 12 ASes depicted in Figure 13. Single-headed plain arrows point from providers to their customers (AS A is the provider of AS C), while double-headed dashed arrows connect peers (AS A and AS B are peers). AS F advertises a single prefix: 12.34.58.0/22 to all its neighbors. All ASes apply the default selection and exportation BGP policies based on their customers, peers and providers.

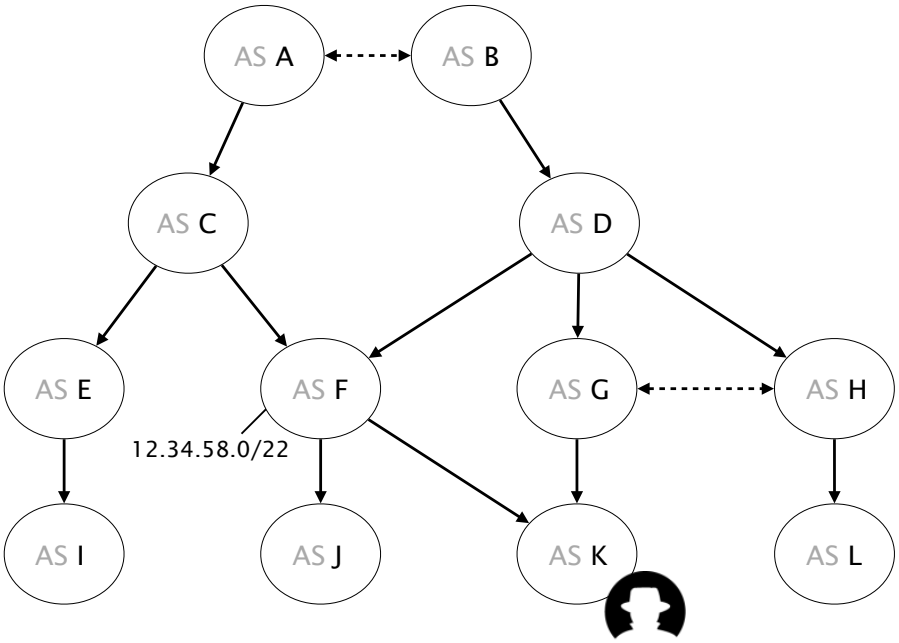


Figure 13: A simple Internet, with a malicious AS.

(i) Assume that AS K is malicious and wants to attract traffic destined to 12.34.58.0/22. Knowing that BGP is purely based on trust, it decides to see how much traffic it can attract by advertising the exact same prefix (12.34.58.0/22) to all its neighbors. List all the ASes for which it manages to divert traffic from. (2 Points)

(ii) AS K advertises more-specific prefixes (12.34.56.0/23 and 12.34.58.0/23) to try to attract more traffic. List all the ASes for which it manages to divert traffic from and explain why there is a difference (or why there is no difference) between announcing the two /23 prefixes and announcing the /22 prefix. (4 Points)

(iii) Assume all non-malicious ASes meet after this attack and agree to create a central database mapping an AS number to the IP prefixes it owns and have their routers systematically query that database before considering any advertisement. That is, whenever an AS receives a route to a prefix P, it first checks that the last AS in the route indeed owns P. For example, upon receiving a path to 12.34.58.0/22, an AS will check that the last AS in the route is F. With this system in place, can AS K (the only malicious AS) still attract traffic for IP address blocks belonging to F? Briefly explain. (4 Points)

c) Back to the (DNS) roots (2 Points)

If you were responsible for the root DNS servers a.root-servers.net, would you configure them to operate in recursive mode or in iterative mode? Explain why.

d) Loading a website (10 Points)

Consider the website hosted at `https://www.your-shop.ch` with the following elements:

- HTML	`https://www.your-shop.ch/index.html`
- Stylesheet	`http://www.your-shop.ch/style.css`
- Image	`http://your-shop.ch/logo.png`
- Image	`http://images.your-shop.ch/product.jpg`
- Facebook like button	`http://static.facebook.com/like.png`
- Facebook "tracking" code	`https://www.facebook.com/track.js`
- Google "tracking" code	`https://www.google.com/track.js`

(i) Assuming that your host is configured to use a local recursive DNS server in your network and all caches are empty. List all the DNS queries that your host sends to this DNS server when you open up `https://www.your-shop.ch/` in your favorite browser. (3 Points)

(ii) After loading the website, you send an email to `contact@your-shop.ch` via a mail server that uses the same DNS server as your host. Does the local recursive DNS server need to run additional queries to other DNS servers if it has all the replies from the queries in the previous task in its cache? Explain why or why not. (2 Points)

(iii) How many TCP connections would an unoptimized browser (also referred to as "naive" in the lecture) open to load `https://www.your-shop.ch`? Briefly explain your answer. (1 Point)

(iv) During your holidays in Australia, you realize that the Facebook like button loads much faster than the logo of the shop even though both images have the same size. Can you explain the reason for this and why you do not observe this behavior in Switzerland? (2 Points)

(v) One hour later (still in Australia), you open the shop's website again. This time, the logo of the shop and the Facebook button appear at the same time. Explain **two distinct reasons** that would justify this behavior. (2 Points)

Reason 1:

Reason 2: