



## **Communication Networks**

Prof. Laurent Vanbever

Exercise 99 - Additional Practice

# Link State and Distance Vector

99.1 Warm-up Questions (Exam Question 2016)

For the following statements, decide if they are *true* or *false*. Motivate your decision.

- **a)** Consider a positively weighted graph *G*. Applying the Bellman-Ford (used by distance-vector protocols) or Dijkstra (used by link-state protocols) algorithm on *G* would lead to the same forwarding state.
- **b)** Link-state protocols (such as OSPF) are guaranteed to compute loop-free forwarding state as long as the link-state databases are consistent on all routers.
- **c)** Link-state protocols (such as OSPF) require routers to maintain less state than distance-vector protocols (such as RIP).
- **d)** Poisoned reverse solves the problem of count-to-infinity.
- **e)** Consider a positively weighted graph *G*. Multiplying all link weights by 2 would change the all-pairs shortest paths computed by the Dijkstra algorithm on *G*.
- **f)** Consider a positively weighted graph *G*. Adding 1 to all link weights would change the all-pairs shortest paths computed by the Dijkstra algorithm on *G*.





Dijkstra's algorithm with a link failure

The routers in the network on the top left use Dijkstra's Algorithm to find the shortest path towards destination d. You can assume that every router knows the entire network graph. In case of a failure, the routers directly affected by it inform the other routers by flooding this information in the network.

The blue arrows indicate how the routers forward the traffic:  $R_2$ , for example, sends packets for destination d to router  $R_3$ .

Now the link between router R3 and d fails (network at the bottom left) and R3 can no longer send packets towards destination d. As R3 is directly connected to the failed link, it detects the failure immediately. It starts to flood this information, such that all the routers can update their network view and recompute Dijkstra's algorithm.

- a) What is the new shortest-path from *R*3 towards destination *d*?
- **b)** Assume now that the computation of the new shortestpath is *very* fast and finishes before *R*3 starts the flooding of the messages announcing the link failure. *R*3 sends a packet towards *d* using the new shortest-path. Will the packet reach its destination? Which path will it take?
- c) Can you find a sequence of link failure messages and shortest-path computations such that the problem discovered in the previous task is observed between *R*1 and *R*2? The *only* link failure is still between router *R*3 and *d*.
- **d)** As we have discovered, the order in which the link failure messages are processed and the new shortest-paths are computed is crucial for a correct forwarding behavior in the network. In which order should the router update their forwarding tables, such that the previously observed problems will not occur? Can you find a more general "rule" for a save ordering of forwarding rule updates?

### 99.3 A very long downtime (Exam Question 2019)

Consider the network in the Figure below which connects a large cluster of 500k servers (right) with a large pool of clients (left). The network uses OSPF internally. Each of the 500k servers is located in a different IP subnet, while all of the clients are located in the same /8 subnet.



Consider that the link between router A and C fails. As router A is adjacent to the failure, it immediately detects it and starts rerouting the traffic for the 500k prefixes to B. The Figure below reports the evolution of the throughput observed for 3 of the 500k server prefixes: 23.0.0./24, 121.0.0./24, and 202.0.0./24. One can see that the downtime experienced by each prefix is different: 23.0.0./24 experiences less downtime than 121.0.0./24, which itself experiences less downtime than 202.0.0./24.



a) Explain why the downtimes experienced by the three prefixes differ.

**b**) Describe two solutions to speed up the convergence for these three prefixes.

# **BGP** Routing and Security

### 99.4 Putting Everything Together (Exam Question 2016)

Consider the ISP network composed of 5 routers (*A*, *B*, *C*, *D*, *E*) depicted in the Figure below. Three of these routers, *A*, *E* and *D*, are connected to routers located in neighboring ASes via eBGP. These neighboring routers are indicated by *X*, *Y* and *Z*. Each of them advertises the same three distinct IP prefixes p1, p2 and p3.

The three tables in the Figure indicate the Local-Preference (LP) associated to each external prefix by A, E and D along with their corresponding AS-PATH length. For instance, A learns a route to p1 from X with an AS-PATH length of 10 to which it associates a LP of 200. Internally, the ISP uses an iBGP full-mesh to distribute the BGP routes and OSPF as intra-domain routing protocol. The weight of each internal link is indicated next to it.



An ISP network which receives BGP routes for 3 external prefixes (p1, p2, p3) from 3 routers (X, Y, Z) in neighboring ASes.

For each router in the ISP, indicate the router ID of the selected egress (A, E, D) along with the router ID of the internal next-hop (A, B, C, D, E or *direct*) used to reach it. For that you can use the tables on the next page. You can assume that A, E and D use the next-hop-self configuration.



Fill in the following tables with the selected egress and internal next-hop.



Where are my packets going?

After passing the Communication Networks exam with flying colors, ETH hires you as a junior network engineer. *Congrats*!

Your first mission is to analyze their BGP configuration. They indeed suspect that something might be wrong, especially since they installed this box from Sisco Systems that automatically configure BGP announcements according to Traffic Engineering objectives. For the sake of simplicity, assume again that ETH has only one prefix: 82.130.64.0/21 and three providers: Swisscom, Deutsche Telekom and Sunrise. The actual announcements are depicted on the left. Customers are drawn below their providers (Swisscom is a customer of Deutsche Telekom), while peers are drawn next to each other (Swisccom is a peer of Sunrise).

Consider the incoming traffic from Swisscom. What path is taken for packets destined to:

- a) 82.130.66.1?
- **b)** 82.130.69.1?
- c) 82.130.70.1?

Are Swisscom, Deutsche Telekom and Sunrise happy about these announcements? Can they do anything about that? Explain briefly.

As the *polyterasse* router is getting older, its reliability starts to suffer. In theory, this should not be a big problem as ETH is *triple*-homed! Yet, in practice, the ETH engineers observe regular network connectivity upon failures.

Assuming the same announcements, what path ends up being taken by the packets destined to the above three IP addresses when:

- a) the link between *polyterasse* and Swisscom goes down?(i) 82.130.66.1?
  - (ii) 82.130.69.1?

(iii) 82.130.70.1?

- b) the *polyterasse* router dies?(i) 82.130.66.1?
  - (ii) 82.130.69.1?
  - (iii) 82.130.70.1?

What would you change in the ETH announcements to improve reliability, without disturbing the inbound Traffic Engineering performed by the Sisco box in the steady case (without failures)? Explain briefly.



Is this network as redundant as it looks?

#### **99.6** BGP Hijack (Exam Question 2018)

Consider the Internet topology consisting of 11 Autonomous Systems (ASes) in the Figure below. Single-headed plain arrows point from providers to their customers (AS *A* is the provider of AS *E*), while double-headed dashed arrows connect peers (AS *A* and AS *B* are peers). Each AS is made up of a single BGP router and applies the default selection and exportation BGP policies based on their customers, peers and providers.

In this task, the routers break ties using the AS number of the neighbor: in case multiple routes are equally good, the router selects the route of the neighbor with the lowest AS number (alphabetical order).

AS *G* is the origin of prefix 20.0.0.0/22 and advertises it to its neighbors. Independently of what the external advertisements are, AS *G* **always** prefers its internal route to reach any IP destination in 20.0.0.0/22.



An Internet topology with 11 ASes. AS K aims at hijacking traffic destined to AS G.

- **a)** AS *K* wants to hijack all the traffic going to AS *G* for 20.0.0/22. It starts advertising the exact same prefix. From which ASes is it able to attract the traffic?
- **b)** AS *K* is not satisfied by the result. What can it do to attract traffic destined to AS *G* from more of the ASes? List the ASes from which it is able to attract the traffic and explain why this works.
- c) The ASes from which AS *K* manages to attract the traffic realize what is happening as all their traffic to 20.0.0/22 goes to a dead-end (AS *K*).

Show how AS K could still deliver the traffic to the real destination (AS G) by poisoning the AS path while attracting as much traffic as possible. In addition, list the ASes from which it can attract the traffic.

**d)** Can you think of a different way for AS *K* to achieve similar results as in *c*) without poisoning the AS path? Explain.

# **Reliable Transport**

## 99.7 Go-Back-N Warm-Up Questions

Sender and receiver keep separate windows and buffers for sent and received segments.

- a) Compare how the sender and the receiver advance their respective windows.
- b) Which segments does the *sender* buffer? When can segments be removed from the buffer?
- c) A *receiver* typically buffers out-of-order segments. What is the advantage of such a buffer?

*Cumulative ACKs* acknowledge that all segments *up to* the acknowledged segment have been received.

d) Why are cumulative ACKs used? Do they help with lost data segments, lost ACKs, or both?

When *Fast Retransmit* is used, the sender retransmits a segment after duplicate ACKs.

- e) How does Fast Retransmit improve performance?
- f) Compare Fast Retransmit in the case of mild congestion (some segment losses) and heavy congestion (nearly all segments lost).

Consider a Go-Back-N (GBN) protocol with the following implementation choices for sender and receiver.

- The sender and receiver window have a size of 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 segmen-t/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: Use the diagram on the next page to 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.



## IP, DHCP, ARP

## 99.9 IPv6 Calculations

IPv6 addresses have a slightly different notation. Because the addresses are 128 bit long, we switch from a decimal notation to a hexadecimal one. In general, IPv6 addresses are represented by eight colon-separated blocks of up to four hexadecimal digits each. In a block, leading zeros can be omitted. Furthermore, we can use the "::" symbol to compress one or more consecutive zero blocks. However, the "::" symbol can only be used once in a single IPv6 address. As an example, the IPv6 address: 2001:0db8:0000:0000:0000:ff00:0042:8329 can be simplified to 2001:db8::ff00:42:8329.

- a) You are the operator of an enterprise network. Your ISP is giving you a /96 subnet 2001::/96. How many addresses do you have available? Is that a reasonable subnet size compared with the currently available IPv4 addresses?
- **b)** In your enterprise network, each host machine is identified by a unique ID starting from 1. Now that you have a lot of IPv6 addresses available, you decide to give each host a unique IP address. The host with ID 1 gets the first IPv6 address in your subnet (2001::1), the host with ID 2 the second IP address and so on. Complete the following table:

IPv6 address	host ID
	5
	14
2001::3A5	
	4 294 967 295
2001::3:0	



Describe everything that happens to the packets sent between the client and the server

Consider the network on the left composed of three Ethernet segments separated by two intermediate routers (A and B). In this network, the server's interface along with the routers' interfaces are configured with static IP addresses. While clients connected to the 3.0.0.0/24 Ethernet segment obtain an IP address via DHCP.

Assuming that the client has just started, with a perfectly empty state, precisely describe all packets that are generated when the command "ping 1.0.0.2" is issued (until the server answers back). Among others, your answer *must* include the content of the Layer 2 and Layer 3 headers.

#### **99.11** Detective work

You just started your first job as a network operator of a small network. To get more familiar with the network, you look at a packet trace captured at a switch. The trace contains packets from multiple hosts and one router connected by a (layer 2) switch. The router acts as default gateway, providing access to the Internet and is assigned the first IP address in the subnet. Each row in the following table represents one packet observed at the switch.

SRC MAC Address	DST MAC Address	SRC IP Address	DST IP Address	
6a:00:02:49:a1:a0	11:05:ab:59:bb:02	65.222.11.1	65.222.8.2	
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.16.1	
da:15:00:00:01:11	11:05:ab:59:bb:02	129.132.103.40	65.222.8.2	
11:05:ab:59:bb:02	40:34:00:7a:00:01	65.222.8.2	65.222.15.254	
11:05:ab:59:bb:02	ac:00:0a:aa:10:05	65.222.8.2	65.222.9.99	
ac:00:0a:aa:10:05	01:05:3c:34:00:02	65.222.9.99	65.222.13.255	
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.8.1	

- a) Can you identify all the hosts that are part of the local network?
- **b)** Can you reconstruct the IP subnet used to address the hosts within that local network?

### 99.12 Changing addresses (Exam Question 2019)

Consider the network depicted in the Figure below which is composed of two hosts along with two routers, one of which acts as Network Address Translator (NAT). Host 1 is located in a private subnet (192.168.1.0/24) and uses 192.168.1.1 as gateway, while host 2 is located in a public subnet (81.0.0.0/24) and uses 81.0.0.1 as gateway. The Figure below also depicts the MAC address of each of the 6 interfaces connected at either end of the three links. The NAT/router performs address translation between the private and the public subnets, translating traffic originating from private IPs to its public one (here, 20.0.0.1), and vice-versa.



A network topology relying on Network Address Translation.

a) Consider that host 1 tries to open a TCP connection with host 2 on port 80 using 1337 as (random) source port. Write down a possible sequence of packet headers observed at each link for the first two packets (i.e., the SYN sent by host 1, and the SYN/ACK sent by host 2). Fill in the table below to answer. Assume that hosts and routers have the required MAC addresses in their ARP table.

	src MAC	dst MAC	src IP	dst IP	src TCP port	dst TCP port
link A						
link B						
link C						
link C						
link B						
link A						

**b)** Could host 2 initiate a TCP connection to host 1? Briefly explain why/why not.

## DNS

### 99.13 Curious students

Consider that ITET has a local DNS server serving the DNS requests for all students' devices connected in the department. How could you determine if an external website has been visited recently by a fellow colleague of yours? Explain.

#### 99.14 Multiple answers

Whenever a client (e.g., your computer) receives multiple IP addresses as answer to a DNS lookup, it picks the very first one. Only if that one does not work, it tries the next one in order.

When you run dig yahoo.com, you receive multiple IP addresses as an answer compared to, for example, dig google.com.

Can you think of a reason for providing multiple IP addresses? Run the lookup for yahoo.com multiple times.

### 99.15 Local vs. authoritative DNS server

Perform a DNS query for uzh.ch using first the authoritative DNS server (ns1.uzh.ch) and then your local server.

Note: When using nslookup on Windows, you need to specify the -debug flag to get the relevant information for this task. For example:

nslookup -debug <Domain Name> <DNS-Server>

- Compare the ANSWER SECTION of the responses. Can you see differences between the answers from your local DNS server and the authoritative server? Run the query to your local server multiple times to make the differences more obvious.
- What is the reason for this difference?
- As you have seen in the lecture, DNS can be used to balance the incoming load. What are the considerations one has to make when using DNS load balancing with respect to the TTL?

## HTTP (and Web)

### **99.16** Loading a website (Exam Question 2017)

The website www.your-shop.ch consists of the following elements:

- HTML www.your-shop.ch/index.html
- Stylesheet www.your-shop.ch/style.css
- image www.your-shop.ch/logo.png
- image images.your-shop.ch/product.jpg
- Facebook like button cdn.facebook.com/like.png
- Facebook "tracking" code www.facebook.com/track.js
- Google "tracking" code www.google.com/track.js
- a) 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.
- **b)** 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.
- c) 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.
- **d)** 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?
- e) 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.