ETH zürich

Networked Systems
ETH Zürich — seit 2015

Communication Networks

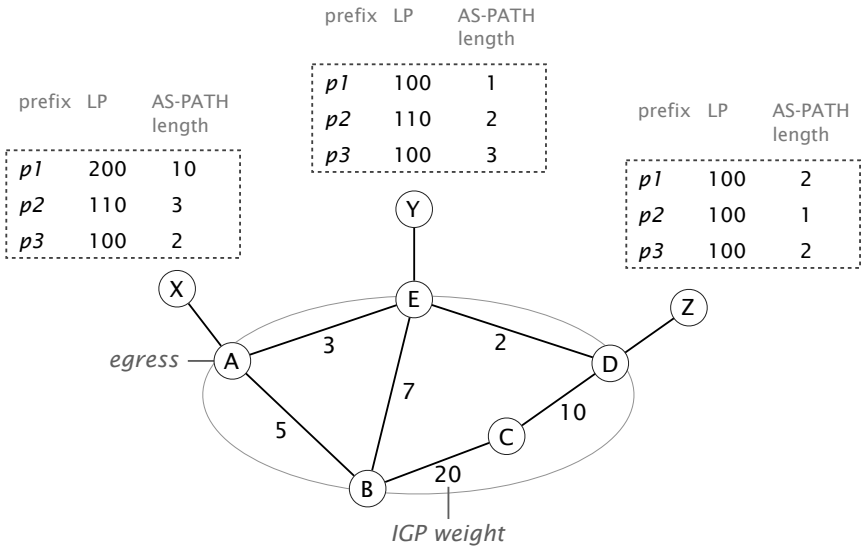
Prof. Laurent Vanbever

Solution: Exercise 8 – BGP - advanced concepts

8.1 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.

Solution:

To solve this question, follow the BGP decision algorithm for the given values. For the same prefix, BGP will prefer the route with the higher local preference. If the preferences are equal, BPG picks the route with the lower AS-PATH

length. Should the path length be equal as well (e.g., **p3** route from *X* and *Z*), the routers will pick the route with the shortest path to the next-hop. This means, the routers will pick the best route based on the IGP (OSPF) path with the lowest cost.

The internal next hop is found by figuring out how the packets are forwarded (shortest path) to reach the corresponding egress point.

A

prefix	egress	internal NH
p1	**A**	**direct**
p2	**E**	**E**
p3	**A**	**direct**

B

prefix	egress	internal NH
p1	**A**	**A**
p2	**E**	**E**
p3	**A**	**A**

C

prefix	egress	internal NH
p1	**A**	**D**
p2	**E**	**D**
p3	**D**	**D**

D

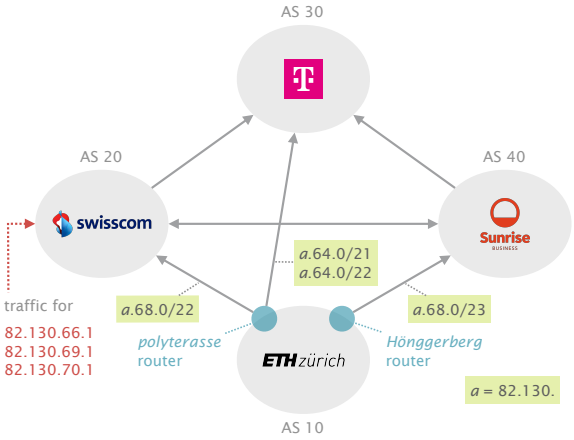
prefix	egress	internal NH
p1	**A**	**E**
p2	**E**	**E**
p3	**D**	**direct**

E

prefix	egress	internal NH
p1	**A**	**A**
p2	**E**	**direct**
p3	**D**	**D**

Solution to task 3 c).

8.2 Traffic (not so much) Engineered



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? **Solution:** [20, 30, 10]

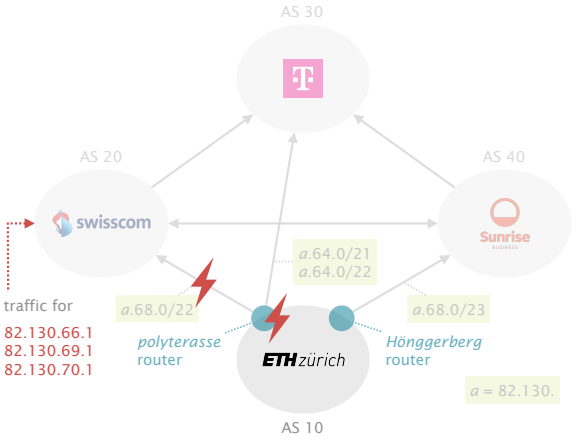
b) 82.130.69.1? **Solution:** [20, 40, 10]

c) 82.130.70.1? **Solution:** [20, 10]

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

Solution: Swisscom and Sunrise are not happy. Should ETH advertise them its full prefix, they could direct all traffic to it directly and earn money instead of paying money to Deutsche Telekom or use their peering (revenue neutral) link. Deutsche Telekom is happy though as it receives extra traffic (and therefore, revenues) from Swisscom and Sunrise for traffic pertaining to the sub-prefixes.

As routing and forwarding in the Internet is done according to the longest destination prefix, there is nothing that Swisscom or Sunrise can do besides convincing ETH to change its announcements.



Is this network as redundant as it looks?

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? **Solution:** [20, 30, 10]
- (ii) 82.130.69.1? **Solution:** [20, 40, 10]
- (iii) 82.130.70.1? **Solution:** [20, 30, 10]

b) the *polyterasse* router dies?
- (i) 82.130.66.1? **Solution:** *drop*
- (ii) 82.130.69.1? **Solution:** [20, 40, 10]
- (iii) 82.130.70.1? **Solution:** *drop*

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.

Solution: A simple solution would be to advertise the full prefix 82.130.64.0/21 to all providers *along* with the more-specific ones. Since traffic is based on the longest-prefix match, announcing additional less-specific prefix will not change the forwarding in the steady state but will enable any of its three providers to continue providing full transit in case of failures.

8.3 BGP Hijack (Exam Question 2018)

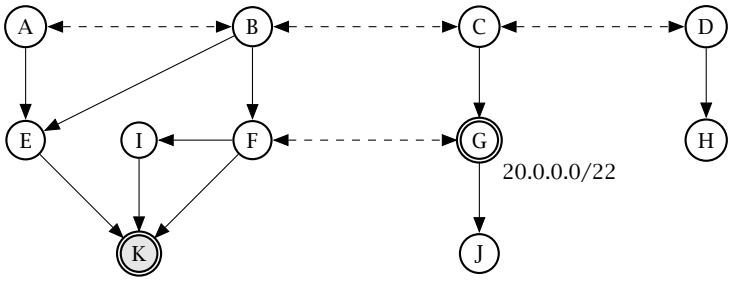
Consider the Internet topology consisting of 11 Autonomous Systems (ASes) in the Figure below. Single-headed plain arrows point from providers to their customers (AS A is the provider of AS E), while double-headed dashed arrows connect peers (AS A and AS B are peers). Each AS is made up of a single BGP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

a) AS K wants to hijack all the traffic going to AS G for 20.0.0.0/22. It starts advertising the exact same prefix. From which ASes is it able to attract the traffic?

Solution:
A, B, E, F, I.



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

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.

Solution:
AS K can break the prefix into more specific prefixes, for example 20.0.0.0/23 and 20.0.2.0/23. We can consider the two prefix lengths (/23 and /22) separately since the traffic will always follow the route for the most specific prefix independent of the other less specific routes:

First, we look at the two /23 prefixes and which ASes receive the route from K for them: A, B, C, E, F, and I. Given that these ASes now know routes for the two /23 prefixes, they will forward their traffic according to the /23 routes as the forwarding decisions are based on the longest-prefix match. Hence, the traffic of A, B, C, E, F, and I is hijacked by K through the more specific routes. Note, C will not advertise the /23 routes to its peer D as it already received the route from its other peer B.

Second, we can look at the original /22 prefix advertised by G. AS C receives an announcement for the /22 route from its customer G. It will advertise this route to both of its peers: B and D. For B, the announcement has no effect as it knows a more specific route. However, D does not have any other route for the address space of G and hence will use the /22 route as its best path. In addition, it will advertise the route to its customer H. D thinks that its traffic goes directly via C to G. In reality, the traffic from D goes to C where it will follow the more specific route and end up at hijacker K. Therefore, K is able to hijack the traffic of A, B, C, D, E, F, H, and I. It only fails to hijack the traffic of G and J.

c) The ASes from which AS K manages to attract the traffic realize what is happening as all their traffic to 20.0.0.0/22 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.

Poisoning the AS path means that AS K would put specific AS numbers in the AS path of the hijacked prefix in order to abuse the BGP loop prevention mechanism. This way, specific ASes will drop the hijacked prefix instead of forwarding it.

Solution:
BGP has a built-in loop detection mechanism. Every router checks the AS path when it receives a route announcement. If the AS path contains its own AS number, the router rejects the route as a routing loop could be created. One can "abuse" this mechanism by adding specific AS numbers to the AS path. This is called poisoning.

The hijacker could therefore add the AS number of certain ASes to the AS path of its hijack advertisements. This will trigger the BGP loop detection of those ASes where the hijack advertisements are immediately dropped. How can that help the hijacker? By doing so, certain ASes will still prefer the route from the original origin. These ASes can then be used to keep a path open to the destination and to forward the hijacked traffic to its original destination.

Applying that to the question, AS K can poison the AS path of the two /23 prefixes with AS F, i.e. prepending F once. Therefore, F will drop the /23 prefixes as it detects a loop in the AS path. F still prefers the original route towards G. K can now simply forward the hijacked traffic to F using the direct link and a static route, for example. Note that all the other ASes still handle the /23 prefixes as before. Therefore, K will be able to hijack traffic from A, B, C, D, E, H, I. All the hijacked traffic is then forwarded to G making the ongoing hijack much more difficult to detect.

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

Solution:
Instead of poisoning, AS *K* could advertise the **/22 prefix** only to a subset of its neighbors ensuring that at least one path remains open towards the original destination. AS *K* could advertise the prefix only to *E* attracting the traffic from *A*, *B* and *E*, while keeping a path via *F* to *G*. It cannot advertise the prefix to *I* as *F* would always prefer the route from *I* over the one from *G* therefore dismissing the return path. Also, AS *K* *cannot* advertise the more specific /23 prefixes as AS F would then send its traffic via *B* to *K*.

In general, please note that in this question the hijacker has complete knowledge of the entire "Internet" as well as all the details about business relationships between the ASes. Clearly that is not the case in the real Internet making such targeted attacks more difficult.