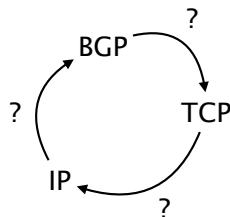# Communication Networks

Prof. Laurent Vanbever

Solution: Exercise 7 – Border Gateway Protocol (BGP)

BGP

7.1 Circularity

Chicken or the egg?

Consider two directly-connected routers A and B that have an eBGP session (running over TCP) between them. Explain how BGP, which is used to propagate *IP-based* routing information, can rely on establishing a TCP connection to exchange routes, which itself relies on IP. How is this circularity resolved? How do the two TCP endpoints manage to reach each other?

Solution: As the two routers A and B are directly connected, they do not need a routing protocol to learn about how to reach each other. Both routers are configured with an IP address belonging to the same subnet (say, 10.0.0.1/30 and 10.0.0.2/30).

When a BGP session is configured on A (resp. B) for 10.0.0.2/30 (resp. 10.0.0.1/30), the router uses its direct link to reach the other end-point. The circularity problem highlighted in the question would only happen if the subnet connecting the two routers itself would have to be learned via BGP, which is not the case.

7.2 Not-so-reliable Internet

Consider the BGP network composed of 5 ASes shown on the left which uses the normal customer-provider and peer-to-peer policies. Providers are connected to their customers with a single-headed arrow pointing to their customers (AS 1 is the provider of AS 4), while peers are connected with double-headed arrows (AS 1 and AS 2 are peers).

Assume that AS 2 is the only one to advertise an IPv4 prefix: 82.130.64.0/21 (to *all* its neighbors) and that the Internet has converged. Which BGP messages are exchanged after the following events happen, one after the other:

Which messages are exchanged?

a) the link between AS 0 and AS 2 fails (event 1)

Solution:

  (i) AS 0 sends a WITHDRAW for 82.130.64.0/21 to AS 3 (optional);

  (ii) AS 0 sends an UPDATE for 82.130.64.0/21 to AS 3 with AS-PATH [0,1,2].

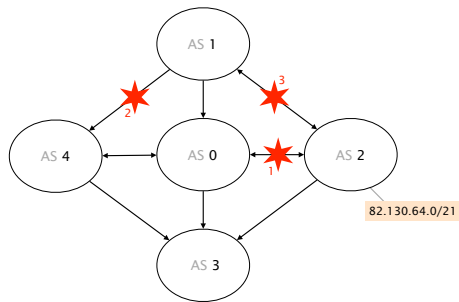**b)** the link between AS 1 and AS 4 fails (event 2)

Solution:

AS 4 sends a WITHDRAW for 82.130.64.0/21 to AS 3.

c) the link between AS 1 and AS 2 fails (event 3)

Solution:

  (i) AS 1 sends a WITHDRAW for 82.130.64.0/21 to AS 0;

 (ii) AS 0 sends a WITHDRAW for 82.130.64.0/21 to AS 3;

Is the network still connected at the end? If not, list the ASes that cannot reach the prefix anymore.

Solution:

No. The BGP network is not connected anymore. Only AS 3 is able to reach 82.130.64.0/21 via its direct link with AS 2. Observe that the physical graph *is* still connected yet as BGP policies prevent paths to be used, blackholes appear nonetheless.

7.3 Visibility (Exam Question 2016)

Consider now the network depicted on the left. 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.
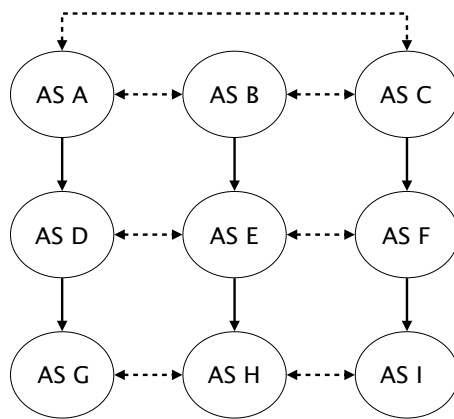
a) What path (sequence of ASes) is followed when AS G sends packets destined to the prefix originated by AS E?

Solution: Path: [G, D, E]

b) What path (sequence of ASes) is followed when AS F sends packets destined to the prefix originated by AS G?

Solution: Path: [F, C, A, D, G]

c) Suppose AS A and AS C give you a "dump" of all the BGP routes they *learn* for every destination. You then extract all links from the AS paths seen in those "dumps" and use them to construct a view of the AS-level topology. Draw the resulting AS-level topology in the figure below.
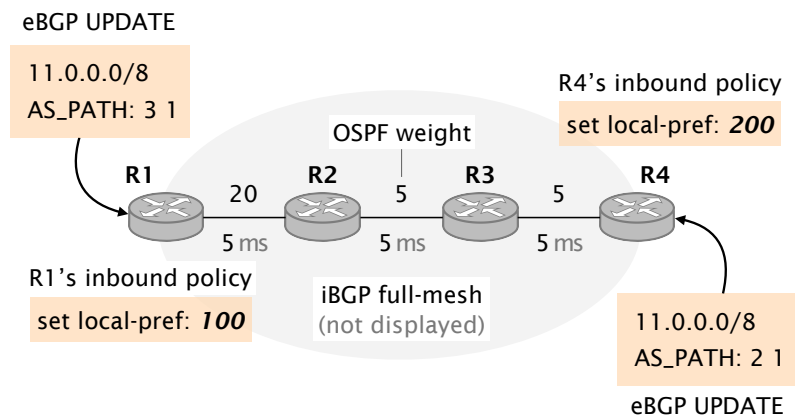
Solution:
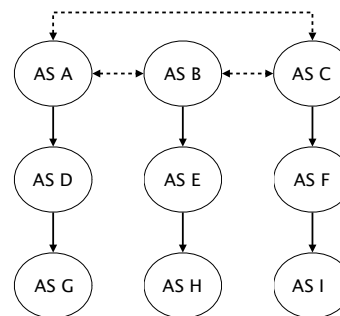
Solution

d) Give the minimum set of ASes that must provide a "dump" of each route they learn s.t. all the edges (the ones in the figure on the left) are visible? Justify your answer.

Solution: $\{A, H\}$ or $\{C, H\}$ are two possible answers (there are more possibilities). The set is of size 2. H "sees" all the links with three exceptions: (i) the top peering link A–C, (ii) C–F, and (iii) A–D. To see them, you need A or C.

A simple BGP network

7.4 Left? Right? Both? (Exam Question 2017)

Consider the BGP network composed of 4 routers depicted in Figure below. 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. The Figure also indicates the propagation delay for each link (e.g., it takes 5ms for a packet to propagate between $R1$ and $R2$).

eBGP UPDATE

11.0.0.0/8
AS_PATH: 3 1

R4's inbound policy
set local-pref: **200**

OSPF weight

R1 20 **R2** 5 **R3** 5 **R4**

5 ms 5 ms 5 ms

R1's inbound policy
set local-pref: **100**

iBGP full-mesh
(not displayed)

11.0.0.0/8
AS_PATH: 2 1

eBGP UPDATE

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

a) 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. Note that we are not looking for the *BGP* next-hop but rather the next-hop a packet would take when being forwarded.

Solution: Since the externally-learned route at $R4$ has a higher local-preference than the one at $R1$ (200 vs. 100), all routers select the route from $R4$. We get the following next-hops:

- R1: $R2$
- R2: $R3$
- R3: $R4$
- R4: <external>

b) 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.

Solution: Before the change R1 has two routes for 11/8 available:

  (i) 11/8: [2, 1] - LP 200 - R4
 (ii) 11/8: [3, 1] - LP 100 - R1

R1 selects the route from R4 as its best route and therefore does not propagate the externally-learned route with lower local-preference.

All other routers have one route available: 11/8: [2, 1] - LP 200 - R4

At first, none of the routers will change their best route, when the local-preference of the externally-learned route at R4 is reduced to 50, as they just have that one route available. However, when R1 learns about the local-preference change, it will select its own, externally-learned route and advertise that route to all the other routers in the network.

Advertisements:

- Timestamp [0 ms] R4 sends the message 11/8 - [2, 1] - LP 50 to R1, R2, R3
- Timestamp [15 ms] R1 sends the message 11/8 - [3, 1] - LP 100 to R2, R3, R4
- Timestamp [30 ms] R4 sends the message 11/8 - withdraw to R1, R2, R3

c) 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).

Solution: No, there was no forwarding loop. The route change happens from left (R1) to right (R4). This is due to the fact that R1 also uses the path through R4 and therefore does not advertise its alternative route until it becomes the best route.

d) 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.

Solution: Since both externally-learned routes are now equally preferred, the routers consider the next criteria in the decision process. Finally, they will select the route with the lower IGP metric to the BGP next-hop. We get the following next-hops:

- R1: <external>
- R2: R3
- R3: R4
- R4: <external>

e) 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.

Solution: Before the failure all routers have two routes for 11/8 available:

  (i) 11/8: [2, 1] - LP 100 - R4
  (ii) 11/8: [3, 1] - LP 100 - R1

R4 detects the failure and switches to the route from R1 while withdrawing its own route.

Advertisements:

- Timestamp [0 ms] R4 sends the message 11/8 - [2, 1] - withdraw to R1, R2, R3

f) 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).

Solution: Yes, there exist forwarding loops for 10 ms in total. The routers receive the withdrawal with a 5 ms time difference from right (R4) to left (R1). This leads to a forwarding loop of 5 ms for the two pairs of routers (R3 and R4, R2 and R3). In both cases, the router closer to R4 changes first its best path, while the router further away is still using the withdrawn path.