

Name 1:

Name 2:

COMPUTER NETWORKING

LAB EXERCISES (TP) 6

INTER-DOMAIN ROUTING: BGP-4

With Solutions

Abstract

This lab covers BGP-4, which is the Inter-Domain Routing Protocol of the Internet. You will be asked to configure peering between dual-stack edge routers, to create filters for the advertisements exchanged by the peers and to set up some simple routing policies.

Similar to the previous labs, you will use the virtual environment to configure a network setup that represents a simplified model of the Internet. This model is constructed by interconnecting the networks of three small Internet Service Providers (ISPs).

1 INTRODUCTION

1.1 BRIEF ORGANIZATION

For this lab, it is assumed that you have mastered the configuration of network interfaces with *zebra* and the configuration of OSPF with *ospfd* and *ospf6d* daemons.

In this document, we immediately start with the basics of BGP configuration using Quagga's *bgpd* daemon. In Section 3, we cover the fundamental concepts and commands used to configure the exchange of IPv4 and IPv6 routes between different autonomous systems (AS's). In the three subsequent sections, we will let you do case study on basic but important BGP operations and properties. In Section 7 we deal with policy routing. In particular, we first consider the design of filters that allow us to accept or deny a route based on its properties. Then, we introduce several advanced policy routing concepts, which permit you to modify the values of path attributes.

1.2 LAB REPORT

In the report, you are expected to answer the questions that accompany different parts of the lab. When answering the questions, fill in your answers directly in the spaces provided in this document. You should hand in one report per group. Don't forget to write your names on the first page of the report.

The deadline is **December 19, 23:55pm**.

2 NETWORK SETUP

For this lab, we will construct a network of 3 hosts, 5 routers and 7 ethernet switches, as given in Figure 1. Note that quagga configuration files and python simulation files are already provided in the package on Moodle. Please download it and unzip it into the proper location. Here, you can find the “proper location” by reading quagga files. If you prefer to work in another folder, please remember to change paths in all related files.

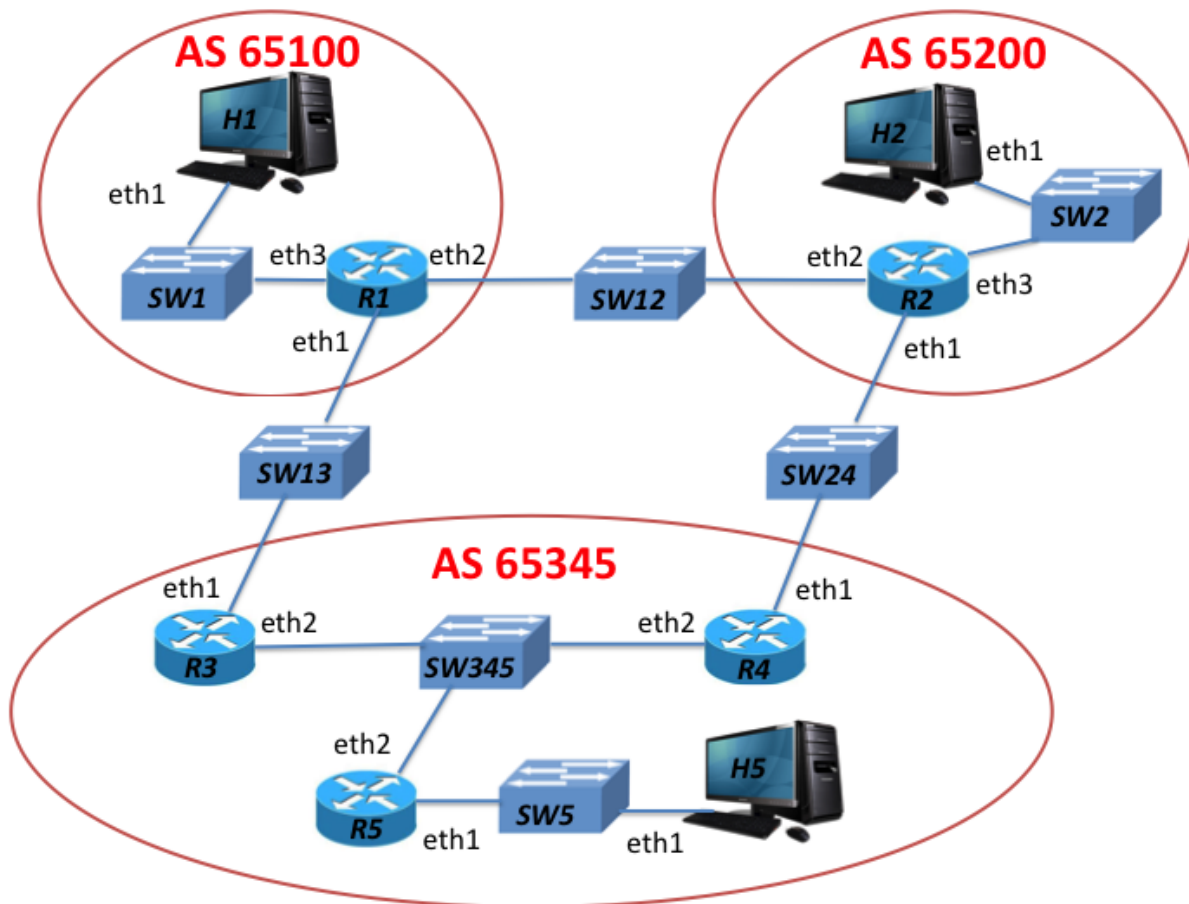


Figure 1: The network topology for this lab. It consists of 5 routers that belong to 3 autonomous systems.

3 BASIC BGP CONFIGURATION

As shown in Figure 1, the network consists of three autonomous systems (AS65345, AS65100 and AS65200). Specifically, AS65345 contains two edge routers (R3 and R4) and an internal router (R5). And each of AS65100 and AS65200 has a single router.

3.1 WARMING-UP REMINDERS

By reading all the provided files, you will notice that:

1. The passwords for *zebra,ospfd,ospf6d,bgpd* processes are no longer quagga;

2. All the above processes (if enabled in “daemon” file) will automatically start whenever you run mininet;
3. Log files will not be cleaned after exit from mininet.

Before continue, make sure that you know how to

1. Check routing tables, link state databases;
2. “On-the-fly” bring up and shut down interfaces.

3.2 QUAGGA’S BGPD PROCESS

Just as *ospfd* and *ospf6d* processes allow you to configure OSPF routing, the Quagga’s *bgpd* process allows you to configure BGP routing. Again, the preferred way of configuring BGP is to use the configuration files. “On-the-fly” configuration is also possible, but not advised, as you are prone to make mistakes when using this method.

The key difference between the *bgpd* process and the other two processes (*i.e.*, *ospfd* and *ospf6d*) is that *bgpd* is dual-stack. This means that you only need one instance of *bgpd* to support both IPv4 and IPv6 routing. You do not need to start two separate processes (daemons).

However, although a single instance of *bgpd* is sufficient to support both IPv4 and IPv6 routing, we need two TCP connections between our dual-stack peers in order to exchange IPv4 and IPv6 updates (prefixes). In the next section we explain how this dual-stack peering should be configured.

3.3 BGPD CONFIGURATION FILE EXAMPLE

The main configuration steps that must be performed in order to run BGP on a router are:

- enabling BGP routing;
- declaring peers;
- choosing certain prefixes to redistribute into BGP;
- configuring policy routing (by setting up filtering rules and/or changing the path attributes).

Let’s have a closer look at the first three bullets. The fourth bullet is covered in Section 7 of this lab. An example of the configuration file for a *bgpd* process is shown as following.

```
! bgpd configuration file example
!
log file /home/lca2/Downloads/Lab6/configs/r1/bgpd.log
debug bgp updates
!debug bgp keepalives
!debug bgp events
!
router bgp 65100
network 192.10.10.0/24
neighbor 192.13.13.3 remote-as 65345
no bgp default ipv4-unicast
neighbor 2001:1:0:1313::3 remote-as 65345
!
address-family ipv6
```

```

network 2001:1:0:1010::/64
neighbor 2001:1:0:1313::3 activate
exit-address-family
!
line vty

```

Let's examine the commands used in the example above:

- `router bgp <as-number>`: enables BGP routing, and specifies the AS number of the domain that the router belongs to (in the example above, the AS number is 65100).
- `network 192.10.10.0/24`: instructs the *bgpd* process on the local router to announce the prefix 192.10.10.0/24 to all BGP peers. The command has the same syntax for both IPv4 and IPv6 prefixes (see the line `network 2001:1:0:1010::/64`). If instead of a particular subnet (prefix), we want to announce all directly connected subnets via BGP, then we can use the command `redistribute connected`. If we want to redistribute into BGP the routes learned via another routing protocol, or statically configured routes, we use `redistribute <protocol>`, where `<protocol>` can be set to `rip`, `ospf`, or `static`, depending on the source of the routes.
- `neighbor <ip address> remote-as <as-number>`: is used to declare a BGP peer (that belongs to the same or a different AS). A TCP connection is established between the local *bgpd* process and the peer. This connection is used to exchange keepalives and BGP routing updates. The command `neighbor <ipv6 address> activate` activates the exchange with the declared IPv6 peer. Note that no distinction between I-BGP and E-BGP is made. This information is learned automatically from the AS numbers.
- `no bgp default ipv4-unicast`: By default the IPv4 prefixes (also called the IPv4 family prefixes) are exchanged, both between the IPv4 peers and between the IPv6 peers. When this command is used in front of the peer declaration (*i.e.*, in front of the `neighbor` command), it prevents the exchange of IPv4 prefixes via the TCP connection established between the BGP IPv6 peers.
- `address-family ipv6` and `exit-address-family`: these two commands enclose the block of commands that configure the IPv6 communication with the declared IPv6 peer.
- `debug bgp updates`, `debug bgp keepalives` and `debug bgp events`: debugging of the data exchanged between BGP peers.

The set of commands used in the configuration file above is just a subset of the BGP configuration commands offered by the Quagga routing suite. For a complete reference, check the documentation on the Quagga website (<http://www.nongnu.org/quagga/docs.html>);

On Moodle, the *bgpd* configuration files for the routers R1, R2 and R3, are already created for you. By looking at the above example as well as the *bgpd* files of R1, R2 and R3, you are expected to write the *bgpd* configuration files for the routers R4 and R5.

Please note that, depending on the exercise, the *bgpd* configuration files may require certain modifications. In some of the exercises, the *bgpd* processes will not be running on all routers. Also, some of the exercises require you to implement policies.

3.4 BGP MONITORING COMMANDS

As you may already know, in order to monitor the activity of a running Quagga process, you can enter the Quagga configuration mode by typing `telnet localhost bgpd` with password `bgpd`.

Use `show ip bgp` to visualize the BGP database at the router. Note that the BGP database is the same as BGP RIB-In with the best route to each destination marked by “>”.

4 STANDARD BGP MODE OF OPERATION

4.1 BGP RUNNING ON ALL ROUTERS

Create the `bgpd` configuration files for the routers R4 and R5. Each router should redistribute the directly connected subnets into BGP. Once you have finished editing the configuration files, run python file `Lab6_Network.py` script to start the network simulation with “`zebra` and `bgpd` processes” on R1 and R2, and “`zebra`, `ospfd`, `ospf6d` and `bgpd` processes” on R3, R4 and R5. (**Recall that you need to enable the processes in “daemon” file.**)



Q1/ Show the BGP database entries of R5, and look for the subnets 192.10.10.0/24, 192.12.12.0/24 and 192.20.20.0/24. How did R5 learn about these subnets?

Solution: *The BGP routes to subnets 192.10.10.0/24, 192.12.12.0/24 and 192.20.20.0/24 are advertised by R1 and R2 (and later by R3 and R4 using I-BGP).*

Start Wireshark on R1 and observe the BGP packets exchanged between R1 and its BGP peers.



Q2/ What is the role of KEEPALIVE messages? Can you find the period of them? Explain briefly how to find.

Solution: *They are used to (periodically) check if the BGP peers are still reachable and active. By observing the BGP packets exchanged between BGP peers, we know that they are sent every 60 seconds.*

4.2 R1 AND R2 NOT BGP PEERS

For this part of the lab, we will modify the `bgpd.conf` files of R1 and R2 so that AS65100 and AS65200 won't be BGP neighbors *i.e.*, R1 and R2 don't exchange BGP routing updates with each other. After you do the necessary modifications, restart the network.



Q3/ Can all routers reach all subnets? You may try pinging between hosts.

Solution: *Yes.*



04/ Observe the BGP database of R1 and R2. What are the entries for subnet 192.20.20.0/24 in R1 and for subnet 192.10.10.0/24 in R2? What are the AS paths for these entries? Furthermore, check also the corresponding subnets and AS paths in IPv6. Is there anything essentially different from IPv4?

Solution: In BGP database at R1: 192.20.20.0 (Network) 192.13.13.3 (Next Hop) 0 (Weight) 65345 65200 (Path); In BGP database at R2: 192.10.10.0 (Network) 192.24.24.4 (Next Hop) 0 (Weight) 65345 65100 (Path). To reach 192.20.20.0/24, R1 should go through AS65345 and AS65200. To reach 192.10.10.0/24, R2 should go through AS65345 and AS65100.

For IPv6, everything is similar to that of IPv4.

5 RUNNING BGP ONLY ON EDGE ROUTERS

In the previous section, BGP runs on all three routers of AS65345.



05/ In practice, running BGP on all routers in an AS is not feasible. Why?

Solution: If we run BGP on all routers in an AS that has a sufficiently large number of routers, there will be a large number of TCP connections as a result of I-BGP's full-mesh requirement.

In order to avoid running BGP on all routers, we ask our friends Leo, Mikey and Don to come up with alternative solutions. They propose three solutions given as following.

5.1 REDISTRIBUTE OSPF

Leo thinks that he can avoid the large number of TCP connections in a large AS **by running BGP only in edge routers**. But he also knows that if only edge routers run BGP, other non-edge routers will not be able to reach subnets in other AS's. Thus he decides to **redistribute BGP into OSPF** so that these routers will learn about such subnets using OSPF. Leo also knows that the subnets that are not directly connected to the edge routers will not be visible to routers in other AS's. Therefore, he also decides to **redistribute OSPF into BGP**. In order to test Leo's solution, we need to make some changes. Assume that we start with the settings in Section 4.2. Also, assume that **we only focus on IPv6 routes in this subsection**.



06/ At R3 and R4, which configuration files should you modify? Write down all the modifications you make? (Hint: pay attention to the bold words in the above paragraph.)

Solution: *bgpd.conf*: redistribute ospf6 in the IPv6 address-family blocks
ospf6d.conf: redistribute bgp

Now, restart the network with “zebra” on all routers, “bgpd” on routers R1, R2, R3 and R4, and *ospf6d*” on all routers in AS65345 (i.e., R3, R4 and R5). If the modifications are correct, all routers should be able to reach all subnets in the network.



7/ Observe the BGP database of R1 and R2. In terms of the entry for subnet 2001:1:0:2020::/64 in R1 and the entry for subnet 2001:1:0:1010::/64 in R2, are they the same as what you get in Q4? From your observation of the AS paths, do you think Leo's solution is appropriate?

Solution: *It's different from Q4. Obviously the path is not complete. So, Leo's solution is not appropriate. In details, the AS that redistributes OSPF will be assumed as the origin for all subnets advertised in its BGP updates by other AS's.*

5.2 USING THE *network* COMMAND

Mikey, like Leo, believes the solution to avoiding a full-mesh I-BGP is to run BGP only on edge routers and to redistribute BGP into OSPF. But different from Leo, Mikey does not want to redistribute OSPF into BGP. Hence he wants to you to test his solution by simply running BGP at edge routers and redistributing BGP into OSPF. To test Mikey's initial solution, make the following changes:

- Undo the changes you made above to `bgpd.conf` files of R1 and R2 so that R1 and R2 are BGP peers.
- Redistribute BGP into OSPF in R3 and R4.
- **DON'T** redistribute OSPF into BGP in R3 and R4.

After making these changes, restart the network with “*zebra*” on all routers, “*bgpd*” on routers R1, R2, R3 and R4, and “*ospfd* and *ospf6d*” on all routers in AS65345 (*i.e.*, R3, R4 and R5).



8/ Observe the routing tables (via *zebra*) at R4. How many entries are there for the prefix 192.10.10.0/24? Which entry is chosen as the best route to this destination? Why?

Solution: *B 192.10.10.0/24 [200/1] via 192.13.13.1, 00:00:06*
*O > * 192.10.10.0/24 [110/20] via 192.34.34.3, r4-eth2, 00:00:15*
The entry that is learned through OSPF is chosen as the best. This is because the administrative distance is 110 for OSPF, but 200 for I-BGP.



9/ Can you ping between all hosts? Is there missing information at any router?

Solution: *No. Subnets 192.50.50.0/24 and 2001:1:0:5050::/64 are not reachable from R1 and R2*

From your experiment, you must have observed that if an AS has subnets that are not directly connected to an edge router (BGP speaker), routers in other AS's won't be able to route packets to such “hidden” subnets.

Mikey has also observed this, he therefore asks Don for help. Don decides to use the `network <prefix>` command on the edge router's `bgpd.conf` files to instruct the edge routers to include the prefixes of “hidden” subnets in their BGP routing updates. He considers this as the best solution because all AS's know the list of prefixes in their network and should not be difficult for them to use `network` commands in the edge routers for all such prefixes.

Don wants you to test his solution and see if it works. To test Don's solution,

- Modify the `bgpd.conf` files of R3 and R4 such that you use `network <prefix>` command to advertise the 192.50.50.0/24 and 2001:1:0:5050::/64 subnet in their BGP routing updates.

Restart the network with “zebra” on all router, “bgpd” on routers R1, R2, R3 and R4, and “ospfd and ospf6d” on all routers in AS65345 (*i.e.*, R3, R4 and R5).



Q10/ By looking at the routing tables of the routers, can all routers reach all subnets in the network? How many entries are there in the BGP database of R1 for the subnet 192.50.50.0/24? Which entry is chosen as the best one? Why?

Solution: Yes. Two entries, with next-hops 192.13.13.3 and 192.12.12.2. Entry with next-hop 192.13.13.3 is chosen because of shortest AS path.

6 BROKEN LINK

In this section we observe how connectivity is affected if a link breaks in the network. Continue working with Don’s configuration *i.e.*,

- use `network <prefix>` command in R3 and R4 to advertise the subnet 192.50.50.0/24 and the subnet 2001:1:0:5050::/64 in their BGP routing updates.
- Redistribute BGP into OSPF in R3 and R4.
- **DON’T** redistribute OSPF into BGP in R3 and R4.
- Run `bgpd` on R1, R2, R3 and R4.
- Run `ospfd` and `ospf6d` on all routers in AS65345.

6.1 EXTERNAL BROKEN LINK

In this experiment, we will simulate a broken link between R3 and SW13 by shutting down the interface `eth1` at R3. Utilise the commands you used in Lab 4 to shut down the interface. Though the BGP connection between AS65100 and AS65345 will be broken as a result of the link failure, the ring topology of the network will let BGP converge after some exchange of messages between the different BGP peers.



Q11/ In the log file of `bgpd` at R4, there are UPDATE messages received from R3 as a result of the link failure. What do they mean?

Solution: R3 sends UPDATE Messages to R4 when there is link failure. Its purpose is to withdraw routes to the following subnetworks: 1. 192.10.10.0/24; 2. 192.12.12.0/24; 3. 192.13.13.0/24. (Similar for IPv6.)



Q12/ How does R3 route packets to 192.10.10.0/24 subnet?

Solution: R3 takes the “Next Hop” as 192.34.34.4 at R4.

6.2 INTERNAL BROKEN LINK

In the previous experiment, we saw what happens when a link connecting two AS’s breaks in the presence of redundancy (*e.g.*, ring topology). In what follows, we will look at a pathological case where a link failure

causes a disconnected AS. In order to do this experiment, first bring up interface *eth1* of R3 that you shut down in the previous experiment. Then, simulate a broken link between R3 and SW345 by shutting down the interface *eth2* at R3. (Again apply the same commands you used previously.)



Q13/ Look at the BGP database of R1. How many entries are there for the subnet 192.50.50.0/24? Which entry is chosen as the best one? Can you ping from R1 to this subnet? Explain briefly the reason.

Solution: *There are two routes:*

192.50.50.0 (Network) 192.13.13.3 (Next Hop) 0 (Metric) 0 (Weight) 65345 i (Path)

192.12.12.2 (Next Hop) 0 (Weight) 65200 65345 i (Path)

The first one is chosen as the best. Cannot ping this subnetwork from R1. This is because the link is broken. Actually, R3 keeps advertising this route.



Q14/ Will R3 learn the subnet 192.34.34.0/24 by BGP? Why or why not?

Solution: *No. This is because when R1 advertises to R3 the route to 192.34.34.0/24, R3 finds its own AS in the AS-path. So, in order to avoid a loop, R3 ignores the route.*

7 POLICY ROUTING (OPTIONAL)

7.1 FILTERING BGP UPDATES (ACCESS LIST)

Filtering BGP updates allows a router to accept some of them and reject the others. A filter can be applied:

- to the BGP updates coming from a BGP peer, before they are added to the BGP database of the router;
- to the updates sent to a BGP peer, before that they are actually sent.

Filtering can be based **on *prefix*** or **on *as-path attribute*** (an AS number or a sequence of AS numbers).

In the case of filtering by *prefix*, the `distribute-list` type of filter is used. The syntax is:

```
neighbor <ip_address> distribute-list <filter_name> in/out
```

- `<ip_address>`: is an IPv4 or IPv6 address of the BGP peer,
- `<filter_name>`: is the name of the filter, and
- `in/out`: specifies whether the filter is applied to the input or to the output of the router.

In the case of *as-path* based filtering, the `filter-list` type of filter is used. The syntax is:

```
neighbor <ip_address> filter-list <filter_name> in/out
```

To apply a filter to the updates sent by a peer, you must define it first. A filter is defined with a list of rules, which is called the `access-list`. The syntax to define an `access-list` rule differs, depending on whether the filtering is performed by *prefix* or by *as-path*.

The syntax to define an `access-list` rule in the case of filtering by *prefix* is the following:

```
(ipv6) access-list <filter_name> permit/deny <prefix>
```

- `ipv6`: this part of the command exists only in case of an IPv6 prefix
- `<filter_name>`: the name of the filter: it can be composed of letters and/or numbers,
- `permit/deny`: specifies whether the interface permits or denies the updates that match the prefix,
- `<prefix>`: the prefix the filter should match. It can be replaced by `any`, in which case the filter applies to any advertised prefix.

The syntax to define an `access-list` rule in the case of filtering by *as-path* is the following:

```
ip as-path access-list <filter_name> permit/deny <as_number_sequence>
```

- `<filter_name>`: is the name of the filter; it can be composed of letters and/or numbers,
- `permit/deny`: specifies whether the interface permits or denies the updates that match the AS path,
- `<as_number_sequence>`: is the sequence of one or more AS numbers. It can be substituted by `.*`, in which case it applies to any AS path.

When a filter is applied to a network interface, each BGP update that passes through the interface (in the indicated direction) is checked against the list of rules that constitute the filter. The rules are checked in the same order in which they appear in the configuration file. If a matching rule is found, the subsequent rules are not checked. **If no matching rule is found, the update is dropped (implicit deny).**

If this sounds a bit complicated, here is an example to help you:

```
router bgp 65101
  bgp router-id 192.45.88.3
  network 192.45.88.0
  neighbor 200.44.0.13 remote-as 65131
  neighbor 200.44.0.13 distribute-list Alpha_1 in
  neighbor 200.44.0.12 remote-as 65121
  neighbor 200.44.0.12 filter-list Beta_2 out
  !
  access-list Alpha_1 permit 192.168.33.0/24
  access-list Alpha_1 permit 133.33.23.0/24
  !
  ip as-path access-list Beta_2 permit 65121
  ip as-path access-list Beta_2 permit 65111 65131
```

In the example above, the filter `Alpha_1` is an input filter. It accepts the updates for prefixes `192.168.33.0/24` and `133.33.23.0/24`, from the BGP peer at address `200.44.0.13` in AS65131.

The filter `Beta_2` is an output filter. It allows the updates that contain AS65121 or the AS sequence `65111 65131` (in this order and at an arbitrary position inside the AS path) to be sent to the BGP peer at address `200.44.0.12` in AS65121.

Note:

- A filter (like `Alpha_1`) has to be assigned to an interface (like for `bgp neighbor 200.44.0.13`) before the filter definition (like `access-list`);
- In the case of IPv4, a filter should be assigned to an interface by declaring the filter just after `bgp neighbor`;
- In the case of IPv6, a filter should be assigned to an interface by declaring the filter just after the interface activation inside the `address-family ipv6` block;
- For IPv4 and IPv6, please give detailed filter definitions after the `address-family ipv6` block.

7.2 MODIFYING PATH ATTRIBUTES (ROUTE MAP)

Filtering is an important tool in the process of route selection. However, they only allow you to accept or reject a route. In order to take more sophisticated actions (*e.g.*, give different preferences to different routes), you have to use a *route map*.

A route map is a more powerful tool than a filter. In addition to accepting/rejecting a route, a route map allows you to modify its attributes. A route map is typically applied to a network interface in the *incoming* or *outgoing* direction. The syntax to do this is the following:

```
neighbor <ip_address> route-map <rm_name> in/out
```

- `<ip_address>`: is an IPv4 or IPv6 address of the BGP peer,
- `<rm_name>`: is the route map name (any sequence of characters and/or numbers),
- `in/out`: specifies whether the route map is applied in the incoming or in the outgoing direction.

The definition of a route map has the following form. It consists of *filter-action constructs*.

```
route-map <rm_name> permit <order>
match <filter_type> <filter_name>
(action)
```

- `<rm_name>` is the name of the route map this *filter-action construct* belongs to,
- `<order>` is an integer that determines in which order the constructs are processed. *Filter-action constructs* with lower order value are processed first.
- `<filter_type>` is the type of the filter used for matching BGP updates (*i.e.*, `as-path` if matching by AS path or `ip(v6) address` if matching by prefix);
- `<filter_name>` is the name of the filter, which has to be defined in the same configuration file,
- `(action)` is the action performed with the updates that match the filter.

Here is an example of a route map (two filters in the `match` commands are also provided):

```
ip as-path access-list Gamma_1 permit 65201 65432
access-list Delta_2 permit 192.54.15.0/24
!
route-map ccw permit 20
match as-path Gamma_1
```

```
set weight 200
!  
route-map ccw permit 10  
match ip address Delta_2  
set weight 500
```

The route map in the example above has two *filter-action constructs*. For each update that passes through the interface to which the route map is applied (in the indicated direction), the list of *filter-action constructs* is scanned in the increasing order. The construct with the lowest *order* value is processed first. Unlike the filtering rules, the order in which the constructs appear in a route map does not matter. Thus, in the example above, the construct with the order value 10 is checked first.

If the filter part of the construct matches the arriving update, the construct action is performed. The remaining constructs are not processed. If the filter part of the construct does not match, the next construct is processed and so on. **If none of the constructs matches, the update is dropped.** Thus, always bear in mind that you might need a default permit construct at the end of the route map.

A number of possible actions can be performed with a route. Here are some of the path attributes that can be modified using the `set` command:

- the *weight* associated to a route: by default, all routes have weight 0. The exception are locally connected subnets that have weight 32768. The router selects the route with the highest weight. The weight can be set to any positive integer. The syntax is:

```
set weight 500
```

- the *local preference* associated to a route: by default, it is equal to 100. Higher local preference is preferred. The difference between weight and local preference is that weight is local to a router, whereas local preference is announced to all BGP routers in the same AS. The syntax is:

```
set local-preference 50
```

- the *AS path sequence* of the advertised route. It can be modified using the `set as-path` command. An AS sequence can be prepended to the advertised AS path, as shown in the following example:

```
set as-path prepend 65289 65453 65789
```

The command inserts the sequence 65289 65453 65789 at the beginning of the advertised AS path.

Note: `route-map` is used in a similar way to `distribute-list`, please make sure that all the declarations and definitions are correctly positioned.

7.3 POLICY ROUTING PLAYGROUND

7.3.1 AVOID TRANSIT TRAFFIC

Recover the configuration in Section 4.2 and imagine that AS65345 in Figure 1 represents an enterprise network of the Foot Clan, which is connected to the Internet via two ISPs—X (AS 65100) and Y (AS 65200). You are the administrator of this enterprise network and you do not want your autonomous system (AS65345) to become a transit AS, *i.e.*, you do not want the traffic that neither originates nor terminates in AS65345 to be routed through this AS. At the same time other autonomous systems have to know about all the prefixes in AS65345.



15/ Give the most important lines of the modified *bgpd* configuration file at R3 and R4. (*Hint: use distribute-list and filtering by prefix*)

Solution: For R3:

```
neighbor 192.13.13.1 distribute-list advertise-own-ipv4 out
neighbor 2001:1:0:1313::1 distribute-list advertise-own-ipv6 out
access-list advertise-own-ipv4 permit 192.34.34.0/24
access-list advertise-own-ipv4 permit 192.50.50.0/24
ipv6 access-list advertise-own-ipv6 permit 2001:1:0:3434::/64
ipv6 access-list advertise-own-ipv6 permit 2001:1:0:5050::/64.
Similar for R4, by replacing 192.13.13.1, 2001:1:0:1313::1 with 192.24.24.2, 2001:1:0:2424::2.
```

If the modifications are correct, you should be able to ping between H1,H5 (and similarly between H2,H5), but not between H1,H2.

7.3.2 POLICY WAR

Recover the configuration in Section 4.1 and imagine that the autonomous systems 65100, 65200 and 65345 represent domains of three internet service providers (ISPs), as indicated in Figure 1. The relations among these ISPs are such that ISP X (AS 65100) and ISP Y (AS65200) pay a fee to ISP Z (AS 65345) to have it carry their traffic (*i.e.*, ISP Z is a service provider for the other two ISPs). ISP X and ISP Y have a peering agreement, *i.e.*, they exchange traffic without paying fees to each other.

Now, let us consider the subnets 192.50.50.0/24 and 2001:1:0:5050::/64 that belong to ISP Z. ISP X carries a lot of traffic to these subnets and it would consequently prefer the traffic destined to these subnets (and only the traffic destined to these subnets, other traffic should flow normally!) to go via ISP Y, if possible. This would be a cheaper solution for ISP X, but of course unfair to ISP Y.



16/ Write down the most significant lines of the modified configuration files at R1. (*Hint: use route-map, weights, and filtering by prefix. Remember to have a default-case access-list.*)

Solution: This can be achieved by using weight or local preference attributes.

```

router bgp 65100
 redistribute connected
 neighbor 192.13.13.3 remote-as 65345
 neighbor 192.12.12.2 remote-as 65200
 neighbor 192.12.12.2 route-map redirect4 in
 no bgp default ipv4-unicast
 neighbor 2001:1:0:1313::3 remote-as 65345
 no bgp default ipv4-unicast
 neighbor 2001:1:0:1212::2 remote-as 65200
 address-family ipv6
 redistribute connected
 neighbor 2001:1:0:1212::2 activate
 neighbor 2001:1:0:1212::2 route-map redirect6 in
 neighbor 2001:1:0:1313::3 activate
 exit-address-family

access-list subnet4 permit 192.50.50.0/24
access-list other4 permit any
ipv6 access-list subnet6 permit 2001:1:0:5050::/64
ipv6 access-list other6 permit any

route-map redirect4 permit 10
match ip address subnet4
set weight 1000
route-map redirect4 permit 20
match ip address other4

route-map redirect6 permit 10
match ipv6 address subnet6
set weight 1000
route-map redirect6 permit 20
match ipv6 address other6

```

ISP Y notices the changes and took this as the beginning of a small “war”. To avoid carrying traffic between ISP X and ISP Z’s subnets 192.50.50.0/24 and 2001:1:0:5050::/64, ISP Y decides to change the configuration at the router R2 (again, other traffic should not be affected by these modifications).

The first plan for ISP Y is to add an output filter towards R1 that discards all updates for the subnets 192.50.50.0/24 and 2001:1:0:5050::/64.



17/ Write down the most significant lines of the modified configuration files at R2. (*Hint: use distribute-list and filtering by prefix.*)

Solution:

```
neighbor 192.12.12.1 distribute-list hide4 out
neighbor 2001:1:0:1212::1 distribute-list hide6 out

access-list hide4 deny 192.50.50.0/24
access-list hide4 permit any
ipv6 access-list hide6 deny 2001:1:0:5050::/64
ipv6 access-list hide6 permit any
```



18/ Can you check that your solution works by inspecting the BGP database at R1?

Solution: Yes, there should be only one route for each of the two prefixes (192.50.50.0/24 and 2001:1:0:5050::/64).

The second plan for ISP Y is to do the same as ISP X. Namely, ISP Y would prefer the traffic destined to 192.50.50.0/24 and 2001:1:0:5050::/64 to go via ISP X.



19/ What will happen?

Solution: ISP X and ISP Y will compete to win the policy war.

8 CONCLUSION

This is your final Lab for the TCP/IP course. We hope you enjoy the Labs during the semester. **All the TA team members wish you happy holidays and good luck to your exams.**

