Spring 2020 Prof. L. Vanbever/ T. Bühler, R. Birkner, T. Holterbach, C. Busse-Grawitz

Building your own mini-Internet

Deadline: April 9 2020 at 11.59pm

In this project, you together with more than 100 of your fellow classmates will build and operate your very own mini-Internet. Your main goal? Enabling end-to-end connectivity across ≈ 80 Autonomous Systems (ASes) composed of hundreds of network devices. In doing so, you will experiment with the most common switching and routing technologies used today in the Internet. You will also face the same challenges actual network operators experience every day.

To reach Internet-wide connectivity, you will first need to enable internal connectivity, within your own AS, before interconnecting your AS with others ASes, managed by other groups of students. To establish connectivity within your AS, you will use the Spanning Tree Protocol (STP) and Open Shortest Path First (OSPF). To establish connectivity across different ASes, you will use the only inter-domain routing protocol available today: the Border Gateway Protocol (BGP). At the end of the project, any end-host should be able to communicate with each other, independently of the AS they are located in.

To help you, we have pre-built a base network topology on top of virtual layer-2 switches, running Open vSwitch [1] and virtual routers, running the FRRouting software routing suite [2]. You will configure the virtual switches and routers through a Command Line Interface (CLI). This interface is virtually identical to the one used by actual network operators.

The rest of this document is organized as follows: Section 1 provides general information about the project, including **submission instructions**. Section 2 gives an overview of the mini-Internet and the network you will be configuring. Section 3 contains the concrete tasks to solve and document in the report, while Section 4 shows you how you can test and verify your configuration. Section 5 gives the answers to frequently asked questions.

In a separate document,[1] we give you a crash course on how to access the mini-Internet and how to configure a FRRouting router and an Open vSwitch.

1 General Information

This section tells you what to do if you have questions, how to backup and submit your work and how it will be graded. Furthermore, it explains our policies on academic integrity and misuse of the resources.

1.1 If you have questions: use Slack or send us an email

Ask your question on the `#routing_project` channel available at `comm-net20.slack.com`. Please do not ask questions in the `#general` channel. You can also ask questions by email.

1.2 Online Q&A sessions on Slack

During the normal exercise timeslots (Thursdays from 10:15AM to 11:55AM), we will be available on Slack to answer questions and help you with the project. We will offer additional Q&A sessions to assist you as needed. The times of these additional sessions will be announced through Slack

[1] `http://comm-net.ethz.ch/pdfs/project_1/tutorial.pdf`

and our website in a timely manner. Outside of these timeslots, you can always post your questions on Slack, but you might not get an immediate answer.

1.3 Regularly backup your work

We provide you with a script that automatically saves all configs of all your routers and switches (`save_configs.sh`) in one place (see also the tutorial pdf). We advise you to use this script regularly and to copy the generated folder to your local machine in order to prevent losing your work in the case of unexpected problems.

1.4 Submit your work by email

Send your final report and configuration by email to Laurent Vanbever (`lvanbever@ethz.ch`), Thomas Holterbach (`thomahol@ethz.ch`) and Tobias Bühler (`buehlert@ethz.ch`). Make sure that your email includes a zip or tar.gz archive containing a PDF report as well as all your configuration files (the directory generated with the `save_configs.sh` script). Please make sure that your PDF report includes your group number as well as the name of the members in your group. The maximum length for your PDF report is 10 A4 pages (including screenshots, traceroutes, looking glass etc.). You can always remove parts from the screenshots, looking glass entries, etc. which are not needed to answer the question or demonstrate the correct functionality of your configuration.

Important: the subject of your email must follow this format: **[comm_net] groupX project 1**, where X is your group number.

1.5 Our grading policy

This assignment will be graded and counts for 20% of your final grade. There are a maximum of 10 points (plus half a bonus point). Each group member will receive the same grade: $\min\left\{1 + \frac{\sum pts}{2}, 6\right\}$

1.6 Academic integrity

We adopt a strict zero tolerance policy when it comes to cheating. Cheating will immediately result in the group failing the assignment and being reported to ETH administration. In particular, you can only do your assignment with the other members of your group. Do not look at other groups' configuration and do not copy configurations from anywhere. It is OK to discuss things or find help online, but you must do the work by yourself.

Your configuration and report may be checked with automated tools so as to discover plagiarism. Again, **do not copy-and-paste** code, text, etc.

1.7 Misuse of the resources and infrastructure

It is prohibited to use and modify the network in other ways than expressly allowed in this task description. The forbidden misuse includes, but is not limited to, BGP hijacks, DDoS attacks, resource-/bandwidth-hungry programs, and the attempt to access the docker containers of other groups. We monitor, investigate, and apply the appropriate disciplinary actions for cases of misuse.

2 Network Topologies

Similarly to real networks, your network spans over layer-2 (using switches) and layer-3 (using routers). Your network also connects (at layer-3) to other networks, creating an *Inter*net. We now describe each aspect of the network topology.

L2 topology Your layer-2 network (called Swiss local network) is composed of three switches (Fig. 1) located at three different locations: ETHZ, EPFL, and CERN. The switch at ETHZ is

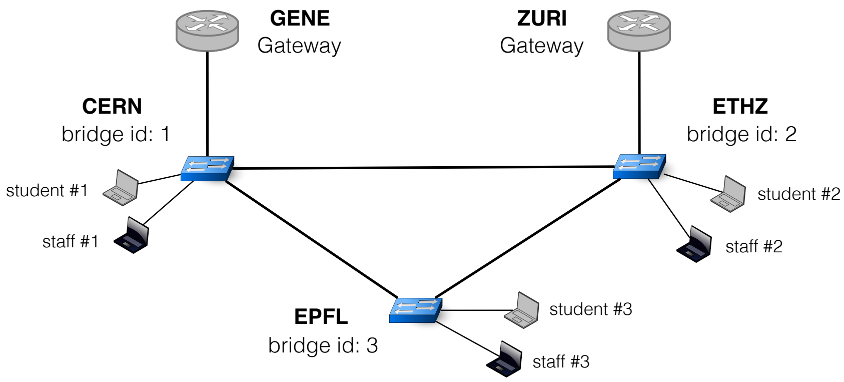

Figure 1: Each group will have to manage its own Swiss local network. This layer-2 network is composed of three Open vSwitches located at different locations. The switch at ETHZ connects to a layer-3 router (ZURI) which acts as the gateway and CERN is connected to GENE.

connected to a layer-3 router (ZURI), the one at CERN is connected to GENE. Both routers can act as a gateway, meaning that a host in the local network must send a packet to it to reach any non-local destination. The router will then take care of sending that packet to the destination.

Two types of users exist in your layer-2 network: students and staff. Each switch is connected to one student and one staff member. Each switch also has a bridge ID which is indicated in Figure 1. For example, the switch at ETHZ has bridge ID 2.

L3 topology For this project, imagine that your layer-2 network is part of an AS spanning across the atlantic that you also manage. Your AS number is your group number: *e.g.,* AS 28 for group 28. Your AS has routers located on two continents: four routers in Europe (Geneva, London, Paris, and Zürich) and four in the US (Atlanta, Boston, New York, and Miami) see Figure 2.

Every AS has been allocated one /8 prefix that it can allocate internally. If you are group X, then the prefix X.0.0.0/8 is yours, meaning that group 48 has the prefix 48.0.0.0/8. You will use this IP space to allocate IP addresses to your hosts and routers.

Finally, one host is connected to each router with the exception of ZURI and GENE as these two routers are connected to the Swiss local network.

Internet topology Every router has an external connection to one of your neighboring ASes. Some are connected to a provider, some to a customer and others to a peer. NEWY is connected to an Internet eXchange Point (IXP). You will have to establish eBGP sessions on these external links. Figure 3 shows the mini-Internet topology you will end up building.

The red ASes (1, 2, etc.) are all Tier1 ASes, meaning their neighboring ASes are either peers or customers. The grey ASes (13, 14, etc.) are stub ASes, meaning their neighboring ASes are either peers or providers but they have no customers. We (the TA team) will take care of the Tier1 ASes as well as the stub ASes.

The Tier2 ASes (blue ASes) have peers, customers and providers. For example, group 5 has two providers (3 and 4), two peers (6 and the IXP 81) and two customers (7 and 8).

There are seven IXPs within our mini-Internet. The primary purpose of an IXP is to allow networks to interconnect directly. One advantage of using an IXP is that an AS can directly peer with another AS through the IXP, instead of reaching it via a provider that it has to pay. Another advantage is that only one physical connection with an IXP is needed to potentially interconnect with all the other IXP participants. An IXP uses a BGP Route Server to advertise prefixes between its participants.

One IXP is connected to all the Tier1 ASes, allowing them to be connected in a full-mesh fashion. The other IXPs are always interconnecting two blocks. This enables these ASes to peer between them (as long as they respect the BGP customer/provider policies), instead of using (and paying!) their providers. The following example illustrates the benefit of being connected to an IXP: AS6 can send traffic to AS105 via the IXP126, instead of paying AS 4 to send the traffic via the path 4-2-101-103-105 if IXP126 is not used.

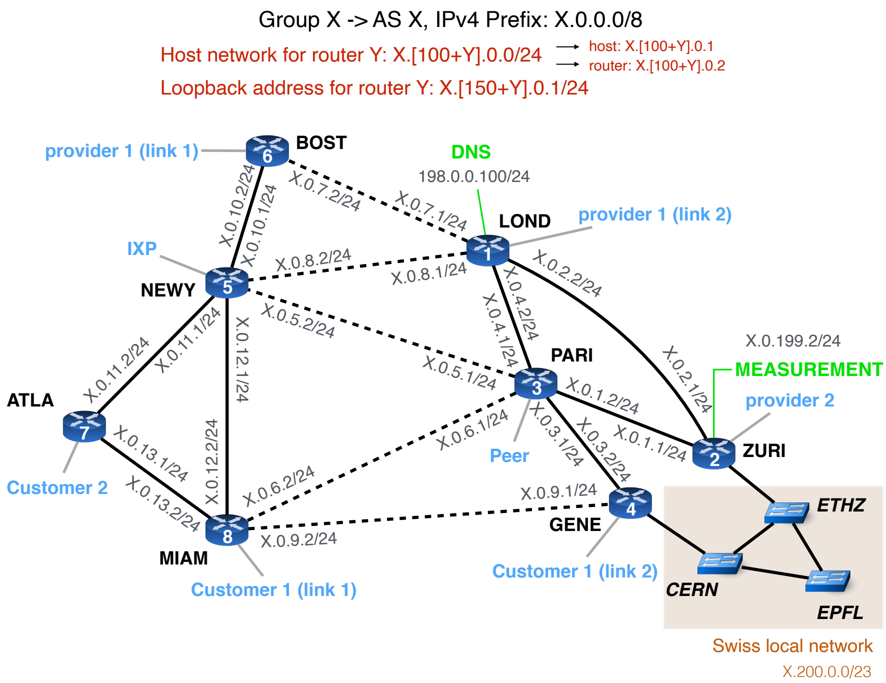

Figure 2: Each group will have to manage an entire AS. Your AS is composed of 8 routers. A /8 prefix has been assigned to each group. You can use it to configure your local networks. One host is also connected to each router, but ZURI and GENE. The subnets you must use are indicated on each interface. ZURI and GENE are connected to the Swiss local network.

3 Questions

The assignment is split in three parts: (i) intra-domain, (ii) inter-domain and (iii) policy routing. In the first part you establish connectivity within your network (intra-domain). It must be finished **within** the first two weeks of the project (before 30.03.2020). You will start by configuring the layer-2 network, followed by setting up your OSPF and iBGP configuration. The second part is all about interconnecting all the networks (inter-domain). It involves bringing your eBGP sessions with your neighboring ASes up and advertising your prefixes. The third and last part involves implementing your BGP policies according to the business relationships that you have with your neighbors. In addition to these three parts, there is also one bonus question related to VPN.

Normally, we organize a hackathon in the middle of the project to interconnect all the networks. This year, we will not have an actual hackathon due to the COVID-19 virus. We therefore adapted the assignment appropriately. For example, normally you would use the hackathon to discuss with your neighboring ASes in order to figure out which IP address to use for your peering connections. This year, we have already provided you with the appropriate IP addresses. Nonetheless, we strongly advise you to finish the first part within the first two weeks. You can then start with the inter-domain part on Monday, March 30th. This will ensure that all groups start to configure their eBGP sessions at a similar time which simplifies debugging.

To help you, we give you a crash course on how to configure FRRouting routers and Open vSwitches in a separate document available at `http://comm-net.ethz.ch/pdfs/project_1/tutorial.pdf`

For each question, we precisely tell you what you must include in your report. In addition to your report, you must also send us your switch and router configurations. To make your life easier, we provide you a script named `save_configs.sh` in the main docker container that puts all the configurations (routers and switches) in a single directory. It also generates a zip out of the directory. Then, you just need to download the zipped directory, add your report (pdf), and send it to us.

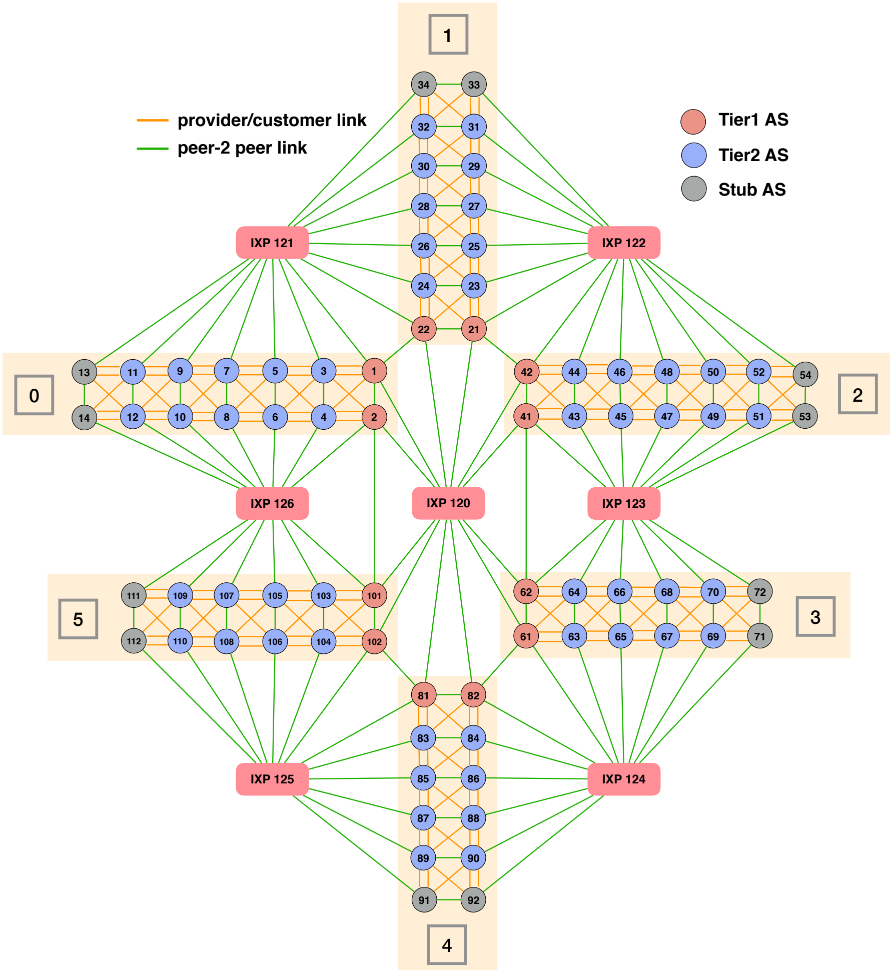

Figure 3: The AS-level topology of our mini-Internet. There are 12 Tier1 ASes, 12 stub ASes, and 54 Tier2 ASes. The topology is divided in 6 blocks (0, ..., 5), which are connected to each other via the Tier1 ASes or an IXP. The students operate the Tier2 ASes, while the Communication Networks TA team takes care of the Tier1 and the Stub ASes.

3.1 Intra-Domain Routing (to solve in the first two weeks) (3.75 points)

Question 1.1 (1.25 point)

Your goal for this question is to enable direct layer-2 connectivity between students, between staff members, but not in between them. Obviously, students and staff members should still be able to communicate between themselves, but via a layer-3 router. This will prevent typical layer-2 attacks such as MAC spoofing used to impersonate a type of user and get access to sensitive data.

To enable end-to-end connectivity, you will need to configure an IP address as well as a default gateway on each host (student and staff). For this question, you must use IP addresses belonging to the Swiss local subnet, which is X.200.0.0/23 where X is your group number. You are free to use any IP address as long as it is in that subnet. To test connectivity, you can use `ping`.

In addition, every host needs to have a standard gateway to be able to reach external destinations. Configure it such that all hosts connected to CERN and EPFL use GENE as standard gateway and the hosts at ETHZ use ZURI.

You also have to configure VLANs: use VLAN 10 for the staff and VLAN 20 for the students. VLAN 30 is reserved for later use. The interface of ZURI connected to ETHZ in VLAN 10 is named `ZURI-L2.10`, and the one in VLAN 20 is named `ZURI-L2.20` (you can see them with a **`show interface brief`** in the FRRouting CLI). The same holds for the interfaces in GENE: use `GENE-L2.10` for VLAN 10 and `GENE-L2.20` for VLAN 20. Do not use the interface `ZURI-L2` and `ZURI-L2.30`.

To include in your report: Explain what IP addresses you assigned to the different hosts. Finally, show the output for one `traceroute` from EPFL-student to EPFL-staff, one from ETHZ-staff to EPFL-student and one from EPFL-student to ETHZ-staff. In a few sentences, explain what you observe.

Question 1.2 (0.5 point)

Configure OSPF network-wide by establishing OSPF adjacencies between neighboring routers. Then, make sure to advertise all your subnets into OSPF so as to enable end-to-end connectivity between all the hosts in your AS.

Before configuring OSPF, you will have to configure all the IP addresses on each interface of your routers and hosts. Unlike for Question 1.1, you must use the IP addresses shown in Figure 2. For the router interfaces between NEWY and PARI, for example, you have to use the subnet X.0.5.0/24. The interface in PARI that is connected to NEWY must have the IP address X.0.5.1 and the interface in NEWY that is connected to PARI must have the IP address X.0.5.2 (where X is your group number).

Every router also has a loopback interface with the name `lo` that you have to configure. The router with ID Y has the loopback address X.[150+Y].0.1/24 where X is your group number (router IDs are shown on each router, for example the ID of BOST is 6). As an example, the loopback address of the router BOST for the group 10 is 10.156.0.1/24.

For the connection between the routers and their corresponding host, you have to use the subnet X.[100+Y].0.0/24, where X is your group number, and Y is the ID of the router. Then, the host gets the IP address X.[100+Y].0.1 and the interface of the router that is connected to this host will have the IP address X.[100+Y].0.2. For example, the subnet used for group 85 between the MIAM router and the corresponding host is 85.108.0.0/24. The interface at the router MIAM that is connected to the host, is called `host` and uses 85.108.0.2/24. The interface of the host connected to the router is called `MIAMrouter` and uses the IP address 85.108.0.1/24.

Be sure that each host can ping its directly connected router. Then, you can start configuring OSPF.

Verify that the subnet of the DNS server and the measurement container are visible in OSPF (for instance with **`show ip route ospf`**). From now on, always prefer to launch `traceroute` from the hosts because they can use the DNS service (routers cannot). If one host cannot access the DNS server because the OSPF configuration is not ready yet, run `traceroute` with the option **`-n`** so that it does not try to translate each IP address found on the path.

Note: Do not modify the `dns`-interface on LOND, the `measurement`-interface on ZURI, and the `matrix`-interface on PARI.

To include in your report: Include the result of a `traceroute` from PARI-host to ATLA-host.

Question 1.3 (0.75 point)

As a network operator, your goal is now to provide the best performance to your customers. In this question, your goal is to minimize latency and prevent traffic congestion.

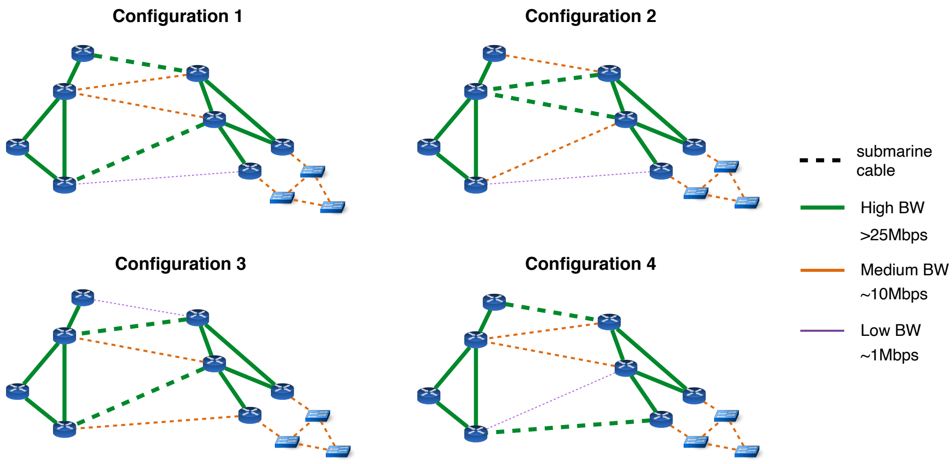

Figure 4: The bandwidth (BW) of your submarine cables follows one of the four depicted configurations.

Your top priority is to minimize latency, and to do so you must configure OSPF weights such that the traffic never traverses two submarine links (dashed links in Figure 2). For example, you do not want the traffic from BOST to MIAM to pass through Europe, but to stay on the same continent. Then, to minimize congestion, you must configure the OSPF weights such that submarine paths with higher bandwidth are preferred whenever it is possible.

For the submarine links, there are four different bandwidth configurations as shown in Figure 4. In a first step, you need to identify the configuration your AS has by using `iperf3`. Once you have identified the bandwidth configuration, you can assign the weights such that the high bandwidth links are preferred.

In addition, you need to make sure that all traffic from MIAM to NEWY is loadbalanced on the two paths MIAM-NEWY, and MIAM-ATLA-NEWY and the traffic from ZURI to LOND is loadbalanced on the two paths ZURI-LOND, and ZURI-PARI-LOND.

As a final requirement, you need to balance the traffic between ATLA and ZURI across the two submarine links with high bandwidth.

To include in your report: List all the OSPF weights you used. Then, include the results of a `traceroute` from ATLA-host to the loopback interface of ZURI. Comment the results of your `traceroutes`: do you see what you expect according to the weights you have configured, why?

Question 1.4 (0.5 point)

The Swiss local network is not able to handle large amounts of traffic. Therefore, make sure that no transit traffic goes from ZURI to GENE (or vice versa) through the Swiss local network. Only traffic destined to one of the hosts in the Swiss local network should go through these interfaces.

However, there is one exception for the traffic from ZURI to the host connected to PARI. This traffic should take the following path: ZURI-GENE-PARI.

To include in your report: Explain what technique you used to achieve the result and discuss potential drawbacks of the solution. Then, show the result of a `traceroute` from ETHZ-staff to the host connected to PARI.

Question 1.5 (0.75 point)

Configure internal BGP sessions (iBGP) between all pairs of routers (full-mesh). Verify that each one of your routers does have an iBGP session with all the other routers with the command `show ip bgp summary`.

When you establish a BGP session, you must use the loopback address for each endpoint of the connection. The loopback address is a virtual address that is always up as long as the router is running. Using the loopback interface instead of any other physical interface prevents the BGP session to go down if a physical interface becomes unavailable. To use loopback addresses for your BGP sessions, you will have to use the `update-source` command when you configure the internal BGP sessions. We explain why and how to configure it in our FRRouting tutorial.

To include in your report: Explain what `update-source` does and why you have to use it. Show the result of a `show ip bgp summary` for the router ATLA.

3.2 Inter-Domain Routing (start on Monday, 30.03.2020) (1 point)

Question 2.1 (0.5 point)

Configure the external BGP sessions (eBGP) with your neighboring ASes (including the IXPs). Normally, you would need to negotiate with your neighboring ASes and agree on which IP addresses should be used by you and your peer during the hackathon. This year, we provide you these IP addresses alongside the information about where and with whom you are supposed to have an eBGP session in the following file: `http://comm-net.ethz.ch/routing_project/as_connections`. For every eBGP session, the file shows its type (peer2peer, customer2provider or provider2customer), which router is connected to the neighboring AS, and what IP address (and subnet) you should use for the interface in your router. Table 1 is an example of what the file `as_connections` tells you if you are group 6 (AS 6).

6	ZURI	customer2provider	3	179.0.51.2/24
6	BOST	customer2provider	4	179.0.53.2/24
6	LOND	customer2provider	4	179.0.54.2/24
6	PARI	peer2peer	5	179.0.59.2/24
6	ATLA	provider2customer	7	179.0.62.1/24
6	GENE	provider2customer	8	179.0.61.1/24
6	MIAM	provider2customer	8	179.0.60.1/24
6	NEWY	peer2peer	126	180.126.0.6/24

Table 1: An example of what you can find in the file `as_connections`

Based on Table 1 we can see that AS 6 has two peers (AS5 and IXP126), two customers (AS7 and AS8) and two providers (AS3 and AS4). As an illustration, AS6 has two connections with AS4, one via its router BOST (where the IP address of the interface is 179.0.53.2/24) and one via its router LOND (where the IP address of the interface is 179.0.54.2/24). AS6 is connected to its customer AS7 via its router ATLA, and uses the IP address 179.0.62.1/24. The neighboring AS, AS 7 uses the IP address 179.0.62.2/24. This you can see when you look at the corresponding line of AS 7: `7 ZURI customer2provider 6 179.0.62.2/24`

AS6 is also connected to the IXP126 via its router NEWY. In this case, you must configure the IP address 180.126.0.6/24 on the interface of NEWY connected to the IXP. In our Internet, the AS number of an IXP is its identification number. For example, IXP126 has the AS number 126. The IP address of the IXP route server is 180.Z.0.Z with Z the IXP number. The route server of IXP126, for example, has the IP address 180.126.0.126.

Once the eBGP sessions are up, advertise your prefix to your peers. You must only advertise the /8 that has been assigned to you. Unfortunately, if you `redistribute ospf` routes into BGP, you will advertise all the /24 prefixes to your peers. In the mean time, your peers should advertise to you their /8 prefix, as well as all the /8 prefixes they have learned (since there are no BGP policies yet).

Hint: to answer this question, you will have to use the `next-hop-self` command when you configure the external BGP sessions. We explain why and how to configure it in our FRRouting tutorial.

Reminder: the IP address of the IXP Route Server is 180.X.0.X with X the IXP number.

Note: to check whether a BGP session is working and a connection has been established, you can use the command `show ip bgp summary`. You will see a list of all BGP neighbors. If there is a time entry in the column `Up/Down`, then the session has successfully been established.

To include in your report: Explain what `next-hop-self` does and why you have to use it using an example in your own network. Also, explain on which BGP sessions `next-hop-self` is required. Then, show us the results of a `show ip bgp` for the router PARI. You should see the prefixes advertised by your neighboring ASes, which would indicate that your eBGP sessions are correctly configured and that the advertisements are correctly propagated through your iBGP sessions. Then, show us that your neighboring ASes do receive the advertisement for your /8 prefix. To do that, show in your report the result of the looking glass for one router located in a neighboring AS. You should see your prefix in the looking glass. Finally, show us that you have data-plane connectivity with your neighbors by showing the result of a `traceroute` from your PARI-host to the PARI-host of one of your neighboring ASes.

Question 2.2 (0.5 point)

By default, we have configured the IXPs to not relay your BGP advertisements to their other peers. To announce a prefix to another peer via an IXP, you must specify it using a BGP community value. IXPs are configured to relay a BGP advertisement to a peer X if the advertisement has a community value equal to N:X with N the IXP number. For example, if you are AS7 and you want to advertise a prefix to AS28 via the IXP121, you must add the community value 121:28 in your BGP advertisements.

Use the community values to send BGP advertisements to the peers connected to you through an IXP.

To include in your report: Take a screenshot of the relevant parts of the `out` route-map you configured on the session from the router in NEWY to the IXP. In a few sentences explain what all the lines in the route-map mean and do. Then, show a looking glass entry of another AS which proves that your prefix has been advertised through the IXP. Finally, use the measurement container to perform a `traceroute` from another AS (in another region) to your AS for a destination where the traffic should go through the IXP. Show the result in your report.

3.3 Policy Routing (5.25 points + 0.5 bonus points)

Question 3.1 (2 points)

Configure your local-preference as well as the exportation rules to implement the customer/provider and peer/peer business relationships with your neighbors [3]. The connections you have through your IXP must be considered as peer-to-peer connections.

Hint: To configure the exportation rules, you can tag incoming routes using BGP communities to keep track of where the routes have been learned, and then match on the tag when exporting the routes. We advice you to verify with `traceroutes` or with the looking glass that the paths used do respect the business relationships.

To include in your report: Briefly explain what BGP communities you used for your peers, customers and providers. Then show a screenshot of one `in` and one `out` route-map and briefly explain the different lines in the route-map. Then, show that your configuration works properly by adding the result of the looking glass of one of your peers, which is supposed to show that this peer does receive the prefixes of your customers, but does not receive the prefixes of your other peers. Finally, launch a `traceroute` from one of your customers towards one of your peers using the measurement container. Verify that your AS forwards the packet directly to your peer and not to your provider. Include the result of the `traceroute` in your report.

Question 3.2 (1 point)

The AS topology (Fig.3) shows six main regions (0, 1, 2, 3, 4, and 5). Configure your BGP policies such that you can leverage your connection with your IXP at NEWY. You do want to peer through this IXP with ASes that are located in another region. However, for business reasons, you do **not** want to peer through this IXP with ASes that are located in the same region. To not peer through the IXP with ASes in the same region, you must (i) not advertise them any prefixes and (ii) deny any advertisements coming from them.

To check whether you properly configured (ii), we have configured the stub ASes to advertise their prefix to all the ASes connected to their IXP.

To include in your report: Again take a screenshot of the relevant parts of the route-map at NEWY and explain what the different lines mean and do. Show that the advertisement from the stub AS in the same region and connected to the same IXP as you is denied by showing the result of a `sow ip bgp` in your router NEWY. For clarity, you do not need to write the full output, just the part that is interesting (*i.e.,* the part which could have the prefix of the stub AS). Then, include in the report the output of the looking glass for the router NEWY of the stub AS in the same region and connected to the same IXP as you. Finally, include in your report the output of the looking glass for a group in another region but connected to the IXP. When you include the output of a looking glass in your report, only keep the parts that prove the correctness of your configuration and omit the irrelevant ones.

Question 3.3 (0.75 point)

In this question, the goal will be to configure your BGP policies in order to influence the **inbound** traffic destined to your **own** prefix. More precisely, your goal is to configure BGP policies such that the inbound traffic coming from a provider and destined to your own prefix uses the provider connected to ZURI in priority.

To include in your report: Explain in a few sentences the technique you used and discuss any potential drawbacks. Then, include the result of the looking glass for both of your providers. You can omit parts of the output that are irrelevant, and only show the part that shows that your configuration is correct (*i.e.,* the part where your own prefix is shown).

Question 3.4 (0.75 point)

In this question, the goal will be to configure your BGP policies such that the **inbound** traffic coming from the provider with whom you have two external BGP sessions (*e.g.,* AS 15 if you are AS 17) arrives in priority via your router BOST (instead of LOND).

To include in your report: Explain in a few sentences the technique you used and discuss any potential drawbacks. Then, include the result of the looking glass of your provider with whom you have two external BGP sessions which shows that your configuration is working as expected (for instance from the router MIAM). You can omit parts of the output that are irrelevant.

Question 3.5 (active from 06.04.2020 on) (0.75 point)

As we have seen in the lecture, not all parties in the Internet play a fair game. You realize that part of your /8 subnet is being hijacked by another AS. In this question, you first have to figure out which prefix is being hijacked. Then, you should try to nullify the ongoing attack and attract the hijacked traffic back to your AS.
Hint: to ensure full connectivity while configuring your eBGP sessions, the hijacks will only be active during the last week of the project (starting from Monday, April 6th).

To include in your report: Explain how you found the hijacked prefix space and describe your countermeasures. Would your current solution to mitigate the hijack work for every possible hijack?

Bonus question 3.6 (0.5 point)

Although you can already connect to the mini-Internet over ssh, it would also be interesting to directly connect one of your devices with the mini-Internet, e.g. to test applications which run locally. To achieve that, you can use a VPN service. Each group has two VPN servers running, one connected to CERN and one to EPFL. Have a look at the tutorial file to understand how you can connect a local VPN client with the mini-Internet.

The VPN traffic uses the VLAN number 30 in the layer 2 network. Therefore, extend the layer 2 configuration as follow: First, make sure that you configure the interface `ZURI-L2.30` (on the ZURI router) with the IP X.200.30.1/24. Similarly, use the IP X.200.30.2/24 on the interface `GENE-L2.30` in the GENE router (X your group number). The two VPN servers already know that they should use these IP addresses as gateway. Afterwards, extend your layer 2 trunk links to include the VLAN 30 and add the VLAN tag 30 to the two interfaces connecting to the VPNs in CERN and EPFL. Finally, follow the tutorial document to get a working VPN connection from a local device.

In the report, answer the following questions:

a) Just after establishing the VPN connection, which prefix space of the mini-Internet are you able to reach from your local device?

b) What do you have to configure locally in order to be able to ping a host outside of your AS (using the VPN connection)?

c) Assuming you want to ping the host connected to ATLA in your network from your local device (using the VPN). How could you confirm that your command `ping X.107.0.1` (run locally) is indeed going over the VPN and reaches the ATLA host inside the mini-Internet and not an actual server in the real Internet with the IP X.107.0.1 (X your group number)?

d) Now a second group member establishes a VPN connection towards the other VPN server on a different local device while the first VPN connection is still running. Perform a traceroute between the two VPN endpoints over the mini-Internet. Explain which path the traffic takes.

4 Testing and verifying your network configurations

As any network operator, you must verify that your configuration does what you want, and debug it in case something goes wrong. We offer you several tools that you can use to verify your configuration. These tools are similar to the ones network operators use in practice.

Measurement container We have setup a measurement container which will enable you to launch `traceroutes` from any Tier2 AS (and not necessarily only your own AS), towards any destination in the mini-Internet. This will help you to know the paths used *towards* your network. The measurement container is connected to each AS via the interface `measurement_X` of the router ZURI. The IP address of this interface is pre-configured and follows the convention X.0.199.1/24 (see Fig. 2), with X your group number. For example if you are group 15, your pre-configured IP address on the interface `measurement_15` at ZURI will be 15.0.199.1/24. The X.0.199.1/24 subnet must be reachable from anywhere in your network. You must therefore add it in your OSPF configuration. To access the measurement container, use the following command:

```
> ssh -p 2099 root@snowball.ethz.ch
```

The password will be available in the `#routing_project` Slack channel. To launch a traceroute, you can use the script `launch_traceroute.sh`, which takes as argument the group number from which the traceroute starts, and the destination IP address (possibly in another AS). For instance, if you want to perform a traceroute from group 11 to 22.107.0.1 (*i.e.,* the host connected to ATLA in group 22), just use the following command in the measurement container:

```
> ./launch_traceroute.sh 11 22.107.0.1
```

Note that the traceroute will start from the router ZURI of group 11, since the measurement container is connected to that router. In practice, network operators can use large-scale Internet measurement platforms such as RIPE Atlas[2] to assess the connectivity of their network from outside.

DNS service To help you decoding your traceroute output, we have setup a DNS server and have pre-configured your hosts to use it. If the DNS server is used, the IP addresses in the traceroute output will be replaced by the corresponding router names. For example, 19.0.2.2 will be translated into LOND-ZURI.group19, because this is the IP address configured on the interface `port_ZURI` of the router LOND in AS19. The DNS server is located in a container connected to the interface `dns` of the router LOND. The IP address on that interface is pre-configured, you do not need to modify it. Also, each host is pre-configured to use the DNS server. Only hosts use the DNS server, routers do not. Of course, hosts can only use the DNS server if they can reach the network 198.0.0.0/24 (where the DNS server is located). As such, do not forget to include this network in your OSPF configuration.

BGP looking glass We have setup a looking glass service. In practice, looking glasses are servers remotely accessible which display the routing information of an IP router. For example, SWITCH, the Swiss educational network, gives public access to its looking glass[3]. This is useful to see how your BGP advertisements look like from a remote point of view. For this assignment, we make publicly available on our website (under the `looking_glass` directory) one file per group and router showing the result of a `show ip bgp`. The files for group X are in the directory GX, there exists one file for each router. For example, if you want to get the result of a `show ip bgp` at PARI for group 23, download the file `http://comm-net.ethz.ch/routing_project/looking_glass/G23/PARI.txt`. These files are updated every minute.

Connectivity matrix We also provide you with a connectivity matrix, which shows you whether two ASes can ping between each other. Before you setup the eBGP sessions, everything will be red. As soon as you will setup the eBGP sessions with your neighbors, the matrix will turn green for some pairs of ASes. At the end of this assignment, we hope to see this matrix completely green! The matrix is available at `http://comm-net.ethz.ch/routing_project/matrix/matrix.html` and is updated every few minutes.

[2] `https://atlas.ripe.net`
[3] `https://www.switch.ch/network/tools/lookingglass/`

5 Frequently Asked Questions

I have a problem when running a traceroute. It takes quite some time and the DNS service does not work.

When you run a `traceroute`, keep in mind that only the hosts connected to the routers are configured to use the DNS service. If you run a `traceroute` from a router, it will not translate the IP addresses. If the DNS service is not reachable (*e.g.,* because you have not configured OSPF yet), run the `traceroute` with the option `-n`. This will tell `traceroute` to not translate the IP addresses, and it will save you some time.

How can I erase a configuration I have done on a FRRouting router?

To erase a configuration on a FRRouting router, it's very easy, just use the same command you used, but add "`no`" at the beginning. For example, if you configured an IP address with `ip address 1.0.0.1/24`, just run `no ip address 1.0.0.1/24` to erase it.

I have the error `"VTY configuration is locked by other VTY"` when I run a `conf t` on a FRRouting router.

Only one VTY session can configure a FRRouting router at a time. If you have this problem, it's either because one member of your group is already configuring this router, or because VTY sessions are still running in the background and block your access to the router (for example because you lost a previous ssh connection). In the latter case (and you are sure that nobody of your group is currently configuring this router), please send us a message on Slack and we will fix it as soon as possible.

My neighboring ASes are not active and I can't get connectivity with the rest of the mini-Internet because of that.

Although each group has two providers, two customers and two peers, it can happen that some of them are inactive or have misconfigured BGP, which makes you unreachable from some parts of the mini-Internet. This can also prevent you to test your configuration or run `ping` or `traceroute`. If you experience this problem when you answer the questions, please describe it in your report and explain what you were not able to do because of it. If this is really a big problem (*e.g.,* both of your providers are inactive and you can't reach the rest of the mini-Internet), please let us know and we will find a way to solve the problem. If you could not show something in your report due to failures of your neighbors, it will not negatively influence your grade!

References

[1] B. Pfaff, J. Pettit, T. Koponen, E. Jackson, A. Zhou, J. Rajahalme, J. Gross, A. Wang, J. Stringer, P. Shelar, K. Amidon, and M. Casado, "The design and implementation of open vswitch," in 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15). Oakland, CA: USENIX Association, 2015, pp. 117–130. [Online]. Available: https://www.usenix.org/conference/nsdi15/technical-sessions/presentation/pfaff

[2] FRRouting. [Online]. Available: https://frrouting.org/

[3] L. Gao and J. Rexford, "Stable internet routing without global coordination," SIGMETRICS Perform. Eval. Rev., vol. 28, no. 1, pp. 307–317, Jun. 2000.