IPv6 Module 16 – An IPv6 Internet Exchange Point

Objective: To investigate methods for connecting to an Internet Exchange Point.

Prerequisites: Modules 12, 14 and 15, and the Exchange Points Presentation

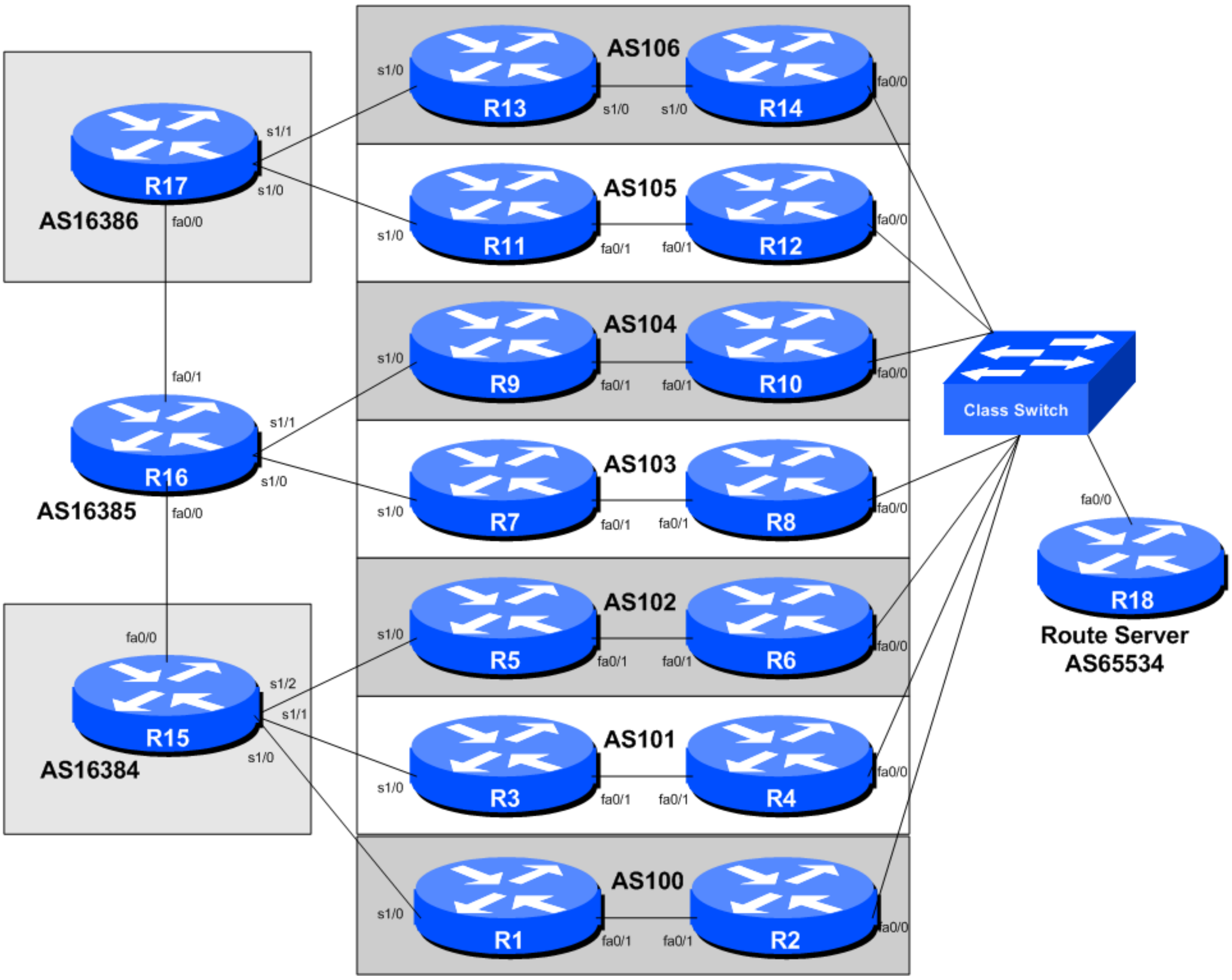

Figure 1 – IXP Configuration

Lab Notes

The purpose of this module is to introduce the concept of an Internet Exchange Point, how to peer at IXPs, and look at some of the recommended configuration practices.

It has seven ASNs participating at the IX, with two routers being assigned to each autonomous system. One router peers at the IX, the other router is internal to the autonomous system (so will have an OSPF and iBGP session with the IX facing router). It also has three routers operating as three transit provider or Tier1 ISPs. This version of the module is incremental over Module 10 and covers correct OSPF and iBGP configuration practices for connecting the IXP facing router to the rest of the ISP network.

Lab Exercises

1. **A More Advanced IXP.** This IXP example has seven participating ASNs with each ASN having two routers. One router peers at the IXP; the other router is internal to the ISP's network and connects to an upstream ISP. The two routers in each ASN present at the IXP use OSPF and iBGP to communicate routing information with the each other. The three upstream ASNs are also interconnected with each other, to represent the high speed Internet core. What we have done here is to simulate an Internet Exchange Point and its participants' connections to upstream ISPs.

2. **Basic Configuration.** Each router team should configure their router to fit into the network layout depicted in Figure 1. Check all connections. Note that all links to the IX switch are by Ethernet. The links between each router in each AS is by serial connection (to represent the WAN connection from the IX back to the ISP's core network). The links from each AS to their upstream is by Ethernet (to represent the WAN link to the upstream).

3. **Addressing Plan.** These address ranges should be used throughout this section of the module. You are welcome to use your own range within an AS if you desire, just so long as you consult with the teams in other ASes to ensure there is no overlap.

AS100	**1001:100::/32**	**AS105**	**1001:105::/32**
AS101	**1001:101::/32**	**AS106**	**1001:106::/32**
AS102	**1001:102::/32**	**AS16384**	**1001:4000::/32**
AS103	**1001:103::/32**	**AS16385**	**1001:4001::/32**
AS104	**1001:104::/32**	**AS16386**	**1001:4002::/32**

4. **Basic Router Setup.** Set up the routers as you would have done in previous modules. That is, basic security, IP addressing etc. The addressing for the point-to-point link between the two routers in each AS is left as an exercise to the Router Teams. But remember, in IPv6 a point-to-point link requires a /64 address block (address it as a /127). Don't forget to configure the loopback interfaces too.

5. **Management Router and IXP LAN.** The address range used for the IXP LAN is 1234:3:2:1::/64 – the route server (if provided in this module) has an IP address of 1234:3:2:1::1. Each of the ASes is assigned a /112 out of the IXP LAN block to use on the exchange point LAN. So, for example, AS100 has 1234:3:2:1::100:1, 1234:3:2:1::100:2, 1234:3:2:1::100:2, and so on. AS101 has 1234:3:2:1::101:1, 1234:3:2:1::101:2, 1234:3:2:1::101:3, etc. And so on.

   Q. Why do you think the address assignment at the IXP has been done like this?

6. **Configure the ethernet of each router at the IXP.** The ethernet interfaces connected to the IXP should be configured appropriately for a public connection. Review the IOS Essentials materials and the IXP presentation. The configuration for AS102 might be:

```
interface fastethernet 0/0
 description Exchange Point LAN
 ipv6 address 1234:3:2:1::102:1/64
!
```

   If you are unclear as to what any of the configuration lines do, please ask the lab instructor.

<u>Checkpoint #1:</u> *When you have properly configured your router, and the other routers at the IXP are reachable (i.e. you can ping the other routers), please let the instructor know.*

7. **Configuring OSPF (part 1).** Obviously each IXP team will need to configure ISIS between the two routers in their AS. The router which is internal to the ASN's network should be straightforward to configure. It has one loopback interface, one active interface pointing to the IXP facing router, and one interface pointing to the upstream. OSPF should only be activated on the internal interface, with the loopback marked as passive. Note we do not configure the interface pointing to the upstream as we will be using *next-hop-self* for our iBGP sessions. Example for Router 11 might be (borrowing from the IPv4 lab):

```
ipv6 router ospf 105
 log-adjacency-changes
 passive-interface default
 no passive-interface serial 1/0
!
interface serial 1/0    ! link to Router 12
 ipv6 ospf 105 area 0
!
interface loopback 0
 ipv6 ospf 105 area 0
!
```

8. **Configuring OSPF (part 2).** The OSPF configuration on the IXP facing router needs more care. It is highly important that the IXP LAN network **does not** ever appear in the ISP's routing table, either in OSPF or in iBGP. If it does, there is the possibility of leaking the IX LAN network by the ISP's BGP to other ASNs, providing transit to the IXP. This might not be a good thing, and many IXPs have rules against this behaviour.

   The OSPF configuration for the IXP router looks something like this, for example for Router 8:

```
ipv6 router ospf 103
 log-adjacency-changes
 passive-interface default
 no passive-interface serial 1/0
!
interface serial 1/0    ! link to Router 7
 ipv6 ospf 103 area 0
!
interface loopback 0
 ipv6 ospf 103 area 0
!
```

```
interface fastethernet 0/0
 ipv6 ospf 103 area 0
!
```

NB. Do NOT mark OSPF as "no passive" on the Ethernet interface. Do **NOT** mark it as "no passive", repeat, do **NOT**! It is **NOT** a mistake.

9. **Information Only: Configuring iBGP on non-IXP facing routers.** There are only two routers per AS in this Module – this step only applies to iBGP peerings with routers other than the IXP router. The iBGP configuration on the internal routers is straightforward enough, and is as has been done in early workshop modules. For example, for Router 7:

```
router bgp 103
 address-family ipv6
  no synchronization
  network 1001:103::/32
  neighbor ibgp-peersv6 peer-group
  neighbor ibgp-peersv6 remote-as 103
  neighbor ibgp-peersv6 update-source loopback0
  neighbor ibgp-peersv6 next-hop-self
  neighbor ibgp-peersv6 password cisco
  neighbor ibgp-peersv6 description iBGP peers
  neighbor ibgp-peersv6 send-community
  neighbor router8-loopback peer-group ibgp-peers
  no auto-summary
!
ipv6 route 1001:103::/32 null0
!
```

10. **Configuring iBGP between upstream facing routers and the IXP facing router.** Any router which is connected to other ASNs (the ISP border routers) such as the ISP's upstreams or ISP's peers needs to have a modified iBGP configuration when peering with the IXP facing router. If all the prefixes learned from the upstream ISP are passed onto the IXP facing router, then there is the possibility that the IXP peers (and non-peers) can point static routes for those destinations at the IXP facing router, thereby gaining outbound transit across the local ASN. This is undesirable, and is basically a security risk. The configuration in the previous step therefore needs to be modified for IXP facing routers (the only router in this lab module example) so that external prefixes are not announced to the IXP facing router. In real life, ISPs create a peer-group which is used only when peering with IXP-facing routers, for example:

```
router bgp 103
 address-family ipv6
  no synchronization
  network 1001:103::/32
  neighbor ibgp-ixpv6 peer-group
  neighbor ibgp-ixpv6 remote-as 103
  neighbor ibgp-ixpv6 update-source loopback0
  neighbor ibgp-ixpv6 next-hop-self       ! set external next-hops to us
  neighbor ibgp-ixpv6 password cisco
  neighbor ibgp-ixpv6 description iBGP peering with IXP routers
  neighbor ibgp-ixpv6 send-community
  neighbor ibgp-ixpv6 filter-list 10 out
  neighbor router8-loopback peer-group ibgp-ixpv6
  no auto-summary
!
ip route 1001:103::/32 null0
!
```

```
ip as-path access-list 10 permit ^$
!
```

Note the addition of the *as-path access-list 10* – this basically allows only locally originated prefixes to reach the IX facing router. For example, Router8 with this configuration will now no longer send externally learned prefixes to Router7, thereby guaranteeing the security of the AS100 network.

11. **Configuring iBGP on IXP facing routers (Part 1).** The iBGP configuration on the IXP facing router needs much more care. We do **NOT** want the IXP LAN to appear in our iBGP. And because the IXP LAN is not in OSPF, we can't use this net for valid next-hops. So that BGP has a valid next-hop, we again use the `next-hop-self` BGP configuration as we have done for our other iBGP sessions earlier on. For example, for Router 6:

```
router bgp 102
 address-family ipv6
  no synchronization
  neighbor ibgp-peersv6 peer-group
  neighbor ibgp-peersv6 remote-as 102
  neighbor ibgp-peersv6 update-source loopback0
  neighbor ibgp-peersv6 next-hop-self     ! set external next-hops to us
  neighbor ibgp-peersv6 password cisco
  neighbor ibgp-peersv6 description iBGP peers
  neighbor ibgp-peersv6 send-community
  neighbor router5-loopback peer-group ibgp-peersv6
!
```

12. **Configuring iBGP on IXP facing routers (Part 2).** The second important difference required for the iBGP at the IXP router is that the local address block must **NOT** be originated here. If the IX router is disconnected from the core backbone for whatever reason, if it originated the ISP's address block, it would end up blackholing all traffic from the IXP participants to the local network. This is highly undesirable, so the correct procedure is to originate the ISP address block only in the core of the network, not the edge, so that normal BGP failover rules can apply.

    Q. What do you think the normal BGP failover rules might be in this case?

    A. If the IXP facing router is disconnected from the core network, the iBGP fails, the OSPF fails. So the router no longer hears the announcement of the ISP's address block from the core, so no longer announces it to the IXP peers, its eBGP neighbours. When this happen, the other ASNs will use fall back paths (alternative to the IXP) to reach the local ASN. These are the normal BGP failover rules.

13. **Configuring iBGP on the IXP facing routers (Part 3).** The second very important requirement at an Internet exchange point is that a default route (or the routes not originated by the local ASN) should not be made available at the peering LAN. If this requirement is not followed, then it is possible for the Exchange Point participants to use the local network to get transit to the rest of the Internet.

    In this Module, each ASN will be receiving a default route from their upstream provider (a common scenario in the Internet today). This default route must **NOT** be made available at the ISP's Exchange Point router. There are two ways to do this. Either block the announcement of the default route by iBGP at the borders router which connect to the upstreams, or set up a static default route to the Null interface at the IXP router (or indeed do both!). For example:

```
ipv6 route ::/0 null0
```

Then if any of the IXP participants point a default route to the local network, the traffic will simply be dumped in the Null interface of their IXP router. Only traffic for specific destinations which available in the routing table on the IXP router will be forwarded to the rest of the network. This is a very important **network security** requirement.

14. **Configuring eBGP on the IXP facing routers.** Next, eBGP needs to be set up on the IXP routers. Create a peer-group and apply that peer-group to each eBGP neighbour. A sample configuration for Router8 might be:

```
ipv6 prefix-list myprefixes permit 1001:103::/32
ipv6 prefix-list peer100 permit 1001:100::/32
..
ipv6 prefix-list peer106 permit 1001:106::/32
!
router bgp 103
 address-family ipv6
  no synchronization
  bgp log-neighbor-changes
  neighbor ixp-peersv6 peer-group
  neighbor ixp-peersv6 remove-private-AS
  neighbor ixp-peersv6 prefix-list myprefixes out
  neighbor ixp-peersv6 route-map set-local-pref in
  neighbor <router2> remote-as 100
  neighbor <router2> description Peering with AS100
  neighbor <router2> peer-group ixp-peersv6
  neighbor <router2> prefix-list peer100 in
..
  neighbor <router14> remote-as 106
  neighbor <router14> description Peering with AS106
  neighbor <router14> peer-group ixp-peersv6
  neighbor <router14> prefix-list peer106 in
!
route-map set-local-pref permit 10
 set local-preference 150
!
```

The configurations for the other routers will be similar to this one. All router teams will have done sufficient BGP configuration throughout this workshop to extrapolate from the above examples. If in any doubt, ask the lab demonstrator for assistance.

Note the prefix-lists. There is a per-peer inbound prefix-list. Some service providers only filter ASes – that has inherent dangers, and does not prevent against inbound leaking of prefixes incorrectly originated by the peer AS. But only filtering on prefixes doesn't scale, especially in larger IXPs with large participating service providers as they are frequently adding to the prefixes they announce. The Internet Routing Registry is usually used to solve this problem.

15. **Set up passwords on the eBGP sessions at the IXP.** Negotiate with each ASN a password which you can use on your BGP session with them. And then agree to cut the eBGP session over to using passwords such that the eBGP session does not fall over due to password mismatches (as in Module 2). An excerpt from Router 8's configuration might be:

```
router bgp 103
```

```
  ...
  neighbor <router10> password peer104
  ...
 !
```

16. **Configuring eBGP between the routers in AS16384, AS16385 and AS16386.** The instructors will operate the three transit/Tier1 ASNs; they will now configure eBGP between each other. No filtering is expected or required between the three ASNs, so the eBGP configuration is quite straightforward, for example as on Router 15:

```
router bgp 16384
 address-family ipv6
  no synchronization
  bgp log-neighbor-changes
  neighbor <router16> remote-as 16385
  neighbor <router16> password as16385-password
  neighbor <router16> description Peering with AS16385
!
```

17. **Configuring eBGP with the routers in AS16384, AS16385 and AS16386.** All IXP members must now configure eBGP sessions with the routers in the three transit ASNs. They must expect to receive the default route from their upstreams, and only send their address block to the upstream. An example configuration for Router 2 might be:

```
ipv6 prefix-list myprefixes permit 1001:101::/32
!
router bgp 101
 address-family ipv6
  no synchronization
  bgp log-neighbor-changes
  neighbor <router15> remote-as 16384
  neighbor <router15> description Peering with AS16384
  neighbor <router15> password as16384-password
  neighbor <router15> prefix-list default in
  neighbor <router15> prefix-list myprefixes out
!
```

18. **Connectivity Test.** Check connectivity throughout the IXP network. Each router team should be able to see all the other routers at the IXP. When you are satisfied that BGP is working correctly, try running traceroutes to check the paths being followed.

```
Router9#trace 1001:101::1

  1 a:b:c:d 4 msec 4 msec 4 msec
  2 1234:3:2:1::103:1 4 msec 4 msec 4 msec
  3 c:d:a:b [AS 106] 4 msec *  4 msec
Router9#
```

Checkpoint #2: *Once the BGP configuration has been completed, check the routing table and ensure that you have complete reachability over the entire network. If there are any problems, work with the other router teams to resolve those.*

STOP AND WAIT HERE

19. **Implementing a Route Server.** With route server functionality added to Cisco IOS as from the 15.2 release, it is now possible to operate Router 18 as a route server. The lab instructors will have configured Router 18 to be a route server.

    The following information is included for the information of each router team member only. For those in the workshop interested in setting up an Internet Exchange Point following the completion of the workshop, this section contains the information necessary to configure and operate a Cisco IOS based Route Server.

    The Route Server has eBGP sessions configured with each member of the IXP in this workshop lab (in real life, the Route Server has an eBGP session configured with those members of the IX who wish to use its facilities.) Each eBGP session is configured with the "route-server-client" command included in the eBGP peer-group.

    This is an excerpt from the configuration of Router 18 after it has been configured to be a route server:

```
router bgp 65534
 bgp log-neighbor-changes
 neighbor v6ixp-peers peer-group
 neighbor v6ixp-peers description Route Server peers
 neighbor v6ixp-peers password cisco
 neighbor 1234:3:2:1::100:1 remote-as 100
 neighbor 1234:3:2:1::100:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::100:1 description AS100 peer
 neighbor 1234:3:2:1::101:1 remote-as 101
 neighbor 1234:3:2:1::101:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::101:1 description AS101 peer
 neighbor 1234:3:2:1::102:1 remote-as 102
 neighbor 1234:3:2:1::102:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::102:1 description AS102 peer
 neighbor 1234:3:2:1::103:1 remote-as 103
 neighbor 1234:3:2:1::103:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::103:1 description AS103 peer
 neighbor 1234:3:2:1::104:1 remote-as 104
 neighbor 1234:3:2:1::104:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::104:1 description AS104 peer
 neighbor 1234:3:2:1::105:1 remote-as 105
 neighbor 1234:3:2:1::105:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::105:1 description AS105 peer
 neighbor 1234:3:2:1::106:1 remote-as 106
 neighbor 1234:3:2:1::106:1 peer-group v6ixp-peers
 neighbor 1234:3:2:1::106:1 description AS106 peer
 !
 address-family ipv6
  neighbor v6ixp-peers route-server-client
  neighbor 1234:3:2:1::100:1 activate
  neighbor 1234:3:2:1::101:1 activate
  neighbor 1234:3:2:1::102:1 activate
  neighbor 1234:3:2:1::103:1 activate
  neighbor 1234:3:2:1::104:1 activate
  neighbor 1234:3:2:1::105:1 activate
  neighbor 1234:3:2:1::106:1 activate
  distance bgp 200 200 200
 exit-address-family
!
```

Notice that in this case there are no BGP filters at all – it is left to the IXP member to do necessary filtering.

This is just a simple example. Most IXPs will insist on the route server doing the basic bogon filtering which has been discussed elsewhere during this workshop. Some IXPs operate their route servers to filter according to the information contained in the Internet Routing Registry.

20. **Set up eBGP session with Router 18.** Each team member should now configure their router to peer eBGP with the route server. We will leave the existing bi-lateral peering sessions in place for the time being. Here is an example configuration:

```
router bgp 103
...
 neighbor 1234:3:2:1::1 remote-as 65534
 neighbor 1234:3:2:1::1 description eBGP with the IX Route Collector
 neighbor 1234:3:2:1::1 password ixp-collector
 neighbor 1234:3:2:1::1 remove-private-AS
 neighbor 1234:3:2:1::1 prefix-list deny-all in
 neighbor 1234:3:2:1::1 prefix-list myprefixes out
...
!
ipv6 prefix-list deny-all deny ::/0 le 128
...
```

Notice that the route server is running in a private AS – there isn't really any need for it to use a public AS as the Server does not need to be directly visible outside of the IXP.

Note also that there is no inbound prefix filter. This is not normally the case for an IXP. Even if an ISP peers with the Route Server, they should still filter inbound BGP prefixes they receive via the Route Server, just in case a peer who uses the Route Server advertises prefixes they should not be advertising.

21. **Modifications to the ISP's IXP routers to talk to the Route Server.** Route servers, as a rule, do not include their AS number in BGP announcements to their clients; they are simply there to distribute the BGP table they have learned from their clients. Historically RSd and routing software platforms such as Zebra and Quagga have always behaved this way when running in route server mode. Some IXPs used routers instead of these software routing implementations, and routers have not had the functionality to drop the local ASN, meaning that the Route Server's AS has appeared first in the AS path.

Cisco routers participating at IXPs with the software based route servers have long had a capability to accept peerings with a full Route Server AS – but it is not IOS's default behaviour. IOS's default behaviour is to check that all prefix announcements from an eBGP peer have the peer AS as the first AS in the AS path. To support an eBGP peering where the Route Server AS is not included in the AS path, the router needs the following configuration:

```
router bgp 103
 no bgp enforce-first-as
...
```

which stops the router rejecting the prefix updates if the peer AS is not first in the AS path. You will notice that before entering this command, the BGP peering will be up, but no prefixes will be received by the ISP’s IXP router – and the router logs will have messages about malformed attributes being sent by the Route Server.

Checkpoint #3: *Now look at your BGP routing table. You should see two paths to each destination, one via the RS, the other via the bi-lateral BGP peer.*

22. **Removing non-RS BGP sessions.** Once the peering session has been established with the Route Server and prefixes from all ASNs are being received from the Route Server, then each router team can remove their direct eBGP configurations with each AS at the IXP. While many ISPs choose to continue with bi-lateral peering configuration as well as the Route Server peering, this step here shows what happens when just the RS is used for the eBGP sessions at the IX:

```
router bgp 100
 no neighbor 1234:3:2:1::101:1
 no neighbor 1234:3:2:1::102:1
 no neighbor 1234:3:2:1::103:1
 no neighbor 1234:3:2:1::104:1
 no neighbor 1234:3:2:1::105:1
 no neighbor 1234:3:2:1::106:1
!
```

23. **Local Preference on RS supplied prefixes.** The prefixes heard from the Route Server all have the default local preference of 100. What each team should do now is use the previous route-map set-local-pref an applied it to the BGP peering with the RS. For example:

```
router bgp 102
 neighbor 1234:3:2:1::1 route-map set-local-pref in
```

and then refresh the BGP session with the route-server. All the prefixes learned from the route-server will now have the local-preference set to 150.

24. **Filtering RS supplied prefixes.** It is generally quite a good idea not to trust what the route-server is sending to you. Each team should now take the individual prefix filters they had for each neighbour and **replace** them with a combined prefix filter for the route-server. This way each AS is still filtering what it receives from each AS. Note that in practice, many IXPs operating a route-server will filter the prefixes they receive from each participant in the route-server, and several more will filter what they send to each participant as well. This all helps with the integrity of the IXP peering mesh. For example:

```
router bgp 104
 neighbor 1234:3:2:1::1 prefix-list RS-v6prefixes in
!
ip prefix-list RS-v6prefixes permit 1001:100::/32
ip prefix-list RS-v6prefixes permit 1001:101::/32
ip prefix-list RS-v6prefixes permit 1001:102::/32
ip prefix-list RS-v6prefixes permit 1001:103::/32
ip prefix-list RS-v6prefixes permit 1001:104::/32
ip prefix-list RS-v6prefixes permit 1001:105::/32
ip prefix-list RS-v6prefixes permit 1001:106::/32
!
```

Don’t forget to tidy up unused configuration by removing the individual prefix-lists used in the previous section for the individual (bi-lateral) peerings.

Checkpoint #4: *Now look at your BGP routing table. You should only see prefixes from each AS, but learned via the route server.*

25. Summary. This module has given example configurations as used by Internet Service Providers at Internet Exchange Points. It has concentrated on using prefix-lists only – more sophisticated configurations are possible by using as-path filters and BGP communities. These examples are left to the reader to consider. If there is time at the end of the workshop, ask the Instructor to test out some other scenarios.