Module 16 – An Internet Exchange Point

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

Prerequisites: Modules 12 and 13, and the Exchange Points Presentation

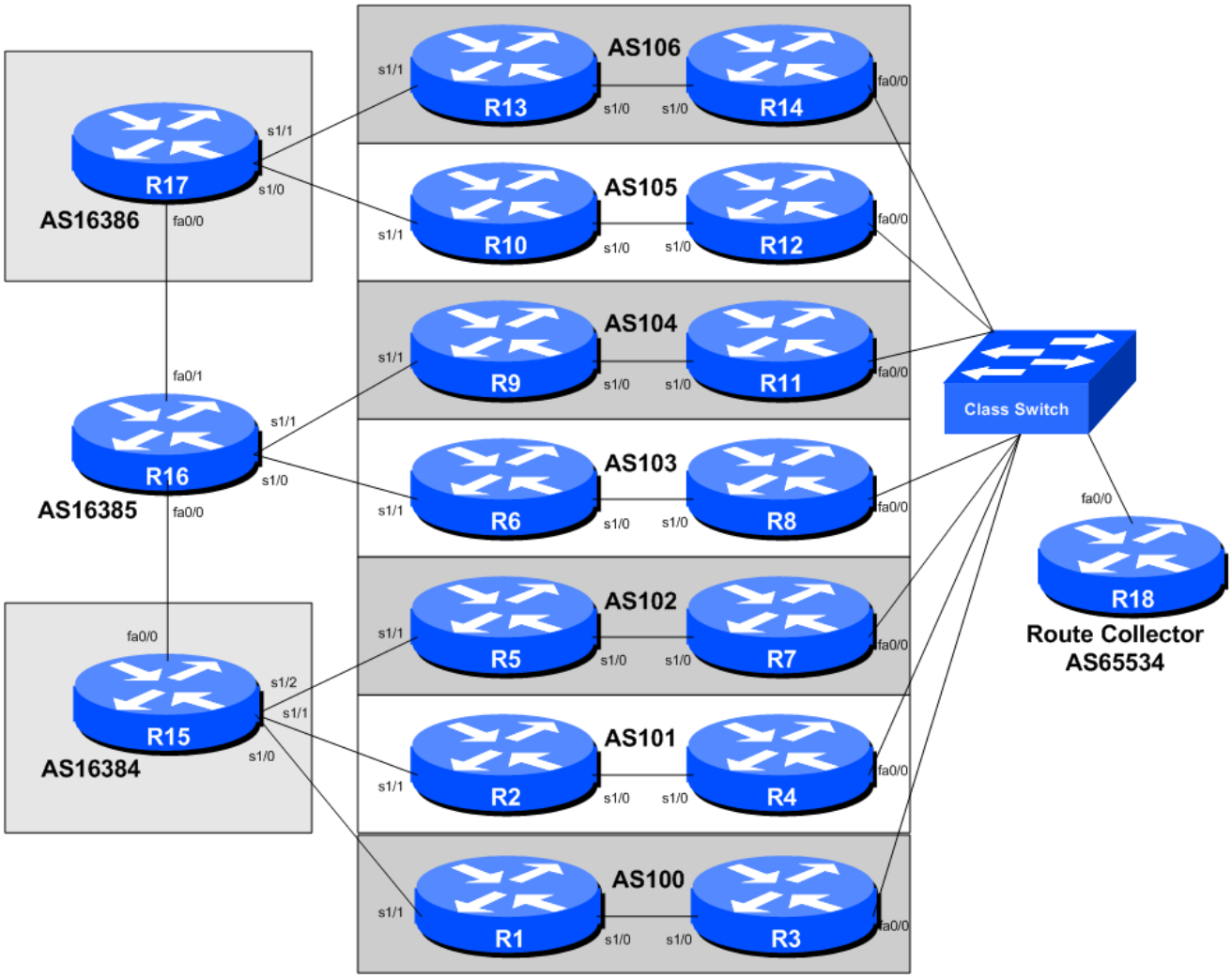

Figure 1 – IXP Configuration

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	**10.1.0.0/20**	**AS105**	**10.6.80.0/20**
AS101	**10.2.16.0/20**	**AS106**	**10.7.96.0/20**
AS102	**10.3.32.0/20**	**AS16384**	**10.8.112.0/20**
AS103	**10.4.48.0/20**	**AS16385**	**10.9.128.0/20**
AS104	**10.5.64.0/20**	**AS16386**	**10.10.144.0/20**

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, a point-to-point link requires a /30 address block. Don't forget to configure the loopback interfaces too.

5. **IXP LAN.** The address range used for the IXP LAN is 172.17.10.0/24 – the route collector (if provided in this module) has an IP address of 172.17.10.254. Each of the ASes is assigned addresses sequentially for use on the exchange point LAN. So, for example, AS100 has 172.17.10.1; AS101 has 172.17.10.2; AS102 has 172.17.10.3. And so on.

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 Router 3 might be:

```
interface fastethernet 0/0
 description Exchange Point LAN
 ip address 172.17.10.1 255.255.255.0
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
```

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

Checkpoint #4: *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 OSPF 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 interface connecting to the router at the IXP, and one interface connecting to the upstream. Therefore OSPF should have two network statements, and one active interface configured. 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 10 might be:

```
router ospf 105
 log-adjacency-changes
 passive-interface default
 no passive-interface serial 1/0
!
interface serial 1/0
 ip ospf 105 area 0
interface loopback 0
 ip ospf 105 area 0
!
```

8. **Configuring OSPF (part 2).** The OSPF configuration on the router connecting to the IXP 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:

```
router ospf 103
 log-adjacency-changes
 passive-interface default
 no passive-interface serial 0/0
!
interface serial 1/0
 ip ospf 103 area 0
interface loopback 0
 ip ospf 103 area 0
!
```

   NB. There is **NO** configuration for the Ethernet interface appearing in the OSPF configuration. Do **NOT** put any in, repeat, do **NOT**! It is **NOT** a mistake.

9. **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
 no synchronization
 network 10.4.48.0 mask 255.255.240.0
 neighbor ibgp-ixp peer-group
 neighbor ibgp-ixp remote-as 103
 neighbor ibgp-ixp update-source loopback0
 neighbor ibgp-ixp next-hop-self
 neighbor ibgp-ixp password cisco
 neighbor ibgp-ixp description iBGP peering with IXP routers
 neighbor ibgp-ixp send-community
 neighbor ibgp-ixp filter-list 10 out
 neighbor router8-loopback peer-group ibgp-ixp
 no auto-summary
!
ip route 10.4.48.0 255.255.240.0 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, Router6 with this configuration will now no longer send externally learned prefixes to Router8, thereby guaranteeing the security of the AS103 network.

10. **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 use the `next-hop-self` BGP configuration as we have done for our other iBGP sessions earlier on. For example, for Router 7:

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

This configuration is also valid for other iBGP peers in the network if there were more than two routers in the ASN. ISPs try and keep configuration differences to the absolute minimum, and having two different peer-groups defined for iBGP is usually sufficient!

11. **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.

12. **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:

    ```
    ip route 0.0.0.0 0.0.0.0 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.

13. **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:

    ```
    ip prefix-list myprefixes permit 10.4.48.0/20
    ip prefix-list peer100 permit 10.1.0.0/20
    ..
    ip prefix-list peer106 permit 10.7.96.0/20
    !
    router bgp 103
     no synchronization
     bgp log-neighbor-changes
     neighbor ixp-peers peer-group
     neighbor ixp-peers remove-private-AS
     neighbor ixp-peers prefix-list myprefixes out
     neighbor ixp-peers route-map set-local-pref in
     neighbor <router3> remote-as 100
    ```

```
 neighbor <router3> description Peering with AS100
 neighbor <router3> peer-group ixp-peers
 neighbor <router3> prefix-list peer100 in
..
 neighbor <router14> remote-as 106
 neighbor <router14> description Peering with AS106
 neighbor <router14> peer-group ixp-peers
 neighbor <router14> prefix-list peer106 in
 no auto-summary
!
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.

14. **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 <router11> password peer104
 ...
!
```

15. **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
 no synchronization
 bgp log-neighbor-changes
 neighbor <router15b> remote-as 16385
 neighbor <router15b> password as16385-password
 neighbor <router15b> description Peering with AS16385
 no auto-summary
!
```

16. **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:

```
ip prefix-list myprefixes permit 10.2.16.0/20
!
```

```
router bgp 101
 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
 no auto-summary
!
```

17. **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.

    Q. Why do the traceroutes from one ASN to another ASN across the IXP show the middle hop as being star'ed out?

```
Router9#trace 10.7.112.1

  1 10.5.64.6 4 msec 4 msec 4 msec
  2  *  *  *
  3 10.7.96.1 [AS 106] 4 msec *  4 msec
Router9#
```

Checkpoint #5: *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.*

18. **(OPTIONAL) Set up eBGP session with Router 18.** The lab instructors will have configured Router 18 to be a route collector. This is a router which collects all the routes available at the IX. It serves purpose other than to be an information repository showing how many routes are available at the IX – quite often the IXP management will operate such a router, connected to a Looking Glass web interface, to increase the marketing value of the IX. The more peers who get attracted by the routes available at the IX, the greater the value proposition the IX is to the rest of the members. It's in everyone's interest to peer with the router collector:

```
router bgp 103
...
 neighbor 172.17.10.254 remote-as 65534
 neighbor 172.17.10.254 description eBGP with the IX Route Collector
 neighbor 172.17.10.254 password ixp-collector
 neighbor 172.17.10.254 remove-private-AS
 neighbor 172.17.10.254 prefix-list deny-all in
 neighbor 172.17.10.254 prefix-list myprefixes out
...
!
ip prefix-list deny-all deny 0.0.0.0/0 le 32
...
```

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

Note also the inbound prefix filter blocking all prefixes on the eBGP session with the Route Collector. The Collector will not advertise any prefixes, by design. However, ISPs should never trust any other AS or its operator, so the inbound prefix filter is provided for safety, just in case of problems at the Route Collector.

19. **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.

20. **Completed!** The IXP is now complete, up and running. The lab instructors will log into the route collector and show the prefixes visible. All 6 announcements should be clearly seen in the output of *sh ip bgp* on the route collector.

Checkpoint #3: *Compare your BGP routing table with that you see on the route collector. If you have missing prefixes, or some other problems, ask the lab demonstrators.*

21. **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.