Module 10 – An Internet Exchange Point

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

Prerequisites: Modules 1, 6 to 8, and the Exchange Points Presentation

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.

This module concentrates more on IXP peering, so setting up the eBGP sessions, implementing correct prefix-filters, and so on. It is more suited to where workshop participants are interested in participating in an IX in their region, and so can get first hand experience at configuring the eBGP sessions with the other participants. However it misses out some important details regarding connecting a fully fledged ISP network to an IX.

Module 16 on the other hand looks at a more realistic IXP. It has six 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. That module is strongly recommended over this one as it covers correct OSPF and iBGP configuration practices for connecting the IXP facing router to the rest of the ISP network.

Lab Exercises

1. **A Simple IXP.** This example is of a very simple IXP. But using this configuration, any participant in this workshop should be able to go away and set up a working IXP in their own economy. Technically they are not hard to implement. Politics & business economics are not covered in this workshop.

   Only prefix lists are used to filter BGP announcements. eBGP peers should be in peer-groups, and route refresh should be used to implement any policy changes as in other modules.

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 are by ethernet.

3. **Addressing Plan.** These address ranges should be used throughout this 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.

AS101	10.1.0.0/20	AS108	10.8.0.0/20
AS102	10.2.0.0/20	AS109	10.9.0.0/20
AS103	10.3.0.0/20	AS110	10.10.0.0/20
AS104	10.4.0.0/20	AS111	10.11.0.0/20
AS105	10.5.0.0/20	AS112	10.12.0.0/20
AS106	10.6.0.0/20	AS113	10.13.0.0/20
AS107	10.7.0.0/20	AS114	10.14.0.0/20

4. **Basic Router Setup.** Set up the routers as you would have done in previous modules. That is, basic security, the BGP outline configuration, IOS Essentials, etc.
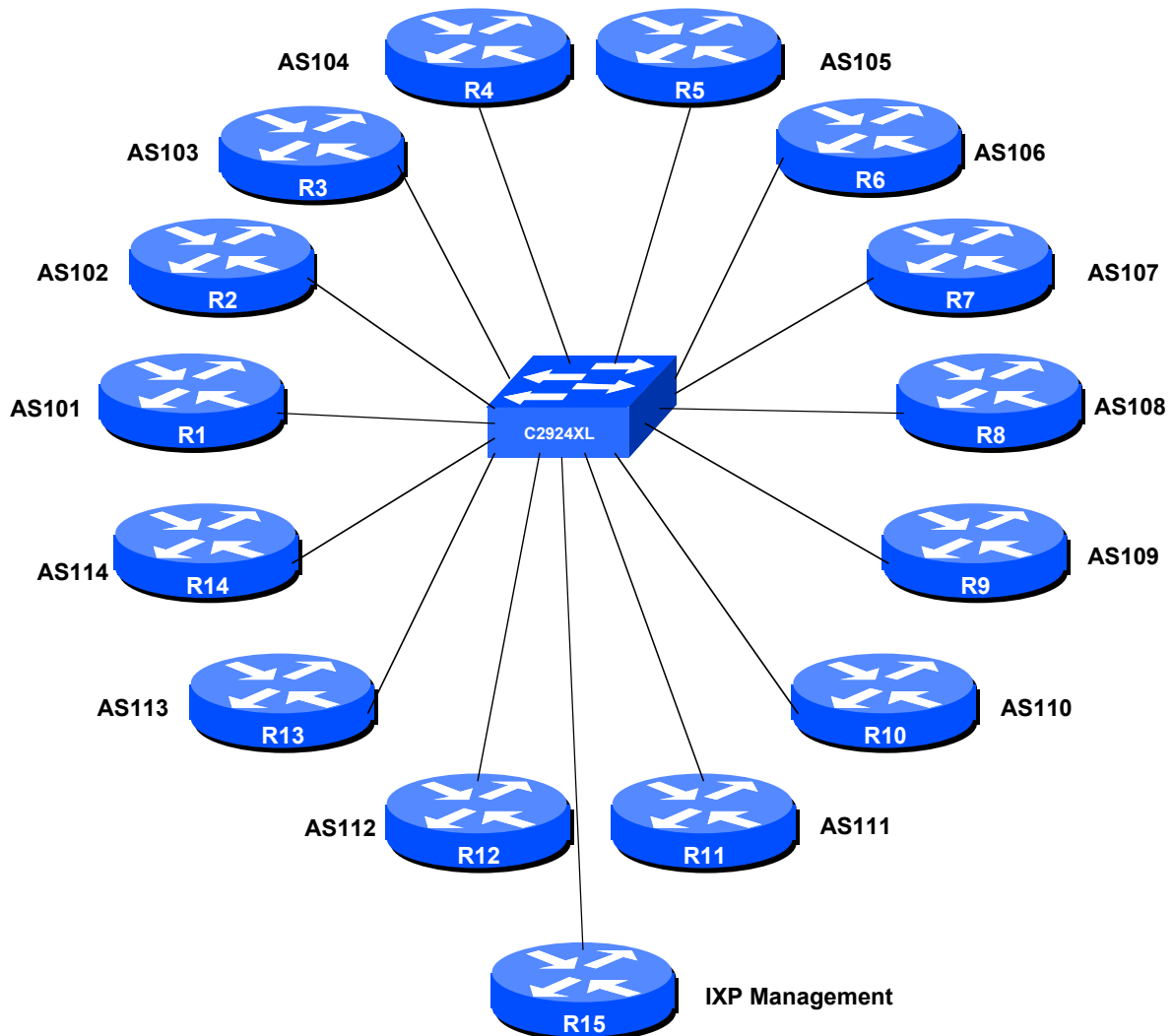

Figure 1 – IXP Configuration

5. **Management Router and IXP LAN.** The lab instructor will have connected another router to the exchange point − this is Router 15 in the figure. Each router team should set up their router to synchronise time off that router using NTP. The password on the NTP session is "cisco" as in previous Modules. The address range used for the IXP LAN is 172.16.0.0/24 − the management router in this module has an IP address of 172.16.0.254. Each of the ASes is assigned a block of 3 addresses to use on the exchange point LAN. So, for example, AS101 has 172.16.0.1, 172.16.0.2 and 172.16.0.3. AS102 has 172.16.0.4, 172.16.0.5 and 172.16.0.6. And so on.

Q. Why do you think three addresses have been assigned to each participant at the IXP?

Checkpoint #1: *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.*

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 14 might be:

```
interface fastethernet 0
 description Exchange Point LAN
 ip address 172.16.0.40 255.255.255.0
 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.

7. **Configuring BGP on the routers.** Next, eBGP needs to be set up on the routers. Create a peer-group and apply that peer-group to each eBGP neighbour. A sample configuration for Router13 might be:

```
ip prefix-list myprefixes permit 10.13.0.0/20
ip prefix-list peer101 permit 10.1.0.0/20
..
ip prefix-list peer114 permit 10.14.0.0/20
!
router bgp 113
 no synchronization
 bgp log-neighbor-changes
 network 10.13.0.0 mask 255.255.240.0
 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 <router1> remote-as 101
 neighbor <router1> description Peering with AS101
 neighbor <router1> peer-group ixp-peers
 neighbor <router1> prefix-list peer101 in
..
 neighbor <router14> remote-as 114
 neighbor <router14> description Peering with AS114
 neighbor <router14> peer-group ixp-peers
 neighbor <router14> prefix-list peer114 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.

8. **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.

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.*

9. **Set up passwords on the eBGP sessions.** 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 6).

```
router bgp 113
 neighbor <router14> password peer114
 !
```

10. **Set up eBGP session with Router 15.** The lab instructors will have now configured Router 15 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 113
...
 neighbor 172.16.0.254 remote-as 65534
 neighbor 172.16.0.254 description eBGP with the IX Route Collector
 neighbor 172.16.0.254 password cisco
 neighbor 172.16.0.254 remove-private-AS
 neighbor 172.16.0.254 prefix-list deny-all in
 neighbor 172.16.0.254 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.

11. **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.

12. **Completed!** The IXP is now complete, up and running. The lab instructors will log into the route collector and show the prefixes visible. All 14 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.*

STOP AND WAIT HERE

13. **Route Server.** Some Internet Exchange Points operate a device called a route server. This is basically a route collector but with the added feature that it announces the prefixes it has learned to all the ISPs who are connected to it. The lab instructors will now modify the route collector's configuration so that it announces prefixes to the ASNs participating in the IXP.

14. **Participating with the Route Server.** Each ISP should now modify their BGP peering with the Route Collector so that it can receive prefixes from that router. This means modifying or simply removing the inbound prefix filter, for example as below:

```
router bgp 113
...
 neighbor 172.16.0.254 remote-as 65534
 neighbor 172.16.0.254 description eBGP with the IX Route Collector
 neighbor 172.16.0.254 password cisco
 neighbor 172.16.0.254 remove-private-AS
 neighbor 172.16.0.254 prefix-list route-server in
 neighbor 172.16.0.254 prefix-list myprefixes out
...
!
ip prefix-list route-server permit 0.0.0.0/0 le 32
...
```

15. **Bi-lateral peering sessions.** Each team will now see that they have BGP prefixes learned by direct or bi-lateral peerings as well as through the route-server. The prefixes learned via the route-server have AS65534 in the AS-PATH – the AS is longer, so the bi-lateral peering is preferred. Each router team will now remove all their bi-lateral peerings, leaving only the peering with the route-server in place.

```
router bgp 113
 no neighbor <router1>
 no neighbor <router2>
 . . .
```

Once this is done, the BGP table will only contain prefixes learned via the route-server.

16. **Summary.** Route servers are more normally used to assist with the scaling of larger IXPs. Managing a large number of eBGP sessions can be quite challenging, especially for ISPs who have routers which are less able to handle large numbers of eBGP sessions. To reduce the administrative overhead, they learn most if not all their peer prefixes via the route-server instead. In the industry today it is normal for ISPs at an Exchange Point to set up bi-lateral peerings with ISPs who do not participate in the route-server, and set up peerings with the route-server for those who do participate. Not ever ISP chooses to participate in the route-server, as was discussed in the presentation.

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.*