Module 9 – Internet Exchange Points

Objective: An optional module to demonstrate the use of BGP at Internet Exchange Points.

Prerequisite: Modules 1 to 7 and the BGP presentations

Topology

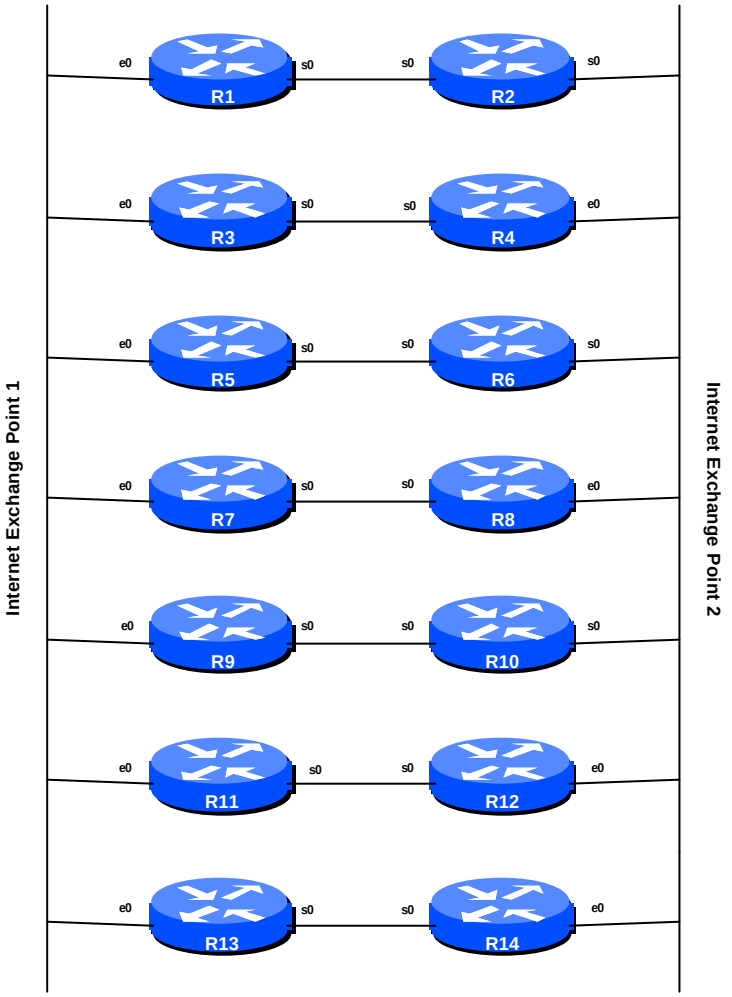

Figure 1 – IXP Layout

INTRODUCTION:
This Module introduces some of the concepts used at many of the Internet exchange points. For the purpose of the workshop, the exchange point is a Catalyst 2924XL 10/100 switch. Each router is a different ISP, with a different AS. Each ISP has a connection to the exchange point (public peering), as well as a private connection to another ISP via a direct link (a not uncommon scenario).

Important points to note:

- You should use your IXP to reach **ONLY** the networks belonging to peers with whom you peer at the IXP.
- For networks belonging to non-IXP peers (the routers belonging to the other IXP in this case), **DO NOT** use IXP connections to send traffic. That traffic should go over your private link.

For example, referring to the Figure, to route traffic to R10, R1 must not use the IXP Ethernet. R1 will use the IXP Ethernet to route packets to networks belonging to R3, R5, R7, R9, R11 and R13 only.

Two methods are introduced here. The first uses the more traditional route and AS-path filters, while the second uses BGP communities. Either option can be chosen, or both can be investigated.

- Route and AS-path filters tend to be more complex to configure and maintain. They do give well defined control of what goes in and out of a network.
- Most ISPs now prefer to use BGP communities for fine-grained control of routing policy. The configuration is much less complex and generally easier to understand and maintain.
- In reality a combination of community, route and AS-path filters are used. Policy within a network and towards peers can be controlled using communities, but a route filter is used in addition to ensure there are no outbound route leaks. Communities are used for peering information between ASes, but most ISPs prefer to control what they hear from peers by using route and AS-path filters.

Both methods presented here achieve the same end result. No attempt is made to demonstrate the operation of a typical IXP because there isn't really a typical IXP. Each in the Internet today uses its own version of infrastructure, peering model, and addressing scheme.

Big Hint: This module uses most of the techniques which have been covered in the other modules in this ISP workshop. Making full use of these techniques will make the configuration much easier! If in doubt, ask the instructors, or refer to the BGP documentation.

1. **Build the network:** The first step is to build the network for this Module. Figure 1 shows what is required here. All **Odd** numbered routers are in assigned to IXP 1 and all **Even** numbered routers are assigned to IXP 2. All routers have two connections, a direct connection to a router in the other IXP and a connection to an Ethernet Switch which connects all the routers assigned to a single IXP.

   The lab instructor will have located two ethernet switches to be used as the two IXPs. Use the supplied ethernet cable to connect your router to your IXP. If you have an existing link to the

Cisco Systems Inc
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

router on the opposite side of the lab, you are now set. If not, locate an ethernet cable, or serial cable, and connect that to your opposite number.

2. **Basic Router Configuration:** Now that the network is physically constructed, configure the router as in Steps 1 to 5 in Module 1. Notice the order – physical connection, basic functionality, followed by LAN and WAN configuration.

Checkpoint #1: *Call the lab instructors and show how the network has been constructed and demonstrate the configuration of your router so far.*

3. **AS assignments and loopback.** The AS assignments are as follows:

R1	**AS 101**	**R2**	**AS 202**
R3	**AS 103**	**R4**	**AS 204**
R5	**AS 105**	**R6**	**AS 206**
R7	**AS 107**	**R8**	**AS 208**
R9	**AS 109**	**R10**	**AS 210**
R11	**AS 111**	**R12**	**AS 212**
R13	**AS 113**	**R14**	**AS 214**

Configure the router with its AS using the *autonomous-system* and *router bgp* commands. The following network blocks are assigned to each router. Each router must have static routes, bgp network statements and a loopback address configured for its network block:

R1	**210.101.4.0/22**	**R8**	**220.208.4.0/22**
R2	**220.202.4.0/22**	**R9**	**210.109.4.0/22**
R3	**210.103.4.0/22**	**R10**	**220.210.4.0/22**
R4	**220.204.4.0/22**	**R11**	**210.111.4.0/22**
R5	**210.105.4.0/22**	**R12**	**220.212.4.0/22**
R6	**220.206.4.0/22**	**R13**	**210.113.4.0/22**
R7	**210.107.4.0/22**	**R14**	**220.214.4.0/22**

Example: Router R1
Configure static routes for four /24 networks out of the network block 210.101.4.0/22. These are used simply to populate the BGP routing table on the router, and in real life would represent connected customer networks, for example.

```
ip route 210.101.4.0 255.255.255.0 Null0
ip route 210.101.5.0 255.255.255.0 Null0
ip route 210.101.6.0 255.255.255.0 Null0
ip route 210.101.7.0 255.255.255.0 Null0
```

Also configure a loopback interface:

```
interface loopback 0
 ip address 210.101.7.224 255.255.255.255
```

and configure BGP:

```
router bgp 101
 network 210.210.4.0
 network 210.210.5.0
 network 210.210.6.0
 network 210.210.7.0
```

Note: The loopback interface isn't used as such in this module, but is used as reference point for ping and trace commands later in this module. (It is good practice to always configure a loopback interface on a router – remember that it never goes away.) Use the final /28 out of the /22 address space assigned to the router.

4. **Private link to neighbouring ISP:** Agree an IP subnet to be used on the private peering between you and your neighbouring AS. For example, R1 and R2 need to agree on IP addresses for the serial link between them. They could use 210.101.5.0/30, for example.

   Configure the interface to your private peer and check connectivity by using ping.

5. **Configuring interfaces at the IXPs:**

   IXP 1 Ethernet: Use IP Subnet 201.201.201.0 255.255.255.240 for address space and assign the following addresses to the Ethernet interface of the router:

   R1: 201.201.201.1,
   R3: 201.201.201.3,
   R5: 201.201.201.5,
   Etc…

   IXP 2 Ethernet: Use IP Subnet 202.202.202.0 255.255.255.240 for address space and assign the following addresses to the Ethernet interface of the router:

   R2: 202.202.202.2,
   R4: 202.202.202.4,
   R6: 202.202.202.6,
   Etc…

6. Determine which IXP your router is assigned to and collect information on all other routers/peers who belong to the same IXP. In other words, you want to find out about all the routes being advertised by your peers on the IXP.

7. Configure the ethernet interface on the IXP and check that you can ping other routers.

8. **Route Flap Dampening:** Configure Route Flap Dampening on the BGP process. Observe the BGP routing table during the remainder of this module, and collect statistics.

Checkpoint #2: *Call the lab instructors and demonstrate that you can ping your neighbouring private peer, and the IXP you are connected to.*

METHOD ONE – Route and AS path filtering

9. **Configuring peer-groups and filters:** Configure an eBGP **peer-group** to include all the peers which belong to the same IXP. This peer-group should include the *soft-reconfiguration* directive to allow non-intrusive resetting of the BGP sessions with your peers.

10. **Announcing local ISP aggregate:** Configure your BGP process to advertise an **aggregate** route which represents all the routes you have statically configured in your router. For example, R1 will generate an aggregate to cover the network block 210.101.4.0/22. **Hint:** use the *network x.x.x.x mask y.y.y.y* bgp command to do this.

11. **Announcing IXP aggregate:** Configure your BGP process to advertise an **aggregate** route which represents all the routes present at the exchange point you are connected to. For example, R4 will generate an aggregate for 220.0.0.0/8. **Hint:** use the *aggregate-address x.x.x.x y.y.y.y* bgp command to do this.

12. **Routing Policy at the IXP:** Once physical connectivity has been established, and the local configuration has been finalised, the next step is to apply some routing policy outbound and inbound at the IXP.

    - Configure an **outbound** filter to peers within the same IXP, such that you **ONLY** advertise routes belonging to you. For example, R1 must **only** advertise 210.101.4.0/22. **Hint:** use the *distribute-list* command to filter out the more specific networks in the /22 block.

    - Configure an **inbound** filter for each of your IXP peers, such that you **ONLY** receive routes that belong to them. For example, R1, R5, R7, R9, R11 and R13 accept **only** 210.103.4.0/22 from R3, etc. **Hint:** use the *distribute-list* command to filter each peer's announcement.

There are two ways of applying these filters. One is to configure a separate line in an access-list for each specific network which has to be blocked. This gets tedious and difficult to maintain as the ISP grows in size. The other, preferred, way is to configure an access-list which matches the exact network and mask which is to be announced.

Important Note:
These distribute-lists require the use of extended access-lists so that an exact match on the network and mask can be achieved. Using simple access-lists will result in all networks matching the access-list being allowed. For example, if:

```
access-list 1 permit 210.101.4.0 0.0.3.255
router bgp 101
 neighbor 201.201.201.3 distribute-list 1 out
```

is configured on Router R1 on its peering with R3, the four /24 networks and the /22 network originated by R1 will be announced to R3. This results in what ISPs term *leaking of specific routes* and is considered anti-social behaviour in the Internet community. (Specific routes can be announced in particular circumstances, for example to achieve load balancing or a particular routing policy. In general, the announcing of an aggregate network, and the more specific subnets is considered bad practice and contributes to the bloat of the Internet routing table.)

However, extended access-lists can be used to specify the network mask of the network being filtered, as in the example below which achieves the desired result:

```
access-list 101 permit ip host 210.101.4.0 host 255.255.252.0
router bgp 101
 neighbor 201.201.201.3 distribute-list 101 out
```

This format of access-list 101 only permits the network 210.101.4.0 with mask matching 255.255.252.0. More specific versions of this network have longer masks, and therefore are dropped. For example, 210.101.4.0 mask 255.255.255.0 doesn't match, so is not announced in the BGP peering.

For more information on extended access-lists consult the Documentation CD.

Aside: the bgp command *aggregate-address x.x.x.x m.m.m.m summary-only* could be used here. Do you see any potential issues with this?

13. **Routing Policy for Private Peers:** With the IXP peering secured, routing policies now need to be applied to all private peers.

Cisco Systems Inc
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

- Configure an **outbound** route filter such that you advertise **ONLY** the aggregate representing routes in the IXP and nothing else. For example, R1 must advertise only 210.0.0.0/8 to R2, nothing else.

- Configure an **inbound** route filter such that you receive **ONLY** an aggregate representing routes in the other IXP and nothing else. For example, R1 must receive only 220.0.0.0/8 from R2, nothing else.

Recall the configuration recommendation regarding access-lists in the previous step. It can be used to good effect here as well.

Checkpoint #3: *Call your lab instructor and display the following information. Each router in an IXP must have in its routing table:*

- *Its own routes*
- *More specific routes for all its other peers in the IXP*
- *An aggregate for all the routes in its IXP*
- *An aggregate for all routes in the other IXP*

i] *Do a traceroute to a destination belonging to a peer of the same IXP and show that the trace uses the IXP to reach its destination. For example, for R1 to reach 210.109.4.1 , the packet must use the IXP Ethernet.*
ii] *Do a traceroute to a destination belonging to a peer of the other IXP and show that the trace does not use your local IXP to reach the destination. For example, for R1 to reach 220.110.4.1, the packet must not use the IXP Ethernet but use its link to R2 to reach its destination.*

METHOD TWO: Community based filtering

Note: If continuing from Method One, remember to delete the route-maps and filters which no longer apply. Remember, all **unused configuration should be removed**.

14. **Configuring peer-groups and communities:** Configure an eBGP **peer-group** to include all the peers which belong to the same IXP. This peer-group should include *soft-reconfiguration* and the *send-community* directive.

    - Each router in IXP 1 will tag routes belonging to it with a community of *AS*:200. In other words, R1 will tag the network 210.101.4.0/22 with a community of 101:200; R3 will tag the network 210.103.4.0/22 with a community of 103:200, etc…

- Each router in IXP 2 will tag routes belonging to it with a community of *AS*:400. In other words, R2 will tag routes 220.202.4.0/22 with a community of 202:400; R4 will tag routes 220.204.4.0/22 with a community of 204:400, etc…

Remember the *bgp community new-format* command, which is required to display communities in the above format.

15. **Announcing local ISP network aggregate:** Configure your BGP process to advertise an **aggregate** route which represents all the routes you have statically configured in your router. For example, R1 will generate an aggregate to cover the network block 210.101.4.0/22. **Hint:** use the *network x.x.x.x mask y.y.y.y route-map <map-name>* to do this.

16. **Announcing IXP aggregate:** Configure your BGP process to advertise an **aggregate** route which represents all the routes present at the exchange point you are connected to. For example, R4 will generate an aggregate for 220.0.0.0/8. **Hint:** use the *aggregate-address x.x.x.x y.y.y.y attribute-map <map-name>* to do this.

    Assign a community of *AS*:100 to the aggregate generated for IXP1 and *AS*:300 to the aggregate generated for IXP2. For example, router R4 connecting to IXP 2 will generate aggregate for 220.0.0.0/8 with community 204:300.

17. **Community summary:** Eventually all the aggregate routes belonging to routers in IXP1 will have a community of *AS*:200 and the aggregate representing all routes in IXP1 [210.0.0.0/8] which is advertised to IXP2 peers will have a community of *AS*:100. Similarly, all the routes belonging routers in IXP2 will have a community of *AS*:400 and the aggregate representing all routes in IXP2 [220.0.0.0/8] which is advertised to IXP2 peers will have a community of *AS*:300.

18. **Routing Policy at the IXP:** Once physical connectivity has been established, and the local configuration has been finalised, the next step is to apply some routing policy outbound and inbound at the IXP.

    - Configure an **outbound** route-map to peers within the same IXP, such that you **ONLY** advertise routes belonging to you. For example, R1 must **only** advertise 210.101.4.0/22.
    - Configure an **inbound** route-map for each of your IXP peers, such that you **ONLY** receive routes that belong to them. For example, R1, R5, R7, R9, R11 and R13 accept **only** 210.103.4.0/22 from R3, etc.

19. **Routing Policy for Private Peers:** With the IXP peering secured, routing policies now need to be applied to all private peers.

Cisco Systems Inc
170 West Tasman Drive.
San Jose, CA 95134-1706
Phone: +1 408 526-4000
Fax: +1 408 536-4100

- Configure an **outbound** route-map such that you advertise **ONLY** the aggregate representing routes in the IXP and nothing else. For example, R1 must advertise only 210.0.0.0/8 to R2, nothing else.

- Configure an **inbound** route-map such that you receive **ONLY** an aggregate representing routes in the other IXP and nothing else. For example, R1 must receive only 220.0.0.0/8 from R2, nothing else.

  Hint: This is a good place to use the community string which was set earlier for the aggregates. For example, aggregate 210.0.0.0/8 has a community of *AS*:100 associated with it. Therefore, configure a route-map to use this community to select the aggregate to implement the above policies.

Checkpoint #4: *Call your lab instructor and display the following information. Each router in an IXP must have in its routing table:*

- *Its own routes*
- *More specific routes for all its other peers in the IXP*
- *An aggregate for all the routes in its IXP with the right community*
- *An aggregate for all routes in the other IXP with the right community*

i] *Do a traceroute to a destination belonging to a peer of the same IXP and show that the trace uses the IXP to reach its destination. For example, for R1 to reach 210.109.4.1, the packet must use the IXP Ethernet.*
ii] *Do a traceroute to a destination belonging to a peer of the other IXP and show that the trace does not use your local IXP to reach the destination. For example, for R1 to reach 220.110.4.1, the packet must not use the IXP Ethernet but use its link to R2 to reach its destination.*

20. Filtering using access-lists and communities: Many ISPs prefer to use inbound route and AS-path filters for IXP peerings, in addition to using communities.

Q. Why do you think this is the case?
A. Inbound filters control what is **received** by an ISP's network. If an ISP relied on community based filtering only, it is relying on the peer ISP setting communities correctly. Most ISPs don't take this chance, so protect any community based filtering and policy control with extra inbound filters. These filters often contain the network the peer ISP is announcing, as well as RFC1918 networks ("private" address space), the default network, their own networks, and any

others they may desire to block. See the ***IOS Essentials for ISPs*** document for more information about RFC1918 networks and inbound filtering practices of ISPs.

Q. What configuration additions are required to implement inbound filters in the above example, in addition to inbound filtering based on communities?
A. Extra *distribute-list* commands should be added to the *bgp neighbor* statement in the eBGP peerings. Below is a possible example for the peering between R1 and R3 on IXP1:

```
router bgp 101
 neighbor IXP1-peers peer-group
 neighbor IXP1-peers send-community
 neighbor IXP1-peers soft-reconfiguration inbound
 neighbor IXP1-peers route-map IXP1-peers-out out
 neighbor 201.201.201.3 remote-as 103
 neighbor 201.201.201.3 peer-group IXP1-peers
 neighbor 201.201.201.3 distribute-list 103 in
 neighbor 201.201.201.3 route-map R3-in in
 . . .
access-list 103 permit ip 210.103.4.0 0.0.3.255 255.255.252.0 0.0.3.255
access-list 103 deny ip any any
```

Checkpoint #5: *Call the lab assistant and demonstrate the configuration changes you needed to make.*

CONFIGURATION NOTES

Documentation is critical! You should record the configuration at each ***Checkpoint***, as well as the configuration at the end of the module.