Module 20 – Overseas Collocation

Objective: To investigate methods for connecting to Internet backbones overseas.

Prerequisites: Modules 12, 13, 18 and (optionally) 19, and the Collocation Presentation

The following will be the common topology used.

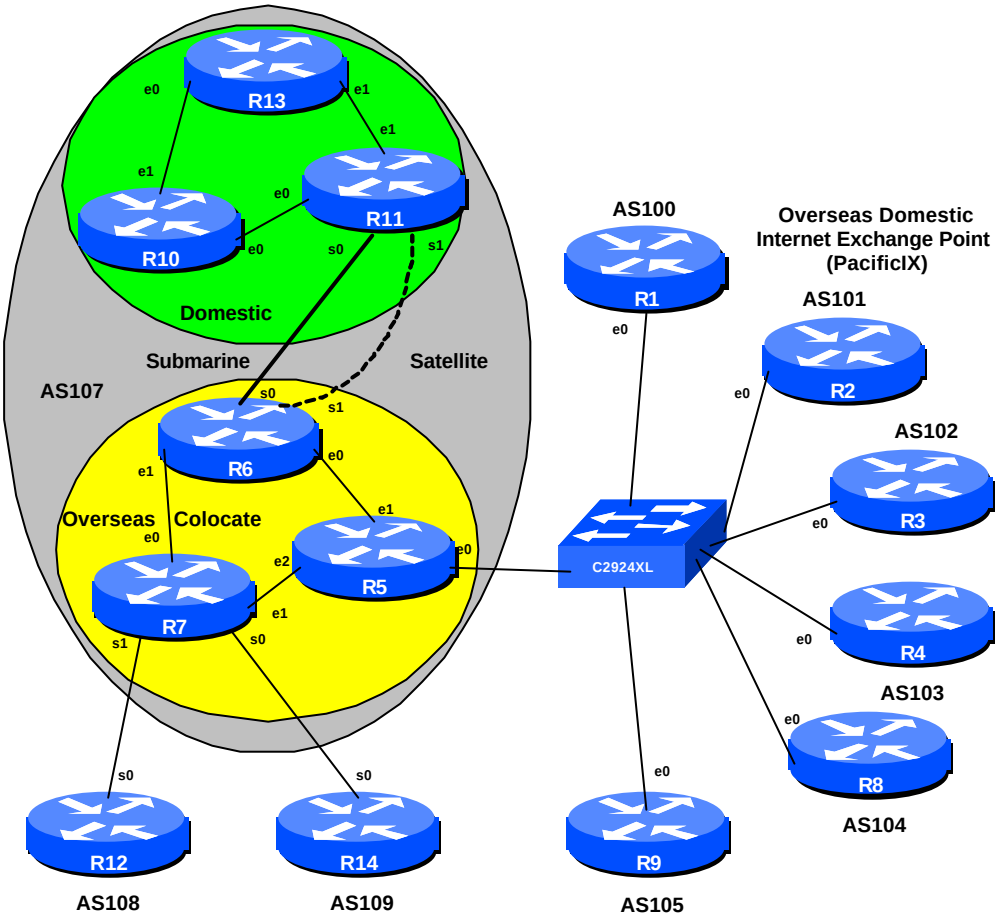

Figure 1 – International Configuration

Lab Notes

The purpose of this module is to investigate the principles and practices surrounding overseas collocation. The BGP presentation should be reviewed during the study of this module as it provides some of the technical motivation behind locating equipment overseas.

This example assumes an Asian ISP looking for collocation space in the west coast of the United States, but it can apply equally well to any ISP operating a network which interconnects to other networks overseas.

Lab Exercise

1. **Basic Configuration.** Each router team should configure their router to fit into the network topology depicted in Figure 1 and physical layout depicted in Figure 2. **Notice that Routers 6, 7 and 11 require more than two serial ports** (the 3620s in the ISP Workshop kit have 4 serial ports). Check all connections.

2. **Address Ranges.** 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.

AS100	**218.11.0.0/19**	**AS105**	**219.64.0.0/19**
AS101	**218.35.0.0/19**	**AS106**	**219.99.0.0/19**
AS102	**218.76.0.0/19**	**AS107**	**220.10.0.0/19**
AS103	**219.13.0.0/19**	**AS108**	**220.19.0.0/19**
AS104	**219.58.0.0/19**	**AS109**	**220.73.0.0/19**

3. **Basic Router Setup.** With the exception of the routers in AS107, set up the routers as you would have done in previous modules. That is, basic security, the BGP outline configuration, IOS Essentials, etc. Routers AS108 and AS109 should do the basic configuration for their routers – don't set up eBGP to AS107 yet.

4. **AS107 routers.** The routers making up AS107 should set up OSPF and iBGP in the AS. Note the two serial links between Routers 6 and 11. Both of these should be activated – they are intended to simulate the transoceanic links which many ISPs install to reach the US Internet from Asia. Ask the lab instructors for any longer cables if they are available. Routers 6 and 11 should be **route reflectors** for the colocate and domestic network respectively. Using a route reflector is more efficient than a full mesh iBGP, certainly at this level. Note that Router 6 and 11 have normal iBGP between them.

Hint: Router6 is a reflector for Router5 and Router7, Router 11 is a reflector for Router 10 and Router 13.

Hint: The overseas routers in AS107 **MUST NOT** originate the AS107 address block. Why not?

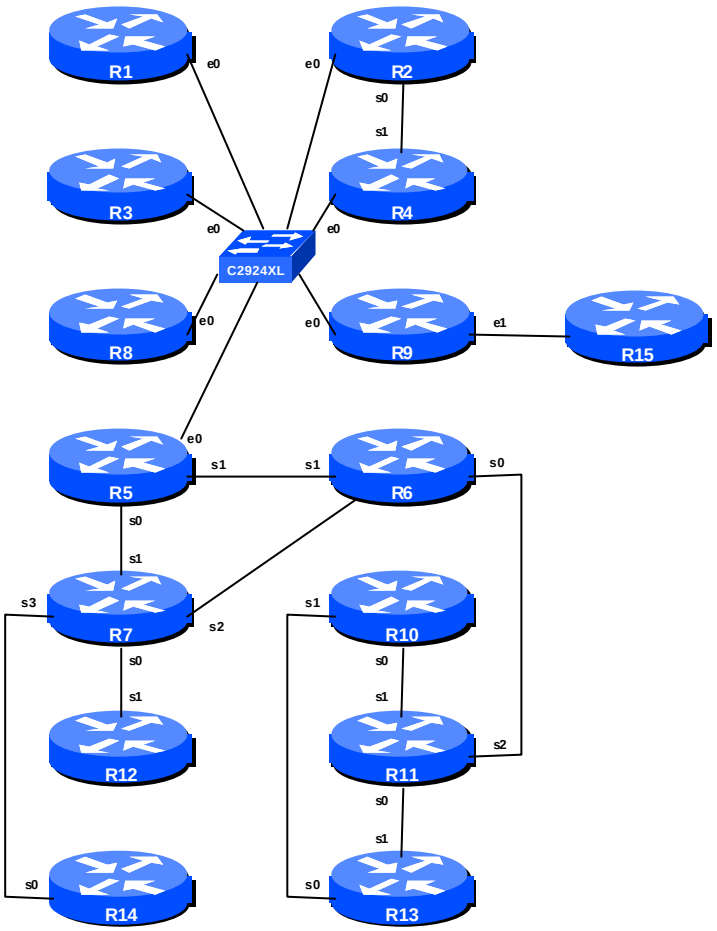

Figure 2 – Lab Physical Layout

5. **IXP Participants.** Those routers participating in the IXP (Routers 1 to 4, 8 and 9) should use the 220.5.10/24 network for the IP addresses of the IXP LAN. As in Module 19, ASes 100 to 105

should set up eBGP between each other. Remember how the configuration was implemented in Module 19? Only announce your prefix, only accept your peer's prefix, uRPF checks, etc. If in doubt, please refer to your notes from Module 19.

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. Routers in AS107 should have iBGP and OSPF set up, and all the routers visible and pingable.*

Configuring the Peerings with the Exchange Point Participants

6. **Configure AS107 relationship with the IXP.** The Domestic Internet Exchange Point provides AS107 with access to the regional domestic market at the collocate site. This assumes of course that all the service providers at the IXP would be willing to peer. (In practice they will be as they gain a valuable access to an overseas market without paying their transit provider, and it is good for the overseas provider not having to pay so much for transit to their upstream. Win-win situation for both AS107 and the IXP participants.)

7. **Router6 Route Reflector configuration detail.** AS107 has to be extremely careful how it peers with the IXP participants. Recall from the presentation that Router5 absolutely must not have a default route or the full routing table on it. (**Why not?**) This is ensured by careful configuration of the route reflector Router6 – it only announces AS107 and it's customer prefixes to Router5, nothing more. The team operating Router5 should now configure their router so that it only sends AS107 prefixes to Router5. The easiest way for this module is to use an outbound prefix-list. An example configuration for Router6 might be:

```
ip prefix-list myprefixes permit 220.10.0.0/19
!
router bgp 107
 neighbor <router11> remote-as 107
 neighbor <router11> description My Home Route Reflector
 neighbor <router5> remote-as 107
 neighbor <router5> description Router at the PacificIX
 neighbor <router5> route-reflector-client
 neighbor <router5> prefix-list myprefixes out    ! NOTE THIS LINE
 neighbor <router7> remote-as 107
 neighbor <router7> description Router connecting to my upstreams
 neighbor <router7> route-reflector-client
!
```

A prefix-list was used rather than a filter-list. **Why?** A filter list on filters ASes – it does not filter prefixes. Remember that AS108 and AS109 were announcing the default route to AS107 – the default route would not be blocked by a filter list.

The alternative to using a single prefix-list as above is to use a combination of prefix-list and filter-list. The prefix-list would block the default prefix, the filter-list would only allow the local and customer ASes. In this case the configuration for Router6 might be:

```
ip prefix-list nodefault deny 0.0.0.0/0
!
ip as-path access-list 10 permit ^$
!
router bgp 107
 neighbor <router11> remote-as 107
 neighbor <router11> description My Home Route Reflector
 neighbor <router5> remote-as 107
 neighbor <router5> description Router at the PacificIX
 neighbor <router5> route-reflector-client
 neighbor <router5> prefix-list nodefault out     ! NOTE THIS LINE
 neighbor <router5> filter-list 10 out       ! NOTE THIS LINE
 neighbor <router7> remote-as 107
 neighbor <router7> description Router connecting to my upstreams
 neighbor <router7> route-reflector-client
!
```

Note that as-path access-list 10 can be added to for however many customer ASes AS107 has. AS108 and AS109 must not be included in this list.

It is important not to announce AS108 or AS109 prefixes to Router8 either. If IXP participants discover that there is a path through AS107 to AS108 or AS109 they may use that rather than their own paths – costing AS107 money.

Finally, all this can be easier to handle/manage using communities. That configuration is left as an exercise to the reader!

8. **Router 5 configuration at the IXP.** As with the IXP Module, care is required configuring a router participating at any IXP. Router6 is only sending local prefixes to Router5, so this ensures some degree of safety. The team operating Router5 should now configure the router to peer with the IXP participants. Remember the concepts from Module 19. Only announce your prefixes to the IXP, only accept the prefixes your peers are entitled to send you.

9. **Connectivity Test.** Check connectivity throughout the IXP network. Each router team in the IXP and AS107 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 for AS107, check the routing table and ensure that you have complete reachability from AS107 to the IXP network. If there are any problems, work with the other router teams to resolve those.*

Configuring the Links to the Backbone Providers

10. **Configure AS108 and AS109 relationship with AS107.** AS108 and AS109 are the upstream US Tier One ISPs of AS107. Basically AS107 has bought Internet transit from these two ISPs. Why two? If one has service problems, the other provides "backup" or redundancy. The configuration for Router 12 and Router 14 are very similar to what we have covered in earlier modules. Basically Router 12 and Router 14 treat AS107 as a customer, so announce only the default to the customer, and only accept the customer's prefixes. The teams operating Router12 and Router14 should now set up the serial interface connecting to AS107 and configure eBGP on their routers. Note that it is common convention that the point to point link between backbone ISP and their customer comes from the ISP address block…

11. **Configure Router7's eBGP peering with AS108 and AS109.** The team operating Router7 should configure eBGP peering with AS108 and AS109 routers. Don't forget the good practices learned earlier. You want to announce only your prefix, and only accept default from the upstream. It is also good practise to disable vulnerable services on serial interfaces of the router. For example:

```
interface serial 0/0
 description 2MBps connection to AS108
 ip address 220.19.31.2 255.255.255.252
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
```

Once the interfaces to AS108 and AS109 are functioning (you can ping the other end of the link), eBGP should be set up. An example might be:

```
ip prefix-list myprefixes permit 220.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
```

```
router bgp 107
 neighbor <router12> remote-as 108
 neighbor <router12> description Connection to AS108 Transit Provider
 neighbor <router12> prefix-list myprefixes out
 neighbor <router12> prefix-list default in
 neighbor <router14> remote-as 109
 neighbor <router14> description Connection to AS109 Transit Provider
 neighbor <router14> prefix-list myprefixes out
 neighbor <router14> prefix-list default in
!
```

Note one thing. **Router 7 DOES NOT originate AS107's prefix.** Indeed, none of the routers in the overseas part of AS107 should originate AS107's prefix. If the transoceanic cable is broken for whatever reason, the routers in the overseas part of AS107 will still announce the aggregate – this will create a blackhole for AS107 traffic if AS107 has another Internet connection elsewhere in its backbone.

12. **Connectivity Test.** Check connectivity throughout the entire network. AS108 and AS109 should only be able to see AS107. AS107 should be able to see everything. The IXP participants should be able to see each other, but not AS108 or AS109. When you are satisfied that BGP is working correctly, try running traceroutes to check the paths being followed.

Checkpoint #3: *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.*

The Transoceanic Circuit(s)

13. **Loadsharing on the Transoceanic Circuits.** If OSPF has been set up properly traffic on the two "transoceanic" links should be load shared. To test this, either try running traceroute from domestic to collocate parts of the network, or connect laptops to various points of the network and send data across the backbone. If loadsharing doesn't seem to be working, check the OSPF configuration – did you remember to set the bandwidth on the interface to match the clockrate on the circuit, for example? Disconnect one of the "transoceanic" cables and see what happens. Routing should failover gracefully.

14. **Satellite links.** One of the circuits will now be converted to simulate a simplex satellite connection (uni-directional link). The dotted cable in Figure 1 will be the satellite connection. Policy Routing will be used to send delay insensitive traffic over that connection, with the remaining traffic going over the "submarine" cable. Refer to Module 8 on policy routing if you

don't remember how to configure policy routing in IOS. Before configuring policy routing, lower the clockrate on the "circuit" from 2000000 to 64000bps. Don't forget to change the "bandwidth" command on the interfaces at either end. And remove the circuit from OSPF – it is unidirectional, and we will be using policy routing to put traffic on to it.

15. **Configuring Policy Routing.** Policy Routing will now be configured on Router 6 to manage the two links back to the domestic network. Basically port **tcp/80** will be redirected over the "satellite" connection. An example configuration might be:

```
access-list 100 permit tcp any any eq www
!
route-map divert-web permit 10
 match ip address access-list 100
 set interface ser 0/1
route-map divert-web permit 20
!
interface ethernet 0/0
 description connection to Router7
 ip policy route-map divert-web
 ip route-cache policy
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
interface ethernet 0/1
 description connection to Router6
 ip policy route-map divert-web
 ip route-cache policy
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
interface serial 0/1
 description Satellite connection Home
 ip address 220.10.31.2 255.255.255.252
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
```

16. **Connectivity Test.** When configured, test the set up by attempting to connect to a webserver which the workshop instructors will have connected to AS109. What happens?

Checkpoint #4: *Once the policy routing 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. Only tcp/80 will be diverted over the satellite link.*

17. **Summary.** This module has given a detailed example of how an ISP would configure an overseas collocate site. It has highlighted (again) the care required when peering at a public exchange point, and given the necessary configuration tips to ensure a successful peering. It has pointed out how to connect to an upstream ISP, and how to configure a transoceanic connection back to the domestic network. It has also briefly looked at some of the possibilities for using a satellite based transoceanic connection for non-delay sensitive traffic.

CONFIGURATION NOTES

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

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