Module 13 – Multihoming to Different ISPs

Objective: To investigate various methods for multihoming onto two different upstream ISPs.

Prerequisites: Module 12 and Multihoming Presentation

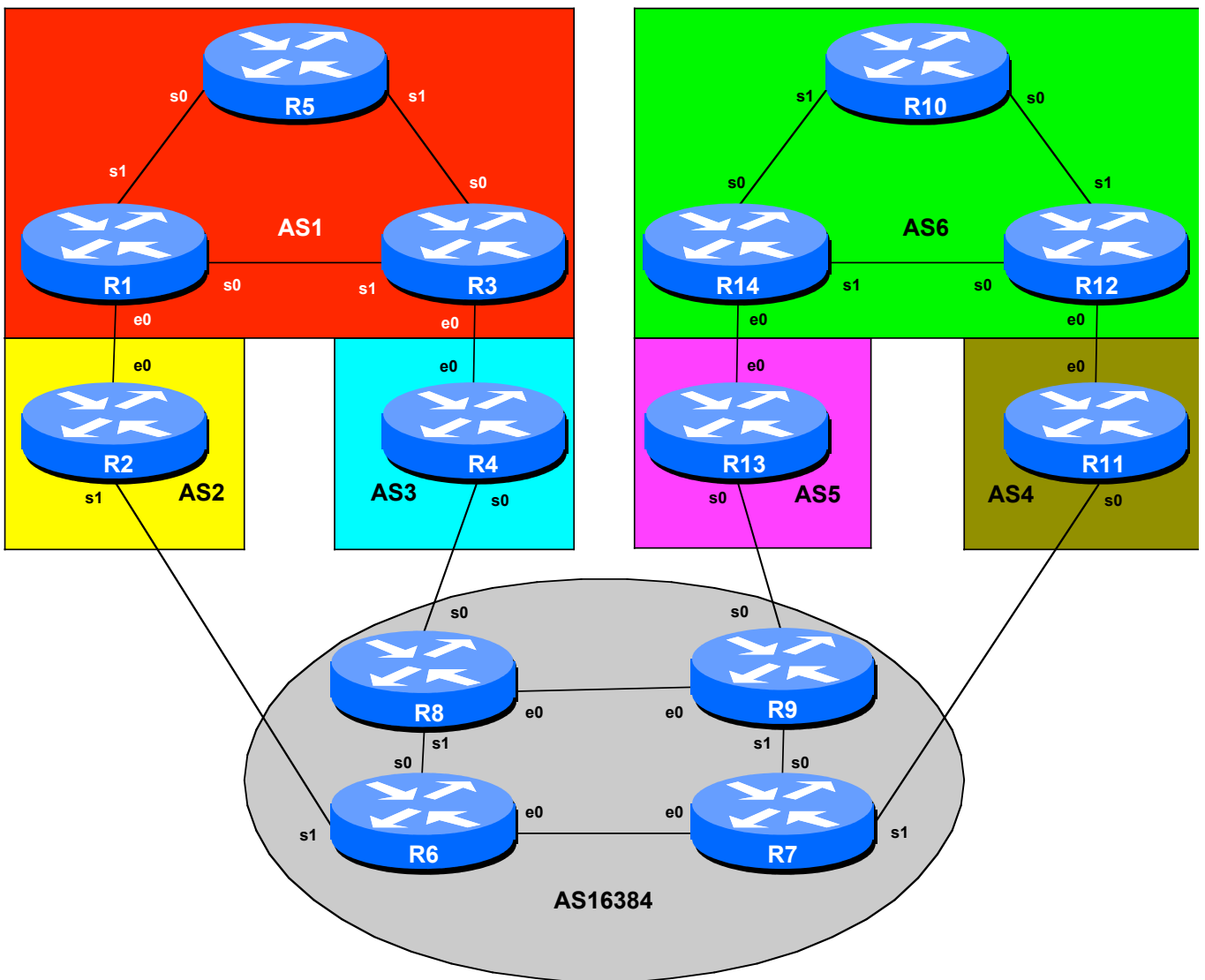The following will be the common topology used.

Figure 1 – ISP Lab Multihoming Configuration

Monday, June 20, 2005

Lab Notes

The purpose of this module is to demonstrate multihoming in the situation where the customer AS has one connection to more than one upstream service provider. There are at least two situations where this is applicable:
- Enterprise or Service Provider customer requires more than one connection to the Internet to provide resiliency, and/or loadsharing.
- Enterprise or Service Provider customer requires more than one Internet connection to give their network service provider redundancy.

It is important that you review the multihoming presentation before you start with this module. Only configuration examples will be given – it will be left to the workshop participant to use the presentation notes to help them configure their routers correctly.

To ensure an understandable and easy to follow configuration, as well as good practice, a few assumptions about configuring BGP will be made. These are:
- **Use prefix-lists to filter prefixes**
- **Use as-path access-lists to filter ASes**
- **Use route-maps to implement policy**

There are rarely any exceptions to this. Prefix lists are very efficient access-lists and make implementation of prefix filtering on AS borders (and elsewhere) very easy. Please review the BGP presentation materials if there is any uncertainty as to how prefix lists work.

Lab Exercise

1. **Basic Configuration.** Each router team should configure their router to fit into the network layout depicted in Figure 1. Check all connections. Note that most links are using serial cables. Remember what was covered in Module 11!

2. **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. In the every day Internet, such address assignment is carried out by the Regional Internet Registry. AS16384 is the transit provider used in this module, and as such represents the Internet at large. A /16 network block has been assigned to that provider.

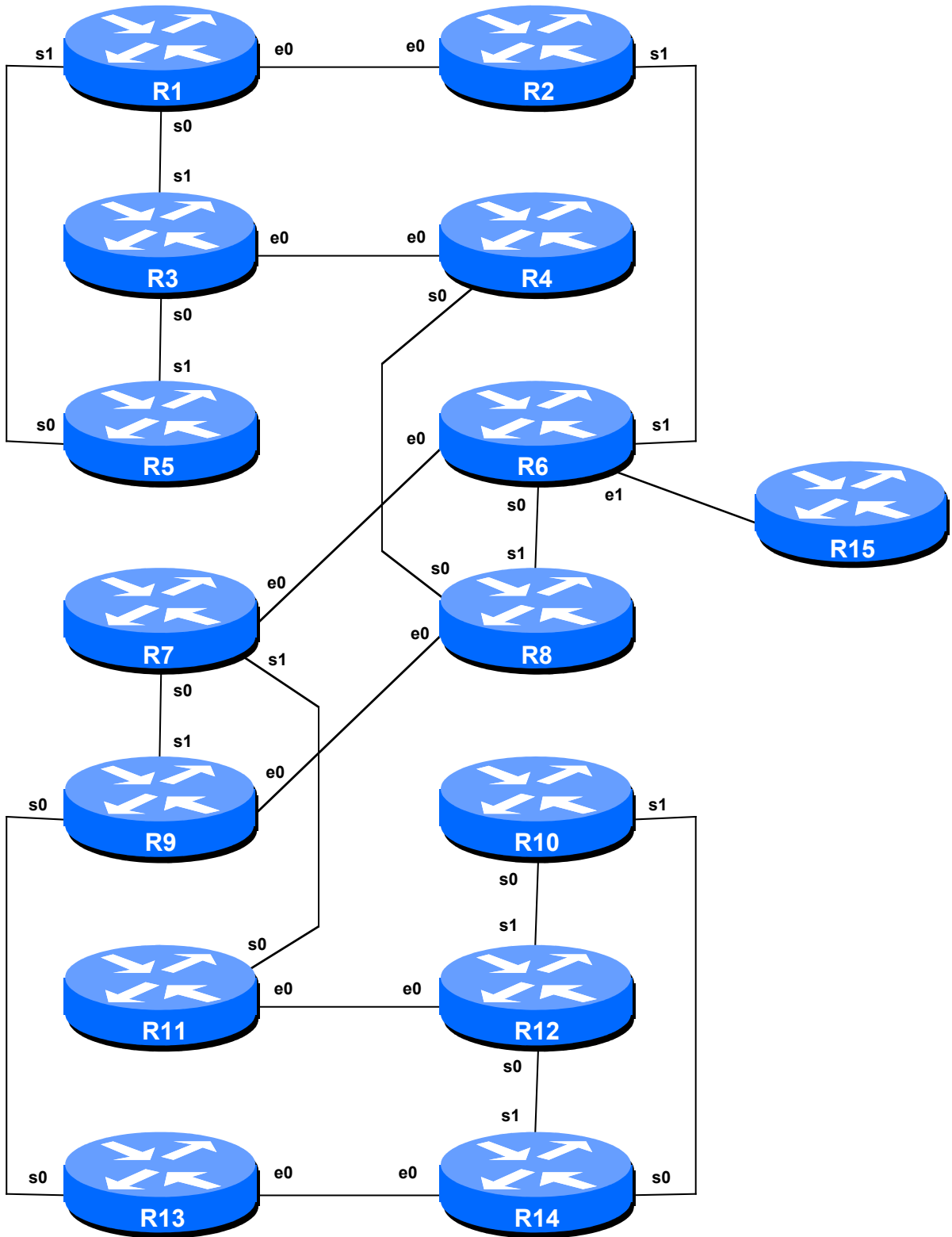|			
AS1	**120.10.0.0/19**	**AS5**	**121.35.0.0/19**
AS2	**120.19.0.0/19**	**AS6**	**121.99.0.0/19**
AS3	**120.73.0.0/19**	**AS16384**	**122.11.0.0/16**
AS4	**121.19.0.0/19**		

Figure 2 – Multihoming Lab Physical Layout

3. **Routing Protocols.** OSPF (area 0 only) and iBGP should now be configured between the routers for each of AS 1, AS6 and AS16384. The other 4 ASes do not require OSPF or iBGP as they only have a single router in them. Any interfaces which should not be running OSPF *MUST* be marked as passive in the configuration. And don't forget to use BGP peer groups for iBGP peers.

Checkpoint #1: *When you have properly configured your router, and the other routers in the AS are reachable (i.e. you can ping the other routers, and see BGP and OSPF prefixes in the routing table), please let the instructor know.*

Scenario One – Basic Configuration, no redundancy

The first scenario is not found in common practice but serves as a good introduction to the concepts of BGP multihoming as found on the Internet. And it will also serve as a reminder of some of the configuration concepts covered in the Basic ISP Workshop.

The customer has two upstream links, one to each ISP. Both links require to be used equally for traffic. The best way of achieving this is to announce half the address space on one link, and the other half of the address space on the other link. (The alternative is to simply announce the whole address block on each link. However, this probably will not achieve any useful degree of loadsharing.)

4. **Enable eBGP between the transit ASes.** AS2 and AS3 should now enable their eBGP link to AS16384. AS4 and AS5 should do likewise. All router teams in these ASes must ensure that they can see all the prefixes of AS2, AS3, AS4, AS5 and AS16384. If they are not there, work with your team members to ensure they appear. Don't forget the static pull-up route when injecting prefixes into BGP! Also, at this stage there is no need to install prefix filters between these ASes – if you would like to, don't forget that you need to allow through the network blocks of the ASes you are providing transit to.

5. **Prepare to enable eBGP between AS1 and its two upstreams.** AS1 should currently be running iBGP within its own network. To announce AS1's prefix to AS2 and AS3 we will take the /19 address block and divide it into two. The aim is to achieve relatively even utilisation of the links between AS1 and AS2/AS3, and common practice is to subdivide the address space. AS2 and AS3 will not announce any prefixes to AS1 – they will simply announce a default route. There is no need for any more routing information to be injected into the customer site.

6. **Create AS1 prefix lists.** First, create the prefix lists on the routers in AS1. For example, Router1 will announce the first subblock, Router3 will announce the second subblock. Both will accept the default route. Example for Router1:

```
ip prefix-list subblock1 permit 120.10.0.0/20
ip prefix-list default permit 0.0.0.0/0
```

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

4

7. **Create AS2 and AS3 prefix-lists.** The routers in AS2 and AS3 should only accept those prefixes which the customer is entitled to announce. So a prefix list needs to be installed on both Router2 and Router4 to do this. Notice in the example how a range of addresses has been listed in the permit statement – this is so that the customer can make changes without the upstream having to reconfigure their filters:

```
ip prefix-list Customer permit 120.10.0.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

8. **Configure eBGP in AS1.** With the prefix lists configured it is now possible to set up eBGP. It is good practice to configure the filters first, then configure BGP, not the other way around. This helps prevent accidents. Example configuration for Router 3:

```
ip prefix-list subblock2 permit 120.10.16.0/20
ip prefix-list default permit 0.0.0.0/0
!
router bgp 1
 neighbor <router4> remote-as 3
 neighbor <router4> description Peering with Router 4 in AS3
 neighbor <router4> prefix-list subblock2 out
 neighbor <router4> prefix-list default in
!
```

9. **Configure eBGP in AS3 with AS1.** AS3 is going to originate the default route in the peering with AS1. The BGP command "default-originate" is used to do this. Example configuration for Router 4:

```
router bgp 3
 neighbor <router3> remote-as 1
 neighbor <router3> description Multihomed Customer
 neighbor <router3> default-originate
 neighbor <router3> prefix-list Customer in
 neighbor <router3> prefix-list default out
!
```

10. **AS2 and AS3 eBGP configuration with AS16384.** Without further configuration changes in AS3, AS1 will be announced by AS2 and AS3 routers to other ASes. It is always good practice to only announce the prefixes you are entitled to announce to the Internet – so now is the time to add prefix filters to the peering with AS16384. The example below is for AS3:

```
ip prefix-list mynets permit 120.10.0.0/19 le 20
ip prefix-list mynets permit 120.73.0.0/19
!
router bgp 3
 neighbor <router8> remote-as 16384
 neighbor <router8> description Peering with AS16384 – The Internet
 neighbor <router8> prefix-list mynets out
!
```

Cisco Systems

11. **AS16384 eBGP configuration.** The configuration for the routers in AS16384 is left as an exercise. You can choose to filter your customer ASes if you wish – the prefix filters will be based on those configured by your "customers" in the previous step. You will announce all prefixes you have to your customers, so in this case there is no need for an outbound prefix filter (although you can construct one if you really wish to do so).

12. **AS6 and its upstreams of AS4 and AS5.** The same types of configuration concepts are also required on AS4, AS5 and AS6. AS6 is a multihomed customer of AS4 and AS5. The teams looking after the routers in these two ASes should use the above configuration examples as hints to set up their own peering sessions.

<u>Checkpoint #2:</u> 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

Scenario Two – Primary link and backup link

The second scenario is more commonly employed, especially where the customer has a large circuit to their upstream, and an inexpensive circuit they use almost exclusively for backup purposes to another ISP. (The case maybe that the primary ISP has good connection to the Internet, and the backup ISP is used only as a last resort – for technical, commercial, or political reasons.)

In this case, the whole address block is announced out of both links. However, the announcement going out the backup link has its AS path length increased so that it is at a lower priority. Likewise, the incoming default route announcement from the ISP is "weighted" using local-preference. (**Hint:** remember the purpose of changing the AS Path length? If in doubt, review the BGP presentation material.)

13. **Cleanup the configuration.** Remove the configuration which split the address blocks into two pieces and inserted them into the BGP table.

14. **Configure the main link.** Configure the main link between the customer AS and the ISP. For AS1, the link between Router1 and Router2 in AS2 is the main link – the link between Router3 and Router4 in AS3 is the backup. For AS6, the main link is between Router 12 and Router 11 in AS4. Example configuration for Router1:

```
ip prefix-list myblock permit 120.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
router bgp 1
 network 120.10.0.0 mask 255.255.224.0
 neighbor <router2> remote-as 2
```

```
 neighbor <router2> description Link to Router2 in AS2
 neighbor <router2> prefix-list myblock out
 neighbor <router2> prefix-list default in
!
ip route 120.10.0.0 255.255.224.0 null0
```

15. **Configure the backup link.** Configure the backup link between the private AS and the ISP. Increase the AS Path Length on outbound announcements to 3, and set local preference on inbound announcements to 80. Remember that the shortest AS Path Length and highest local-preference win during the BGP path selection process. To do this, use a route-map on the peering – you will require an inbound and outbound route-map. Example configuration for Router14:

```
ip prefix-list myblock permit 121.99.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map outfilter permit 10
 match ip address prefix-list myblock
 set as-path prepend 6 6 6
route-map outfilter permit 20
!
route-map infilter permit 10
 match ip address prefix-list default
 set local-preference 80
route-map infilter permit 20
!
router bgp 6
 network 121.99.0.0 mask 255.255.224.0
 neighbor <router13> remote-as 5
 neighbor <router13> description Link to Router13 in AS5
 neighbor <router13> prefix-list myblock out
 neighbor <router13> prefix-list default in
 neighbor <router13> route-map outfilter out
 neighbor <router13> route-map infilter in
!
ip route 121.99.0.0 255.255.224.0 null0
```

16. **Connectivity Test.** Check connectivity throughout the lab network. Each router team should be able to see all other routers in the room. When you are satisfied that BGP is working correctly, try running traceroutes to ensure that the primary paths are being followed. When you are satisfied this is the case, check that the backup functions (do this by disconnecting the cable between the two routers on the primary path) – you will see that the backup path is now used.

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.*

<div align="center">

STOP AND WAIT HERE

</div>

Cisco Systems

Scenario Three – More Controlled Loadsharing

The third scenario is the most commonly deployed. Most multihomed sites want to implement some kind of loadsharing on the circuits they have to their upstream provider. The example here discusses only two circuits, but the techniques work equally well for a greater number.

To do this, the whole address block is announced out of both links. On the first link it is left untouched, but on the second link it attracts an AS Path Length increase. Also on the second link, the address block is split into two pieces, with one subprefix being announced with no modification to AS Path Length.

17. **Clean up the configuration of AS1 and AS6.** Remove the configuration which set the weighting for the previous example – specifically the route-maps. They must be removed from the BGP configuration, and from the main configuration.

18. **Configure the address block and subprefixes in AS1 and AS6.** Modify the router configuration so that the /19 address block and one /20 subprefix is present in the BGP table. Also set up prefix lists to cater for these blocks. For example:

```
ip prefix-list aggregate permit 120.10.0.0/19
!
ip prefix-list subblocks permit 120.10.0.0/19 le 20
!
ip prefix-list default permit 0.0.0.0/0
!
router bgp 1
 network 120.10.0.0 mask 255.255.224.0
 network 120.10.0.0 mask 255.255.240.0
!
ip route 120.10.0.0 255.255.224.0 null0
ip route 120.10.0.0 255.255.240.0 null0
```

19. **Configure BGP in AS1 and AS6.** For AS1, the link between Router1 and Router2 in AS2 is the first link – the link between Router3 and Router4 in AS3 is the second (and is the one which should also announce the subprefix). For AS6, the first link is between Router 12 and Router 11 in AS4. Configure BGP on the border routers in the customer ASes so that the prefix and one sub prefix is announced to the direct peer as described earlier. For example, Router1 could announce *aggregate* as above, whereas Router3 could announce *aggregate* with a lengthened AS Path, and announce *subblock1* as is. For example on Router3:

```
route-map outfilter permit 10
 match ip address prefix-list aggregate
 set as-path prepend 1 1 1
route-map outfilter permit 20
!
route-map infilter permit 10
 match ip address prefix-list default
```

```
 set local-preference 80
route-map infilter permit 20
!
router bgp 1
 network 120.19.0.0 mask 255.255.224.0
 neighbor <router4> remote-as 3
 neighbor <router4> description Link to Router4 in AS3
 neighbor <router4> prefix-list subblocks out
 neighbor <router4> prefix-list default in
 neighbor <router4> route-map outfilter out
 neighbor <router4> route-map infilter in
```

20. **Connectivity test.** Check connectivity throughout the lab network. Each router team should be able to see all other routers in the room. When you are satisfied that BGP is working correctly, try running traceroutes to check the path being followed. Also check that backup via the alternative path still functions (do this by disconnecting the cable between the two routers on the primary path) – you will see that the backup path is now used.

__Checkpoint #4:__ 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.

21. **Check the network paths.** Run traceroutes between your router and other routers in the classroom. Ensure that all routers are reachable. If any are not, work with the other router teams to establish what might be wrong.

22. **Summary.** This module has covered the major situations where a customer requires to multihomed onto more than one service provider backbone. It has demonstrated how to implement this multihoming using prefix-lists, AS Path Length modifications and local-preference where appropriate.

CISCO SYSTEMS

Monday, June 20, 2005

CONFIGURATION NOTES

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