Module 12 – Multihoming to the Same ISP

Objective: To investigate various methods for multihoming onto the same upstream's backbone

Prerequisites: Module 11 and Multihoming Presentation

The following will be the common topology used.

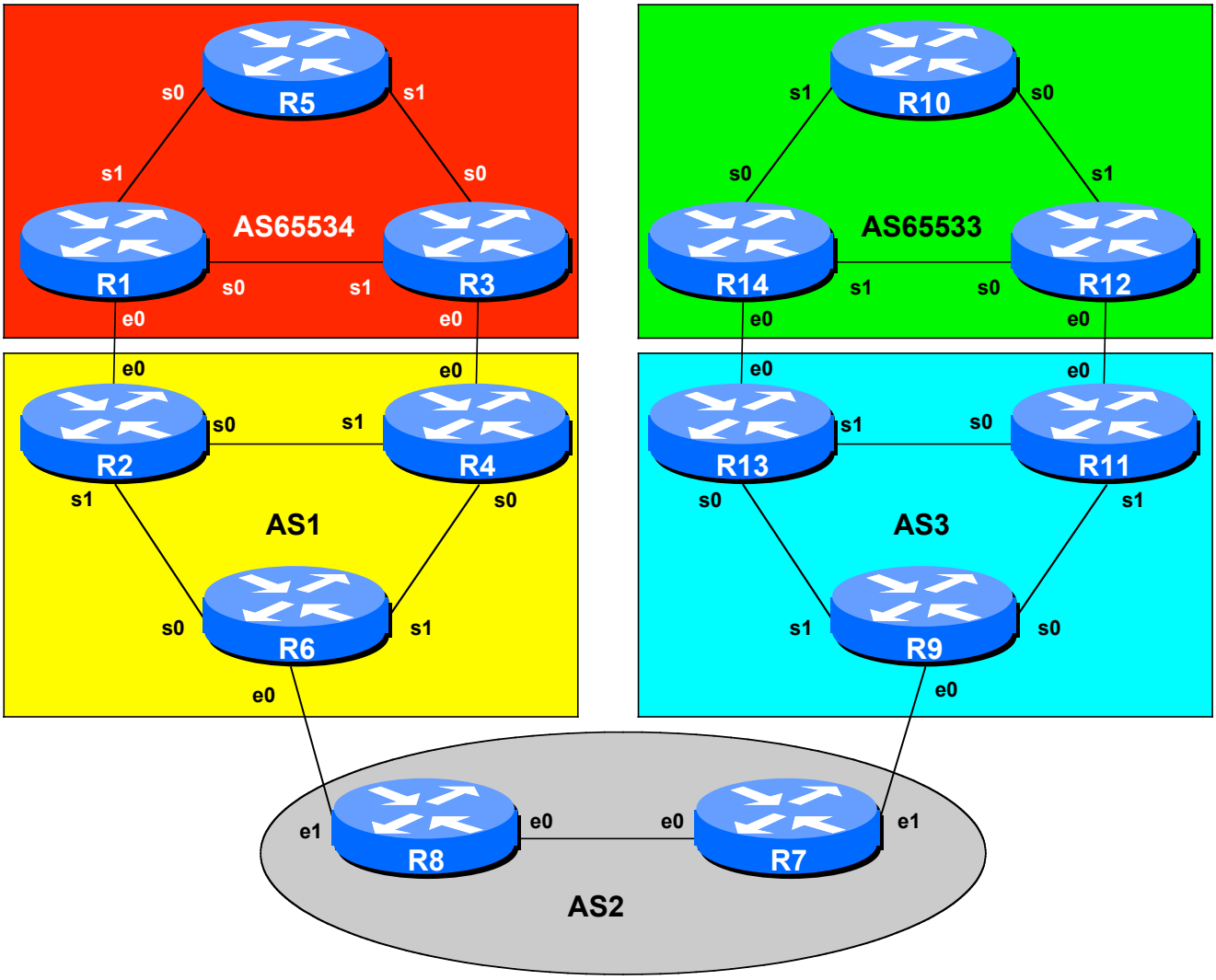


Figure 1 – ISP Lab Configuration

Lab Notes

The purpose of this module is to demonstrate multihoming in the situation where the customer has more than one connection to their upstream service provider. There are several situations where this is applicable:

- Enterprise customer requires more than one connection to the service provider to provide resiliency, and/or loadsharing.
- Enterprise customer has multiple sites which require connection to the Internet.
- Start-up ISP requires more than a single link to the Internet, but has little requirement to connect to more than one upstream ISP

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.

The accepted way to multihome to a single upstream ISP on the Internet today is to use a private AS number. The IANA defines the range 64512 to 65534 as being private ASes, in the same way that RFC1918 defines private address space. These ASes should **never** be visible on the Internet.

Finally, 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 more detailed policy**

There are rarely any exceptions to this. Prefix lists are very efficient access-lists and they make the 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. **Physical Layout.** 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.

   IMPORTANT: Each router team should ensure that their router has the basic configuration as covered in the first steps of Module 11. Figure 2 shows the physical layout of the lab. Note that you will have to move routers around to ensure you have sufficient ports. Also, try and minimise the disruption as you re-cable the lab.

2. **Basic Configuration.** Before starting to configure anything for this module, each router team should clean up the old configuration on the router they are using. The easiest way is probably to do a “write erase” and start afresh. The alternative is to remove any interface ip addressing, OSPF and all BGP configuration. The latter does need care – remember that all unused configuration must at all times be removed from the router.

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

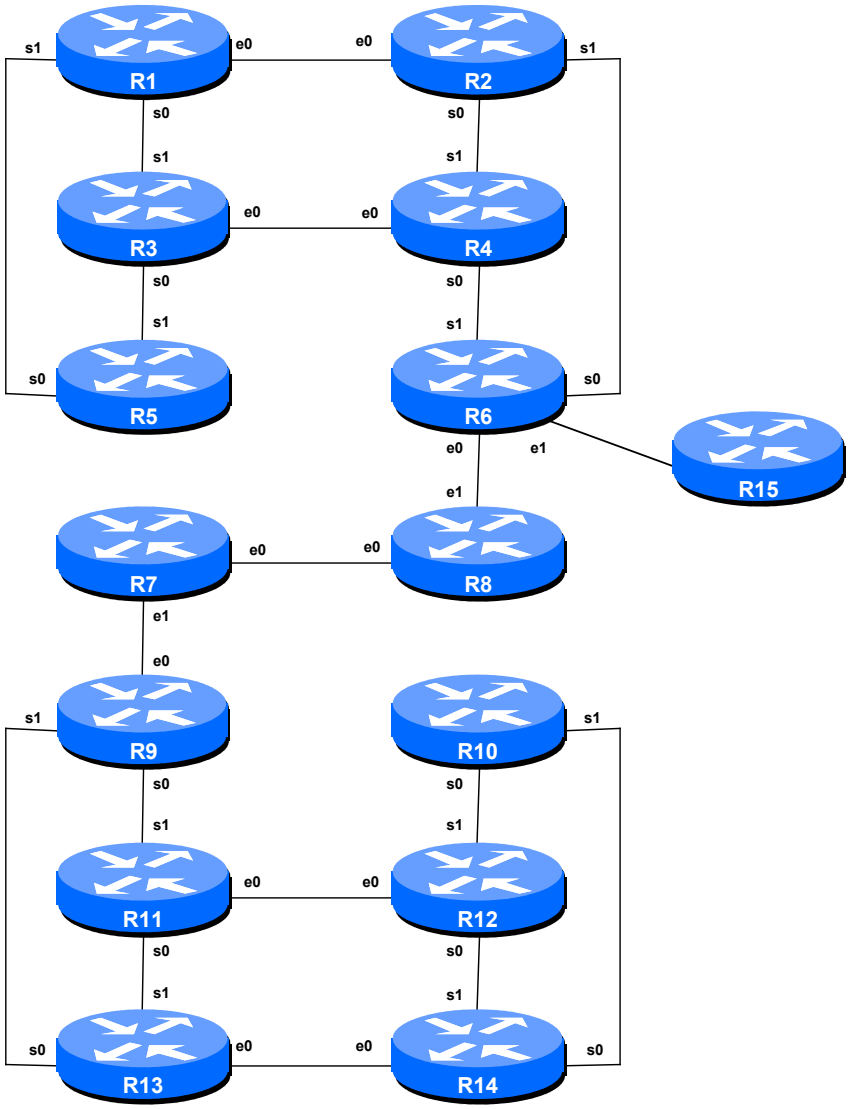


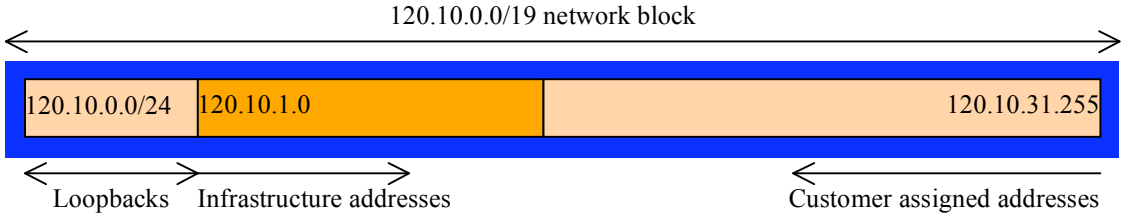
Figure 2 – Multihoming Lab Physical Layout

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. In the every day Internet such address assignment is carried out by the Regional Internet Registry.

AS65534	**120.10.0.0/19**	**AS2**	**121.19.0.0/19**
AS65533	**120.19.0.0/19**	**AS3**	**121.35.0.0/19**
AS1	**120.73.0.0/19**		

When constructing an addressing plan, don't forget to use a small block for loopback interfaces and another small block for point to point links. Also, agree between yourselves and your neighbouring ASes which addresses will be used for the point to point links between you. Remember, it is Internet convention that addresses from the upstream ISP's address block are used for point to point link addresses to their customers.

Suggestion:



4. **Routing Protocols.** OSPF (area 0 only) and iBGP should now be configured between the routers in each AS. 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 – Primary link and backup link

This first scenario is more commonly employed where the customer has two circuits to their upstream: a large circuit which is used for all the inbound and outbound traffic, and an inexpensive circuit they use almost exclusively for backup purposes.

In this case the whole address block is announced out of both links. However, the announcement going out the backup link is "weighted" using MEDs 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 MEDs and local-preference? If in doubt, review the BGP presentation material.)

5. **Enable eBGP between AS1, AS2 and AS3.** AS1 has a connection to AS2, and AS2 has a connection to AS3. The eBGP sessions between the relevant routers in those ASes should now be configured. All router teams in these ASes should ensure that they are able to see all the prefixes of AS1, AS2 and AS3 (basically there will be entries for 120.73.0.0/19, 121.19.0.0/19 and 121.35.0.0/19 in the BGP table). If they are not there, work with your team members to ensure they appear. Don't forget the static pull-up route!

6. **Prepare to enable eBGP between AS1 and AS65534.** Both AS65534 and AS1 should now be running iBGP within their ASes. To announce AS65534's prefix to AS1 we will take the /19 address block and announce it on both BGP peerings between the ASNs. AS1 will not announce any specific prefixes to AS65534 – it will simply announce a default route. There is no need for any more routing information to be injected into the customer site.

   Note: some customers request/demand a full Internet routing table to be announced to their network – if after education they still insist, give it to them, but creative ISPs often charge for this service. Plus the customer needs to be aware they will need a router with at least 64Mbytes of useable memory given today's size of the Internet routing table.

7. **Prepare to enable eBGP between AS3 and AS65533.** The process for enabling eBGP in this case is the same as for between AS1 and AS65534.

8. **Create AS65534 prefix lists.** First, create the prefix lists on the routers in AS65534. Both Router1 and Router3 will announce the /19. Both will accept the default route. Example for Router1:

```
ip prefix-list myblock permit 120.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

9. **Create AS1 prefix-lists.** The routers in AS1 are the customer aggregation routers and should only accept those prefixes which the customer is entitled to announce. So prefix lists need to be installed on Routers 2 and 4 to do this. The example given is for Router 4:

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

10. **Configure the main link.** Configure the main link between the private AS and the ISP. For AS65534, the link between Router1 and Router2 in AS1 is the main link – the link between Router3 and Router5 is the backup. For AS65533, the main link is between Router14 and Router13 in AS3. Example configuration for Router1:

```
router bgp 65534
 network 120.10.0.0 mask 255.255.224.0
 neighbor <router2> remote-as 1
 neighbor <router2> description Link to Router2 in AS1
 neighbor <router2> prefix-list myblock out
 neighbor <router2> prefix-list default in
!
ip route 120.10.0.0 255.255.224.0 null0
```

11. **Configure the backup link.** Configure the backup link between the private AS and the ISP. Set the metric on outbound announcements to 20, and set local preference on inbound announcements to 90. Remember that lowest metric 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 Router12:

```
ip prefix-list myblock permit 121.19.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map outfilter permit 10
 match ip address prefix-list myblock
 set metric 20
route-map outfilter permit 20
!
route-map infilter permit 10
 match ip address prefix-list default
 set local-preference 90
route-map infilter permit 20
!
router bgp 65533
 network 121.19.0.0 mask 255.255.224.0
 neighbor <router11> remote-as 3
 neighbor <router11> description Link to Router11 in AS3
 neighbor <router11> prefix-list myblock out
 neighbor <router11> prefix-list default in
```

```
 neighbor <router11> route-map outfilter out
 neighbor <router11> route-map infilter in
!
ip route 121.19.0.0 255.255.224.0 null0
```

12. **Configure eBGP in AS1.** AS1 is going to originate the default route in the peering with AS65534. The BGP command "default-originate" is used to do this. Example configuration for Router 2:

```
router bgp 1
 neighbor <router1> remote-as 65534
 neighbor <router1> description Dualhomed Customer
 neighbor <router1> default-originate
 neighbor <router1> prefix-list Customer in
 neighbor <router1> prefix-list default out
!
```

13. **Strip the private AS from external announcements from AS1.** Without further configuration changes in AS1, the private AS65534 will be announced by AS1 routers to other ASes. To stop private ASes from being announced, AS1 will need to configure the BGP command *remove-private-AS*. This needs to be done on all border routes in AS1 – in this case it is Router6.

```
ip prefix-list mynets permit 120.10.0.0/19 le 20
ip prefix-list mynets permit 120.73.0.0/19
!
router bgp 1
 neighbor <router8> remote-as 2
 neighbor <router8> description Peering with AS2
 neighbor <router8> remove-private-AS
 neighbor <router8> prefix-list mynets out
!
```

    Note: As a general rule, the remove-private-AS configuration should be included in all eBGP configurations on any router in the public Internet. Doing this ensures that any internal private ASN topology is not accidentally leaked to the Internet.

14. **AS3 and AS65533.** The same types of configuration concepts are also required on AS3 and AS65533. AS65533 is a multihomed customer of AS3. 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.

15. **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 #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. Notice that you still should* ***not*** *see any private ASes in the BGP table of AS2.*

STOP AND WAIT HERE

Scenario Two – Loadsharing (Method One)

Most dualhomed 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.

In this case, the whole address block is announced out of both links. Also, the address block is split into two pieces, with one subprefix being announced out of one link, and the other being announced out of the other link. The result of this is that traffic for the first /20 comes in one path, and traffic for the second /20 comes in the other path. If either path fails, the advertisement of the /19 address block (aggregate) ensures continued connectivity.

16. **Clean up the private ASes.** 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.

17. **Configure the address block and subprefixes in the private ASes.** Modify the router configuration so that the /19 address block and two /20 subprefixes are present in the BGP table. Also set up prefix lists to cater for these blocks. For example:

```
ip prefix-list subblock1 permit 120.10.0.0/19
ip prefix-list subblock1 permit 120.10.0.0/20
!
ip prefix-list subblock2 permit 120.10.0.0/19
ip prefix-list subblock2 permit 120.10.16.0/20
!
ip prefix-list default permit 0.0.0.0/0
!
router bgp 65534
 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
```

18. **Configure BGP in the private ASes.** Configure BGP on the border routers in the private ASes so that the prefix and one sub prefix is announced to the direct peer. For example, Router1 could announce *subblock1* as above, whereas Router3 could announce an equivalent *subblock2*. For example, for AS65533:

```
router bgp 65533
 neighbor <router11> remote-as 3
 neighbor <router11> description Link to Router11 in AS3
 neighbor <router11> prefix-list subblock1 out
 neighbor <router11> prefix-list default in
!
```

19. **Check filters in AS1 and AS3.** Ensure that the customer facing routers in AS1 and AS3 have appropriate filters to allow the /19 and /20 prefixes in from the private AS customers. The teams operating Routers 2, 4, 11 and 13 will need to verify that their filters will allow the aggregate and sub-prefix in, for example:

```
router bgp 1
 neighbor <router1> remote-as 65534
 neighbor <router1> description Link to Router1 in AS65534
 neighbor <router1> prefix-list Customer in
...
!
ip prefix-list Customer permit 120.10.0.0/19 le 20
!
```

20. Announcements to transit AS, AS2. Routers which connect to the transit provider, AS2, need to ensure that the subprefixes leaked by the customer private ASNs are not sent to the transit provider. While remove-private-as will strip out the private ASNs in the AS-PATH, the teams operating Router 6 and Router 9 will need an outbound filter on their peering with Router 8 and 7 respectively to block the /20s coming from their customer. For example:

```
router bgp 1
 neighbor <router8> remote-as 2
 neighbor <router8> description Link to Router8 in AS2
 neighbor <router8> prefix-list myprefixes out
...
!
ip prefix-list myprefixes permit 120.10.0.0/19
ip prefix-list myprefixes permit 120.73.0.0/19
!
```

21. 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 #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. Notice that you still should* ***not*** *see any private ASes in the BGP table of AS2.*

STOP AND WAIT HERE

Scenario Three – Loadsharing (Method Two)

The third scenario is a variation on the second scenario and provides another example.

As before the whole address block is announced out of both links. In fact, one of the key features of multihoming, and providing redundancy, is that the ISP's address blocks are always announced out of each external link. The key to loadbalancing is how those external announcements are made. In this example, one /20 is taken out of the /19 address block and announced on one link between the customer and the ISP as well as the /19. The other link sees just the standard announcement of the /19.

22. Clean up the private ASes. Remove the configuration which subdivided the address space for the previous example. Remember it is always very important to remove any configuration which isn't being used from the router.

23. **Configure the address block and subprefixes in the private ASes.** Modify the configuration of the routers in AS65534 and AS65533 so that the /19 address block and one /20 subprefix are 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 subblock permit 120.10.0.0/19
ip prefix-list subblock permit 120.10.0.0/20
!
ip prefix-list default permit 0.0.0.0/0
!
router bgp 65534
 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
```

24. **Configure BGP in the private ASes.** Configure BGP on the border routers in the private ASes so that the prefix and one sub prefix is announced to the direct peer on one link, and just the aggregate is announced on the other link. The router teams in AS65534 and AS65533 respectively should discuss amongst themselves who will leak the /20 prefix. For example, Router1 could announce *subblock* as above, whereas Router3 could announce just the aggregate. A configuration example for AS65533 might look like:

```
router bgp 65533
 neighbor <router11> remote-as 3
 neighbor <router11> description Link to Router11 in AS3
 neighbor <router11> prefix-list subblock out
 neighbor <router11> prefix-list default in
!
```

25. **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. Notice that you still should* ***not*** *see any private ASes in the BGP table of AS2.*

STOP AND WAIT HERE

Scenario Four – Scaling to support multiple dualhomed customers

The final scenario shows how to scale the third scenario described above. ISPs will offer multiple location connections as a service, so it is important to consider how to scale the configuration of the ISP's aggregation routers.

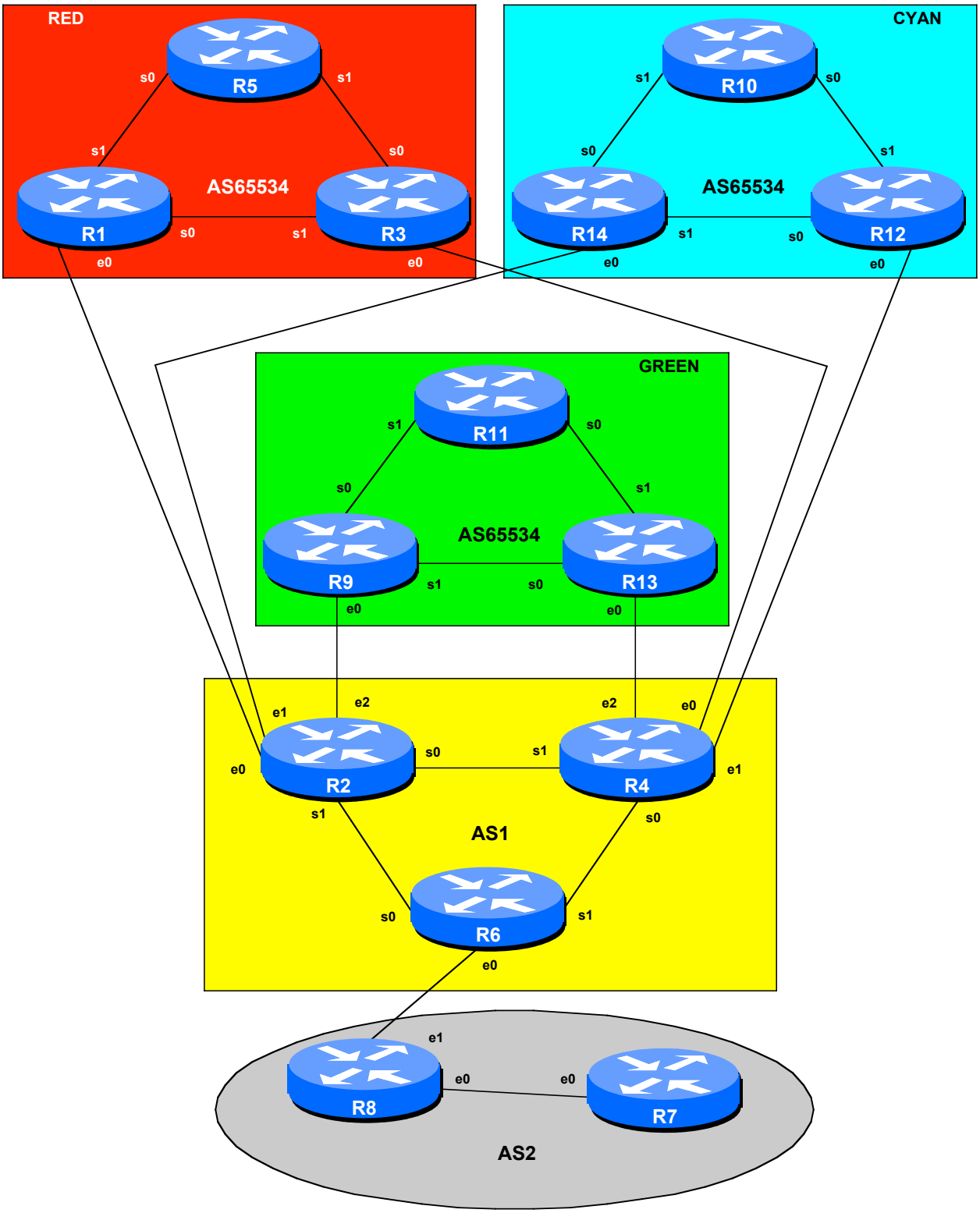


Figure 3 – Multiple Dualhomed Customers

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

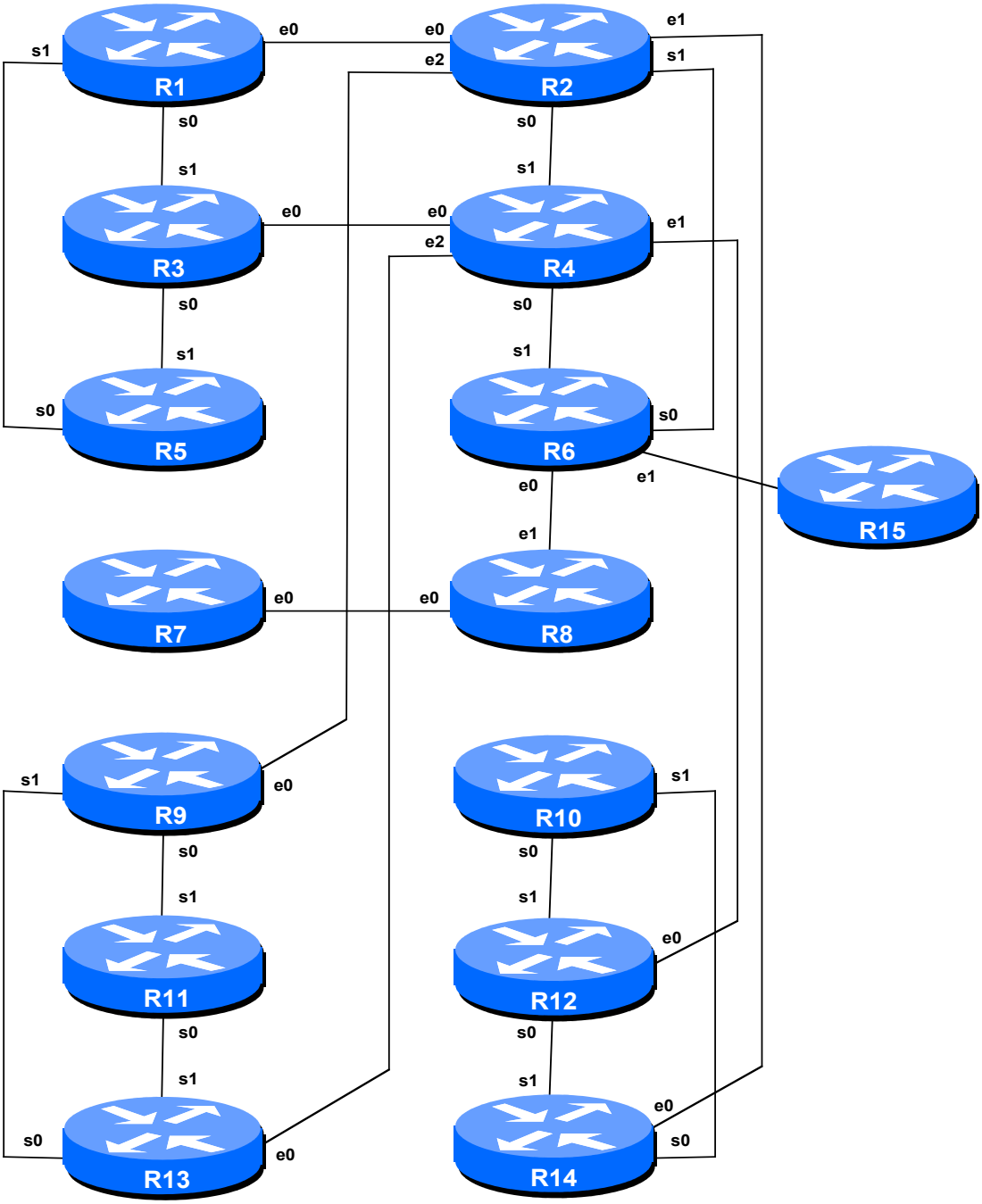


Figure 4 – RFC2270 Lab Physical Layout

The customer configuration is unchanged from the previous step – both the customer address block and its subprefixes are announced to the upstream. However, the customers can all use the same private ASN – the ASN information is not transited by the ISP, the customer simply point default at the upstream, so BGP loop detection is not an issue. This device is used to great effect by many ISPs for their multihoming customers – not only does it ensure that the ISP doesn't have multiple private ASes in their backbone, it also ensures that their configuration tools have less complexity to deal with. Simplicity is always the design goal in successful ISP operations.

Note: RFC2270 describes this type of multihoming in more detail.

26. **Reconfigure the network.** Routers in AS3 and AS65533 should be reconfigured to become customers of AS1. Please refer to Figure 4 for connection details. As previously, each router team will need to set up OSPF and iBGP within their own AS. So, for example, in the Green (middle) network, Routers 9, 11 and 13 will need to set up OSPF and iBGP within their own network.

27. **Configure the address blocks and subblocks within each private AS.** The address blocks to use are as follows:

AS2	**121.19.0.0/19**	**AS65534 (R9,11,13)**	**121.35.0.0/19**
AS1	**120.73.0.0/19**	**AS65534 (R10,12,14)**	**120.19.0.0/19**
AS65534 (R1,3,5)	**120.10.0.0/19**		

28. **Configure eBGP between each AS65534 customer and AS1.** Following the configuration hints in the previous section, each router team in AS65534 should configure their border routers to peer eBGP with AS1. Hint – the configuration should look something like:

```
ip prefix-list subblock1 permit y.y.0.0/19
ip prefix-list subblock1 permit y.y.0.0/20
!
ip prefix-list default permit 0.0.0.0/0
!
router bgp 65534
 network y.y.0.0 mask 255.255.224.0
 network y.y.0.0 mask 255.255.240.0
 neighbor x.x.x.x remote-as 3
 neighbor x.x.x.x description Link to RouterX in Aszzz
 neighbor x.x.x.x prefix-list subblock1 out
 neighbor x.x.x.x prefix-list default in
!
ip route y.y.0.0 255.255.224.0 null 0
ip route y.y.0.0 255.255.240.0 null 0
```

29. **Configure eBGP on AS1 border routers.** Scalable eBGP configuration on Routers 2 and 4 is required. If AS1 has multiple BGP customers, it ensures that the growth of the AS1 network is not hindered by having to handcraft a configuration for every new customer.

    The first first step is to use peer-groups for this. All the customers have the same outbound configuration, basically announce a default route. Remember that inbound policy can still be modified per peergroup neighbour – peergroups must have uniform **outbound** policy.

```
router bgp 1
 neighbor bgp-customers peer-group
 neighbor bgp-customers remote-as 65534
 neighbor bgp-customers default-originate
 neighbor bgp-customers prefix-list default out
!
```

    After creating the peer-group, it can be applied to every BGP customer connecting to the router. Don't forget to create a prefix-list to filter the customer's inbound announcements. This is still required on a per customer basis.

```
ip prefix-list default permit 0.0.0.0/0
ip prefix-list RedCustomer permit 120.10.0.0/19 le 20
ip prefix-list GreenCustomer permit 121.35.0.0/19 le 20
```

```
ip prefix-list CyanCustomer permit 120.19.0.0/19 le 20
!
router bgp 1
 neighbor x.x.x.x peer-group bgp-customers
 neighbor x.x.x.x description Red AS customer
 neighbor x.x.x.x prefix-list RedCustomer in
 neighbor x.x.x.x peer-group bgp-customers
 neighbor x.x.x.x description Green AS customer
 neighbor x.x.x.x prefix-list GreenCustomer in
 neighbor x.x.x.x peer-group bgp-customers
 neighbor x.x.x.x description Cyan AS customer
 neighbor x.x.x.x prefix-list CyanCustomer in
!
```

30. **Configuring AS1 border router to AS2.** The configuration of the AS1 border router connecting to AS2 (Router6) should be little changed from previous examples. It still requires the configuration to remove the private AS. And notice that it should only be allowing the customer blocks through, not the subprefixes of the customer blocks. As a reminder, the configuration of Router6 should look something like:

```
ip prefix-list mynets permit 120.10.0.0/19
ip prefix-list mynets permit 120.19.0.0/19
ip prefix-list mynets permit 120.73.0.0/19
ip prefix-list mynets permit 121.35.0.0/19
!
router bgp 1
 neighbor x.x.x.x remote-as 2
 neighbor x.x.x.x description Peering with AS2
 neighbor x.x.x.x remove-private-AS
 neighbor x.x.x.x prefix-list mynets out
```

If the prefix-list is omitted, AS1 will leak the sub-prefixes of its multihomed customers to AS2. As there is no need to leak these sub-prefixes, this is frowned upon as bad practise on the Internet today.

31. **Announcing Prefixes – Important Note.** In the above step, the outbound prefix-list "mynets" listed all the prefixes which needed to be announced to the Internet. This case is correct **ONLY** when the customer address blocks do not come from the upstream ISP – in other words, the customer has so called provider independent (PI) space. This situation is becoming more rare on the Internet, and more often customers will have provider aggregatable (PA) space – in other words, the customer's address space comes from the upstream ISP address block. In that situation, the upstream ISP **MUST NOT** announce the sub-prefixes – he should only announce his own block.

    For example, consider the situation where an ISP has the 121.19.0.0/19 address block. All his multihomed customers come out of this address block. His router configuration should be:

```
ip prefix-list myblock permit 120.10.0.0/19
!
router bgp 1
 network 120.10.0.0 mask 255.255.224.0
 neighbor x.x.x.x remote-as 2
 neighbor x.x.x.x description Peering with AS2
 neighbor x.x.x.x remove-private-AS
 neighbor x.x.x.x prefix-list myblock out
```

Notice how the "remove-private-AS" directive has been retained – it is an extra precaution in case of accidents with the prefix-list.

32. **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.

33. **Summary.** This module has covered the major situations where a customer requires to multihomed onto the service provider backbone. It has demonstrated how to implement this multihoming using prefix-lists, MEDs and local-preference where appropriate. It has also demonstrated the *remove-private-AS* BGP command, which ensures that private ASes are stripped out of any announcements to the wider Internet.

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

CONFIGURATION NOTES

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