Multihoming: Inbound Traffic Engineering

ISP Workshops

Last updated 11th May 2021

Acknowledgements

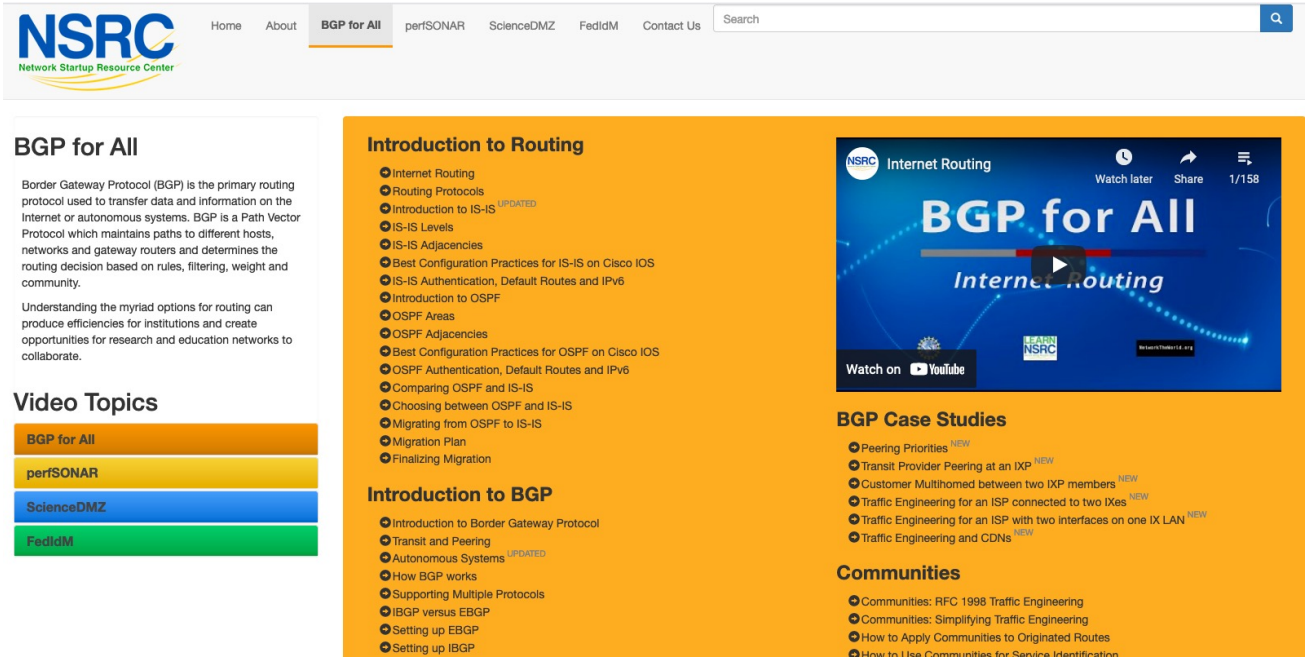
- This material originated from the Cisco ISP/IXP Workshop Programme developed by Philip Smith & Barry Greene

- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

BGP Videos

- NSRC has produced a library of BGP presentations (including this one), recorded on video, for the whole community to use
  - https://learn.nsrc.org/bgp



Basic Multihoming

- No frills multihoming
  - Only will look at inbound traffic engineering
    - (Traffic engineering of inbound traffic)
    - Of interest to most edge/access networks on the Internet
- Will look at two cases:
  - Multihoming with the same AS
  - Multihoming to different ASes
- Will keep the examples easy
  - Understanding easy concepts will make the more complex scenarios easier to comprehend
  - All assume that the site multihoming has a /19 address block

Multihoming: Inbound Traffic Engineering

- This type is most commonplace at the edge of the Internet
  - Networks here are usually concerned with inbound traffic flows
  - Outbound traffic flows being “nearest exit” is usually sufficient
- Can apply to the leaf Network Operator as well as Enterprise networks

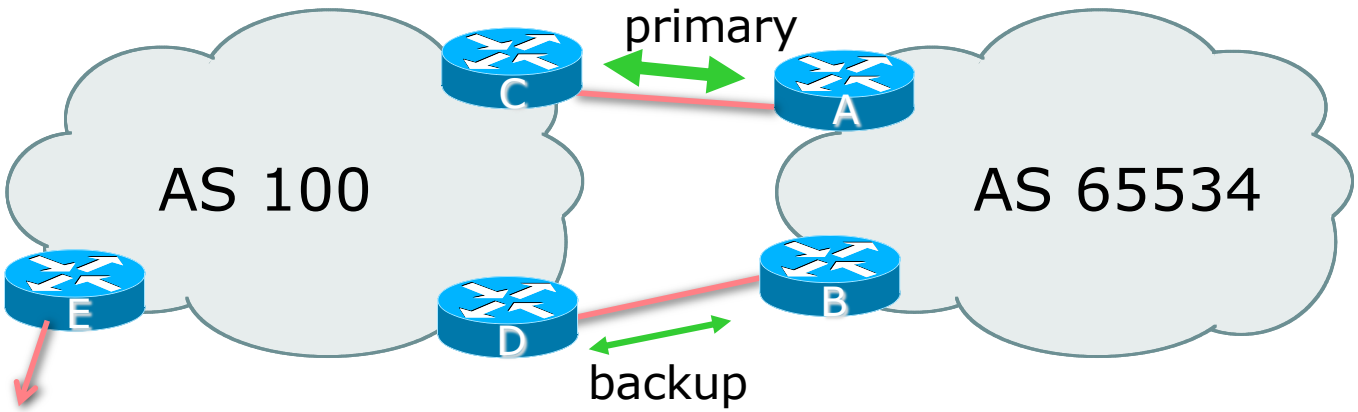
Two links to the same AS

One link primary, the other link backup only

Two links to the same AS (one as backup only)

- Applies when end-site has bought a large primary WAN link to their upstream and a small secondary WAN link as the backup
  - For example, primary path might be 20Mbps, backup might be 5Mbps

Two links to the same AS (one as backup only)



- AS100 removes private AS and any customer subprefixes from Internet announcement

Two links to the same AS (one as backup only)

- Announce /19 aggregate on each link
  - primary link:
    - Outbound – announce /19 unaltered
    - Inbound – receive default route
  - backup link:
    - Outbound – announce /19 with increased metric
    - Inbound – received default, and reduce local preference
- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

Two links to the same AS (one as backup only)

- Router A Configuration

```
router bgp 65534
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  neighbor 100.66.10.2 remote-as 100
  neighbor 100.66.10.2 description RouterC
  neighbor 100.66.10.2 prefix-list AGGREGATE out
  neighbor 100.66.10.2 prefix-list DEFAULT in
  neighbor 100.66.10.2 activate
!
ip prefix-list AGGREGATE permit 100.64.0.0/19
ip prefix-list DEFAULT permit 0.0.0.0/0
!
ip route 100.64.0.0 255.255.224.0 null0
```

Two links to the same AS (one as backup only)

- Router B Configuration

```
router bgp 65534
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  neighbor 100.66.10.6 remote-as 100
  neighbor 100.66.10.6 description RouterD
  neighbor 100.66.10.6 prefix-list AGGREGATE out
  neighbor 100.66.10.6 route-map MED10-out out
  neighbor 100.66.10.6 prefix-list DEFAULT in
  neighbor 100.66.10.6 route-map LP-LOW-in in
  neighbor 100.66.10.6 activate
!
```

..next slide

Two links to the same AS (one as backup only)

```
ip prefix-list AGGREGATE permit 100.64.0.0/19
ip prefix-list DEFAULT permit 0.0.0.0/0
!
ip route 100.64.0.0 255.255.224.0 null0
!
route-map MED10-out permit 10
 set metric 10
!
route-map LP-LOW-in permit 10
 set local-preference 90
!
```

Two links to the same AS (one as backup only)

- Router C Configuration (main link)

```
router bgp 100
 address-family ipv4
  neighbor 100.66.10.1 remote-as 65534
  neighbor 100.66.10.1 default-originate
  neighbor 100.66.10.1 prefix-list CUSTOMER in
  neighbor 100.66.10.1 prefix-list DEFAULT out
  neighbor 100.66.10.1 activate
!
ip prefix-list CUSTOMER permit 100.64.0.0/19
ip prefix-list DEFAULT permit 0.0.0.0/0
```

Two links to the same AS (one as backup only)

- Router D Configuration (backup link)

```
router bgp 100
 address-family ipv4
  neighbor 100.66.10.5 remote-as 65534
  neighbor 100.66.10.5 default-originate
  neighbor 100.66.10.5 prefix-list CUSTOMER in
  neighbor 100.66.10.5 prefix-list DEFAULT out
  neighbor 100.66.10.5 activate
!
ip prefix-list CUSTOMER permit 100.64.0.0/19
ip prefix-list DEFAULT permit 0.0.0.0/0
```

Two links to the same AS (one as backup only)

- Router E Configuration

```
router bgp 100
 address-family ipv4
  neighbor 100.66.10.17 remote-as 110
  neighbor 100.66.10.17 remove-private-AS
  neighbor 100.66.10.17 prefix-list CUSTOMER out
  neighbor 100.66.10.17 activate
!
ip prefix-list CUSTOMER permit 100.64.0.0/19
```

- Router E removes the private AS and customer's subprefixes from external announcements
- Private AS still visible inside AS100

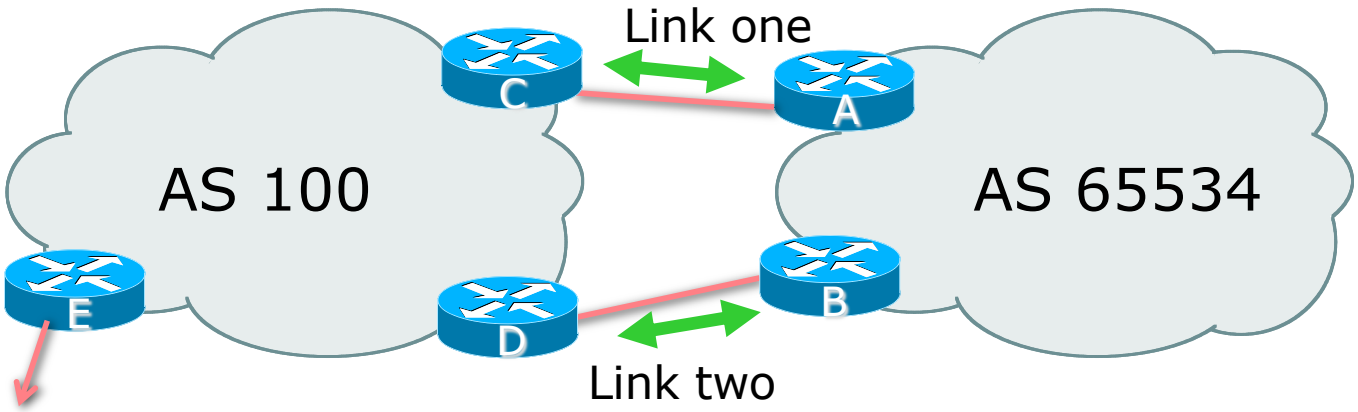
Two links to the same AS

With Loadsharing

Loadsharing to the same AS

- More common case
- End sites tend not to buy circuits and leave them idle, only used for backup as in previous example
- This example assumes equal capacity circuits
  - Unequal capacity circuits requires more refinement – see later

Loadsharing to the same AS



- Border router E in AS100 removes private AS and any customer subprefixes from Internet announcement

Loadsharing to the same AS (with redundancy)

- Announce /19 aggregate on each link
- Split /19 and announce as two /20s, one on each link
  - Basic inbound loadsharing
  - Assumes equal circuit capacity and even spread of traffic across address block
- Vary the split until "perfect" loadsharing achieved
- Accept the default from upstream
  - Basic outbound loadsharing by nearest exit
  - Okay in first approximation as most Network Operator and end-site traffic is inbound

Loadsharing to the same AS (with redundancy)

- Router A Configuration

```
router bgp 65534
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.0.0 mask 255.255.240.0
  neighbor 100.66.10.2 remote-as 100
  neighbor 100.66.10.2 prefix-list AS100-LINK1 out
  neighbor 100.66.10.2 prefix-list DEFAULT in
  neighbor 100.66.10.2 activate
!
ip prefix-list DEFAULT permit 0.0.0.0/0
ip prefix-list AS100-LINK1 permit 100.64.0.0/20
ip prefix-list AS100-LINK1 permit 100.64.0.0/19
!
ip route 100.64.0.0 255.255.240.0 null0
ip route 100.64.0.0 255.255.224.0 null0
```

Loadsharing to the same AS (with redundancy)

- Router B Configuration

```
router bgp 65534
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.16.0 mask 255.255.240.0
  neighbor 100.66.10.6 remote-as 100
  neighbor 100.66.10.6 prefix-list AS100-LINK2 out
  neighbor 100.66.10.6 prefix-list DEFAULT in
  neighbor 100.66.10.6 activate
!
ip prefix-list DEFAULT permit 0.0.0.0/0
ip prefix-list AS100-LINK2 permit 100.64.16.0/20
ip prefix-list AS100-LINK2 permit 100.64.0.0/19
!
ip route 100.64.16.0 255.255.240.0 null0
ip route 100.64.0.0 255.255.224.0 null0
```

Loadsharing to the same AS (with redundancy)

- Router C Configuration

```
router bgp 100
 address-family ipv4
  neighbor 100.66.10.1 remote-as 65534
  neighbor 100.66.10.1 default-originate
  neighbor 100.66.10.1 prefix-list CUSTOMER in
  neighbor 100.66.10.1 prefix-list DEFAULT out
  neighbor 100.66.10.1 activate
!
ip prefix-list CUSTOMER permit 100.64.0.0/19 le 20
ip prefix-list DEFAULT permit 0.0.0.0/0
```

- Router C only allows in /19 and /20 prefixes from customer block
- Router D configuration is identical

Loadsharing to the same AS (with redundancy)

- Router E Configuration

```
router bgp 100
 address-family ipv4
  neighbor 100.66.10.17 remote-as 110
  neighbor 100.66.10.17 remove-private-AS
  neighbor 100.66.10.17 prefix-list CUSTOMER out
  neighbor 100.66.10.17 activate
!
ip prefix-list CUSTOMER permit 100.64.0.0/19
```

- Private AS still visible inside AS100

Loadsharing to the same AS (with redundancy)

- Default route for outbound traffic?
  - Originate the default route in the IGP on the Border routers
    - Rely on IGP metrics for nearest exit
    - IGP originates default route as long as BGP puts default route in RIB
  - e.g. on router A using OSPF:

```
router ospf 65534
 default-information originate
```

  - e.g. on router A using IS-IS:

```
router isis as65534
 default-information originate route-map DEFAULT-ORIG
```

  - See the "BGP Case Studies" presentation for more details

Loadsharing to the same AS (with redundancy)

- Loadsharing configuration is only on customer router
- Upstream provider has to
  - Remove customer subprefixes from external announcements
  - Remove private AS from external announcements
- Could also use BGP communities
  - See the “BGP Communities” presentation for an example

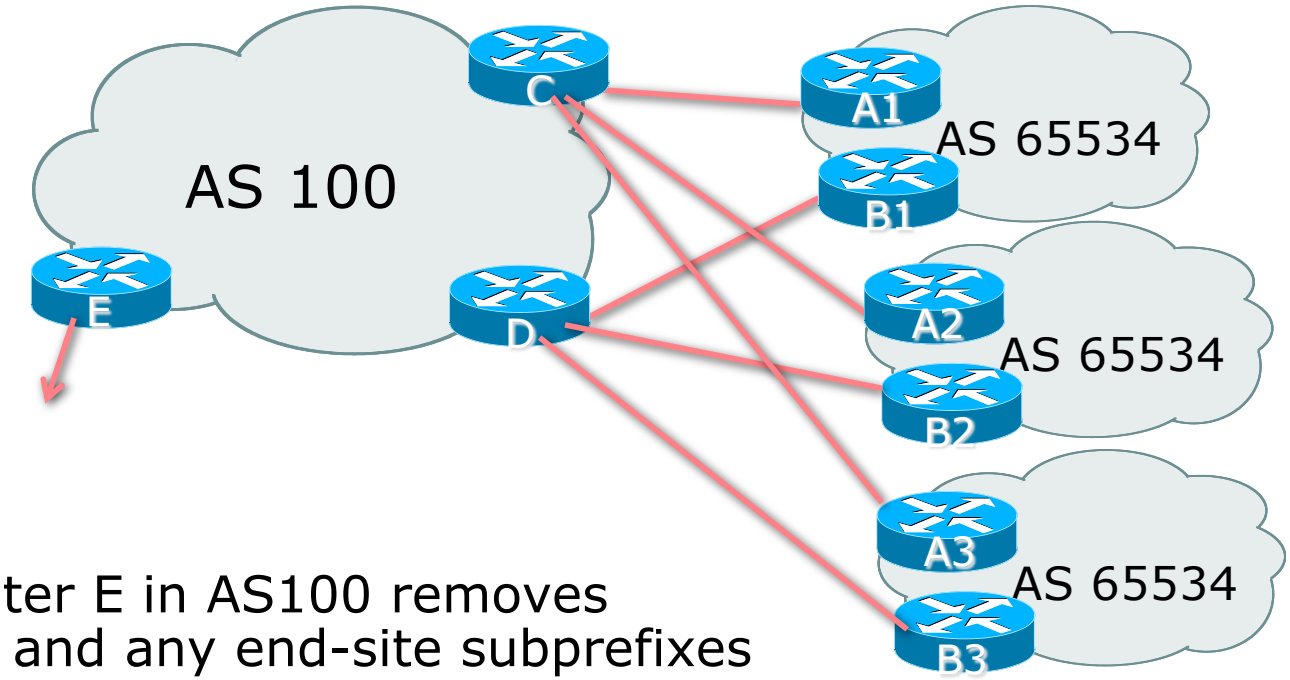
Two links to the same AS

Multiple Dualhomed End-sites
(RFC2270)

Multiple Dualhomed End-Sites (RFC2270)

- Unusual for any Network Operator just to have one dualhomed end-site
  - Valid/valuable service offering for an operator with multiple PoPs
  - Better for the operator than having end-site multihome with another provider!
- Look at scaling the configuration
  - ⇒ Simplifying the configuration
  - Using templates, peer-groups, etc
  - Every customer has the same configuration (basically)

Multiple Dualhomed End-Sites (RFC2270)



- Border router E in AS100 removes private AS and any end-site subprefixes from Internet announcement

Multiple Dualhomed End-sites (RFC2270)

- End-site announcements as per previous example
- Use the same private AS for each end-site
  - Documented in RFC2270
  - Address space is not overlapping
  - Each end-site hears default only
- Router A*n* and B*n* configuration same as Router A and B previously

Multiple Dualhomed End-Sites (RFC2270)

- Router A1 Configuration

```
router bgp 65534
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.0.0 mask 255.255.240.0
  neighbor 100.66.10.2 remote-as 100
  neighbor 100.66.10.2 prefix-list AS100-LINK1
out
  neighbor 100.66.10.2 prefix-list DEFAULT in
  neighbor 100.66.10.2 activate
!
ip prefix-list DEFAULT permit 0.0.0.0/0
ip prefix-list AS100-LINK1 permit 100.64.0.0/20
ip prefix-list AS100-LINK1 permit 100.64.0.0/19
!
ip route 100.64.0.0 255.255.240.0 null0
ip route 100.64.0.0 255.255.224.0 null0
```

Multiple Dualhomed End-Sites (RFC2270)

- Router B1 Configuration

```
router bgp 65534
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.16.0 mask 255.255.240.0
  neighbor 100.66.10.6 remote-as 100
  neighbor 100.66.10.6 prefix-list AS100-LINK2 out
  neighbor 100.66.10.6 prefix-list DEFAULT in
  neighbor 100.66.10.6 activate
!
ip prefix-list DEFAULT permit 0.0.0.0/0
ip prefix-list AS100-LINK2 permit 100.64.16.0/20
ip prefix-list AS100-LINK2 permit 100.64.0.0/19
!
ip route 100.64.0.0 255.255.224.0 null0
ip route 100.64.16.0 255.255.240.0 null0
```

Multiple Dualhomed End-Sites (RFC2270)

- Router C Configuration

```
router bgp 100
 address-family ipv4
  neighbor BGP-CUSTOMERS peer-group
  neighbor BGP-CUSTOMERS remote-as 65534
  neighbor BGP-CUSTOMERS default-originate
  neighbor BGP-CUSTOMERS prefix-list DEFAULT out
  neighbor 100.66.10.1 peer-group BGP-CUSTOMERS
  neighbor 100.66.10.1 description Customer One
  neighbor 100.66.10.1 prefix-list CUSTOMER1 in
  neighbor 100.66.10.1 activate
  neighbor 100.66.10.9 peer-group BGP-CUSTOMERS
  neighbor 100.66.10.9 description Customer Two
  neighbor 100.66.10.9 prefix-list CUSTOMER2 in
  neighbor 100.66.10.9 activate
```

Multiple Dualhomed End-Sites (RFC2270)

```
  neighbor 100.66.10.17 peer-group BGP-CUSTOMERS
  neighbor 100.66.10.17 description Customer Three
  neighbor 100.66.10.17 prefix-list CUSTOMER3 in
  neighbor 100.66.10.17 activate
!
ip prefix-list CUSTOMER1 permit 100.64.0.0/19 le 20
ip prefix-list CUSTOMER2 permit 100.67.64.0/19 le 20
ip prefix-list Customer3 permit 100.65.192.0/19 le 20
ip prefix-list DEFAULT permit 0.0.0.0/0
```

- Router C only allows in /19 and /20 prefixes from end-site block

Multiple Dualhomed End-Sites (RFC2270)

- Router D Configuration

```
router bgp 100
 address-family ipv4
  neighbor BGP-CUSTOMERS peer-group
  neighbor BGP-CUSTOMERS remote-as 65534
  neighbor BGP-CUSTOMERS default-originate
  neighbor BGP-CUSTOMERS prefix-list DEFAULT out
  neighbor 100.66.10.5 peer-group BGP-CUSTOMERS
  neighbor 100.66.10.5 description Customer One
  neighbor 100.66.10.5 prefix-list CUSTOMER1 in
  neighbor 100.66.10.5 activate
  neighbor 100.66.10.13 peer-group BGP-CUSTOMERS
  neighbor 100.66.10.13 description Customer Two
  neighbor 100.66.10.13 prefix-list CUSTOMER2 in
  neighbor 100.66.10.13 activate
```

Multiple Dualhomed End-Sites (RFC2270)

```
  neighbor 100.66.10.21 peer-group BGP-CUSTOMERS
  neighbor 100.66.10.21 description Customer Three
  neighbor 100.66.10.21 prefix-list CUSTOMER3 in
  neighbor 100.66.10.21 activate
!
ip prefix-list CUSTOMER1 permit 100.64.0.0/19 le 20
ip prefix-list CUSTOMER2 permit 100.67.64.0/19 le 20
ip prefix-list CUSTOMER3 permit 100.65.192.0/19 le 20
ip prefix-list DEFAULT permit 0.0.0.0/0
```

- Router D only allows in /19 and /20 prefixes from end-site block

Multiple Dualhomed End-Sites (RFC2270)

- Router E Configuration
  - Assumes end-site address space is not part of upstream's address block

```
router bgp 100
 address-family ipv4
  neighbor 100.66.10.17 remote-as 110
  neighbor 100.66.10.17 remove-private-AS
  neighbor 100.66.10.17 prefix-list CUSTOMERS out
  neighbor 100.66.10.17 activate
!
ip prefix-list CUSTOMERS permit 100.64.0.0/19
ip prefix-list CUSTOMERS permit 100.67.64.0/19
ip prefix-list CUSTOMERS permit 100.65.192.0/19
```

- Private AS still visible inside AS100

Multiple Dualhomed End-Sites (RFC2270)

- If end-sites' prefixes come from Network Operator's address block
  - Do NOT announce them to the Internet
  - Announce Network Operator aggregate only
- Router E configuration:

```
router bgp 100
 neighbor 100.66.10.17 remote-as 110
 neighbor 100.66.10.17 prefix-list AGGREGATE out
!
ip prefix-list AGGREGATE permit 100.64.0.0/12
```

Multihoming Summary

- Use private AS for multihoming to the same upstream
- Leak subprefixes to upstream only to aid loadsharing
- Upstream router E configuration is identical across all situations

Basic Multihoming

Multihoming to Different ASes

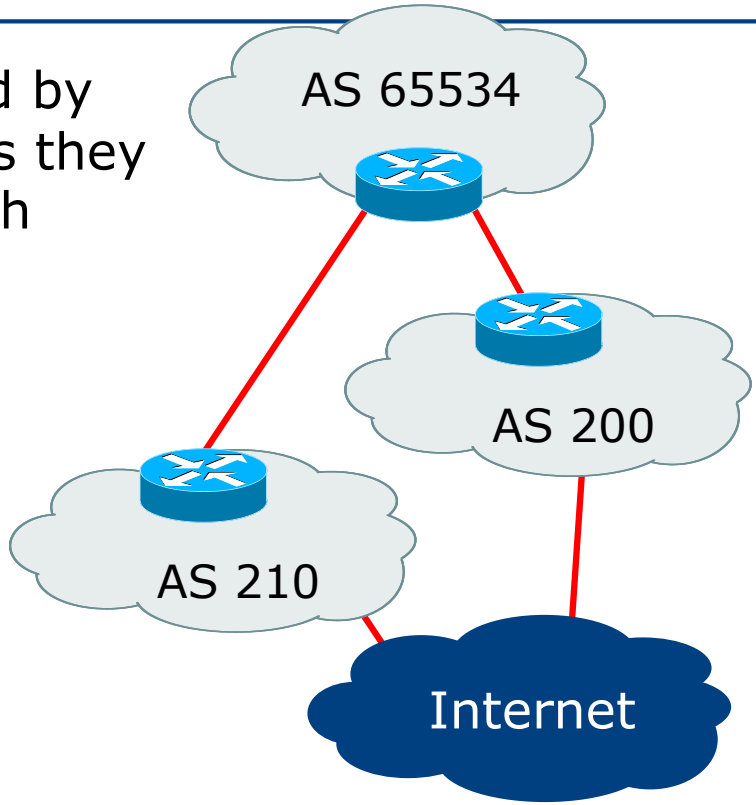
Two links to different ASes

- Use a Public AS number
  - Or use private AS number if agreed with the other Network Operator
  - But some people don't like the "inconsistent-AS" which results from use of a private AS number
- Address space comes from
  - Both upstreams *or*
  - Regional Internet Registry
  - NB. Very hard to multihome with address space from both upstreams due to typical operational policy in force to day
- Configuration concepts very similar to those used for two links to the same AS

Inconsistent-AS?

- Viewing the prefixes originated by AS65534 in the Internet shows they appear to be originated by both AS210 and AS200
  - This is NOT bad
  - Nor is it illegal
- Cisco IOS command is

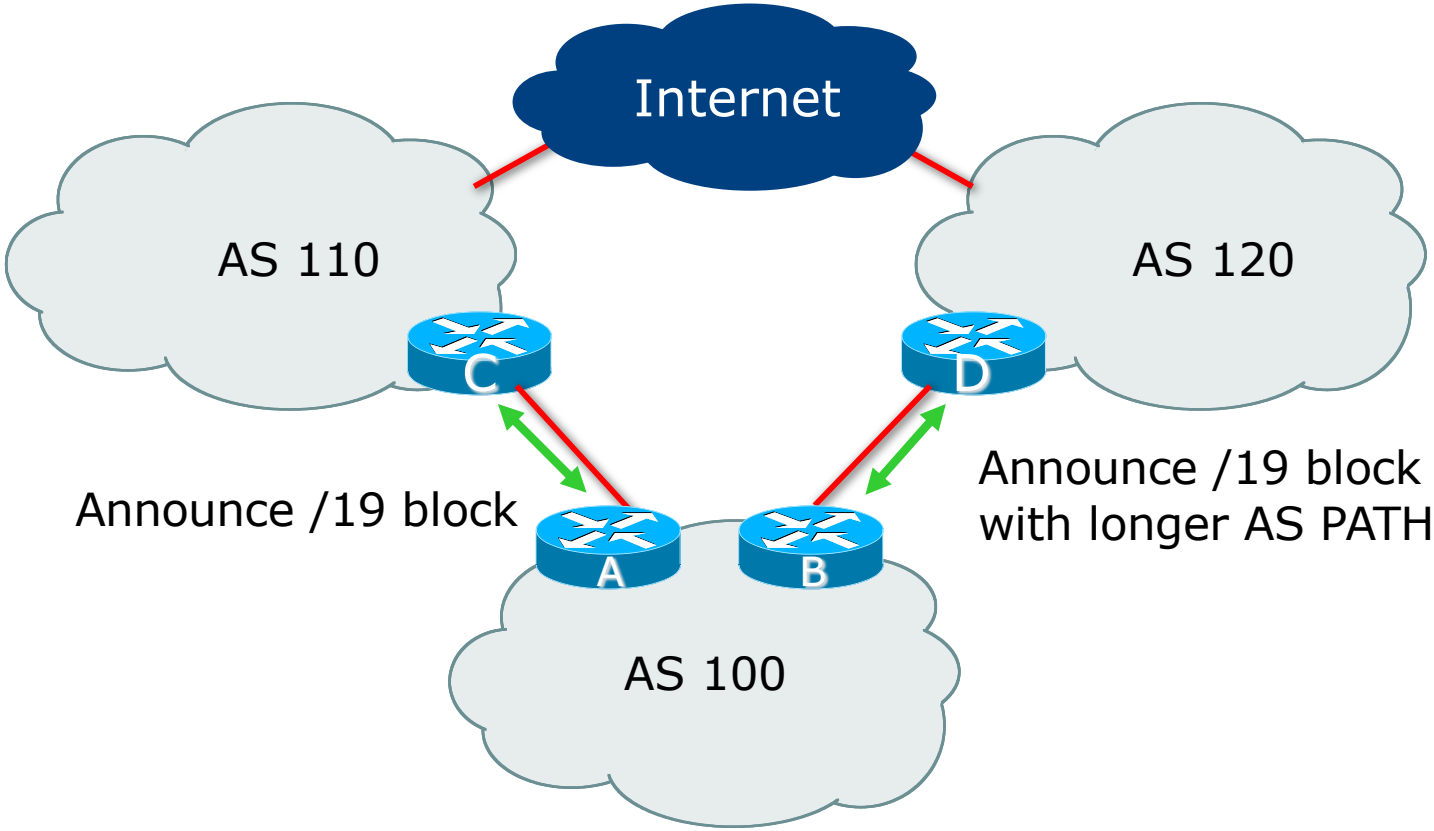
```
show ip bgp inconsistent-as
```



Two links to different ASes

One link primary, the other link backup only

Two links to different ASes (one as backup only)



Two links to different ASes (one as backup only)

- Announce /19 aggregate on each link
  - Primary link makes standard announcement
  - Backup link lengthens the AS PATH by using AS PATH prepend
- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

Two links to different ASes (one as backup only)

- Router A Configuration

```
router bgp 100
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  neighbor 100.66.10.1 remote-as 110
  neighbor 100.66.10.1 prefix-list AGGREGATE out
  neighbor 100.66.10.1 prefix-list DEFAULT in
  neighbor 100.66.10.1 activate
!
ip prefix-list AGGREGATE permit 100.64.0.0/19
ip prefix-list DEFAULT permit 0.0.0.0/0
!
ip route 100.64.0.0 255.255.224.0 null0
```

Two links to different ASes (one as backup only)

- Router B Configuration

```
router bgp 100
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  neighbor 100.67.5.1 remote-as 120
  neighbor 100.67.5.1 prefix-list AGGREGATE out
  neighbor 100.67.5.1 route-map AS120-PREPEND out
  neighbor 100.67.5.1 prefix-list DEFAULT in
  neighbor 100.67.5.1 route-map LP-LOW in
  neighbor 100.67.5.1 activate
!
...next slide...
```

Two links to different ASes (one as backup only)

```
ip route 100.64.0.0 255.255.224.0 null0
!
ip prefix-list AGGREGATE permit 100.64.0.0/19
ip prefix-list DEFAULT permit 0.0.0.0/0
!
route-map AS120-PREPEND permit 10
 description Three prepends to AS120
 set as-path prepend 100 100 100
!
route-map LP-LOW permit 10
 description All routes local pref 80
 set local-preference 80
!
```

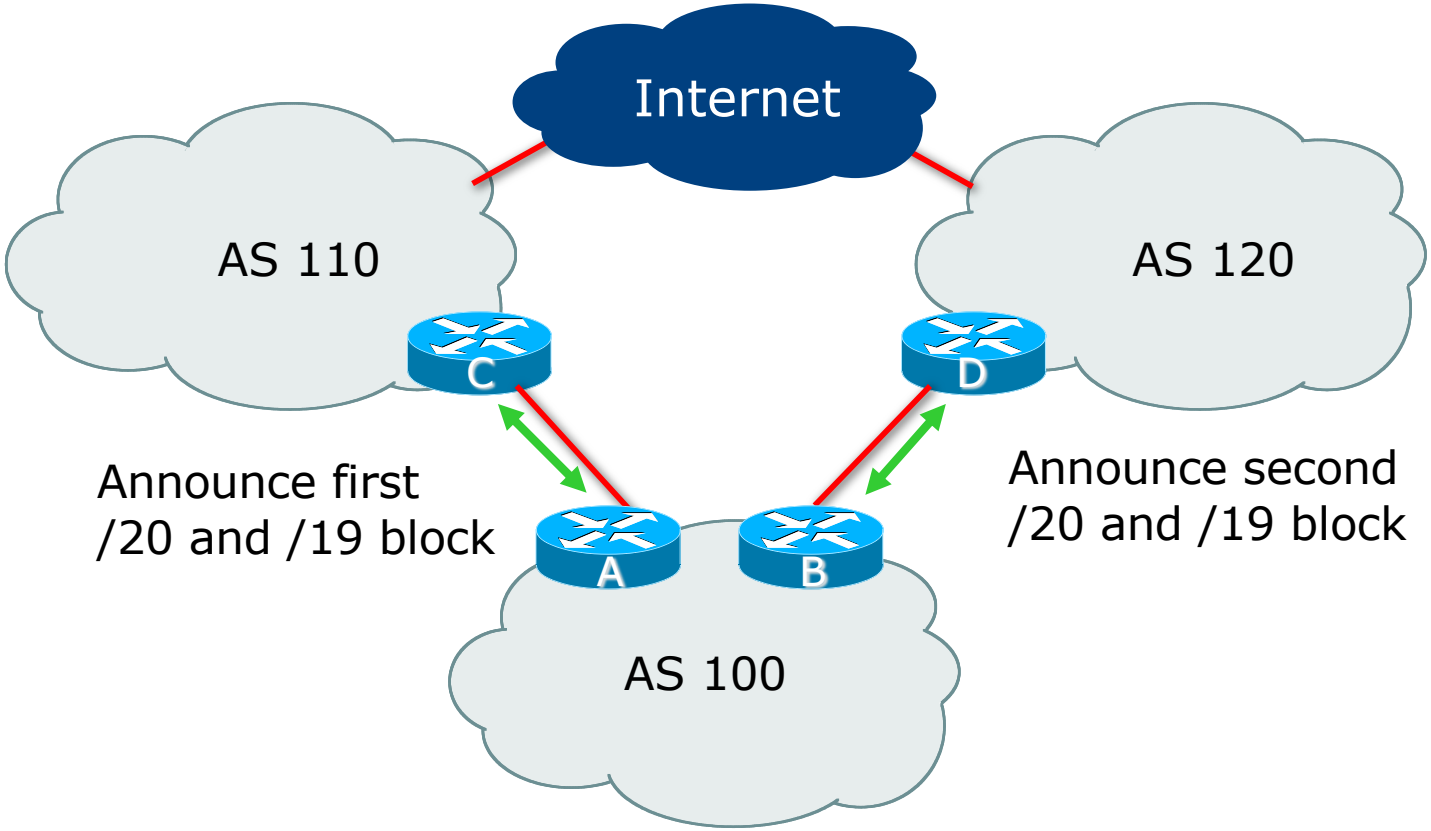
Two links to different ASes (one as backup only)

- Not a common situation as most sites tend to prefer using whatever capacity they have
  - (Useful when two competing ISPs agree to provide mutual backup to each other)
- But it shows the basic concepts of using local-prefs and AS-path prepends for engineering traffic in the chosen direction

Two links to different ASes

With Loadsharing

Two links to different ASes (with loadsharing)



Two links to different ASes (with loadsharing)

- Announce /19 aggregate on each link
- Split /19 and announce as two /20s, one on each link
  - Basic inbound loadsharing
- When one link fails, the announcement of the /19 aggregate via the other ISP ensures continued connectivity

Two links to different ASes (with loadsharing)

- Router A Configuration

```
router bgp 100
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.0.0 mask 255.255.240.0
  neighbor 100.66.10.1 remote-as 110
  neighbor 100.66.10.1 prefix-list AS110-OUT out
  neighbor 100.66.10.1 prefix-list DEFAULT in
  neighbor 100.66.10.1 activate
!
ip route 100.64.0.0 255.255.224.0 null0
ip route 100.64.0.0 255.255.240.0 null0
!
ip prefix-list DEFAULT permit 0.0.0.0/0
ip prefix-list AS110-OUT permit 100.64.0.0/20
ip prefix-list AS110-OUT permit 100.64.0.0/19
```

Two links to different ASes (with loadsharing)

- Router B Configuration

```
router bgp 100
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.16.0 mask 255.255.240.0
  neighbor 100.67.5.1 remote-as 120
  neighbor 100.67.5.1 prefix-list AS120-OUT out
  neighbor 100.67.5.1 prefix-list DEFAULT in
  neighbor 100.67.5.1 activate
!
ip route 100.64.0.0 255.255.224.0 null0
ip route 100.64.16.0 255.255.240.0 null0
!
ip prefix-list DEFAULT permit 0.0.0.0/0
ip prefix-list AS120-OUT permit 100.64.0.0/19
ip prefix-list AS120-OUT permit 100.64.16.0/20
```

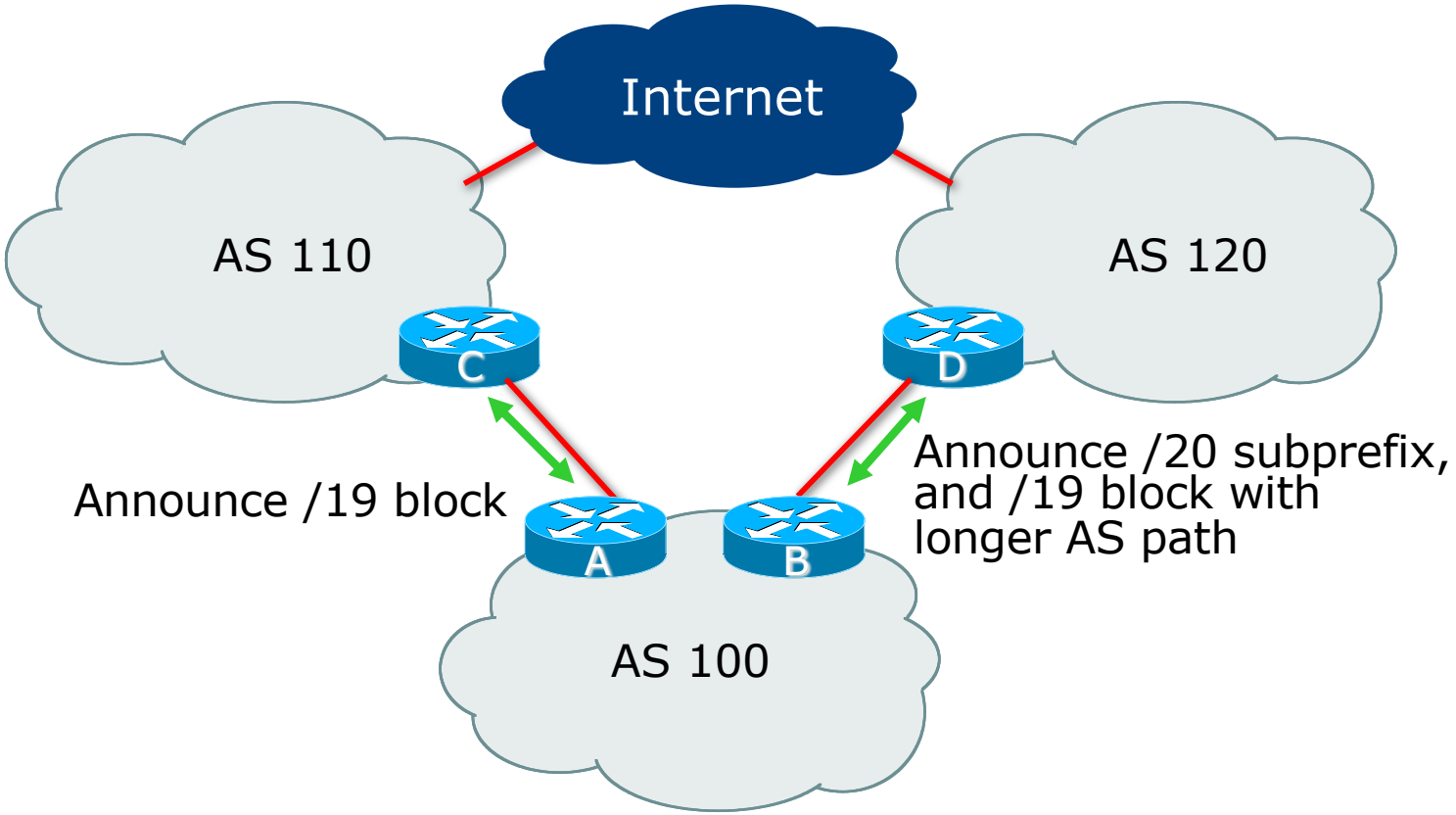
Two links to different ASes (with loadsharing)

- Loadsharing in this case is very basic
- But shows the first steps in designing a load sharing solution
  - Start with a simple concept
  - And build on it…!

Two links to different ASes

More Controlled Loadsharing

Loadsharing with different ASes



Loadsharing with different ASes

- Announce /19 aggregate on each link
  - On first link, announce /19 as normal
  - On second link, announce /19 with longer AS PATH, and announce one /20 subprefix
    - Controls loadsharing between upstreams and the Internet
- Vary the subprefix size and AS PATH length until "perfect" loadsharing achieved
- Still require redundancy!

Loadsharing with different ASes

- Router A Configuration

```
router bgp 100
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  neighbor 100.66.10.1 remote-as 110
  neighbor 100.66.10.1 prefix-list DEFAULT in
  neighbor 100.66.10.1 prefix-list AS110-OUT out
  neighbor 100.66.10.1 activate
!
ip route 100.64.0.0 255.255.224.0 null0
!
ip prefix-list AS110-OUT permit 100.64.0.0/19
!
ip prefix-list DEFAULT permit 0.0.0.0/0
```

Loadsharing with different ASes

- Router B Configuration

```
router bgp 100
 address-family ipv4
  network 100.64.0.0 mask 255.255.224.0
  network 100.64.16.0 mask 255.255.240.0
  neighbor 100.67.5.1 remote-as 120
  neighbor 100.67.5.1 prefix-list DEFAULT in
  neighbor 100.67.5.1 prefix-list AS120-OUT out
  neighbor 100.67.5.1 route-map AGGREGATE-PREPEND out
  neighbor 100.67.5.1 activate
!
ip route 100.64.0.0 255.255.224.0 null0
ip route 100.64.16.0 255.255.240.0 null0
!
...next slide...
```

Loadsharing with different ASes

```
route-map AGGREGATE-PREPEND permit 10
 description Find aggregate and set three prepends
 match ip address prefix-list AGGREGATE
 set as-path prepend 100 100
!
route-map AGGREGATE-PREPEND permit 20
 description All other routes are untouched
!
ip prefix-list DEFAULT permit 0.0.0.0/0
!
ip prefix-list AS120-OUT permit 100.64.0.0/19
ip prefix-list AS120-OUT permit 100.64.16.0/20
!
ip prefix-list AGGREGATE permit 100.64.0.0/19
!
```

Loadsharing with different ASes

- This example is more commonplace
- Shows how Network Operators and end-sites subdivide address space frugally, as well as use the AS-PATH prepend concept to optimise the load sharing between different ISPs
- Notice that the /19 aggregate block is ALWAYS announced

Summary

Summary

- Previous examples dealt with simple case
- Load balancing inbound traffic flow
  - Achieved by modifying outbound routing announcements
  - Aggregate is always announced
- We have not looked at outbound traffic flow
  - For now this is left as “nearest exit”

Multihoming: Inbound Traffic Engineering

ISP Workshops