BGP Multihoming

ISPCON 2K Tutorial, Melbourne

15 August 2000

Cisco Systems

Introduction

- **Presenter:**

  Philip Smith, Consulting Engineer

  Office of the CTO, Cisco Systems

  e-mail: pfs@cisco.com

- **Please ask questions**

Agenda

- Multihoming
- Configuration Examples
- Community Usage
- Transit
- International Collocation

Multihoming Definition

- More than one link external to the local network
  - two or more links to the same ISP
  - two or more links to different ISPs
- Usually two external facing routers
  - one router gives link and provider redundancy only

Multihoming

- **The scenarios described here apply equally well to end sites being customers of ISPs and ISPs being customers of other ISPs**
- **Implementation detail may be different**

end site ® ISP	**ISP controls config**
ISP1 ® ISP2	**ISPs share config**

AS Numbers

- An Autonomous System Number is required by BGP
- Obtained from upstream ISP or Regional Registry
- Necessary when you have links to more than one ISP or exchange point

Configuring Policy (Cisco IOS)

- **Assumptions:**

  prefix-lists are used throughout

  easier/better/faster than access-lists

- **Three BASIC Principles**

  prefix-lists to filter prefixes

  filter-lists to filter ASNs

  route-maps to apply policy

Originating Prefixes

- **Basic Assumptions**

  MUST announce assigned address block to Internet

  MAY also announce subprefixes - reachability is not guaranteed

  RIR minimum allocation is /20 - several ISPs filter RIR blocks on this boundary - “Net Police”

Part of the "Net Police" prefix list

```
!! RIPE
ip prefix-list FILTER permit 62.0.0.0/8 ge 12 le 20
ip prefix-list FILTER permit 193.0.0.0/8 ge 12 le 20
ip prefix-list FILTER permit 194.0.0.0/7 ge 12 le 20
ip prefix-list FILTER permit 212.0.0.0/7 ge 12 le 20
!! APNIC
ip prefix-list FILTER permit 61.0.0.0/8 ge 12 le 20
ip prefix-list FILTER permit 202.0.0.0/7 ge 12 le 20
ip prefix-list FILTER permit 210.0.0.0/7 ge 12 le 20
!! ARIN
ip prefix-list FILTER permit 63.0.0.0/8 le 20
ip prefix-list FILTER permit 64.0.0.0/8 le 20
ip prefix-list FILTER permit 199.0.0.0/8 le 20
ip prefix-list FILTER permit 200.0.0.0/8 le 20
ip prefix-list FILTER permit 204.0.0.0/6 le 20
ip prefix-list FILTER permit 208.0.0.0/7 le 20
ip prefix-list FILTER permit 216.0.0.0/8 le 20
```

“Net Police” prefix list issues

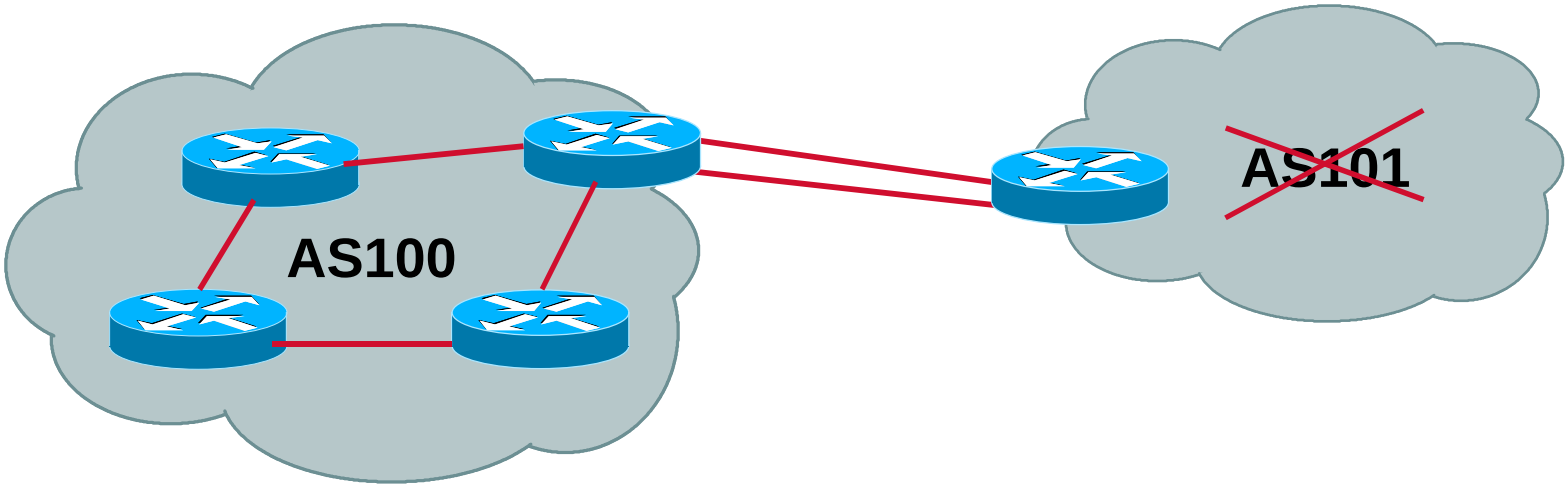
- meant to “punish” ISPs who won’t and don’t aggregate
- impacts legitimate multihoming
- impacts regions where domestic backbone is unavailable or costs $$$ compared with international bandwidth
- hard to maintain - requires updating when RIRs start allocating from new address blocks
- don’t do it unless consequences understood

Multihoming Options

Multihoming Scenarios

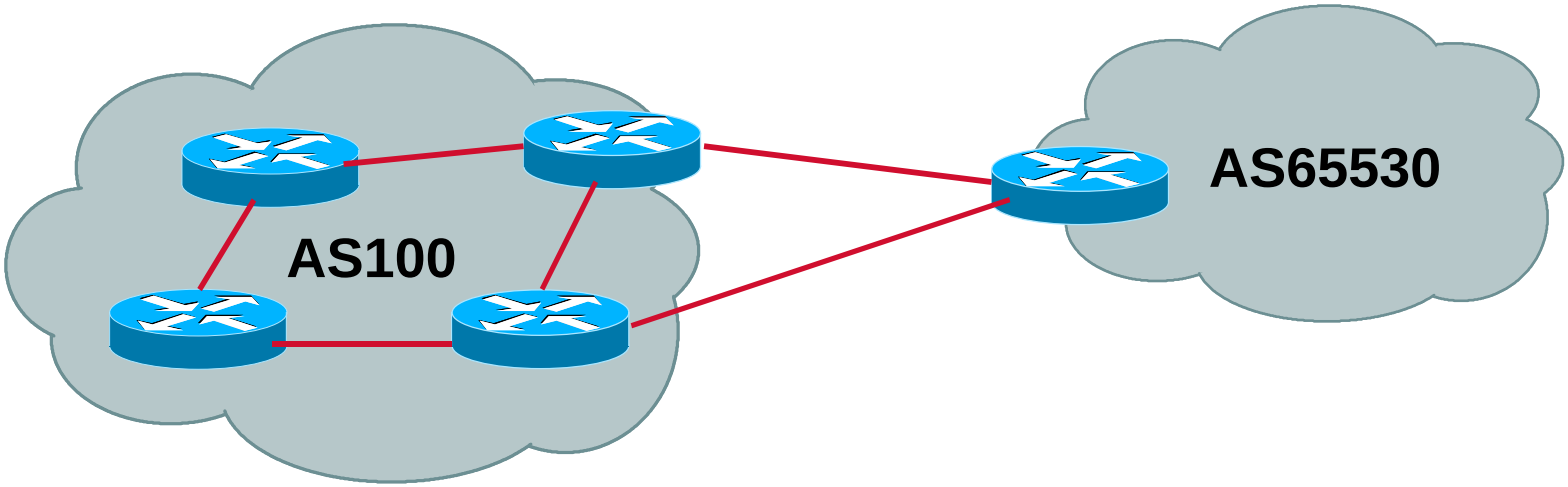
- Stub network
- Multi-homed stub network
- Multi-homed network
- Configuration Options

Stub Network



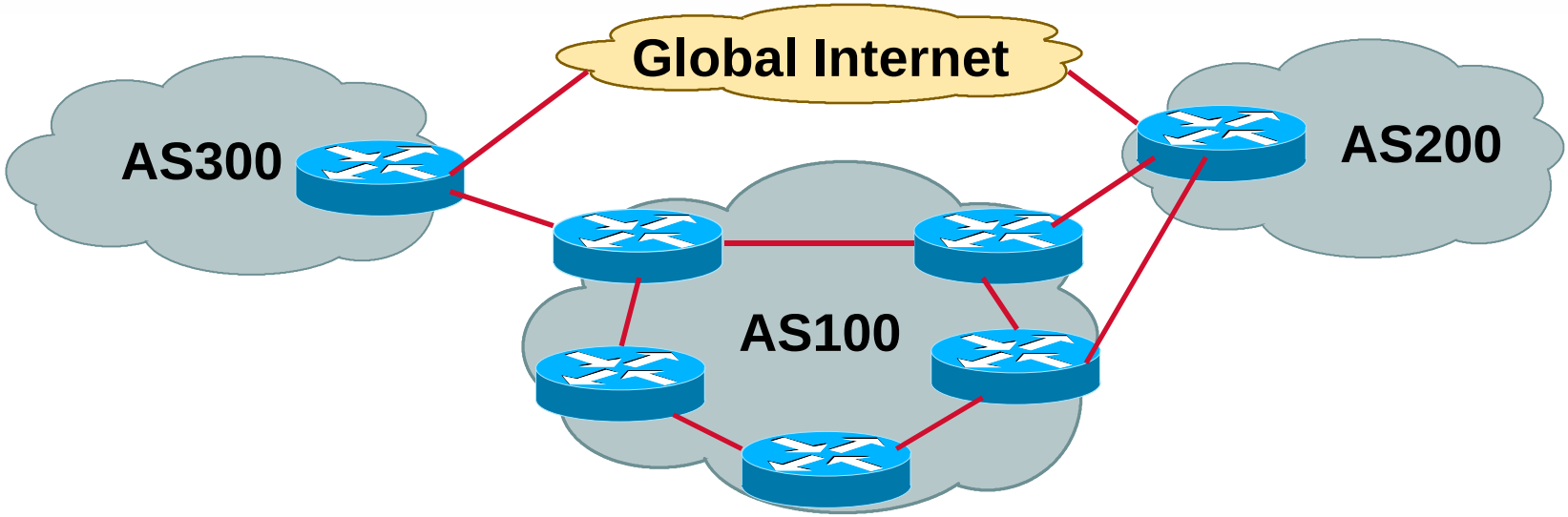
- No need for BGP
- Point static default to upstream ISP
- Upstream ISP advertises stub network
- Policy confined within upstream ISP's policy

Multi-homed Stub Network



- Use BGP (not IGP or static) to loadshare
- Use private AS (ASN > 64511)
- Upstream ISP advertises stub network
- Policy confined within upstream ISP's policy

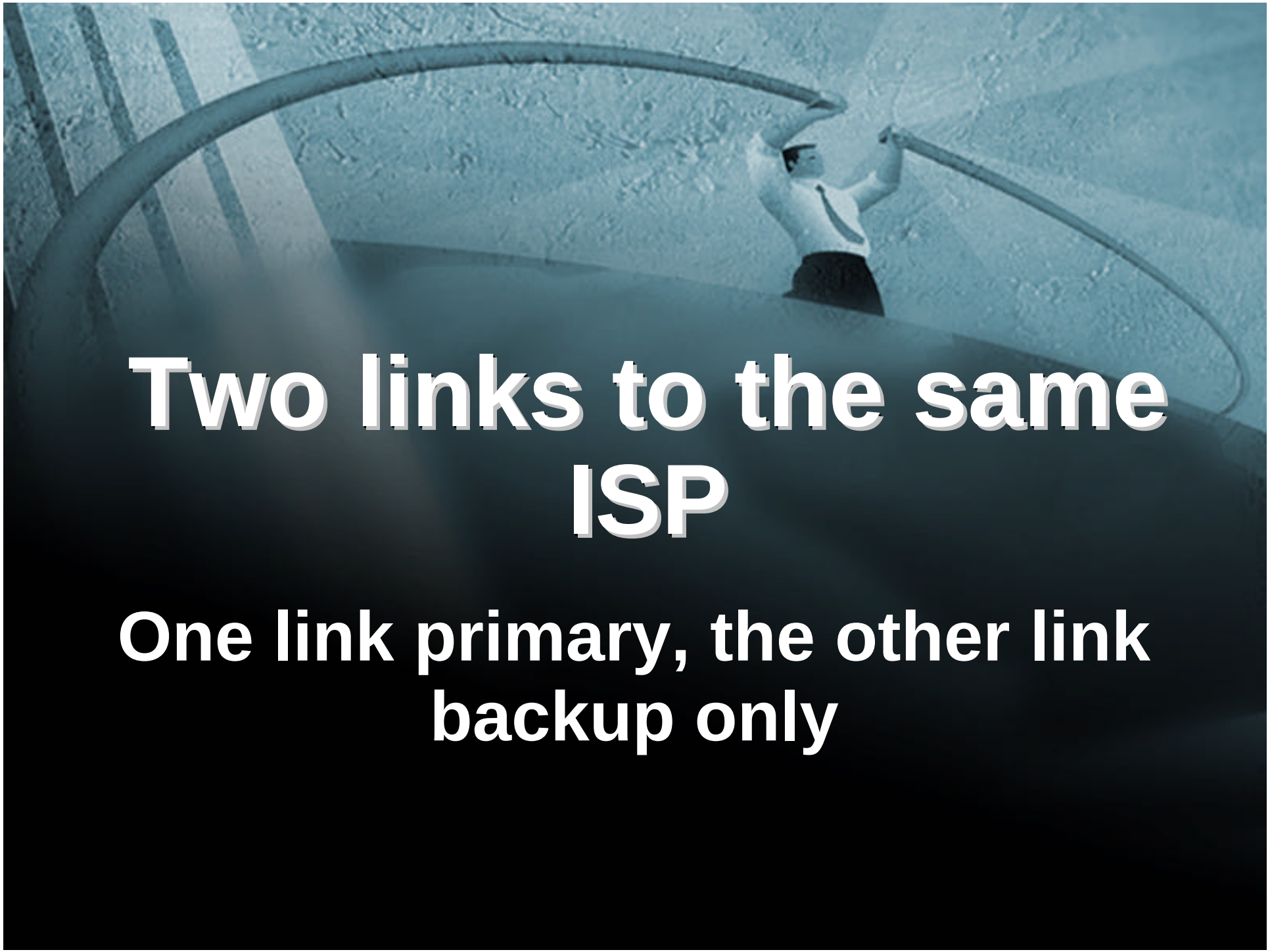
Multi-Homed Network



- **Many situations possible**
  - **multiple sessions to same ISP**
  - **secondary for backup only**
  - **load-share between primary and secondary**
  - **selectively use different ISPs**

Multiple Sessions to ISPs

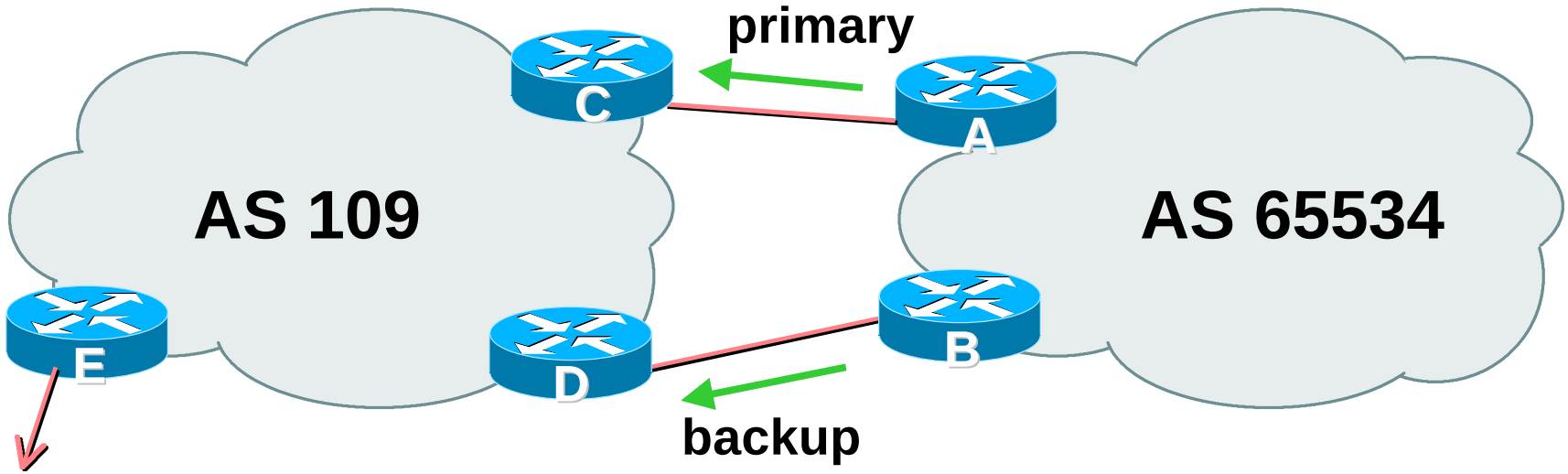
- Planning and some work required to achieve load sharing
  - Point default towards one ISP
  - Learn selected prefixes from second ISP
  - Modify the number of prefixes learnt to achieve acceptable load sharing
  - Think about network addressing scheme
- No magic solution



Two links to the same ISP

One link primary, the other link backup only

Two links to the same ISP



- **AS109 removes private AS and any customer subprefixes from Internet announcement**

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

- **Announce /19 aggregate on each link**

  primary link makes standard announcement

  backup link increases metric on outbound, and reduces local-pref on inbound

- **When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity**

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

- **Router A Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.2 remote-as 109
 neighbor 222.222.10.2 description RouterC
 neighbor 222.222.10.2 prefix-list aggregate out
 neighbor 222.222.10.2 prefix-list default in
!
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
```

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

- **Router B Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.6 remote-as 109
 neighbor 222.222.10.6 description RouterD
 neighbor 222.222.10.6 prefix-list aggregate out
 neighbor 222.222.10.6 route-map routerD-out out
 neighbor 222.222.10.6 prefix-list default in
 neighbor 222.222.10.6 route-map routerD-in in
!
..next slide
```

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

```
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
 match ip address prefix-list aggregate
 set metric 10
route-map routerD-out permit 20
!
route-map routerD-in permit 10
 set local-preference 90
!
```

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

- **Router C Configuration (main link)**

```
router bgp 109
 neighbor 222.222.10.1 remote-as 65534
 neighbor 222.222.10.1 default-originate
 neighbor 222.222.10.1 prefix-list Customer in
 neighbor 222.222.10.1 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

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

- **Router D Configuration (backup link)**

```
router bgp 109
 neighbor 222.222.10.5 remote-as 65534
 neighbor 222.222.10.5 default-originate
 neighbor 222.222.10.5 prefix-list Customer in
 neighbor 222.222.10.5 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

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

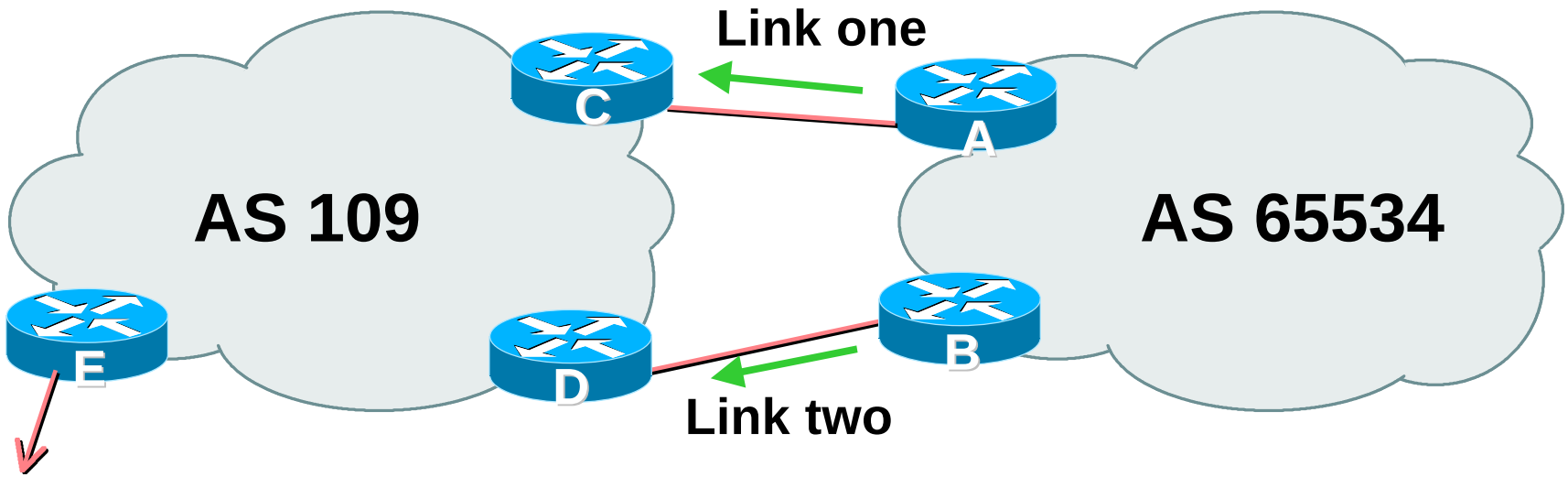
- **Router E Configuration**

```
router bgp 109
 neighbor 222.222.10.17 remote-as 110
 neighbor 222.222.10.17 remove-private-AS
 neighbor 222.222.10.17 prefix-list Customer out
!
ip prefix-list Customer permit 221.10.0.0/19
```

- **Router E removes the private AS and customer's subprefixes from external announcements**
- **Private AS still visible inside AS109**


Two links to the same ISP
With Redundancy and Loadsharing

Two links to the same ISP (with redundancy)



- **AS109 removes private AS and any customer subprefixes from Internet announcement**

Loadsharing to the same ISP

- 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

Two links to the same ISP

- **Router A Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.0.0 mask 255.255.240.0
 neighbor 222.222.10.2 remote-as 109
 neighbor 222.222.10.2 prefix-list routerC out
 neighbor 222.222.10.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerC permit 221.10.0.0/20
ip prefix-list routerC permit 221.10.0.0/19
!
ip route 221.10.0.0 255.255.240.0 null0
ip route 221.10.0.0 255.255.224.0 null0
```

Two links to the same ISP

- **Router B Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.16.0 mask 255.255.240.0
 neighbor 222.222.10.6 remote-as 109
 neighbor 222.222.10.6 prefix-list routerD out
 neighbor 222.222.10.6 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerD permit 221.10.16.0/20
ip prefix-list routerD permit 221.10.0.0/19
!
ip route 221.10.0.0 255.255.224.0 null0
ip route 221.10.16.0 255.255.240.0 null0
```

Two links to the same ISP

- **Router C Configuration**

```
router bgp 109
 neighbor 222.222.10.1 remote-as 65534
 neighbor 222.222.10.1 default-originate
 neighbor 222.222.10.1 prefix-list Customer in
 neighbor 222.222.10.1 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

- **Router C only allows in /19 and /20 prefixes from customer block**

Two links to the same ISP

- **Router D Configuration**

```
router bgp 109
 neighbor 222.222.10.5 remote-as 65534
 neighbor 222.222.10.5 default-originate
 neighbor 222.222.10.5 prefix-list Customer in
 neighbor 222.222.10.5 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

- **Router D only allows in /19 and /20 prefixes from customer block**

Two links to the same ISP

- **Router E is AS109 border router**

  removes subprefixes in the private AS from external announcements

  removes the private AS from external announcement of the customer /19

Two links to the same ISP (with redundancy)

- **Router E Configuration**

```
router bgp 109
 neighbor 222.222.10.17 remote-as 110
 neighbor 222.222.10.17 remove-private-AS
 neighbor 222.222.10.17 prefix-list Customer out
!
ip prefix-list Customer permit 221.10.0.0/19
```

- **Private AS still visible inside AS109**

Loadsharing to the same ISP

- Loadsharing configuration is only on customer router
- Upstream ISP has to

  remove customer subprefixes from external announcements

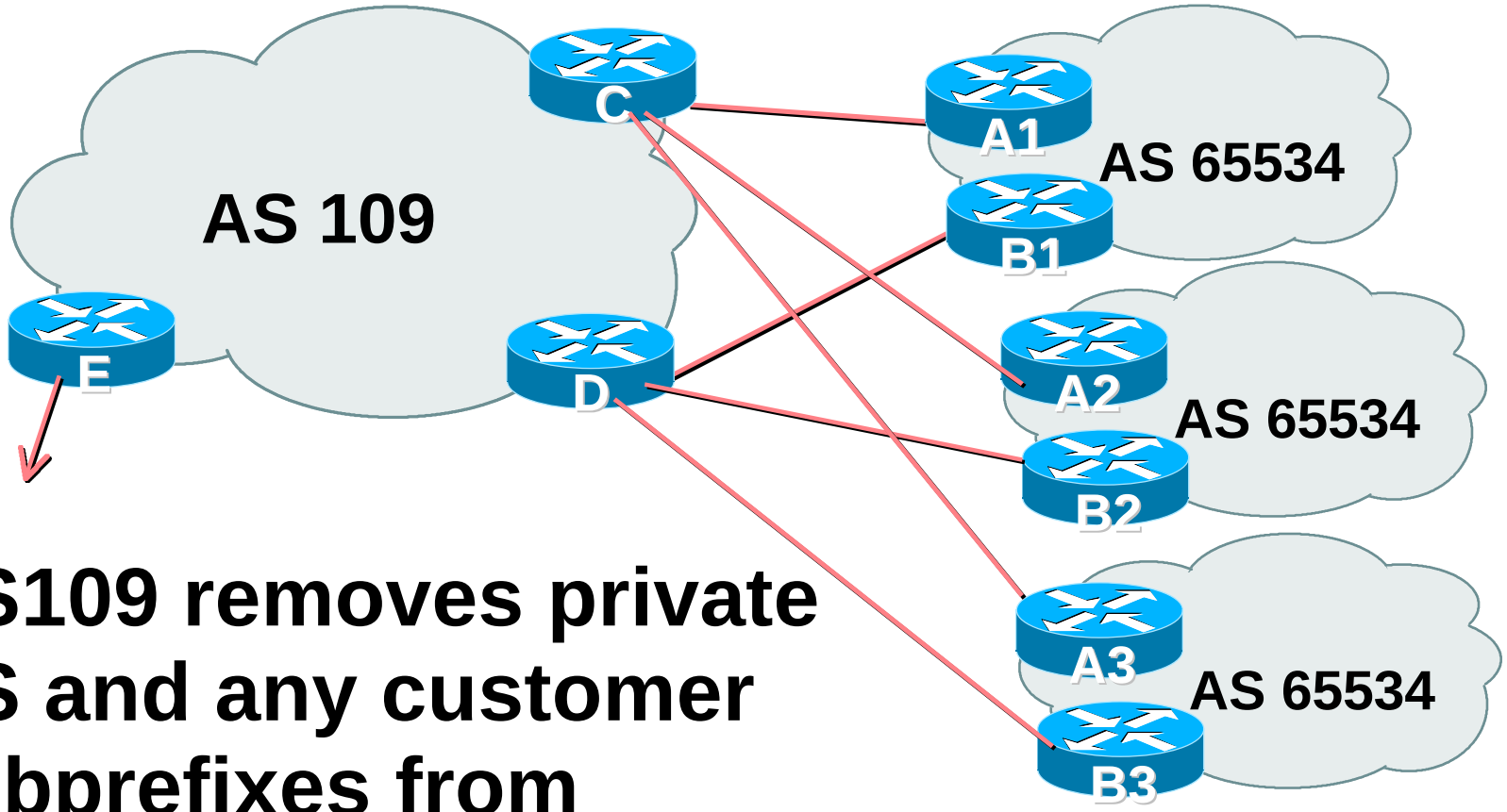
  remove private AS from external announcements
- Could also use BGP communities

Two links to the same ISP

Multiple Dualhomed Customers

(RFC2270)

Multiple Dualhomed Customers (RFC2270)



- AS109 removes private AS and any customer subprefixes from Internet announcement

Multiple Dualhomed Customers

- Customer announcements as per previous example
- Use the *same* private AS for each customer
  - documented in RFC2270
  - address space is not overlapping
  - each customer hears default only
- Router A*n* and B*n* configuration same as Router A and B previously

Two links to the same ISP

- **Router A1 Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.0.0 mask 255.255.240.0
 neighbor 222.222.10.2 remote-as 109
 neighbor 222.222.10.2 prefix-list routerC out
 neighbor 222.222.10.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerC permit 221.10.0.0/20
ip prefix-list routerC permit 221.10.0.0/19
!
ip route 221.10.0.0 255.255.240.0 null0
ip route 221.10.0.0 255.255.224.0 null0
```

Two links to the same ISP

- **Router B1 Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.16.0 mask 255.255.240.0
 neighbor 222.222.10.6 remote-as 109
 neighbor 222.222.10.6 prefix-list routerD out
 neighbor 222.222.10.6 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerD permit 221.10.16.0/20
ip prefix-list routerD permit 221.10.0.0/19
!
ip route 221.10.0.0 255.255.224.0 null0
ip route 221.10.16.0 255.255.240.0 null0
```

Multiple Dualhomed Customers

- **Router C Configuration**

```
router bgp 109
 neighbor bgp-customers peer-group
 neighbor bgp-customers remote-as 65534
 neighbor bgp-customers default-originate
 neighbor bgp-customers prefix-list default out
 neighbor 222.222.10.1 peer-group bgp-customers
 neighbor 222.222.10.1 description Customer One
 neighbor 222.222.10.1 prefix-list Customer1 in
 neighbor 222.222.10.9 peer-group bgp-customers
 neighbor 222.222.10.9 description Customer Two
 neighbor 222.222.10.9 prefix-list Customer2 in
```

Multiple Dualhomed Customers

```
 neighbor 222.222.10.17 peer-group bgp-customers
 neighbor 222.222.10.17 description Customer Three
 neighbor 222.222.10.17 prefix-list Customer3 in
!
ip prefix-list Customer1 permit 221.10.0.0/19 le 20
ip prefix-list Customer2 permit 221.16.64.0/19 le 20
ip prefix-list Customer3 permit 221.14.192.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

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

Multiple Dualhomed Customers

- **Router D Configuration**

```
router bgp 109
 neighbor bgp-customers peer-group
 neighbor bgp-customers remote-as 65534
 neighbor bgp-customers default-originate
 neighbor bgp-customers prefix-list default out
 neighbor 222.222.10.5 peer-group bgp-customers
 neighbor 222.222.10.5 description Customer One
 neighbor 222.222.10.5 prefix-list Customer1 in
 neighbor 222.222.10.13 peer-group bgp-customers
 neighbor 222.222.10.13 description Customer Two
 neighbor 222.222.10.13 prefix-list Customer2 in
```

Multiple Dualhomed Customers

```
 neighbor 222.222.10.21 peer-group bgp-customers
 neighbor 222.222.10.21 description Customer Three
 neighbor 222.222.10.21 prefix-list Customer3 in
!
ip prefix-list Customer1 permit 221.10.0.0/19 le 20
ip prefix-list Customer2 permit 221.16.64.0/19 le 20
ip prefix-list Customer3 permit 221.14.192.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

- **Router D only allows in /19 and /20 prefixes from customer block**

Multiple Dualhomed Customers

- **Router E Configuration is as previously**

  assumes customer address space is not part of upstream's address block

```
router bgp 109
 neighbor 222.222.10.17 remote-as 110
 neighbor 222.222.10.17 remove-private-AS
 neighbor 222.222.10.17 prefix-list Customers out
!
ip prefix-list Customers permit 221.10.0.0/19
ip prefix-list Customers permit 221.16.64.0/19
ip prefix-list Customers permit 221.14.192.0/19
```

- **Private AS still visible inside AS109**

Multiple Dualhomed Customers

- **If customers' prefixes come from ISP's address block**

  do NOT announce them to the Internet

  announce ISP aggregate only

- **Router E configuration:**

```
router bgp 109
 neighbor 222.222.10.17 remote-as 110
 neighbor 222.222.10.17 prefix-list my-aggregate out
!
ip prefix-list my-aggregate permit 221.8.0.0/13
```

Multihoming Summary

- Use private AS for multihoming to upstream
- Leak subprefixes to upstream only to aid loadsharing
- Upstream router E configuration is uniform across all scenarios

Two links to different ISPs

Two links to different ISPs

- **Use Public ASes**

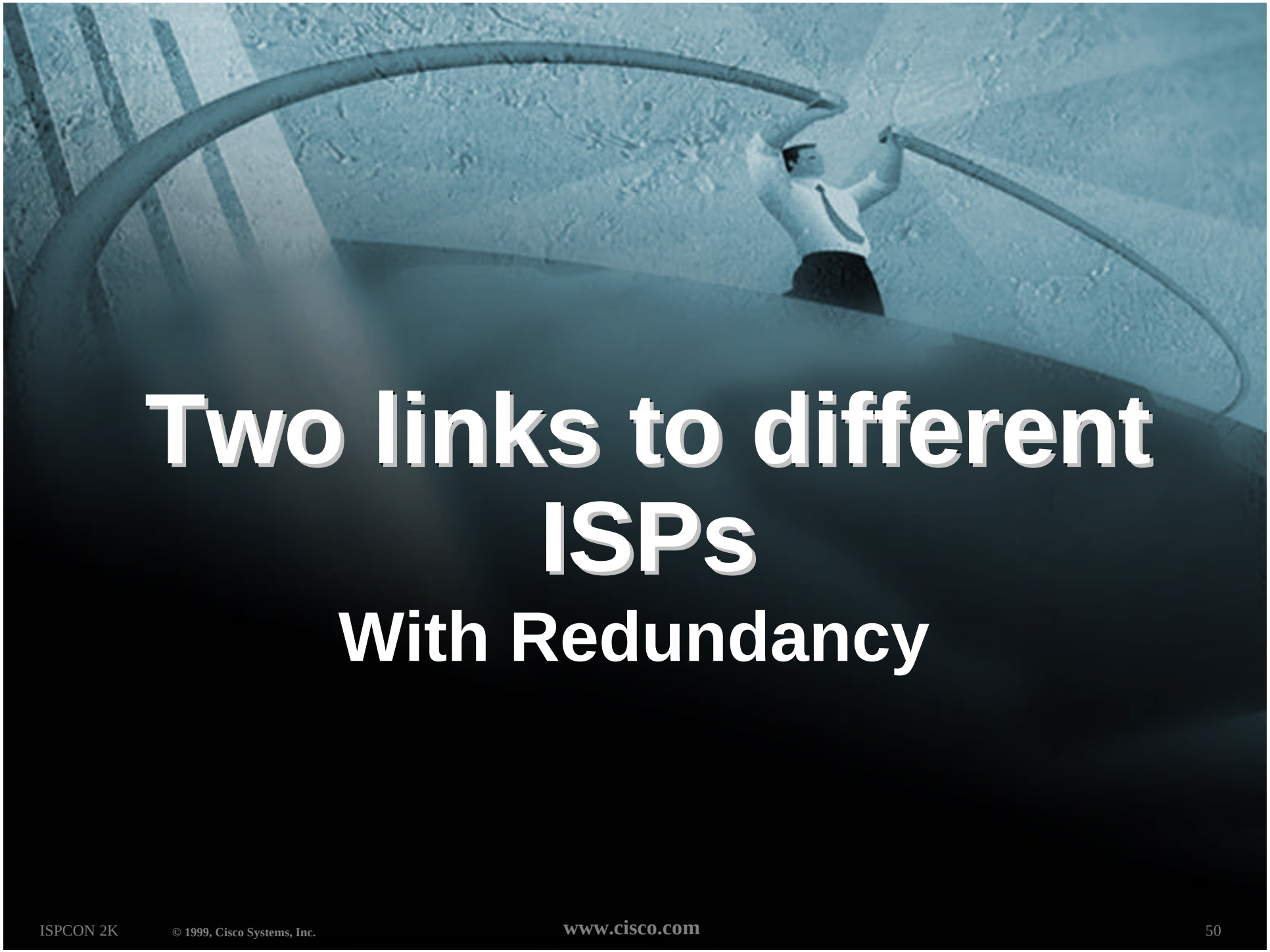
  or use private AS if agreed with the other ISP

- **Address space comes from**

  both upstreams or

  Regional Internet Registry

- **Configuration concepts very similar**



Two links to different ISPs

With Redundancy

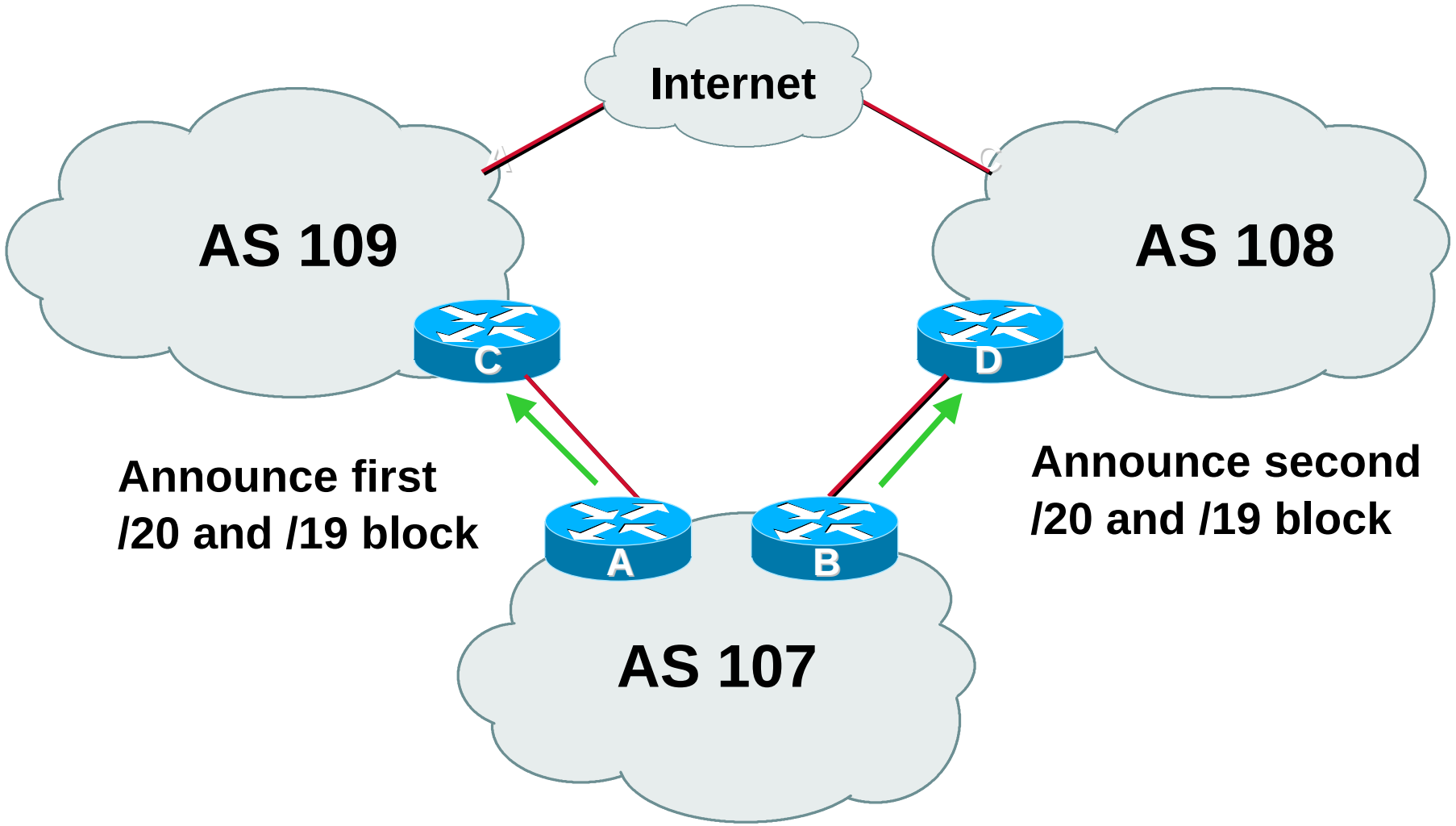
Two links to different ISPs (with redundancy)

- Example: ISP network, or large enterprise site
- 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 ISPs (with redundancy)



Two links to different ISPs (with redundancy)

- **Router A Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.0.0 mask 255.255.240.0
 neighbor 222.222.10.1 remote-as 109
 neighbor 222.222.10.1 prefix-list firstblock out
 neighbor 222.222.10.1 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
!
ip prefix-list firstblock permit 221.10.0.0/20
ip prefix-list firstblock permit 221.10.0.0/19
```

Two links to different ISPs (with redundancy)

- **Router B Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.16.0 mask 255.255.240.0
 neighbor 220.1.5.1 remote-as 108
 neighbor 220.1.5.1 prefix-list secondblock out
 neighbor 220.1.5.1 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
!
ip prefix-list secondblock permit 221.10.16.0/20
ip prefix-list secondblock permit 221.10.0.0/19
```

Two links to different ISPs (with redundancy)

- **Router C Configuration**

```
router bgp 109
 neighbor 221.10.1.1 remote-as 107
 neighbor 221.10.1.1 default-originate
 neighbor 221.10.1.1 prefix-list AS107cust in
 neighbor 221.10.1.1 prefix-list default-out out
!
```

- **Router C only announces default to AS 107**
- **Router C only accepts AS107's prefix block**

Two links to different ISPs (with redundancy)

- **Router D Configuration**

```
router bgp 108
 neighbor 220.1.5.1 remote-as 107
 neighbor 220.1.5.1 default-originate
 neighbor 220.1.5.1 prefix-list AS107cust in
 neighbor 220.1.5.1 prefix-list default-out out
!
```

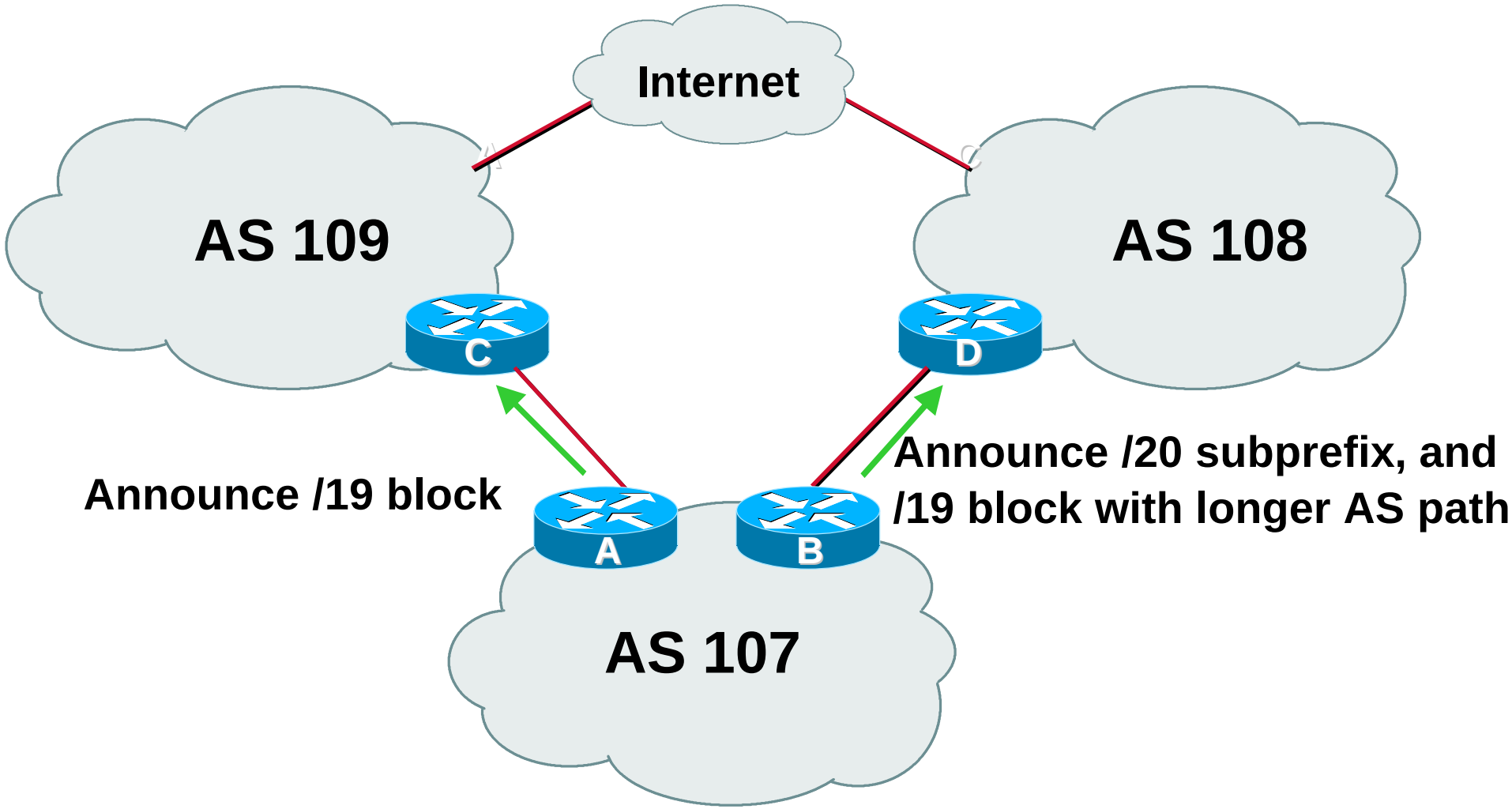
- **Router D only announces default to AS 107**
- **Router D only accepts AS107's prefix block**


Two links to different ISPs
More Controlled Loadsharing

Loadsharing with different ISPs

- 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 ISPs



Loadsharing with different ISPs

- **Router A Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.1 remote-as 109
 neighbor 222.222.10.1 prefix-list default in
 neighbor 222.222.10.1 prefix-list aggregate out
!
ip prefix-list aggregate permit 221.10.0.0/19
```

Loadsharing with different ISPs

- **Router B Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 network 221.10.16.0 mask 255.255.240.0
 neighbor 220.1.5.1 remote-as 108
 neighbor 220.1.5.1 prefix-list default in
 neighbor 220.1.5.1 prefix-list subblocks out
 neighbor 220.1.5.1 route-map routerD out
!
..next slide..
```

Loadsharing with different ISPs

```
route-map routerD permit 10
 match ip address prefix-list aggregate
 set as-path prepend 107 107
route-map routerD permit 20
!
ip prefix-list subblocks permit 221.10.0.0/19 le 20
ip prefix-list aggregate permit 221.10.0.0/19
```

Enterprise Multihoming

Address Space from different ISPs

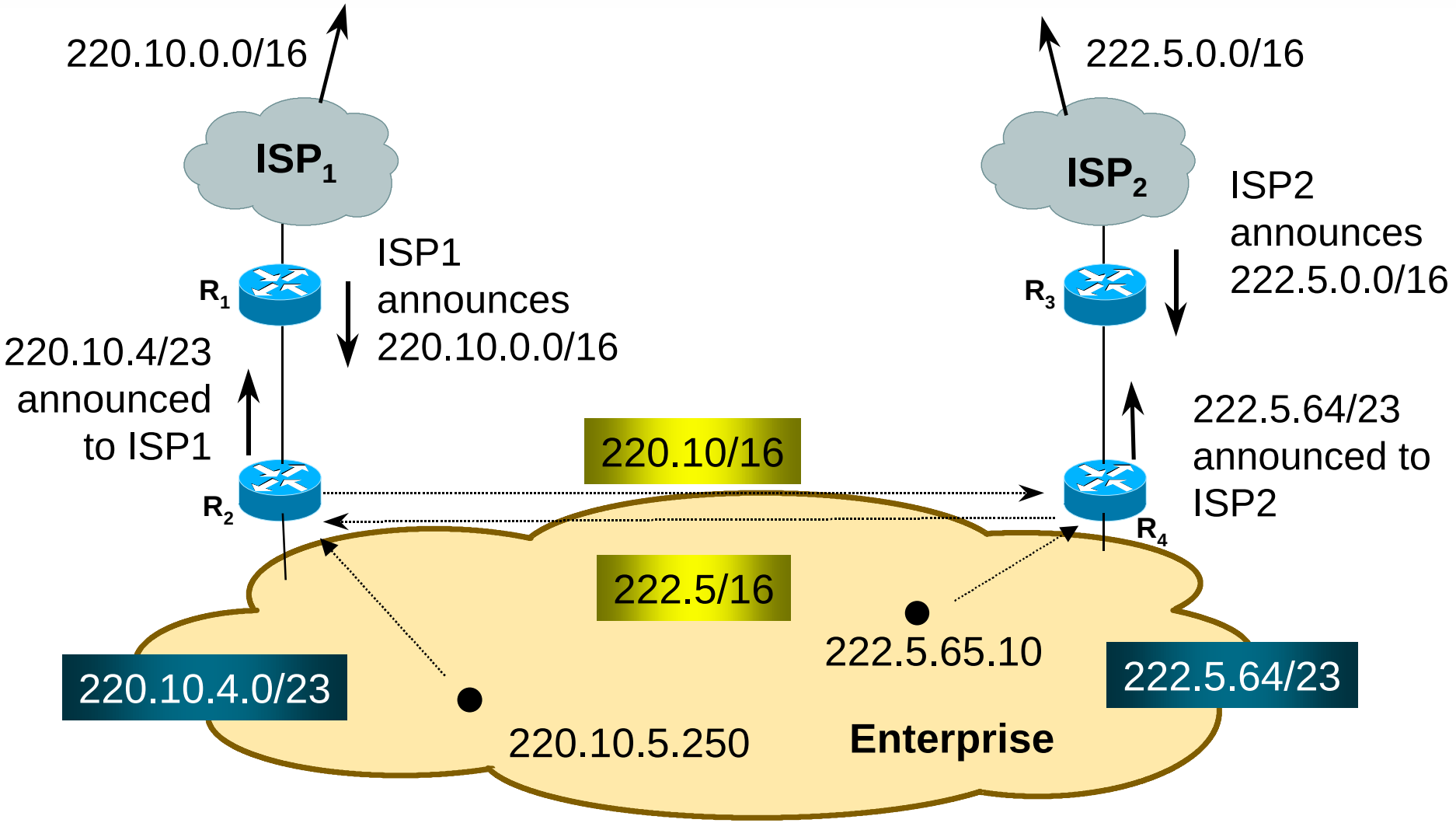
Enterprise Multihoming

- Common situation is enterprise multihoming
  - address space used by enterprise comes from both upstream ISPs
  - multihoming and loadsharing more difficult
  - want to avoid leaking subprefixes of upstream provider address space

Enterprise Multihoming

- **Conditional advertisement feature in IOS**
  - **loadsharing under normal conditions**
  - **subprefixes only announced in failure scenarios**
  - **requires upstreams to announce only one prefix to enterprise border network**

Steady State



Steady State

- ISP1 has 220.10.0.0/16 address block
- ISP2 has 222.5.0.0/16 address block
- Enterprise customer multihomes
  - upstreams don't announce subprefixes
  - can use private AS (ASN>64511)
  - R2 and R4 originate default in their IGP
    - outbound traffic uses nearest exit (IGP metrics)

Steady State

- **Router2 configuration:**

```
router bgp 65534
 network 220.10.4.0 mask 255.255.254.0
 network 222.5.64.0 mask 255.255.254.0
 neighbor <R1> remote-as 150
 neighbor <R1> prefix-list isp1-in in
 neighbor <R1> prefix-list isp1-out out
 neighbor <R1> advertise-map isp2-sb non-exist-map isp2-bb
 neighbor <R4> remote-as 65534
 neighbor <R4> update-source loopback 0
!
ip route 220.10.4.0 255.255.254.0 null0 250
..next slide
```

Steady State

```
ip route 222.5.64.0 255.255.254.0 null0 250
!
ip prefix-list isp1-out permit 220.10.4.0/23
ip prefix-list isp2-out permit 222.5.64.0/23
!
ip prefix-list isp1-in permit 220.10.0.0/16
ip prefix-list isp2-in permit 222.5.0.0/16
!
route-map isp2-sb permit 10
 match ip address prefix-list isp2-out
!
route-map isp2-bb permit 10
 match ip address prefix-list isp2-in
!
```

Steady State

- **Router2 peers iBGP with Router4**

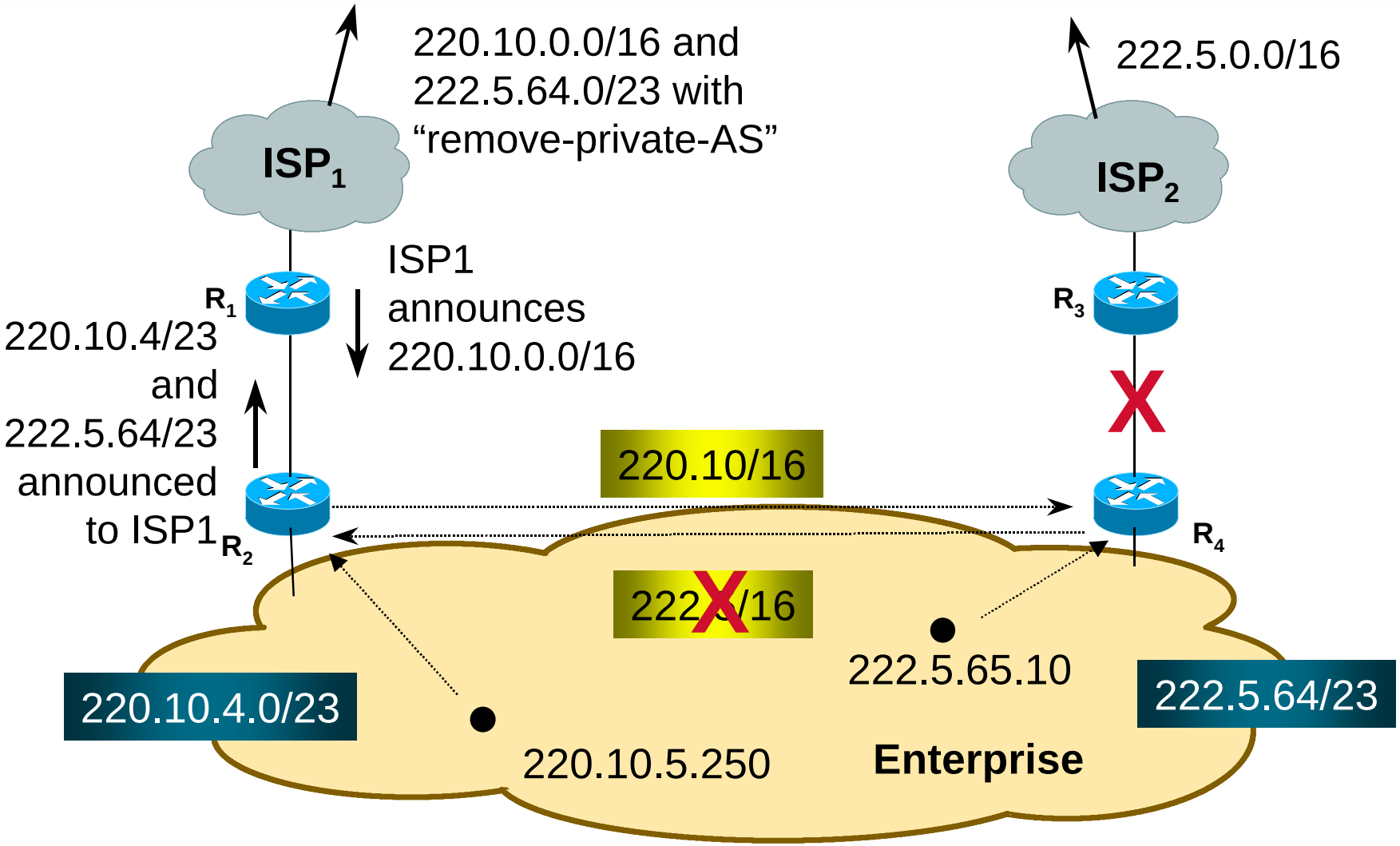
  hears ISP2's /16 prefix

- **Router2 peers eBGP with Router1**

  hears ISP1's /16 prefix only

  announces 220.10.4.0/23 only

Link Failure



Link Failure

- Peering between Router 4 and Router3 (ISP2) goes down

  222.5.0.0/16 prefix withdrawn

- Conditional advertisement process activated

  Router2 starts to announce 222.5.64.0/23 to Router1

- Connectivity for Enterprise maintained

Enterprise Multihoming

- Conditional advertisement useful when address space comes from both upstreams
  - no subprefixes leaked to Internet unless in failure situation
- Alternative backup mechanism would be to leak /23 prefixes with longer AS path
  - routing table bloat, reachability issues

Summary

- Private vs Public ASes
- Multihoming to Same ISP
- Multihoming to Different ISPs
- Community based multihoming
- Enterprise multihoming

Advanced Community usage

RFC1998

- Informational RFC
- Describes how to implement loadsharing and backup on multiple inter-AS links
  - BGP communities used to determine local preference in upstream's network
- Gives control to the customer
- Simplifies upstream's configuration
  - simplifies network operation!

RFC1998

- **Community values defined to have particular meanings:**

ASx:100	**set local pref 100**	**preferred route**
ASx:90	**set local pref 90**	**backup route if dualhomed on ASx**
ASx:80	**set local pref 80**	**main link is to another ISP with same AS path length**
ASx:70	**set local pref 70**	**main link is to another ISP**

RFC1998

- **Sample Customer Router Configuration**

```
router bgp 107
 neighbor x.x.x.x remote-as 109
 neighbor x.x.x.x description Backup ISP
 neighbor x.x.x.x route-map config-community out
 neighbor x.x.x.x send-community
!
ip as-path access-list 20 permit ^$
ip as-path access-list 20 deny .*
!
route-map config-community permit 10
 match as-path 20
 set community 109:90
```

RFC1998

- **Sample ISP Router Configuration**

```
! Homed to another ISP
ip community-list 70 permit 109:70
! Homed to another ISP with equal ASPATH length
ip community-list 80 permit 109:80
! Customer backup routes
ip community-list 90 permit 109:90
!
route-map set-customer-local-pref permit 10
 match community 70
 set local-preference 70
```

RFC1998

- **Sample ISP Router Configuration**

```
route-map set-customer-local-pref permit 20
 match community 80
 set local-preference 80
!
route-map set-customer-local-pref permit 30
 match community 90
 set local-preference 90
!
route-map set-customer-local-pref permit 40
 set local-preference 100
```

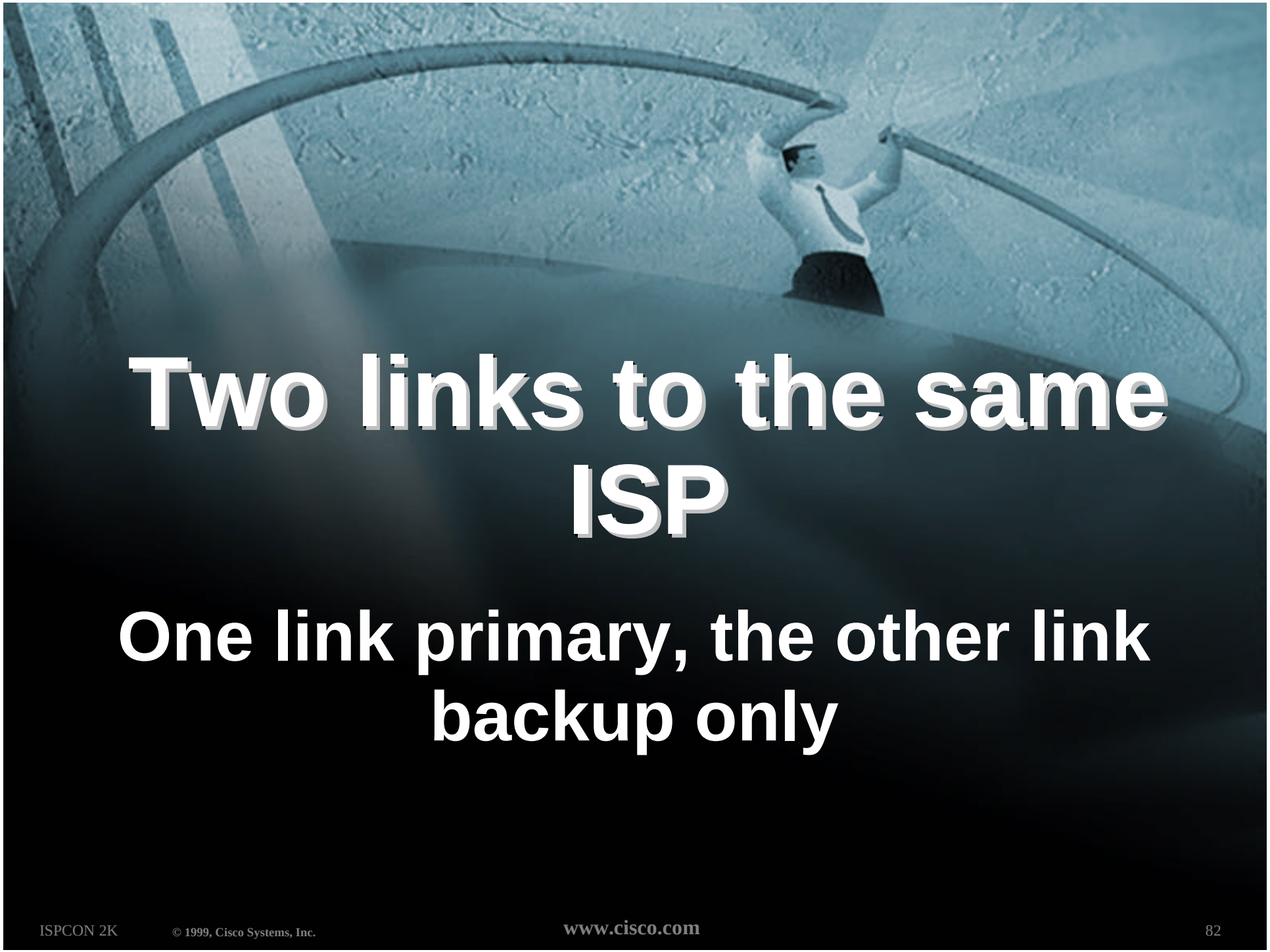
RFC1998

- **Supporting RFC1998**

  many ISPs do, more should

  check AS object in the Internet Routing Registry

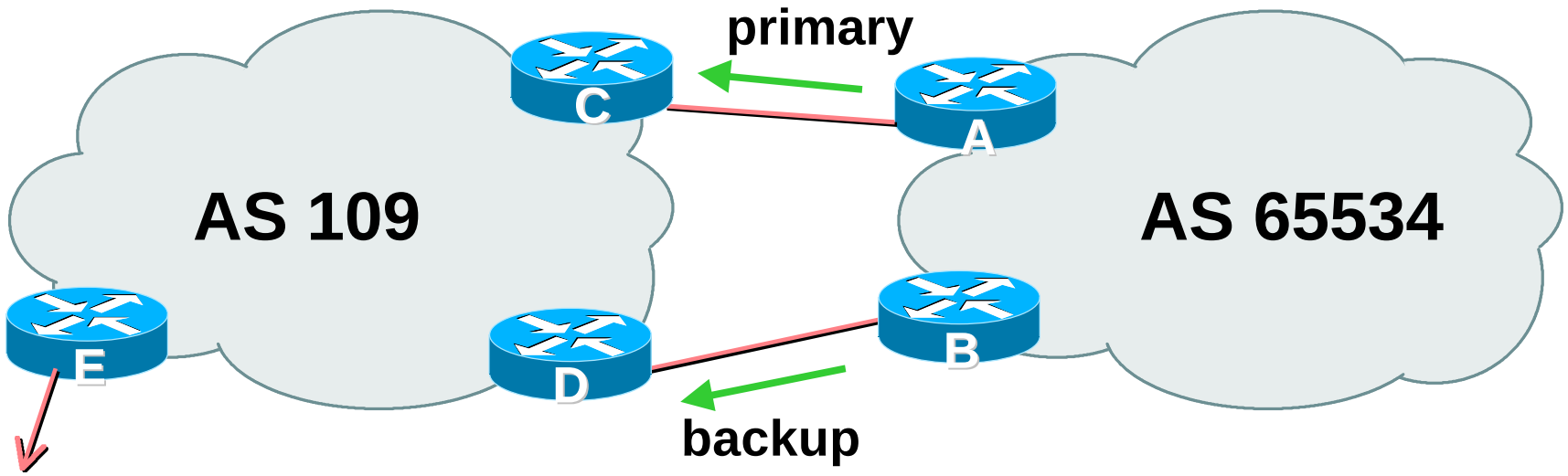
  if you do, insert comment in AS object in the IRR



Two links to the same ISP

One link primary, the other link backup only

Two links to the same ISP



- **AS109 proxy aggregates for AS 65534**

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

- Announce /19 aggregate on each link

  primary link makes standard announcement

  backup link sends community

- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

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

- **Router A Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.2 remote-as 109
 neighbor 222.222.10.2 description RouterC
 neighbor 222.222.10.2 prefix-list aggregate out
 neighbor 222.222.10.2 prefix-list default in
!
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
```

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

- **Router B Configuration**

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.6 remote-as 109
 neighbor 222.222.10.6 description RouterD
 neighbor 222.222.10.6 send-community
 neighbor 222.222.10.6 prefix-list aggregate out
 neighbor 222.222.10.6 route-map routerD-out out
 neighbor 222.222.10.6 prefix-list default in
 neighbor 222.222.10.6 route-map routerD-in in
!
..next slide
```

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

```
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
 match ip address prefix-list aggregate
 set community 109:90
route-map routerD-out permit 20
!
route-map routerD-in permit 10
 set local-preference 90
!
```

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

- **Router C Configuration (main link)**

```
router bgp 109
 neighbor 222.222.10.1 remote-as 65534
 neighbor 222.222.10.1 default-originate
 neighbor 222.222.10.1 prefix-list Customer in
 neighbor 222.222.10.1 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

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

- **Router D Configuration (backup link)**

```
router bgp 109
 neighbor 222.222.10.5 remote-as 65534
 neighbor 222.222.10.5 default-originate
 neighbor 222.222.10.5 prefix-list Customer in
 neighbor 222.222.10.5 route-map bgp-cust-in in
 neighbor 222.222.10.5 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
..next slide
```

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

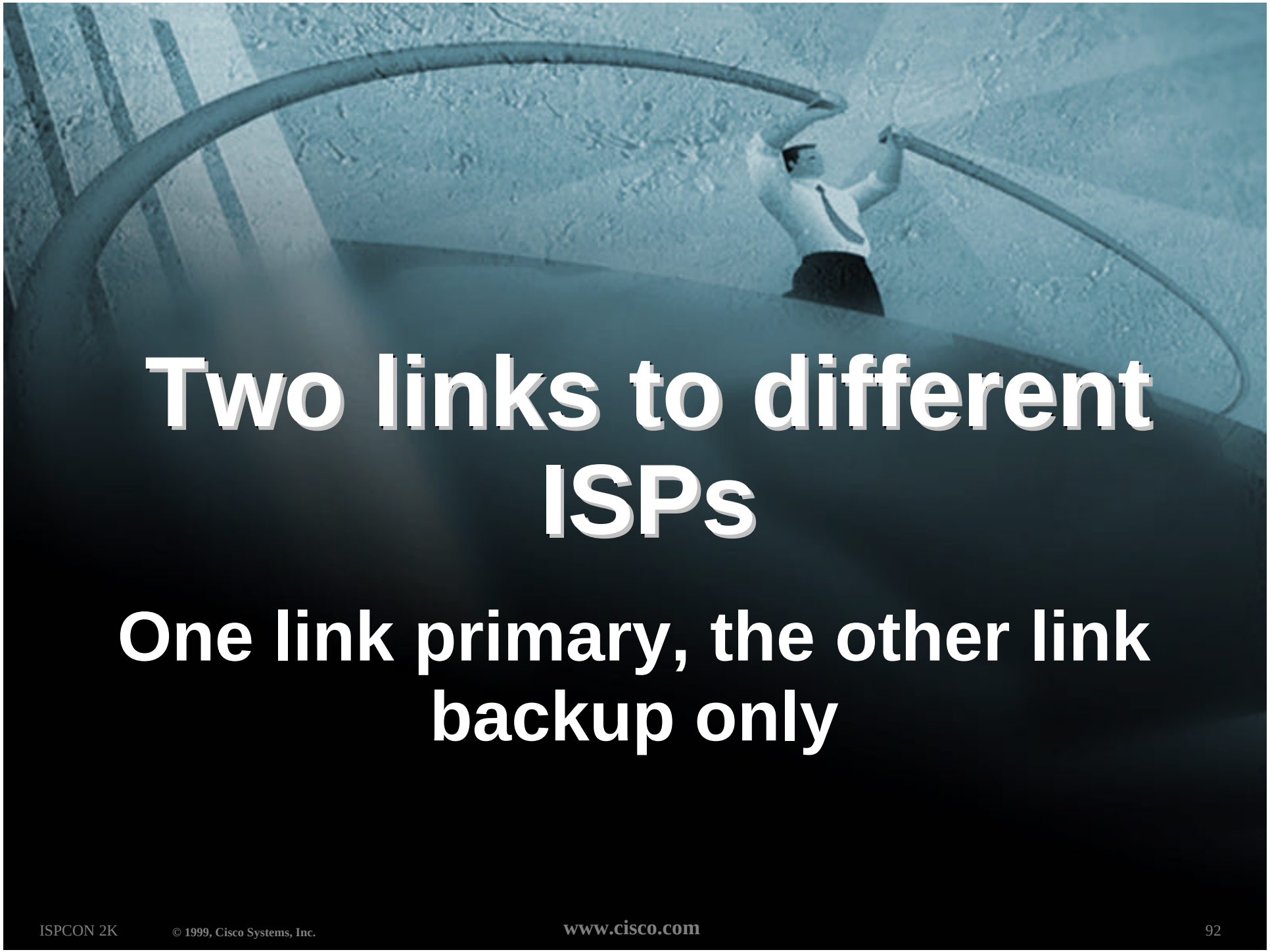
```
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip community-list 90 permit 109:90
!
<snip>
route-map bgp-cust-in permit 30
 match community 90
 set local-preference 90
route-map bgp-cust-in permit 40
 set local-preference 100
```

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

- **Router E Configuration**

```
router bgp 109
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.17 remote-as 110
 neighbor 222.222.10.17 filter-list 1 out
!
ip as-path access-list 1 deny ^(65534_)+$
ip as-path access-list 1 permit ^$
ip route 221.10.0.0 255.255.224.0 null0
```

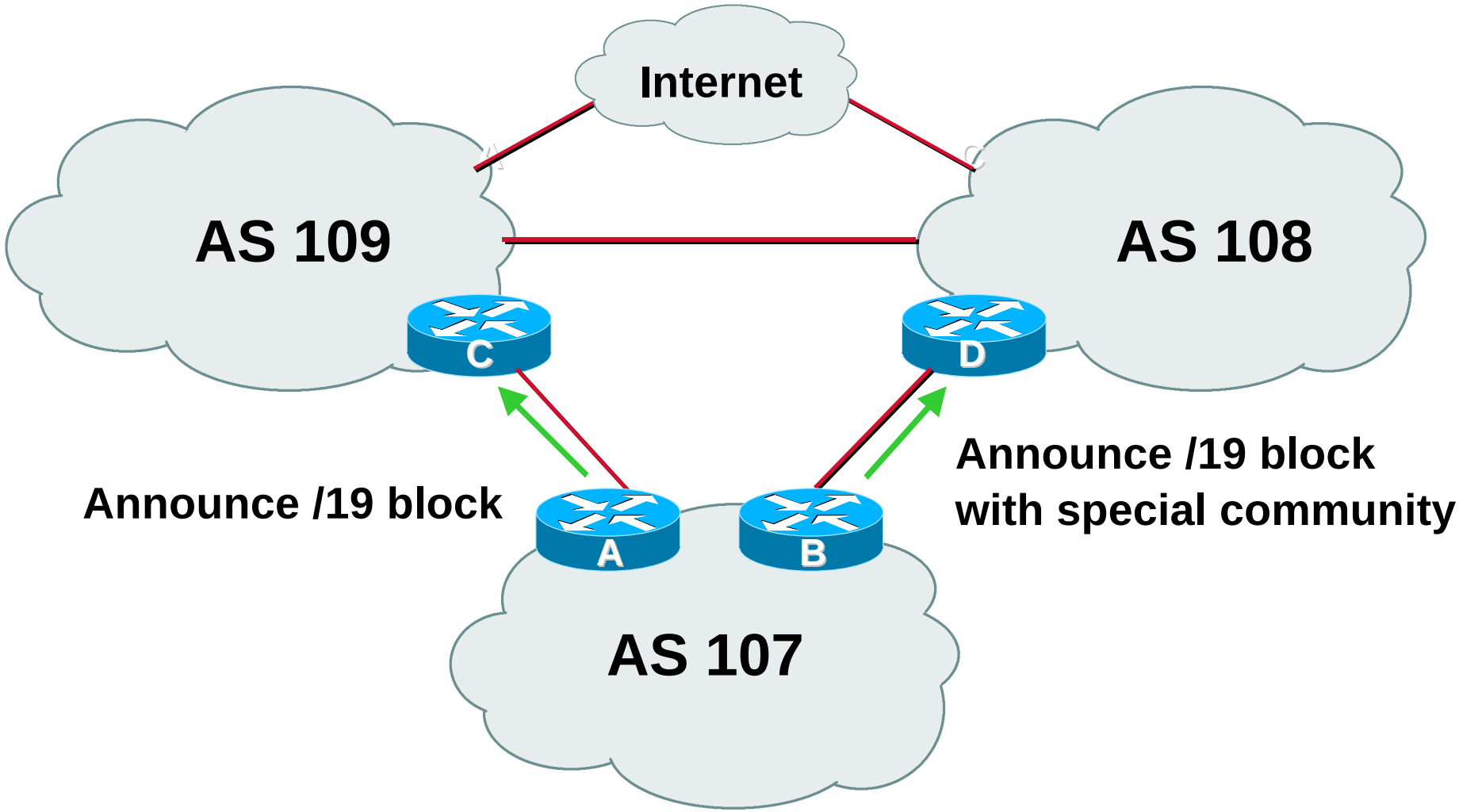
- **Router E removes prefixes in the private AS from external announcements**
- **Private AS still visible inside AS109**



Two links to different ISPs

One link primary, the other link backup only

Two links to different ISPs (one as backup only)



Two links to different ISPs (one as backup only)

- Announce /19 aggregate on each link

  main link sends community 109:100 - this sets local pref in AS109 to 100

  backup link sends community 108:80 - this sets local pref in AS108 to 80

- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

Two links to different ISPs (one as backup only)

- Note that this assumes that AS109 and AS108 are interconnected
- If they are not, AS path length "stuffing" has to be used too

  but that can be done on a per community basis also

  custom additions to RFC1998

Two links to different ISPs (one as backup only)

- **Router A Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.1 remote-as 109
 neighbor 222.222.10.1 prefix-list aggregate out
 neighbor 222.222.10.1 route-map routerC-out out
 neighbor 222.222.10.1 prefix-list default in
!
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerC-out permit 10
 set community 109:100
```

Two links to different ISPs (one as backup only)

- **Router B Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 220.1.5.1 remote-as 108
 neighbor 220.1.5.1 prefix-list aggregate out
 neighbor 220.1.5.1 route-map routerD-out out
 neighbor 220.1.5.1 prefix-list default in
 neighbor 220.1.5.1 route-map routerD-in in
..next slide
```

Two links to different ISPs (one as backup only)

```
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
 set community 108:80
!
route-map routerD-in permit 10
 set local-preference 80
```

Two links to different ISPs (one as backup only)

- **Router D**

  sees path from router B with community 108:80 set - sets local preference to 80

  sees path from peering with AS109 - default local preference is 100

  local-pref comes before AS Path length

  highest local-pref wins

  traffic for AS107 is sent to AS109

Two links to different ISPs (one as backup only)

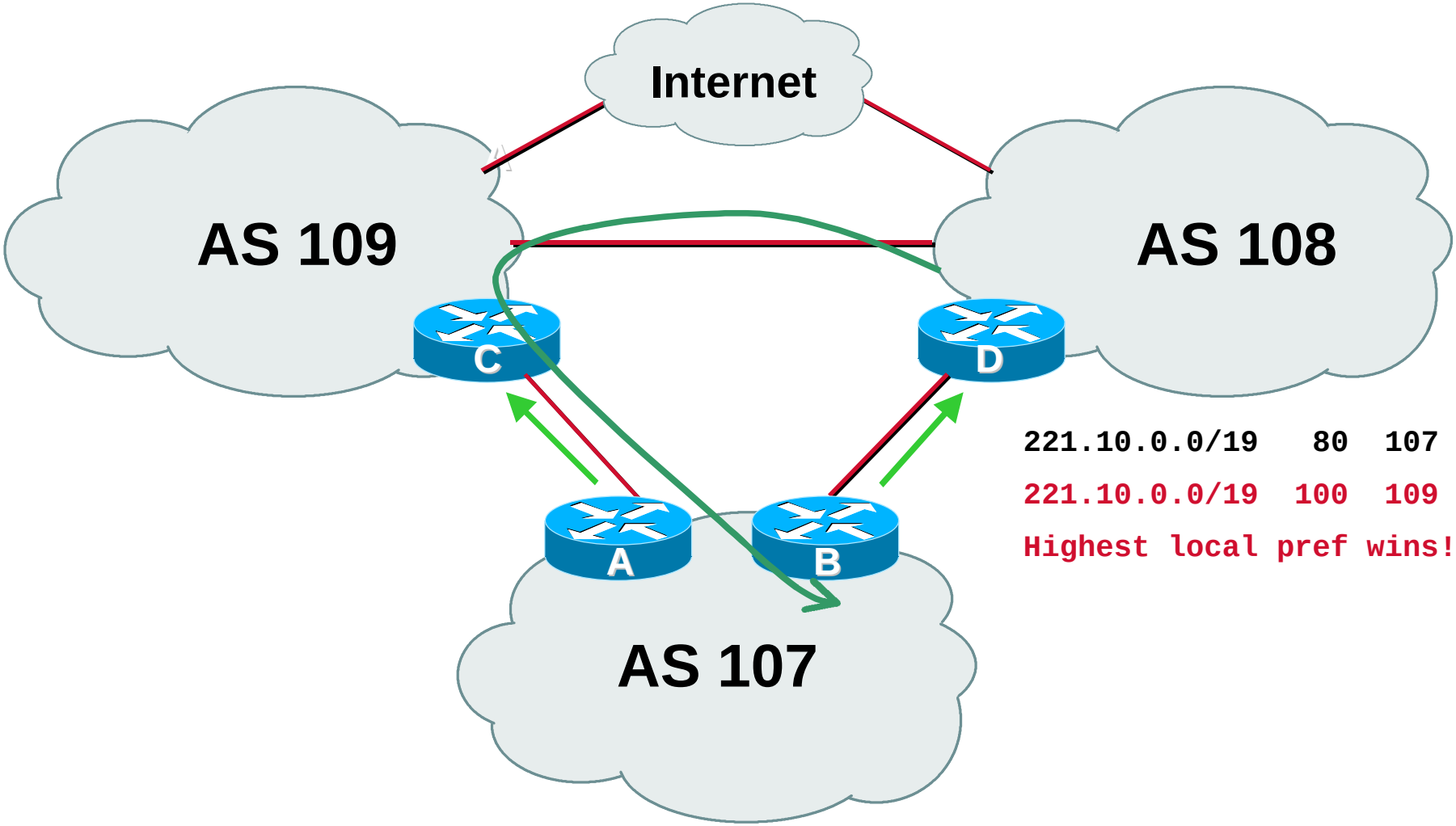
- **Router D**

  Only requires RFC1998 configuration

  no per customer configuration

  scalability!

Two links to different ISPs (one as backup only)




```
221.10.0.0/19    80  107
221.10.0.0/19   100  109 107
Highest local pref wins!
```

Two links to different ISPs (one as backup only)

- If AS107 wants to make the link to AS108 the main link

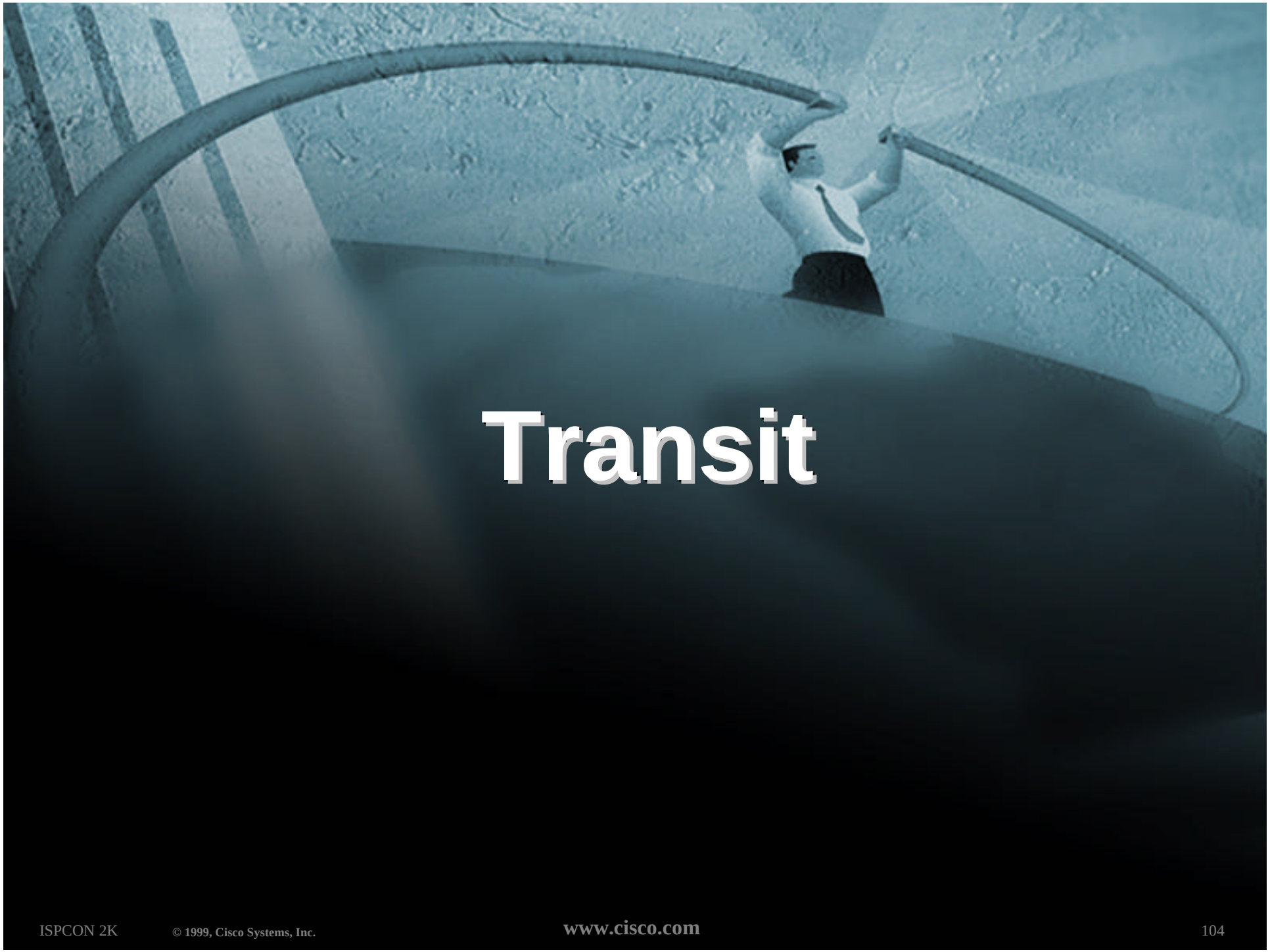
  sends community 108:100 to router C

  sends community 109:80 to router B

- AS108 and AS109 NOC intervention not required

Summary

- **Communities are fun! ☺**
- **And they are extremely powerful tools**
- **Think about community policies**
- **Supporting extensive community usage makes customer configuration easy**
- ***Watch out for routing loops!***



ISP Transit

- **Transit - carrying traffic across a network, usually for a fee**

  traffic and prefixes originating from one AS are carried across an intermediate AS to reach their destination AS

- **Only announce default to your BGP customers unless they need more prefixes**
- **Only accept the prefixes which your customer is entitled to originate**
- **If your customer hasn't told you he is providing transit, don't accept anything else**

ISP Transit Issues

Many mistakes are made on the Internet today due to incomplete understanding of how to configure BGP for transit

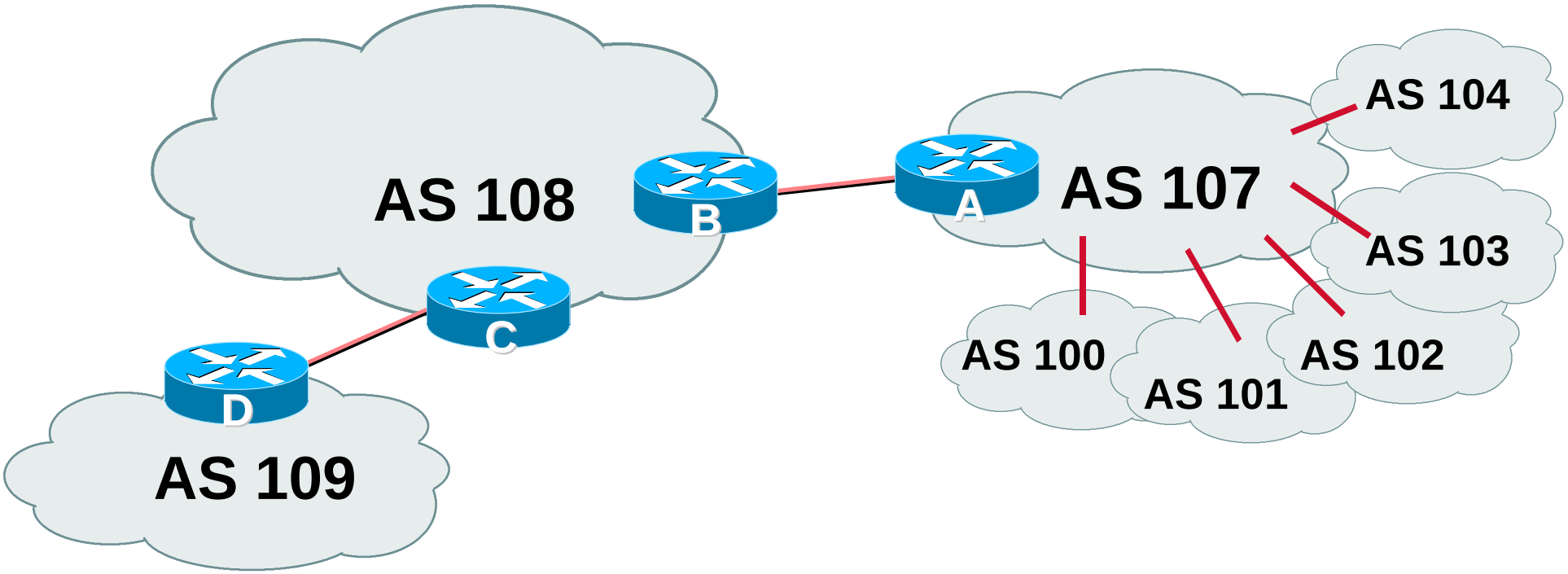
ISP Transit Provider

Configuration Example

ISP Transit

- AS107 and AS109 are stub/customer ASes of AS108
  - AS107 has many customers with their own ASes
    - AS104 doesn't get announced to AS108
  - AS108 provides transit between AS107 and AS109

ISP Transit



- **AS107 has several customer ASes connecting to its backbone**

ISP Transit

- **Router A Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.2 remote-as 108
 neighbor 222.222.10.2 prefix-list upstream-out out
 neighbor 222.222.10.2 filter-list 5 out
 neighbor 222.222.10.2 prefix-list upstream-in in
!
ip route 221.10.0.0 255.255.224.0 null0 250
!
..next slide
```

ISP Transit

```
!
! As-path filters..
ip as-path access-list 5 permit ^$
ip as-path access-list 5 permit ^(100_)+$
ip as-path access-list 5 permit ^101$
ip as-path access-list 5 permit ^102$
ip as-path access-list 5 permit ^103$
ip as-path access-list 5 deny ^104_
ip as-path access-list 5 permit ^105$
! Etc...
!
..next slide
```

ISP Transit

- **Prefix-list upstream-out blocks all Martian prefixes to the upstream**

  also blocks subprefixes from AS107

- **Prefix-list upstream-in blocks all Martian prefixes from the upstream**

  also blocks AS107 prefix and subprefixes

ISP Transit

- **Router B Configuration**

```
router bgp 108
 neighbor 222.222.10.1 remote-as 107
 neighbor 222.222.10.1 prefix-list rfc1918-dsua in
 neighbor 222.222.10.1 prefix-list rfc1918-dsua out
 neighbor 222.222.10.1 filter-list 10 in
 neighbor 222.222.10.1 filter-list 15 out
!
ip as-path access-list 15 permit ^$
ip as-path access-list 15 permit ^109$
```

Router B announces AS108 and AS109 prefixes to Router A, and accepts all AS107 customer ASes

ISP Transit

- **Router C Configuration**

```
router bgp 108
 neighbor 222.222.20.1 remote-as 109
 neighbor 222.222.20.1 default-originate
 neighbor 222.222.20.1 prefix-list Customer109 in
 neighbor 222.222.20.1 prefix-list default out
!
ip prefix-list Customer109 permit 219.0.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

- **Router C announces default to Router D, only accepts customer /19**

ISP Transit

- **Router D Configuration**

```
router bgp 109
 network 219.0.0.0 mask 255.255.224.0
 neighbor 222.222.20.2 remote-as 108
 neighbor 222.222.20.2 prefix-list upstream out
 neighbor 222.222.20.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list upstream permit 219.0.0.0/19
!
ip route 219.0.0.0 255.255.224.0 null0
```

ISP Transit

- **AS107 only hears AS108 and AS109 prefixes**

  inbound AS path filter on Router A is optional, but good practice (never trust a peer)

  DSUA prefix-list filters are mandatory on all Internet peerings

Collocation Practises

How and Why to place equipment overseas

Why Collocate Overseas?

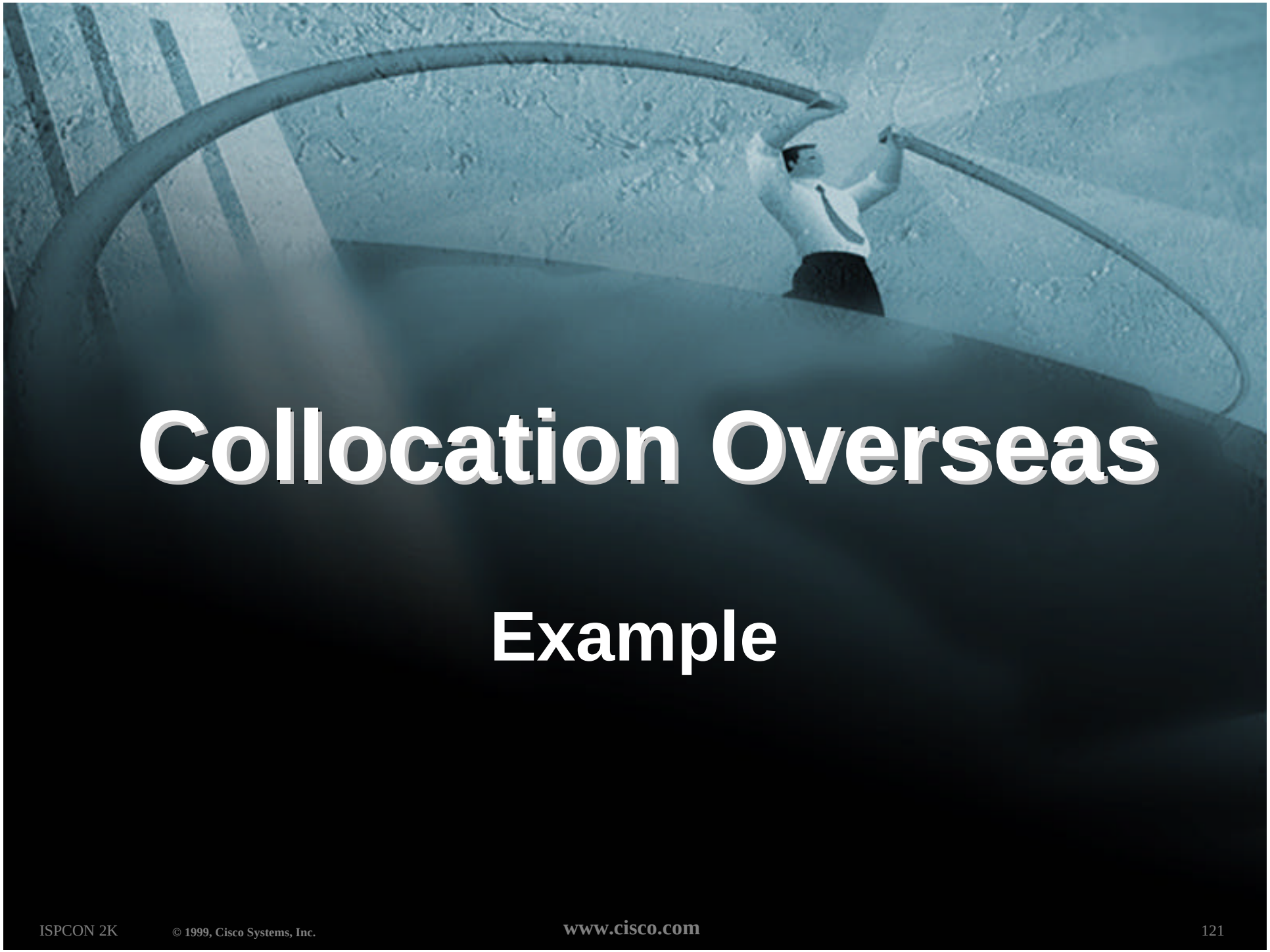
- **Hard to re-terminate transoceanic circuit in case of "issues" with upstream ISP**
- **No Quality of Service**
- **No Control over infrastructure**
- **No Monitoring of link performance**

Collocation Overseas

- Many AP region ISPs collocate equipment in the US
  - install their own router(s) and other hardware (servers, caches,…)
  - establish peering relationships with US NSPs and domestic ISPs
  - buy facilities management services
    - usually hardware maintenance, installation management

Collocation Overseas

- **Many AP region ISPs collocate equipment in the US**
  - **US domestic circuits are “cheap”**
  - **Easy to change your upstream**
  - **Easy to have multiple upstreams**
  - **Easy to implement QoS related features, service differentiation, etc...**



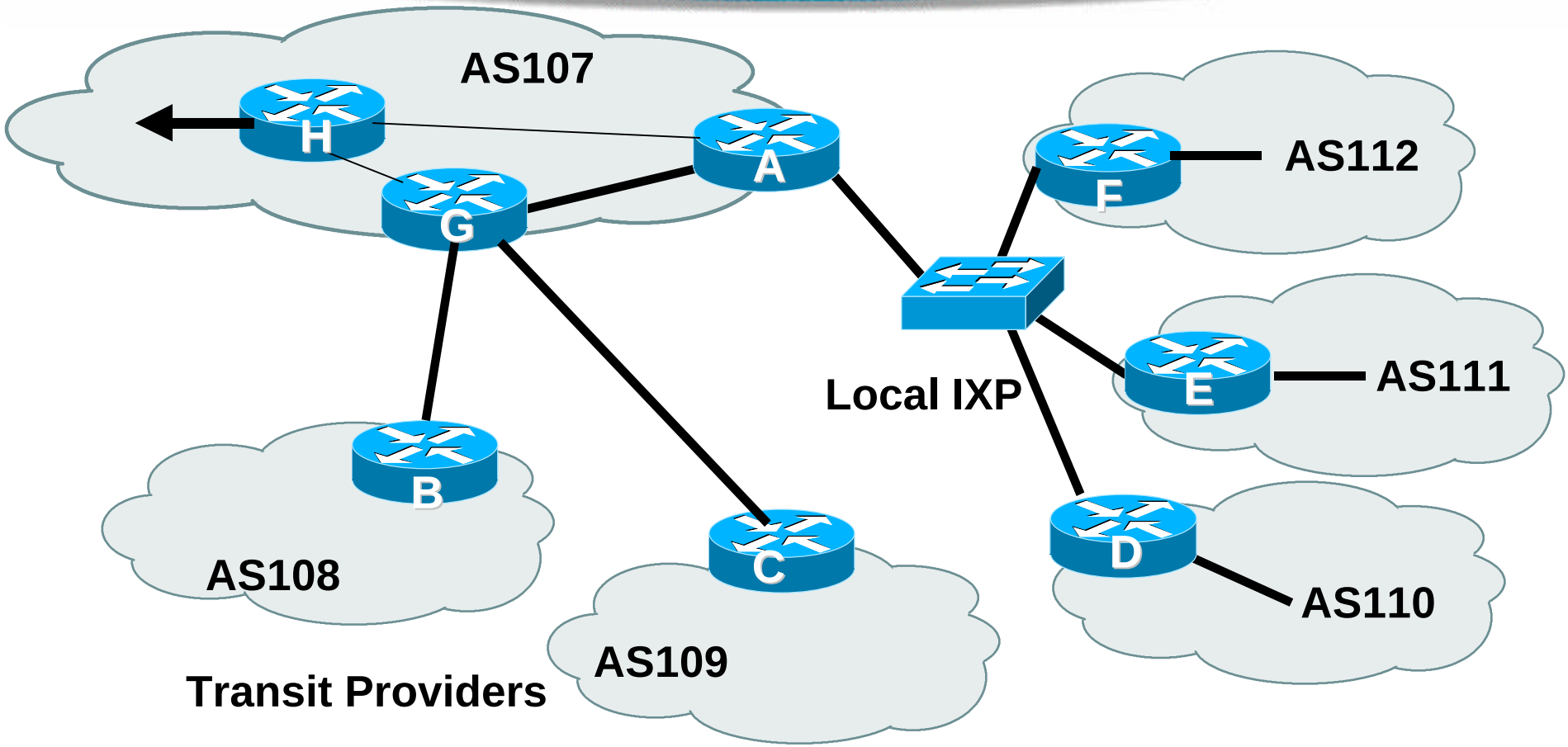
Collocation Overseas

Example

Collocation

- **Common Scenario:**
  - **AS107 has collocate space in the US**
  - **AS108 and AS109 are transit providers for AS107**
  - **AS107 is also present at the local exchange point for regional peers**

Collocation



Collocation

- **AS107**

  Router A is dedicated to peering at local IXP

  Router G is dedicated to links with the transit providers

  Router H is dedicated to the transoceanic link

Collocation Router A Configuration

```
router bgp 107
 neighbor ixp-peers peer-group
 neighbor ixp-peers soft-reconfiguration in
 neighbor ixp-peers prefix-list myprefixes out
 neighbor 221.0.0.2 remote-as 107
 neighbor 221.0.0.2 description Router G - Upstream Peers
 neighbor 221.0.0.2 update-source loopback 0
 neighbor 221.0.0.3 remote-as 107
 neighbor 221.0.0.3 description Router H - transpacific router
 neighbor 221.0.0.3 update-source loopback 0
 neighbor 221.0.0.4 remote-as 107
 neighbor 221.0.0.4 description Router at HQ
 neighbor 221.0.0.4 update-source loopback 0
..next slide
```

Collocation Router A Configuration

```
 neighbor 220.5.10.4 remote-as 110
 neighbor 222.5.10.4 peer-group ixp-peers
 neighbor 222.5.10.4 prefix-list peer110 in
 neighbor 220.5.10.5 remote-as 111
 neighbor 222.5.10.5 peer-group ixp-peers
 neighbor 222.5.10.5 prefix-list peer111 in
 neighbor 220.5.10.3 remote-as 112
 neighbor 222.5.10.3 peer-group ixp-peers
 neighbor 222.5.10.3 prefix-list peer112 in
!
ip prefix-list myprefixes permit 221.10.0.0/19
ip prefix-list peer110 permit 222.12.0.0/19
ip prefix-list peer111 permit 222.18.128.0/19
ip prefix-list peer112 permit 222.1.32.0/19
```

Collocation Router A Configuration

- **Router A does NOT originate AS107's prefix block**

  if router is disconnected from AS107 either locally or across the ocean, announcement could cause blackhole

- **Prefix-list filtering is the minimum required**

  usually include AS path filtering too

Collocation Router G Configuration

```
router bgp 107
 neighbor 221.0.0.1 remote-as 107
 neighbor 221.0.0.1 description Router A - US Local IXP
 neighbor 221.0.0.1 update-source loopback 0
 neighbor 221.0.0.1 prefix-list myprefixes out
 neighbor 221.0.0.3 remote-as 107
 neighbor 221.0.0.3 description Router H - transpacific router
 neighbor 221.0.0.3 update-source loopback 0
 neighbor 221.0.0.4 remote-as 107
 neighbor 221.0.0.4 description Router at HQ
 neighbor 221.0.0.4 update-source loopback 0
..next slide
```

Collocation Router G Configuration

```
 neighbor 222.0.0.1 remote-as 108
 neighbor 222.0.0.1 prefix-list myprefixes out
 neighbor 222.0.0.1 prefix-list rfc1918-dsua in
 neighbor 218.6.0.1 remote-as 109
 neighbor 218.6.0.1 prefix-list myprefixes out
 neighbor 218.6.0.1 prefix-list rfc1918-dsua in
!
ip prefix-list myprefixes permit 221.10.0.0/19
```

Collocation
Router G Configuration

- Router G accepts full BGP prefixes from both AS108 and AS109
- Router G announces AS107 prefix to upstreams
- Simple Example - policy may also be required for loadsharing etc

Collocation
Router H Configuration

```
router bgp 107
 neighbor 221.0.0.1 remote-as 107
 neighbor 221.0.0.1 description Router A - US Local IXP
 neighbor 221.0.0.1 update-source loopback 0
 neighbor 221.0.0.2 remote-as 107
 neighbor 221.0.0.2 description Router G - peering router
 neighbor 221.0.0.2 update-source loopback 0
 neighbor 221.0.0.4 remote-as 107
 neighbor 221.0.0.4 description Router at HQ
 neighbor 221.0.0.4 update-source loopback 0
!
```

Collocation Router H Configuration

- Router H is dedicated to transoceanic link

  part of ISP core iBGP mesh

- More complex configuration likely

  CAR, RED, etc

- More complex links likely

  e.g satellite uplink for low revenue latency insensitive traffic

Collocation Summary

- **Richer interconnectivity possible**
- **Better redundancy possible**
- **Overall advantage - control!**

Summary

- Multihoming Examples
- Advanced Community Usage
- Transit
- Overseas Collocation

- Any questions?