

BGP and the Internet

Enterprise Multihoming

Enterprise Multihoming

Cisco.com

- **Common scenario in Internet today**
- **More and more non-SPs multihoming for:**
 - service provider redundancy**
 - link redundancy**
- **Issues on Internet today:**
 - Routing Table size accelerating**
 - more and more /24 prefixes appearing in Internet Routing Table**
 - ASN consumption accelerating**

Enterprise Multihoming

Cisco.com

- **The following examples**
 - apply to smaller ISPs who don't yet have their own address block**
 - require BGP but a private AS (ASN >64511) can and should be used**
 - are good for the health of the Internet**

Medium/Large ISP Multihoming

Cisco.com

- **ISPs **should** obtain their own address block and ASN**

Get it from RIR

Makes multihoming easier

Makes changing upstreams easier

Does not cause so much fragmentation in Internet Routing Table

Enterprise Multihoming

Example One Provider Redundancy

Enterprise Multihoming

Cisco.com

- **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 when possible**
 - require provider redundancy (not just link redundancy)**

Enterprise Multihoming

Cisco.com

- **Address space from upstream should match link bandwidth to upstream, e.g.**
 - ISP1 ® Enterprise = 256kbps ® /22**
 - ISP2 ® Enterprise = 128kbps ® /23**
 - assumes address space is uniformly distributed across network**
 - assumes that there is a requirement for 3x /23 in the Enterprise backbone**
- **Next example assumes equal bandwidth links from Enterprise to ISP1 and ISP2**

Enterprise Multihoming

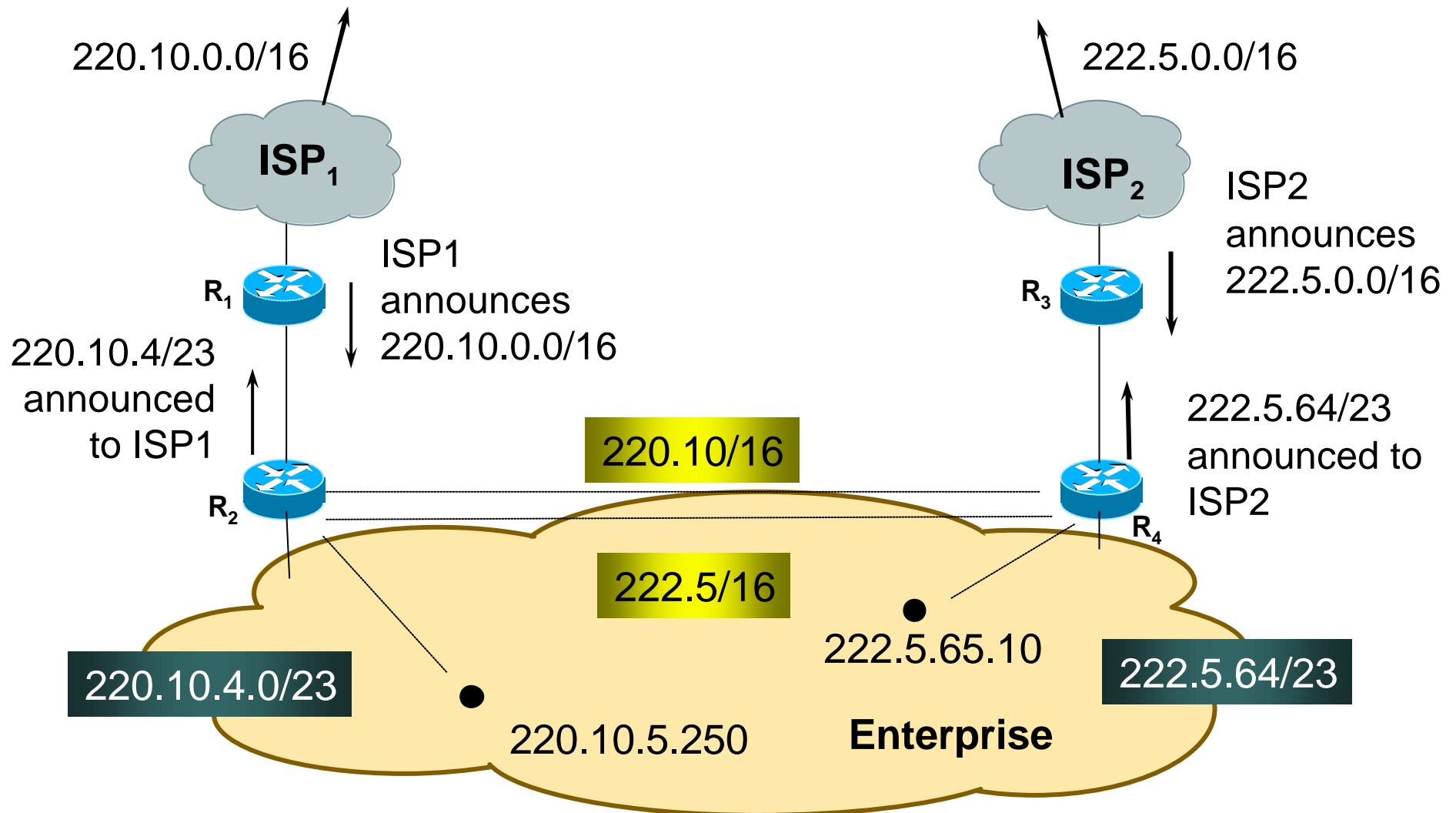
Conditional Advertisement

Cisco.com

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

Steady State

Cisco.com



Steady State

Cisco.com

- **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
!
```

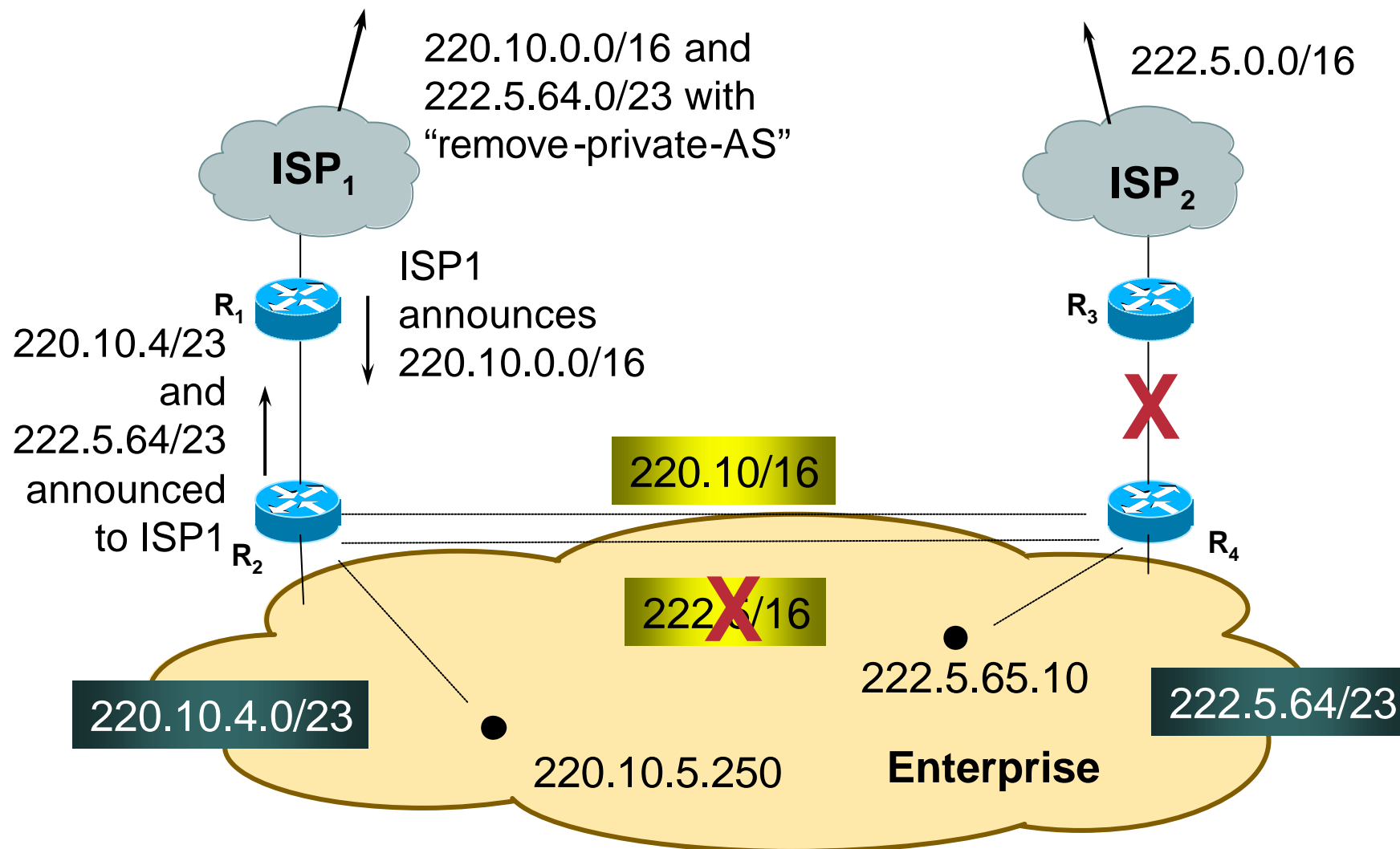
Steady State

Cisco.com

- **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

Cisco.com



Link Failure

Cisco.com

- **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

Cisco.com

- **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**

What goes in the Internet Routing Registry?

Cisco.com

- **ISP1 and ISP2 obviously put their own address blocks as route objects in the IRR**
- **ISP1 will put the ISP1 subprefix which Enterprise will announce into the IRR with origin-as of ISP2**
- **ISP2 will put the ISP2 subprefix which Enterprise will announce into the IRR with origin-as of ISP1**
- **No inconsistent origin AS, no “problem”**

Enterprise Multihoming

Example Two Link Redundancy

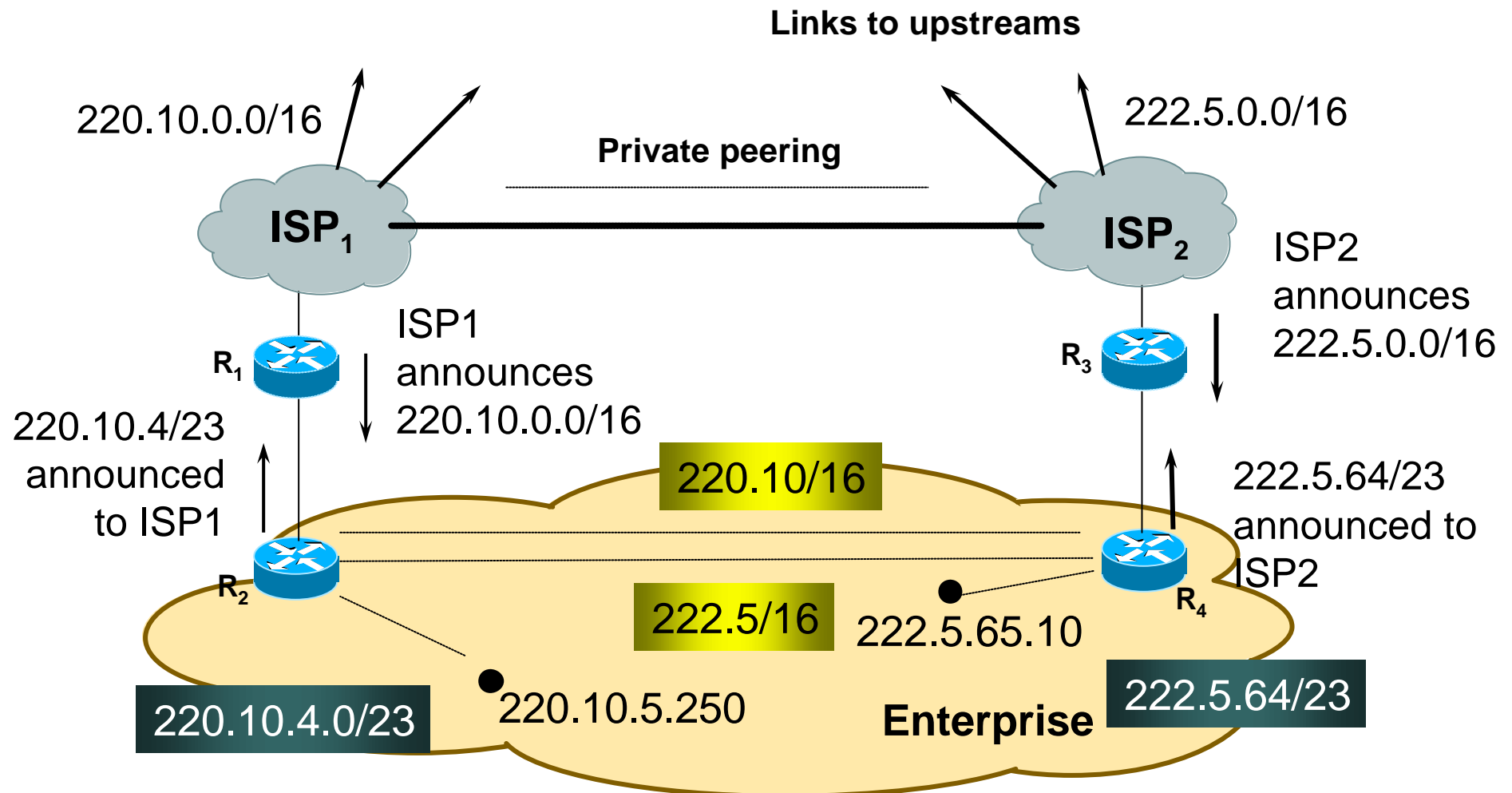
Enterprise Multihoming

Cisco.com

- **Situation similar to previous example**
 - address space used by enterprise comes from **both** upstream ISPs**
 - use conditional advertisement**
 - want to avoid leaking subprefixes of upstream provider address space into the Internet**

Steady State

Cisco.com



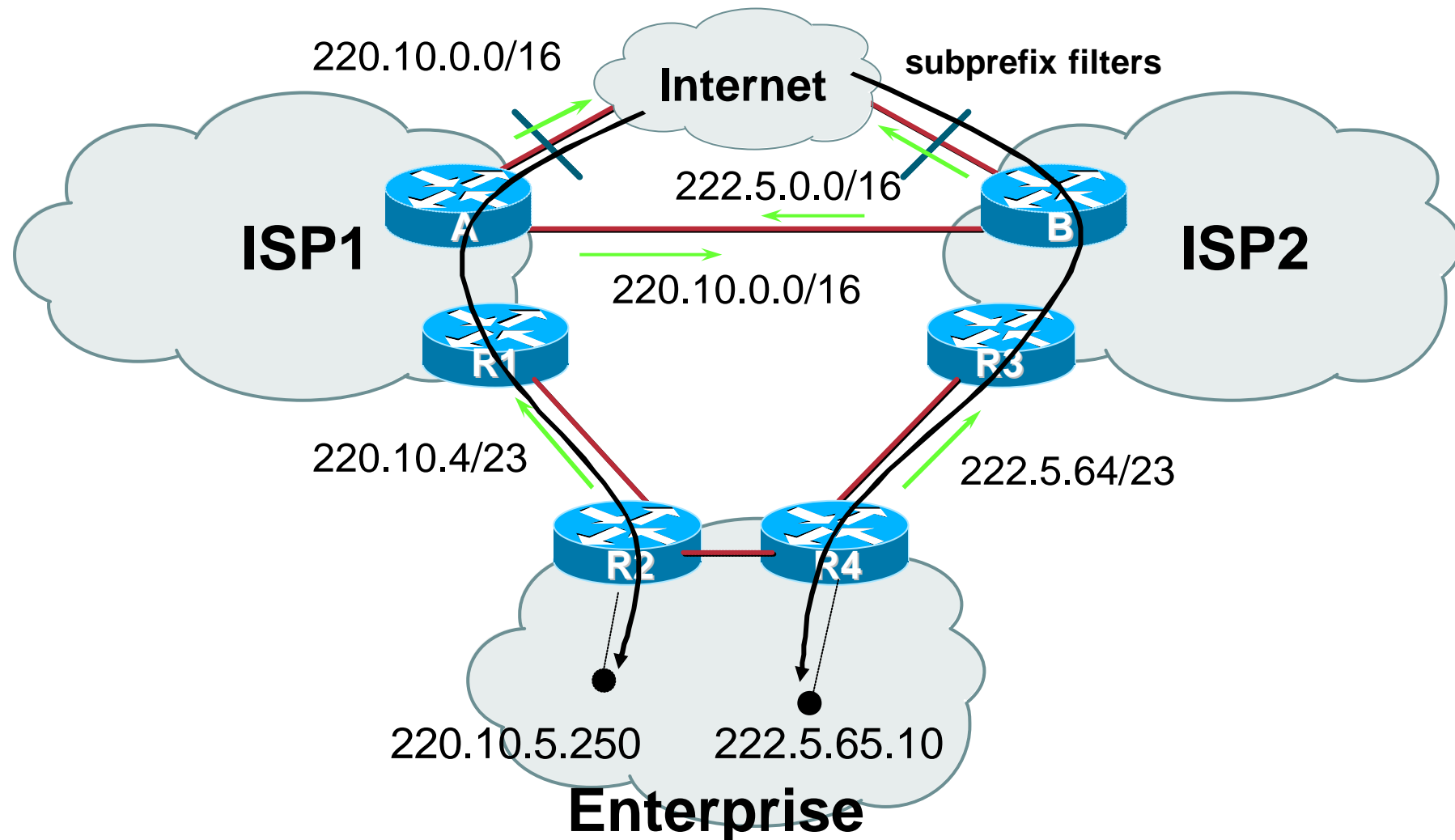
Steady State

Cisco.com

- **ISP1 and ISP2 have private peering
exchange each other's prefixes
enterprise customer is looking for link
redundancy only
no subprefixes leaked to Internet**
- **Configuration of R2 as in previous
example**

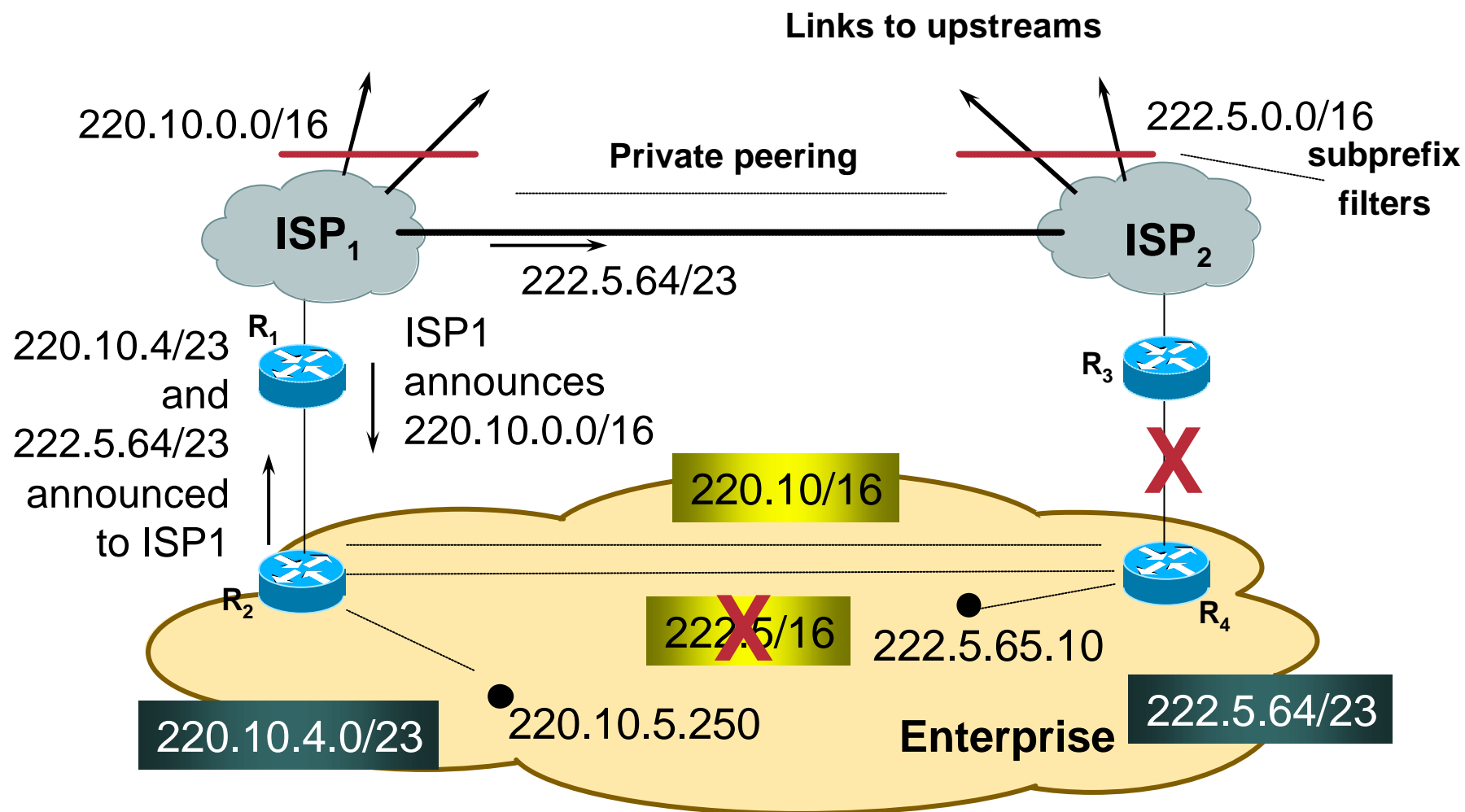
Traffic Flow Steady State

Cisco.com



Link Failure

Cisco.com



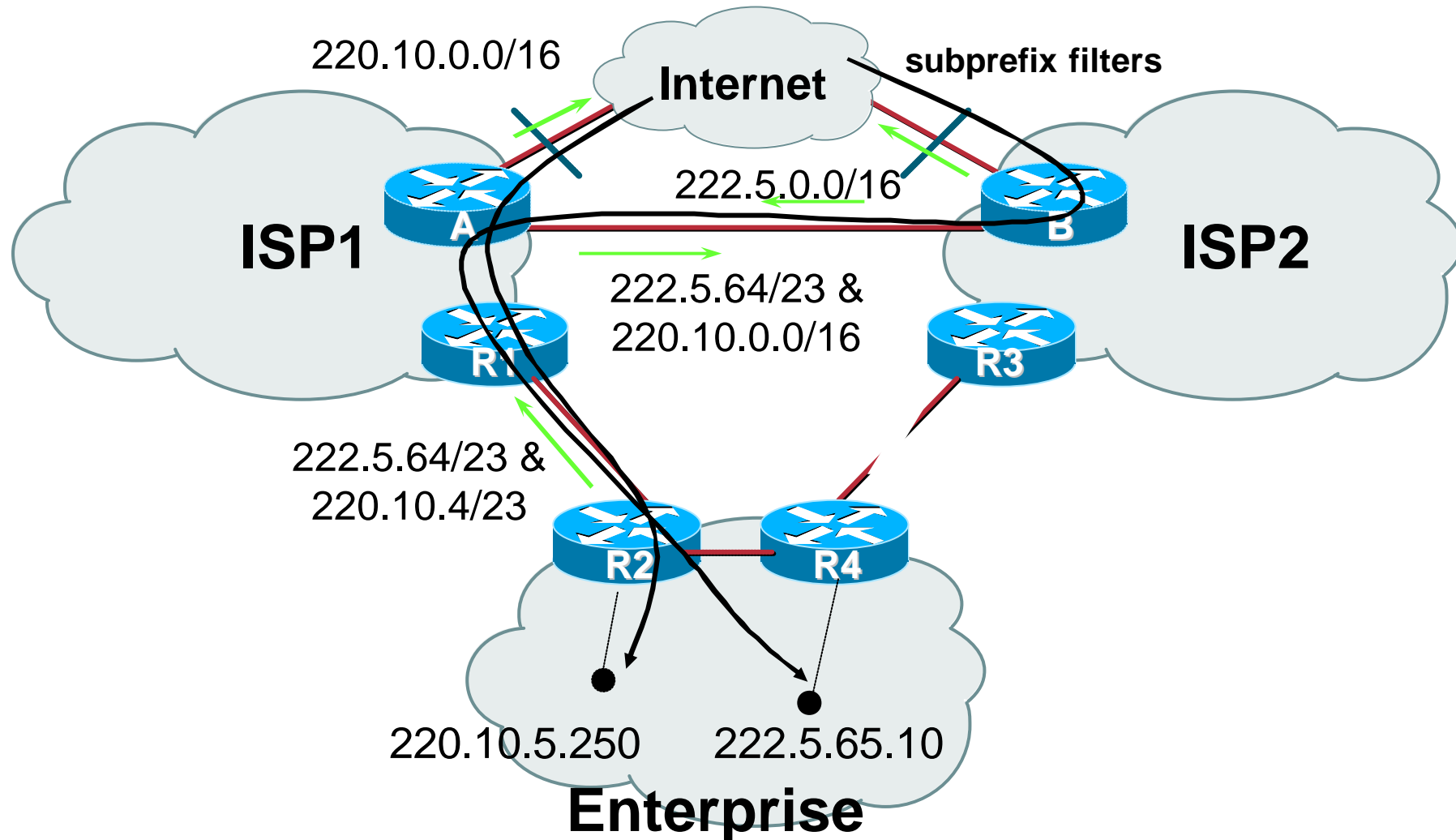
Link Failure

Cisco.com

- **R3 ® R4 link goes down**
 - conditional advertisement effective**
 - 222.5.64/23 announced by R2 to R1**
 - 222.5.64/23 announced by ISP1 to ISP2**
- **Filters!**
 - ISP1 and ISP2 filter subprefixes from their blocks outbound to Internet**
 - backup yet no subprefixes leaked to Internet**

Link Failure

Cisco.com



Configuration

- **RouterA ISP1 border router configuration:**

```
router bgp 150
  network 220.10.0.0 mask 255.255.0.0
  neighbor <routerB> remote-as 140
  neighbor <routerB> prefix-list isp2-in in
  neighbor <routerB> prefix-list isp2-out out
  neighbor <upstream> remote-as 110
  neighbor <upstream> prefix-list rfc1918-dsua in
  neighbor <upstream> prefix-list myblock out
!
ip route 220.10.0.0 255.255.0.0 null0
..next slide
```

Configuration

Cisco.com

```
ip prefix-list isp2-out permit 220.10.0.0/16
ip prefix-list isp2-out permit 222.5.64.0/23
!
ip prefix-list isp2-in permit 222.5.0.0/16
ip prefix-list isp2-in permit 220.10.4.0/23
!
ip prefix-list myblock permit 220.10.0.0/16
!
```

- The “myblock” prefix list ensures that no subprefixes are leaked to the Internet routing table

Recommendations

Cisco.com

- **Address space for Enterprise network should be obtained from **both** upstreams**
according to link bandwidths
- **Address space should be distributed according to utilisation**
loadsharing is about address assignment policies, monitoring bandwidth utilisation, as well as BGP attribute manipulation
- **Use a private AS – no need for a public AS**
needs agreement between two upstreams

What goes in the Internet Routing Registry?

Cisco.com

- **ISP1 and ISP2 obviously put their own address blocks as route objects in the IRR**
- **No need for any other entries as no subprefixes appear in the global internet routing table**
- **No inconsistent origin AS, no “problem”**

BGP and the Internet

Enterprise Multihoming