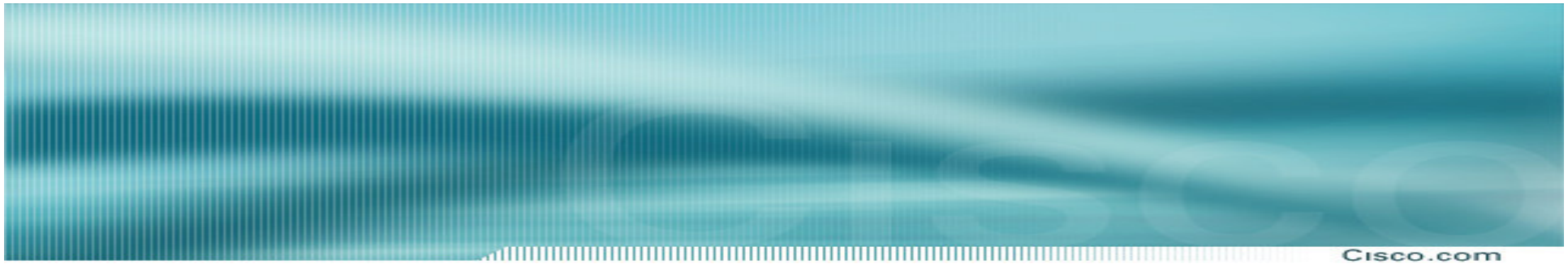
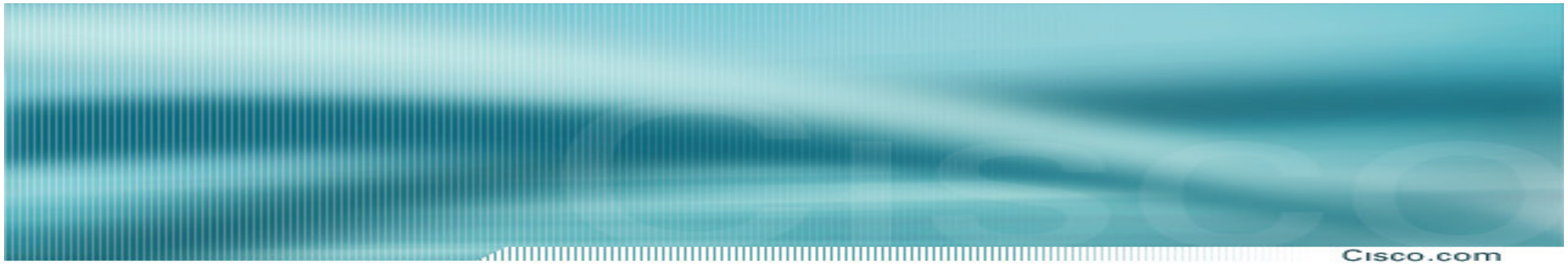


Preparation on the Control Plane

Agenda

Cisco.com

- IP Address management
- Routing Protocols and Architectures
- Routing Protocol Convergence Speed and Security
- Securing the Routing Protocol
- Protocol Authentication
- BGP BCPs That Help Build Security Resistance
- BGP BCPs that help add Resistance
- Default Routes, ISPs, and Security
- Route Flap Damping
- Prefix Filtering
  - ✓ How?
  - ✓ What?
  - ✓ Where?

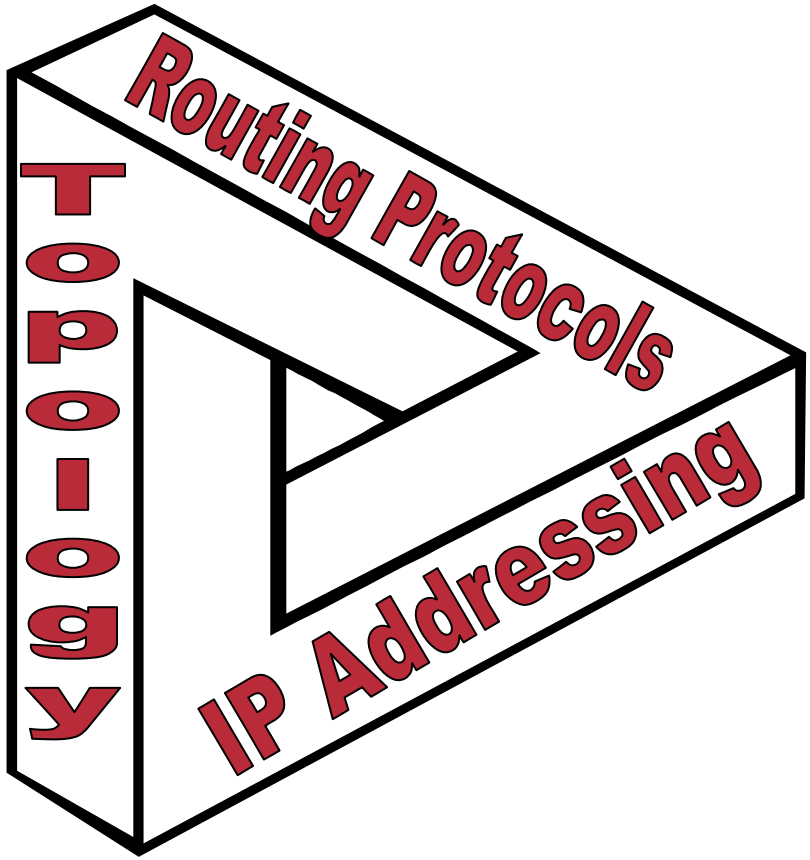


IP Address Management

How IP Addressing Effect the ISP's Architecture

Why is IP Address Essential to ISP Security?

- Effective Addressing with stability, routing aggregation, and security are key to keeping your Operational Security Overhead low.
- Critical to the scalability and success of an ISP's network
- Ignored, bad addressing policies will result in:
  - ✓ Higher cost
  - ✓ Longer deployment times
  - ✓ Complex troubles, and increased routing table convergence times



Principles of Addressing

- **Separate customer and infrastructure address pools**
  - ✓**Manageability**

    Different personnel manage infrastructure and assignments to customers
  - ✓**Scalability**

    Easier renumbering—customers are difficult, infrastructure is easy

Principles of Addressing

- **Further separate infrastructure**
  - ✓**In the IGP:**
    - **P2P addresses of backbone connections**
    - **Router loopback addresses**
  - ✓**Not in the IGP:**
    - **RAS server address pools**
    - **Virtual web and content hosting LANs**
    - **Mail, DNS servers**

Principles of Addressing

- Customer networks
  - ✓ Carry in iBGP
  - ✓ Do not put in IGP—ever
- Do not need to aggregate customer assigned address space
  - ✓ iBGP can carry in excess of unique 200,000 prefixes, no IGP is designed to do this.

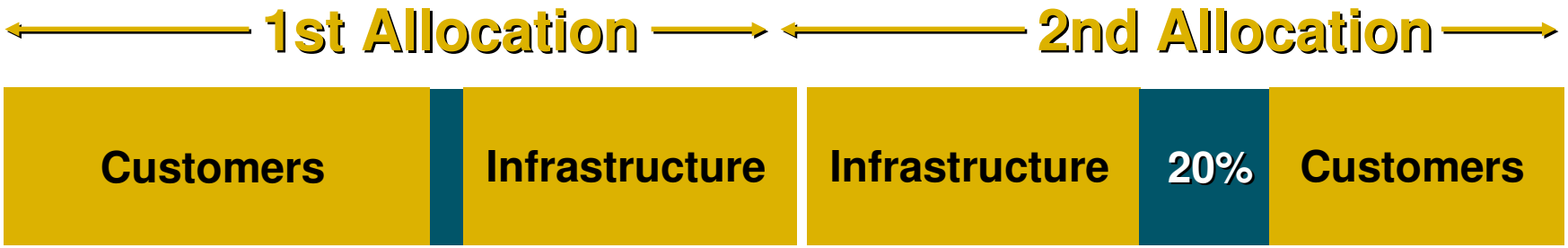
Management—Simple Network

- First allocation from APNIC
  - ✓Infrastructure is known, customers are not
  - ✓20% free is trigger for next request

  - ✓Grow usage of blocks from edges
  - ✓Assign customers sequentially

Management—Simple Network

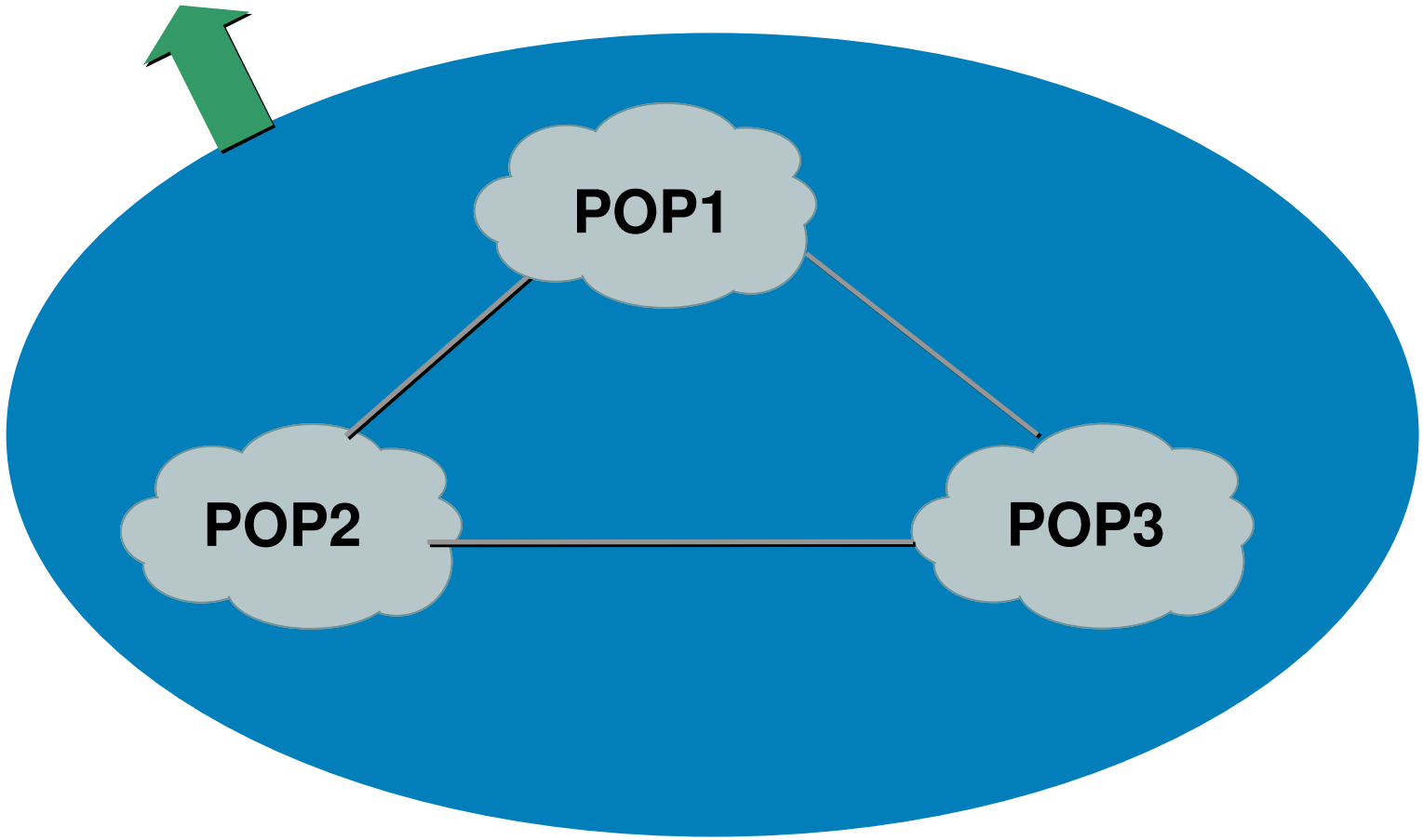
- If second allocation is contiguous



- ✓Reverse order of division of first block
- ✓Maximize contiguous space for infrastructure
  - Easier for debugging
- ✓Customer networks can be discontiguous

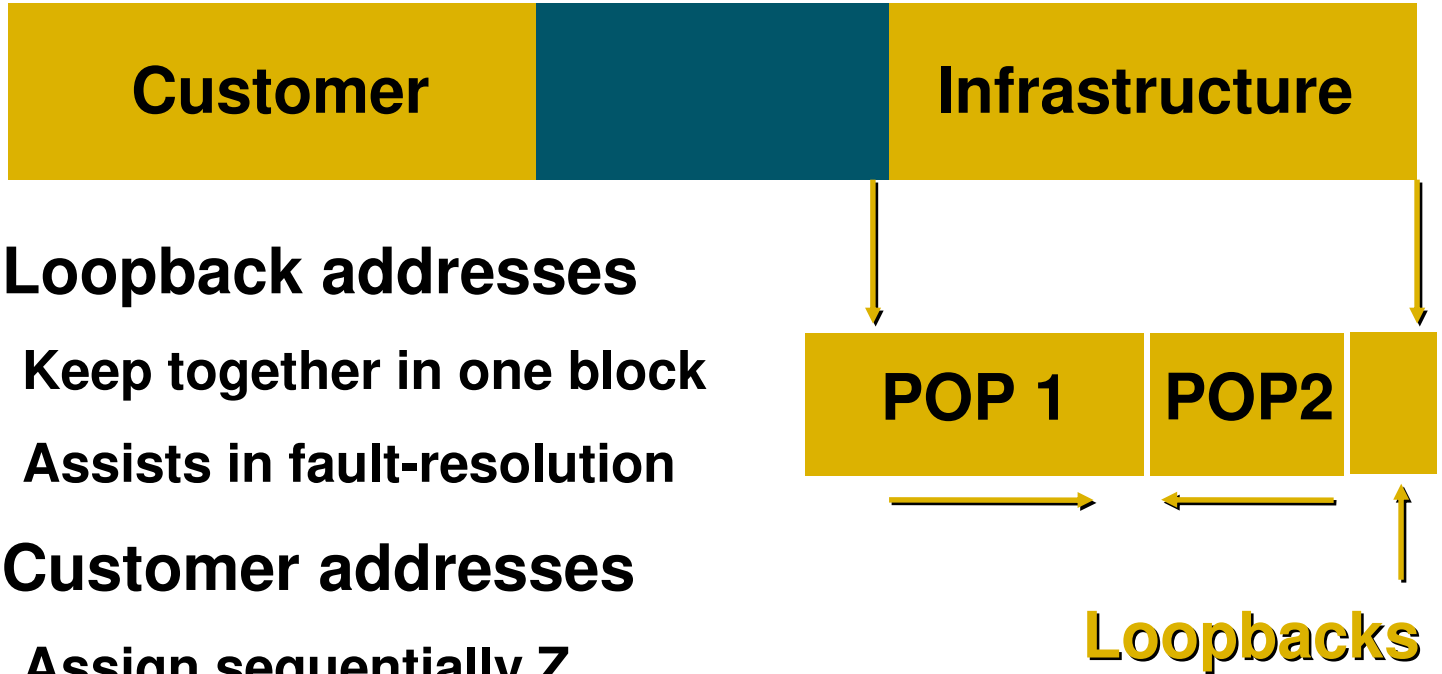
Management—Many POPs

WAN Link to Single Transit ISP



Management—Many POPs

- **POP sizes**
  - ✓**Choose address pool for each POP according to need**



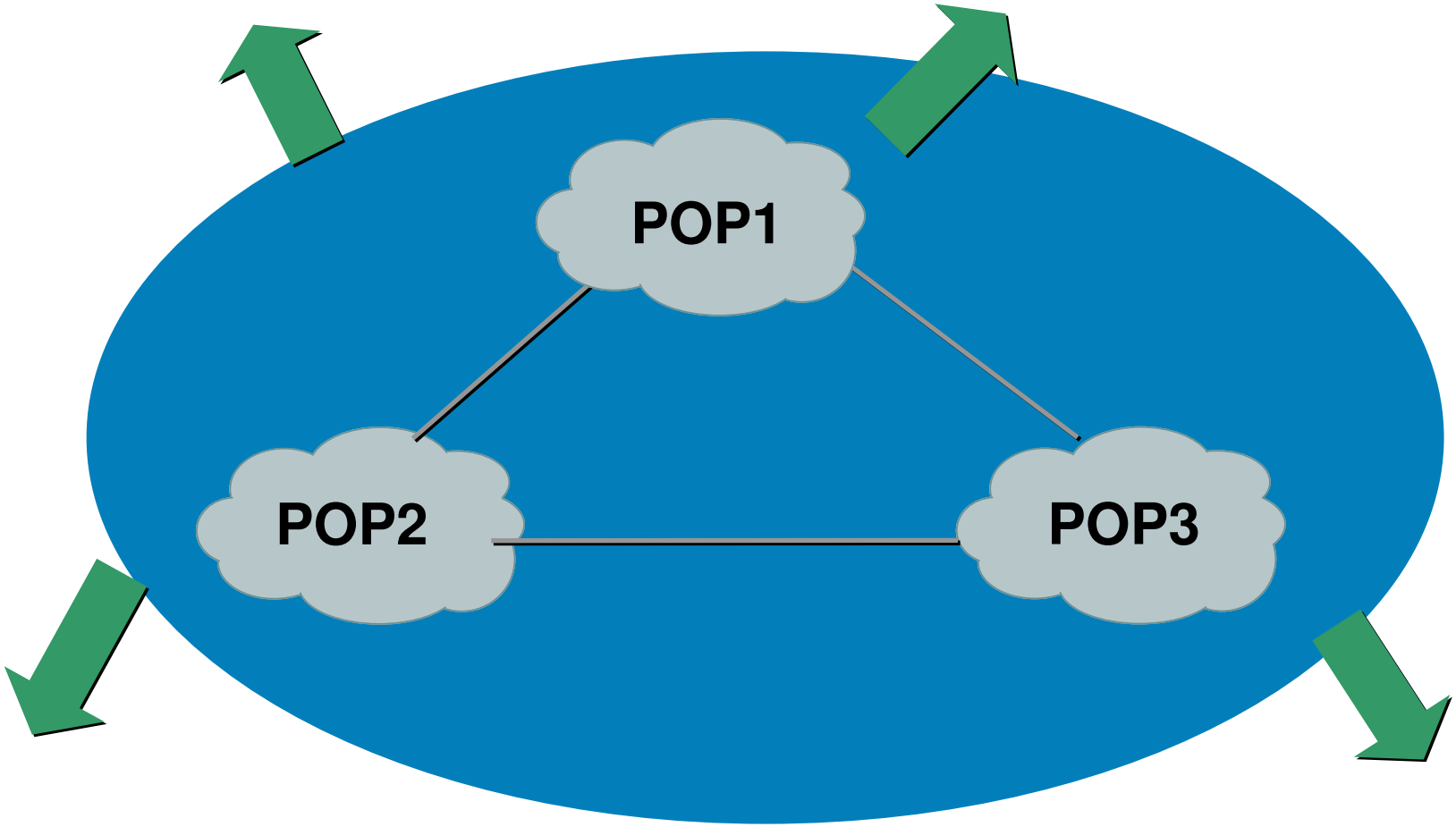
  - ✓**Loopback addresses**
    - **Keep together in one block**
    - **Assists in fault-resolution**
  - ✓**Customer addresses**
    - **Assign sequentially Z**

Management—Many POPs

- /20 minimum allocation is not enough for all your POPs
  - ✓Deploy addresses on infrastructure first
- Common mistake:
  - ✓Reserving customer addresses on a per POP basis
- Do not constrain network plans due to lack of address space
  - ✓Re-apply once address space has been used
  - ✓There is plenty of it

Management—Multiple Exits

WAN Links to Different ISPs



Management—Multiple Exits

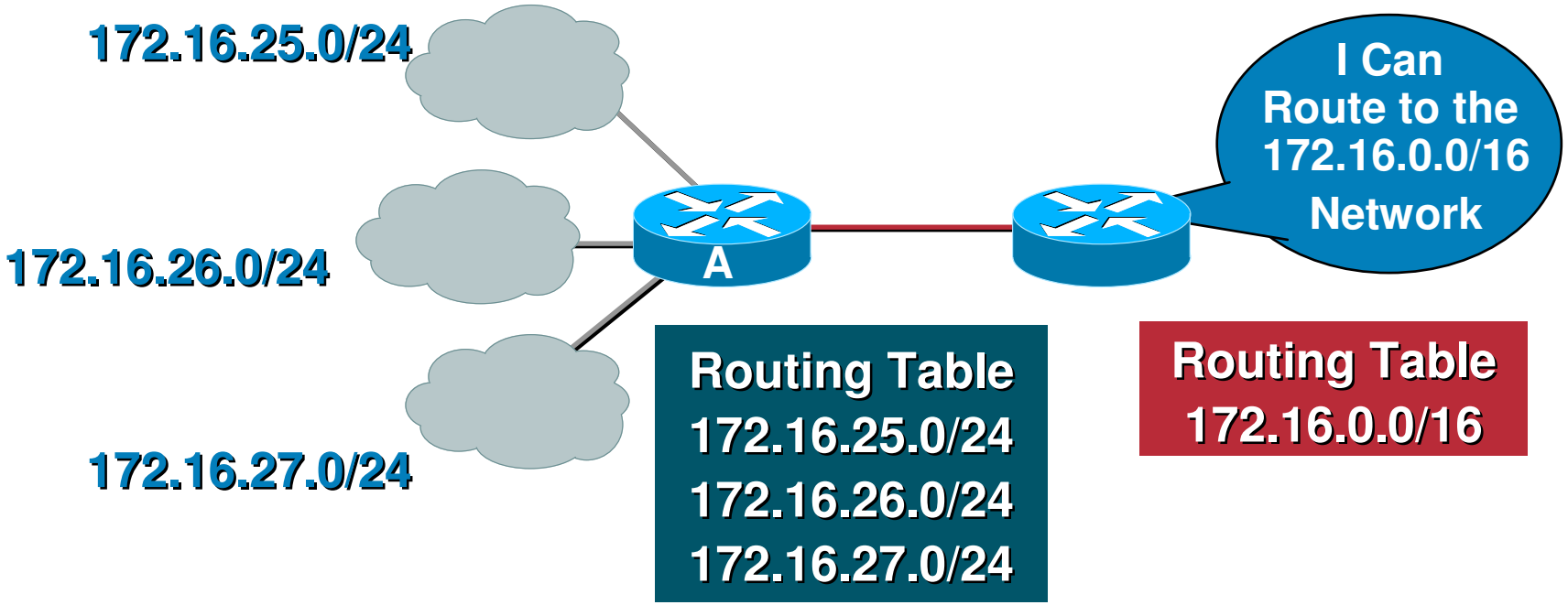
- Create a ‘national’ infrastructure pool

- Carry in IGP
  - ✓e.g. loopbacks, p2p links, infrastructure connecting routers and hosts which are multiply connected
- On a per POP basis
  - ✓Consider separate memberships if requirement for each POP is very large from day one

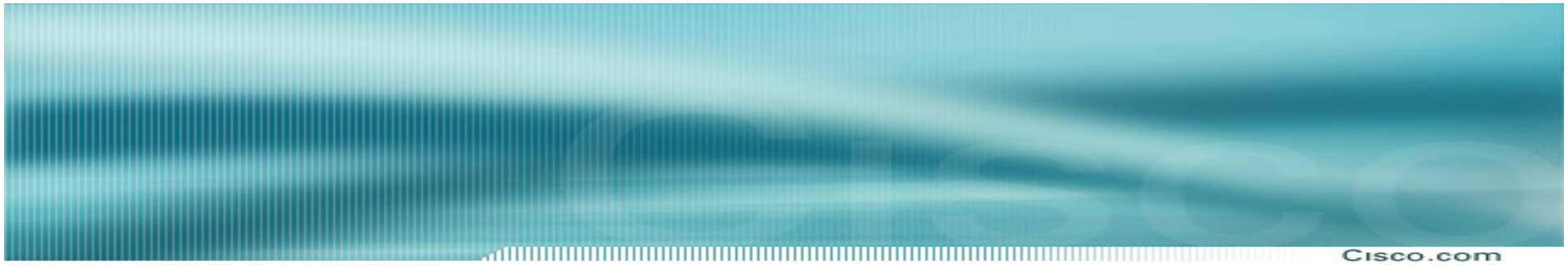
Information Reduction

- How do you manage 10,000 devices on a network
- Grouping, aggregation, and summarization of data is an important scaling tool
- Networks that do not, run into problems

Information Reduction



- **Example—Summarizing IPv4 addresses in a network—Leaner RIBs mean faster convergence**

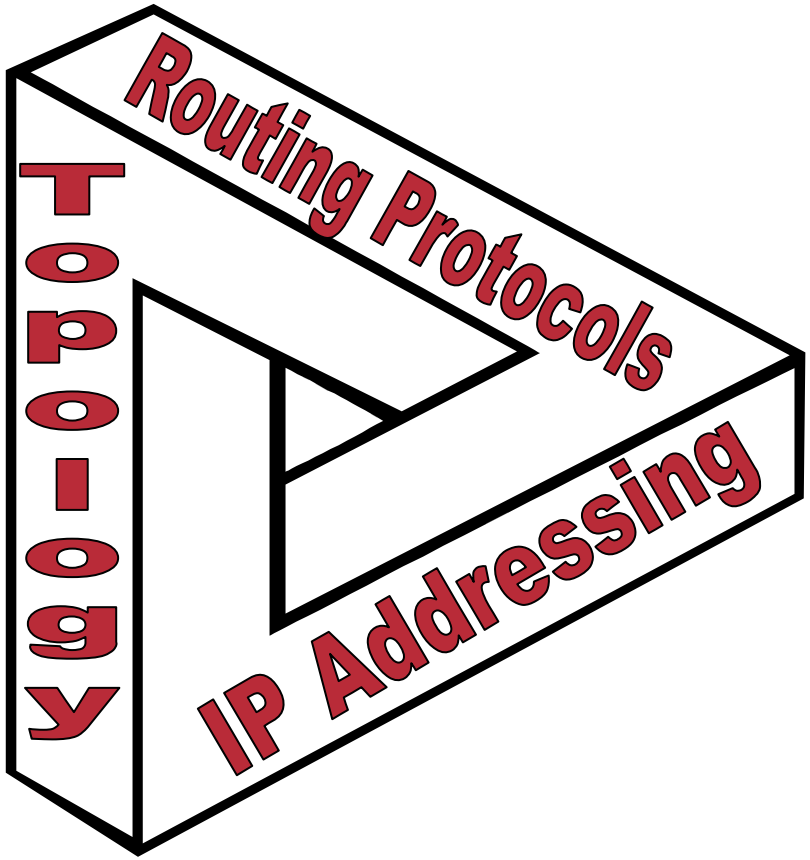


Routing Protocols and Architectures

How Routing Protocols Effect the ISP Architecture

Routing Protocols Effect on ISP Architecture

- Routing Protocols inter-act with IP addressing and Topology to provide the glue that makes it all work
- Poor routing protocol planning is often the first indication that a network if having scaling problems



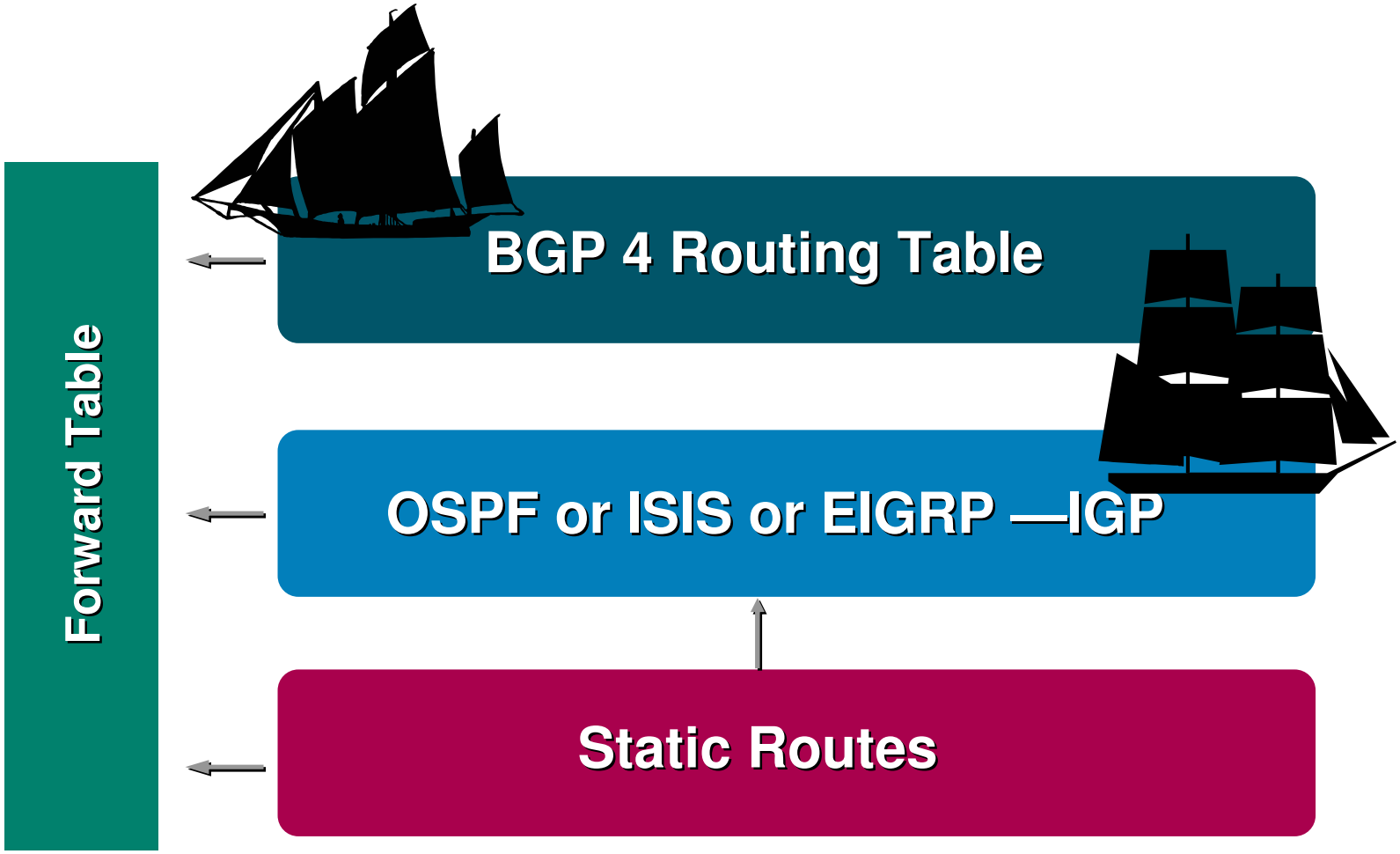
Interior vs Exterior Routing Protocols

IGP	EGP
Automatic Neighbor Discovery	Configured Peers
Generally Trust Your Peers	Outside Connections; Peer Trust Minimal
Most Reachability Information Shared	Administratively Bound Reachability
Binds Routers in A Network Together	Binds Networks Together (Internets)

Interior vs Exterior Routing Protocols

IGP	EGP
Carries ISP Infrastructure Information Only	Carries Customer Routes and Reachability to Other AS'
Smaller Tables for Effeciency and Speed of Convergence	Large Tables with Stability as a Priority
Tied to the Network Topology	Relatively Independent of Network Topology

Ships in the Night



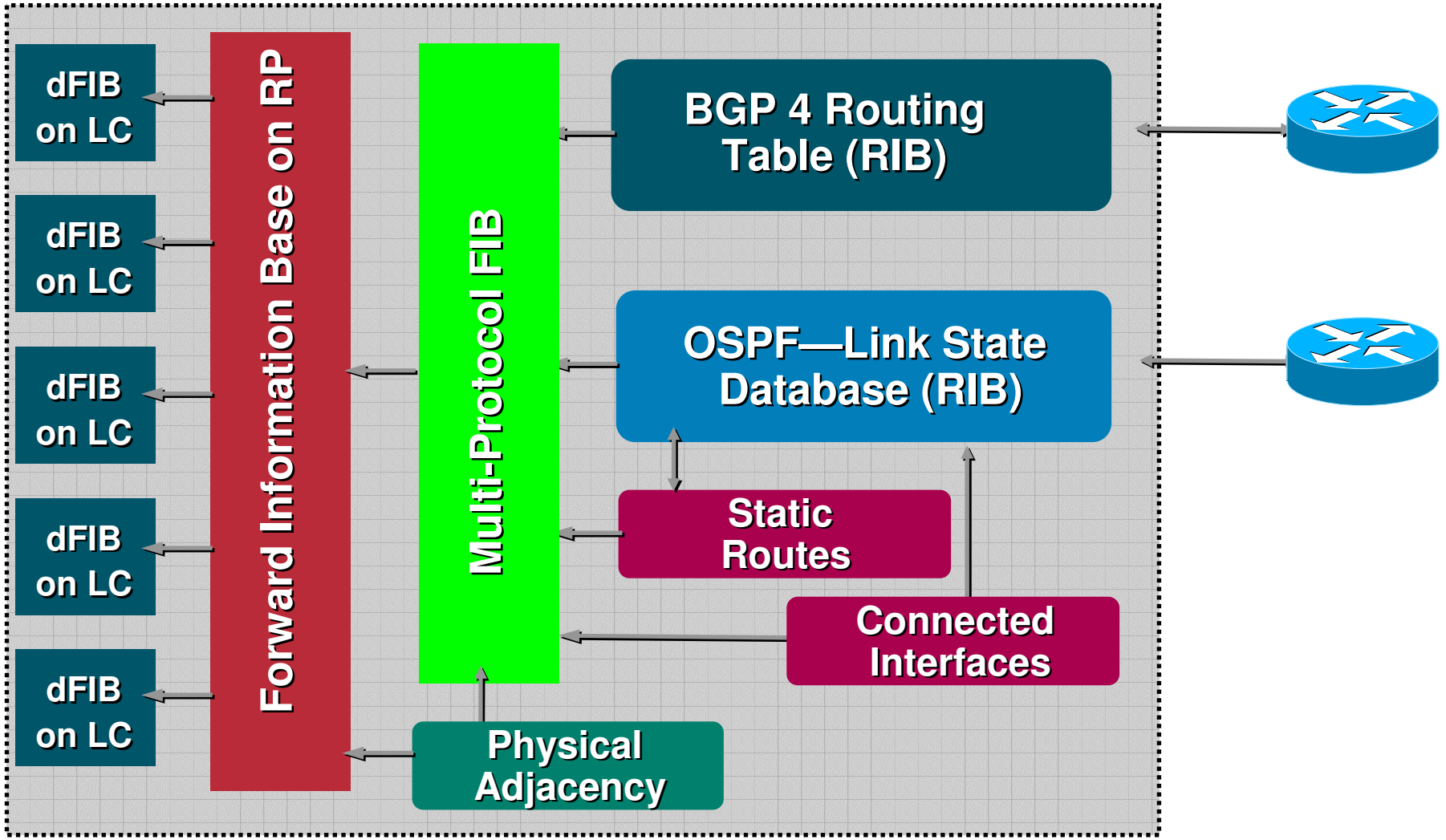
IGPs and EGPs Have Two Different Functions inside an ISP

- **Interior (OSPF or ISIS)**
  - ✓Glues the internal network together
  - ✓Propagates next-hop locations
  - ✓Fast Convergence and redundancy
- **Exterior (BGPv4)**
  - ✓Glues the internet together
  - ✓Carries customer networks in iBGP
  - ✓Stability on the Internet

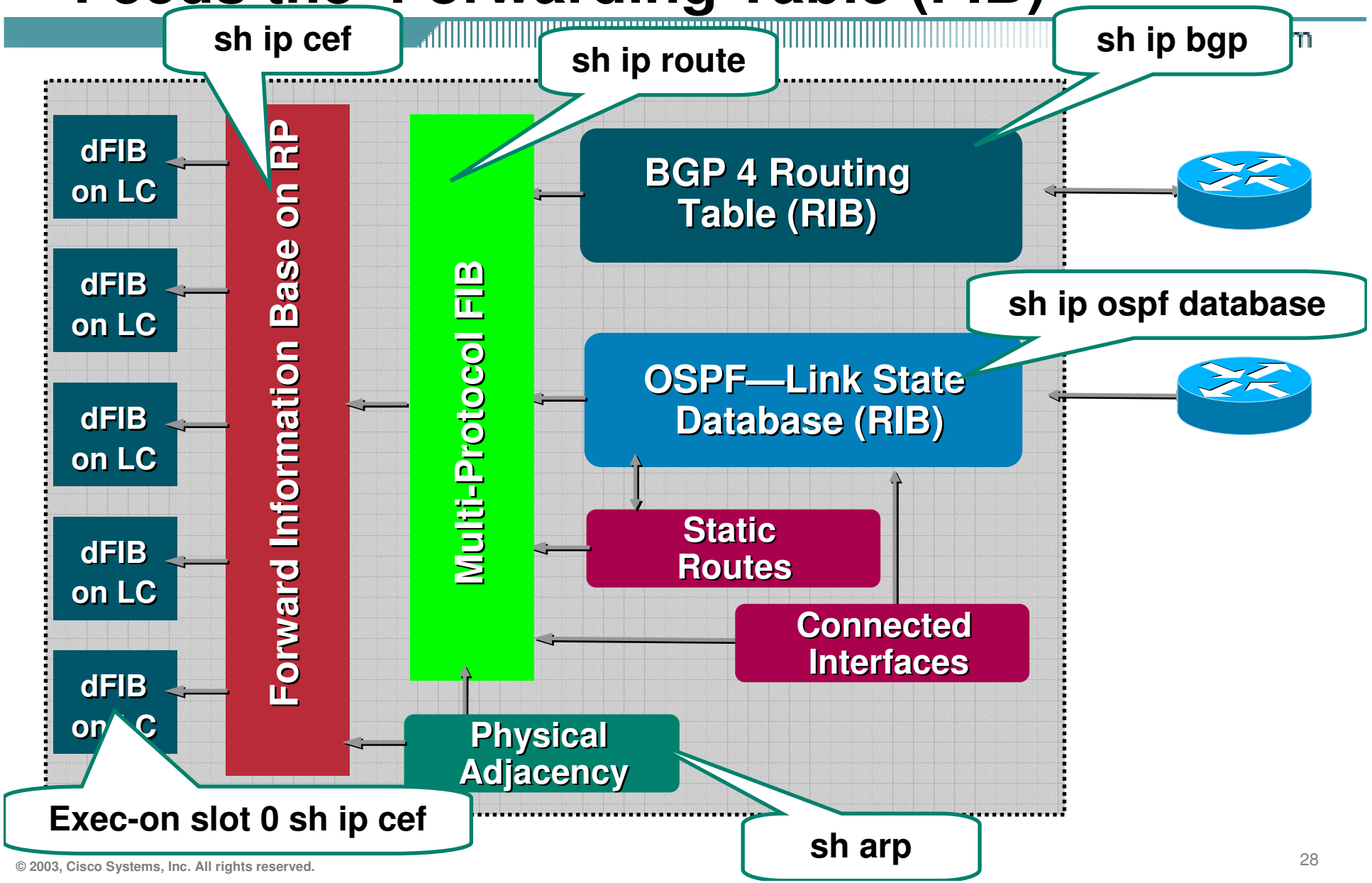
New Terminology

- **Routing Information Base (RIB)**
  - ✓Generated by each routing protocol
- **Forwarding Information Base (FIB)**
  - ✓Network layer routing information
  - ✓New term used to describe the forwarding table
- **Adjacency table (Adj)**
  - ✓ Next hop link layer information
- **Distributed FIB**
  - ✓FIB push out to the line cards on a router so that forwarding can be done locally on each line card

Routing Tables (RIB) Feeds the Forwarding Table (FIB)



Routing Tables (RIB) Feeds the Forwarding Table (FIB)



BGP vs. OSPF/ISIS

- **Internal Routing Protocols (IGPs)**
  - ✓ **Examples are ISIS and OSPF**
  - ✓ **Used for carrying infrastructure addresses**
  - ✓ **Not used for carrying Internet prefixes or customer prefixes**

BGP vs. OSPF/ISIS

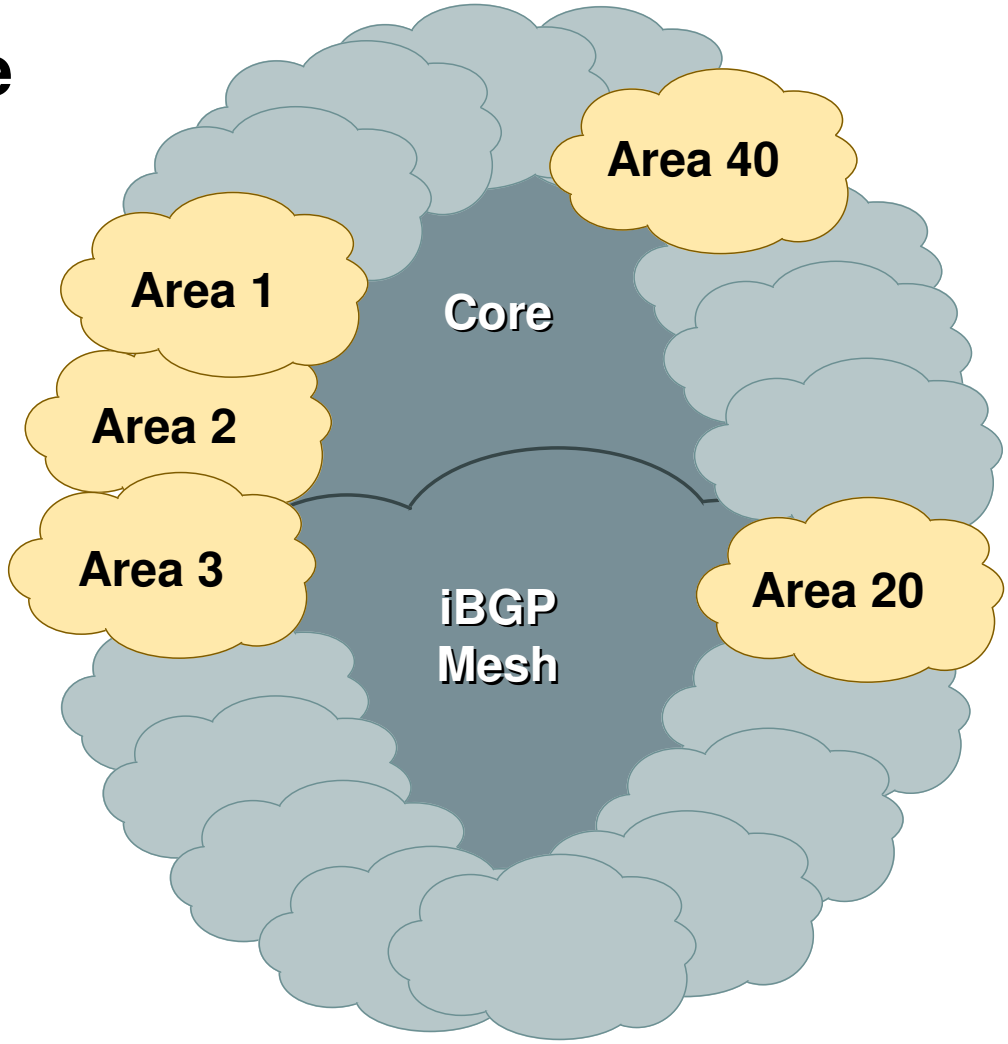
- BGP used internally (iBGP) and externally (eBGP)
- iBGP used to carry
  - ✓Some/all Internet prefixes across backbone
  - ✓Customer prefixes
- eBGP used to
  - ✓Exchange prefixes with other ASes
  - ✓Implement routing policy

BGP vs. OSPF/ISIS

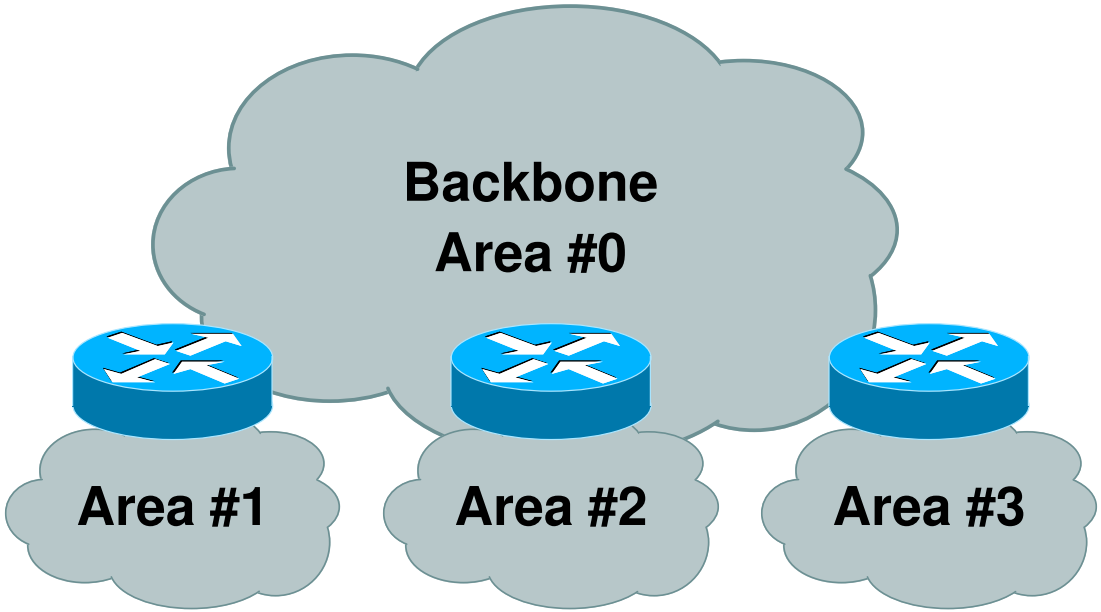
- **Do not:**
  - X Distribute BGP prefixes into an IGP
  - X Distribute IGP routes into BGP
  - X Use an IGP to carry customer prefixes
- **YOUR NETWORK WILL NOT SCALE**

Today's IGP/EGP Hierarchy

- IGP carries only core links plus peering address information
- BGP carries all the routes
- Reset BGP Next-hop to loopback.
  - ✓ Security, stability, and efficiency
- Increased stability



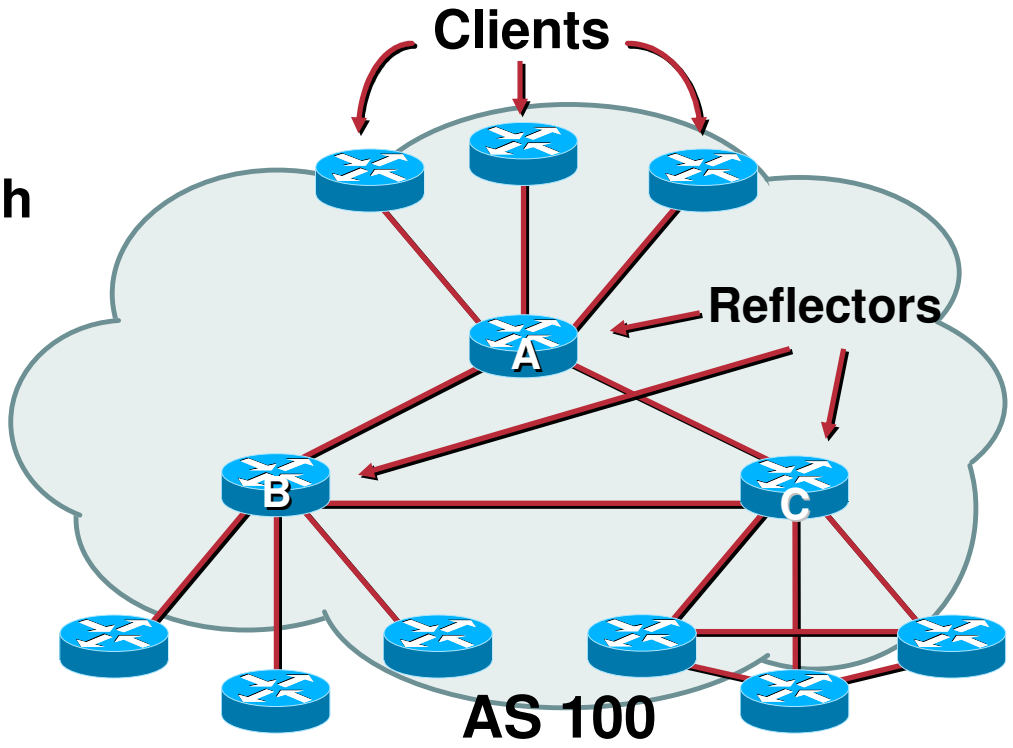
Routing Protocols and Architecture

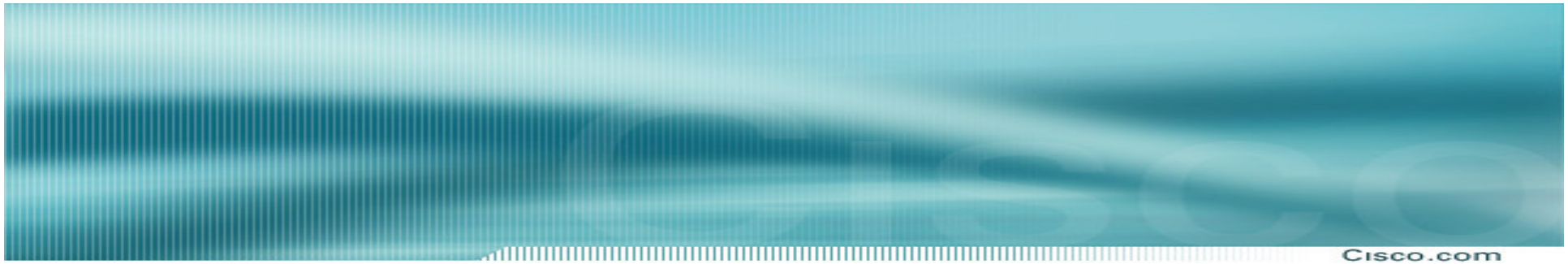


- OSPF/ISIS Hierarchy should, where possible, match topology hierarchy
- Results in marked reduction in routing traffic

Routing Protocols and Architecture

- BGP route reflector or Confederation Hierarchies should where possible match the network's topologies
- Good synergy would be productive to a healthy network
- No congruency could mean complex problems that are hard to resolve





Routing Protocol Convergence Speed and Security

Tune for Fast Convergence

- **Faster Convergence means the network can recover from a security incident.**
  - ✓ **Interface Settings:**

    Increase the interface "hold-queue 1500" for each interface (default is 75). Do check your Line Card/VIP memory before increasing the hold-queue.
  - ✓ **TCP Settings (improves BGP convergence):**

    ip tcp selective-ack

    ip tcp mss 1460

    ip tcp window-size 65535

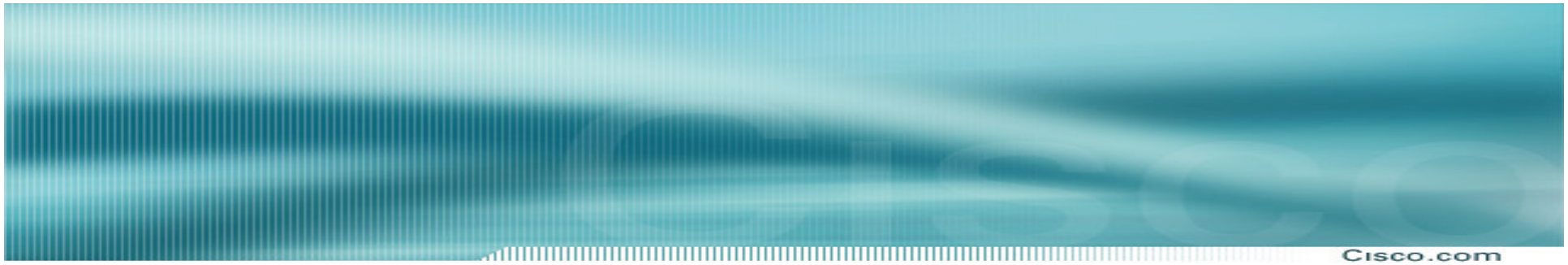
    ip tcp queuemax 50

    ip tcp path-mtu-discovery

Tune for Fast Convergence

- **Cisco 12000 Line Cards:**

  Increase "ip cef linecard ipc memory 10000" to improve the FIB download process.



Securing the Routing Protocol

Routing Protocol Security

- Routing protocol can be attacked
  - ✓ Denial of service
  - ✓ Smoke screens
  - ✓ False information
  - ✓ Reroute packets

May Be Accidental or Intentional

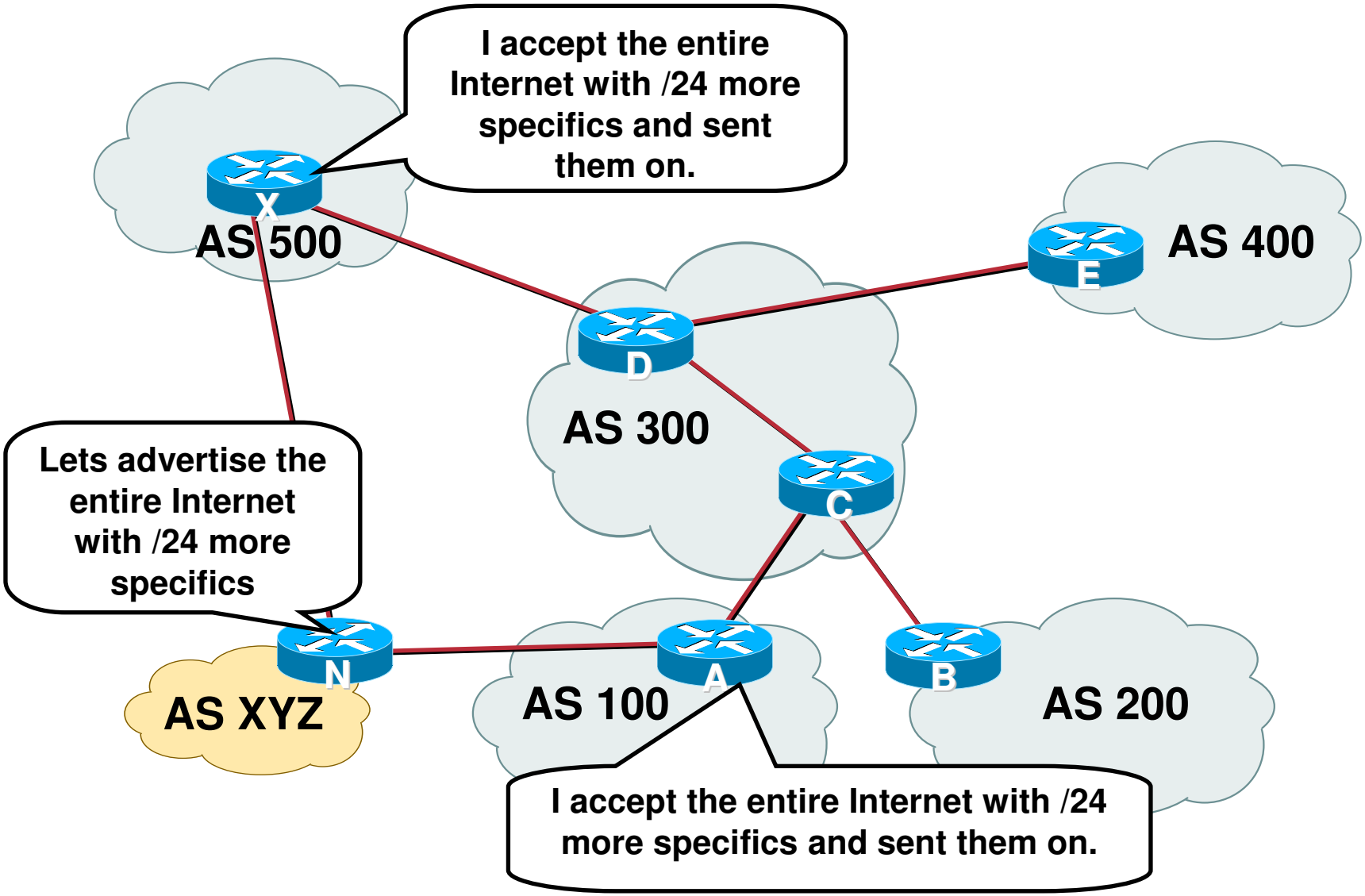
Malicious Route Injection
Perceive Threat

- **Bad Routing Information does leak out. This has been from mistakes, failures, bugs, and intentional.**
- **Intruders are beginning to understand that privileged access to a router means route tables can be altered**
- **CERT/CC is aware of a small number of incidents involving malicious use of routing information.**
- **Perceived Threat is that this will be a growth area for attackers.**

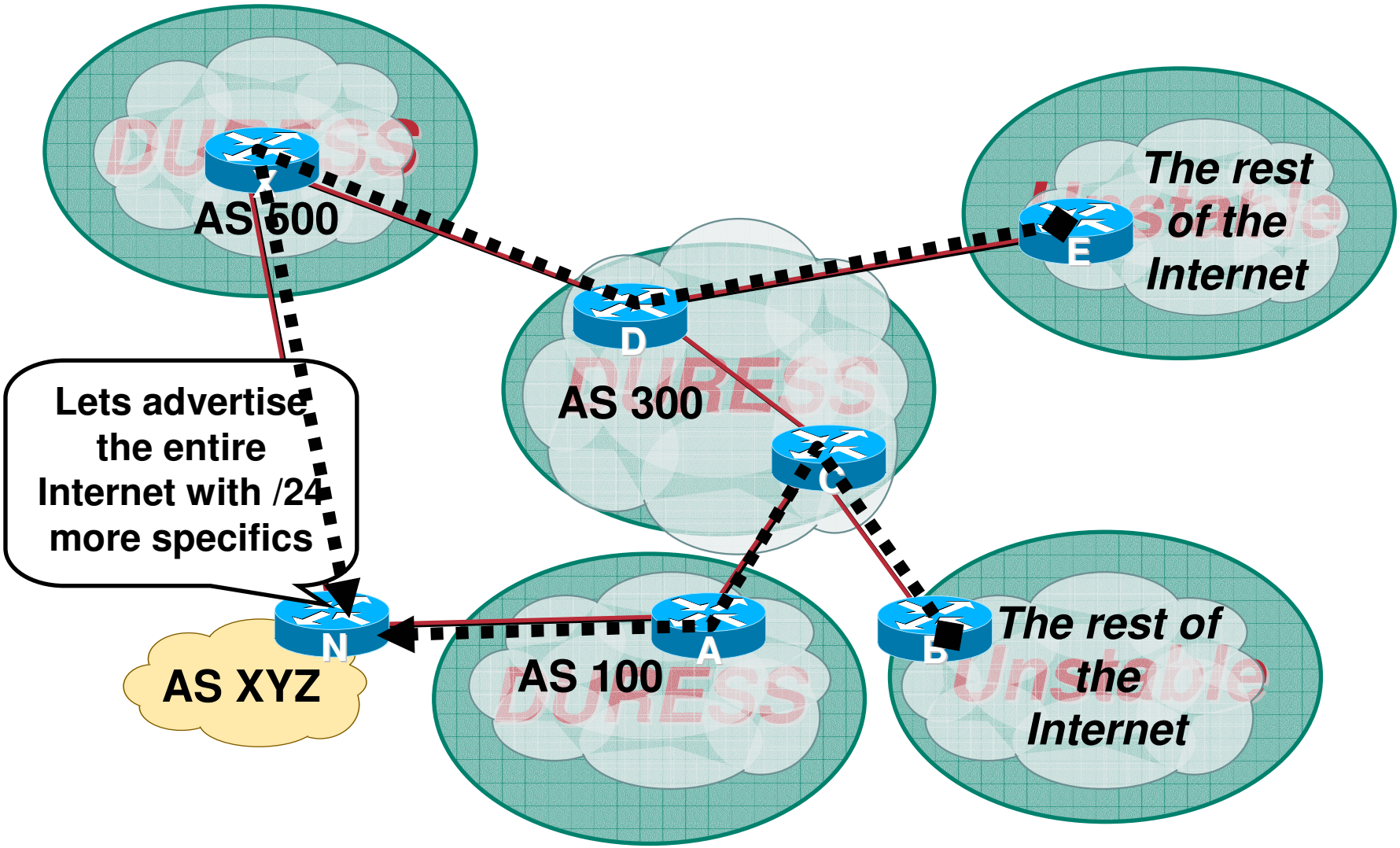
Malicious Route Injection
Reality – an Example

- **AS 7007 incident used as an attack.**
- **Multihomed CPE router is violated and used to “de-aggregate” large blocks of the Internet.**
- **Evidence collected by several CERTs that hundreds of CPEs are violated.**

Garbage in – Garbage Out: What is it?

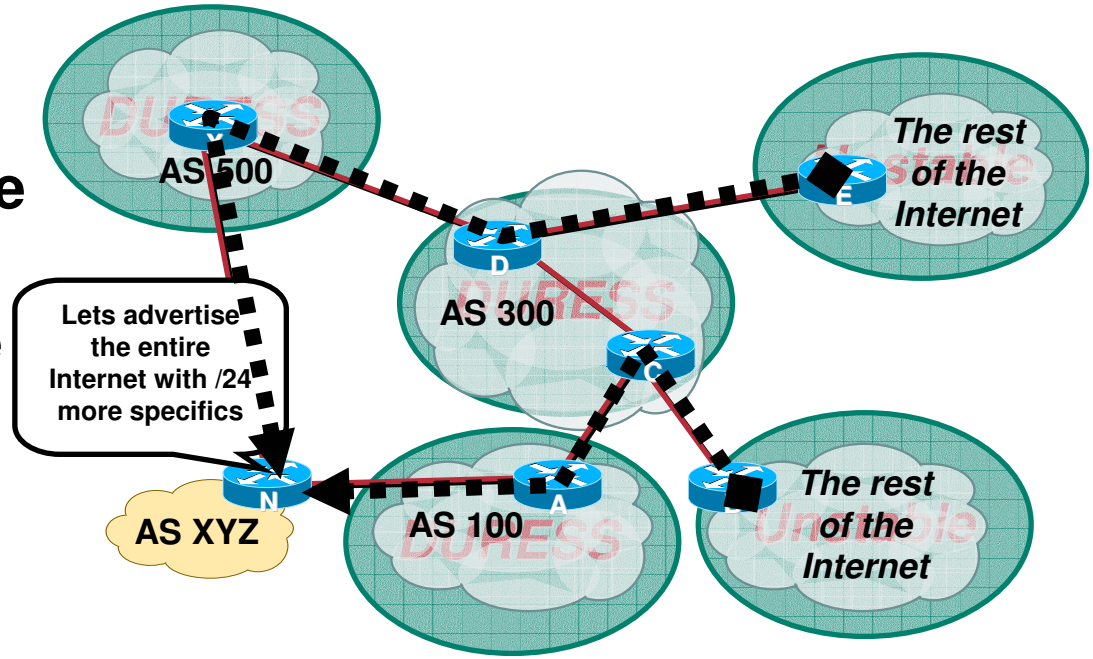


Garbage in – Garbage Out: Results



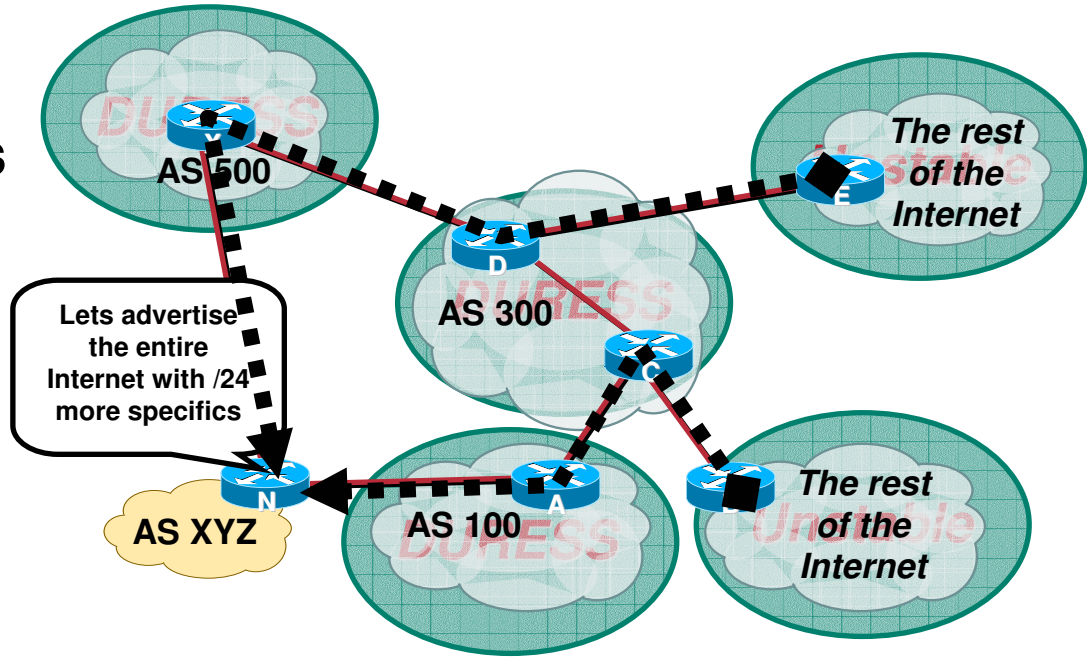
Garbage in – Garbage Out: Impact

- Garbage in – Garbage out does happen on the Net
- AS 7007 Incident (1997) was the most visible case of this problem.
- Key damage are to those ISPs who pass on the garbage.
- Disruption, Duress, and Instability has been an Internet wide effect of Garbage in – Garbage out.



Garbage in – Garbage Out: What to do?

- **Take care of your own Network.**
  - ✓ **Filter your customers**
  - ✓ **Filter you advertisements**
- **Net Police Filtering**
  - ✓ **Mitigate the impact when it happens**
- **Prefix Filtering and Max Prefix Limits**



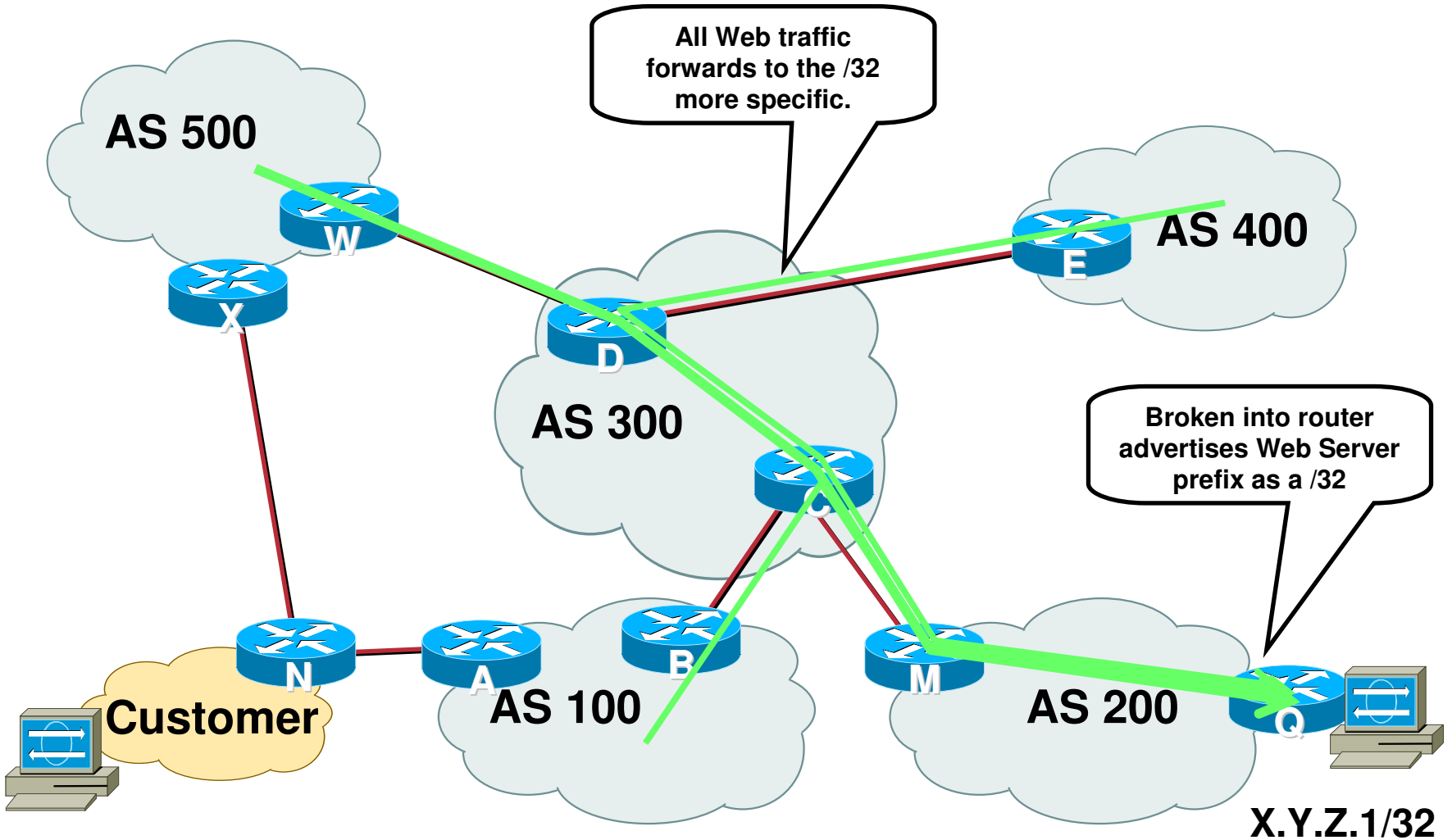
Malicious Route Injection
Attack Methods

- **Good News – Risk is mainly to BGP speaking Routers.**
- **Bad News – Multihomed BGP Speaking customers are increasing!**
- **Really Bad News – Many of these routers have no passwords!**
- **Local layer 3 configuration alteration on compromised router**
- **Intra-AS propagation of bad routing information**
- **Inter-AS propagation of bad routing information**

Malicious Route Injection
Impact

- **Denial-Of-Service to Customer(s), ISP(s), and the Internet.**
- **Traffic Redirection / Interception**
- **Prefix Hijacking**
- **AS Hijacking**

What is a prefix hijack?



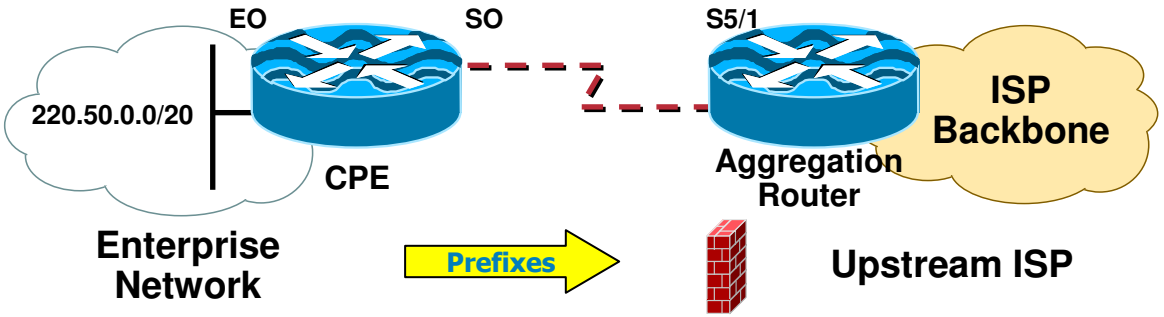
Malicious Route Injection
What can ISPs Do?

- **Customer Ingress Prefix Filtering!**
- **ISPs should only accept customer prefixes which have been assigned or allocated to their downstream customers.**
- **For example**
  - ✓ **Downstream customer has 220.50.0.0/20 block.**
  - ✓ **Customer should only announce this to peers.**
  - ✓ **Upstream peers should only accept this prefix.**

Malicious Route Injection
What can ISPs Do?

- **Cisco Configuration Example on Upstream**

```
router bgp 100
 neighbor 222.222.10.1 remote-as 101
 neighbor 222.222.10.1 prefix-list customer in
!
ip prefix-list customer permit 220.50.0.0/20
ip prefix-list customer deny 0.0.0.0/0 le 32
```



Malicious Route Injection
What can ISPs Do?

- **Containment Filters!**
  - ✓ **Design your network with the principles of of survivability.**
  - ✓ **Murphy's Law of Networking implies that the customer ingress prefix filter will fail.**
  - ✓ **Remember 70% to 80% of ISP problems are maintenance injected trouble (MIT).**
  - ✓ **Place Egress Prefix Filters on the Network to contain prefix leaks.**

What can ISPs Do?
Containment Egress Prefix Filters

- **It is not rocket science!**
- **Just create a hard list of your RIR allocated prefixes.**
- **Cisco Configuration Example**

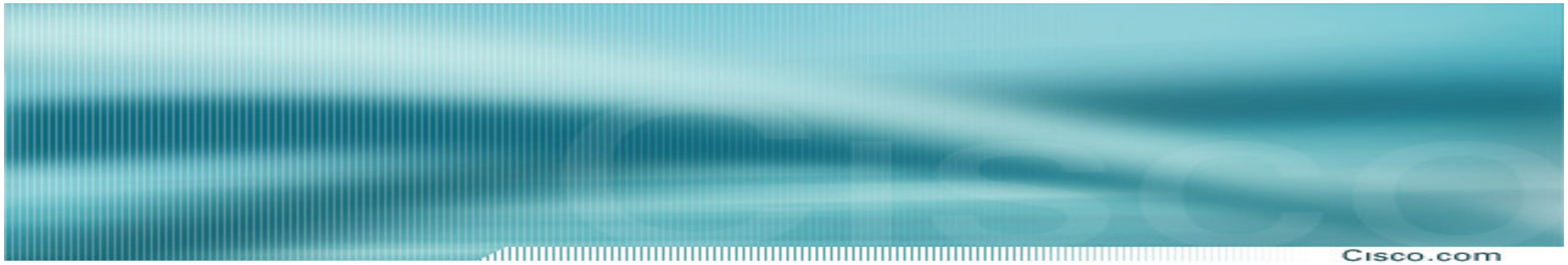
```
router bgp 100
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.1 remote-as 101
 neighbor 222.222.10.1 prefix-list out-filter out
!
ip route 221.10.0.0 255.255.224.0 null0
!
ip prefix-list out-filter permit 221.10.0.0/19
ip prefix-list out-filter deny 0.0.0.0/0 le 32
```

What can ISPs Do?
Containment Egress Prefix Filters

- **What about all my multihomed customers with prefixes from other ISPs?**
- **Add them to the customer ingress prefix filter.**
  - ✓ **You <u>should know</u> what you will accept.**
- **Add them to the master egress prefix-filter.**
  - ✓ **You <u>should know</u> what you're advertising to everyone else.**
  - ✓ ***Bigness* is not an excuse.**

Malicious Route Injection
What can ISPs Do?

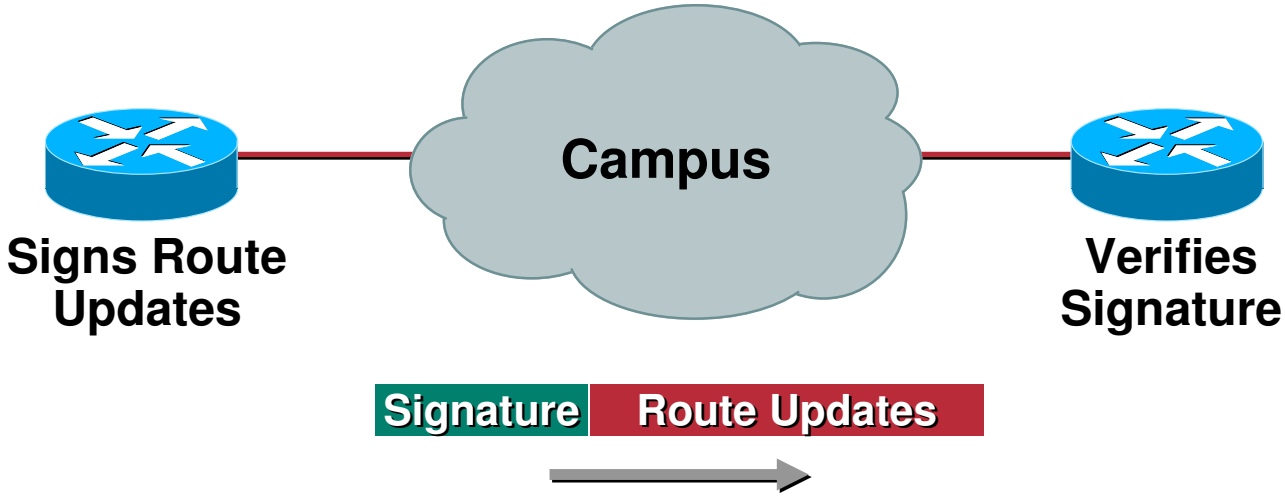
- **Customer Ingress Prefix Filtering**
- **Prefix filtering between intra-AS trust zones**
- **Route table monitoring to detect alteration of critical route paths**
- **SPAMers are using route-hijacking.**



Protocol Authentication

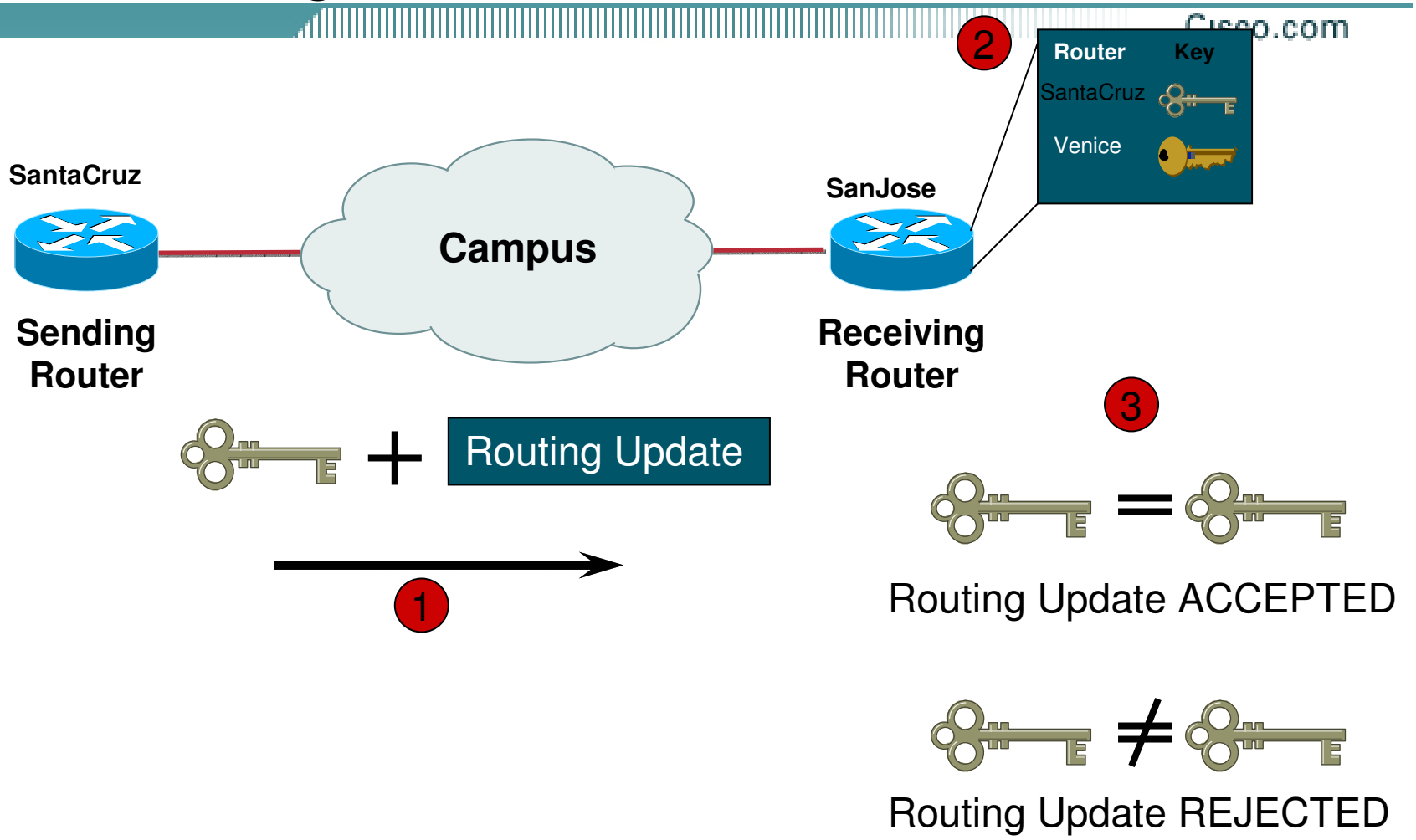
Secure Routing
Route Authentication

Configure Routing Authentication

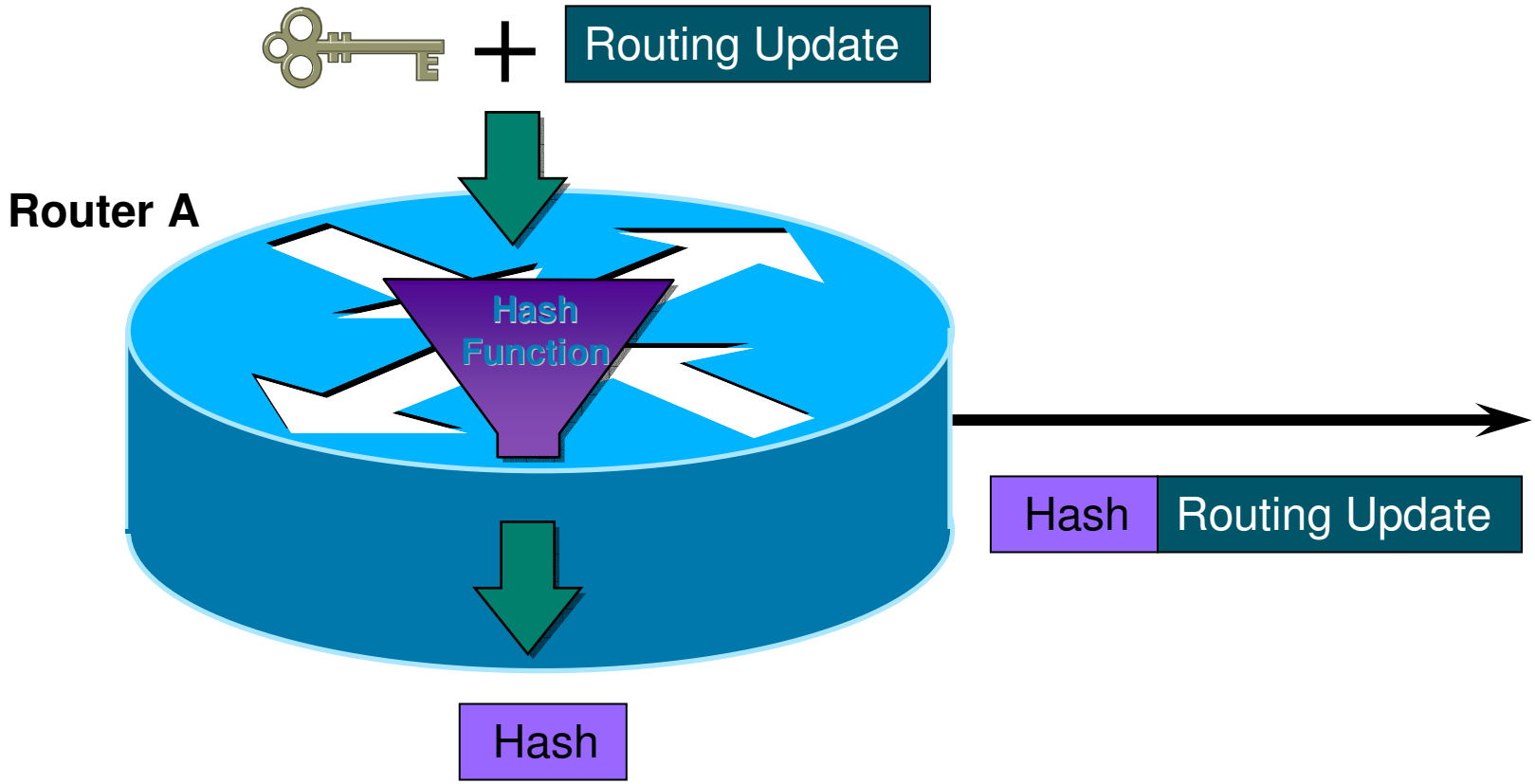


Certifies Authenticity of Neighbor and Integrity of Route Updates

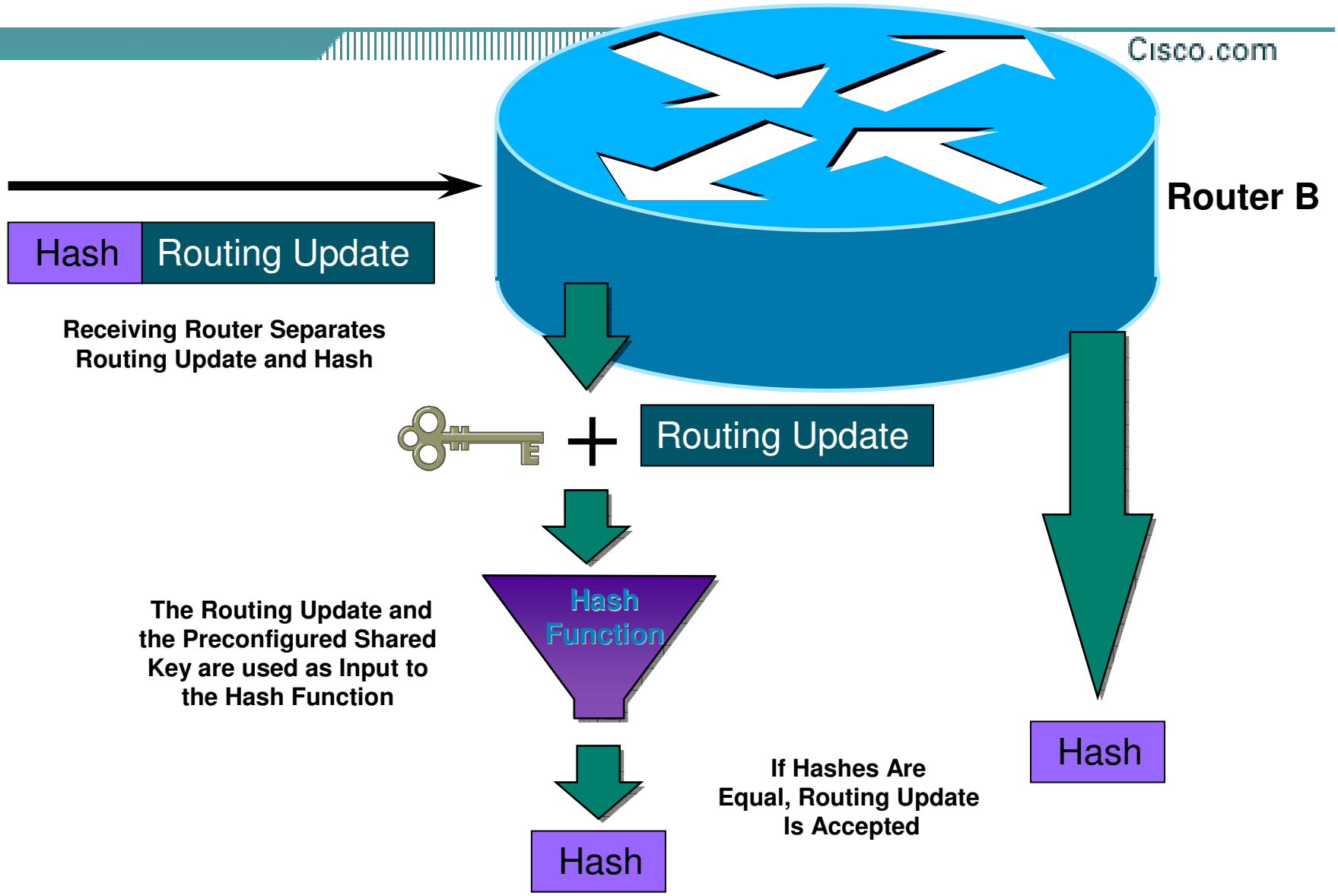
Plain-text neighbor authentication



MD-5 Neighbor Authentication: Originating Router



MD-5 Neighbor Authentication: Receiving Router



Peer Authentication

- Authenticates routing update packets
- Shared key included in routing updates
  - ✓ Plain text—Protects against accidental problems only
  - ✓ Message Digest 5 (MD5)—Protects against accidental and intentional problems

Peer Authentication

- **Multiple keys supported**
  - ✓ **Key lifetimes based on time of day**
  - ✓ **Only first valid key sent with each packet**
- **Supported in: BGP, IS-IS, OSPF, RIPv2, and EIGRP(11.2(4)F)**
- **Syntax differs depending on routing protocol**

OSPF Peer Authentication

Cisco.com

- **OSPF area authentication**
  - ✓ **Two types**
    - **Simple password**
    - **Message Digest (MD5)**

ip ospf authentication-key *key* (this goes under the specific interface)
area *area-id* **authentication** (this goes under "router ospf <process-id>")

ip ospf message-digest-key *keyid* **md5** *key* (used under the interface)
area *area-id* **authentication message-digest** (used under "router ospf <process-id>")

OSPF and ISIS Authentication Example

- **OSPF**

```
interface ethernet1
ip address 10.1.1.1 255.255.255.0
ip ospf message-digest-key 100
    md5 cisco
!
router ospf 1
network 10.1.1.0 0.0.0.255 area 0
area 0 authentication message-
    digest
```

- **ISIS**

```
interface ethernet0
ip address 10.1.1.1
    255.255.255.0
ip router isis
isis password cisco level-2
```

BGP Peer Authentication

Cisco.com

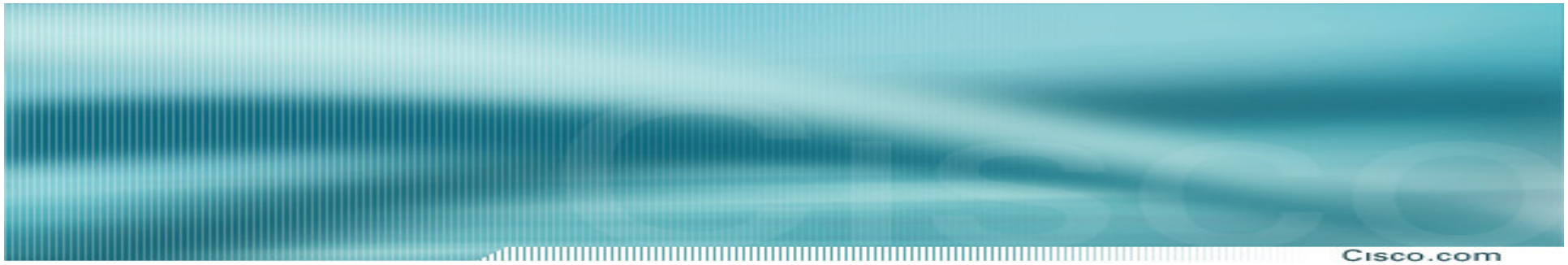
```
router bgp 200
 no synchronization
 neighbor 4.1.2.1 remote-as 300
 neighbor 4.1.2.1 description Link to Excalabur
 neighbor 4.1.2.1 send-community
 neighbor 4.1.2.1 version 4
 neighbor 4.1.2.1 soft-reconfiguration inbound
 neighbor 4.1.2.1 route-map Community1 out
 neighbor 4.1.2.1 password 7 cisco
```

BGP Peer Authentication

- **Works per neighbor or for an entire peer-group**
- **Two routers with password mis-match:**
  - ✓ %TCP-6-BADAUTH: Invalid MD5 digest from [peer's IP address]:11004 to [local router's IP address]:179
- **One router has a password and the other does not:**
  - ✓ %TCP-6-BADAUTH: No MD5 digest from [peer's IP address]:11003 to [local router's IP address]:179

BGP MD5's Problem

- BGP MD5 is not deemed to be operationally deployable.
- Any password needs to be regularly changed. This is a core security fundamental.
- BGP MD5 requires two parties to meet at the same time and make the change.
- If the change happens while there is BGP traffic and the keys mis-match, then the TCP session will be dropped.
  - ✓ Dropping the TCP session drop the eBGP session.
  - ✓ Dropping the eBGP session drops packet forwarding.
- IETF community is looking for a resolution.



BGP BCPs That Help Build Security Resistance

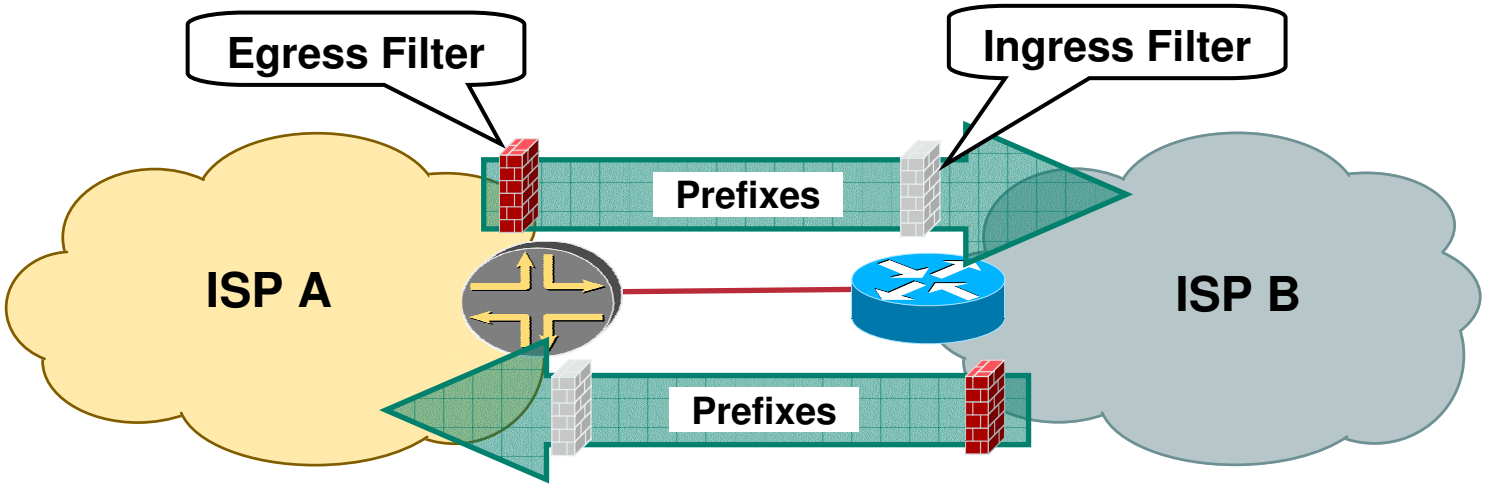
BGP BCPs That Help Build Security Resistance

- Peering Fundamentals
- Maximum Prefix Tracking
- BGP Damping
- BGP Log Neighbor Change

BGP Peering Fundamentals

- BGP Peering assumes that something could go wrong with the policy filters between the neighboring routers.
- Filters are all created to mutually reinforce each other. If one policy filter fails, the policy filter on the neighboring router will take over – providing redundancy to the policy filters.
- This mutually reinforced concept used BGP peering filters are created are also called guarded trust, mutual suspicion, or Murphy Filtering.

Guarded Trust



- ISP A trust ISP B to send X prefixes from the Global Internet Route Table.
- ISP B Creates a egress filter to insure only X prefixes are sent to ISP A.
- ISP A creates a mirror image ingress filter to insure ISP B only sends X prefixes.
- ISP A's ingress filter reinforces ISP B's egress filter.

BGP Maximum Prefix Tracking

- Allow configuration of the maximum number of prefixes a BGP router will receive from a peer
- Two level control
  - ✓ Warning threshold: Log warning message
  - ✓ Maximum: Tear down the BGP peering, manual intervention required to restart

BGP Maximum Prefix Tracking


```
neighbor <x.x.x.x> maximum-prefix <max>
  [<threshold>] [warning-only]
```

- **Threshold is an optional parameter between 1 to 100 percent**
  - ✓ **Specify the percentage of <max> that a warning message will be generated; Default is 75%**
- **Warning-only is an optional keyword which allows log messages to be generated but peering session will not be torn down**

BGP Maximum Prefix Tracking

- **Sample logs:**
  - ✓ **The number of prefixes received from a peer reaches 75% of the maximum configured:**

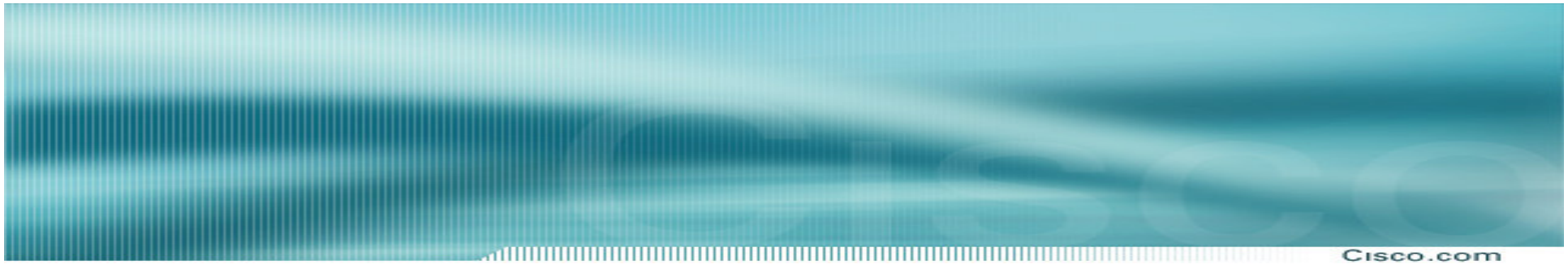
    %BGP-4-MAXPFX: No. of prefix received from 44.1.1.2 reaches 3, max 4

  - ✓ **The number of prefix exceeds the maximum number of prefixes configured:**

    %BGP-3-MAXPFXEXCEED: No. of prefix received from 44.1.1.2: 4 exceed limit 3

BGP Log-Neighbor-Changes

- Log neighbor up/down events, and the reason for the last neighbor peering reset
- In 11.1 CC and 12.0 releases
- Syntax (router subcommand):
  - ✓ [no] log-neighbor-changes
- Typical log messages:
  - ✓ %BGP-6-ADJCHANGE: neighbor x.x.x.x Up
  - ✓ %BGP-6-RESET: neighbor x.x.x.x reset (User reset request)

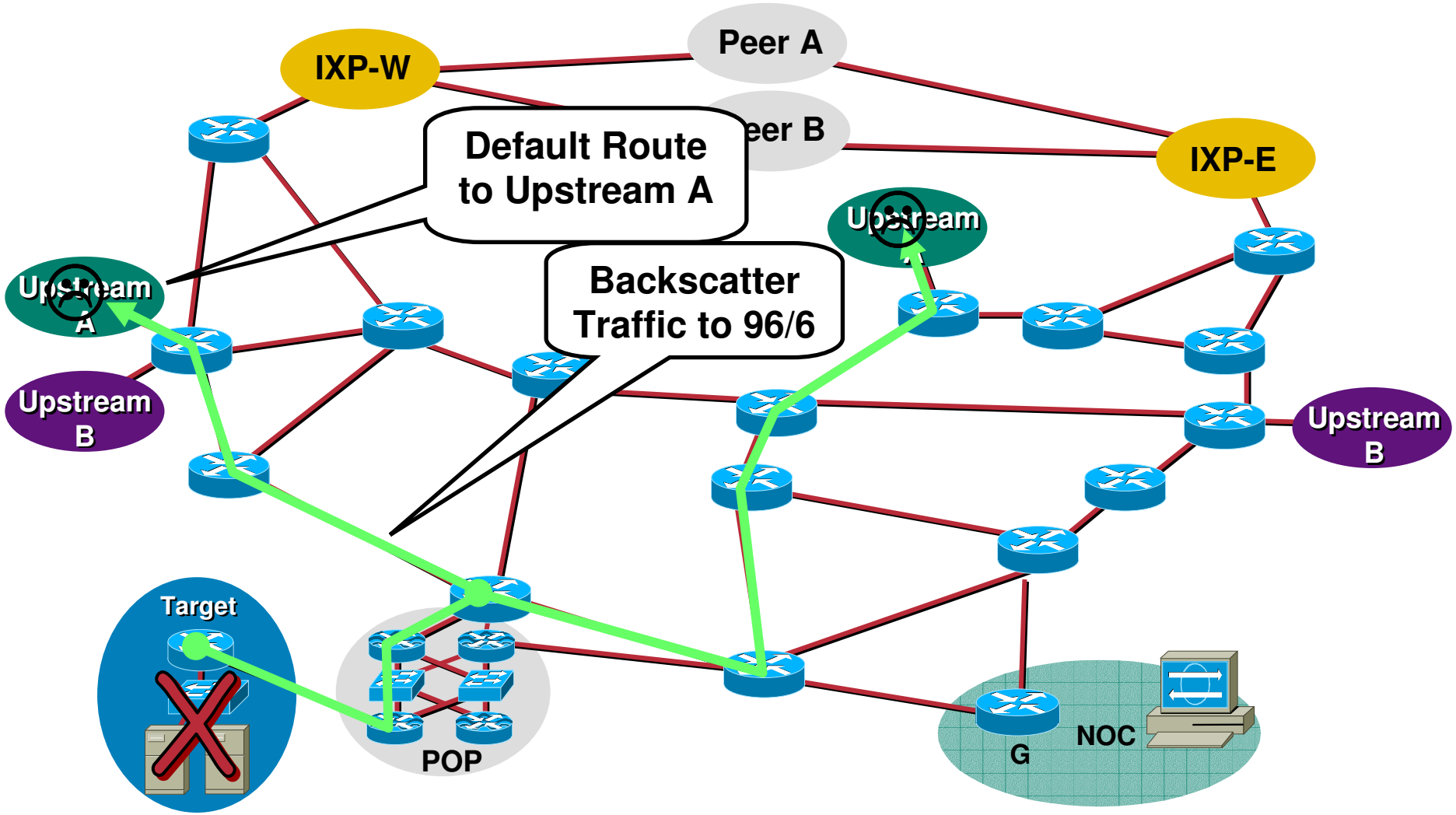


Default Routes, ISPs, and Security

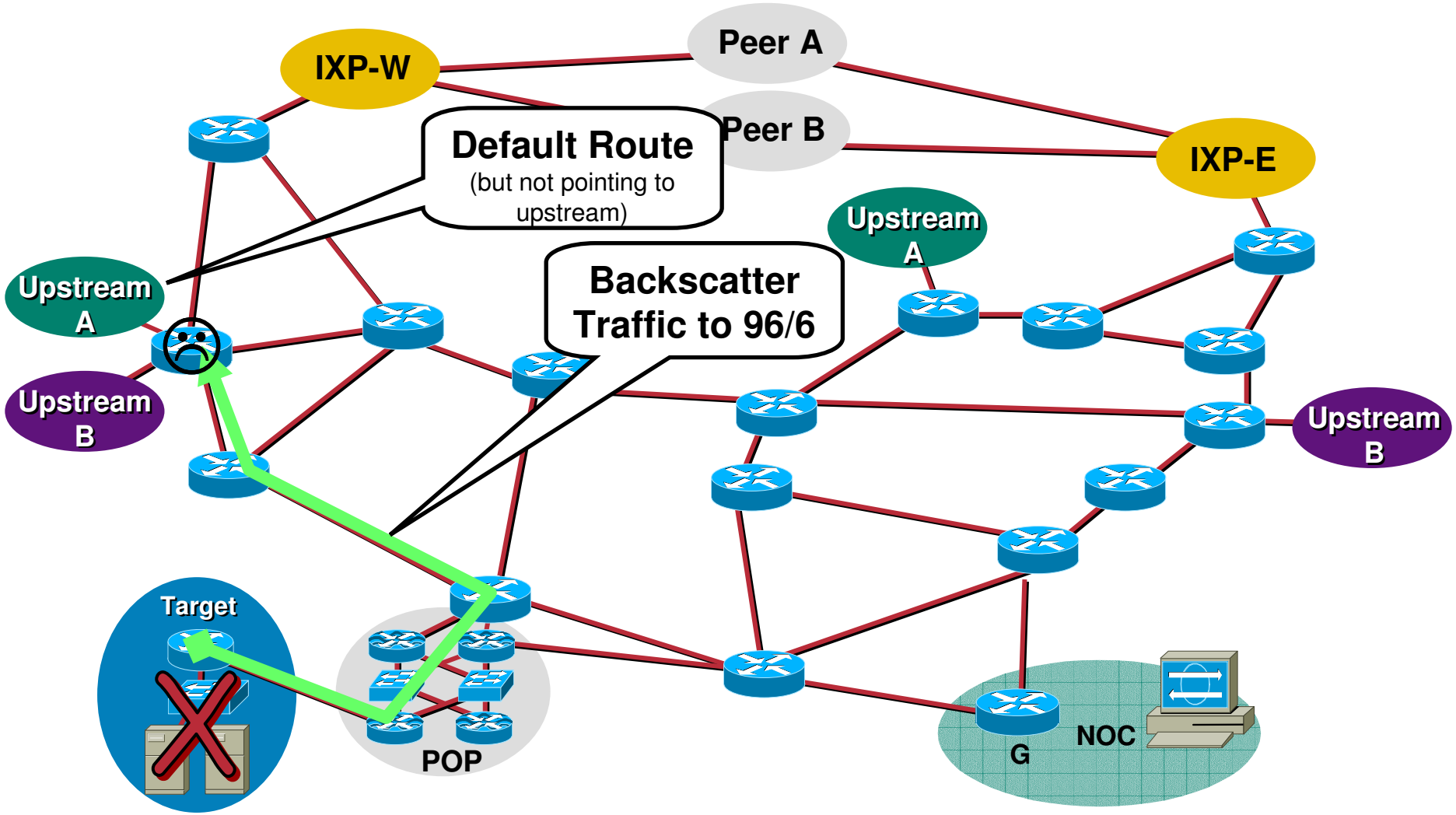
Avoid Default Routes

- ISPs with full BGP feeds should avoid default routes.
- DOS/DDOS attack use spoofed addresses from the un-allocated IPV4 space.
  - ✓ See http://www.iana.org/assignments/ipv4-address-space for the latest macro allocations.
- Backscatter traffic from DOS/DDOS targets need to go somewhere. If there is a default, then this traffic will do to this one router and get dropped.
- Dropping backscatter traffic might overload the router.

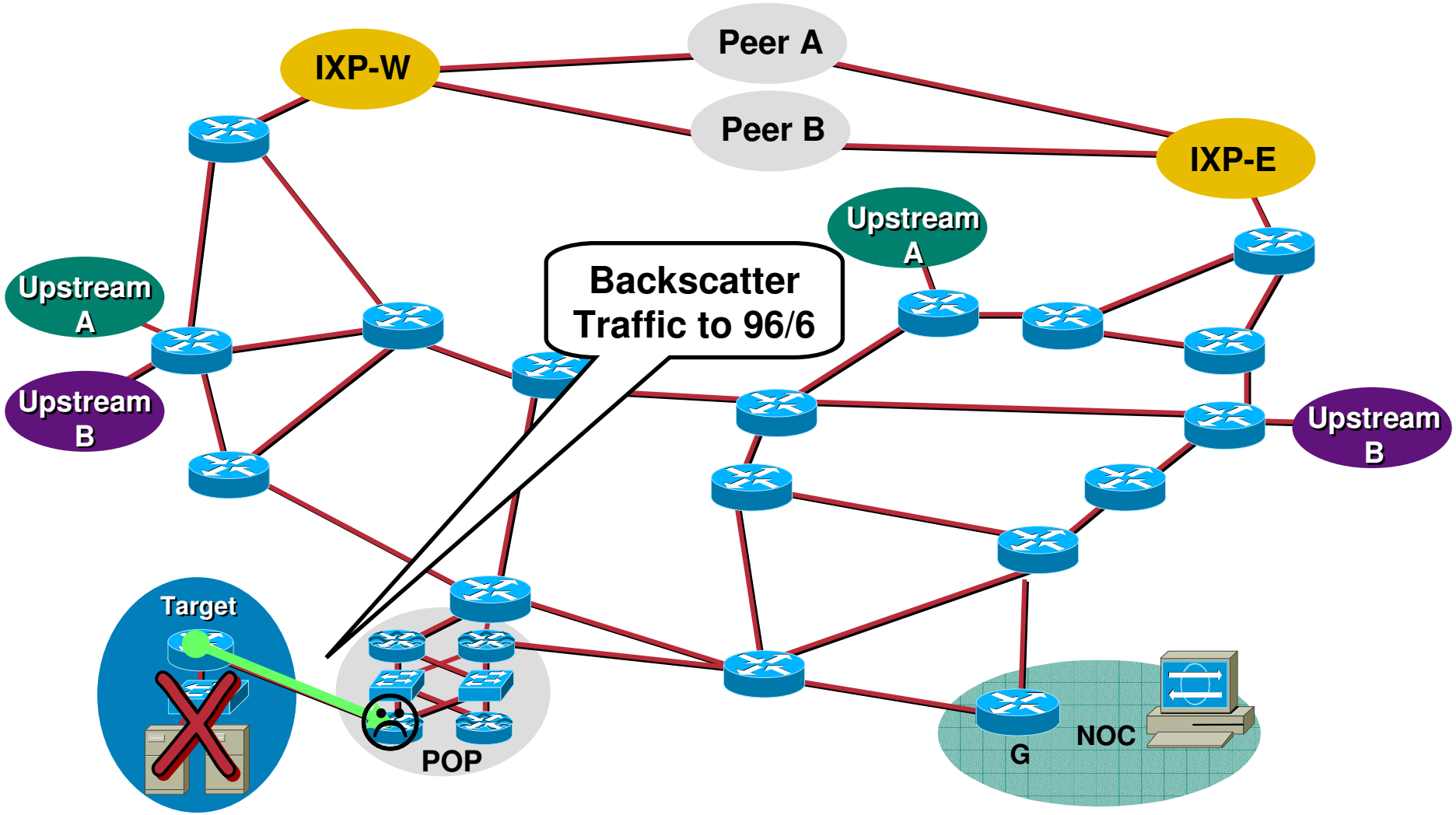
Network with Default Route – Pointing to Upstream A



Network with Default Route – But not Pointing to Upstream



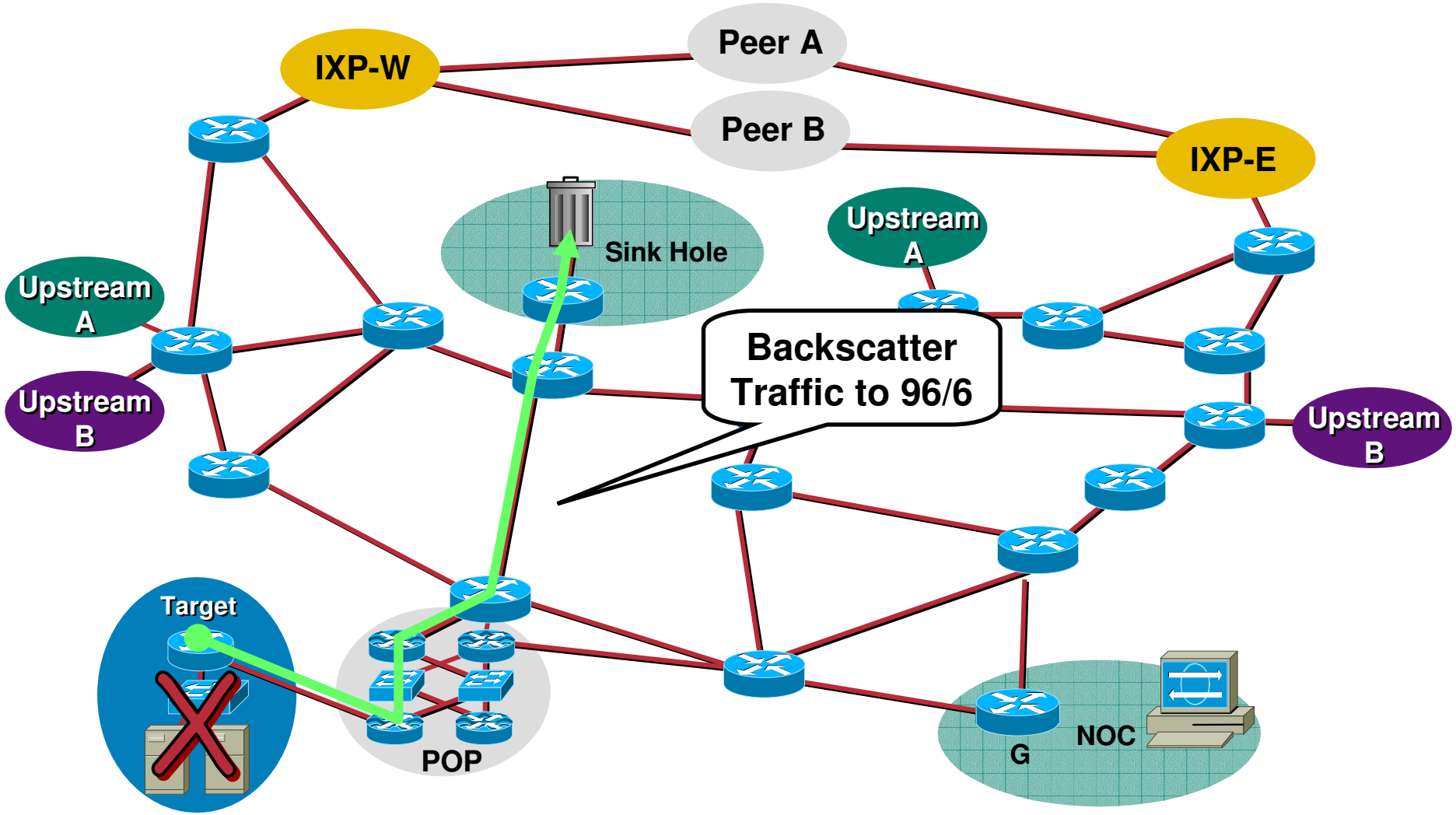
Network with No Default Route

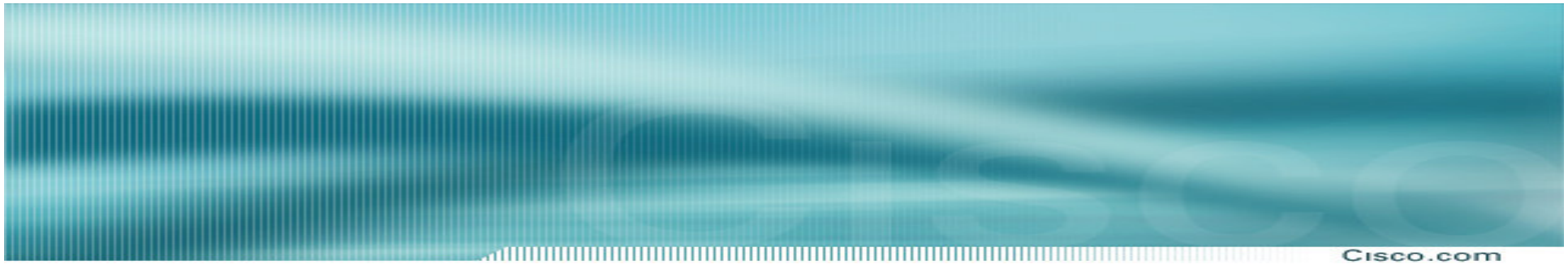


Default Route and ISP Security - Guidance

- **Engineer Default Route with ISP Security as one of the factors.**
  - ✓ **Most just engineer default with routing/forwarding as the only factor**
- **If you need to use default, best to forward it upstream or to a Sink-Hole network engineered for packet drops.**

Default to a Sink-Hole Router/Network





Route Flap Damping

Stabilising the Network

Route Flap Damping

- Route flap
  - ✓Going up and down of path or change in attribute
    - BGP WITHDRAW followed by UPDATE = 1 flap
    - eBGP neighbour peering reset is NOT a flap
  - ✓Ripples through the entire Internet
  - ✓Wastes CPU
- Damping aims to reduce scope of route flap propagation

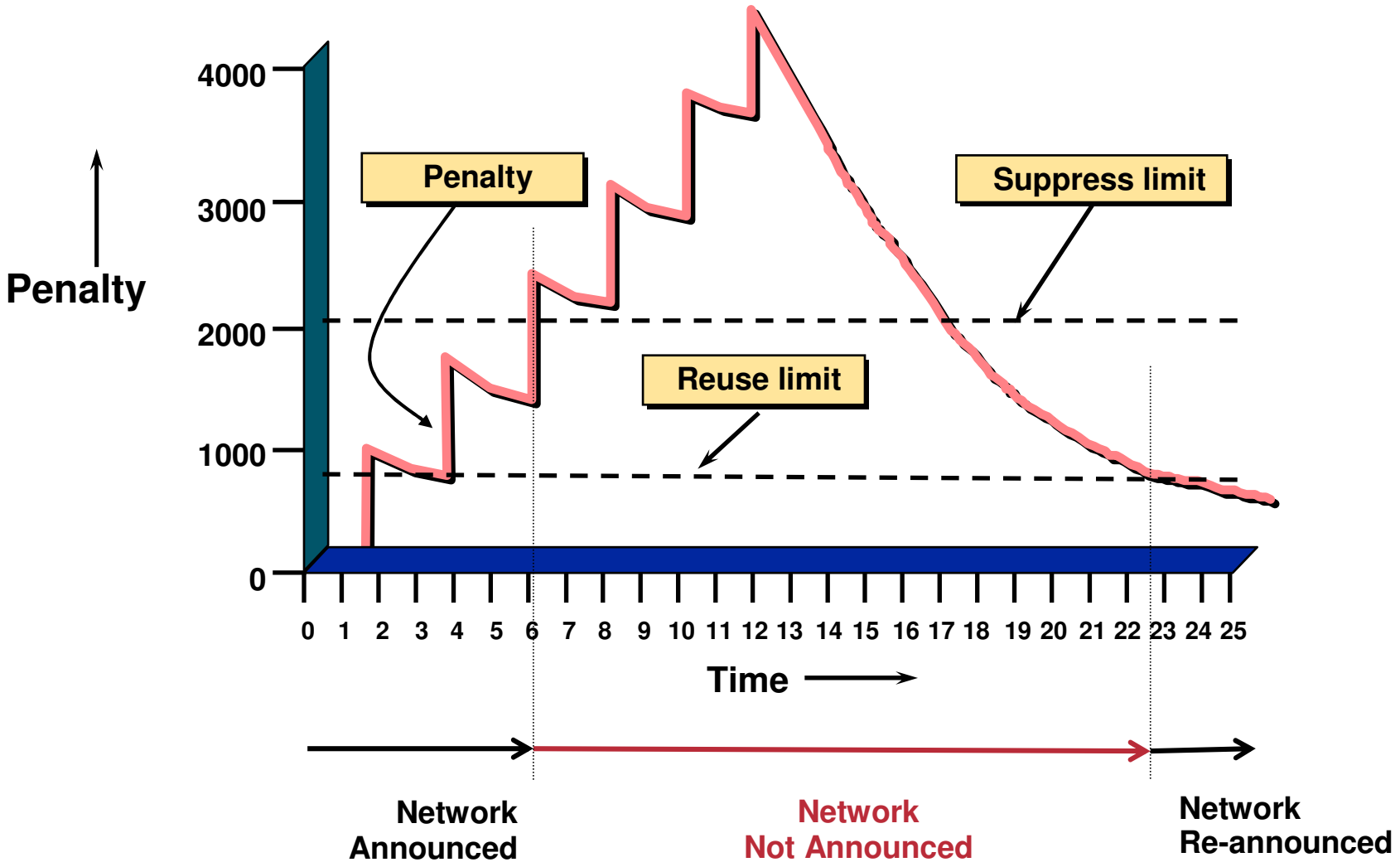
Route Flap Damping (continued)

- **Requirements**
  - ✓ **Fast convergence for normal route changes**
  - ✓ **History predicts future behaviour**
  - ✓ **Suppress oscillating routes**
  - ✓ **Advertise stable routes**
- **Documented in RFC2439**

Operation

- **Add penalty (1000) for each flap**
  - ✓**Change in attribute gets penalty of 500**
- **Exponentially decay penalty**
  - •**half life determines decay rate**
- **Penalty above suppress-limit**
  - •**do not advertise route to BGP peers**
- **Penalty decayed below reuse-limit**
  - •**re-advertise route to BGP peers**
  - •**penalty reset to zero when it is half of reuse-limit**

Operation



Operation

- Only applied to inbound announcements from eBGP peers
- Alternate paths still usable
- Controlled by:
  - ✓Half-life (default 15 minutes)
  - ✓reuse-limit (default 750)
  - ✓suppress-limit (default 2000)
  - ✓maximum suppress time (default 60 minutes)

Configuration

Fixed damping

```
router bgp 100
 bgp dampening [<half-life> <reuse-value> <suppress-
  penalty> <maximum suppress time>]
```

Selective and variable damping

```
 bgp dampening [route-map <name>]
```

Variable damping

- recommendations for ISPs
- http://www.ripe.net/docs/ripe-229.html

Operation

- Care required when setting parameters
- Penalty must be less than reuse-limit at the maximum suppress time
- Maximum suppress time and half life must allow penalty to be larger than suppress limit

Maths!

- **Maximum value of penalty is**

$$\text{max-penalty} = \text{reuse-limit} \times 2^{\left(\frac{\text{max-suppress-time}}{\text{half-life}}\right)}$$

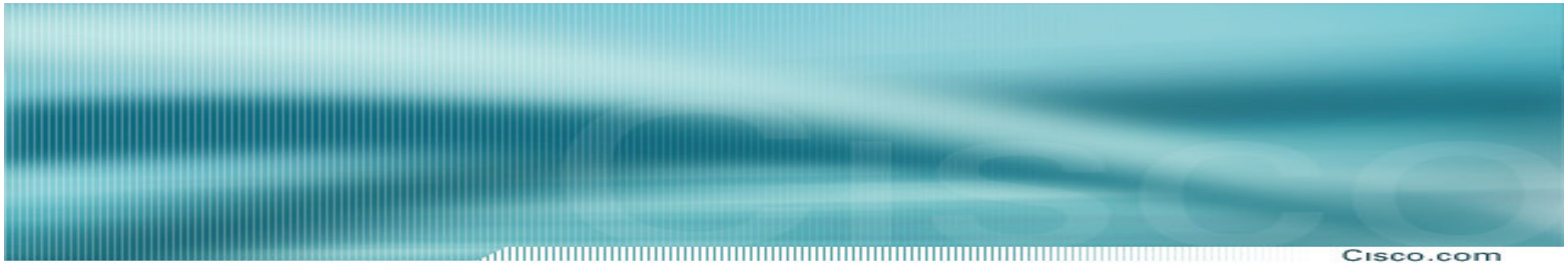
- **Always make sure that suppress-limit is LESS than max-penalty otherwise there will be no flap damping**

Configuration

- **Examples - ✘**
  - ✓**bgp dampening 30 750 3000 60**

    reuse-limit of 750 means maximum possible penalty is 3000 – no prefixes suppressed as penalty cannot exceed suppress-limit
- **Examples - ✓**
  - ✓**bgp dampening 30 2000 3000 60**

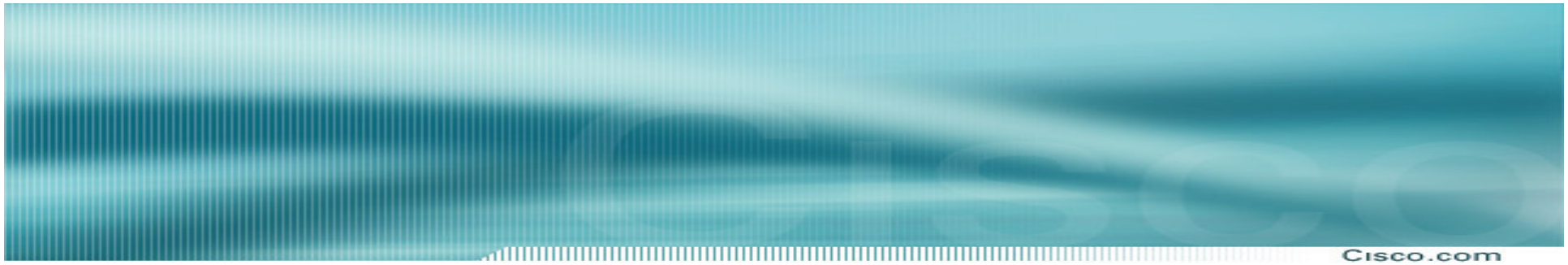
    reuse-limit of 2000 means maximum possible penalty is 8000 – suppress limit is easily reached



Prefix Filters

Agenda

- How to Prefix Filter
- Where to Prefix Filter
- What to Prefix Filter
- Detailed Filterfing
  - ✓Prefix Filter on Customers
  - ✓Egress Filter to Peers
  - ✓Ingress Filter from Peers
  - ✓Net Police Route Filtering



How to Prefix Filter?

Ingress and Egress Route Filtering

- Three flavors of route filtering:
  - ✓ Distribute list—Decreasingly used
  - ✓ Prefix list— Widely used
  - ✓ BGP Communities – Used with the other two
- Two filtering techniques:
  - ✓ Explicit Permit (permit then deny any)
  - ✓ Explicit Deny (deny then permit any)

Ingress and Egress Route Filtering

Extended ACL for a BGP Distribute List

```
access-list 150 deny ip host 0.0.0.0 any
access-list 150 deny ip 10.0.0.0 0.255.255.255 255.0.0.0 0.255.255.255
access-list 150 deny ip 127.0.0.0 0.255.255.255 255.0.0.0 0.255.255.255
access-list 150 deny ip 169.254.0.0 0.0.255.255 255.255.0.0 0.0.255.255
access-list 150 deny ip 172.16.0.0 0.15.255.255 255.240.0.0 0.15.255.255
access-list 150 deny ip 192.0.2.0 0.0.0.255 255.255.255.0 0.0.0.255
access-list 150 deny ip 192.168.0.0 0.0.255.255 255.255.0.0 0.0.255.255
access-list 150 deny ip 224.0.0.0 31.255.255.255 224.0.0.0 31.255.255.255
access-list 150 permit ip any any
```

Ingress and Egress Route Filtering

BGP with Distribute List Flavor of Route Filtering

```
router bgp 200
no synchronization
bgp dampening
 neighbor 220.220.4.1 remote-as 210
 neighbor 220.220.4.1 version 4
 neighbor 220.220.4.1 distribute-list 150 in
 neighbor 220.220.4.1 distribute-list 150 out
 neighbor 222.222.8.1 remote-as 220
 neighbor 222.222.8.1 version 4
 neighbor 222.222.8.1 distribute-list 150 in
 neighbor 222.222.8.1 distribute-list 150 out
 no auto-summary
!
```

Ingress and Egress Route Filtering

Prefix-List for a for a BGP Prefix List

```
ip prefix-list rfc1918-dsua deny   0.0.0.0/8 le 32
ip prefix-list rfc1918-dsua deny   10.0.0.0/8 le 32
ip prefix-list rfc1918-dsua deny   127.0.0.0/8 le 32
ip prefix-list rfc1918-dsua deny   169.254.0.0/16 le 32
ip prefix-list rfc1918-dsua deny   172.16.0.0/12 le 32
ip prefix-list rfc1918-dsua deny   192.0.2.0.0/24 le 32
ip prefix-list rfc1918-dsua deny   192.168.0.0/16 le 32
ip prefix-list rfc1918-dsua deny   224.0.0.0/3 le 32
ip prefix-list rfc1918-dsua permit 0.0.0.0/0 le 32
```

Ingress and Egress Route Filtering

BGP with Prefix-List Flavor of Route Filtering

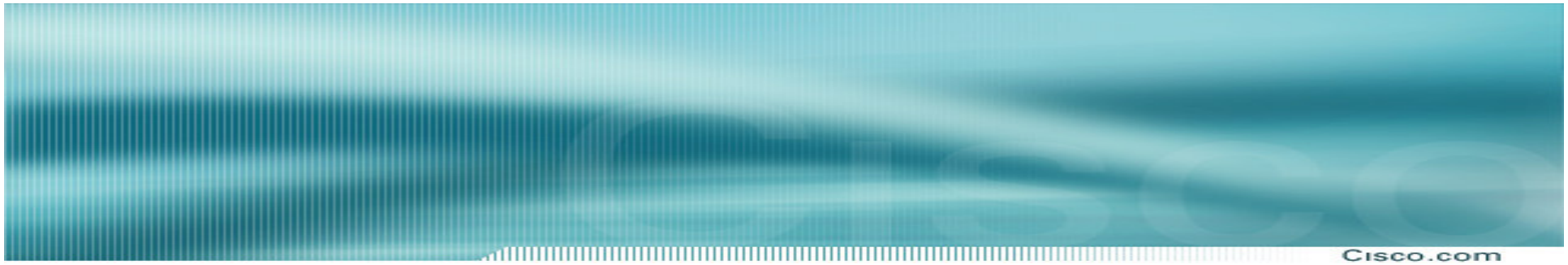
```
router bgp 200
 no synchronization
 bgp dampening
 neighbor 220.220.4.1 remote-as 210
 neighbor 220.220.4.1 version 4
 neighbor 220.220.4.1 prefix-list rfc1918-dsua in
 neighbor 220.220.4.1 prefix-list rfc1918-dsua out
 neighbor 222.222.8.1 remote-as 220
 neighbor 222.222.8.1 version 4
 neighbor 222.222.8.1 prefix-list rfc1918-dsua in
 neighbor 222.222.8.1 prefix-list rfc1918-dsua out
 no auto-summary
!
```

Two Filtering Techniques

- There are two fundamental ways to create a drop list:
  - ✓ Explicit Permit. Permit specific networks, then deny everything else.
  - ✓ Explicit Deny. Deny specific networks, then permit everything else.
- Which technique depends on the situation and the filtering objectives.
- Applying the wrong technique to the filtering objective will usually cause large drop filters.

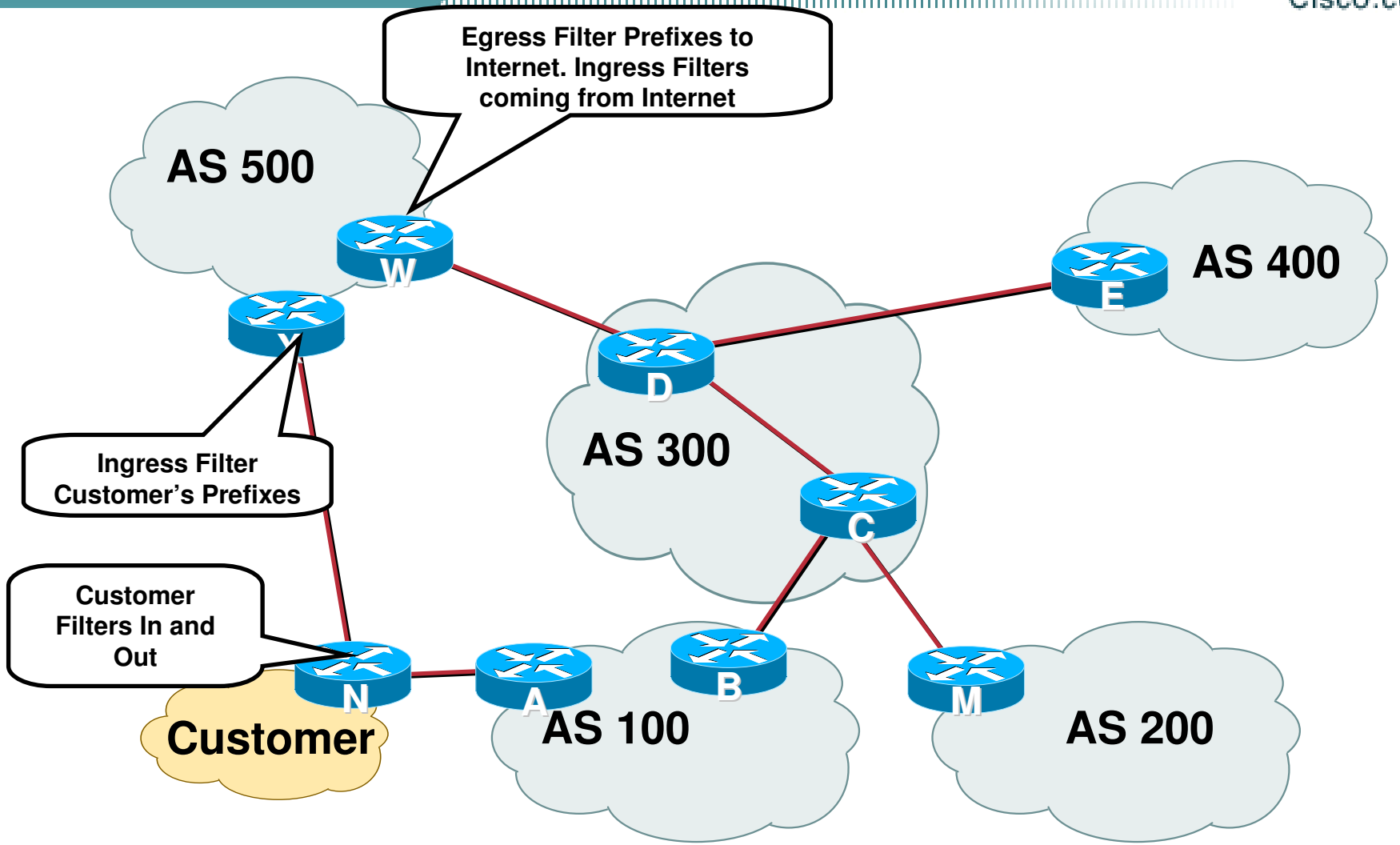
Ideal Customer Ingress/Egress Route Filtering ….

- **Assume Murphy's Law of Networking**
  - ✓ **Do not assume the ingress route filter on the customer will always work. Murphy's Law assumes that it will break at the worst possible time.**
  - ✓ **Customer Egress filtering to your peers/upstreams will have two layers of protection.**



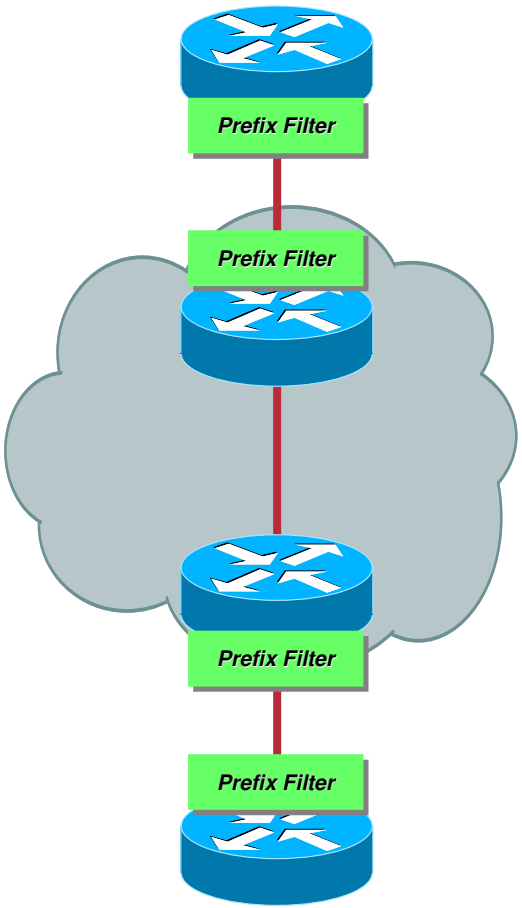
Where to Prefix Filter?

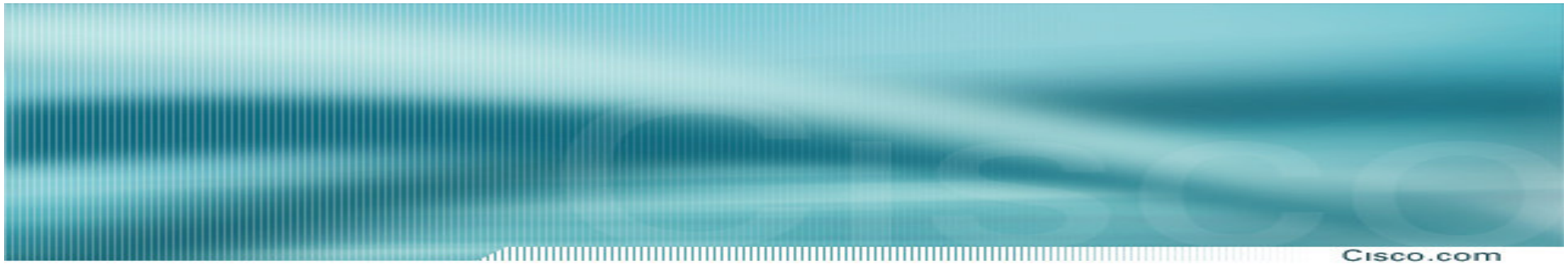
Where to Prefix Filter?



Where to Prefix Filter?

- Customer's Ingress/Egress
- ISP Ingress on Customer (may Egress to Customer)
- ISP Egress to Peer and Ingress from Peer
- Peer Ingress from ISP and Egress to ISP





What to Prefix Filter?

Documenting Special Use Addresses (DUSA) and Bogons

Documenting Special Use Addresses (DUSA)

- There are routes that should NOT be routed on the Internet
  - ✓ RFC 1918 and “Martian” networks (DUSA)
  - ✓ 127.0.0.0/8 and multicast blocks (DUSA)
  - ✓ See Bill Manning’s ID for background information:

    ftp://ftp.ietf.org/internet-drafts/draft-manning-dsua-07.txt
- BGP should have filters applied so that these routes are not advertised to or propagated through the Internet

Documenting Special Use Addresses (DUSA)

- Quick review
  - ✓ 0.0.0.0/8 and 0.0.0.0/32—Default and broadcast
  - ✓ 127.0.0.0/8—Host loopback
  - ✓ 192.0.2.0/24—TEST-NET for documentation
  - ✓ 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16—RFC 1918 private addresses
  - ✓ 169.254.0.0/16—End node auto-config for DHCP

Documenting Special Use Addresses (DUSA)


```
ip prefix-list rfc1918-dsua deny 0.0.0.0/8 le 32
ip prefix-list rfc1918-dsua deny 10.0.0.0/8 le 32
ip prefix-list rfc1918-dsua deny 127.0.0.0/8 le 32
ip prefix-list rfc1918-dsua deny 169.254.0.0/16 le 32
ip prefix-list rfc1918-dsua deny 172.16.0.0/12 le 32
ip prefix-list rfc1918-dsua deny 192.0.2.0/24 le 32
ip prefix-list rfc1918-dsua deny 192.168.0.0/16 le 32
ip prefix-list rfc1918-dsua deny 224.0.0.0/3 le 32
ip prefix-list rfc1918-dsua deny 0.0.0.0/0 ge 25
ip prefix-list rfc1918-dsua permit 0.0.0.0/0 le 32
```

Bogons

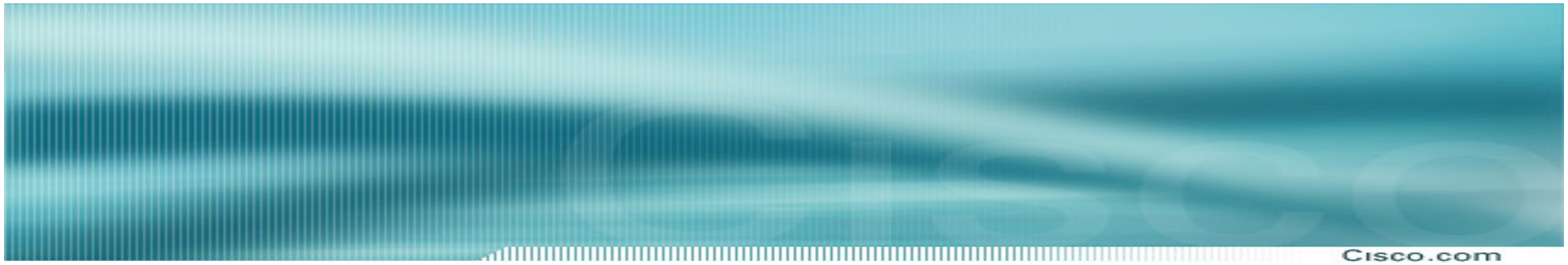
- IANA has reserved several blocks of IPv4 that have yet to be allocated to a RIR:
  - ✓ http://www.iana.org/assignments/ipv4-address-space
- These blocks of IPv4 addresses should never be advertised into the global Internet Route Table.
- Filters should be applied on the AS border for all inbound and outbound advertisements.

Ingress Prefix Filter Template

- “It is hard to build the list.” --- “OK, we’ll build the community a template. Next excuse.”
- Bogon List by CYMRU Bogon Team
  - ✓ http://www.cymru.com/Bogons/
  - ✓ Starting point for putting together the Bogon Filtering.
  - ✓ Supplies up to date templates for Cisco and Juniper

Ingress Prefix Filter Template

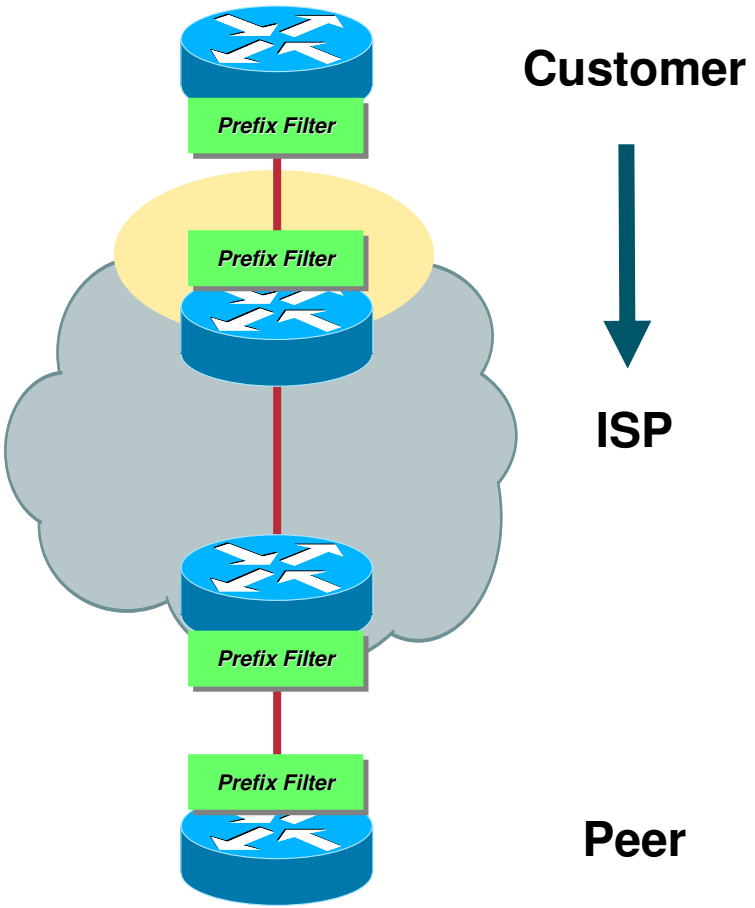
- **Cisco Template by Barry Greene**
  - ✓ **ftp://ftp-eng.cisco.com/cons/isp/security/Ingress-Prefix-Filter-Templates/**
- **Juniper Template by Steven Gill**
  - ✓ **http://www.qorbit.net/documents.html**



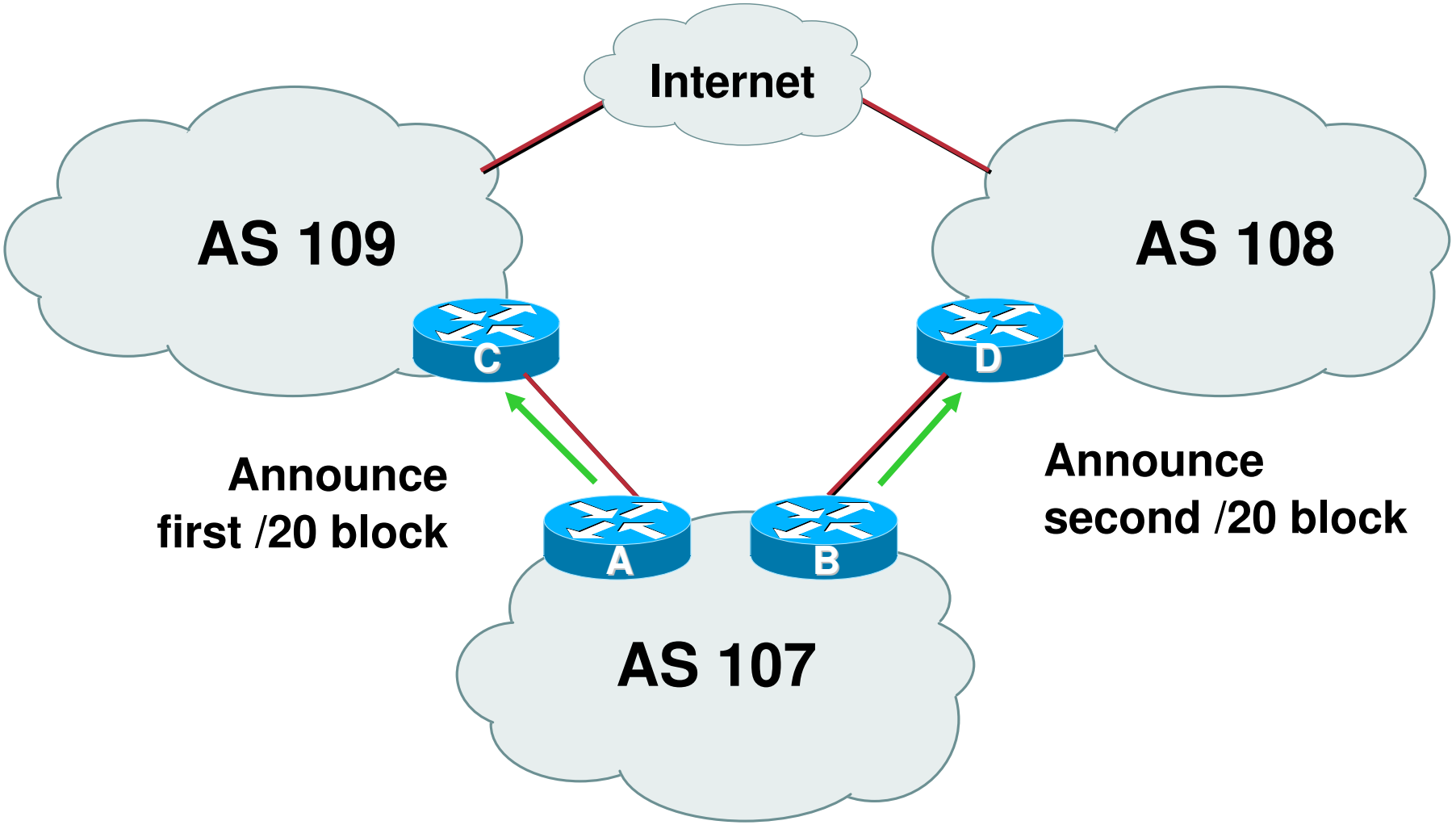
Prefix Filters on Customers

Prefix Filters on Customers

- **Prefix filter all routes from your customers!**



BGP with Customer Infers Multihoming



Receiving Customer Prefixes

- ISPs should only accept prefixes which have been assigned or allocated to their downstream peer/customer.
- For example
  - ✓ Downstream has 220.50.0.0/20 block
  - ✓ Should only announce this to peers
  - ✓ Peers should only accept this from them
  - ✓ Explicitly permit prefixes from other ISPs (i.e. multihomed to two or more ISPS).

Receiving Customer Prefixes

- **Configuration example on upstream:**

```
router bgp 100
  neighbor 222.222.10.1 remote-as 101
  neighbor 222.222.10.1 prefix-list customer in
  !
ip prefix-list customer permit 220.50.0.0/2
ip prefix-list customer deny 0.0.0.0/0 le 32
```

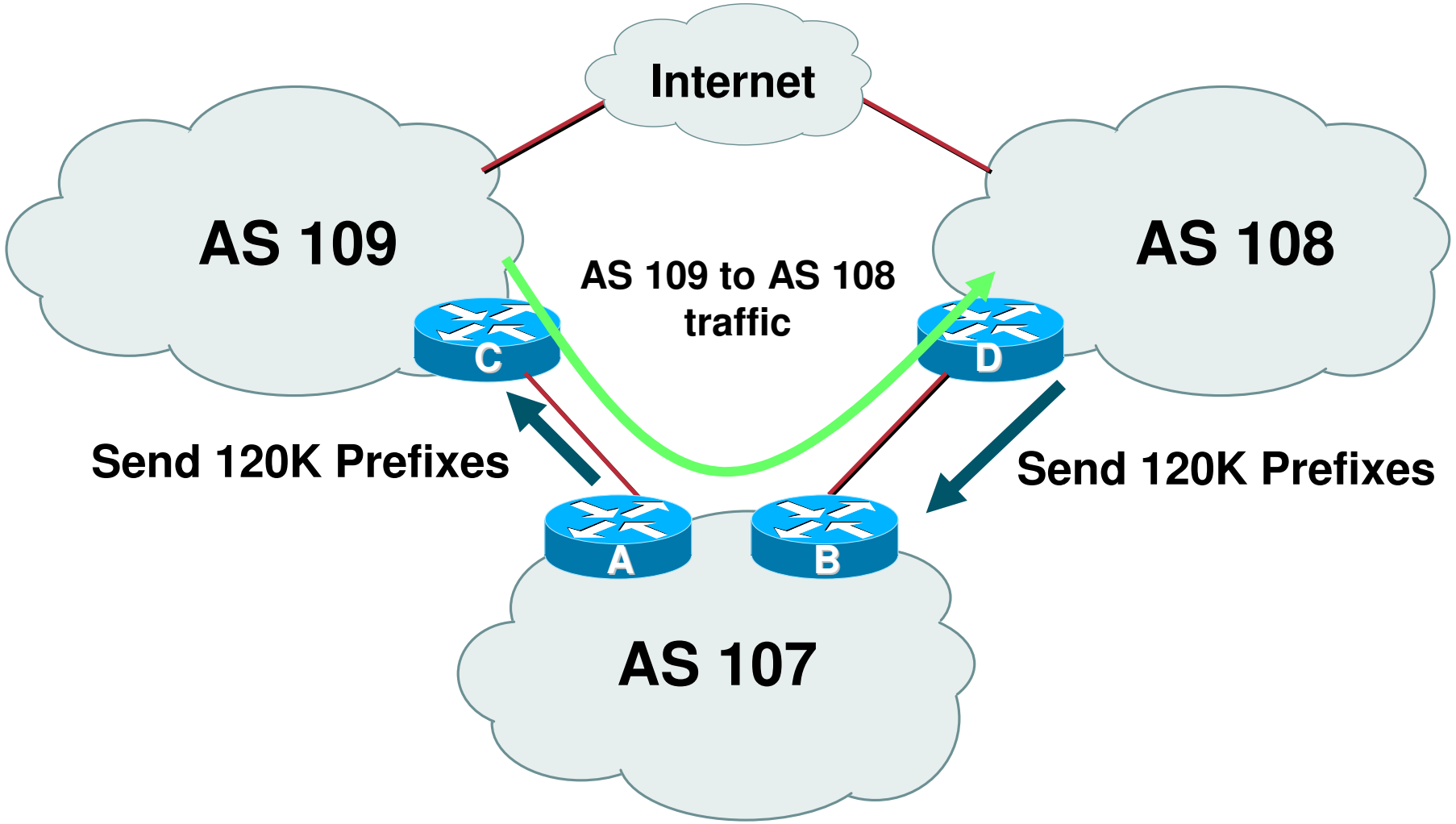
Excuses – Why providers are not prefix filtering customers.

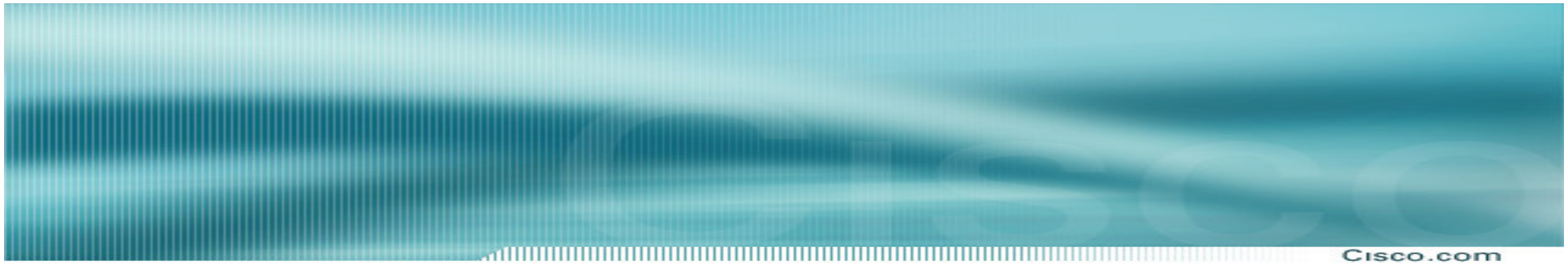
- “Some of my customers are multihomed, so they want to advertise more specifics.”
- “These are down stream ISPs, so their advertisements will change.”

What if you do not filter your customer?

- **Not filtering your customers puts your network at risk to:**
  - ✓ **Bogon Prefix Insertion (sucks down backscatter)**
  - ✓ **Un-Authorized Route Insertion (sucks down traffic)**
  - ✓ **Re-advertise other ISP's routes (customer's T1 becomes the peering link).**

What if you do not filter your customer?

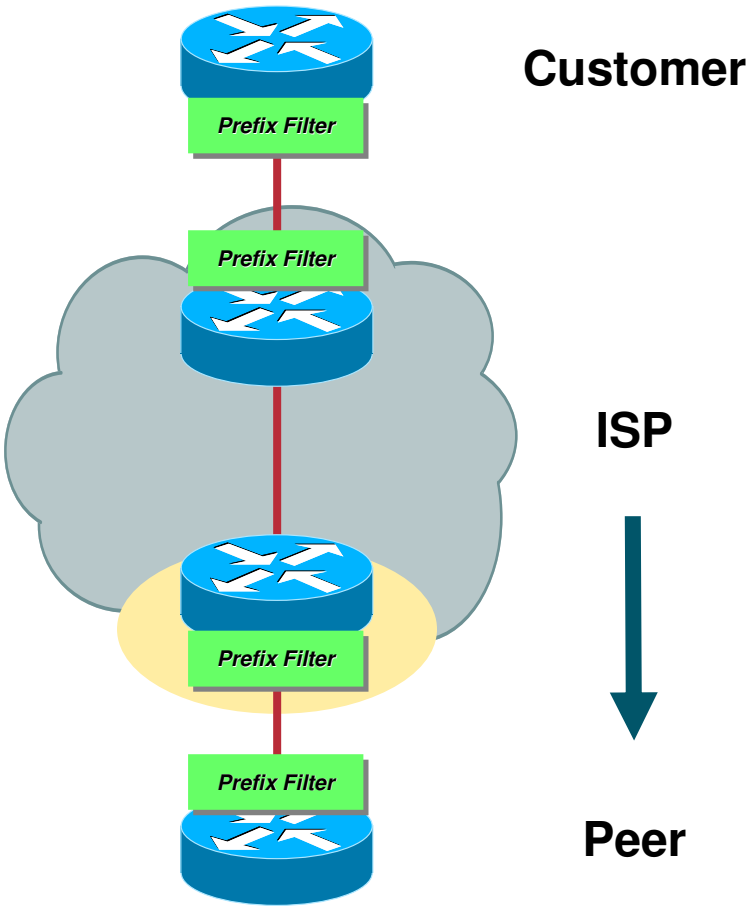




Prefixes to Peers

Prefixes to Peers

- **Prefix filter all routes to your peers!**



Prefixes to Peers

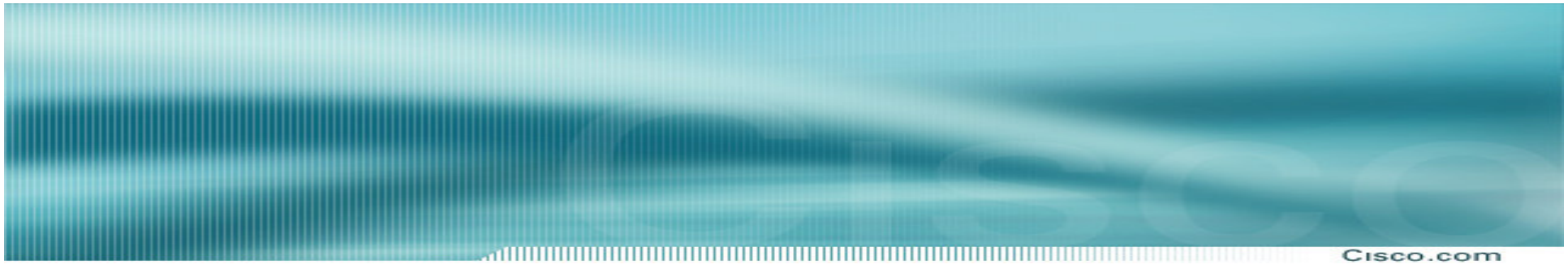
- What do you send to the Internet?
  - ✓ Your prefixes.
  - ✓ More specific customers prefixes (customers who are multihoming)
- What do you *not* send to the Internet?
  - ✓ DUSA Prefixes – assume junk will leak into your iBGP.
  - ✓ Bogons – assume garbage will leak into your iBGP.
  - ✓ Lower Prefix Boundary – Unless absolutely necessary, Do no allow anything in the /25 - /32 range.

Egress Filter to ISP Peers - Issues

- **The egress filter list can grow to be very large:**
  - ✓ **More specifics for customers.**
  - ✓ **Specific blocks from other ISPs**

Policy Questions

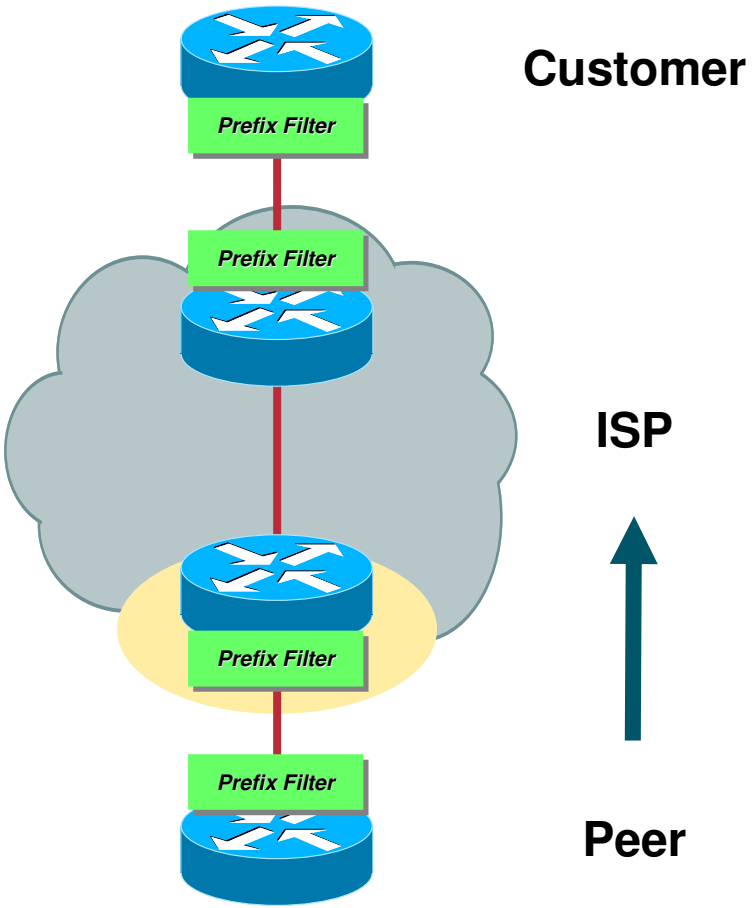
- Will you allow customers to announce IP prefixes from other ISPs?
- Will the customer be required to tell you these prefixes?
- Will you advertise these prefixes back to the ISP?
- Will you advertise these prefixes to the entire Internet?



Ingress Prefix Filtering from Peers

Prefixes from Peers

- **Prefix filter all routes from your peers!**



Ingress Routes from Peers or Upstream

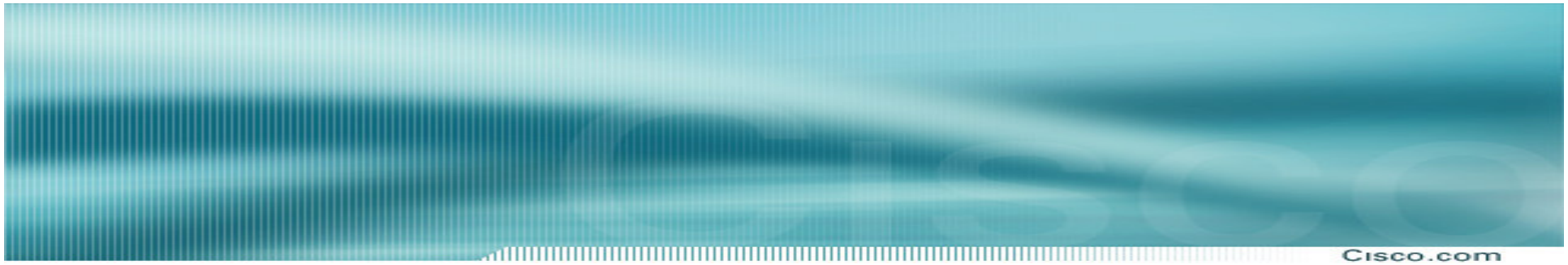
- Ingress Routes from Peers and/or the Upstream ISP are the nets of the Internet.
- Ideally, the peering policy should be specific so that exact filters can be put in place.
  - ✓ Dynamic nature of the peering makes it hard to maintain specific route filters.

Receiving Prefixes from Upstream & Peers (ideal case)

- ✓Don't accept RFC1918 etc prefixes
- ✓Don't accept your own prefix
- ✓Don't accept default (unless you need it)
- ✓Don't accept prefixes longer than /24
- ✓Don't accept prefixes on IXPs your whom you have membership
- ✓Consider *Net Police* Filtering

Receiving Prefixes — Cisco IOS


```
router bgp 100
 network 221.10.0.0 mask 255.255.224.0
 neighbor 221.5.7.1 remote-as 101
 neighbor 221.5.7.1 prefix-list in-filter in
!
ip prefix-list in-filter deny 0.0.0.0/0                ! Block default
ip prefix-list in-filter deny 0.0.0.0/8 le 32
ip prefix-list in-filter deny 10.0.0.0/8 le 32
ip prefix-list in-filter deny 127.0.0.0/8 le 32
ip prefix-list in-filter deny 169.254.0.0/16 le 32
ip prefix-list in-filter deny 172.16.0.0/12 le 32
ip prefix-list in-filter deny 192.0.2.0/24 le 32
ip prefix-list in-filter deny 192.168.0.0/16 le 32
ip prefix-list in-filter deny 221.10.0.0/19 le 32      ! Block local prefix
ip prefix-list in-filter deny 224.0.0.0/3 le 32
ip prefix-list in-filter deny 0.0.0.0/0 ge 25          ! Block prefixes >/24
ip prefix-list in-filter permit 0.0.0.0/0 le 32
```



Net Police Route Filtering

“Net Police” Route Filtering

- ***Net Police*** **route filtering describes ingress peering filtering that only allows the *minimum practical allocation* from a RIR (Regional Internet Registry).**
  - ✓ **So if APNIC’s minimum practical allocation is a /20, then the Net Police filter will only allow a /8 to a /20/. Any prefix larger than a /20 (I.e. a /21) will get dropped by the filter.**
- **Net Police Filtering has two effects:**
  - ✓ **Reduces the number of prefixes in an ISP’s RIB.**
  - ✓ **Protects the ISP from *Garbage in Garbage* out problems/incidents on the Net.**

“Net Police” Route Filtering

- **Three Techniques:**
  - ✓ **Permit only prefixes on the RIR’s *minimum practical allocation*s.**
  - ✓ **Permit prefixes allocated by the RIRs with a lower boundary set by the ISP (i.e. /24 vs a /20).**
  - ✓ **Deny prefixes that have not been allocated by IANA.**

Net Police Filter Technique #1

- **Permit Only Allocated IPv4 Blocks**
- **Need to check with each of the RIR's for details on which networks they are allocating from and what the specific *minimum practical allocation* for each block.**
  - ✓ **RIRs are announcing changes to the Internet Operations Aliases.**
- **ARIN - http://www.arin.net/statistics/index.html#cidr**
- **RIPE - http://www.ripe.net/ripe/docs/smallest-alloc-sizes.html**
- **APNIC - http://www.apnic.net/db/min-alloc.html**

Technique #1 Net Police Prefix List

(check for update)


```
!! APNIC
ip prefix-list FILTER permit 61.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 202.0.0.0/7 ge 9 le 20
ip prefix-list FILTER permit 210.0.0.0/7 ge 9 le 20
ip prefix-list FILTER permit 218.0.0.0/7 ge 9 le 20
!! ARIN
ip prefix-list FILTER permit 63.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 64.0.0.0/7 ge 9 le 20
ip prefix-list FILTER permit 66.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 199.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 200.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 204.0.0.0/6 ge 9 le 20
ip prefix-list FILTER permit 208.0.0.0/7 ge 9 le 20
ip prefix-list FILTER permit 216.0.0.0/8 ge 9 le 20
!! RIPE NCC
ip prefix-list FILTER permit 62.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 80.0.0.0/7 ge 9 le 20
ip prefix-list FILTER permit 193.0.0.0/8 ge 9 le 20
ip prefix-list FILTER permit 194.0.0.0/7 ge 9 le 20
ip prefix-list FILTER permit 212.0.0.0/7 ge 9 le 20
```

Net Police Prefix List Deployment Issues

- Objective – protect the network from ISPs who won't and don't aggregate
- Impacts *more specific* style multihoming
- Impacts regions where domestic backbone is unavailable or costs $$$ compared with international bandwidth
- Maintenance Overhead – requires updating when RIRs start allocating from new address blocks
- Understand the Consequences!

Technique #2 Net Police Prefix List Alternative

- Permit Only Allocated IPv4 Blocks
- Move the minimal allocation prefix to a /24
- Most Operators agree that blocks longer than /24 should not be seen on the Net.
- This minimizes some of the operational impact to customer multihoming.

Technique #2 Net Police Prefix List Alternative (check for update)

Cisco.com

```
!! APNIC
ip prefix-list FILTER permit 61.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 202.0.0.0/7 ge 9 le 24
ip prefix-list FILTER permit 210.0.0.0/7 ge 9 le 24
ip prefix-list FILTER permit 218.0.0.0/7 ge 9 le 24
!! ARIN
ip prefix-list FILTER permit 63.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 64.0.0.0/7 ge 9 le 24
ip prefix-list FILTER permit 66.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 199.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 200.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 204.0.0.0/6 ge 9 le 24
ip prefix-list FILTER permit 208.0.0.0/7 ge 9 le 24
ip prefix-list FILTER permit 216.0.0.0/8 ge 9 le 24
!! RIPE NCC
ip prefix-list FILTER permit 62.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 80.0.0.0/7 ge 9 le 24
ip prefix-list FILTER permit 193.0.0.0/8 ge 9 le 24
ip prefix-list FILTER permit 194.0.0.0/7 ge 9 le 24
ip prefix-list FILTER permit 212.0.0.0/7 ge 9 le 24
```

Net Police Filter – Technique #3

- Deny All Non-Allocated IPv4 Blocks
- Uses IANA's master allocation table to deny any block that is not yet allocated to one of the RIRs.

  http://www.iana.org/assignments/ipv4-address-space
- These non-allocated addresses are also referred to as *bogons*.

Technique #3 Net Police Prefix List

(check for update)


```
ip prefix-list DUSA-Bogons description Bogon networks we
      won't accept.
ip prefix-list DUSA-Bogons seq 5 deny 0.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 10 deny 1.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 15 deny 2.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 20 deny 5.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 25 deny 7.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 30 deny 10.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 35 deny 23.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 40 deny 27.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 45 deny 31.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 50 deny 36.0.0.0/7 le 32
ip prefix-list DUSA-Bogons seq 60 deny 39.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 65 deny 41.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 70 deny 42.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 75 deny 49.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 80 deny 50.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 85 deny 58.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 90 deny 59.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 95 deny 60.0.0.0/8 le 32
Cont …
```

```
ip prefix-list DUSA-Bogons seq 110 deny 69.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 115 deny 70.0.0.0/7 le 32
ip prefix-list DUSA-Bogons seq 125 deny 72.0.0.0/5 le 32
ip prefix-list DUSA-Bogons seq 165 deny 82.0.0.0/7 le 32
ip prefix-list DUSA-Bogons seq 175 deny 84.0.0.0/6 le 32
ip prefix-list DUSA-Bogons seq 195 deny 88.0.0.0/5 le 32
ip prefix-list DUSA-Bogons seq 235 deny 96.0.0.0/6 le 32
ip prefix-list DUSA-Bogons seq 255 deny 100.0.0.0/6 le 32
ip prefix-list DUSA-Bogons seq 275 deny 104.0.0.0/5 le 32
ip prefix-list DUSA-Bogons seq 320 deny 112.0.0.0/4 le 32
ip prefix-list DUSA-Bogons seq 395 deny 169.254.0.0/16 le 32
ip prefix-list DUSA-Bogons seq 400 deny 172.16.0.0/12 le 32
ip prefix-list DUSA-Bogons seq 405 deny 192.0.2.0/24 le 32
ip prefix-list DUSA-Bogons seq 410 deny 192.168.0.0/16 le 32
ip prefix-list DUSA-Bogons seq 415 deny 197.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 420 deny 201.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 435 deny 221.0.0.0/8 le 32
ip prefix-list DUSA-Bogons seq 440 deny 222.0.0.0/7 le 32
ip prefix-list DUSA-Bogons seq 450 deny 224.0.0.0/3 le 32
```

Net Police Filter – Technique #3

- **One plus to Technique #2 is the ability to track the prefix-list drops.**
  - ✓ **Allow you to keep track of the noise and/or problems on the Net.**

.

```
GSR-1#sh ip prefix-list det
Prefix-list with the last deletion/insertion: DUSA-Bogons
ip prefix-list DUSA-Bogons:
   Description: Bogon networks we won't accept.
   count: 33, range entries: 33, sequences: 5 - 450, refcount: 4
.
seq 125 deny 72.0.0.0/5 le 32 (hit count: 1, refcount: 1)
```

Bottom Line

- **Net Police filtering effectively protects networks from garbage in garbage out problems on the Net.**
  - ✓ **ISPs using Net Police filters did not have the AS 7007 or 129/8 incidents effect their network.**
- **While Net Police filters are controversial, their use as a security tool has been proven.**

Looking for examples?

- **Cisco Template by Barry Greene**
  - ✓ **ftp://ftp-eng.cisco.com/cons/isp/security/Ingress-Prefix-Filter-Templates/**
- **Juniper Template by Steven Gill**
  - ✓ **http://www.qorbit.net/documents.html**