IOS Essentials

Essential Features every ISP should Consider

Cisco Systems

Overview

- **IOS Software and Router Management**
- **General Features**
- **Securing the Router**
- **Securing the Network**
- **Routing Configuration Guidelines**

Which IOS version?

- **Platforms**
  - **GSR, 7500 series, 7200 series**
- **Recommended release is 12.0S train**
  - **Current version is 12.0(10)S1**
  - **Available on CCO**
- **Has all of latest ISP supported features**

Which IOS version?

- **Platforms**

  4x00, 3600, 2600 and 2500 series

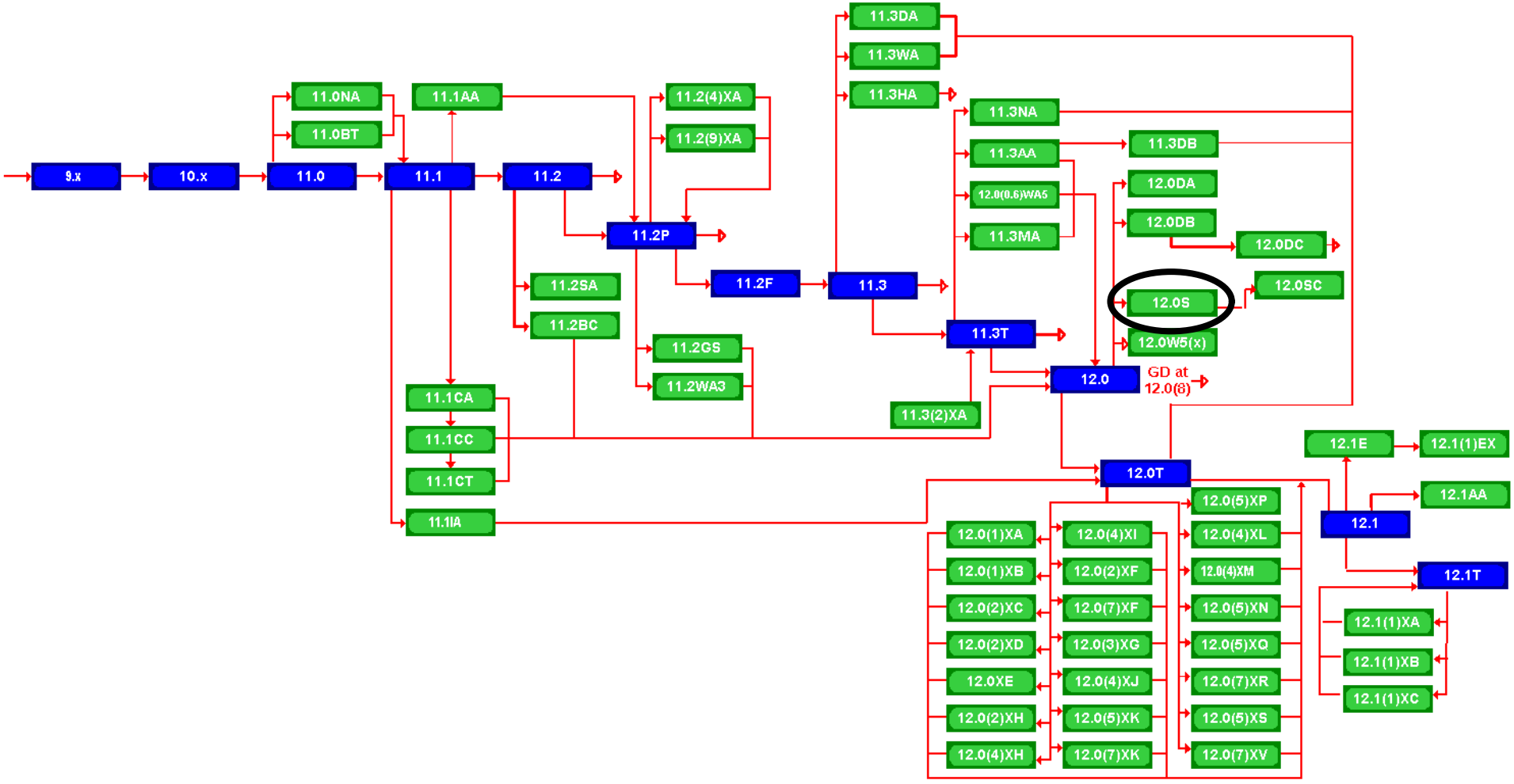
- **Recommended release is the 12.0 mainline train**

  Current version is 12.0(11)

  Has many of the features found in 11.1CC, 11.2P and 11.3T

  Available on CCO

Cisco IOS Roadmap



http://www.cisco.com/warp/public/620/roadmap.shtml

IOS Software Management Flash Memory

- **Good practice is to have at least two distinct flash memory volumes**
  - **allows backup image(s)**
  - **back out path in case of upgrade problems**
- **Partition the built-in flash**

```
partition flash 2 8 8
```

- **Install a PCMCIA flash card in external slot(s)**

IOS Software Management Flash Memory

- **Ensure that there is a configured backup to selected IOS image**

  backup image is previous “good” image

```
boot system flash slot0:rsp-pv-mz.120-10.S
boot system flash slot1:rsp-pv-mz.111-32.CC
boot system flash
```

  which means “boot quoted image from slot0:. If it isn’t there, boot the quoted image in slot1:. If that isn’t there, try the first image available in flash

IOS Software Management System Memory

- **Good practice is to maximise router memory**
  - **allows for the rapidly growing Internet**
- **128Mbytes needed for full Internet routing table**
  - **will (just) work with 64Mbytes, but BGP inefficient**
- **Recognised that equipment works best when “left alone”**

IOS Software Management When to Upgrade

- **Upgrades needed when:**
  - **bug fixes released**
  - **new hardware support**
  - **new software features required**
- **Otherwise:**

If it isn't broken, don't fix it!

Configuration Management

- **Backup NVRAM configuration off the router:**
  - **write configuration to TFTP server**
  - **TFTP server files kept under revision control**
  - **router configuration built from master database**
- **Allows rapid recovery in case of emergency**

Larger Configurations

- **Compress Configuration**

  Used when configuration required is larger than configuration memory (NVRAM) available.

  ```
  service compress-config
  ```

- **FLASH or remote server**

  Used when NVRAM compression is not enough

Use detailed logging

- **Off load logging information to a logging server.**
- **Use the full detailed logging features to keep exact details of the activities.**

```
service timestamps debug datetime msec localtime show-timezone
service timestamps log datetime msec localtime show-timezone
logging buffered 16384
logging trap debugging
logging facility local7
logging 169.223.32.1
logging 169.223.35.8
logging source-interface loopback0
```

Network Time Protocol

- If you want to cross compare logs, you need to synchronize the time on all the devices.
- Use NTP

  from external time source

    Upstream ISP, Internet, GPS, atomic clock

  from internal time source

    router can act as stratum 1 time source

Network Time Protocol

- **Set timezone**

```
clock timezone <name> [+/-hours [mins]]
```

- **Router as source**

```
ntp master 1
```

- **External time source (master)**

```
ntp server a.b.c.d
```

- **External time source (equivalent)**

```
ntp peer e.f.g.h
```

Network Time Protocol

- **Example Configuration:**

```
clock timezone SST 8
ntp update-calendar
ntp source loopback0
ntp server <other time source>
ntp peer <other time source>
ntp peer <other time source>
```

SNMP

- **Remove any SNMP commands if SNMP is not going to be used.**
- **If SNMP is going to be used:**

```
access-list 98 permit 169.223.1.1
access-list 98 deny   any
snmp-server community 5nmc02m RO 98
snmp-server trap-source Loopback0
snmp-server trap-authentication
snmp-server host 169.223.1.1 5nmc02m
```

HTTP Server

- **HTTP Server in IOS from 11.1CC and 12.0S**

  router configuration via web interface

- **Disable if not going to be used:**

```
no ip http server
```

- **Configure securely if going to be used:**

```
ip http server
ip http port 8765
ip http authentication aaa
ip http access-class <1-99>
```

Core Dumps

- Cisco routers have a *core dump* feature that will allow ISPs to transfer a copy of the core dump to a specific FTP server.
- Set up a FTP account on the server the router will send the core dump to.
- The server should NOT be a public server

  use filters and secure accounts

  locate in NOC with network operations staff access only

Core Dumps

- **Example configuration:**

```
ip ftp username cisco
ip ftp password 7 045802150C2E
ip ftp source-interface loopback 0
exception protocol ftp
exception dump 169.223.32.1
```

General Features

Command Line Interface Features

- **Some Convenient Editing Keys**

TAB	**command completion**
arrow keys	**scroll history buffer**
ctrl A	**beginning of line**
ctrl E	**end of line**
ctrl K	**delete all chars to end of line**
ctrl X	**delete all chars to beginning of line**
ctrl W	**delete word to left of cursor**
esc B	**back one word**
esc F	**forward one word**

Command Line Interface Features

- **CLI now has string searches**

  ```
  show configuration | [begin|include|exclude] <regexp>
  ```

- **Pager "--more--" now has string searches**

  ```
  /<regexp>, -<regexp>, +<regexp>
  ```

- **"More" command has string searches**

  ```
  more <filename> | [begin|include|exclude] <regexp>
  ```

Interface Configuration

- **“ip unnumbered”**
  - **no need for an IP address on point-to-point links**
  - **keeps IGP small**
- **“description”**
  - **customer name, circuit id, cable number, etc**
  - **on-line documentation!**
- **“bandwidth”**
  - **used by IGP**
  - **documentation!**

Interface Configuration - Example

- **ISP router**

```
!
interface loopback 0
description Loopback interface on GW2 Router
ip address 215.17.3.1 255.255.255.255
!
interface Serial 5/0
description 128K HDLC link to Galaxy
Publications Ltd [galpub1] WT50314E  R5-0
bandwidth 128
ip unnumbered loopback 0
!
ip route 215.34.10.0 255.255.252.0 Serial 5/0
```

- **Customer router**

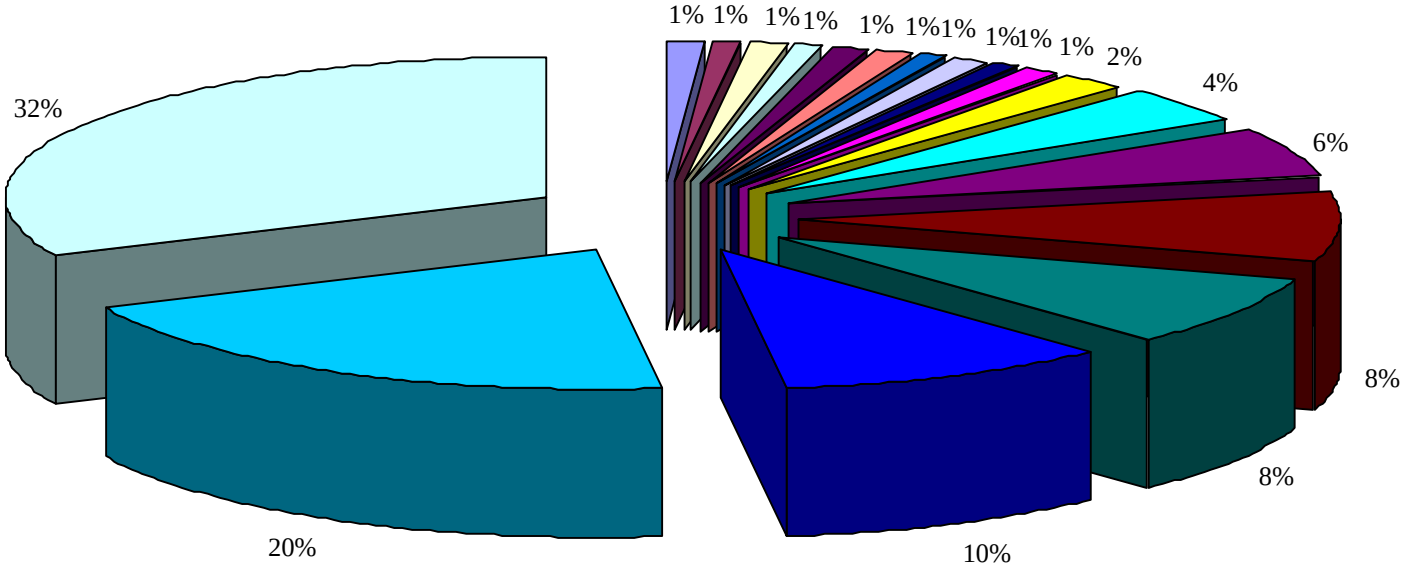
```
!
interface Ethernet 0
description Galaxy Publications LAN
ip address 215.34.10.1 255.255.252.0
!
interface Serial 0
description 128K HDLC link to Galaxy
Internet Inc WT50314E  C0
bandwidth 128
ip unnumbered ethernet 0
!
ip route 0.0.0.0 0.0.0.0 Serial 0
```

NetFlow

- Provides network administrators with "packet flow" information
- Allows:
  - security monitoring
  - network management and planning
  - customer billing
  - traffic flow analysis
- Available from 11.1CC for 7x00 and 12.0 for remaining router platforms

Netflow - Capacity Planning

Public Routers 1 , 2, 3 Month of September Outbound Traffic



WEC	WebTV	ABSNET	AOL	Compuserve
SURAnet	IBM	OARNet	NIH	PacBell Internet Service
JHU	C&W	UMD	AT&T	BBN
Erols	Digex	Other		

NetFlow

- **Configuration example:**

```
interface serial 5/0
 ip route-cache flow
```

- **If CEF not configured, NetFlow enhances existing switching path**
- **If CEF configured, NetFlow becomes a flow information gatherer**

NetFlow

- **Information export:**

  router to collector system

  ```
  ip flow-export version 5 [origin-as|peer-as]
  ip flow-export destination x.x.x.x <udp-port>
  ```

- **Flow aggregation (new in 12.0S):**

  router sends aggregate records to collector system

  ```
  ip flow-aggregation cache as|prefix|dest|source|proto
    enabled
    export destination x.x.x.x <udp-port>
  ```

NetFlow

- **Sample Output on router:**

```
Beta-7200-2>sh ip cache flow
IP packet size distribution (17093 total packets):
   1-32   64   96  128  160  192  224  256  288  320  352  384  416  448  480
   .000 .735 .088 .054 .000 .000 .008 .046 .054 .000 .009 .000 .000 .000 .000

    512  544  576 1024 1536 2048 2560 3072 3584 4096 4608
   .000 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000

IP Flow Switching Cache, 1257536 bytes
  3 active, 15549 inactive, 12992 added
  210043 ager polls, 0 flow alloc failures
  last clearing of statistics never
Protocol         Total    Flows   Packets Bytes  Packets Active(Sec) Idle(Sec)
--------         Flows     /Sec     /Flow  /Pkt     /Sec     /Flow     /Flow
TCP-Telnet          35      0.0        80    41      0.0      14.5      12.7
UDP-DNS             20      0.0         1    67      0.0       0.0      15.3
UDP-NTP           1223      0.0         1    76      0.0       0.0      15.5
UDP-other        11709      0.0         1    87      0.0       0.1      15.5
ICMP                 2      0.0         1    56      0.0       0.0      15.2
Total:           12989      0.0         1    78      0.0       0.1      15.4

SrcIf         SrcIPaddress    DstIf         DstIPaddress    Pr SrcP DstP  Pkts
Et1/1         144.254.153.10  Null          144.254.153.127 11 008A 008A     1
Et1/1         144.254.153.112 Null          255.255.255.255 11 0208 0208     1
Et1/1         144.254.153.50  Local         144.254.153.51  06 701D 0017    63
```

Using DNS

- Map names to addresses
- Descriptive names

```
ip domain-name
ip name-server
```

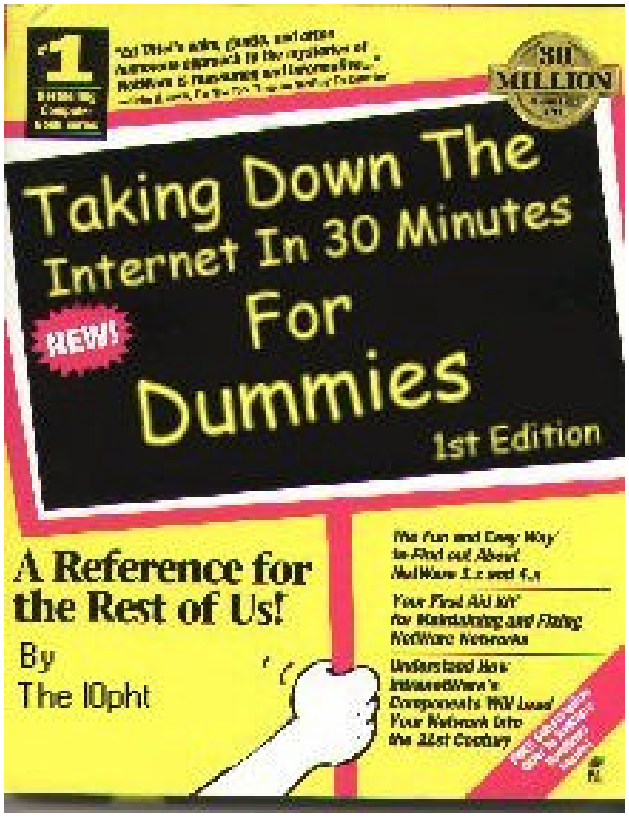
- Sample trace through network:

```
4:Received echo from sj-wall-2.cisco.com [198.92.1.138] in 440 msec.
5:Received echo from barrnet-gw.cisco.com [192.31.7.37] in 335 msec.
6:Received echo from paloalto-cr1.bbnplanet.net [131.119.26.9] in 335 msec.
7:Received echo from paloalto-br2.bbnplanet.net [131.119.0.194] in 327 msec.
8:Received echo from core6-hssi6-0.SanFrancisco.mci.net [206.157.77.21] in 468 msec.
9:Received echo from bordercore1-loopback.Washington.mci.net [166.48.36.1] in 454 msec.
10:Received 48 bytes from www.getit.org [199.233.200.55] in 466 msec
```

Securing the Router

ISP Security

- **ISPs need to:**
  - **Protect themselves**
  - **Help protect their customers from the Internet**
  - **Protect the Internet from their customers**



ISP Security

ISP Security

- **Where to start …..**

  Cisco Internet Security Advisories

  http://www.cisco.com/warp/public/779/largeent/security/advisory.html

  Cisco IOS documentation for 12.0

  http://www.cisco.com/univercd/data/doc/software/11_2/2cbook.html

  RFC2196 (Site Security Handbook)

  Networker's Security Sessions

Global Services You Turn OFF

- **Some services turned on by default, should be turned off to save memory and prevent security breaches/attacks**

```
no service finger
no service pad
no service udp-small-servers
no service tcp-small-servers
no ip bootp server
```

Interface Services You Turn OFF

- **Some IP features are great for Campus LANs, but do not make sense on a ISP backbone.**
- **All interfaces on an ISP's backbone router should have the follow as a *default:***

```
no ip redirects
no ip directed-broadcast
no ip proxy-arp
```

Cisco Discovery Protocol

- Lets network administrators discover neighbouring Cisco equipment, model numbers and software versions
- Should not be needed on ISP network

```
no cdp run
```

- Should not be activated on any public facing interface: IXP, customer, upstream ISP
- Disable per interface

```
no cdp enable
```

Login Banner

- **Use a good login banner, or nothing at all:**

```
banner login ^
  Authorised access only
  This system is the property of Galactic Internet
  Disconnect IMMEDIATELY if you are not an authorised user!
  Contact noc@net.galaxy +99 876 543210 for help.
^
```

Exec Banner

- **Useful to remind logged in users of local conditions:**

```
banner exec ^
  PLEASE NOTE - THIS ROUTER SHOULD NOT HAVE A DEFAULT ROUTE!
  It is used to connect paying peers. These 'customers'
should not be able to default to us.
  The config for this router is NON-STANDARD
  Contact Network Engineering +99 876 543234 for more info.
^
```

Use Enable Secret

- Encryption '7' on a Cisco is reversible.
- The “enable secret” password encrypted via a one-way algorithm.

```
enable secret <removed>

no enable password

service password-encryption
```

Turn on Nagle

- **Telnet was designed to do one character, one packet dialog.**
- **John Nagle's algorithm (RFC 896) helps alleviate the small-packet problem in TCP.**

```
service nagle
```

ident Feature

- **Identification (ident) support allows you to query a Transmission Control Protocol (TCP) port for identification.**
- **This feature enables an insecure protocol, described in RFC 1413, to report the identity of a client initiating a TCP connection and a host responding to the connection. No attempt is made to protect against unauthorized queries.**

```
ip ident
```

VTY and Console port timeouts

- **Default idle timeout on async ports is 10 minutes 0 seconds**

```
exec-timeout 10 0
```

- **Timeout of 0 means permanent connection**
- **TCP keepalives on incoming network connections**

```
service tcp-keepalives-in
```

VTY security

- **Access to VTYs should be controlled, not left open. Consoles should be used for last resort admin only:**

```
access-list 3 permit 215.17.1.0 0.0.0.255
access-list 3 deny   any
line vty 0 4
 access-class 3 in
 exec-timeout 5 0
 transport input telnet ssh
 transport output none
 transport preferred none
 password 7 045802150C2E
```

VTY Access and SSH

- **Secure Shell Supported as from IOS 12.0S**
- **Obtain, load and run appropriate crypto images on router**
- **Set up SSH on router**

```
Beta7200(config)#crypto key generate rsa
```

- **Add it as input transport**

```
line vty 0 4
  transport input telnet ssh
```

User Authentication

- **Account per user, with passwords**

```
aaa new-model
aaa authentication login neteng local
username joe password 7 1104181051B1
username jim password 7 0317B21895FE
line vty 0 4
 login neteng
 access-class 3 in
```

User Authentication

- **Use distributed authentication system**

  RADIUS (not recommended for system security)

  TACACS+

```
aaa new-model
aaa authentication login default tacacs+ enable
aaa authentication enable default tacacs+ enable
aaa accounting exec start-stop tacacs+
ip tacacs source-interface Loopback0
tacacs-server host 215.17.1.1
tacacs-server key CKr3t#
line vty 0 4
 access-class 3 in
```

User Authentication

TACACS+ Provides a detailed audit trail of what is happening on the network devices.

User-Name	Group-N	cmd	priv-lvl	service	NAS-Portname	task_id	NAS-IP-	reason
bgreene	NOC	enable <cr>	0	shell	tty0	4	210.210.51.224	
bgreene	NOC	exit <cr>	0	shell	tty0	5	210.210.51.224	
bgreene	NOC	no aaa accounting exec Workshop	0	shell	tty0	6	210.210.51.224	
bgreene	NOC	exit <cr>	0	shell	tty0	8	210.210.51.224	
pfs	NOC	enable <cr>	0	shell	tty0	11	210.210.51.224	
pfs	NOC	exit <cr>	0	shell	tty0	12	210.210.51.224	
bgreene	NOC	enable <cr>	0	shell	tty0	14	210.210.51.224	
bgreene	NOC	show accounting <cr>	15	shell	tty0	16	210.210.51.224	
bgreene	NOC	write terminal <cr>	15	shell	tty0	17	210.210.51.224	
bgreene	NOC	configure <cr>	15	shell	tty0	18	210.210.51.224	
bgreene	NOC	exit <cr>	0	shell	tty0	20	210.210.51.224	
bgreene	NOC	write terminal <cr>	15	shell	tty0	21	210.210.51.224	
bgreene	NOC	configure <cr>	15	shell	tty0	22	210.210.51.224	
bgreene	NOC	aaa new-model <cr>	15	shell	tty0	23	210.210.51.224	
bgreene	NOC	aaa authorization commands 0 de	15	shell	tty0	24	210.210.51.224	
bgreene	NOC	exit <cr>	0	shell	tty0	25	210.210.51.224	
bgreene	NOC	ping <cr>	15	shell	tty0	32	210.210.51.224	
bgreene	NOC	show running-config <cr>	15	shell	tty66	35	210.210.51.224	
bgreene	NOC	router ospf 210 <cr>	15	shell	tty66	45	210.210.51.224	
bgreene	NOC	debug ip ospf events <cr>	15	shell	tty66	46	210.210.51.224	

Securing the Network

Ingress and Egress Route Filtering

- **There are routes that should NOT be routed on the Internet.**

  RFC 1918 and "Martian" Networks

  127.0.0.0/8 and Multicast blocks

  See ftp://ftp.ietf.org/internet-drafts/draft-manning-dsua-03.txt for background information

- **BGP should have filters applied so that these routes are not advertised to or propagated through the Internet.**

Ingress and Egress Route Filtering

BGP Configuration

```
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
!
```

Ingress and Egress Route Filtering

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 & Egress Route Filtering

Your customers should not be sending *any* IP packets out to the Internet with a source address other then the address you have allocated to them!

Ingress & Egress Packet Filtering

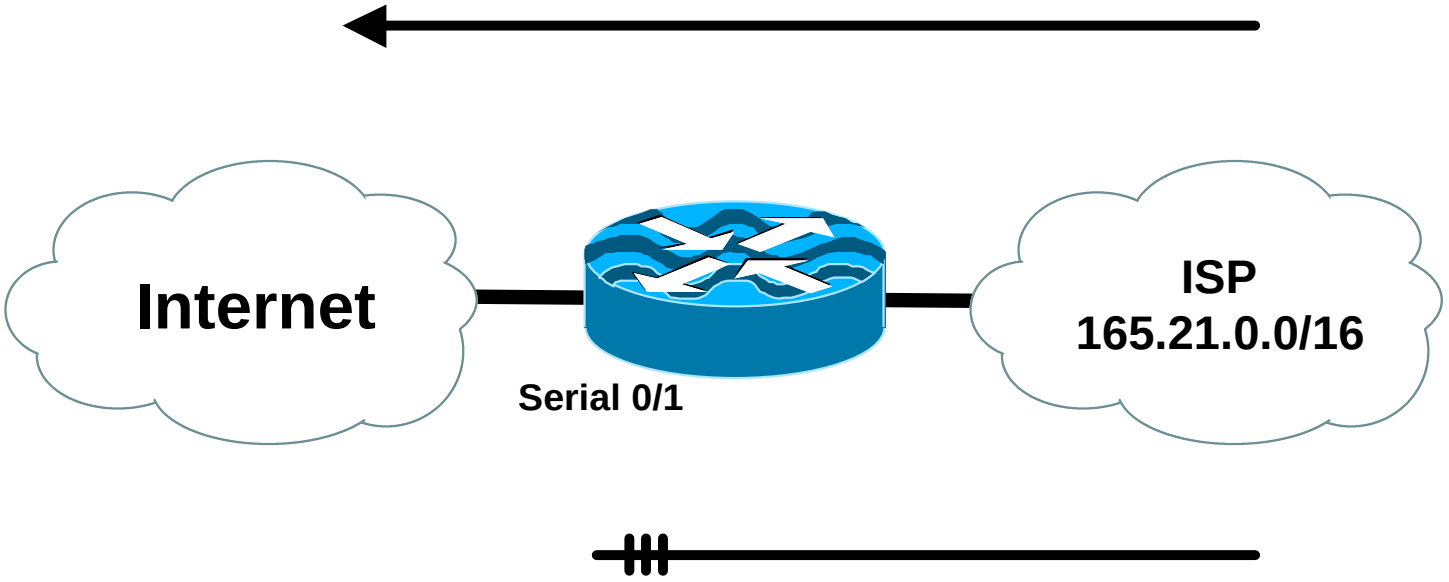
- BCP 38/ RFC 2827
- Title: Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing
- Author(s): P. Ferguson, D. Senie

Packet Filtering

- **Static Access List on the edge of the Network.**
- **Dynamic Access List with AAA Profiles**
- **Unicast RPF**

Ingress Route Filtering

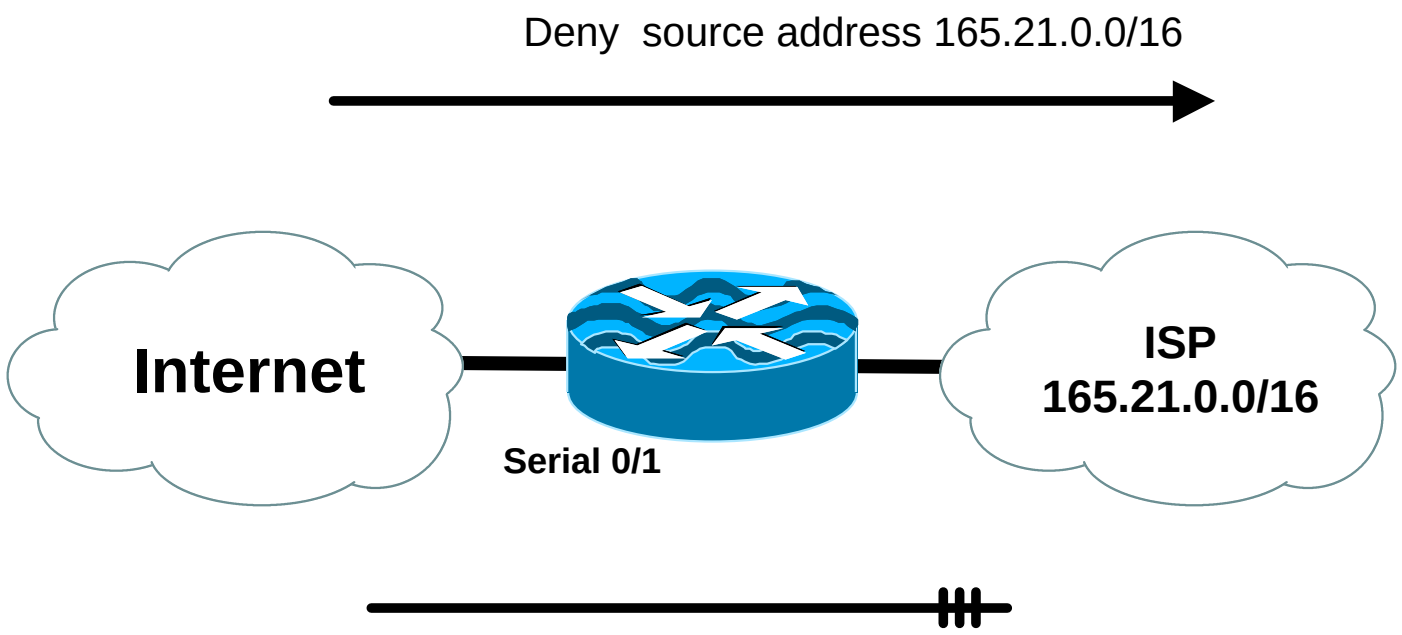
Allow source address 165.21.0.0/16



Block source address from all other networks

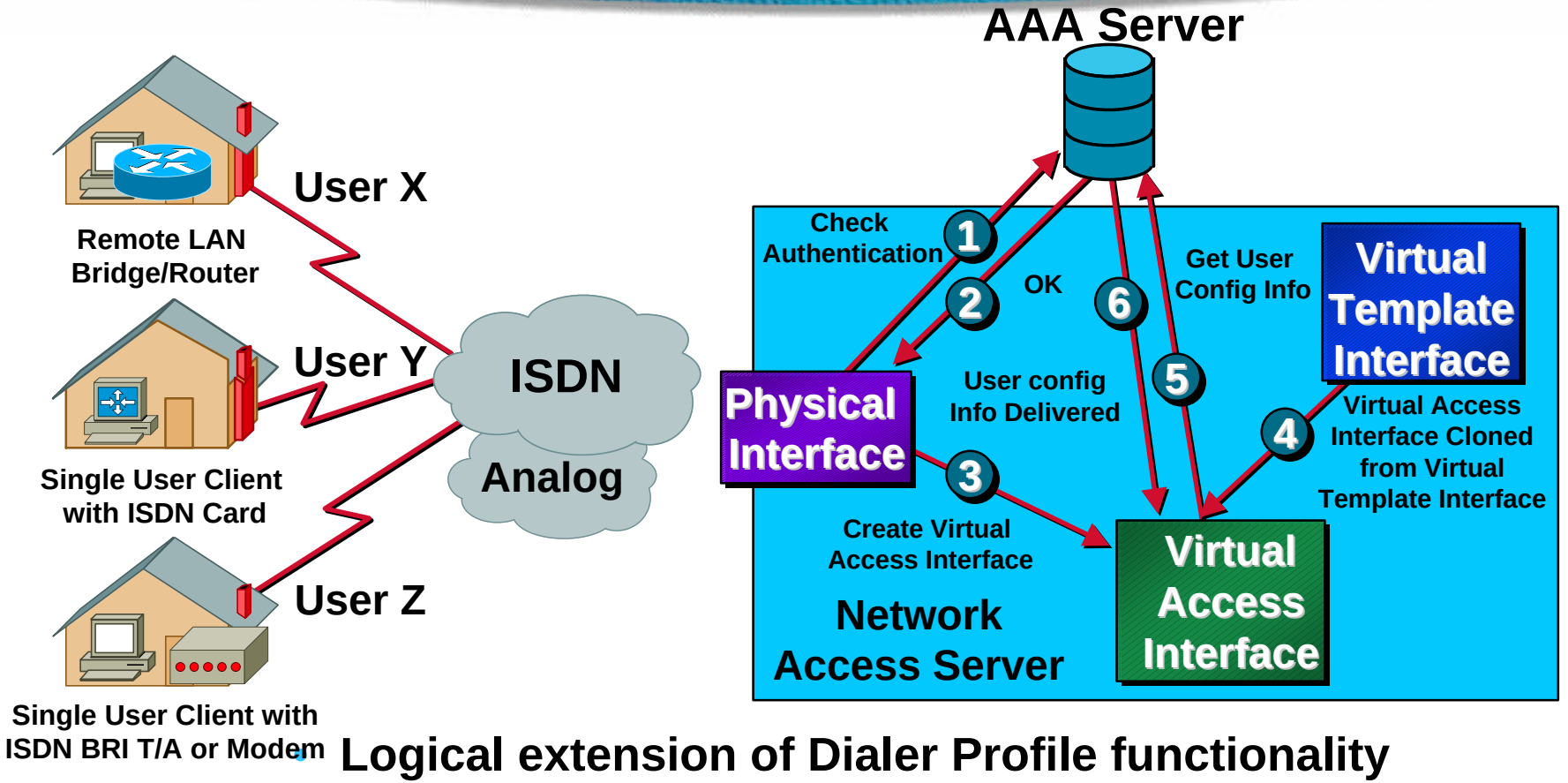
Ex. IP addresses with a source of 10.1.1.1 would be blocked

Egress Route Filtering



Ex. IP addresses with a source of 10.1.1.1 would be blocked

Dynamic ACLs with AAA Virtual Profiles

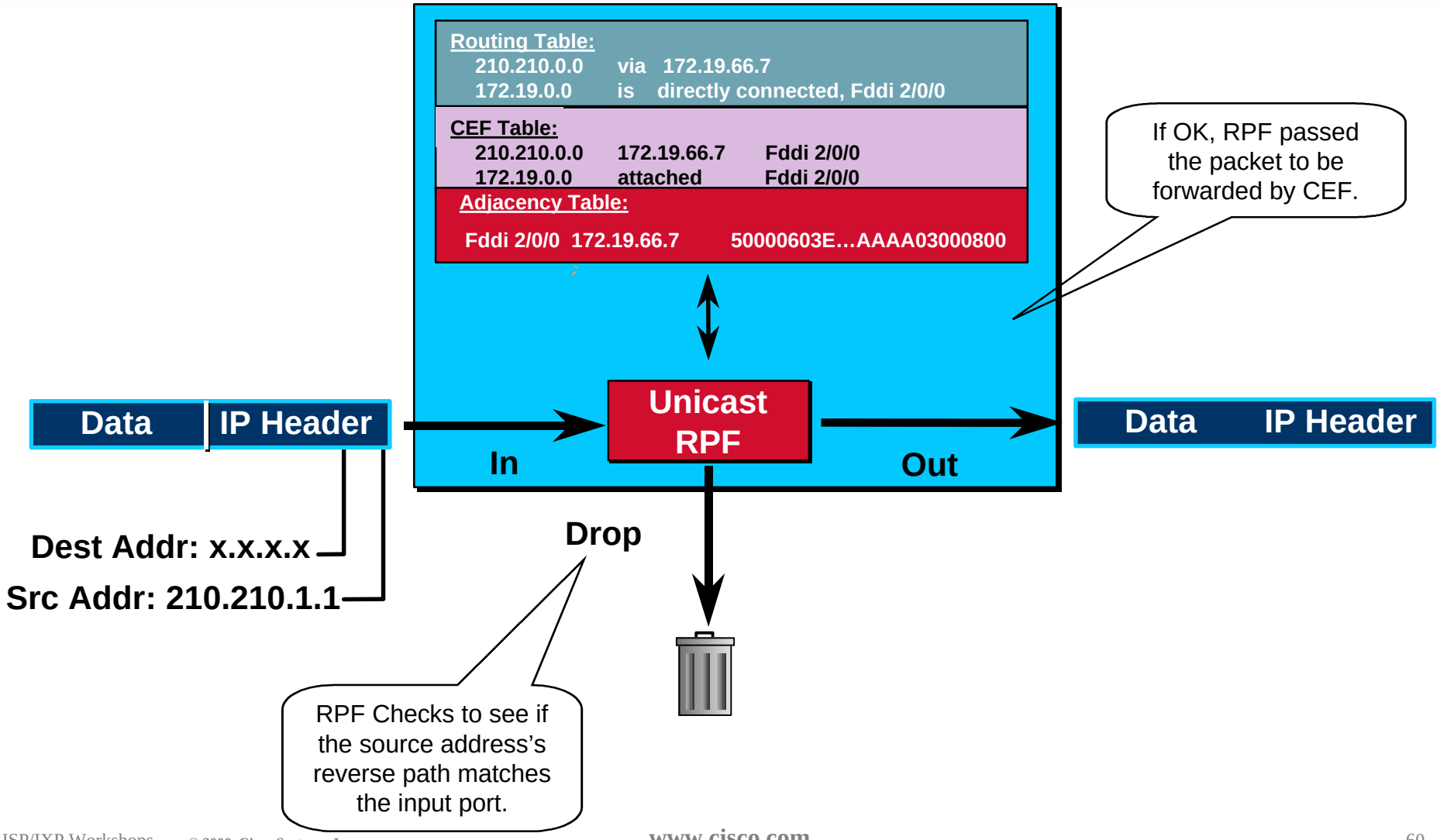


- Logical extension of Dialer Profile functionality
- ACLs stored in the Central AAA Server
- Supports both Radius and Tacacs+

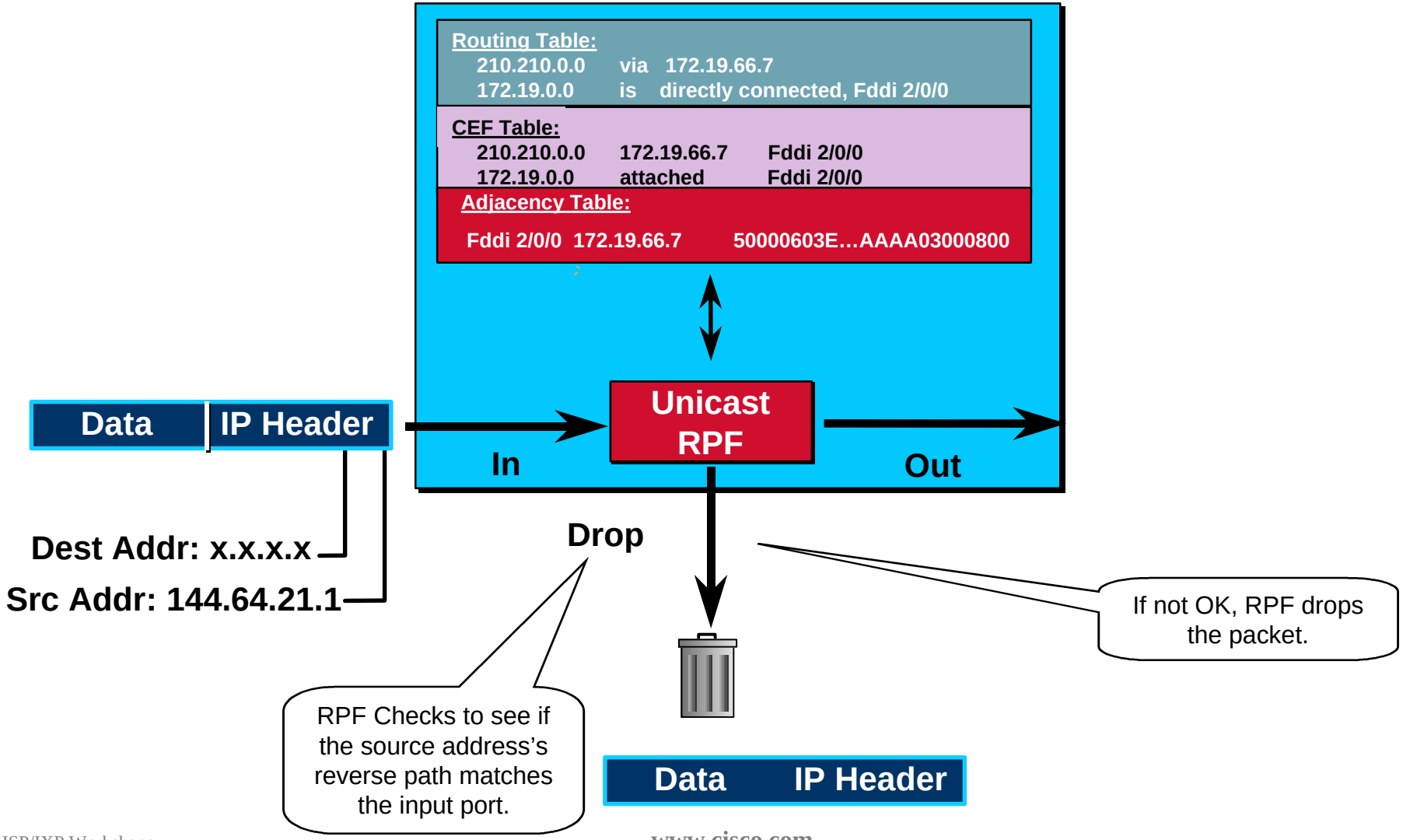
Reverse Path Forwarding

- Supported from 11.1(17)CC images
- CEF switching must be enabled
- Source IP packets are checked to ensure that the route back to the source uses the same interface
- Thought/planning required in multihoming situations

CEF Unicast RPF



CEF Unicast RPF



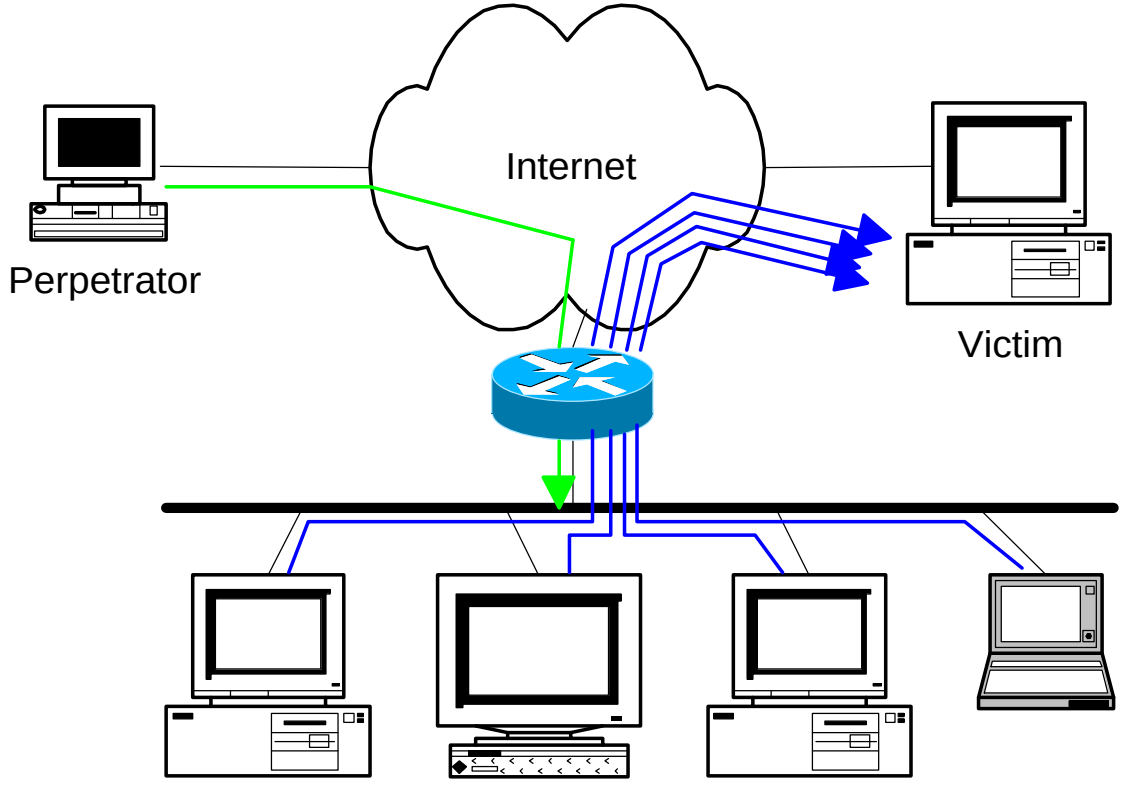
Description of "Smurfing"

- **Smurf is Denial of Service attack**
  - **Network-based, fills access pipes**
  - **Uses ICMP echo/reply packets with broadcast networks to multiply traffic**
  - **Requires the ability to send spoofed packets**
- **Abuses "bounce-sites" to attack victims**
  - **Traffic multiplied by a factor of 50 to 200**

Description of “Smurfing”

ICMP echo (spoofed source address of victim)
Sent to IP broadcast address

ICMP echo reply



Multiplied Bandwidth - Example

- Perpetrator has T1 bandwidth available (typically a cracked account), and uses half of it (768 Kbps) to send spoofed packets, half to bounce site 1, half to bounce site 2
- Bounce site 1 has a switched co-location network of 80 hosts and T3 connection to net
- Bounce site 2 has a switched co-location network of 100 hosts and T3 connection to net

Multiplied Bandwidth - Consequences

- (384 Kbps * 80 hosts) = 30 Mbps outbound traffic for bounce site 1
- (384 Kbps * 100 hosts) = 37.5 Mbps outbound traffic for bounce site 2
- Victim is pounded with 67.5 Mbps (!) from half a T1!

Profiles of Participants

- **Typical Perpetrators**
  - Cracked superuser account on well-connected enterprise network
  - Superuser account on university residence hall network (Ethernet)
  - Typical PPP dial-up account (for smaller targets)
- **Typical Bounce Sites**
  - Large co-location subnets
  - Large switched enterprise subnets
  - Typically scanned for large numbers of responding hosts
- **Typical Victims**
  - IRC Users, Operators, and Servers
  - Providers who eliminate troublesome users' accounts

Prevention Techniques

- **How to prevent your network from being the source of the attack:**
  - **Apply filters to each customer network**
    - **Ingress: Allow only those packets with source addresses within the customer's assigned netblocks**
  - **Apply filters to your upstreams**
    - **Egress: Allow only those packets with source addresses within your netblocks to protect others**
    - **Ingress: Deny those packets with source addresses within your netblocks to protect yourself**

Prevention Techniques

- **Filters will also prevent other forms of attacks as well**
- **If you do become a bounce site:**

  Trace the traffic streams to the edge of your network, and work with your upstream or peer in order to track the stream further

  MCI's DoSTracker tool

  Manual tracing/logging tips

Prevention Techniques

- How to suppress an attack if you're the victim:
  - Implement ACL's at network edges to block ICMP echo responses to your high-visibility hosts, such as IRC servers
    - Will impair troubleshooting -- "ping" breaks
    - Will still allow your access pipes to fill
  - Work with upstream providers to determine the help they can provide to you
    - Blocking ICMP echoes for high-visibility hosts from coming through your access pipes
    - Tracing attacks

Prevention Techniques

- **Technical help tips for Cisco routers - One:**
  - **BugID CSCdj35407 - “fast drop” ACL code**
    - **This bug fix optimizes the way that packets denied by an ACL are dropped within IOS, reducing CPU utilization for large amounts of denied traffic.**
    - **First major release of integration is 11.1(14)CA**
    - **Not available in 11.2 yet, but coming**

Prevention Techniques

- **Technical help tips for Cisco routers - Two:**

  BugID CSCdj35856 - ACL logging throttles

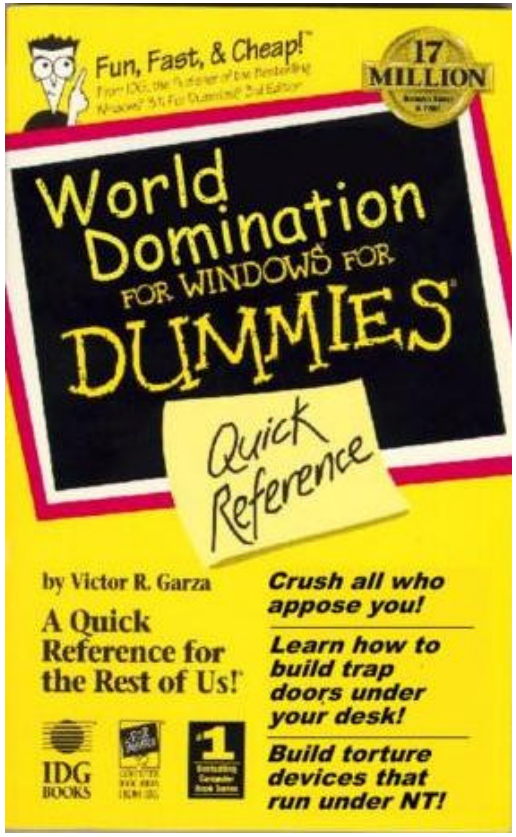
  This bug fix places a throttle in IOS which will allow a user to specify the rate at which logging will take place of packets which match a condition in an ACL where “log” or “log-input” is specified.

  First maintenance release of integration is 11.1(14.1)CA

  Not available in 11.2 yet, but coming

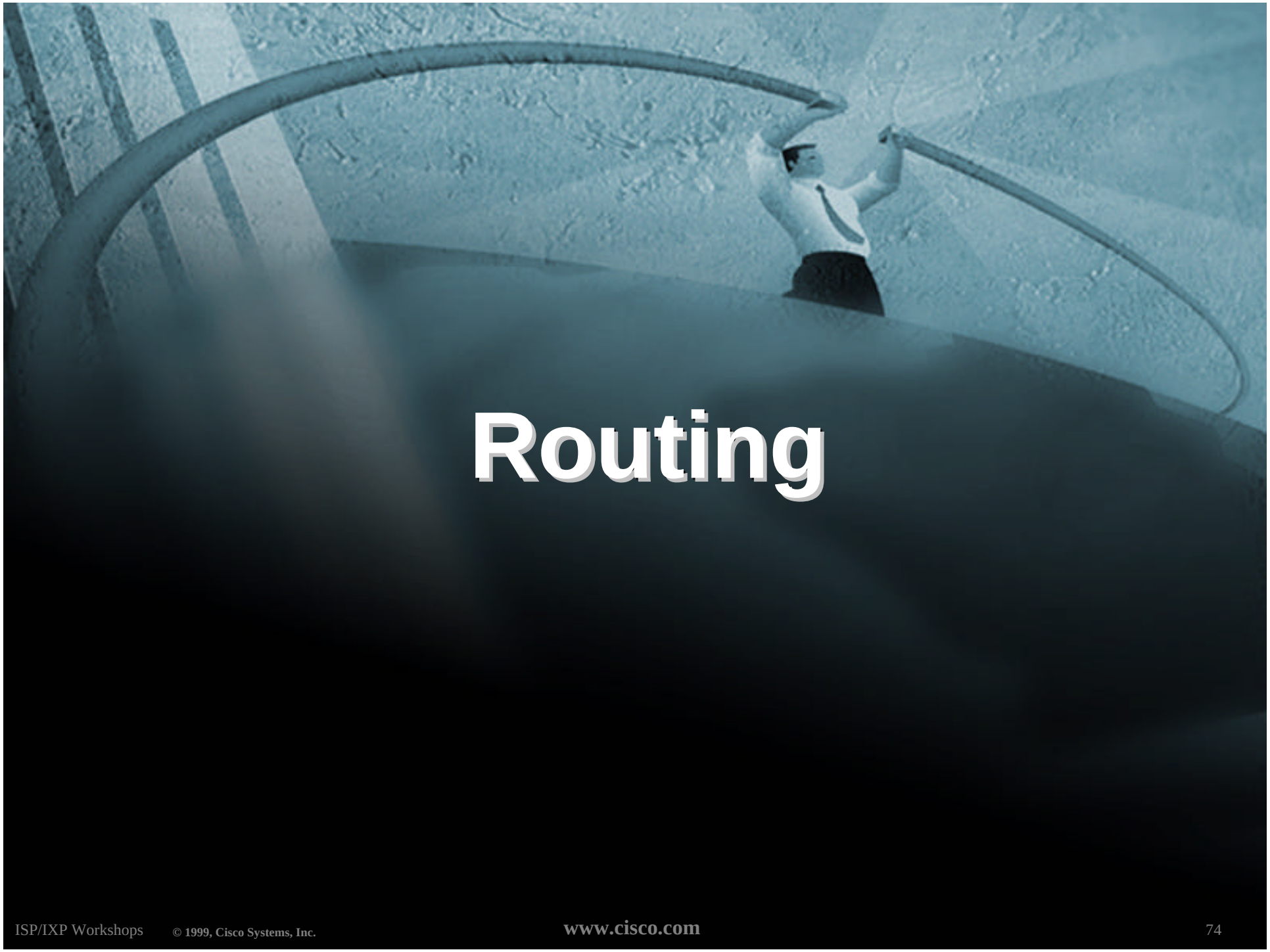
DDoS versus DoS

- Same methods and tools as DoS
- Much larger scale attacks - Elephant hunting
- Uses hundreds or even thousands of attacking points to overwhelm target
- Very difficult to determine difference between DDoS and normal network outage



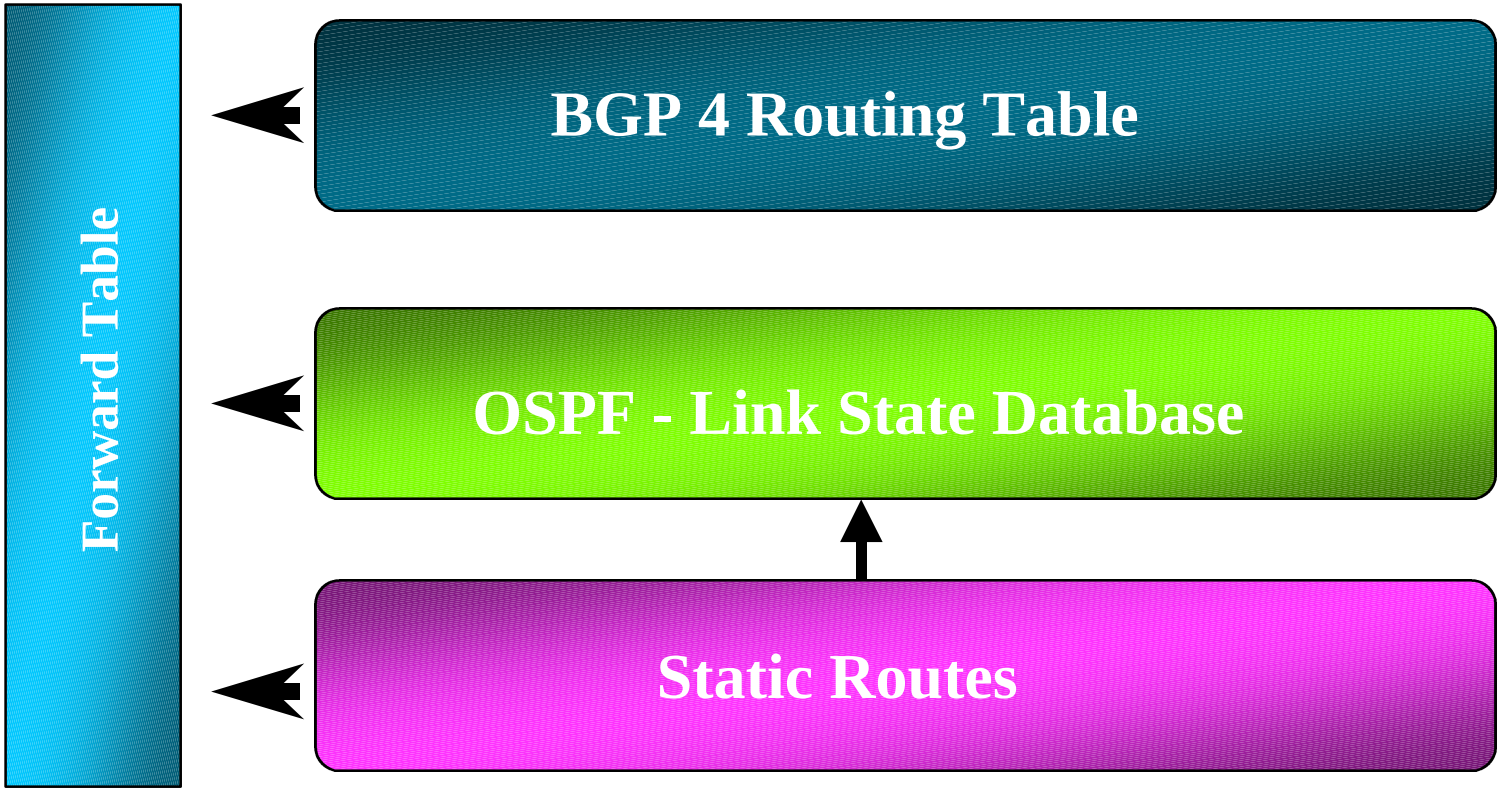
DDoS Links

- http://www.denialinfo.com/
- http://www.staff.washington.edu/dittrich
- http://www.fbi.gov/nipc/trinoo.htm
- http://www.sans.org/y2k/DDoS.htm
- http://www.nanog.org/mtg-9910/robert.html
- http://cve.mitre.org/
- http://packetstorm.securify.com/distributed/



Routing

Routing Tables Feed the Forwarding Table



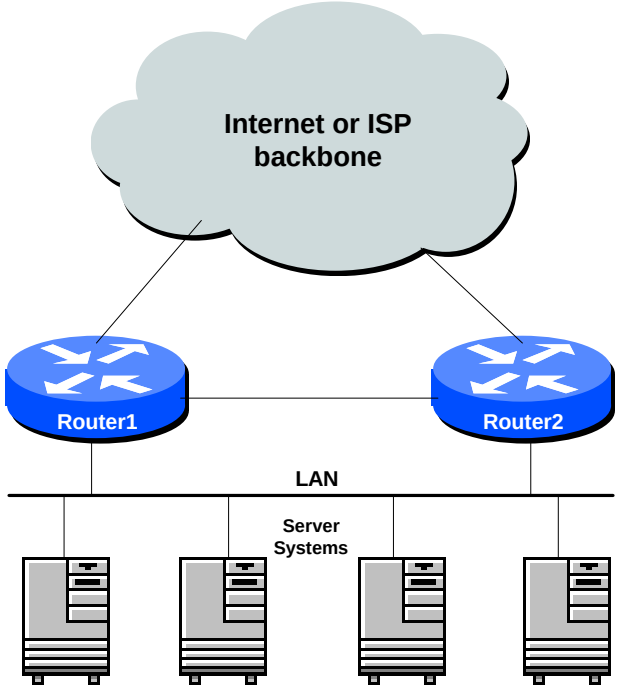
HSRP

- **Hot Standby Routing Protocol**

  virtual default gateway for dumb system LAN

  transparent cut-over in case of failure

```
Router1:
   interface ethernet 0/0
   description Service LAN
   ip address 169.223.10.1 255.255.255.0
   standby 10 ip 169.223.10.254
Router2:
   interface ethernet 0/0
   description Service LAN
   ip address 169.223.10.2 255.255.255.0
   standby 10 priority 150
   standby 10 preempt
   standby 10 ip 169.223.10.254
```



CIDR Features

- The Internet is a classless world. All routers connect to the Internet must be CIDR compliant, else there will be problems with the network connection to the Internet.
- All Cisco routers should have the following commands configured for CIDR:

```
ip subnet-zero
ip classless
```

- These are default from IOS 12.0 onwards

Selective Packet Discard

- When a link goes to a saturated state, you will drop packets. The problem is that you will drop any type of packets - including your routing protocols.
- Selective Packet Discard (SPD) will attempt to drop non-routing packets instead of routing packets when the link is overloaded.

```
ip spd enable
```

- Enabled by default from 11.2(5)P and later releases, available option in 11.1CA/CC.

Source Routing

- IP has provision to allow source IP host to specify route through Internet
- ISPs should turn this off, unless it is specifically required:

```
no ip source-route
```

BGP

- **There are key BGP features that should be configured by ISPs:**

```
update-source loopback 0
ip bgp-community new-format
no synchronization
bgp dampening
no auto-summary
bgp neighbor authentication
bgp neighbor maximum-prefix
```

BGP

- **More helpful features:**

```
bgp neighbor soft-reconfiguration
bgp neighbor shutdown
bgp log-neighbor-changes
no bgp fast-external-fallover
bgp peer-groups
ip prefix-lists
```

iBGP configuration

- **Use loopback interface**
  - **it never goes away**
  - **routers have multiple external paths**
  - **has multiple uses**

```
interface loopback 0
 ip address 215.17.1.34 255.255.255.255
router bgp 200
 neighbor 215.17.1.35 remote-as 200
 neighbor update-source loopback 0
 neighbor 215.17.1.36 remote-as 200
 neighbor update-source loopback 0
```

BGP Community Format

- **Communities are used extensively**
- **Cisco IOS supports two formats**
  - **One 32 bit integer eg 13107210**
  - **Two 16 bit integers eg 200:10**
- **RFC1998 recommends 16:16 format**
  - **Format AS:xxxx**

```
ip bgp-community new-format
```

BGP Synchronization

- **BGP does not advertise a route before all routers in the AS have learned it via an IGP**
- **Disable synchronization if:**
  - AS doesn't pass traffic from one AS to another
  - All transit routers in AS run BGP
  - iBGP is used across backbone

```
no synchronization
```

BGP Neighbour Shutdown

- **Shutdown BGP peering**

  previously required to delete configuration

  now can simply “shutdown” the peering

- **Configuration example:**

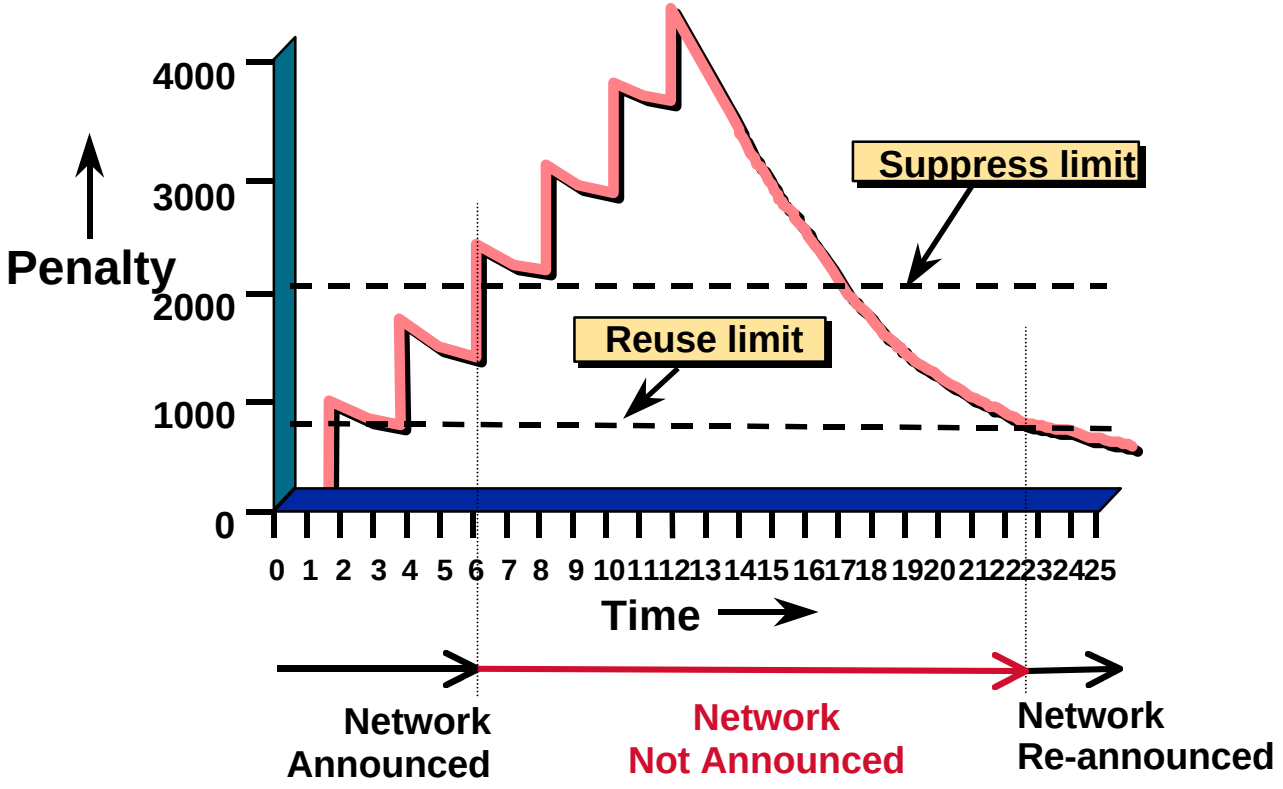
```
router bgp 200
 neighbor 215.7.1.1 remote-as 210
 neighbor 215.7.1.1 shutdown
```

- **Can be reactivated with**

```
no neighbor 215.7.1.1 shutdown
```

BGP Dampening

- **Route flap dampening to minimise instability in local network and Internet**



BGP Dampening

- **Recommended values and sample configurations for ISPs at:**

  http://www.ripe.net/docs/ripe-210.html

- **Example techniques:**

  Internet Routing Architecture - Bassam Halabi

`bgp dampening`

BGP Auto Summarisation

- **Automatically summarises subprefixes to the classful network.**
- **Must be turned off for any Internet connected site using BGP.**
- **Internet is classless - class A, class B and class C are no more.**

```
no auto-summary
```

BGP Neighbour Authentication

- **MD5 authentication between two peers**

  password must be known to both peers

- **peer-group can be used to apply to multiple peerings**

```
neighbor 169.222.10.1 password v61ne0qkel33&
```

Clear BGP Sessions per AS

- Ability to clear the BGP sessions of all the neighbors configured with a specific AS number
- Syntax:

  clear ip bgp <as number>

- Availability 11.1(14)CA, 11.1CC, 11.2(9), 11.3(2)

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 log-neighbor-changes

- **Log neighbour up/down events, and the reason for the last neighbour 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)

Reason for Last Peer Reset

- Router keeps reason for the last BGP peer reset for each of its peers. Useful to analyze BGP session resets.
- Available as part of the **show ip bgp neighbor** command output. Accessible also through SNMP.
- Availability 11.1CC, 11.2(12), 11.3(2)

Reset Reasons

"BGP protocol initialization”

"No memory for path entry"
"No memory for attribute entry"
"No memory for prefix entry"
"No memory for aggregate entry"
"No memory for dampening info"
"No memory for BGP updates"

"BGP Notification received"

Reset Reasons (Cont.)

"Erroneous BGP Update received”
Connection is in error state - generally waiting for TCP close.

"User reset"

"Peer timeout"

"Password change"

"Error during connection collision"

"Peer closed the session”

"Peer over prefix limit"

Reset Reasons (Cont.)

"Interface flap"

"Router ID changed"

"Neighbor deleted"

"Member added to peergroup"

"Admin. shutdown"

"Remote AS changed"

"NLRI changed"

"RR client config change"

"Soft reconfig change"

BGP Peering

- **By default, peerings are reset immediately the line protocol to an external neighbour goes down**
  - **bad for high latency, unreliable, long distance, or congested links**
- **IOS option to disable this**
  - **recommended in RIPE-210**
  - **uses standard keepalive/hold timers (60s/180s)**
  - **`no bgp fast-external-fallover`**

BGP peer groups

- **Reduces CPU load and memory**
  - **update generation processed once**
  - **BGP configuration simplified**

```
router bgp 109
 neighbor internal peer-group
 neighbor internal remote-as 109
 neighbor internal update-source loopback 0
 neighbor 131.108.10.1 peer-group internal
 neighbor 131.108.20.1 peer-group internal
```

Prefix Lists

- High performing access-list
- Faster loading of large lists
- Incremental configuration

  sequence numbers optional

  `no ip prefix-list sequence-number`
- Available from 11.1(17)CC and 12.0
- Configured by:

  `ip prefix-list <list-name>`

Prefix-list Command

[no] ip prefix-list <*list-name*> [seq <*seq-value*>] deny | permit <*network*>/<*len*> [ge <*ge-value*>] [le <*le-value*>]

<*network*>/<*len*>: The prefix and its length

ge <*ge-value*>: "greater than or equal to"

le <*le-value*>: "less than or equal to"

Both "ge" and "le" are optional. Used to specify the range of the prefix length to be matched for prefixes that are more specific than <*network*>/<*len*>

Prefix Lists - Examples

- **Deny default route**

```
ip prefix-list EG deny 0.0.0.0/0
```

- **Permit the prefix 35.0.0.0/8**

```
ip prefix-list EG permit 35.0.0.0/8
```

- **In 192/8 allow up to /24**

```
ip prefix-list EG permit 192.0.0.0/8 le 24
```

- **In 192/8 deny /25 and above**

```
ip prefix-list EG deny 192.0.0.0/8 ge 25
```

- **Permit all**

```
ip prefix-list EG permit 0.0.0.0/0 le 32
```

Prefix Lists in BGP

- **Prefix-list can be used as alternative to distribute-list**

```
router bgp 200
 neighbor 169.222.1.1 remote-as 200
 neighbor 169.222.1.1 prefix-list FILTER-IN in
 neighbor 169.222.1.1 prefix-list FILTER-OUT out
```

- **Prefix-lists and access-lists are mutually exclusive**

Prefix-list route-map command

```
route-map <name> permit|deny <seq-num>
  match ip address prefix-list <name>
  [<name> ...]
```

- Used for route filtering, originating default, and redistribution in other routing protocols as well
- Not for packet filtering

Prefix-List ORF

- Outbound Route Filter Capability when using prefix-lists

  new from 12.0(5)S release

- If remote BGP peer supports ORF capability, local BGP router can send inbound prefix-list to remote router
- Remote router installs received prefix-list in addition to its own outbound filters
- Reduces unwanted routing updates from peers

BGP Conditional Advertisement

- Reduce the number of prefixes advertised when there is no failure
- Prefix injected when there is a failure to restore connectivity

  For multihoming customers or backup scenario

- Help scale the Internet backbone

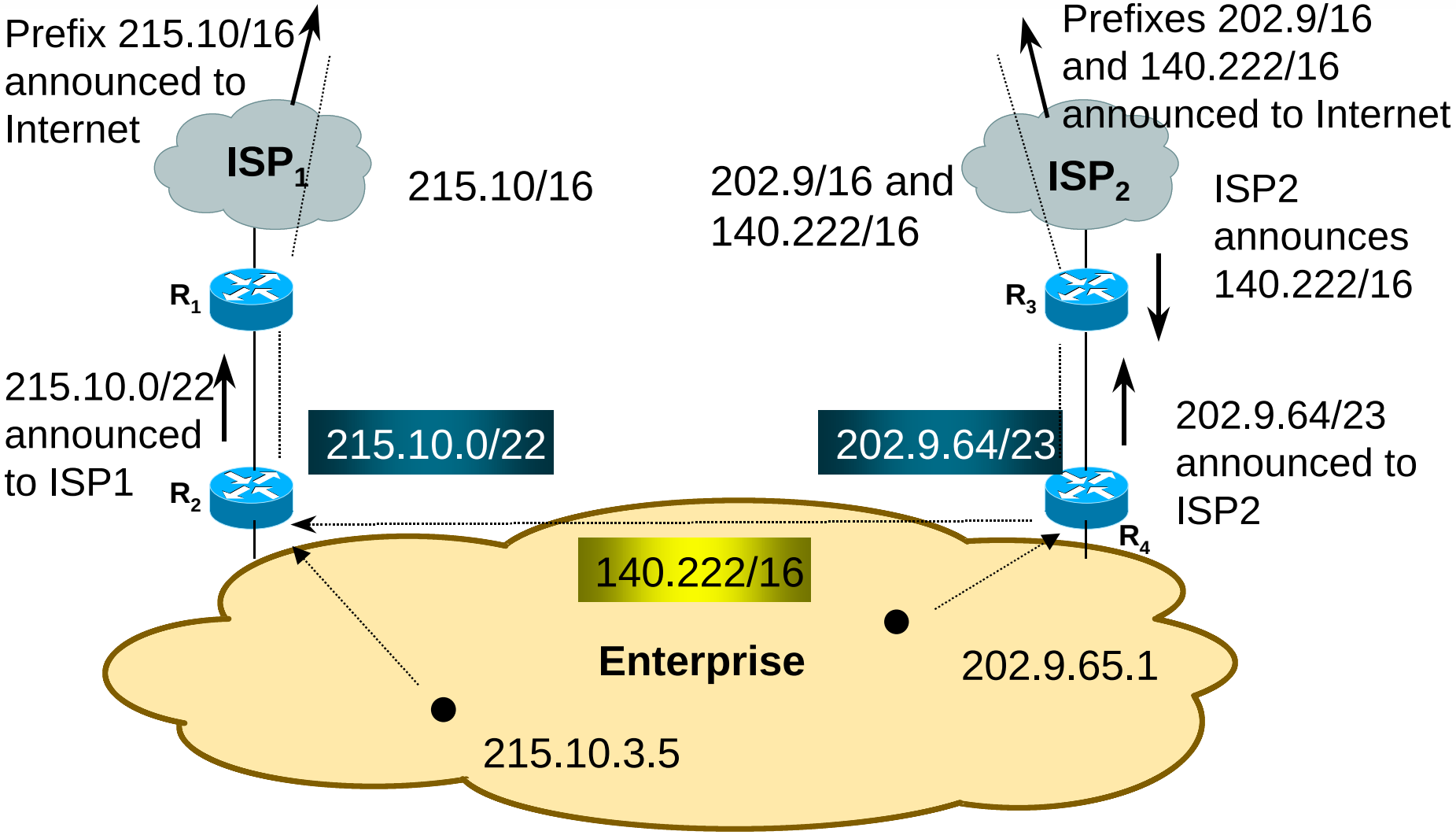
  It is in everybody's best interest…

BGP Conditional Advertisement: configuration

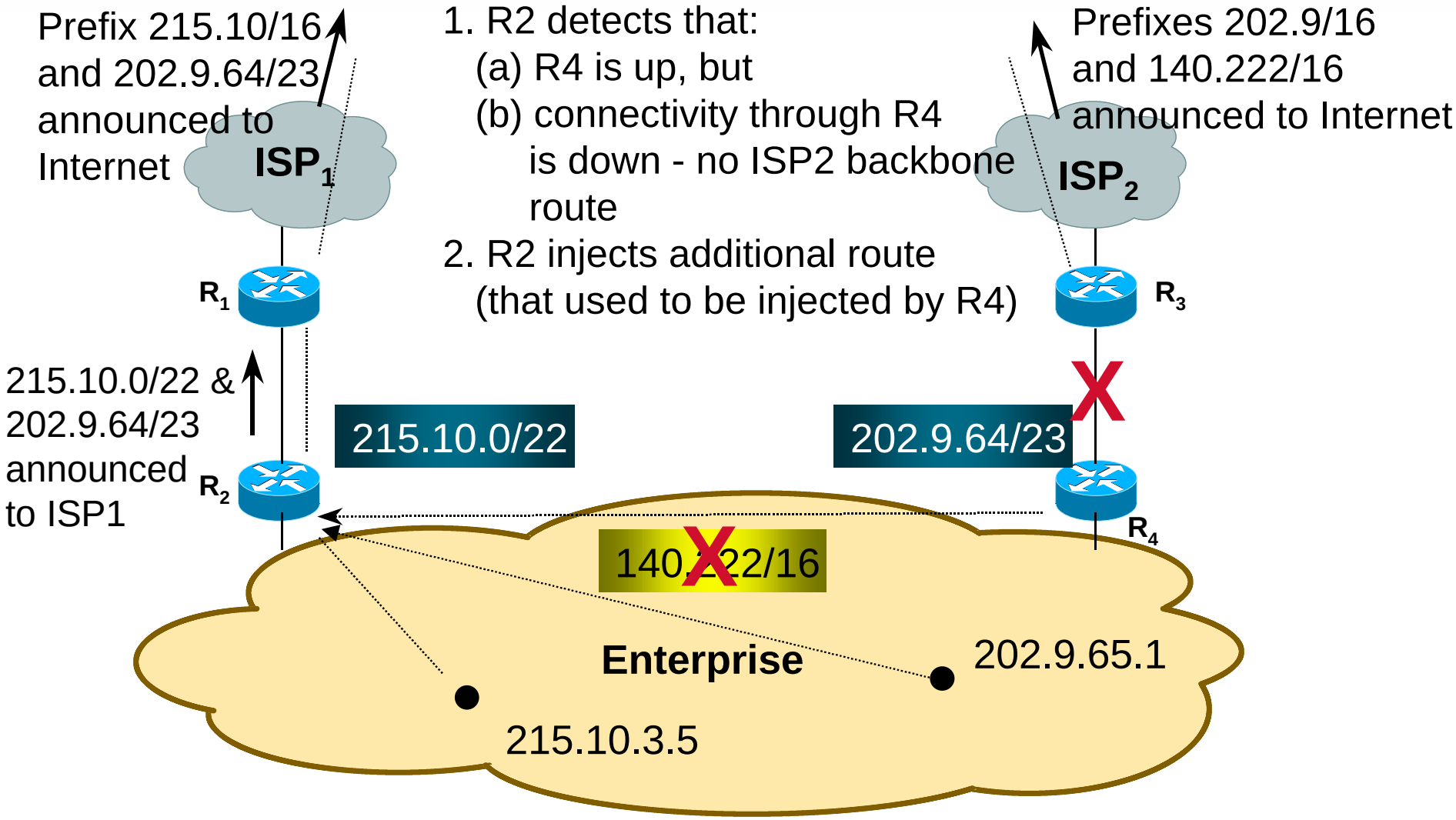
```
neighbor <x.x.x.x> advertise-map <route-map>
  non-exist-map <route-map>
```

- **<*route-map*> is a standard route-map**
- **non-exist-map specifies prefix that BGP speaker will track**
- **advertise-map specifies prefix that will be advertised when prefix in non-exist-map no longer exist**

Example - steady state



Example - link failure



Example Configuration

On router R2:

```
router bgp 100
  neighbor <R1> advertise-map ISP2-subblock non-exist-map ISP2-backbone
route-map ISP2-subblock permit 10
   match ip address prefix-list ISP2-sub        ! <ISP2-subblock-prefix>
route-map ISP2-backbone permit 10
   match ip address prefix-list ISP2-bb         ! <ISP2-backbone-prefix>
ip prefix-list ISP2-sub permit 202.9.64.0/23    ! <ISP2-subblock-prefix>
ip prefix-list ISP2-bb permit 140.222.0.0/16    ! <ISP2-backbone-prefix>
```

Where to get more information

- **Supporting *IOS Essentials* WhitePaper**

  `http://www.cisco.com/public/cons/isp/documents/IOSEssentialsPDF.zip`

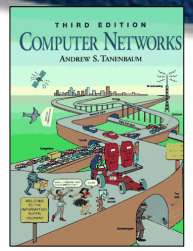
- **Check the CTO Consulting Engineering ISP Resources page:**

  `http://www.cisco.com/public/cons/isp/`

- **Join the cisco-nsp mailing list - set up by ISPs for ISPs**

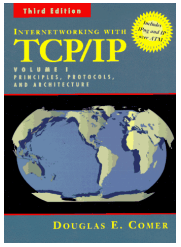
  `send e-mail to majordomo@puck.nether.net with the words “subscribe cisco-nsp” in the body`

For Further Reference…

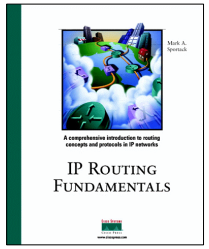


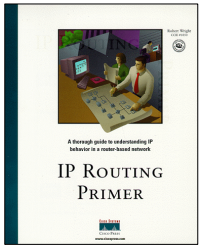
- **Computer Networks, Third Edition**
  by Andrew Tanenbaum (ISBN: 0-13349-945-6)

- **Interconnections : Bridges and Routers (second Ed)**
  by Radia Perlman (ISBN: 0-20163-448-1)



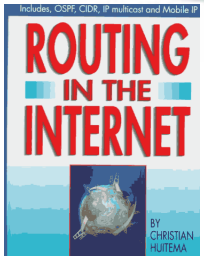
- **Internetworking with TCP / IP, Volume 1: Principles, Protocols, and Architecture**
  by Douglas Comer (ISBN: 0-13216-987-8)



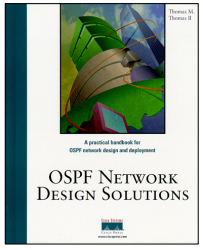


- **IP Routing Fundamentals**
  by Mark Sportack (ISBN: 1-57870-071-x)
- **IP Routing Primer**
  by Robert Wright (ISBN: 1-57870-108-2)

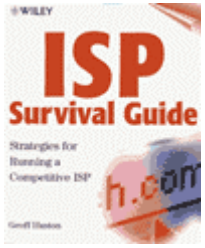
For Further Reference…



- **Routing in the Internet**
  by Christian Huitema (ISBN: 0-13132-192-7)



- **OSPF Network Design Solutions**
  by Thomas, Thomas M. (ISBN: 1-57870-046-9)



- **ISP Survival Guide : Strategies for Running a Competitive ISP**
  by Geoff Huston (ISBN:0-47131-499-4)

- **Internet Routing Architectures**
  by Bassam Halabi (ISBN: 1-56205-652-2)


Cisco Systems
Empowering the
Internet Generation℠