

Troubleshooting BGP

Philip Smith <pfs@cisco.com>

AfNOG 2003, Kampala, Uganda

Presentation Slides

Cisco.com

- Available on

<ftp://ftp-eng.cisco.com/pfs/seminars/AfNOG2003-BGP-Troubleshooting.pdf>

Assumptions

Cisco.com

- **Presentation assumes working knowledge of BGP**
Beginner and Intermediate experience of protocol
- **Knowledge of Cisco CLI**
Hopefully you can translate concepts into your own router CLI
- **If in any doubt, please ask!**

Fundamentals of Troubleshooting

Cisco.com

- **Before we begin/Troubleshooting is about:**
 - Not panicking**
 - Creating a checklist**
 - Working to that checklist**
 - Starting at the bottom and working up**
- **This presentation will have references throughout to checklists**
 - They are the best way to work to a solution**
 - They are what many NOC staff follow when diagnosing and solving network problems**

Agenda

Cisco.com

- **Peer Establishment**
- **Missing Routes**
- **Internet Reachability Problems**

Peer Establishment

Cisco.com

- **Routers establish a TCP session**

Port 179—Permit in ACLs

IP connectivity (route from IGP)

- **OPEN messages are exchanged**

Peering addresses must match the TCP session

Local AS configuration parameters

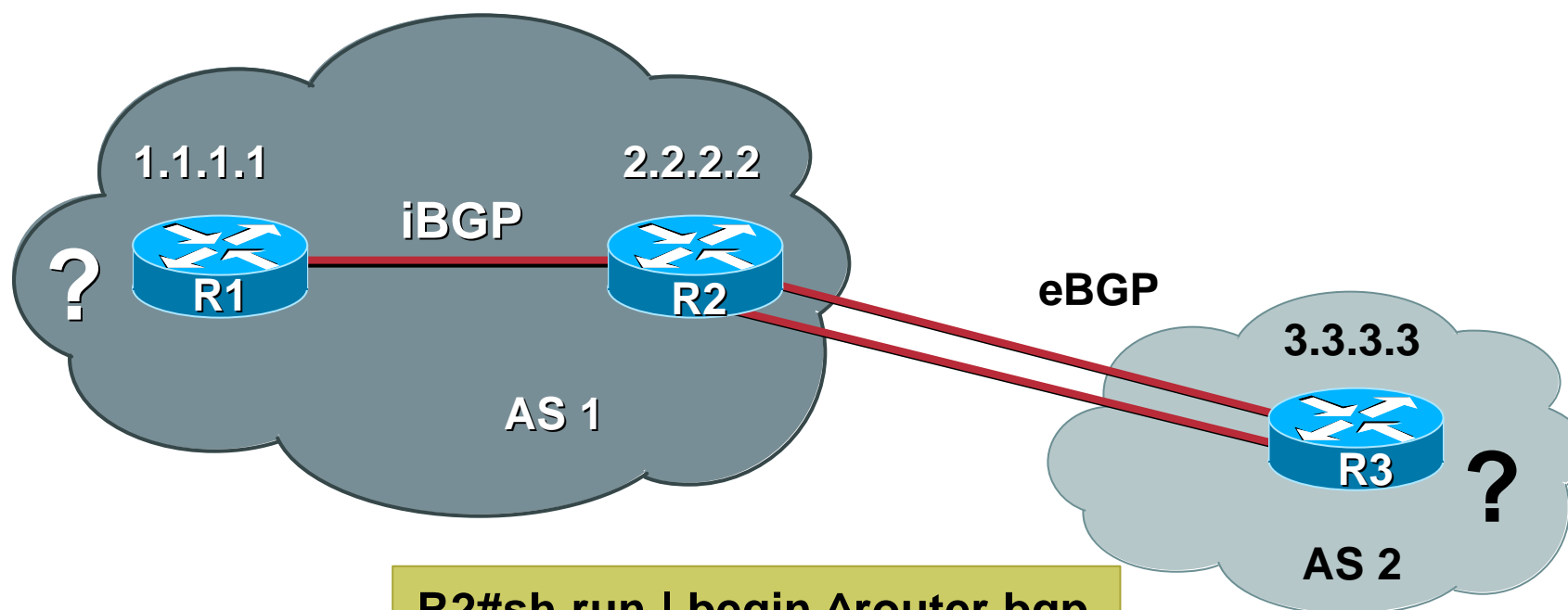
Common Problems

Cisco.com

- **Sessions are not established**
 - No IP reachability**
 - Incorrect configuration**
- **Peers are flapping**
 - Layer 2 problems**

Peer Establishment—Diagram

Cisco.com



```
R2#sh run | begin ^router bgp
router bgp 1
  bgp log-neighbor-changes
  neighbor 1.1.1.1 remote-as 1
  neighbor 3.3.3.3 remote-as 2
```


Peer Establishment—Symptoms

Cisco.com

R2#show ip bgp summary

BGP router identifier 2.2.2.2, local AS number 1

BGP table version is 1, main routing table version 1

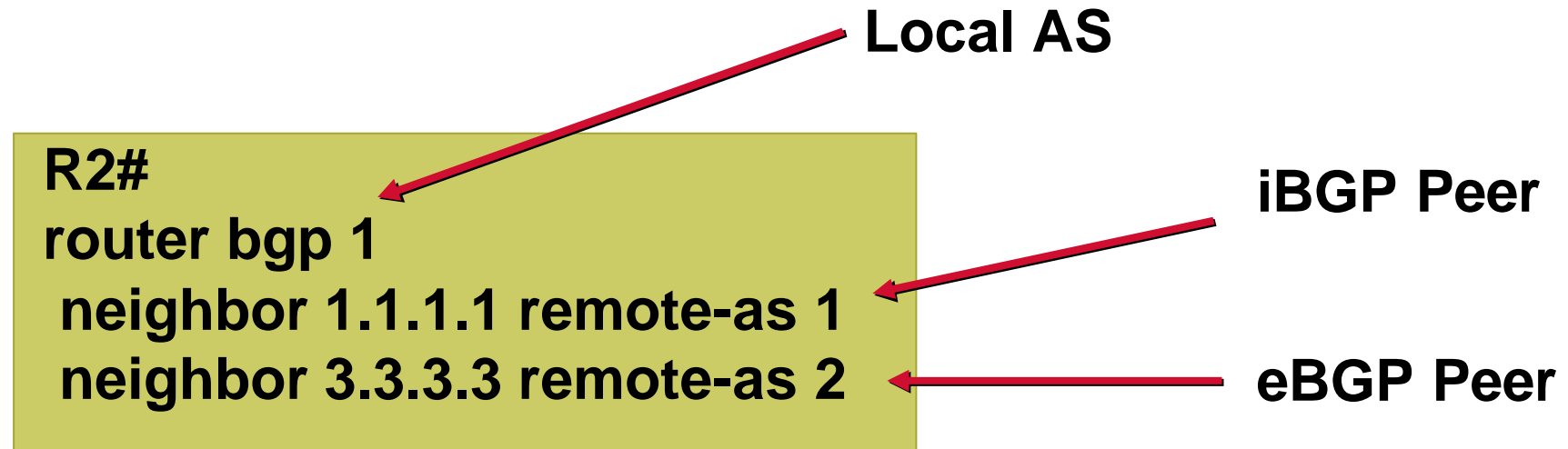
Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State
1.1.1.1	4	1	0	0	0	0	0	never	Active
3.3.3.3	4	2	0	0	0	0	0	never	Idle

- Both peers are having problems

State may change between Active, Idle and Connect

Peer Establishment

- Is the Local AS configured correctly?
- Is the remote-as assigned correctly?
- Verify with your diagram or other documentation!



Peer Establishment—iBGP

Cisco.com

- Assume that IP connectivity has been checked
- Check TCP to find out what connections we are accepting

```
R2#show tcp brief all
```

TCB	Local Address	Foreign Address	(state)
005F2934	*.179	3.3.3.3.*	LISTEN
0063F3D4	*.179	1.1.1.1.*	LISTEN

We Are Listening for TCP Connections for Port 179 for the Configured Peering Addresses Only!

```
R2#debug ip tcp transactions
TCP special event debugging is on
R2#
TCP: sending RST, seq 0, ack 2500483296
TCP: sent RST to 4.4.4.4:26385 from 2.2.2.2:179
```

Remote Is Trying to Open the Session from 4.4.4.4 Address...

Peer Establishment—iBGP

Cisco.com

What about Us?

```
R2#debug ip bgp
BGP debugging is on
R2#
BGP: 1.1.1.1 open active, local address 4.4.4.5
BGP: 1.1.1.1 open failed: Connection refused by remote host
```

We Are Trying to Open the Session from 4.4.4.5 Address...

```
R2#sh ip route 1.1.1.1
Routing entry for 1.1.1.1/32
  Known via "static", distance 1, metric 0 (connected)
    * directly connected, via Serial1
      Route metric is 0, traffic share count is 1

R2#show ip interface brief | include Serial1
Serial1          4.4.4.5      YES manual  up    up
```

Peer Establishment—iBGP

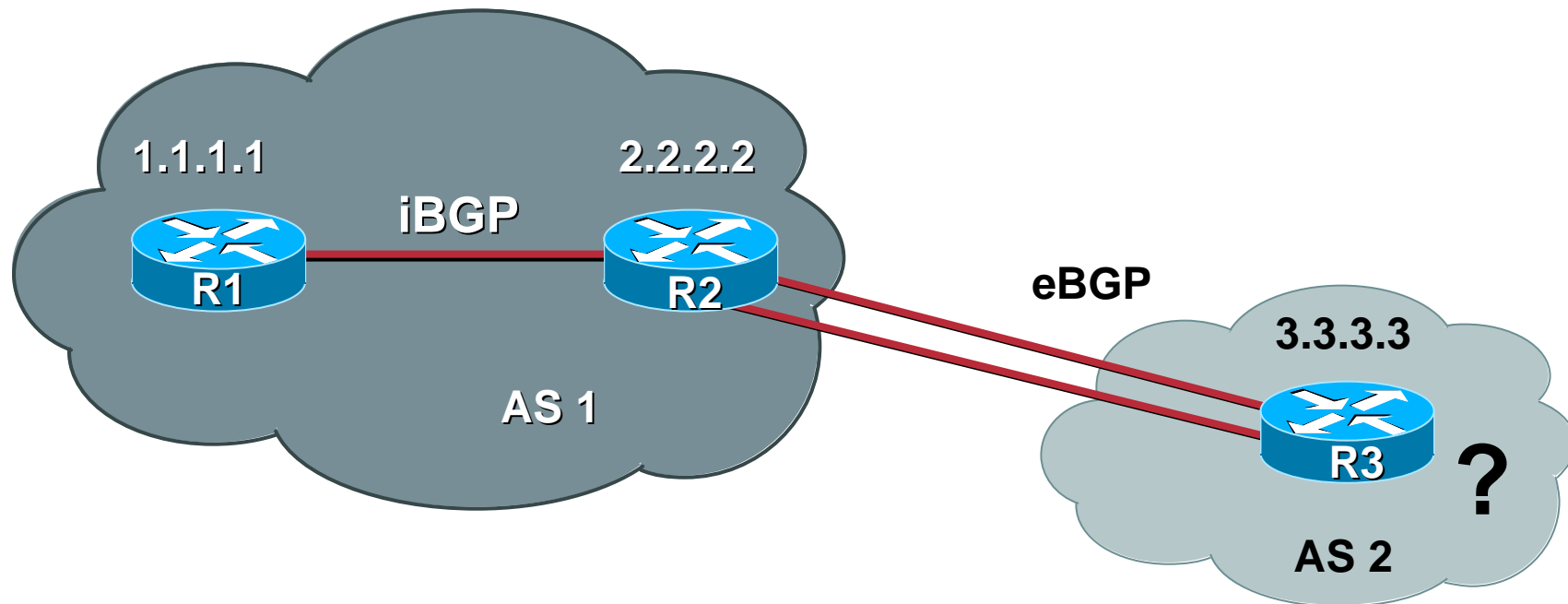
Cisco.com

- Source address is the outgoing interface towards the destination but peering in this case is using loopback interfaces!
- Force both routers to source from the correct interface
- Use “update-source” to specify the loopback when loopback peering

```
R2#  
router bgp 1  
  neighbor 1.1.1.1 remote-as 1  
  neighbor 1.1.1.1 update-source Loopback0  
  neighbor 3.3.3.3 remote-as 2  
  neighbor 3.3.3.3 update-source Loopback0
```

Peer Establishment—Diagram

Cisco.com



- R1 is established now
- The eBGP session is still having trouble!

Peer Establishment—eBGP

Cisco.com

- Trying to load-balance over multiple links to the eBGP peer
- Verify IP connectivity

Check the routing table

Use ping/trace to verify two way reachability

```
R2#ping 3.3.3.3
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 3.3.3.3, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 4/4/8 ms
```

- Routing towards destination correct, but...

Peer Establishment—eBGP

Cisco.com

```
R2#ping ip
Target IP address: 3.3.3.3
Extended commands [n]: y
Source address or interface: 2.2.2.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 3.3.3.3, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

- Use extended pings to test loopback to loopback connectivity
- R3 does not have a route to our loopback, 2.2.2.2

Peer Establishment—eBGP

Cisco.com

- Assume R3 added a route to 2.2.2.2
- Still having problems...

```
R2#sh ip bgp neigh 3.3.3.3
BGP neighbor is 3.3.3.3, remote AS 2, external link
  BGP version 4, remote router ID 0.0.0.0
  BGP state = Idle
  Last read 00:00:04, hold time is 180, keepalive interval is 60 seconds
  Received 0 messages, 0 notifications, 0 in queue
  Sent 0 messages, 0 notifications, 0 in queue
  Route refresh request: received 0, sent 0
  Default minimum time between advertisement runs is 30 seconds
For address family: IPv4 Unicast
  BGP table version 1, neighbor version 0
  Index 2, Offset 0, Mask 0x4
  0 accepted prefixes consume 0 bytes
  Prefix advertised 0, suppressed 0, withdrawn 0
  Connections established 0; dropped 0
  Last reset never
  External BGP neighbor not directly connected.
  No active TCP connection
```

Peer Establishment—eBGP

Cisco.com

```
R2#  
router bgp 1  
  neighbor 3.3.3.3 remote-as 2  
  neighbor 3.3.3.3 ebgp-multihop 2  
  neighbor 3.3.3.3 update-source Loopback0
```

- **eBGP peers are normally directly connected**
By default, TTL is set to 1 for eBGP peers
If not directly connected, specify ebgp-multihop
- **At this point, the session should come up**

Peer Establishment—eBGP

Cisco.com

```
R2#show ip bgp summary
```

```
BGP router identifier 2.2.2.2, local AS number 1
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
3.3.3.3	4	2	10	26	0	0	0	never	Active

- **Still having trouble!**

**Connectivity issues have already
been checked and corrected**

Peer Establishment—eBGP

Cisco.com

```
R2#debug ip bgp events
14:06:37: BGP: 3.3.3.3 open active, local address 2.2.2.2
14:06:37: BGP: 3.3.3.3 went from Active to OpenSent
14:06:37: BGP: 3.3.3.3 sending OPEN, version 4
14:06:37: BGP: 3.3.3.3 received NOTIFICATION 2/2
                (peer in wrong AS) 2 bytes 0001
14:06:37: BGP: 3.3.3.3 remote close, state CLOSEWAIT
14:06:37: BGP: service reset requests
14:06:37: BGP: 3.3.3.3 went from OpenSent to Idle
14:06:37: BGP: 3.3.3.3 closing
```

- If an error is detected, a **notification** is sent and the session is closed
- R3 is configured incorrectly
 - Has “neighbor 2.2.2.2 remote-as 10”
 - Should have “neighbor 2.2.2.2 remote-as 1”
- After R3 makes this correction the session should come up

eBGP summary

Cisco.com

- **Remember to allow TCP/179 through filters**

Common eBGP implementation error

```
access-list 100 permit tcp host 3.3.3.3 eq 179 host 2.2.2.2  
access-list 100 permit tcp host 3.3.3.3 host 2.2.2.2 eq 179
```

- **Need to be careful with ebgp-multihop**

Peer between loopback interfaces

Needed to loadshare

Remember `update-source loopback 0`

TTL must be at least 2 for ebgp-multihop between directly connected neighbours

Use TTL value carefully

Peer Establishment—Passwords

Cisco.com

- **Using passwords on iBGP and eBGP sessions**

Link won't come up

Been through all the previous troubleshooting steps

```
R2#show ip bgp summary
```

```
BGP router identifier 2.2.2.2, local AS number 1
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
3.3.3.3	4	2	10	26	0	0	0	never	Active

Peer Establishment—Passwords

Cisco.com

```
R2#  
router bgp 1  
  neighbor 3.3.3.3 remote-as 2  
  neighbor 3.3.3.3 ebgp-multihop 2  
  neighbor 3.3.3.3 update-source Loopback0  
  neighbor 3.3.3.3 password 7 05080F1C221C
```

- Configuration on R2 looks fine!
- Check the log messages – enable “log-neighbor-changes”

```
%TCP-6-BADAUTH: No MD5 digest from 3.3.3.3:179 to  
2.2.2.2:11272  
%TCP-6-BADAUTH: No MD5 digest from 3.3.3.3:179 to  
2.2.2.2:11272  
%TCP-6-BADAUTH: No MD5 digest from 3.3.3.3:179 to  
2.2.2.2:11272
```

Peer Establishment—Passwords

Cisco.com

```
R3#  
router bgp 2  
  neighbor 2.2.2.2 remote-as 1  
  neighbor 2.2.2.2 ebgp-multihop 2  
  neighbor 2.2.2.2 update-source Loopback0
```

- **Check configuration on R3**

Password is missing from the eBGP configuration

- **Fix the R3 configuration**

Peering should now come up!

But it does not

Peer Establishment—Passwords

Cisco.com

- Let's look at the log messages again for any clues

R2#

```
%TCP-6-BADAUTH: Invalid MD5 digest from  
3.3.3.3:11024 to 2.2.2.2:179
```

```
%TCP-6-BADAUTH: Invalid MD5 digest from  
3.3.3.3:11024 to 2.2.2.2:179
```

```
%TCP-6-BADAUTH: Invalid MD5 digest from  
3.3.3.3:11024 to 2.2.2.2:179
```

- We are getting invalid MD5 digest messages – password mismatch!

Peer Establishment—Passwords

Cisco.com

- **We must have typo'd the password on one of the peering routers**

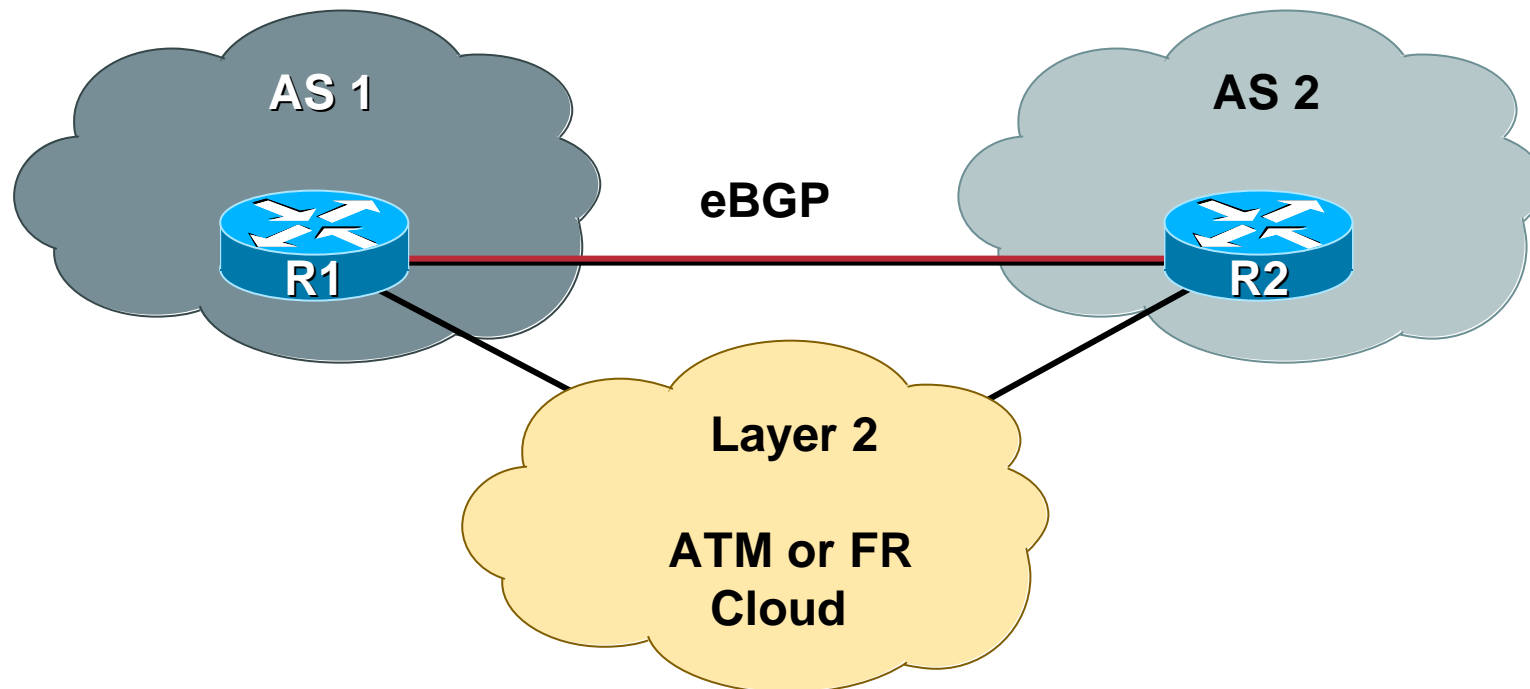
Fix the password – best to re-enter password on both routers

eBGP session now comes up

```
%TCP-6-BADAUTH: Invalid MD5 digest from  
3.3.3.3:11027 to 2.2.2.2:179  
%BGP-5-ADJCHANGE: neighbor 3.3.3.3 Up
```

Flapping Peer—Diagram

Cisco.com



- Symptoms – the eBGP session flaps
- eBGP peering establishes, then drops, re-establishes, then drops,...

Flapping Peer

Cisco.com

- Enable “bgp log-neighbor-changes” so you get a log message when a peer flaps
- R1 and R2 are peering over ATM cloud

R2#

```
%BGP-5-ADJCHANGE: neighbor 1.1.1.1 Down BGP  
Notification sent
```

```
%BGP-3-NOTIFICATION: sent to neighbor 1.1.1.1 4/0  
(hold time expired) 0 bytes
```

```
R2#show ip bgp neighbor 1.1.1.1 | include Last reset
```

```
Last reset 00:01:02, due to BGP Notification sent,  
hold time expired
```

- We are not receiving keepalives from the other side!

Flapping Peer

- Let's take a look at our peer!

```
R1#show ip bgp sum
```

BGP router identifier 172.16.175.53, local AS number 1

BGP table version is 10167, main routing table version 10167

10166 network entries and 10166 paths using 1352078 bytes of memory

1 BGP path attribute entries using 60 bytes of memory

0 BGP route-map cache entries using 0 bytes of memory

0 BGP filter-list cache entries using 0 bytes of memory

BGP activity 10166/300 prefixes, 10166/0 paths, scan interval 15 secs

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
2.2.2.2	4	2	53	284	10167	0	97	00:02:15	0

```
R1#show ip bgp summary | begin Neighbor
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
2.2.2.2	4	2	53	284	10167	0	98	00:03:04	0

- Hellos are stuck in OutQ behind update packets!
- Notice that the MsgSent counter has not moved

Flapping Peer

Cisco.com

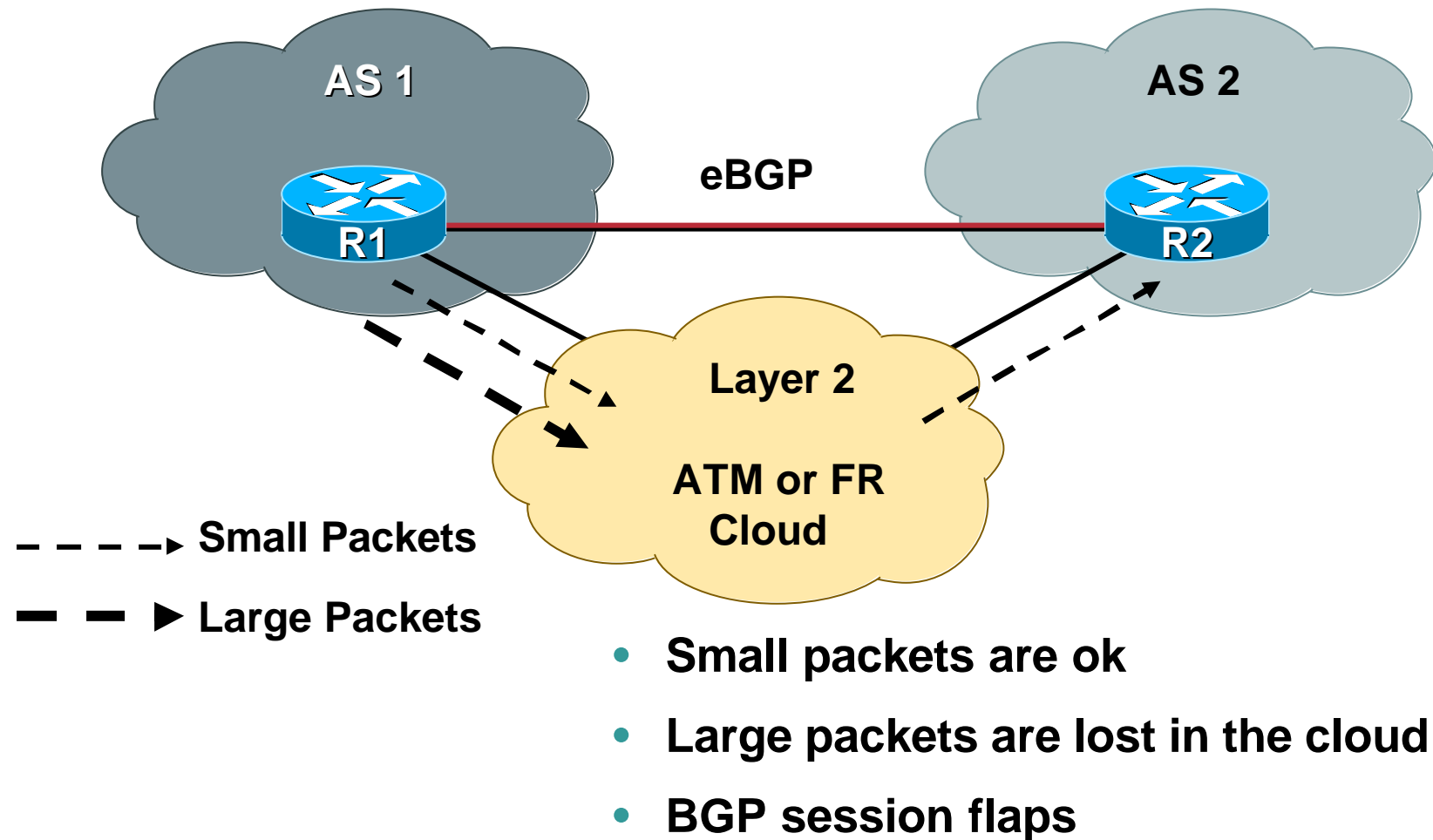
```
R1#ping 2.2.2.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2.2.2.2, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 16/21/24 m
```

```
R1#ping ip
Target IP address: 2.2.2.2
Repeat count [5]:
Datagram size [100]: 1500
Timeout in seconds [2]:
Extended commands [n]:
Sweep range of sizes [n]:
Type escape sequence to abort.
Sending 5, 1500-byte ICMP Echos to 2.2.2.2, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

- Normal pings work but a ping of 1500 fails?

Flapping Peer—Diagram

Cisco.com



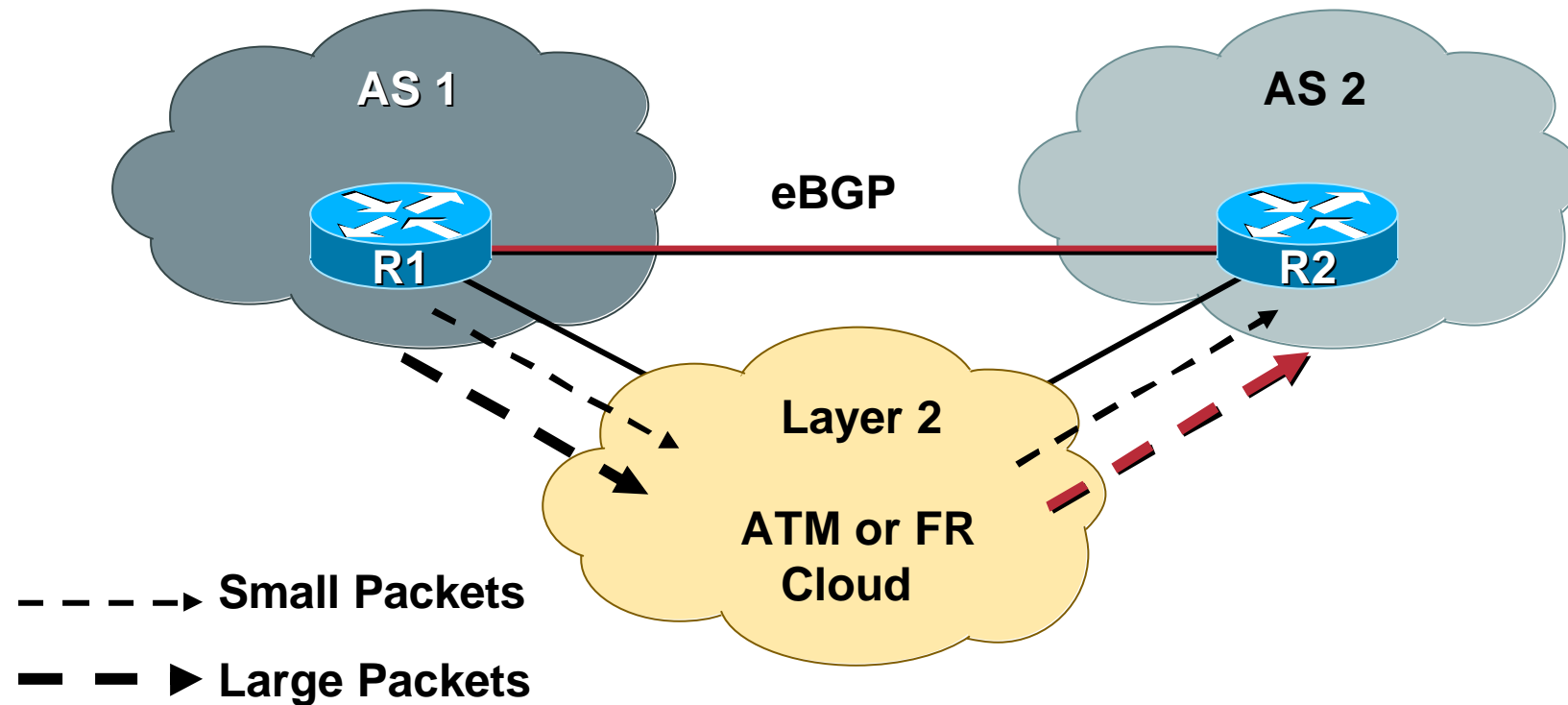
Flapping Peer

Cisco.com

- **Things to check**
 - MTU values**
 - Traffic shaping**
 - Rate-limiting parameters**
- **Looks like a Layer 2 problem**
- **At this point we have verified that BGP is not at fault**
- **Next step is to troubleshoot layer 2...**

Flapping Peer—Diagram

Cisco.com



- Large packets are ok now
- BGP session is stable!

Troubleshooting Tips

Cisco.com

- **Extended ping/traceroute allow you to verify**
 - Loopback to loopback IP connectivity**
 - TTL issues**
- **“show ip bgp summary”**
 - Displays the state of all peers**
- **“show ip bgp neighbor”**
 - Gives a lot of information regarding the peer**

Troubleshooting Tips

Cisco.com

- **“debug ip bgp”**
Should give you a good hint as to why a peer will not establish
- **“debug ip bgp events”**
Displays state transitions for peers
- **“show ip bgp neighbor | include Last reset”**
Will show you the last reset reason for all peers

Agenda

Cisco.com

- **Peer Establishment**
- **Missing Routes**
- **Internet Reachability Problems**

Quick Review

- **Once the session has been established, UPDATES are exchanged**

All the locally known routes

Only the bestpath is advertised

- **Incremental UPDATE messages are exchanged afterwards**

Quick Review

- **Bestpath received from eBGP peer**
Advertise to all peers
- **Bestpath received from iBGP peer**
Advertise only to eBGP peers
A full iBGP mesh must exist

Missing Routes—Agenda

Cisco.com

- **Route Origination**
- **UPDATE Exchange**
- **Filtering**
- **iBGP mesh problems**

Route Origination—Example I

Cisco.com

- **Network statement**

```
R1# show run | include 200.200.0.0
network 200.200.0.0 mask 255.255.252.0
```

- **BGP is not originating the route???**

```
R1# show ip bgp | include 200.200.0.0
R1#
```

- **Do we have the **exact** route?**

```
R1# show ip route 200.200.0.0 255.255.252.0
% Network not in table
```


Route Origination—Example I

Cisco.com

- Nail down routes you want to originate

```
ip route 200.200.0.0 255.255.252.0 Null0 254
```

- Check the RIB

```
R1# show ip route 200.200.0.0 255.255.252.0
      200.200.0.0/22 is subnetted, 1 subnets
S      200.200.0.0 [1/0] via Null 0
```

- BGP originates the route!!

```
R1# show ip bgp | include 200.200.0.0
*> 200.200.0.0/22          0.0.0.0          0          32768
```

Route Origination—Example II

Cisco.com

- Trying to originate an aggregate route

```
aggregate-address 7.7.0.0 255.255.0.0 summary-only
```

- The RIB has a component but BGP does not create the aggregate???

```
R1# show ip route 7.7.0.0 255.255.0.0 longer
      7.0.0.0/32 is subnetted, 1 subnets
C      7.7.7.7 [1/0] is directly connected, Loopback 0
```

```
R1# show ip bgp | i 7.7.0.0
R1#
```

Route Origination—Example II

Cisco.com

- Remember, to have a BGP aggregate you need a **BGP component**, not a RIB (Routing Information Base, a.k.a. the routing table) component

```
R1# show ip bgp 7.7.0.0 255.255.0.0 longer
R1#
```

- Once BGP has a component route we originate the aggregate

```
network 7.7.7.7 mask 255.255.255.255

R1# show ip bgp 7.7.0.0 255.255.0.0 longer
*> 7.7.0.0/16 0.0.0.0 32768 i
s> 7.7.7.7/32 0.0.0.0 0 32768 i
```

- s** means this component is suppressed due to the “summary-only” argument

Troubleshooting Tips

Cisco.com

- **BGP Network statement rules**
Always need an exact route (RIB)
- **aggregate-address looks in the BGP table, not the RIB**
- **“show ip route x.x.x.x y.y.y.y longer”**
Great for finding RIB component routes
- **“show ip bgp x.x.x.x y.y.y.y longer”**
Great for finding BGP component routes

Missing Routes

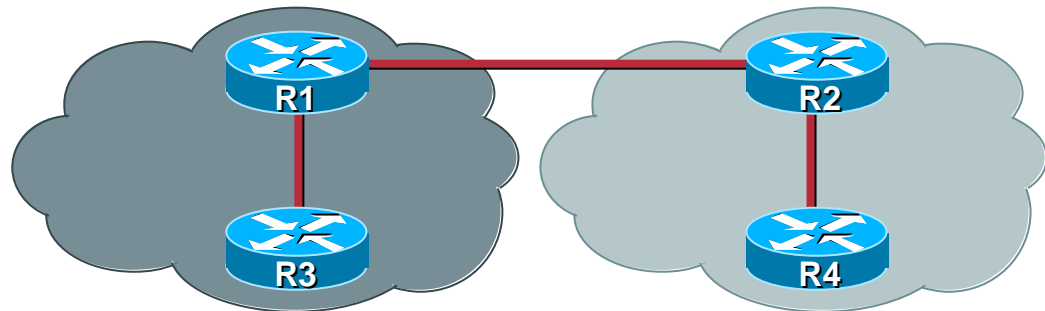
Cisco.com

- **Route Origination**
- **UPDATE Exchange**
- **Filtering**
- **iBGP mesh problems**

Missing Routes—Example I

Cisco.com

- Two RR clusters
- R1 is a RR for R3
- R2 is a RR for R4
- R4 is advertising 7.0.0.0/8
- R2 has the route but R1 and R3 do not?



Missing Routes—Example I

Cisco.com

- First, did R2 advertise the route to R1?

```
R2# show ip bgp neighbors 1.1.1.1 advertised-routes
```

```
BGP table version is 2, local router ID is 2.2.2.2
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i7.0.0.0	4.4.4.4	0	100	0	1

- Did R1 receive it?

```
R1# show ip bgp neighbors 2.2.2.2 routes
```

```
Total number of prefixes 0
```

Missing Routes—Example I

Cisco.com

- Time to debug!!

```
access-list 100 permit ip host 7.0.0.0 host 255.0.0.0
R1# debug ip bgp update 100
```

- Tell R2 to resend his UPDATES

```
R2# clear ip bgp 1.1.1.1 out
```

- R1 shows us something interesting

```
*Mar 1 21:50:12.410: BGP(0): 2.2.2.2 rcv UPDATE w/ attr:
nexthop 4.4.4.4, origin i, localpref 100, metric 0,
originator 100.1.1.1, clusterlist 2.2.2.2, path , community
, extended community
*Mar 1 21:50:12.410: BGP(0): 2.2.2.2 rcv UPDATE about
7.0.0.0/8 -- DENIED due to: ORIGINATOR is us;
```

- Cannot accept an update with our Router-ID as the ORIGINATOR_ID. Another means of loop detection in BGP

Missing Routes—Example I

- R1 and R4 have the same Router-ID

```
R1# show ip bgp summary | include identifier.
```

```
BGP router identifier 100.1.1.1, local AS number 100.
```

```
R4# show ip bgp summary | include identifier.
```

```
BGP router identifier 100.1.1.1, local AS number 100.
```

- Can be a problem in multicast networks; for RP (Rendezvous Point) purposes the same address may be assigned to multiple routers
- Specify a unique Router-ID

```
R1#show run | include router-id  
bgp router-id 1.1.1.1
```

```
R4#show run | include router-id  
bgp router-id 4.4.4.4
```

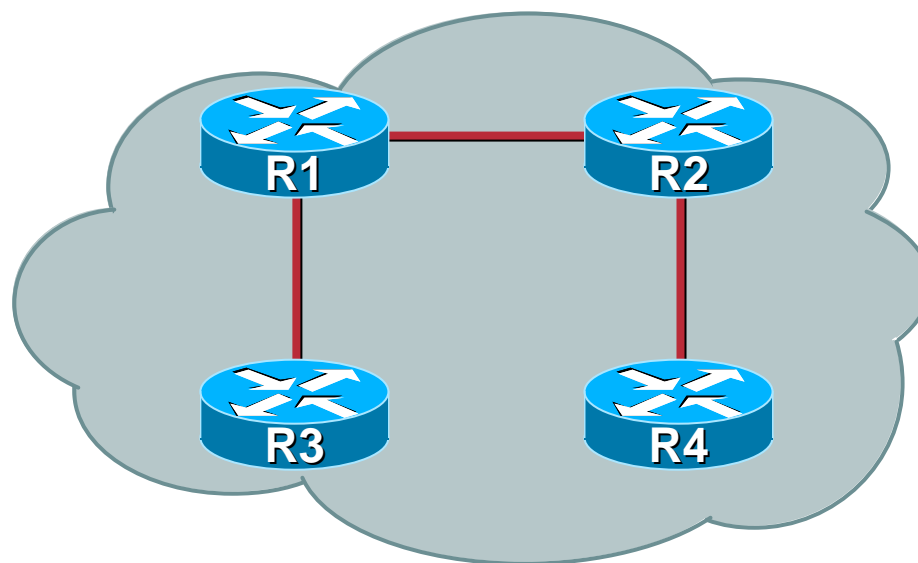
Missing Routes—Example II

Cisco.com

- One RR cluster
- R1 and R2 are RRs
- R3 and R4 are RRCs
- R4 is advertising 7.0.0.0/8

R2 has it

R1 and R3 do not



```
R1#show run | include cluster
bgp cluster-id 10
R2#show run | include cluster
bgp cluster-id 10
```

Missing Routes—Example II

Cisco.com

- Same steps as last time!
- Did R2 advertise it to R1?

```
R2# show ip bgp neighbors 1.1.1.1 advertised-routes
```

```
BGP table version is 2, local router ID is 2.2.2.2
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i7.0.0.0	4.4.4.4	0	100	0	i

- Did R1 receive it?

```
R1# show ip bgp neighbor 2.2.2.2 routes
```

```
Total number of prefixes 0
```

Missing Routes—Example II

Cisco.com

- Time to debug!!

```
access-list 100 permit ip host 7.0.0.0 host 255.0.0.0
R1# debug ip bgp update 100
```

- Tell R2 to resend his UPDATES

```
R2# clear ip bgp 1.1.1.1 out
```

- R1 shows us something interesting

```
*Mar  3 14:28:57.208: BGP(0): 2.2.2.2 rcv UPDATE w/ attr: nexthop
4.4.4.4, origin i, localpref 100, metric 0, originator 4.4.4.4,
clusterlist 0.0.0.10, path , community , extended community
*Mar  3 14:28:57.208: BGP(0): 2.2.2.2 rev UPDATE about 7.0.0.0/8 --
DENIED due to: reflected from the same cluster;
```

- Remember, all RRCs must peer with all RRs in a cluster; allows R4 to send the update directly to R1

Troubleshooting Tips

Cisco.com

- **“show ip bgp neighbor x.x.x.x advertised-routes”**

Lets you see a list of NLRI that you sent a peer

Note: The attribute values shown are taken from the BGP table; attribute modifications by outbound route-maps will not be shown

- **“show ip bgp neighbor x.x.x.x routes”**

Displays routes x.x.x.x sent to us that made it through our inbound filters

- **“show ip bgp neighbor x.x.x.x received-routes”**

Can only use if “soft-reconfig inbound” is configured

Displays all routes received from a peer, even those that were denied

Troubleshooting Tips

Cisco.com

- More on usefulness of “soft-reconfiguration”

Ideal for troubleshooting problems with inbound filters and attributes

- “show ip bgp neighbor x.x.x.x routes”

```
alpha#sh ip bgp neigh 192.168.12.1 routes
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i1.0.0.0	192.168.12.1	0	50	0	i
*>i222.222.0.0/19	192.168.5.1		200	0	3 4 i

- “show ip bgp neighbor x.x.x.x received-routes”

```
alpha#sh ip bgp neigh 192.168.12.1 received-routes
```

Network	Next Hop	Metric	LocPrf	Weight	Path
* i1.0.0.0	192.168.12.1	0	100	0	i
* i169.254.0.0	192.168.5.1	0	100	0	3 i
* i222.222.0.0/19	192.168.5.1		100	0	3 4 i

Troubleshooting Tips

Cisco.com

- **“clear ip bgp x.x.x.x in”**
Ask x.x.x.x to resend his UPDATEs to us
- **“clear ip bgp x.x.x.x out”**
Tells BGP to resend UPDATEs to x.x.x.x
- **“debug ip bgp update”**
Always use an ACL to limit output
Great for troubleshooting “Automatic Denies”
- **“debug ip bgp x.x.x.x update”**
Allows you to debug updates to/from a specific peer
Handy if multiple peers are sending you the same prefix

Missing Routes

Cisco.com

- **Route Origination**
- **UPDATE Exchange**
- **Filtering**
- **iBGP mesh problems**

Update Filtering

Cisco.com

- **Type of filters**
 - Prefix filters**
 - AS_PATH filters**
 - Community filters**
 - Route-maps**
- **Applied incoming and/or outgoing**

Missing Routes—Update Filters

Cisco.com

- **Determine which filters are applied to the BGP session**

show ip bgp neighbors x.x.x.x

show run | include neighbor x.x.x.x

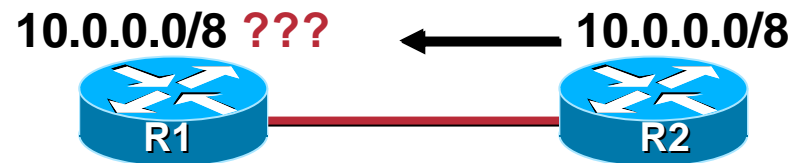
- **Examine the route and pick out the relevant attributes**

show ip bgp x.x.x.x

- **Compare the attributes against the filters**

Missing Routes—Update Filters

Cisco.com



- Missing 10.0.0.0/8 in R1 (1.1.1.1)
- Not received from R2 (2.2.2.2)

```
R1#show ip bgp neigh 2.2.2.2 routes  
  
Total number of prefixes 0
```

Missing Routes—Update Filters

Cisco.com

- R2 originates the route
- Does not advertise it to R1

```
R2#show ip bgp neigh 1.1.1.1 advertised-routes
```

Network	Next Hop	Metric	LocPrf	Weight	Path
---------	----------	--------	--------	--------	------

```
R2#show ip bgp 10.0.0.0
```

```
BGP routing table entry for 10.0.0.0/8, version 1660
```

```
Paths: (1 available, best #1)
```

```
Not advertised to any peer
```

```
Local
```

```
0.0.0.0 from 0.0.0.0 (2.2.2.2)
```

```
Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best
```

Missing Routes—Update Filters

Cisco.com

- Time to check filters!
- ^ matches the beginning of a line
- \$ matches the end of a line
- ^\$ means match any empty AS_PATH
- Filter “looks” correct

```
R2#show run | include neighbor 1.1.1.1
neighbor 1.1.1.1 remote-as 3
neighbor 1.1.1.1 filter-list 1 out
```

```
R2#sh ip as-path 1
AS path access list 1
  permit ^$
```

Missing Routes—Update Filters

Cisco.com

```
R2#show ip bgp filter-list 1
```

```
R2#show ip bgp regexp ^$
```

BGP table version is 1661, local router ID is 2.2.2.2

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.0.0.0	0.0.0.0	0	32768	i	

- Nothing matches the filter-list???
- Re-typing the regexp gives the expected output

Missing Routes—Update Filters

Cisco.com

- **Copy and paste** the entire regexp line from the configuration

```
R2#show ip bgp regexp ^$
```

Nothing matches again! Let's use the up arrow key to see where the cursor stops

```
R2#show ip bgp regexp ^$
```



End of Line Is at the Cursor

- There is a trailing white space at the end
- It is considered part of the regular expression

Missing Routes—Update Filters

Cisco.com

- Force R2 to resend the update after the filter-list correction
- Then check R1 to see if it has the route

```
R2#clear ip bgp 1.1.1.1 out
```

```
R1#show ip bgp 10.0.0.0  
% Network not in table
```

- R1 still does not have the route
- Time to check R1's inbound policy for R2

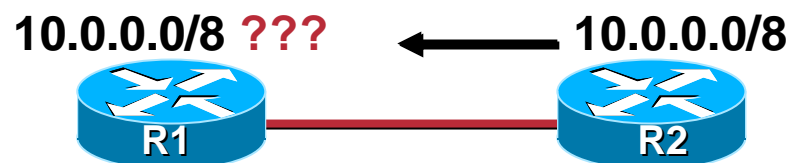
Missing Routes—Update Filters

Cisco.com

```
R1#show run | include neighbor 2.2.2.2
neighbor 2.2.2.2 remote-as 12
neighbor 2.2.2.2 route-map POLICY in
R1#show route-map POLICY
route-map POLICY, permit, sequence 10
  Match clauses:
    ip address (access-lists): 100 101
    as-path (as-path filter): 1
  Set clauses:
    Policy routing matches: 0 packets, 0 bytes
R1#show access-list 100
Extended IP access list 100
    permit ip host 10.0.0.0 host 255.255.0.0
R1#show access-list 101
Extended IP access list 101
    permit ip 200.1.0 0.0.0.255 host 255.255.255.0
R1#show ip as-path 1
AS path access list 1
    permit ^12$
```

Missing Routes—Update Filters

Cisco.com



- **Confused? Let's run some debugs**

```
R1#show access-list 99
```

```
Standard IP access list 99
```

```
permit 10.0.0.0
```

```
R1#debug ip bgp 2.2.2.2 update 99
```

```
BGP updates debugging is on for access list 99 for neighbor 2.2.2.2
```

```
R1#
```

```
4d00h: BGP(0): 2.2.2.2 rcvd UPDATE w/ attr: nexthop 2.2.2.2, origin i,  
metric 0, path 12
```

```
4d00h: BGP(0): 2.2.2.2 rcvd 10.0.0.0/8 -- DENIED due to: route-map;
```

Missing Routes—Update Filters

Cisco.com

```
R1#sh run | include neighbor 2.2.2.2
neighbor 2.2.2.2 remote-as 12
neighbor 2.2.2.2 route-map POLICY in
R1#sh route-map POLICY
route-map POLICY, permit, sequence 10
  Match clauses:
    ip address (access-lists): 100 101
    as-path (as-path filter): 1
  Set clauses:
    Policy routing matches: 0 packets, 0 bytes
R1#sh access-list 100
Extended IP access list 100
  permit ip host 10.0.0.0 host 255.255.0.0
R1#sh access-list 101
Extended IP access list 101
  permit ip 200.1.1.0 0.0.0.255 host 255.255.255.0
R1#sh ip as-path 1
AS path access list 1
  permit ^12$
```

Missing Routes—Update Filters

Cisco.com

- **Wrong mask! Needs to be /8 and the ACL allows a /16 only!**

Extended IP access list 100

```
permit ip host 10.0.0.0 host 255.255.0.0
```

- **Should be**

Extended IP access list 100

```
permit ip host 10.0.0.0 host 255.0.0.0
```

- **Use prefix-list instead, more difficult to make a mistake**

```
ip prefix-list my_filter permit 10.0.0.0/8
```

- **What about ACL 101?**

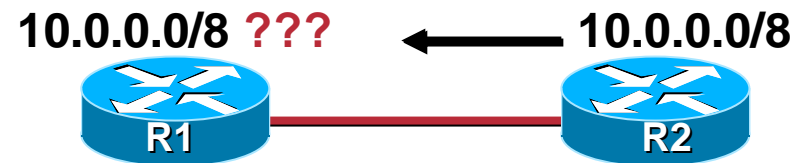
Multiple matches on the same line are ORed

Multiple matches on different lines are ANDed

- **ACL 101 does not matter because ACL 100 matches which satisfies the OR condition**

Missing Routes—Community Problems

Cisco.com



- Missing 10.0.0.0/8 in R1 (1.1.1.1)
- Not received from R2 (2.2.2.2)

```
R1#show ip bgp neigh 2.2.2.2 routes  
  
Total number of prefixes 0
```

Missing Routes—Community Problems

Cisco.com

- R2 originates the route

```
R2#show ip bgp 10.0.0.0
```

```
BGP routing table entry for 10.0.0.0/8, version 1660
```

```
Paths: (1 available, best #1)
```

```
Not advertised to any peer
```

```
Local
```

```
0.0.0.0 from 0.0.0.0 (2.2.2.2)
```

```
Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best
```

- But the community is not set

Would be displayed in the “sh ip bgp” output

Missing Routes—Community Problems

Cisco.com

- Fix the configuration so community is set

```
R2#show run | begin bgp
router bgp 2
  network 10.0.0.0 route-map set-community
...
route-map set-community permit 10
  set community 2:2 1:50
```

```
R2#show ip bgp 10.0.0.0
```

```
BGP routing table entry for 10.0.0.0/8, version 1660
```

```
Paths: (1 available, best #1)
```

```
Not advertised to any peer
```

```
Local
```

```
0.0.0.0 from 0.0.0.0 (2.2.2.2)
```

```
Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best
```

```
Community 2:2 1:50
```

Missing Routes—Community Problems

Cisco.com

- R2 now advertises prefix with community to R1
- But R1 still doesn't see the prefix

R1 insists there is nothing wrong with their configuration

```
R1#show ip bgp neigh 2.2.2.2 routes  
  
Total number of prefixes 0
```

- Configuration verified on R2
- No filters blocking announcement on R2
- So what's wrong?

Missing Routes—Community Problems

Cisco.com

- Check R2 configuration again!

```
R2#show run | begin bgp
router bgp 2
  network 10.0.0.0 route-map set-community
  neighbor 1.1.1.1 remote-as 1
  neighbor 1.1.1.1 prefix-list my-agg out
  neighbor 1.1.1.1 prefix-list their-agg in
!
ip prefix-list my-agg permit 10.0.0.0/8
ip prefix-list their-agg permit 20.0.0.0/8
!
route-map set-community permit 10
  set community 2:2 1:50
```

- Looks okay – filters okay, route-map okay
- But forgotten “neighbor 1.1.1.1 send-community”

Cisco IOS does NOT send communities by default

Missing Routes—Community Problems

Cisco.com

- R2 now advertises prefix with community to R1
- But R1 still doesn't see the prefix

Nothing wrong on R2 now, so turn attention to R1

```
R1#show run | begin bgp
router bgp 1
  neighbor 2.2.2.2 remote-as 2
  neighbor 2.2.2.2 route-map R2-in in
  neighbor 2.2.2.2 route-map R1-out out
!
ip community-list 1 permit 1:150
!
route-map R2-in permit 10
  match community 1
  set local-preference 150
```

Missing Routes—Community Problems

Cisco.com

- Community match on R1 expects 1:150 to be set on prefix
- But R2 is sending 1:50

Typo or miscommunication between operations?

- R2 is also using the route-map to filter

If the prefix does not have community 1:150 set, it is dropped
– there is no next step in the route-map

Watch the route-map rules in Cisco IOS – they are basically:

if <match> then <set> and exit route-map

else if <match> then <set> and exit route-map

else if <match> then <set> etc...

Blank route-map line means **match** everything, **set** nothing

Missing Routes—Community Problems

Cisco.com

- Fix configuration on R2 to set community 1:150 on announcements to R1
- Fix configuration on R1 to also permit prefixes not matching the route-map – troubleshooting is easier with prefix-filters doing the filtering

```
R1#show run | begin ^route-map
route-map R2-in permit 10
  match community 1
  set local-preference 150
route-map R2-in permit 20
```

```
R1#show ip bgp neigh 2.2.2.2 routes
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	10.0.0.0	2.2.2.2	0		0	2 i

```
Total number of prefixes 1
```

Missing Routes—Community Problems

Cisco.com

- **Watch route-maps**

Route-map rules often catch out operators when they are used for filtering

Absence of an appropriate match means the prefix will be **discarded**

- **Don't forget to configure “send-community”**

Include it in your default template for iBGP

It should be iBGP default in a Service Provider Network

Remember that it is required to send communities for eBGP too

Missing Routes—General Problems

Cisco.com

- **Stick to simple policy rules:**
 - Prefix-lists** ® filter prefix announcements
 - Filter-lists** ® filter on AS-paths
 - Route-maps** ® apply policies
- **By applying policies I mean setting attributes on groups of prefixes, rather than simply filtering**

Missing Routes

Cisco.com

- **Route Origination**
- **UPDATE Exchange**
- **Filtering**
- **iBGP mesh problems**

Missing Routes—iBGP

Cisco.com

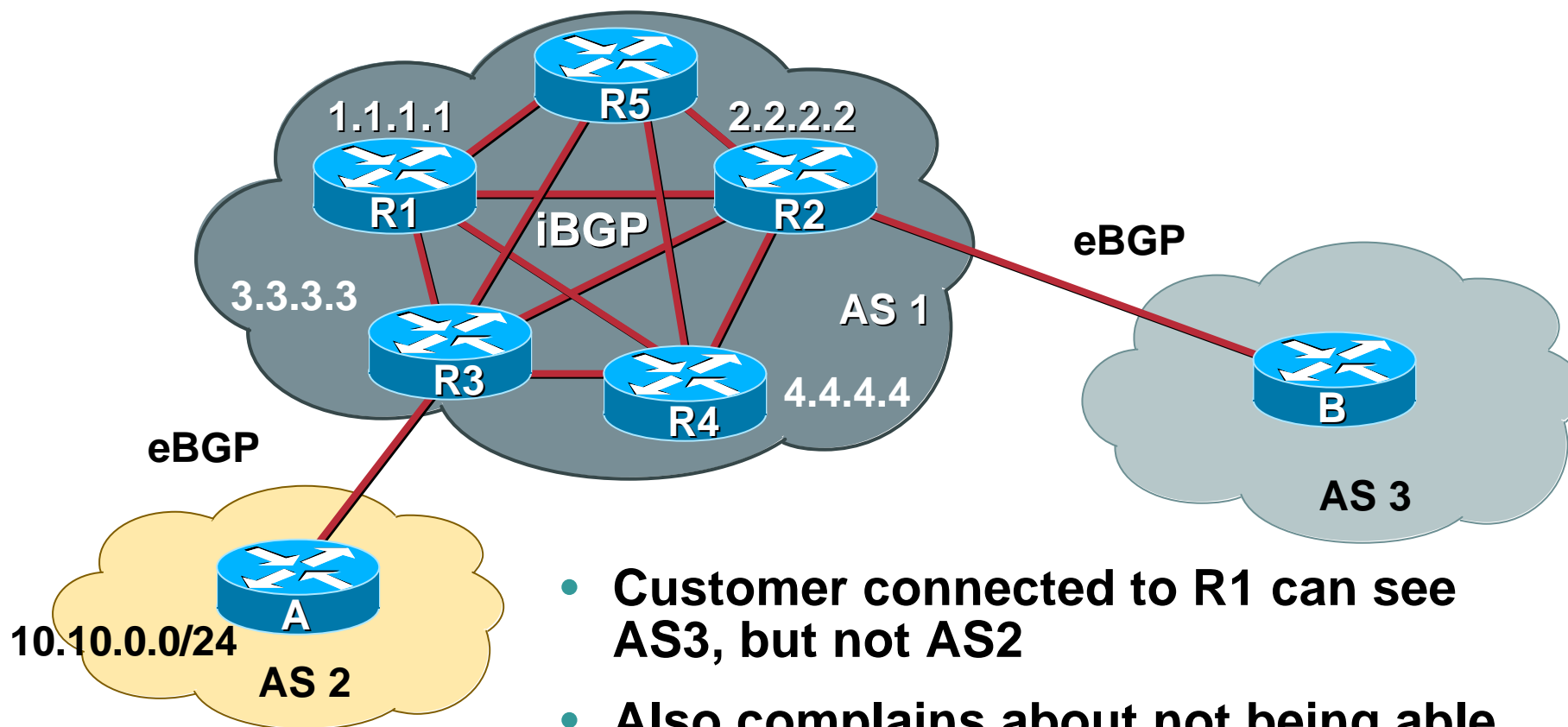
- **Symptom: customer complains about patchy Internet access**

Can access some, but not all, sites connected to backbone

Can access some, but not all, of the Internet

Missing Routes—iBGP

Cisco.com



- Customer connected to R1 can see AS3, but not AS2
- Also complains about not being able to see sites connected to R5
- No complaints from other customers

Missing Routes—iBGP

Cisco.com

- **Diagnosis: This is the classic iBGP mesh problem**

The full mesh isn't complete – how do we know this?

- **Customer is connected to R1**

Can't see AS2 \Rightarrow R3 is somehow not passing routing information about AS2 to R1

Can't see R5 \Rightarrow R5 is somehow not passing routing information about sites connected to R5

But can see rest of the Internet \Rightarrow his prefix is being announced to some places, so not an iBGP origination problem

Missing Routes—iBGP

Cisco.com

```
R3#sh ip bgp sum | begin ^Neigh
Neighbor      V      AS MsgRcvd MsgSent  TblVer  InQ  OutQ Up/Down  State/PfxRcd
1.1.1.1        4       1     200     20      32    0    0 3d10h   Active
2.2.2.2        4       1     210     25      32    0    0 3d16h     15
4.4.4.4        4       1     213     22      32    0    0 3d16h     12
5.5.5.5        4       1     215     19      32    0    0 3d16h      0
10.10.10.10    4       2    2501    2503     32    0    0 3d16h    100
R3#
```

- **BGP summary shows that the peering with router R1 is down**

Up/Down is 3 days 10 hours, yet active

Which means it was last up 3 days and 10 hours ago

So something has broken between R1 and R3

Missing Routes—iBGP

Cisco.com

- Now check configuration on R1

```
R1#sh conf | b bgp
router bgp 1
  neighbor iBGP-ipv4-peers peer-group
  neighbor iBGP-ipv4-peers remote-as 1
  neighbor iBGP-ipv4-peers update-source Loopback0
  neighbor iBGP-ipv4-peers send-community
  neighbor iBGP-ipv4-peers prefix-list ibgp-prefixes out
  neighbor 2.2.2.2 peer-group iBGP-ipv4-peers
  neighbor 4.4.4.4 peer-group iBGP-ipv4-peers
  neighbor 5.5.5.5 peer-group iBGP-ipv4-peers
```

- Where is the peering with R3?
- Restore the missing line, and the iBGP with R3 comes back up

Missing Routes—iBGP

Cisco.com

```
R3#sh ip bgp sum | begin ^Neigh
Neighbor      V      AS  MsgRcvd  MsgSent   TblVer   InQ  OutQ  Up/Down   State/PfxRcd
1.1.1.1        4       1     200      20        32    0    0  00:00:50      8
2.2.2.2        4       1     210      25        32    0    0  3d16h       15
4.4.4.4        4       1     213      22        32    0    0  3d16h       12
5.5.5.5        4       1     215      19        32    0    0  3d16h        0
10.10.10.10    4       2    2501    2503        32    0    0  3d16h      100
R3#
```

- **BGP summary shows that no prefixes are being heard from R5**

This could be due to inbound filters on R3 on the iBGP with R5

But there were no filters in the configuration on R3

This must be due to outbound filters on R5 on the iBGP with R3

Missing Routes—iBGP

Cisco.com

- Now check configuration on R5

```
R5#sh conf | b neighbor 3.3.3.3
neighbor 3.3.3.3 remote-as 1
neighbor 3.3.3.3 update-source loopback0
neighbor 3.3.3.3 prefix-list ebgp-filters out
neighbor 4.4.4.4 remote-as 1
neighbor 4.4.4.4 update-source loopback0
neighbor 4.4.4.4 prefix-list ibgp-filters out
!
ip prefix-list ebgp-filters permit 20.0.0.0/8
ip prefix-list ibgp-filters permit 10.0.0.0/8
```

- Error in prefix-list in R3 iBGP peering

ebgp-filters has been used instead of **ibgp-filters**

Typo – another advantage of using peer-groups!

Missing Routes—iBGP

Cisco.com

- Fix the prefix-list on R5
- Check the iBGP again on R3

Peering with R1 is up

Peering with R5 has prefixes

- Confirm that all is okay with customer

```
R3#sh ip bgp sum | begin ^Neigh
Neighbor      V      AS  MsgRcvd  MsgSent   TblVer   InQ  OutQ  Up/Down   State/PfxRcd
1.1.1.1        4        1     200      20        32     0    0  00:01:53         8
2.2.2.2        4        1     210      25        32     0    0  3d16h         15
4.4.4.4        4        1     213      22        32     0    0  3d16h         12
5.5.5.5        4        1     215      19        32     0    0  3d16h          6
10.10.10.10    4        2    2501    2503        32     0    0  3d16h        100
R3#
```

Troubleshooting Tips

Cisco.com

- **Watch the iBGP full mesh**

Use peer-groups both for efficiency and to avoid making policy errors within the iBGP mesh

Use route reflectors to avoid accidentally missing iBGP peers, especially as the mesh grows in size

Troubleshooting Tips

Cisco.com

- **“show ip as-path-access-list”**
Displays the filter
- **“show ip bgp filter-list”**
Displays BGP paths that match the filter
- **“show ip bgp regexp”**
Displays BGP paths that match the as-path regular expression; handy for troubleshooting filter-list issues

Troubleshooting Tips

Cisco.com

- **“show ip community-list”**
Displays the filter
- **“show ip bgp community-list”**
Displays BGP paths that match the filter
- **“show ip prefix-list”**
Displays the filter
Prefix-lists are generally easier to use than ACLs
- **“show ip bgp prefix-list”**
Displays BGP paths that match the filter

Troubleshooting Tips

Cisco.com

- **“show route-map”**
Displays the filter
- **“show ip bgp route-map”**
Displays BGP paths that match the filter
- **“show access-list”**
Displays the filter
- **debug ip bgp update ACL**
After going through the config, debug!
Don't forget the ACL

Agenda

Cisco.com

- **Peer Establishment**
- **Missing Routes**
- **Internet Reachability Problems**

Internet Reachability Problems

Cisco.com

- **BGP Attribute Confusion**

To Control Traffic in ® Send MEDs and AS-PATH prepends on outbound announcements

To Control Traffic out ® Attach local-preference to inbound announcements

- **Troubleshooting of multihoming and transit is often hampered because the relationship between routing information flow and traffic flow is forgotten**

Internet Reachability Problems

Cisco.com

- **BGP Path Selection Process**

Each vendor has “tweaked” the path selection process

Know it, learn it, for your router equipment – saves time later

- **MED confusion**

Default MED on Cisco IOS is ZERO – it may not be this on your router, or your peer’s router

Internet Reachability Problems

Cisco.com

- **Community confusion**

set community does just that – it overwrites any other community set on the prefix

Use **additive** keyword to add community to existing list

Use Internet format for community (AS:xx) not the 32-bit IETF format

Cisco IOS never sends community by default

Other implementations may send community by default for iBGP and/or eBGP

Never assume that your neighbouring AS will honour your **no-export** community – ask first!

Internet Reachability Problems

Cisco.com

- **AS-PATH prepends**

20 prepends won't lessen the priority of your path any more than 10 prepends will – check it out at a Looking Glass

The Internet is on average only 5 ASes deep, maximum AS prepend most ISPs have to use is around this too

Know your BGP path selection algorithm

Some ISPs use **bgp maxas-path 15 to drop prefixes with ridiculously long AS-paths**

Internet Reachability Problems

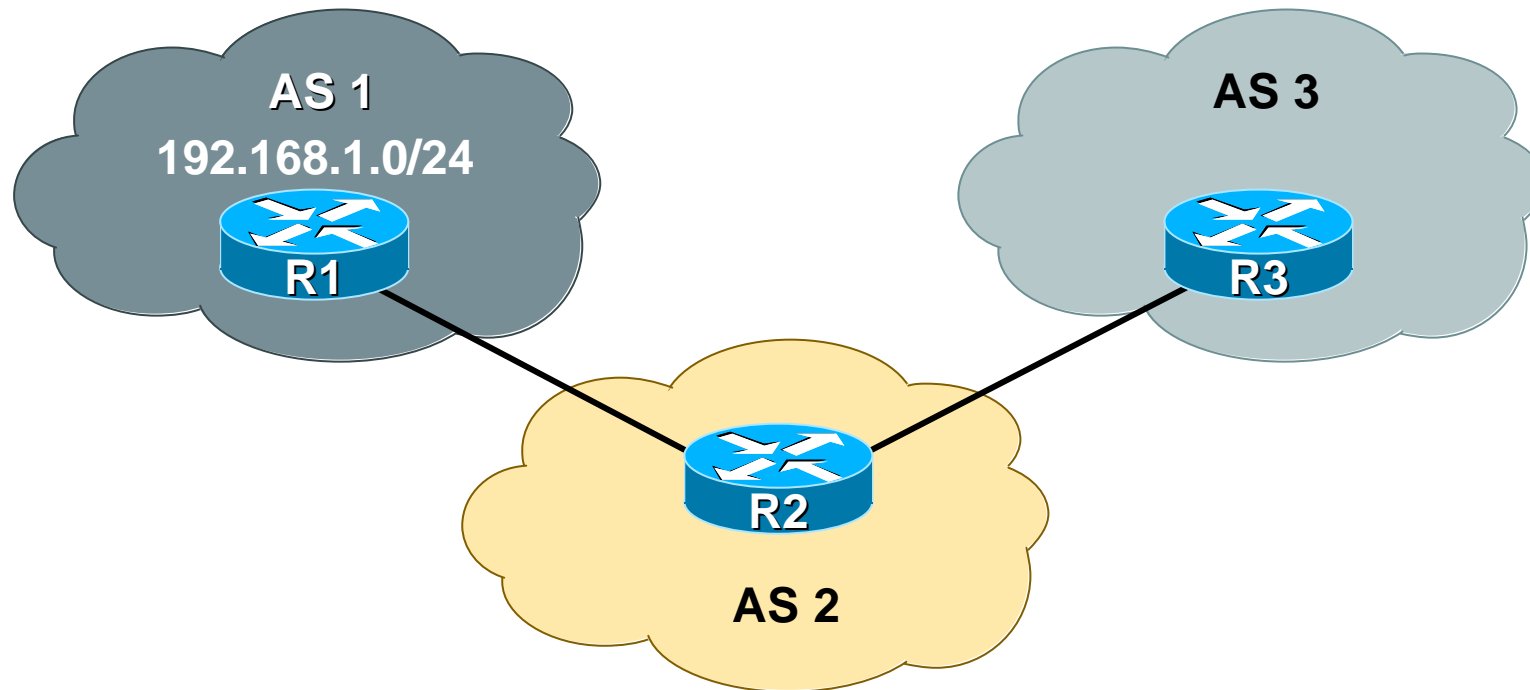
Cisco.com

- Private ASes should not ever appear in the Internet
- Cisco IOS **remove-private-AS** command does not remove every instance of a private AS
 - e.g. won't remove private AS appearing in the middle of a path surrounded by public ASNs

www.cisco.com/warp/public/459/32.html
- Apparent non-removal of private-ASNs may not be a bug, but a configuration error somewhere else

Troubleshooting Connectivity – Example I

Cisco.com



- **Symptom: AS1 announces 192.168.1.0/24 to AS2 but AS3 cannot see the network**

Troubleshooting Connectivity – Example I

Cisco.com

- **Checklist:**

AS1 announces, but does AS2 see it?

We are checking eBGP filters on R1 and R2. Remember that R2 access will require cooperation and assistance from your peer

Does AS2 see it over entire network?

We are checking iBGP across AS2's network (unnecessary step in this case, but usually the next consideration). Quite often iBGP is misconfigured, lack of full mesh, problems with RRs, etc.

Troubleshooting Connectivity – Example I

Cisco.com

- **Checklist:**

Does AS2 send it to AS3?

We are checking eBGP configuration on R2. There may be a configuration error with as-path filters, or prefix-lists, or communities such that only local prefixes get out

Does AS3 see all of AS2's originated prefixes?

We are checking eBGP configuration on R3. Maybe AS3 does not know to expect prefixes from AS1 in the peering with AS2, or maybe it has similar errors in as-path or prefix or community filters

Troubleshooting Connectivity – Example I

Cisco.com

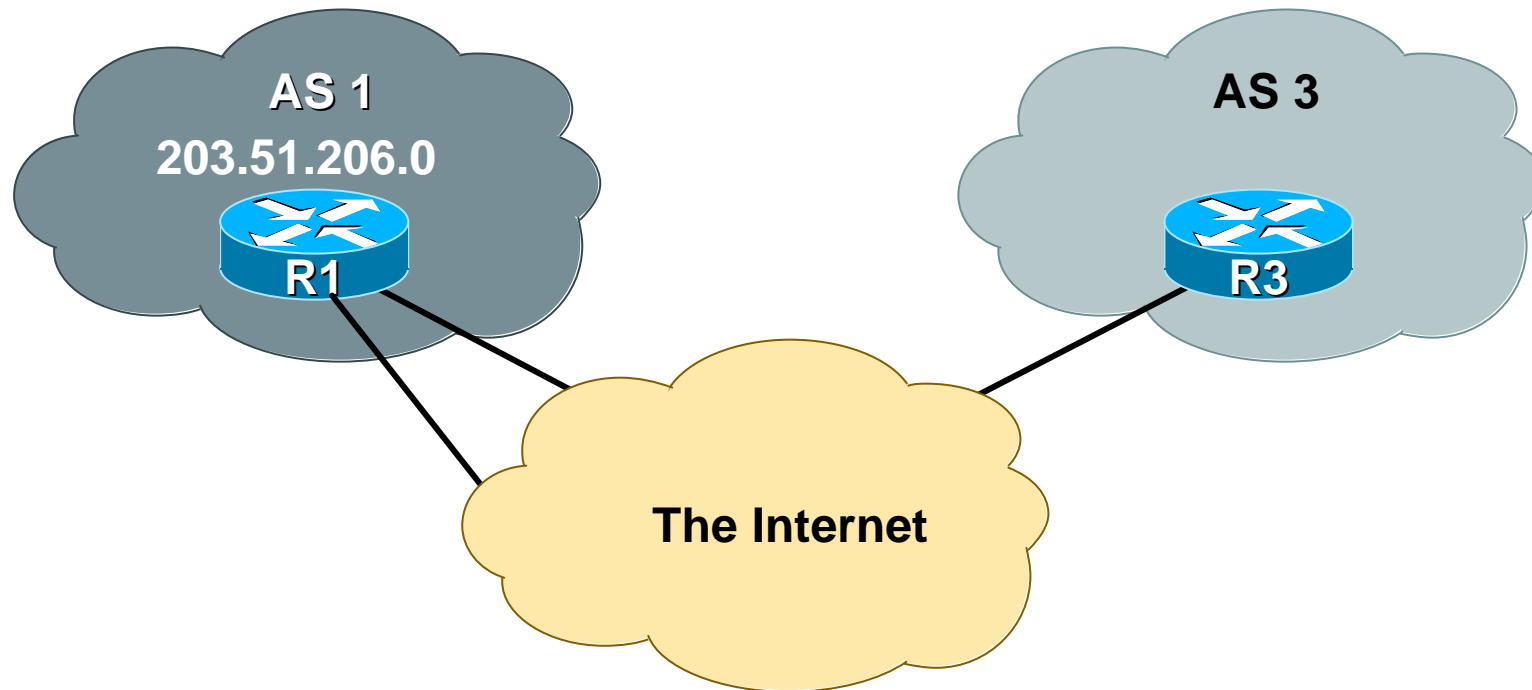
- **Troubleshooting connectivity beyond immediate peers is much harder**

Relies on your peer to assist you – they have the relationship with their BGP peers, not you

Quite often connectivity problems are due to the private business relationship between the two neighbouring ASNs

Troubleshooting Connectivity – Example II

Cisco.com



- **Symptom: AS1 announces 203.51.206.0/24 to its upstreams but AS3 cannot see the network**

Troubleshooting Connectivity – Example II

Cisco.com

- **Checklist:**

AS1 announces, but do its upstreams see it?

We are checking eBGP filters on R1 and upstreams.
Remember that upstreams will need to be able to help
you with this

Is the prefix visible anywhere on the Internet?

We are checking if the upstreams are announcing the
network to anywhere on the Internet. See next slides
on how to do this.

Troubleshooting Connectivity – Example II

Cisco.com

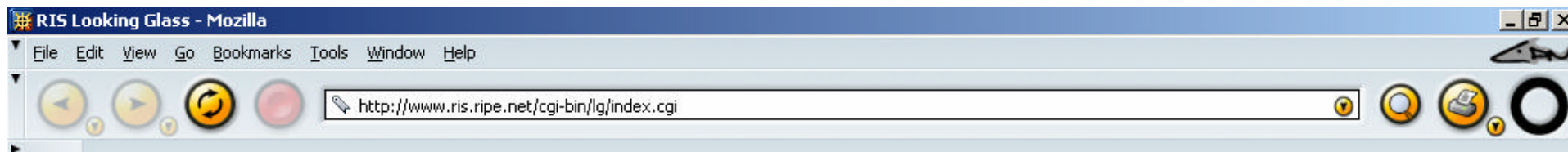
- Help is at hand – the Looking Glass
- Many networks around the globe run Looking Glasses

These let you see the BGP table and often run simple ping or traceroutes from their sites

www.traceroute.org for IPv4

www.traceroute6.org for IPv6

- Many still use the original: nitrous.digex.net
- Next slides have some examples of a typical looking glass in action



[homepage](#) | [what's new](#) | [whois db](#) | [search](#) | [site map](#) | [f.a.q.](#)

RIS - Looking Glass

RRC Box:

Query:

- ☐ bgp
- ☒ bgp sum
- ☐ bgp neig
- ☐ bgp rege
- ☐ bgp path
- ☐ version
- ☐ traceroute
- ☐ ping

RRC00, Amsterdam
RRC01, LINX
RRC02, SFINX
RRC03, AMS-IX
RRC04, CIXP
RRC05, VIX
RRC06, NSPIX2
RRC07, Netnod
RRC08, MAE-West

Argument:

Multi-Router Looking Glass version 3.3.2 Beta

Written by: John Fraizer - [EnterZone, Inc](#)

[Contact Webmaster](#)

Copyright © RIPE NCC

[Mail RIPE NCC](#)



Document: Done (5.128 secs)

RIS Looking Glass - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://www.ris.ripe.net/cgi-bin/lg/index.cgi

RRC Box: RRC01, LINX

Query:

- ☒ bgp
- ☐ bgp summary
- ☐ bgp neighbors
- ☐ bgp regexp
- ☐ bgp paths
- ☐ version
- ☐ traceroute
- ☐ ping

Argument: 203.51.206.0 Execute

BGP routing table entry for 203.48.0.0/14
Paths: (3 available, best #2, table Default-IP-Routing-Table)
Not advertised to any peer
13237 1 4637 1221, (aggregated by 1221 203.62.252.26)
195.66.224.99 from 195.66.224.99 (80.245.35.6)
Origin IGP, localpref 100, valid, external, atomic-aggregate
Community: 13237:44693
Last update: Fri Oct 18 09:24:43 2002

286 209 4637 1221, (aggregated by 1221 203.62.252.26)
195.66.224.54 from 195.66.224.54 (134.222.86.174)
Origin IGP, localpref 100, valid, external, atomic-aggregate, best
Last update: Wed Oct 16 18:16:29 2002

8406 8210 1239 4637 1221, (aggregated by 1221 203.62.252.26)
195.66.226.71 from 195.66.226.71 (62.72.156.25)
Origin IGP, metric 0, localpref 100, valid, external, atomic-aggregate
Last update: Tue Oct 15 14:33:06 2002

rrc01

Multi-Router Looking Glass version 3.3.2 Beta
Written by John Fraizer, EnterZone, Inc.

Document: Done (2.033 secs)

Troubleshooting Connectivity – Example II

Cisco.com

- **Hmmm....**
- **Looking Glass can see 203.48.0.0/14**
This includes 203.51.206.0/24
So the problem must be with AS3, or AS3's upstream
- **A traceroute confirms the connectivity**

RIS Looking Glass - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://www.ris.ripe.net/cgi-bin/lg/index.cgi

RIS - Looking Glass

RRC Box:

Query:

- ☐ bgp
- ☐ bgp summary
- ☐ bgp neighbors
- ☐ bgp regexp
- ☐ bgp paths
- ☐ version
- ☒ traceroute
- ☐ ping

Argument:

Traceroute from **RRC011** to **203.51.206.206**.

```
traceroute to 203.51.206.206 (203.51.206.206), 30 hops max, 40 byte packets
 1  collector.linx.net (195.66.225.254)  0.629 ms  0.580 ms  0.607 ms
 2  195.66.224.166 (195.66.224.166)  0.615 ms  0.464 ms  0.487 ms
 3  pos3-0.lnx01.London.net.reach.com (202.40.148.34)  0.685 ms  0.680 ms  0.585 ms
 4  202.84.143.138 (202.84.143.138)  72.142 ms  72.283 ms  72.363 ms
 5  202.84.143.85 (202.84.143.85)  104.273 ms  104.375 ms  104.274 ms
 6  202.84.143.57 (202.84.143.57)  170.109 ms  170.195 ms  169.894 ms
 7  202.84.143.22 (202.84.143.22)  357.587 ms  357.811 ms  357.514 ms
 8  GigabitEthernet1-2.pad-core4.Sydney.telstra.net (203.50.13.245)  357.674 ms  359.339 ms  357.256 ms
 9  Pos12-0.ken-core4.Sydney.telstra.net (203.50.6.21)  357.431 ms  357.326 ms  357.311 ms
10  Pos4-0.woo-core1.Brisbane.telstra.net (203.50.6.222)  368.377 ms  369.083 ms  371.041 ms
11  GigabitEthernet1-2.cha23.Brisbane.telstra.net (203.50.50.33)  369.113 ms  368.800 ms  370.978 ms
12  tcbroad1.lnk.telstra.net (139.130.193.118)  368.616 ms  369.616 ms  368.675 ms
13  gigabit-msfc1.qld-remote.bigpond.net.au (61.9.209.4)  368.455 ms  368.517 ms  370.156 ms
14  * * *
15  * * *
16  CPE-203-51-206-206.qld.bigpond.net.au (203.51.206.206)  387.364 ms  387.103 ms *
```

Document: Done (73.836 secs)

Troubleshooting Connectivity – Example II

Cisco.com

- **Help is at hand – RouteViews**
- **The RouteViews router has BGP feeds from around 60 peers**

www.routeviews.org explains the project

Gives access to a real router, and allows any provider to find out how their prefixes are seen in various parts of the Internet

Complements the Looking Glass facilities

- **Anyway, back to our problem...**

Troubleshooting Connectivity – Example II

Cisco.com

- **Checklist:**

Does AS3's upstream send it to AS3?

We are checking eBGP configuration on AS3's upstream. There may be a configuration error with as-path filters, or prefix-lists, or communities such that only local prefixes get out. This needs AS3's assistance.

Does AS3 see any of AS1's originated prefixes?

We are checking eBGP configuration on R3. Maybe AS3 does not know to expect the prefix from AS1 in the peering with its upstream, or maybe it has some errors in as-path or prefix or community filters

Troubleshooting Connectivity – Example II

Cisco.com

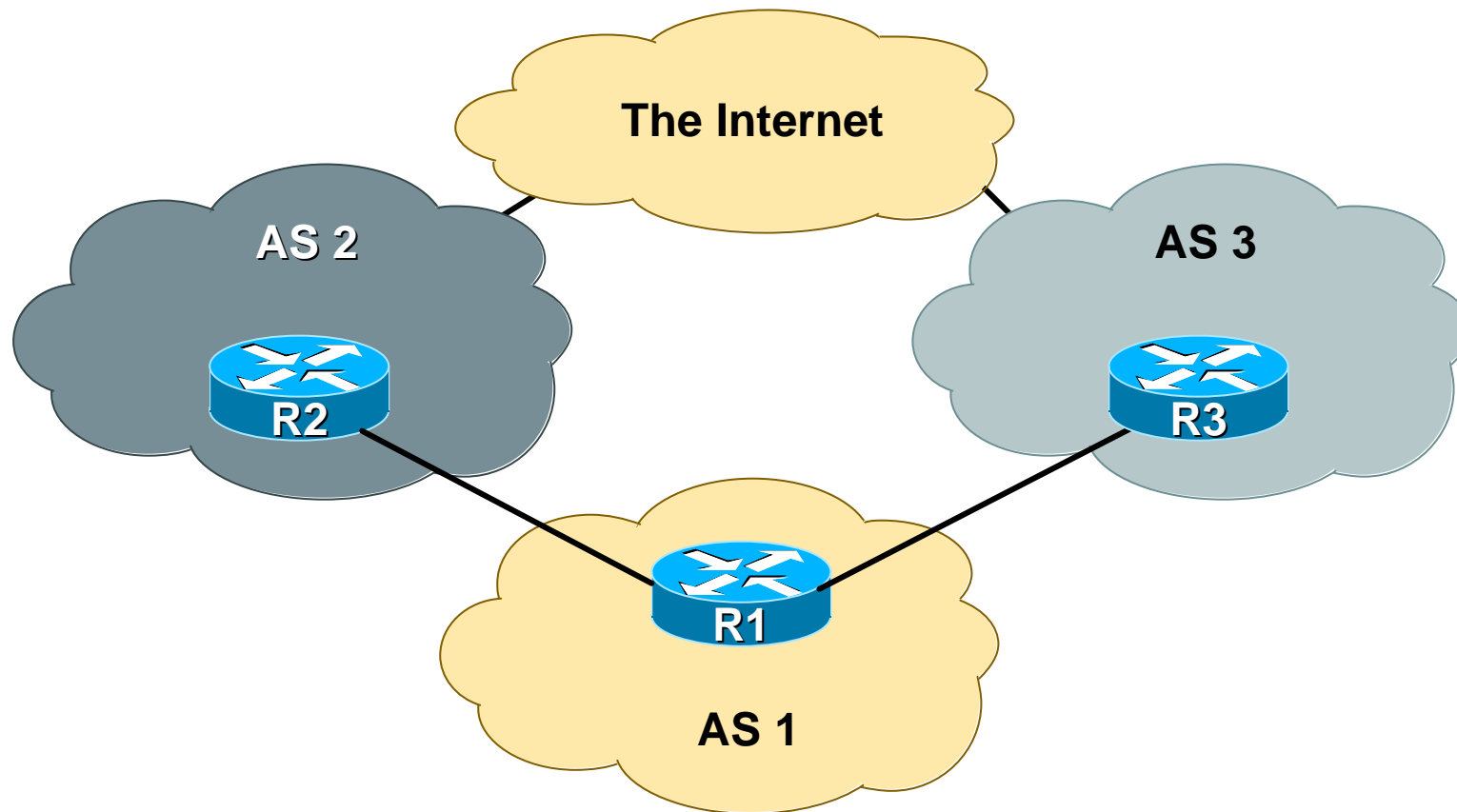
- **Troubleshooting across the Internet is harder**
But tools are available
- **Looking Glasses, offering traceroute, ping and BGP status are available all over the globe**

Most connectivity problems seem to be found at the edge of the network, rarely in the transit core

Problems with the transit core are usually intermittent and short term in nature

Troubleshooting Connectivity – Example III

Cisco.com



- **Symptom: AS1 is trying to loadshare between its upstreams, but has trouble getting traffic through the AS2 link**

Troubleshooting Connectivity – Example III

Cisco.com

- **Checklist:**

What does “trouble” mean?

- **Is outbound traffic loadsharing okay?**

Can usually fix this with selectively rejecting prefixes, and using local preference

Generally easy to fix, local problem, simple application of policy

- **Is inbound traffic loadsharing okay?**

Errummm, bigger problem if not

Need to do some troubleshooting if configuration with communities, AS-PATH prepends, MEDs and selective leaking of subprefixes don't seem to help

Troubleshooting Connectivity – Example III

Cisco.com

- **Checklist:**

AS1 announces, but does AS2 see it?

We are checking eBGP filters on R1 and R2. Remember that R2 access will require cooperation and assistance from your peer

Does AS2 see it over entire network?

We are checking iBGP across AS2's network. Quite often iBGP is misconfigured, lack of full mesh, problems with RRs, etc.

Troubleshooting Connectivity – Example III

Cisco.com

- **Checklist:**

Does AS2 send it to its upstream?

We are checking eBGP configuration on R2. There may be a configuration error with as-path filters, or prefix-lists, or communities such that only local prefixes get out

Does the Internet see all of AS2's originated prefixes?

We are checking eBGP configuration on other Internet routers. This means using looking glasses. And trying to find one as close to AS2 as possible.

Troubleshooting Connectivity – Example III

Cisco.com

- **Checklist:**

Repeat all of the above for AS3

- **Stopping here and resorting to a huge prepend towards AS3 won't solve the problem**
- **There are many common problems – listed on next slide**

And tools to help decipher the problem

Troubleshooting Connectivity – Example III

Cisco.com

- **No inbound traffic from AS2**

AS2 is not seeing AS1's prefix, or is blocking it in inbound filters

- **A trickle of inbound traffic**

Switch on NetFlow (if the router has it) and check the origin of the traffic

If it is just from AS2's network blocks, then is AS2 announcing the prefix to its upstreams?

If they claim they are, ask them to ask their upstream for a “show ip bgp” output – or use a Looking Glass to check

Troubleshooting Connectivity – Example III

Cisco.com

- **A light flow of traffic from AS2, but 50% less than from AS3**

Looking Glass comes to the rescue

LG will let you see what AS2, or AS2's upstreams are announcing

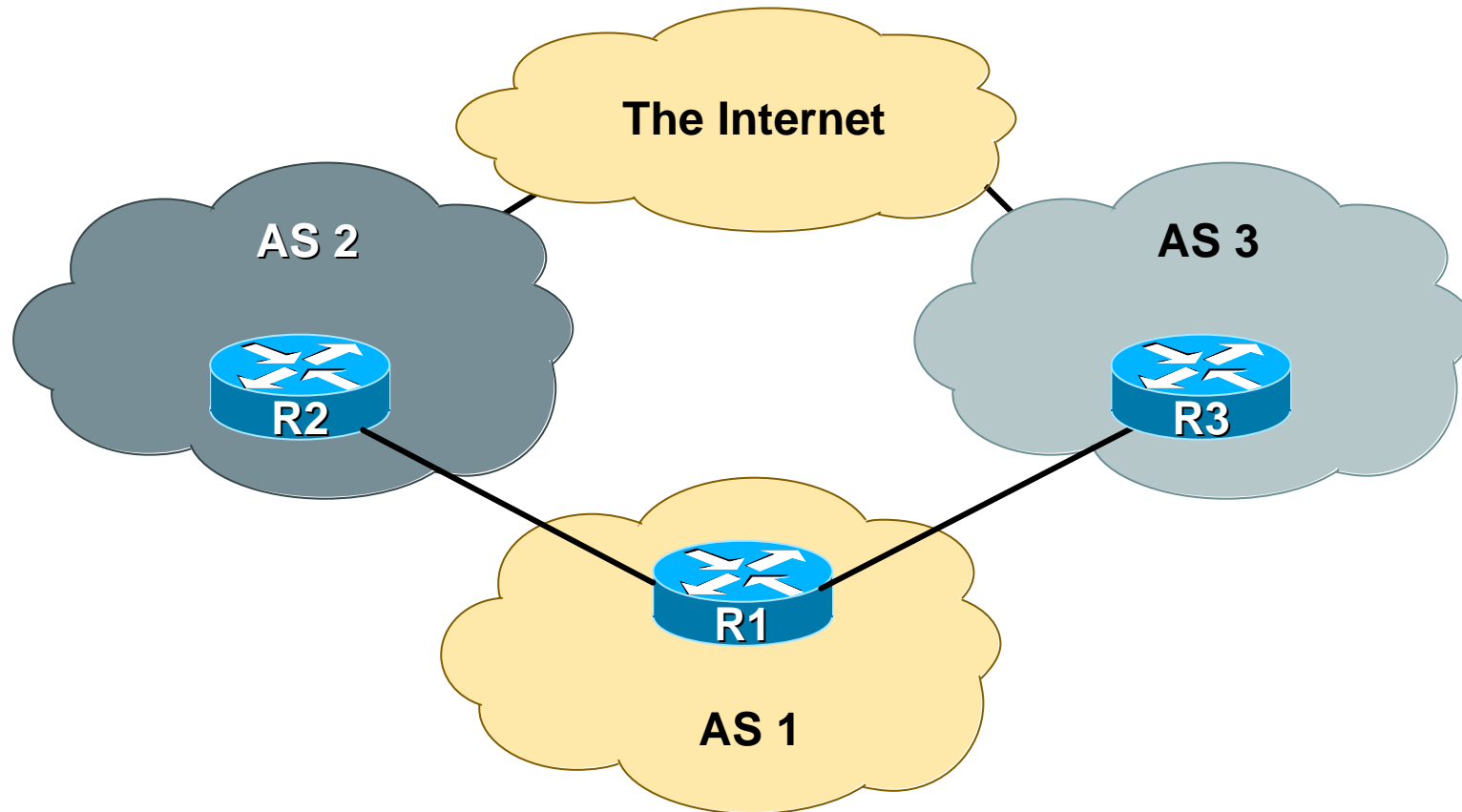
AS1 may choose this as primary path, but AS2 relationship with their upstream may decide otherwise

NetFlow comes to the rescue

Allows AS1 to see what the origins are, and with the LG, helps AS1 to find where the prefix filtering culprit might be

Troubleshooting Connectivity – Example IV

Cisco.com



- **Symptom: AS1 is loadsharing between its upstreams, but the traffic load swings randomly between AS2 and AS3**

Troubleshooting Connectivity – Example IV

Cisco.com

- **Checklist:**

Assume AS1 has done everything in this tutorial so far

All the configurations look fine, the Looking Glass outputs look fine, life is wonderful... Apart from those annoying traffic swings every hour or so

L2 problem? Route Flap Damping?

Since BGP is configured fine, and the net has been stable for so long, can only be an L2 problem, or Route Flap Damping side-effect

Troubleshooting Connectivity – Example IV

Cisco.com

- **L2 – upstream somewhere has poor connectivity between themselves and the rest of the Internet**

Only real solution is to impress upon upstream that this isn't good enough, and get them to fix it

Or change upstreams

Troubleshooting Connectivity – Example IV

Cisco.com

- **Route Flap Damping**

Many ISPs implement route flap damping

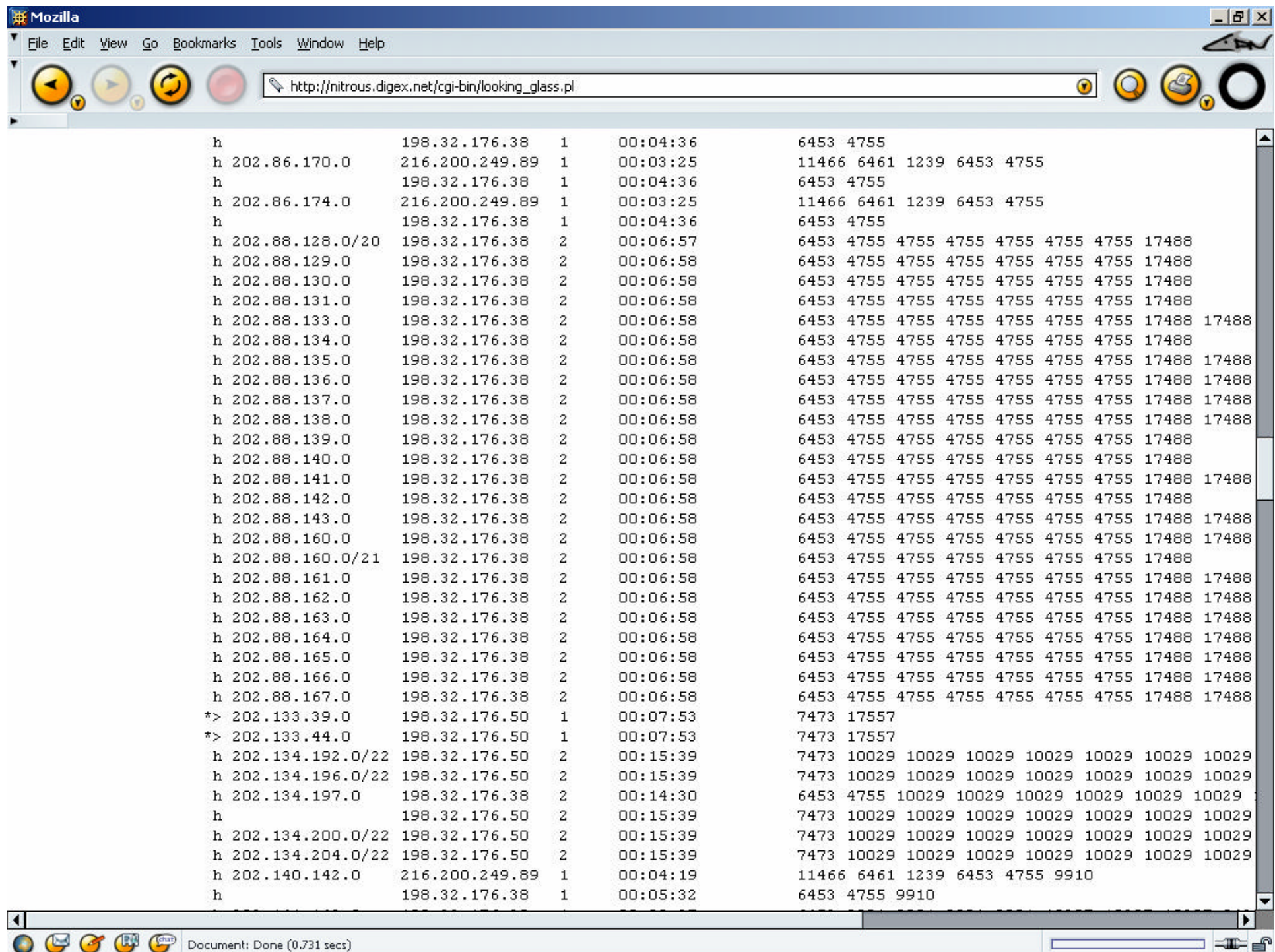
Many ISPs simply use the vendor defaults

Vendor defaults are generally far too severe

There is even now some real concern that the “more lenient” RIPE-229 values are too severe

www.cs.berkeley.edu/~zmao/Papers/sig02.pdf

- **Again Looking Glasses come to the operator's assistance**



Troubleshooting Connectivity – Example IV

Cisco.com

- **Most Looking Glasses allow the operators to check the flap or damped status of their announcements**

Many oscillating connectivity issues are usually caused by L2 problems

Route flap damping will cause connectivity to persist via alternative paths even though primary paths have been restored

Quite often, the exponential back off of the flap damping timer will give rise to bizarre routing

Common symptom is that bizarre routing will often clear away by itself

Troubleshooting Summary

Cisco.com

- **Most troubleshooting is about:**
- **Experience**
Recognising the common problems
- **Not panicking**
- **Logical approach**
Check configuration first
Check locally first before blaming the peer
Troubleshoot layer 1, then layer 2, then layer 3, etc

Troubleshooting Summary

Cisco.com

- **Most troubleshooting is about:**
- **Using the available tools**

The debugging tools on the router hardware

Internet Looking Glasses

Colleagues and their knowledge

Public mailing lists where appropriate

Agenda

Cisco.com

- **Peer Establishment**
- **Missing Routes**
- **Internet Reachability Problems**

Troubleshooting BGP

The End! 😊