



BGP Attributes and Policy Control

ISP/IXP Workshops

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

1

Agenda

- BGP Attributes
- BGP Path Selection
- Applying Policy

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

2



BGP Attributes

The “tools” available for the job

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

3

What Is an Attribute?



- Describes the characteristics of prefix
- Transitive or non-transitive
- Some are mandatory

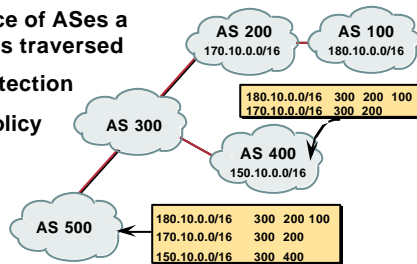
Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

4

AS-Path

- Sequence of ASes a route has traversed
- Loop detection
- Apply policy

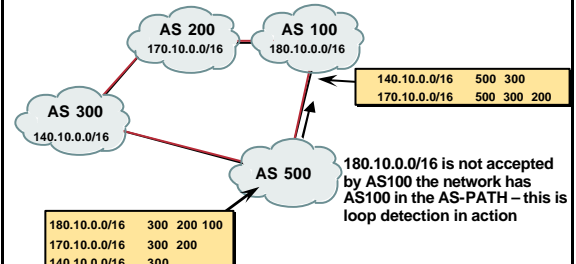


Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

5

AS-Path loop detection

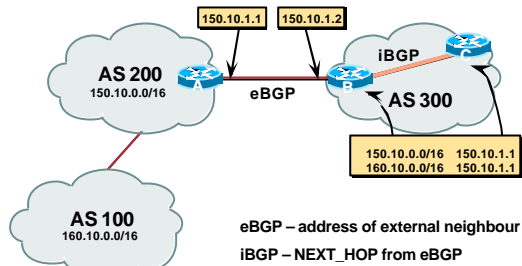


Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

6

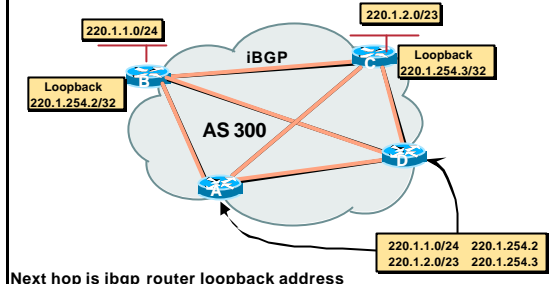
Next Hop



Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

7

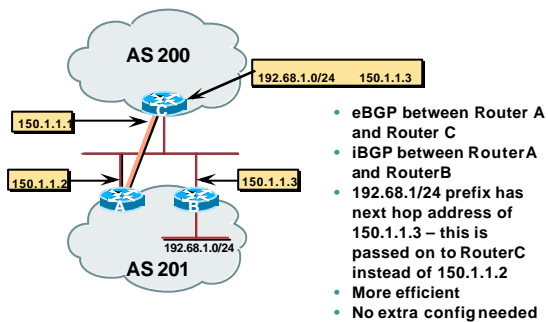
iBGP Next Hop



Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

8

Third Party Next Hop



Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

9

Next Hop (summary)

- IGP should carry route to next hops
- Recursive route look-up
- Unlinks BGP from actual physical topology
- Allows IGP to make intelligent forwarding decision

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

10

Origin

- Conveys the origin of the prefix
- “Historical” attribute
- Influences best path selection
- Three values: IGP, EGP, incomplete
 - IGP – generated by BGP network statement
 - EGP – generated by EGP
 - incomplete – redistributed from another routing protocol

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

11

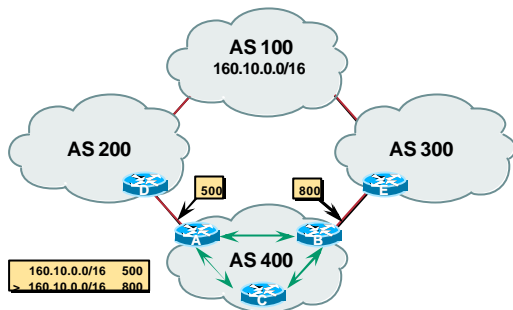
Aggregator

- Useful for debugging purposes
- Conveys the IP address of the router/BGP speaker generating the aggregate route
- Does not influence path selection

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

12

Local Preference



Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

13

Local Preference

- Local to an AS – non-transitive
local preference set to 100 when heard from neighbouring AS
- Used to influence BGP path selection
determines best path for **outbound** traffic
- Path with highest local preference wins

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

14

Local Preference

- Configuration of Router B:

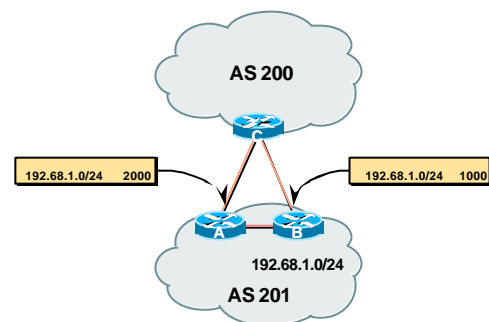
```
router bgp 400
 neighbor 220.5.1.1 remote-as 300
 neighbor 220.5.1.1 route-map local-pref in
!
route-map local-pref permit 10
 match ip address prefix-list MATCH
 set local-preference 800
!
ip prefix-list MATCH permit 160.10.0.0/16
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

15

Multi-Exit Discriminator (MED)



Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

16

Multi-Exit Discriminator

- Inter-AS – non-transitive
metric attribute not announced to next AS
- Used to convey the relative preference of entry points
determines best path for **inbound** traffic
- Comparable if paths are from same AS
- IGP metric can be conveyed as MED
set metric-type internal in route-map

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

17

MED & IGP Metric

- **set metric-type internal**
enable BGP to advertise a MED which corresponds to the IGP metric values
changes are monitored (and re-advertised if needed) every 600s
bgp dynamic-med-interval <secs>

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

18

Multi-Exit Discriminator

- Configuration of Router B:

```
router bgp 400
neighbor 220.5.1.1 remote-as 200
neighbor 220.5.1.1 route-map set-med out
!
route-map set-med permit 10
match ip address prefix-list MATCH
set metric 1000
!
ip prefix-list MATCH permit 192.68.1.0/24
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

19

Weight

- Not really an attribute – local to router
- Highest weight wins
- Applied to all routes from a neighbour


```
neighbor 220.5.7.1 weight 100
```
- Weight assigned to routes based on filter

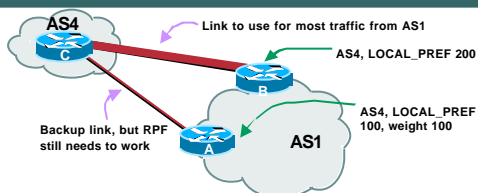

```
neighbor 220.5.7.3 filter-list 3 weight 50
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

20

Weight – Used to help Deploy RPF



- Best path to AS4 from AS1 is always via B due to local-pref
- But packets arriving at A from AS4 over the direct C to A link will pass the RPF check as that path has a priority due to the weight being set
- If weight was not set, best path back to AS4 would be via B, and the RPF check would fail

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

21

Community

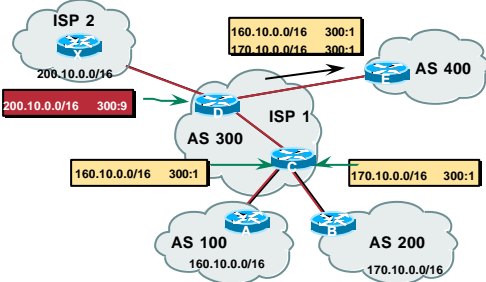
- Communities described in RFC1997
- 32 bit integer
 - Commonly represented as two 16bit integers (RFC1998)
- Used to group destinations
 - Each destination could be member of multiple communities
- Community attribute carried across AS's
- Very useful in applying policies

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

22

Community



Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

23

Well-Known Communities

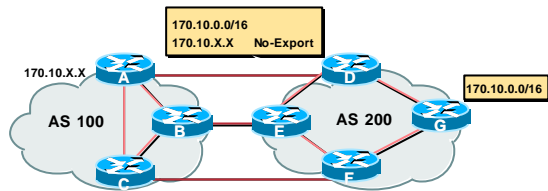
- no-export
 - do not advertise to eBGP peers
- no-advertise
 - do not advertise to any peer
- local-AS
 - do not advertise outside local AS (only used with confederations)

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

24

No-Export Community



- AS100 announces aggregate and subprefixes
aim is to improve loadsharing by leaking subprefixes
- Subprefixes marked with **no-export** community
- Router G in AS200 does not announce prefixes with **no-export** community set

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

25



BGP Path Selection Algorithm

Why is this the best path?

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

26

BGP Path Selection Algorithm

- Do not consider path if no route to next hop
- Do not consider iBGP path if not synchronised
- Highest weight (local to router)
- Highest local preference (global within AS)
- Prefer locally originated route
- Shortest AS path

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

27

BGP Path Selection Algorithm (continued)

- Lowest origin code
IGP < EGP < incomplete
- Lowest Multi-Exit Discriminator (MED)
If **bgp deterministic-med**, order the paths before comparing
If **bgp always-compare-med**, then compare for all paths
otherwise MED only considered if paths are from the same AS (default)

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

28

BGP Path Selection Algorithm (continued)

- Prefer eBGP path over iBGP path
- Path with lowest IGP metric to next-hop
- For eBGP paths:
If multipath is enabled, install N parallel paths in forwarding table
If router-id is the same, go to next step
If router-id is not the same, select the oldest path

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

29

BGP Path Selection Algorithm (continued)

- Lowest router-id (originator-id for reflected routes)
- Shortest cluster-list
Client **must** be aware of Route Reflector attributes!
- Lowest neighbour address

Cisco/ISPWorkshops © 2004, Cisco Systems, Inc. All rights reserved.

30



Applying Policy with BGP

How to use the "tools"

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

31

Applying Policy with BGP

- Policy-based on AS path, community or the prefix
- Rejecting/accepting selected routes
- Set attributes to influence path selection
- Tools:
 - Prefix-list (filters prefixes)
 - Filter-list (filters ASes)
 - Route-maps and communities

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

32

Policy Control – Prefix List

- Per neighbour prefix filter incremental configuration
- High performance access-list
- Inbound or Outbound
- Based upon network numbers (using familiar IPv4 address/mask format)

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

33

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>

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

34

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
```
- Deny the prefix 172.16.0.0/12

```
ip prefix-list EG deny 172.16.0.0/12
```
- In 192/8 allow up to /24

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

This allows all prefix sizes in the 192.0.0.0/8 address block, apart from /25, /26, /27, /28, /29, /30, /31 and /32.

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

35

Prefix Lists – Examples

- In 192/8 deny /25 and above

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

This denies all prefix sizes /25, /26, /27, /28, /29, /30, /31 and /32 in the address block 192.0.0.0/8.

It has the same effect as the previous example
- In 193/8 permit prefixes between /12 and /20

```
ip prefix-list EG permit 193.0.0.0/8 ge 12 le 20
```

This denies all prefix sizes /8, /9, /10, /11, /21, /22, ... and higher in the address block 193.0.0.0/8.
- Permit all prefixes

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

0.0.0.0 matches all possible addresses, "0 le 32" matches all possible prefix lengths

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

36

Policy Control – Prefix List

- Example Configuration

```
router bgp 200
 network 215.7.0.0
 neighbor 220.200.1.1 remote-as 210
 neighbor 220.200.1.1 prefix-list PEER-IN in
 neighbor 220.200.1.1 prefix-list PEER-OUT out
!
ip prefix-list PEER-IN deny 218.10.0.0/16
ip prefix-list PEER-IN permit 0.0.0.0/0 le 32
ip prefix-list PEER-OUT permit 215.7.0.0/16
ip prefix-list PEER-OUT deny 0.0.0.0/0 le 32
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

37

Policy Control – Filter List

- Filter routes based on AS path
- Inbound or Outbound
- Example Configuration:

```
router bgp 100
 network 215.7.0.0
 neighbor 220.200.1.1 filter-list 5 out
 neighbor 220.200.1.1 filter-list 6 in
!
ip as-path access-list 5 permit ^200$
ip as-path access-list 6 permit ^150$
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

38

Policy Control – Regular Expressions

- Like Unix regular expressions

- Match one character
- Match any number of preceding expression
- Match at least one of preceding expression
- Beginning of line
- End of line
- Beginning, end, white-space, brace
- Or
- brackets to contain expression

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

39

Policy Control – Regular Expressions

- Simple Examples

- match anything
- match at least one character
- match routes local to this AS
- originated by AS1800
- received from AS1800
- via AS1800
- via AS1800 and AS790
- multiple AS1800 in sequence (used to match AS-PATH prepends)
- via AS65530 (confederations)

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

40

Policy Control – Regular Expressions

- Not so simple Examples

- Match AS_PATH length of one
- Match AS_PATH length of two
- Match AS_PATH length of one or two
- Match AS_PATH length of one or two (will also match zero)
- Match AS_PATH length of three
- Match anything which has gone through AS701 or AS1800
- Match anything of origin AS12163 and passed through AS1849

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

41

Policy Control – Route Maps

- A route-map is like a “programme” for IOS
- Has “line” numbers, like programmes
- Each line is a separate condition/action
- Concept is basically:
 - if *match* then do *expression* and *exit*
 - else
 - if *match* then do *expression* and *exit*
 - else *etc*

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

42

Route Maps – Caveats

- Lines can have multiple set statements but only one match statement
- Line with only a set statement
all prefixes are matched and set
any following lines are ignored
- Line with a match/set statement and no following lines
only prefixes matching go through
the rest are dropped

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

43

Route Maps – Caveats

• Example

omitting the third line below means that prefixes not matching **list-one** or **list-two** are dropped

```
route-map sample permit 10
match ip address prefix-list list-one
set local-preference 120
!
route-map sample permit 20
match ip address prefix-list list-two
set local-preference 80
!
route-map sample permit 30 ! Don't forget this
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

44

Policy Control – Route Maps

• Example Configuration – route map and prefix-lists

```
router bgp 100
neighbor 1.1.1.1 route-map infiltrer in
!
route-map infiltrer permit 10
match ip address prefix-list HIGH-PREF
set local-preference 120
!
route-map infiltrer permit 20
match ip address prefix-list LOW-PREF
set local-preference 80
!
ip prefix-list HIGH-PREF permit 10.0.0.0/8
ip prefix-list LOW-PREF permit 20.0.0.0/8
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

45

Policy Control – Route Maps

• Example Configuration – route map and filter lists

```
router bgp 100
neighbor 220.200.1.2 remote-as 200
neighbor 220.200.1.2 route-map filter-on-as-path in
!
route-map filter-on-as-path permit 10
match as-path 1
set local-preference 80
!
route-map filter-on-as-path permit 20
match as-path 2
set local-preference 200
!
ip as-path access-list 1 permit _150$
ip as-path access-list 2 permit _210_
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

46

Policy Control – Route Maps

• Example configuration of AS-PATH prepend

```
router bgp 300
network 215.7.0.0
neighbor 2.2.2.2 remote-as 100
neighbor 2.2.2.2 route-map SETPATH out
!
route-map SETPATH permit 10
set as-path prepend 300 300
```

- Use your **own** AS number when prepending
Otherwise BGP loop detection may cause disconnects

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

47

Policy Control – Route Maps

• Route Map MATCH Articles

as-path	ip next-hop
clns address	ip route-source
clns next-hop	length
clns route-source	metric
community	nlri
interface	route-type
ip address	tag

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

48

Policy Control – Route Maps

- Route map SET Articles

as-path	dampening
automatic-tag	default interface
clns	interface
comm-list	ip default next-hop
community	ip next-hop

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

49

Policy Control – Route Maps

- Route map SET Articles

ip precedence	next-hop
ip qos-group	nlri multicast
ip tos	nlri unicast
level	origin
local preference	tag
metric	traffic-index
metric-type	weight

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

50

Policy Control – Matching Communities

- Example Configuration

```
router bgp 100
  neighbor 220.200.1.2 remote-as 200
  neighbor 220.200.1.2 route-map filter-on-community in
  !
  route-map filter-on-community permit 10
    match community 1
    set local-preference 50
  !
  route-map filter-on-community permit 20
    match community 2 exact-match
    set local-preference 200
  !
  ip community-list 1 permit 150:3 200:5
  ip community-list 2 permit 88:6
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

51

Policy Control – Setting Communities

- Example Configuration

```
router bgp 100
  network 215.7.0.0
  neighbor 220.200.1.1 remote-as 200
  neighbor 220.200.1.1 send-community
  neighbor 220.200.1.1 route-map set-community out
  !
  route-map set-community permit 10
    match ip address prefix-list NO-ANNOUNCE
    set community no-export
  !
  route-map set-community permit 20
    match ip address prefix-list EVERYTHING
  !
  ip prefix-list NO-ANNOUNCE permit 172.168.0.0/16 ge 17
  ip prefix-list EVERYTHING permit 0.0.0.0/0 le 32
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

52

Aggregation Policies

- Suppress Map

Used to suppress selected more-specific prefixes (e.g. defined through a route-map) in the absence of the **summary-only** keyword.

- Unsuppress Map

Used to unsuppress selected more-specific prefixes per BGP peering when the **summary-only** keyword is in use.

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

53

Aggregation Policies – Suppress Map

- Example

```
router bgp 100
  network 220.10.10.0
  network 220.10.11.0
  network 220.10.12.0
  network 220.10.33.0
  network 220.10.34.0
  aggregate-address 220.10.0.0 255.255.0.0 suppress-map block-net
  neighbor 222.5.7.2 remote-as 200
  !
  route-map block-net permit 10
    match ip address prefix-list SUPPRESS
  !
  ip prefix-list SUPPRESS permit 220.10.8.0/21 le 32
  ip prefix-list SUPPRESS deny 0.0.0.0/0 le 32
  !
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

54

Aggregation Policies – Suppress Map

- show ip bgp on the local router

```
router1#sh ip bgp
BGP table version is 11, local router ID is 222.5.7.1
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
*> 220.10.0.0/16 0.0.0.0          0         32768 i
s> 220.10.10.0   0.0.0.0          0         32768 i
s> 220.10.11.0   0.0.0.0          0         32768 i
s> 220.10.12.0   0.0.0.0          0         32768 i
*> 220.10.33.0   0.0.0.0          0         32768 i
*> 220.10.34.0   0.0.0.0          0         32768 i
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

55

Aggregation Policies – Suppress Map

- show ip bgp on the remote router

```
router2#sh ip bgp
BGP table version is 90, local router ID is 222.5.7.2
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
*> 220.10.0.0/16 222.5.7.1        0         0 100 i
*> 220.10.33.0   222.5.7.1        0         0 100 i
*> 220.10.34.0   222.5.7.1        0         0 100 i
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

56

Aggregation Policies – Unsuppress Map

- Example

```
router bgp 100
 network 220.10.10.0
 network 220.10.11.0
 network 220.10.12.0
 network 220.10.33.0
 network 220.10.34.0
 aggregate-address 220.10.0.0 255.255.0.0 summary-only
 neighbor 222.5.7.2 remote-as 200
 neighbor 222.5.7.2 unsuppress-map leak-net
!
route-map leak-net permit 10
 match ip address prefix-list LEAK
!
ip prefix-list LEAK permit 220.10.8.0/21 le 32
ip prefix-list LEAK deny 0.0.0.0/0 le 32
!
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

57

Aggregation Policies – Unsuppress Map

- show ip bgp on the local router

```
router1#sh ip bgp
BGP table version is 11, local router ID is 222.5.7.1
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
*> 220.10.0.0/16 0.0.0.0          0         32768 i
s> 220.10.10.0   0.0.0.0          0         32768 i
s> 220.10.11.0   0.0.0.0          0         32768 i
s> 220.10.12.0   0.0.0.0          0         32768 i
s> 220.10.33.0   0.0.0.0          0         32768 i
s> 220.10.34.0   0.0.0.0          0         32768 i
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

58

Aggregation Policies – Unsuppress Map

- show ip bgp on the remote router

```
router2#sh ip bgp
BGP table version is 90, local router ID is 222.5.7.2
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
*> 220.10.0.0/16 222.5.7.1        0         0 100 i
*> 220.10.10.0   222.5.7.1        0         0 100 i
*> 220.10.11.0   222.5.7.1        0         0 100 i
*> 220.10.12.0   222.5.7.1        0         0 100 i
```

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

59

Aggregation Policies – Aggregate Address

- Summary-only used
- Absence of summary-only

all subprefixes
suppressed

unsuppress-map to
selectively leak
subprefixes

bgp per neighbour
configuration

no subprefixes
suppressed

suppress-map to
selectively suppress
subprefixes

bgp global configuration

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

60



BGP Attributes and Policy Control

ISP/IXP Workshops

Cisco/ISPWorkshops

© 2004, Cisco Systems, Inc. All rights reserved.

61