BGP and the Internet

Advanced Community Usage

RFC1998

- Informational RFC
- Describes how to implement loadsharing and backup on multiple inter-AS links
  - BGP communities used to determine local preference in upstream's network
- Gives control to the customer
- Simplifies upstream's configuration
  - simplifies network operation!

RFC1998

- Community values defined to have particular meanings:

ASx:100	set local pref 100	preferred route
ASx:90	set local pref 90	backup route if dualhomed on ASx
ASx:80	set local pref 80	main link is to another ISP with same AS path length
ASx:70	set local pref 70	main link is to another ISP

RFC1998

- Sample Customer Router Configuration

```
router bgp 107
 neighbor x.x.x.x remote-as 109
 neighbor x.x.x.x description Backup ISP
 neighbor x.x.x.x route-map config-community out
 neighbor x.x.x.x send-community
!
ip as-path access-list 20 permit ^$
ip as-path access-list 20 deny .*
!
route-map config-community permit 10
 match as-path 20
 set community 109:90
```


RFC1998

- Sample ISP Router Configuration

```
! Homed to another ISP
ip community-list 70 permit 109:70
! Homed to another ISP with equal ASPATH length
ip community-list 80 permit 109:80
! Customer backup routes
ip community-list 90 permit 109:90
!
route-map set-customer-local-pref permit 10
 match community 70
 set local-preference 70
```


RFC1998

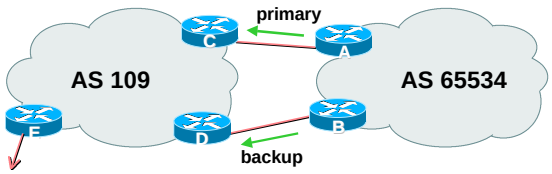
- Sample ISP Router Configuration

```
route-map set-customer-local-pref permit 20
 match community 80
 set local-preference 80
!
route-map set-customer-local-pref permit 30
 match community 90
 set local-preference 90
!
route-map set-customer-local-pref permit 40
 set local-preference 100
```

RFC1998

- Supporting RFC1998
  - many ISPs do, more should
  - check AS object in the Internet Routing Registry
  - if you do, insert comment in AS object in the IRR

Two links to the same ISP



- AS109 proxy aggregates for AS 65534

Two links to the same ISP (one as backup only)

- Announce /19 aggregate on each link
  - primary link makes standard announcement
  - backup link sends community
- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

Two links to the same ISP (one as backup only)

- Router A Configuration

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.2 remote-as 109
 neighbor 222.222.10.2 description RouterC
 neighbor 222.222.10.2 prefix-list aggregate out
 neighbor 222.222.10.2 prefix-list default in
!
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
```


Two links to the same ISP (one as backup only)

- Router B Configuration

```
router bgp 65534
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.6 remote-as 109
 neighbor 222.222.10.6 description RouterD
 neighbor 222.222.10.6 send-community
 neighbor 222.222.10.6 prefix-list aggregate out
 neighbor 222.222.10.6 route-map routerD-out out
 neighbor 222.222.10.6 prefix-list default in
 neighbor 222.222.10.6 route-map routerD-in in
!
..next slide
```

Two links to the same ISP (one as backup only)

```
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
 match ip address prefix-list aggregate
 set community 109:90
route-map routerD-out permit 20
!
route-map routerD-in permit 10
 set local-preference 90
!
```


Two links to the same ISP (one as backup only)

- **Router C Configuration (main link)**

```
router bgp 109
 neighbor 222.222.10.1 remote-as 65534
 neighbor 222.222.10.1 default-originate
 neighbor 222.222.10.1 prefix-list Customer in
 neighbor 222.222.10.1 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```


Two links to the same ISP (one as backup only)

- **Router D Configuration (backup link)**

```
router bgp 109
 neighbor 222.222.10.5 remote-as 65534
 neighbor 222.222.10.5 default-originate
 neighbor 222.222.10.5 prefix-list Customer in
 neighbor 222.222.10.5 route-map bgp-cust-in in
 neighbor 222.222.10.5 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
..next slide
```


Two links to the same ISP (one as backup only)

```
 ip prefix-list Customer permit 221.10.0.0/19
 ip prefix-list default permit 0.0.0.0/0
 !
 ip community-list 90 permit 109:90
 !
<snip>
 route-map bgp-cust-in permit 30
  match community 90
  set local-preference 90
 route-map bgp-cust-in permit 40
  set local-preference 100
```

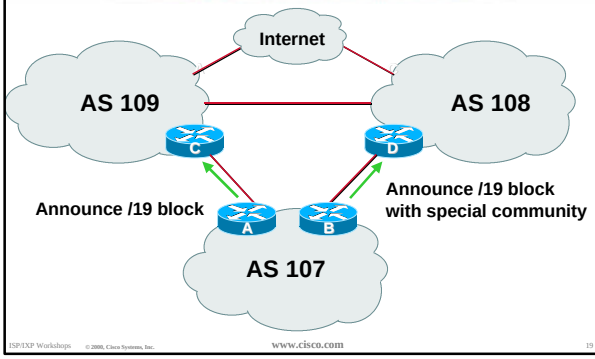

Two links to the same ISP (one as backup only)

- **Router E Configuration**

```
router bgp 109
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.17 remote-as 110
 neighbor 222.222.10.17 filter-list 1 out
!
ip as-path access-list 1 deny ^(65534_)+$
ip as-path access-list 1 permit ^$
ip route 221.10.0.0 255.255.224.0 null0
```

- **Router E removes prefixes in the private AS from external announcements**
- **Private AS still visible inside AS109**

Two links to different ISPs (one as backup only)



Two links to different ISPs (one as backup only)

- Announce /19 aggregate on each link
  - main link makes sends community 109:100 - this sets local pref in AS109 to 100
  - backup link sends community 108:80 - this sets local pref in AS108 to 80
- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

Two links to different ISPs (one as backup only)

- Note that this assumes that AS109 and AS108 are interconnected
- If they are not, AS path length "stuffing" has to be used too
  - but that can be done on a per community basis also
  - RFC1998++

Two links to different ISPs (one as backup only)

- Router A Configuration

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.1 remote-as 109
 neighbor 222.222.10.1 prefix-list aggregate out
 neighbor 222.222.10.1 route-map routerC-out out
 neighbor 222.222.10.1 prefix-list default in
!
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerC-out permit 10
 set community 109:100
```


Two links to different ISPs (one as backup only)

- Router B Configuration

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 220.1.5.1 remote-as 108
 neighbor 220.1.5.1 prefix-list aggregate out
 neighbor 220.1.5.1 route-map routerD-out out
 neighbor 220.1.5.1 prefix-list default in
 neighbor 220.1.5.1 route-map routerD-in in
..next slide
```


Two links to different ISPs (one as backup only)

```
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
 set community 108:80
!
route-map routerD-in permit 10
 set local-preference 80
```

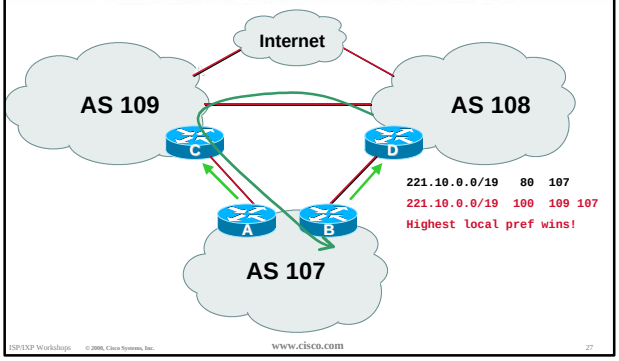
Two links to different ISPs (one as backup only)

- Router D
  - sees path from router B with community 108:80 set - sets local preference to 80
  - sees path from peering with AS109 - default local preference is 100
  - local-pref comes before AS Path length
  - highest local-pref wins
  - traffic for AS107 is sent to AS109

Two links to different ISPs (one as backup only)

- Router D
  - Only requires RFC1998 configuration
  - no per customer configuration
  - scalability!

Two links to different ISPs (one as backup only)



Two links to different ISPs (one as backup only)

- If AS107 wants to make the link to AS108 the main link
  - sends community 108:100 to router C
  - sends community 109:80 to router B
- AS108 and AS109 NOC intervention not required

RFC1998++

- RFC1998 is okay for “simple” multihomed customers
  - assumes that upstreams are interconnected
- RFC1998++
  - unofficial but often used additions by many ISPs
  - assumes nothing!

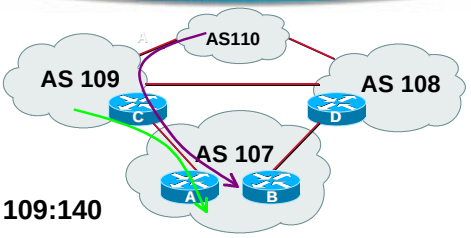
RFC1998++

- More community definitions:

ASx:140	set local pref 140	set local pref high on upstreams
ASx:130	set local pref 130	set local pref low on upstreams
ASx:120	set local pref 120	more preferred (opposite to ASx:80)
<RFC1998 definitions>		
ASx:60	set local pref 60	ASx:90 but add 2 times AS PATH
ASx:50	set local pref 50	don't announce to any peer
ASx:40	set local pref 40	and set local pref high on upstreams
ASx:30	set local pref 30	and set local pref low on upstreams

(and variations on this theme depending on local conditions, e.g. IXPs, domestic vs. international transit, etc.)

RFC1998++ example

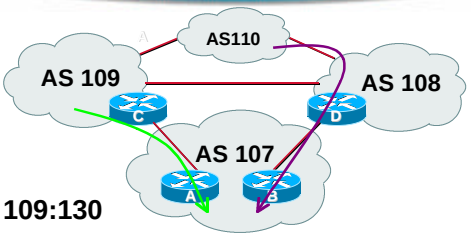


- 109:140

  traffic in AS109 comes directly to you

  traffic in AS110 sent to AS109 rather than best path

RFC1998++ example



- 109:130

  traffic in AS109 comes directly to you

  traffic in AS110 sent to AS108 rather than best path

RFC1998++ example

- 109:60

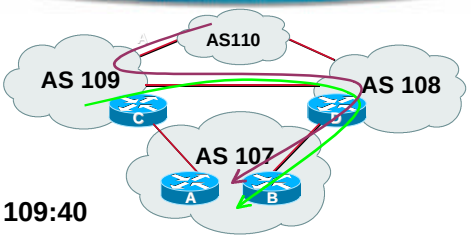
  set local pref low in AS109

  prepend any announcements to peers of AS109 with 109_109 - AS109 is my backup transit AS

- 109:50

  don't announce to any peer - used when you have good local connections to AS109 and better long distance via AS108

RFC1998++ example

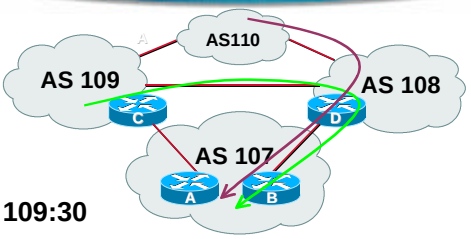


- 109:40

  traffic in AS109 sent to AS108

  traffic in AS110 sent to AS109 rather than best path

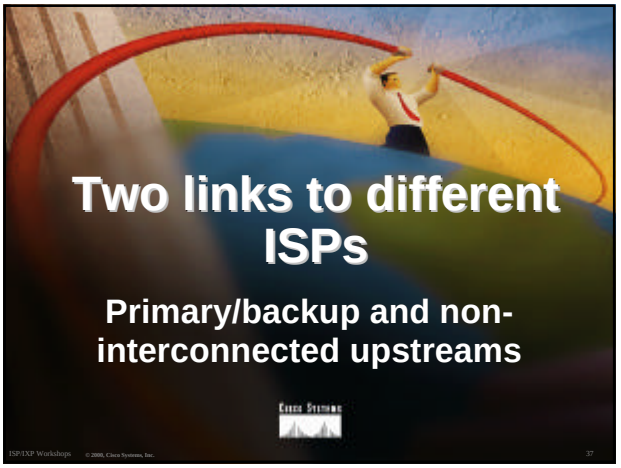
RFC1998++ example



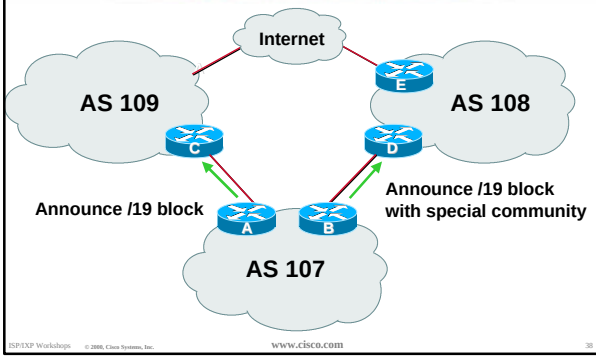
- 109:30

  traffic in AS109 sent to AS108

  traffic in AS110 sent to AS108 rather than best path



Two links to different ISPs (RFC1998++)



Two links to different ISPs (RFC1998++)

- Announce /19 aggregate on each link
  - main link makes sends community 109:100 - this sets local pref in AS109 to 100
  - backup link sends community 108:60 - this sets local pref in AS108 to 60
- When one link fails, the announcement of the /19 aggregate via the other link ensures continued connectivity

Two links to different ISPs (RFC1998++)

- Router A Configuration

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.1 remote-as 109
 neighbor 222.222.10.1 send-community
 neighbor 222.222.10.1 prefix-list aggregate out
 neighbor 222.222.10.1 route-map routerB-out out
 neighbor 222.222.10.1 prefix-list default in
!
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerB-out permit 10
 set community 109:100
```

Two links to different ISPs (RFC1998++)

- Router B Configuration

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 220.1.5.1 remote-as 108
 neighbor 220.1.5.1 prefix-list aggregate out
 neighbor 220.1.5.1 route-map routerD-out out
 neighbor 220.1.5.1 prefix-list default in
 neighbor 220.1.5.1 route-map routerD-in in
..next slide
```

Two links to different ISPs (RFC1998++)

```
ip prefix-list aggregate permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
 set community 108:60
!
route-map routerD-in permit 10
 set local-preference 80
```

Two links to different ISPs (RFC1998++)

- Router D Configuration

```
router bgp 108
 neighbor 220.1.5.2 remote-as 109
 neighbor 220.1.5.2 default-originate
 neighbor 220.1.5.2 prefix-list Customer in
 neighbor 220.1.5.2 route-map bgp-cust-in in
 neighbor 220.1.5.2 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
..next slide
```


Two links to different ISPs (RFC1998++)

```
 ip prefix-list Customer permit 221.10.0.0/19
 ip prefix-list default permit 0.0.0.0/0
 !
 ip community-list 60 permit 108:60
 !
<snip>
 route-map bgp-cust-in permit 10
  match community 60
  set local-preference 60
  set community 108:4000
<snip>
```


Two links to different ISPs (RFC1998++)

- Router D
  - sees path from router B with community 108:60 set:
    - sets local pref to 60
    - changes community to AS108 community which prepends two times AS108

Two links to different ISPs (RFC1998++)

- Router D (contd)
  - sees path from AS109 via Internet - default local preference is 100
  - local-pref comes before AS Path length
  - highest local-pref wins
  - traffic for AS107 is sent via Internet

Two links to different ISPs (RFC1998++)

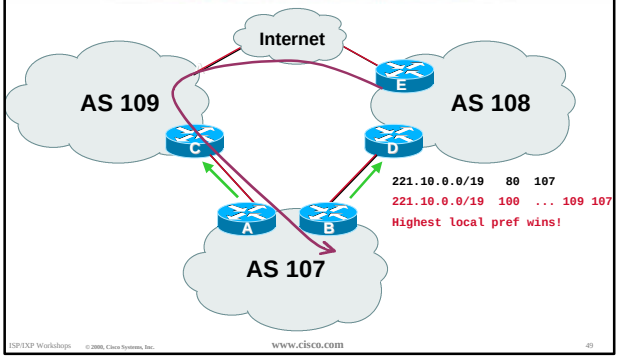
- Router E Configuration

```
router bgp 108
 neighbor x.x.x.x remote-as UP-ASN
 neighbor x.x.x.x prefix-list Upstream in
 neighbor x.x.x.x route-map upstream-in in
 neighbor x.x.x.x prefix-list AS108-list out
 neighbor x.x.x.x route-map upstream-out out
!
..next slide
```


Two links to different ISPs (RFC1998++)

```
! Customer peers who want AS-PATH prepend
ip community-list 1 permit 108:4000
! Customer peers who want control one set away from us
ip community-list 2 permit 108:4010
!
<snip>
  route-map upstream-out permit 10
   match community 1
   set as-path prepend 108 108
  route-map upstream-out permit 20
   match community 2
   set community UP-ASN:80
<snip>
```

Two links to different ISPs (RFC1998++)



Two links to different ISPs (RFC1998++)

- If AS107 wants to make the link to AS108 the main link
  - sends community 108:100 to router D
  - sends community 109:60 to router C
- AS108 and AS109 NOC intervention not required

Communities

- Communities are fun! ☺
- And they are extremely powerful tools
- Think about community policies, e.g. like RFC1998++
- Supporting extensive community usage makes customer configuration easy
- *Watch out for routing loops!*

Cisco Systems

Empowering the Internet Generation℠