



BGP and the Internet

Using Communities for Multihoming

Multihoming and Communities

- The BGP community attribute is a very powerful tool for assisting and scaling BGP Multihoming
- Most major ISPs make extensive use of BGP communities:
 - Internal policies
 - Inter-provider relationships (MED replacement)
 - Customer traffic engineering

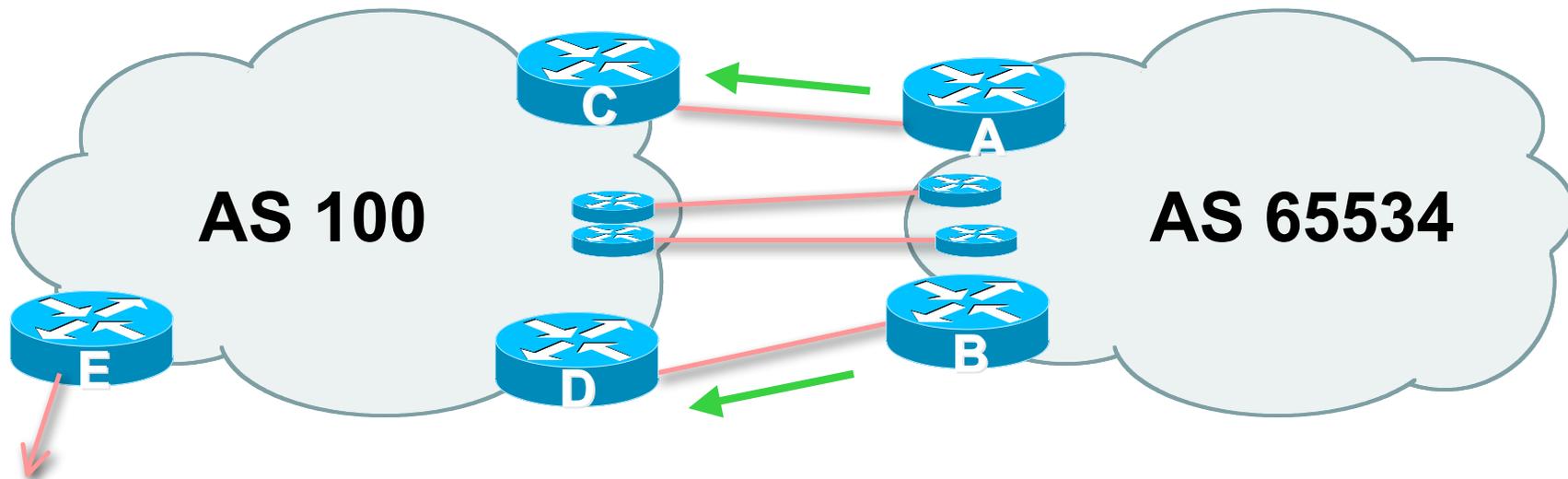


Loadsharing Using Communities

4 links – Private AS

Communities

Private AS



- AS100 removes private AS and any customer subprefixes from Internet announcement

Communities

Private AS

- Announce /19 aggregate on each link
- Split /19 and announce as four /21s, one on each link
 - basic inbound loadsharing
 - assumes equal circuit capacity and even spread of traffic across address block
- Vary the split until “perfect” loadsharing achieved
 - use the no-export community for subprefixes

Communities

Private AS

- Router A Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.0.0 mask 255.255.248.0
  neighbor 122.102.10.2 remote-as 100
  neighbor 122.102.10.2 send-community
  neighbor 122.102.10.2 prefix-list subblocks1 out
  neighbor 122.102.10.2 route-map routerC-out out
  neighbor 122.102.10.2 prefix-list default in
!
..next slide
```

Communities

Private AS

```
ip prefix-list subblocks1 permit 121.10.0.0/19
ip prefix-list subblocks1 permit 121.10.0.0/21
!
ip prefix-list firstblock permit 121.10.0.0/21
ip prefix-list default permit 0.0.0.0/0
!
route-map routerC-out permit 10
  match ip address prefix-list firstblock
  set community no-export
route-map routerC-out permit 20
```

Communities

Private AS

- Router B Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.24.0 mask 255.255.248.0
  neighbor 122.102.20.2 remote-as 100
  neighbor 122.102.20.2 send-community
  neighbor 122.102.20.2 prefix-list subblocks2 out
  neighbor 122.102.20.2 route-map routerD-out out
  neighbor 122.102.20.2 prefix-list default in
!
..next slide
```

Communities

Private AS

```
ip prefix-list subblocks2 permit 121.10.0.0/19
ip prefix-list subblocks2 permit 121.10.24.0/21
!
ip prefix-list secondblock permit 121.10.24.0/21
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
  match ip address prefix-list secondblock
  set community no-export
route-map routerD-out permit 20
```

Communities

Private AS

- Router E Configuration

```
router bgp 100
  neighbor 122.102.10.17 remote-as 110
  neighbor 122.102.10.17 remove-private-AS
!
```

- Router E removes the private AS from external announcements
- Router E automatically removes subprefixes with no-export community set
- Private AS still visible inside AS100

Communities

Private AS

- Router C and D configuration is as previously
- AS100 routers will not advertise prefixes marked with community no-export to other ASes
- AS100 routers still need to filter the private AS
- Only a single /19 prefix is announced to the Internet - no routing table bloat! :-)

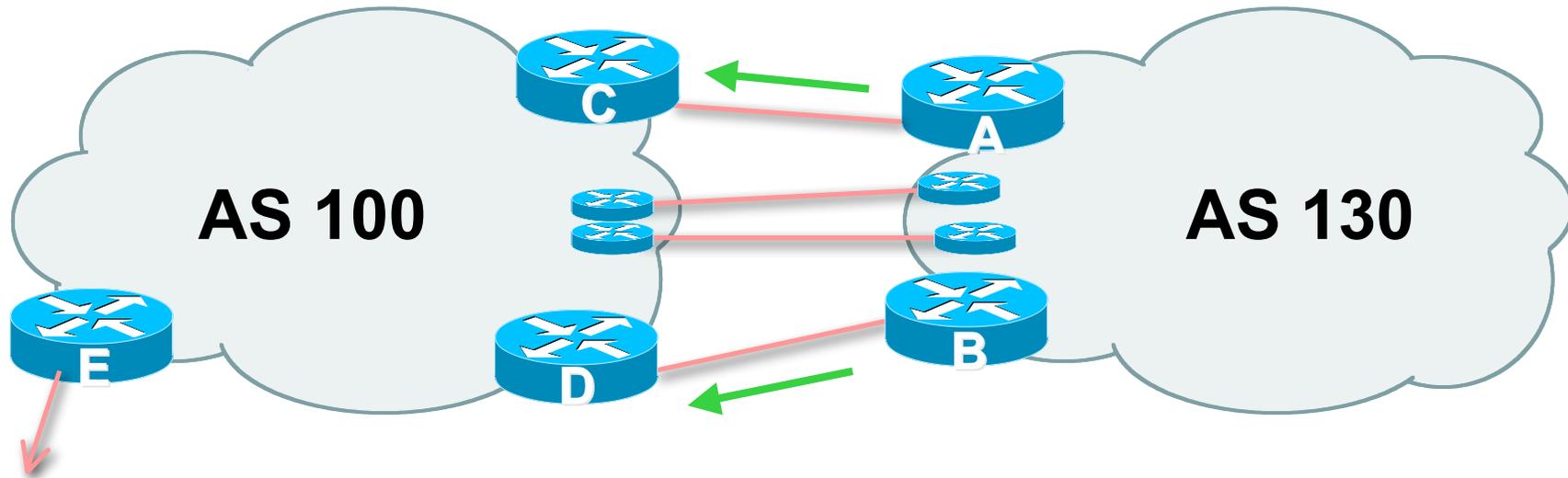


Loadsharing Using Communities

4 links – Public AS

Communities

Public AS



- 4 links between AS130 and AS100

Communities

Public AS

- Announce /19 aggregate on each link
- Split /19 and announce as four /21s, one on each link
 - basic inbound loadsharing
 - assumes equal circuit capacity and even spread of traffic across address block
- Vary the split until “perfect” loadsharing achieved
 - use the no-export community for subprefixes

Communities

Public AS

- Router A Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.0.0 mask 255.255.248.0
  neighbor 122.102.10.2 remote-as 100
  neighbor 122.102.10.2 send-community
  neighbor 122.102.10.2 prefix-list subblocks1 out
  neighbor 122.102.10.2 route-map routerC-out out
  neighbor 122.102.10.2 prefix-list default in
!
..next slide
```

Communities

Public AS

```
ip prefix-list subblocks1 permit 121.10.0.0/19
ip prefix-list subblocks1 permit 121.10.0.0/21
!
ip prefix-list firstblock permit 121.10.0.0/21
ip prefix-list default permit 0.0.0.0/0
!
route-map routerC-out permit 10
  match ip address prefix-list firstblock
  set community no-export
route-map routerC-out permit 20
```

Communities

Public AS

- Router B Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.24.0 mask 255.255.248.0
  neighbor 122.102.20.2 remote-as 100
  neighbor 122.102.20.2 send-community
  neighbor 122.102.20.2 prefix-list subblocks2 out
  neighbor 122.102.20.2 route-map routerD-out out
  neighbor 122.102.20.2 prefix-list default in
!
..next slide
```

Communities

Public AS

```
ip prefix-list subblocks2 permit 121.10.0.0/19
ip prefix-list subblocks2 permit 121.10.24.0/21
!
ip prefix-list secondblock permit 121.10.24.0/21
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
  match ip address prefix-list secondblock
  set community no-export
route-map routerD-out permit 20
```

Communities

Public AS

- Router C Configuration

```
router bgp 100
  neighbor 122.102.10.1 remote-as 130
  neighbor 122.102.10.1 default-originate
  neighbor 122.102.10.1 prefix-list Customer in
  neighbor 122.102.10.1 prefix-list default out
!
ip prefix-list Customer permit 121.10.0.0/19 le 21
ip prefix-list default permit 0.0.0.0/0
```

Communities

Public AS

- Router D Configuration

```
router bgp 100
  neighbor 122.102.10.5 remote-as 130
  neighbor 122.102.10.5 default-originate
  neighbor 122.102.10.5 prefix-list Customer in
  neighbor 122.102.10.5 prefix-list default out
!
ip prefix-list Customer permit 121.10.0.0/19 le 21
ip prefix-list default permit 0.0.0.0/0
```

Communities

Public AS

- Router E Configuration

```
router bgp 100
  neighbor 122.102.10.17 remote-as 110
  neighbor 122.102.10.17 filter-list 1 out
!
ip as-path access-list 1 permit ^130$
ip as-path access-list 1 permit ^$
```

- Router E only has to announce AS130 in the same way it announces other ASes

Communities

Public AS

- AS100 routers will not advertise prefixes marked with community no-export to other ASes
- AS100 ISP has no configuration work to do
AS130 ISP can control his own loadsharing
- Only a single /19 prefix is announced to the Internet -
no routing table bloat! :-)



RFC1998

An example of how ISPs use communities...

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 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 130
  neighbor x.x.x.x remote-as 100
  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 100:90
```

RFC1998

- Sample ISP Router Configuration

```
! Homed to another ISP
ip community-list 70 permit 100:70
! Homed to another ISP with equal ASPATH length
ip community-list 80 permit 100:80
! Customer backup routes
ip community-list 90 permit 100:90
!
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

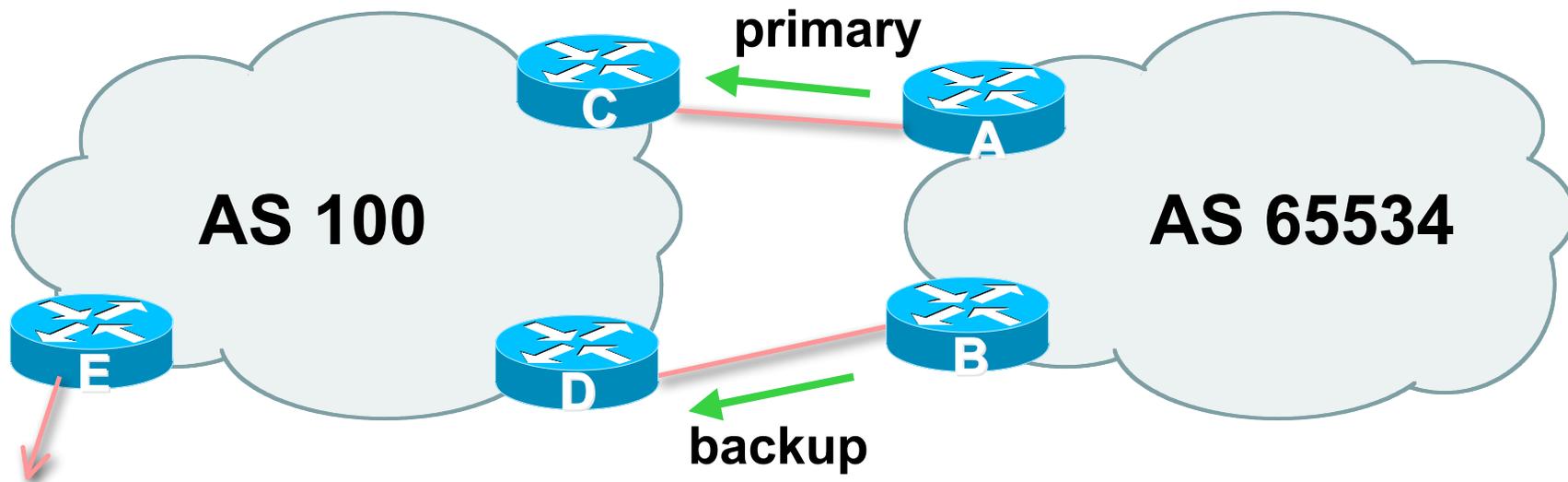
if you do, insert comment in AS object in the IRR



Two links to the same ISP

One link primary, the other link backup only

Two links to the same ISP



- AS100 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 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.2 remote-as 100
  neighbor 122.102.10.2 description RouterC
  neighbor 122.102.10.2 prefix-list aggregate out
  neighbor 122.102.10.2 prefix-list default in
!
ip prefix-list aggregate permit 121.10.0.0/19
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 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.6 remote-as 100
  neighbor 122.102.10.6 description RouterD
  neighbor 122.102.10.6 send-community
  neighbor 122.102.10.6 prefix-list aggregate out
  neighbor 122.102.10.6 route-map routerD-out out
  neighbor 122.102.10.6 prefix-list default in
  neighbor 122.102.10.6 route-map routerD-in in
!
```

..next slide

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

```
ip prefix-list aggregate permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
  match ip address prefix-list aggregate
  set community 100:90
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 100
  neighbor 122.102.10.1 remote-as 65534
  neighbor 122.102.10.1 default-originate
  neighbor 122.102.10.1 prefix-list Customer in
  neighbor 122.102.10.1 prefix-list default out
!
ip prefix-list Customer permit 121.10.0.0/19
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 100
  neighbor 122.102.10.5 remote-as 65534
  neighbor 122.102.10.5 default-originate
  neighbor 122.102.10.5 prefix-list Customer in
  neighbor 122.102.10.5 route-map bgp-cust-in in
  neighbor 122.102.10.5 prefix-list default out
!
ip prefix-list Customer permit 121.10.0.0/19
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 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip community-list 90 permit 100:90
!
<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 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.17 remote-as 110
  neighbor 122.102.10.17 filter-list 1 out
!
ip as-path access-list 1 deny ^(65534_)+$
ip as-path access-list 1 permit ^$
ip route 121.10.0.0 255.255.224.0 null0
```

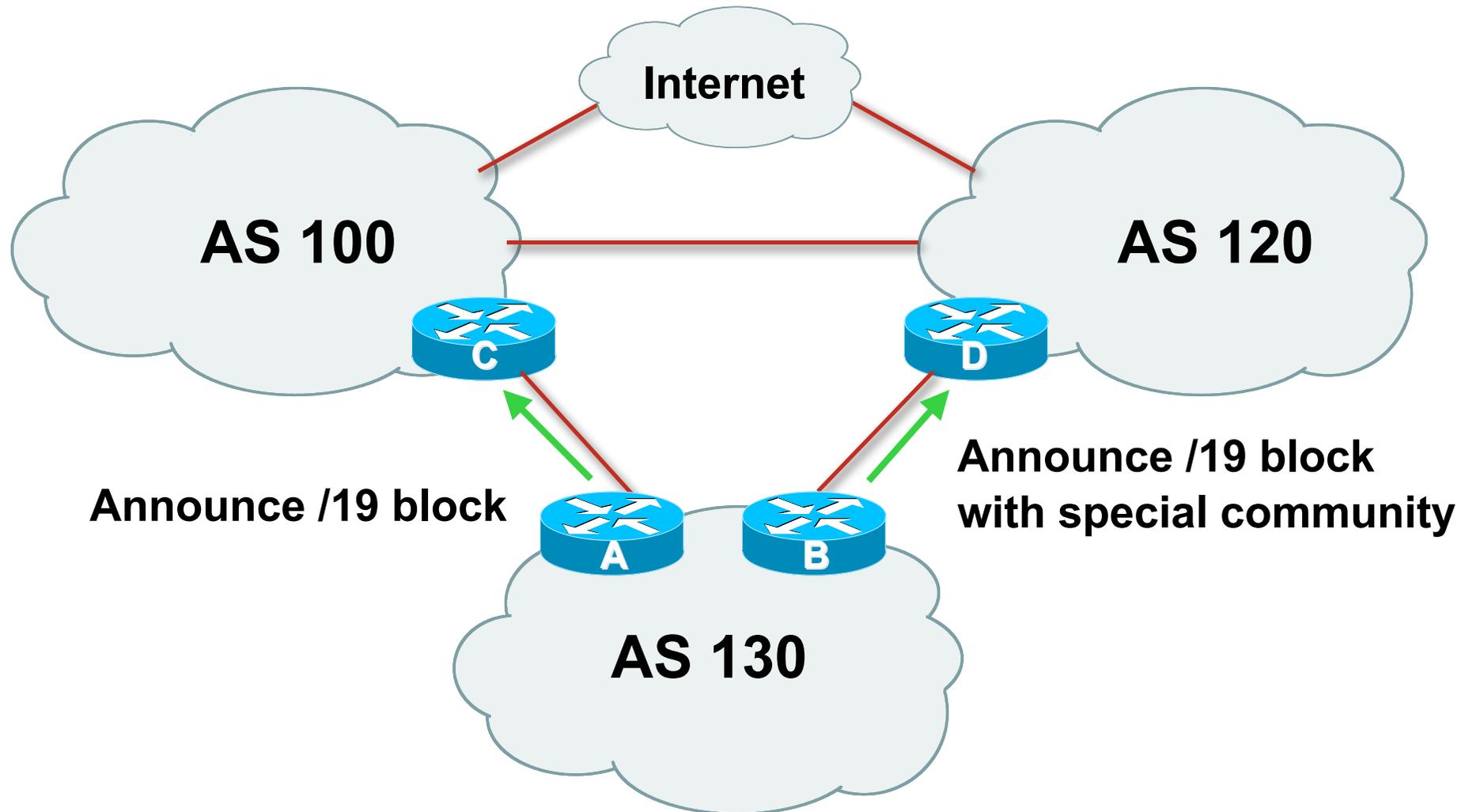
- Router E removes prefixes in the private AS from external announcements
- Private AS still visible inside AS100



Two links to different ISPs

One link primary, the other link backup only

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 sends community 100:100 – this sets local pref in AS100 to 100
 - backup link sends community 120:80 – this sets local pref in AS120 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 AS100 and AS120 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

Two links to different ISPs (one as backup only)

- Router A Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 100
  neighbor 122.102.10.1 prefix-list aggregate out
  neighbor 122.102.10.1 route-map routerC-out out
  neighbor 122.102.10.1 prefix-list default in
!
ip prefix-list aggregate permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerC-out permit 10
  set community 100:100
```

Two links to different ISPs (one as backup only)

- Router B Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  neighbor 220.1.5.1 remote-as 120
  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 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
  set community 120:80
!
route-map routerD-in permit 10
  set local-preference 80
```

Two links to different ISPs (one as backup only)

- Router D Configuration

 - Sees path from router B with community 120:80 set – sets local preference to 80

 - Sees path from peering with AS100 – default local preference is 100

 - Local-pref comes before AS Path length

 - Highest local-pref wins

 - Traffic for AS130 is sent to AS100

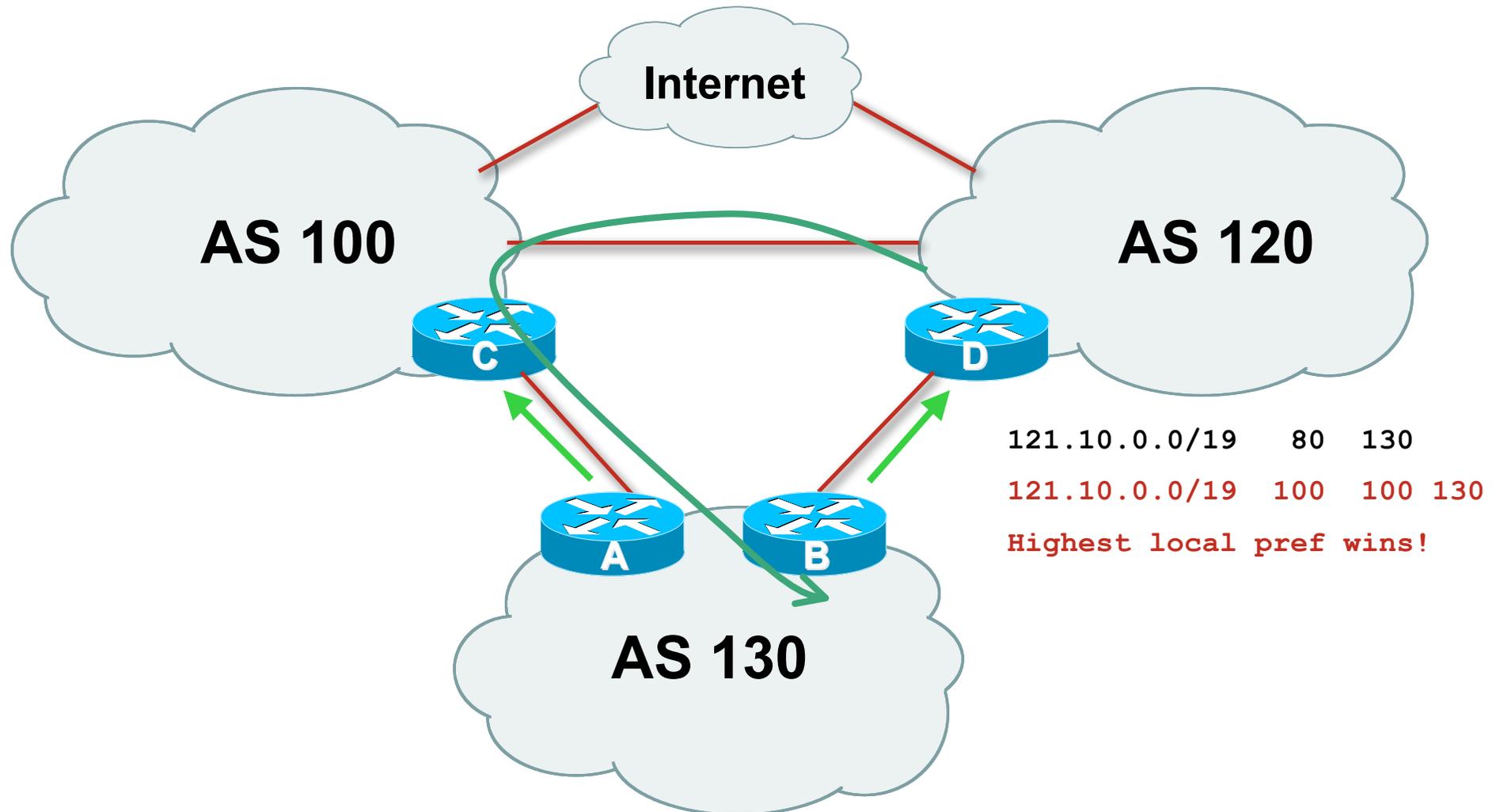
- Router D Summary

 - 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 AS130 wants to make the link to AS120 the main link
 - sends community 120:100 to router C
 - sends community 100:80 to router B
- AS120 and AS100 NOC intervention not required



Service Provider use of Communities

Some working examples

Background

- RFC1998 is okay for “simple” multihomed customers
 - assumes that upstreams are interconnected
- ISPs create many other communities to handle more complex situations
 - Simplify ISP BGP configuration
 - Give customer more policy control

ISP BGP Communities

- There are no recommended ISP BGP communities apart from RFC1998
The five standard communities
www.iana.org/assignments/bgp-well-known-communities
- Efforts have been made to document from time to time
totem.info.ucl.ac.be/publications/papers-elec-versions/draft-quoitin-bgp-comm-survey-00.pdf
But so far... nothing more... ☹️
Collection of ISP communities at www.onesc.net/communities
www.nanog.org/meetings/nanog40/presentations/BGPcommunities.pdf
- ISP policy is usually published
On the ISP's website
Referenced in the AS Object in the IRR

Some ISP Examples: Sprintlink

Welcome to Sprint.net

http://www.sprint.net/index.php?module=policies/bgp_policy

Apple .Mac Amazon eBay Yahoo! News (144) Apple (40)

WHAT YOU CAN CONTROL

AS-PATH PREPENDS

Sprint allows customers to use AS-path prepending to adjust route preference on the network. Such prepending will be received and passed on properly without notifying Sprint of your change in announcements.

Additionally, Sprint will prepend AS1239 to eBGP sessions with certain autonomous systems depending on a received community. Currently, the following ASes are supported: 1668, 209, 2914, 3300, 3356, 3549, 3561, 4635, 701, 7018, 702 and 8220.

String	Resulting AS Path to ASXXX
65000:XXX	Do not advertise to ASXXX
65001:XXX	1239 (default) ...
65002:XXX	1239 1239 ...
65003:XXX	1239 1239 1239 ...
65004:XXX	1239 1239 1239 1239 ...

String	Resulting AS Path to ASXXX in Asia
65070:XXX	Do not advertise to ASXXX
65071:XXX	1239 (default) ...
65072:XXX	1239 1239 ...
65073:XXX	1239 1239 1239 ...
65074:XXX	1239 1239 1239 1239 ...

String	Resulting AS Path to ASXXX in Europe
65050:XXX	Do not advertise to ASXXX
65051:XXX	1239 (default) ...
65052:XXX	1239 1239 ...
65053:XXX	1239 1239 1239 ...

More info at www.sprintlink.net/policy/bgp.html

Some ISP Examples

AAPT

```
aut-num:          AS2764
as-name:          ASN-CONNECT-NET
descr:           AAPT Limited
admin-c:          CNO2-AP
tech-c:           CNO2-AP
remarks:          Community support definitions
remarks:          Community Definition
remarks:          -----
remarks:          2764:2 Don't announce outside local POP
remarks:          2764:4 Lower local preference by 15
remarks:          2764:5 Lower local preference by 5
remarks:          2764:6 Announce to customers and all peers
                   (incl int'l peers), but not transit
remarks:          2764:7 Announce to customers only
remarks:          2764:14 Announce to AANX
notify:           routing@connect.com.au
mnt-by:           CONNECT-AU
changed:          nobody@connect.com.au 20050225
source:           CCAIR
```

More at <http://info.connect.com.au/docs/routing/general/multi-faq.shtml#q13>

Some ISP Examples

Verizon Business Europe

```
aut-num: AS702
descr: Verizon Business EMEA - Commercial IP service provider in Eur
remarks: VzBi uses the following communities with its customers:
 702:80      Set Local Pref 80 within AS702
 702:120    Set Local Pref 120 within AS702
 702:20     Announce only to VzBi AS'es and VzBi customers
 702:30     Keep within Europe, don't announce to other VzBi AS
 702:1      Prepend AS702 once at edges of VzBi to Peers
 702:2      Prepend AS702 twice at edges of VzBi to Peers
 702:3      Prepend AS702 thrice at edges of VzBi to Peers
Advanced communities for customers
 702:7020   Do not announce to AS702 peers with a scope of
            National but advertise to Global Peers, European
            Peers and VzBi customers.
 702:7001   Prepend AS702 once at edges of VzBi to AS702
            peers with a scope of National.
 702:7002   Prepend AS702 twice at edges of VzBi to AS702
            peers with a scope of National.
```

(more)

Some ISP Examples

VzBi Europe

(more)

```
702:7003 Prepend AS702 thrice at edges of VzBi to AS702
        peers with a scope of National.
702:8020 Do not announce to AS702 peers with a scope of
        European but advertise to Global Peers, National
        Peers and VzBi customers.
702:8001 Prepend AS702 once at edges of VzBi to AS702
        peers with a scope of European.
702:8002 Prepend AS702 twice at edges of VzBi to AS702
        peers with a scope of European.
702:8003 Prepend AS702 thrice at edges of VzBi to AS702
        peers with a scope of European.
```

Additional details of the VzBi communities are located at:
<http://www.verizonbusiness.com/uk/customer/bgp/>

```
mnt-by: WCOM-EMEA-RICE-MNT
source: RIPE
```

Some ISP Examples

BT Ignite

```
aut-num:      AS5400
descr:        BT Ignite European Backbone
remarks:
remarks:      Community to                               Community to
remarks:      Not announce                               To peer:                               AS prepend 5400
remarks:
remarks:      5400:1000 All peers & Transits                               5400:2000
remarks:
remarks:      5400:1500 All Transits                               5400:2500
remarks:      5400:1501 Sprint Transit (AS1239)                               5400:2501
remarks:      5400:1502 SAVVIS Transit (AS3561)                               5400:2502
remarks:      5400:1503 Level 3 Transit (AS3356)                               5400:2503
remarks:      5400:1504 AT&T Transit (AS7018)                               5400:2504
remarks:      5400:1506 GlobalCrossing Trans (AS3549) 5400:2506
remarks:
remarks:      5400:1001 Nexica (AS24592)                               5400:2001
remarks:      5400:1002 Fujitsu (AS3324)                               5400:2002
remarks:      5400:1004 C&W EU (1273)                               5400:2004
<snip>
notify:       notify@eu.bt.net
mnt-by:       CIP-MNT
source:       RIPE
```



Some ISP Examples

Carrier1

```
aut-num: AS8918
descr: Carrier1 Autonomous System
<snip>
remarks: Community Definition
remarks: *
remarks: 8918:2000 Do not announce to C1 customers
remarks: 8918:2010 Do not announce to C1 peers, peers+ and transit
remarks: 8918:2015 Do not announce to C1 transit providers
remarks: *
remarks: 8918:2020 Do not announce to Teleglobe (AS 6453)
remarks: 8918:2035 Do not announce to UUNet (AS 702)
remarks: 8918:2040 Do not announce to Cogent (AS 174)
remarks: 8918:2050 Do not announce to T-Systems (AS 3320)
remarks: 8918:2060 Do not announce to Sprint (AS 1239)
remarks: *
remarks: 8918:2070 Do not announce to AMS-IX peers
remarks: 8918:2080 Do not announce to NL-IX peers
remarks: 8918:2090 Do not announce to Packet Exchange Peers
<snip>
notify: inoc@carrier1.net
mnt-by: CARRIER1-MNT
source: RIPE
```



Some ISP Examples

Level 3

```
aut-num:          AS3356
descr:           Level 3 Communications
<snip>
remarks:         -----
remarks:         customer traffic engineering communities - Suppression
remarks:         -----
remarks:         64960:XXX - announce to AS XXX if 65000:0
remarks:         65000:0   - announce to customers but not to peers
remarks:         65000:XXX - do not announce at peerings to AS XXX
remarks:         -----
remarks:         customer traffic engineering communities - Prepending
remarks:         -----
remarks:         65001:0   - prepend once  to all peers
remarks:         65001:XXX - prepend once  at peerings to AS XXX
<snip>
remarks:         3356:70   - set local preference to 70
remarks:         3356:80   - set local preference to 80
remarks:         3356:90   - set local preference to 90
remarks:         3356:9999 - blackhole (discard) traffic
<snip>
mnt-by:          LEVEL3-MNT
source:          RIPE
```



And many
many more!

Creating your own community policy

- Consider creating communities to give policy control to customers
 - Reduces technical support burden
 - Reduces the amount of router reconfiguration, and the chance of mistakes
 - Use the previous examples as a guideline

Communities

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



BGP and the Internet

Using Communities for Multihoming