

BGP and the Internet

Using Communities for Multihoming

Multihoming and Communities

Cisco.com

- **The BGP community attribute is a very powerful tool for assisting and scaling BGP Multihoming**

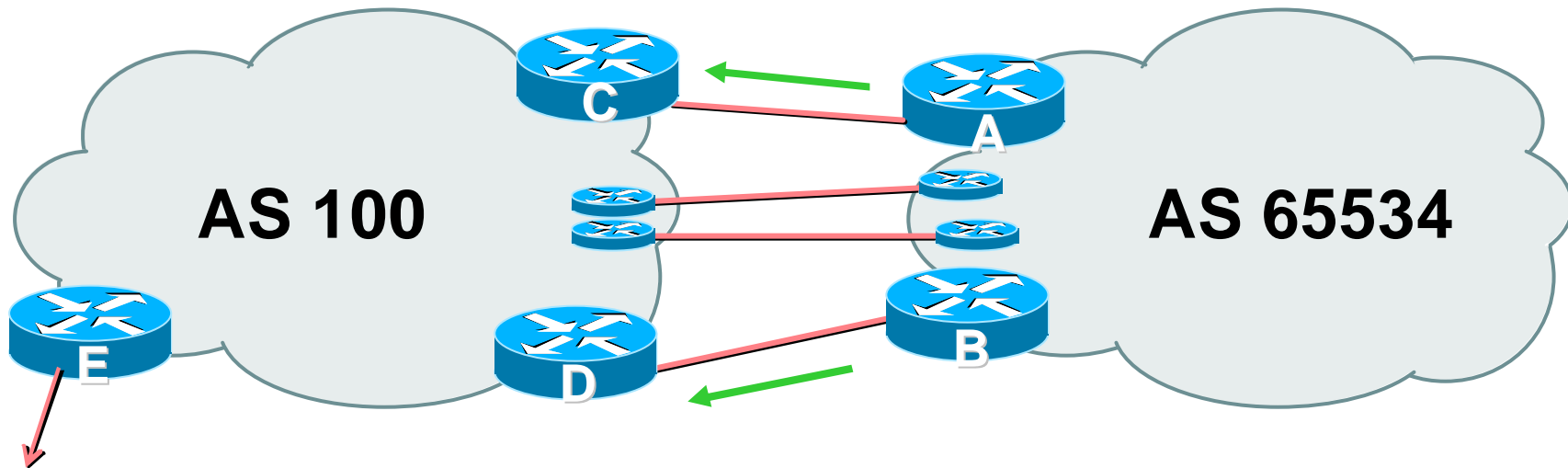
Loadsharing Using Communities

4 links – Private AS

Communities

Private AS

Cisco.com



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

Communities

Private AS

Cisco.com

- **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

Cisco.com

- **Router A Configuration**

```
router bgp 65534
  network 221.10.0.0 mask 255.255.224.0
  network 221.10.0.0 mask 255.255.248.0
  neighbor 222.222.10.2 remote-as 100
  neighbor 222.222.10.2 send-community
  neighbor 222.222.10.2 prefix-list subblocks1 out
  neighbor 222.222.10.2 route-map routerC-out out
  neighbor 222.222.10.2 prefix-list default in
!
..next slide
```

Communities

Private AS

Cisco.com

```
ip prefix-list subblocks1 permit 221.10.0.0/19
ip prefix-list subblocks1 permit 221.10.0.0/21
!
ip prefix-list firstblock permit 221.10.0.0/21
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

Cisco.com

- **Router B Configuration**

```
router bgp 65534
  network 221.10.0.0 mask 255.255.224.0
  network 221.10.24.0 mask 255.255.248.0
  neighbor 222.222.20.2 remote-as 100
  neighbor 222.222.20.2 send-community
  neighbor 222.222.20.2 prefix-list subblocks2 out
  neighbor 222.222.20.2 route-map routerD-out out
  neighbor 222.222.20.2 prefix-list default in
!
..next slide
```


Communities

Private AS

Cisco.com

```
ip prefix-list subblocks2 permit 221.10.0.0/19
ip prefix-list subblocks2 permit 221.10.24.0/21
!
ip prefix-list secondblock permit 221.10.24.0/21
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

Cisco.com

- **Router E Configuration**

```
router bgp 100
  neighbor 222.222.10.17 remote-as 110
  neighbor 222.222.10.17 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

Cisco.com

- 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! :-)

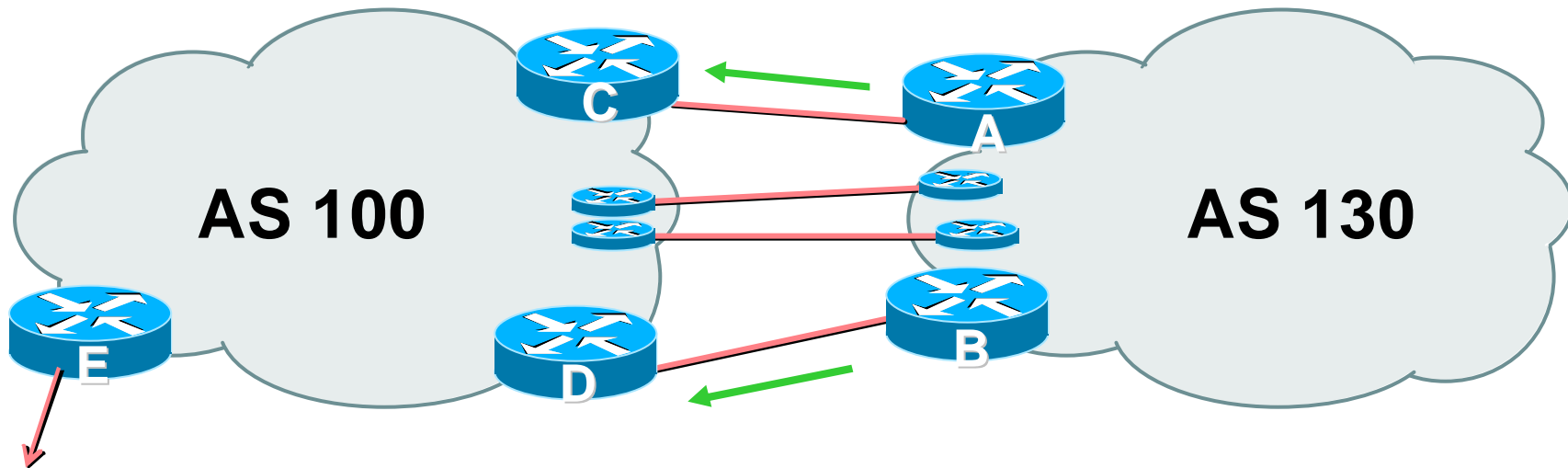
Loadsharing Using Communities

4 links – Public AS

Communities

Public AS

Cisco.com



- 4 links between AS130 and AS100

Communities

Public AS

Cisco.com

- **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

Cisco.com

- **Router A Configuration**

```
router bgp 130
  network 221.10.0.0 mask 255.255.224.0
  network 221.10.0.0 mask 255.255.248.0
  neighbor 222.222.10.2 remote-as 100
  neighbor 222.222.10.2 send-community
  neighbor 222.222.10.2 prefix-list subblocks1 out
  neighbor 222.222.10.2 route-map routerC-out out
  neighbor 222.222.10.2 prefix-list default in
!
..next slide
```

Communities

Public AS

Cisco.com

```
ip prefix-list subblocks1 permit 221.10.0.0/19
ip prefix-list subblocks1 permit 221.10.0.0/21
!
ip prefix-list firstblock permit 221.10.0.0/21
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

Cisco.com

- **Router B Configuration**

```
router bgp 130
  network 221.10.0.0 mask 255.255.224.0
  network 221.10.24.0 mask 255.255.248.0
  neighbor 222.222.20.2 remote-as 100
  neighbor 222.222.20.2 send-community
  neighbor 222.222.20.2 prefix-list subblocks2 out
  neighbor 222.222.20.2 route-map routerD-out out
  neighbor 222.222.20.2 prefix-list default in
!
..next slide
```

Communities

Public AS

Cisco.com

```
ip prefix-list subblocks2 permit 221.10.0.0/19
ip prefix-list subblocks2 permit 221.10.24.0/21
!
ip prefix-list secondblock permit 221.10.24.0/21
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

Cisco.com

- **Router C Configuration**

```
router bgp 100
  neighbor 222.222.10.1 remote-as 130
  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 le 21
ip prefix-list default permit 0.0.0.0/0
```

Communities

Public AS

Cisco.com

- **Router D Configuration**

```
router bgp 100
  neighbor 222.222.10.5 remote-as 130
  neighbor 222.222.10.5 default-originate
  neighbor 222.222.10.5 prefix-list Customer in
  neighbor 222.222.10.5 prefix-list default out
!
ip prefix-list Customer permit 221.10.0.0/19 le 21
ip prefix-list default permit 0.0.0.0/0
```

Communities

Public AS

Cisco.com

- **Router E Configuration**

```
router bgp 100
  neighbor 222.222.10.17 remote-as 110
  neighbor 222.222.10.17 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

Cisco.com

- **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

Cisco.com

- **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!**

- **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

- **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
```

- **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
```

- **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
```

- **Supporting RFC1998**

many ISPs do, more should

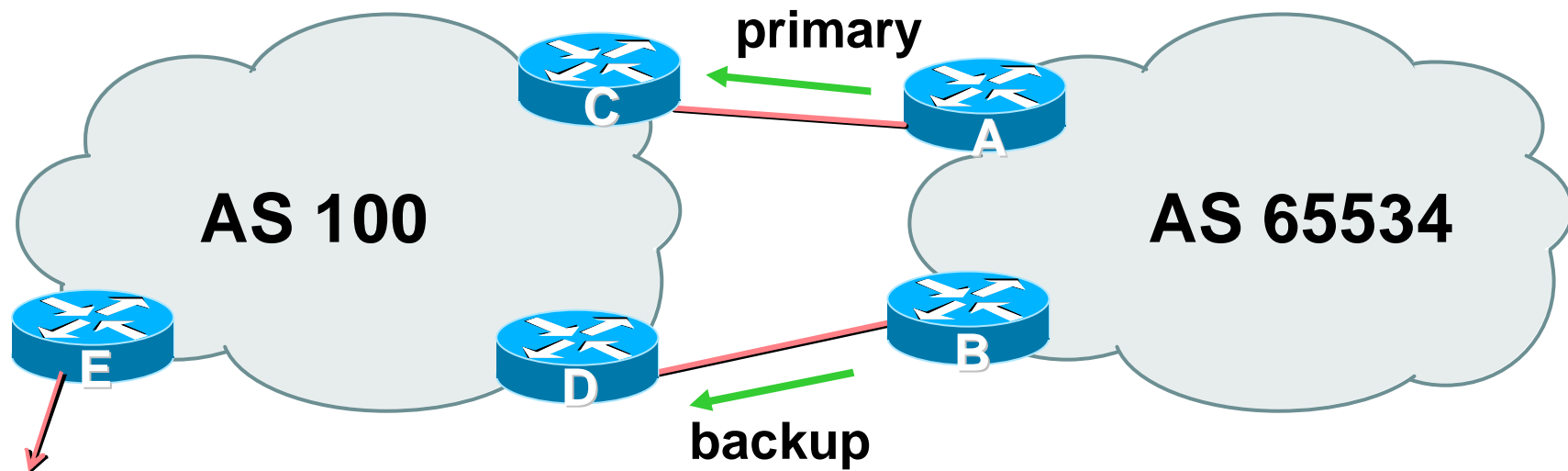
check AS object in the Internet Routing Registry

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)

Cisco.com

- **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)

Cisco.com

- Router A Configuration

```
router bgp 65534
  network 221.10.0.0 mask 255.255.224.0
  neighbor 222.222.10.2 remote-as 100
  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)

Cisco.com

- **Router B Configuration**

```
router bgp 65534
  network 221.10.0.0 mask 255.255.224.0
  neighbor 222.222.10.6 remote-as 100
  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)

Cisco.com

```
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 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)

Cisco.com

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

```
router bgp 100
  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)

Cisco.com

- Router D Configuration (backup link)

```
router bgp 100
  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)

Cisco.com

```
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 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)

Cisco.com

- **Router E Configuration**

```
router bgp 100
  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
```

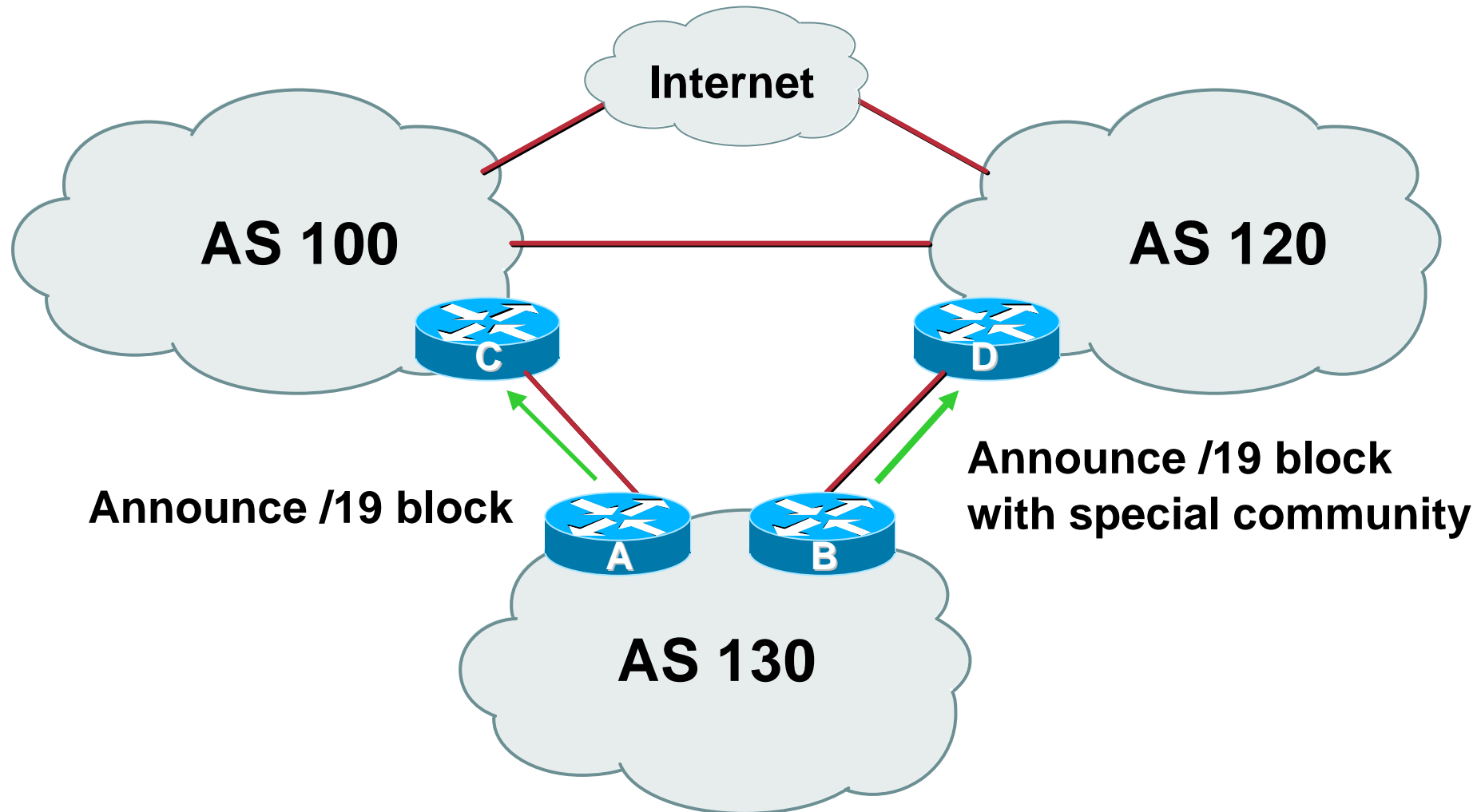
- **Router E removes prefixes in the private AS from external announcements**
- **Private AS still visible inside AS100**

Two links to different ISPs

One link primary, the other link backup only

Two links to different ISPs (one as backup only)

Cisco.com



Two links to different ISPs (one as backup only)

Cisco.com

- **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)

Cisco.com

- **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)

Cisco.com

- **Router A Configuration**

```
router bgp 130
  network 221.10.0.0 mask 255.255.224.0
  neighbor 222.222.10.1 remote-as 100
  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 100:100
```

Two links to different ISPs (one as backup only)

Cisco.com

- **Router B Configuration**

```
router bgp 130
  network 221.10.0.0 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)

Cisco.com

```
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 120:80
!
route-map routerD-in permit 10
    set local-preference 80
```

Two links to different ISPs (one as backup only)

Cisco.com

- **Router D**

**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

Two links to different ISPs (one as backup only)

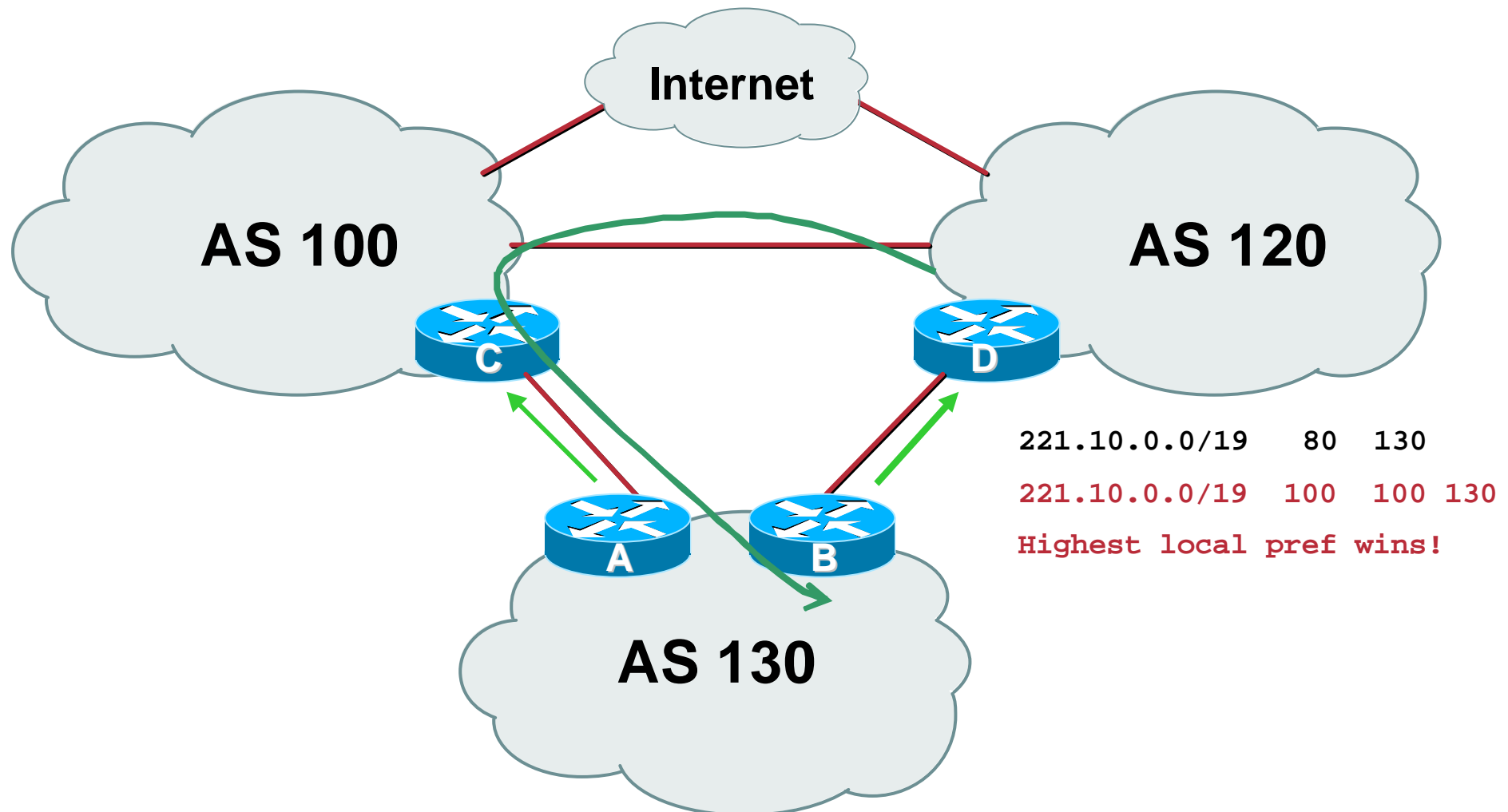
Cisco.com

- **Router D**

**Only requires RFC1998 configuration
no per customer configuration
scalability!**

Two links to different ISPs (one as backup only)

Cisco.com



Two links to different ISPs (one as backup only)

Cisco.com

- **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

Cisco.com

- **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

Some ISP Examples

Cisco.com

- **Public policy is usually listed in the IRR**

Following examples are all in the IRR or referenced from the AS Object in the IRR

- **Consider creating communities to give policy control to customers**

Reduces technical support burden

Reduces the amount of router reconfiguration, and the chance of mistakes

Some ISP

Connect

```
aut-num:          AS2764
as-name:          ASN-CONNECT-NET
descr:            connect.com.au pty ltd
admin-c:          CC89
tech-c:           MP151
remarks:          Community Definition
remarks:          -----
remarks:          2764:1 Announce to "domestic" rate ASes only
remarks:          2764:2 Don't announce outside local POP
remarks:          2764:3 Lower local preference by 25
remarks:          2764:4 Lower local preference by 15
remarks:          2764:5 Lower local preference by 5
remarks:          2764:6 Announce to non customers with "no-export"
remarks:          2764:7 Only announce route to customers
remarks:          2764:8 Announce route over satellite link
notify:           routing@connect.com.au
mnt-by:           CONNECT-AU
changed:          mrp@connect.com.au 19990506
source:           CCAIR
```

Some IS

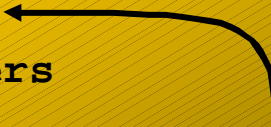
UUNET

```
aut-num: AS702
as-name: AS702
descr: UUNET - Commercial IP service provider in Europe
remarks: -----
remarks: UUNET uses the following communities with its customers:
remarks: 702:80 Set Local Pref 80 within AS702
remarks: 702:120 Set Local Pref 120 within AS702
remarks: 702:20 Announce only to UUNET AS'es and UUNET customers
remarks: 702:30 Keep within Europe, don't announce to other UUNET AS's
remarks: 702:1 Prepend AS702 once at edges of UUNET to Peers
remarks: 702:2 Prepend AS702 twice at edges of UUNET to Peers
remarks: 702:3 Prepend AS702 thrice at edges of UUNET to Peers
remarks: Details of UUNET's peering policy and how to get in touch with
remarks: UUNET regarding peering policy matters can be found at:
remarks: http://www.uu.net/peering/
remarks: -----
mnt-by: UUNET-MNT
changed: eric-apps@eu.uu.net 20010928
source: RIPE
```


Some ISP

BT Ignite

```
aut-num: AS5400
as-name: CIPCORE
descr: BT Ignite European Backbone
remarks: The following BGP communities can be set by BT Ignite
remarks: BGP customers to affect announcements to major peers.
remarks:
remarks: Community to Community to
remarks: Not announce To peer: AS prepend 5400
remarks:
remarks: 5400:1000 European peers 5400:2000
remarks: 5400:1001 Sprint (AS1239) 5400:2001
remarks: 5400:1003 Unisource (AS3300) 5400:2003
remarks: 5400:1005 UUnet (AS702) 5400:2005
remarks: 5400:1006 Carrier1 (AS8918) 5400:2006
remarks: 5400:1007 SupportNet (8582) 5400:2007
remarks: 5400:1008 AT&T (AS2686) 5400:2008
remarks: 5400:1009 Level 3 (AS9057) 5400:2009
remarks: 5400:1010 RIPE (AS3333) 5400:2010
<snip>
remarks: 5400:1100 US peers 5400:2100
notify: notify@eu.ignite.net
mnt-by: CIP-MNT
source: RIPE
```



**And many
many more!**

Some ISP Carrier

```
aut-num:      AS8918
descr:        Carrier1 Autonomous System
<snip>
remarks:      Community Support Definitions:
remarks:      Communities that determine the geographic
remarks:      entry point of routes into the Carrier1 network:
remarks:      *
remarks:      Community      Entry Point
remarks:      -----
remarks:      8918:10         London
remarks:      8918:15         Hamburg
remarks:      8918:18         Chicago
remarks:      8918:20         Amsterdam
remarks:      8918:25         Milan
remarks:      8918:28         Berlin
remarks:      8918:30         Frankfurt
remarks:      8918:35         Zurich
remarks:      8918:40         Geneva
remarks:      8918:45         Stockholm
<snip>
notify:       inoc@carrier1.net
mnt-by:       CARRIER1-MNT
source:       RIPE
```

And many
many more!

Some ISP 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
remarks:          65002:0   - prepend twice  to all peers
remarks:          65002:XXX - prepend twice  at peerings to AS XXX
remarks:          65003:0   - prepend 3x     to all peers
remarks:          65003:XXX - prepend 3x     at peerings to AS XXX
remarks:          65004:0   - prepend 4x     to all peers
remarks:          65004:XXX - prepend 4x     at peerings to AS XXX
<snip>
mnt-by:           LEVEL3-MNT
source:           RIPE
```

**And many
many more!**

Communities

Cisco.com

- **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