BGP Configuration for IXPs

ISP/IXP Workshops

Background

- This presentation covers the BGP configurations required for a participant at an Internet Exchange Point

  It does not cover the technical design of an IXP

  Nor does it cover the financial and operational benefits of participating in an IXP

  See the IXP Design Presentation that is part of this Workshop Material set for financial, technical and operational details

Recap: Definitions

- Transit – carrying traffic across a network, usually for a fee

    Traffic and prefixes originating from one AS are carried across an intermediate AS to reach their destination AS

- Peering – private interconnect between two ASNs, usually for no fee

- Internet Exchange Point – common interconnect location where several ASNs exchange routing information and traffic

IXP Peering Issues

- Only announce your prefixes and your customer prefixes at IXPs

- Only accept the prefixes which your peer is entitled to originate

- Never carry a default route on an IXP (or private) peering router

ISP Transit Issues

Many mistakes are made on the Internet today due to incomplete understanding of how to configure BGP for peering at Internet Exchange Points
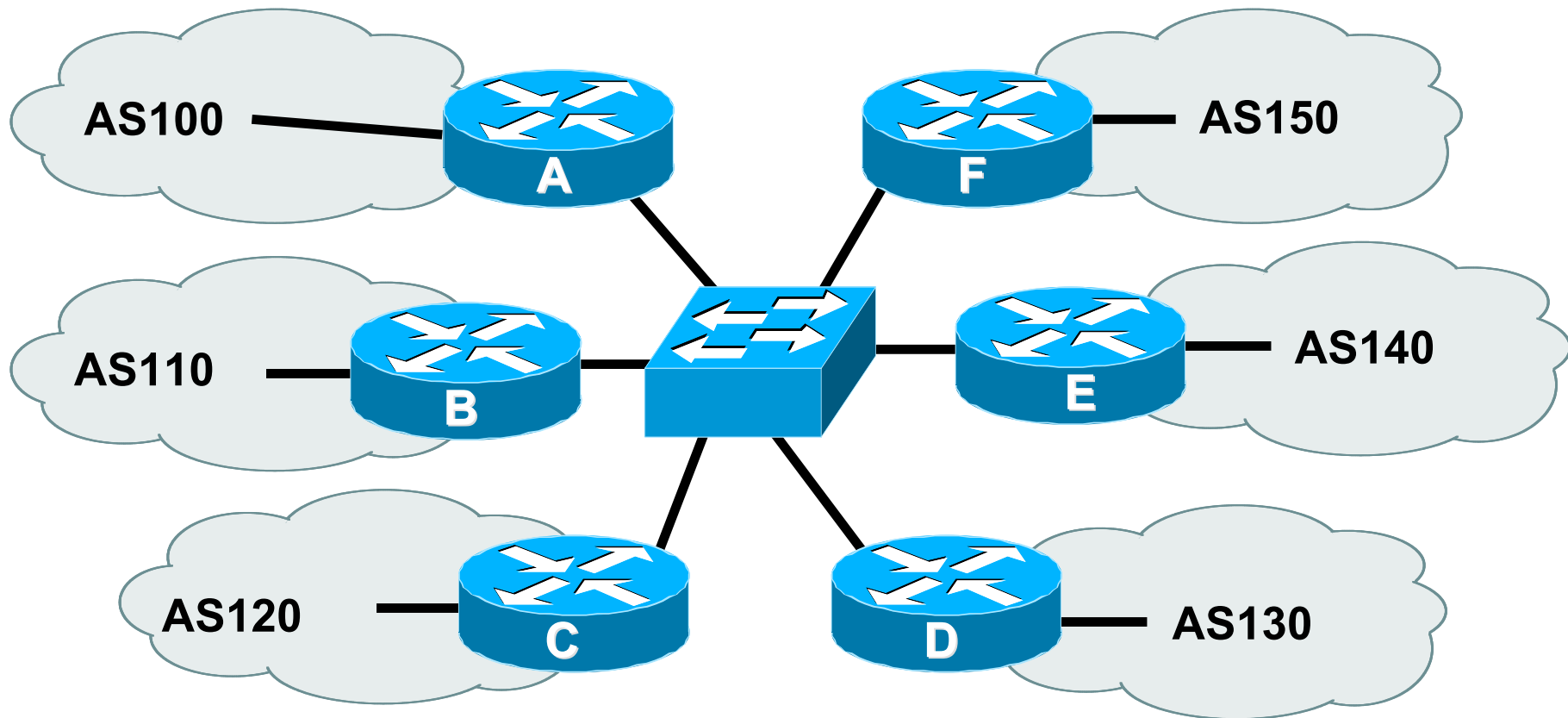
Simple BGP Configuration example

Exchange Point Configuration

Exchange Point Example

- Exchange point with 6 ASes present
  - Layer 2 – ethernet switch

- Each ISP peers with the other
  - NO transit across the IXP allowed

Exchange Point

AS100 — A

AS150 — F

AS110 — B

AS140 — E

AS120 — C

AS130 — D

- Each of these represents a border router in a different autonomous system

Router configuration

- IXP router is usually located at the Exchange Point premises

    So configuration needs to be such that disconnecting it from the backbone does not cause routing loops or traffic blackholes

- Create a peer-group for IXP peers

    All outbound policy to each peer will be the same

- Ensure the router is not carrying the default route

    Or the full routing table (for that matter)

Creating a peer-group & route-map

```
router bgp 100
 neighbor ixp-peer peer-group
 neighbor ixp-peer send-community
 neighbor ixp-peer prefix-list my-prefixes out
 neighbor ixp-peer route-map set-local-pref in
!
ip prefix-list my-prefixes permit 121.10.0.0/19
!
route-map set-local-pref permit 10
 set local-preference 150
!
```
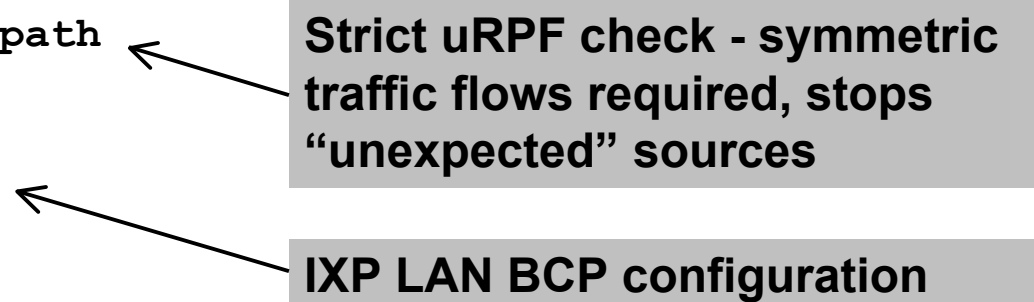
Only allow AS100 address block to IXP peers

Prefixes heard from IXP peers have highest preference

Interface and BGP configuration (1)

```
interface fastethernet 0/0
 description Exchange Point LAN
 ip address 120.5.10.1 mask 255.255.255.224
 ip verify unicast reverse-path
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
router bgp 100
 neighbor 120.5.10.2 remote-as 110
 neighbor 120.5.10.2 peer-group ixp-peer
 neighbor 120.5.10.2 prefix-list peer110 in
 neighbor 120.5.10.3 remote-as 120
 neighbor 120.5.10.3 peer-group ixp-peers
 neighbor 120.5.10.3 prefix-list peer120 in
```

Strict uRPF check - symmetric traffic flows required, stops "unexpected" sources

IXP LAN BCP configuration

Interface and BGP Configuration (2)

```
neighbor 120.5.10.4 remote-as 130

neighbor 120.5.10.4 peer-group ixp-peers

neighbor 120.5.10.4 prefix-list peer130 in

neighbor 120.5.10.5 remote-as 140

neighbor 120.5.10.5 peer-group ixp-peers

neighbor 120.5.10.5 prefix-list peer140 in

neighbor 120.5.10.6 remote-as 150

neighbor 120.5.10.6 peer-group ixp-peers

neighbor 120.5.10.6 prefix-list peer150 in

!

ip route 121.10.0.0 255.255.224.0 null0

!

ip prefix-list peer110 permit 122.0.0.0/19

ip prefix-list peer120 permit 122.30.0.0/19

ip prefix-list peer130 permit 122.12.0.0/19

ip prefix-list peer140 permit 122.18.128.0/19

ip prefix-list peer150 permit 122.1.32.0/19
```

Peer-group applied to each peer

Each peer has own inbound filter

Exchange Point

- Configuration of the other routers in the AS is similar in concept

- Notice inbound and outbound prefix filters

     outbound announces myprefixes only

     inbound accepts peer prefixes only

- Notice inbound route-map

     Set local preference higher than default ensures that if the same prefix is heard via AS100 upstream, the best path for traffic is via the IXP

Exchange Point

- Ethernet port configuration

    Use `ip verify unicast reverse-path`

    Helps prevent "stealing of bandwidth"

    (Only traffic sourced from address space announced by the IXP peers will be permitted in this interface)

- IXP border router must NOT carry prefixes with origin outside local AS and IXP participant ASes

    Helps prevent "stealing of bandwidth"

Exchange Point

- Issues:

    AS100 needs to know all the prefixes its peers are announcing

    New prefixes requires the prefix-lists to be updated

- Alternative solutions

    Use the Internet Routing Registry to build prefix list
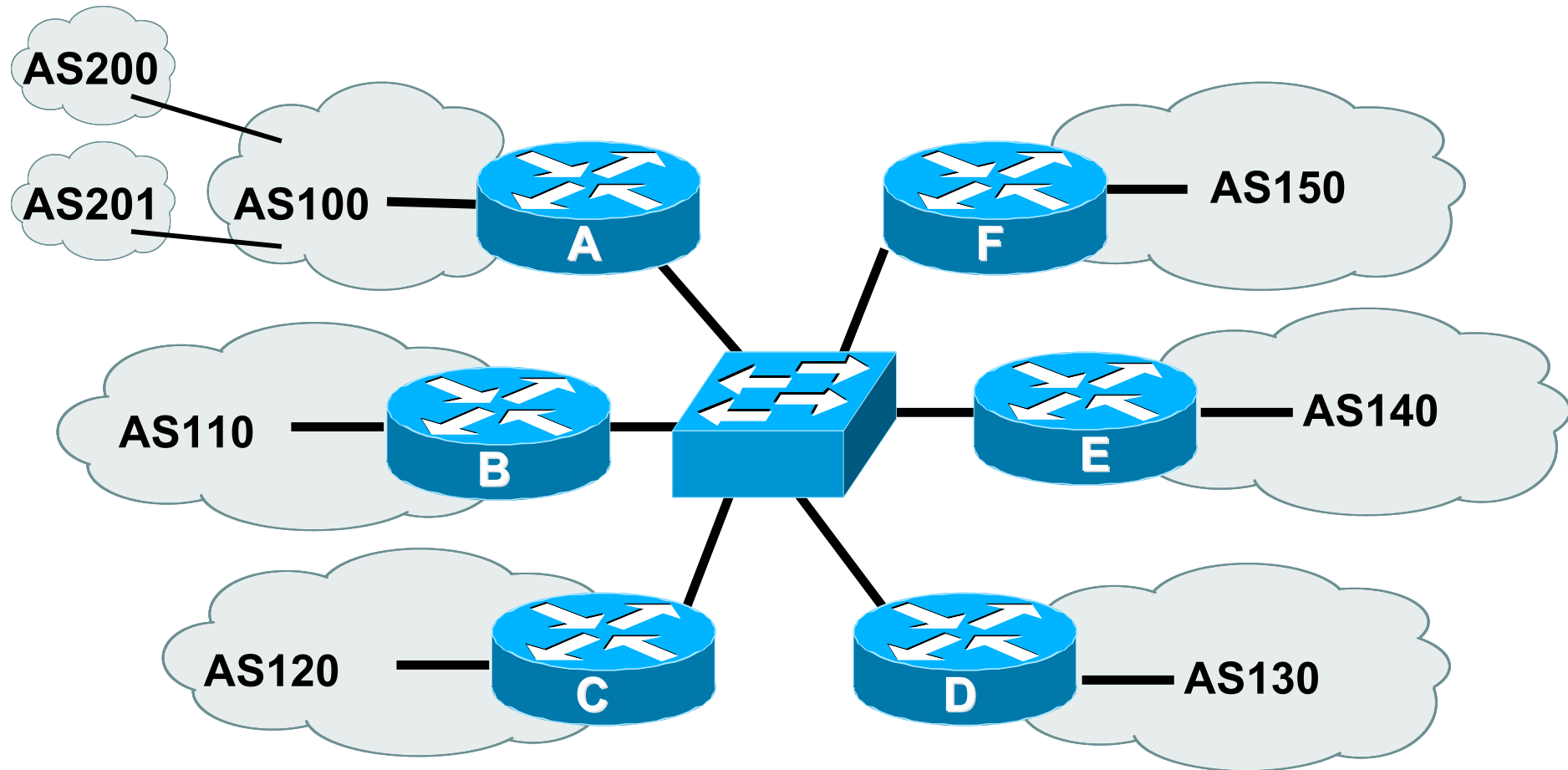
    Use AS Path filters (could be risky)

More Complex BGP example

Exchange Point Configuration

Exchange Point Example

- Exchange point with 6 ASes present
  - Layer 2 – ethernet switch

- Each ISP peers with the other
  - NO transit across the IXP allowed
  - ISPs at exchange points provide transit to their BGP customers

Exchange Point

- Each of these represents a border router in a different autonomous system

Exchange Point Router A configuration

```
interface fastethernet 0/0
 description Exchange Point LAN
 ip address 120.5.10.2 mask 255.255.255.224
 ip verify unicast reverse-path
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
router bgp 100
 neighbor ixp-peers peer-group
 neighbor ixp-peers send-community
 neighbor ixp-peers prefix-list bogons out
 neighbor ixp-peers filter-list 10 out
 neighbor ixp-peers route-map set-local-pref in
..next slide
```

Filter by ASN rather than by prefix - and block bogons too

Exchange Point

```
neighbor 120.5.10.2 remote-as 110
neighbor 120.5.10.2 peer-group ixp-peers
neighbor 120.5.10.2 prefix-list peer110 in
neighbor 120.5.10.3 remote-as 120
neighbor 120.5.10.3 peer-group ixp-peers
neighbor 120.5.10.3 prefix-list peer120 in
neighbor 120.5.10.4 remote-as 130
neighbor 120.5.10.4 peer-group ixp-peers
neighbor 120.5.10.4 prefix-list peer130 in
neighbor 120.5.10.5 remote-as 140
neighbor 120.5.10.5 peer-group ixp-peers
neighbor 120.5.10.5 prefix-list peer140 in
neighbor 120.5.10.6 remote-as 150
neighbor 120.5.10.6 peer-group ixp-peers
neighbor 120.5.10.6 prefix-list peer150 in
```

Exchange Point

```
ip route 121.10.0.0 255.255.224.0 null0
!
ip as-path access-list 10 permit ^$
ip as-path access-list 10 permit ^200$
ip as-path access-list 10 permit ^201$
!
ip prefix-list peer110 permit 122.0.0.0/19
ip prefix-list peer120 permit 122.30.0.0/19
ip prefix-list peer130 permit 122.12.0.0/19
ip prefix-list peer140 permit 122.18.128.0/19
ip prefix-list peer150 permit 122.1.32.0/19
!
route-map set-local-pref permit 10
 set local-preference 150
```

Exchange Point

- Notice the change in router A's configuration

    Filter-list instead of prefix-list permits local and customer ASes out to exchange

    Prefix-list blocks Special Use Address prefixes – rest get out, could be risky

- Other issues as previously

- This configuration will not scale as more and more BGP customers are added to AS100

    As-path filter has to be updated each time

    Solution: BGP communities
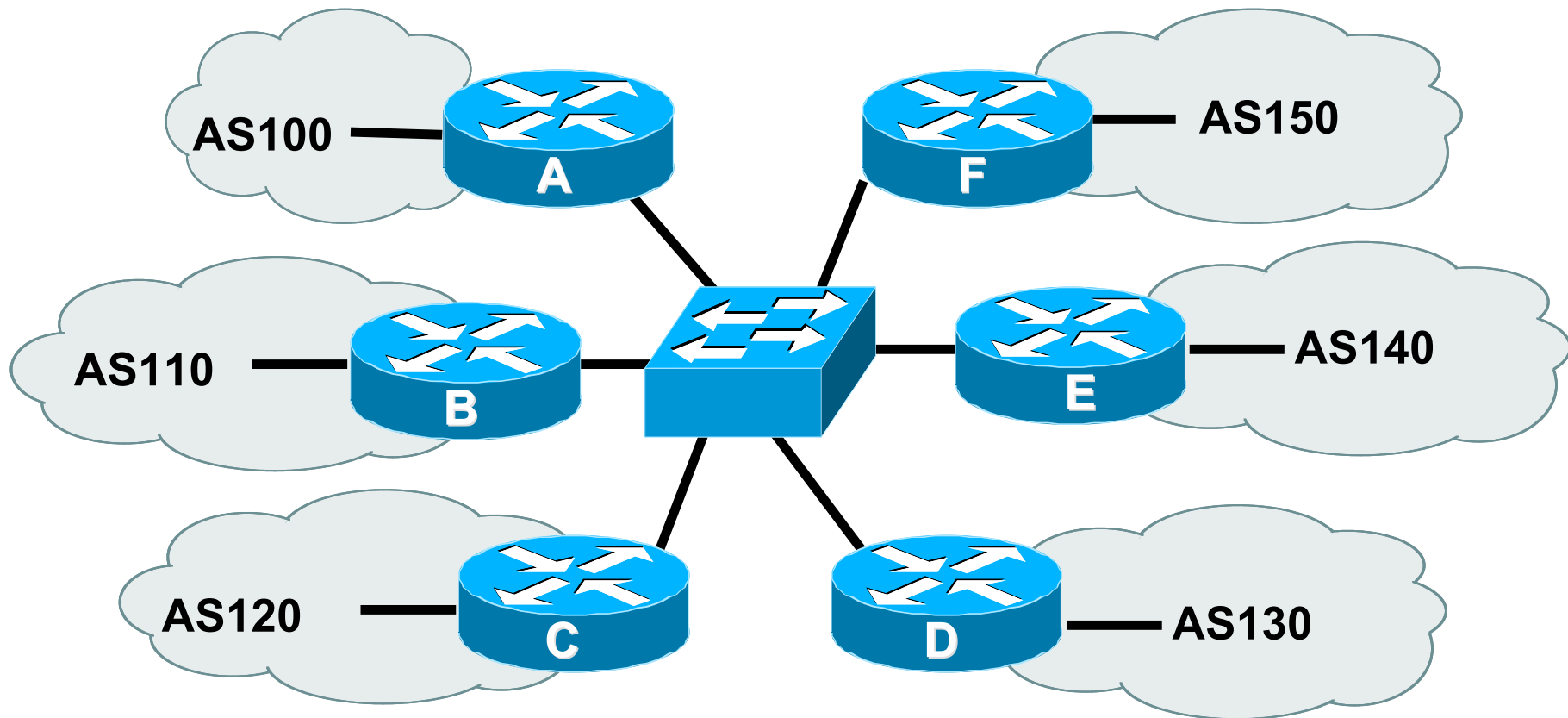
More scalable BGP example

Exchange Point Configuration

Exchange Point Example (Scalable)

- Exchange point with 6 ASes present
    - Layer 2 – ethernet switch

- Each ISP peers with the other
    - NO transit across the IXP allowed
    - ISPs at exchange points provide transit to their BGP customers

- (Scalable solution is presented here)

Exchange Point

- Each of these represents a border router in a different autonomous system - each ASN has BGP customers of their own

Router configuration

- Take AS100 as an example

    Has 15 BGP customers, in AS501 to AS514

- Create a peer-group for IXP peers

    All outbound policy to each peer will be the same

- Communities will be used

    AS-path filters will not scale well

- Community Policy

    AS100 aggregate put into 100:1000

    All BGP customer prefixes go into 100:1100

Creating a peer-group & route-map

```
router bgp 100
 neighbor ixp-peer peer-group
 neighbor ixp-peer send-community
 neighbor ixp-peer route-map ixp-peers-out out
 neighbor ixp-peer route-map set-local-pref in
!
ip community-list 10 permit 100:1000          ← AS100 aggregate
ip community-list 11 permit 100:1100          ← AS100 BGP customers
!
route-map ixp-peers-out permit 10
 match community 10 11
!
route-map set-local-pref permit 10            Prefixes heard from IXP peers
 set local-preference 150                     have highest preference
!
```

BGP configuration for IXP router

```
router bgp 100
 neighbor 120.5.10.2 remote-as 110
 neighbor 120.5.10.2 peer-group ixp-peer
 neighbor 120.5.10.2 prefix-list peer110 in
 neighbor 120.5.10.3 remote-as 120
 neighbor 120.5.10.3 peer-group ixp-peers
 neighbor 120.5.10.3 prefix-list peer120 in
...etc
```

- Remaining configuration is the same as earlier
- Note the reliance again on inbound prefix-lists for peers
  - Peers need to update the ISP if filters need to be changed
  - And that's what the IRR is for (otherwise use email)

BGP configuration for customer aggregation router

```
router bgp 100
 network 121.10.0.0 mask 255.255.192.0 route-map set-comm
 neighbor 121.10.4.2 remote-as 501
 neighbor 121.10.4.2 prefix-list as510-in in
 neighbor 121.10.4.2 prefix-list default out
 neighbor 121.10.4.2 route-map set-cust-policy in
...etc
!
route-map set-comm permit 10
 set community 100:1000
!
route-map set-cust-policy permit 10
 set community 100:1100
!
```

Set community on AS100 aggregate

Set community on BGP customer routes

Scalable IXP policy

- ISP now relies on communities to determine what is announced at the IXP

    No need to update any as-path filters, prefix-lists, &c

- If BGP customer announces more prefixes, only the filters at the aggregation edge need to be updated

    And those new prefixes will automatically be tagged with the community to allow them through to AS100's IXP peers

- Consult the BGP community presentation for more extensive examples

Route Servers

- IXP operators quite often provide a Route Server to assist with scaling the BGP mesh

  - All prefixes sent to a Route Server are usually distributed to all ASNs that peer with the Route Server

  - (although some IXPs offer ISPs the facility to configure specific policies on their Route Server)

- BGP configuration to peer with a Route Server is the same as for any other ordinary peer

  - But note that the route server will offer prefixes from several ASNs (the IXP membership who choose to participate)

  - Inbound filter should be constructed appropriately

Route Servers

- Route Server software suppresses the ASN of the RS so that it doesn't appear in the AS-path

- IOS by default will **not** accept prefixes from a neighbouring AS unless that AS is first in the AS-path

Needed so that IOS can receive prefixes without AS65534 being first in path

```
router bgp 100
 no bgp enforce-first-as
 neighbor x.x.x.a remote-as 65534
 neighbor x.x.x.a route-map IXP-RS-in in
 neighbor x.x.x.a route-map ixp-peers-out out
```

Summary

Exchange Point Configuration

Summary

- Ensure that BGP is scalable on your IXP peering router

    Manually updating filters every time a new customer connects is tiresome and has potential to cause errors

- Only carry local ASN prefixes and customer routes on the IXP peering router

    Anything else (eg default or full BGP table) has the potential to result in bandwidth theft

- Filter IXP peer announcements

    Inbound - use the IRR if maintaining prefix-lists is difficult

    Outbound - use communities for scalability

BGP Configuration for IXPs

ISP/IXP Workshops