



# Introduction to BGP

**ISP/IXP Workshops**

# Border Gateway Protocol

- **Routing Protocol used to exchange routing information between networks**  
**exterior gateway protocol**
- **Described in RFC4271**  
**RFC4276 gives an implementation report on BGP-4**  
**RFC4277 describes operational experiences using BGP-4**
- **The Autonomous System is BGP's fundamental operating unit**  
**It is used to uniquely identify networks with common routing policy**

# BGP

- **Path Vector Protocol**
- **Incremental Updates**
- **Many options for policy enforcement**
- **Classless Inter Domain Routing (CIDR)**
- **Widely used for Internet backbone**
- **Autonomous systems**

# Path Vector Protocol

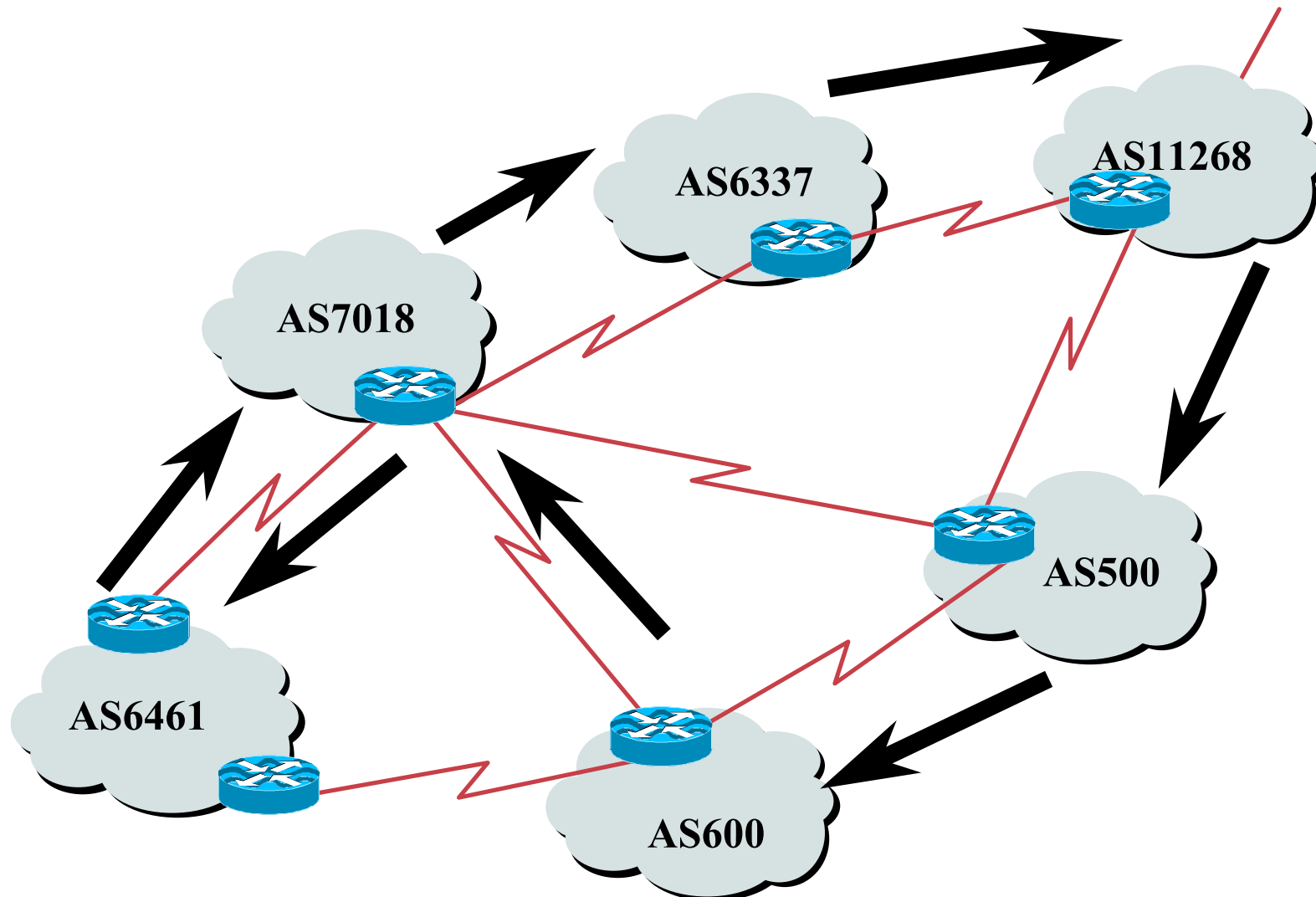
- **BGP is classified as a *path vector* routing protocol**  
(see RFC 1322)

**A path vector protocol defines a route as a pairing between a destination and the attributes of the path to that destination.**

12.6.126.0/24 207.126.96.43 1021 0 6461 7018 6337 11268 i

**AS Path**

# Path Vector Protocol



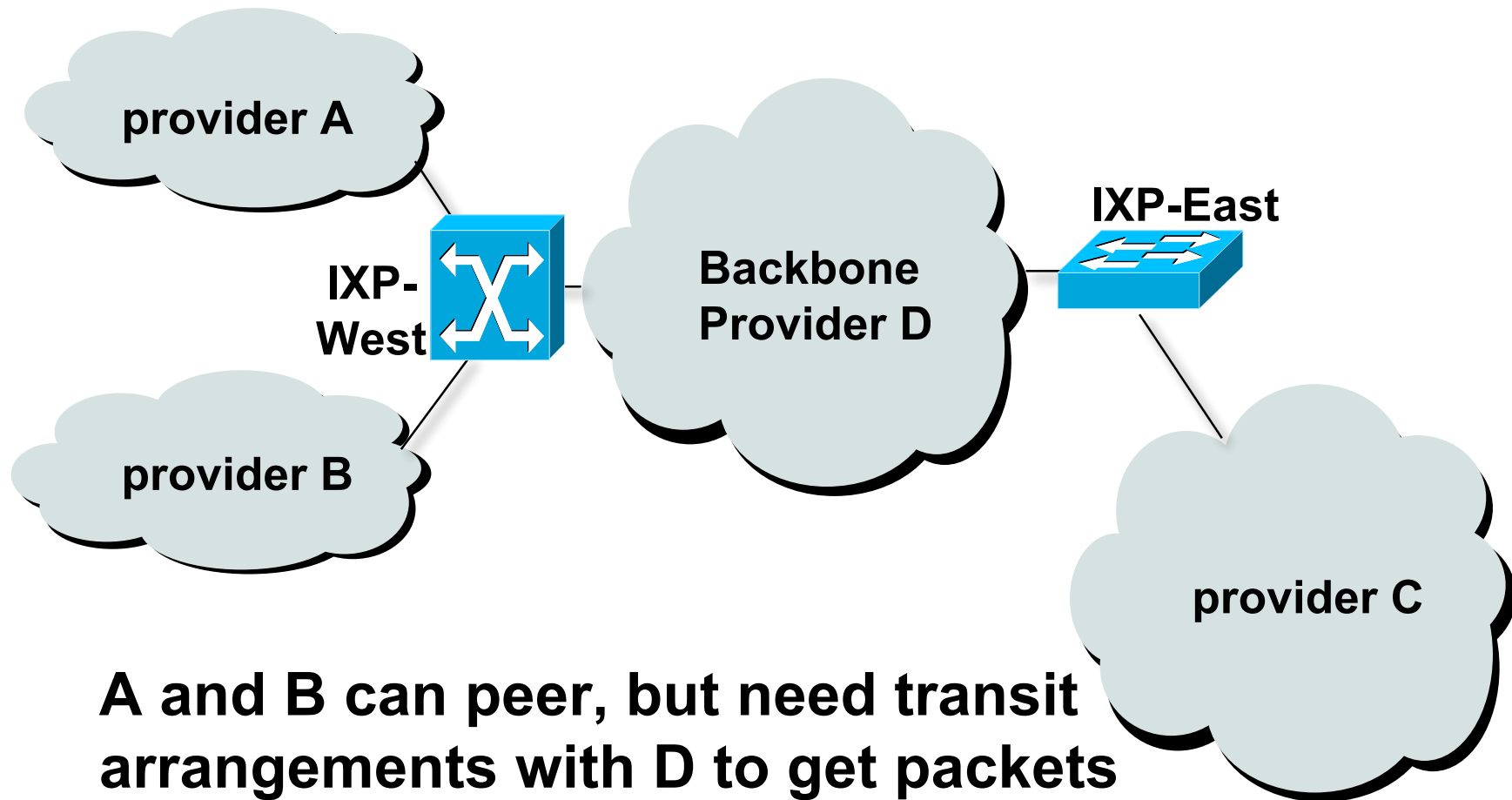
# Definitions

- **Transit** – carrying traffic across a network, usually for a fee
- **Peering** – exchanging routing information and traffic
- **Default** – where to send traffic when there is no explicit match in the routing table

## Default Free Zone

**The default free zone is made up of Internet routers which have explicit routing information about the rest of the Internet, and therefore do not need to use a default route.**

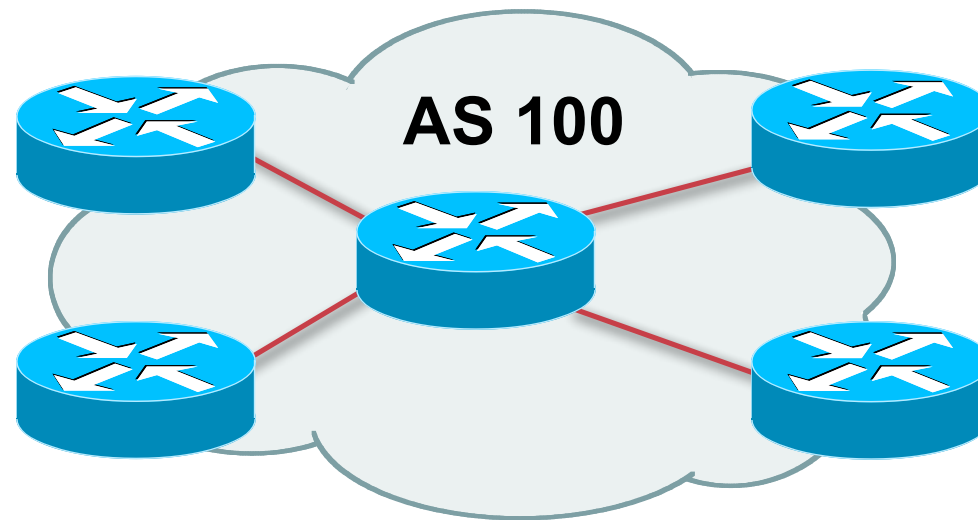
# Peering and Transit example



**A and B can peer, but need transit arrangements with D to get packets to/from C**



# Autonomous System (AS)

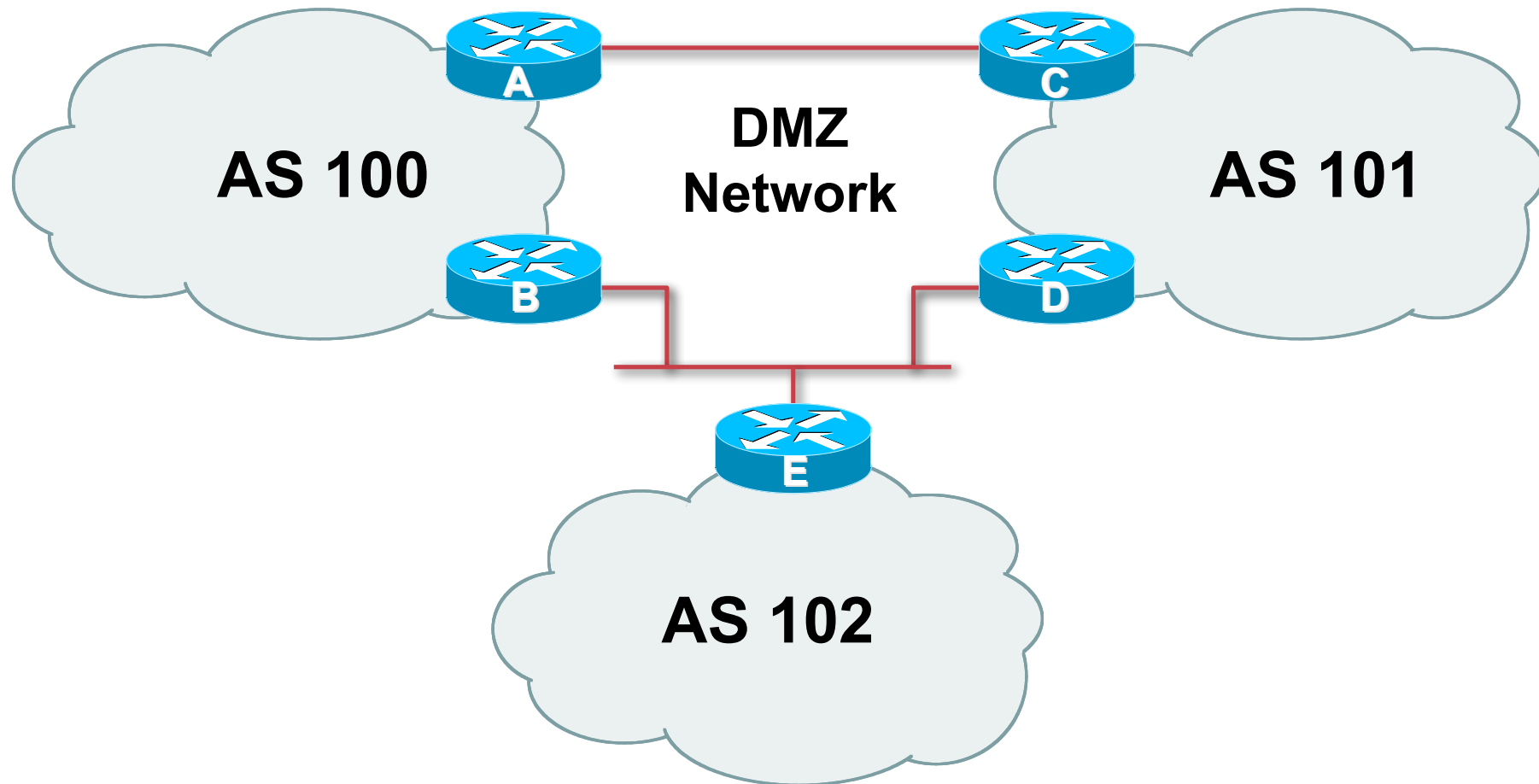


- **Collection of networks with same routing policy**
- **Single routing protocol**
- **Usually under single ownership, trust and administrative control**
- **Identified by a unique number**

# Autonomous System Number (ASN)

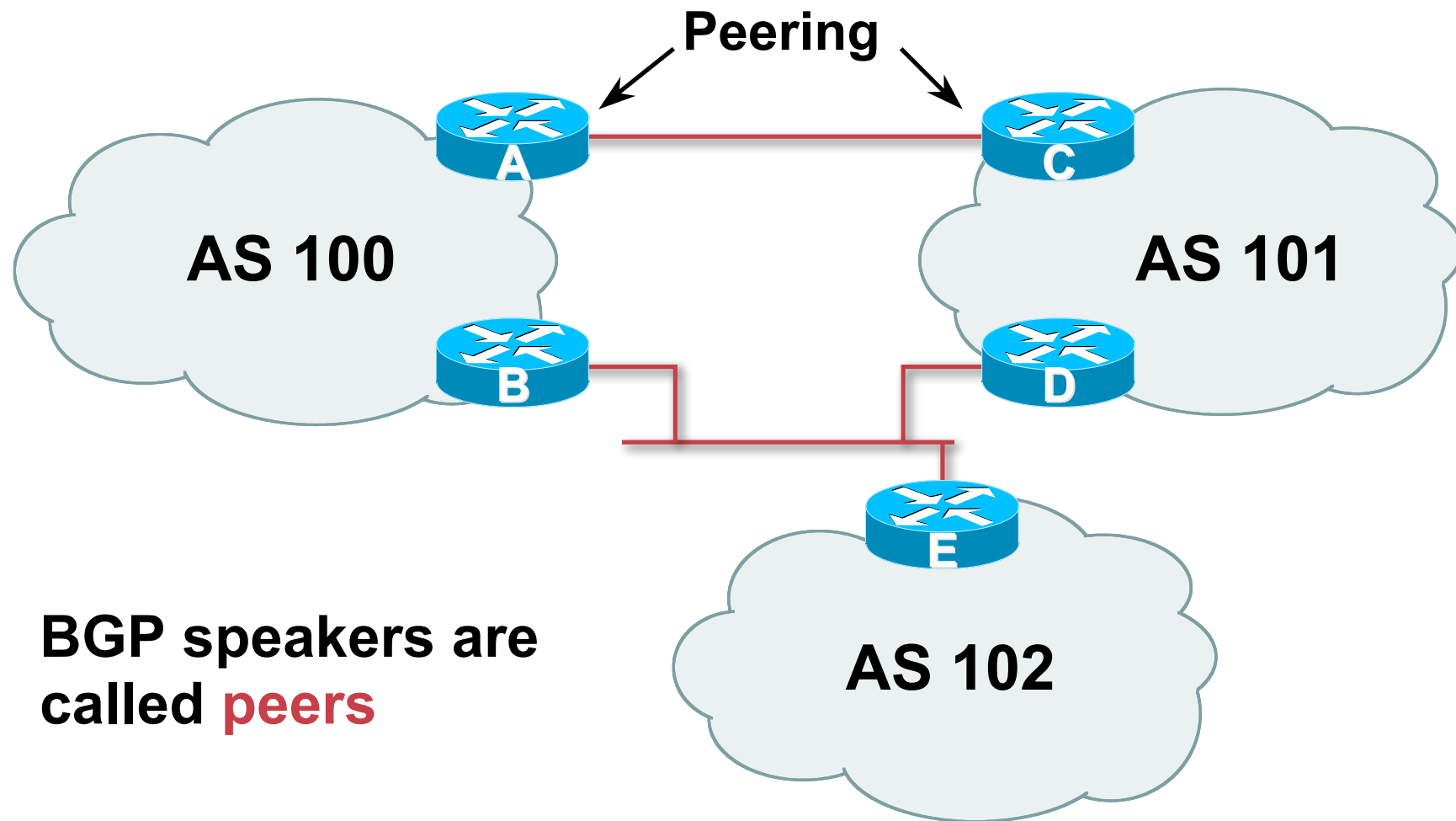
- An ASN is a 16 bit number
  - 1-64511 are assigned by the RIRs
  - 64512-65534 are for private use and should never appear on the Internet
  - 0 and 65535 are reserved
- 32 bit ASNs are coming soon
  - [www.ietf.org/internet-drafts/draft-ietf-idr-as4bytes-12.txt](http://www.ietf.org/internet-drafts/draft-ietf-idr-as4bytes-12.txt)
- ASNs are distributed by the Regional Internet Registries
  - Also available from upstream ISPs who are members of one of the RIRs
  - Current ASN allocations up to 39935 have been made to the RIRs

# Demarcation Zone (DMZ)



- Shared network between ASes

# BGP Basics



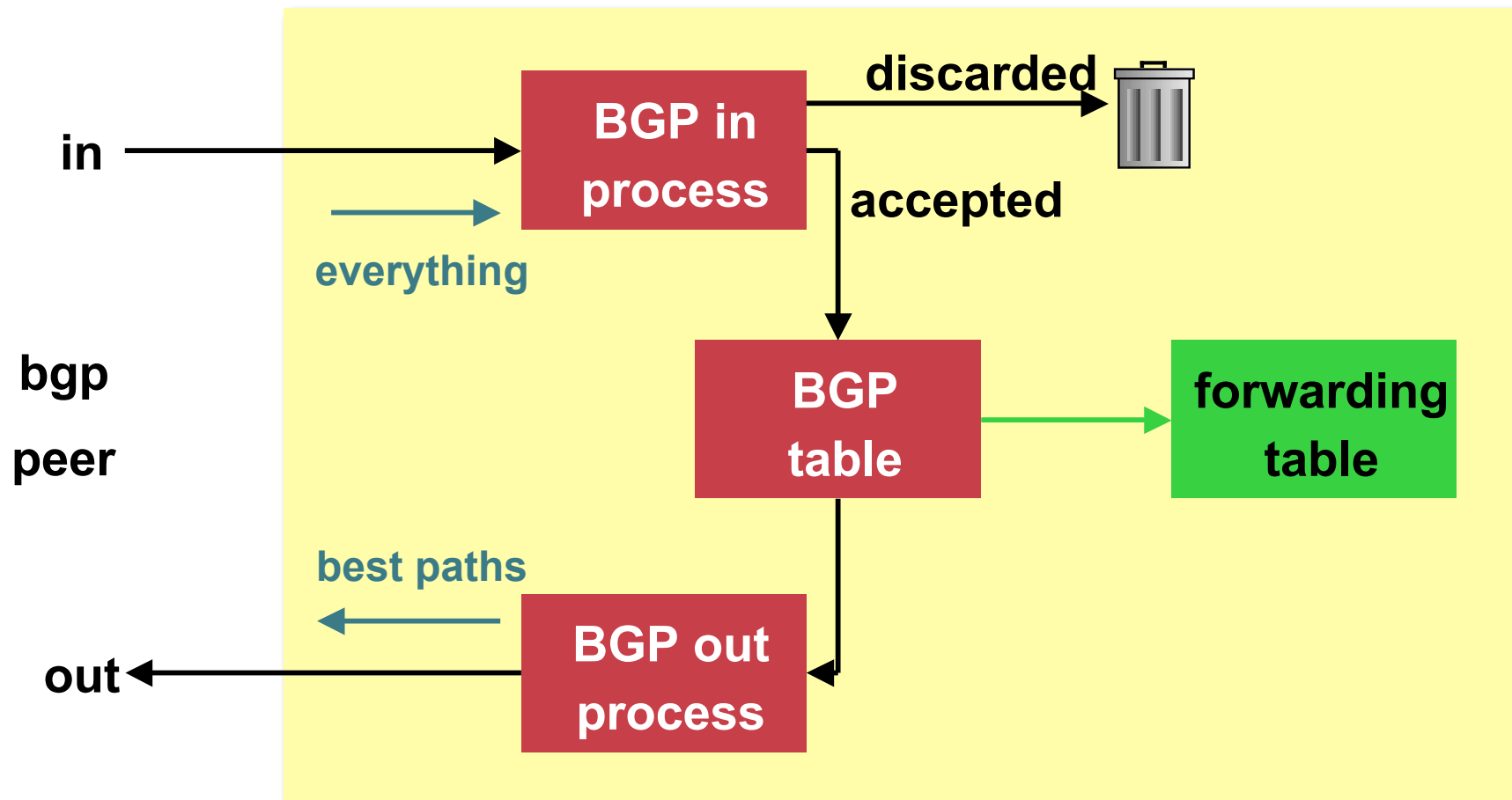
# BGP General Operation

- **Learns multiple paths via internal and external BGP speakers**
- **Picks the best path and installs in the forwarding table**
- **Policies applied by influencing the best path selection**

# Constructing the Forwarding Table

- **BGP “in” process**
  - receives path information from peers
  - results of BGP path selection placed in the BGP table
  - “best path” flagged
- **BGP “out” process**
  - announces “best path” information to peers
- **Best paths installed in forwarding table if:**
  - prefix and prefix length are unique
  - lowest “protocol distance”

# Constructing the Forwarding Table



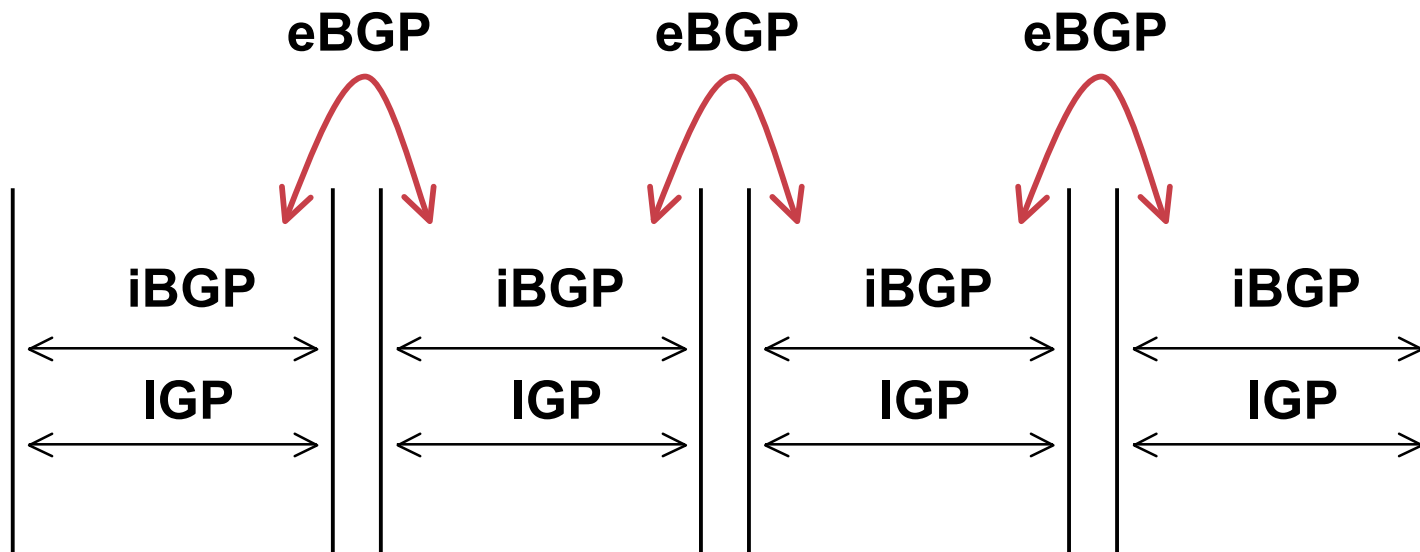
# eBGP & iBGP

- **BGP used internally (iBGP) and externally (eBGP)**
- **iBGP used to carry**
  - some/all Internet prefixes across ISP backbone**
  - ISP's customer prefixes**
- **eBGP used to**
  - exchange prefixes with other ASes**
  - implement routing policy**

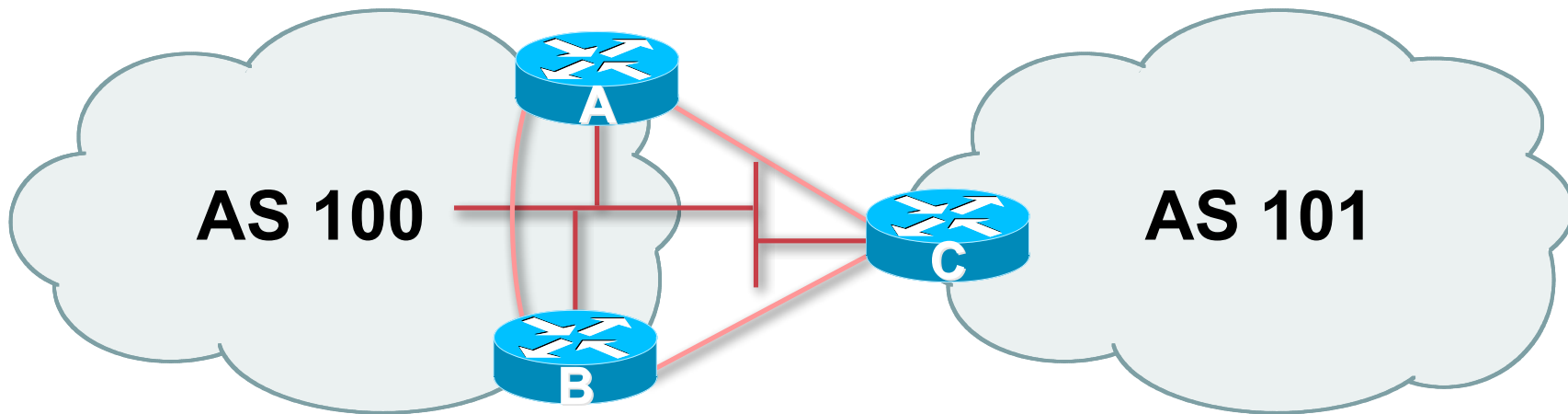


# BGP/IGP model used in ISP networks

- **Model representation**



# External BGP Peering (eBGP)



- Between BGP speakers in different AS
- Should be directly connected
- **Never** run an IGP between eBGP peers

# Configuring External BGP

## Router A in AS100

```
interface ethernet 5/0
  ip address 102.102.10.2 255.255.255.240
!
router bgp 100
  network 100.100.8.0 mask 255.255.252.0
  neighbor 102.102.10.1 remote-as 101
  neighbor 102.102.10.1 prefix-list RouterC in
  neighbor 102.102.10.1 prefix-list RouterC out
!
```

ip address on  
ethernet interface

Local ASN

Remote ASN

ip address of Router C  
ethernet interface

Inbound and  
outbound filters

# Configuring External BGP

## Router C in AS101

```
interface ethernet 1/0/0
  ip address 102.102.10.1 255.255.255.240
!
router bgp 101
  network 100.100.8.0 mask 255.255.252.0
  neighbor 102.102.10.2 remote-as 100
  neighbor 102.102.10.2 prefix-list RouterA in
  neighbor 102.102.10.2 prefix-list RouterA out
!
```

ip address on  
ethernet interface

Local ASN

Remote ASN

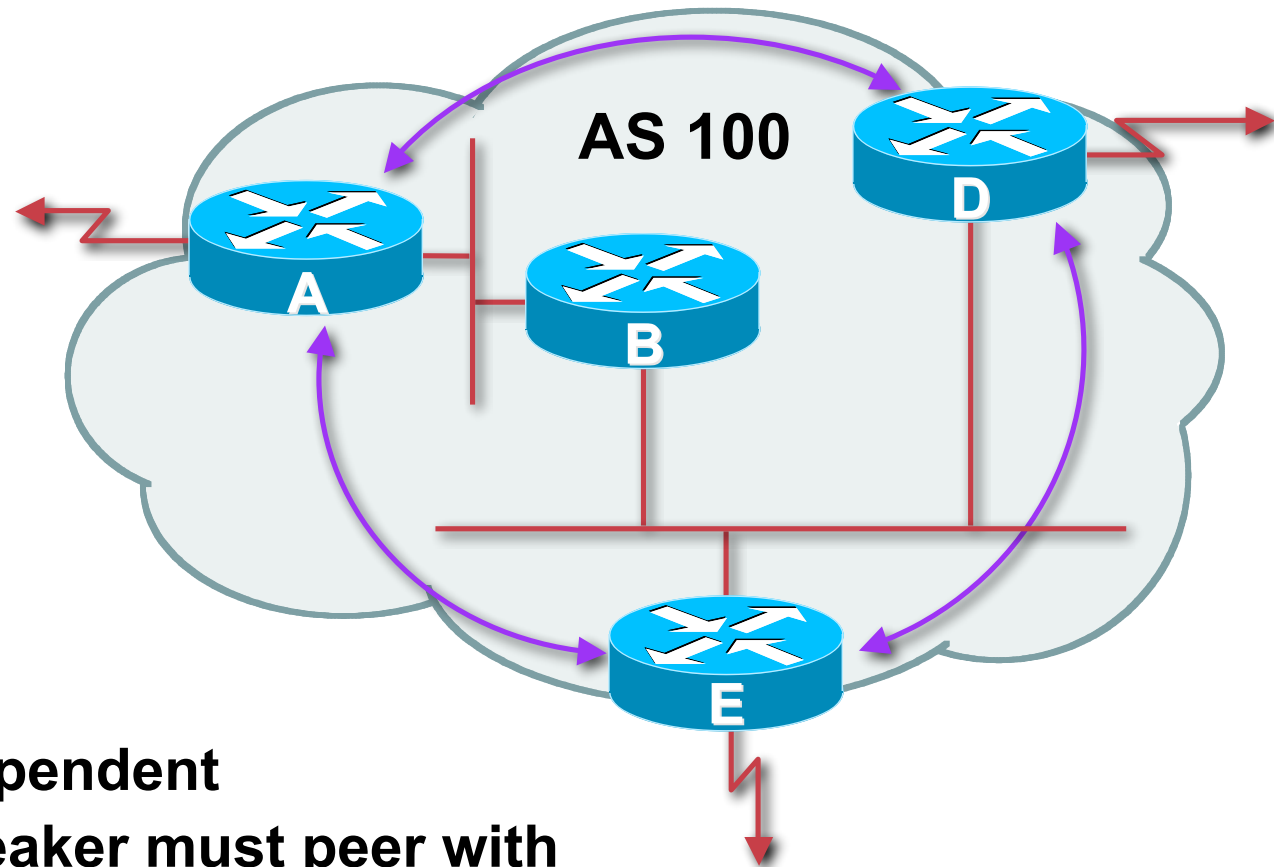
ip address of Router A  
ethernet interface

Inbound and  
outbound filters

# Internal BGP (iBGP)

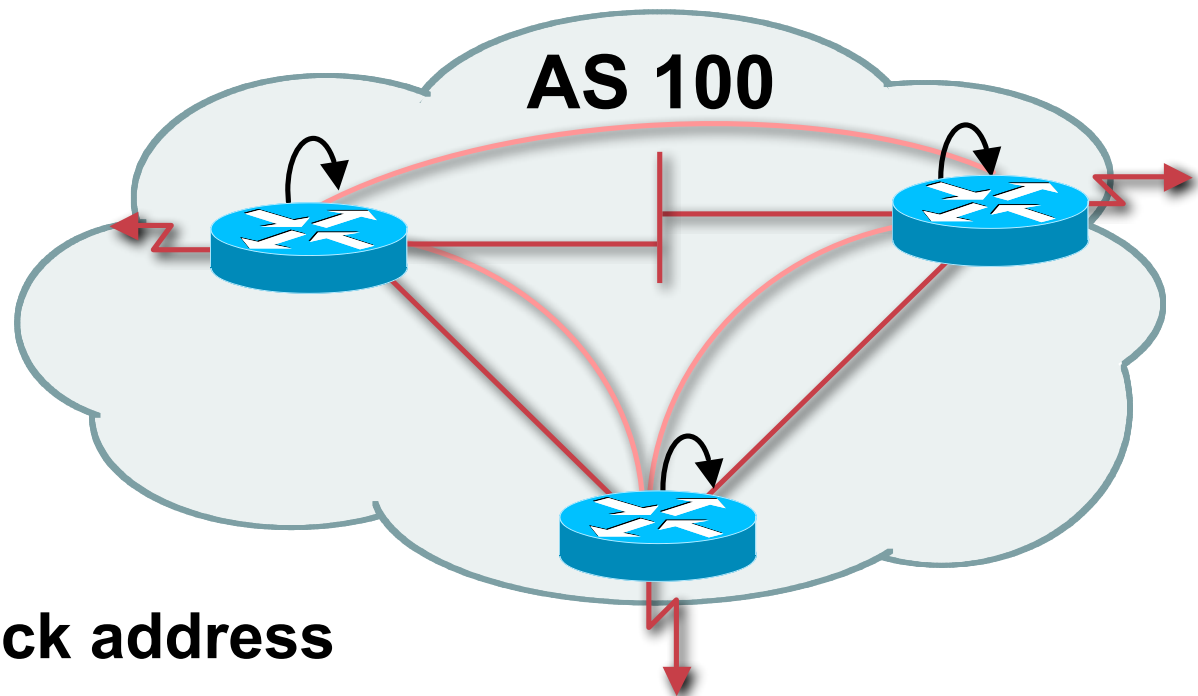
- **BGP peer within the same AS**
- **Not required to be directly connected**  
IGP takes care of inter-BGP speaker connectivity
- **iBGP speakers need to be fully meshed**  
they originate connected networks  
they do not pass on prefixes learned from other iBGP speakers

# Internal BGP Peering (iBGP)



- **Topology independent**
- **Each iBGP speaker must peer with every other iBGP speaker in the AS**

# Peering using Loop-back Address



- **Peer with loop-back address**  
Loop-back interface does not go down – ever!
- **iBGP session is not dependent on state of a single interface**
- **iBGP session is not dependent on physical topology**

# Configuring Internal BGP

## Router A in AS100

```
interface loopback 0
  ip address 105.3.7.1 255.255.255.255
!
router bgp 100
  network 100.100.1.0
  neighbor 105.3.7.2 remote-as 100
  neighbor 105.3.7.2 update-source loopback0
  neighbor 105.3.7.3 remote-as 100
  neighbor 105.3.7.3 update-source loopback0
!
```

ip address on  
loopback interface

Local ASN

Local ASN

ip address of Router B  
loopback interface



# Configuring Internal BGP

## Router B in AS100

```
interface loopback 0
  ip address 105.3.7.2 255.255.255.255
!
router bgp 100
  network 100.100.1.0
  neighbor 105.3.7.1 remote-as 100
  neighbor 105.3.7.1 update-source loopback0
  neighbor 105.3.7.3 remote-as 100
  neighbor 105.3.7.3 update-source loopback0
!
```

ip address on  
loopback interface

Local ASN

Local ASN

ip address of Router A  
loopback interface

# Inserting prefixes into BGP

- **Two ways to insert prefixes into BGP**
  - redistribute static**
  - network command**

# Inserting prefixes into BGP – redistribute static

- **Configuration Example:**

```
router bgp 100
```

```
    redistribute static
```

```
ip route 102.10.32.0 255.255.254.0 serial0
```

- **Static route must exist before redistribute command will work**
- **Forces origin to be “incomplete”**
- **Care required!**

# Inserting prefixes into BGP – redistribute static

- Care required with **redistribute**!

**redistribute <routing-protocol> means everything in the <routing-protocol> will be transferred into the current routing protocol**

**Will not scale if uncontrolled**

**Best avoided if at all possible**

**redistribute normally used with “route-maps” and under tight administrative control**

# Inserting prefixes into BGP – network command

- **Configuration Example**

```
router bgp 100
```

```
network 102.10.32.0 mask 255.255.254.0
```

```
ip route 102.10.32.0 255.255.254.0 serial0
```

- **A matching route must exist in the routing table before the network is announced**
- **Forces origin to be “IGP”**

# Configuring Aggregation

- **Three ways to configure route aggregation**
  - redistribute static**
  - aggregate-address**
  - network command**

# Configuring Aggregation

- **Configuration Example:**

```
router bgp 100
```

```
  redistribute static
```

```
ip route 102.10.0.0 255.255.0.0 null0 250
```

- **static route to “null0” is called a pull up route**

**packets only sent here if there is no more specific match in the routing table**

**distance of 250 ensures this is last resort static**

**care required – see previously!**

# Configuring Aggregation – Network Command

- **Configuration Example**

```
router bgp 100
```

```
network 102.10.0.0 mask 255.255.0.0
```

```
ip route 102.10.0.0 255.255.0.0 null0 250
```

- **A matching route must exist in the routing table before the network is announced**
- **Easiest and best way of generating an aggregate**



# Configuring Aggregation – aggregate-address command

- **Configuration Example**

```
router bgp 100
  network 102.10.32.0 mask 255.255.252.0
  aggregate-address 102.10.0.0 255.255.0.0 [ summary-only ]
```

- **Requires more specific prefix in BGP table before aggregate is announced**
- **{summary-only} keyword**  
optional keyword which ensures that only the summary is announced if a more specific prefix exists in the routing table

# Historical Defaults – Auto Summarisation

- **Disable historical default 1**
- **Automatically summarises subprefixes to the classful network when redistributing to BGP from another routing protocol**

Example:

61.10.8.0/22 → 61.0.0.0/8

- **Must** be turned off for any Internet connected site using BGP

```
router bgp 100  
no auto-summary
```

# Historical Defaults – Synchronisation

- **Disable historical default 2**
- **In Cisco IOS, BGP does not advertise a route before all routers in the AS have learned it via an IGP**
- **Disable synchronisation if:**
  - AS doesn't pass traffic from one AS to another, or**
  - All transit routers in AS run BGP, or**
  - iBGP is used across backbone**

```
router bgp 100  
no synchronization
```

# Summary

## BGP neighbour status

```
Router1>sh ip bgp sum
```

```
BGP router identifier 100.1.15.224, local AS number 10
```

```
BGP table version is 27, main routing table version 27
```

```
14 network entries using 1582 bytes of memory
```

```
14 path entries using 672 bytes of memory
```

```
3/2 BGP path/bestpath attribute entries using 324 bytes of memory
```

```
0 BGP route-map cache entries using 0 bytes of memory
```

```
0 BGP filter-list cache entries using 0 bytes of memory
```

```
BGP using 2578 total bytes of memory
```

```
BGP activity 17/3 prefixes, 22/8 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
100.1.31.224	4	10	195	193	27	0	0	03:09:48	1
100.1.63.224	4	10	71	70	27	0	0	01:05:31	1
100.2.15.224	4	10	46	47	27	0	0	00:17:00	1
...									

**BGP Version**

**Updates sent  
and received**

**Updates waiting**

# Summary

- **BGP4 – path vector protocol**
- **iBGP versus eBGP**
- **stable iBGP – peer with loopbacks**
- **announcing prefixes & aggregates**
- **no synchronization & no auto-summary**



# Introduction to BGP

**ISP/IXP Workshops**