

# University of Rome Tor Vergata ICT and Internet Engineering

## Network and System Defense

Alessandro Pellegrini, Angelo Tulumello

A.A. 2023/2024

# Lecture 8: BGP basics

Angelo Tulumello

# Routing Protocols at a Glance

#### Routing protocols

- ☐ Goal: setup IP routing tables according to configurable administrative rules
- Routing protocols populate the Routing Tables
  - Possibly more routes for the same destination prefix
- An internal algorithm choses **best routes per prefix** and inject this choice in the Forwarding Table used during packet forwarding operation

modifica automaticamente le routing tables

#### Interior and Exterior Routing Protocols

- ☐ IP routing protocols fall into one of two major categories: *interior gateway*protocols (IGP) or exterior gateway protocols (EGP)
  - ☐ **IGP**: A routing protocol that was designed and intended for use inside a single autonomous system (AS)
  - **EGP**: A routing protocol that was designed and intended for use between different autonomous systems
- An AS is a network under the administrative control of single organization
- Each AS can be assigned a number called AS number (ASN)
- Each AS has one (or more) public IP address range(s)

#### Interior Gateway Protocols

- Organizations have several options when choosing an IGP for their enterprise network (but most companies today use *OSPF*)
- Two main branches of routing protocol algorithms exist for IGP routing protocols:
  - Distance Vector (e.g. RIP)
  - ☐ Link-state (e.g. OSPF, IS-IS)
- In a distance vector protocol the routers only exchange the routing table with their neighbors (i.e. routers on the same links)
- Link state protocols require the synchronization and the knowledge of the entire network topology on every router in the AS

## Interior Gateway Protocols

Organizations have several option network (but most com-

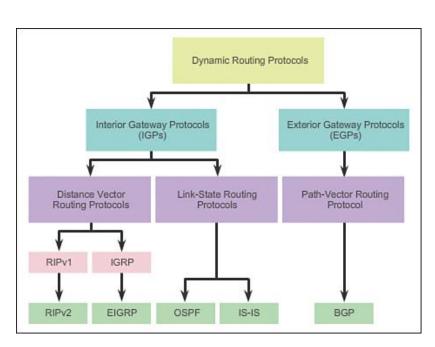
rprise

- Although IGPs are a fundamental building block of IP networks and their vulnerabilities should be carefully addressed, IGPs security is outside Two main bra prot
- the scope of this course outing table with the
- Link zation the knowledge of the entire netwo

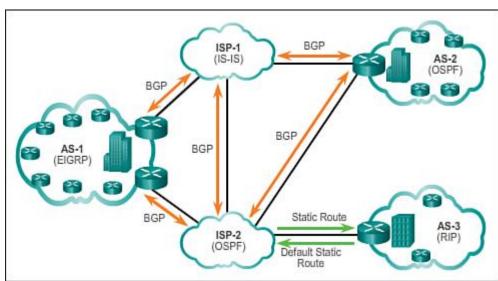
#### Exterior Gateway Protocols

- Typically, EGPs are used to exchange routing information between ISPs, or in some cases between a customer's AS and the provider's network.
- **Border Gateway Protocol (BGP)**, version 4 (BGP4) is the most common EGP and is considered the Internet standard.
- An internetwork is a confederation of smaller, independent networks, called autonomous systems, owned and operated by a different organization: a company, university, government agency, or some other group.
- Each AS typically represents an independent organization, and applies its own unique routing and security policies.
- EGPs facilitate the sharing of routing information between autonomous systems.

#### The Big Picture



#### questo potrebbe essere un AS di una grande azienda



si possono avere molte connessioni all'interno della rete. Lo scambio tra AS è fatto tramite BGP. all'interno dell'AS si hanno Interior Gateway Protocol.

parametro per stabilire la priorità di una certa rete

- Multiple routing protocols in a same domain
- Routing protocols might learn routes to the same subnets
- When a node must choose between routes learned using different routing protocols, the so-called *administrative distance* is compared
- Administrative distance is a number that denotes how believable an entire routing protocol is on a single router.
- The lower the number, the better, or more believable, the routing protocol.
- The administrative distance values are configured on a single router and are not exchanged with other routers

#### Administrative distance on CISCO IOS

Route Type	Administrative Distance
Connected	0
Static	1 ip route add, può essere configurato staticamente dall'amministra
BGP (external routes)	20
EIGRP (internal routes)	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
EIGRP (external routes)	170
BGP (internal routes)	200
Unusable	255

## Border Gateway Protocol (BGP)

Warning: BGP would require a lot of time to be exhaustively covered. This is going to be more of an overview required to understand the the next two topics: BGP/MPLS VPNs and BGP security

#### **BGP Basics**

	BGP is a distance vector routing protocol (more specifically a path vector routing protocol)	
	It relies on downstream neighbors to pass along routes from their routing table.	
	The node makes its route calculations based on the advertised routes and passes the result	
	to upstream neighbors.	
BGP uses a list of AS numbers through which a packet must pass to reach a destination distance metric to minimize		
	about attributes later)	
	□ A <mark>network address</mark>	
	The path vector carrying the list of ASes traversed to reach the destination	
	The next hop to reach the destination	
	Each AS has an identifying number, assigned by an Internet registry or a service provider, between 1	
	and 65535.	
	☐ Private AS numbers: Between 64512 through 65535	
	AS numbers, as IPv4 addresses have begun to <b>run out</b> → RFC 6793 introduced 4-bytes AS numb	
	☐ Private AS numbers: Between 4200000000 to 4294967294 estende l'address space dell'AS	

#### BGP Basics

changed.

When two routers establish a TCP-enabled BGP connection between each other, they are called **neighbors** or **peers**. Each router running BGP is called a **BGP speaker**. When two neighbors first establish a BGP connection, they exchange their entire BGP routing tables. After that, they exchange incremental, partial updates with only the information that has 

**BGP runs on TCP** (so differently from OSPF, IP/TCP connectivity is required)

- Peers exchange keepalive messages to ensure the connection is maintained.
- If three keepalive intervals pass the peer declares its neighbor down and all the routes advertised by the peer are withdrawn withdrawn=scompaiono; è un punto fondamentale per gli attacchi, in

particolare per i DoS attack.

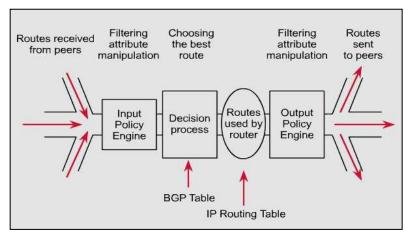
#### **BGP Basics**

- Before establishing a BGP peer connection the two neighbors must perform the standard TCP three-way handshake and open a TCP connection to port 179.
- After the TCP session is established, BGP peers exchanges several messages to open and confirm connection parameters and to send BGP routing information.
- All BGP messages are unicast to the one neighbor over the TCP connection.
- There are four BGP message types:
  - ☐ Type 1: OPEN
  - ☐ Type 2: KEEPALIVE
  - **☐** Type 3: UPDATE
  - **☐** Type 4: NOTIFICATION

#### **BGP** Routing process

- BGP is so flexible because it is a fairly simple protocol.
- Routes are exchanged between BGP peers via UPDATE messages.
- BGP routers receive the UPDATE messages, run some policies or filters over the updates, and then pass on the routes to other BGP peers.
- Different implementations of BGP keep track of all BGP updates in a BGP table separate from the IP routing table.

#### **BGP** speaker



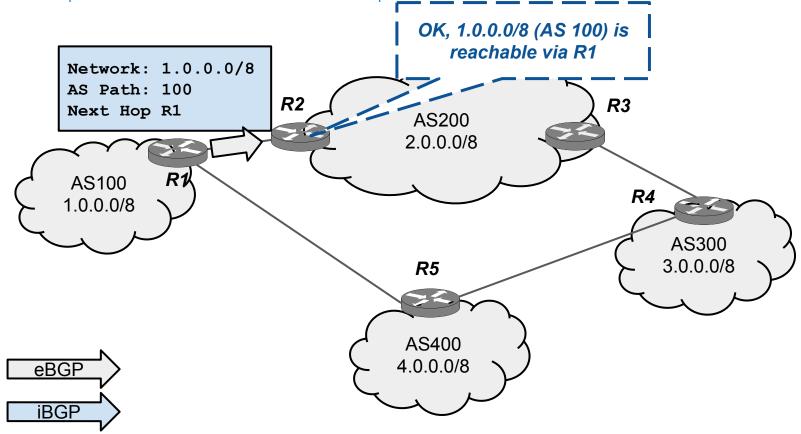
BGP permette di filtrare tramite input policy engine alcuni pacchetti

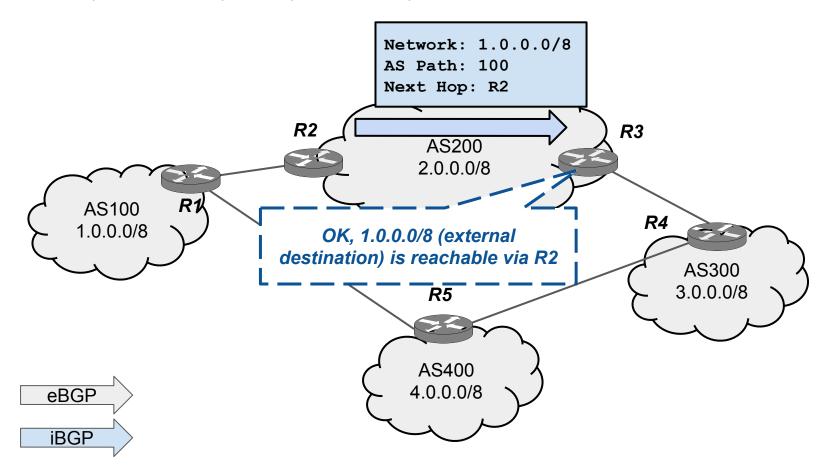
#### But BGP is not only used among different ASes

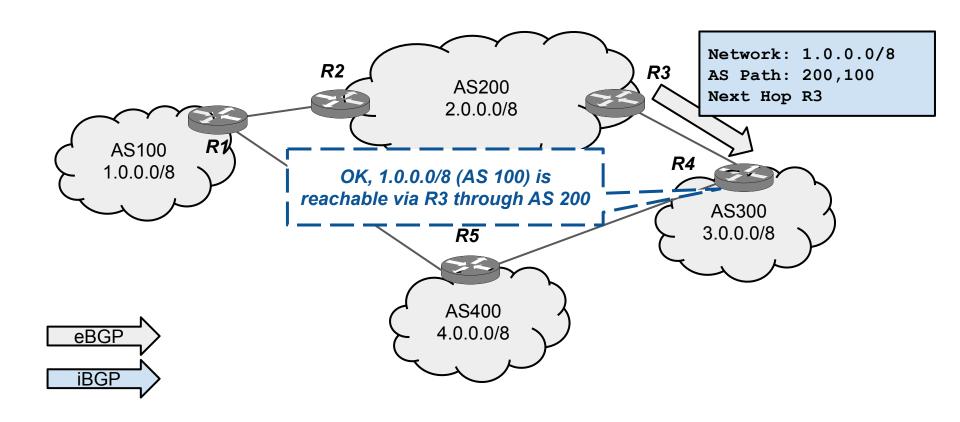
- BGP is also used to distribute routing information for external destination among routers in the same AS
- When BGP is running inside an AS, it is referred to as Internal BGP (IBGP).
  - If a BGP router role is to route IBGP traffic, it is called a transit router.
- When BGP runs between autonomous systems, it is called External BGP (EBGP).
  - Routers that sit on the boundary of an AS and use EBGP to exchange information with the ISP are called border routers.
- → "With very few exceptions, interior BGP (IBGP) BGP between peers in the same AS is used only in multihomed scenarios." Doyle

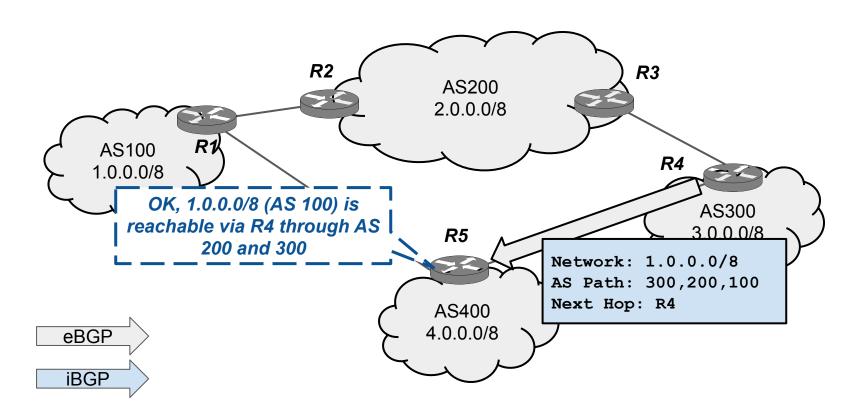
4 AS, AS 200 ha due end router differenti connessi a diversi AS.

Con EBGP AS 100 manda un "announcement" dicendo che per raggiungere la destinazione 1.0.0.0/8 bisogna passare come next-hop per R1. AS200 salva questa informazione nella sua tabella e condivide que<u>sta informazione con R3.</u>







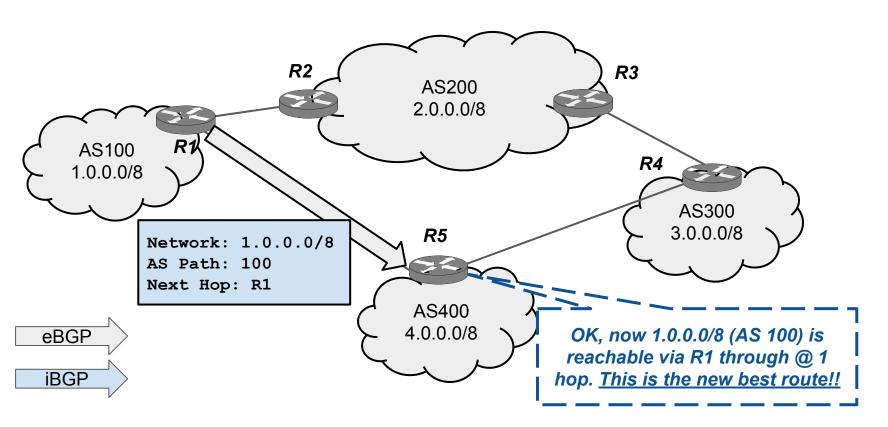


#### How BGP chooses the best path

- Network topology has nothing to do with geographical topology
  - of course, it is more probable that a AS at 1 hop is most likely geographically closer AS
  - □ but ASes can be distributed over the entire world...
  - and peering also depends on commercial agreement
- Let's imagine that the BGP topology changes
  - ☐ For example, R1 and R5 become neighbours
- ☐ In this case R5 would receive an advertisement for 1.0.0.0/8 with AS\_PATH=100 and next\_hop R1
- ☐ The current routing entry to 1.0.0.0/8 is next\_hop R4, AS\_PATH=100,200,300
- ☐ The new entry is preferred (because it has a shorter AS\_PATH attribute) and installed in the global IP routing table

R1 ad un certo punto invia un announcement a R5 per dirgli che attraverso lui può raggiungere 1.0.0.0/8 attraverso solo la AS100, R5 quindi dato che il path è più corto sceglie questa informazione e la mantiene insieme a quella vecchia. Quest'ultima rotta sarà preferita. La scelta viene fatta in base al path più corto.

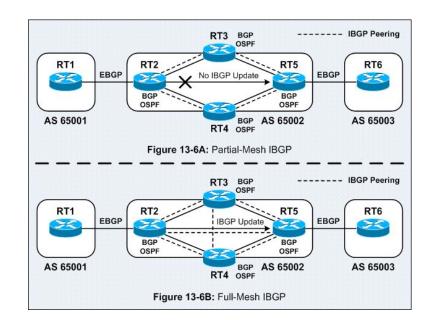
Questo è importante anche per fault tollerance, se la rotta verso R1 si rompe raggiunge l'indirizzo tramite R4.



#### BGP loop avoidance

come BGP evita loops

- AS number is added to the AS path only when a BGP UPDATE goes out from the AS
  - Possible loops can be easily detected
  - AS path check can not be used to avoid loops within a AS
- BGP Split Horizon Rule used to avoid loop inside the AS
  - A route learned via IBGP is never propagated to other IBGP peers.
  - A route is not advertised back to the EBGP peer from which the route was received
- Consequence: if all internal routers need the full routing table, there should be a full-mesh of IBGP sessions
  - unless we configure a so-called **BGP Route**Reflector



#### **BGP Attributes**

- Part of a BGP Update
  - Describes the characteristics of prefix
  - **Each route has its own set of defined attributes,** which can include path information, route preference, next-hop, and aggregation information.
  - The destination network, the AS path and the next hop are mandatory BGP attributes
- Administrators use these attributes to enforce routing policy.
- Based on attribute values, you can configure BGP to filter routing information, prefer certain paths, or otherwise customize its behavior.
- Every UPDATE message has a variable-length sequence of path attributes in the form <attribute type, attribute length, attribute value>.
- Since you will use path attributes extensively when configuring routing policy, you should note that not all vendor implementations of BGP recognize the same attributes. In fact, path attributes come in four different types: (1) **Well-known mandatory**, (2) **Well-known discretionary**; (3) **Optional transitive**; (4) **Optional non-transitive**

#### BGP Attributes and Path Selection

- □ BGP selects only one path as the best path.
- When the path is selected, BGP puts the selected path in its (BGP) routing table and propagates the path to its neighbors.
- ☐ The AS Path attribute is not the only one used in the selection
- BGP uses the following criteria, in the order presented(\*), to select a path for a destination:
  - "We Love Oranges AS Oranges Mean Pure Refreshment"
  - → Weight (Highest)
  - □ L LOCAL PREF (Highest)
  - ☐ Originate (local)
  - □ AS AS PATH (shortest)
  - ORIGIN Code (IGP > EGP > Incomplete)
  - ☐ M MED (lowest)
  - P Paths (External > Internal)
  - □ R RID (lowest)
- ☐ Each attribute has its own semantic and usage

(\*) actually depends on the specific router BGP implementation

### BGP Attributes and Path Selection

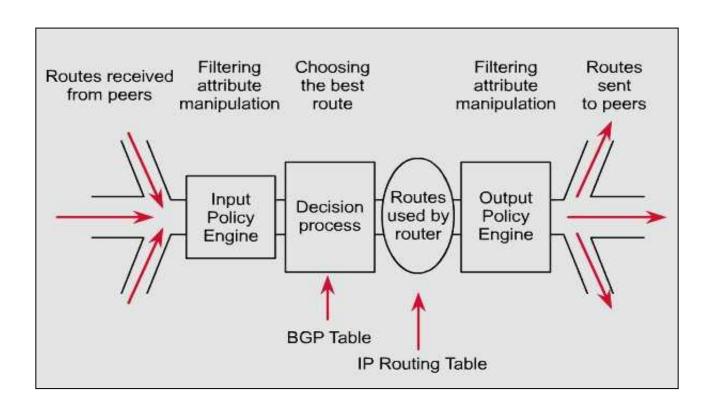
gates

mation:

- A comprehensive study of the BGP attributes is outside the scope of this aı > Internal) (lowest)
  - Each attribute has its own semantic and usage

(\*) actually depends on the specific router BGP implementation

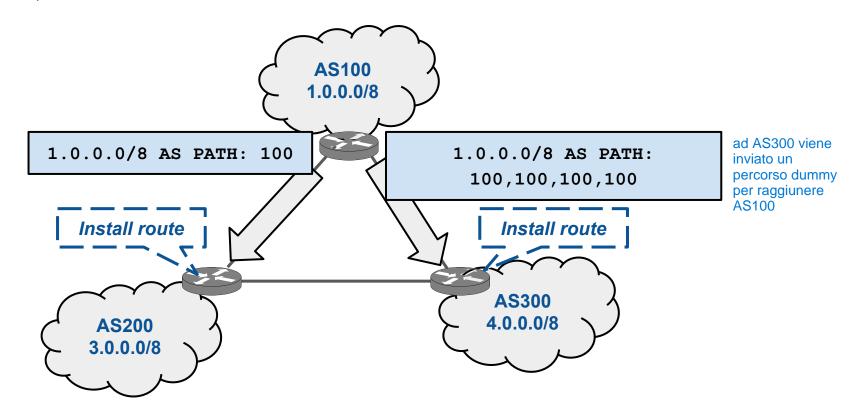
#### The BGP processing pipeline

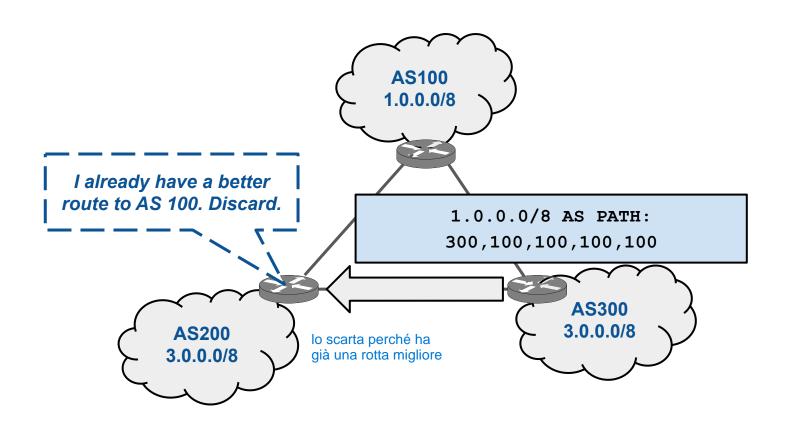


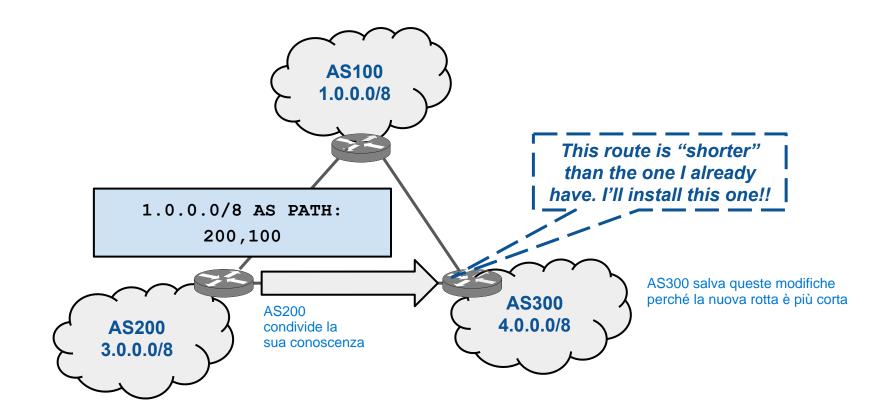
#### **BGP Attributes Manipulation**

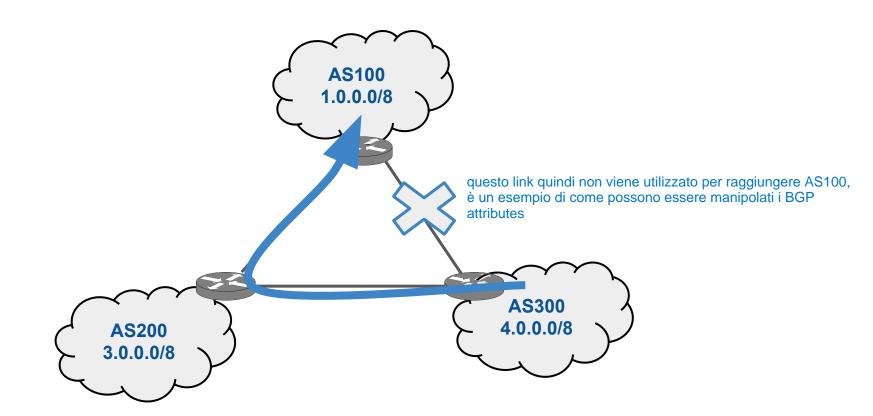
- □ BGP attributes can be manipulated (modified, added, removed...)
  - in incoming BGP updates, before the actual route selection
  - ☐ in outgoing BGP updates
- For example, a well known mechanism to influence the path selection is the so-called *AS-PATH prepend mechanism* 
  - path information are changed by including dummy AS path numbers that would increase the path length and influence the traffic trajectory one way or the other
- BGP attributes are a fundamental BGP tool
- With BGP attributes routes can be chosen in order not only to minimize the number of AS traversed to reach the destination

BGP update sono unicast verso solo il nodo a cui sono inviati









#### **COMMUNITY** attribute

- ☐ The BGP community attribute is an *optional transitive attribute* 
  - we'll see that it can be used to control the BGP/MPLS VPNs topologies..
- It represents a *numerical value* that can be assigned to a specific prefix and advertised to other neighbors
- When the neighbour receives the prefix it will examine the community value and take proper action whether it is filtering or modifying other attributes
- A BGP community is a 32-bit number that can be included with a route. A BGP community can be displayed as a full 32-bit number (0-4,294,967,295) or as two 16-bit numbers (0-65535):(0-65535) commonly referred to as new-format.
- ☐ A number of well known "standard" community values exists (see next slide)

### Well known community values

- **□** NO EXPORT (0xFFFFF01)
  - A route carrying this community value should not be advertised to peers outside a confederation (or the AS if it is the only AS in the confederation).
- ☑ NO ADVERTISE (0xFFFFFF02)
  - A route carrying this community value, when received, should not be advertised to any BGP peer
- ☐ Internet (0xFFFFFF03)
  - A route carrying this community value, when received, should be advertised to all other routers.
- ☐ Local-as (0xFFFFF04)
  - A route carrying this community value, when received, should be advertised to peers within the AS, but not advertised to peers in an external system
- ☐ Custom community values can be also used
  - ☐ The private BGP community pattern could vary from organization to organization, do not need to be registered, and could signify geographic locations for one AS while signifying a method of route advertisement in another AS

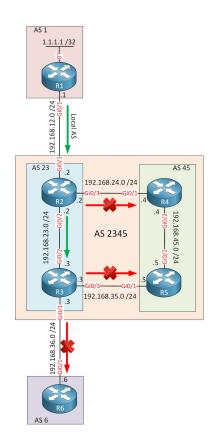


Fig. example of the Local-AS community

**BGP** configuration

(Free Range Routing)

#### AS number

- To begin configuring a BGP process, issue the following familiar command: in generale basta 1 BGP process, un router può essere parte di un solo AS (ci sono scenari in cui un router può far parte di diversi AS, non ci interessa)

  Router (config) # router bgp AS-number
- BGP configuration commands appear on the surface to mirror the syntax of familiar IGP (for example, RIP, OSPF) commands.
- Although the syntax is similar, the function of these commands is significantly different.

## Manually advertise a network

Router(config-router) #network network-number [mask network-mask]

- The **network** command is used with IGPs, such as OSPF, to determine the interfaces on which to send and receive updates, as well as which directly connected networks to advertise.
- When configuring BGP, the network command does not affect what interfaces BGP runs on.
- In BGP, the network command tells the BGP process what locally learned networks to advertise.
- The networks can be connected routes, static routes, or routes learned via a dynamic routing protocol, such as OSPF

## Manually advertise a network

- These networks must also exist in the local router's routing table (show ip route), or they will not be sent out in updates.
- → You can use the mask keyword with the network command to specify individual subnets.
- Routes learned by the BGP process are propagated by default, but are often filtered by a routing policy.

# BGP peering

Router(config-router) #neighbor ip-address remote-as AS-number

- In order for a BGP router to **establish a neighbor relationship with another BGP router**, you must issue the above configuration command.
- This command serves to identify a peer router with which the local router will establish a session.
- The **AS-number** argument determines whether the neighbor router is an EBGP or an IBGP neighbor.

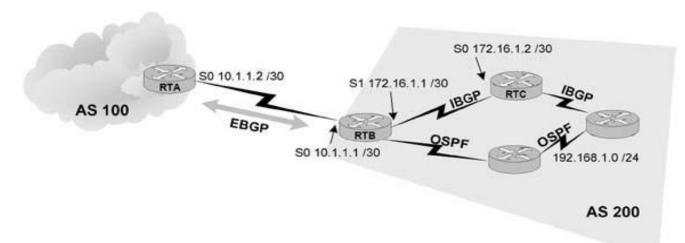
   equindi un modo per discriminare tra EBGP e IBGP

## BGP and Loopback interfaces

assegnare indirizzi, se si ricevono pacchetti BGP peering session may use loopback interfaces per questa interfaccia si può rispondere da qualsiasi porta Loopback interface ensures that the neighbor stays up and is not affected by malfunctioning hardware of a specific interface Useful when there are multiple paths between the BGP peers iBGP peering are among internal routers generally connected by different IGP (OSPF) routes Useful also in case of multilink, for load balancing **Best Common Practice IBGP** on loopback EBGP on physical interface (unless multilink) 

dummy interfaccia nella quale si possono

# eBGP peering



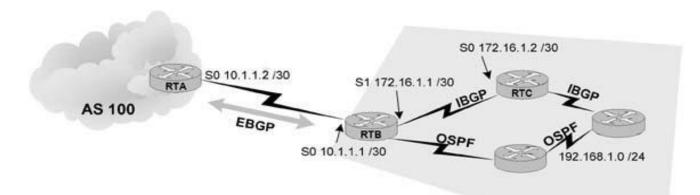
RTA(config) #router bgp 100

RTA(config-router) #neighbor 10.1.1.1 remote-as 200

RTB(config) #router bgp 200

RTB(config-router) #neighbor 10.1.1.2 remote-as 100

# iBGP peering



RTB(config) #router bgp 200

AS 200

RTB(config-router) #neighbor 172.16.1.2 remote-as 200

RTB(config-router) #neighbor 172.16.1.2 update-source loopback 0

RTC(config) #router bgp 200

RTC(config-router) #neighbor 172.16.1.1 remote-as 200

RTC(config-router) #neighbor 172.16.1.1 update-source loopback 0

## Loopback source

- ☐ The *update-source loopback 0* command is used to instruct the router to use any operational interface for TCP connections (as long as Lo0 is up and configured with an IP address).
- □ Without the *update-source loopback 0* command, BGP routers can use only the closest IP interface to the peer.
- ☐ The ability to use any operational interface provides BGP with robustness in the event the link to the closet interface fails.
  - Since eBGP sessions are typically point-to-point, there is no need to use this command with eBGP.

#### Reset the BGP process

- Finally, whenever you are configuring BGP, you will notice that changes you make to an existing configuration may not appear immediately.
- To force BGP to clear its table and reset BGP sessions, use the clear ip bgp command. The easiest way to enter this command is as follows:

Router# clear ip bqp \* questo comando è il motivo che ha mandato down meta

- ☐ Use this command with CAUTION, better yet not at all, in a production network.
  - remember October 4, 2021 Meta Inc. blackout?!
    - https://engineering.fb.com/2021/10/05/networking-traffic/outage-details/
    - https://blog.cloudflare.com/october-2021-facebook-outage/

## Verifying BGP Configurations

RTA#show ip bgp

BGP table version is 3, local router ID is 10.2.2.2

Status codes: s suppressed, d damped, h history, \* valid, > best, i - internal

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path	ı	
* i1.0.0.0	192.168.1.6	0	100	0	200	400	е
*>i10.1.1.1/32	10.1.1.1	0	100	0	i		
*>i172.16.1.0/24	10.1.1.1	0	100	0	i		
* i192.168.1.32/27	192.168.1.6	0	100	0	200	i	

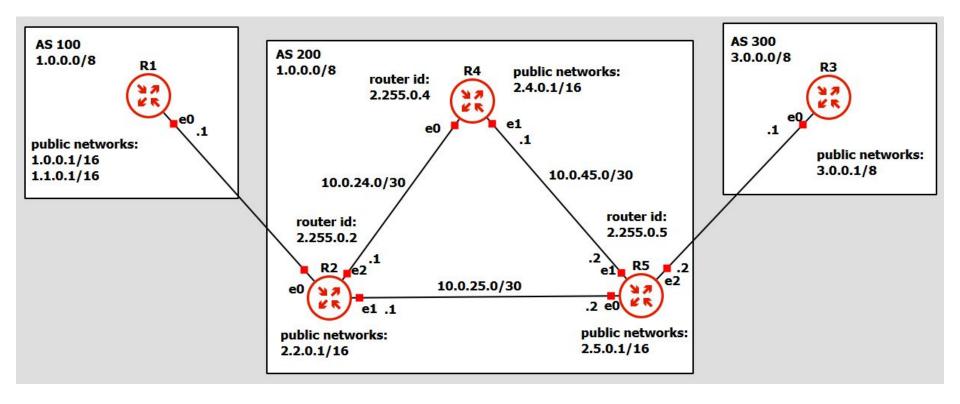
# Verifying BGP Configurations

```
RTA#show ip bgp neighbors
BGP neighbor is 172.24.1.18, remote AS 200, external link
 BGP version 4, remote router ID 172.16.1.1
 BGP state = Established, up for 00:03:25
 Last read 00:00:25, hold time is 180, keepalive interval is 60 seconds
 Neighbor capabilities:
   Route refresh: advertised and received
   Address family IPv4 Unicast: advertised and received
 Received 7 messages, 0 notifications, 0 in queue
  Sent 8 messages, 0 notifications, 0 in queue
 Route refresh request: received 0, sent 0
 Minimum time between advertisement runs is 30 seconds
```

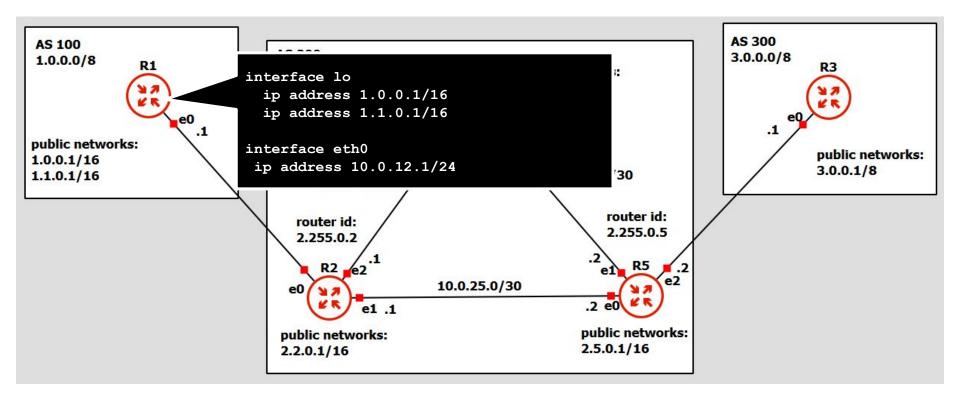
it would be interesting to show you how to manipulate BGP attributes to enforce some kind of routing policies. Unfortunately we don't have time for

# Laboratory: a simple BGP scenario

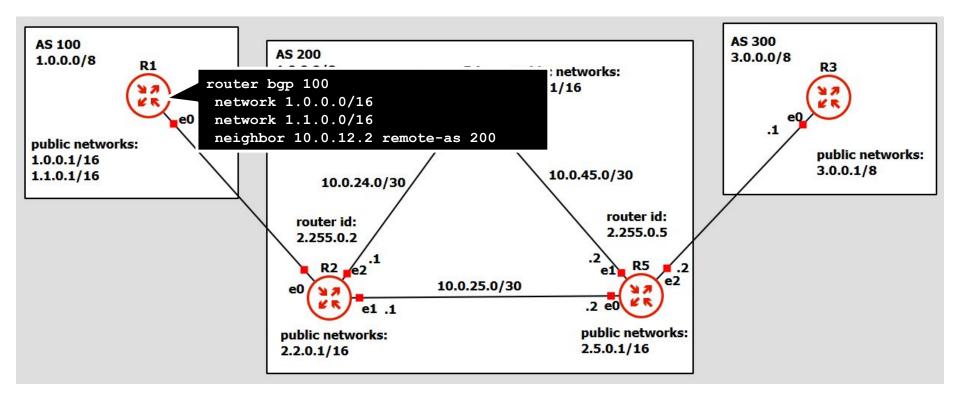
## Topology



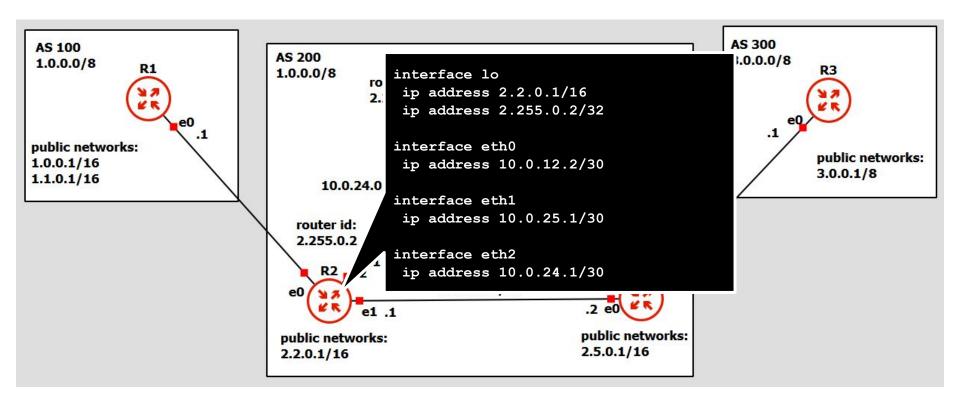
# IP configuration on R1



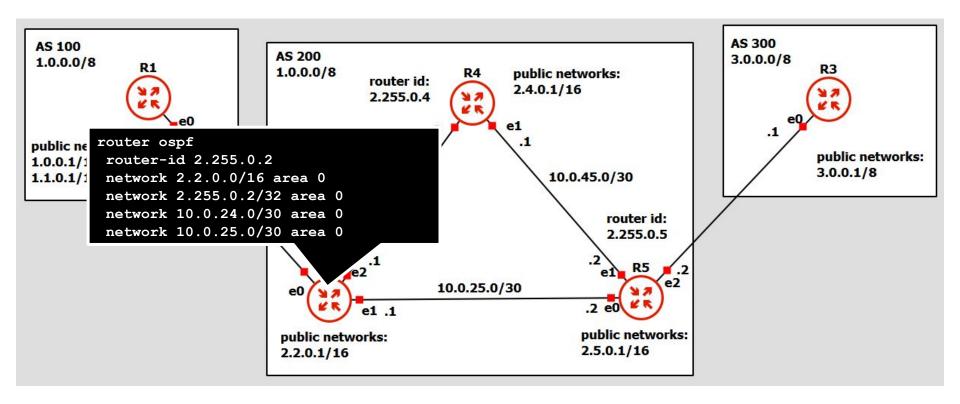
# BGP configuration on R1



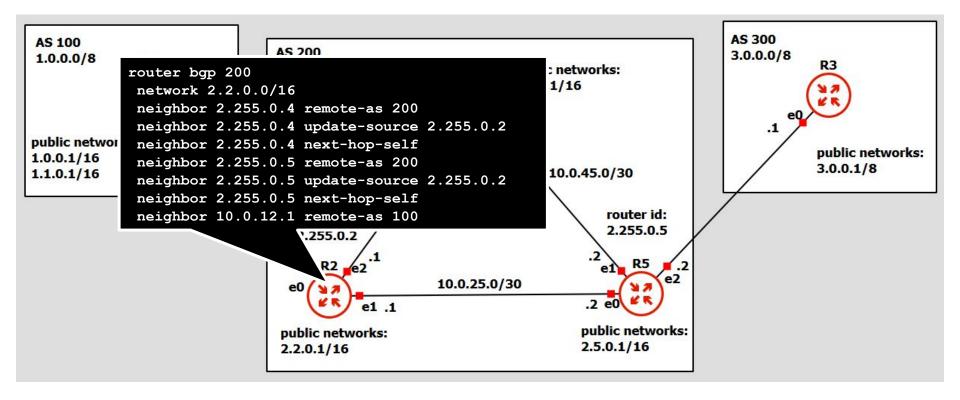
# IP configuration on R2



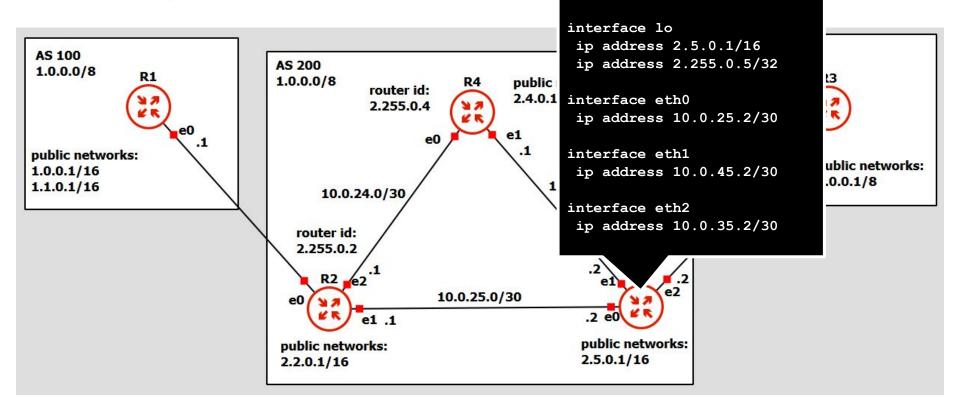
# OSPF configuration on R2



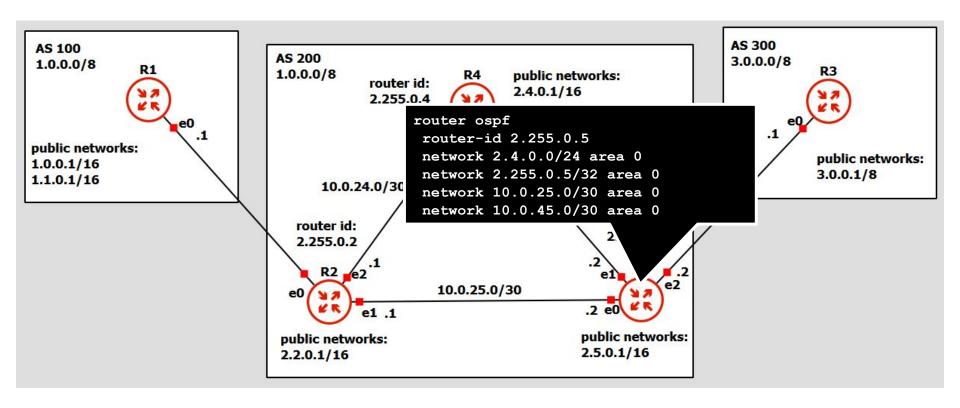
#### BGP configuration on R2



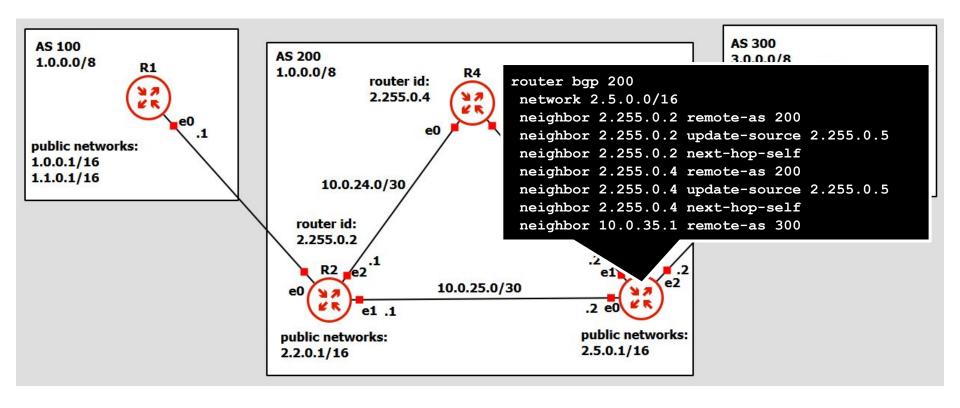
# IP configuration on R5



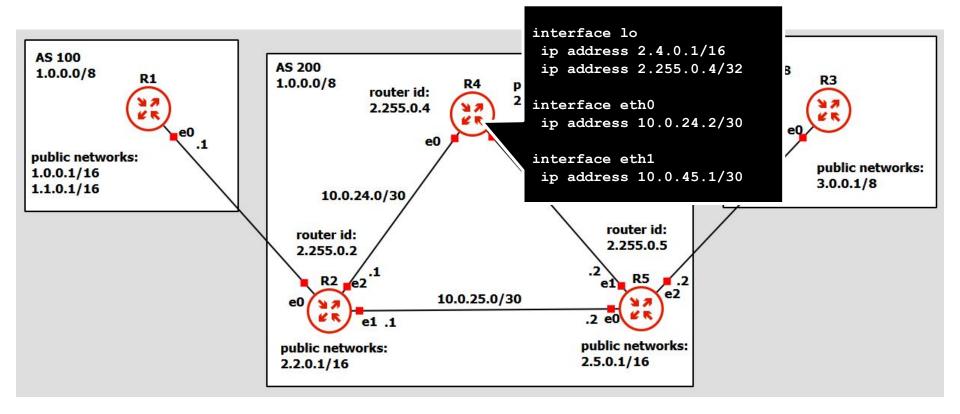
# OSPF configuration on R5



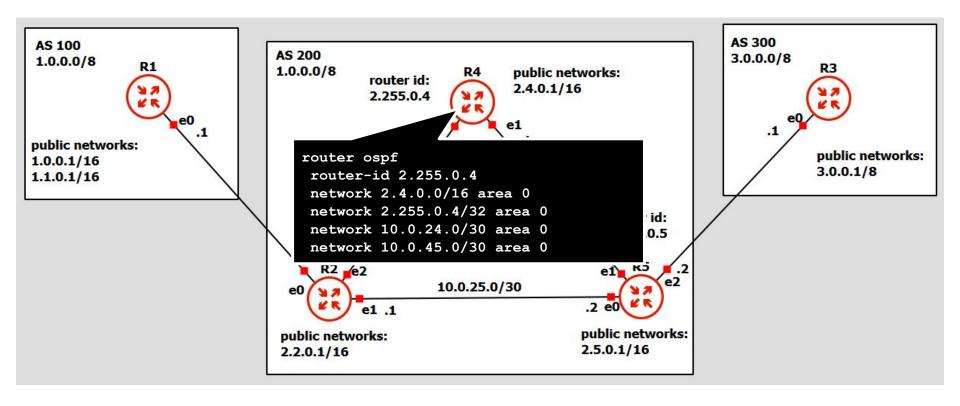
## BGP configuration on R5



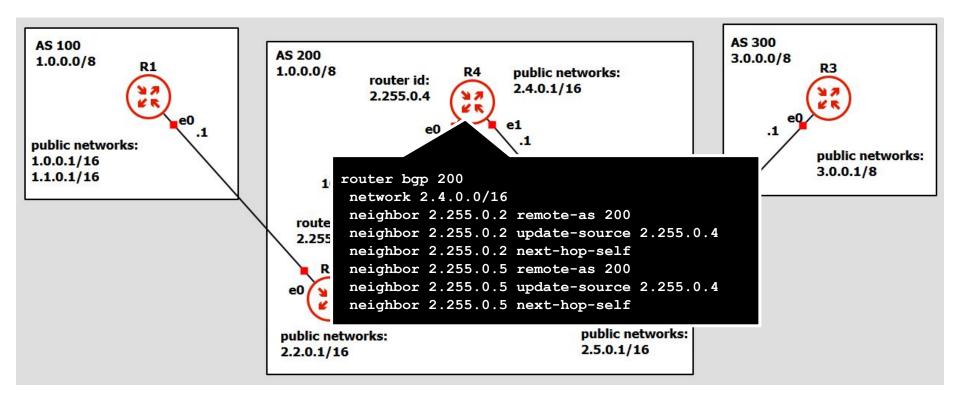
# IP configuration on R4



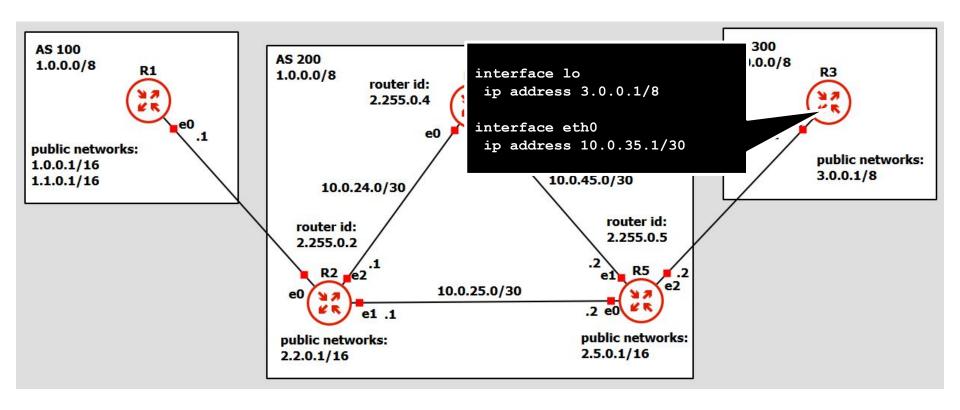
# OSPF configuration on R4



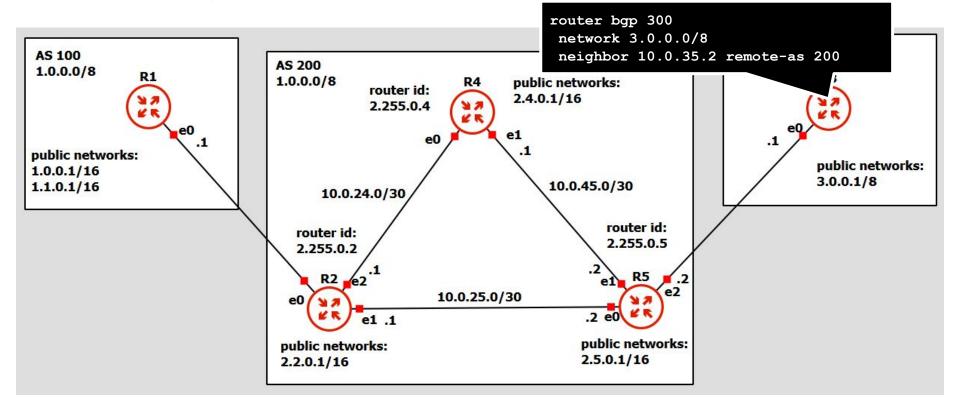
## BGP configuration on R4



# IP configuration on R3



# BGP configuration on R3



# Test 1: check BGP peers and table in R3

AS 100 1.0.0.0/8

1.0.0.1/16

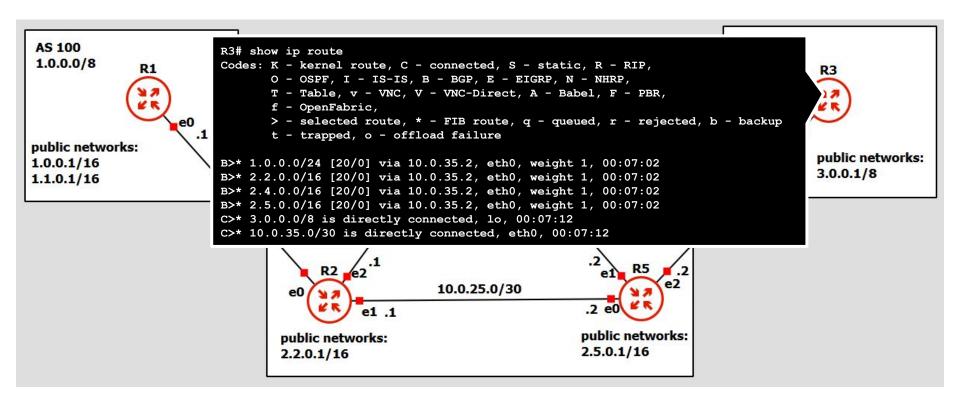
1.1.0.1/16

```
R3# show ip bgp summary
             IPv4 Unicast Summary (VRF default):
             BGP router identifier 3.0.0.1, local AS number 300 vrf-id 0
             BGP table version 5
              RIB entries 9, using 1728 bytes of memory
             Peers 1, using 13 KiB of memory
              Neighbor
                                                      MsqSent TblVer InQ OutQ Up/Down State/PfxRcd
                                            MsaRcvd
              PfxSnt Desc
              R5(10.0.35.2)
                                      200
                                                114
                                                          112
                                                                     5
                                                                               0 00:05:22
              5 N/A
             Total number of neighbors 1
public network
             R3# show ip bgp
             BGP table version is 5, local router ID is 3.0.0.1, vrf id 0
             Default local pref 100, local AS 300
              Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
                            i internal, r RIB-failure, S Stale, R Removed
              Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
              Origin codes: i - IGP, e - EGP, ? - incomplete
             RPKI validation codes: V valid, I invalid, N Not found
                  Network
                                  Next Hop
                                                      Metric LocPrf Weight Path
               *> 1.0.0.0/24
                                  10.0.35.2(R5)
                                                                         0 200 100 i
               *> 2.2.0.0/16
                                  10.0.35.2 (R5)
                                                                         0 200 i
               *> 2.4.0.0/16
                                  10.0.35.2 (R5)
                                                                         0 200 i
               *> 2.5.0.0/16
                                  10.0.35.2(R5)
                                                                         0 200 i
                                                           0
               *> 3.0.0.0/8
                                  0.0.0.0(R3)
                                                                     32768 i
                                                           0
             Displayed 5 routes and 5 total paths
```

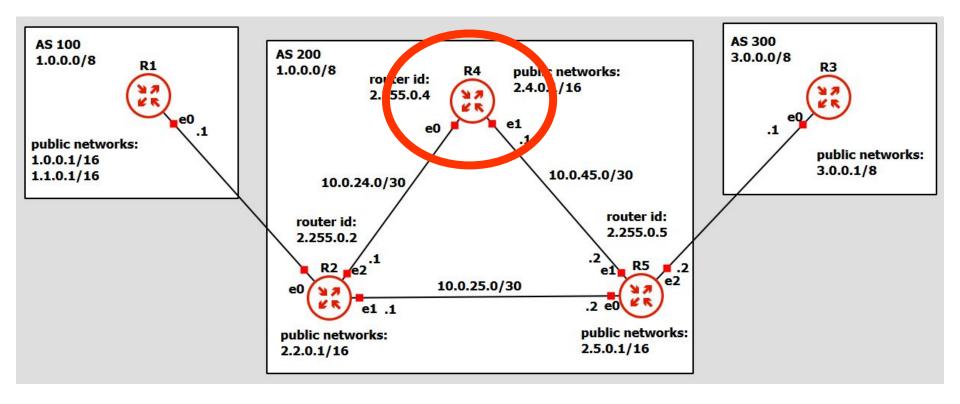
public networks:

3.0.0.1/8

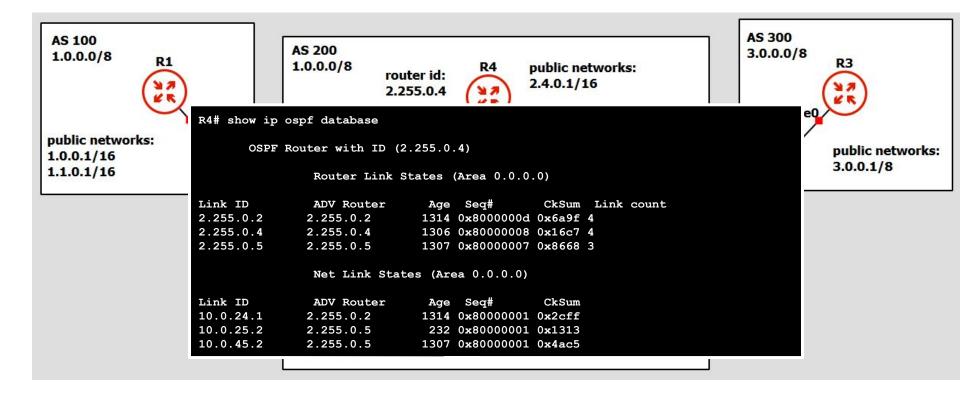
# Test case 1: check BGP peers and table in R3



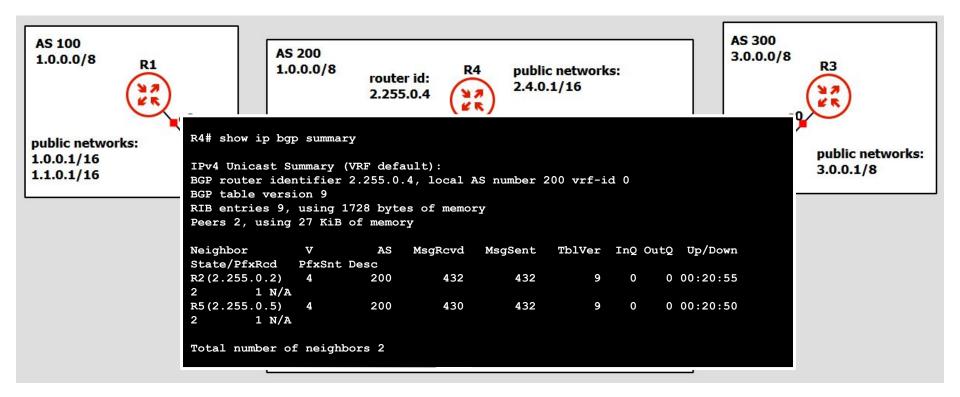
#### Test case 3: check R4



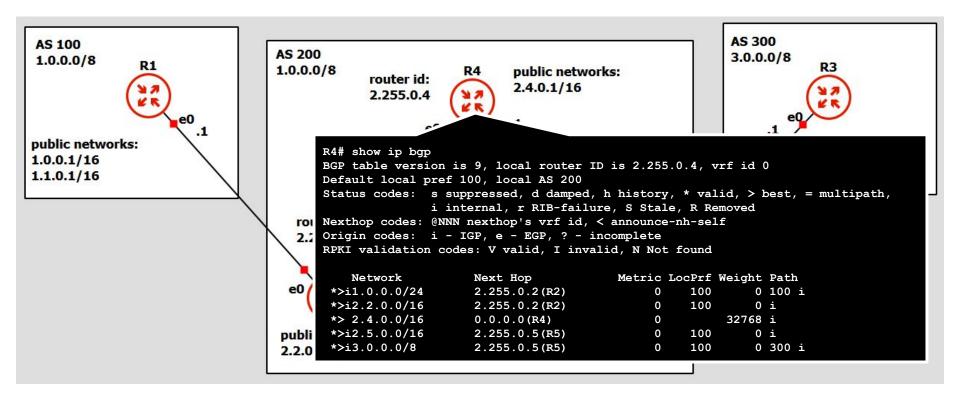
#### Test case 3: OSPF DB on R4



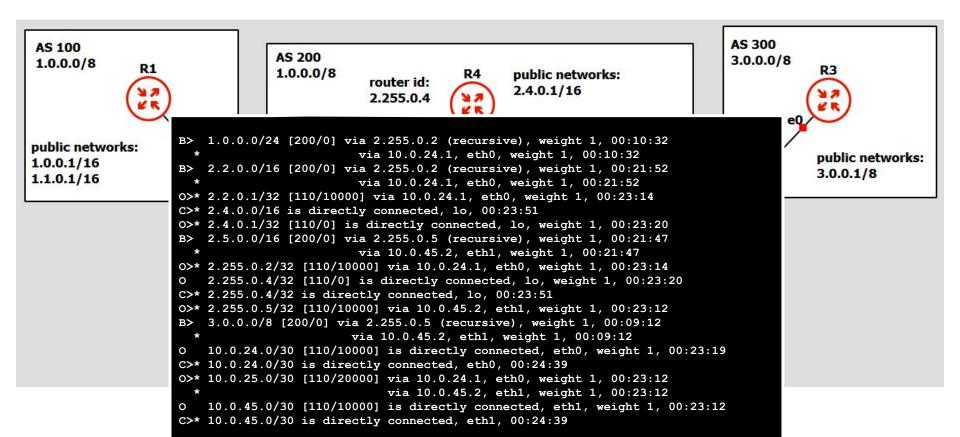
## Test case 3: BGP summary on R4



#### Test case 3: BGP table on R4



# Test 3: Routing table on R4



# Test 3: Ping from R4

