

# BGP路径属性详解

朱仕耿

[www.huawei.com](http://www.huawei.com)

Author / Email : Zhushigeng 261992 / zhushigeng@huawei.com

Version 1.1

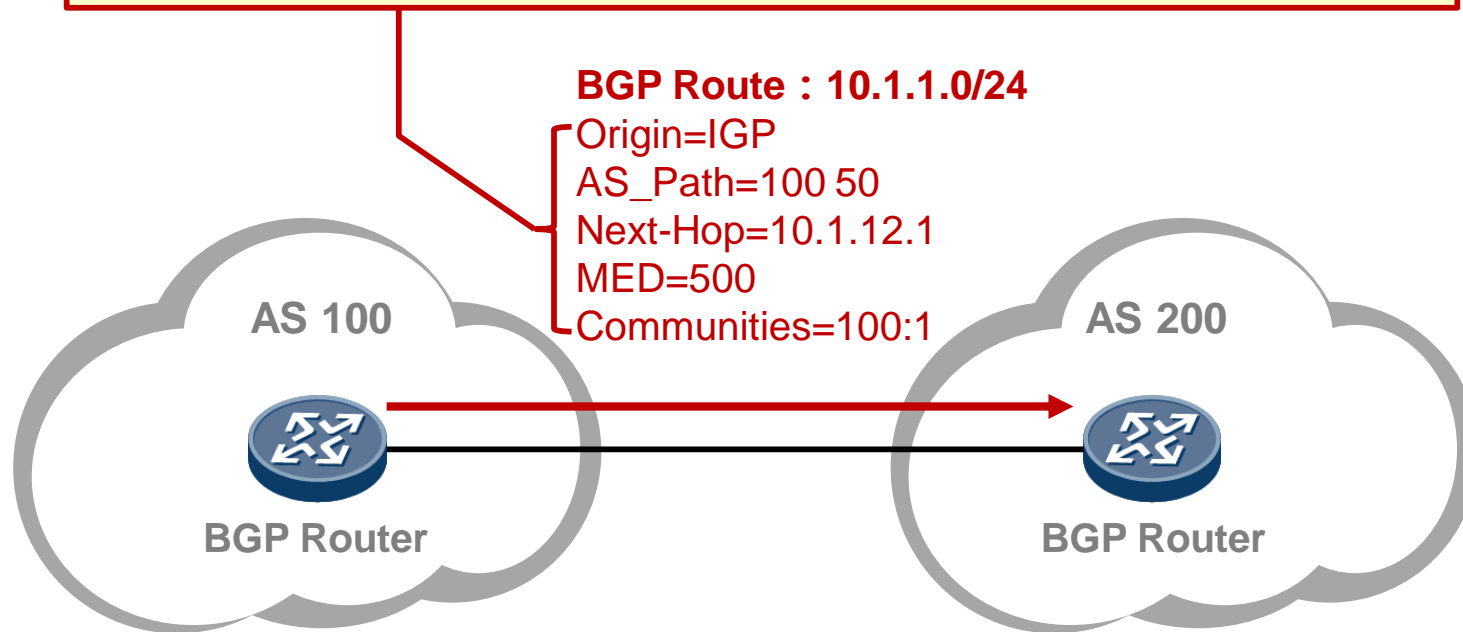


# 课程目标

- 理解BGP路径属性的概念及意义。
- 熟悉BGP常用路径属性：Preferred-value，Local\_Preference，AS\_Path，Origin，MED，Next-Hop，Community，Atomic\_Aggregate及aggregator。

# BGP路径属性

当一条BGP路由被BGP路由器更新给其对等体时，这条BGP路由会携带多个路径属性值（Path Attributes）一并传递给对等体。BGP的这些路径属性，将影响BGP的路由优选。它们的存在使得BGP的路由策略能力异常强大。



# 路径属性分类

- 公认属性 Well-Known
  - 公认必遵属性 Well-known mandatory
  - 公认自由决定属性 Well-known discretionary
- 可选属性 Optional
  - 可选传递的 Optional non-transitive
  - 可选非传递的 Optional non-transitive

## 路径属性分类

公认属性	公认必遵	所有的BGP实现都必须都能识别，且在Update报文中必须携带。	Origin AS_Path Nexthop
	公认自决	所有的BGP实现都必须都能识别，但不要求必须包含在Update报文中。	Local-Preference ATOMIC_Aggregate
可选属性	可选传递	设备可以不支持该属性，但即使不支持，也应当接收包含该属性的路由并传递给其他对等体。	Community Aggregator
	可选非传递	设备可以不支持该属性，不识别的BGP进程忽略包含这个属性的路由更新，并且不传递给其他BGP对等体。	MED Originator_ID Cluster_list *pre_value

# BGP update报文

BGP update报文  
中的路径属性

BGP路由前缀

## Border Gateway Protocol

### [-] UPDATE Message

Marker: 16 bytes

Length: 55 bytes

Type: UPDATE Message (2)

Unfeasible routes length: 0 bytes

Total path attribute length: 28 bytes

### [-] Path attributes

[+] ORIGIN: IGP (4 bytes)

[+] AS\_PATH: empty (3 bytes)

[+] NEXT\_HOP: 1.1.1.1 (7 bytes)

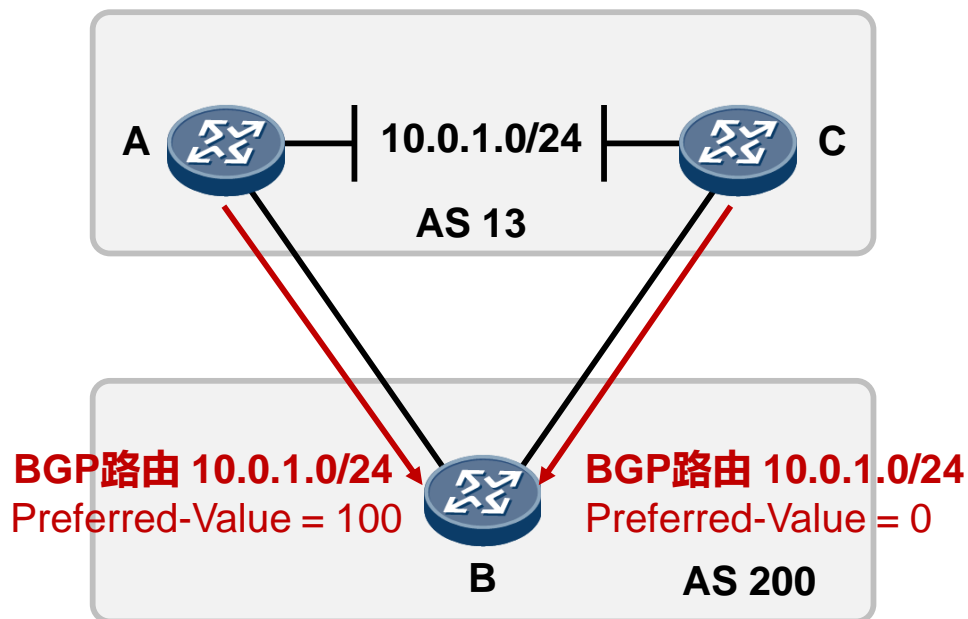
[+] MULTI\_EXIT\_DISC: 0 (7 bytes)

[+] LOCAL\_PREF: 100 (7 bytes)

[+] Network layer reachability information: 4 bytes

# Preferred-Value

- 华为私有的路径属性，相当于路由的权重值，取值范围：0~65535；该值越大，则路由越优先。
- Preferred-Value只能在路由器本地配置，而且只影响本设备的路由优选。该属性不会传播给任何BGP对等体。
- 路由器本地始发的BGP路由默认的Preferred-Value为0，从其他BGP对等体学习到的路由默认Preferred-Value也为0。



在B上部署路由策略（Import策略），将A传递过来的10.0.1.0/24路由的Preferred-Value值设定为100，而C传递过来的路由的Preferred-Value值则保持缺省，也就是0。如此一来关于10.0.1.0/24，B会优选A传递过来的路由。

# 在BGP表中查看Preferred-Value

[B] display bgp routing-table

BGP Local router ID is 2.2.2.2

Status codes: \* - valid, > - best, d - damped,  
h - history, i - internal, s - suppressed, S - Stale  
Origin : i - IGP, e - EGP, ? - incomplete

Total Number of Routes: 4

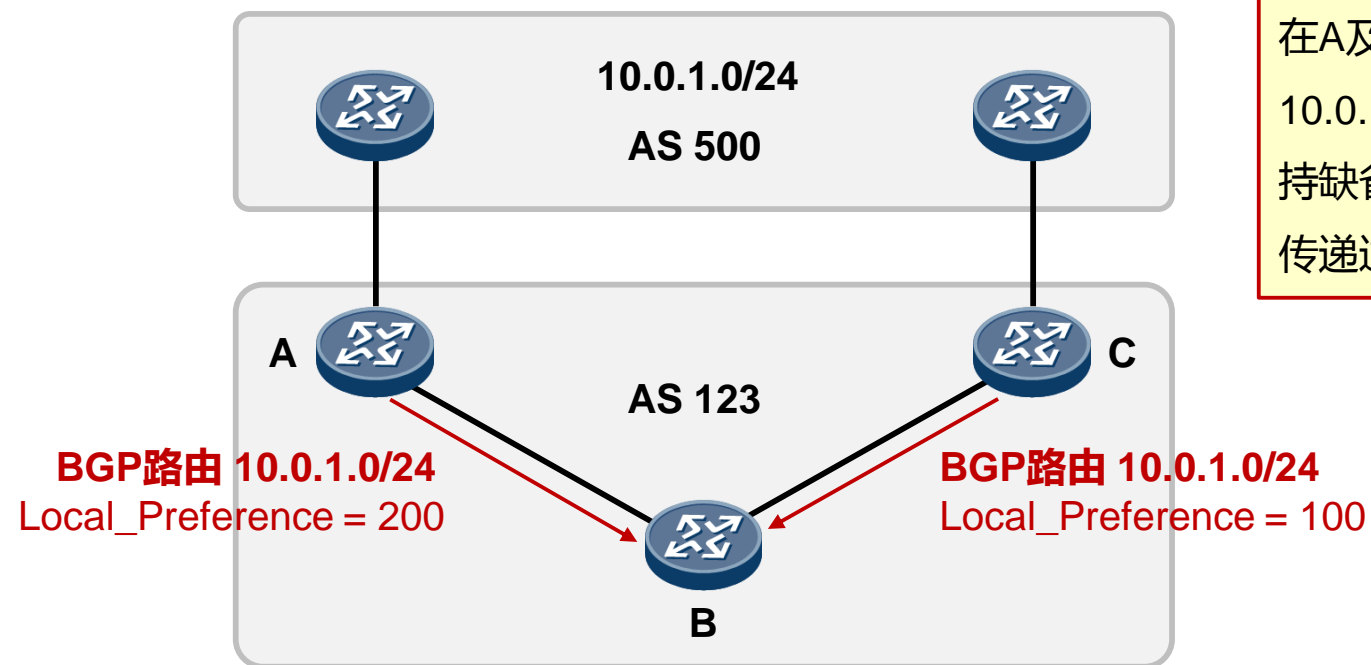
	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	10.0.1.0/24	10.1.12.1	0		100	13 i
*		10.1.23.3	0		0	13 i

Preferred-value



# Local\_Preference

- 也即本地优先级属性，是公认自决属性，用于告诉AS中的路由器，哪条路径是离开AS的首选路径。
- Local\_Preference属性值越大则BGP路由越优。缺省的Local\_Preference值为100。
- 该属性只能被传递给IBGP对等体，而不能传递给EBGP对等体。



在A及C上分别对B部署路由策略，使得A发送给B的10.0.1.0/24路由的Local\_Preference为200，而C则保持缺省，那么对于B而言，关于10.0.1.0/24，会优选A传递过来的BGP路由。

# Local\_Preference

[B] display bgp routing-table

BGP Local router ID is 2.2.2.2

Status codes: \* - valid, > - best, d - damped,

h - history, i - internal, s - suppressed, S - Stale

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

Total Number of Routes: 4

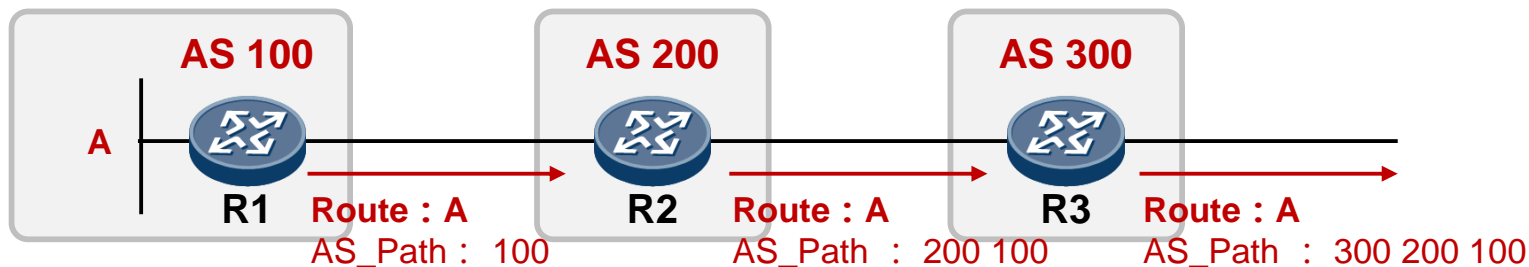
Network	NextHop	MED	Local_Preference LocPrf	PrefVal	Path/Ogn
*>i 10.0.1.0/24	10.1.12.1	0	200	0	i
* i	10.1.23.3	0	100	0	i

## Local\_Preference注意事项

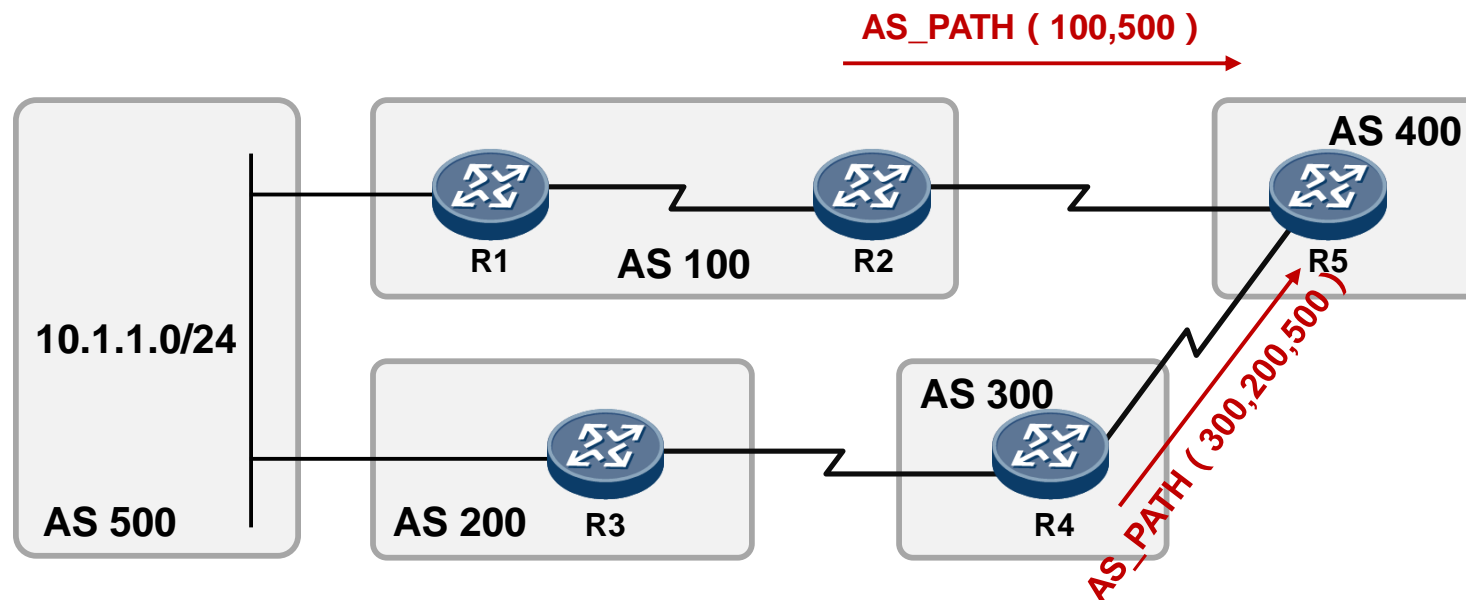
- Local\_Preference属性只能在IBGP对等体间传递（除非做了策略否则Local\_Preference值在IBGP对等体间传递过程中不会丢失），而不能在EBGP对等体间传递，如果在EBGP对等体间收到的路由的路径属性中携带了Local\_Preference，则会触发Notification报文，造成会话中断；
- 但是可以在AS边界路由器上使用Import方向的策略来修改Local\_Preference属性值。也就是在收到路由之后，在本地为路由赋予Local\_Preference。
- 使用bgp default local-preference命令修改缺省Local\_Preference值，该值缺省为100。
- 路由器在向其EBGP对等体发送路由更新时，不能携带Local\_Preference属性，但是对方接收路由之后，会在本地为这条路由赋一个缺省Local\_Preference值（100），然后再将路由传递给自己的IBGP对等体。
- 本地使用network命令引入及重发布引入的路由，Local\_Preference为缺省值100，并能在AS内向其他IBGP对等体传递，传递过程中除非受路由策略影响，否则Local\_Preference不变。

## AS\_Path

- 该属性为公认必遵属性，是前往目标网络的路由经过的AS号列表；
- 作用：确保路由在EBGP对等体之间传递无环；另外也作为路由优选的衡量标准之一；
- 路由在被通告给EBGP对等体时，路由器会在该路由的AS\_Path中追加上当地的AS号；路由被通告给IBGP对等体时，AS-path不会发生改变。



# AS\_Path

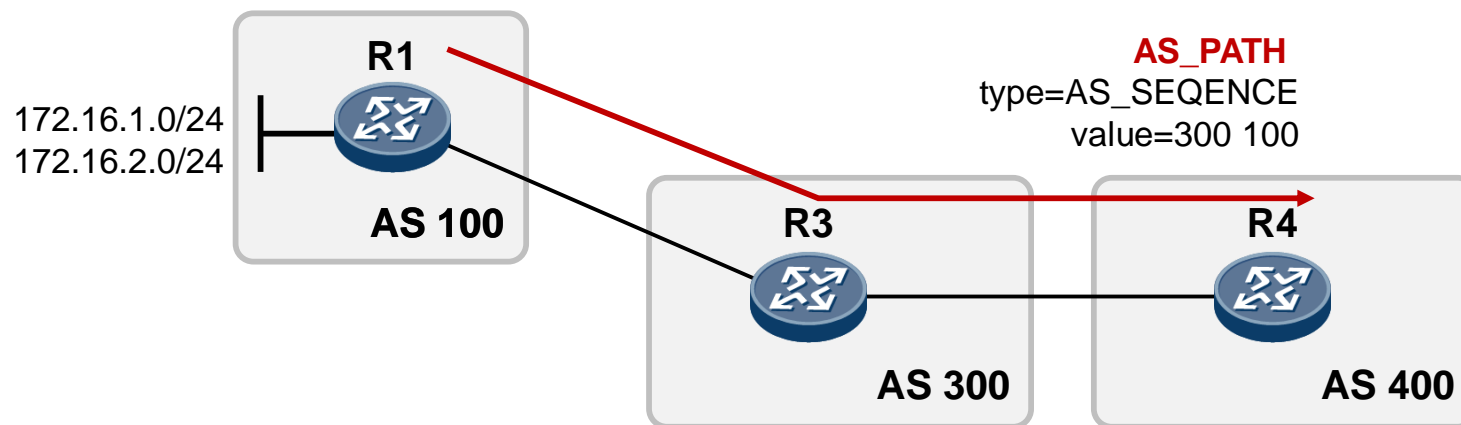


AS\_Path的重要作用之一便是影响BGP路由的优选，在上图中，R5同时从R2及R4学习到去往10.1.1.0/24网段的BGP路由，在其他条件相同的情况下，R5会优选R2通告的路由，因为该条路由的AS\_Path属性值较短，也即AS号的个数更少。

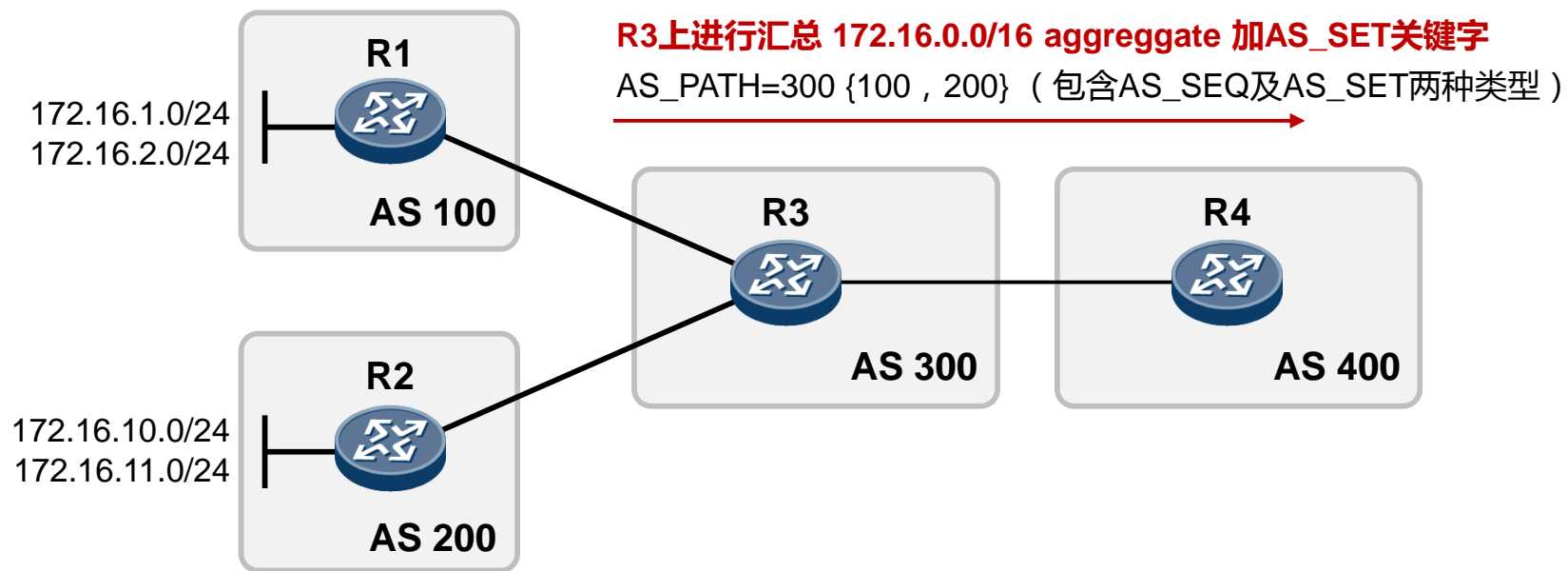
## AS\_Path的四种类型

- **AS\_SET** : 一个去往特定目的地所经路径上的无序AS号列表。
- **AS\_SEQUENCE** : 一个有序的AS号列表。
- **AS\_CONFED\_SEQUENCE** : 一个去往特定目的地所经路径上的有序AS 号列表，其用法与AS\_SEQUENCE完全一样，区别在于该列表中的AS号属于本地联邦中的AS。
- **AS\_CONFED\_SET** : 一个去往特定目的地所经路径上的无序AS号列表，其用方法与AS\_SET完全一样，区别在于列表中的AS号属于本地联邦中的AS。

## AS\_Path的四种类型：AS\_SEQUENCE

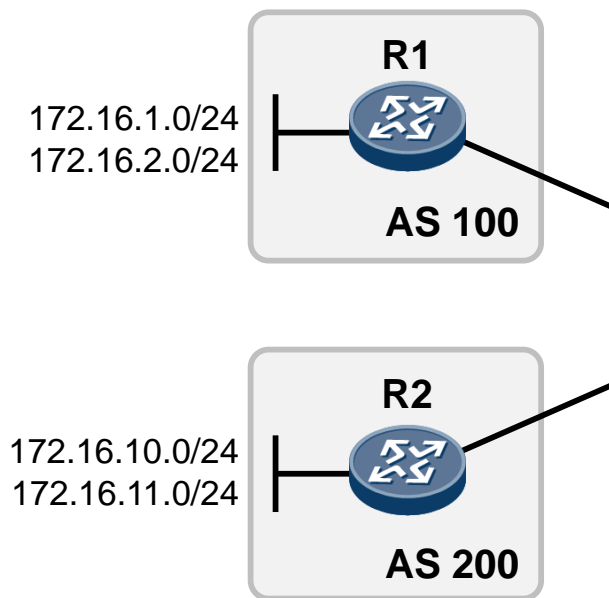


## AS\_Path的四种类型：AS\_SET





# AS\_Path的四种类型：AS\_SET



## Border Gateway Protocol

### UPDATE Message

Marker: 16 bytes

Length: 66 bytes

Type: UPDATE Message (2)

Unfeasible routes length: 0 bytes

Total path attribute length: 40 bytes

### Path attributes

ORIGIN: IGP (4 bytes)

AS\_PATH: 300 {100, 200} (13 bytes)

Flags: 0x40 (well-known, Transitive, Complete)

Type code: AS\_PATH (2)

Length: 10 bytes

AS path: 300 {100, 200}

AS path segment: 300

Path segment type: AS\_SEQUENCE (2)

Path segment length: 1 AS

Path segment value: 300

AS\_SEQ

AS path segment: {100, 200}

Path segment type: AS\_SET (1)

Path segment length: 2 ASs

Path segment value: 100 200

AS\_SET, 是无序的

NEXT\_HOP: 10.1.34.3 (7 bytes)

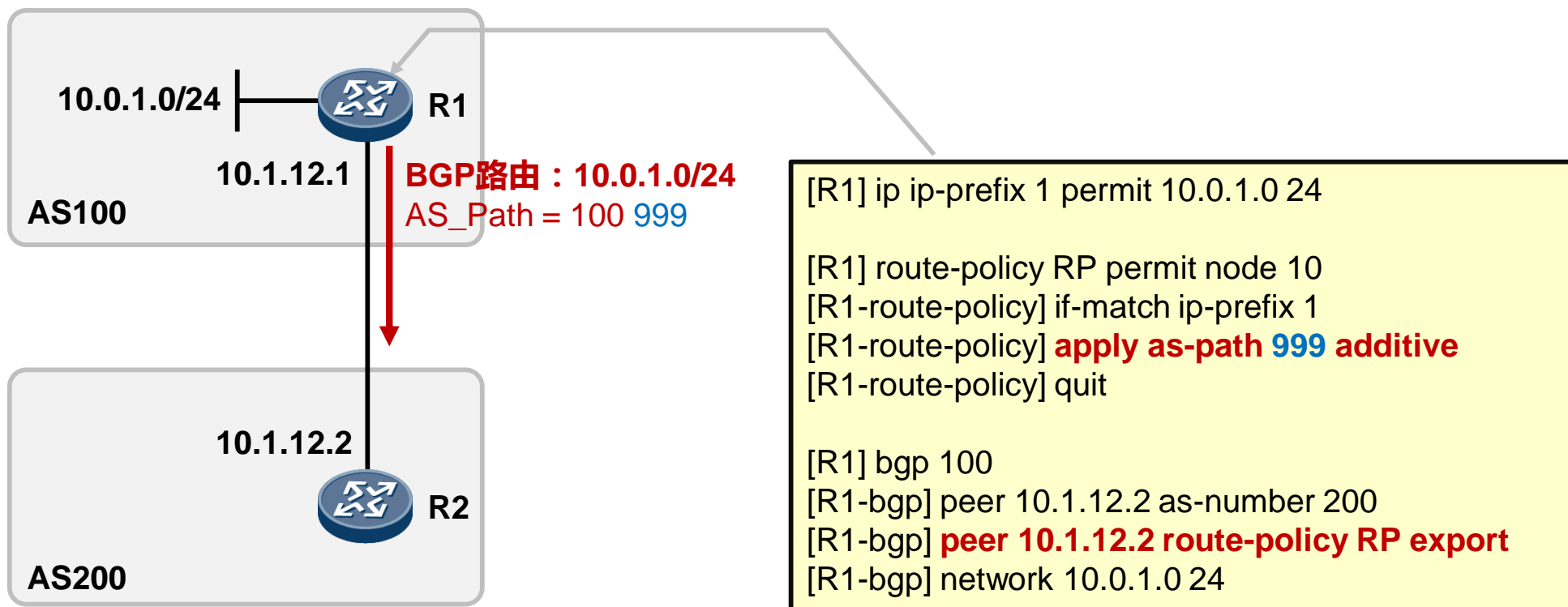
MULTI\_EXIT\_DISC: 0 (7 bytes)

AGGREGATOR: AS: 300 origin: 10.1.34.3 (9 bytes)

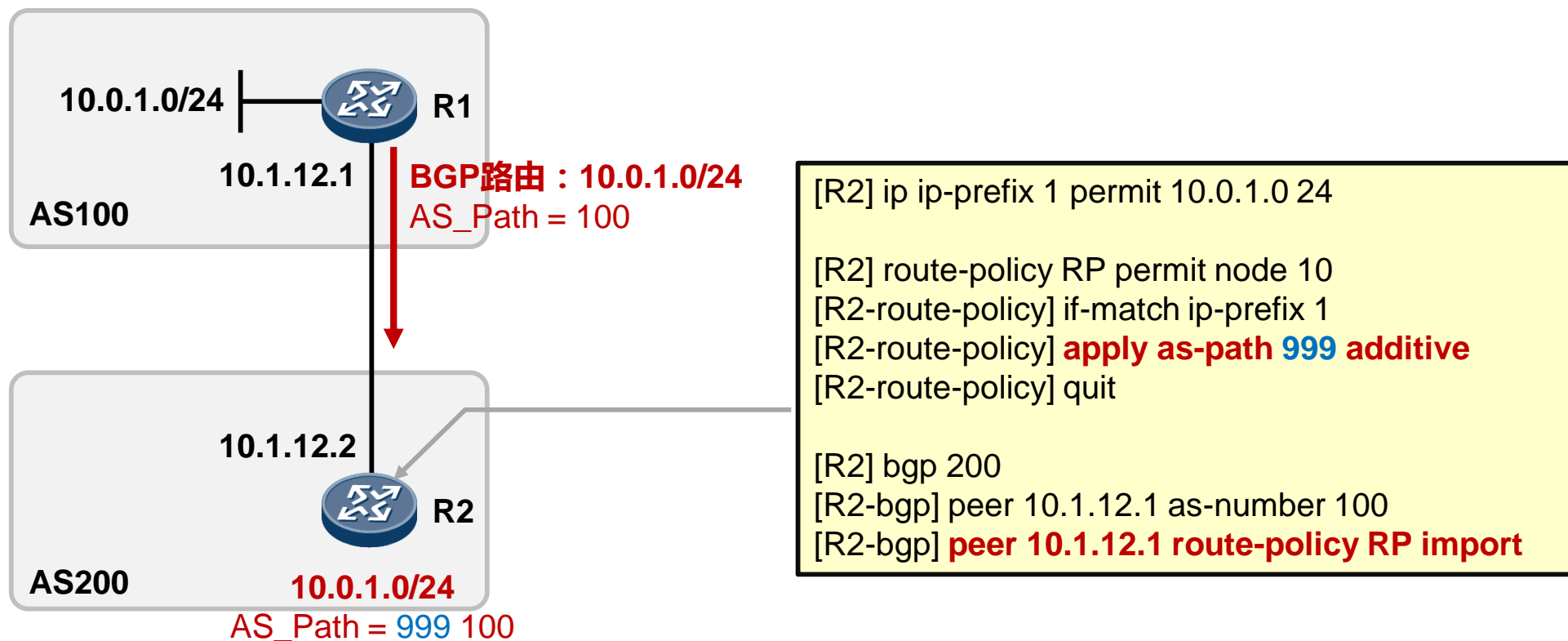
Network layer reachability information: 3 bytes

172.16.0.0/16

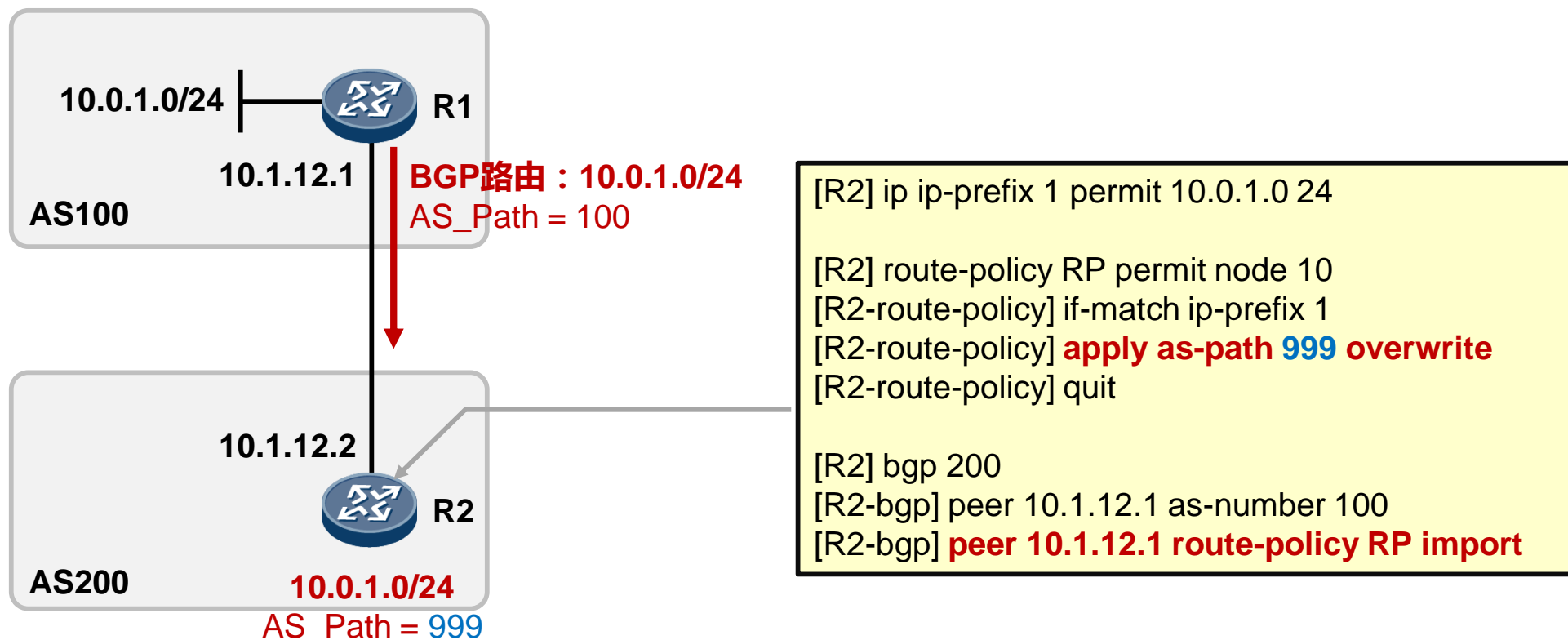
## 使用route-policy修改AS\_Path属性 示例1



## 使用route-policy修改AS\_Path属性 示例2



## 使用route-policy修改AS\_Path属性 示例3



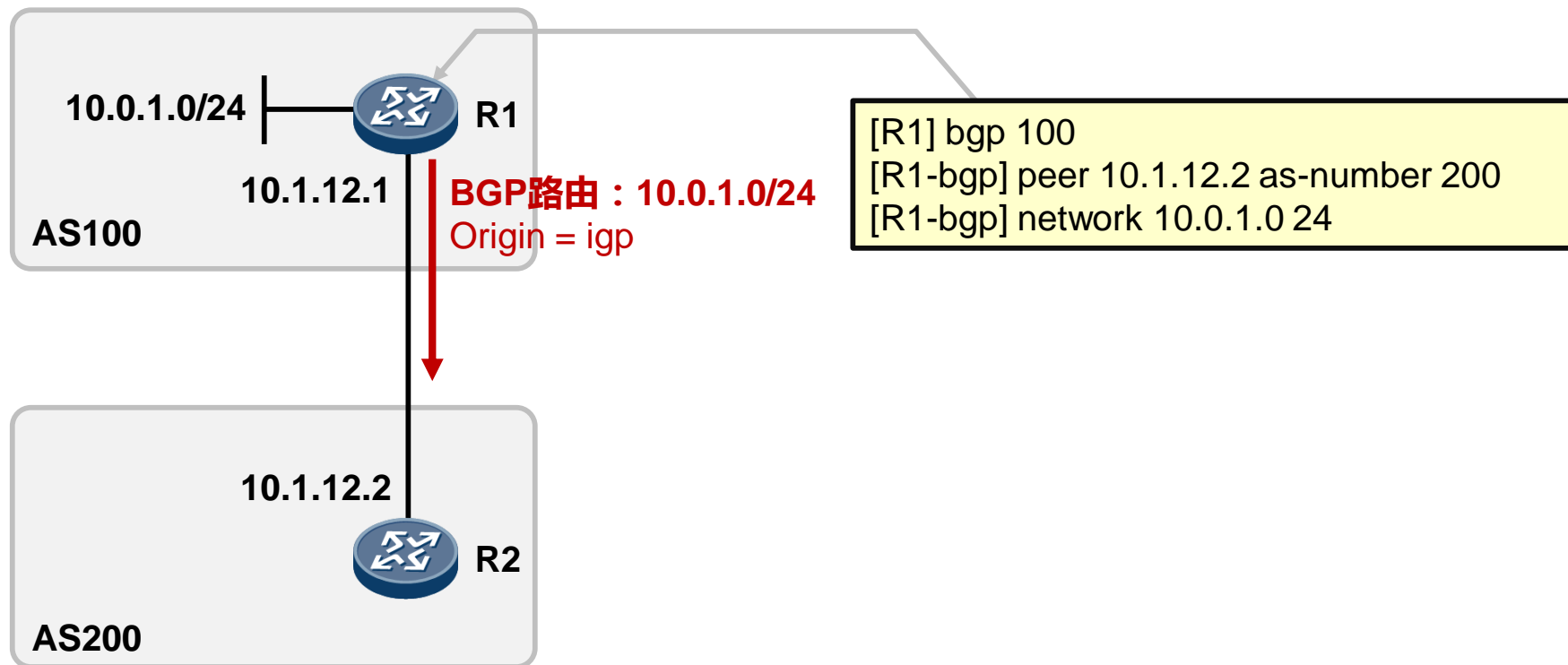
## Origin

- 该属性为公认必遵属性，它标识了BGP路由的起源。如下表所示，根据路由被引入BGP的方式不同，存在三种类型的Origin。
- 当去往同一个目的地存在多条不同Origin属性的路由时，在其他条件都相同的情况下，BGP将按如Origin的下顺序优选路由：IGP > EGP > Incomplete。

名称	标记	描述
igp	i	通过BGP network的路由，也就是起源于IGP的路由，Origin为igp。因为BGP network必须保证该网络在路由表中。
egp	E	如果BGP路由是由EGP 这种早期的协议重发布而来，那么其Origin为egp。
Incomplete	?	通过Import命令，从其他协议引入到BGP的路由，其Origin为Incomplete（确认该路由来源的信息不完全）。

# Origin

- R1采用network的方式将直连路由10.0.1.0/24引入BGP，则路由的Origin为igp。



## Origin在BGP表中的显示

[R2] display bgp routing-table

BGP Local router ID is 10.1.12.1

Status codes: \* - valid, > - best, d - damped,

h - history, i - internal, s - suppressed, S - Stale

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

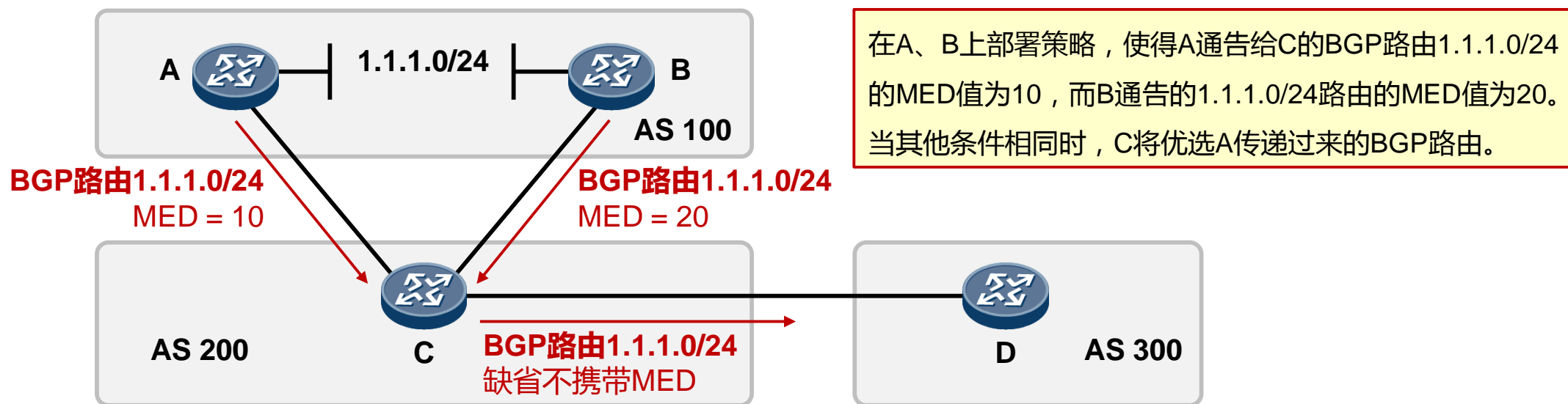
Total Number of Routes: 3

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	10.0.1.0/24	10.1.12.1	0		0	100 <i>i</i>

↑  
**Origin**

# MED

- MED ( Multi Exit Discriminator ) 是可选非传递属性，是一种度量值，用于向外部对等体指出进入本AS的首选路径，即当进入本AS的入口有多个时，AS可以使用MED动态地影响其他AS选择进入的路径。
- MED属性值越小则BGP路由越优。
- MED主要用于在AS之间影响BGP的选路。MED被传递给EBGP对等体后，对等体在其AS内传递路由时，携带该MED值，但将路由传递给其EBGP对等体时，缺省不会携带MED属性。

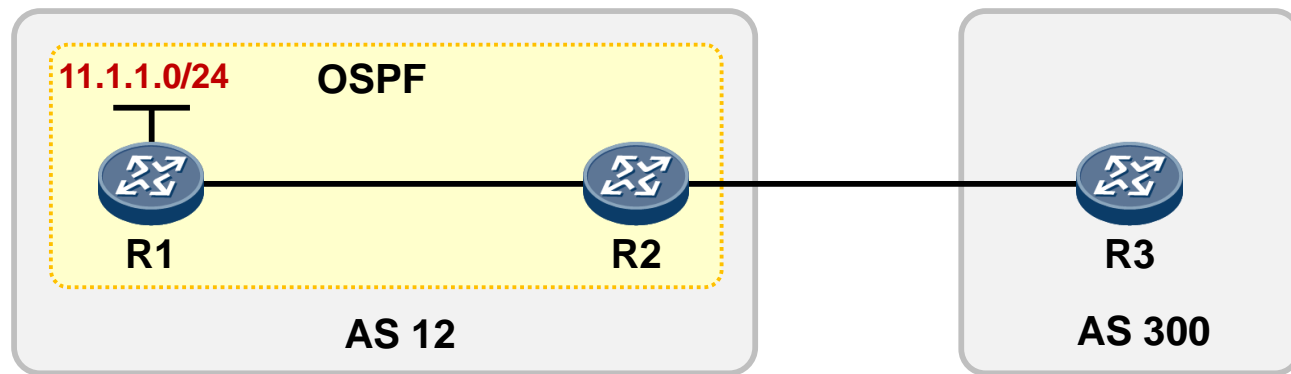




## 关于MED的一些注意事项

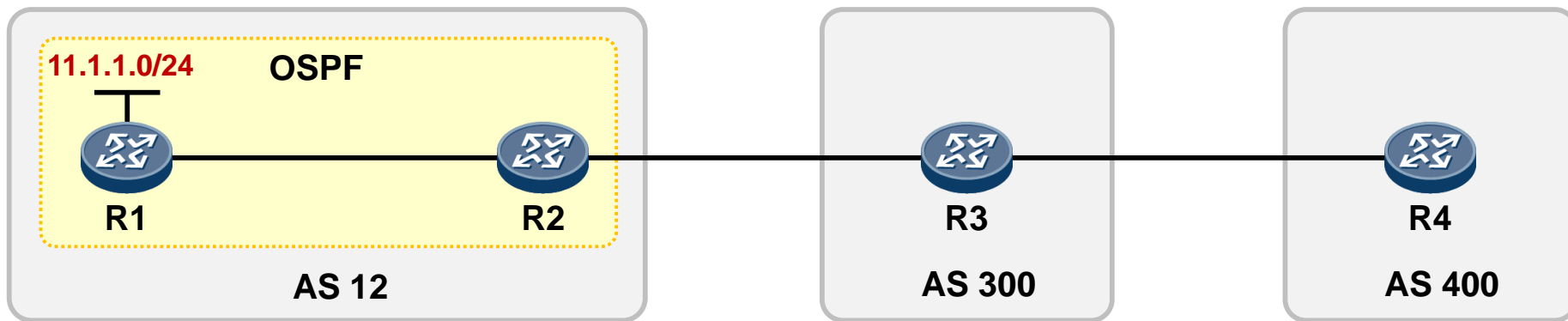
- 缺省情况下，路由器只比较来自同一相邻AS的BGP路由的MED值，也就是说如果去往同一个目的地的两条路由来自不同的相邻AS，则不进行MED值的比较。
- 一台BGP路由器将路由通告给EBGP对等体时，是否携带MED属性，需要根据以下条件进行判断（不对EBGP对等体使用策略的情况下）：
  - 如果该BGP路由是本地始发（本地通过network或import-route命令引入）的，则缺省携带MED属性发送给EBGP对等体。
  - 如果该BGP路由是从其他BGP对等体学习过来的，那么将该路由通告给EBGP对等体时不携带MED。
- 在IBGP对等体之间传递路由时，MED值会被保留并传递，除非部署了策略，否则MED值在传递过程中不发生改变也不会丢失。

## MED的默认操作



- 如果路由器通过IGP学习到一条路由，并通过network或import-route的方式将路由引入BGP，产生的BGP路由的MED值继承路由在IGP中的metric。例如上图中如果R2通过OSPF学习到了11.1.1.0/24路由，并且该路由在R2的全局路由表中OSPF Cost=100，那么当R2将路由network进BGP后，产生的BGP路由的MED值为100。
- 如果路由器将本地直连路由通过network或import-route的方式引入BGP，那么这条BGP路由的MED为0，因为直连路由cost为0。
- 如果路由器将本地静态路由通过network或import-route的方式引入BGP，那么这条BGP路由的MED为0，因为静态路由cost为0。

## MED的默认操作



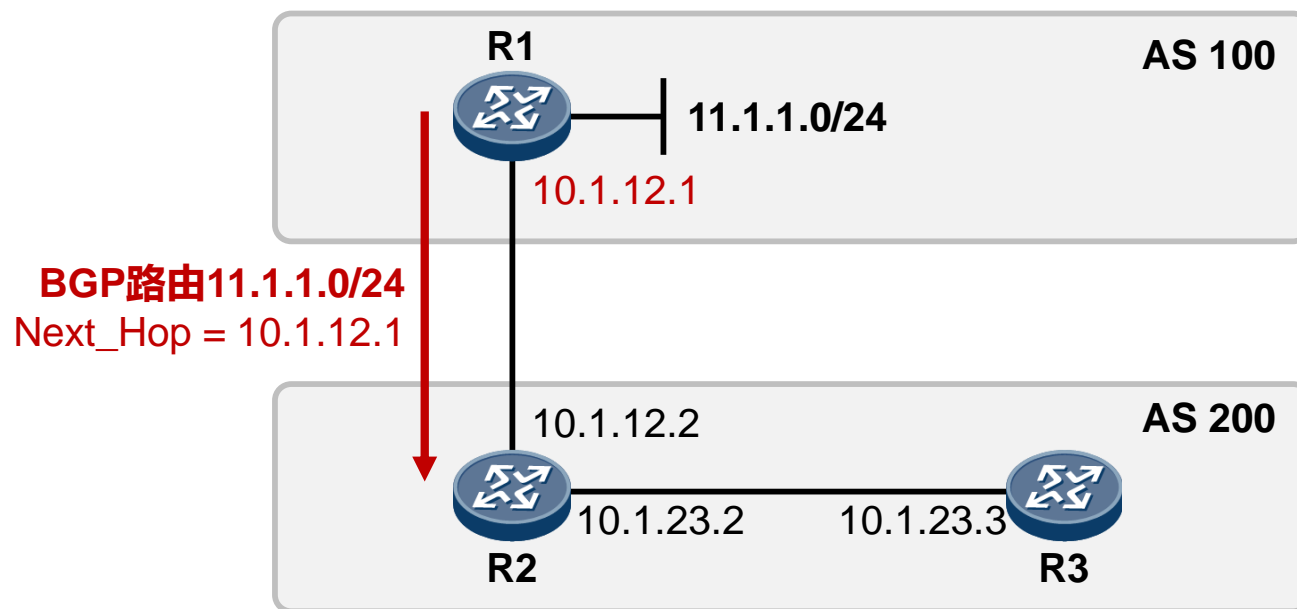
- 如果路由器通过BGP学习到其他对等体传递过来的一条路由，那么将路由更新给自己的EBGP对等体时，默认是不携带MED的。这就是所谓的：“MED不会跨AS传递”。例如在上图中，R2将携带MED属性的BGP路由传递给R3，则R3将该路由传递给R4时，不会携带MED。
- 可以使用default med命令修改缺省的MED值，default med命令只对本设备上用import-route命令引入的路由和BGP的聚合路由生效。例如在R2上配置default med 999，那么R2通过import-route及aggregate命令产生的路由传递给R3时，路由携带的MED为999。

## Next\_Hop

- 该属性是一个公认必遵属性，用于指定到达目标网络的下一跳地址。
- 当路由器学习到BGP路由后，需对BGP路由的Next\_Hop属性值进行检查，该属性值（IP地址）必须在本地路由可达，如果不可达，则这条BGP路由不可用。
- 在EBGP及IBGP对等体的场景中，Next\_Hop的缺省操作是存在差异的。

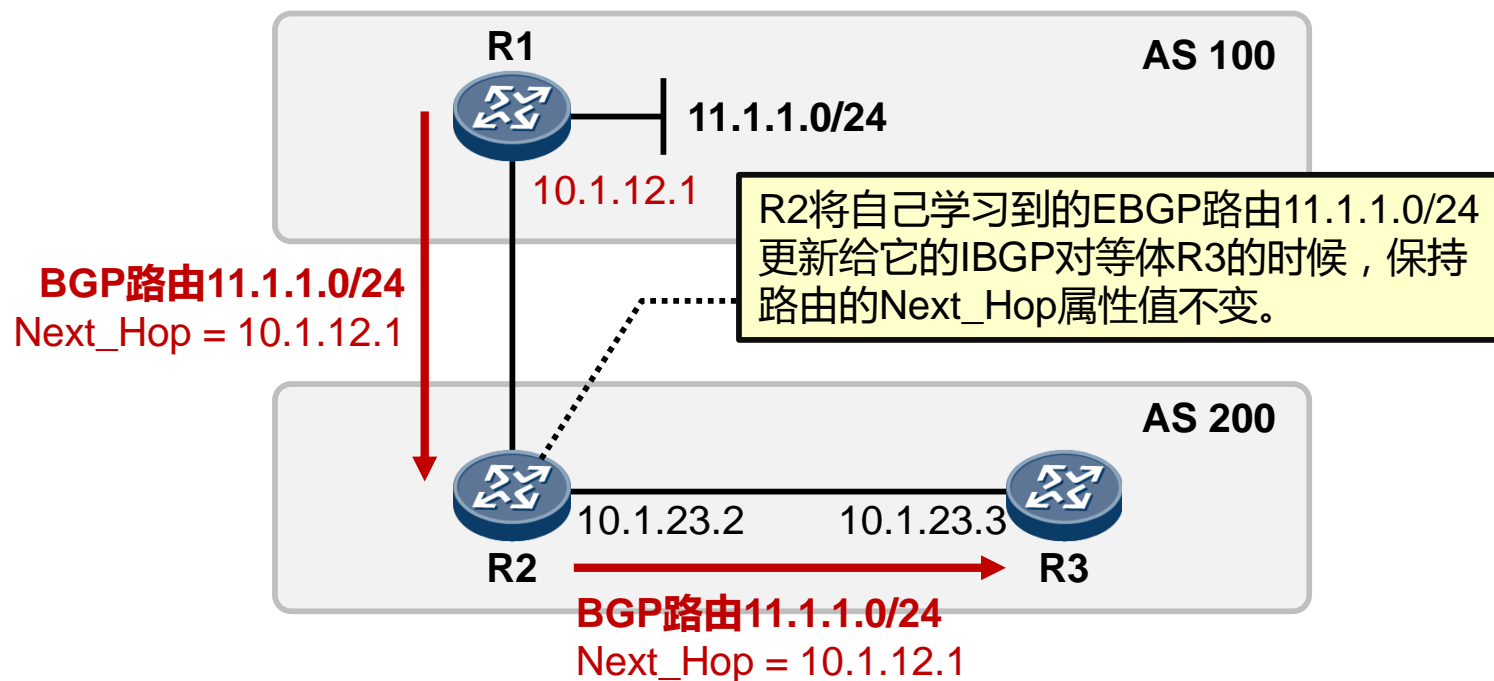
## Next\_Hop的缺省操作

- 路由器将BGP路由通告给自己的EBGP对等体时，将该路由的Next\_Hop设置为自己的更新源IP地址。

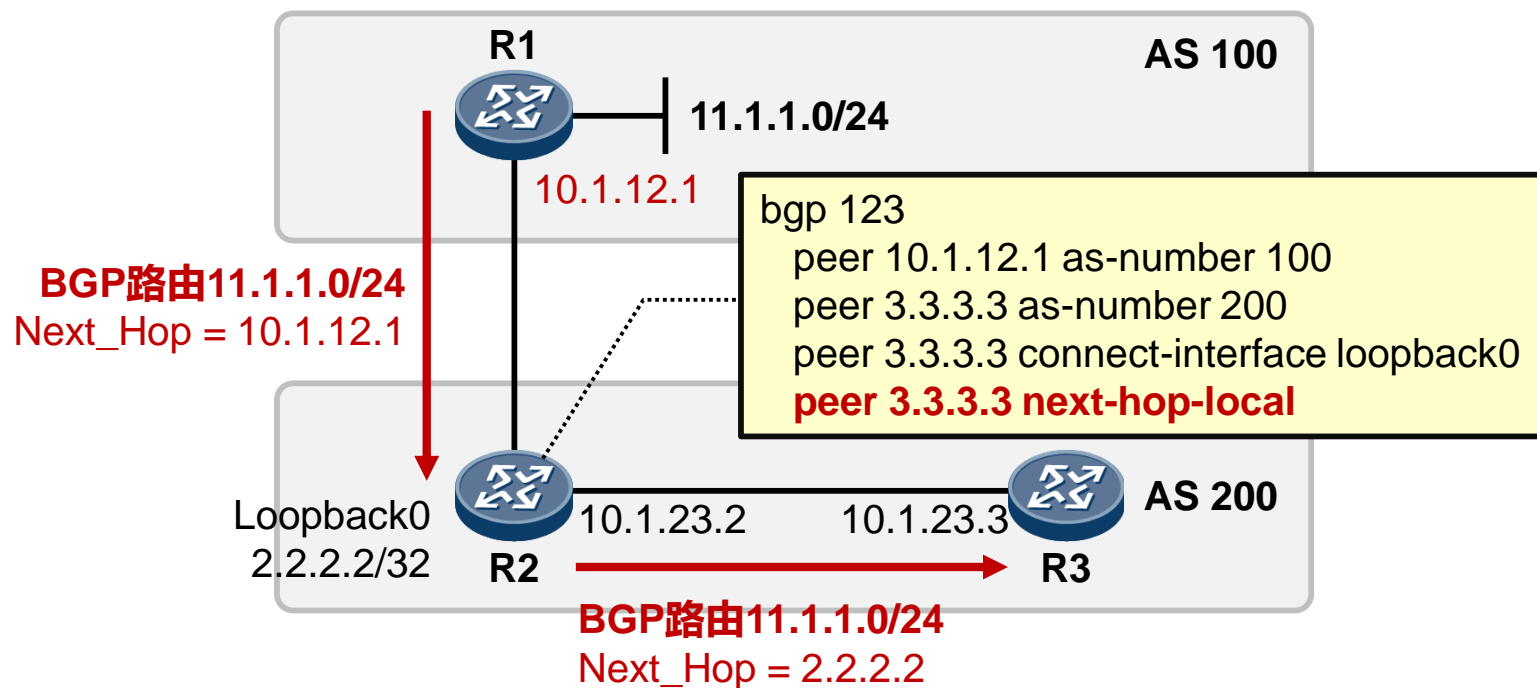


## Next\_Hop的缺省操作

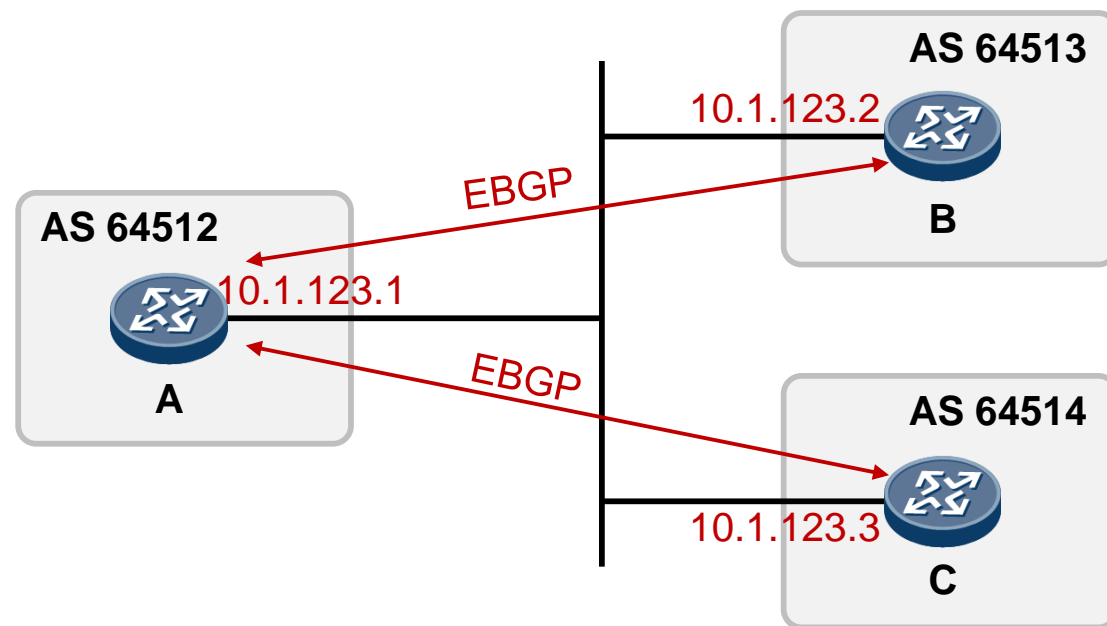
- 路由器在收到EBGP对等体所通告的BGP路由后，在将路由传递给自己的IBGP对等体时，会保持路由的Next\_Hop属性值不变。



## 通过next-hop-local变更next-hop属性



## Next\_Hop的缺省操作



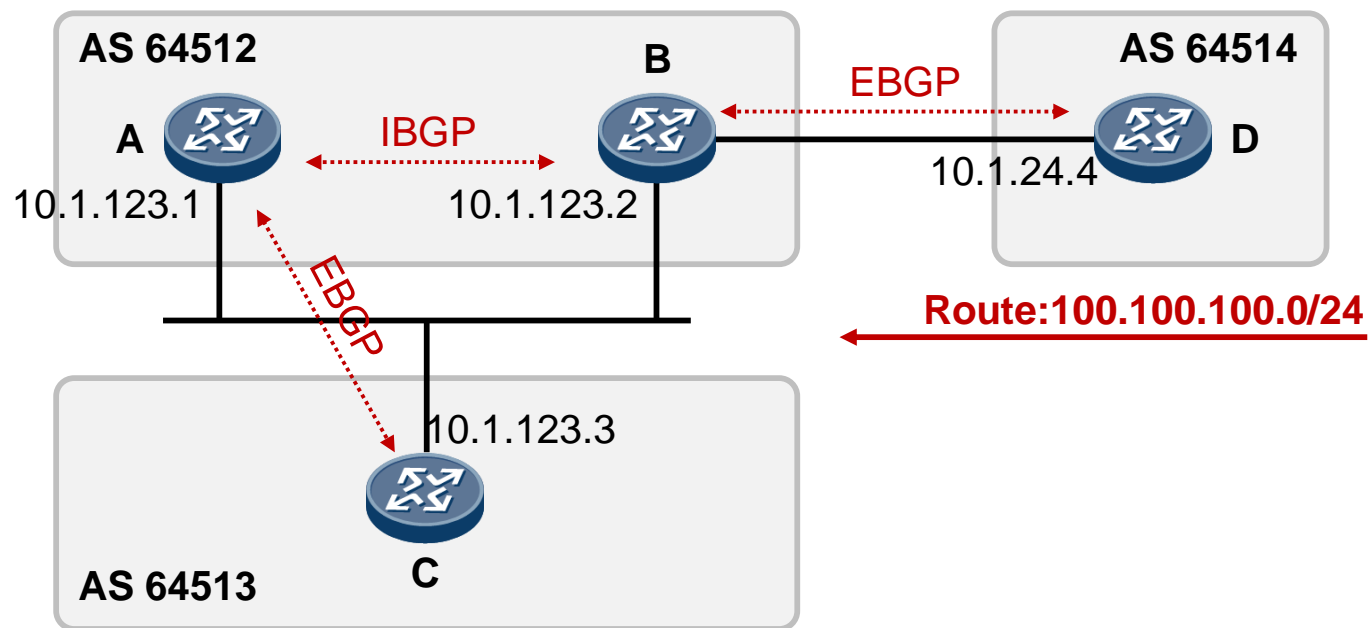
B将路由100.100.100.0/24传递给A，Next\_Hop为10.1.123.2；

A将路由100.100.100.0/24传递给C，此时Next\_Hop保持不变；

如果路由器收到某条BGP路由，该路由的Next\_Hop地址值与EBGP对等体（更新对象）同属一个网段，那么该条路由的Next\_Hop地址将保持不变并传递给它的BGP对等体。

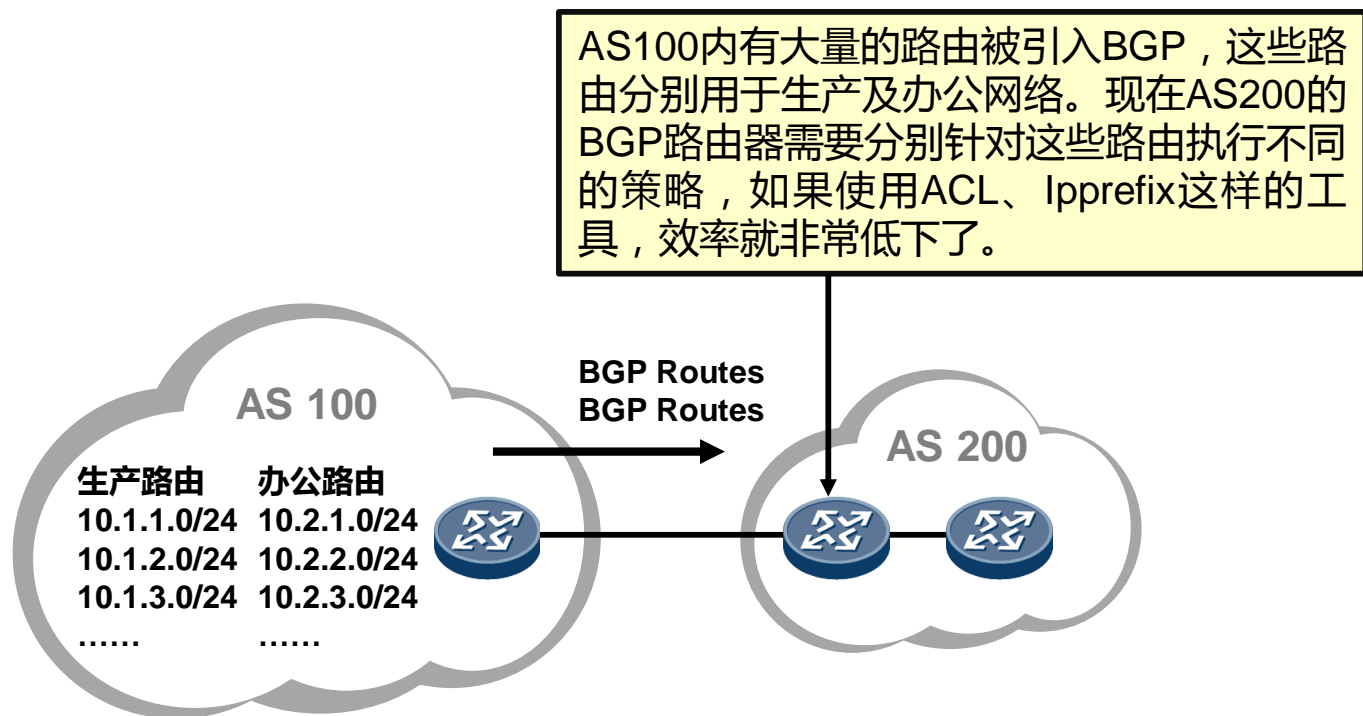


## Next\_Hop的缺省操作

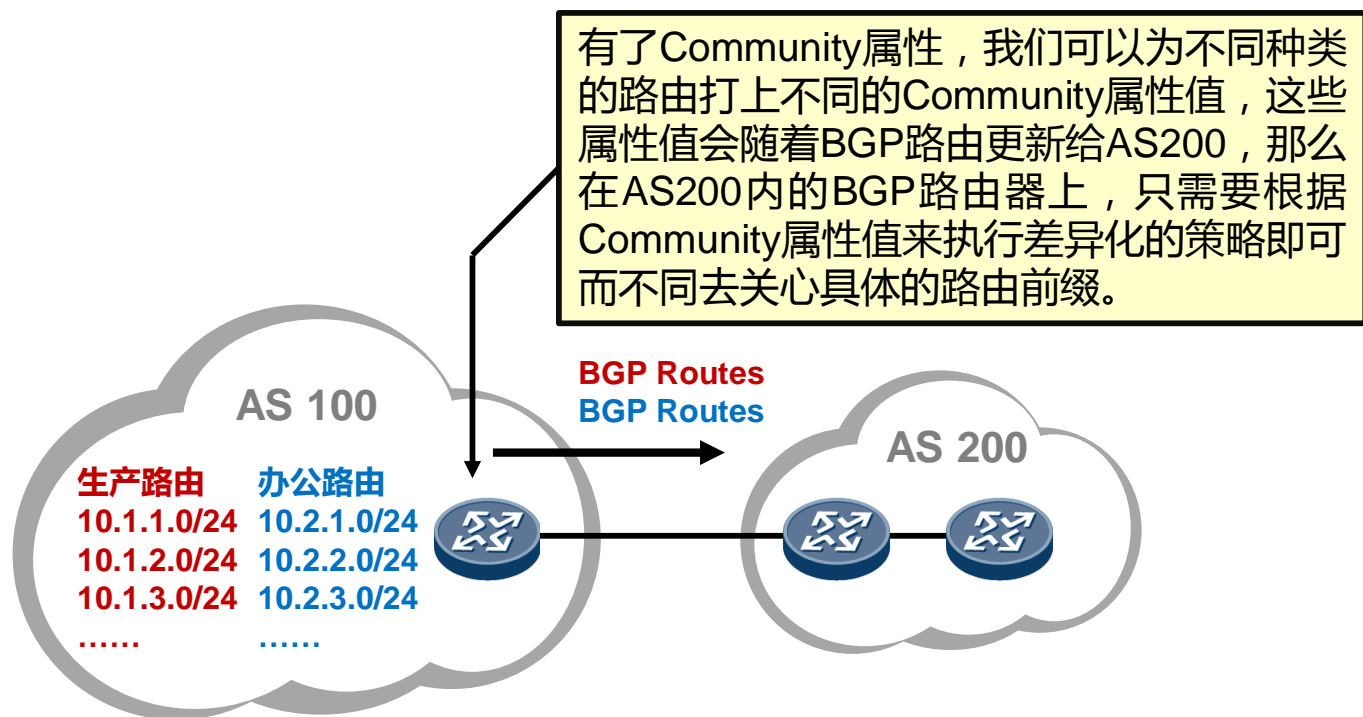


Router C 上关于100.100.100.0/24的路由，NEXT\_HOP属性为10.1.123.2

# Community 技术背景

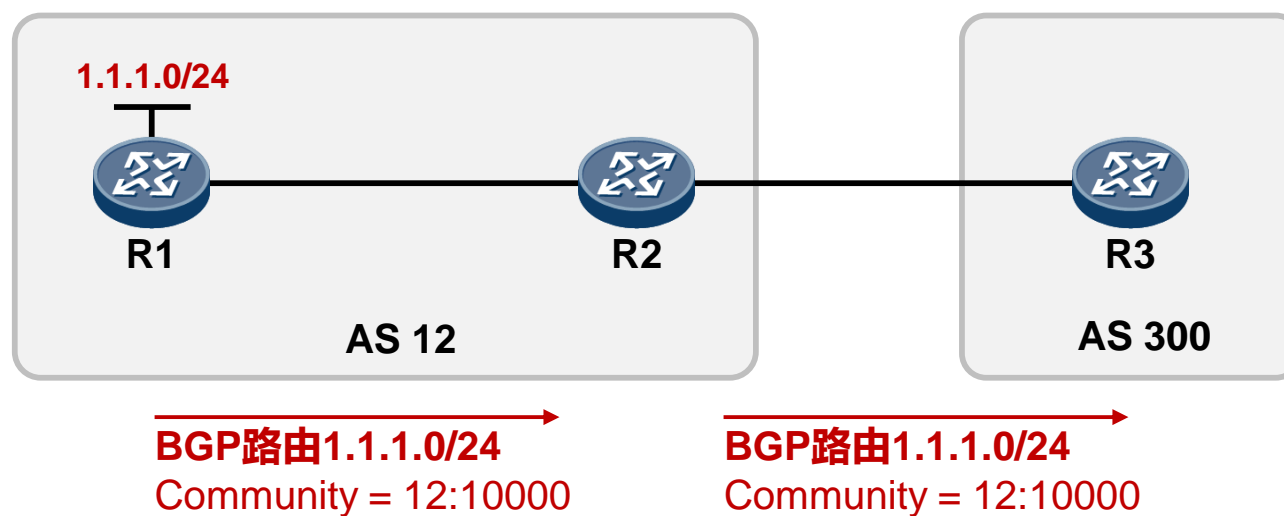


# Community 技术背景



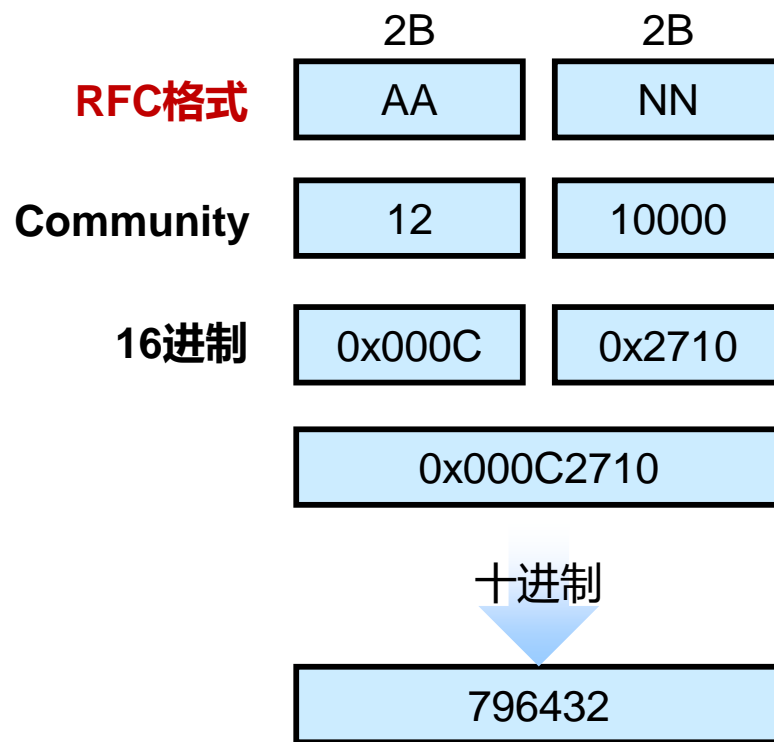
# Community

- 该属性为可选传递属性，是一种路由标记，用于简化路由策略的执行。
- 可以将某些路由分配一个特定的Community属性值，之后就可以基于Community值而不是网络号/掩码信息来抓取路由并执行相应的策略了。



# Community

- Community属性值长度为32个比特，也就是4个字节。可使用两种形式呈现，一是单个十进制整数格式，二是AA：NN格式，其中AA表示AS号，NN是自定义的编号。



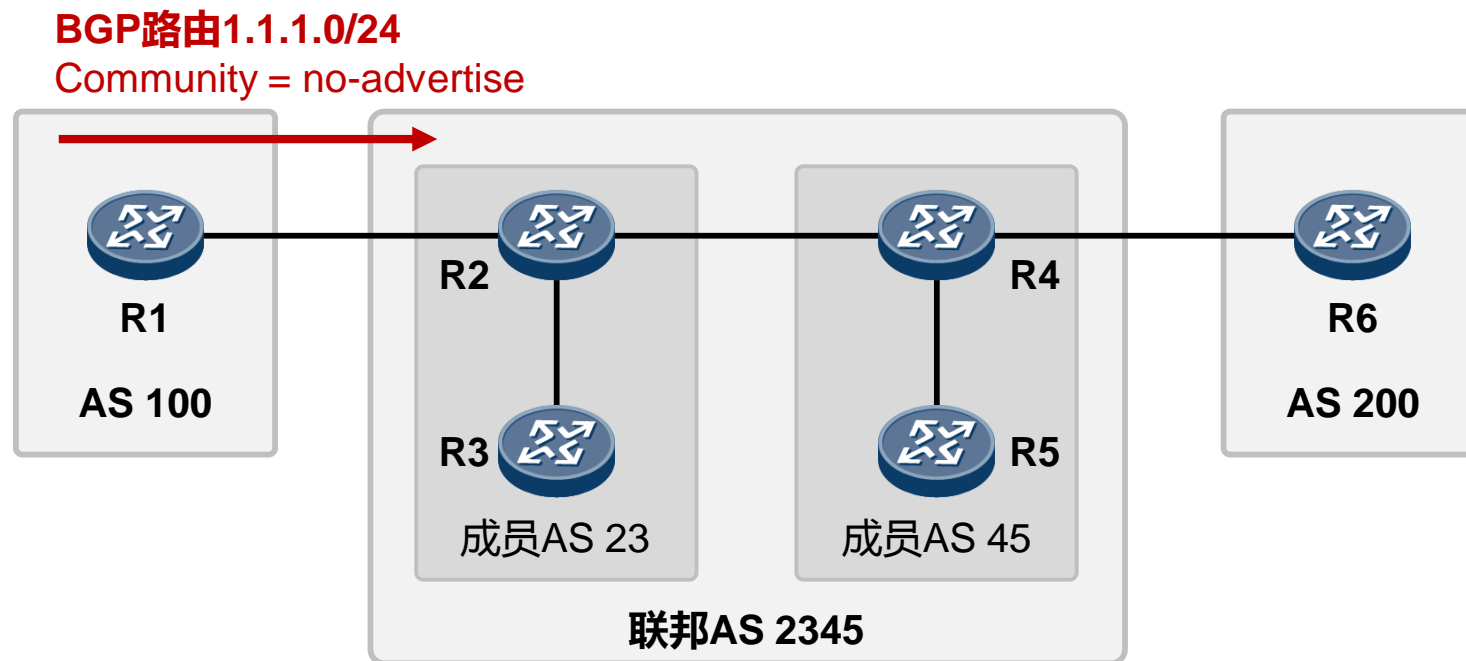
# 众所周知 ( Well-Known ) 的Community值

**route-policy xxx permit node 10**

**apply community ?**

INTEGER <0-4294967295>	Specify community number
STRING <3-11>	Specify aa<0-65535>:nn<0-65535>
internet	Internet (well-known community attributes)
no-advertise	Do not advertise to any peer (well-known community attributes)
no-export	Do not export to external peers(well-known community attributes)
no-export-subconfed	Do not send outside a sub-confederation(well-known community attributes)
none	No community attribute

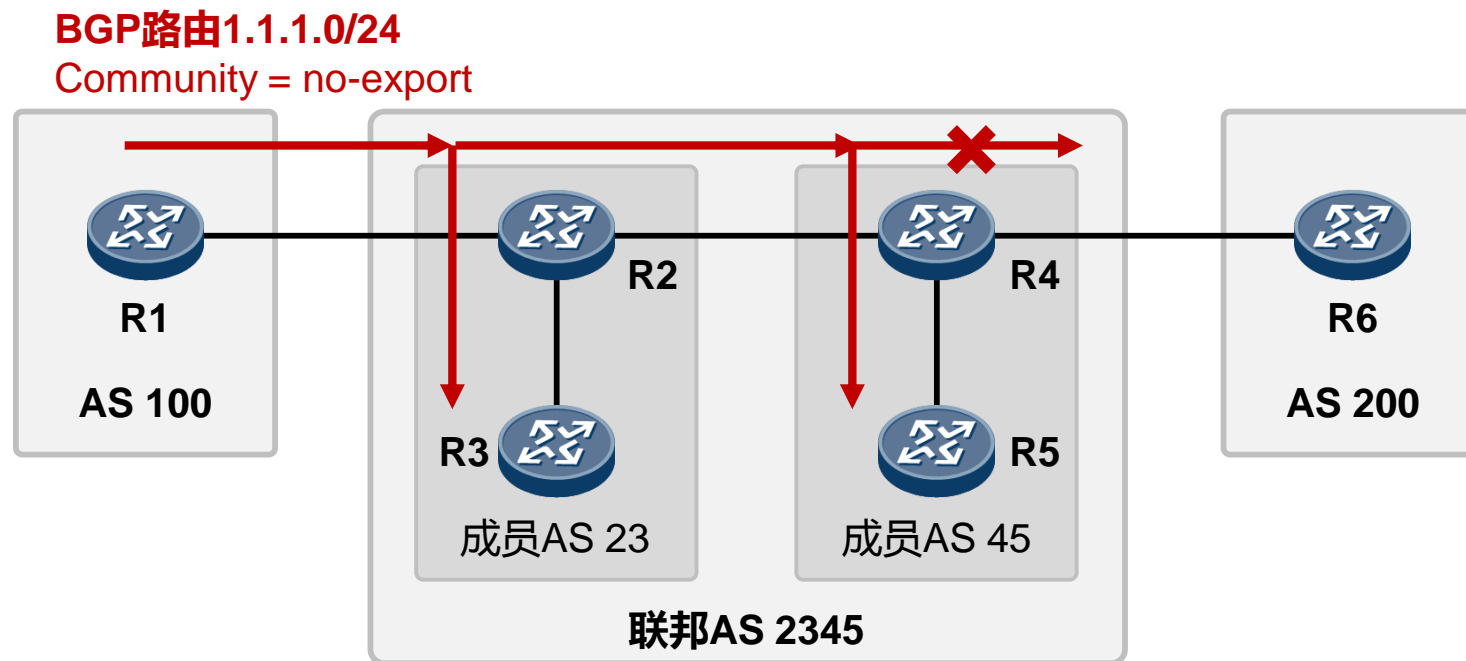
## 众所周知的Community值：no-advertise



如果路由器收到一条携带Community属性的BGP路由，且其中包含no-advertise属性值，那么该条路由将不能再传递给任何其他BGP对等体。

在本例中，R2不会将1.1.1.0/24路由再通告给自己任何BGP对等体，也就是路由只能自己用。

## 众所周知的Community值：no-export

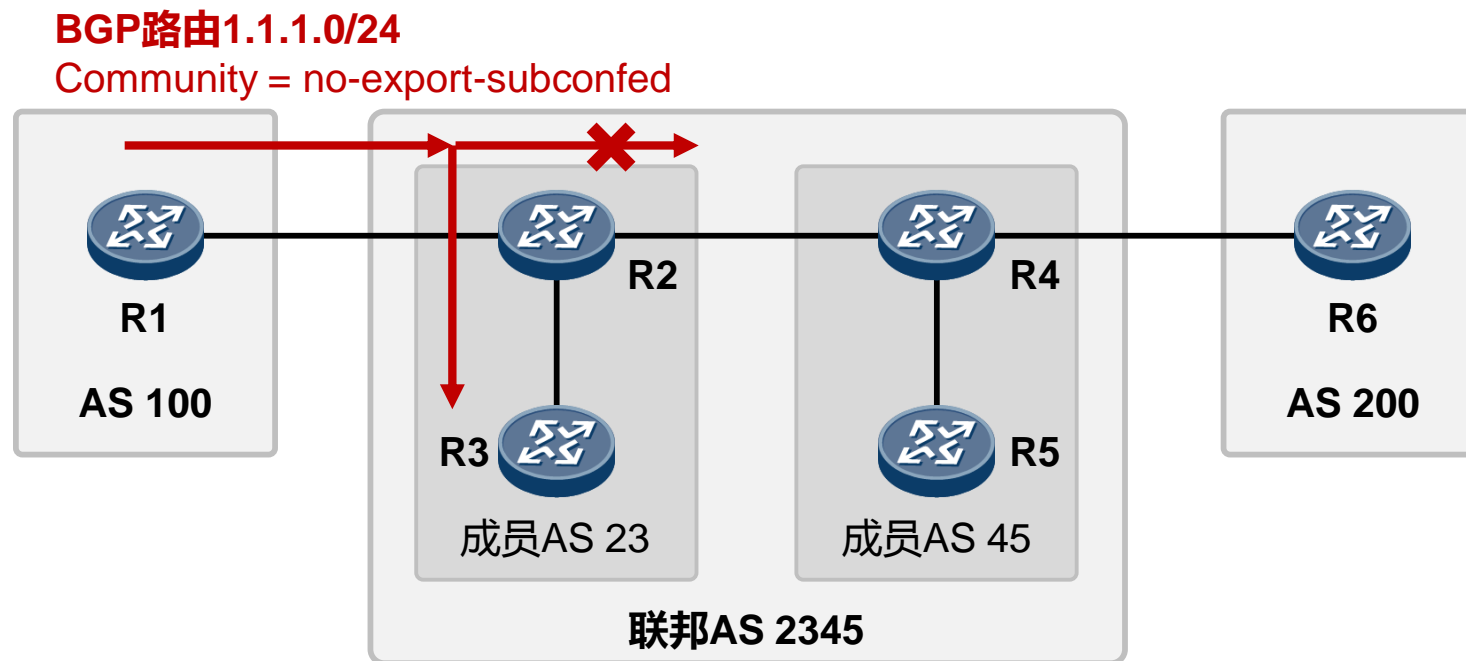


如果路由器收到一条携带Community属性的BGP路由，且其中包含no-export属性值，那么该条路由将不能再传递给任何其他EBGP对等体（联邦EBGP对等体除外）。

在本例中，R2不会将1.1.1.0/24路由再通告给EBGP对等体R6，但是依然会传递给联邦EBGP对等体R4。

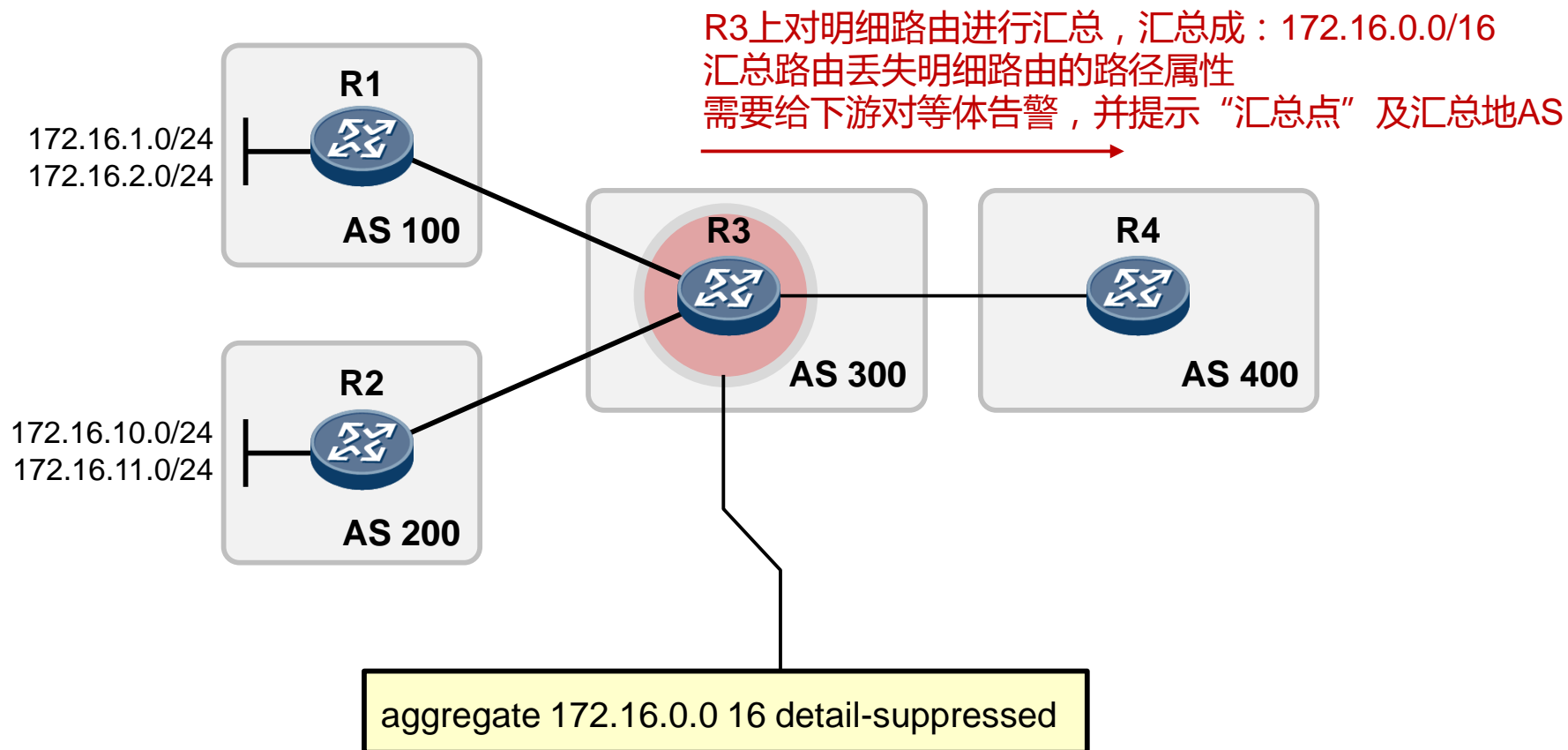


## 众所周知的Community值：no-export-subconfed



如果路由器收到一条携带Community属性的BGP路由，且其中包含no-export-subconfed属性值，那么该条路由将不能再传递给任何其他EBGP对等体（包括联邦EBGP对等体）。  
在本例中，R2不会将1.1.1.0/24路由再通告给EBGP对等体R6，以及联邦EBGP对等体R4。

# Atomic\_Aggregate及aggregator



# Atomic\_Aggregate及aggregator

## R4#display bgp routing-table 172.16.0.0

BGP local router ID : 4.4.4.4

Local AS number : 400

Paths: 1 available, 1 best, 1 select

BGP routing table entry information of 172.16.0.0/16:

From: 10.1.34.3 (3.3.3.3)

Route Duration: 00h00m21s

Direct Out-interface: GigabitEthernet0/0/0

Original nexthop: 10.1.34.3

Qos information : 0x0

AS-path 300, origin igp, pref-val 0, valid, external, best, select, active, pre 255

**Aggregator: AS 300, Aggregator ID 3.3.3.3, Atomic-aggregate**

Not advertised to any peer yet

# Atomic\_Aggregate及aggregator

- [-] UPDATE Message
  - Marker: 16 bytes
  - Length: 63 bytes
  - Type: UPDATE Message (2)
  - Unfeasible routes length: 0 bytes
  - Total path attribute length: 37 bytes
- [-] Path attributes
  - [+] ORIGIN: IGP (4 bytes)
  - [+] AS\_PATH: 300 (7 bytes)
  - [+] NEXT\_HOP: 10.1.34.3 (7 bytes)
  - [+] MULTI\_EXIT\_DISC: 0 (7 bytes)
  - [-] ATOMIC\_AGGREGATE (3 bytes)
    - [+] Flags: 0x40 (well-known, Transitive, complete)
    - Type code: ATOMIC\_AGGREGATE (6)
    - Length: 0 bytes
  - [-] AGGREGATOR: AS: 300 origin: 3.3.3.3 (9 bytes)
    - [+] Flags: 0xc0 (Optional, Transitive, Complete)
    - Type code: AGGREGATOR (7)
    - Length: 6 bytes
    - Aggregator AS: 300
    - Aggregator origin: 3.3.3.3 (3.3.3.3)
- [-] Network layer reachability information: 3 bytes
  - [+] 172.16.0.0/16

# Thank you

[www.huawei.com](http://www.huawei.com)

Copyright ©2014 Huawei Technologies Co.,Ltd. All Rights Reserved.

The information contained in this document is for reference purpose only, and is subject to change or withdrawal according to specific customer requirements and conditions.

©2014 华为技术有限公司 版权所有  
本资料仅供参考，不构成任何承诺及保证