

# BGP NEXT\_HOP Attribute

Shishio Tsuchiya

[shtsuchi@arista.com](mailto:shtsuchi@arista.com)

## はじめに

- IRSあとにRFC4271 5.1.3を読み直してみた
  - やったことあるなー/よく使うなー
  - を皆さんに共有したいと思います。
- 
- つまり30分 BGP NEXTHOPの話だけです

# RFC4271 5.1.3

## • 5.1.3. NEXT\_HOP

に全て例が書かれている

Withdrawn Routes Length (2 octets)
Withdrawn Routes (variable)
Total Path Attribute Length (2 octets)
Path Attributes (variable)
Network Layer Reachability Information (variable)

### 5.1.3. NEXT\_HOP

The NEXT\_HOP is a well-known mandatory attribute that defines the IP address of the router that SHOULD be used as the next hop to the destinations listed in the UPDATE message. The NEXT\_HOP attribute is calculated as follows:

- 1) When sending a message to an internal peer, if the route is not locally originated, the BGP speaker SHOULD NOT modify the NEXT\_HOP attribute unless it has been explicitly configured to announce its own IP address as the NEXT\_HOP. When announcing a

Rakhter, et al.

Standards Track

[Page 24]

RFC 4271

BGP-4

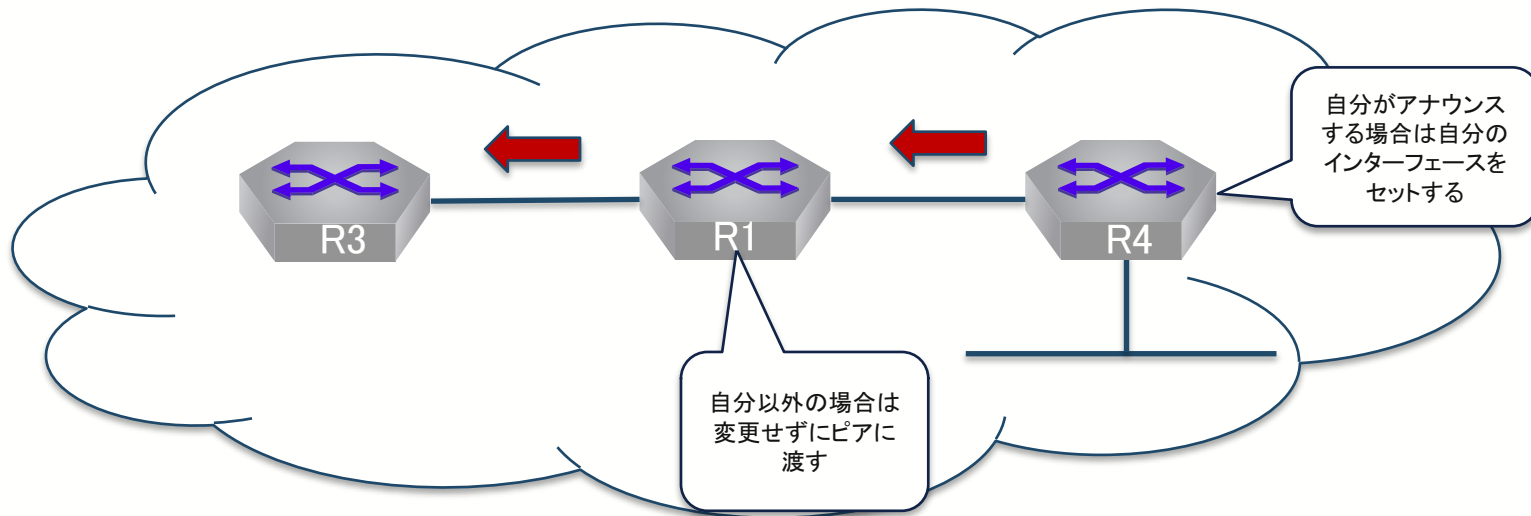
January 2004

locally-originated route to an internal peer, the BGP speaker SHOULD use the interface address of the router through which the announced network is reachable for the speaker as the NEXT\_HOP. If the route is directly connected to the speaker, or if the interface address of the router through which the announced network is reachable for the speaker is the internal peer's address, then the BGP speaker SHOULD use its own IP address for the NEXT\_HOP attribute (the address of the interface that is used to reach the peer).

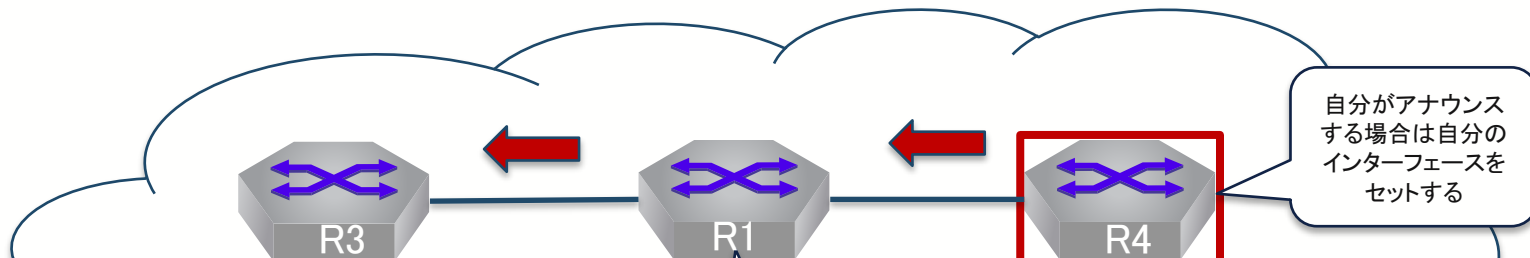
- 2) When sending a message to an external peer, X, and the peer is one IP hop away from the speaker:

- If the route being announced was learned from an internal peer or is locally originated, the BGP speaker can use an interface address of the internal peer router (or the internal router) through which the announced network is reachable for the speaker for the NEXT\_HOP attribute, provided that peer X shares a common subnet with this address. This is a form of "third party" NEXT\_HOP attribute.
- Otherwise, if the route being announced was learned from an external peer, the speaker can use an IP address of any adjacent router (known from the received NEXT\_HOP attribute)

## 5.1.3(1) iBGP



## 5.1.3(1) iBGP



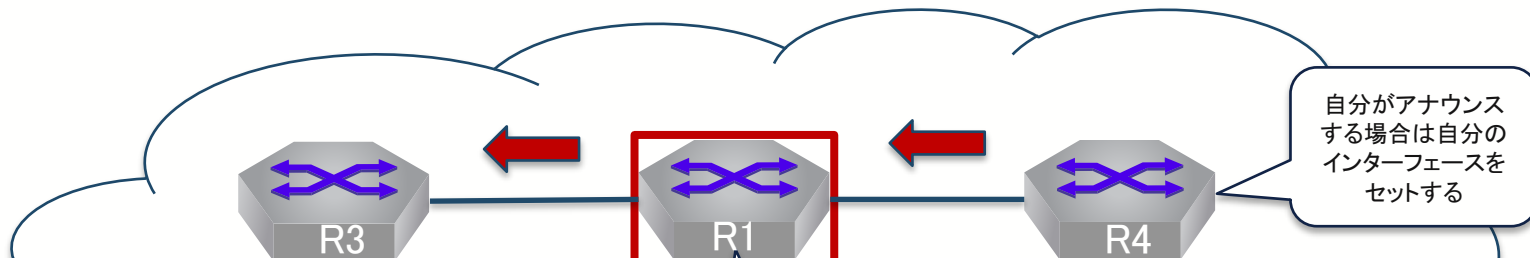
```
R4#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.4, local AS number 65500
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop           Metric  LocPref  Weight  Path
* >  10.100.1.0/24   10.255.255.1       0       100      0       i
* >  10.100.4.0/24  -                   1        0       -       i

R4#show ip bgp 10.100.4.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.4, local AS number 65500
BGP routing table entry for 10.100.4.0/24
  Paths: 1 available
    Local
  - from - (10.255.255.4)
    Origin IGP, metric 1, localpref 0, IGP metric -, weight -, received 00:11:56 ago, valid, local, best, redistributed (Connected)
    Rx SAFI: Unicast

R4#
```

## 5.1.3(1) iBGP



```
R1#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.1, local AS number 65500
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop
```

	Network	Next Hop	Metric	LocPref	Weight	Path
* >	10.100.1.0/24	-	1	0	-	i
* >	<b>10.100.4.0/24</b>	<b>10.255.255.4</b>	<b>0</b>	<b>100</b>	<b>0</b>	<b>i</b>

```
R1#show ip bgp 10.100.4.0/24
```

```
BGP routing table information for VRF default
Router identifier 10.255.255.1, local AS number 65500
```

```
BGP routing table entry for 10.100.4.0/24
```

```
Paths: 1 available
```

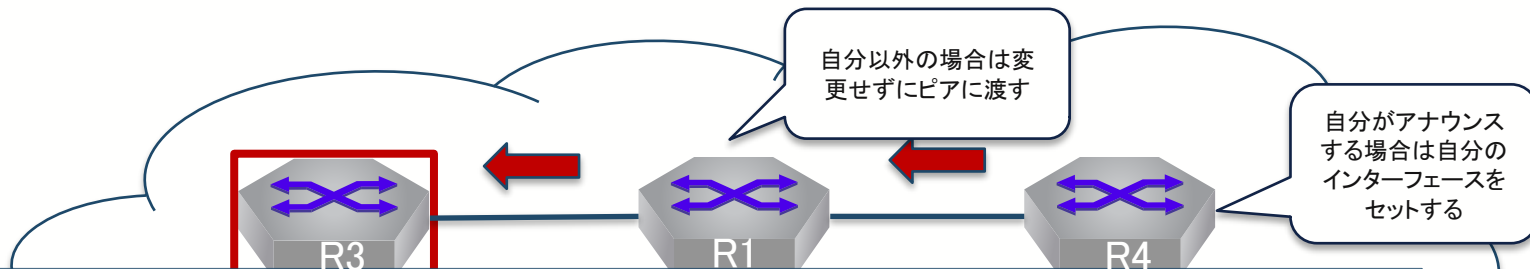
```
Local (Received from a RR-client)
```

```
10.255.255.4 from 10.255.255.4 (10.255.255.4)
```

```
Origin IGP, metric 0, localpref 100, IGP metric 20, weight 0, received 00:13:30 ago, valid, internal, best
```

```
Rx SAFI: Unicast
```

## 5.1.3(1) iBGP

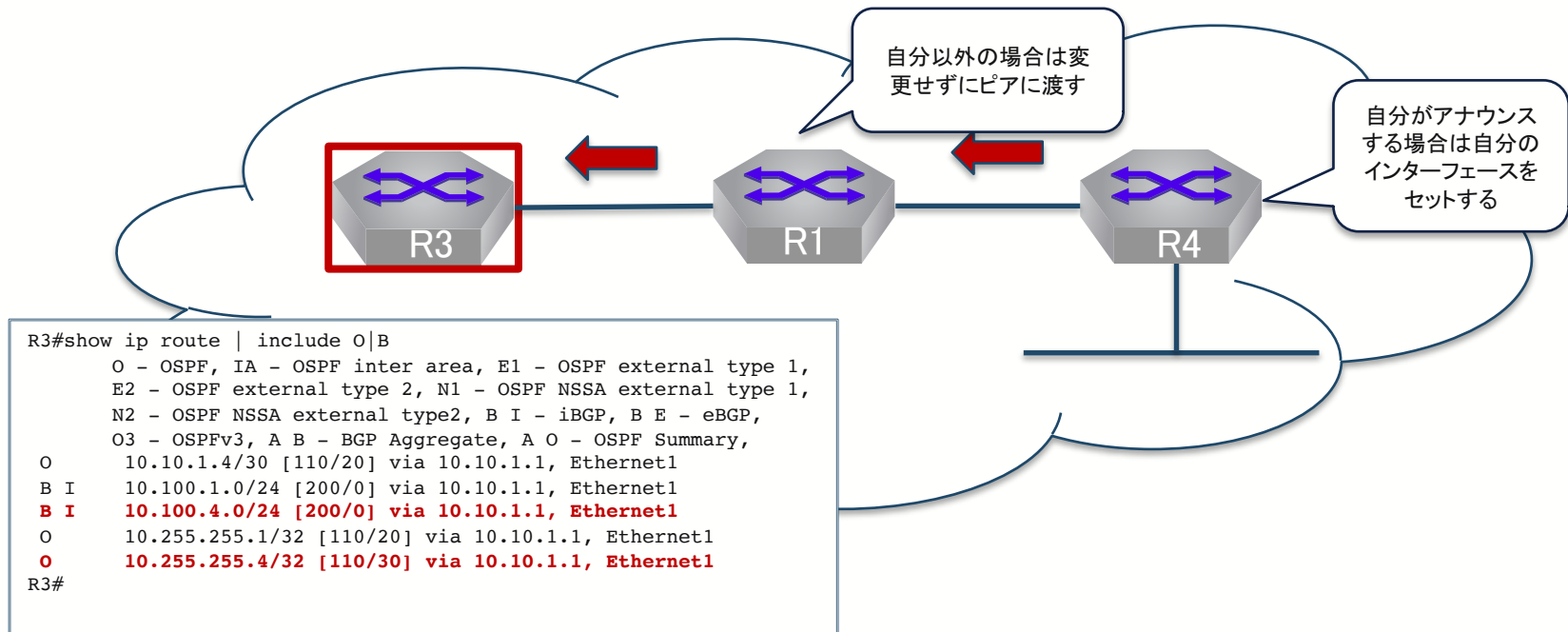


```
R3#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 65500
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop           Metric  LocPref  Weight  Path
* >  10.100.1.0/24    10.255.255.1       0       100     0       i
* >  10.100.4.0/24    10.255.255.4       0       100     0       i Or-ID: 10.255.255.4 C-LST: 10.255.255.1

R3#show ip bgp 10.100.4.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 65500
BGP routing table entry for 10.100.4.0/24
Paths: 1 available
Local
  10.255.255.4 from 10.255.255.1 (10.255.255.1)
  Origin IGP, metric 0, localpref 100, IGP metric 30, weight 0, received 00:17:39 ago, valid, internal, best
  Originator: 10.255.255.4, Cluster list: 10.255.255.1
  Rx SAFI: Unicast
```

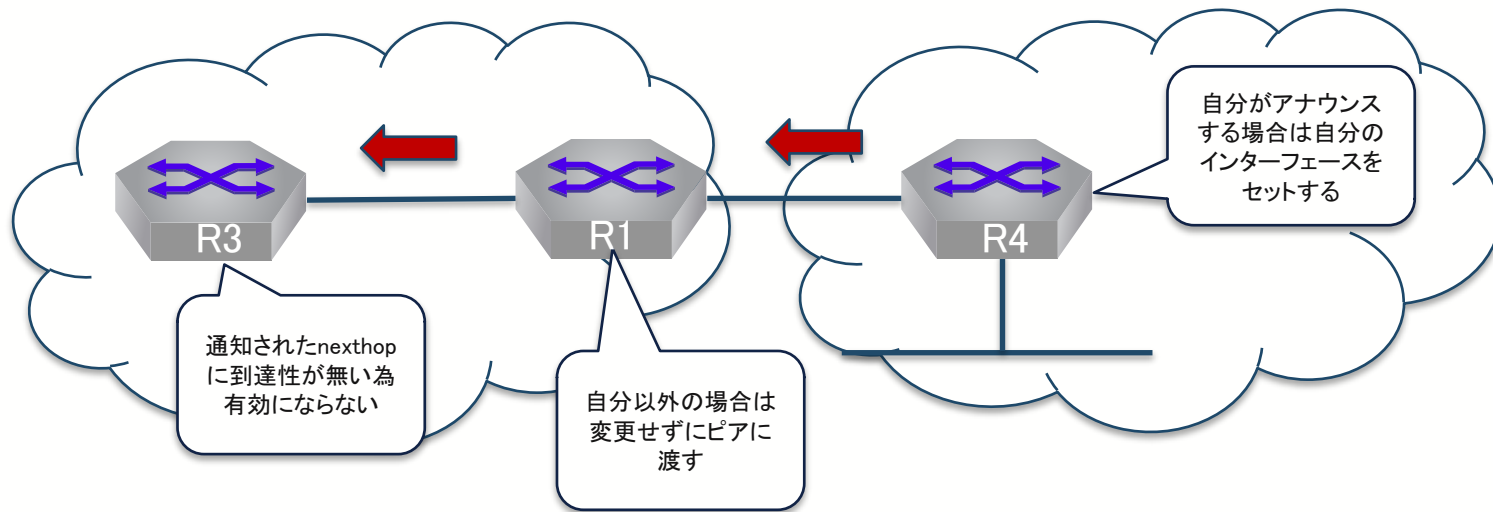
## 5.1.3(1) iBGP



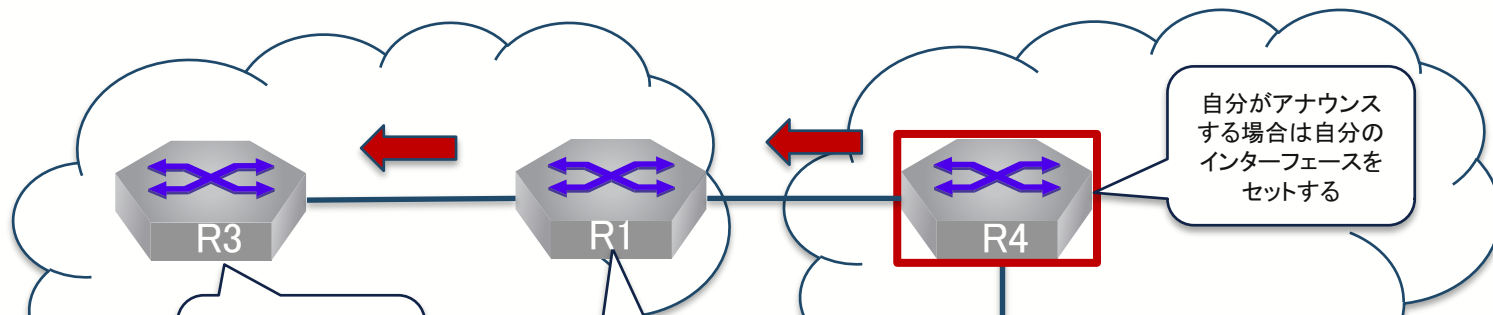
- BGP next hopをルーティングテーブルより解決して、ルーティングテーブルにインストール



## 5.1.3(1) iBGP



## 5.1.3(1) iBGP

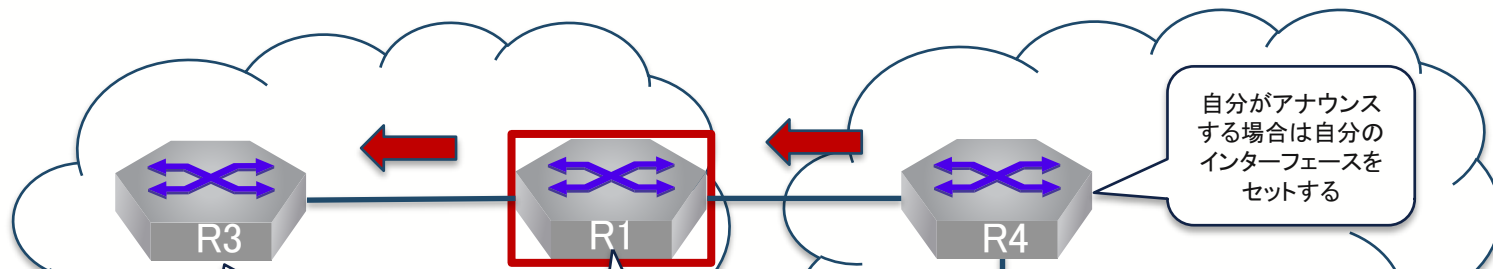


```
R4#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.4, local AS number 65504
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop           Metric  LocPref Weight Path
   * > 10.100.1.0/24  10.10.1.5         0       100    0   65500 i
   * > 10.100.4.0/24  -                  1       0     -    i

R4#show ip bgp 10.100.4.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.4, local AS number 65504
BGP routing table entry for 10.100.4.0/24
  Paths: 1 available
    Local
  - from - (10.255.255.4)
    Origin IGP, metric 1, localpref 0, IGP metric -, weight -, received 00:53:59 ago, valid, local, best, redistributed (Connected)
    Rx SAFI: Unicast
```

## 5.1.3(1) iBGP

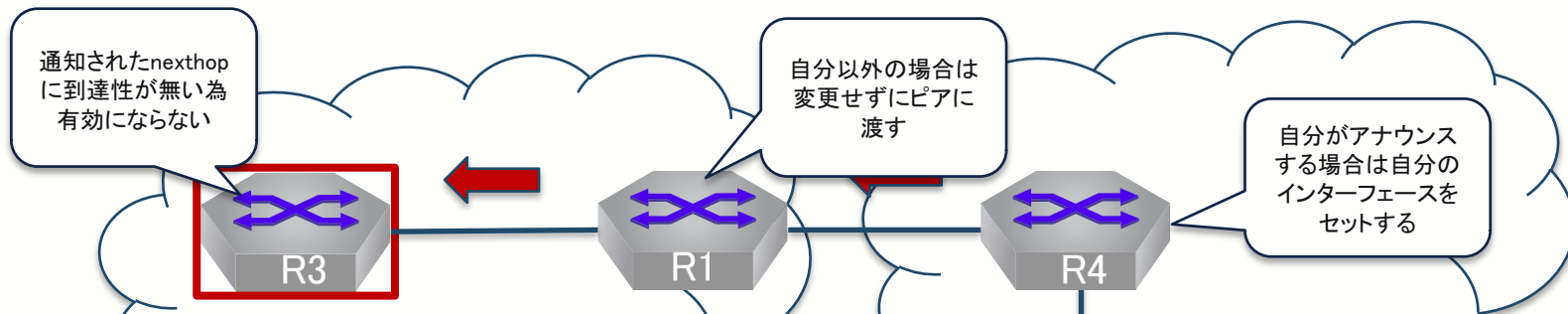


```
R1#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.1, local AS number 65500
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop           Metric  LocPref Weight Path
* >  10.100.1.0/24   -                  1       0      -    i
* >  10.100.4.0/24  10.10.1.6         0       100    0    65504 i

R1#show ip bgp 10.100.4.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.1, local AS number 65500
BGP routing table entry for 10.100.4.0/24
Paths: 1 available
 65504
 10.10.1.6 from 10.10.1.6 (10.255.255.4)
   Origin IGP, metric 0, localpref 100, IGP metric 1, weight 0, received 00:25:17 ago, valid, external, best
   Rx SAFI: Unicast
R1#
```

## 5.1.3(1) iBGP



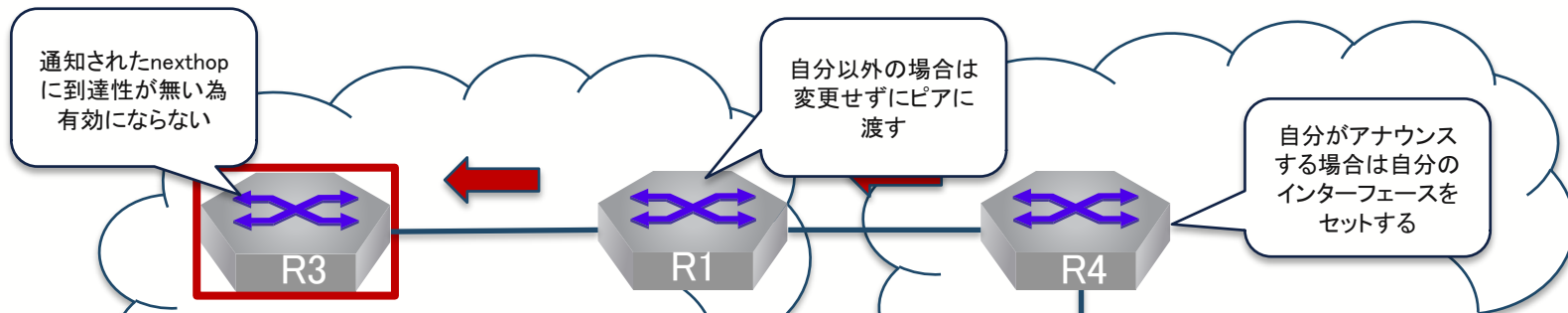
```
R3#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 65500
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop           Metric  LocPref Weight Path
* >  10.100.1.0/24   10.255.255.1       0      100    0    i
* >  10.100.4.0/24  10.10.1.6          0      100    0    65504 i

R3#show ip bgp 10.100.4.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 65500
BGP routing table entry for 10.100.4.0/24
Paths: 1 available
 65504
 10.10.1.6 from 10.255.255.1 (10.255.255.1)
Origin IGP, metric 0, localpref 100, IGP metric 0, weight 0, received 00:55:11 ago, invalid, internal
Rx SAFI: Unicast

R3#
```

## 5.1.3(1) iBGP

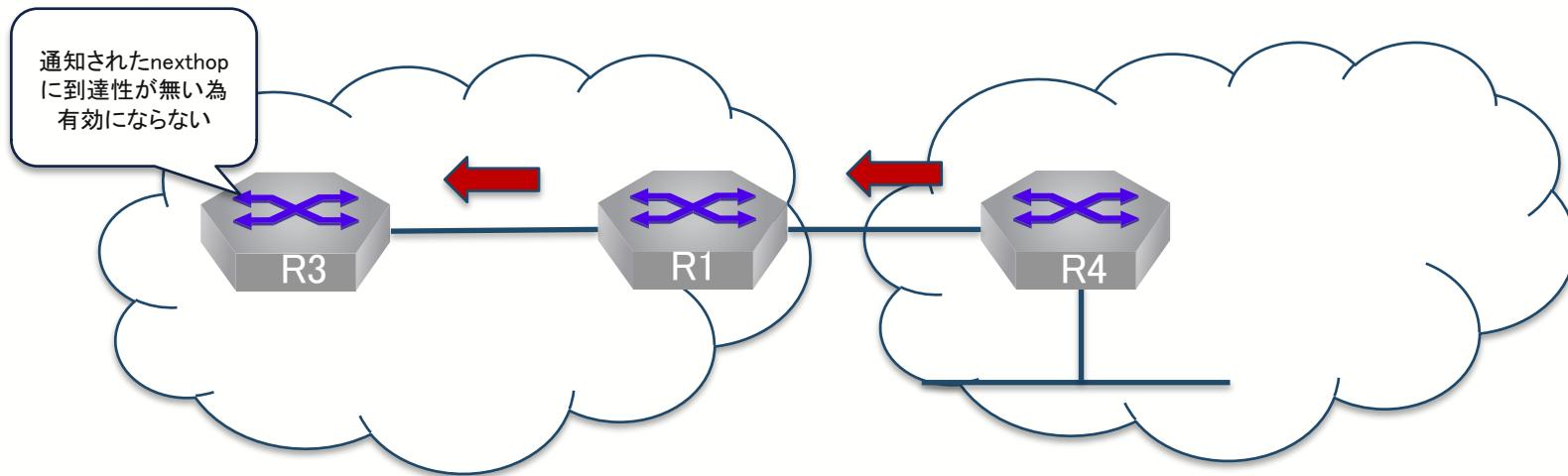


```
R3#show ip route bgp

VRF: default
Codes: C - connected, S - static, K - kernel,
O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
N2 - OSPF NSSA external type2, B I - iBGP, B E - eBGP,
R - RIP, I L1 - IS-IS level 1, I L2 - IS-IS level 2,
O3 - OSPFv3, A B - BGP Aggregate, A O - OSPF Summary,
NG - Nexthop Group Static Route, V - VXLAN Control Service,
DH - DHCP client installed default route, M - Martian

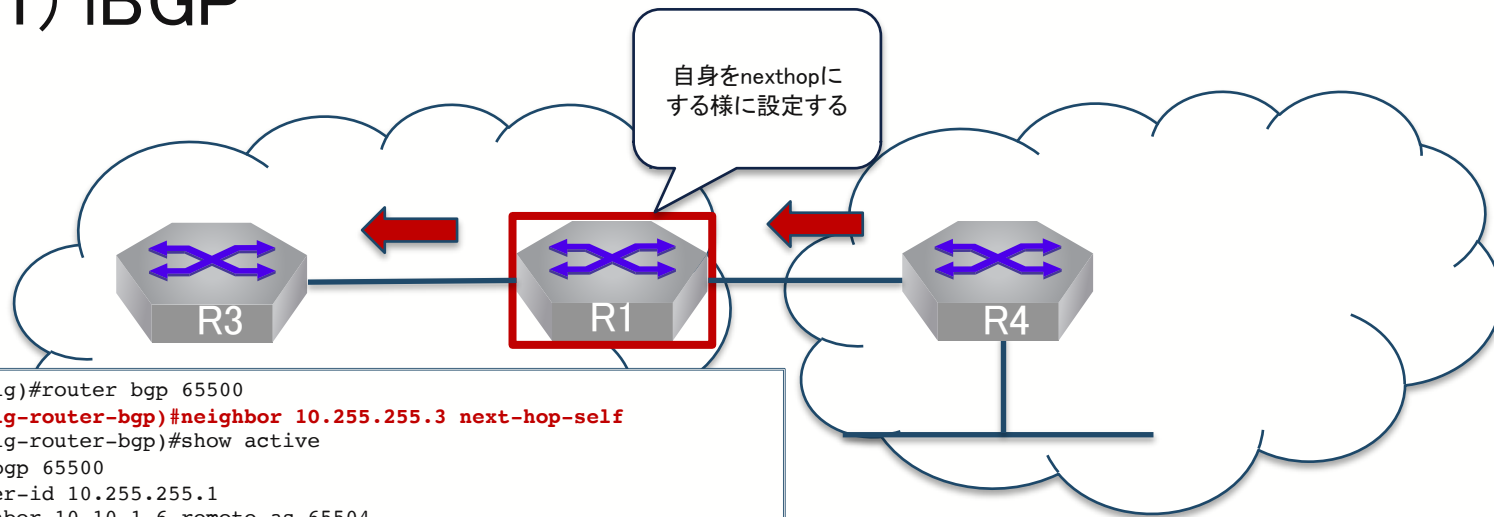
B I    10.100.1.0/24 [200/0] via 10.10.1.1, Ethernet1
```

## 5.1.3(1) iBGP



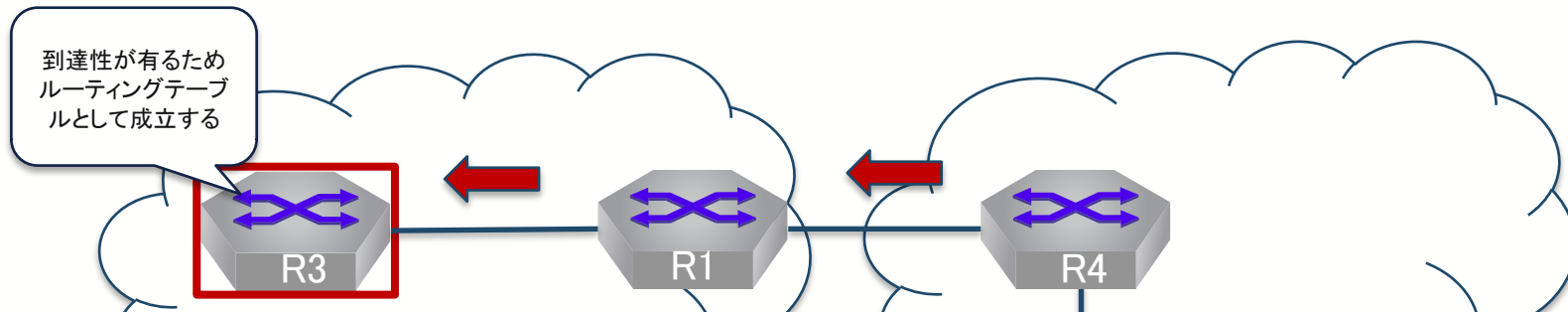
- だからnext-hop-selfしてね

## 5.1.3(1) iBGP



```
R1(config)#router bgp 65500
R1(config-router-bgp)#neighbor 10.255.255.3 next-hop-self
R1(config-router-bgp)#show active
router bgp 65500
  router-id 10.255.255.1
  neighbor 10.10.1.6 remote-as 65504
  neighbor 10.10.1.6 maximum-routes 12000
  neighbor 10.255.255.3 remote-as 65500
  neighbor 10.255.255.3 next-hop-self
  neighbor 10.255.255.3 update-source Loopback0
  neighbor 10.255.255.3 route-reflector-client
  neighbor 10.255.255.3 maximum-routes 12000
  network 10.100.1.0/24
```

## 5.1.3(1) iBGP



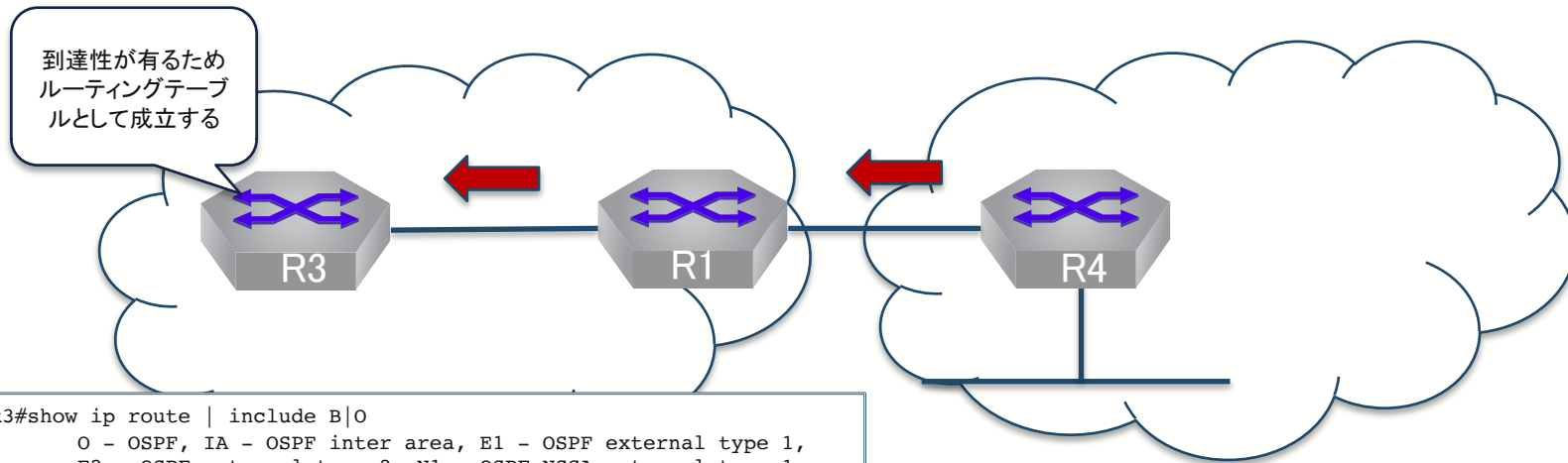
```
R3#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 65500
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                   S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop           Metric  LocPref  Weight  Path
* >  10.100.1.0/24   10.255.255.1       0       100     0       i
* >  10.100.4.0/24   10.255.255.1       0       100     0       65504 i

R3#show ip bgp 10.100.4.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 65500
BGP routing table entry for 10.100.4.0/24
Paths: 1 available
65504
  10.255.255.1 from 10.255.255.1 (10.255.255.1)
    Origin IGP, metric 0, localpref 100, IGP metric 20, weight 0, received 01:03:27 ago, valid,
    internal, best
    Rx SAFI: Unicast
```



## 5.1.3(1) iBGP

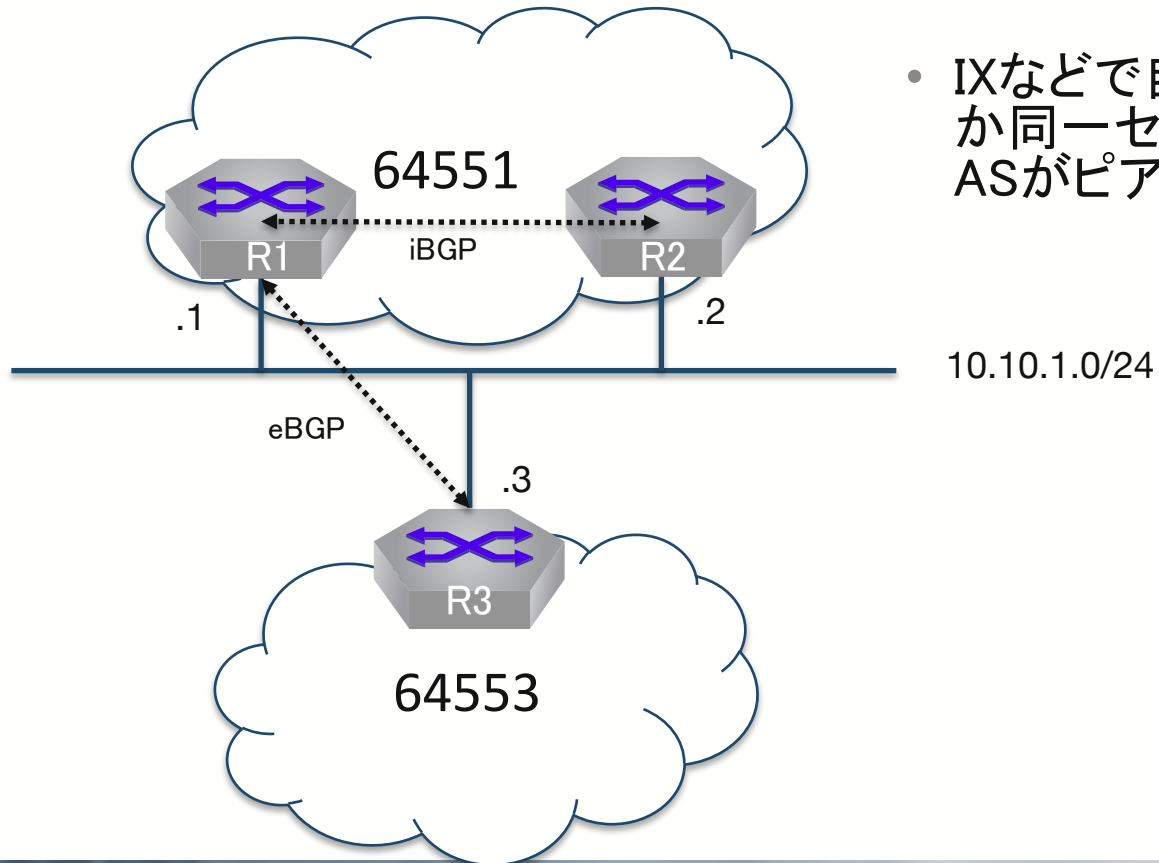


```
R3#show ip route | include B|O
O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
N2 - OSPF NSSA external type2, B I - iBGP, B E - eBGP,
O3 - OSPFv3, A B - BGP Aggregate, A O - OSPF Summary,
B I 10.100.1.0/24 [200/0] via 10.10.1.1, Ethernet1
B I 10.100.4.0/24 [200/0] via 10.10.1.1, Ethernet1
O 10.255.255.1/32 [110/20] via 10.10.1.1, Ethernet1
R3#
```

## とにかく基本は

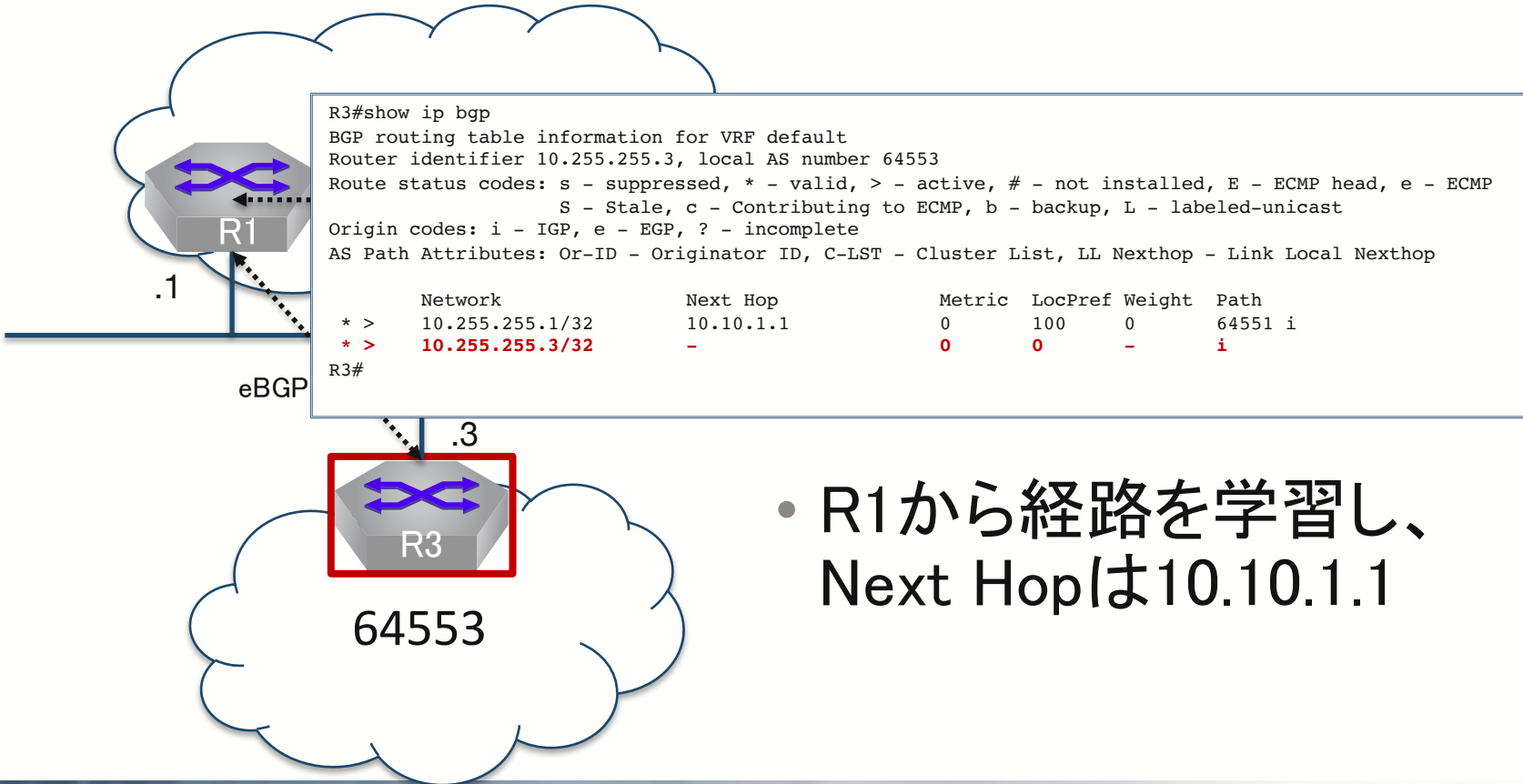
- iBGPはNEXTHOPを保持する
- eBGPはNEXTHOPを自分に変更する

## 5.1.3(2) 3<sup>rd</sup> party next hop



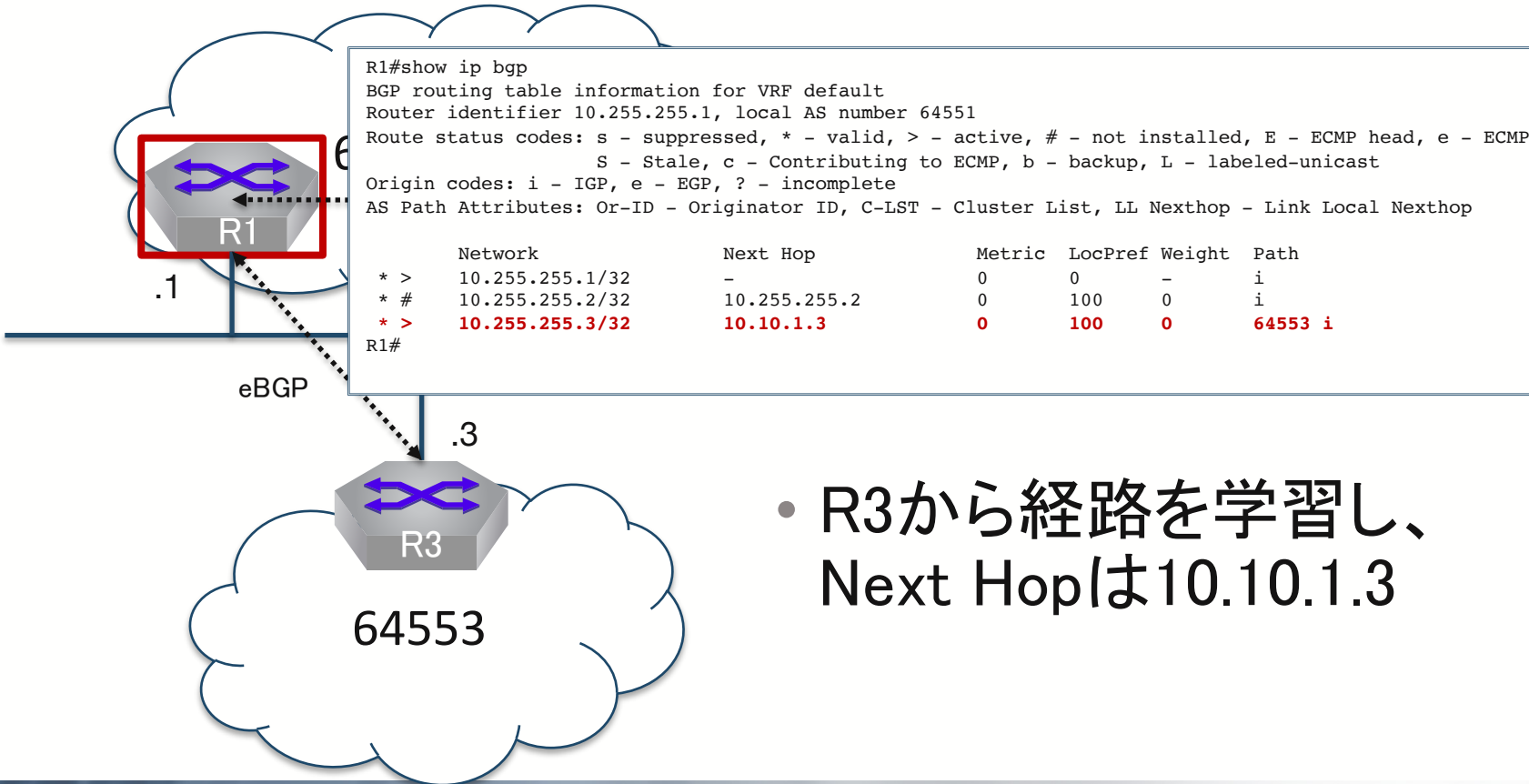
- IXなどで自ASノードがいくつか同一セグメントにあり、外部ASがピアリングしてるケース

## 5.1.3(2) 3<sup>rd</sup> party next hop



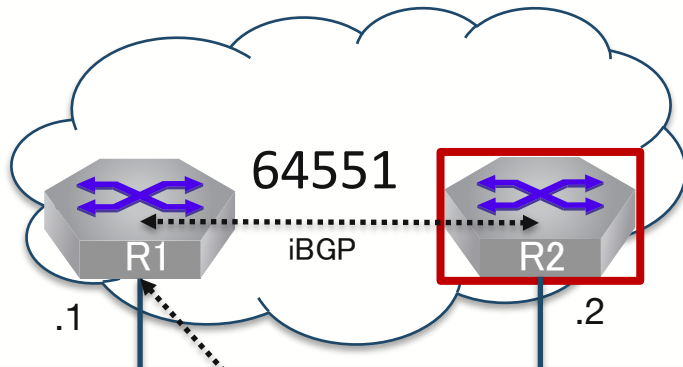
- R1から経路を学習し、Next Hopは10.10.1.1

## 5.1.3(2) 3<sup>rd</sup> party next hop



- R3から経路を学習し、Next Hopは10.10.1.3

## 5.1.3(2) 3<sup>rd</sup> party next hop

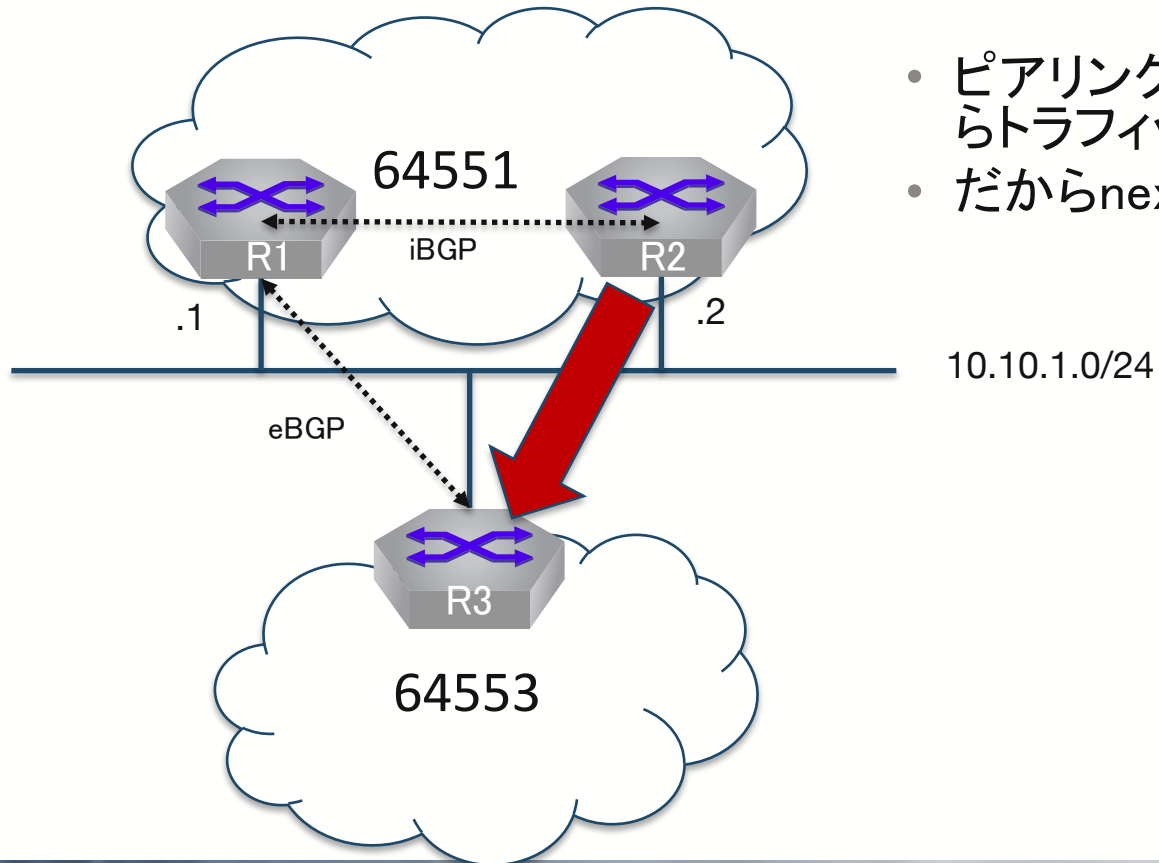


- R1から経路を学習し、なおかつ到達性のあるインターフェースがnexthopであるために経路として成立する

```
R2#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.2, local AS number 64551
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

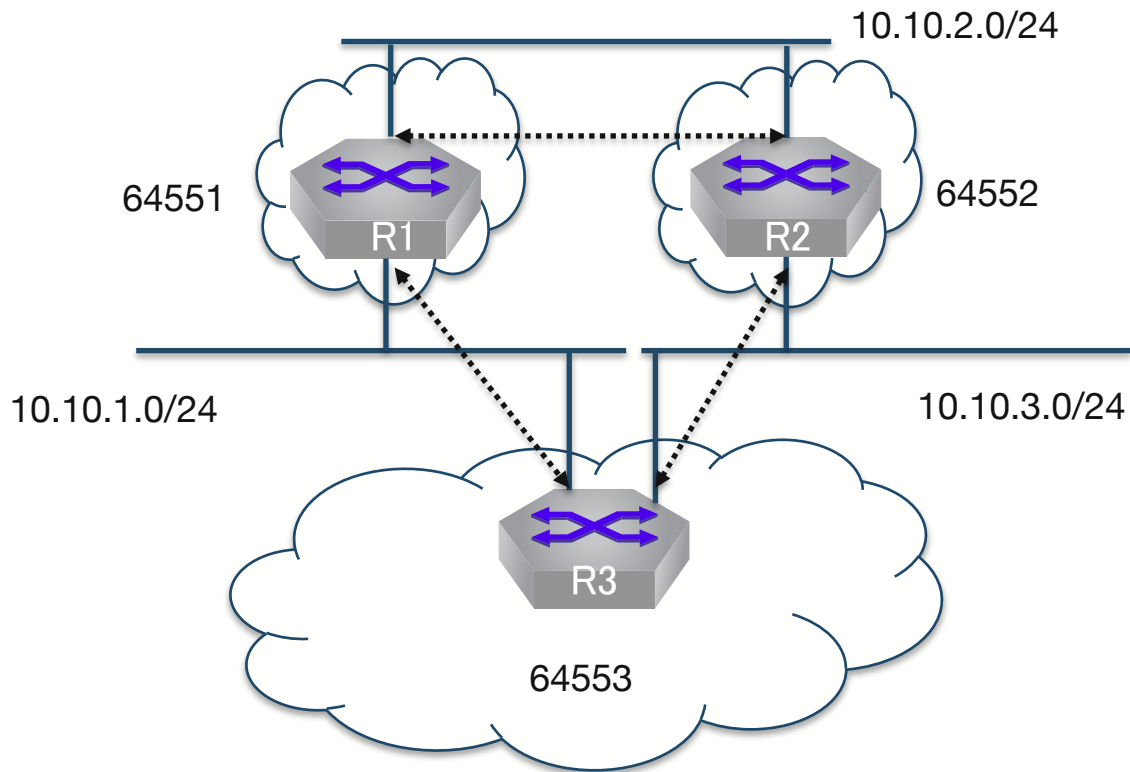
   Network          Next Hop          Metric  LocPref  Weight  Path
* #  10.255.255.1/32   10.255.255.1     0       100     0       i
* >  10.255.255.2/32   -                 0        0       -       i
* >  10.255.255.3/32  10.10.1.3        0       100     0       64553 i
R2#
```

## 5.1.3(2) 3<sup>rd</sup> party next hop



- ピアリングしていないノードからトラフィックが来る
- だからnext hop selfしてね

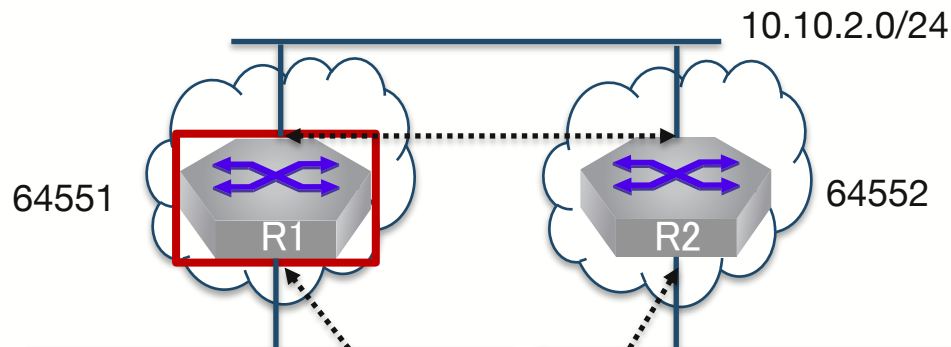
## 5.1.3(2) 3<sup>rd</sup> party nexthop



- R1とR2は10.10.2.0/24
- R2とR3は10.10.3.0/24
- R1とR3は10.10.1.0/24でそれぞれピアリング



## 5.1.3(2) 3<sup>rd</sup> party next hop



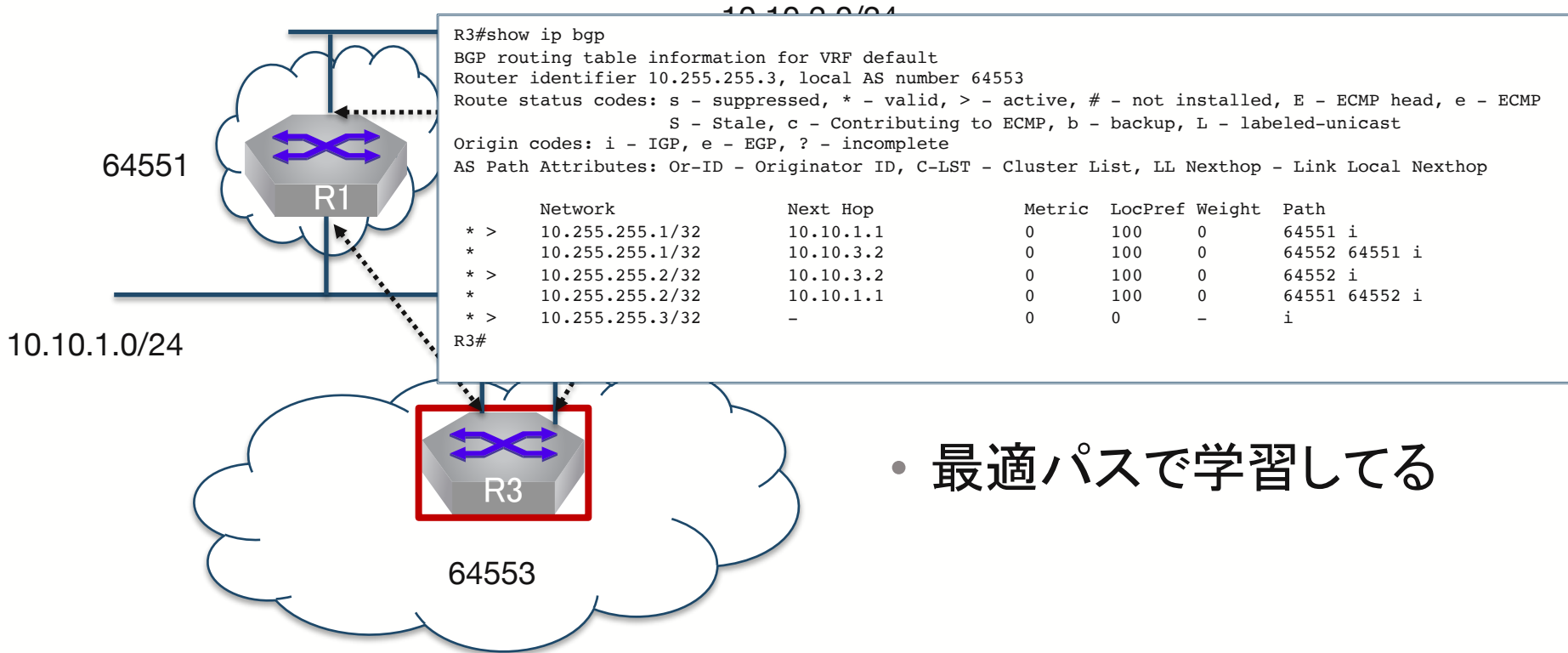
- ピアリングしてるノードが他ともピアリングしてるのがわかる

10.10.1.0

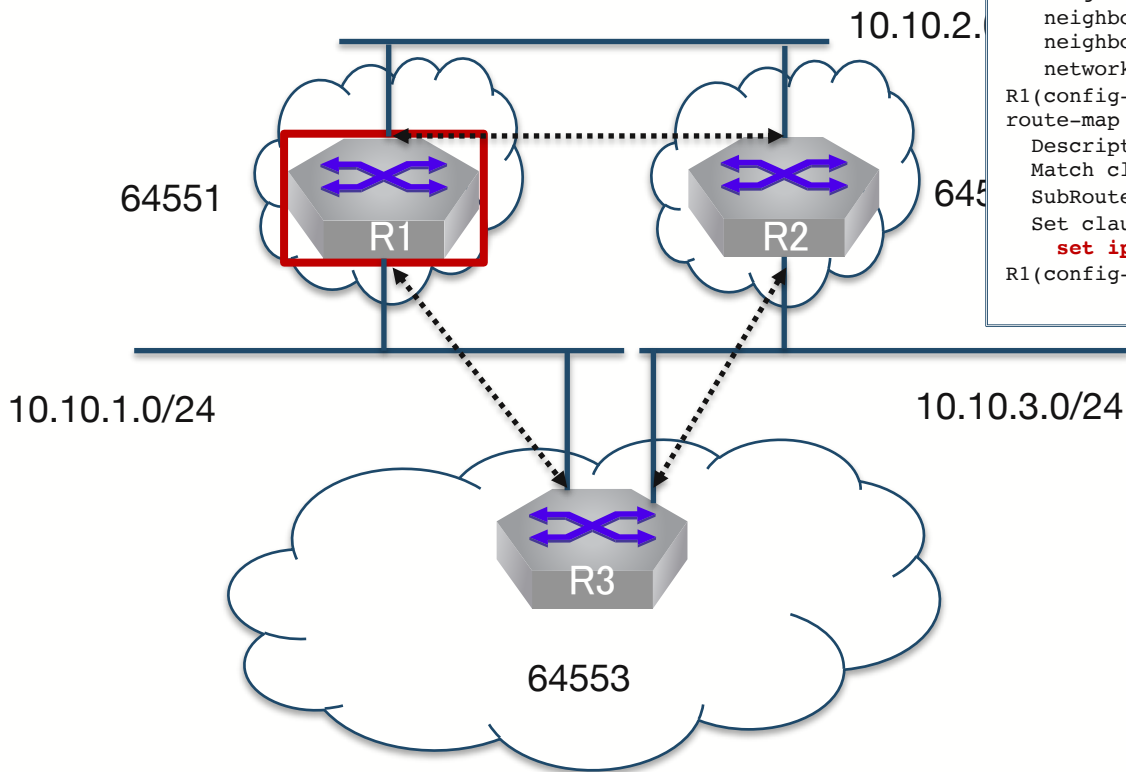
```
R1#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.1, local AS number 64551
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

   Network          Next Hop          Metric  LocPref  Weight  Path
* >  10.255.255.1/32  -                  0        0        -       i
* >  10.255.255.2/32  10.10.2.2         0        100       0       64552 i
*   10.255.255.2/32  10.10.1.3         0        100       0       64553 64552 i
* >  10.255.255.3/32  10.10.1.3         0        100       0       64553 i
*   10.255.255.3/32  10.10.2.2         0        100       0       64552 64553 i
R1#
```

## 5.1.3(2) 3<sup>rd</sup> party next hop



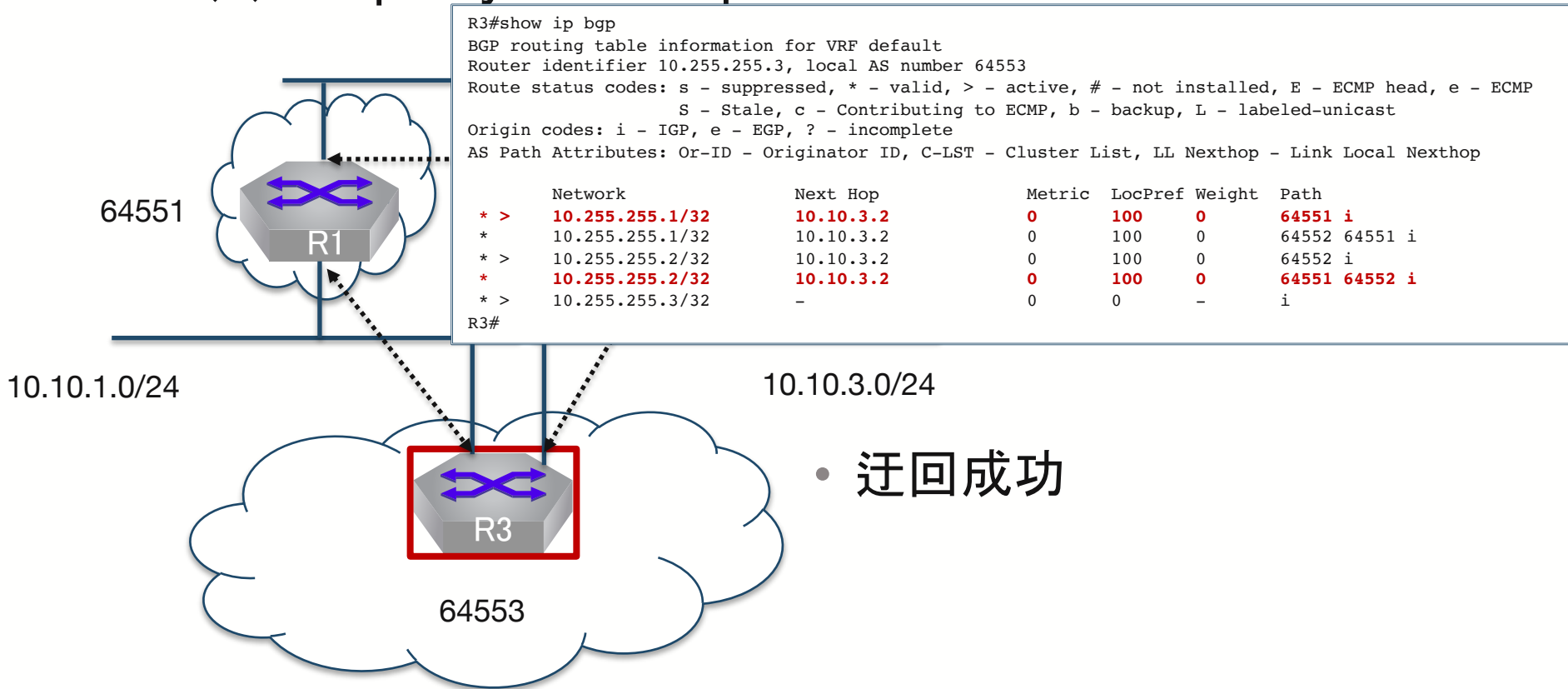
## 5.1.3(2) 3<sup>rd</sup> party next hop



```
R1(config-router-bgp)#show active
router bgp 64551
  router-id 10.255.255.1
  neighbor 10.10.1.3 remote-as 64553
  neighbor 10.10.1.3 route-map 3rdpartyNexthop out
  neighbor 10.10.1.3 maximum-routes 12000
  neighbor 10.10.2.2 remote-as 64552
  neighbor 10.10.2.2 maximum-routes 12000
  network 10.255.255.1/32
R1(config-router-bgp)#show route-map
route-map 3rdpartyNexthop permit 10
  Description:
  Match clauses:
  SubRouteMap:
  Set clauses:
    set ip next-hop 10.10.3.2
R1(config-router-bgp)#
```

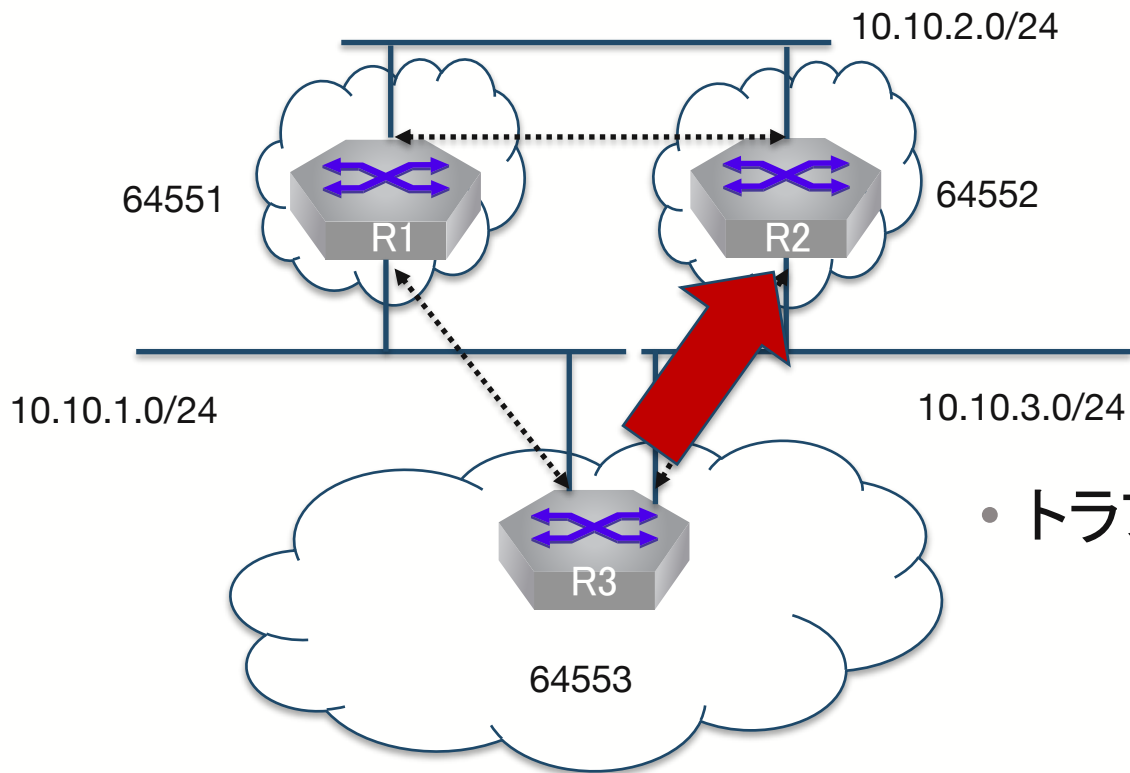
- next hopを変更しトラフィックをR2に寄せる

# 5.1.3(2) 3<sup>rd</sup> party next hop



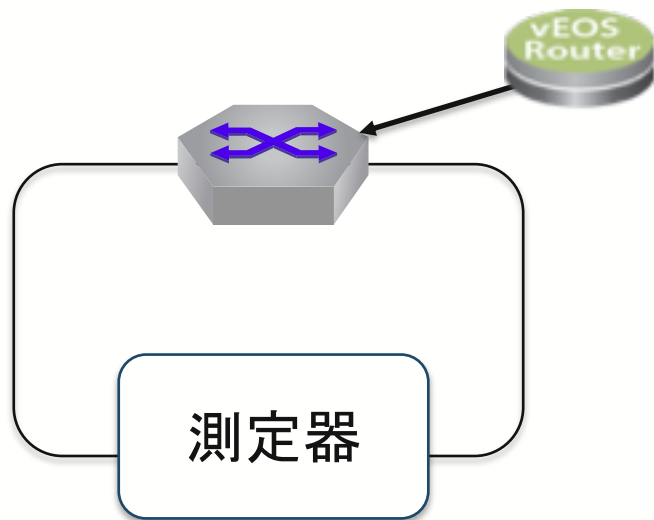
• 迂回成功

## 5.1.3(2) 3<sup>rd</sup> party next hop



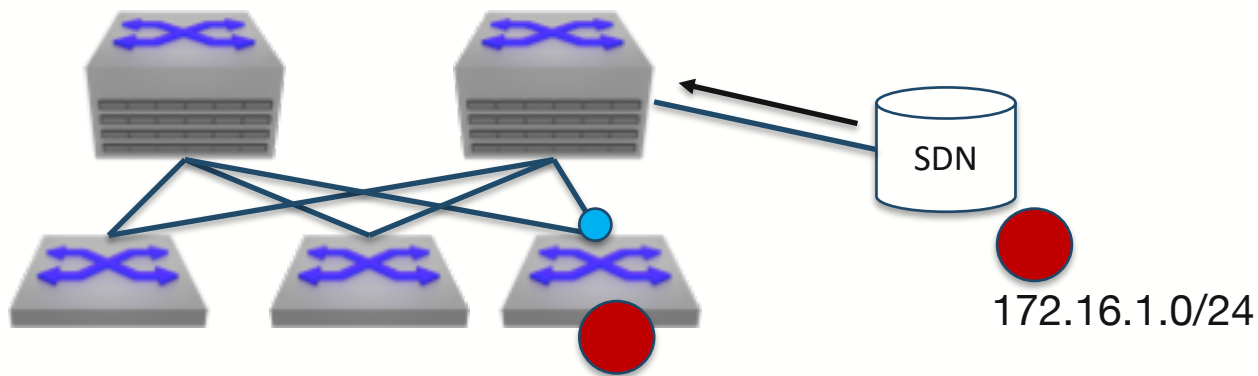
- トラフィックは全て寄せられる

# いつ使うのか？ 3<sup>rd</sup> Party next hop



- MRTなどで仮想ルータにフルルートを読み込みます
- 全てNext Hopを測定器にしてフルルートでのパフォーマンス測定
- 安価にフルルートでのパフォーマンス測定が可能

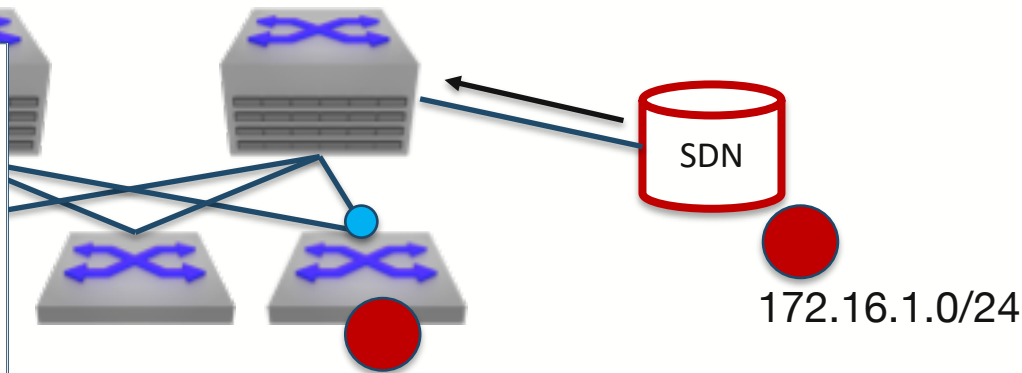
# いつ使うのか？ 3<sup>rd</sup> Party next hop BGP SDNデータセンター



- レイヤー3データセンターではeBGPで構成する
- BGPコントローラーで各ルートをコントロールするのにnext hopを操作

# いつ使うのか？ 3<sup>rd</sup> Party next hop BGP SDNデータセンター

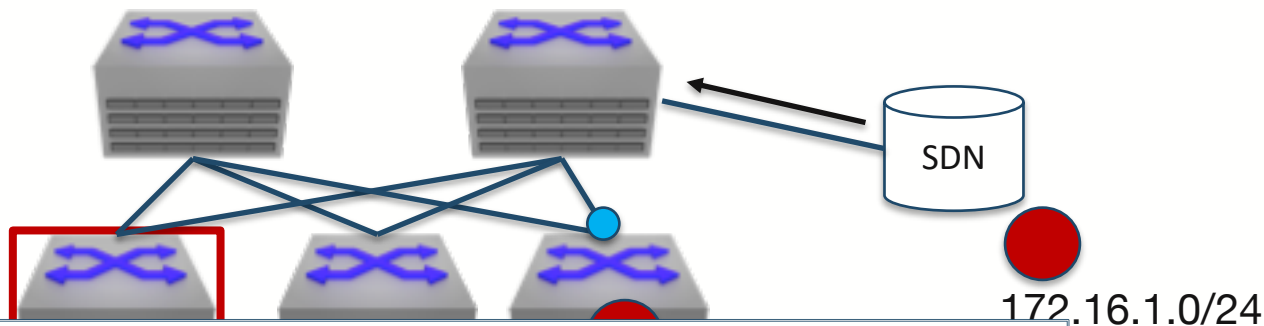
```
SDN(config)#router bgp 64550
SDN(config-router-bgp)#show active
router bgp 64550
  neighbor 10.10.110.1 remote-as 64551
  neighbor 10.10.110.1 route-map SDN out
  neighbor 10.10.110.1 maximum-routes 12000
  network 172.16.1.0/24
SDN(config-router-bgp)#show route-map
route-map SDN permit 10
  Description:
  Match clauses:
    match ip address prefix-list SDN
  SubRouteMap:
  Set clauses:
    set ip next-hop 10.10.1.2
SDN(config-router-bgp)#show ip prefix-list
ip prefix-list SDN seq 10 permit 172.16.1.0/24
SDN(config-router-bgp)#
```



- レイヤー3データセンターではeBGPで構成する
- BGPコントローラーで各ルートをコントロールするのに next hopを操作



# いつ使うのか？ 3<sup>rd</sup> Party next hop BGP SDNデータセンター

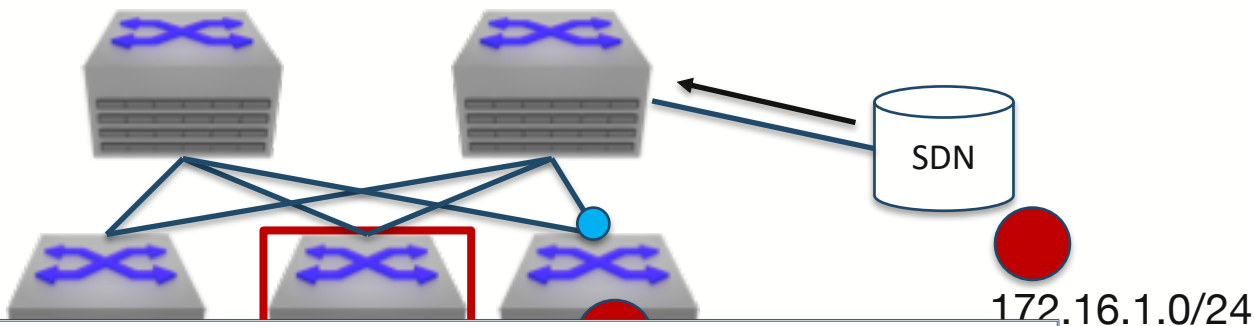


```
R4#show ip bgp 172.16.1.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.4, local AS number 64554
BGP routing table entry for 172.16.1.0/24
  Paths: 2 available
    64551 64550
      10.10.1.5 from 10.10.1.5 (10.255.255.1)
        Origin IGP, metric 0, localpref 100, IGP metric 1, weight 0, received 00:16:43 ago, valid, external, best
        Rx SAFI: Unicast
    64552 64553 64551 64550
      10.10.1.17 from 10.10.1.17 (10.255.255.2)
        Origin IGP, metric 0, localpref 100, IGP metric 1, weight 0, received 00:16:43 ago, valid, external
        Rx SAFI: Unicast
R4#
```

172.16.1.0/24

する  
しするのに

# いつ使うのか？ 3<sup>rd</sup> Party next hop BGP SDNデータセンター

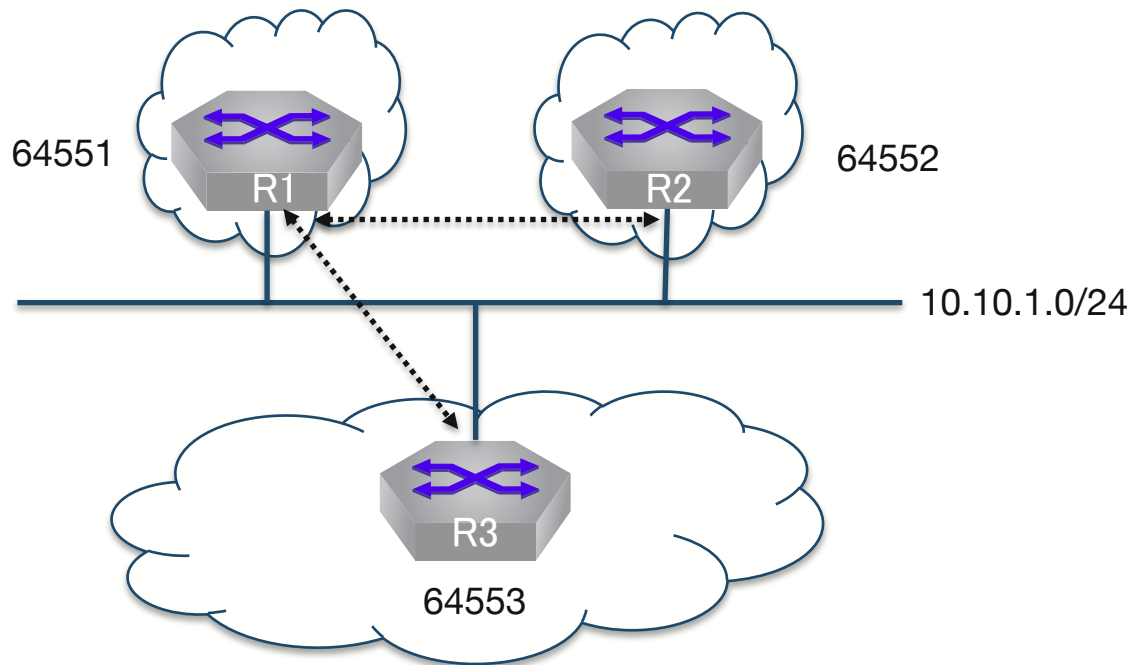


```
R5#show ip bgp 172.16.1.0/24
BGP routing table information for VRF default
Router identifier 10.255.255.5, local AS number 64555
BGP routing table entry for 172.16.1.0/24
Paths: 2 available
 64551 64550
   10.10.1.9 from 10.10.1.9 (10.255.255.1)
     Origin IGP, metric 0, localpref 100, IGP metric 1, weight 0, received 00:17:08 ago, valid, external, best
     Rx SAFI: Unicast
 64552 64553 64551 64550
   10.10.1.21 from 10.10.1.21 (10.255.255.2)
     Origin IGP, metric 0, localpref 100, IGP metric 1, weight 0, received 00:17:08 ago, valid, external
     Rx SAFI: Unicast
```

172.16.1.0/24

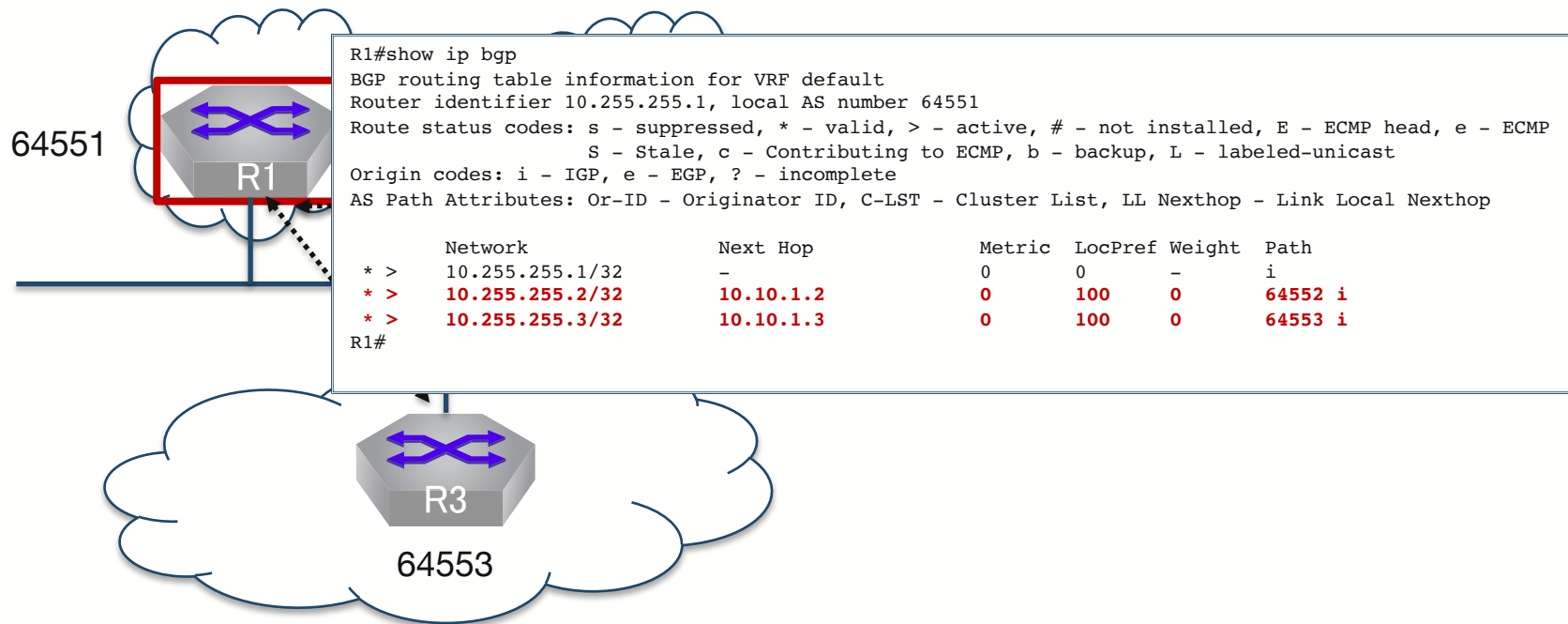
する  
しするのに

## 5.1.3(2) 1st party next hop



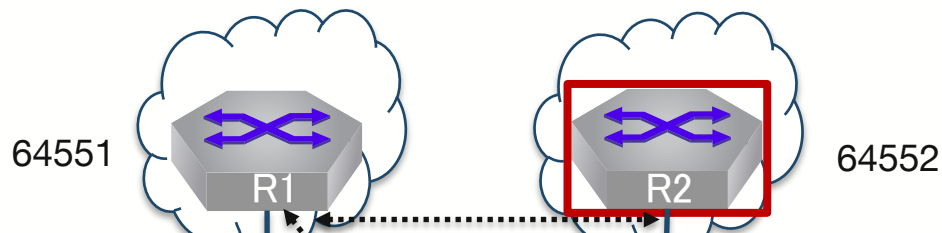
- R1とR2,R1とR3が10.10.1.0/24でピアリング

## 5.1.3(2) 1st party next hop



- R1はそれぞれから学習

## 5.1.3(2) 1st party next hop

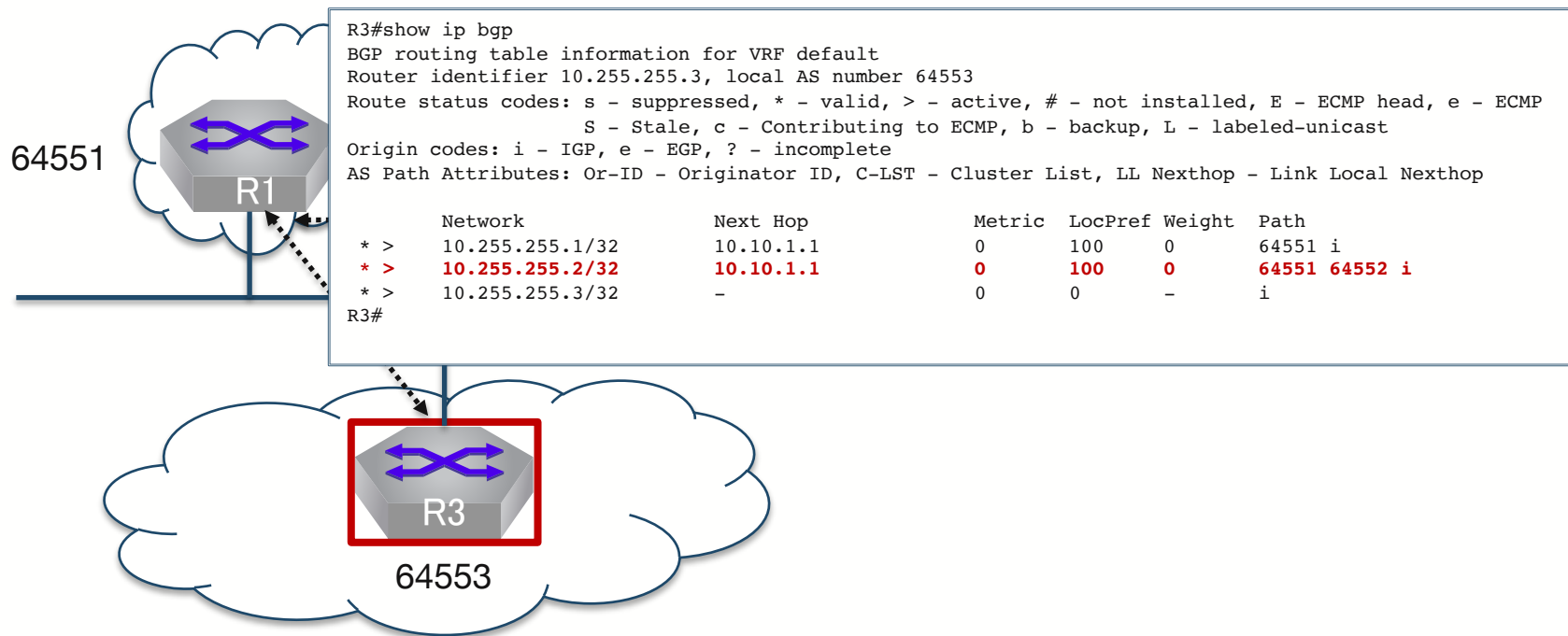


```
R2#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.2, local AS number 64552
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

      Network                Next Hop                Metric  LocPref  Weight  Path
* >   10.255.255.1/32         10.10.1.1                0       100     0       64551 i
* >   10.255.255.2/32         -                          0         0     -         i
* >   10.255.255.3/32         10.10.1.1                0       100     0       64551 64553 i
R2#
```

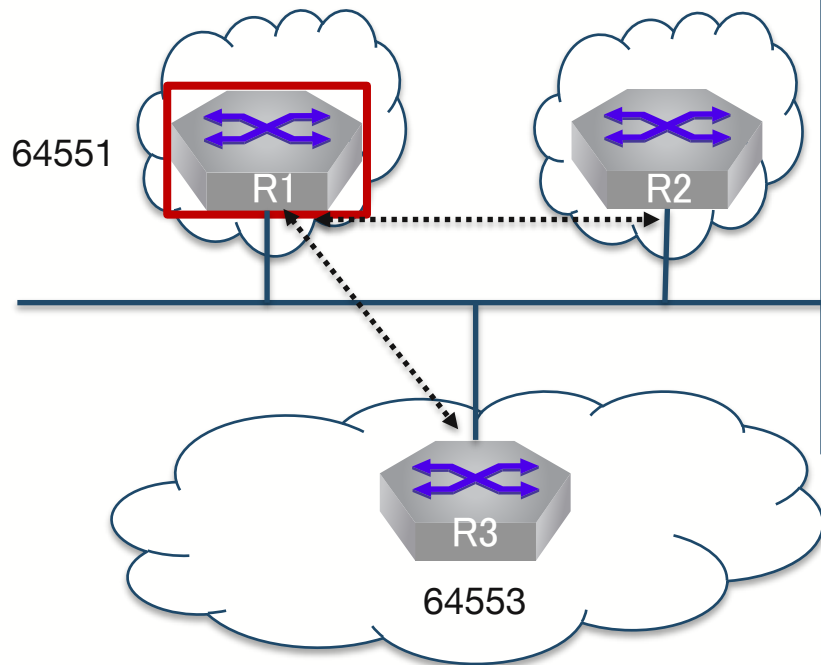
- R3の経路もR1経由で学習

## 5.1.3(2) 1st party next hop



- R2の経路もR1経由で学習

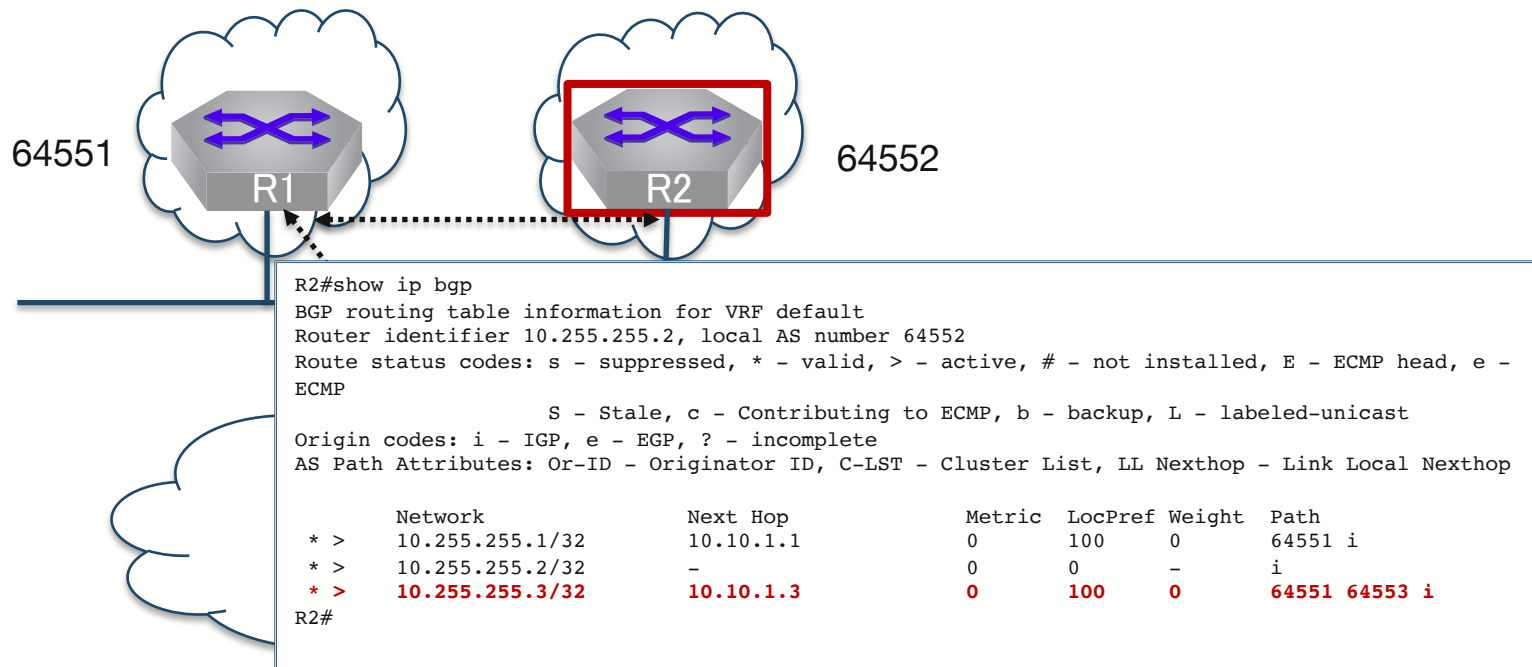
## 5.1.3(2) 1st party next hop



```
R1(config-router-bgp)#show active
router bgp 64551
  router-id 10.255.255.1
  neighbor 10.10.1.2 remote-as 64552
  neighbor 10.10.1.2 route-map 1stpartyNexthop out
  neighbor 10.10.1.2 maximum-routes 12000
  neighbor 10.10.1.3 remote-as 64553
  neighbor 10.10.1.3 route-map 1stpartyNexthop out
  neighbor 10.10.1.3 maximum-routes 12000
  network 10.255.255.1/32
R1(config-router-bgp)#show route-map 1stpartyNexthop
route-map 1stpartyNexthop permit 10
Description:
Match clauses:
SubRouteMap:
Set clauses:
  set ip next-hop unchanged
```

- 経路を最適化する為に1st party next hopを実施

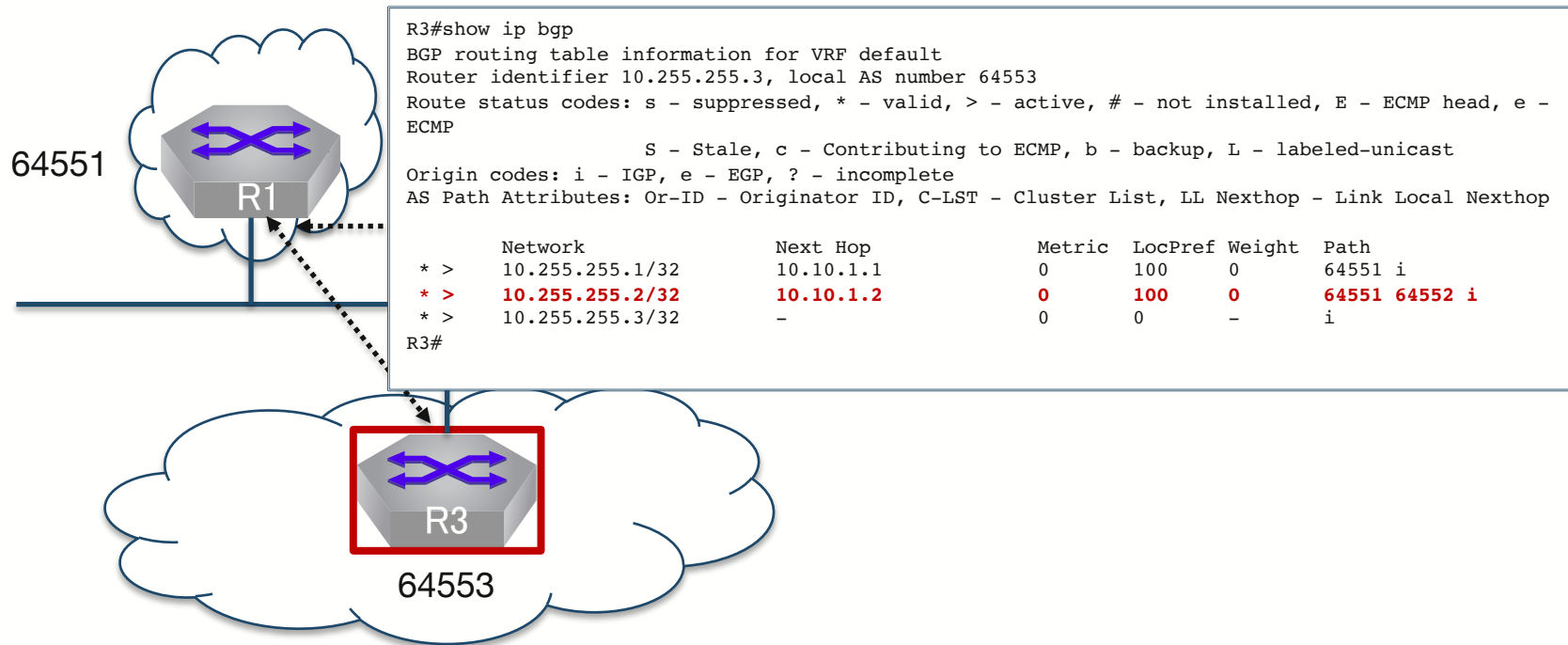
## 5.1.3(2) 1st party next hop



- R3がピアリングしていないR3経由に

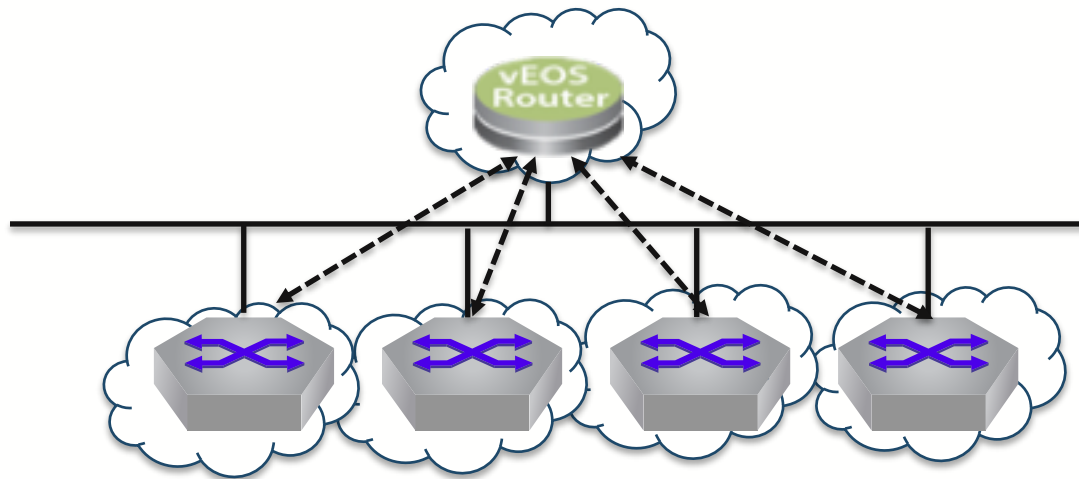


## 5.1.3(2) 1st party next hop



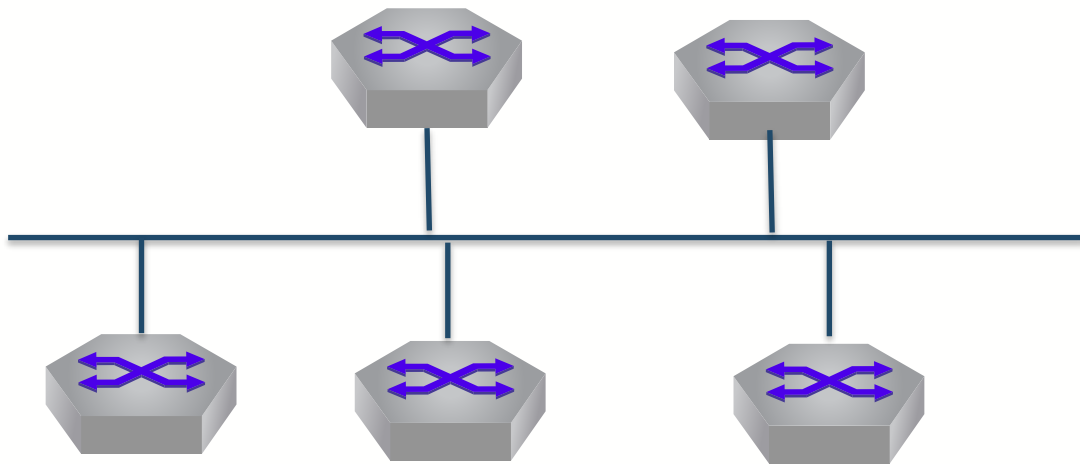
- R2がピアリングしていないR2経由に

# いつ使うのか？ 1<sup>st</sup> Party next hop



- IXでのマルチリテラルピアリング
- ルートサーバー([RFC7947](#)) (AS PATH透過性なども必要)

# いつ使うのか? 1st hop nexthop



- 広域イーサネットにおける企業ネットワーク構築
- マルチアクセスにする為にOSPFなどを使うと最大何ネイバー?議論になってしまう
- センターRRにして、拠点をRRクライアント
- センターをSpineの様に見立て、eBGPでnext hop unchangedも可

# いつ使うのか? 1st hop nexthop



```
R5#show ip ospf neighbor
Neighbor ID      VRF      Pri State                    Dead Time   Address      Interface
10.255.255.1    default  1  2 WAYS/DROTHER           00:00:29   10.10.1.1   Ethernet1
10.255.255.3    default  1  FULL/BDR                  00:00:35   10.10.1.3   Ethernet1
10.255.255.2    default  1  2 WAYS/DROTHER           00:00:34   10.10.1.2   Ethernet1
10.255.255.4    default  1  FULL/DR                   00:00:32   10.10.1.4   Ethernet1
R5#show ip route ospf
```

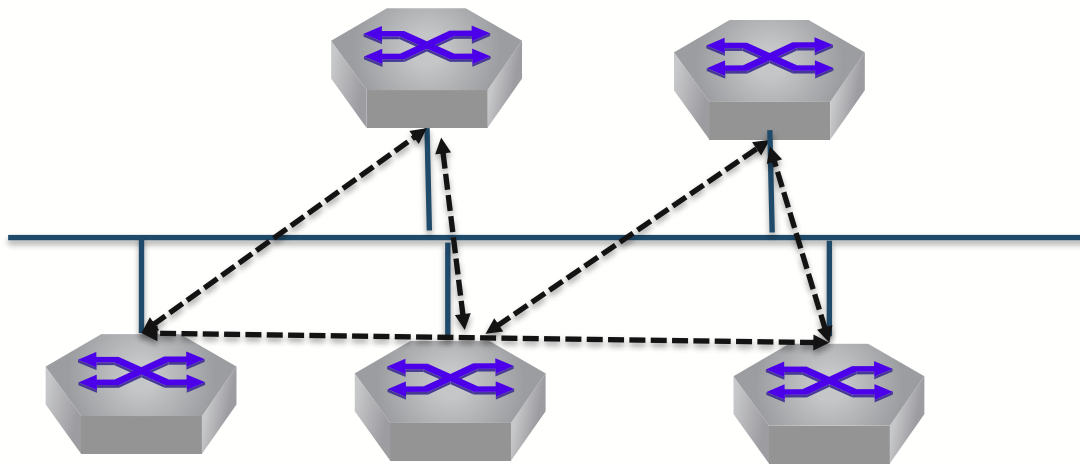
```
VRF: default
Codes: C - connected, S - static, K - kernel,
O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
N2 - OSPF NSSA external type2, B I - iBGP, B E - eBGP,
R - RIP, I L1 - IS-IS level 1, I L2 - IS-IS level 2,
O3 - OSPFv3, A B - BGP Aggregate, A O - OSPF Summary,
NG - Nexthop Group Static Route, V - VXLAN Control Service,
DH - DHCP client installed default route, M - Martian
```

```
O    10.255.255.1/32 [110/20] via 10.10.1.1, Ethernet1
O    10.255.255.2/32 [110/20] via 10.10.1.2, Ethernet1
O    10.255.255.3/32 [110/20] via 10.10.1.3, Ethernet1
O    10.255.255.4/32 [110/20] via 10.10.1.4, Ethernet1
```

- 広域イーサネット
- マルチアクセス
- センターRRにし
- センターをSpine

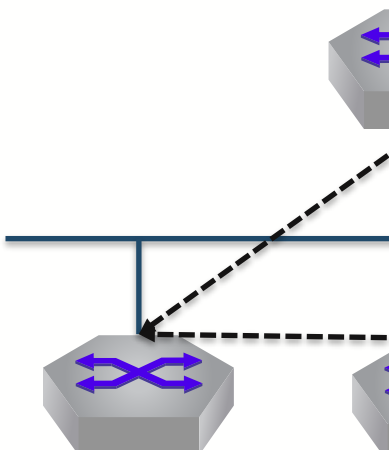
う

# いつ使うのか? 1st hop nexthop



- こんな感じにeBGP with nexthop unchanged
- 2ネイバーくらいでフルメッシュ通信が可能

# いつ使うのか? 1st hop nexthop



```
R5(config)#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.5, local AS number 64555
Route status codes: s - suppressed, * - valid, > - active, # - not installed, E - ECMP head, e - ECMP
                    S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop
```

	Network	Next Hop	Metric	LocPref	Weight	Path
* >	10.255.255.1/32	10.10.1.1	0	100	0	64552 64551 i
*	10.255.255.1/32	10.10.1.1	0	100	0	64554 64553 64551 i
* >	10.255.255.2/32	10.10.1.2	0	100	0	64552 i
* >	10.255.255.3/32	10.10.1.3	0	100	0	64554 64553 i
*	10.255.255.3/32	10.10.1.3	0	100	0	64552 64551 64553 i
* >	10.255.255.4/32	10.10.1.4	0	100	0	64554 i
* >	10.255.255.5/32	-	0	0	-	i

```
R5(config)#show ip route bgp
```

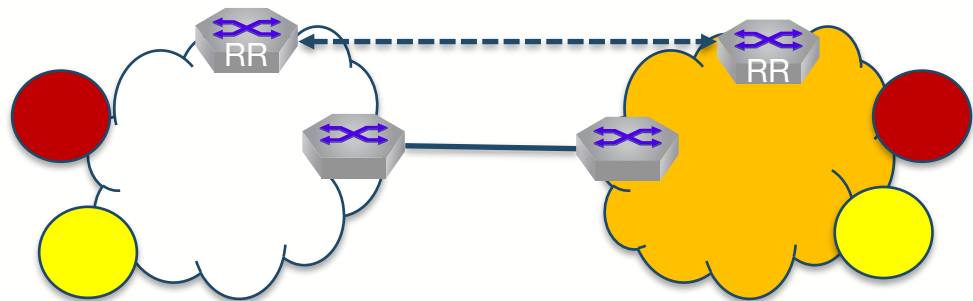
```
VRF: default
```

```
Codes: C - connected, S - static, K - kernel,
O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
N2 - OSPF NSSA external type2, B I - iBGP, B E - eBGP,
R - RIP, I L1 - IS-IS level 1, I L2 - IS-IS level 2,
O3 - OSPFv3, A B - BGP Aggregate, A O - OSPF Summary,
NG - Nexthop Group Static Route, V - VXLAN Control Service,
DH - DHCP client installed default route, M - Martian
```

```
B E 10.255.255.1/32 [200/0] via 10.10.1.1, Ethernet1
B E 10.255.255.2/32 [200/0] via 10.10.1.2, Ethernet1
B E 10.255.255.3/32 [200/0] via 10.10.1.3, Ethernet1
B E 10.255.255.4/32 [200/0] via 10.10.1.4, Ethernet1
```

- こんな感じ
- 2ネイバー

## Inter AS option c)



- AS間を分けたVPNサービスの場合、RRが全てのVPN経路情報を持っている
- マルチホップBGPでVPN経路をアナウンス
- このとき自宛になっても困るので、next hopを変えずにアナウンス



# Thank You

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