ルーティングのためのIPアドレス?

Shishio Tsuchiya
shtsuchi@arista.com

Public. Copyright © Arista 2018. All rights reserved.

ARISTA

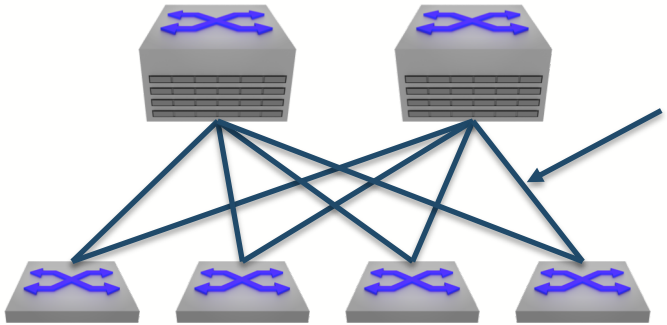
はじめに

- IPv4アドレス枯渇やIPv6への移行とかでIPアドレスのお話は色々話はあるけれど、ん？それってこんなので良くない？とか色々話はあります。
- メンテナンスのためにとかもあるけど、うーん？ということも。
- 本セッションではルータにつける/ルータが持つIPアドレスを中心にお話したいと思います。

Public. Copyright © Arista 2018. All rights reserved.

ARISTA

ルータにつけるIPアドレス



つけるIPアドレスは?
/30?

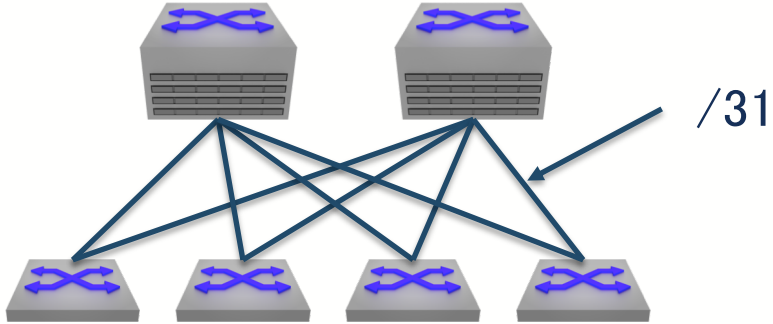
10.	1.	1	0	0	0	0	0	0	0	0
10.	1.	1.	0	0	0	0	0	0	0	1
10.	1.	1.	0	0	0	0	0	0	1	0
10.	1.	1.	0	0	0	0	0	0	1	1

30ビット

10.1.1.0/30 ネットワーク
10.1.1.1/30
10.1.1.2/30
10.1.1.3/30 ブロードキャスト

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

ルータにつけるIPアドレス



31ビット

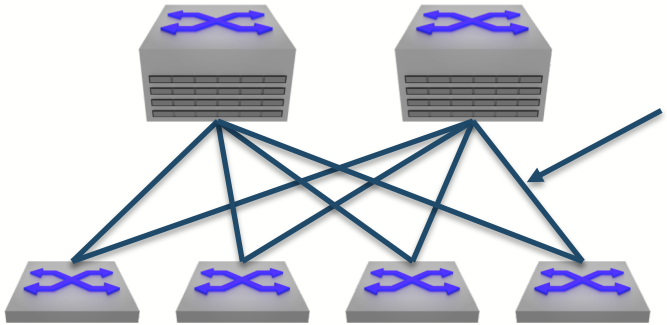
10.	1.	1	0	0	0	0	0	0	0	0
10.	1.	1.	0	0	0	0	0	0	0	1

10.1.1.0/31
10.1.1.1/31

- RFC3021 31-Bit Prefixes on IPv4 Links
- ホストアドレス2個
- かなり古いRFC(2000年)で、既に誰でもサポートしてると思うけど、つかってないケースも

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

ルータにつけるIPアドレス（IPv6）

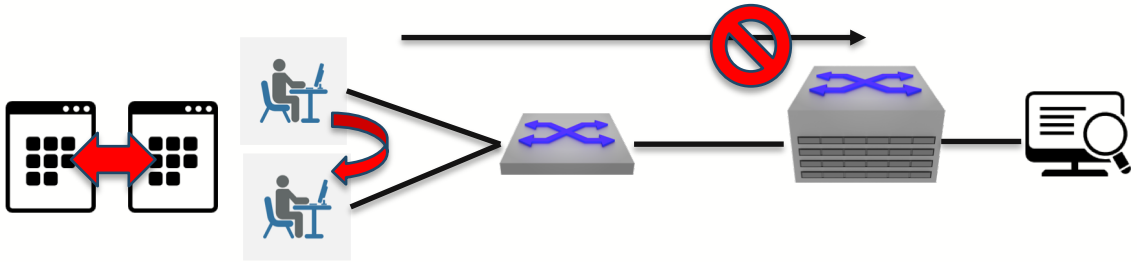


リンクローカル
or
リンクローカル and
グローバルアドレス

- IPv6ではリンクローカルアドレスが存在する
- グローバルアドレスは必要か？

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

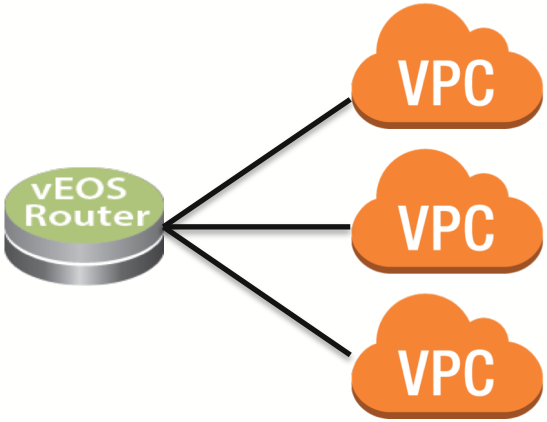
ところでIPv4のリンクローカルアドレス



- DHCPなどのIPアドレスを動的に払い出す方法があるが、いつでもサーバーが有効である事は無く、またワイヤレスなどでアドホックなコミュニケーションが必要になった時の為にRFC3929 IPv4 Link-Local 169.254/16を定義
- Microsoft Windows 98以降やMac OS 8.5以降で既にサポートされていたが、IETF標準として文書化
- ARPプローブによりアドレスの重複を防ぐ

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

ちょっと?なケース

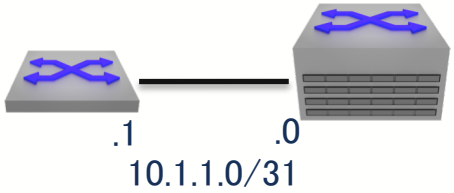


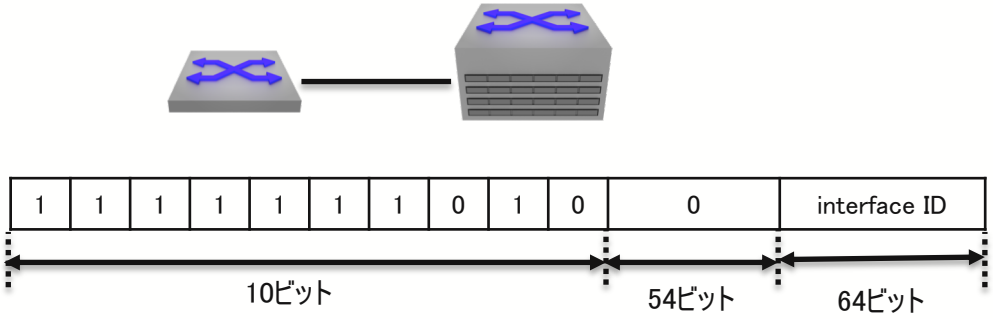
- Q. 仮想インターフェイスの両端には、それぞれどの IP アドレスが割り当てられますか?
- パブリック AWS クラウドへの仮想インターフェイスを構成する場合は、お客様が所有するパブリック IP スペースからの IP アドレスが接続の両端に割り当てられている必要があります。仮想インターフェイスの接続先が VPC であり、AWS でピア IP CIDR を自動生成することにした場合は、AWS によって接続の両端の IP アドレススペースが **169.254.0.0/16** の範囲内で割り当てられます。

https://aws.amazon.com/jp/directconnect/faqs/

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

IPv6アドレス種類 リンクローカル



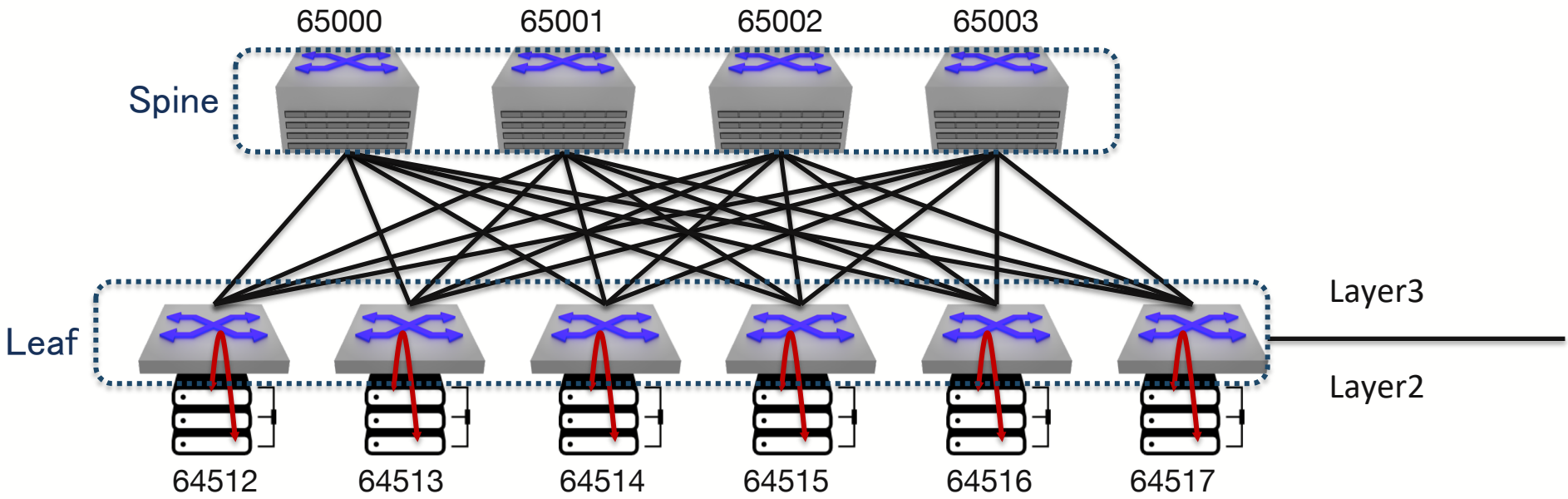


FE80::/10

- RFC4291 IPv6 Addressing Architecture
- シングルリンクでの自動アドレス設定/ネイバー発見/またはルータがいない時に使われる
- ルータはこのアドレスが宛先/送信元である時には転送してはならない
- すべてのルーティングプロトコルはこのアドレスで隣接関係を確立できる

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

データセンターでのネットワーク



- 非常に密接したトポロジーBGPを利用したIP Closデザイン
- IPv6リンクローカルアドレスとRFC5549 v4 NLRI with v6 NHの活用でアドレスを最小限度にする

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

RFC5549 Capability Negotiation

```
+-------------------------------+
| Capability Code (1 octet:5 ) |
+-------------------------------+
| Capability Length (1 octet)  |
+-------------------------------+
| Capability Value (variable)  |
+-------------------------------+
```

```
+-----------------------------------------------------+
| NLRI AFI (2 octets)   1:IPv4                        |
+-----------------------------------------------------+
| NLRI SAFI(2 octets)   1:Unicast 2:Multicast         |
|                       4:Label 128:MPLS-VPN          |
+-----------------------------------------------------+
| Nexthop AFI(2 octets)  2:IPv6                       |
+-----------------------------------------------------+
```

- RFC5492 のCapability Code 5 Extended Next Hop Encodingでネゴシエーション
- Valueにてサポート可能なAFI/SFIおよび Nexthop AFIを通知(IPv6:2)

Public. Copyright © Arista 2018. All rights reserved.

RFC5549 UPDATE/WITHDRAWN

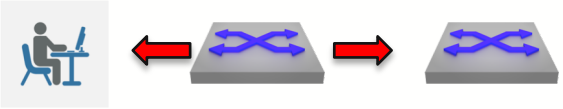
```
+-----------------------------------------------------------+
| Address Family Identifier (2 octets)                      |
+-----------------------------------------------------------+
| Subsequent Address Family Identifier (1 octet)            |
+-----------------------------------------------------------+
| Length of Next Hop Network Address (1 octet)              |
+-----------------------------------------------------------+
| Network Address of Next Hop (variable)                    |
+-----------------------------------------------------------+
| Reserved (1 octet)                                        |
+-----------------------------------------------------------+
| Network Layer Reachability Information (variable)         |
+-----------------------------------------------------------+
```

- RFC4760 のMPBGPのMP_REACH_NLRIとMP_UNREACH_NLRIをそのまま使用可能

Public. Copyright © Arista 2018. All rights reserved.

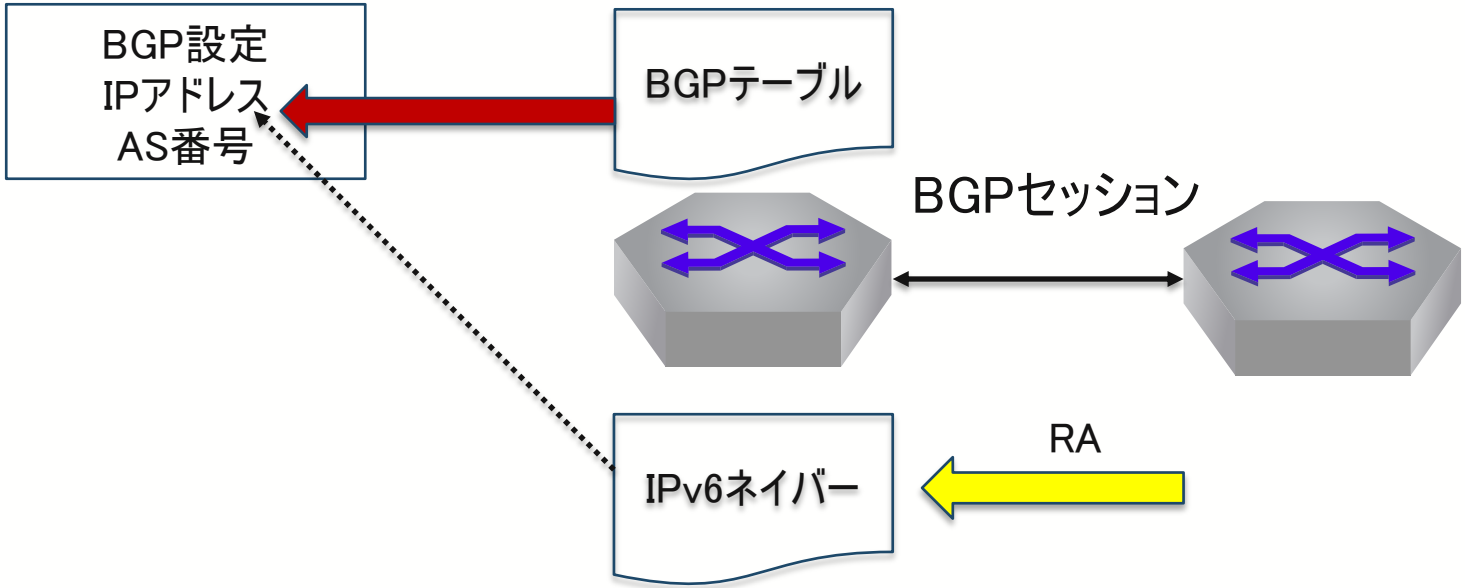
BGPとIPv6の特徴

- 他のIGPとは違いダイナミックディスカバリーが無い
  - neighbor x.x.x.x remote-as X
- IPv6リンクローカルアドレスはシステムMACアドレスなどより自動的に作成される
- ルータはルータアドバタイズメントを定期的行う
  - https://tools.ietf.org/html/rfc4861#section-4.2



Public. Copyright © Arista 2018. All rights reserved.

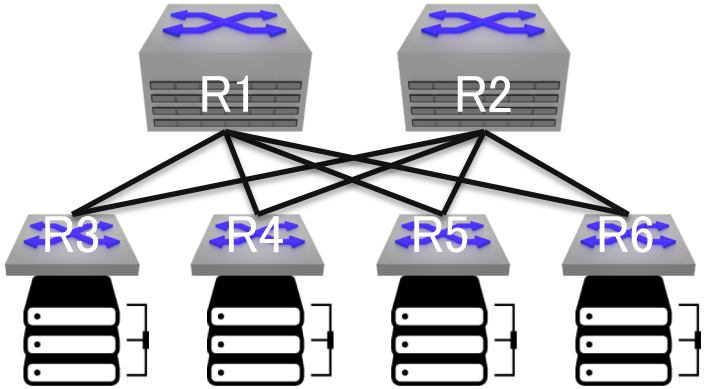
通常のBGPとBGP IPv6 Link Local Peer Discovery



- 通常BGPはIPアドレスやAS番号の設定情報よりBGPのセッション確立を行う
- IPv6 NDでRAを受けたノードに関しては設定情報と同様な扱いにする

Public. Copyright © Arista 2018. All rights reserved.

RFC5549 with BGP IPv6 Link Local Peer Discovery



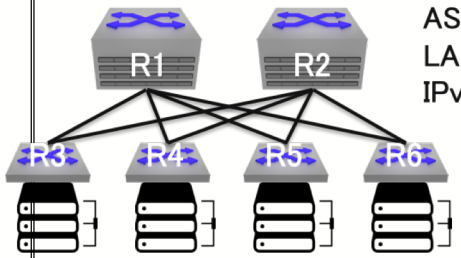
loopback:10.255.255.x/32
AS番号:6455x
LAN側:10.10.x.0/24
IPv6 link-local:fe80::x/64

- RFC5549とIPv6 Link Local Peer Discoveryを設定する
- LeafではSpineは決め打ちの設定を行う

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

RFC5549 with BGP IPv6 Link Local Peer Discovery

```
router bgp 64553
   bgp default ipv4-unicast transport ipv6
   maximum-paths 2
   neighbor fe80::1%Et1 remote-as 64551
   neighbor fe80::1%Et1 maximum-routes 12000
   neighbor fe80::2%Et2 remote-as 64552
   neighbor fe80::2%Et2 maximum-routes 12000
   redistribute connected
   !
   address-family ipv4
      neighbor fe80::1%Et1 activate
      neighbor fe80::1%Et1 next-hop address-family ipv6 originate
      neighbor fe80::2%Et2 activate
      neighbor fe80::2%Et2 next-hop address-family ipv6 originate
   !
   address-family ipv6
      neighbor fe80::1%Et1 activate
      neighbor fe80::2%Et2 activate
!
```



loopback:10.255.255.x/32
AS番号:6455x
LAN側:10.10.x.0/24
IPv6 link-local:fe80::x/64

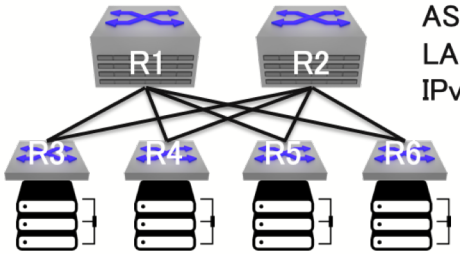
全てのIPv6ネイバーでIPv4ユニキャストの受信を可能に

指定ネイバーでnexthop IPv6送信を行う

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

RFC5549 with BGP IPv6 Link Local Peer Discovery

```
router bgp 64551
   bgp default ipv4-unicast transport ipv6
   neighbor R3 peer group
   neighbor R3 maximum-routes 12000
   neighbor R4 peer group
   neighbor R4 maximum-routes 12000
   neighbor R5 peer group
   neighbor R5 maximum-routes 12000
   neighbor R6 peer group
   neighbor R6 maximum-routes 12000
   redistribute connected
   neighbor interface Et1 peer-group R3 remote-as 64553
   neighbor interface Et2 peer-group R4 remote-as 64554
   neighbor interface Et3 peer-group R5 remote-as 64555
   neighbor interface Et4 peer-group R6 remote-as 64556
   !
   address-family ipv4
      neighbor R3 activate
      neighbor R3 next-hop address-family ipv6 originate
-snip-
```



loopback:10.255.255.x/32
AS番号:6455x
LAN側:10.10.x.0/24
IPv6 link-local:fe80::x/64

指定のインターフェースのリンクローカルのBGPピアと接続する

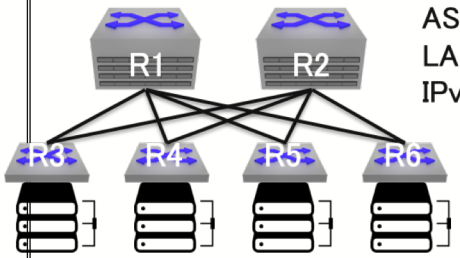
AS番号の個別指定もしたくない時は?

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

RFC5549 with BGP Listen Range

受け入れ可能なAS番号を設定する

```
peer-filter PrivateAS
   10 match as-range 64513-64550 result reject
   20 match as-range 64551-64556 result accept
!
router bgp 64552
   bgp default ipv4-unicast transport ipv6
   bgp listen range fe80::/10 peer-group Leaf peer-filter PrivateAS
   neighbor Leaf peer group
   neighbor Leaf maximum-routes 12000
   redistribute connected
   !
   address-family ipv4
      neighbor Leaf activate
      neighbor Leaf next-hop address-family ipv6 originate
   !
   address-family ipv6
      neighbor Leaf activate
!
```



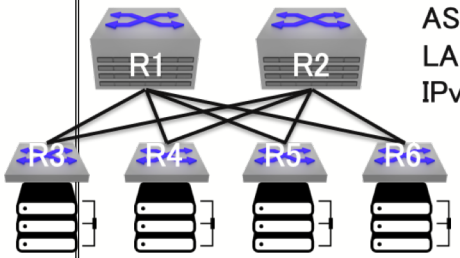
loopback:10.255.255.x/32
AS番号:6455x
LAN側:10.10.x.0/24
IPv6 link-local:fe80::x/64

受け入れ可能なIPアドレスの範囲とAS番号を設定する

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

RFC5549 with BGP Listen Range

```
R3#show ipv6 bgp peers
BGP neighbor is fe80::1%Et1, remote AS 64551, external link
  BGP version 4, remote router ID 10.255.255.1, VRF default
  Negotiated BGP version 4
  Member of update group 1
  Last read 00:00:55, last write 00:00:55
  Hold time is 180, keepalive interval is 60 seconds
  Configured hold time is 180, keepalive interval is 60 seconds
  Connect timer is inactive
  Idle-restart timer is inactive
  BGP state is Established, up for 03:45:55
  Number of transitions to established: 2
  Last state was OpenConfirm
  Last event was RecvKeepAlive
  Last rcvd notification:Cease/peer de-configured, Last time 03:56:57
  Neighbor Capabilities:
    Multiprotocol IPv4 Unicast: advertised and received and negotiated
    Multiprotocol IPv6 Unicast: advertised and received and negotiated
    Extended Next-Hop Capability:
      IPv4 Unicast: advertised and received and negotiated
    Four Octet ASN: advertised and received and negotiated
    Route Refresh: advertised and received and negotiated
   -snip-
```

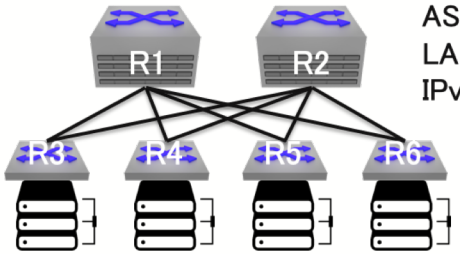


loopback:10.255.255.x/32
AS番号:6455x
LAN側:10.10.x.0/24
IPv6 link-local:fe80::x/64

Public. Copyright © Arista 2018. All rights reserved.

show ip route

```
R3#show ip route bgp
-snip-
 B E      10.10.4.0/24 [200/0]
           via fe80::1, Ethernet1
           via fe80::2, Ethernet2
 B E      10.10.5.0/24 [200/0]
           via fe80::1, Ethernet1
           via fe80::2, Ethernet2
 B E      10.10.6.0/24 [200/0]
           via fe80::1, Ethernet1
           via fe80::2, Ethernet2
 B E      10.255.255.1/32 [200/0]
           via fe80::1, Ethernet1
 B E      10.255.255.2/32 [200/0]
           via fe80::2, Ethernet2
 B E      10.255.255.4/32 [200/0]
           via fe80::1, Ethernet1
           via fe80::2, Ethernet2
 B E      10.255.255.5/32 [200/0]
           via fe80::1, Ethernet1
           via fe80::2, Ethernet2
 B E      10.255.255.6/32 [200/0]
           via fe80::1, Ethernet1
           via fe80::2, Ethernet2
```

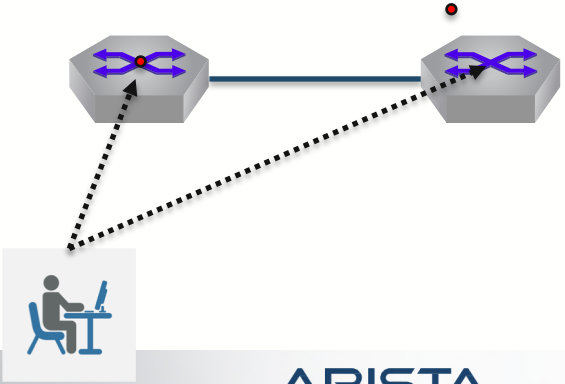


loopback:10.255.255.x/32
AS番号:6455x
LAN側:10.10.x.0/24
IPv6 link-local:fe80::x/64

Public. Copyright © Arista 2018. All rights reserved.

リンクローカルだけで運用出来ないのか？

- もともとIPv6ルーティングプロトコルはリンクローカルだけで動作
  - AWSでもやってるよね?☺
- SSH/Telnet/ICMP/SNMP/はループバックを使用すればよい
- ICMPエラーはループバックを使用する
- アドレス設計/設定の簡素化/セキュリティという意味でもメリットが多い
- RFC7404 Using Only Link-Local Addressing inside an IPv6 Network
  - https://tools.ietf.org/html/rfc7404



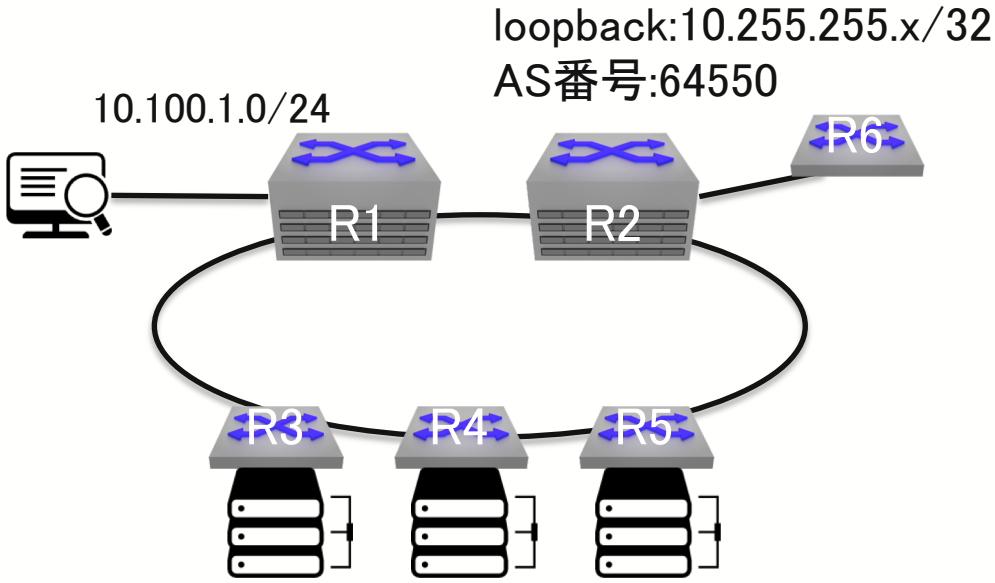
Public. Copyright © Arista 2018. All rights reserved.

そもそもP2Pにアドレスが必要か？(IPv4)

- PPPではループバックなどのアドレスを借りてきて、そのアドレスを識別子として、隣接関係を構築する
- IPv6では同一リンクのみ利用可能なリンクローカルが使用される
- もともとブロードキャストと比べるとP2Pでは単純
- RFC5309 Point-to-Point Operation over LAN in Link State Routing Protocols
  - https://tools.ietf.org/html/rfc5309

Public. Copyright © Arista 2018. All rights reserved.

RFC5309 unnumbered Link State Protocol

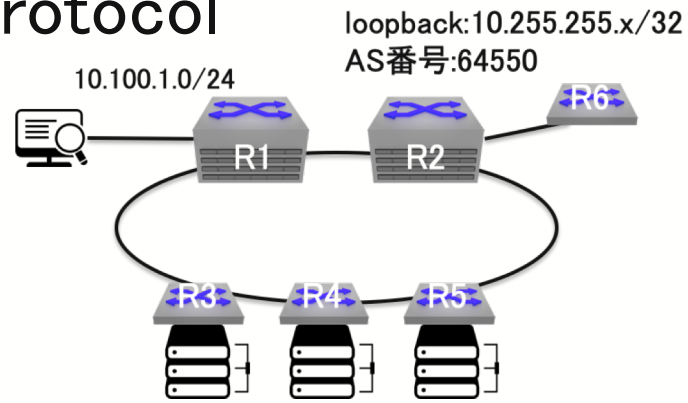


- メトロサービスを意識したトポロジーでRFC5309 P2P over LAN

Public. Copyright © Arista 2018. All rights reserved.

RFC5309 unnumbered Link State Protocol

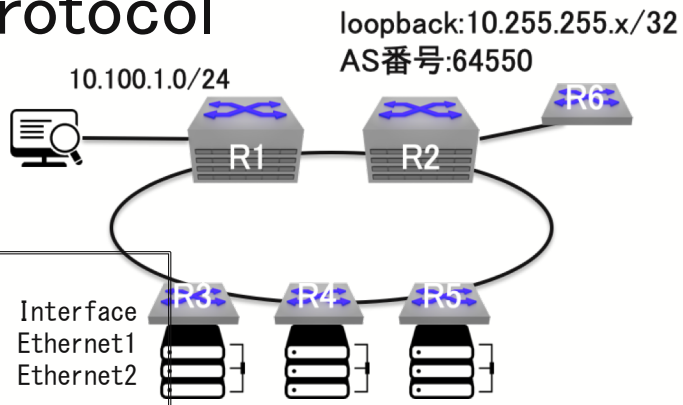
```
R1#show startup-config | section ospf
interface Ethernet1
   no switchport
   ip address unnumbered Loopback0
   ip ospf network point-to-point
   ip ospf area 0.0.0.0
!
interface Ethernet2
   no switchport
   ip address unnumbered Loopback0
   ip ospf network point-to-point
   ip ospf area 0.0.0.0
!
router ospf 1
   passive-interface Ethernet3
   network 10.100.1.0/24 area 0.0.0.0
   network 10.255.255.1/32 area 0.0.0.0
   max-lsa 12000
R1#
```



- LANインターフェースでP2Pの設定を行う
- インターフェース内でareaなどの設定

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

RFC5309 unnumbered Link State Protocol

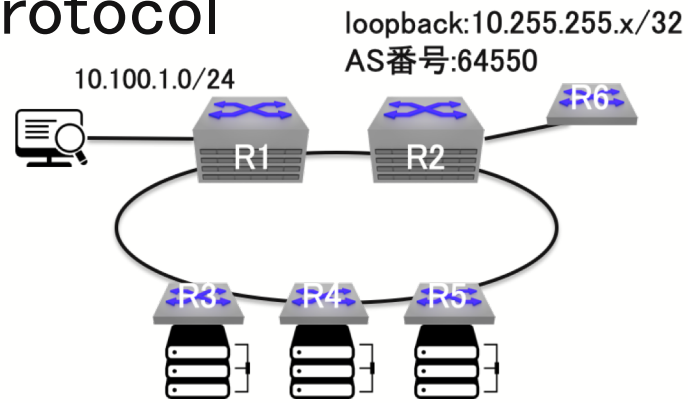



```
R1#show ip ospf neighbor
Neighbor ID     Instance VRF      Pri State                   Dead Time   Address         Interface
10.255.255.2    1        default  0   FULL                    00:00:34    10.255.255.2    Ethernet1
10.255.255.3    1        default  0   FULL                    00:00:36    10.255.255.3    Ethernet2
R1#show ip ospf database

            OSPF Router with ID(10.255.255.1) (Instance ID 1) (VRF default)


                 Router Link States (Area 0.0.0.0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.255.255.2    10.255.255.2    149         0x80000006 0x8106   4
10.255.255.4    10.255.255.4    186         0x80000005 0x5d4e   3
10.255.255.6    10.255.255.6    164         0x80000004 0x4888   2
10.255.255.3    10.255.255.3    150         0x80000005 0x9a8    3
10.255.255.5    10.255.255.5    161         0x80000005 0x2b7f   3
10.255.255.1    10.255.255.1    145         0x80000005 0xdb52   4
R1#
```


Public. Copyright © Arista 2018. All rights reserved.

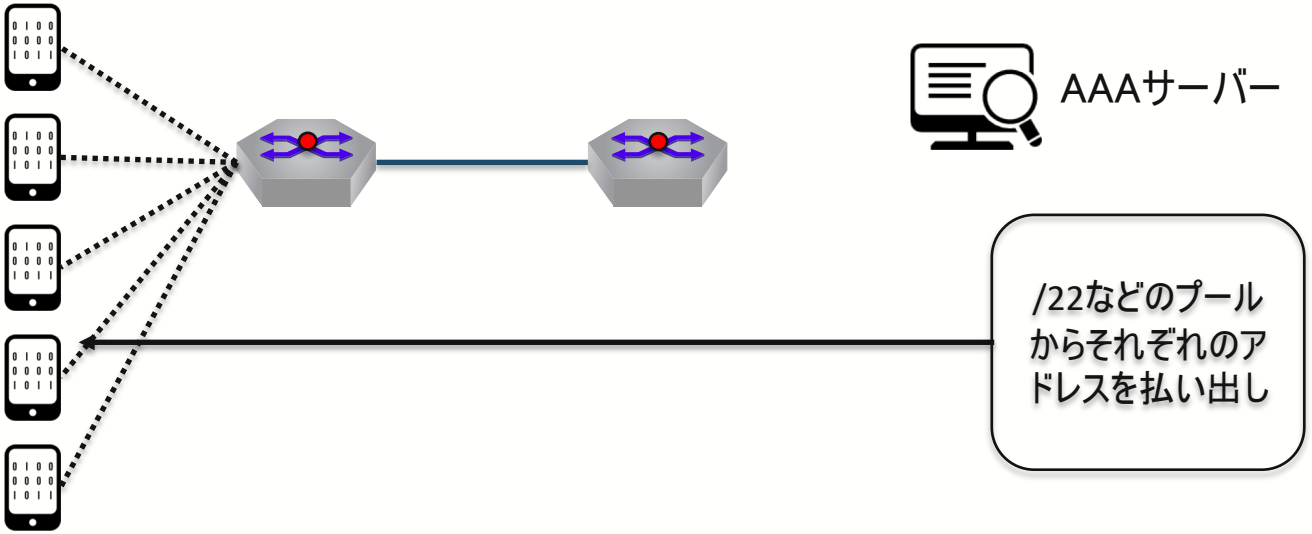
RFC5309 unnumbered Link State Protocol

```
R1#show ip ospf database  detail 10.255.255.6

            OSPF Router with ID(10.255.255.1) (Instance ID 1) (VRF default)

  LS Age: 607
  Options: (E DC)
  LS Type: Router Links
  Link State ID: 10.255.255.6
  Advertising Router: 10.255.255.6
  LS Seq Number: 0x80000004
  Checksum: 0x4888
  Length: 48
  Number of Links: 2

    Link connected to: a Stub Network
     (Link ID) Network/subnet number: 10.255.255.6
     (Link Data) Network Mask: 255.255.255.255
      Number of TOS metrics: 0
       TOS 0 Metrics: 10


    Link connected to: a Point-to-point Network
     (Link ID) Neighboring Router ID: 10.255.255.2
     (Link Data)  0.0.0.1
      Number of TOS metrics: 0
       TOS 0 Metrics: 10
```



Public. Copyright © Arista 2018. All rights reserved.

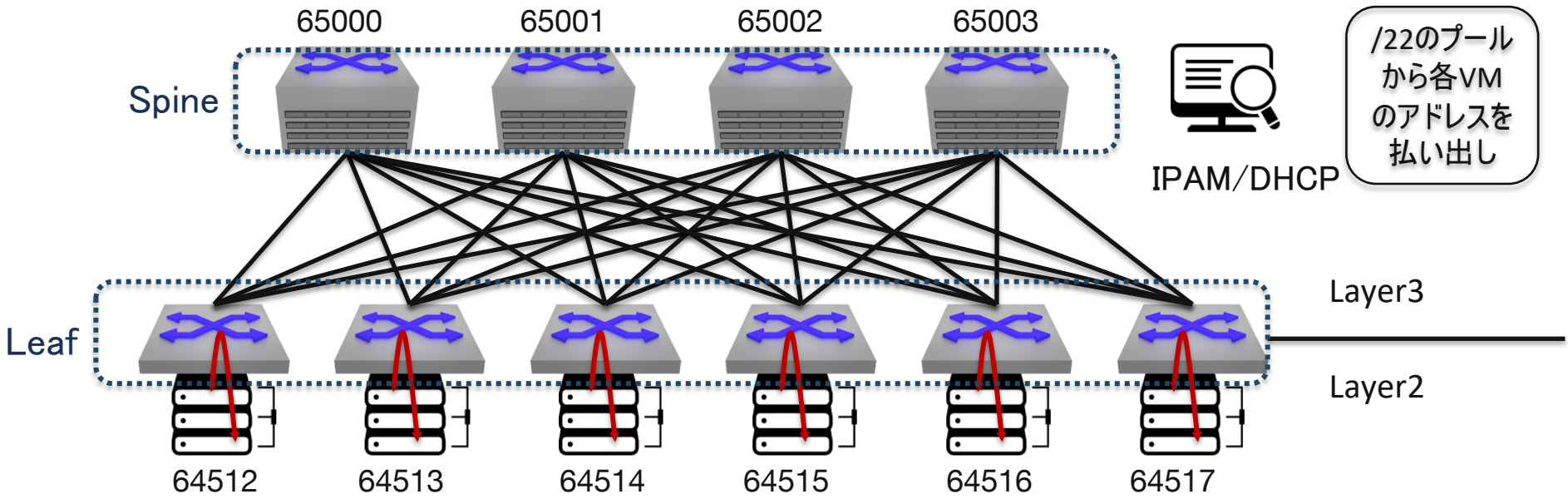
ブロードバンド環境でのIPアドレス



- ルータには/32ルートがインストールされ、subscriberルートとしてアドバタイスされルーティングされる

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

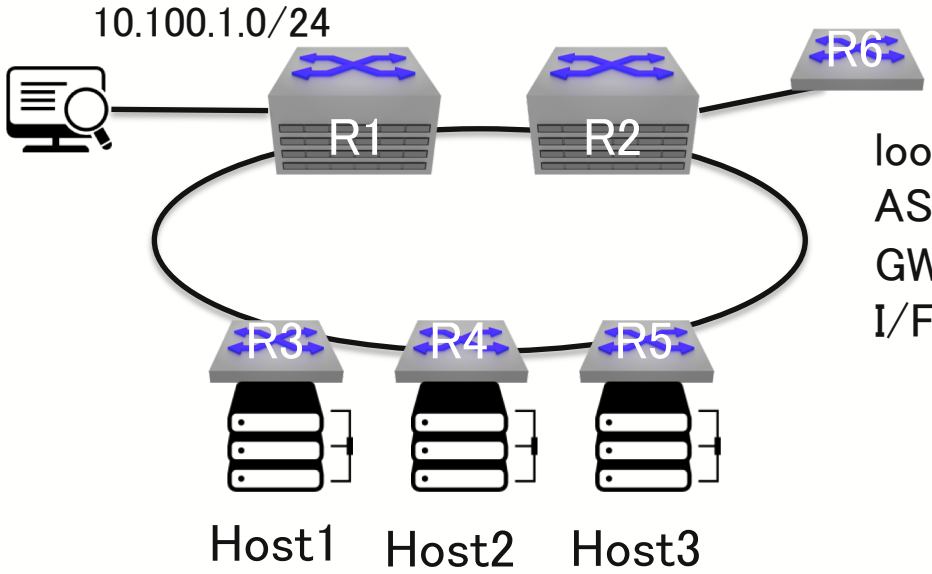
データセンターやイーサネット環境でのSubscribers ルート



- ブロードバンドと同じようにホストルートをインストールできれば効率のよいアドレス運用が出来るかもしれない
- ARPやNDPで解決されたホストルートをルーティングテーブルに

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

ARP converted Host Routes injection into BGP



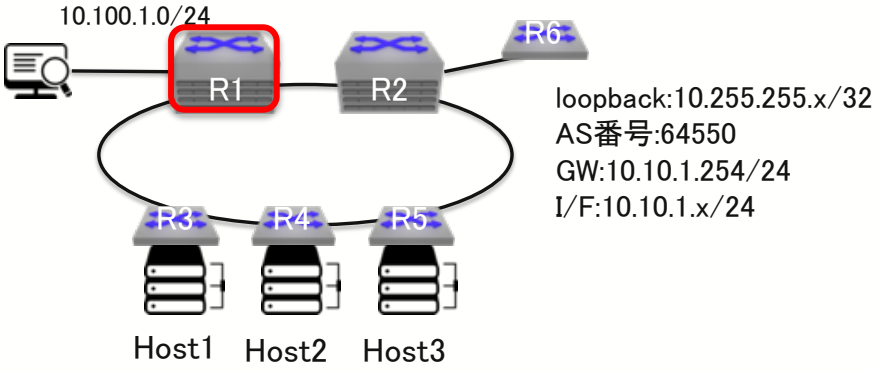
loopback:10.255.255.x/32
AS番号:64550
GW:10.10.1.254/24
I/F:10.10.1.x/24

- R3-R5にてHost Route Injectionを設定する
- R6のBGP RRにてすべてのノードに分配される

Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

```
R1#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 - BGP, 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,
       DP - Dynamic Policy Route, L - VRF Leaked

 B I      10.10.1.3/32 [200/0] via 10.255.255.3, Ethernet2
 B I      10.10.1.4/32 [200/0] via 10.255.255.3, Ethernet2
 B I      10.10.1.5/32 [200/0] via 10.255.255.2, Ethernet1
```

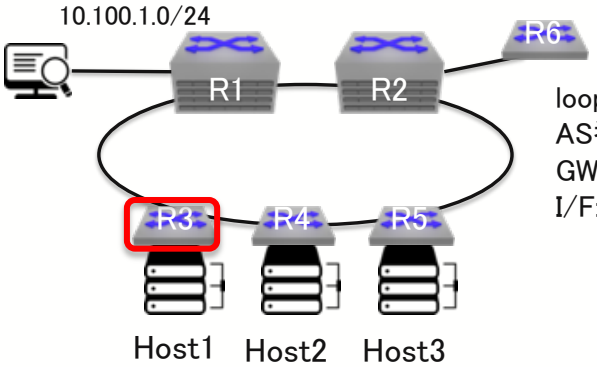


- 事前にR1は各ノードのGWアドレスを学習している

Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

```
!
interface Vlan100
   ip address 10.10.1.3/24
   ip helper-address 10.100.1.1
   ip attached-host route export
   ip virtual-router address 10.10.1.254
!
ip virtual-router mac-address 00:1c:73:00:00:99
!
ip routing
!
router bgp 64500
   neighbor 10.255.255.6 remote-as 64500
   neighbor 10.255.255.6 maximum-routes 12000
   network 10.10.1.3/32
   redistribute attached-host
!
```



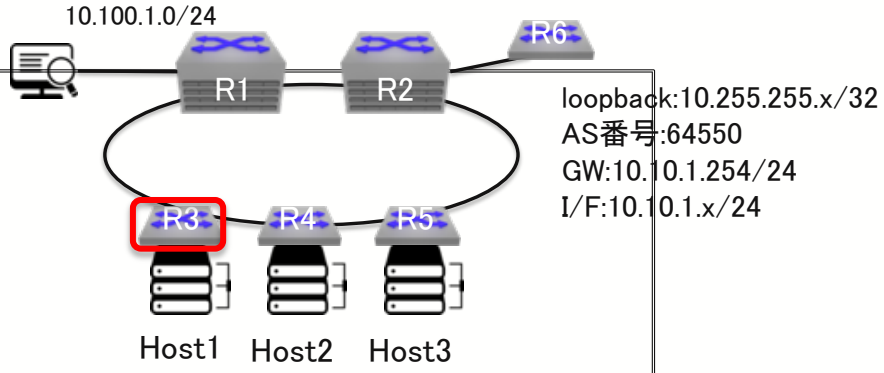
loopback:10.255.255.x/32
AS番号:64550
GW:10.10.1.254/24
I/F:10.10.1.x/24

接続されたホストテーブルをルーティングテーブルにexportをする

exportされたホストテーブルをBGPに再配信する

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

ARP converted Host Routes injection into BGP

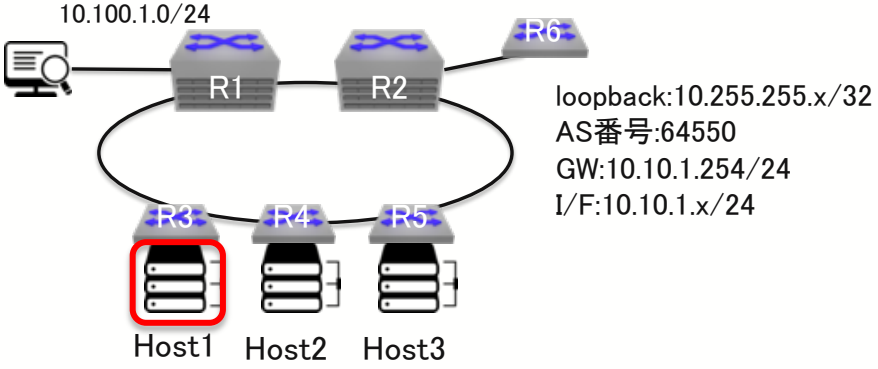



```
R3#show ip arp
Address         Age (sec)  Hardware Addr   Interface
10.255.255.1      0:00:39  001c.73c0.c601  Ethernet1
10.255.255.4      2:38:46  001c.73c3.c601  Ethernet2
10.10.1.10        2:38:54  001c.73c7.c611  Vlan100, not learned
R3#show ip attached-host route export
   IP Address       VLAN       Interface       Prefix    VRF
---------------- ---------- --------------- ------------ ---

R3#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 64500
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
 * >     10.10.1.3/32           -                     0       0       -       i
 * >     10.10.1.4/32           10.255.255.4          0       100     0       i Or-ID: 10.255.255.4 C-LST: 10.255.255.6
 * >     10.10.1.5/32           10.255.255.5          0       100     0       i Or-ID: 10.255.255.5 C-LST: 10.255.255.6
R3#
```


Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

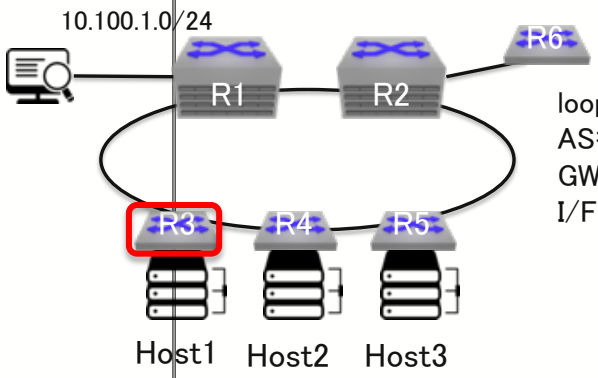
```
Host1#show startup-config interfaces ethernet 1
interface Ethernet1
   no switchport
   ip address 10.10.1.10/24
   ip address 10.10.1.11/24 secondary
   ip address 10.10.1.12/24 secondary
   ip address 10.10.1.13/24 secondary
   ip address 10.10.1.14/24 secondary
   ip address 10.10.1.15/24 secondary
   ip address 10.10.1.16/24 secondary
   ip address 10.10.1.17/24 secondary
   ip address 10.10.1.18/24 secondary
--snip-
   ip address 10.10.1.94/24 secondary
   ip address 10.10.1.95/24 secondary
   ip address 10.10.1.96/24 secondary
   ip address 10.10.1.97/24 secondary
   ip address 10.10.1.98/24 secondary
   ip address 10.10.1.99/24 secondary
Host1#
```



Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

```
R3#ping 10.10.1.10
PING 10.10.1.10 (10.10.1.10) 72(100) bytes of data.
80 bytes from 10.10.1.10: icmp_seq=1 ttl=64 time=14.6 ms
80 bytes from 10.10.1.10: icmp_seq=2 ttl=64 time=4.56 ms
80 bytes from 10.10.1.10: icmp_seq=1 ttl=64 time=18.6 ms (DUP!)
80 bytes from 10.10.1.10: icmp_seq=2 ttl=64 time=16.1 ms (DUP!)
80 bytes from 10.10.1.10: icmp_seq=3 ttl=64 time=4.20 ms
80 bytes from 10.10.1.10: icmp_seq=3 ttl=64 time=8.31 ms (DUP!)
80 bytes from 10.10.1.10: icmp_seq=4 ttl=64 time=2.08 ms
80 bytes from 10.10.1.10: icmp_seq=4 ttl=64 time=5.82 ms (DUP!)
80 bytes from 10.10.1.10: icmp_seq=5 ttl=64 time=2.05 ms

--- 10.10.1.10 ping statistics ---
5 packets transmitted, 5 received, +4 duplicates, 0% packet loss, time 47ms
rtt min/avg/max/mdev = 2.051/8.508/18.666/6.003 ms, pipe 2, ipg/ewma 11.791/9.453 ms
R3#ping 10.10.1.12
PING 10.10.1.12 (10.10.1.12) 72(100) bytes of data.
80 bytes from 10.10.1.12: icmp_seq=1 ttl=64 time=4.83 ms
80 bytes from 10.10.1.12: icmp_seq=2 ttl=64 time=1.31 ms
80 bytes from 10.10.1.12: icmp_seq=1 ttl=64 time=9.45 ms (DUP!)
80 bytes from 10.10.1.12: icmp_seq=2 ttl=64 time=7.05 ms (DUP!)
80 bytes from 10.10.1.12: icmp_seq=3 ttl=64 time=5.05 ms
80 bytes from 10.10.1.12: icmp_seq=3 ttl=64 time=8.24 ms (DUP!)
80 bytes from 10.10.1.12: icmp_seq=4 ttl=64 time=4.34 ms
80 bytes from 10.10.1.12: icmp_seq=4 ttl=64 time=8.24 ms (DUP!)
80 bytes from 10.10.1.12: icmp_seq=5 ttl=64 time=3.97 ms
```

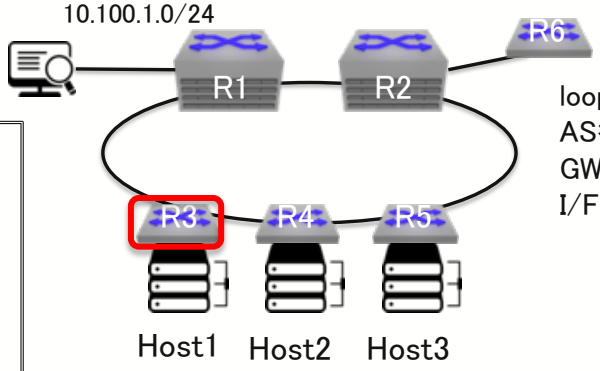


loopback:10.255.255.x/32
AS番号:64550
GW:10.10.1.254/24
I/F:10.10.1.x/24

Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

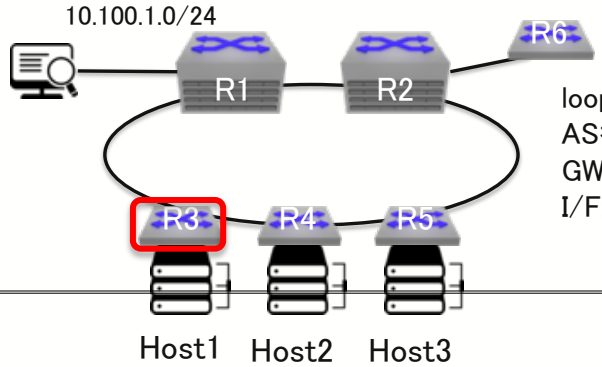
```
R3#show ip arp
Address         Age (sec)  Hardware Addr   Interface
10.255.255.1      0:00:02  001c.73c0.c601  Ethernet1
10.255.255.4      2:50:22  001c.73c3.c601  Ethernet2
10.10.1.10        2:50:31  001c.73c7.c611  Vlan100, Ethernet3
10.10.1.11        0:00:03  001c.73c7.c611  Vlan100, Ethernet3
10.10.1.12        0:00:05  001c.73c7.c611  Vlan100, Ethernet3
R3#show ip attached-host route export
   IP Address       VLAN       Interface       Prefix           VRF
---------------- ---------- --------------- ------------------- -------
   10.10.1.12       100        Ethernet3       10.10.1.12/32    default
   10.10.1.10       100        Ethernet3       10.10.1.10/32    default
   10.10.1.11       100        Ethernet3       10.10.1.11/32    default
```



loopback:10.255.255.x/32
AS番号:64550
GW:10.10.1.254/24
I/F:10.10.1.x/24

Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP



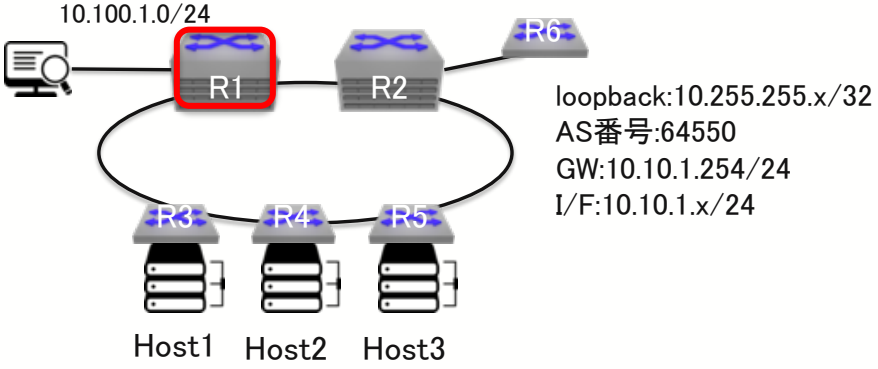
loopback:10.255.255.x/32
AS番号:64550
GW:10.10.1.254/24
I/F:10.10.1.x/24

```
R3#show ip bgp
BGP routing table information for VRF default
Router identifier 10.255.255.3, local AS number 64500
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
 * >      10.10.1.3/32           -                     0       0       -       i
 * >      10.10.1.4/32           10.255.255.4          0       100     0       i Or-ID: 10.255.255.4 C-LST: 10.255.255.6
 * >      10.10.1.5/32           10.255.255.5          0       100     0       i Or-ID: 10.255.255.5 C-LST: 10.255.255.6
 * >      10.10.1.10/32          -                     0       0       -       ?
 * >      10.10.1.11/32          -                     0       0       -       ?
 * >      10.10.1.12/32          -                     0       0       -       ?
R3#
```


Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

```
R1#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 - BGP, 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,
       DP - Dynamic Policy Route, L - VRF Leaked

 B I      10.10.1.3/32 [200/0] via 10.255.255.3, Ethernet2
 B I      10.10.1.4/32 [200/0] via 10.255.255.3, Ethernet2
 B I      10.10.1.5/32 [200/0] via 10.255.255.2, Ethernet1
 B I      10.10.1.10/32 [200/0] via 10.255.255.3, Ethernet2
 B I      10.10.1.11/32 [200/0] via 10.255.255.3, Ethernet2
 B I      10.10.1.12/32 [200/0] via 10.255.255.3, Ethernet2
```

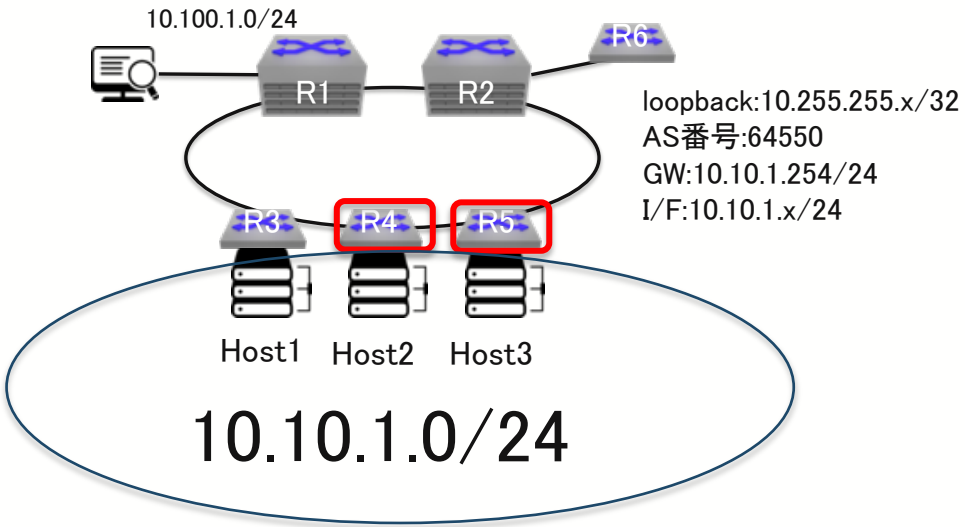


Public. Copyright © Arista 2018. All rights reserved.

ARP converted Host Routes injection into BGP

```
R4#show startup-config section virtual-router
interface Vlan100
   ip address 10.10.1.4/24
   ip helper-address 10.100.1.1
   ip attached-host route export
   ip virtual-router address 10.10.1.254
!
ip virtual-router mac-address 00:1c:73:00:00:99
```

```
R5#show startup-config section virtual-router
interface Vlan100
   ip address 10.10.1.5/24
   ip helper-address 10.100.1.1
   ip attached-host route export
   ip virtual-router address 10.10.1.254
!
ip virtual-router mac-address 00:1c:73:00:00:99
```



- GWはDHCP Relayの為にアドバタイズしておく
- セグメントは違う為に同一GWアドレスを使用可能であるため、VARPもしくはVRRPなどで共有化出来、IPアドレスを無駄なく使用できる

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

おまけ ルータリンク127ビットのお話

RFC6164
RFC 3627 to Historic Status

Network Working Group
Request for Comments: 3627
Category: Informational

Use of /127 Prefix Length Between Routers Considered

Status of this Memo

This memo provides information for the Internet communit
not specify an Internet standard of any kind. Distribut
memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights R

Abstract

In some cases, the operational decision may be to use IP
prefix lengths, especially on point-to-point links betwe
Under certain situations, this may lead to one router cl
addresses due to subnet-router anycast being implemented
document discusses the issue and offers a couple of solu
problem; nevertheless, /127 should be avoided between tw

1. Introduction

Internet Engineering Task Force (IETF)
Request for Comments: 6547
Obsoletes: 3627
Updates: 6164
Category: Informational
ISSN: 2070-1721

W. George
Time Warner Cable
February 2012

RFC 3627 to Historic Status

Abstract

This document moves "Use of /127 Prefix Length Between Routers Considered Harmful" (RFC 3627) to Historic status to reflect the updated guidance contained in "Using 127-Bit IPv6 Prefixes on Inter-Router Links" (RFC 6164). A Standards Track document supersedes an informational document; therefore, guidance provided in RFC 6164 is to be followed when the two documents are in conflict. This document links the two RFCs so that the IETF's updated guidance on this topic is clearer.

Internet Engineering Task Force (IETF)
Request for Comments: 6164
Track

M. Kohno
Juniper Networks, Keio University
B. Nitzan
Juniper Networks
R. Bush
Y. Matsuzaki
Internet Initiative Japan
L. Colitti
Google
T. Narten
IBM Corporation
April 2011

27-Bit IPv6 Prefixes on Inter-Router Links

point-to-point links, it is useful, for security and
o use 127-bit IPv6 prefixes. Such a practice
se of 31-bit prefixes in IPv4. This document
otivation for, and usages of, 127-bit IPv6 prefix
-router point-to-point links.

Status of This Memo

RFC3627
/127 Prefix Length Considered Harmful

RFC6167
Using 127-Bit IPv6 Prefixes on Inter-Router Links

- RFC3627がある為、問題に？
- RFC6167を新たに作成
- 紛らわしいので、RFC3627をヒストリックに

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

まとめ

- ルータにつけるIPアドレス、扱うIPアドレスの話をした
- 実際の運用例などで考慮するべきことなどあればさらに議論をすすめたい
- リンクローカルのみで運用してますか？グローバルアドレス付けてますか？
- そのアドレス有効的に使えてますか？

Public. Copyright © Arista 2018. All rights reserved.

参考URL

- BGP IPv6 Link Local Peers Discovery –BGP Unnumberedに対するAristaの解–
  - https://eos.arista.com/ja/bgp-ipv6-link-local-peers-discovery/
- BGP IPv6 Link Local Peers Discovery
  - https://eos.arista.com/eos-4-21-3f/bgp-ipv6-link-local-peers-discovery/
- AS Ranges for Dynamic BGP Peer Groups
  - https://eos.arista.com/eos-4-18-1f/bgp-as-ranges/
- IPv4 Unnumbered Interfaces
  - https://eos.arista.com/eos-4-23-0f/ipv4-unnumbered-interfaces/
- ARP converted Host Routes injection into BGP
  - https://eos.arista.com/eos-4-20-1f/hostinject/
- IPv6 and VRF support for ARP converted Host routes injection into BGP
  - https://eos.arista.com/eos-4-21-1f/ipv6-and-vrf-support-for-arp-converted-host-routes-injection-into-bgp/
- デモンストレーション RFC5309 unnumbered Link State Protocol/ARP converted Host Routes injection into BGP
  - https://www.youtube.com/watch?v=tff-ixgE0Ds&feature=youtu.be

Public. Copyright © Arista 2018. All rights reserved.
ARISTA

Thank You

www.arista.com

Public. Copyright © Arista 2018. All rights reserved.
ARISTA