



# JANOGで話題になったBGPネタのその後

Shishio Tsuchiya

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

# 本セッションの狙い

- JANOGで過去に議論
- ソフトウェアなどが拡張
- 便利になったのに、誰も気づいてない。。。
- をまとめて説明

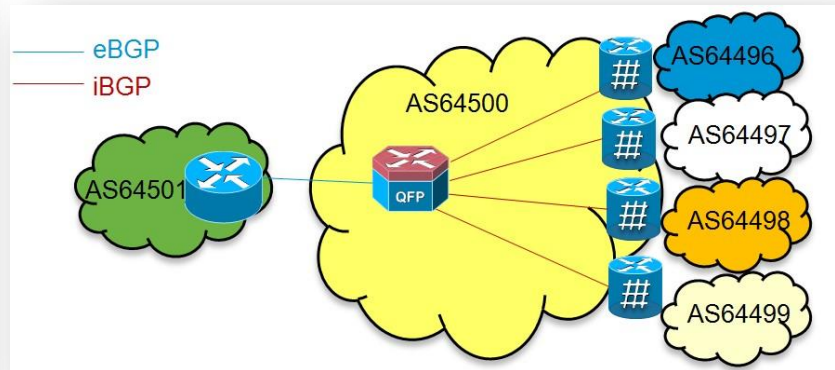
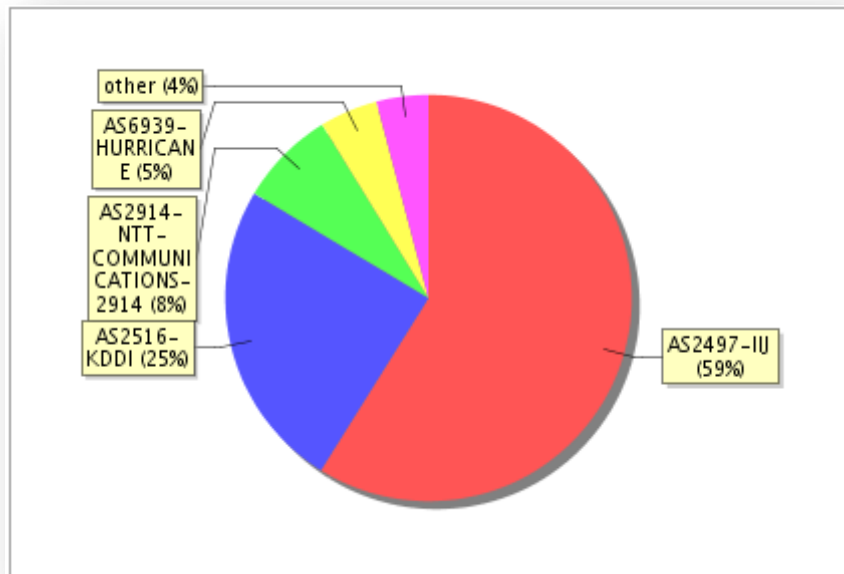
# Agenda

- **BGP Selective Download**
- BGP 4byte Private AS
- BMPサーバー
- BGP Flowspec

# 関連トピック

- JANOG27
  - FIB table saving technique (with Simple Virtual Aggregation)
- JANOG34
  - Fullroute やめたら人生変わった@AS55394
- JANOG36 BoF
  - フローコレクターとルーティングの連携
- JANOG LT night#1
  - うちもフルルートやめてみたその後

# どうい話？フルルートは本当に必要か？

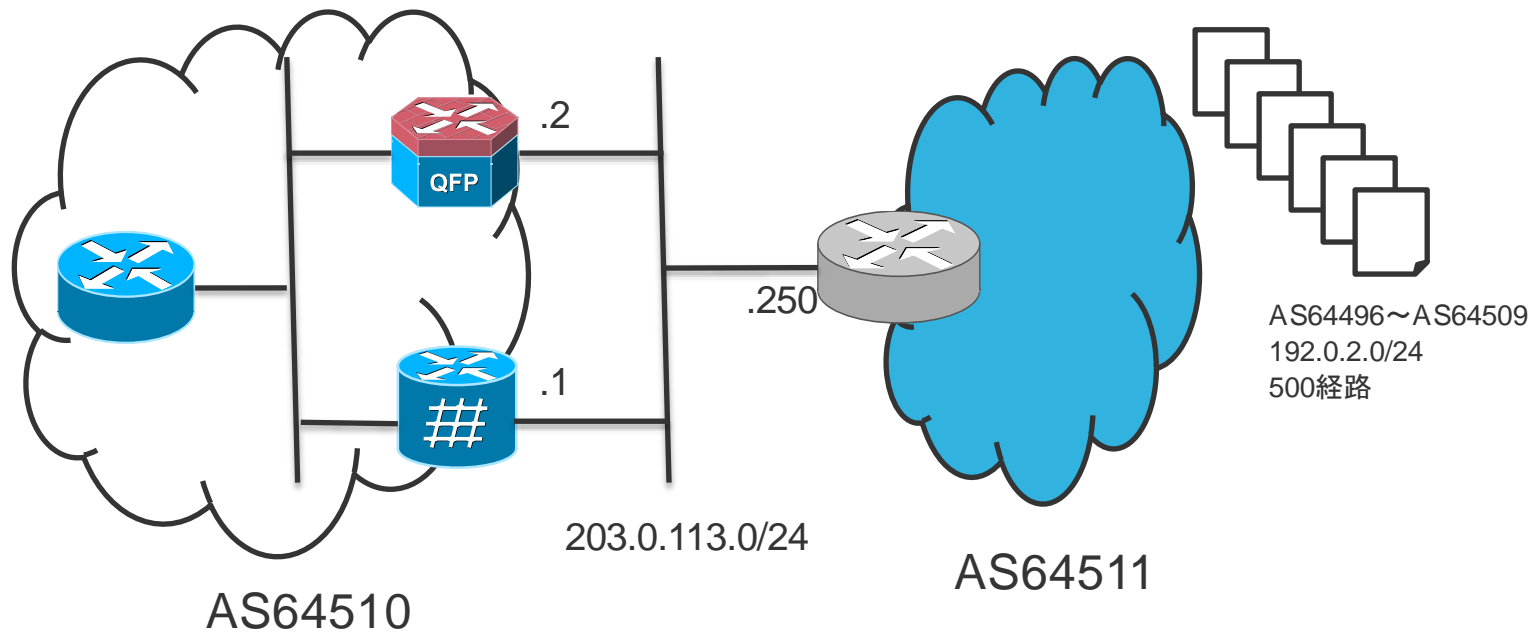


- 下位プロバイダーへのフルルート提供の為には必要
- [RIPE AS dashboard](#)のtransits distributionではパス情報から、偏りを見る事が出来る。

# JANOGで発表があった各社のアプローチ

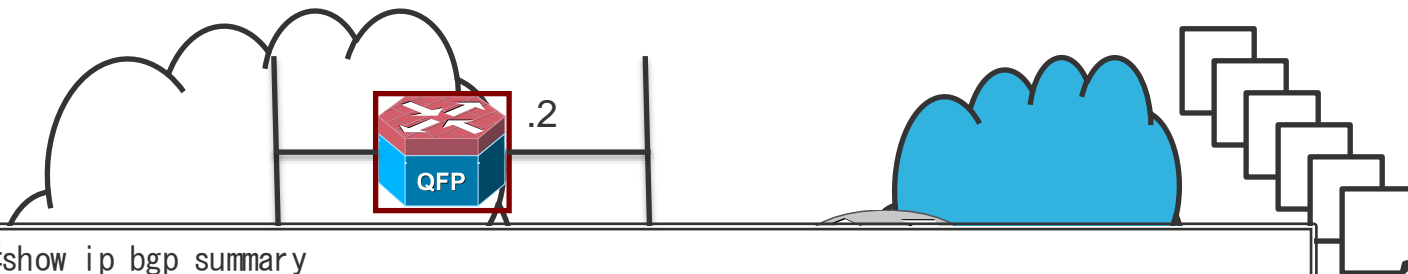
JANOG	発表者	ステータス
JANOG27 <a href="#">FIB table saving technique (with Simple Virtual Aggregation)</a>	小川 奈夏 (シスコシステムズ合同会社)	<a href="#">Grow WG=&gt;Dead=&gt;RFC6769</a>
JANOG34 <a href="#">Fullrouteやめたら人生変わった@AS55394</a>	黒河内 倫 (フリー株式会社)	上位プロバイダーへ依頼 5AS:1800経路
JANOG36 <a href="#">フローコントローラータイプの連携(Spotly use-case)</a>	Paolo Lucente (pmacct Project)	TopN Flowに対して bgp selective downloadを実施
JANOG LT Night#1 <a href="#">うちもフルルートやめてみたその後</a>	吉野 純平 (株式会社ミクシィ)	bgp no-advertiseでASBR以外は Partial Routeしか持たない

# BGP Selective Download テストポロジ



# BGP Selective Download

## *show ip bgp summary*



```
IOS-XE#show ip bgp summary
BGP router identifier 10.255.255.2, local AS number 64510
BGP table version is 7501, main routing table version 7501
500 network entries using 124000 bytes of memory
500 path entries using 60000 bytes of memory
25/25 BGP path/bestpath attribute entries using 6200 bytes of memory
25 BGP AS-PATH entries using 1240 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 191440 total bytes of memory
BGP activity 1500/1000 prefixes, 4000/3500 paths, scan interval 60 secs
```

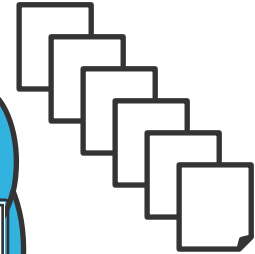
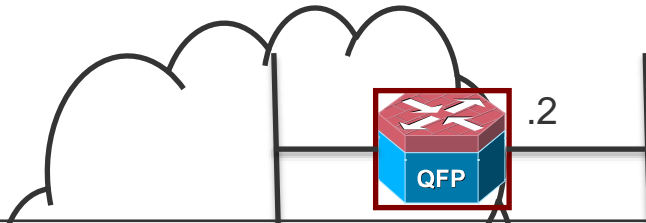
Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
10.255.255.255	4	64510	180	579	7501	0	0	02:41:15	0
203.0.113.250	4	64511	71	31	7501	0	0	00:25:17	500

4496~AS64509  
0.2.0/24  
怪路



# BGP Selective Download

## *show ip route summary*



AS64496~AS64509  
192.0.2.0/24  
500経路

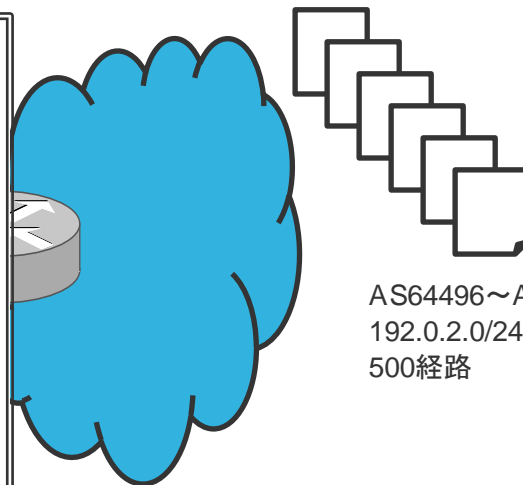
```
IOS-XE#show ip route summary
IP routing table name is default (0x0)
IP routing table maximum-paths is 32
Route Source    Networks    Subnets    Replicates    Overhead    Memory (bytes)
application      0           0           0             0           0
connected        0           5           0            480        1440
static           0           0           0             0           0
ospf 1           0           3           0            288        876
  Intra-area: 3 Inter-area: 0 External-1: 0 External-2: 0
  NSSA External-1: 0 NSSA External-2: 0
bgp 64510      0         500       0          48000     144000
  External: 500 Internal: 0 Local: 0
internal         4           508         0            48768     188388
Total            4           508         0            48768     188388
```

# BGP Selective Download

## *show ip bgp*

```
IOS-XE#show ip bgp
BGP table version is 7501, local router ID is 10.255.255.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.0.2.0/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.1/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.2/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.3/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.4/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.5/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.6/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.7/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.8/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.9/32	203.0.113.250	1657		0	64511 64499 64506 64512 e
*> 192.0.2.10/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.11/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.12/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.13/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.14/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.15/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.16/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.17/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e
*> 192.0.2.18/32	203.0.113.250	2073		0	64511 64508 64502 64504 64503 64509 e



AS64496~AS64509  
192.0.2.0/24  
500経路

# BGP Selective Download

## *show ip bgp \_64500\$*

```
IOS-XE#show ip bgp regexp _64500$
BGP table version is 7501, local router ID is 10.255.255.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

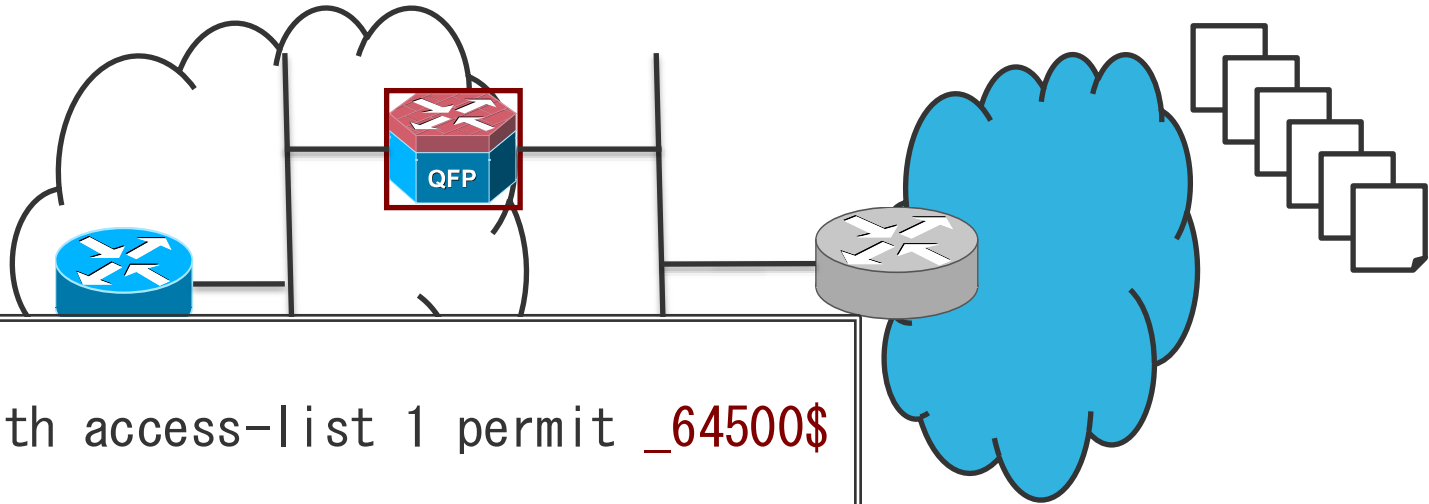
   Network          Next Hop           Metric LocPrf Weight Path
*> 192.0.2.110/32    203.0.113.250      1214             0 64511 64504 64500 e
*> 192.0.2.111/32    203.0.113.250      1214             0 64511 64504 64500 e
*> 192.0.2.112/32    203.0.113.250      1214             0 64511 64504 64500 e
*> 192.0.2.113/32    203.0.113.250      1214             0 64511 64504 64500 e
-snip-
*> 198.51.100.115/32
                   203.0.113.250      1214             0 64511 64504 64500 e
*> 198.51.100.116/32
                   203.0.113.250      1214             0 64511 64504 64500 e
*> 198.51.100.117/32
                   203.0.113.250      1214             0 64511 64504 64500 e
*> 198.51.100.118/32
                   203.0.113.250      1214             0 64511 64504 64500 e
*> 198.51.100.119/32
                   203.0.113.250      1214             0 64511 64504 64500 e
```



A S64496~A S64509  
192.0.2.0/24  
500経路

特定のAS64500をoriginにしてるもの  
のみにすると20経路になる

# BGP Selective Download route-map



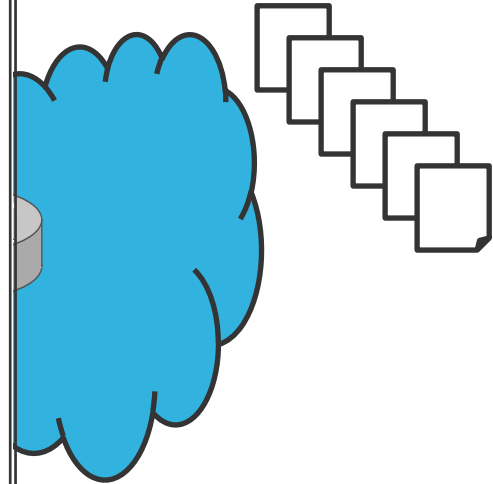
```
!  
ip as-path access-list 1 permit _64500$  
!  
route-map AS64500origin permit 10  
  match as-path 1  
!  
route-map AS64500origin deny 20
```

# BGP Selective Download

## *table-map*反映前

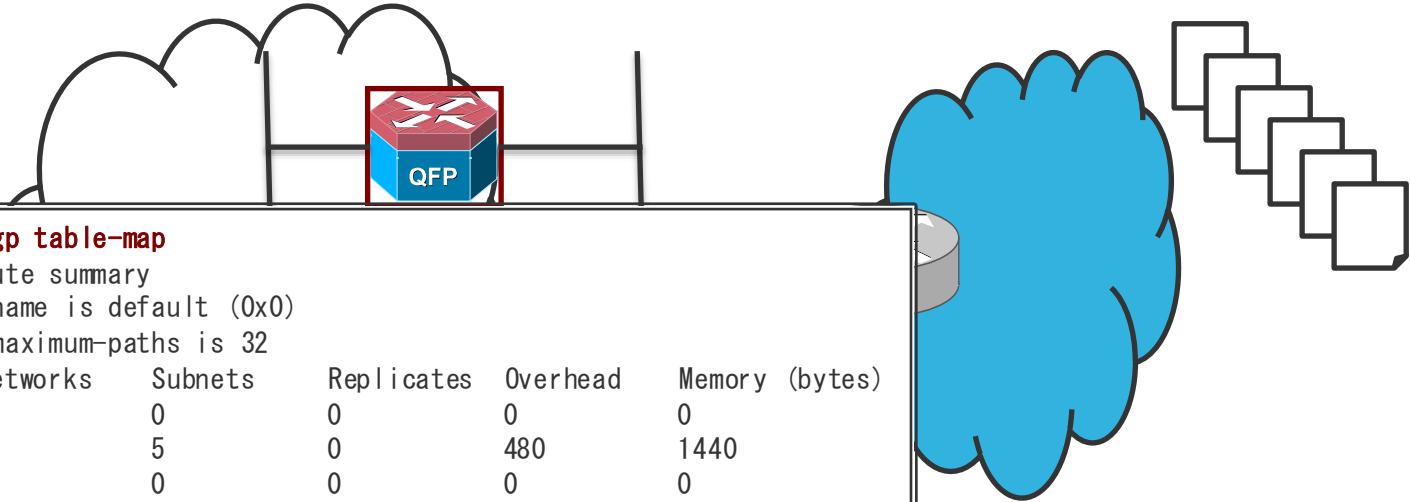
```
IOS-XE#conf t
Enter configuration commands, one per line.  End with CNTL/Z.
IOS-XE(config)#router bgp 64510
IOS-XE(config-router)#address-family ipv4 unicast
IOS-XE(config-router-af)#table-map AS64500origin filter
IOS-XE(config-router-af)#end
IOS-XE#show ip route summary
IP routing table name is default (0x0)
IP routing table maximum-paths is 32
```

Route Source	Networks	Subnets	Replicates	Overhead	Memory (bytes)
application	0	0	0	0	0
connected	0	5	0	480	1440
static	0	0	0	0	0
ospf 1	0	3	0	288	876
Intra-area: 3 Inter-area: 0 External-1: 0 External-2: 0					
NSSA External-1: 0 NSSA External-2: 0					
bgp 64510	0	500	0	48000	144000
External: 500 Internal: 0 Local: 0					
internal	4				42072
Total	4	508	0	48768	188388



# BGP Selective Download

## *table-map*反映後

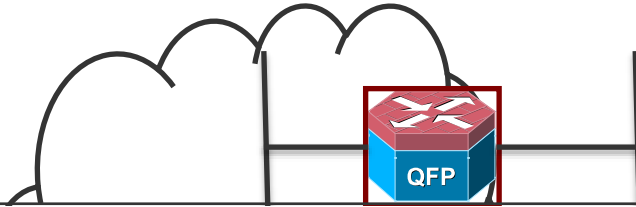


```
IOS-XE#clear ip bgp table-map
IOS-XE#show ip route summary
IP routing table name is default (0x0)
IP routing table maximum-paths is 32
```

Route Source	Networks	Subnets	Replicates	Overhead	Memory (bytes)
application	0	0	0	0	0
connected	0	5	0	480	1440
static	0	0	0	0	0
ospf 1	0	3	0	288	876
Intra-area: 3 Inter-area: 0 External-1: 0 External-2: 0					
NSSA External-1: 0 NSSA External-2: 0					
<b>bgp 64510</b>	<b>0</b>	<b>20</b>	<b>0</b>	<b>1920</b>	<b>5760</b>
External: 20 Internal: 0 Local: 0					
internal	4				3672
Total	4	28	0	2688	11748

# BGP Selective Download

## show ip bgp summary



```
IOS-XE#show ip bgp summary
BGP router identifier 10.255.255.2, local AS number 64510
BGP table version is 7501, main routing table version 7501
500 network entries using 124000 bytes of memory
500 path entries using 60000 bytes of memory
25/25 BGP path/bestpath attribute entries using 6200 bytes of memory
25 BGP AS-PATH entries using 1240 bytes of memory
25 BGP route-map cache entries using 1600 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 193040 total bytes of memory
BGP activity 1500/1000 prefixes, 4000/3500 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
10.255.255.255	4	64510	181	581	7501	0	0	02:42:25	0
203.0.113.250	4	64511	72	32	7501	0	0	00:26:27	500

# 影響の大きさ

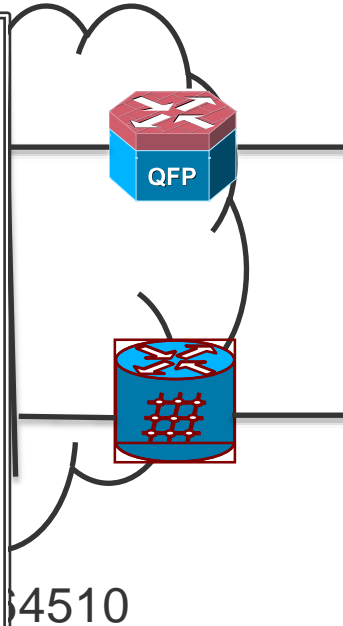


- BGP→RIB→FIBとなるので、FIB容量が少ないものでフォワーディング性能が良いものをエッジにする事が出来る
- 1M RIBが持てれば、FIBは数Kで良い
- 単純計算本テストでは1RIB辺り、 $144000/500=288$ byte
- メモリも少量で済む



# BGP Selective Download IOS-XR

```
RP/0/0/CPU0:IOS-XR#show rpl
Mon Jan  4 16:07:56.701 JST
as-path-set AS64500
  ios-regex '_64500$'
end-set
!
route-policy All-accept
  pass
end-policy
!
route-policy AS64500origin
  if as-path in AS64500 then
    pass
  else
    drop
  endif
end-policy
!
```



```
RP/0/0/CPU0:IOS-XR#show running-config router bgp
Mon Jan  4 16:08:08.396 JST
router bgp 64510
  address-family ipv4 unicast
    table-policy AS64500origin
  !
  neighbor 203.0.113.250
    remote-as 64511
    address-family ipv4 unicast
      route-policy All-accept in
      route-policy All-accept out
  !
  neighbor 10.255.255.255
    remote-as 64510
    update-source Loopback0
    address-family ipv4 unicast
      route-policy All-accept in
      route-policy All-accept out
    next-hop-self
  !
  !
  !
```

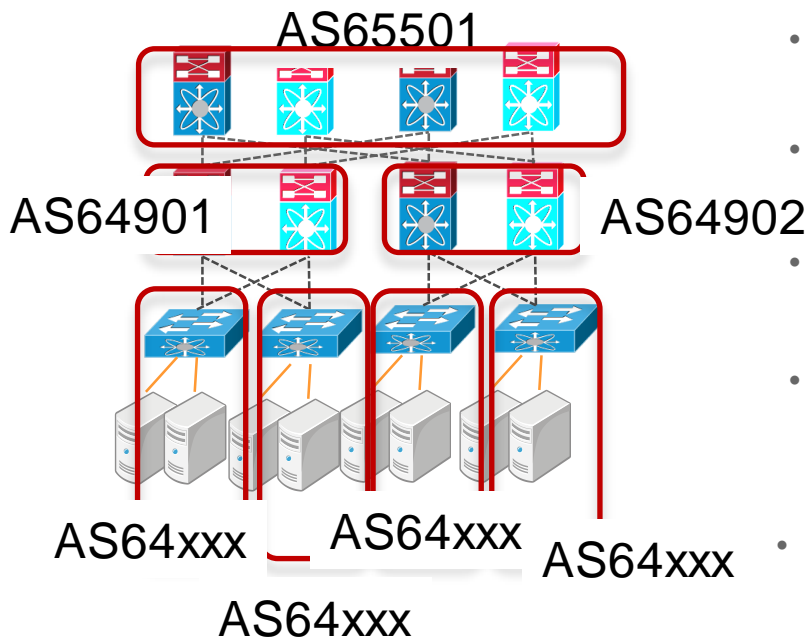
# まとめ

- table-map/table-policyはBGPテーブルに影響は与えずに、尚且つルートマップと同様なフォーマットでBGP RIBからFIBへのインストール制御が可能
- Simple-VAは[RFC6769](#)になった。(ほぼ同様)
- xFlow TopNやサービスモデルから考え、適切なプラットフォームを選考する事が可能

# Agenda

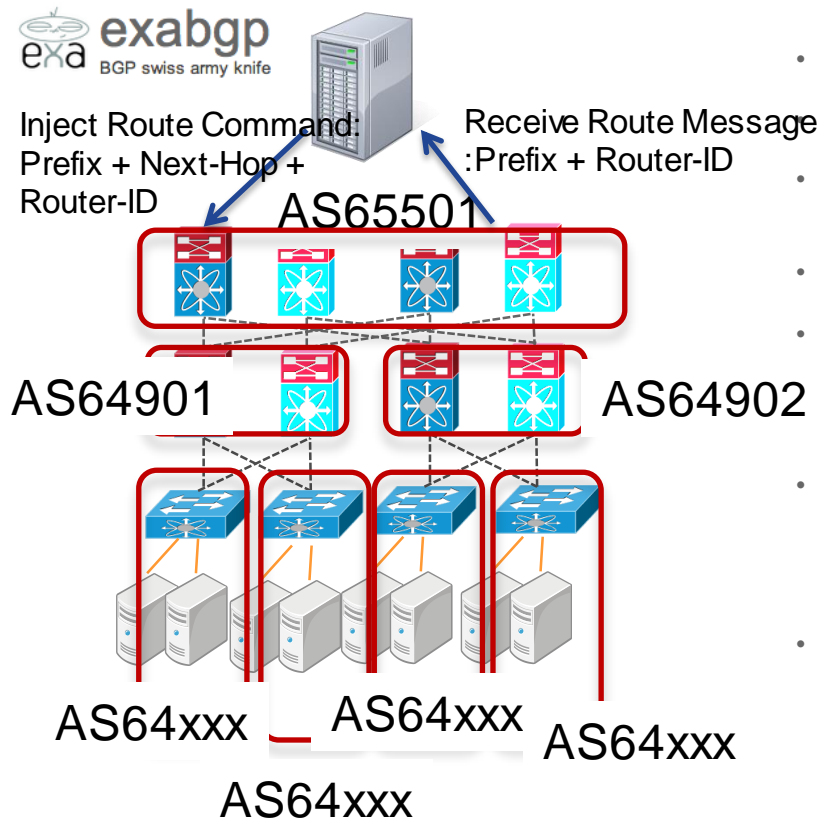
- BGP Selective Download
- **BGP 4byte Private AS**
- BMPサーバー
- BGP Flowspec

# Microsoft Routing Design for Large Scale Data Centers



- Microsoft Bingネットワークでは安定性とトラフィックエンジニアリングを目的とし、ToRまでのBGPを設定
- 各Rack毎にAS番号を使用し、トラブルシューティングを容易にした
- コンフィグが複雑!?
  - 自動化するので問題無い
- Use of BGP for routing in large-scale data centers
  - <https://tools.ietf.org/html/draft-ietf-rtgwg-bgp-routing-large-dc>
- NANOG55  
<http://www.nanog.org/meetings/nanog55/presentations/Monday/Lapukhov.pdf>

# BGP-Only SDN for Large-Scale Data-Centers



- ExaBGPを用いてコントローラーを作成
- BGP Peerとしてパス情報を受信
- 3rd Party next-hopsにて経路をアナウンスし、トラフィック・コントロール
- 複数のコントローラーではOpen Replicaを使用
- コントローラーのコードは<1000 LoC
- 中央制御に失敗しても、BGPレギュラールーティングにfail-back出来る
- Centralized Routing Control in BGP Networks Using Link-State Abstraction
- <http://tools.ietf.org/html/draft-lapukhov-bgp-sdn>
- NANOG58
- <http://www.nanog.org/sites/default/files/wed.general.brain.slug.lapukhov.20.pdf>

# 関連トピック

- JANOG33

- Experiences with BGP in large Scale Data Centers

<http://www.janog.gr.jp/meeting/janog33/program/bgp.html>

- Q: 32ビットのAS番号はまだ出来ないというコメントがあったが、実は既に使ったりするの？
- A: 実際的にはベンダーの実装に縛られるので、32ビットAS番号はまだ使っていない。

# RFC6996 32ビットPrivate AS

- RFC6996 Autonomous System (AS) Reservation for Private Use
- 16ビット64512 – 65534 1023個に 32ビット4200000000 – 4294967294 94,967,295追加
- IHANETも大喜び
  - <http://www.bgp.moe/technical-info/ASsAndPrefixes>
- 機器実装に何が必要だったか？

# 4. Operational Considerations

- RFC6793 BGP Support for 4-Octet AS Number Spaceをサポートする必要がある
- グローバルインターネットに出る際にはプライベートASNは削除される必要がある

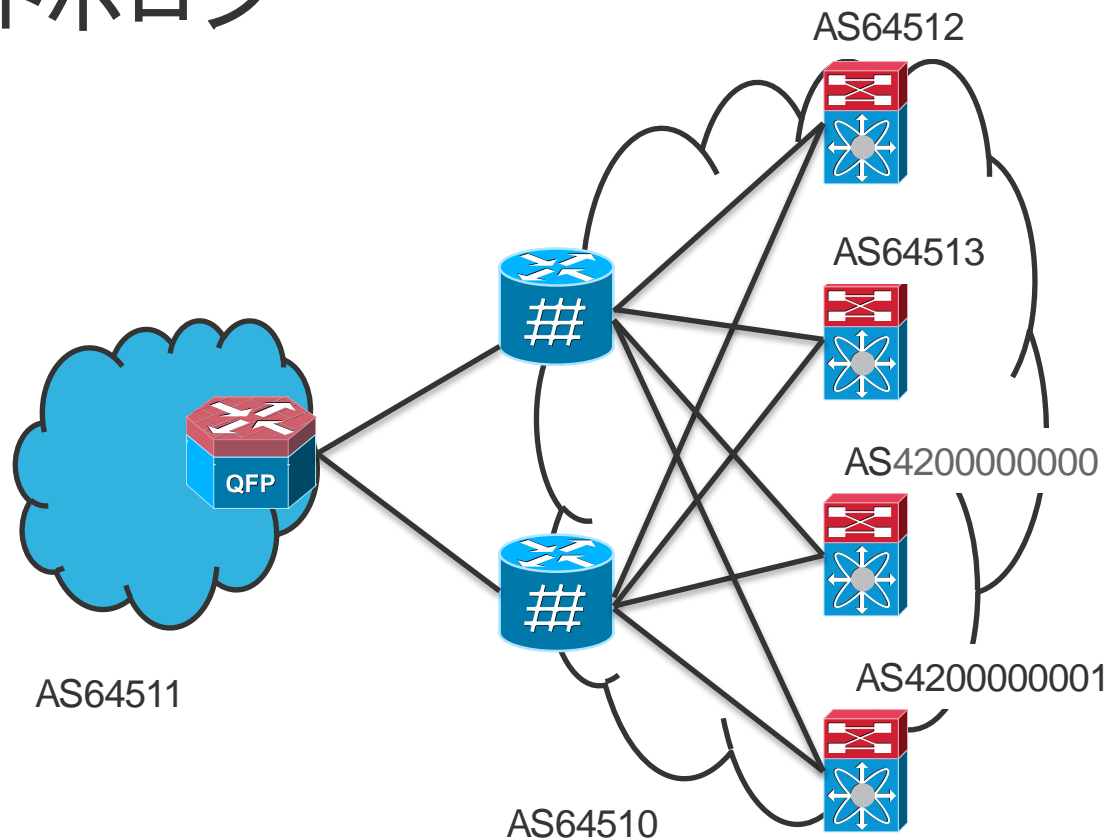


# private-as removal

```
!  
router bgp 65536  
!  
neighbor 172.20.1.1  
!  
remote-as 65551  
address-family ipv4 unicast  
remove-private-as  
!
```

- ASパスの中に含むPrivateASを自動的に削除
- この32bit対応が必要 (route-mapで代用可)

# Private AS Removal テストポロジ



# ここで復習

## JANOG26 IPアドレスの行方

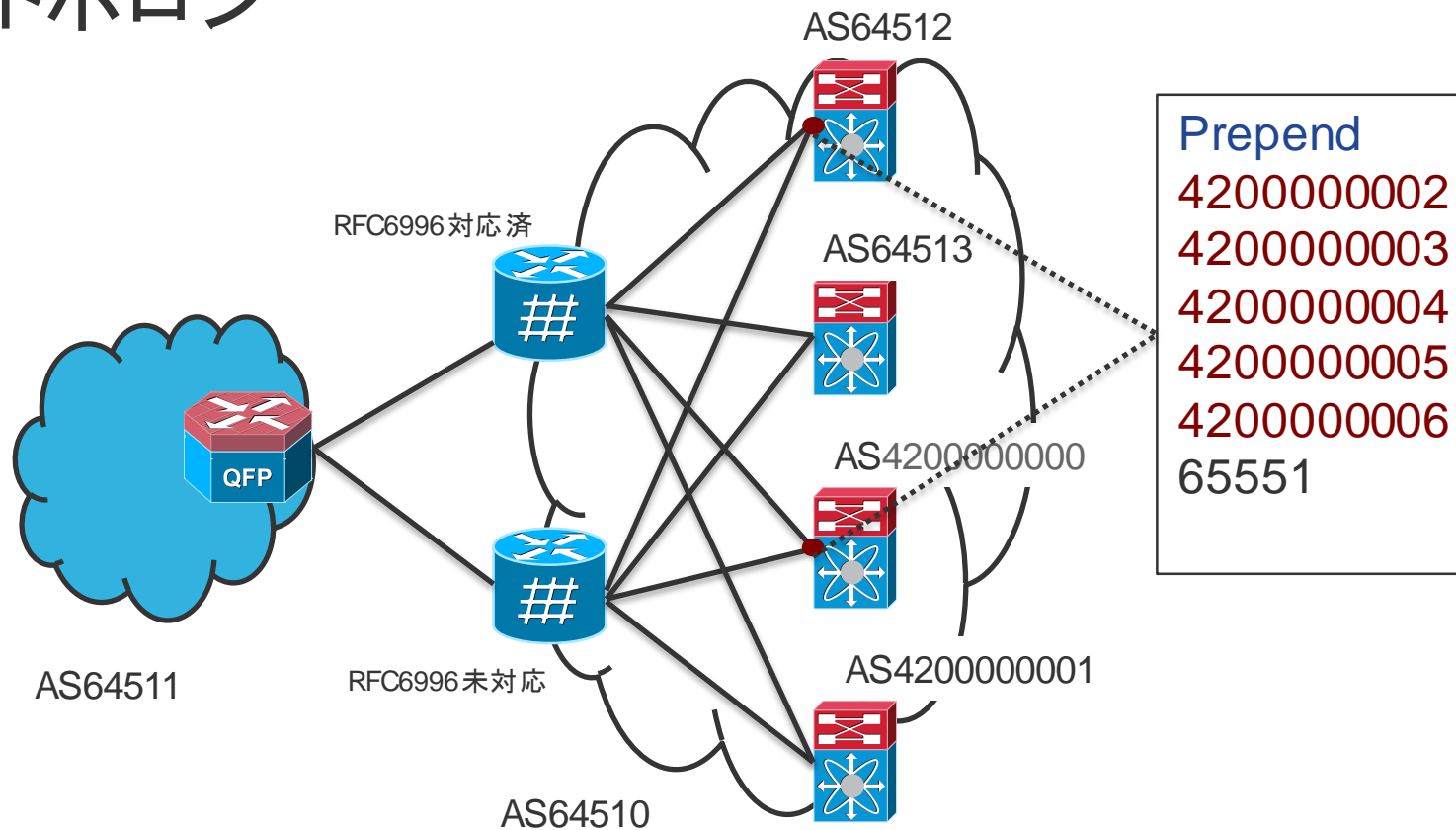
### ドキュメンテーションの為のガイドライン

用途	例	リファレンス
IPv4アドレス (グローバル)	192.0.2.0/24 (TEST-NET-1) 198.51.100.0/24(TEST-NET-2) 203.0.113.0/24(TEST-NET-3)	<a href="#">RFC5737</a>
IPv4アドレス (プライベート)	10.0.0.0/8 172.16.0.0/12 192.168.0.0/16	<a href="#">RFC1918</a>
IPv6アドレス	<a href="#">2001:db8::/32</a>	<a href="#">RFC3849</a>
DNS Name	.example example.com example.net example.org	<a href="#">RFC2606</a>
JPドメイン	EXAMPLE.JP EXAMPLE1.JP ドメイン名例.JP	<a href="#">JPRS_FAQ</a>
BGP ASN(グローバル)	64496 – 64511(16bit) 65536 – 65551(32bit)	<a href="#">RFC5398</a>
BGP ASN(プライベート)	64512-65534(16bit)	<a href="#">RFC1930</a>

# IPアドレス/AS番号アップデート

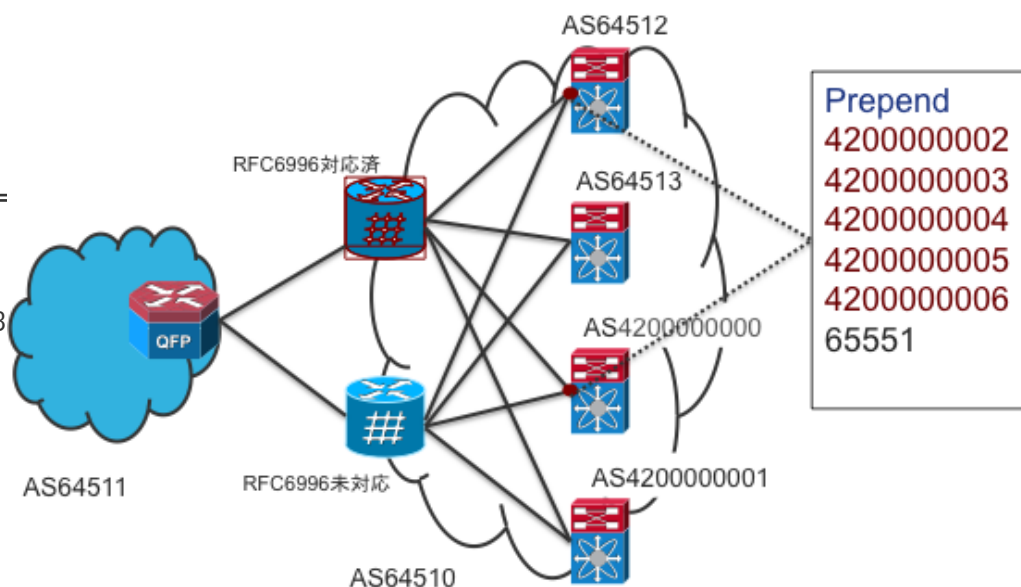
用途	例	リファレンス
IPv4アドレス (グローバル:文書用)	192.0.2.0/24 (TEST-NET-1) 198.51.100.0/24(TEST-NET-2) 203.0.113.0/24(TEST-NET-3)	<a href="#">RFC5737</a>
IPv4アドレス(プライベート)	10.0.0.0/8 172.16.0.0/12 192.168.0.0/16	<a href="#">RFC1918</a>
IPv4アドレス (ISP共有アドレス)	100.64.0.0/10.	<a href="#">RFC6598</a>
IPv6アドレス (文書用)	2001:db8::/32	<a href="#">RFC3849</a>
BGP ASN (グローバル:文書用)	64496 – 64511(16bit) 65536 – 65551(32bit)	<a href="#">RFC5398</a>
BGP ASN (プライベート)	64512-65534(16bit) 4200000000 – 4294967294 (32bit)	<a href="#">RFC1930</a> <a href="#">RFC6996</a>
APNICリサーチ用	1.0.0.0/24 1.1.1.0/24	<a href="#">APNIC Prop-109</a>

# Private AS Removal テストポロジ



# Private AS Removal BGPテーブル

```
RP/0/0/CPU0:XR1#show version | include IOS
Tue Jan  5 16:06:05.067 UTC
Cisco IOS XR Software, Version 6.0.0[Default]
cisco IOS XRv Series (Pentium Celeron Stepping 3) processor with 3
IOS XRv Chassis
RP/0/0/CPU0:XR1#show bgp
Tue Jan  5 16:06:10.627 UTC
BGP router identifier 10.255.255.1, local AS number 64510
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0xe0000000  RD version: 15
BGP main routing table version 15
BGP NSR Initial initsync version 2 (Reached)
BGP NSR/ISSU Sync-Group versions 0/0
BGP scan interval 60 secs
```

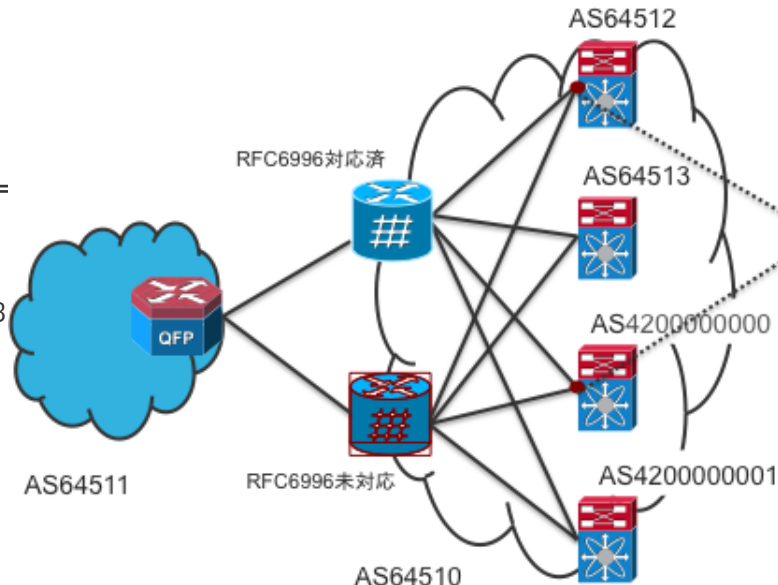


```
Status codes: s suppressed, d damped, h history, * valid, > best
                i - internal, r RIB-failure, S stale, N Nexthop-discard
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.255.255.101/32	10.10.10.2	0	64512	4200000002	4200000003 4200000004 4200000005 4200000006 65511 i
*> 10.255.255.102/32	10.10.10.10	0	64513	i	
*> 10.255.255.103/32	10.10.10.18	0	4200000000	4200000002	4200000003 4200000004 4200000005 4200000006 65511 i
*> 10.255.255.104/32	10.10.10.26	0	4200000001	I	

# Private AS Removal BGPテーブル

```
RP/0/0/CPU0:XR2#show version | include IOS
Tue Jan  5 16:06:46.394 UTC
Cisco IOS XR Software, Version 5.3.0[Default]
cisco IOS XRv Series (Pentium Celeron Stepping 3) processor with 3
IOS XRv Chassis
RP/0/0/CPU0:XR2#show bgp
Tue Jan  5 16:06:56.343 UTC
BGP router identifier 10.255.255.2, local AS number 64510
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0xe0000000  RD version: 12
BGP main routing table version 12
BGP NSR Initial initsync version 2 (Reached)
BGP NSR/ISSU Sync-Group versions 0/0
BGP scan interval 60 secs
```

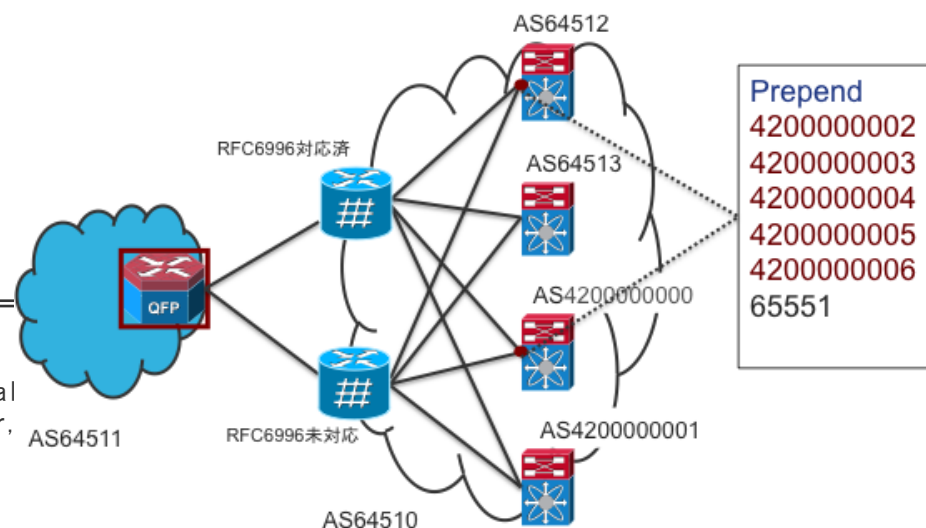


```
Prepend
4200000002
4200000003
4200000004
4200000005
4200000006
65551
```

```
Status codes: s suppressed, d damped, h history, * valid, > best
i - internal, r RIB-failure, S stale, N Nexthop-discard
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.255.255.101/32	10.10.10.6	0	64512	4200000002	4200000003 4200000004 4200000005 4200000006 65551 i
*> 10.255.255.102/32	10.10.10.14	0	64513	i	
*> 10.255.255.103/32	10.10.10.22	0	4200000000	4200000002	4200000003 4200000004 4200000005 4200000006 65551 i
*> 10.255.255.104/32	10.10.10.30	0	4200000001	i	

# Private AS Removal BGPテーブル

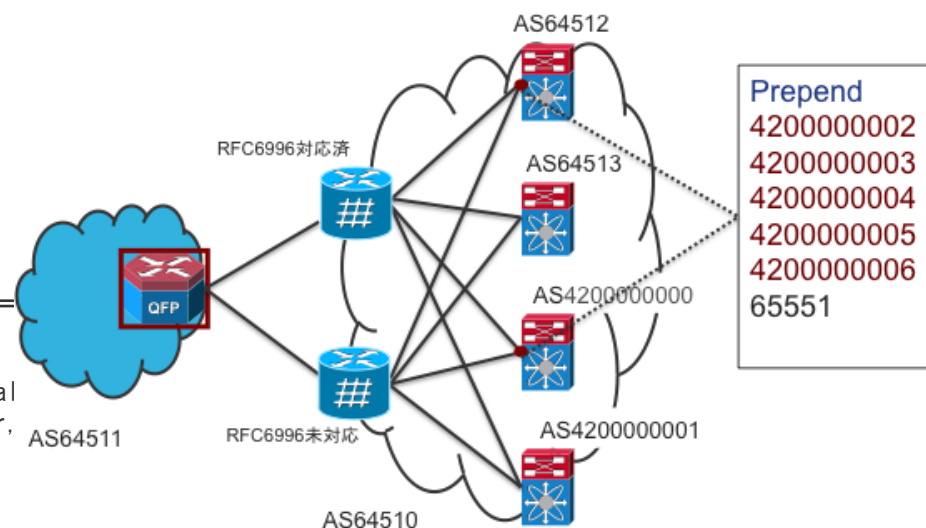


```
IOS-XE#show bgp
BGP table version is 25, local router ID is 10.255.255.255
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.255.255.101/32	10.10.1.6	0	64510	64512	4200000002 4200000003 4200000004 4200000005 4200000006 65551 i
* 10.255.255.102/32	10.10.1.2	0	64510	64512	4200000002 4200000003 4200000004 4200000005 4200000006 65551 i
*> 10.255.255.103/32	10.10.1.6	0	64510	64513	i
* 10.255.255.104/32	10.10.1.2	0	64510	64513	i
*> 10.255.255.105/32	10.10.1.6	0	64510	4200000000	4200000002 4200000003 4200000004 4200000005 4200000006 65551 i
* 10.255.255.106/32	10.10.1.2	0	64510	4200000000	4200000002 4200000003 4200000004 4200000005 4200000006 65551 i
*> 10.255.255.107/32	10.10.1.6	0	64510	4200000001	i
* 10.255.255.108/32	10.10.1.2	0	64510	4200000001	i



# Private AS Removal 実施



```
IOS-XE#show bgp
BGP table version is 29, local router ID is 10.255.255.255
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

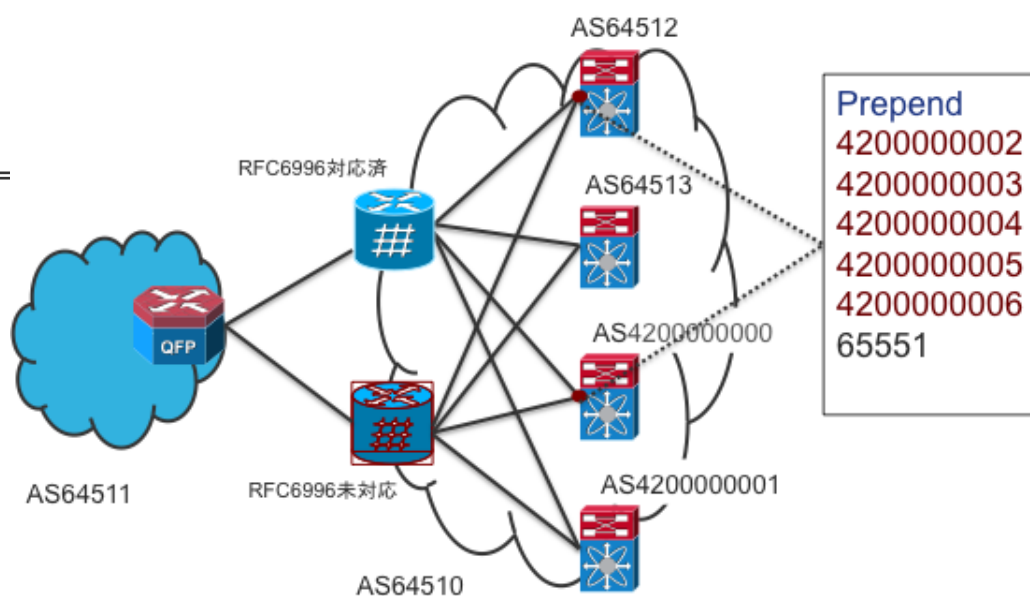
Network	Next Hop	Metric	LocPrf	Weight	Path
* 10.255.255.101/32	10.10.1.6	0	64510	4200000002	4200000003 4200000004 4200000005 4200000006 65551 i
*>	10.10.1.2	0	64510	65551	i
* 10.255.255.102/32	10.10.1.6	0	64510		i
*>	10.10.1.2	0	64510		i
* 10.255.255.103/32	10.10.1.6	0	64510	4200000000	4200000002 4200000003 4200000004 4200000005 4200000006 65551 i
*>	10.10.1.2	0	64510	65551	i
* 10.255.255.104/32	10.10.1.6	0	64510	4200000001	i
*>	10.10.1.2	0	64510		i

未対応側では16ビットは削除されているが  
32ビットは削除されてない

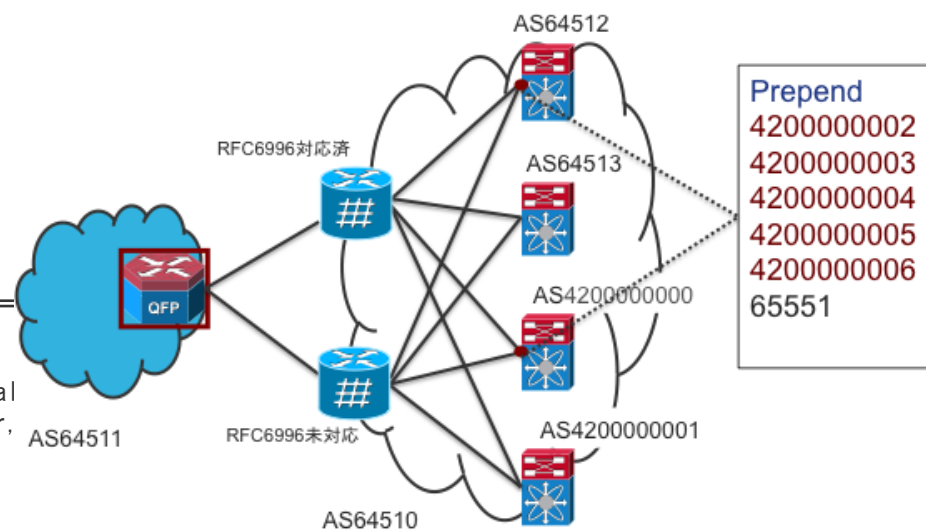
# Private AS Removal ワークアラウンド!?

```
P/0/0/CPU0:XR2#show rpl
Tue Jan  5 16:33:03.156 UTC
route-policy All
  pass
end-policy
!
route-policy RFC6996
  replace as-path '[4200000000..4294967294]'
end-policy
!
```

```
RP/0/0/CPU0:XR2#conf t
Tue Jan  5 16:33:04.636 UTC
RP/0/0/CPU0:XR2(config)#router bgp 64510
RP/0/0/CPU0:XR2(config-bgp)# neighbor 10.10.1.5
RP/0/0/CPU0:XR2(config-bgp-nbr)#address-family ipv4 unicast
RP/0/0/CPU0:XR2(config-bgp-nbr-af)#route-policy RFC6996 out
RP/0/0/CPU0:XR2(config-bgp-nbr-af)#commit
Tue Jan  5 16:33:24.815 UTC
RP/0/0/CPU0:XR2(config-bgp-nbr-af)#
```



# Private AS Removal ワークアラウンド!?



```
IOS-XE#show bgp
BGP table version is 29, local router ID is 10.255.255.255
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
* 10.255.255.101/32	10.10.1.6	0	64510	64510	64510 64510 64510 64510 65551 i
*>	10.10.1.2	0	64510	65551	i
* 10.255.255.102/32	10.10.1.6	0	64510	i	
*>	10.10.1.2	0	64510	i	
* 10.255.255.103/32	10.10.1.6	0	64510	64510	64510 64510 64510 64510 65551 i
*>	10.10.1.2	0	64510	65551	i
* 10.255.255.104/32	10.10.1.6	0	64510	64510	i
*>	10.10.1.2	0	64510	i	

見えてはいない

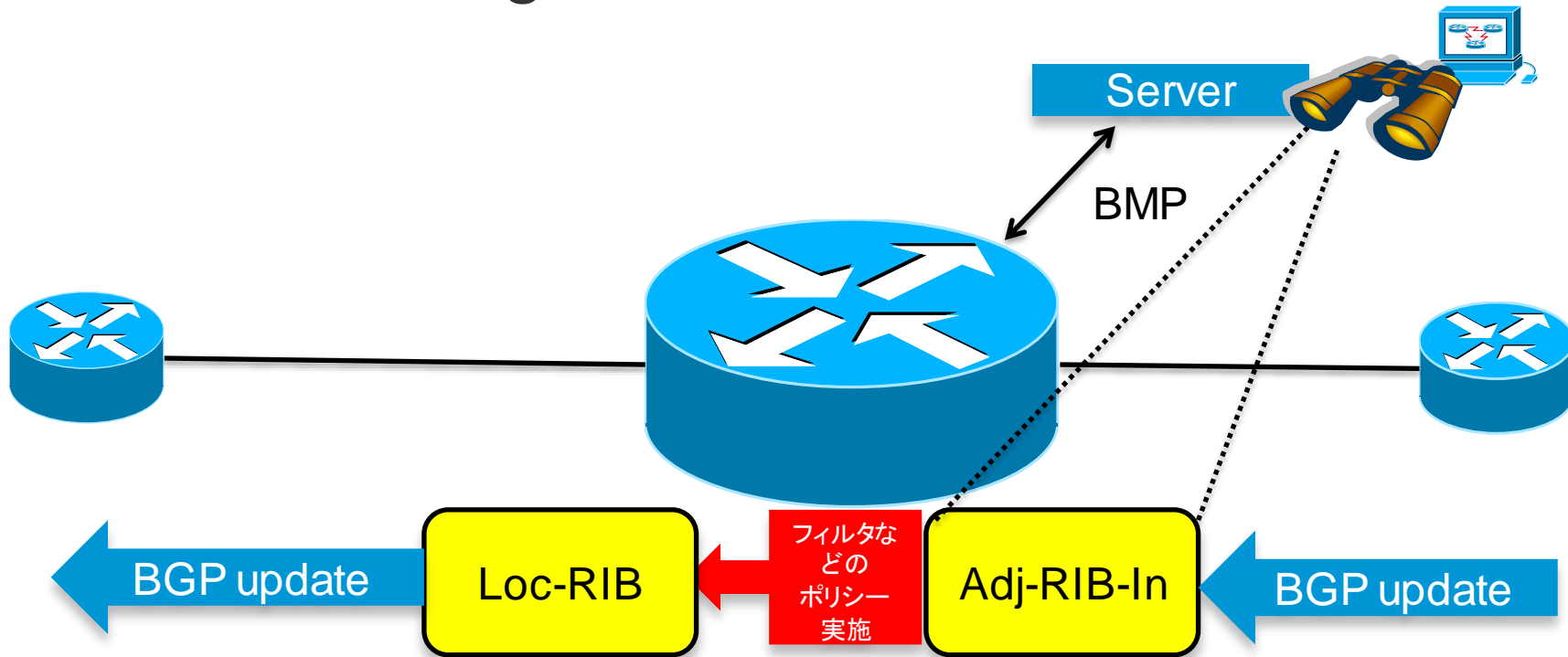
# まとめ

- マイクロソフトの要求により4バイトASのプライベート番号はアサインされた
- RFC6996にて定義 4200000000 – 4294967294
- 4バイトAS対応機器であれば設定は特に意識する必要は無い
- インターネット接続の際には削除もしくは置き換えを忘れずに

# Agenda

- BGP Selective Download
- BGP 4byte Private AS
- **BMPサーバー**
- BGP Flowspec

# BGP Monitoring Protocol



# 関連トピック

- JANOG33.5

BGP Monitoring Protocol

<http://www.janog.gr.jp/meeting/janog33.5/program/bmp.html>

- JANOG34

ルーティングテーブルを覗きたい | \_・)ジー

<http://www.janog.gr.jp/meeting/janog34/program/rtabl.html>

IOS-XEとJUNOSの相互接続を実施

BMPサーバーが動く実装が無い事を報告

# その後

- Ryu BMPサーバーの登場

<http://osrg.github.io/bmp/>

- APNIC38 BMP(BGP Monitoring Protocol) Testing by \*JANOGers

<http://www.slideshare.net/apnic/bmp-shishio-1410405612>

<http://www.slideshare.net/apnic/bmpbgp-monitoring-protocol-testing-by-janogers-2-1410809850>

さらに

Open BGP Monitoring Protocol (OpenBMP) Collector

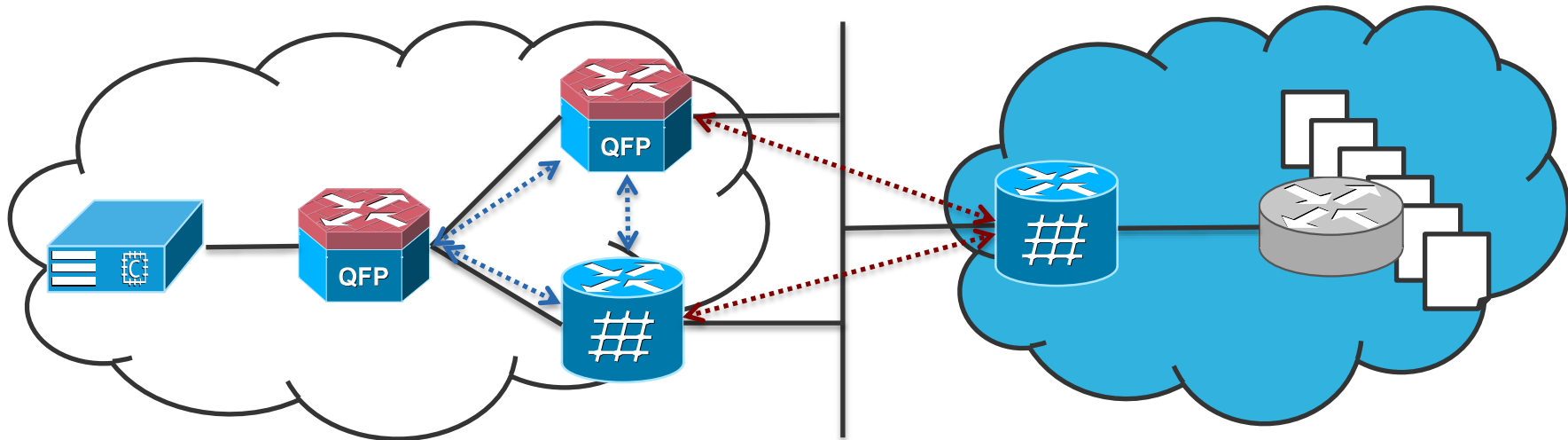
<http://www.openbmp.org/>

pmacct

<http://www.pmacct.net/>

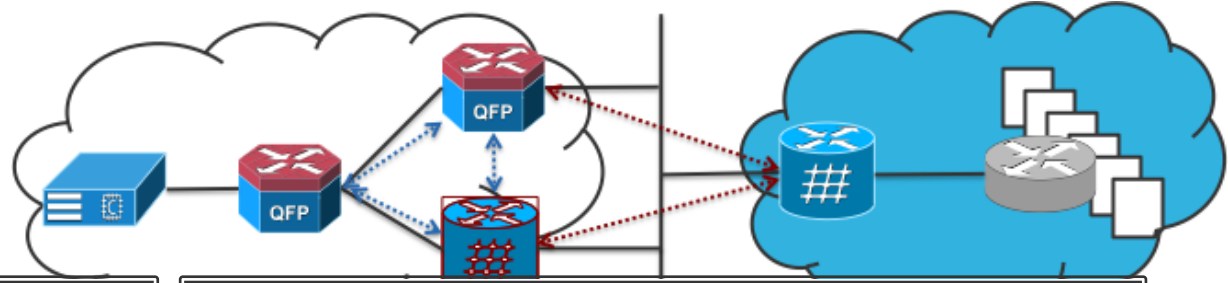


# BMP テストポロジ



```
openbmp:172.16.1.253 5555  
ryu      :172.17.42.1 11019
```

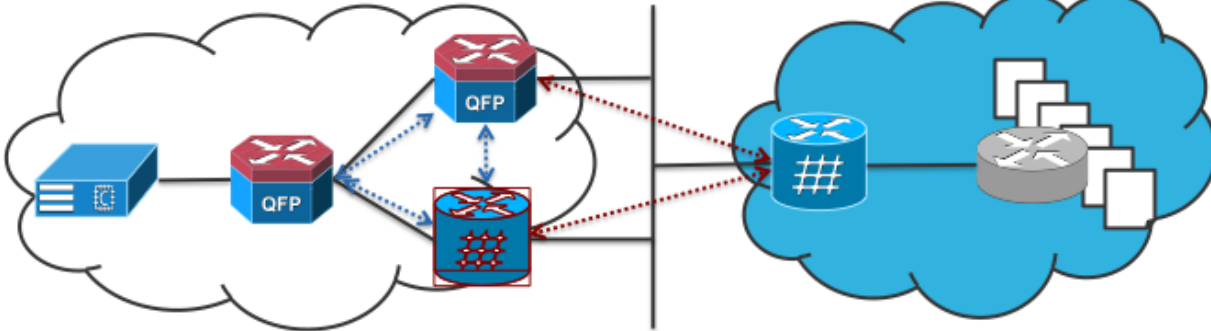
# BMP ルータコンフィグ



```
RP/0/0/CPU0:IOS-XR#show running-config bmp
Sun Jan 10 07:38:57.487 UTC
bmp server 1
  host 172.17.42.1 port 11019
  description RyuBMP
  update-source Loopback0
  initial-delay 30
  stats-reporting-period 50
!
bmp server 2
  host 172.16.1.253 port 5555
  description OpenBMP
  update-source Loopback0
  initial-delay 30
  stats-reporting-period 50
!
```

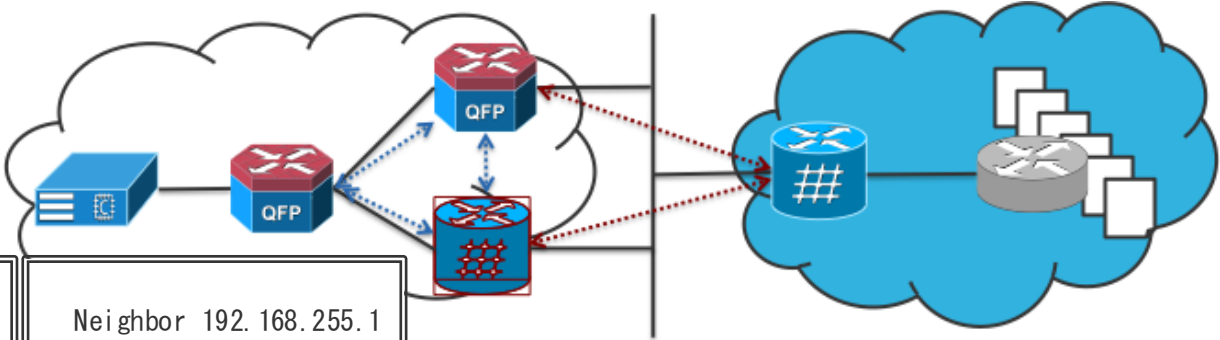
```
RP/0/0/CPU0:IOS-XR#show running-config router bgp
Sun Jan 10 07:38:51.927 UTC
router bgp 64510
  address-family ipv4 unicast
  !
  neighbor 203.0.113.1
  remote-as 64511
  bmp-activate server 1
  bmp-activate server 2
  address-family ipv4 unicast
  route-policy all in
  route-policy all out
  !
  !
--snip--
```

# BMP



```
RP/0/0/CPU0:IOS-XR#show bgp bmp summary
Sun Jan 10 07:39:01.567 UTC
ID  Host                Port    State  Time      NBRs
 1  172.17.42.1          11019  ESTAB  06:13:22  3
 2  172.16.1.253        5555   ESTAB  05:11:29  3
```

# BMP



```
RP/0/0/CPU0:IOS-XR#show bgp bmp server 1 detail
Sun Jan 10 07:39:07.206 UTC
BMP server 1
Host 172.17.42.1 Port 11019
Connected for 06:13:28
Precedence: internet
BGP neighbors: 3
VRF: - (0x60000000)
Update Source: 192.168.255.3 (Lo0)
Update Source Vrf ID: 0x60000000

Message Stats:
Total messages sent: 9990
  INITIATION: 5
  TERMINATION: 0
  STATS-REPORT: 4179
PER-PEER messages: 5806

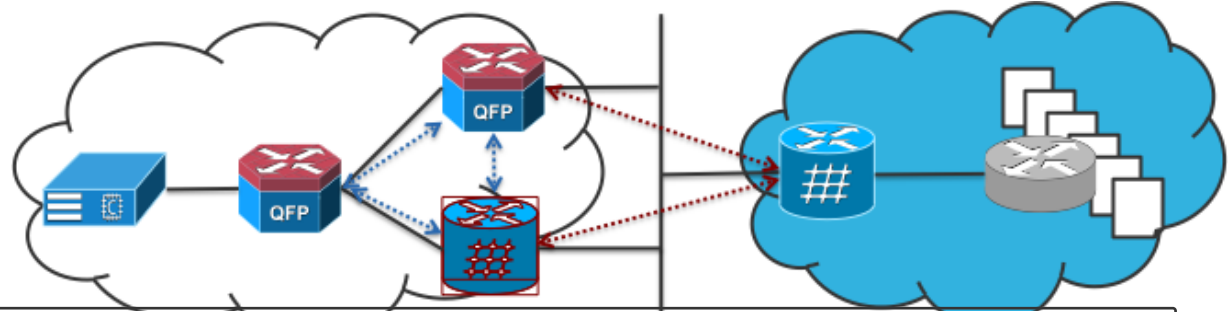
Neighbor 192.168.255.2
Messages pending: 0
Messages sent: 223
  PEER-UP: 5
  PEER-DOWN: 0
  ROUTE-MON: 218
```

```
Neighbor 192.168.255.1
Messages pending: 0
Messages sent: 43
  PEER-UP: 5
  PEER-DOWN: 0
  ROUTE-MON: 38

Neighbor 203.0.113.1
Messages pending: 0
Messages sent: 5540
  PEER-UP: 5
  PEER-DOWN: 0
  ROUTE-MON: 5535
```

# Ryu BMP Server

<http://osrg.github.io/bmp/>



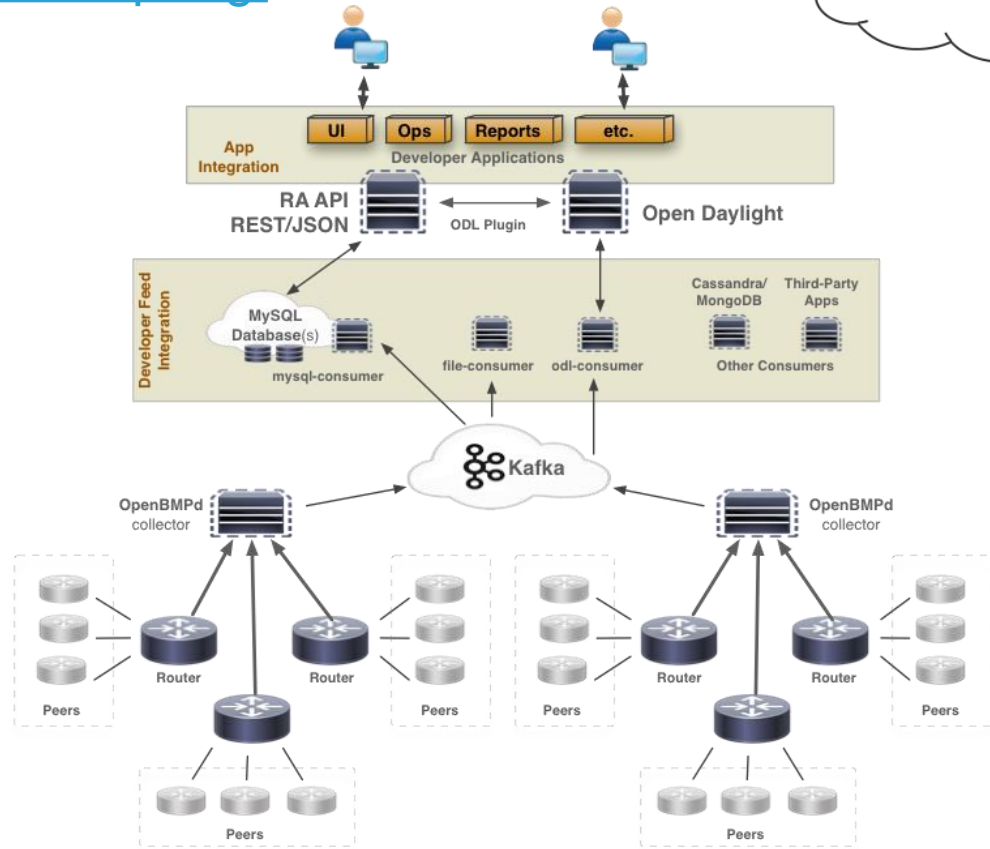
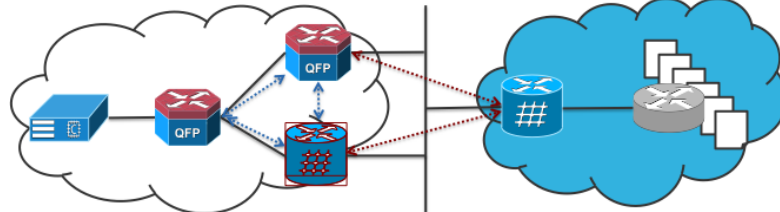
```
instantiating app ./ryu-master/ryu/app/bmpstation of BMPStation
BRICK bmpstation
listening on 0.0.0.0:11019
BMP client connected, ip=192.168.255.3, port=38161
2016 Jan 10 22:47:16 | 192.168.255.3 | BMPInitiation(info=[{'type': 1, 'value': bytearray(b' Cisco IOS XR Software, Version 5.3.2 [Default] \n\n Copyright (c) 2015 by Cisco Systems, Inc.'), 'len': 87}, {'type': 2, 'value': bytearray(b' IOS-XR'), 'len': 6}], len=107, type=4, version=3)

2016 Jan 10 22:47:16 | 192.168.255.3 |
BMPPeerUpNotification(is_post_policy=False, len=184, local_address='0.0.0.0', local_port=179, peer_address='0.0.0.0', peer_as=64510, peer_bgp_id='192.168.255.2', peer_distinguisher=0, peer_type=0, received_open_message=BGPOpen(bgp_identifier='192.168.255.2', hold_time=180, len=57, my_as=64510, opt_param=[BGPOptParamCapabilityMultiprotocol(afi=1, cap_code=1, cap_length=4, length=6, reserved=0, safi=1, type=2), BGPOptParamCapabilityCiscoRouteRefresh(cap_code=128, cap_length=0, length=2, type=2), BGPOptParamCapabilityEnhancedRouteRefresh(cap_code=70, cap_length=0, length=2, type=2), BGPOptParamCapabilityFourOctetAsNumber(as_number=64510, cap_code=65, cap_length=4, length=6, type=2)], opt_param_len=28, type=1, version=4), remote_port=43673, sent_open_message=BGPOpen(bgp_identifier='192.168.255.3', hold_time=180, len=59, my_as=64510, opt_param=[BGPOptParamCapabilityMultiprotocol(afi=1, cap_code=1, cap_length=4, length=6, reserved=0, safi=1, type=2), BGPOptParamCapabilityCiscoRouteRefresh(cap_code=128, cap_length=0, length=2, type=2), BGPOptParamCapabilityRouteRefresh(cap_code=2, cap_length=0, length=2, type=2), BGPOptParamCapabilityFourOctetAsNumber(as_number=64510, cap_code=65, cap_length=4, length=6, type=2), BGPOptParamCapabilityGracefulRestart(cap_code=64, cap_length=2, flags=0, length=4, time=120, tuples=[], type=2)], opt_param_len=30, type=1, version=4), timestamp=1452499071.75092, type=3, version=3)

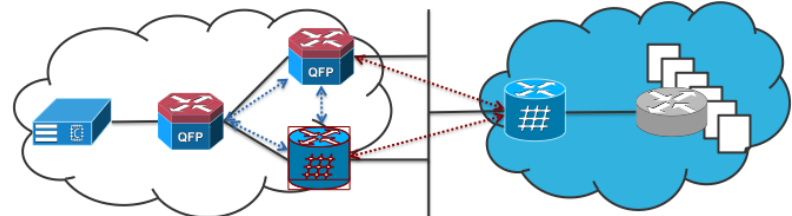
2016 Jan 10 22:47:16 | 192.168.255.3 |
BMPPeerUpNotification(is_post_policy=False, len=184, local_address='0.0.0.0', local_port=32785, peer_address='0.0.0.0', peer_as=64510, peer_bgp_id='192.168.255.1', peer_distinguisher=0, peer_type=0, received_open_message=BGPOpen(bgp_identifier='192.168.255.1', hold_time=180, len=57, my_as=64510, opt_param=[BGPOptParamCapabilityMultiprotocol(afi=1, cap_code=1, cap_length=4, length=6, reserved=0, safi=1, type=2), BGPOptParamCapabilityCiscoRouteRefresh(cap_code=128, cap_length=0, length=2, type=2), BGPOptParamCapabilityEnhancedRouteRefresh(cap_code=70, cap_length=0, length=2, type=2), BGPOptParamCapabilityFourOctetAsNumber(as_number=64510, cap_code=65, cap_length=4, length=6, type=2)], opt_param_len=28, type=1, version=4), remote_port=179, sent_open_message=BGPOpen(bgp_identifier='192.168.255.3', hold_time=180, len=59, my_as=64510, opt_param=[BGPOptParamCapabilityMultiprotocol(afi=1, cap_code=1, cap_length=4, length=6, reserved=0, safi=1, type=2), BGPOptParamCapabilityCiscoRouteRefresh(cap_code=128, cap_length=0, length=2, type=2), BGPOptParamCapabilityRouteRefresh(cap_code=2, cap_length=0, length=2, type=2), BGPOptParamCapabilityGracefulRestart(cap_code=64, cap_length=2, flags=8, length=4, time=120, tuples=[], type=2)], opt_param_len=30, type=1, version=4), timestamp=1452499071.75092, type=3, version=3)
```

# Open BMP

<http://www.openbmp.org/>



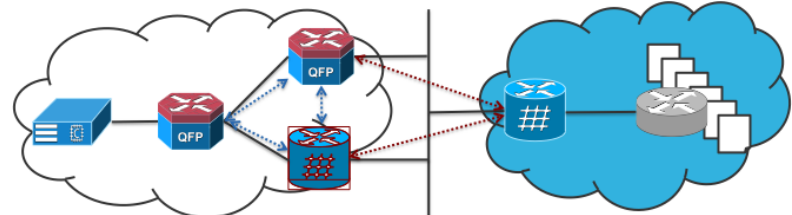
# Open BMP MySQL UI



```
MariaDB [openBMP]> select Prefix, Origin_AS, AS_Path, PeerAddress, PeerASN from v_routes;
```

Prefix	Origin_AS	AS_Path	PeerAddress	PeerASN
192.0.2.81	64502	64511 64499 64508 64500 64505 64501 64504 64502	192.168.255.3	64510
192.0.2.49	64500	64511 64498 64507 64496 64499 64500	192.168.255.3	64510
192.0.2.89	64502	64511 64499 64508 64500 64505 64501 64504 64502	192.168.255.3	64510
192.0.2.85	64502	64511 64499 64508 64500 64505 64501 64504 64502	192.168.255.3	64510
192.0.2.77	64508	64511 64497 64506 64503 64500 64501 64498 64508	192.168.255.3	64510
192.0.2.18	64498	64511 64505 64499 64507 64504 64503 64498	192.168.255.3	64510
192.0.2.86	64502	64511 64499 64508 64500 64505 64501 64504 64502	192.168.255.3	64510
192.0.2.79	64508	64511 64497 64506 64503 64500 64501 64498 64508	192.168.255.3	64510
192.0.2.24	64501	64511 64508 64496 64497 64498 64499 64501	192.168.255.3	64510
192.0.2.40	64505	64511 64498 64502 64503 64507 64497 64504 64505	192.168.255.3	64510
192.0.2.34	64505	64511 64498 64502 64503 64507 64497 64504 64505	192.168.255.3	64510
192.0.2.35	64505	64511 64498 64502 64503 64507 64497 64504 64505	192.168.255.3	64510
192.0.2.64	64501	64511 64498 64502 64503 64499 64500 64501	192.168.255.3	64510
192.0.2.50	64500	64511 64498 64507 64496 64499 64500	192.168.255.3	64510
192.0.2.52	64506	64511 64496 64505 64502 64506	192.168.255.3	64510
192.0.2.94	64507	64511 64505 64497 64506 64507	192.168.255.3	64510
192.0.2.29	64501	64511 64508 64496 64497 64498 64499 64501	192.168.255.3	64510
192.0.2.45	64500	64511 64498 64507 64496 64499 64500	192.168.255.3	64510
192.0.2.2	64506	64511 64504 64503 64500 64506	192.168.255.3	64510

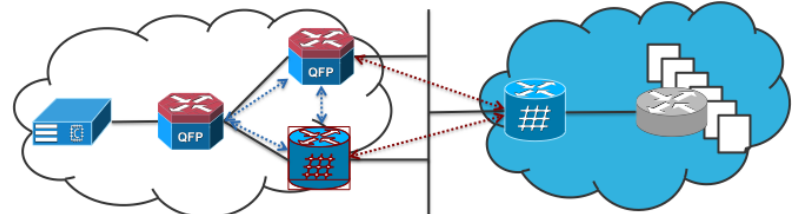
# Open BMP REST UI



```
SHTSUCHI-M-51K1:OpenBMP shtsuchi$ curl -X GET http://10.70.65.196:8001/db_rest/v1/routers
{"routers": {"cols":10,"data":[{"RouterName":"IOS-
XR","Router IP":"192.168.255.3","RouterAS":0,"description":"Cisco IOS XR Software, Version
5.3.2[Default]¥nCopyright (c) 2015 by Cisco Systems,
Inc.", "isConnected":1,"isPassive":0,"LastTermCode":0,"LastTermReason":"","InitData":"","LastModified":"20
16-01-11
10:29:20.617171"}, {"RouterName":"","Router IP":"192.168.255.2","RouterAS":0,"description":"","isConnected"
:1,"isPassive":0,"LastTermCode":0,"LastTermReason":"","InitData":"","LastModified":"2016-01-11
10:33:03.454087"}, {"RouterName":"","Router IP":"192.168.255.1","RouterAS":0,"description":"","isConnected"
:1,"isPassive":0,"LastTermCode":0,"LastTermReason":"","InitData":"","LastModified":"2016-01-11
10:24:19.579993"}],"size":3,"queryTime_ms":1,"fetchTime_ms":0} }
SHTSUCHI-M-51K1:OpenBMP shtsuchi$
```



# Open BMP REST UI



```
SHTSUCHI-M-51K1:OpenBMP shtsuchi$ curl -X GET http://10.70.65.196:8001/db_rest/v1/routers | python -m json.tool
```

```
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 769 100 769 0 0 40995 0 --:--:-- --:--:-- --:--:-- 40473
```

```
{
  "routers": [
    {
      "cols": 10,
      "data": [
        {
          "InitData": "",
          "LastModified": "2016-01-11 10:29:20.617171",
          "LastTermCode": 0,
          "LastTermReason": "",
          "RouterAS": 0,
          "RouterIP": "192.168.255.3",
          "RouterName": "IOS-XR",
          "description": "Cisco IOS XR Software, Version 5.3.2[Default]\nCopyright (c) 2015 by Cisco Systems, Inc.",
          "isConnected": 1,
          "isPassive": 0
        },
        {
          "InitData": "",
          "LastModified": "2016-01-11 10:33:03.454087",
          "LastTermCode": 0,
          "LastTermReason": "",
          "RouterAS": 0,
          "RouterIP": "192.168.255.2",
          "RouterName": "",
          "description": "",
          "isConnected": 1,
          "isPassive": 0
        }
      ]
    }
  ]
}
```

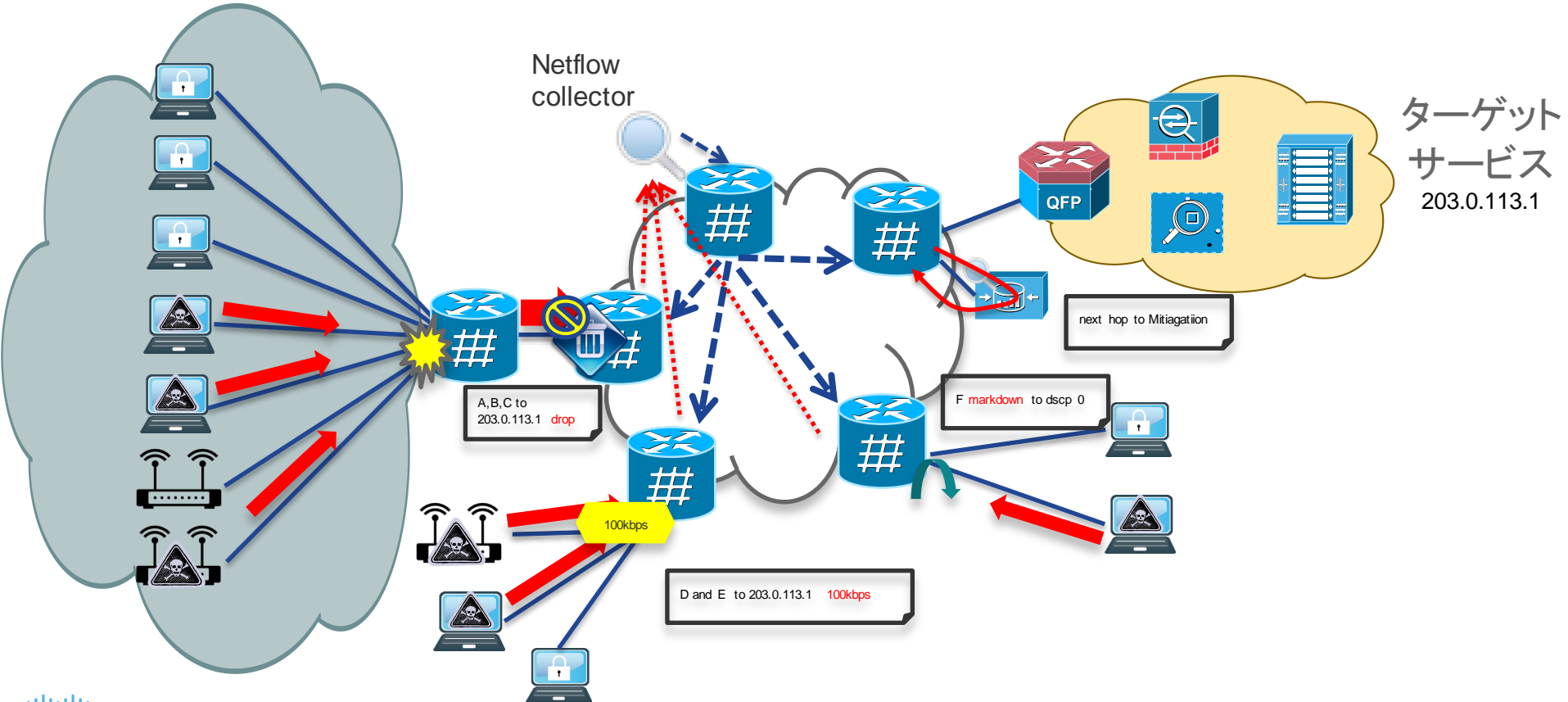
# まとめ

- BMPはもうすぐRFC(現在 IESG Evaluation)
  - ルータの実装も増えてきた(IOS/JUNOSに続き、IOS-XR/GoBGPなど)
  - BMPサーバーの実装も多く、データベースを用いた実用的なものもある
- 
- 動けば変わる

# Agenda

- BGP Selective Download
- BGP 4byte Private AS
- BMPサーバー
- **BGP Flowspec**

# BGP Flowspec(RFC5575)



# 関連トピック

- JANOG35
  - BGP Flowspec(RFC5575)  
<https://www.janog.gr.jp/meeting/janog35/program/bgpfs/>
- JANOG35.5
  - GRNET (Greek Research and Technology Network) FireCircle ギリシャの広域ファイアーウォールのお話  
<http://www.janog.gr.jp/meeting/janog35.5/flowspec>
  - 相互接続実績が少ない事を報告
  - 外部ファイアーウォールサービスを紹介

- Interop TokyoでのBGP Flowspec相互接続検証結果を報告
- Juniper vMX/MX480,Hauwei NE5000E,Cisco ASR9kとの相互接続を実施
- またSAMURAIでの攻撃検知→RR(vMX)にFlow Specでの経路広告→上記3機種にBGP Flowspecでの経路分配

# DDOS対策サービスのリリース

- 「OCN DDoS対策サービス」を大幅リニューアル  
～独自開発トラフィック解析ツールを活用した検知・防御・レポート機能を、  
より安価な料金で提供～

[http://www.ntt.com/release/monthNEWS/detail/20151001\\_3.html](http://www.ntt.com/release/monthNEWS/detail/20151001_3.html)

完全自動のDDOSミチゲーション/リアルタイムレポート

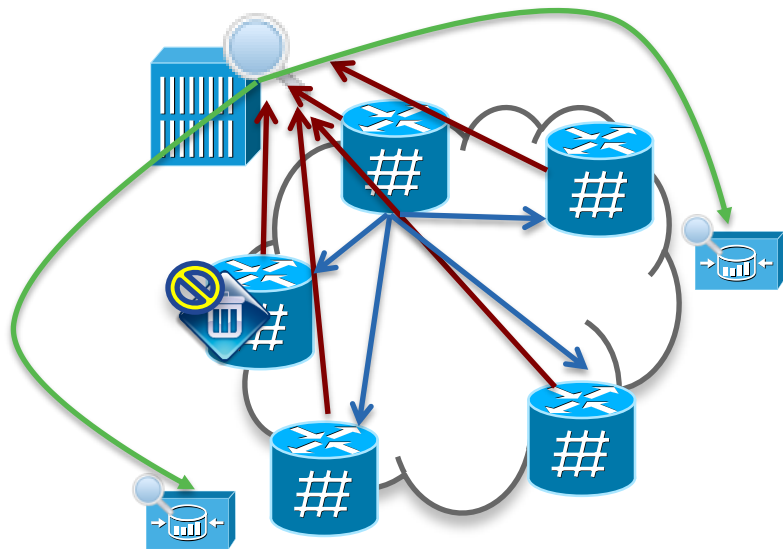
IPアドレスに基づいたパケットフィルタリング

- BIGLOBEがDDoS攻撃対策としてデータセンタ向け「Flow分析プラットフォームサービス」の提供を開始～世界初、Genie社のシステムをクラウド型で提供～

<http://www.biglobe.co.jp/pressroom/release/2016/01/160120-a>

トラフィック状況をリアルタイムに分析

# IETF DOTS(DDoS Open Threat Signaling) WG



- フォワーダーからコレクタ:xFlow
- 各フォワーダへのルール分配:BGP
- サーバーとミチゲーション間のやり取り???

- DOTS Agent(Sever/Client)間のプロトコルを定義
- 現在のベンダー独自技術からの相互接続や拡張を目指す



# Flowspecの拡張

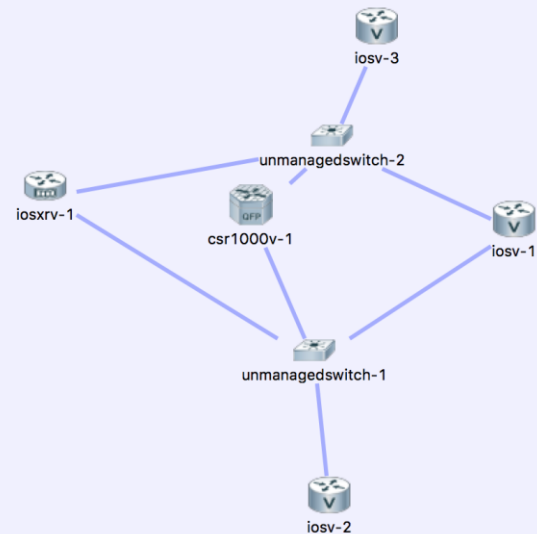
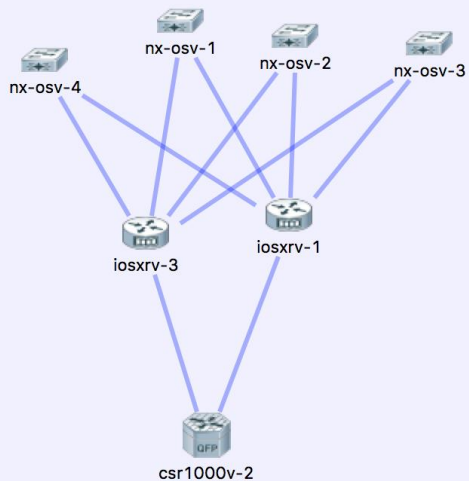
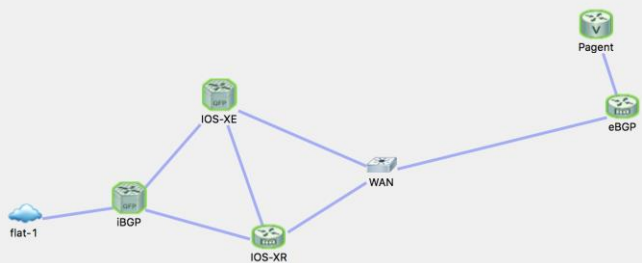
- draft-wu-idr-flowspec-yang-cfg
  - Flowspec configのyangモデル
- draft-ietf-idr-flowspec-l2vpn
  - L2VPN(EVPN/BGP-VPLS)への拡張
- OSPF Extensions for Flow Specification
  - draft-ietf-ospf-flowspec-extensions
- IS-IS Extensions for Flow Specification
  - draft-you-isis-flowspec-extensions
- IETF94横浜 IDR WGではほぼ祭り

# まとめ

- BGP Flowspecは相互接続も増えてきた
- 外部向けへのサービスなども始まりつつあり(BGP Flowspecに限らず)
- DOTSサーバーとクライアントのやり取り部分の標準化やサービスとしてもモデルの標準化はこれから
- Flowspecの拡張も考えられてる

# 全てこれでやりました

<http://virl.cisco.com/>



- JANOG35 仮想ルータの使い方

<http://www.janog.gr.jp/meeting/janog33/program/vr.html>

# まとめ

- 問題提議から実装などは時間がかかる
- アップデートは重要
- また運用者の議論内容などを、標準化団体/開発者にフィードバックをしていくのは、本当に重要



*TOMORROW starts here.*