



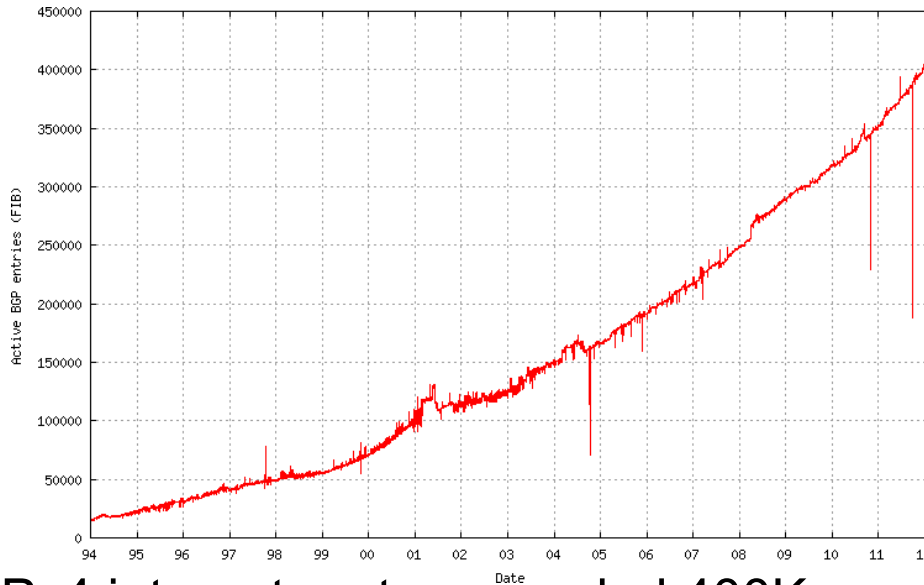
Consideration of route exhaustion –new protocol deployment – Simple Virtual Aggregation

Shishio Tsuchiya

shtsuchi@cisco.com

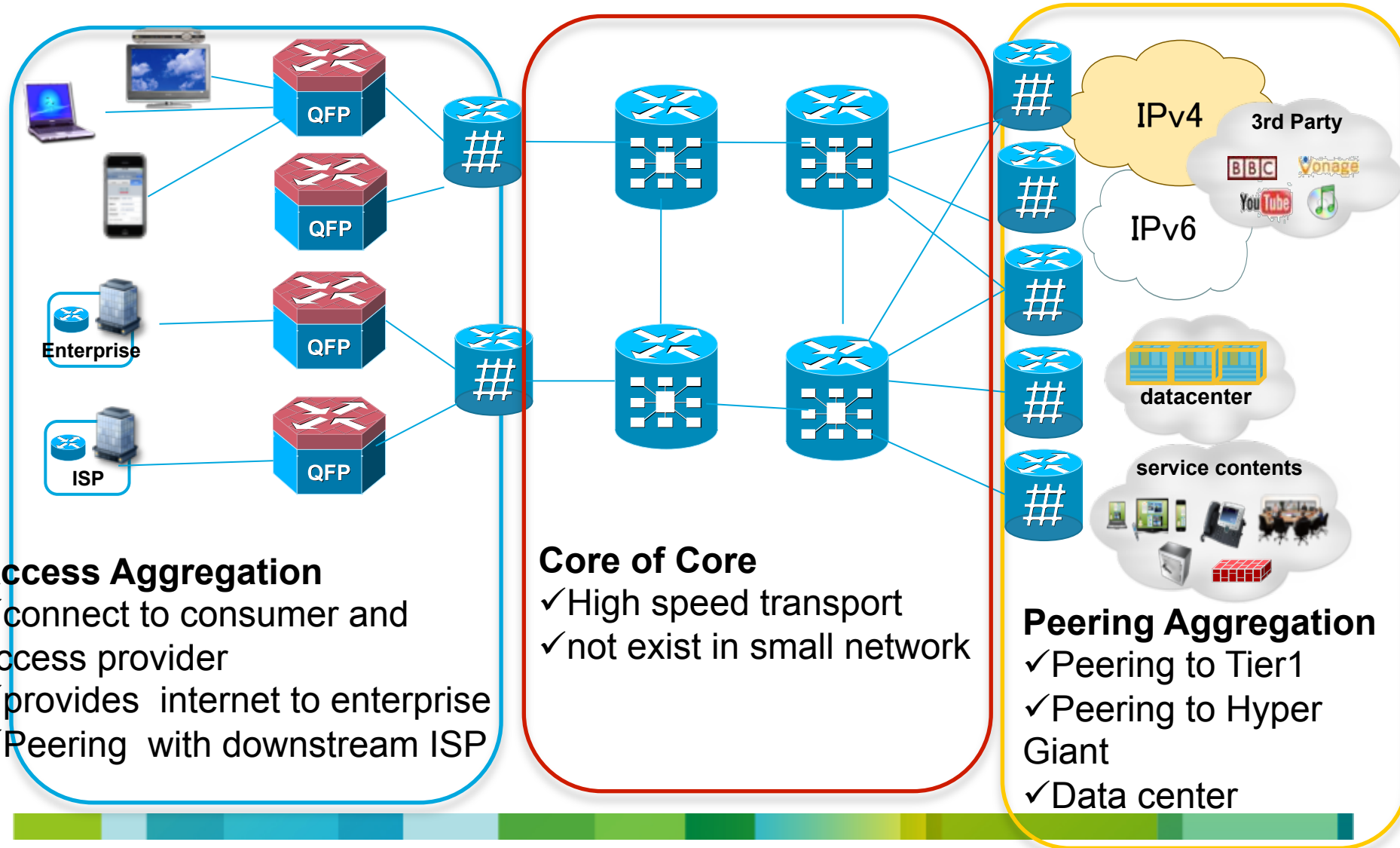
route exhaustion

<http://bgp.potaroo.net/as6447/>



- Today's IPv4 internet route exceeded 400K.
- Modern router's capacity has enough memory and hardware resources.
- But route is growing and there are old/poor capacity routers on ISP network, sometimes.
- Simple VA provides scalability, convergence improvement and simple bgp operation.

network topology of internet service provider



Requirement of each layer

	Access	Core of Core	Peering
interface	variable	100GE/40GE/10GE	10GE/1GE
Number of BGP route (advertise)	full route	None	depend on customer /service route
Number of BGP route(receive)	full route	full route	full route
FIB	Huge	Huge(due to transit)	Huge
Dual Stack	Need	Need(due to transit)	Need
Cost of equipment	\$	\$\$\$	\$\$

Requirement of Access today's focus point

	Access	Core of Core	Peering
interface	variable	100GE/40GE/10GE	10GE/1GE
Number of BGP route (advertise)	full route	None	depend on customer /service route
Number of BGP route(receive)	full route	full route	full route
FIB	Huge	Huge(due to transit)	Huge
Dual Stack	Need	Need(due to transit)	Need
Cost	\$	\$\$\$	\$\$

- Full route capability is required on all of routers, to provide internet full route to customer and downstream ISP.

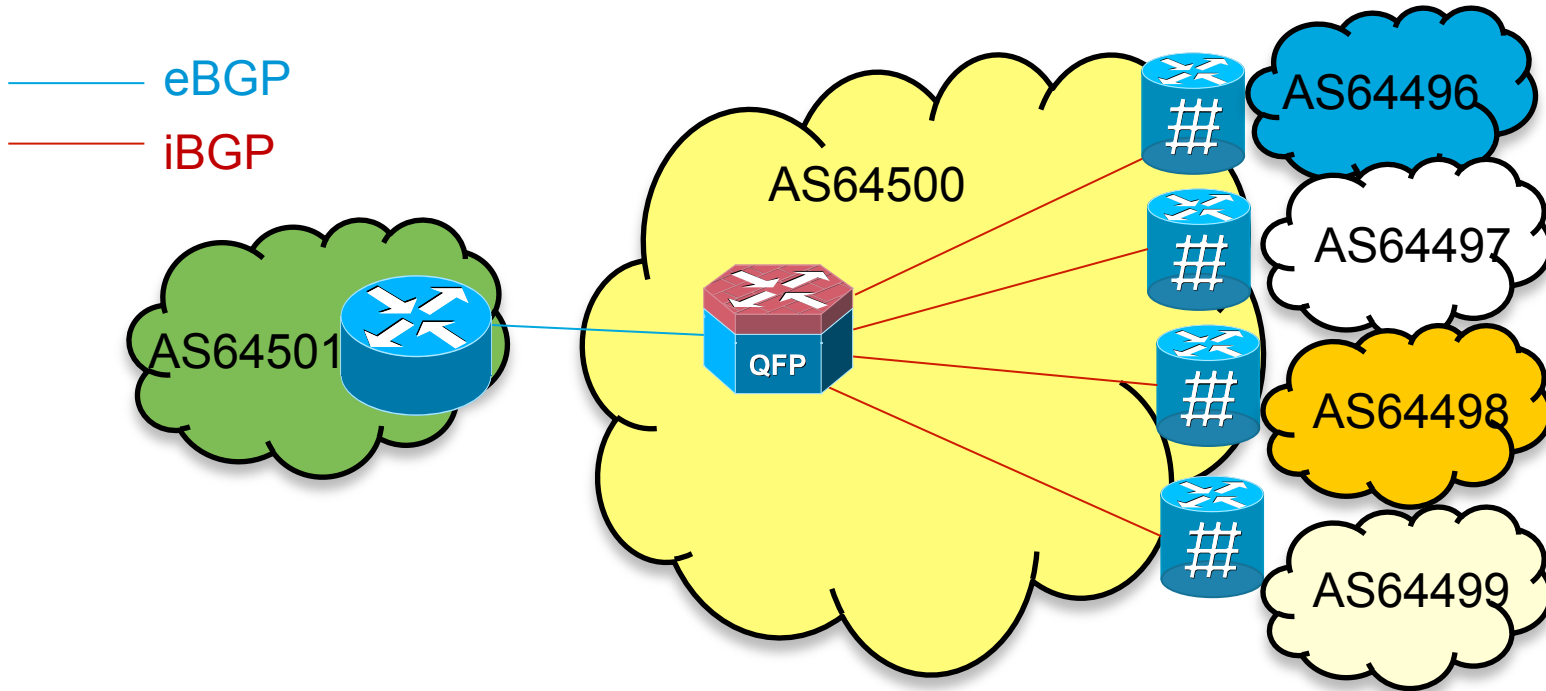
How to create FIB on BGP environment



FIB / RIB Table	Data
Active BGP entries (FIB)	396,184
All BGP entries (RIB)	12,561,626
RIB/FIB ratio	31.7065

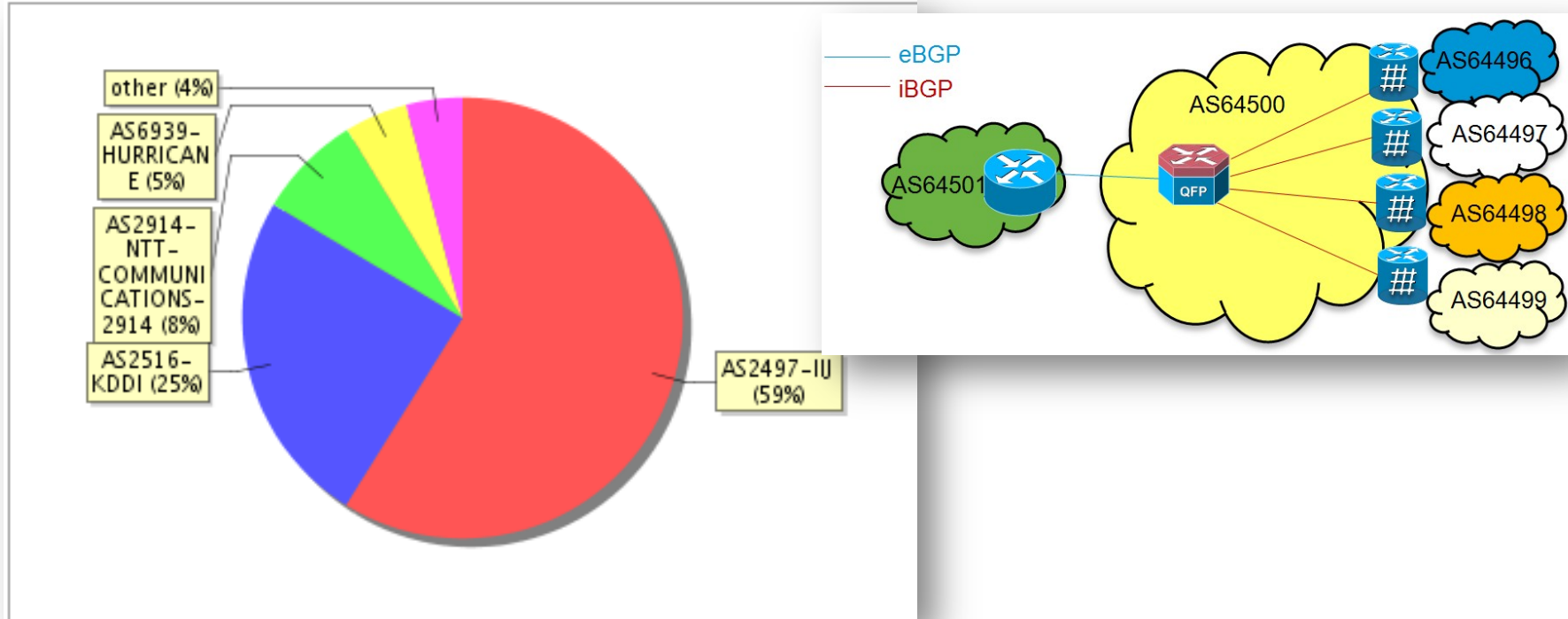
- create routing table(RIB) from BGP table
- create FIB from RIB(copies information to TCAM/NP)
- forward packet based on FIB
- If BGP would be exhaustion then all of resources will be consumed.

topology example



- peering with 4 ISPs and provides full router to customer
- exchange route by iBGP in intraAS.

Do you really need full route?

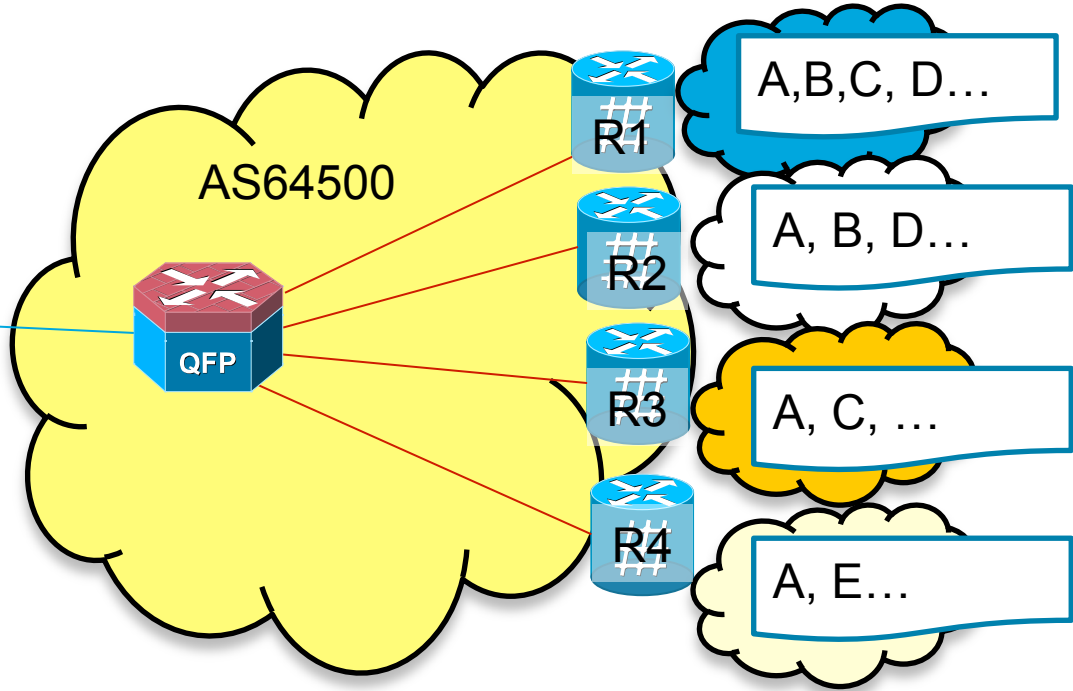


- Yes! need full route for provide full route to down stream ISP
- There is deviation of traffic distribution.
- can confirm distribution of traffic on [RIPE AS dashboard](#)
- IIJ occupies about 60% of total of traffic distribution.

Simple Virtual Aggregation(S-VA)

[draft-ietf-grow-simple-va](#)

BGP Table	
Destination	Nexthop
0.0.0.0	*R1
A	*R1 R2 R3 R4
B	*R1 R2
C	R1 *R3
D	*R1 R2
E	*R4



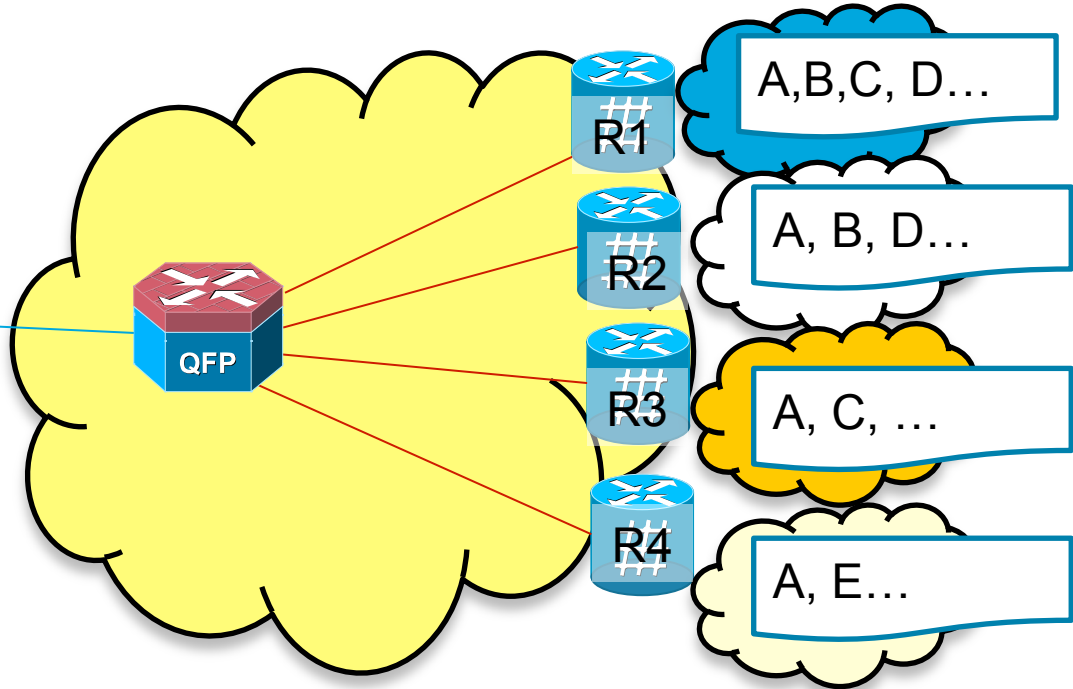
- Normal BGP table
- If there is multiple paths, it selects best path.

Simple Virtual Aggregation(S-VA)

[draft-ietf-grow-simple-va](#)

cont'd

routing table	
Destination	Nextthop
0.0.0.0	*R1
A	*R1
B	*R1
C	*R3
D	*R1
E	*R4



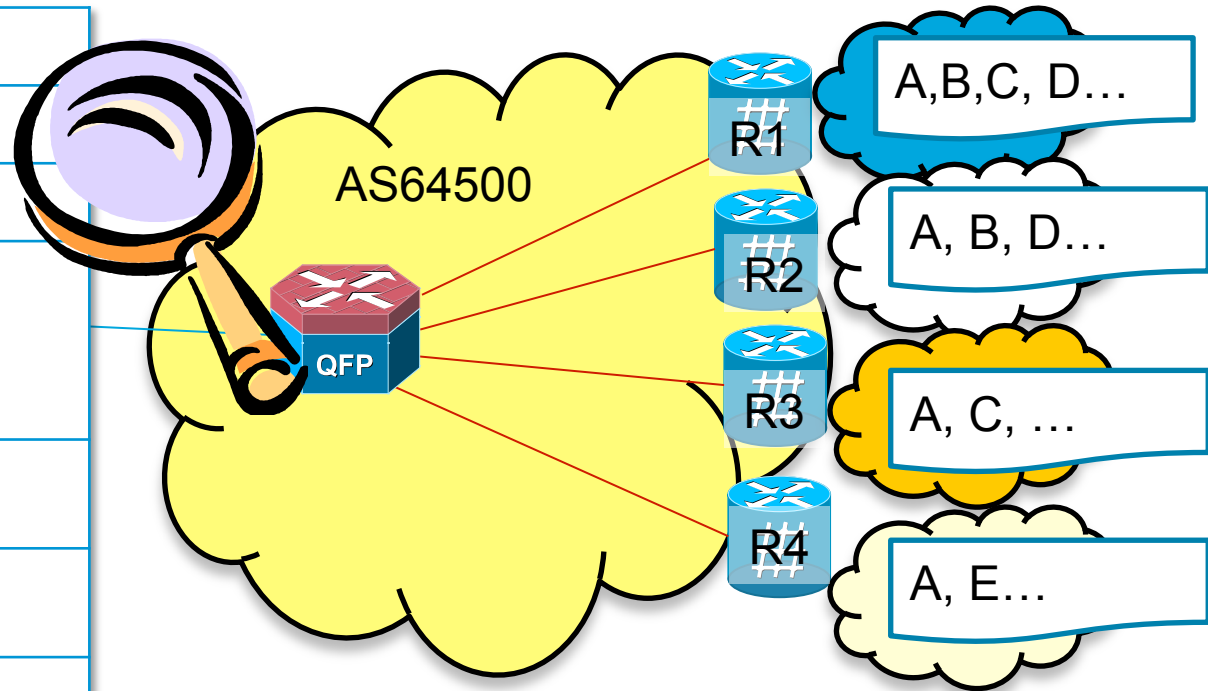
- create RIB
- each of destination route has each of nextthop

Simple Virtual Aggregation(S-VA)

[draft-ietf-grow-simple-va](#)

cont' d

BGP table	
Destination	Nexthop
0.0.0.0	*R1
A	*R1 R2 R3 R4
B	*R1 R2
C	R1 *R3
D	*R1 R2
E	*R4



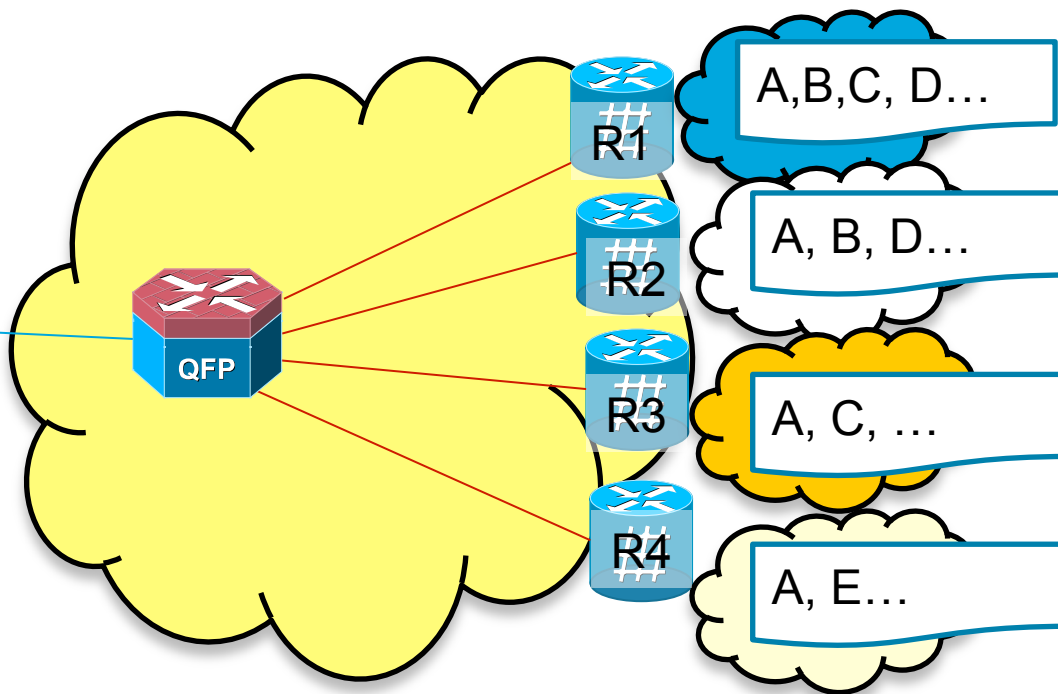
- S-VA:calculates VA Prefix 0/0 at first
- Suppress route which has same next hop as VA Prefix

Simple Virtual Aggregation(S-VA)

[draft-ietf-grow-simple-va](#)

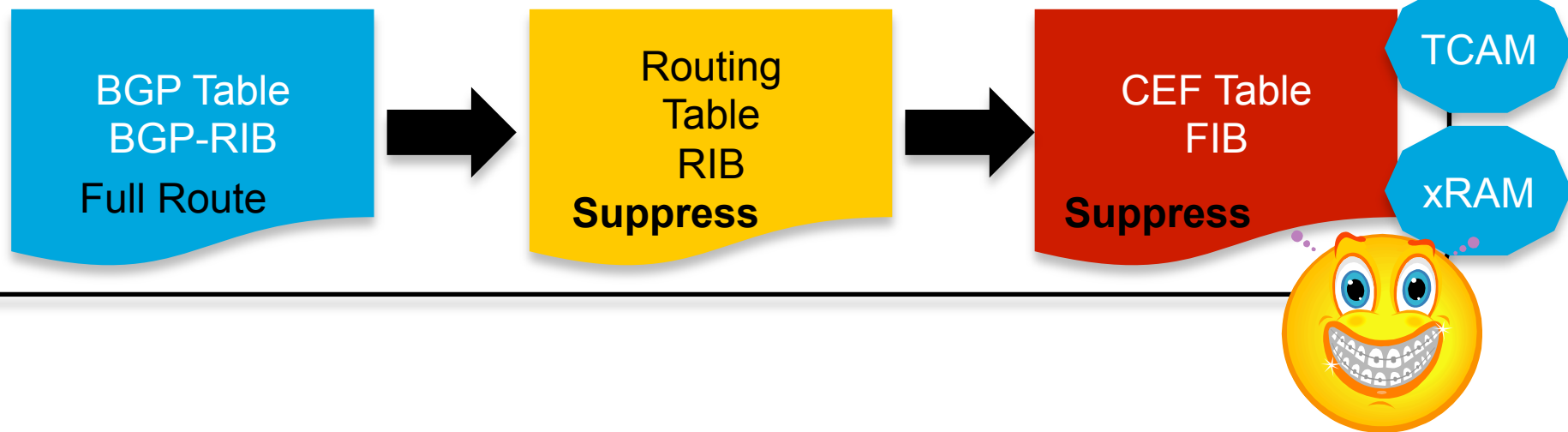
cont'd

routing table	
Destination	Nexthop
0.0.0.0	*R1
C	*R3
E	*R4



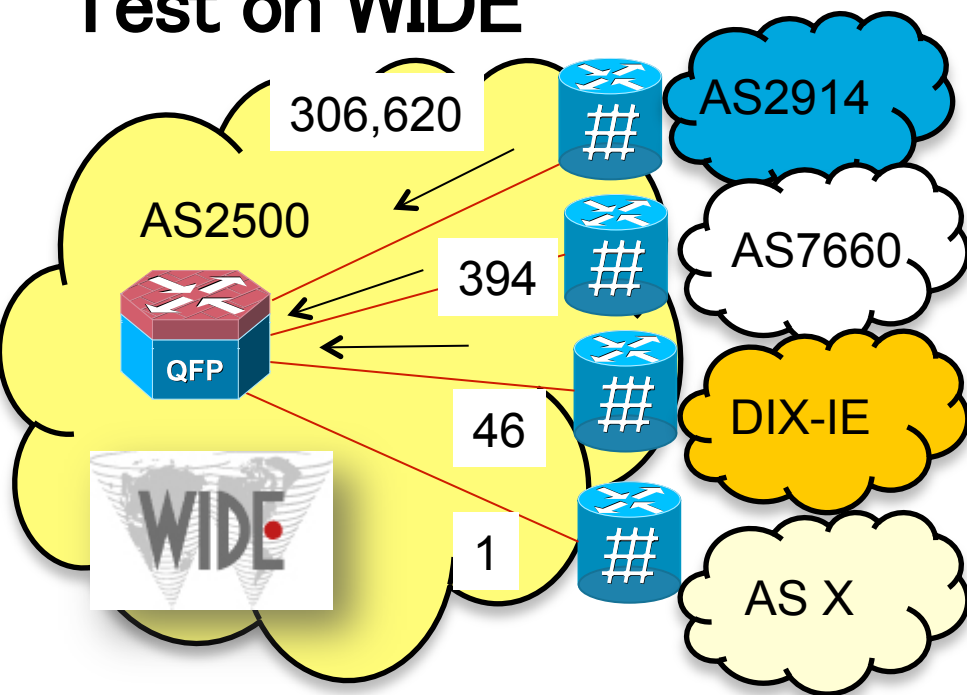
- BGP table is same size as normal.
- But reduce routing table and FIB

How to create FIB in S-VA



- reduce used memory of RIB/FIB
- as result reduce entry of TCAM/xRAM
- routing lookup also would be easy

Test on WIDE



	BGP RIB	RIB	BGP Memory	RIB Memory
Normal	307,061	307,041	61MB	169 MB
S-VA	307,061	429	61MB	12MB

- Test result on WIDE(AS2500)
- Routing table reduce to 0.14% (300K->400)
- 92% compress usage of memory (169MB->12MB)

Summary of Simple Virtual Aggregation(S-VA)

- S-VA is technique which can save utilization of FIB in the current network.
- Not require enhancement of BGP protocol,it can do by only enhancement of Edge Router(FSR)

IOS Implementation

```
!  
router bgp xxx  
  address-family ipv4 unicast  
  bgp va <prefix> mask <mask>  
!
```

- simple configuration!
- All VA compressed route are marked with 'C' in "sh ip bgp".
- IOS-XE 3.5.1 15.2(1)S supported : ASR1K/7600
- ISR G2 will support 15.2(3)T(next release)

Key word of each layer

	Access	Core of Core	Peering
interface	variable	100GE/40GE/10GE	10GE/1GE
Number of BGP route (advertise)	full route	None	depend on customer /service route
Number of BGP route(receive)	full route	Not need (MPLS)	full route(IESG finding solution)
FIB	Reduce by S-VA	Small(LFIB based)	Huge(IESG finding solution)
Dual Stack	Need	Not Need(use 6PE)	Need
Cost	\$	Lean Core	\$\$

- not enough time to discuss of all of layer
- But key words are :S-VA,MPLS,ILNP,LISP and so on

Thank you.

