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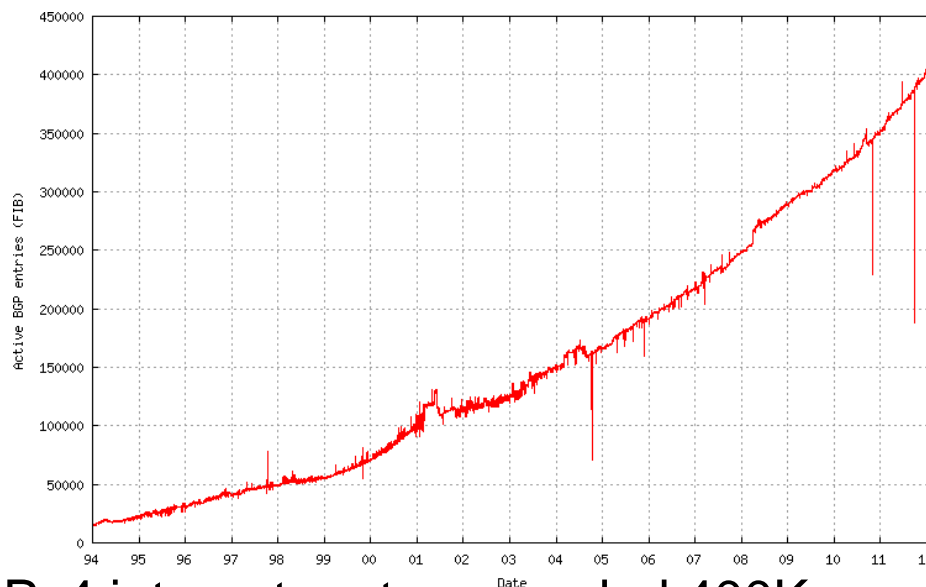
Shishio Tsuchiya

shtsuchi@cisco.com

route exhaustion

- 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

IPv4
3rd Party
BBC Vonage
YouTube
IPv6

datacenter

service contents
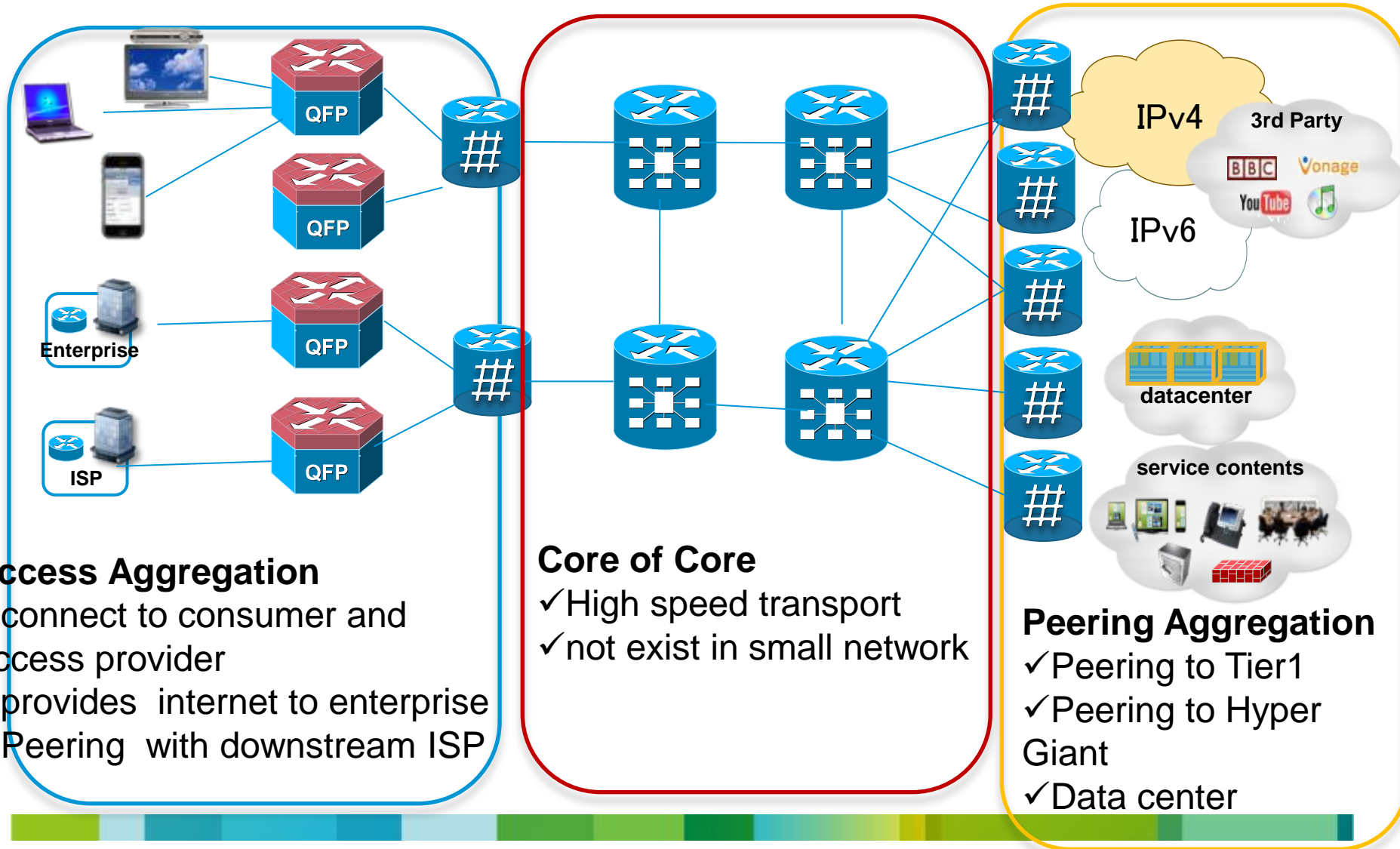
Access Aggregation
✓connect to consumer and access provider
✓provides internet to enterprise
✓Peering with downstream ISP

Enterprise

ISP

QFP
QFP
QFP
QFP

Core of Core
✓High speed transport
✓not exist in small network

Peering Aggregation
✓Peering to Tier1
✓Peering to Hyper Giant
✓Data center

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
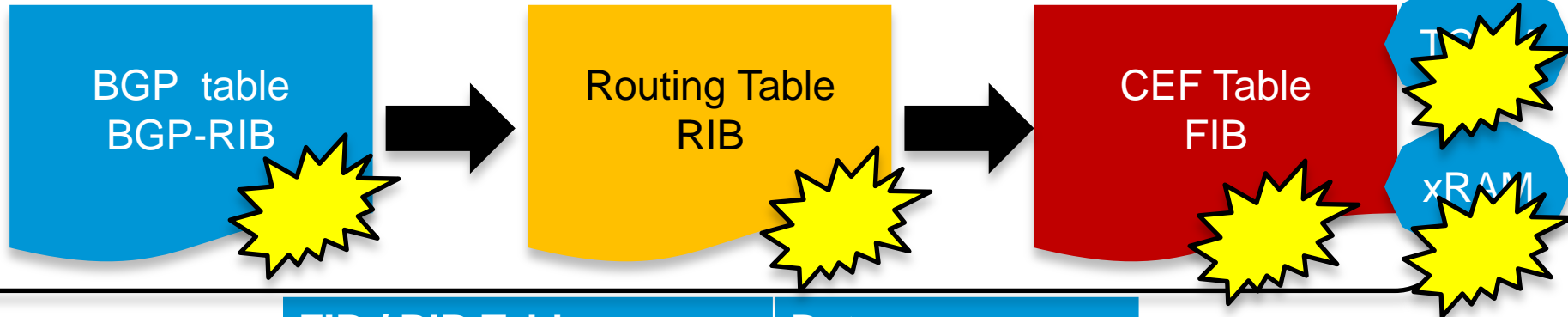
| BGP table BGP-RIB | → | Routing Table RIB | → | CEF Table FIB | TCAM xRAM |

| 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
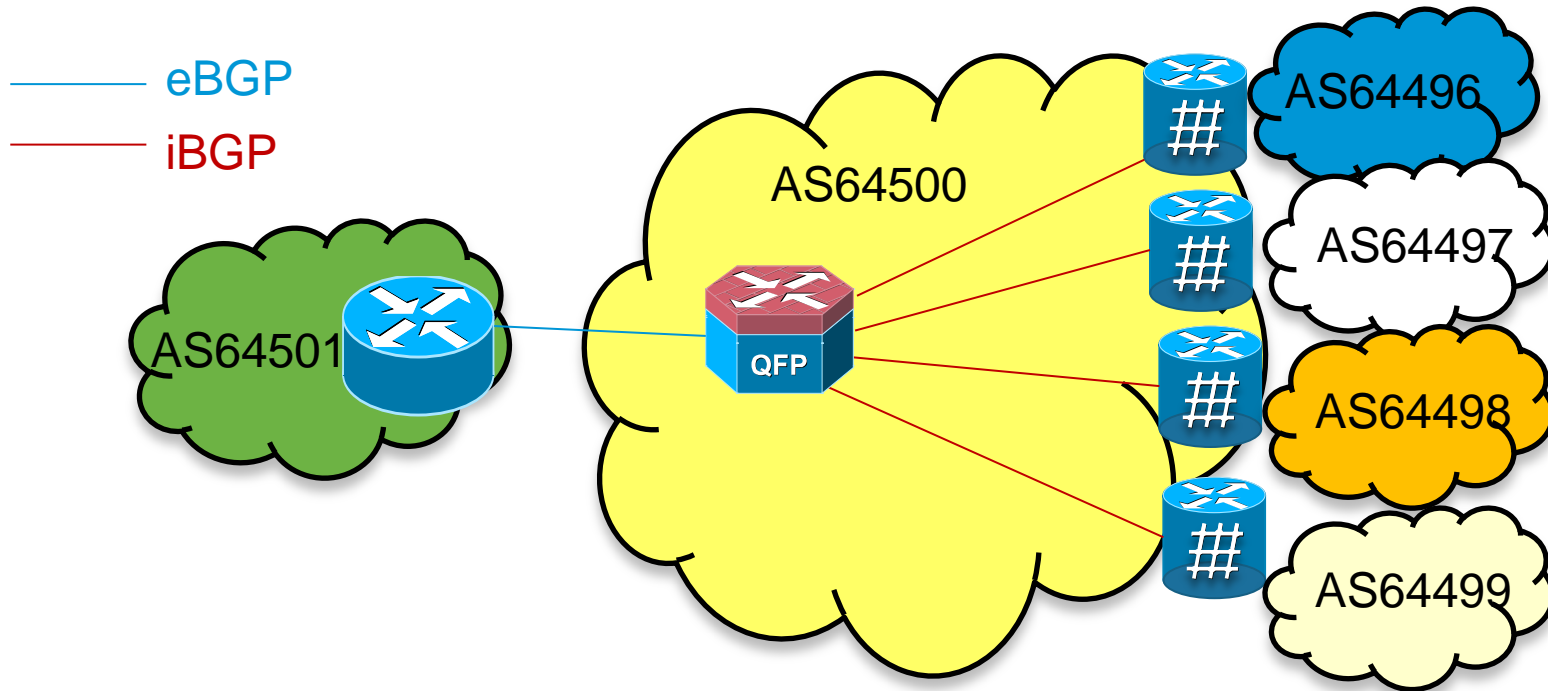
- peering with 4 ISPs and provides full router to customer

- exchange route by iBGP in intraAS.

Do you really need full route?
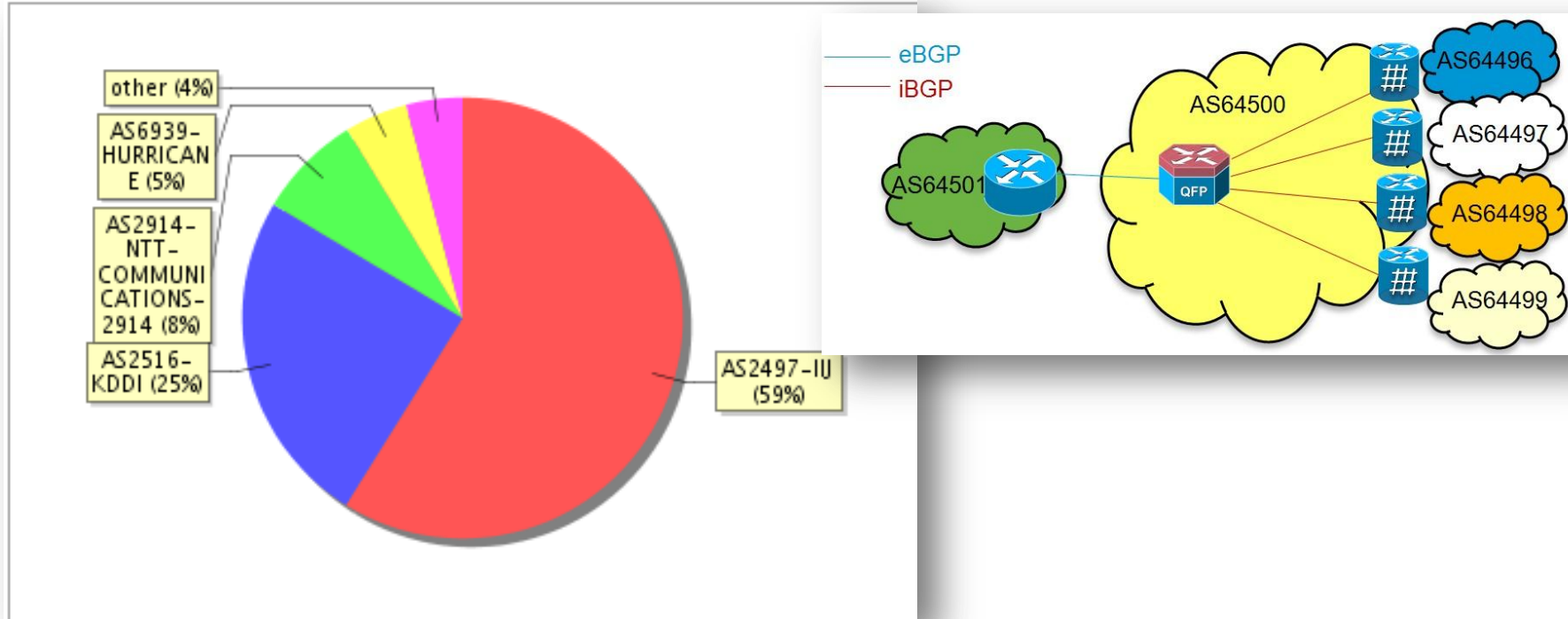
- Yes!need full route for provide full route to down stream ISP

- 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

AS64500

R1 — A,B,C, D…

R2 — A, B, D…

R3 — A, C, …

R4 — A, E…
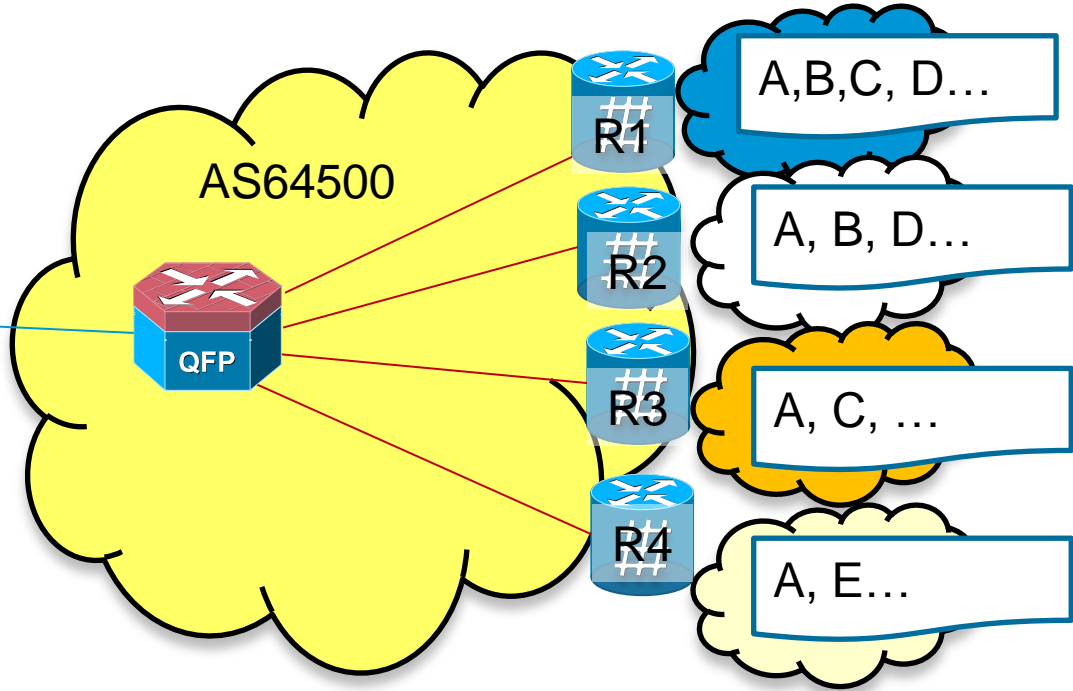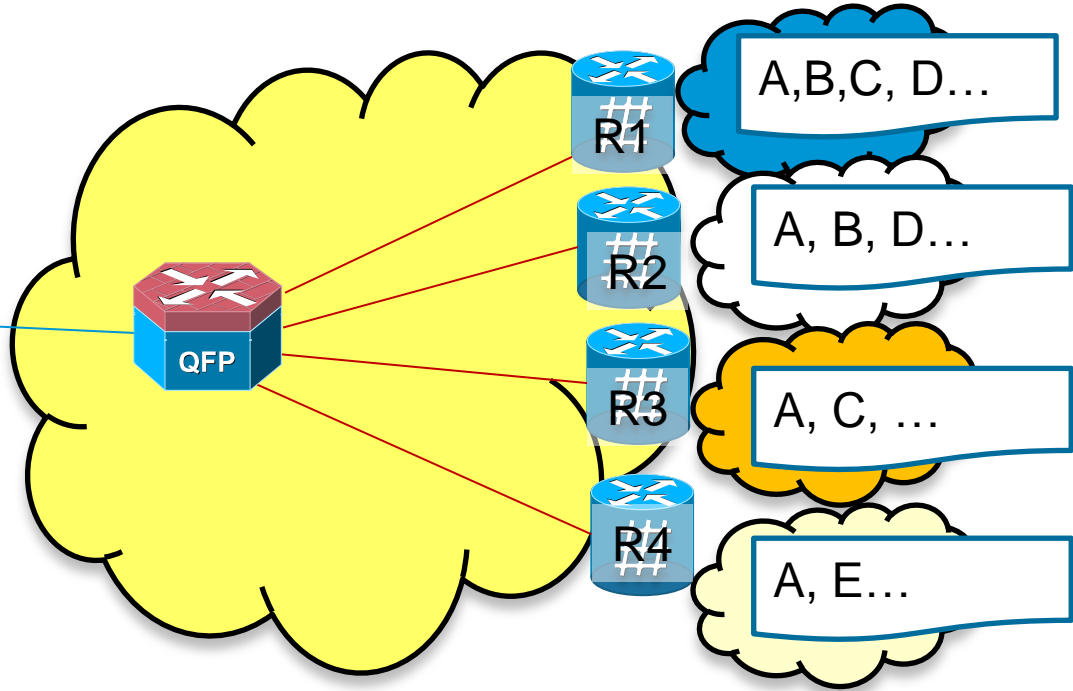
- Normal BGP table

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

Simple Virtual Aggregation(S-VA)
draft-ietf-grow-simple-va
cont'd

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

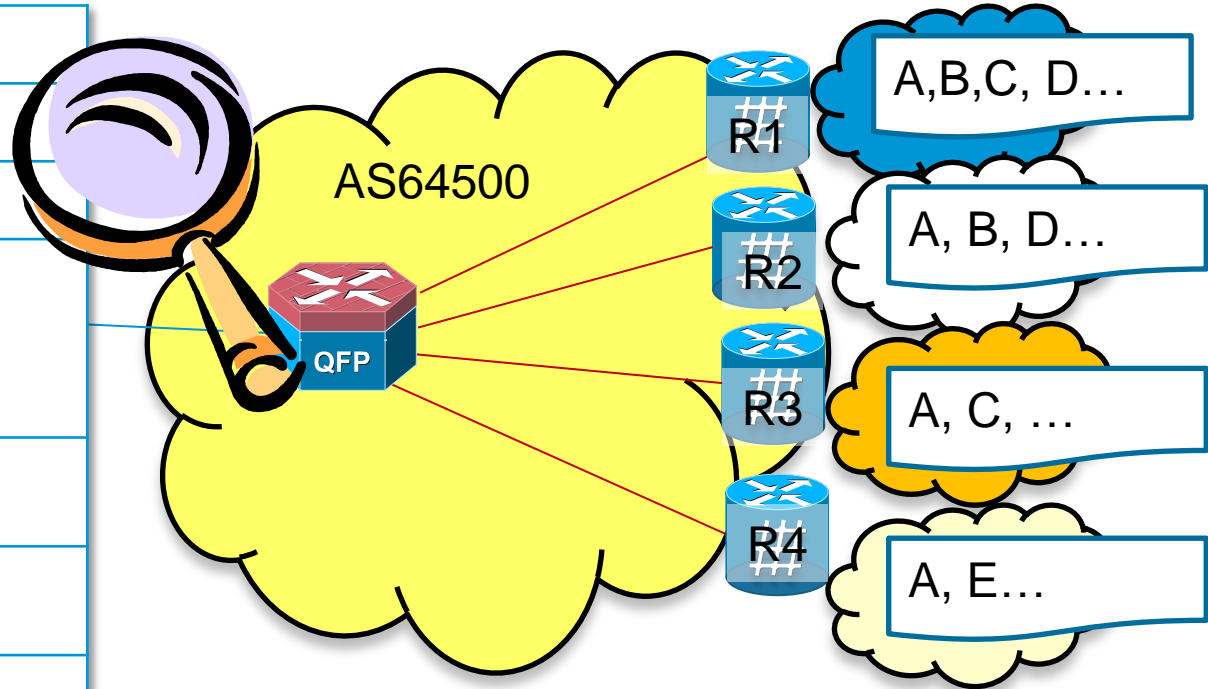A,B,C, D…

A, B, D…

A, C, …

A, E…

- create RIB

- each of destination route has each of nexthop

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

AS64500

QFP

R1 — A,B,C, D…

R2 — A, B, D…
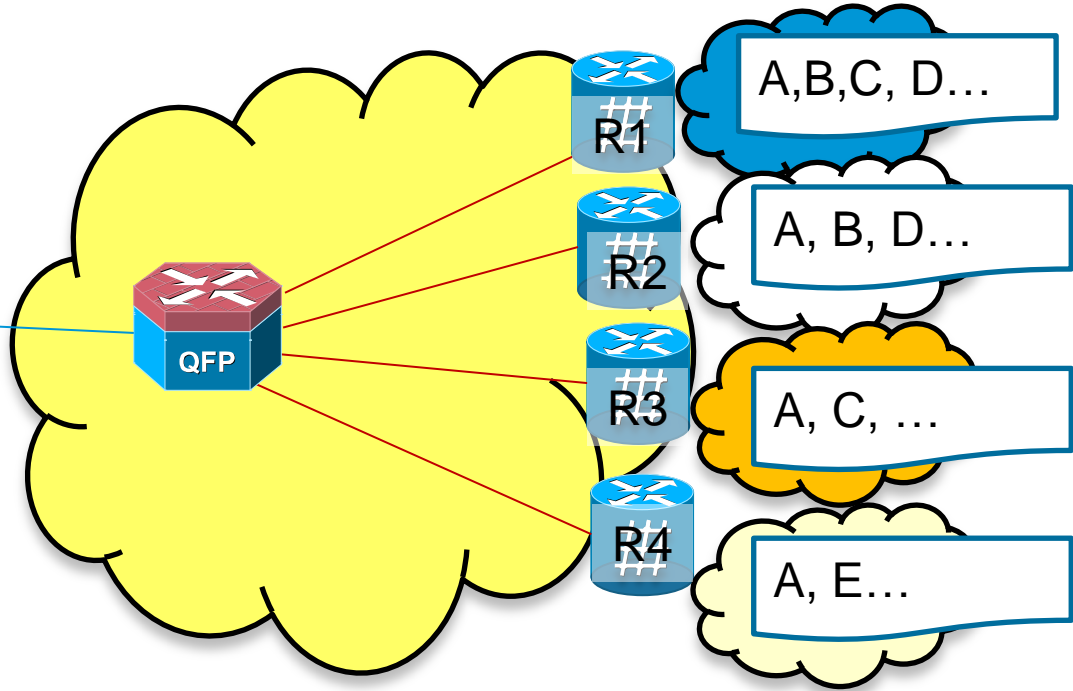
R3 — A, C, …

R4 — A, E…

- S-VA:calculates VA Prefix 0/0 at first

- Suppress route which has same next hop as VA Prefix

Simple Virtual Aggregation(S-VA)
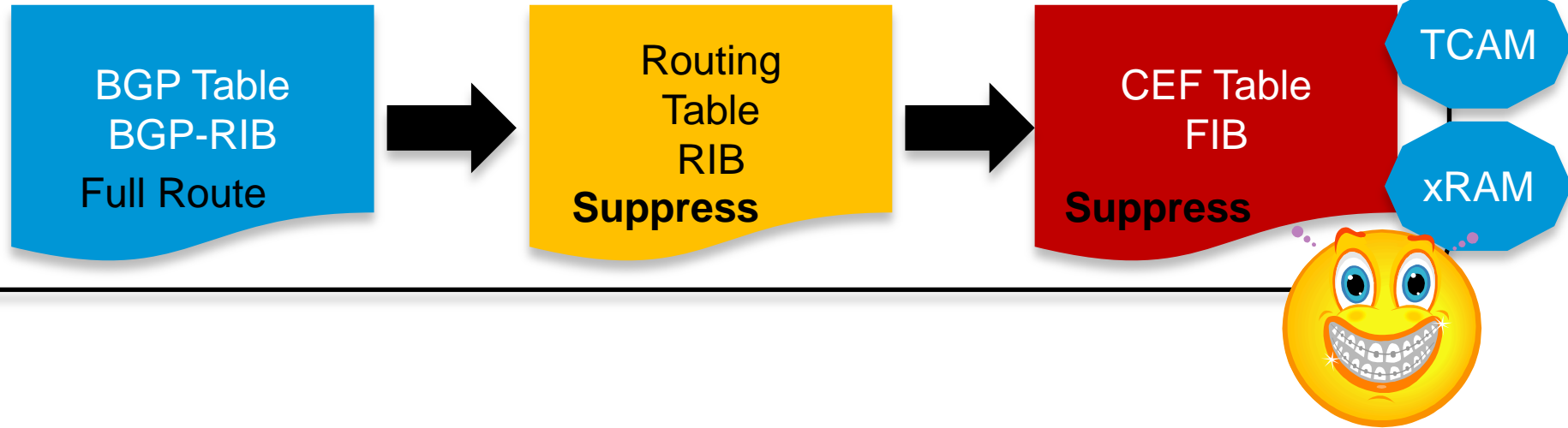## draft-ietf-grow-simple-va
cont'd

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

A,B,C, D…

A, B, D…

A, C, …

A, E…

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

How to create FIB in S-VA

| BGP Table BGP-RIB Full Route | → | Routing Table RIB **Suppress** | → | CEF Table FIB **Suppress** | TCAM / xRAM |

- reduce used memory of RIB/FIB

- as result reduce entry of TCAM/xRAM

- routing lookup also would be easy

Test on WIDE

306,620

AS2500

QFP

AS2914

AS7660

DIX-IE

AS X

394

46

1
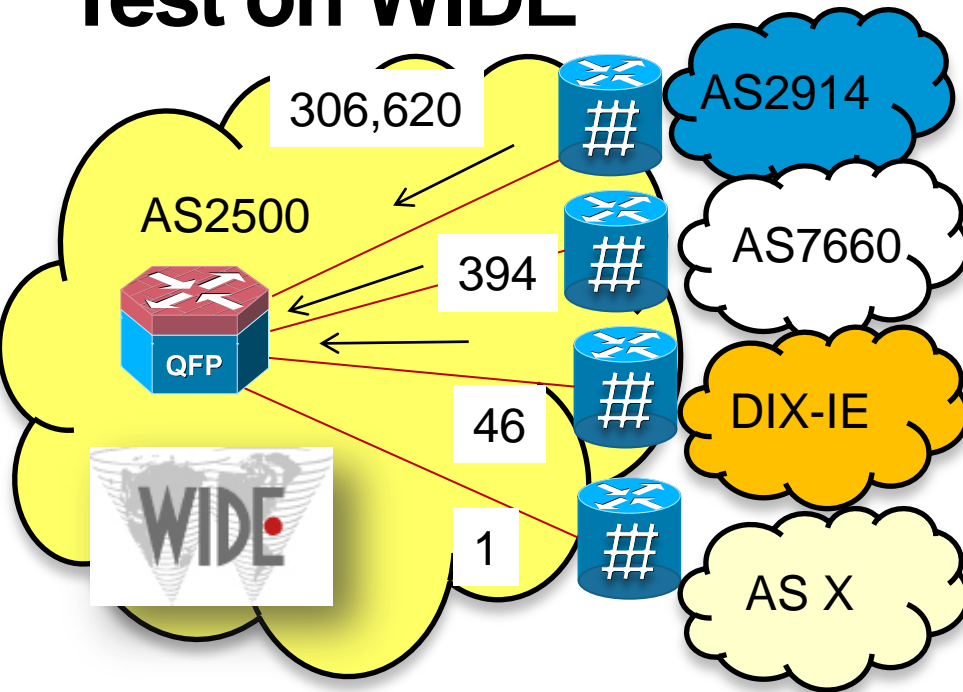
	BGP RIB	RIB	BGP Memory	RIB Memory
Normal	307,061	307,041	61MB	169 MB
S-VA	307,061	429	61MB	12MB

- Test result on WIDE(AS2500),using early beta code.

- 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)

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.