

ROAD TO CLOUD PEERING WITH SDN CONTROLLER

Shaowen Ma, Director, APAC Product, Juniper, mashao@juniper.net

Oct 10th, 2016

AGENDA

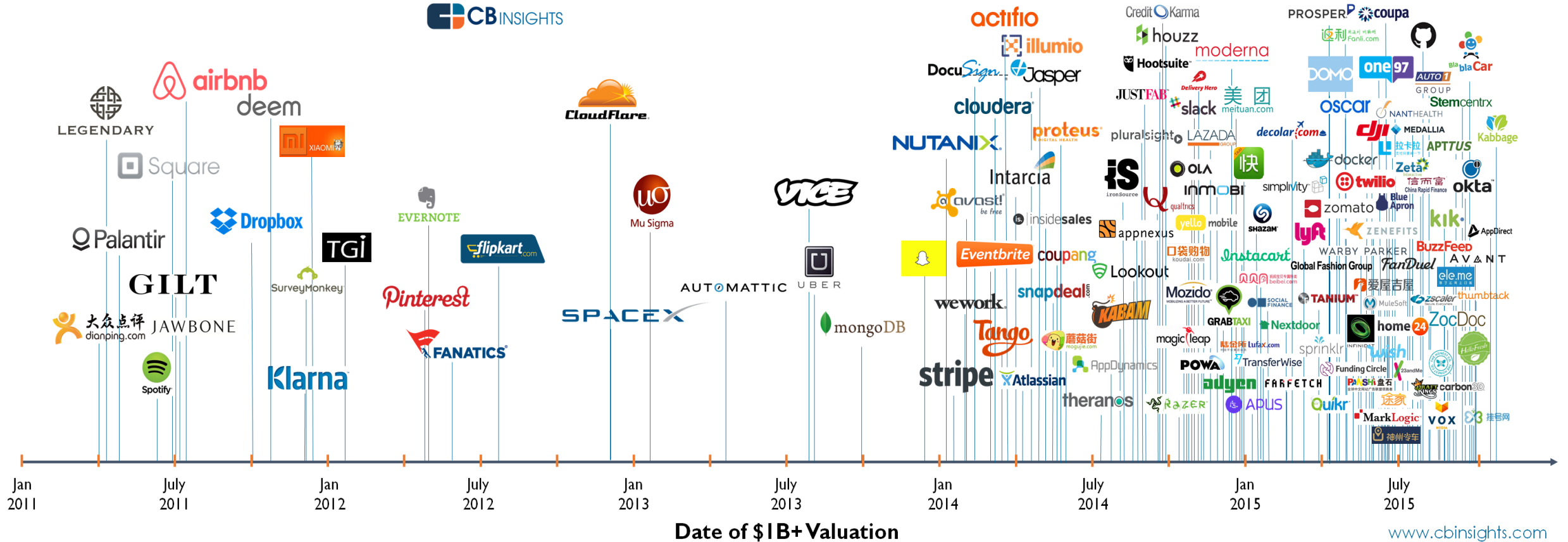
- 1 Inter-Domain Traffic Engineer
- 2 EPE and SPRING End-to-End solution
- 3 Use Case and Benefits
- 4 Summary

\$1B+ UNICORN COMPANY, INNOVATION NEVER SO FAST

The Increasingly Crowded Unicorn Club

Created by:

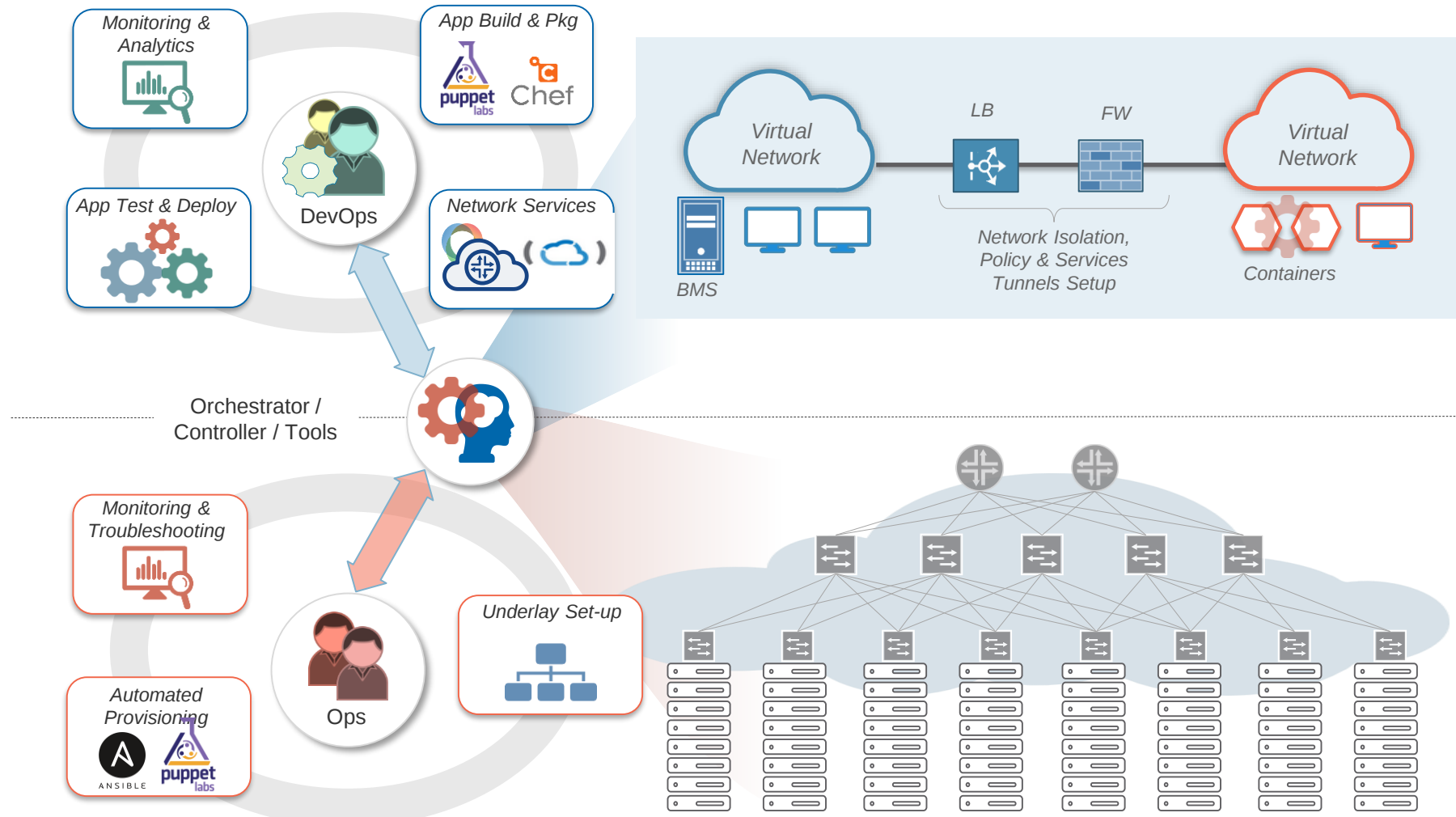
Private Unicorns since 2011



Source: <http://www.visualcapitalist.com/the-increasingly-crowded-unicorn-club/>

CLOUD – TELCO OR PRIVATE/PUBLIC

INFRASTRUCTURE AS CODE FOR DEVOPS AND AUTOMATION FOR INFRA OPS



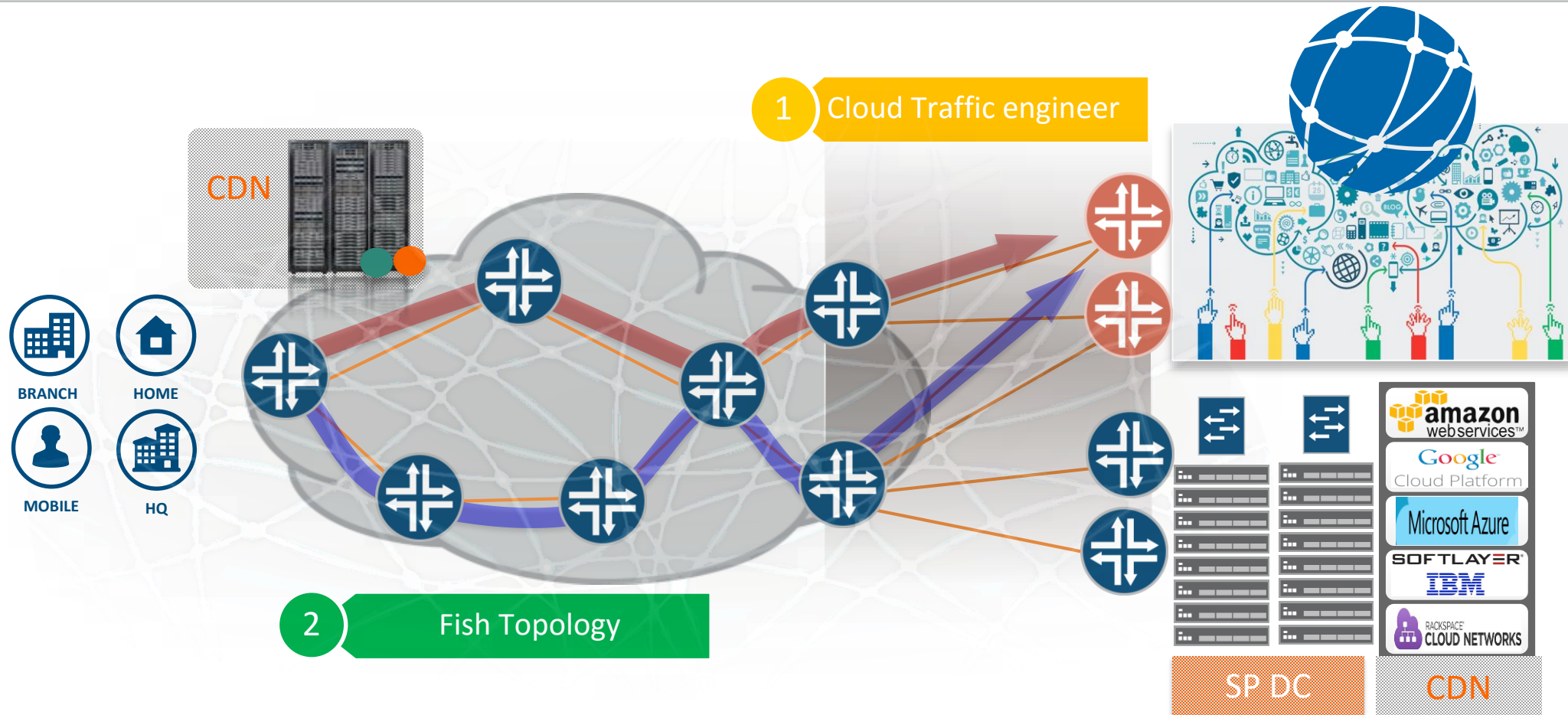
Developer Operations (DevOps)

- **Deploy Network & Network Services when you deploy Apps**
- **Tightly Integrated Orchestration of Compute, Storage, and the Network**
- **All Networking Services – VLANs, Routing, Policy, Load Balancing, IPAM/DHCP, DNS – is provided in software and virtualized**
- **No App specific knowledge in the underlying hardware**

Infrastructure Operations

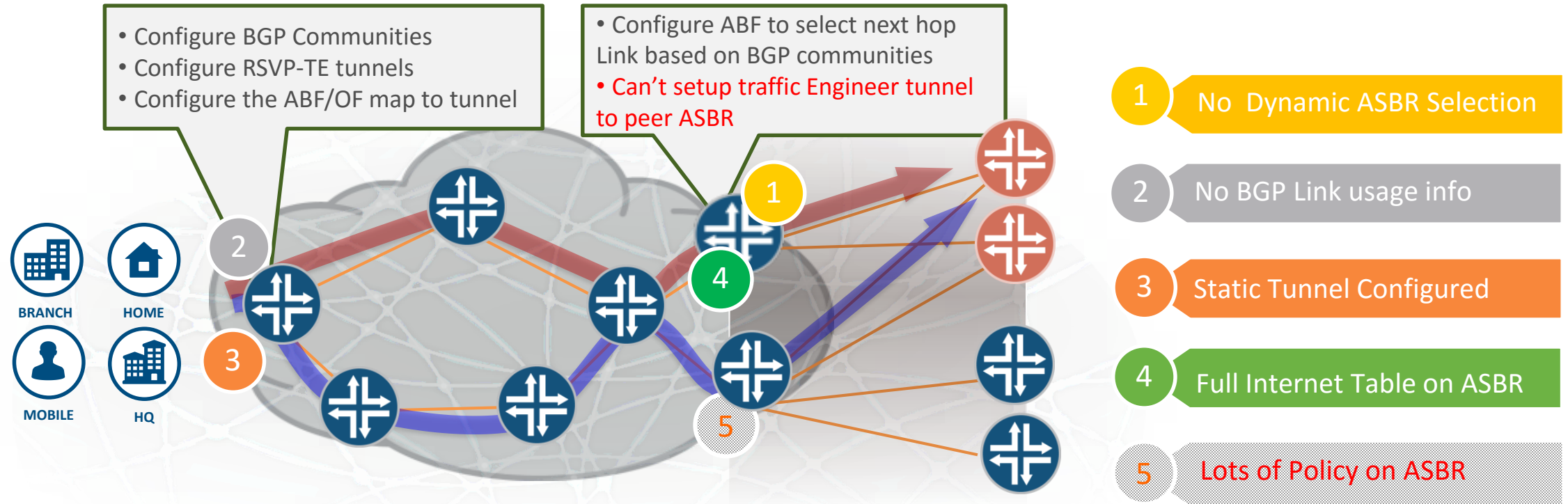
- **Cookie cutter approach to deploying Physical Hardware**
- **Automated Provisioning of Hardware - Gateway Routers, Spine Switches, TORs, Servers, JBODs, DAS ..**
- **Secure the Infrastructure, not the Apps**
- **Integrated Monitoring and Alerting of all hardware components**

INTER-DOMAIN CLOUD TRAFFIC ENGINEER



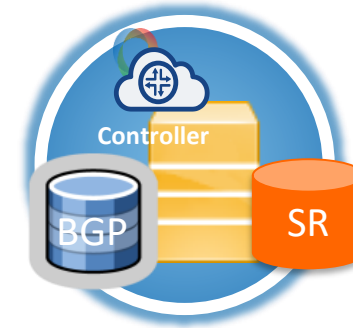
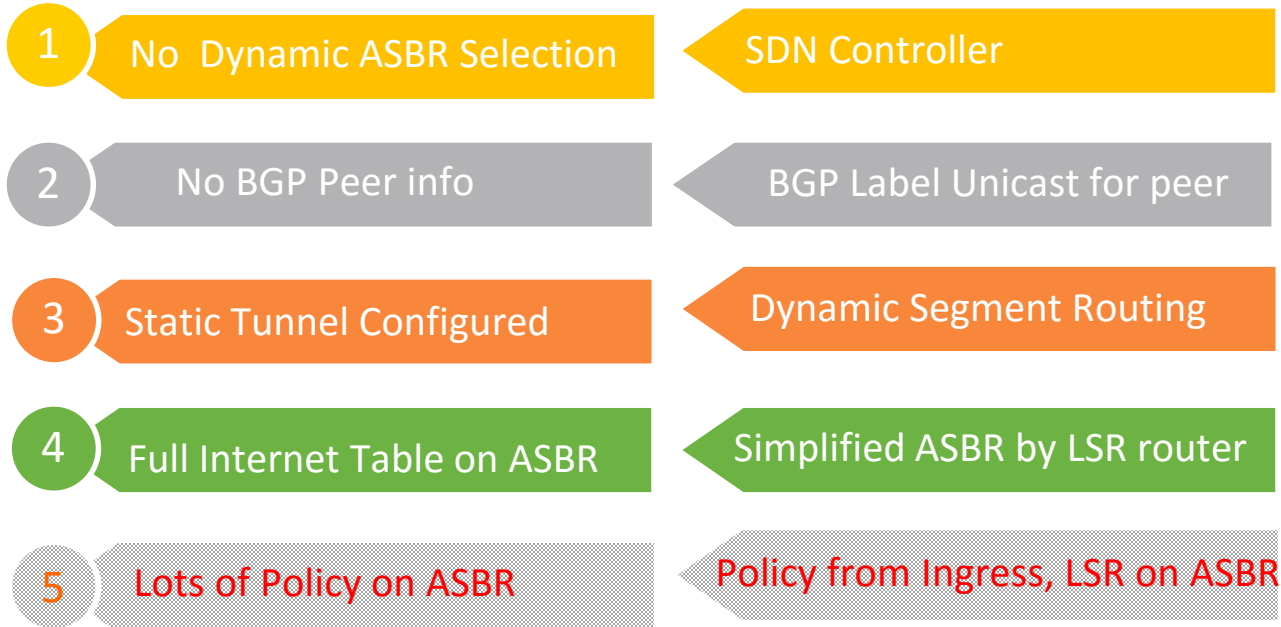
Easy to optimize End-To-End Traffic for SP Owned Network.
How to optimize VIP Customer for Internet/Cloud connection?

CURRENTLY SOLUTION AND LIMITATIONS



Current Solution can't meet Cloud Traffic Engineer Requirement

CLOUD TRAFFIC ENGINEER SOLUTION COMPONENTS



Egress Peering Engineer

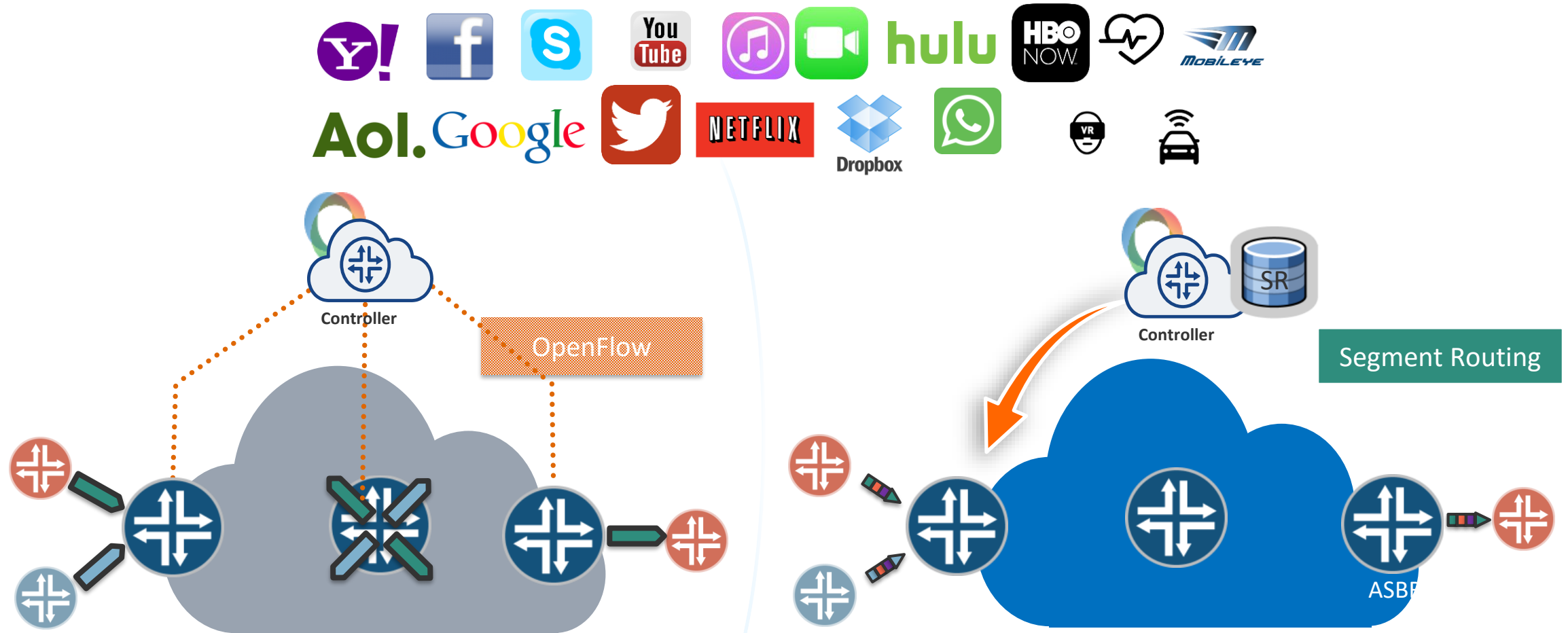
- draft-gredler-idr-bgplu-epe-04
- draft-ietf-idr-bgpls-segment-routing-epe-02
- draft-ietf-spring-segment-routing-central-epe-00

Segment Routing

- draft-ietf-spring-segment-routing-0x

With Controller, Segment Routing and LSR Switch can build
Cheaper and Optimized Cloud traffic Engineer

SDN 2.0 ERA



Segment Routing, RSVP-TE Enable SDN 2.0
Edge Intelligence, Stateless CORE

BGP EPE DESIGN PHILOSOPHY

How to Select Which Peer to send

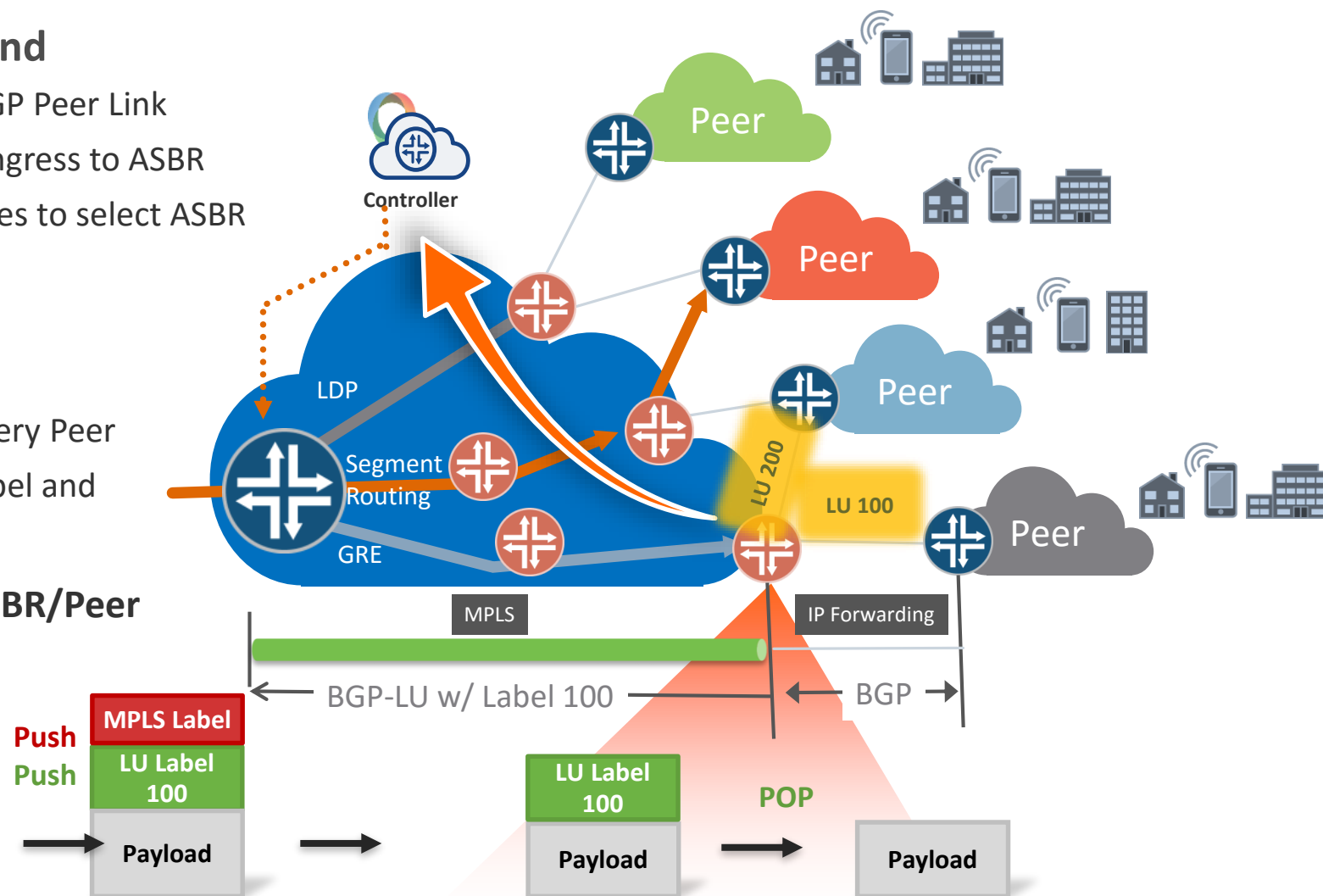
- Controller/RR may monitoring the BGP Peer Link
- Controller/RR find a tunnel from Ingress to ASBR
- Controller/RR based on certain rules to select ASBR

How ASBR identify a Peer

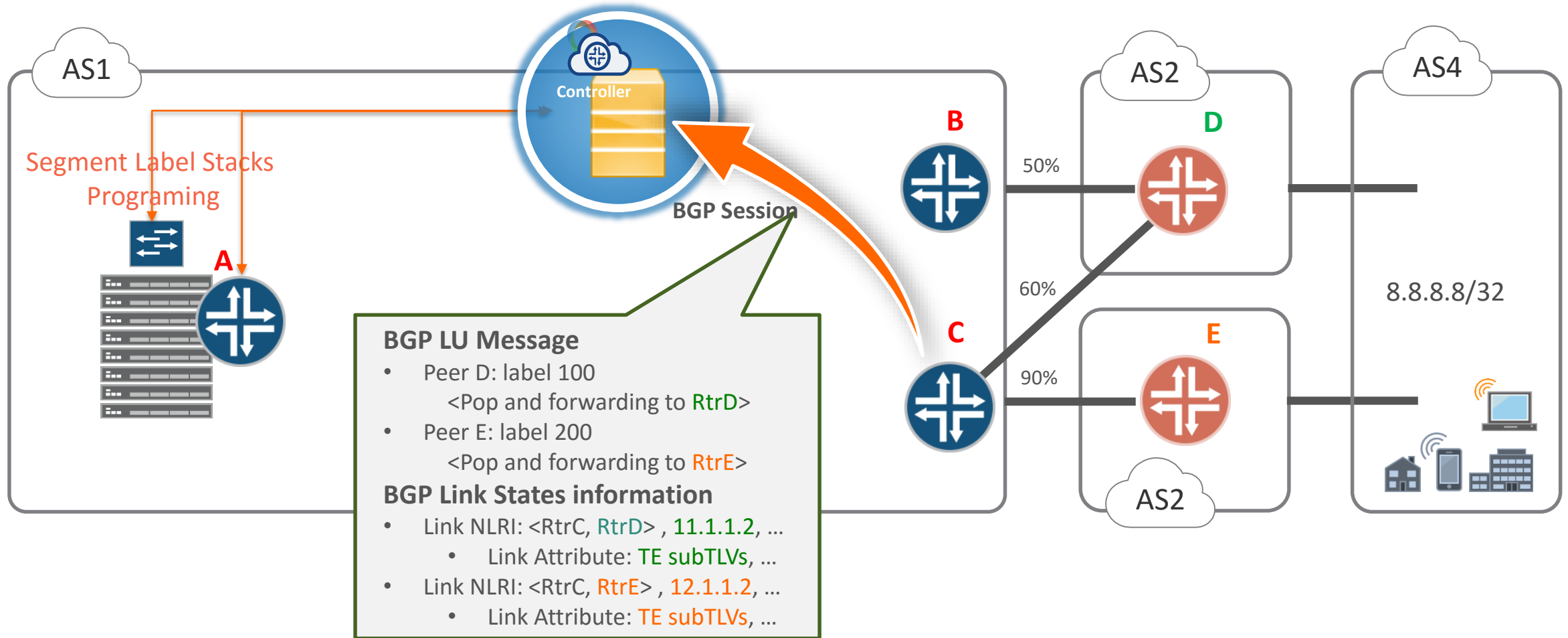
- Per Peer /32 address per label
- Install the MPLS Label POP for every Peer
- When ASBR received different label and send traffic to specific Peer

How Ingress mapping traffic to ASBR/Peer

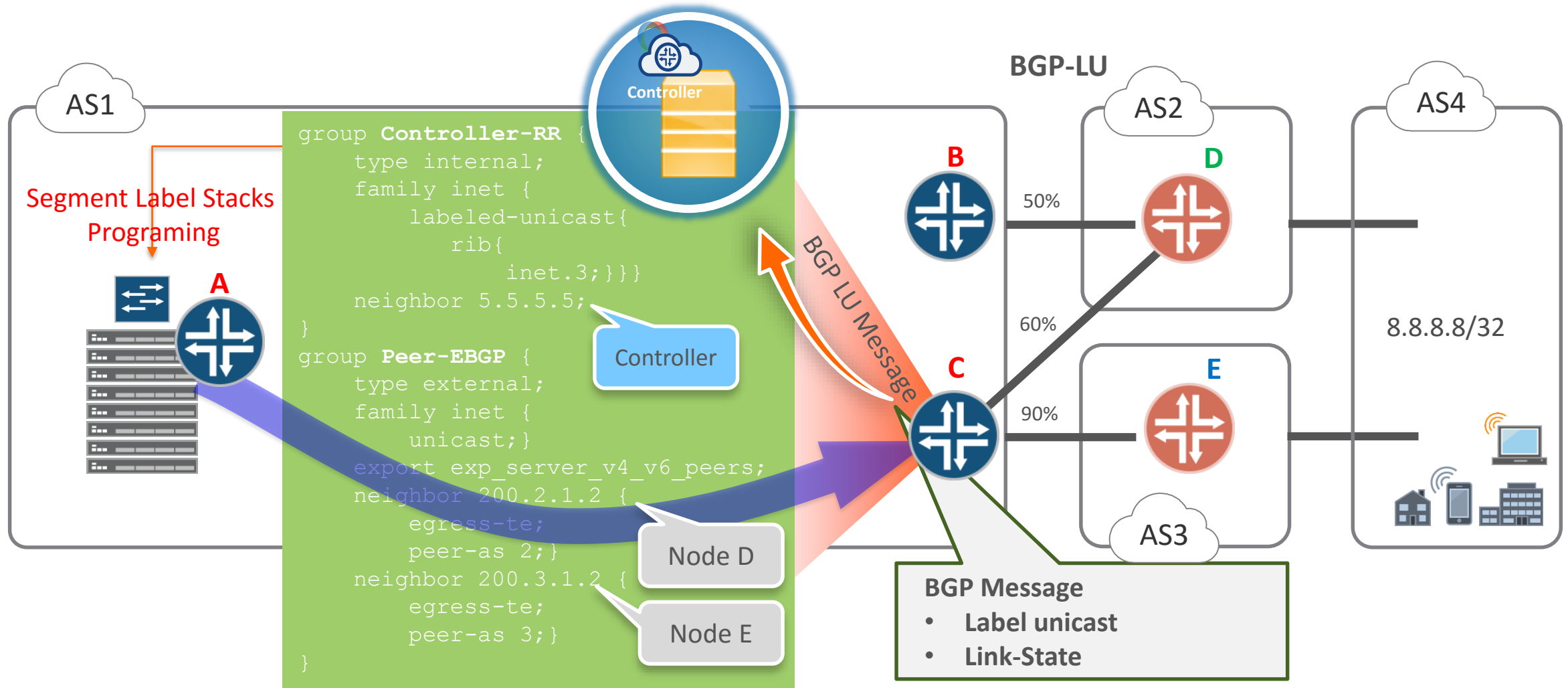
- Ingress push tunnel label to ASBR
- Ingress push BGP-LU label



BGP EGRESS PEER ENGINEERING/BGP-LU DETAIL



BGP EGRESS PEER ENGINEERING/BGP-LU DETAIL



BGP-LU WITH SEGMENT ROUTING

draft-rosen-idr-rfc3107bis-00.txt

BGP-LU Session between Controller/Router

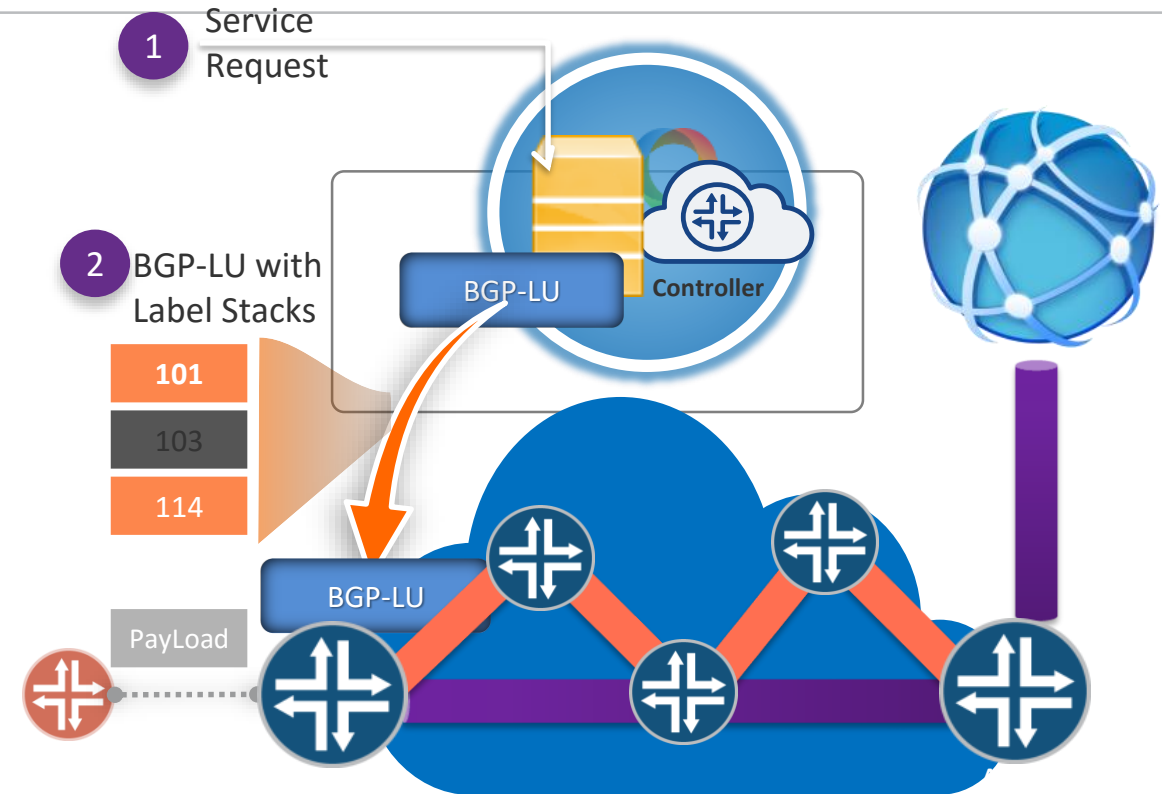
- BGP LU carrier the label stack for SR/LSP
- BGP-LU carrier the Label stack for LSP + VPN Service

BGP-LS get the network information

- TEDB information with label send back to Controller
- draft-gredler-idr-bgp-ls-segment-routing-ext-xx.txt

BGP is the only protocol for Service and Tunnel

- QPPB/BGP FlowSpec
- With additional Openflow/PBR



Example from ExaBGP

```
bespalov@CentOS-1 ~/exabgp-3.4.16/sbin>cat ~bespalov/config/exabgp neighbor 192.168.255.12 {  
  local-address 192.168.255.2;  
  peer-as 65000; local-as 65000;  
  family { ipv4 nlri-mpls; }  
  static {  
    route 10.255.255.8/32 {  
      next-hop 10.0.0.2;  
      label [ 101 103 104 800008 ];  
    }  
  }  
}
```

BGP LU AND FRR BACKUP

BGP LU no IP lookup on ASBR

Normally MPLS label POP and forwarding

- Per Peer /32 address per label

Use FRR in same ASBR

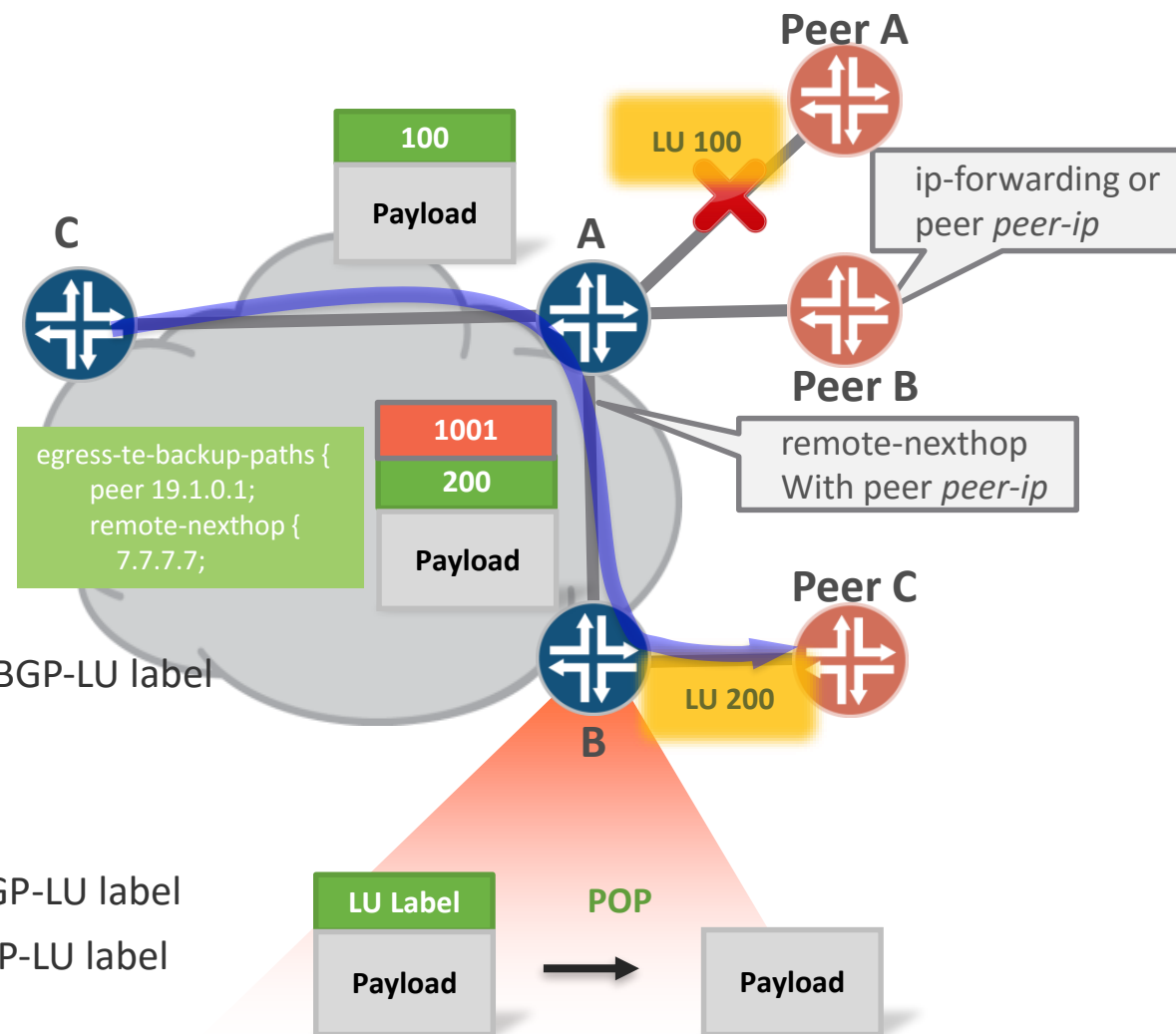
- IP-forwarding option, remove the LU label and then IP lookup
- Peer *peer-ip*, send to backup peer directly

For remote ASBR, leverage remote-nextthop for FRR

- Setup tunnel to remote ASBR
- Get rid of BGP-LU label and replace by the remote neighbor's BGP-LU label

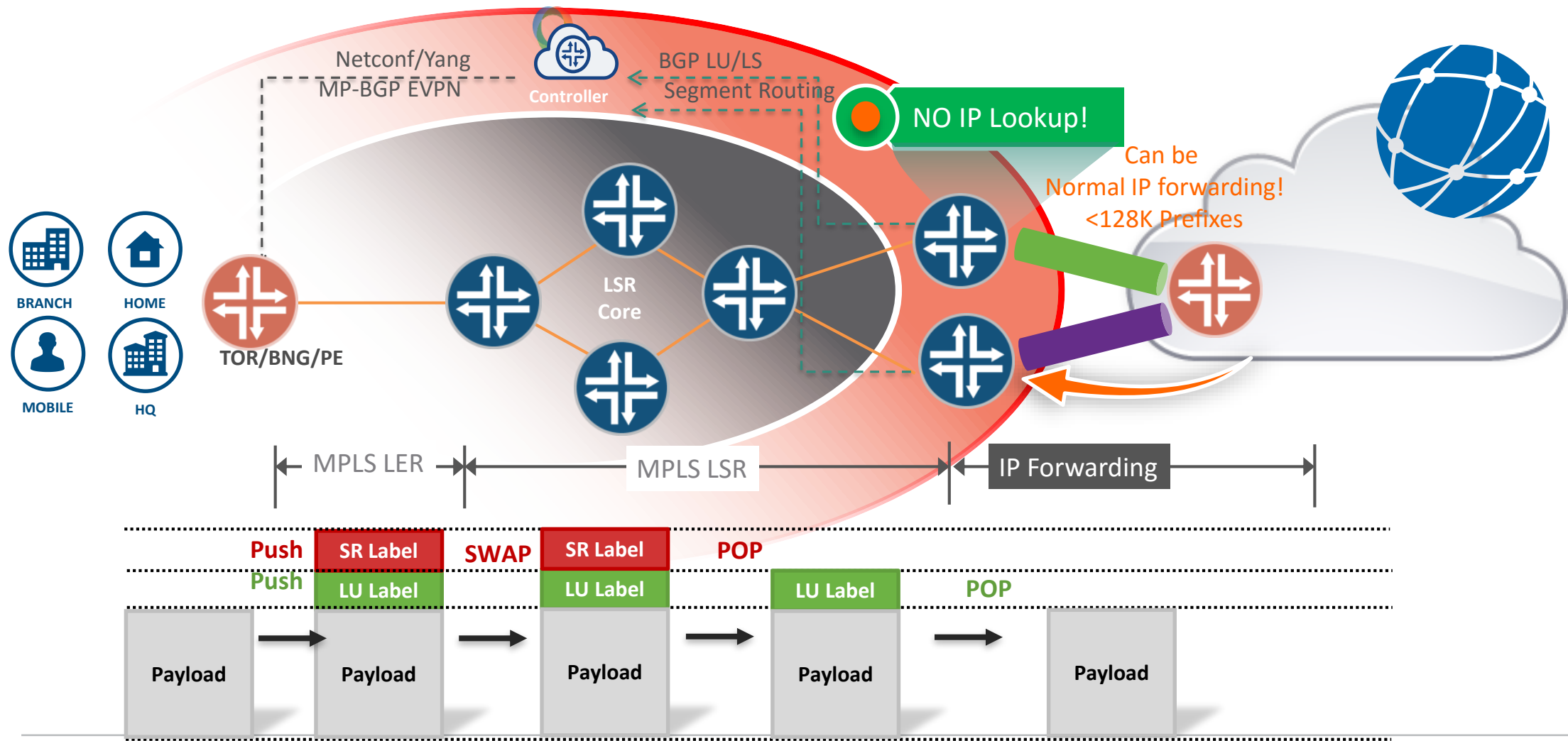
For ASBR failure, will need Egress Protection

- Remote Anycast ASBR need to understand the Failure ASBR BGP-LU label
- Or Remote Anycast ASBR will just do IP-lookup, ignore the BGP-LU label
- For future implementation



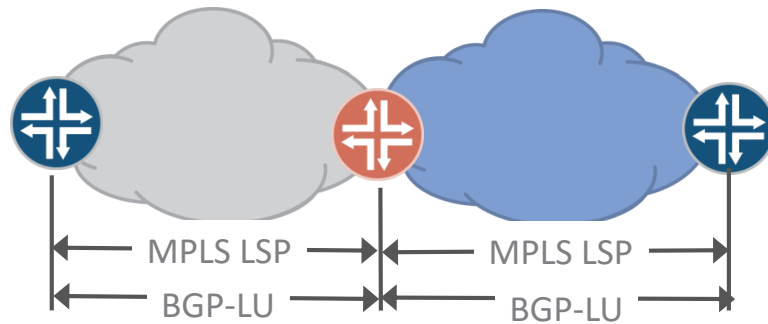
BGP-LU EPE & MPLS KEY BENEFITS

EXTEND HOLLOW CORE/LSR TO PEERING, CHEAPER PEERING SOLUTION



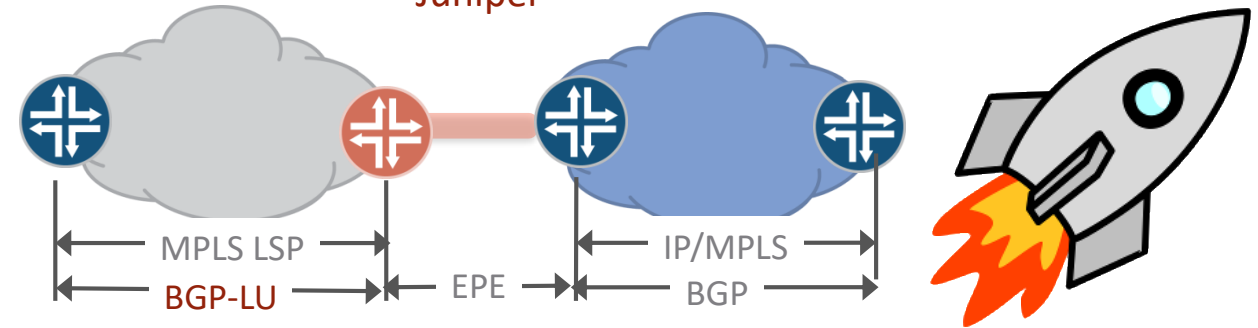
BGP-LU VS BGP-EPE FOR PEER TRAFFIC ENGINEER

BGP-LU for Seamless MPLS



BGP-LU for Peer Traffic Engineer

Juniper

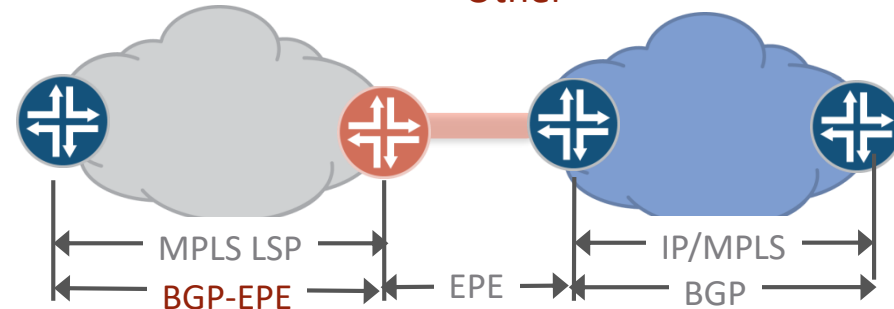


Egress Peering Engineer, 2 different Methods

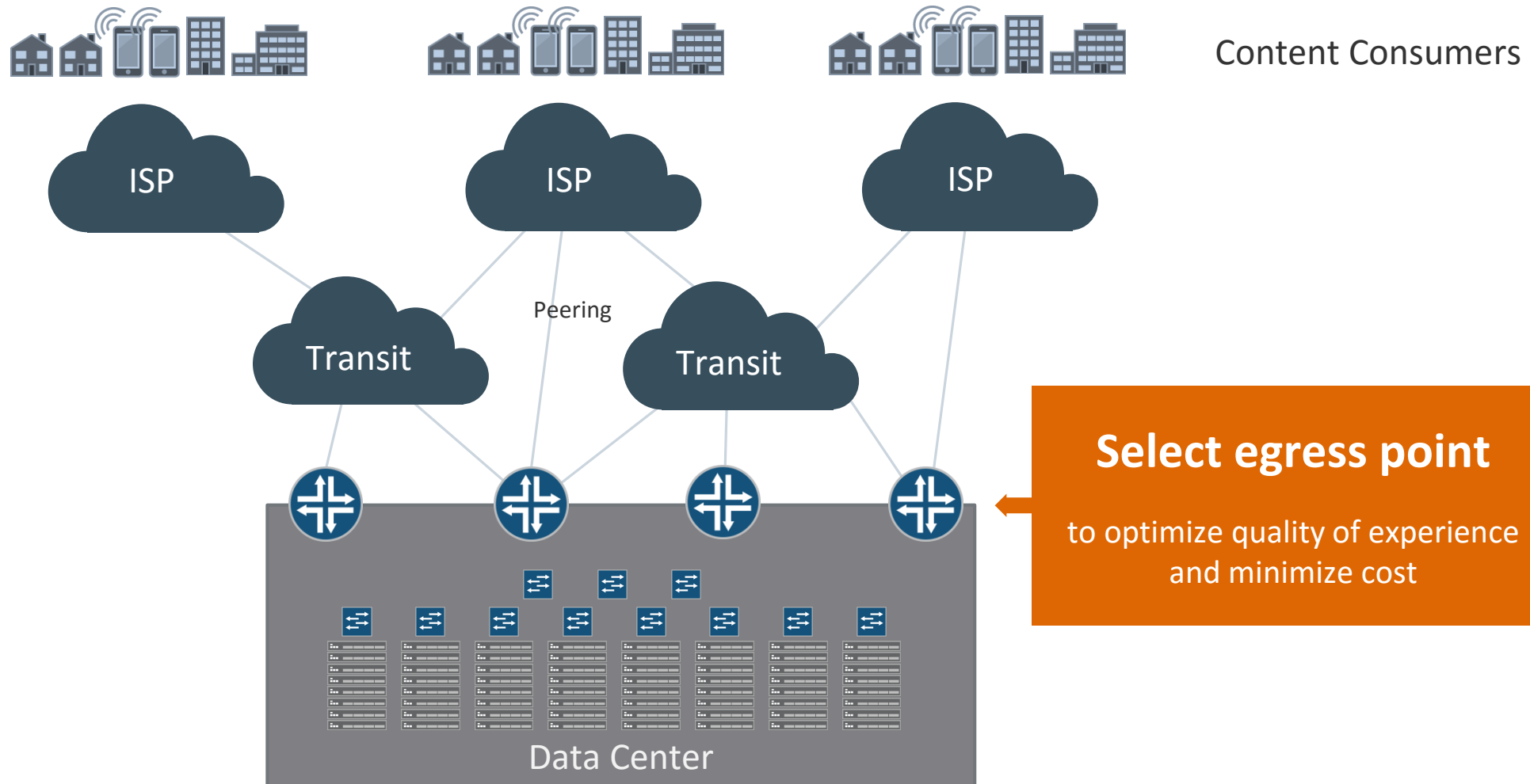
- **[Juniper]** draft-gredler-idr-bgplu-epe-04
 - No New Address Family, Any tunnel can apply
 - build in FRR method for Peer failure
 - Existing Solution with enhance, Fast Time to Market
- **[other]** draft-ietf-idr-bgpls-segment-routing-epe-02
 - New Address Family, request SR tunnels
- Both Assign a Label for Peer, no need upgrade Peer router, peer can be IP or MPLS forwarding

BGP-EPE for Peer Traffic Engineer

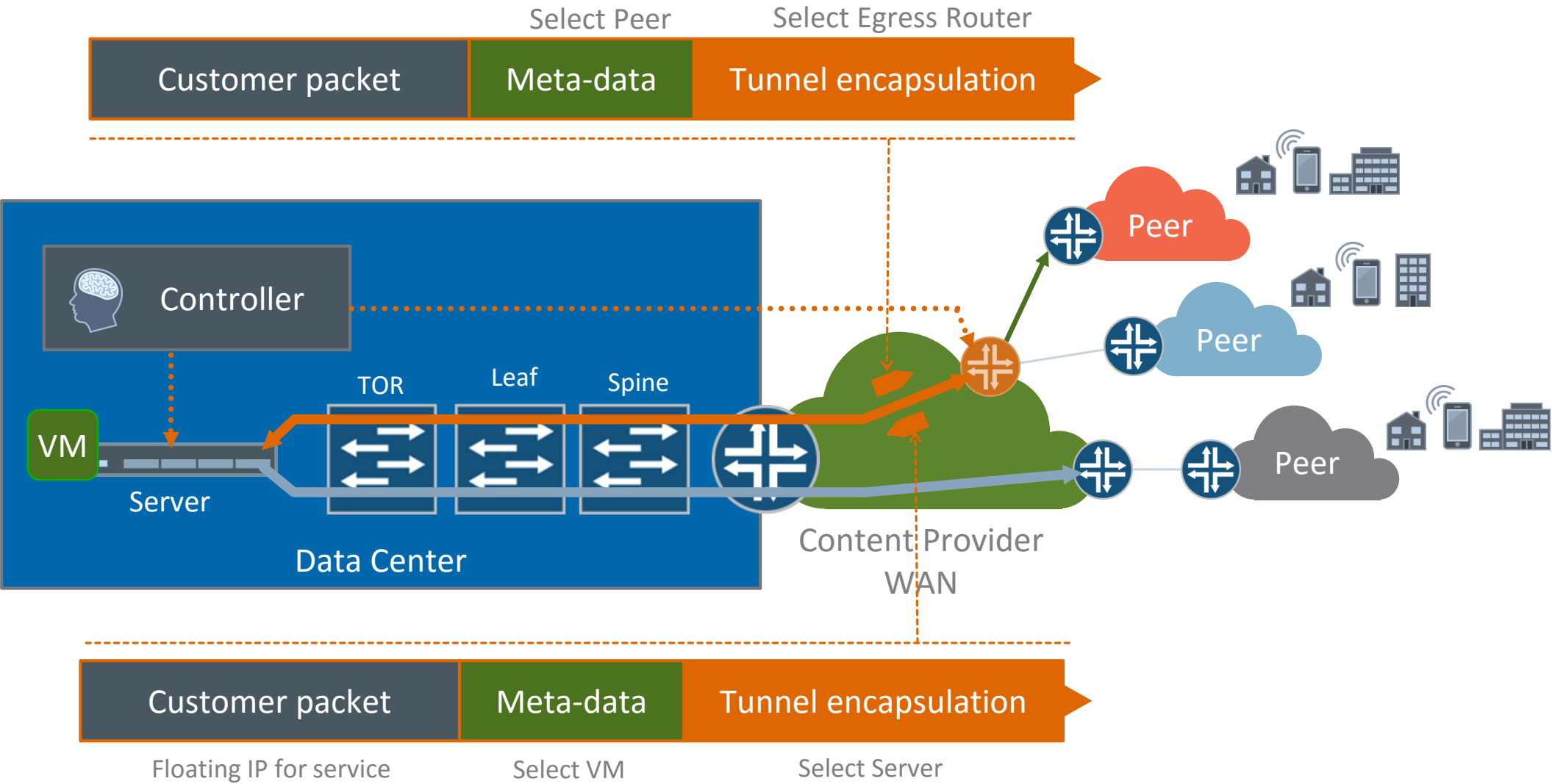
Other



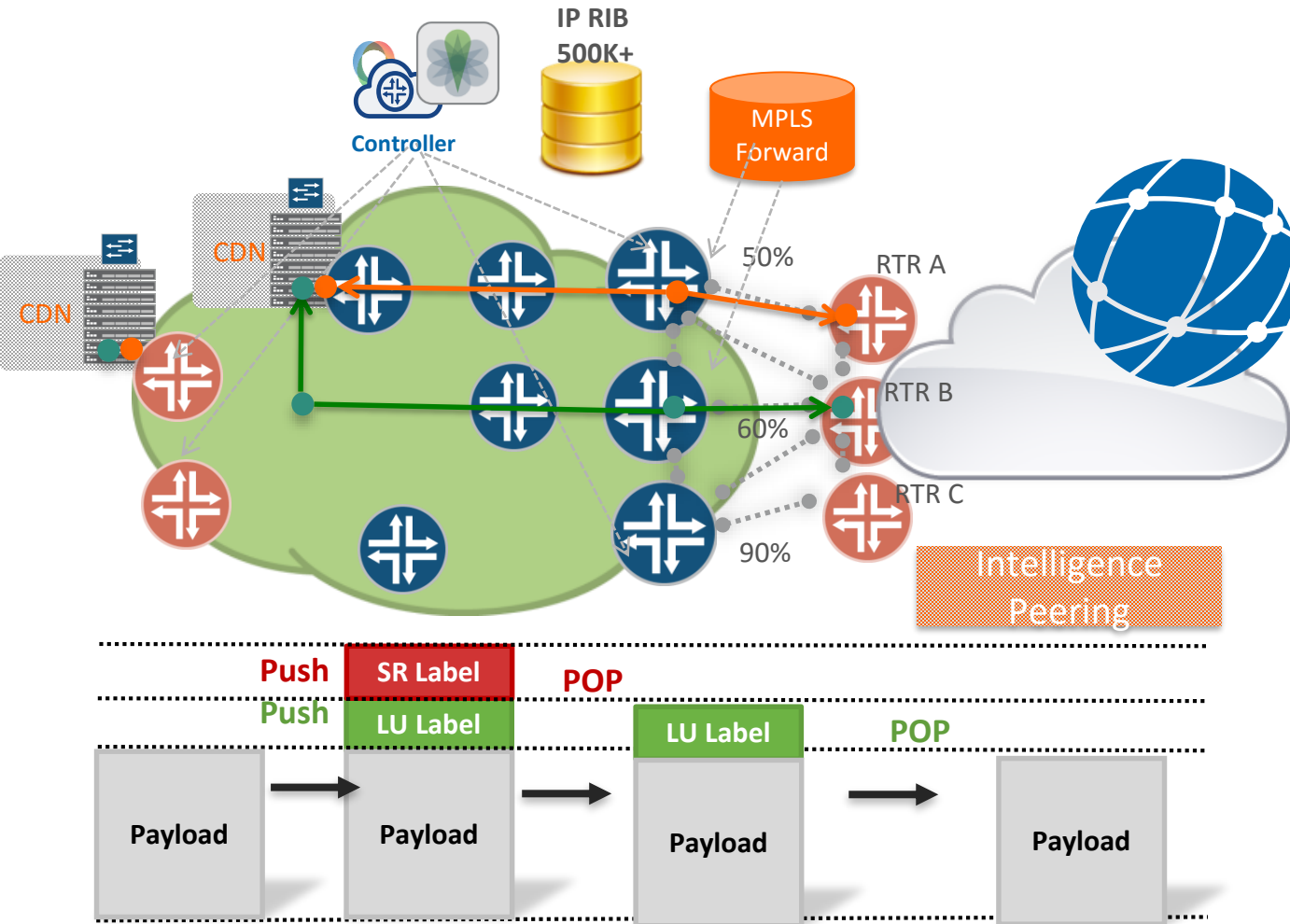
EGRESS PEER ENGINEERING (EPE) USE CASE IN DC



CONTROLLER AND EPE USE CASE



USE CASE, CDN PEERING



ASBR Setup BGP session and pass BGP-LU and BGP-LS information to Controller.

Controller Calculate the Path

- Controller select which Peer A/B/C send traffic to with LU label.
- Controller and ASBR take part in the Segment routing domain, and know to send traffic to ASBR adding a IGP/SR label or tunnels
- Controller will send MPLS label Stack to Ingress Router or Host

Controller keep monitor path and Egress link

- When Congestion happens, will automatically redirect traffic to another ASBR/Peer by changing the label stack

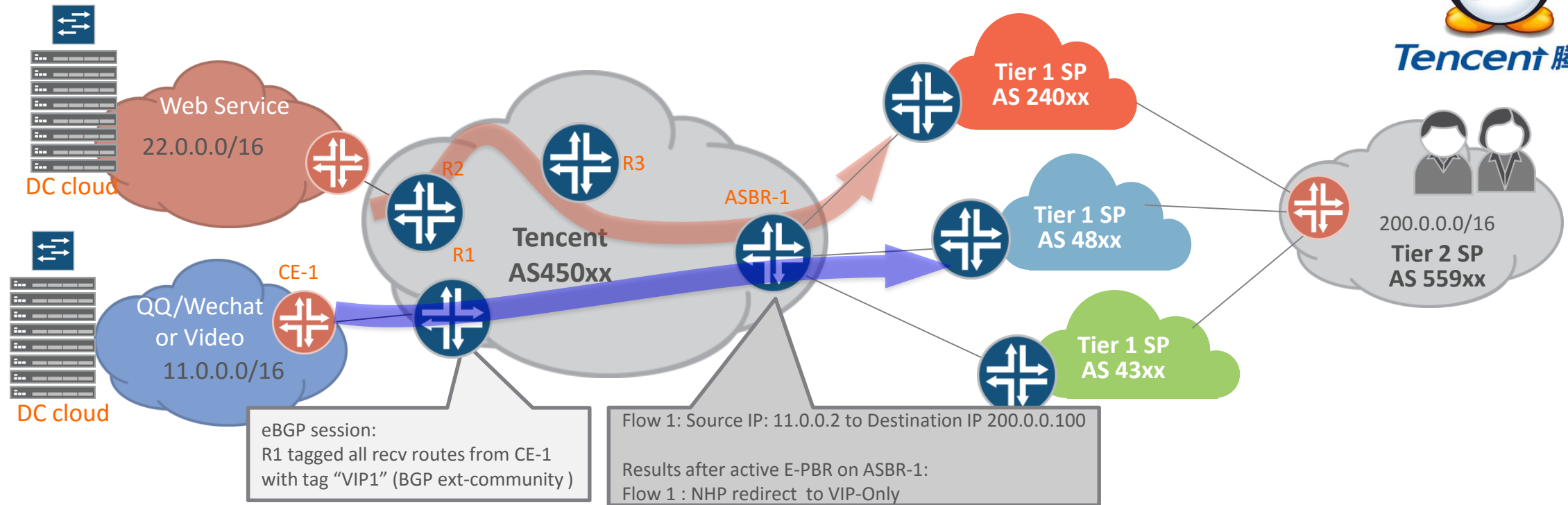
Separate Control/Forwarding

- Controller Full Internet Table, RIB, Control Plane only.
- ASBR only Keep LSR label switching, Forwarding Plane, No IP lookup

Policy start from Ingress

- Linux Host/Hypervisor/switch/router

WITHOUT EPE, CURRENT SOLUTION (TENCENT)



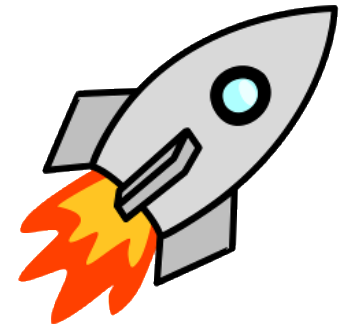
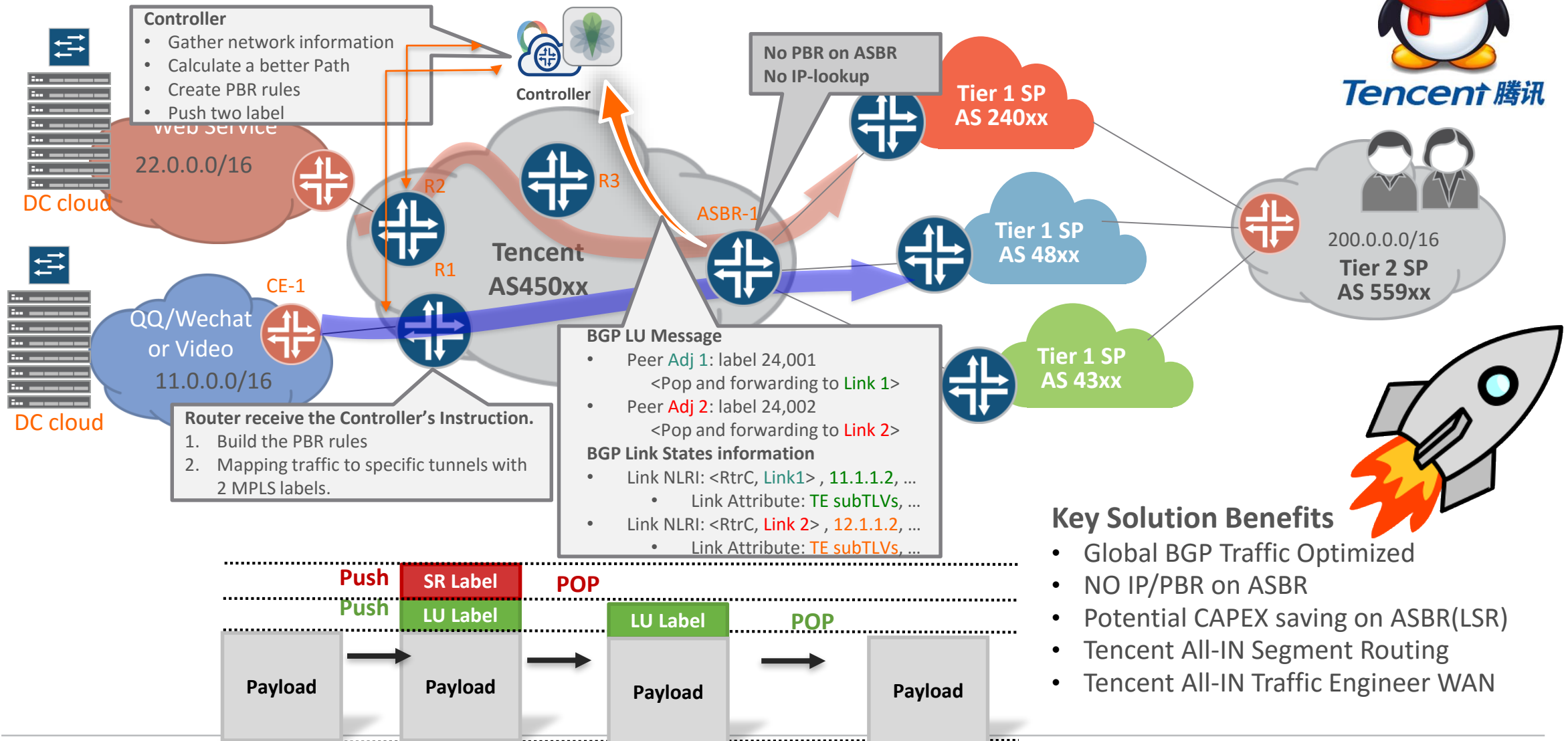
Tencent Peering Situation

- Peering with many Tier 1 and 2 SP, around 20+ peer AS.
- Peering from 4-5 cities across China, Beijing/Shanghai/Shenzhen/Guangzhou etc.
- Peering with Hongkong for international directly

Key Pain Points

- No Global BGP traffic Engineering optimization
- Static RSVP tunnel, A lot of Policies on ASBR.
- Peering Traffic Grow so fast, how to save CAPEX on ASBR?

WITH EPE, PLANNING SOLUTION (TENCENT)

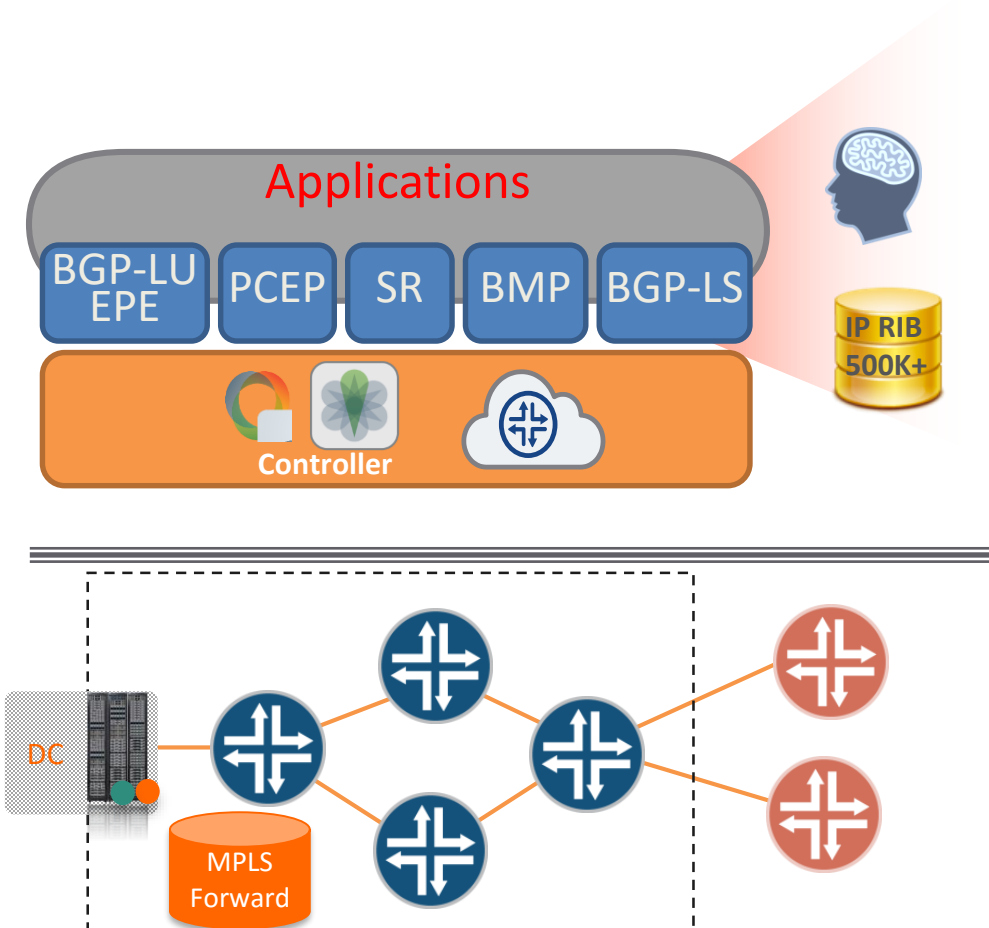


Key Solution Benefits

- Global BGP Traffic Optimized
- NO IP/PBR on ASBR
- Potential CAPEX saving on ASBR(LSR)
- Tencent All-IN Segment Routing
- Tencent All-IN Traffic Engineer WAN

APPLICATIONS DETAILS

SAME TECHNOLOGY FOR DC & WAN



Application is the Network Brian:

- BGP-LU EPE information from ASBR for peer label and internet prefix table.
- BGP-LS/Netflow information for all link TE TLV, and BMP for Prefix
- PCEP, Calculate Segment routing tunnel and apply 2+ labels in network
- Traffic Steering/mapping to tunnels, and monitor tunnels
- Easier to calculate Latency based routing for network wide optimized.

Controller for Segment routing Traffic Tunnel setup/monitor

Separate Control/Forwarding

- Controller Full Internet Table, RIB, Control Plane only.
- ASBR only Keep LSR label switching, Forwarding Plane, No IP lookup

SUMMARY

1

Extend Traffic Engineer to Cloud, Global Network Optimized

2

SDN Controller Solution, Automatic Congestion/Latency Optimized

3

Simplified ASBR Design, no IP, no Policy, LSR only

4

Controller/Application support full intelligence RIB/Traffic Telemetry

5

Standard Based solution, work with existing peer ASBR

THANK YOU