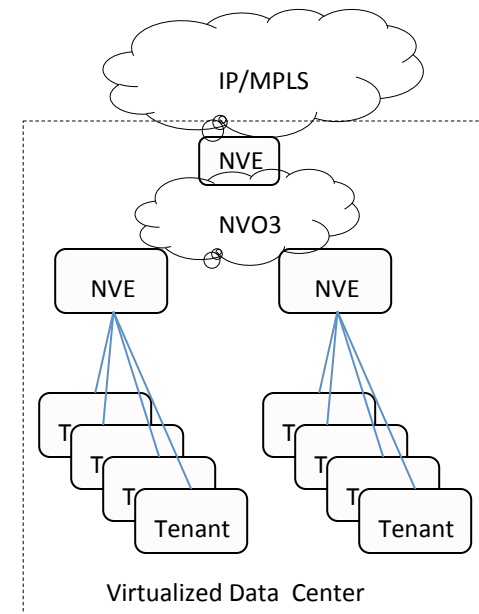# 87th IETF @ Berlin

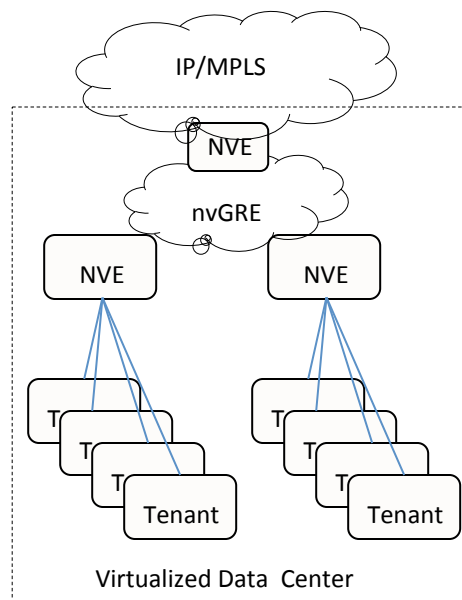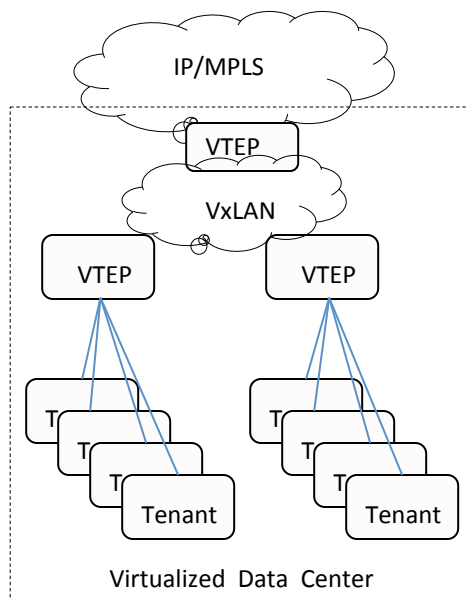## MPLS Big Label and BGP Signaling
draft-renwei-mpls-big-label-00.txt
draft-renwei-mpls-bgp-big-label-00.txt

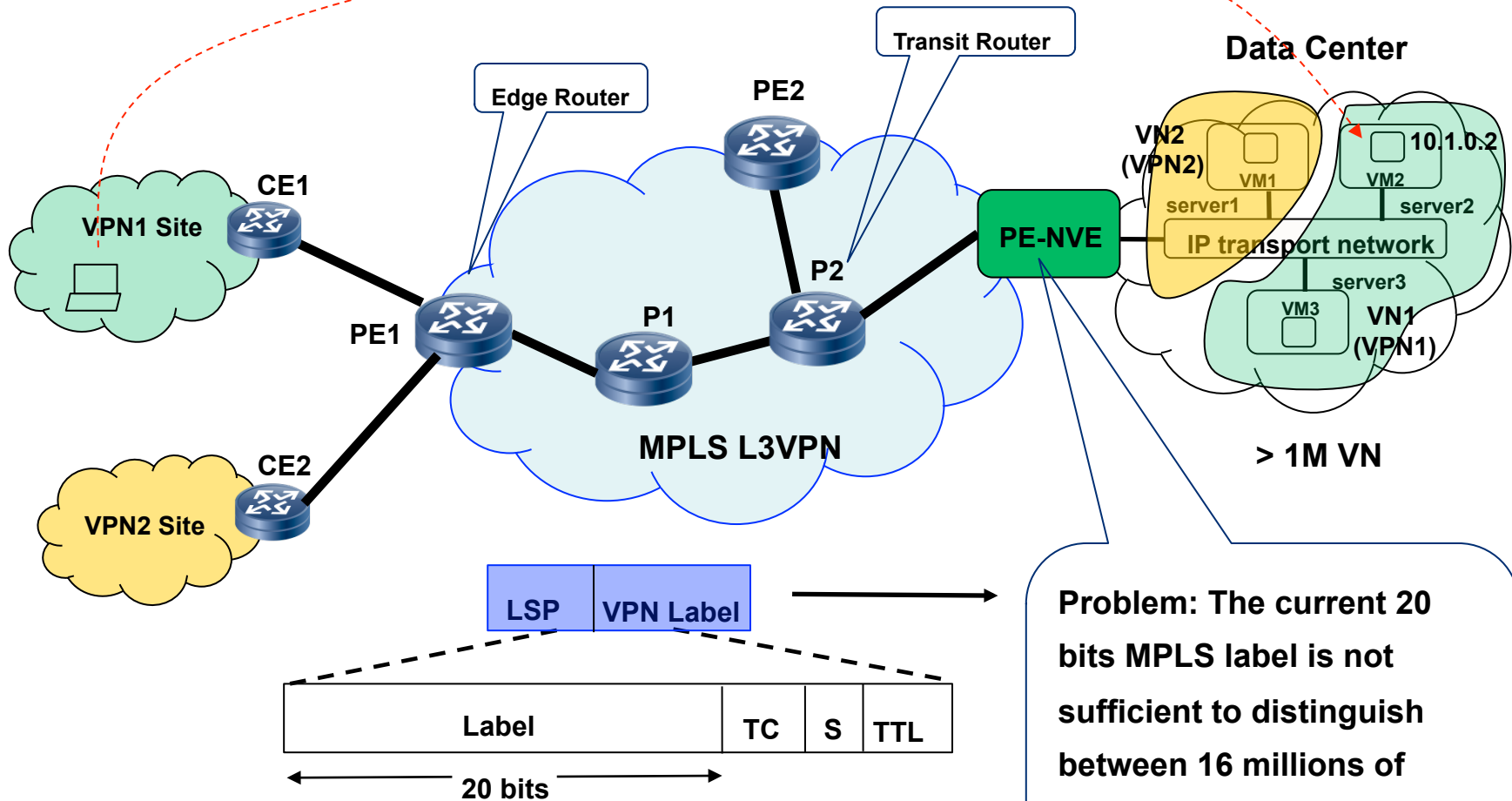Richard Li (renwei.li@huawei.com)

Background

- Supporting large scale network virtualization in data center has become a commonly agreed requirement
- VXLAN, NVGRE and NVO3 are currently being standardized, with all supporting up to 16 millions of virtual overlaid networks

Problem Statement

How to access VM2 in VN1 using L3VPN

Transit Router

Data Center

PE2

Edge Router

PE-NVE

VN2 (VPN2)

10.1.0.2

CE1

VM1

VM2

VPN1 Site

server1

server2

IP transport network

PE1

P2

server3

P1

VM3

VN1 (VPN1)

MPLS L3VPN

CE2

VPN2 Site

> 1M VN

LSP	VPN Label

| Label	TC	S	TTL

← 20 bits →

Problem: The current 20 bits MPLS label is not sufficient to distinguish between 16 millions of virtual networks hosted in a Data Center

NVE : Network Virtualization Edge

VN: Virtual Network

Context-Based Solution

Context-1	TC	0	TTL
VPN Label	TC	1	TTL

■ ■ ■

Context-n	TC	0	TTL
VPN Label	TC	1	TTL

■ ■ ■

Context-16	TC	0	TTL
VPN Label	TC	1	TTL
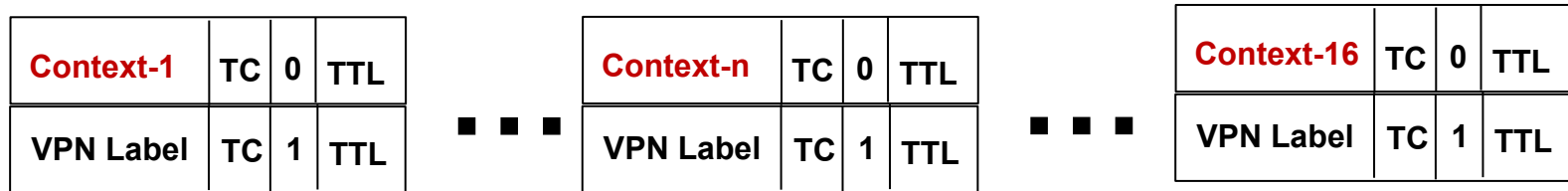
- Need a total of 16 contexts for 16M of VPNs
- The same VPN labels may appear in different contexts for different VPNs
- Two labels jointly serve to simulate one label
- It needs **two** labels, **two** NHLFEs, **two** table lookups
  - Implementation complexity in both hardware and software: a hierarchical or embedded data structure is in need.
  - Slow down the data plane because of two label lookups and the associated actions.
- In disagreement with old L3VPN (old L3VPN requires one VPN label, and at least three commercial implementations assumes so)
- Association between contexts and VPN labels are arbitrary: one can swap VPN labels and their associated NHLFEs between different contexts
- BGP needs to distribute both contexts and VPN labels, and thus adversely affect the re-convergence time.
- A 20 bits value is generally represented by a 32 bits integer, and thus 64 bits are needed in the VPN ILM for a single VPN (remember we are talking about 16M VPNs)

Proposed Solution: Big Labels

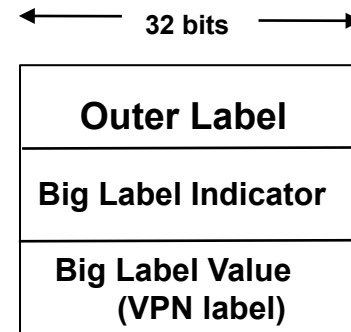| 20 bits	3 bits	1 bit	8 bits
Big Label Indicator	TC	S	TTL
Big Label Value			
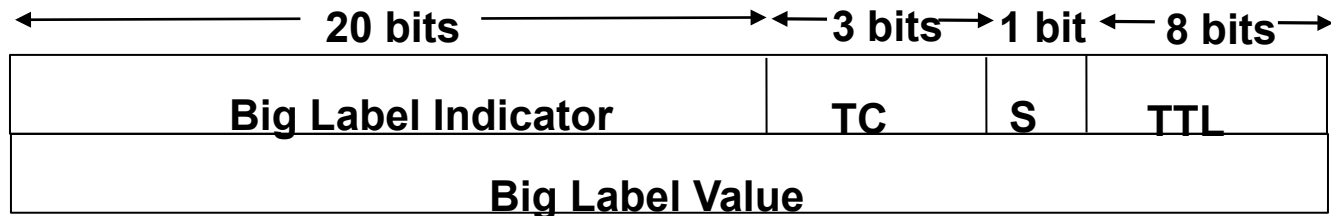
- They are viewed as one and only one label

- No need for two NHLFEs and two table lookups.

- Compliant with the BGP/MPLS L3VPN architecture: the BGP/MPLS L3VPN uses ONE label to represent the VPN/VRF but doesn't use TWO labels for that purpose.

- BGP just needs to distribute the Big Label Value instead of two labels needed in the context-based solution

- A 20 bits value is generally represented by a 32 bits integer. And thus the above format doesn't add complexity to the data structures in software (I am aware that three commercial operating systems represent the 20 bits label with 32 bits integer)

Big Label in MPLS IP VPN

- **The outer label is unchanged**

- **The inner label (VPN label) can be either the common 20 bits label or the current proposed big label**

- **The Big Label Format is compatible with the MPLS Architecture and L3VPN model, but is slightly different from the current one with respect to the de-capsulation of the VPN label.**

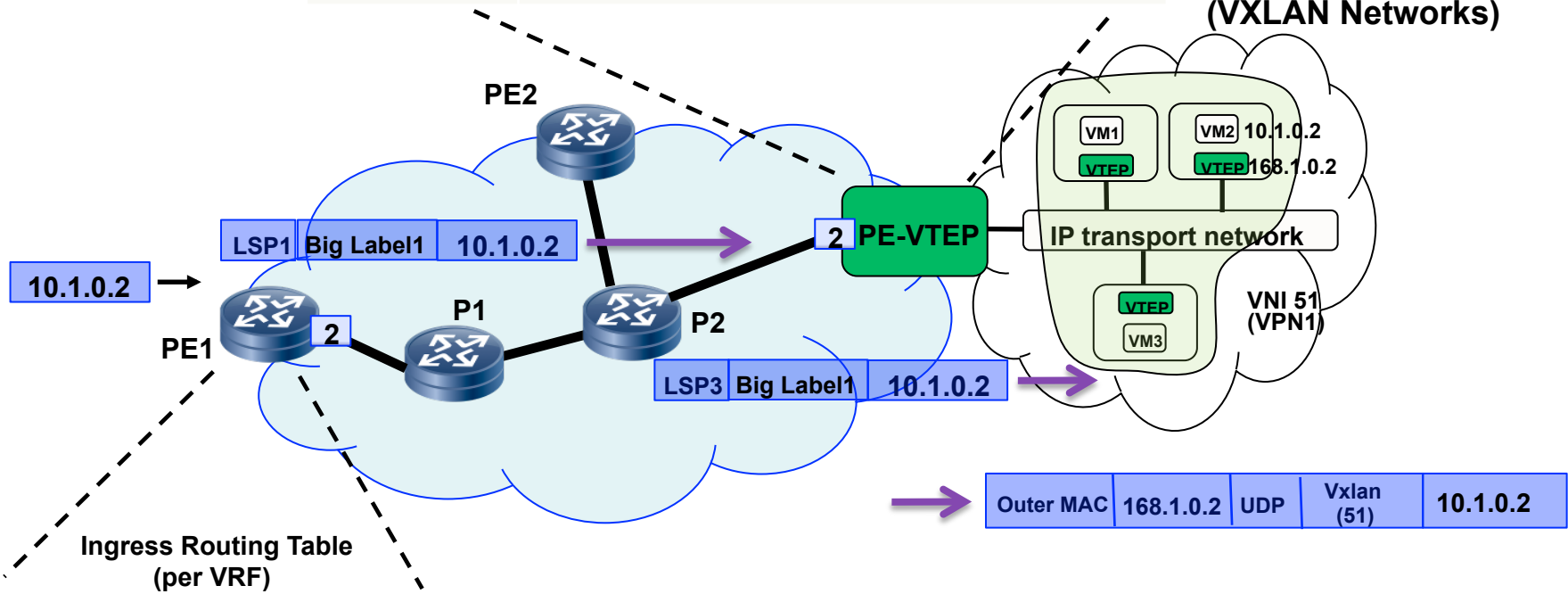| ← 32 bits → |
|---|
| **Outer Label** |
| **20bits Label Value (VPN label)** |
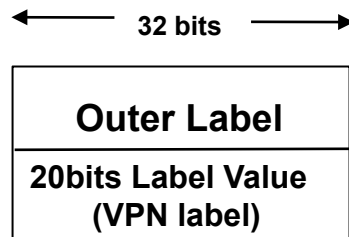
Two Entries with Common Labels

← 32 bits →
Outer Label
Big Label Indicator
Big Label Value (VPN label)

Two Entries with Big Label

Packet Life: From Customer to DC

Egress PE-VTEP FW Table

Destination	Out
10.1.0.0/24	(168.1.0.2, VTEP_MAC1. VNI1(51), VM_MAC1)

Data Center

(VXLAN Networks)

PE2

LSP1	Big Label1	10.1.0.2

10.1.0.2

2	PE-VTEP

VM1

VM2 10.1.0.2

VTEP

VTEP 168.1.0.2

IP transport network

PE1

P1

P2

VTEP

VM3

VNI 51
(VPN1)

LSP3	Big Label1	10.1.0.2

Outer MAC	168.1.0.2	UDP	Vxlan (51)	10.1.0.2

Ingress Routing Table
(per VRF)

Destination	Next Hop
10.1.0.0/24	(2, LSP1, Big Label1)

Extension of MP-BGP

- **New SAFI is introduced for big label signaling**
- **BGP capability advertisement for new SAFI if MPLS big label is used for NLRI**
- **New NLRI Format**

length
Common Label (3 octets)
Prefix (variable)

length
Big Label Value (4 octets)
Prefix (variable)
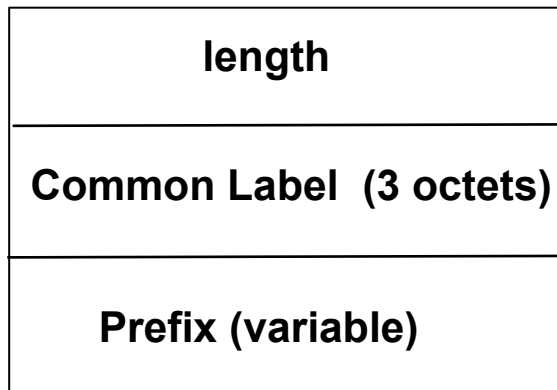
Current NLRI for Common Label   **New NLRI for Big Label**
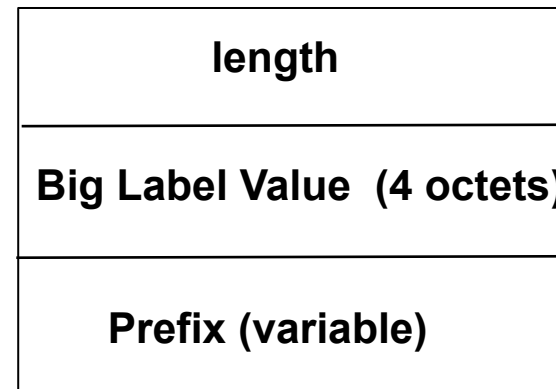
Next Step

- **Seeking comments from you**
- **Experimenting and prototyping**
- **Exploring and comparing with other solutions**