IS-IS VPLS for Data Center Network

draft-xu-l2vpn-vpls-isis-02

Xiaohu Xu (xuxh@huawei.com)

Himanshu Shah (hshah@ciena.com)

IETF82, TAIWAN

Cloud Data Center Network Requirements

- **Flat Layer 2 networking**
  - VM mobility requires extending the Layer2 domains across multiple PODs.
  - Some cluster services also expect Layer2 connectivity.
- **Scalability**
  - Multi-tenancy capability (Beyond 4K VLANs).
  - MAC table scalability (Millions of VMs within a data center) .
- **Maximize available bandwidth**
  - ECMP forwarding capability.
  - Shortest path forwarding capability.
- **Fast convergence**
  - After network failure and VM move.
- **Simplified provisioning and operation**

Deploy VPLS in Data Center: Good News

- **VPLS could meet most requirements:**
  - Flat Layer 2 networking
  - Scalability
    - Multi-tenancy capability->adequate VPN instances.
    - MAC table scalability->PBB+VPLS
  - Maximize available bandwidth
    - ECMP forwarding capability.
    - Shortest path forwarding capability.
  - Fast convergence
- **In addition, VPLS is a much proven L2VPN technology till now.**

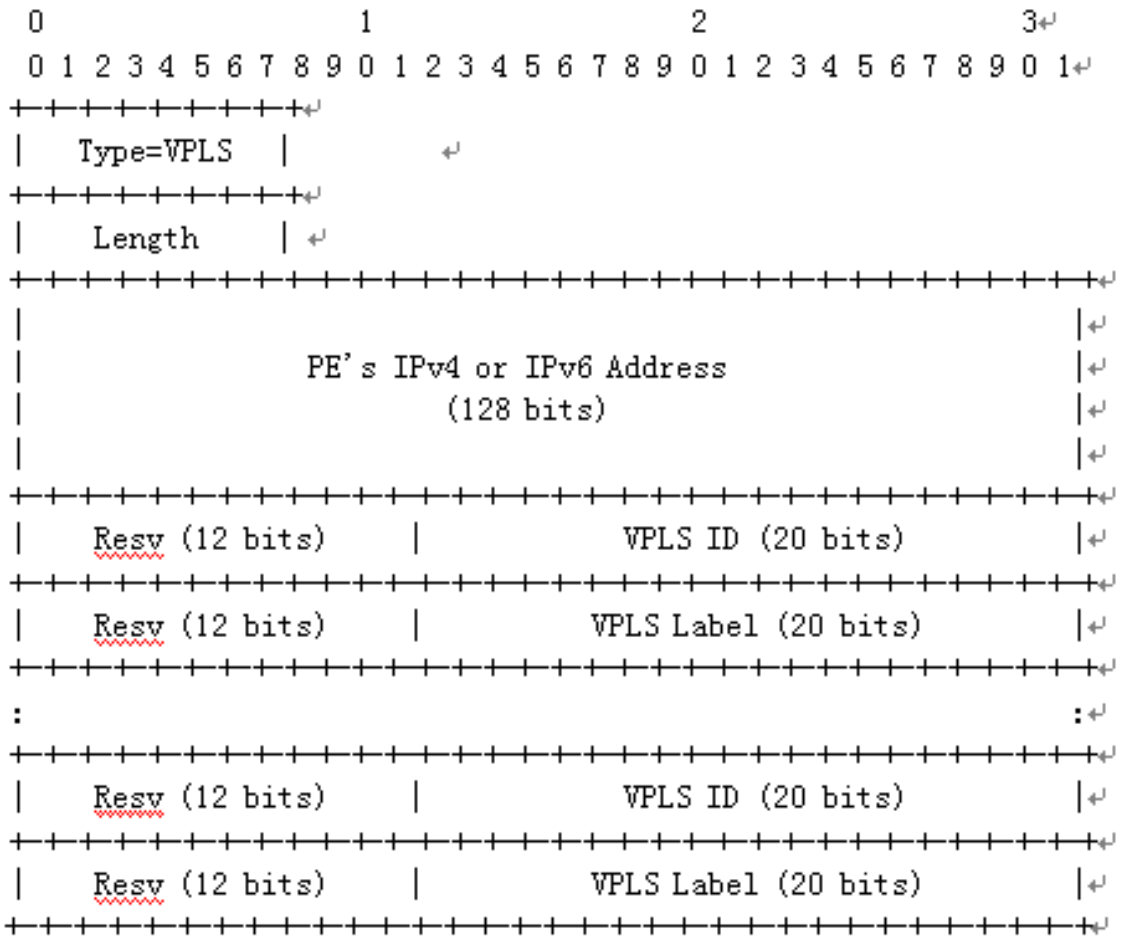
Deploy VPLS in Data Center: Bad News

- **However, VPLS can't meet the requirement of simplified provisioning and operation very well.**
  - Separate protocol(s) for VPLS (LDP and/or BGP)
  - Full-mesh PWs
  - VPLS peer configuration in LDP VPLS (w/o VPLS auto-discovery)
  - BGP peer configuration in BGP VPLS or LDP VPLS (with VPLS auto-discovery)
    - Image deploying PE at hundreds even thousands of ToRs within a single data center.

Why not a Light-weight VPLS

- **Could the already deployed IGP (e.g., IS-IS) be extended a bit so as to deliver a light-weight VPLS which remains the advantages of VPLS while removing the shortcomings of VPLS?**
  - Flat Layer 2 networking ✓
  - Scalability ✓
  - Maximize available bandwidth ✓
  - Fast convergence ✓
  - Simplified provisioning and operation ✓
    - No separate protocol(s) for VPLS
    - No PWs
    - No VPLS peer configuration
    - No BGP peer configuration

IS-IS TLV for VPLS



VPLS Auto-discovery and Signaling

- **Auto-Discovery**
  - Each PE router could automatically discover which other PE routers are part of a given VPLS instance identified by the globally unique VPLS ID.
  - PE router's configuration consists only of the identities of the VPLS instances established on this PE router, not the identities of any other PE routers belonging to that VPLS instance.
- **Signaling**
  - PE router assigns the same MPLS label for a given VPLS instance to any other PE routers.
  - The VPLS label doesn’t need to be globally unique.

Implications on the Control Plane

- **The extended IS-IS TLV for VPLS is partially transparent to P routers.**
  - P routers don't need to process the VPLS membership information contained in that IS-IS TLV, but only need to synchronize the Link State PDUs with their IS-IS neighbors.

Implications on the Data Plane

- **Data encapsulation and data forwarding are not changed.**
- **The only change is to the data-driven MAC learning:**
  - The VPLS label in the received VPLS packet is only intended to identify a given VPLS instance on the egress PE. Hence, the source IP address in the IP-based tunnel header should be resorted to identify the ingress PE of the received VPLS packet.
  - Alternatively, MAC reachability could be distributed among PE routers on the control plane so as to eliminate unknown unicast flood.

How to Deliver Mcast/Bcast/Unknown Unicast

- **Two options:**
  - Ingress Replication
    - No state needed in the core, However, sub-optimal bandwidth utilization.
  - P-Multicast Tree Mode
    - Optimal bandwidth utilization. However, states required in the core.
- **Operators could make the tradeoff flexibly on basis of per tenant instance.**

How to Address the MAC Scalability Issue on PE Routers

- **PBB+VPLS**
  - PBB could be done at ToR switches or even at the servers.
  - VPLS PE routers only need to learn B-MAC addresses.

Comments