INTERNET-DRAFT
Intended Status: Informational Track

Dennis Cai
Sami Boutros
Samer Salam
Reshad Rahman

Expires: December 30, 2012
June 28, 2012

VLAN Aware EVPN services
draft-cai-l2vpn-evpn-vlan-aware-bundling-00.txt

Abstract

This document specifies E-VPN extensions to support the new VLAN aware bundling service interface type defined in [EVPN-REQ]. The new service interface type provides advantages in reducing provisioning overhead as well as E-VPN instances scale in environments where a large number of VLANs need to be extended over an MPLS/IP network, while maintaining traffic segregation among those VLANs. The VLAN aware bundling service interface can handle the high scale requirements of today's Data Centers by bundling different VLANs over a single WAN E-VPN instance used to interconnect those Data Center sites.

Table of Contents

1 Introduction

The high scale requirements of Layer 2 data center interconnect services mandate the signaling of a large number of WAN E-VPN instances. As such, network operators are looking for solutions whereby they can extend multiple Ethernet VLANs over a WAN using a single E-VPN instance, while maintaining traffic segregation among these VLANs in the data-plane. This gives rise to a requirement for a new service interface types: the VLAN aware bundling service interfaces.

These new VLAN aware bundling service interfaces MUST: - Provide the ability to bundle multiple customer VLANs - Guarantee customer VLAN transparency end-to-end.- Maintain data-plane separation between the customer VLANs by creating a dedicated bridge-domain per VLAN.- Support customer VLAN translation to handle the scenario where different VLAN Identifiers (VIDs) are used on different sites to designate the same customer VLAN.

As discussed in [EVPN-REQ], two new service interface types are defined for VLAN aware bundling: with and without translation. The new service interfaces maintain data-plane separation, per VLAN, while sharing one L2VPN E-VPN instance. This document describes the use of different E-VPN routes as defined in [E-VPN] for implementing the VLAN-aware bundling service.

1.1 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

MAC: Media Access Control

MPLS: Multi Protocol Label Switching.

OAM: Operations, Administration and Maintenance.

PE: Provide Edge Node.

CE: Customer Edge device e.g., host or router or switch.

EVI: E-VPN Instance.

2. VLAN-aware-bundling E-VPN

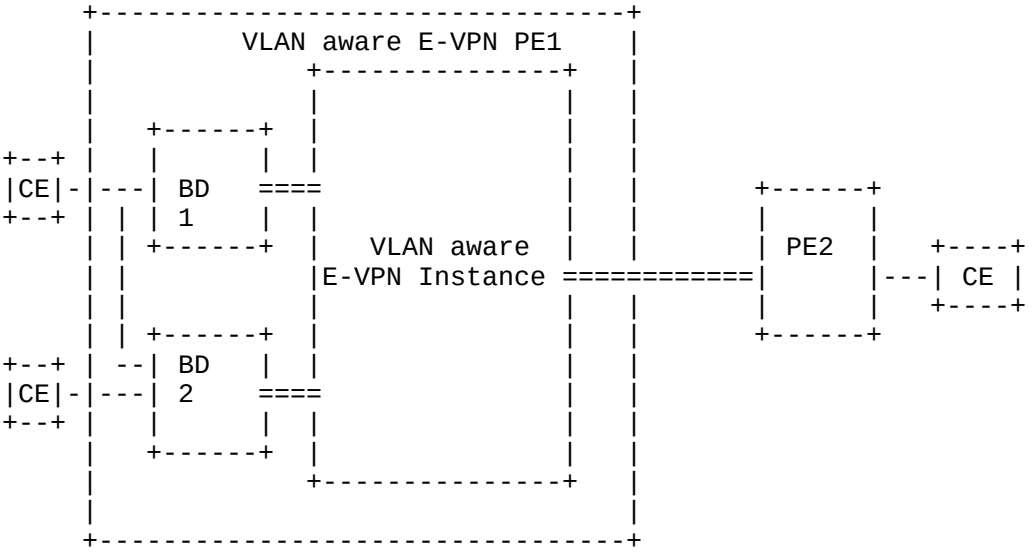
[E-VPN] uses a new BGP NLRI for advertising different route types for

E-VPN operation.

This document discusses how the Ethernet Tag field in the Ethernet Auto-Discovery Route, Mac Advertisement route, and inclusive multicast Ethernet tag route can be used to multiplex several VLANs over the same EVI.

3. Operation

The following figure shows an example of how a VLAN aware Bundling service type over E-VPN could be deployed.



One E-VPN instance has been set up between two sites to extend multiple customer VLANs. On each site, multiple CE devices could be connected to the PE. The link between the CE and the PE could be C-tag or S-tag interface per [802.1Q], carrying several VLANs.Only a single E-VPN instance has been set up to carry customer VLANs between the two sites. The use of two sites in the above figure is for illustration; however, this could be extended to many sites. In order to quantify the benefit of the approach, let’s assume N data center sites, with M customer VLANs. With the new VLAN aware service interface type, the solution would require one E-VPN instance, instead of M E-VPN instances. To maintain data-plane separation among the customer VLANs, each PE will create a bridge-domain per customer VLAN. As well, a customer VLAN on each CE port will represent a unique bridge port in the customer bridge-domain. Only one E-VPN instance would be signaled in the core and will be used to carry multiple customer bridge-domains (or customer VLANs) as long as those

customer VLANs need to be extended to the same set of sites. On the egress PE, the E-VPN label + the VLAN-tag would identify the customer-bridge domain.

3.1 Packet forwarding, MAC learning, aging and flushing

Given the data-plane separation, packet forwarding in the scope of one bridge-domain will remain unchanged. When sending traffic over the E-VPN instance, a qualifying VLAN tag MUST be present on the packet. This VLAN tag has global significance across all sites connected to the E-VPN instance and is used to identify the customer bridge domain in all sites. MAC learning, aging and flushing per bridge-domain will remain un-changed. A mass withdraw for MAC routes learned over the EVPN instance can be done by withdrawing the Ethernet AD route with the tag ID corresponding to the bridge domain.

3.2 Multicast Pruning

Efficient multicast replication in the core can be achieved via the use of the Inclusive Multicast Ethernet Tag Route, to prune the flooding on a per VLAN basis. It is possible to only replicate traffic to PEs that have advertised the Inclusive Multicast Ethernet Tag Route with the Tag value set to the VLAN value. If VLAN value is set to zero, then a single multicast LSM is setup to be used for all VLAN traffic for that E-VPN instance. Multicast snooping protocols such as IGMP and PIM MAY be used to further prune the replication scope for a given multicast group in one customer bridge-domain.

3.3 OAM

Customer Ethernet OAM frames (e.g. CFM [802.1ag]) will be carried transparently over the shared E-VPN instance by the customer’s bridge-domains. Current MPLS OAM mechanisms need to be extended to verify connectivity in the E-VPN instance shared by the customer bridge-domains, service level OAM monitoring should be performed according to [RFC-6136], MPLS OAM extensions is out of scope of the document.

3.4 VLAN translation

As mentioned above, the VLAN tag carried across the E-VPN instance for the new VLAN aware bundling E-VPN instance MUST have network wide significance within the scope of the E-VPN instance. As such, VLAN translation may be performed at each PE attached to the E-VPN instance to translate between the global VLAN tag identifying the customer bridge-domain and the local VLAN tag used by the customer

bridge-domain on this PE.

4 Security Considerations

This document does not introduce any additional security constraints.

4 IANA Considerations

TBD

5 References

5.1 Normative References

[KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC1776] Crocker, S., "The Address is the Message", RFC 1776, April 1 1995.

[TRUTHS] Callon, R., "The Twelve Networking Truths", RFC 1925, April 1 1996.

5.2 Informative References

[EVPN-REQ] A. Sajassi, R. Aggarwal et. al., "Requirements for Ethernet VPN", draft-ietf-l2vpn-evpn-req-00.txt.

[EVPN] A. Sajassi, R. Aggarwal et. al., "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-00.txt.

[RFC-6136] Layer 2 Virtual Private Network (L2VPN) Operations, Administration, and Maintenance (OAM) Requirements and Framework.

6 Appendix Vlan Aware VPLS

It is possible to extend VPLS to support VLAN aware bundling type service, a new PW VLAN Vector TLV to be included the LDP PW FEC label mapping messages for the VPLS service, using the mechanisms specified in RFC 4762, as well as a new LDP capability by which a PE can specify its ability to support this new VLAN aware bundling service interface type. The new PW VLAN Vector TLV would allow multiple VLANs

to share a single VPLS instance, while maintaining data plane segregation among these VLANs. This document defines extension to the PWE3 control protocol [RFC4447] to set up the new VLAN aware bundling type service in MPLS networks. An extension to the MAC Withdrawal mechanisms would allow per VLAN service MAC flushing for this new VLAN aware bundling service.

6.1 VLAN-aware-bundling PW

[RFC4447] uses LDP Label Mapping message [RFC5036] for advertising the FEC-to-PW Label binding. Two types of PW FEC, FEC-128 and FEC-129, can be used for this purpose. Both types of PW FEC contain a PW type Field.

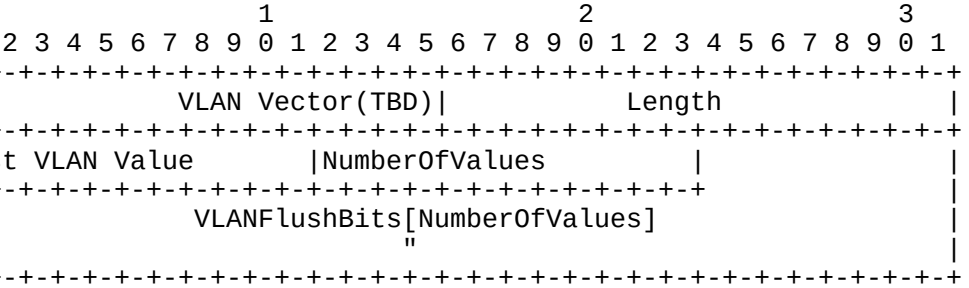
PW type port or raw mode will be used for the VLAN aware bundling interface type service.

Use of control word is optional and frame encapsulation follows the same rules as in [RFC4448].

A new PW VLAN vector TLV is defined, the new PW VLAN Vector TLV will be included in LDP PW label mapping messages, as well it can be included in the MAC flush message.

6.2 PW VLAN Vector TLV

The PW VLAN Vector TLV is described as below:



The U and F bits are set to forward if unknown so that potential intermediate VPLS PEs unaware of the new TLV can just propagate it transparently.

The MAC Flush VLAN Vector TLV type is to be assigned by IANA from the LDP standard [RFC5036] "TLV type name space", as described in section 7.

The TLV value field is of variable length. The first 12 bits encode

the starting VLAN value. The second 12 bits contain the number of values. The VLANFlushBits is an array of bits of length = NumberOfValues, each bit in the array represents a VLAN flush state starting from the 1st VLAN value. A bit value of 1 means flush and a bit value of 0 means don't flush

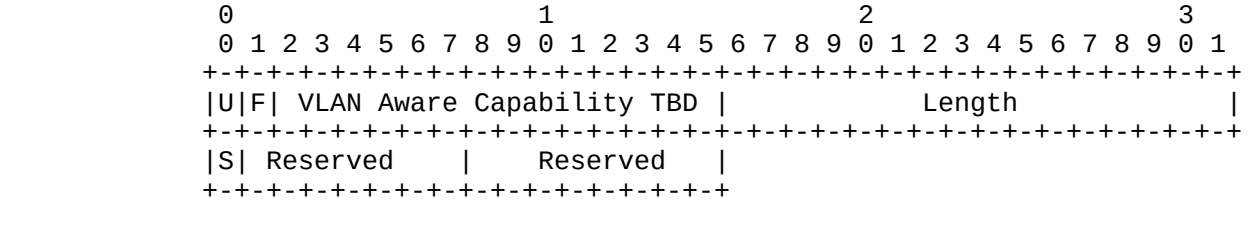
A Starting VLAN value of 0, SHOULD mean include all VLANs, in this case the NumberOfValues SHOULD be 0.

The PW VLAN Vector TLV SHOULD be placed after the PW FEC TLV in the label mapping message as specified in [RFC4447], and SHOULD be placed after the existing TLVs in MAC Flush message as specified in [RFC4762].

6.3 LDP Capability Negotiation

The capability of supporting VLAN Aware Bundling interface type Service MUST be advertised to all LDP peers. This is achieved by using the methods in [RFC5561] and advertising the LDP "VLAN aware Bundling Capability" TLV. If an LDP peer supports the dynamic capability advertisement, it can send a new Capability message with the S bit set for the VLAN Aware Bundling capability TLV. If the peer does not supports dynamic capability advertisement, then the VLAN aware Bundling Capability TLV MUST be included in the LDP Initialization message during the session establishment. An LSR having VLAN Aware Bundling capability MUST recognize the new PW VLAN Vector TLV in LDP label messages.

In line with requirements listed in [RFC5561], the following TLV is defined to indicate the VLAN Aware Bundling capability:



Note: TLV number pending IANA allocation.

* U-bit: SHOULD be 1 (ignore if not understood).

* F-bit: SHOULD be 0 (don't forward if not understood).

* VLAN Aware Bundling Capability TLV Code Point:

  The TLV type, which identifies a specific capability. The VLAN Aware capability code point is requested in the IANA allocation section below.

* S-bit:

The State Bit indicates whether the sender is advertising or withdrawing the VLAN Aware capability. The State bit is used as follows:

1 - The TLV is advertising the capability specified by the TLV Code Point.

0 - The TLV is withdrawing the capability specified by the TLV Code Point.

* Length: MUST be set to 2 (octet).

6.4 Multicast Pruning

Efficient multicast replication in the core can be achieved via the use of the new VLAN vector TLV, to prune the flooding on a per VLAN basis. It is possible to only replicate traffic to PEs that have advertised a given VLAN in their Vector TLV. Multicast snooping protocols such as IGMP and PIM MAY be used to further prune the replication scope for a given multicast group in one customer bridge-domain.

Authors' Addresses

Dennis Cai
Cisco Systems

EMail: dcai@cisco.com

Sami Boutros
Cisco Systems

EMail: sboutros@cisco.com

Samer Salam
Cisco Systems

EMail: ssalam@cisco.com

Reshad Rahman
Cisco Systems

EMail: rrahman@cisco.com