draft-raggarwa-mac-vpn-01.txt
draft-sajassi-l2vpn-rvpls-bgp-01.txt

R. Aggarwal (Juniper), A. Sajassi (Cisco)

W. Hendericx (Alcatel-Lucent), A. Isaac (Bloomberg), J. Uttaro (AT&T), N. Bitar(Verizon), F. Balus(Alcatel-Lucent), K. Patel(Cisco), C. Filsfils(Cisco), R. Shekhar (Juniper)

July 26th 2010

IETF Maastricht

Agenda

- Requirements
- Overview of the solution
- Status & Next Steps

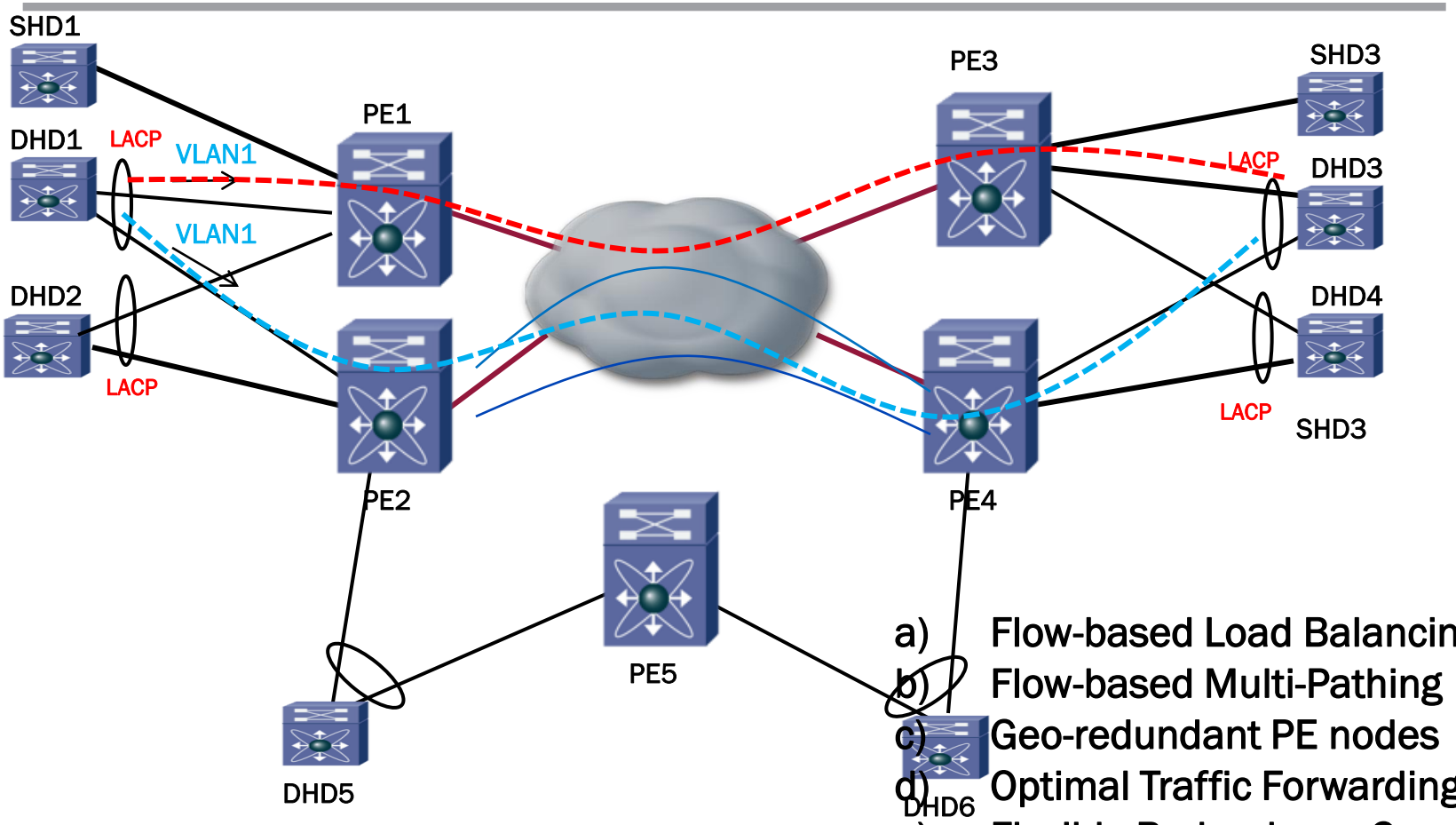
Requirements

1. All-Active Multi-Homing
   a) Flow-based Load Balancing
   b) Flow-based Multi-Pathing
   c) Geo-redundant PE nodes
   d) Optimal Traffic Forwarding
   e) Flexible Redundancy Grouping Support
2. Multi-homed Network
3. Multicast Optimization with MP2MP LSP (in addition to P2MP LSP)

Requirements – Cont.

4. Ease of Provisioning
5. New Service Interface
6. Fast Convergence
7. Flood Suppression
8. Finer control over MAC address learning
    - Control which nodes learn which MAC
    - Support of hub-and-spoke and extranet topologies

Req 1: All-Active Multi-Homing



a) Flow-based Load Balancing
b) Flow-based Multi-Pathing
c) Geo-redundant PE nodes
d) Optimal Traffic Forwarding
e) Flexible Redundancy Grouping Support

Req 4: Ease of Provisioning

- For deployments where VLAN IDs are global across the MPLS network, the MPLS attribute such as (VPN ID, VPN RT, etc.) must be derived automatically
- Where possible, Site-ID must be derived automatically
- Where possible multi-home auto-discovery shall be performed automatically

Req 5: New Service Interface

- Existing VPLS Services (with additional enhancements)
  - Port mode (unqualified learning)
  - VLAN mode (qualified learning)
- As well as the new New Service
  - VLAN-aware port mode
    - All-to-one bundling of customer VLANs
    - VLAN transparency
    - maintain data-plane separation among the VLANs
    - No VLAN configuration shall be required on the port

Req 6: Fast Convergence

- Fast convergence for a multi-homed device requires convergence time for core-to-site traffic be independent of
  - # of MACs addresses being affected by the site AC failure
  - # of service instances (VLANs) being affected by the site AC failure
- Fast convergence for both AC failure as well as MES failure

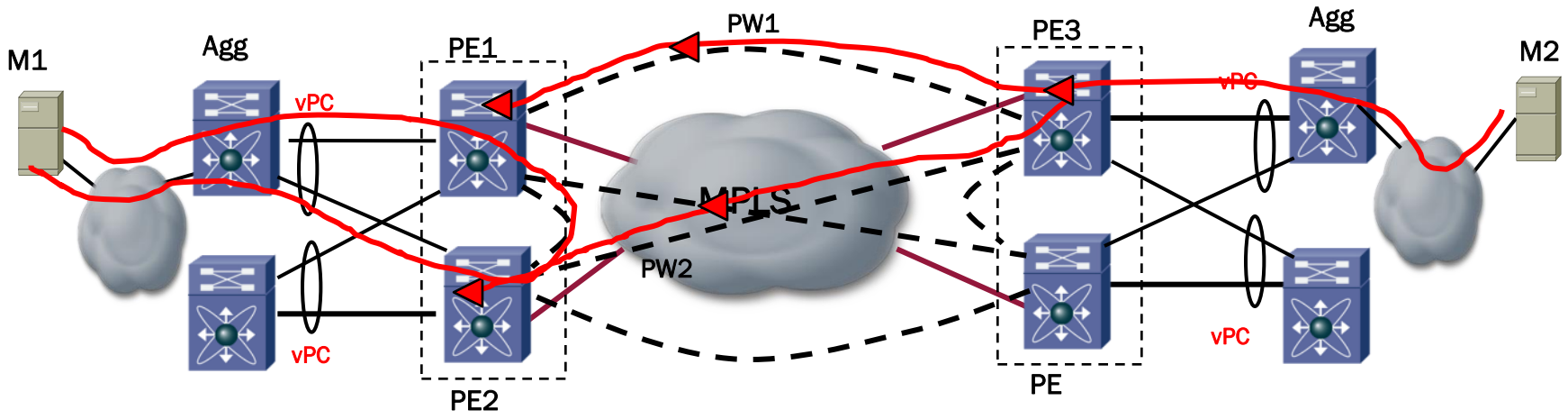
Req 7: Flood Suppression

- To be able to turn off flooding of unknown unicast frames on a per service instance basis
- To avoid unnecessary flooding of unicast frames upon topology change where PEs have prior knowledge of the backup paths for a given MAC

VPLS Issues

Current VPLS Issues in meeting the above requirements

1. Forwarding Loops
2. Duplicate Frame Delivery
3. MAC Forwarding Table Instability
4. Identifying Source PE in MP2MP LSP



Agenda

- Requirements
- Overview of the solution
- Status & Next Steps

Solution: Terminology

- Referred to as “BGP/MPLS MAC VPNs” in draft-raggarwa-mac-vpn and as “Routed VPLS” in draft-sajassi-l2vpn-rvpls-bgp
- The authors are currently discussing terminology
- For this presentation we will use the term “Ethernet VPN” or “E-VPN” for short

E-VPN (1)

- BGP/MPLS based technology for meeting the requirements described earlier
- Reuses several building blocks from existing BGP/MPLS based technologies
- Requires extensions to existing BGP/MPLS based technologies…

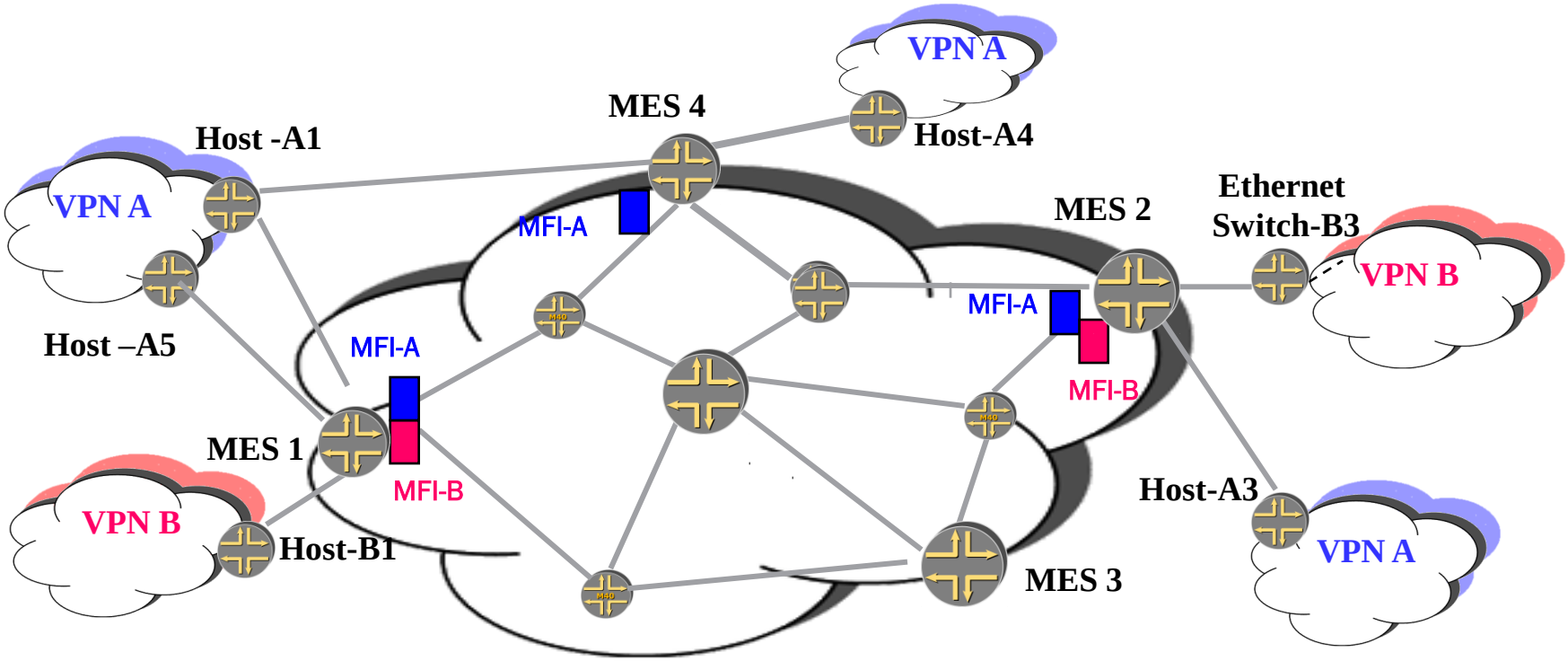
E-VPN (2)

- Introduces the ability to advertise MAC addresses in the control plane in BGP using principles borrowed from IP VPNs
- Advertisement of MAC addresses in BGP provides the building block for the following (not an exhaustive list)
  - Ability to control who learns what along with flexible topologies
  - All-active points of attachment
  - Host mobility
  - Scale
  - Fast service restoration from edge failures leveraging BGP mechanisms

E-VPN: Network Elements

- An E-VPN comprises CEs that are connected to MPLS Edge Switches (MES) aka PE devices
  - CEs may be hosts or switches
  - MESes comprise the edge of the provider network or intra-data center MPLS infrastructure
- The MESes provide layer 2 virtual bridge connectivity between the CEs
- Hosts may be directly connected to MESes
- Hosts may be indirectly connected over a bridge network to MESes
- The MESes are connected by a MPLS LSP infrastructure
  - P2P or MP2P MPLS LSPs and optionally P2MP or MP2MP MPLS LSPs for unicast, multicast and broadcast
  - Provides the benefits of MPLS such as fast-reroute, resiliency etc.

Ethernet VPN: Network Elements



E-VPN: Policy Attributes

- Route Targets (RT) to define the membership of a MAC VPN
  - The MESes, ethernet interfaces/VLANs connecting CEs to MESes
- RTs can be auto-derived from a VLAN ID
  - Particularly applicable if there is a one to one mapping between a MAC VPN and a VLAN
  - Removes the need to configure RTs

Ethernet VPN: MAC Address Learning

- MESes must be able to learn how to reach a given destination unicast MAC address
  - MESes forward packets that they receive based on the destination MAC address
- Local learning
  - A particular MES must be able to learn the MAC addresses of the hosts that are connected directly or indirectly to it
- Remote learning
  - A particular MES must be able to learn the MAC addresses of hosts that are behind other MESes

Ethernet VPN
Local MAC Address Learning

- A MES must support local data plane learning using vanilla ethernet learning procedures
  - When a host generates a data plane packet such as an ARP request
- MESes may learn the MAC addresses of hosts in the control plane using control plane protocols that run between the MES and the CEs
  - Extensions to existing layer 2 protocols
  - Alternate protocols

Ethernet VPN
Remote MAC Address Learning

- Ethernet VPN requires an MES to learn the MAC addresses of hosts connected to other MESes in the *control plane*
  - *Remote MAC addresses are not learned in the data plane*
- Ethernet VPN introduces the ability for an MES to advertise the MAC addresses of the hosts that are connected to it using BGP to other MESes

Remote MAC Address Learning in the Control Plane

- Architectural building block to enable load balancing
  - Allows CEs to connect to multiple active points of attachment
- Improves convergence in the event of certain network failures
  - As convergence does not rely on data plane re-learning
- Allow hosts to relocate within the same subnet without requiring renumbering
- Minimizes flooding of unknown unicast packets
  - Particularly if local MAC address learning can be performed in the control plane
- Control over which MAC addresses are learned by which devices enabling flexible topologies

Other Building Blocks of Ethernet VPN

- Auto-discovery by each MES, in a given Ethernet VPN, of the Ethernet Segment membership of all the other MESes in the MAC VPN
- Ethernet Segment Identifier
  - Used for multi-homing.E.g., derived from the LAG identifer
- Designated forwarder election
  - To ensure that multicast, broadcast and unknown unicast packets are sent to a multi-homed CE by a single MES
- “Split horizon”
  - To ensure that a multicast, broadcast or unknown unicast packet that is sent by Host-A1 to MES1, and then sent by MES1 to all other MESes in the Ethernet VPN, isn’t sent back by MES4 to Host-A1
    - Host A-1 is dual homed to MES1 and MES4
- Auto-discovery of Inclusive Trees for multicast, broadcast and unknown unicast traffic

Ethenet VPN: Multicast

- Leverages significant development and strides made in MPLS multicast which is fairly widely deployed today
- Inclusive trees
  - Traffic for a multicast flow is sent to all MESes in the Ethernet VPN
- Selective trees
  - Traffic for a multicast flow is sent to a subset of the MESes in the Ethernet VPN to minimize sending the traffic to those MESes which do not have receivers in the flow
- Tunneling technologies
  - Ingress replication using RSVP-TE P2P or LDP MP2P LSPs
  - P2MP RSVP-TE or P2MP LDP LSPs
  - MP2MP LDP LSPs

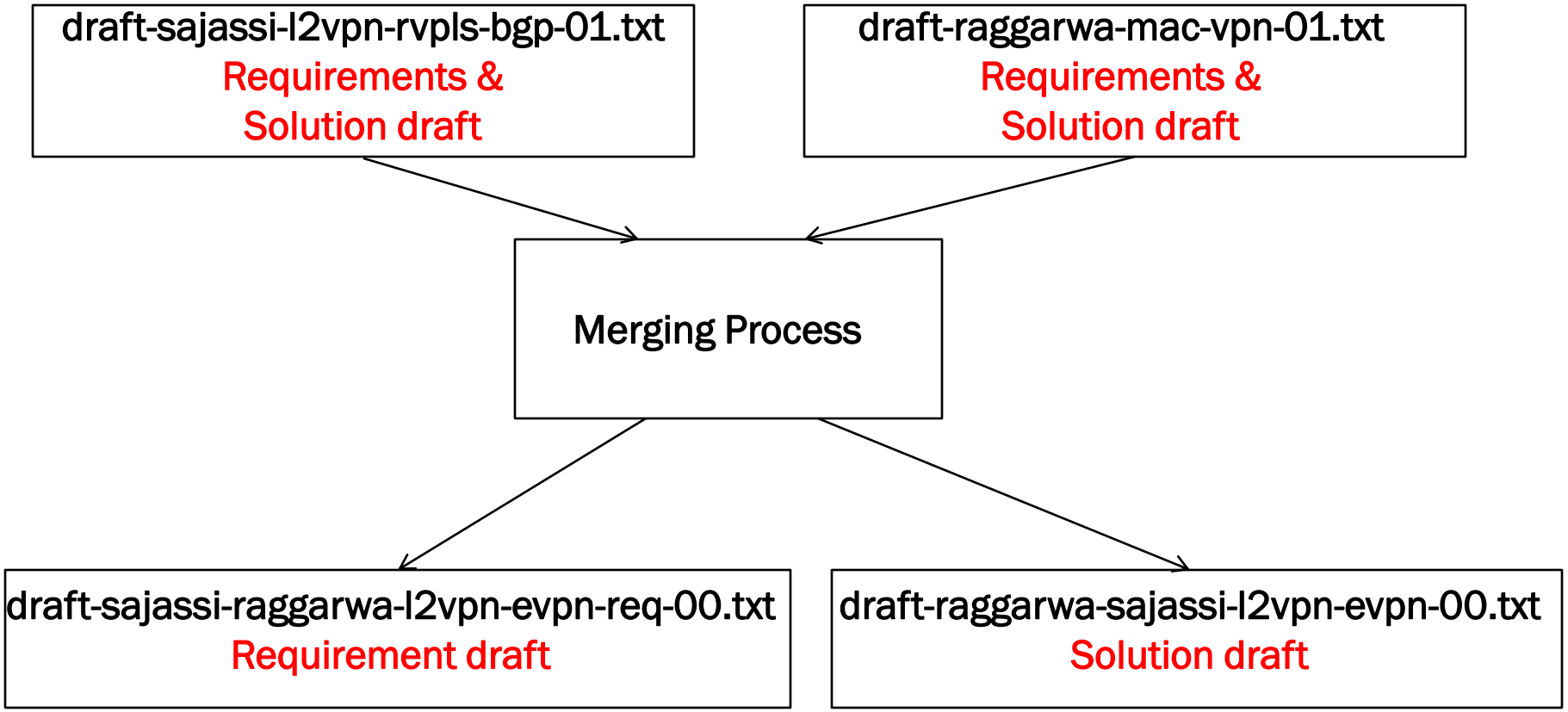
Applicability of “Ethernet VPNs” VPLS as a Service

- The term VPLS today means two things
  - Virtual Private LAN ***Service***
  - BGP-VPLS/LDP-VPLS as a technology to deliver Virtual Private LAN service
- Ethernet VPNs may be used as technology to deliver VPLS as a service when one or more of the “enhanced features” described earlier are required
  - As a result the L2VPN WG is a suitable home for Ethernet VPNs

Agenda

- Requirements
- Overview of the solution
- Status & Next Steps

Status


draft-sajassi-l2vpn-rvpls-bgp-01.txt
Requirements &
Solution draft
draft-raggarwa-mac-vpn-01.txt
Requirements &
Solution draft
Merging Process
draft-sajassi-raggarwa-l2vpn-evpn-req-00.txt
Requirement draft
draft-raggarwa-sajassi-l2vpn-evpn-00.txt
Solution draft

Requirement Draft:
draft-sajassi-raggarwa-l2vpn-evpn-req-00.txt

- Combines the requirements sections of the drafts below
  - draft-sajassi-l2vpn-rvpls-bgp-01.txt &
  - draft-raggarwa-mac-vpn-01.txt
- These requirements were listed at the start of this presentation
- Captures & describes the VPLS issues as listed at the start of this presentation

Solution Draft: draft-raggarwa-sajassi-l2vpn-evpn-00.txt

- Many commonalities in terms of principles of operation
- Converging on a single set of BGP encoding
- Merge of these drafts are being done along the following lines:
  - Non-overlapping procedures (such local versus remote DF election) are carried over from each draft
  - Overlapping functions that are essentially the same (e.g., MAC distribution) are being combined into one
  - Overlapping functions that are different (such as Tag AD versus Site AD) are being currently discussed

BGP Routes

MAC-VPN Routes	R-VPLS Routes
Eth Tag Route (Ethernet Tag A-D Route)	-
MAC Route (MAC Advertisement Route)	MAC Route
Inclusive Multicast Ethernet Tag Route	-
Ethernet Segment Route	RG Route
-	MH-ID Route
Selective Multicast A-D Route	VPLS S-PMSI A-D Route
Leaf A-D Route	VPLS Leaf A-D Route

Alignments

- Working on the alignment of the following remaining functions:
  - DF Election
  - Per-flow load balancing toward the core
  - Optimized MAC withdraw
  - Auto-derived RTs
  - MP2MP
  - Egress PE forwarding w/o MAC lookup

Next Steps

- Lots of interest in this work by L2VPN participants as can be seen by the list of co-authors and contributors
- It seems like the work falls within L2VPN charter as the charter includes multi-homing (although it doesn't call out all-active multi-homing explicitly)
- Therefore, we believe L2VPN WG is the appropriate WG to work on this technology
- A requirement draft will be issued soon and would like to request for a WG call (this is the 2nd IETF meeting that we have discussed the requirements)
- A unified solution draft with the alignments of the above areas will be targeted for next IETF
- Comments ?