VPLS Multicast Challenges and Options

Marc Lasserre

Riverstone Networks

MPLS Japan 2005

Problem Statement

Multicast Optimizations

- Traffic Delivery
  - To receivers only
    - Tracking Joins/Prunes
- Bandwidth Usage
  - Minimize number of copies

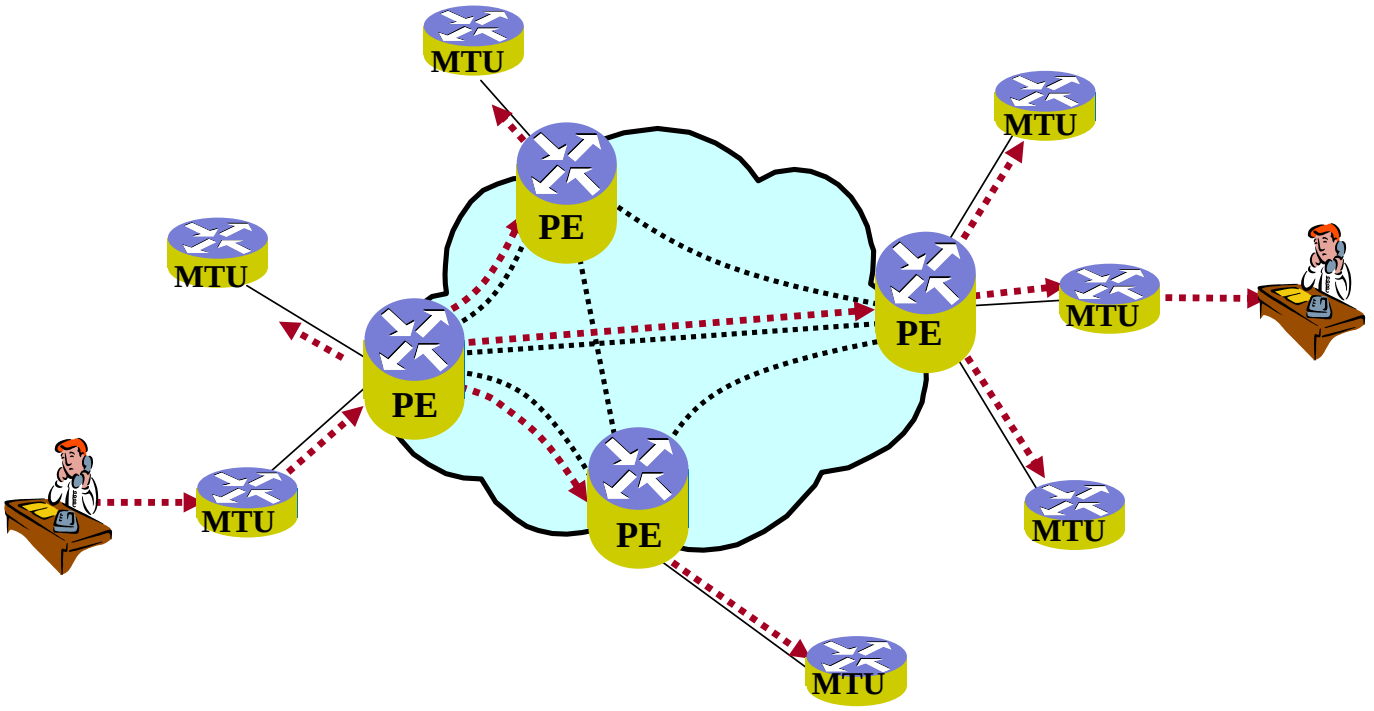
Traffic Delivery Optimization

- IGMP/PIM snooping
  - Snooping on ACs not an issue
    - Amount of (S,G) state to be maintained is bounded
  - Can be performed at different points of the ACs
    - E.g. PE, MTU, DSLAM, CPE
  - Snooping on core PWs can lead to a large amount of state to be maintained per PE
    - Processor load for snooping
    - Memory to maintain (S,G)
  - IGMP Snooping
    - Defined in draft-ietf-magma-snoop
  - PIM Snooping
    - Defined in draft-hemige-serbest-l2vpn-vpls-pim-snooping

Multicast Options

- Trade-off between:
  - State maintained in the core
  - Optimization of bandwidth usage
  - Optimality of multicast routes
- L3 Multicast solutions aim to optimize b/w usage
  - draft-rosen-vpn-mcast
- L2 Multicast solutions aim to keep the core stateless
  - draft-hemige-serbest-l2vpn-vpls-pim-snooping
  - draft-ietf-magma-snoop
- Hybrid model
  - draft-raggarwa-l2vpn-vpls-mcast

VPLS Multicast



Replication Limitations

- Ingress replication as defined in current VPLS drafts can lead to inefficient bandwidth usage in VPLS core
  - Multiple copies of the same multicast frame sent
    - Over shared physical paths
- Number of replications performed in h/w
  - Introduces latency & jitter
  - Limits the maximum number of VPN members that can efficiently receive multicast

Bandwidth Optimization

- So far, replication within metro networks has not been an issue
  - Simple topologies (rings or very few P routers between PEs)
  - Average number of sites per VPN typically small (between 5 and 20)
  - Hierarchical VPLS constructs distribute replication across multiple nodes
- With more complex topologies, multicast applications, use of p2mp LSPs leads to better bandwidth utilization

VPLS Multicast Drivers

- Broadcast video/radio delivery
  - Carrier based services
  - E.g. regular TV channels
- Dedicated multicast streams
  - Business based services
  - Customer video feeds
    - E.g. Bank video advertisements in branch offices
  - Financial information
    - E.g Reuters, TIBCO
  - Video conferencing
    - E.g. NetMeeting

VPLS Multicast Enhancements

Broadcast Trees

Shared Broadcast Tree

- One broadcast tree across VPLS instances
  - Rooted at each VPLS PE
- Used to carry all customers' bcast & mcast traffic
- Applicable to both L2 bcast/mcast and L3 mcast
- Minimizes amount of multicast state in the core
- VPLS/VC agnostic
- Requires support of *draft-ietf-mpls-rsvp-te-p2mp*

Ideally suited for residential Broadcast Video/Radio delivery

Dedicated Broadcast Tree

- One broadcast tree per VPLS instance
  - Rooted at each VSI
- Used to carry one customer's bcast & mcast traffic
- Applicable to both L2 bcast/mcast and L3 mcast
- Minimizes amount of multicast state in the core
- VPLS/VC agnostic
- Requires support of *draft-ietf-mpls-rsvp-te-p2mp*

Ideally suited for Business Broadcast Video delivery

Multicast Trees

Dedicated Multicast Trees

- Several trees per VPLS instance
  - Rooted at each source
- Used to carry efficiently customer's specific mcast traffic
- Applicable to IP mcast only
- Requires support of *draft-ietf-mpls-rsvp-te-p2mp*
- Requires a discovery procedure of multicast membership in core
  - To map (S,G) to correct multicast tree

Suited to business customers with multiple multicast streams with high b/w requirements

Aggregate Multicast Trees

- Use of p2mp trees to a defined set of PEs across VPLS instances
- Requires label coordination (upstream allocation)
  - Per mcast group VC label for demultiplexing
- Only applicable to IP traffic
- Requires a discovery procedure of multicast membership in core
  - PIM/IGMP snooping
  - "signaling" protocol to advertize membership (LDP or BGP)

Suited to business customers that need to exchange multicast streams

Discovery Protocol
for Multicast Trees

- Several options available
  - Use of IGMP/PIM snooping on core PWs
  - Use of LDP extensions to carry mcast membership information
    - *draft-qiu-serbest-l2vpn-vpls-mcast-ldp*
  - Use of BGP or PIM as defined in
    - *draft-raggarwa-l2vpn-vpls-mcast*

VPLS Dataplane Changes

- Use of default p2mp tree instead of ingress replication for:
  - All customer broadcast and multicast *data* traffic
- Non-Use of default p2mp tree for
  - All customer broadcast & mcast *control* traffic
- Use of multicast FECs to map customer mcast traffic to appropriate mcast data trees (Options 3 & 4 only)
- Encoding of mcast VC label for aggregate trees
- PHP turned off except for aggregate trees

VPLS Control Plane Changes

- IGMP/PIM snooping on ACs
- RSVP-TE mcast extensions
- IGMP/PIM snooping on PWEs for multicast trees
  - Optional if use of “signaling” protocol
- PIM refresh reduction turned off for multicast trees
- PIM support in Ps for multicast trees
- BGP/PIM/LDP mcast state signaling over PWEs for aggregated trees

Issues to be resolved

- Additional complexity for multicast trees
- Aggregate trees have add'l requirements:
  - Protocols for discovery and upstream allocation
    - Current PWE/VPLS use downstream unsolicited label advertisement
  - Complex heuristics and/or knobs to decide when to set up new aggregate trees
- Which discovery protocol is best suited?
  - Snooping, PIM, BGP, LDP?
  - Subject for future discussion

Q & A