Multicast in RFC2547 VPNs

- *Ordinary* multicast over IP backbone requires core routers to:
  - maintain one or more distribution trees for each multicast group
  - route to root of each dist. tree
- Problems in VPN scenario:
  - State in core for each enterprise dist. tree undesirable
  - Routes to enterprise roots not known in backbone
- What to do, what to do?
  - Outline three alternatives
  - Select one

Common Features

- PIM assumed
- Multicast VRFs
- No global uniqueness for source addresses or group address

Alternative 1: Extended *Ordinary* Multicast

- When PE gets PIM Join from CE, augment as follows:
  - lookup root in VRF
    - get RD from root's VPN-IP address and insert in Join msg
    - get BGP next hop (egress PE) of root aned insert in Join msg
- Send augmented Join through core:
  - set up PIM route in ordinary way for enterprise dist tree
  - at egress PE, RD selects VRF
  - optimal multicast routing in core
- Problems: still state per enterprise dist tree, not bounded

Alternative 2: Ingress PE replicates

- PEs attached to VPN become PIM adjacencies
- Use point-to-point (tunneled) or NBMA PIM techniques
- Ingress PE replicates packets
- Advantages:
  - No state at all in core
- Disadvantages:
  - Premature replication and its consequences (non-optimal routing)
    - But still better than with L2 VPNs, because no replication by CE

Alternative 3: Core has one Dist Tree per VPN

- Multicast Tunnel: Use PIM to set up one dist tree per VPN, containing all PEs on that VPN
- All PEs in a VPN become PIM adjacencies over this Tunnel
- Use PIM LAN procedures: (Tunnel as LAN interface)
- Advantages:
  - one state in core per VPN (bounded),
  - enterprise multicasts aggregated
- Disadvantages:
  - non-optimal multicast routing in core

State vs. Optimality

- Overall theme: if we want "stateless in core" combined with "optimal in core", we have no solution
- In alternative 3, can increase optimality by increasing state:
  - multiple tunnels per VPN
  - not all have full membership
  - high throughput groups with low membership can be aggregated separately
- Alternative 3 seems best trade-off for now.