Routing in L1VPN Networks

MPLS Japan 2006
Tokyo, October 2006

Dimitri Papadimitriou
NSG/CTO, Alcatel

Outline

Motivation

Service Characterization

Service Models

Discovery / Routing

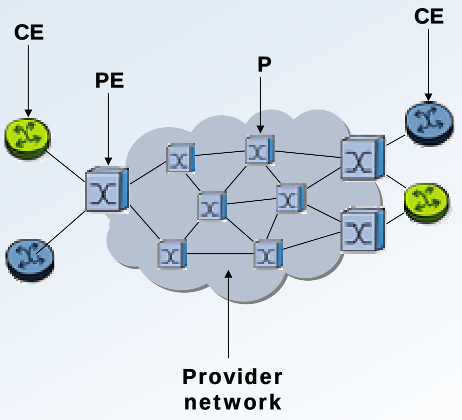
Analysis

Conclusion

Motivations for Layer 1 VPN

- Key Business Drivers for customers
  - Outsourcing
    - –Third party control: outsource direct management of layer 1 network
    - –Off-loading CE-to-CE configuration: no need to configure and manage full connectivity between CEs
  - Cost
    - –Sites can be inter-connected without bearing full cost of deploying and managing the layer 1 network
- Key Business Drivers for network providers
  - Network Resource saving
    - –Sharing layer 1 network infrastructure with many customers
    - –More flexible usage of spare resource (higher sharing ratio)

The BIG Picture



Customer Edge (**CE**) device

- receives L1VPN service from the provider network
- connected to at least one PE device (terminates L1 signal)
- examples: L2 switch, IP/MPS LSR, IP router

Provider Edge (**PE**) device

- provides L1VPN service to the customer
- connected to at least one CE device
- examples: TDM switch, or OXC

Provider device (**P**)

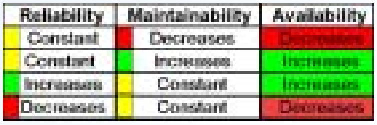
- connected only to other provider devices (P or PE devices)
- Example: TDM switch, or OXC

Membership information

- list of CE-PE TE link belonging to the same VPN

Layer 1 VPN Service Characterization

- Connectivity
  - Source-Destination
  - Capacity
  - Other traffic parameters
- Availability
  - Function f(reliability, maintanability)
  - Availability = uptime (MTTF) / [uptime (MTTF) + downtime (MTTR)]
    - –Reliability = probability P_r that system or component fails within a given period of time (MTTF ~ $1/P_r$)
    - –Maintanability = probability P_m that system or component will be retained in or restored to a specified condition within a given period of time (MTTR ~ $1/P_m$)



Reliability	Maintainability	Availability
🟨 Constant	🟥 Decreases	🟥 Decreases
🟨 Constant	🟩 Increases	🟩 Increases
🟩 Increases	🟨 Constant	🟩 Increases
🟥 Decreases	🟨 Constant	🟥 Decreases

Layer 1 VPN Service Characterization

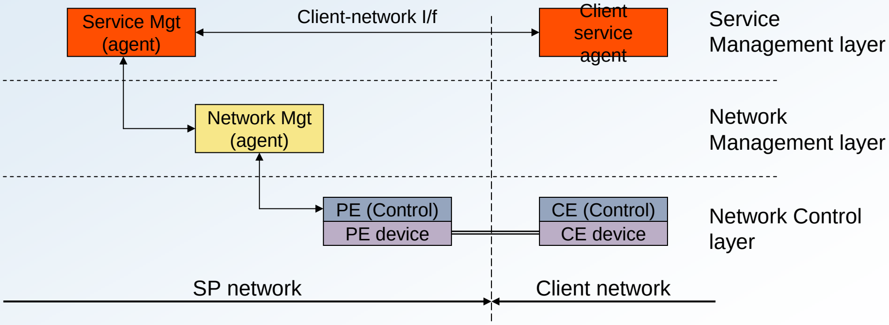
- Scalability
  - Number of sites per VPN / per PE
  - Number of PEs per VPN
  - Number of VPN per SP network (linearly)
  - Maintain VPN specific information on PE (CE)
- Robustness
  - For a given set of features, with a given set of perturbations, the system continues to operate correctly (as expected)
  - Dependence
    - – Configuration/maintenance operation between VPNs
    - – Isolation in case error/failure
- Efficiency (performance)
- Flexibility/adaptivity (evolutivity, migration)
- Manageability (configuration, accounting, monitoring/measurement)
- Security

Layer 1 VPN Service Models

- **Management based** ⇒ Signaling information within provider network
- **Signaling based** ⇒ Routing information within provider network
  - VPN membership information between PEs
  - Provider network routing information
- **Signaling and Routing based** ⇒ Routing information between CE-PE
  - VPN membership information between PEs
  - Provider network routing information
  - Customer network routing information

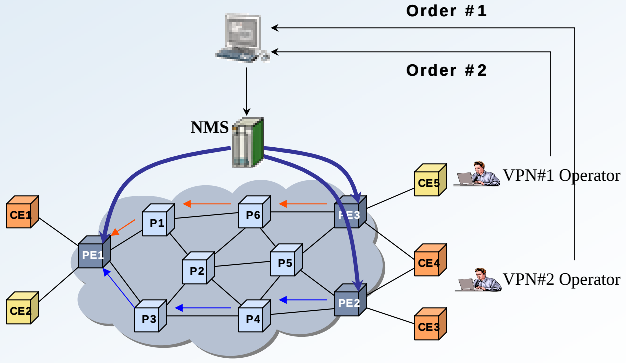
Layer 1 VPN Service Models - Management-based

- Management based ⇒ Signaling information within provider network



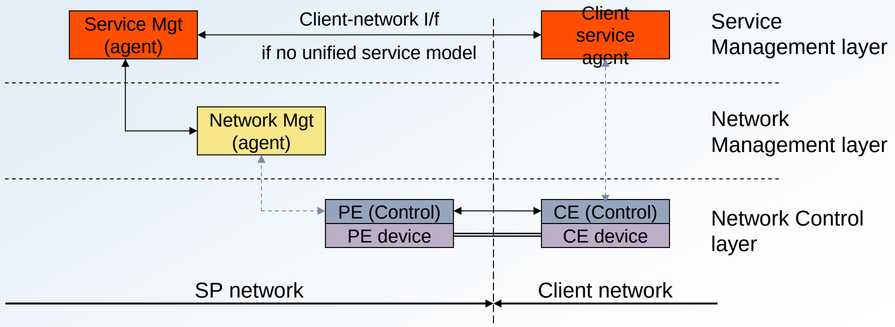
Layer 1 VPN Service Models - Management-based

- Management based ⇒ Signaling information within provider network



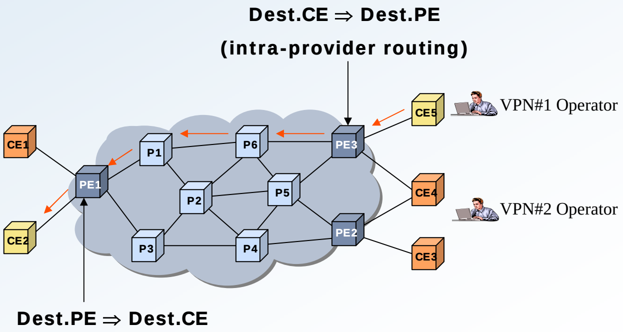
Layer 1 VPN Service Models - Signaling-based

- Signaling based ⇒ Routing information within provider network
  - VPN membership information between PEs
  - Provider network routing information



Layer 1 VPN Service Models - Signaling-based

- Signaling based ⇒ Routing information within provider network
  - VPN membership information between PEs
  - Provider network routing information
- Equivalent to the overlay model operations

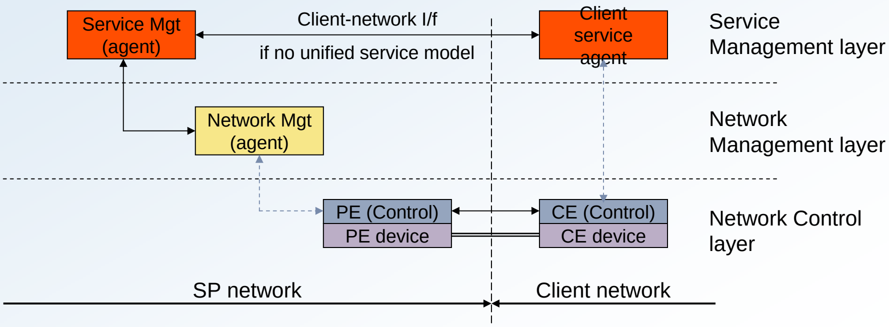

Dest.CE ⇒ Dest.PE

(intra-provider routing)

Dest.PE ⇒ Dest.CE

Layer 1 VPN Service Models - Routing-based

- (Signaling and) Routing based ⇒ Routing information between CE-PE
  - VPN membership information between PEs
  - Provider network routing information
  - Customer network routing information


Service Mgt (agent)

Client-network I/f

if no unified service model

Client service agent

Service Management layer

Network Mgt (agent)

Network Management layer

PE (Control)

PE device

CE (Control)

CE device

Network Control layer

SP network

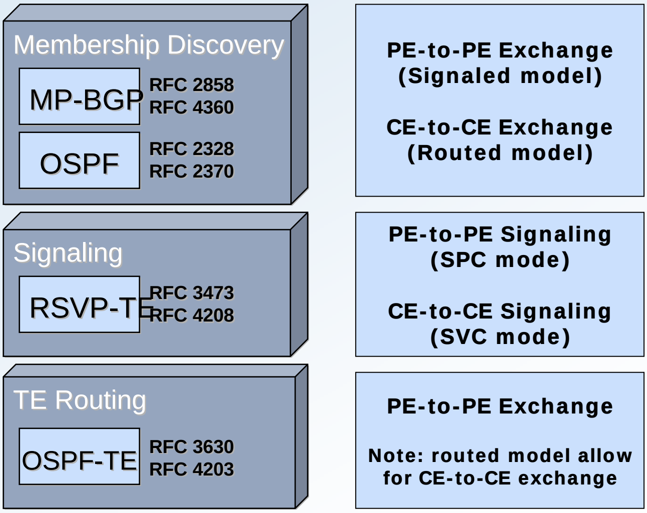
Client network

L1VPN Steps

1. Run discovery mechanism
   - PEs to discover about remote VPN ports and their corresponding provider addressing (PPI)
   - CE to discover (when needed) set of remote CE port addresses (to be used for signaling) - only routed model

2. CE initiates signaling request to attached PE for a given remote CE
   - Using private address space
   - During signaling, at local/remote PE switch private with provider addresses (referred as “shuffling” approach)
   - No need for VPN-ids in signaling between CE-PEs

L1VPN Building Blocks (Control Plane)



Membership Discovery

MP-BGP RFC 2858 RFC 4360

OSPF RFC 2328 RFC 2370

PE-to-PE Exchange (Signaled model)

CE-to-CE Exchange (Routed model)

Signaling

RSVP-TE RFC 3473 RFC 4208

PE-to-PE Signaling (SPC mode)

CE-to-CE Signaling (SVC mode)

TE Routing

OSPF-TE RFC 3630 RFC 4203

PE-to-PE Exchange

Note: routed model allow for CE-to-CE exchange

VPN Membership Information

Port Information Table (PIT) localized on each PE

- Contains list of <CPI,PPI> tuples per VPN

Customer Port Identifier (CPI)

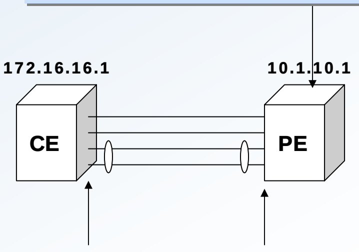
- Numbered link: IPv4/IPv6 address
- Unnumbered link: <port index, CE IPv4/IPv6 address>

Provider Port Identifier (PPI)

- Numbered link: IPv4/IPv6 address
- Unnumbered link: <port index, PE IPv4/IPv6 address>

Note: on PE side, PPI maps VPN-PPI (to maintain address space isolation)

Port Information Table (PIT)
Link #1: <172.16.1.1, 10.1.1.1>
Link #2: <172.16.1.2, 10.1.1.2>
Link #3: <3,172.16.16.1; 3,10.1.10.1>



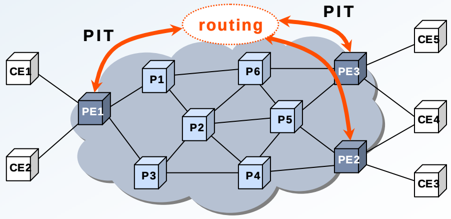
CPI
I/f #1: 172.16.1.1
I/f #2: 172.16.1.2
I/f #3: <3,172.16.16.1>

PPI (provider only)
I/f #1: 10.1.1.1
I/f #2: 10.1.1.2
I/f #3: <3,10.1.10.1>

Discovery

Discovery: Piggybacking of VPN membership info in routing protocol

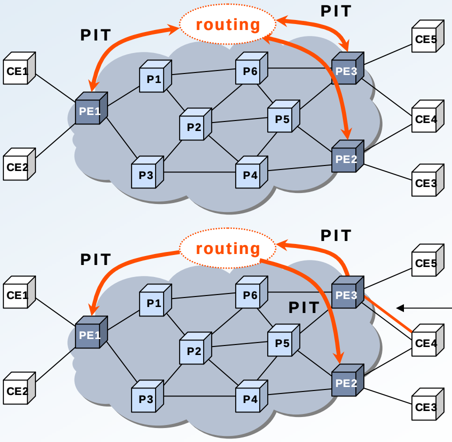
- Each PE advertises (to other PEs)
  - own IP address
  - list of local <CPI, PPI> tuples
  - GUID (Global Unique Identifier) associated to the VPN
- Remote PEs
  - identifies list of common VPN members with advertising PEs
  - perform address resolution during signaling phase



Discovery

Single-end provisioning

- Add new [CE,PE] link to VPN A ⇒ localized configuration action only



Add link #2 [CE4,PE3]

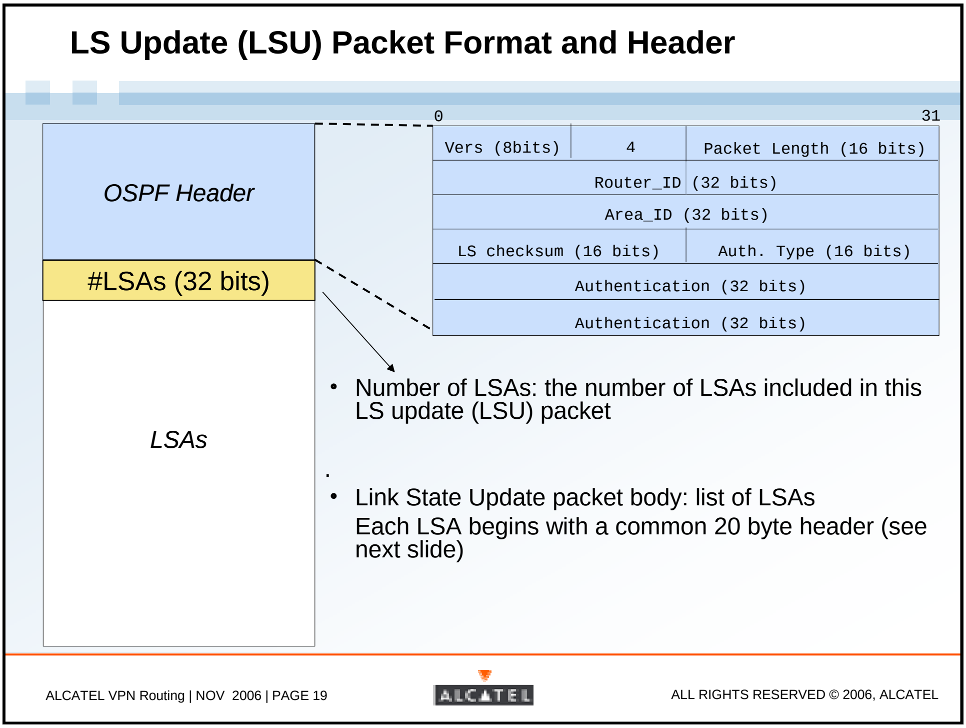
OSPF as VPN Membership Discovery Protocol

Rationales

- Widely used for TE routing (GMPLS protocol suite)
- Facilitate [CE,PE] TE link information exchange
  - Maintain single TE routing protocol
  - Well-defined TE link attributes
- GMPLS L1 networks have usually
  - reasonable dimensions (#PEs, #CEs, etc.)
  - smaller number of VPNs than L3 case (O(10k)) ~ O(10) to O(100)

Issues

- Single AS (single or multi-area)
- AS-wide scope opaque LSAs require extended processing from RFC 2370
- Scalability
  - P participates in the flooding of L1VPN opaque LSAs and maintains LSAs in LSDB

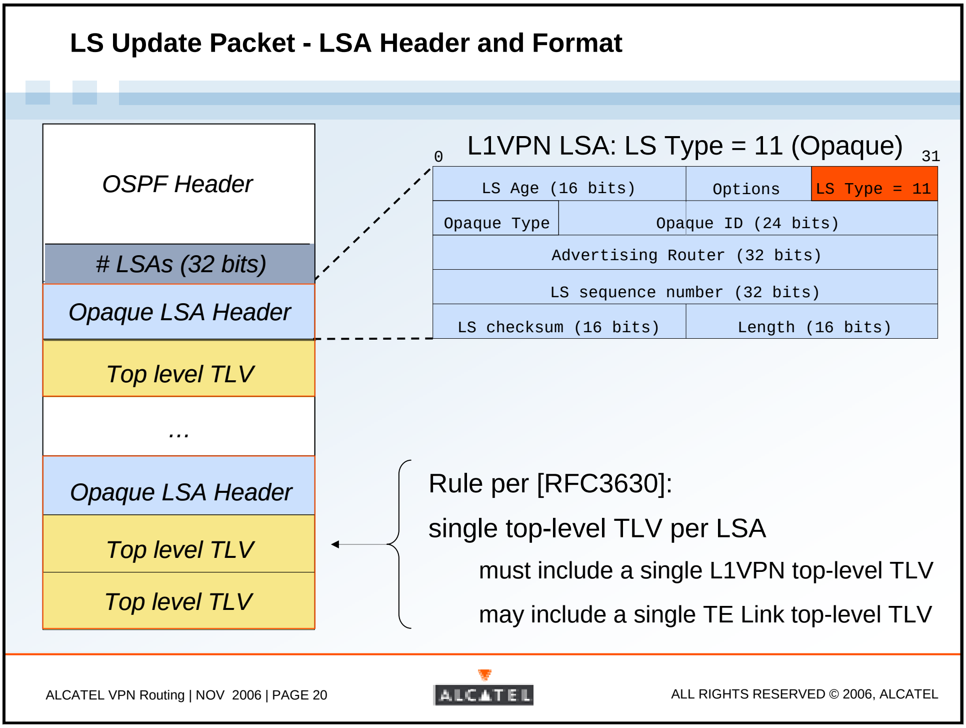


LS Update (LSU) Packet Format and Header

Vers (8bits)	4	Packet Length (16 bits)
Router_ID (32 bits)		
Area_ID (32 bits)		
LS checksum (16 bits)		Auth. Type (16 bits)
Authentication (32 bits)		
Authentication (32 bits)		

- Number of LSAs: the number of LSAs included in this LS update (LSU) packet

- Link State Update packet body: list of LSAs
  Each LSA begins with a common 20 byte header (see next slide)



LS Update Packet - LSA Header and Format

L1VPN LSA: LS Type = 11 (Opaque)

LS Age (16 bits)	Options	LS Type = 11
Opaque Type	Opaque ID (24 bits)	
Advertising Router (32 bits)		
LS sequence number (32 bits)		
LS checksum (16 bits)	Length (16 bits)	

Rule per [RFC3630]:

single top-level TLV per LSA

must include a single L1VPN top-level TLV

may include a single TE Link top-level TLV

L1VPN TLV Structure

L1VPN Globally unique identifier

- <PPI, CPI> tuples association with a particular VPN
- Encode a VPN-ID, a route target, or equivalent

PE IP Address e.g. TE Router or local numbered TE link ID

L1VPN auto-discovery information: set of <CPI,PPI> tuples

L1VPN Information TLV in L1VPN LSA

0	31	
Length (16 bits)	Type (16 bits)	
Globally Unique Identifier		32bits
Globally Unique Identifier		
Reachable IP Address (PE)		
<CPI,PPI> Tuple #1		
...		
<CPI,PPI> Tuple #n		

Theory of Operation

PEs origination and flooding

- local <CPI, PPI> tuples in L1VPN info TLV of LSAs
- Each PE must originate
  - separate L1VPN LSA for each locally configured CE-PE link
  - separate L1VPN LSA for each VPN (no TE link TLV used)
- L1VPN LSA originated
  - PE restart
  - change in PIT entry (*no PIT per CE-PE link but per VPN*)

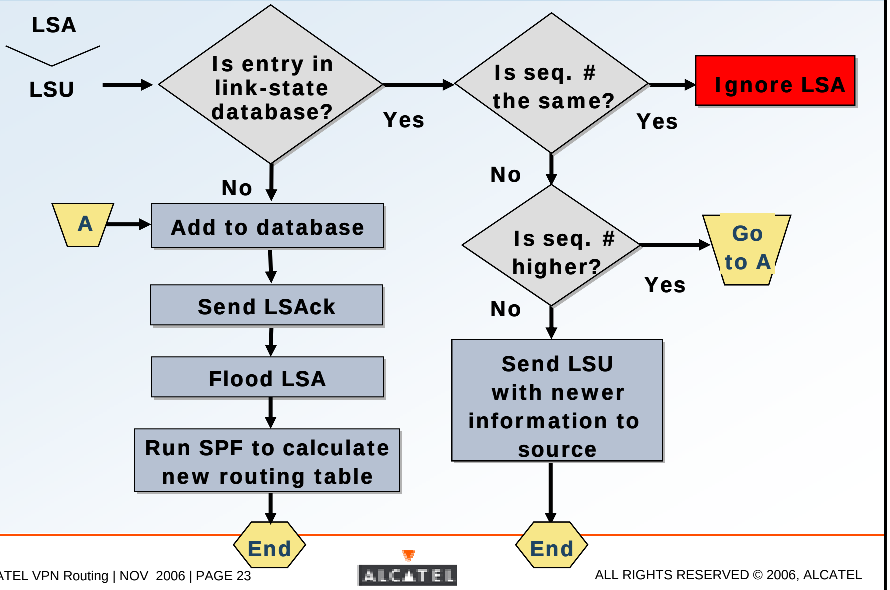
AS wide scope flooding

- Flooding to all PEs within the AS
- Receiving PE to check whether
  - PIT associated with the L1VPN specified in the L1VPN GUID
  - Add/remove or modify the corresponding PIT entry

Policy

- PE local policy with respect to PIT management

LSA Selection Process (simplified)



BGP as VPN Membership Discovery Protocol

Rationales

- Widely used for BGP/MPLS L3VPN routing [RFC4364]
  - Single routing protocol for LxVPN (GVPN) with common mechanisms
- Support any topology (since BGP works across multiple routing domains, it supports L1VPNs that span multiple routing domains)
  - Single AS
  - Multi AS (single or multi-carrier)

Issues

- BGP rarely used in non-packet environments
- (in certain cases) need to augment reachability with CE-PE link information
  - optional non-transitive attribute
- TE information processing

BGP (MP BGP + Route Filtering)

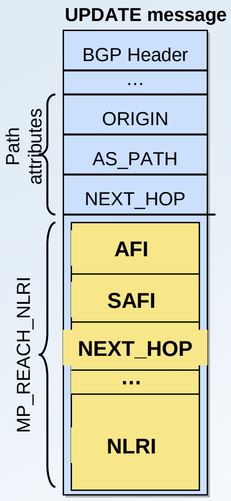
Route propagation: Multi-Protocol extensions to BGP [RFC2858]

- Propagation of local information to other PEs
- At provisioning time, when adding a new CE-PE link between
  - Corresponding PE port associated with a PIT on that PE
  - This PIT is associated (at provisioning time) with its configured VPN

Route filtering: Route Target Extended Community [RFC4360]

- Restrict exchanges to only PITs within a given L1VPN
- Tagging exported local information
  - each PE PIT configured with one or more Route Target Communities, (export Route Targets)
- Filtering imported information
  - each PE PIT configured with one or more Route Target Communities (import Route Targets)
  - set of imported routes into the PIT restricted to only routes that have at least one of these communities

BGP UPDATE message with MP_REACH_NLRI



UPDATE message

BGP Header

…

ORIGIN

AS_PATH

NEXT_HOP

AFI

SAFI

NEXT_HOP

…

NLRI

UPDATE message with MP_REACH_NLRI: announce routes

BGP NEXT_HOP: meaningless in case of an MP_REACH_NLRI and can be left out

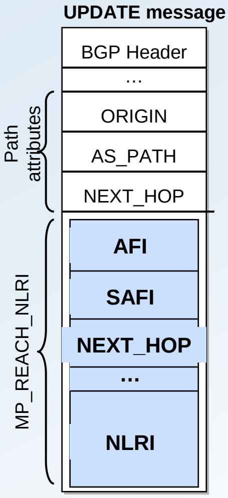
AFI: Specifies the type of network layer protocol associated with the network address of the next hop (encoded in a separate field). Specifies address type of next hop

SAFI: Specifies the type of the NLRI. New L1 VPN SAFI

Next_Hop: address of the next hop for the routes in the NLRI field. The address type is defined by the AFI

NLRI: carrying the CPI and PPI information i.e. one or more <PPI, CPI> tuples

BGP UPDATE message with MP_REACH_NLRI



UPDATE message

BGP Header

…

ORIGIN

AS_PATH

NEXT_HOP

AFI

SAFI

NEXT_HOP

…

NLRI

Path attributes

MP_REACH_NLRI

UPDATE message with MP_REACH_NLRI: announce routes

NLRI: carrying the CPI and PPI information i.e. one or more <PPI, CPI> tuples

Length (1 octet)
PPI AFI (2 octet)
PPI Length (1 octet)
Variable
CPI AFI (2 octet)
CPI Length (1 octet)
Variable

Analysis

Which routing protocol to choose ?

Depends on requirements

- Topology (single vs multi-domain)
- Number of PEs, CEs, VPNs, etc.
- Traffic engineering
  - Coupled to intra-domain routing (single TEDB)
  - De-coupled from intra-domain routing
- Operational
  - Service provisioning (CE-to-CE connectivity)
  - Network provisioning (PE-to-PE connections)

Analysis

Timing: depending on need's importance vs urgency

- Dynamic PTI tables population may be rather static (… static PIT population)
- When dynamic frequency of changes/modification

Performance

- Extend OSPF with “external TE reachability” ⇒ impact on P scaling
- Extend BGP with “TE information” ⇒ impact on PE scaling

Cost

- Single protocol for LxVPN (x = 1, 2, 3)
- Single protocol for L1/TE operations
- Note: are operators looking for integrating their TE operations (including VPN or not) or VPN operations (including TE or not)

Conclusion

VPN membership information discovery in signaled/routed model

- OSPF
- BGP

BGP did not receive lots of attention in the GMPLS context

- Makes the choice of routing protocol less obvious
- Both OSPF and BGP require extensions

Choice of VPN membership information discovery

- Inter-carrier VPN will require BGP (inter-domain routing protocol)
- Single future proof / interoperable solution always better
- … but this future still seems really far away

In Memoriam

of Emmanuel Desmet

deceased suddenly the 17th of October 2006