VPN Toolkit for Service Providers

Yakov Rekhter
Distinguished Engineer
Juniper Networks

On things that do and don't matter...

Layer 2 vs Layer 3 VPN Services

Layer 2:

- Service Provider participates in Layer 2 routing and addressing of VPN customers
  - Within a single Layer 3 segment
- Service Provider does NOT participate in the Layer 3 routing and addressing of VPN customers
- From the Service Provider point of view Layer 2 VPN services are <u>arguably</u> simpler to operate than Layer 3 VPN services
  - As with Layer 3 VPN services the Service Provider has to participate in customers Layer 3 routing
- Etc...

Layer 3:

- Service Provider participates in Layer 3 routing and addressing of each VPN customer
  - Service Provider has to maintain Layer 3 routing information for each VPN customer
- From the VPN customers' point of view Layer 3 VPN services are simpler to use than Layer 2 VPN services
  - As using Layer 3 VPN services reduces the level of IP routing / addressing expertise required by the VPN customers
- Layer 3 VPNs scale better than Layer 2 VPNs
- Etc...

On being pragmatic...

Layer 2 vs Layer 3 VPN comparison based on technology arguments is of little pragmatic relevance:

- **Layer 3 VPN services address different market segment than Layer 2 VPN services:**
  - **Layer 3 VPN services are suitable for the VPN customers who want to run their businesses not their networks**
    - **outsourcing Layer 3 VPN services minimizes the need for IP routing/addressing expertise**
  - **Layer 2 VPN services are suitable for the VPN customers who want (and capable of) full control of their Layer 3 routing**
- **There is a market demand for both Layer 2 and Layer 3 VPN services**
- **For <u>profit-oriented</u> Service Providers profit considerations dominate over technology arguments**

VPN toolkit requirements

- Support both (1) point-to-point Layer 2 VPN, (2) Virtual Private LAN Service (VPLS) and (3) IP VPN services
  - Even if a service provider presently offers only one VPN service (e.g., Layer 2 VPN), the service provider would benefit from the toolkit that can support expanding service offering to other VPN services (e.g., IP VPN or VPLS) with minimum additional effort/cost to the provider
- Have as few tools as possible (but no less than needed to support both point-to-point Layer 2, VPLS, and IP VPN services)
  - A single *operational infrastructure* and a *small set of basic tools* mean cost savings in terms of:
    - Educating the NOC staff
    - Building tools/expertise to monitor VPNs
    - Building tools/expertise to debug and manage the VPNs

VPN toolkit requirements (cont.)

- Support large scale VPN services
  - Large number of VPN customers
- Support multi-AS/multi-provider operations
  - As a particular VPN client may span more than one AS/provider

VPN toolkit requirements – how ?

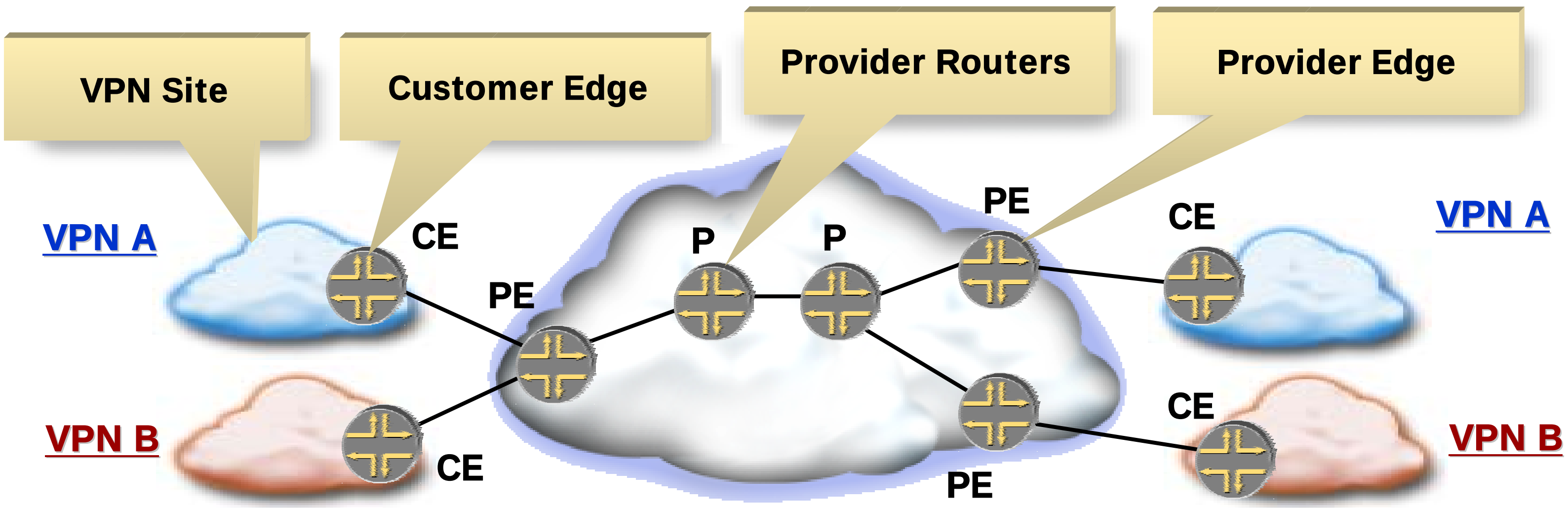
- By taking advantage of the commonalities between point-to-point Layer 2 VPN, VPLS, and IP VPN
  - e.g., auto-discovery, traffic separation, etc…
- By using tools that are general and easily extendable
  - To support VPN-specific (e.g., IP VPN specific, VPLS-specific) extensions
- By using tools that have good scaling properties
- By using tools that can operate in a distributed fashion
  - Including the ability to operate across multiple service providers

The rest of the talk is about the VPN Toolkit built around BGP and MPLS

BGP/MPLS VPN Toolkit - bits of history

- In the beginning…
  - IP VPN services (aka RFC2547 VPN)
    - RFC2547, draft-ietf-ppvnp-2547bis
- Later (re-using parts of RFC2547)…
  - BGP for VPN auto-discovery
    - draft-ietf-ppvpn-bgpvpn-auto
  - Layer 2 (point-to-point) VPN services
    - draft-kompella-ppvpn-l2vpn
- Most recent…
  - Virtual Private LAN Service (VPLS)
    - draft-kompella-ppvpn-vpls

Terminology



- Customer Edge device: device located on customer premises
- Provider Edge device: maintains VPN-related information, exchanges VPN information with other Provider Edge devices, encapsulates/decapsulates VPN traffic
- Provider router: forwards traffic VPN-unaware

VPN Control Plane functionality

- Constrained distribution of VPN information
  - Based on a well-established and widely used BGP Community-based route filtering mechanism
- Exchanging demultiplexor (VPN Label) by piggybacking it on top of the VPN information carried by BGP
  - VPN Label is used for traffic separation in the forwarding plane:
    - VPN Label is attached to data by ingress PE
    - Used by egress PE to determine to which VPN does a given packet belong
      - Also for Layer 2 (point-to-point) VPNs, which source site does a given packet belong
      - Also for VPLS, which source site has a given MAC address (used for MAC address learning)
- The additional overhead of carrying VPN Labels in BGP is insignificant, yet the benefits are obvious

VPN Control Plane functionality (cont.)

- Takes advantage of all the scalability enhancements available in BGP:
  - Route Reflectors, Refresh, etc…
- Supports Multi-AS/Multi-provider operations
  - As BGP is a protocol designed to support exchange of information across multiple ASs/providers
- Piggybacking VPN Label on top of the VPN information, and using BGP to carry both eliminates the need for a separate protocol to carry VPN Labels
  - No need to worry about scalability, multi-provider operations, etc… of the separate protocol
  - No need for complex inter-protocol interactions

Toolkit: support RFC2547 VPN

- CE-PE IP route exchange using a variety of protocols – static, RIP, OSPF, BGP
- PE-PE route distribution using Multi-Protocol BGP (RFC 2858)
  - Route Distinguisher: “uniquifies” routes
  - Route Target: determines VPN topology
- VPN Routing and Forwarding table (VRF) on PE holds all the routes for a VPN
  - Both received from the local CEs and from other PEs
  - Each VPN on a PE has its own VRF
- VPN labels carried by Multi-Protocol BGP (RFC3107)
  - Either label per VRF, or per route

Provisioning RFC2547 VPN on a PE

- Configure Route Distinguisher for VPN
- Configure CE-PE connection
  - Interface(s) to CE
  - Layer 3 Routing Protocol between CE and PE
- Configure import/export Route Targets

RD	1234:5.6.7.8
Layer 3 Protocol	RIP
Imp RT	1234:8765
Exp RT	1234:8765

Toolkit: Layer 2 (point-to-point) VPNs

- CE-PE: VPN Connection Table (VCT) is configured with:
  - Customer Port Identifier (aka CE ID) used by the (local) CE-PE connection
    - for each local ports of that VPN
  - Estimated total number of customer ports within the VPN
  - Analogous to CE-PE routes in RFC2547 VPNs
- PE-PE VCT distribution using Multi-Protocol BGP (RFC 2858)
  - Route Distinguisher: “uniquifies” VCT information
  - Route Target: determines VPN topology
- VPN Forwarding Table (VFT) on PE holds all the VCTs information
  - Both local as well as received from other PEs
  - Analogous to rfc2547 VRFs
- VPN labels carried by Multi-Protocol BGP
  - Label block instead of a single label

Provisioning Layer 2 VPN on a PE

- Configure Route Distinguisher for VPN
- Configure CE-PE connection
  - Interface to CE
  - Layer 2 encapsulation
  - Unique ID for the (local) CE port (CE ID, aka “Site Identifier”)
- Configure total number of customer ports in VPN (estimated)
- Configure import/export Route Targets

RD	1234:5.6.7.8
Layer 2 Protocol	Frame Relay
CE ID	3
# ports (# sites)	20
Imp RT	1234:8765
Exp RT	1234:8765

Toolkit: VPLS

- **CE-PE: VPN Connection Table (VCT) is configured with:**
  - **VPLS Edge Identifier (VE ID)**
    - **One per VPLS per PE (irrespective of how many local ports belong to that VPLS)**
    - **Degenerate case of Customer Port Identifier (CE ID)**
  - **Estimated total number of PEs that have ports belonging to that VPLS**
  - **Analogous to CE-PE routes in RFC2547 VPNs**
- **PE-PE VCT distribution using Multi-Protocol BGP (RFC 2858)**
  - **Route Distinguisher: “uniquifies” VCT information**
  - **Route Target: determines VPN topology**

Toolkit: VPLS (cont.)

- VPN Forwarding Table (VFT) on PE holds all the VCTs information
  - Both local as well as received from other PEs
  - Also contains MAC forwarding information
    - Created via a combination of MAC address learning and the VCT information
  - Analogous to rfc2547 VRFs
- VPN labels carried by Multi-Protocol BGP
  - Label block instead of a single label
    - Just like with Layer 2 (point-to-point) VPNs

Provisioning VPLS on a PE

- Remarkably similar to provisioning for Layer 2 VPNs, except:
  - Used VPLS Edge ID (VE ID) instead of CE ID
  - Layer 2 encapsulation is set to VPLS
  - VPN should be a full mesh
    - Import RT always the same as Export RT

RD	1234:5.6.7.8
Layer 2 Protocol	VPLS
VE ID	3
# sites	20
Imp RT	1234:8765
Exp RT	1234:8765

Configuration Fragment for RFC2547 VPN

```
routing-instances vpnA {   // Configuration for VPN A
    instance-type vrf;         // RFC 2547 VPN
    route-distinguisher 1234:5.6.7.8;
    route-target 1234:8765; // set Route Target to 1234:8765
    protocols {                // PE-CE protocol
        rip {
            version-2;             // RIPv2
            group to-CE-A3 {
                export default;
                interface so-0/0/0.0;  // sub-interface for RIPv2
            }
        }
    }
}
```

Configuration Fragment for L2 VPN

```
routing-instances vpnA {    // Configuration for VPN A
   instance-type l2vpn;      // L2 VPN
   route-distinguisher 1234:5.6.7.8;
   route-target 1234:8765; // set Route Target to 1234:8765
   protocols {               // PE-CE protocol
      l2vpn {
         encapsulation-type frame-relay;
         site CE-A3 {
            site-identifier 3;
            interface so-0/0/0.0;  // sub-interface to A1
            interface so-0/0/0.1;  // other sub-interfaces
         }
      }
   }
}
```

Configuration Fragment for VPLS

```
routing-instances vpnA {   // Configuration for VPN A
    instance-type l2vpn;      // L2 VPN
    route-distinguisher 1234:5.6.7.8;
    route-target 1234:8765; // set Route Target to 1234:8765
    protocols {                // PE-CE protocol
        l2vpn {
            encapsulation-type vpls;
            site CE-A3 {
                site-identifier 3;
                interface ge-0/0/0.0;  // multipoint Ethernet interface
            }
        }
    }
}
```

The Rest Is Up to Multi-Protocol BGP

- Allocates a VPN label (for IP VPNs) or label block (for Layer 2 and VPLS) as the demultiplexor
- Distributes BGP VPN advertisements (plus VPN labels) with Export Route Target to other PEs
- Receives BGP VPN advertisements from all other PEs
- Decides if received advertisement belongs to given VPN based on Import Route Target; if so, uses this advertisement to build VRF/VFT for this VPN

Applies to RFC2547 VPN, L2 VPN, and VPLS !!!

Inter-AS/Inter-provider operations

- Exchange VPN information + VPN labels across AS/provider boundary by using BGP between BGP Route Reflectors in each AS/provider
  - Route Reflectors preserve the next hop information and the VPN label across the AS/provider
- PEs learn routes and label information of the PEs in the neighboring ASes through ASBRs
  - Using labeled IPv4 routes
- No VPN information (e.g., VRF, VFT) on ASBRs

Applies to RFC2547 VPN, L2 VPN, and VPLS !!!

Scalability - "divide and conquer"

(1) Two levels of labels to keep P routers free of all the VPN routing information

(2) PE router has to maintain VPN information only for VPNs whose sites are directly connected to the PE router

(3) Partition BGP Route Reflectors within the VPN Service Provider among VPNs served by the Provider

⇒ No single component within the system is required to maintain information for all the VPNs

⇒ Routing capacity of the system isn't bounded by the capacity of an individual component

Applies to RFC2547 VPN, L2 VPN, and VPLS !!!

Juniper NETWORKS

Overloading BGP ?

- Common concern: {public Internet + RFC2547 VPNs + L2 VPNs + VPLS} will overload BGP, causing it to crash
- Real question: should a single box provide all of the above services ?
  - If so, doesn't matter which protocol – either the box can take it, or not
  - Putting these services in different protocols doesn't reduce overall stress on the box!
  - Existence proof that some boxes *can* take it!

Overloading BGP ? (cont.)

- Other concern: BGP is complex to implement
  - The (perceived) vendor complexity of implementing Multi-Protocol BGP is well worth the greatly simplified operations for the Service Providers
    - Using Multi-Protocol BGP for both auto-discovery and distributing VPN labels means fewer protocols to operate, manage and debug

Why not BGP + LDP ?

- BGP vs LDP for signaling (distributing VPN labels) is a wrong comparison
  - ignores the fact that VPLS includes not just signaling, but autodiscovery – need to look at the whole system
- Need to compare BGP for both autodiscovery and signaling vs BGP for autodiscovery + LDP for signaling
- With BGP + LDP approach signaling of VPN labels requires a completely separate protocol (LDP)
- With BGP/MPLS toolkit signaling of VPN labels is a "side effect" of (BGP-based) autodiscovery
- Additional overhead of carrying labels in BGP is negligible
- The overhead of using BGP for both autodiscovery and signaling is about the same as using BGP just for autodiscovery -> certainly less than using BGP for autodiscovery + LDP for signaling

Why not BGP + LDP ? (cont.)

- Because using label blocks (as required for both Layer 2 VPNs and VPLS) doesn't introduce any significant <u>practical</u> drawbacks
- BGP for auto-discovery, LDP for distributing VPN labels
  - Doesn't work for RFC2547 VPNs – for 2547VPNs BGP carries both VPN information and labels
- For the same reasons why BGP was extended to carry labels, instead of using (BGP+LDP)
- Why using two protocols (BGP+LDP) to accomplish a given task is any better than accomplishing the same task with just one (BGP) ?

Quote From a Convinced Service Provider

"You roll out a protocol thinking you understand it. You spend the first year learning how little you really understood – being paged at 3 a.m. is a wonderful educational tool! You then realize that you don't want to do this again (or as little as you possibly can)."

Summary

Customers want:

- Point-to-point Layer 2 VPNs
- Virtual Private LAN Service (VPLS)
- IP VPNs (RFC 2547 VPN)

Service Providers can offer all of the above:

- over a common infrastructure (MPLS)
- with a common framework (Multi-Protocols BGP / MPLS)
  - Taking advantage of BGP scalability and multi-AS / multi-provider support
- with common concepts (Route Distinguisher, Route Target, VRF / VFTs, ...)

A single operational infrastructure and a small set of basic mechanisms means considerable savings!

References

- RFC 2547 “BGP/MPLS VPNs”
- draft-ietf-ppvpn-rfc2547bis
- draft-ietf-ppvpn-bgpvpn-auto
- draft-kompella-ppvpn-l2vpn
- draft-kompella-ppvpn-vpls

Thank you!

http://www.juniper.net
http://www.juniper.co.jp