Network Working Group
Internet Draft
draft-ietf-l3vpn-vpn-vr-03.txt
Expires: September 6, 2006

Paul Knight (Editor)
Hamid Ould-Brahim
Nortel Networks

Bryan Gleeson
Nokia

March 6, 2006

Network based IP VPN Architecture Using Virtual Routers

Status of this Memo

Copyright Notice

Abstract

This document describes a network-based Virtual Private Network (VPN) architecture using the virtual router (VR) concept. Multiple VRs can exist in a single physical device. A VR emulates all the functionality of a physical router, and therefore inherits all existing mechanisms and tools for configuration, operation,

accounting, and maintenance. Any routing protocol can be used to

distribute VPN reachability information among VRs, and no VPN-related modifications or extensions are needed to the routing protocol for achieving VPN reachability. Direct VR-to-VR connectivity may be configured through layer-2 links or through IP- or MPLS-based tunnels. Traffic from VRs belonging to different VPNs may be aggregated over a "backbone VR" network, which greatly simplifies VPN provisioning. This architecture accommodates various backbone deployment scenarios, both where the VPN service provider owns the backbone, and where the VPN service provider obtains backbone service from one or more other service providers.

Table of Contents

1. Introduction

This document describes a network-based VPN architecture using virtual routers. The objective is to provide per-VPN routing, forwarding, quality of service, and service management capabilities. The VPN service is based on the virtual router concept. The VR VPN architecture is compatible with the Layer 3 PPVPN framework described in [RFC-4110], as well as the generic PPVPN requirements [RFC-3809] and the service requirements for Layer 3 PPVPNs [RFC-4031].

A virtual router (VR) has exactly the same mechanisms as a physical router, and therefore can inherit all existing mechanisms and tools for configuration, deployment, operation, troubleshooting, monitoring, and accounting. Multiple VRs can exist in a single physical device. Virtual routers can be deployed in various VPN configurations. Direct VR to VR connectivity may be configured through layer-2 links or through a variety of tunnel mechanisms, using IP- or MPLS-based tunnels. Multiple VRs may be aggregated over a "backbone VR." This architecture accommodates various backbone deployment scenarios, including where the VPN service provider owns the backbone, and where the VPN service provider obtains backbone service from one or more other service providers.

This informational document does not specify a protocol, and therefore does not specify the manner in which VRs interoperate. Instead, VRs are interconnected using existing IETF specifications for tunneling mechanisms such as IPsec [RFC-4301], GRE [RFC-2784] (optionally with key and sequence number extensions [RFC-2890]), IP-in-IP [RFC-2003], and MPLS [RFC-3031] [RFC-3035] tunnels. Interaction between VRs across these tunnels make use of the same mechanisms and routing protocol standards which would normally be used for interoperation between physical routers [RFC-1812], such as BGP [RFC-4271], OSPF [STD-54], IS-IS [RFC-1195] [RFC-3787], as well as others.

There are several ways of implementing the VR architecture. Although this document does not guarantee interoperability by specifying explicit requirements, the nature of the VR approach makes it likely that interoperability can be achieved among various VR implementations, just as interoperability of VRs with normal routers is straightforward. Section 14, "Interoperability," addresses this in more detail.

In the VR architecture, an instance of routing is used to distribute VPN reachability information among the VRs supporting each VPN. Any routing protocol can be used, and no VPN-related modifications or extensions are needed to the routing protocol for achieving VPN reachability. VPN reachability information to and from customer sites can be dynamically learned from the CE using standard routing protocols, or it can be statically provisioned on the VR. The routing protocol between the virtual routers and CEs is independent of the routing used in the VPN backbone, between the VRs. That is, the routing protocol between the VRs may be the same or it might be different than the routing mechanism used between the CE and VR, or it may be a different instance of the same protocol. Likewise, since the VR-to-VR connectivity can use tunnels, the inter-VR routing protocol can be independent of the routing used in the backbone network(s) over which the VR-based VPN runs.

There are two fundamental architectures for implementing network-based IP VPNs: virtual routers (VR) and aggregated routing. The main difference between the two architectures resides in the model used to achieve VPN reachability and membership functions.

In the VR model, each VR in the VPN domain is running an instance of routing protocol responsible for disseminating VPN reachability information between VRs. Therefore, VPN membership and VPN reachability are treated as separate functions, and separate mechanisms are used to implement these functions. In [RFC-4110], this is referred to as using per-VPN routing. VPN reachability is carried out by a per-VPN instance of routing, and a range of mechanisms is possible for determining membership (see section 6.0).

In the aggregated routing model [RFC-4110], the VPN site's routing protocol is terminated at the edge of the backbone, and a backbone routing protocol (i.e., extended BGP-4) is responsible for disseminating the VPN membership and reachability information between provider edge routers (PE) for all the VPNs configured on the PE. "BGP/MPLS VPNs" [RFC-4364] describes an example of an aggregated routing VPN architecture.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC-2119].

2. Virtual Router VPN Architecture Guidelines

The following guidelines are intended for designers of VR implementations. Since this document does not describe a specific interoperable protocol, there may be alternative ways to achieve these guidelines.

2.1 Membership

The VR Architecture requires a way to determine VPN membership. This can be accomplished by configuration, by the use of an autodiscovery mechanism such as [VPN-BGP], or by other means. This is discussed in detail in Section 6, "VPN Membership and Topology Auto-Discovery".

The use of a VPN identifier (VPN-ID) provides a way to simplify most VPN configuration and operational tasks. All virtual routers that are members of a specific VPN should share the same VPN-ID. This may be the VPN-ID format defined in [RFC-2685].

2.2 Scalability

In this architecture, the backbone internal nodes (e.g., P routers [RFC-4110]) are not VPN- or VR-aware, and therefore they don't keep any VPN state within the backbone. Thus the VR architecture avoids any significant contribution to problems of backbone scalability.

The PE on which the VRs run (and the VRs themselves) should be able to accommodate rapid growth in the number of routes per VR, since this number can change suddenly as membership changes. The PE should be able to accommodate substantial growth in the number of VRs and CEs supported, to avoid reconfiguration that could disrupt existing connectivity.

The optional use of the "backbone VR" improves the scalability of the VR approach, since multiple VRs on a PE may share a single backbone VR connection to their peer VRs on another PE, rather than establishing multiple separate per-VR or per-VPN connections between PEs. The backbone VR is described in more detail in section 5.3.

2.3 Quality of Service

Existing quality of service mechanisms developed for physical routers should all be available to be used on a per-VR basis. Therefore, quality of service (policing, shaping, classification, and scheduling) should be configurable on a per-VPN basis.

2.4 Auto-discovery

It should be possible for the VRs to automatically discover each other, set up tunnels to each other, and exchange private routing information across the backbone. The auto-discovery mechanism must take into consideration the case where the VPNs are implemented across administrative domains. We assume in this document that an auto-discovery mechanism which provides services similar to BGP (as described in [VPN-BGP]) is used as the mechanism to distribute membership, topology, and tunnel information among VRs which are members of the same VPN.

2.5 Routing

2.5.1 Routing between CE and PE

Any existing routing protocol may be used between the CE and the VR running on the PE. Typically, the routing protocol of the specific VPN site will be used. Static routes may be used. The routing protocol between the CE and the VR running on the PE may be independent of the PE-to-PE routing. That is, they may be different routing protocols, or different instances of the same routing protocol.

2.5.2 Routing in the Service Provider Network (Backbone)

The choice of the backbone routing protocol should not be constrained by the VPNs. This is one of the key attractions of the VR model, since it does not require VPN service providers to run and manage any specific backbone protocol or technology.

2.5.3 Routing between VRs in a VPN

Any existing routing protocol may be used between VRs in a VPN. The routing protocol between the VRs may be independent of the CE-to-PE routing.

VRs belonging to the same VPN may construct tunnels providing connections to each other, using information from the backbone routing protocol. They may then exchange routing information and VPN traffic over these tunnels.

A backbone VR network may be constructed among some or all PEs. VRs of customer VPNs may use the backbone VR for routing across the backbone.

It is strongly recommended that care be taken when multiple routing protocols are used, due to differences in metrics, detail of information, etc.

2.5.4 Multicast routing using VRs

VR-based VPNs should provide the same mechanisms for IP multicast routing as ordinary routers. The processes of multicast tree construction and packet replication should be configurable in the same way as for ordinary routers. In addition, VR-based VPNs may be able to employ mechanisms to optimize multicast, depending on the specific VPN configuration.

2.6 Security

The VR architecture accommodates a variety of approaches to ensure security for VPN data, routing, and other control information.

Different levels of security are possible, using the security features of routing and tunneling protocols. The architecture should provide authentication and encryption services for VPNs requiring strong security capabilities. VR-based VPN implementations should support the VPN Security Framework [RFC-4111].

2.7 Topology

VPN topologies such as a hub and spoke, and full mesh can be supported by the VR architecture. It is also possible to build arbitrary VPN topologies.

For example, a PE device with VRs supporting certain VPNs may be able to act as a P (Provider backbone) device with respect to other VPNs. This increases provisioning flexibility in many topologies.

2.8 Tunneling

The VR architecture should not be limited to a single tunneling mechanism. It may allow the use of IPsec [RFC-4301], GRE [RFC-2784], IP in IP [RFC-2003], and MPLS [RFC-3031], [RFC-3035] tunnels. It should also allow multiple VPNs to share a tunnel across a backbone. Within a single VPN, different types of tunnels should be allowed.

2.9 Management

The VR architecture should provide mechanisms to make it easy to configure, deploy, operate and troubleshoot each VPN independently, using existing mechanisms and tools. Tools commonly used for operating, managing and debugging IP networks should be able to be used without any modification.

Most aspects of the management of the multiple VRs on the PE by the Service Provider are implementation-specific, and beyond the scope of this document.

2.10 Additional Characteristics

The following are some additional general characteristics of the VR architecture:

1) The architecture accommodates different sizes of VPNs, and one VPN should not impact other VPNs on the PE.
2) The architecture supports overlapping VPN address spaces in separate VPNs.
3) The architecture supports direct paths between VPN sites that bypass the service provider backbone (backdoor links). Traffic can be directed to the backdoor link, or injected to the backbone with the flexibility of using both the backbone access, and the backdoor link as internal or external paths.
4) The architecture works over different deployment scenarios, e.g.

where the service provider owns its own backbone, and where the service provider obtains backbone service from one or more other service providers.

3. Network Reference Model

A VPN customer site is connected to the provider backbone by means of a connection between a Customer Edge (CE) device, (which can be one or more hosts and/or routers) and a virtual router (VR). CE devices are preconfigured to connect to one or more VRs. Multiple VRs may coexist on the same service provider edge device (PE).

CE devices can be attached to VRs over any type of access link (e.g. ATM, frame relay, Ethernet, PPP [STD-51], L2TP [RFC-2661] or IP tunneling mechanisms such as IPsec or GRE tunnels).

```
                           +---+    +---+
                           | P |....| P |
                           +---+    +---+
                     PE   /              \  PE
          +-----+  +-------+              +-------+  +----+
          | CEs|--|-{VRs}|               |{VRs}-|--|CEs|
          +-----+  +-------+              +-------+  +----+
                          \              /
                           +---+    +---+
                           | P |....| P |
                           +---+    +---+
```

Figure 1: Network Reference Model

CE sites can be statically connected to the provider network via dedicated circuits, or can use dial-up links. Routing tables associated with each virtual router define the site-to-site reachability for each VPN. The internal backbone provider routers (P) are not VPN aware and do not keep VPN state.

3.1 Backbone

In general the backbone is a shared network infrastructure, which represents either:
1) A layer-2 ATM or frame relay network.
2) An IP network.
3) An MPLS network.

Not all VPNs existing on the same PE are necessarily connected via the same backbone. A single PE can be connected to multiple backbones. Individual VRs on the PE may also connect to multiple backbones. Thus a single VPN can be built from multiple transport technologies in the VR architecture.

4. Virtual Router Definition

A virtual router (VR) is an emulation of a physical router at the software and/or hardware levels. Virtual routers have independent IP routing and forwarding tables, and they are isolated from each other. This means that two VRs on a PE can serve two different VPNs which may have overlapping address space. The addresses need only be unique within a VPN domain.

A virtual router has two main functions:
1) Constructing routing tables for the paths between VPN sites using any routing technologies (e.g., static, OSPF, RIP, or BGP).
2) Forwarding packets to the next hops within the VPN domain.

From the VPN user point of view, a virtual router provides the same functionality as a physical router. Separate routing, and forwarding capabilities provide each VR with the appearance of a dedicated router that guarantees isolation from the traffic of other VPNs, while running on shared forwarding and transmission resources.

Virtual routers belonging to the same VPN domain should have the same Virtual Private Network Identifier (VPN-ID). The VPN-ID may use the format described in [RFC-2685]. As noted in [VPN-BGP], when the VRs in a given VPN use BGP as the backbone routing protocol, the VPN-ID can be carried in the NLRI to make the addresses of VRs globally unique. Since globally unique addresses are necessary if BGP is used for auto-discovery, the use of a consistent VPN-ID is a key element in supporting auto-discovery and improving scalability of VR-based VPN services.

To the CE access device, the virtual router appears as a neighbor router in the CE based network. The CE sends all traffic for non-local VPN destinations to the VR, unless the specific VPN topology provides alternate routes. Each CE access device must learn the set of destinations reachable through its connection to the virtual router; this may be as simple as a default route. Virtual routers participating in a single VPN domain are responsible for learning and disseminating VPN reachability information among themselves. A given VR holds the routes only for the specific VPN of which that VR is a member. Any routing protocol can be used between the VRs and the CEs.

5. How VPNs are Built and Deployed using VRs

Three main VR deployment scenarios can be used for building VPNs:
1) VR to VR connectivity over a layer 2 connection.
2) VR to VR connectivity tunneled over an IP or MPLS network.
3) Aggregating multiple virtual routers over a "backbone virtual router," which will provide connectivity over a layer 2, IP, or

MPLS network.

These VR deployment scenarios can coexist on a single PE or within a single VPN.

5.1 VR to VR Connectivity over Layer 2 Connections

As illustrated in Figure 2, virtual routers can be deployed over direct layer-2 frame relay or ATM connections or other layer-2 transport technology.

```
                 PE                             PE
           +---------------+            +---------------+
   +-----+ |               |            |               | +-----+
   |VPN-A| | +----+        Layer-2 connections   +----+ | |VPN-A|
   |sites|-|-|VR-A|<---------------------------->|VR-A|-|-|sites|
   +-----+ | +----+        |  --------  |        +----+ | +-----+
           |               |-( Layer-2)-|               |
   +-----+ | +----+        | (Backbone) |        +----+ | +-----+
   |VPN-B|-|-|VR-B|        |  --------  |        |VR-B|-|-|VPN-B|
   |sites| | +----+<--------------------|------->+----+ | |sites|
   +-----+ |               |            |               | +-----+
           +---------------+            +---------------+
```

Figure 2: VR to VR connectivity over a layer-2 backbone

This type of VR deployment allows direct quality of service engineering on a per-VPN connection basis. The connections can be statically configured or dynamically established.

5.2 VR to VR Connectivity through IP or MPLS tunnels

In addition to connecting via layer-2 transport technologies, virtual routers can connect over an IP or MPLS backbone. In a manner analogous to layer-2 transport, they can use the backbone to support tunneled connections among the VRs. The topology can be described similar to that for layer-2 transport, as in figure 2.

VPN data and routing information is tunneled through the use of IP or MPLS based tunnels (e.g., IPsec, GRE, IP in IP, MPLS). The use of tunnels between VRs is addressed in more detail in the discussion of backbone VRs in the following section of this document.

Although it is clearly possible to use a topology similar to the layer-2 model over an IP or MPLS backbone, the VR capability also provides a highly scalable alternative to the use of individual tunnels between VRs. This alternative is the creation (on each participating PE) of another VR facing into the backbone network, which is used to build a kind of backbone VPN that may be shared among multiple customer VPNs. This is described below as the "backbone VR."

5.3 Virtual Router Backbone Aggregation

Another typical VPN configuration consists of connecting multiple virtual routers to the backbone through the use of a single virtual router in each PE (figure 3). In the following sections we call this single virtual router "the backbone virtual router" or "the backbone VR." The backbone VR is a mechanism to enhance scalability. The use of backbone VRs is optional in VR-based VPNs. When backbone VRs are used, they should be configured on all PEs which participate in VPNs carried over the backbone VRs.

The backbone virtual router is not functionally different than other virtual routers. It is only a virtual router that is configured and deployed in a special configuration.

The backbone VR connects each PE to a shared backbone infrastructure. Backbone VRs can be deployed over ATM, FR, IP, or MPLS networks. Since the backbone VR allows the aggregation of VRs from multiple VPNs, backbone configuration can remain unaffected as new VPNs or VPN sites are added. The relationship between the VRs and the backbone VR is an overlay relationship.

```
                     PE-1                          PE-2
              +---------------+            +---------------+
              |               |            |               |
      +-----+ | +----+     MPLS/IP based Tunnels    +----+ | +-----+
      |VPN-A| | |VR-A|........|<---------->|........|VR-A| | |VPN-A|
      |sites|-|-|(1) |        |            |        |(2) |-|-|sites|
      +-----+ | +----+\+----+ | ---------  | +----+/+----+ | +-----+
              |        |VR-1|-|-(IP/MPLS )-|-|VR-2|        |
      +-----+ | +----+/+----+ |(Backbones) | +----+\+----+ | +-----+
      |VPN-B|-|-|VR-B|        | ---------  |        |VR-B|-|-|VPN-B|
      |sites| | |(1) |        |            |        |(2) | | |sites|
      +-----+ | +----+........|<---------->|........+----+ | +-----+
              |               |            |               |
              +---------------+            +---------------+
```

Figure 3: VR-1 and VR-2 used as backbone VRs

The relationship between the "ordinary" VPN VRs and the backbone VRs is conceptually similar to the relationship between separate routers, even though they coexist in the same device. The individual VRs in a PE, representing different VPNs, can relate to the backbone VR as if they were the CEs of a single VPN, with the backbone VR acting as a PE to them. Thus the VPNs can be multiplexed in a hierarchical fashion, using IP encapsulation or stacked labels, depending on the tunnel technology used between the backbone VRs.

The use of the backbone VR provides multiplexing across the backbone for multiple VPNs, while still allowing individually-engineered connections where desired. Note that Figure 3 depicts both a backbone connection between backbone VRs (VR-1 to VR-2) and also connections between the customer VPN VRs (VR-A(1) to VR-A(2) and VR-B(1) to VR-B(2)) which do not pass through the backbone VRs. Both types of connections may be used simultaneously, e.g., to provide differentiated services to different classes of traffic. Best-effort traffic between VR-A(1) and VR-A(2) may be routed through the shared backbone VRs, while high-priority traffic between these same VRs might be routed through the direct connection, which could be engineered with higher Quality-of-Service parameters. This illustrates how a service provider can trade off greater scalability offered by the backbone VR against higher value "personalized service" for VPN customers.

Note that although the backbone VR concept is described above using a single backbone VR per PE, there may be multiple backbone VRs per PE.

5.3.1 Tunneling

VPN data and routing information is tunneled through the use of IP or MPLS based tunnels (e.g., IPsec, GRE, IP in IP, MPLS). Depending on the tunnel technology used, the tunnels can be statically configured or dynamically established. The tunnel appears to VRs as a point-to-point link. Traffic sent through the tunnel and forwarded by the backbone VR is opaque to the underlying backbone technology used.

A tunnel can be established per VPN or shared among many VPNs (VRs). The tunnel can originate from the backbone virtual router or from the VRs. This can provide an opportunity for service differentiation, in which a service provider can offer a higher level of service (at a higher price point) for individually mapped VPN connections among a customer's VRs.

The backbone VR makes it appear as if each VR within a VPN is directly connected. It supports both full and partial mesh configurations. Each VR within the VPN exchanges routing information directly with the adjacent VRs in the VPN. Note that adjacency in this case is determined by the overlay topology of the particular VPN, as determined by configuration or discovery.

A single VR-based VPN may use different type of tunnels for inter-VR connectivity. Some sites may use MPLS, while other sites (which transit through non-secure domains) may choose to use IPsec to encrypt their data.

The scalability and security of dynamic tunnel establishment between VRs is enhanced by the ability to exchange a VPN-ID. [VPN-BGP] supports auto-discovery of the VPN-ID within BGP-based networks. Further work beyond the scope of this document is needed to determine the requirements and usage of the VPN-ID exchange within most tunneling scenarios.

5.3.1.1 MPLS Tunnels

The VR architecture can use MPLS tunneling in various forwarding scenarios. Individual VRs of some VPNs may be configured to participate in BGP/MPLS IP VPNs as described in [RFC-4364].

In some scenarios, a hierarchy of two labels can be used. One simple forwarding scenario is where the inner label identifies the VR intended to receive the private packet (to be forwarded to the CE).

Another forwarding scenario is to distribute the inner label on a per-VPN basis across the tunnels, after the tunnel endpoints (VRs) have been discovered. The label and reachability distribution is done through the tunnels. In this case the inner label distribution process can be achieved using BGP or an existing label distribution protocol on a per-VPN basis. The inner label relates to the private VPN prefixes. On the egress side traffic will be directed to the egress interface by looking up the inner label.

5.3.1.2 IPsec Tunnels

IPsec [RFC-4301] is needed when there is a requirement for strong encryption or strong authentication. It also supports multiplexing and a signaling protocol - IKE. IPsec tunnels can be established between two VPN sites across the backbone (originating from the backbone VRs).

5.3.2 Routing

The backbone VR exchanges backbone routing information with other backbone entities (P routers and possibly other backbone VRs). The backbone routing is separated from the customer VPN routing.

Virtual routers can run any routing protocol on their local VPN domain. Both static routes and dynamic routing protocols such as RIP, OSPF, and BGP-4 can be used. The VRs of a given VPN exchange routing information with adjacent VRs through the tunnels over the backbone.

If a backdoor link is used between VPN sites running any IGP, then by adjusting the backdoor link costs appropriately, the backbone link can be favored for forwarding VPN traffic. By lowering the weight, the backdoor link can be used as a backup link in case the backbone path fails.

5.3.3 Relationship between the VRs and the Backbone VR

The routing domain of a set of VRs participating in a single VPN has no relation to the routing domain of the backbone VR. The backbone VR is not necessarily aware of the routing instances running on each private virtual router. However, because the backbone VR is also a virtual router, it can build routing relationships with other VRs if needed.

5.3.4 Multiple Backbones Connected to a Single PE

Figure 4 illustrates an example where multiple backbones are connected to the same PE. This type of configuration can be used when the PE is connected to multiple service provider backbones, or when the service provider offers different VPN services for different types of backbones.

```
                  PE                            PE
             +---------------+            +---------------+
     +-----+ |               |            |               | +-----+
     |VPN-A|-|-+----+        |            |        +----+-|-|VPN-A|
     |sites| | |VR-A|\       |            |        |VR-A| | |sites|
     +-----+ | +----+ +----+ |  --------- | +----+/+----+ | +-----+
             |        |VR-1|-|-(Backbone )|-|VR-2|        |
     +-----+ | +----+/+----+ | (    1    )| +----+\+----+ | +-----+
     |VPN-B|-|-|VR-B|        |  --------- |        |VR-B|-|-|VPN-B|
     |sites| | +----+        |            |        +----+ | |sites|
     +-----+ |               |            |               | +-----+
             |               |            |               |
     +-----+ |               |            |               | +-----+
     |VPN-C| | +----+        |            |        +----+ | |VPN-C|
     |sites|-|-|VR-C|\       |            |        |VR-C|-|-|sites|
     +-----+ | +----+ +----+ |  --------  | +----+/+----+ | +-----+
             |        |VR-3|-|-(Backbone)-|-|VR-4|        |
     +-----+ | +----+/+----+ | (  2 & 3 ) | +----+\+----+ | +-----+
     |VPN-D|-|-|VR-D|        |  --------  |        |VR-D|-|-|VPN-D|
     |sites| | +----+        |            |        +----+ | |sites|
     +-----+ |               |            |               | +-----+
             +---------------+            +---------------+
```

Figure 4: Multiple Backbones Connected to a Single PE

6. VPN Membership and Topology Auto-Discovery

The virtual router approach explicitly separates the mechanisms used for distributing reachability information from mechanisms used for distributing VPN topology and membership information. VPN membership information refers to the set of PEs (and the VRs on those PEs) that have customers in a particular VPN. VPN topology represents the set of VRs configured on PEs and their interconnectivity within the VPN. The topology can be a full-mesh of VRs, a hub and spoke, or anything in between. Dynamic topology due to on-demand VPN customers can also be handled.

VPN discovery can be achieved through a variety of different mechanisms, for example:

- Directory server approach, in which VRs query a server to determine their neighbors.
- Explicit configuration via a management platform.
- Piggybacking VPN membership and topology information using existing routing protocols (e.g., BGP) [VPN-BGP].
- Other VPN membership and topology auto-discovery approaches.

The above mechanisms can be combined on a single PE, with different mechanisms used on a per-VPN basis. As an example, for some VPNs topology discovery is done only through a management platform. For others, dynamic topology discovery is achieved using existing routing protocols.

In this document it is assumed that a mechanism that provides services similar to BGP is used to achieve auto-discovery of VPN members. A robust auto-discovery mechanism provides the scalability needed in large provider-provisioned VPNs. In the approach described in [VPN-BGP], VR addresses are exchanged, along with the information needed to enable the PEs to determine which VRs are in the same VPN ("membership"), and which of those VRs are to have VPN connectivity ("topology"). Once the VRs are reachable through the tunnels, routes ("reachability") are then exchanged by running existing routing protocols on a per-VPN basis across the tunnels.

It is important to note that, for the VR architecture, the auto-discovery mechanism is only used to automatically exchange VPN control information between VRs and/or PEs. It is not intended for piggybacking VPN private reachability information onto the backbone routing instance, as is done in [RFC-4364], for example.

7. VRs and Extranets

Extranets are commonly used to refer to a scenario whereby a company has network access to a limited part of another company's corporate network. An important feature of extranets is the control of who can access what data, and this is essentially a policy decision. Policy decisions are enforced at the interconnection points between different domains. The enforcement may be done via a firewall, a router with access list functionality, or any device capable of applying policy decisions to transit traffic.

In the VR architecture, policy can be enforced between two VPNs, or between a VPN and the Internet, in exactly the same manner as is done today without VPNs. For example, two VRs (of different VPNs) could be interconnected, with each VR locally imposing its own policy controls via a firewall or other enforcement mechanism on all traffic that enters its VPN from the outside (whether from another VR or from the Internet). Combining firewalls and exchanging private routes between VRs (members of different VPNs) provide a flexible mechanism to build different flavors of extranets.

8. VPNs across Domains

It is possible that a VPN may cross multiple domains administered by different service providers. In the VR model, tunnels are used to provide intra-VPN connectivity across the backbones. The main requirement for the service provider in order to achieve end-to-end cross-domain VPN connectivity is the ability for both domains to support a common tunnel technology, plus the ability to support a common membership and topology discovery technology. Once the tunnel is established, private data (e.g., routing information, and private customer data) can flow from one domain to the other with the same level of security or isolation as that tunnel mechanism provides when used within a single service provider network.

Another scenario for supporting VPNs with multiple service providers is to use two virtual routers configured on PEs at the interconnection points. Each VR will use policy decisions and firewalling to control VPN traffic transiting from one domain to the other. The two "gateway VRs" have some similarities to the "backbone VRs," specifically with respect to being able to handle multiple VPNs. The individual VPN traffic is not terminated on these "gateway VRs". They provide ingress/egress filtering for any or all the bidirectional tunneled VPN traffic crossing the boundary. The VPN traffic will normally be opaque at the boundary, and typical inter-provider agreements apply to all traffic within individual VPNs, so the inter-provider VPN traffic is typically filtered all-or-nothing (by VPN) based on the visible packet identifiers or labels.

When there are VPN links crossing intervening domains which are not VPN-aware, tunnels should be configured across the intervening domains, and the "gateway VR" approach can be employed at the tunnel endpoints to provide security services appropriate to the circumstances. Some aspects of this are discussed in more detail in the "Carrier's Carrier" section.

The ability to use a standard, globally-unique VPN-ID format also supports the implementation of unambiguous VPN traffic identification mechanisms across domains.

9. Internet Access

The same link attaching the CE to the VR can be used to provide Internet access to the VPN sites. The VR operations can be decoupled from the mechanisms used by the customer sites to access the Internet.

There are a number of ways to provide Internet access to a VPN using the VR model. One way of providing VPN Internet access is to configure a "backbone VR" to steer private traffic to the VPN VR, and Internet traffic to the normal backbone/Internet forwarding table. The backbone VR can hold the Internet routes (so it will not be necessary for the VPN VRs to handle them). Firewall functionality should be used to secure the Internet backbone VR access. Network address translation services can also be configured on the backbone VR or on VPN VRs where needed for Internet access.

There are a number of other options, since the VR architecture reflects the flexibility of normal router architecture. An additional approach is to configure a particular VR to handle Internet access only (rather than going to the backbone VR). Another approach is to use a default route to an Internet gateway (which could be a VR).

10. Carrier's Carrier Case

In some cases, the customer of a VPN is a service provider or carrier offering VPN services for its own customers. We can describe this as a VPN hierarchy, with the "carrier's carrier" providing backbone services to a "sub-carrier." This is sometimes called "VPN wholesaling." The carrier's carrier may support multiple sub-carriers within a single PE device. The VR model provides several approaches to implement this VPN hierarchy.

In one approach, tunnels are built from the VRs of the carrier's carrier to the CEs of the customers of the sub-carrier ("remote CEs"). In this case, the VRs of the carrier's carrier provide VPN service to the remote CEs. The sub-carrier provides transport but does not participate in the VPN services. This can be particularly useful in cases where the sub-carrier's PE or P devices are themselves VRs (which may be instantiated within the same device as the VRs of the carrier's carrier, handling the connections from the remote CEs) and where the sub-carrier is outsourcing the management of its customers' VPN services.

Another approach is where the sub-carrier's VPN services are completely transparent to the VRs of the carrier's carrier. This is the default case. It is up to the sub-carrier's VPN service to distribute VPN reachability among the CEs of its customers.

11. Operations and Management

Each VR operates independently, and can be individually reconfigured without affecting other VRs on the same PE. In some implementations, it may be possible for a VR to be "rebooted" without affecting other VRs. In case of PE failure (e.g., migration, upgrades, etc.), the service provider may want to control and decide what VPN services get reestablished first. This particular point is important when a large number of VPNs is supported on the PE where each VPN service has different service availability requirements.

Since each VR operates as an independent router, it is possible for the management of the VRs to be outsourced. VPN customers may choose to configure (or perhaps only to monitor) the VRs that make up their VPN. It is also possible that the backbone VRs could be managed by a separate entity.

11.1 Backbone Migration

One benefit in using multiple backbone virtual routers is the ability for the backbone network administrator to migrate its backbone from one core technology to another with minimal disruption to VPN services. Conversely, a VPN configuration change or a VPN-software upgrade is totally transparent to the backbone protocol and policies (this is due to decoupling the VPN routing protocol from the provider backbone routing protocol).

11.2 Troubleshooting

The service provider or the VPN customer can use all existing troubleshooting tools on a per-VPN basis (e.g. ping and traceroute). As an example, a VPN customer may be able to perform some troubleshooting operations on its own VR. In this particular case, the service provider can configure restricted privileges for each VPN customer over the VR associated with the customer's VPN network. This access may provide only the privilege to monitor (with no privilege to change) the layer 3 status of the customer's VPN, as seen by the VR. The service provider may be able to offer VPN customers an SNMP-based method for read-only access to information about their own VPN. However, backbone topology information is completely hidden to the VPN VR, and therefore to the service provider's customer.

12. Quality of Service

This architecture can utilize a variety of Quality of Service mechanisms. QoS mechanisms developed for physical routers can be used with VRs, on a per-VR basis, including classification, policing, drop policies, traffic shaping and scheduling/bandwidth reservation. The architecture allows separate quality of service engineering of the VPNs and the backbone.

13. Scalability

The VR VPN architecture shares the scalability advantages of other provider-provisioned VPN architectures. Only the PEs need to handle the VPN type information. The internal backbone routers (the P routers) are not VPN aware. Virtual routers allow multiple private CE-based networks to connect to a single PE.

One advantage of the ability to contain the VPN address space and VPN routing and forwarding capabilities within the virtual router entity is the possibility to distribute PE system resources on a per-VPN basis. Indeed, as an example, different scheduling mechanisms can be used for processing each VPN activity within the PE. This type of per-VPN resource management contributes to establishing a wide range of priority schemes among the VPNs within the PE, and contributes to the ability to support a wide range of VPN scales (high traffic and/or many member sites) in the VR architecture.

As noted earlier in this document, the use of the "backbone VR" provides significant scalability advantages, allowing very straightforward multiplexing of multiple VPNs across PE-PE tunnels or connections. The individual VPNs and their VRs need not participate in the discovery and maintenance of the topology of the backbone network, essentially seeing the backbone as a single large router to which they are all connected.

14. Interoperability

There are several ways of implementing the VR architecture. Although this document does not guarantee interoperability by specifying explicit requirements, the nature of the VR approach makes it likely that interoperability can be achieved among various VR implementations, just as interoperability of VRs with normal routers is straightforward.

The VR architecture doesn't specify any protocol. Rather it defines how one physical device (a router) can support multiple logical devices (virtual routers), where each virtual router is interconnected with other normal routers or virtual routers either via physical links or by tunnels (of pretty much any kind defined by other IETF standards). Thus interoperation between a virtual router and other routers (whether virtual or logical) makes use of the same IETF standards (and proposed standards) that are used to allow interoperation between normal routers.

VR implementations will either interoperate or not interoperate depending upon their implementation of other IETF standards. In particular, for two VR implementations to interoperate, they simply need a common data link or tunnel technology to link them together and common routing protocols. If they are using MPLS, then MPLS has to work over the common link or tunnels, and they need compatible MPLS signaling.

Given that normal physical links terminating on a physical router can be assigned to a specific VR, and given that routers can in general interoperate at the data link level, then finding a common link to hook VRs with VRs, or VRs with routers, is straightforward. Also, there is a short list of relatively well known tunneling techniques (MPLS, IPsec, IP-in-IP tunnels, GRE, etc.) which are commonly used in multi-vendor networks.

Naturally VRs will also interoperate with normal physical routers in almost any case. Even the "backbone VR" mechanism can be terminated at the other end on normal routers. This would involve one remote "backbone router" which would be the last node of a tunnel, inside of which multiple other tunnels are multiplexed, with these multiple tunnels terminating on multiple real routers, each of which terminates a tunnel which it treats as a normal link. These real routers can run routing protocols over the tunnel, with the VR on the other end of the tunnel being the peer router for the routing protocol.

15. Security Considerations

From a security viewpoint, the virtual router VPN architecture is an extension of existing router architectures in which multiple VRs, each with the same mechanisms of a physical router, can be configured in a PE device. Thus the VRs inherit the security concerns and security capabilities of individual routers, which are largely beyond the scope of this document. Many of those elements are discussed in some detail in the routing protocol security document [RP-SEC]. The provider-provisioned VPN framework in general also has a number of security considerations due to the shared infrastructure, which are addressed in the PPVPN security framework document [RFC-4111]. This section addresses security considerations which are more specific to the VR architecture.

The VR architecture provides an inherently high level of security against many types of attacks against individual VPNs, since individual VPN routing information does not propagate throughout the backbone network. The VRs usually do not exchange routing information directly through the backbone routing protocol, but through tunnels, through layer 2 connections, or (in the case of backbone VRs supporting ordinary VRs) through communication internal to the PE device. The tunnels can use the security mechanisms available to the backbone network, such as IPsec in an IP backbone network, to protect both the routing exchange and the VPN data.

Since the VR architecture concentrates multiple VRs in a single device, there is a potential for disruption of one VR to affect other VRs within the same device. It is highly desirable for implementations to provide mechanisms to isolate problems to a single VR within a PE, or to a single VPN.

If physical or logical network links are shared among VRs, it is possible that bandwidth depletion attacks against one VPN may affect other VPNs. VR implementations should provide mechanisms to mitigate the effect of excessive traffic being received for individual VPNs on shared links. In addition, VR implementations should provide mechanisms to control the bandwidth usage on a per-VPN basis for traffic transmitted by the PE device. The VPN service provider should ensure that both access networks and backbone networks are engineered to reduce the likelihood of this kind of attack.

Since the backbone VR(s) may carry traffic from multiple VPNs, the implementation of backbone VR mechanisms should provide redundancy mechanisms. They should provide protection against hostile or inadvertent resource exhaustion attacks, originating either within or outside the VPNs.

If the auto-discovery mechanism used in determining membership to the VPN is subverted, it could potentially be possible for an attacker to join a VPN without authorization. Likewise, if the VPN-ID of a VR is erroneously configured, a VPN site could potentially be joined to the wrong VPN. These issues can both be addressed by the use of tunnel mechanisms between VRs which include other means of authentication, such as a shared secret. Other proposals for VPN membership verification, such as [MPLSVPN-AUTH], offer mechanisms which may also be useful to mitigate this potential issue.

Various levels of data, routing and configuration security can be implemented in the VR architecture. Any existing security-related mechanisms supported by existing routing protocols (e.g. authentication) can be used unmodified. If IPsec tunneling is used as the tunneling protocol, then both the control and data traffic that travels over the tunnel can be secured; so that routing specific security enhancements are not needed. Any private routing, forwarding and addressing manipulation is done within the virtual router context. Direct layer-2 connections (ATM, FR), or specific tunneling mechanisms can also provide various levels of data security.

16. IANA Considerations

This document has no actions for IANA.

17. Normative References

[RFC-1812] F. Baker, Ed., "Requirements for IP Version 4 Routers", RFC 1812, June 1995.

[RFC-2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, BCP 14, March 1997.

[RFC-3031] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC-3035] B. Davie, J. Lawrence, K. McCloghrie, E. Rosen, G. Swallow, Y. Rekhter, P. Doolan, "MPLS using LDP and ATM VC Switching", RFC 3035, January 2001.

[RFC-3809] Nagarajan, A., Ed., "Generic Requirements for Provider Provisioned Virtual Private Networks (PPVPN)", RFC 3809, June 2004.

[RFC-4031] Carugi, M., Ed. and D. McDysan, Ed., "Service Requirements for Layer 3 Provider Provisioned Virtual Private

Networks (PPVPNs)", RFC 4031, April 2005.

[RFC-4110] R. Callon, M. Suzuki, "A Framework for Layer 3 Provider-Provisioned Virtual Private Networks (PPVPNs)", RFC 4110, July 2005.

[RFC-4111] Fang, L., Ed., "Security Framework for Provider Provisioned Virtual Private Networks", RFC 4111, July 2005.

[RFC-4271] Y. Rekhter, Ed., T. Li, Ed., S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC-4301] Kent, S., Seo, K., "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[RFC-4364] Rosen, E., et al, "BGP/MPLS IP Virtual Private networks (VPNs)", RFC 4364, February, 2006.

[STD-54] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

18. Informative References

[MPLSVPN-AUTH] Behringer, M., Guichard, J., Marques, P. R., "Layer-3 VPN Import/Export Verification" (draft-ietf-l3vpn-vpn-verification-00.txt), work in progress.

[RFC-1195] Callon, R., "OSI IS-IS for IP and Dual Environment," RFC 1195, December 1990.

[RFC-2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

[RFC-2661] Townsley, W., et al, "Layer Two Tunneling Protocol L2TP", RFC 2661, August 1999.

[RFC-2685] Fox, B., et al, "Virtual Private Networks Identifier", RFC 2685, September 1999.

[RFC-2784] Hanks, S., Li, T., Farinacci, D. and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RFC-2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, September 2000.

[RFC-3787] Parker, J., Ed., "Interoperable IP Networks using IS-IS", RFC 3787, May 2004

[RP-SEC] Barbir, A., Murphy, S., and Yang, Y., "Generic Threats to Routing Protocols" (draft-ietf-rpsec-routing-threats-07.txt), work in progress.

[STD-51] Simpson, W., "The Point-to-Point Protocol (PPP)", STD 51, RFC 1661, July 1994.

[VPN-BGP] Ould-Brahim, H., et al., "Using BGP as an Auto-Discovery Mechanism for Layer-3 and Layer-2 VPNs" (draft-ietf-l3vpn-bgpvpn-auto-06.txt), work in progress.

19. Contributors and Acknowledgments

The authors would like to acknowledge the active contributions of the following individuals, all of whom would be included as authors if allowed by IETF processes:
Rainer Bach
Rick Bubenik
Luyuan Fang
Isaac Negusse
Chandru Sargor
Timon Sloan
Dr. Christian Weber
Gregory Wright
Abraham Young
Jieyun Jessica Yu

The authors would also like to acknowledge the following individuals for their helpful comments and suggestions: Peter Ashwood-Smith, Ron Bonica, Ross Callon, David Drynan, Mark Duffy, Don Fedyk, David Hudson, Bilel Jamoussi, Ahmad Khalid, Scott Larrigan, Keerti Melkote, Martin Pepin, Benson Schliesser, Jerry Sydir, and Ru Wadasinghe.

20. Authors' Addresses

(change /at/ to @ for email)

Paul Knight (Editor)
Nortel Networks
600 Technology Park Drive
Billerica, MA 01821 USA
paul.knight/at/nortel.com
Phone: +1 (978) 288 6414

Hamid Ould-Brahim
Nortel Networks
P O Box 3511 Station C
Ottawa, ON K1Y 4H7 Canada
Phone: +1 (613) 765 3418
Email: hbrahim/at/nortel.com

Bryan Gleeson
Nokia
313 Fairchild Drive
Mountain View CA 94043 USA

bryan.gleeson/at/nokia.com

Intellectual Property Statement

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.