C. Wang
M. Beadles
A. Khetan
SmartPipes

Internet Draft
Document: draft-wang-cevpn-routing-00.txt
Expires: April 2002

October 2001

Routing Support in CE-based IPsec VPNs

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at
http://www.ietf.org/shadow.html.

Abstract

When IPsec tunneling is used to provide VPN connection, it is important to support site-to-site customer network routing. This document describes the scenario, requirement, and several potential solutions to provide customer network (intranet) routing support for IPsec-based VPNs.

1.0 Introduction

IPsec-based VPN has been one of the options for service providers to offer VPN services. IPsec based VPN can achieve a high level of data security and dramatically reduce cost in comparison with private leased lines.

When IPsec tunnel is used to build a large-scale network, user data as well as network control data of the customer network are carried through the tunnels linking various sites. Although the VPN networks may vary from hub-and-spoke to a full mesh in topology, the basic requirement of providing and maintaining reliable site-to-site data connectivity remains the same.

IPsec is a layer 3 tunneling protocol, which operates purely on IP layer. The IETF IPsec Working Group specifies the IPsec standards [RFC2401, RFC 2402, RFC2406, RFC2407, RFC2408, and RFC2409]. The interaction between IPsec and layer 3 routing has not been specified. Depending on individual implementation, difficulty may arise when an IPsec user wants to support robust routing across IPsec VPNs sites.

This draft intends to identify and analyze the interactions between IPsec and IP routing, when IPsec tunneling is used to build CE-based IPsec VPN. Further, several potential solutions to support IPsec VPN routing are proposed using existing protocols.

1.1 Terminology used in this draft

Customer Edge (CE)
: This is the customer premise equipment, which provides the connection between a customer network and a service provider network. In the case of CE-based VPN, a CE device also terminates VPN tunnels.

CE based VPN
: The term "CE-based VPN" refers to an approach in which knowledge of the customer network is limited to customer premise equipment. The service provider takes on the task of managing and provisioning the Customer Edge equipment, on behalf of its customers. In CE-based VPNs, the customer network is connected by tunnels set up between CE devices. In the case of IPsec VPN, the VPN tunnels use IP encapsulations to sent traffic over the service provider IP networks.

Customer IP Network

A customer network is a network, which a service provider customer manages itself. However, different parts of a customer network may be inter-connected through the service provider network via VPN services.

Provider Edge (PE)

This is the provider edge (PE) equipment, which is attached to the customer network, usually a CE device. For a layer 3 connection, the PE device is an IP router.

Service Provider IP Network

A service provider network is a network administered by a single service provider. A service provider network is the core network which links customer networks on its edge and provide Internet connections to its customer networks.

VPN tunnel

A VPN tunnel is a logical link by encapsulating packets and transmitting the encapsulated packets using the encapsulating header between two CE devices. In the case of IPsec VPN, the encapsulation carried out by using IPsec.

2.0 Scenarios of Running IPsec VPN

An IPsec VPN consists of a number of sites securely inter-connected through IPsec tunnels. Each site consists of the CE router, the customer network behind the CE router, and the link to the PE router to connect to the service provider network. Through the service provider networkÆs IP infrastructure, CE routers are inter-connected via IPsec tunnels to reach each other.

The CE router sits at the boundary between the service provider network and the customer networks. The CE router may have one or more public routable addresses and is linked to the PE router.

An IPsec tunnel links two sites together through the service providerÆs IP network. The tunnel end points use the public address of the CE routers. User data of the customer network are encapsulated by the IPsec header and tunneled through the service provider network. These tunneled packets are routed normally through

the service provider network as other IP packets between CE routers.

To link the various CE sites, different VPN connection topology can be used. For a small VPN with a few sites, simple topologies such as a hub-and-spoke or a full mesh can be used. For a large VPN, a layered, hierarchical approach may be taken.

On the customer network side, a CE router connects to internal networks of an enterprise, where one or more subnets can reside. Many times, the CE router may interact with another internal router.

The CE device could be an integrated device providing both routing and IPsec tunnel termination. Sometimes, a dedicated VPN terminator may be used. Implementations in which the VPN terminator resides on a firewall are also very common. For the sake of simplicity, we assume that the CE router is an integrated device and terminates tunnels.

3.0 Routing Requirement

IPsec is a point-to-point tunneling protocol linking two sites across a Service ProviderÆs IP network. The VPN traffic between the CE routers is routed through the service provider network like normal IP packets. The traffic traversing the IPSec tunnels (tunneled traffic) runs in the customer network space. For the customer enterprise network (intranet) connected by VPN tunnels, route information needs to be updated and distributed dynamically among all sites (CE router plus networks behind it) participating in the VPN, in order to support network connectivity through theses IPsec tunnels.

3.1 Customer network routing û Intranet

In the intranet case all of the sites to be interconnected belong to the same administration (for example, the same company). The options for routing within a single customer network include:

- A single IGP area (using OSPF, IS-IS, or RIP)
- Multiple areas within a single IGP
- A separate IGP within each site, with routes redistributed either statically or via BGP. The use of BGP is applicable in scenarios when the IPSec VPN is used to configure the backbone of a very large enterprise with each site running its own IGP.

If a site consists of only a few LAN segments that are all attached to the CE router, then the CE router is aware of all the routes at

that site. If a site consists of multiple networks that reside one or more hops away from that CE router, the CE router could

participate in a routing protocol with other routers at that site to learn about these other networks. In either case, the IPSec VPNs must be able to transport route advertisements learnt by the CE router to all other CE routers participating in that VPN.

In a VPN network where a VPN site can come and go, the site-to-site routing also needs to responds to site changes, including site addition and deletion.

3.2 Customer network routing û Extranet

In the extranet case the sites to be interconnected belong to multiple different administrations. In this case IGPs (such as OSPF, IS-IS, or RIP) are normally not used across the sites between organizations. Either static routes or BGP may be used between sites to communicate reach-ability information. Since extranets are configured between communities of interests to share access to specific resources only, the reach-ability information shared between sites is limited to the advertisement of the shared resource only.

The requirements for extranet customer network routing remain the same as those identified in Section 3.1. The differences are that under the extranet case, route distribution and update MUST be limited to the shared resources only. In other words, routing support must be designed carefully so that internal network reach-ability information will not be leaked to un-intended partners. In the extranet case, it is likely that a site may join different VPNs. A CE router needs to support separate routing for each VPN. This draft limits its scope to the intranet discussion.

4.0 Limitations of Routing Support through IPsec VPN Tunnels

The current IPsec standards (RFC 2401, RFC 2402, RFC 2406, RFC 2407, RFC 2408, and RFC 2409) have not addressed the issue of providing routing support through IPsec tunneling. The IPsec tunnel merely provides a point-to-point connection. Each end of tunnel may attach to a single host or a complex network. Packets enter the tunnel by IPsec encapsulation and leave the tunnel by IPsec header de-capsulation. At both ends, tunnel access-list controls what can be sent into the tunnel and what can be received from the tunnel. In terms of how these packets are first routed to the IPsec gateway and then into the tunnel and how the de-capsulated packets are then routed to final destination are left to implementations. In other

words, IPsec tunneling is specified independently of any packet routing. In reality, from source to destination, packets need to be routed to the local IPsec device, tunneled to the remote IPsec device, and then routed again to final destination. Tunneling is only one segment that a packet is traversing. To deliver packets to their destination, VPN tunneling and routing need to work together.

To support site-to-site connection through IPsec tunnels, routing information must be exchanged between sites.

The current IPsec standards donÆt require IPsec tunnel to be on a logical interface that packets can be routed to. A non-interface IPsec tunnel end point canÆt participate in packet routing and forwarding. In that case, the IPsec tunnel end point may attach to a physical interface.

The other practical limitation for running routing protocols through IPsec tunnel is that the existing IPsec standards donÆt support multi-cast and broadcast traffic. IPsec tunneling canÆt carry any routing protocols using multicast or broadcast natively.

To support CE-based IPsec VPN, this draft addresses the gap between running tunneling and routing and discusses several potential solutions to support IPsec VPN routing.

5.0 Routing Supporting for Running IPsec VPNs

Depending on how IPsec is implemented several options are available to support routing in IPsec-based VPN.

5.1 IPsec Implemented as a Virtual Interface

IPsec tunnel end point can be implemented as a virtual interface. By doing that, the IPsec tunnel can participate in the CE routerÆs routing table. Having a virtual interface allows assigning of IP addresses separate from the physical interface. This allows establishment of IGP routing peer relationships across the tunnel to other CE routers.

However, since IPSec doesnÆt support non-unicast traffic, routing protocols using multicast or broadcast will not be able to get tunneled. This could be circumvented by configuring CE router with routing peer information as opposed to having the peers learnt automatically via broadcast or multicast messages. When extra-net VPN is involved, establishment of BGP routing peer between CEs can be supported successfully, since BGP only uses uni-cast traffic.

5.2 IPsec not implemented as a Virtual Interface

When IPsec is not implemented as a virtual interface, it cannot participate in the site-to-site routing directly. To support routing across the tunnel, special treatments are needed.

5.2.1 Using Managed Route Update

Managed route update can be used to distribute route update across VPN sites. This approach requires each CE device remain in contact with a centralized management center.

With the managed route update scheme, each site may still run its local routing protocol to maintain a dynamic routed local network. The remote site reach-ability information is collected and distributed through a management solution.

The management center is similar to a route server, which serves all the CE routers participating in a dedicated VPN. The management center monitors and maintains the active membership of a VPN. When a new VPN member joins the group, its reach-ability information is delivered to its VPN peers via management center, after all the corresponding IPsec tunnels have been established. When an active member has left the VPN group, the management center needs to update all affected CE routers so that the related route entries can be updated or deleted. In addition, the management center is also responsible to distribute route update of a connected site to all other connected sites. A CE router is required to inform the management center when a local route update has happened. The management center is then responsible for obtaining the new route update and injecting the new reach-ability information to the related sites. For example, when a CE router in a fully meshed VPN network has a new subnet route added, the CE routerÆs peers need to be updated with the new reach-ability information.

Effectively, the management center manages each CE routerÆs site-to-site static routing. The route update happens when either the VPN membership changes or when a siteÆs local routing table has an update (due to local site routing changes).

Since the management center usually has the capability of monitoring the status of site-to-site VPN tunnel status, it is able to respond quickly when a tunnel connection is lost. The management center can update reach-ability information and provide alternative routing path for those affected sites. The response time for this approach is usually much quicker than that a normal dynamic routing protocol can offer.

It is worth pointing out that the managed route update may have a

scalability issue. Managing a large VPN with many sites may require complex management software to manage these route updates. However,

as with the case of routing, a complex VPN may be designed using a layered approach. In that case, managing route with each sub-VPN networks is still feasible and with good scalability.

5.2.2 Using Encapsulation Protocols

An alternative to managed route update is to run dynamic routing protocols across IPsec tunnels. The IPsec tunnel end needs to appear as some kind of virtual interface in order to run generic routing protocols. In addition multicast or broadcast routing messages need to be encapsulated to become uni-cast packets.

To send routing traffic through an IPsec tunnel, an extra layer of encapsulation protocol can be used, when the encapsulation protocol is able to provide two services: 1) The encapsulation protocol can be implemented as a virtual interface; 2) The encapsulation protocol can tunnel multicast/broadcast packets.

One choice is to use GRE encapsulation when it is implemented as a virtual interface. GRE is defined in an informational RFC 1701 initially and then later turned into a standard track RFC 2784. RFC 1702 Generic Routing Encapsulation over IPv4 networks describes the GRE usage for transporting an arbitrary network layer protocol over IPv4.

GRE encapsulation adds a GRE header and an IP delivery header to the original packet. Using GRE encapsulation, multicast IP packets used by dynamic routing such as RIP can be sent across the IPsec VPN tunnel. The point-to-point IPsec VPN only sees the IP GRE packets in this case.

Other encapsulation protocols include L2TP [RFC2661] and PPTP [RFC2637], when the tunnel end is implemented as a virtual interface.

PPTP and L2TP have a built-in control channel, which is used to establish, manage, and terminate its corresponding data channel. The L2TP packets are carried by UDP protocol. The PPTP control channel is carried over TCP while a GRE tunnel is used to carry the PPTP data packets.

Both L2TP and PPTP have a keep-alive mechanism in its control channel. This can provide valuable information about the status of both the L2TP or PPTP tunnel and the IPsec tunnel that carries the L2TP or PPTP tunnel.

Compared with GRE encapsulation, L2TP or PPTP encapsulation runs a heavier protocol stack and is more complex. However, the keep-alive

feature provided by L2TP or PPTP provides useful tunnel status feedback.

Using encapsulation, routing protocol messages can be tunneled across the IPsec tunnels. Site-to-site reach-ability information is distributed and updated using the native routing protocol exchange.

The virtual interface provided by the encapsulation protocol participates directly in the local routing table. The virtual interface status reflects the underlying IPsec tunnel status. When the IPsec tunnel goes down, the virtual interface needs to report that change back to the routing engine. To be able to detect the tunnel failure, some kind of keep-alive message is needed. The routing protocol itself has a keep-alive mechanism, in the form of a hello message. When the CE discovers that it has lost connectivity to its neighbor, the routing table will be adjusted accordingly. Alternatively, the keep-alive mechanism provided by the encapsulation protocol (L2TP or PPTP) can be used to detect IPsec tunnel status. However, GRE doesnÆt have a keep-alive capability.

6.0 Interaction between IPsec Tunnel and Routing

With CE-based IPsec VPN, IPsec tunneling interacts with two routing spaces. The IPsec packets are routed through the service provider network as normal IP packets. The CE router connects to the PE router to gain Internet access. The PE router is not directly involved in the IPsec VPN, other than forwarding the IPsec packets from the CE router.

The IPsec packet payload carries the customer network IP packet. The customer network linked by site-to-site VPN connection runs its own routing. In that sense the IPsec tunneling interacts closely with the customer network space routing.

IPsec tunneling does not provide a connection-oriented link. An IPsec tunnel consists of just two encasulation and decapsulation modules at each tunnel end. An IPsec tunnelÆs life is always determined by the life of the tunnel key. An IPsec tunnel never outlives its corresponding key. Although IPsec can be configured as a static tunnel with a static key, in practice few service providers will do that due to the security risk associated with the static key. Instead dynamic keyed IPsec tunnel using IKE is the preferred way. An IPsec tunnel should use time-based re-key so that it stays active and still gets key refreshed when there is no or little traffic crossing the tunnel.

If the IPsec tunnels are implemented as packet driven and donÆt stay up all the time, dynamic routing may be affected. There is usually a latency between packet arrival and the time a tunnel has been

established. If the latency is long enough, packet loss may happen. A lost route update message due to tunnel setup latency has to be

re-transmitted. As a comparison, permanent or semi-permanent tunnels with key refresh is more stable and provides a PVC type of connection.

Existing IPsec standards lacks a keep-alive mechanism, making it difficult to determine the tunnel status. The routing table may not receive timely update even when the tunnel has gone down. To overcome this limitation until IPsec establish a keep-alive standard, some kind of tunnel status monitoring is required. When running a dynamic routing protocol across a tunnel, the routing protocolÆs keep-alive mechanism may be used indirectly to check for the tunnel status. For managed route update solution, some kind of tunnel keep-alive mechanism needs to be built into the tunnel management.

7.0 Interaction with NAT

NAT provides address translation between two realms, such as between the private address space and the public address space. For site-to-site VPN, the user data remain in the same customer network address space. NAT is usually not required, unless two sites have overlapping address space. In that case, NAT can be used to solve the address-overlapping problem.

The only scenario when NAT applies is when the CE router serves a split stack function and provides both VPN connection and direct Internet access. In this case, private address space traffic is separated into two streams, one part for site-to-site private traffic and one part for direct Internet traffic. In this case CE performs NAT function for the direct Internet traffic.

8.0 Security Issues

One of the most important features of using IPsec-based VPN is security. To make sure that IPsec VPN is implemented securely, a careful analysis of security issues is required.

1) Encapsulation
Section 5.2.2 discusses using an extra packet encapsulation. If site-to-site traffic is encapsulated (such as by GRE) before entering an IPsec tunnel, the IPsec tunnel traffic filtering loses its inspection capability. Thus unwanted traffic, which can be blocked by an IPsec tunnel filter, may be tunneled through under the cover of encapsulation. It is important to enforce the same set of IPsec traffic filters to the user traffic before it gets

encapsulated and after it gets de-capsulated, when such
encapsulation is used inside an IPsec tunnel.

2) Connection to the Management Center
A CE device usually needs to connect to the management center in order to get provisioned, managed, and monitored. It is important to secure the communication so that an adversary canÆt use this management channel to compromise the security of the VPN network by gaining privileged access to the CE device.

3) Injection of Route
For managed route update, site-to-site route information is injected from the management center. False route injection can severely disrupt network service. Therefore the route distribution and update to each CE must be securely protected. If the management channel used to make the update is secure, then no addition precaution may be needed. Otherwise, each route update must be at least authenticated at least.

With dynamic routing, the route distribution is done by the routing protocols. The site-to-site VPN infrastructure can be used to protect the route messages.

9. Summary for Sub-IP Area

9.1. Summary

The PPVPN WG currently supports three types of VPNs: Provider Provisioned Network Based Layer 3 VPNs, Provider Provisioned Layer 2 VPNs and Provider Provisioned CE-based VPNs. This draft discusses the issue of supporting routing for CE-based IPsec VPN.

9.2. Where does it fit in the Picture of the Sub-IP Work

This work fits squarely in the PPVPN box.

9.3. Why is it Targeted at this WG

This draft describes the scenario, requirement, and several potential solutions to provide private space routing support for CE-based IPsec-based VPNs.

Under the current PPVPN WG charter, Provider Provisioned CE-based VPNs fits the scope of the WG, as stated from the following charter extract:
"This working group is responsible for defining and specifying a limited number of sets of solutions for supporting provider-provisioned virtual private networks (PPVPNs). The work effort will

include the development of a framework document, a service requirements document and several individual technical approach

documents that group technologies together to specify specific VPN service offerings. The framework will define the common components and pieces that are needed to build and deploy a PPVPN. Deployment scenarios will include provider-managed VPN components located on customer premises."

9.4. Justification

This draft is justified since it targets the routing issue of CE-based VPNs, which are identified as a specific type of PPVPNs both in the WG charter and the general framework I-D. CE-based VPN has its own characteristics and operation requirements, among which routing support is one.

10. Reference

[FRAMEWORK] Callon, R. et al., A Framework for Provider Provisioned Virtual Private Networks, draft-ietf-ppvpn-framework-00.txt, Work in progress

[CEFRAMEWORK] Jeremy De Clercq, Olivier Paridaens, Mahadevan Iyer, Andrew Krywaniuk , A Framework for Provider Provisioned CE-based Virtual Private Networks using IPsec, draft-ietf-ppvpn-ce-based-00.txt, Work in progress

[RFC1702] S. Hanks, T. Li, D. Farinacci, P. Traina Generic Routing Encapsulation over IPv4 networks, October 1994

[RFC2401] S. Kent, R. Atkinson Security Architecture for the Internet Protocol, November 1998

[RFC2402] S. Kent, R. Atkinson, IP Authentication Header, November 1998

[RFC2406] Kent, S., and R. Atkinson, "IP Encapsulating Security Payload (ESP)", RFC 2406, November 1998.

[RFC2407] D. Piper, The Internet IP Security Domain of Interpretation for ISAKMP, November 1998

[RFC2408] D. Maughan, M. Schertler, M. Schneider, J. Turner, Internet Security Association and Key Management Protocol (ISAKMP), November 1998

[RFC2409] D. Harkins, D. Carrel, The Internet Key Exchange (IKE),

[RFC2637] K. Hamzeh, G. Pall, W. Verthein, J. Taarud, W. Little, G. Zorn, Point-to-Point Tunneling Protocol (PPTP), July 1999

[RFC2661] Townsley W., et al., "Layer Two Tunneling Layer Two Tunneling Protocol (L2TP)", August 1999.

[RFC2784] D. Farinacci, T. Li, S. Hanks, D. Meyer, P. Traina Generic Routing Encapsulation (GRE), March 2000

Author's Addresses

Cliff Wang
SmartPipes
565 Metro Place South
Dublin, OH 43017, USA
Phone: 1-614-923-6241
Email: cwang@smartpipes.com

Mark Beadles
SmartPipes
565 Metro Place South
Dublin, OH 43017 USA
Phone: 1-614-923-5657
Email: mbeadles@smartpipes.com

Archana Khetan
SmartPipes
555 Twin Dolphin Drive
Suite 650
Redwood city, CA 94065 USA
Phone: 1-650-232-172
Email: akhetan@smartpipes.com