Network Working Group
INTERNET DRAFT
<draft-ietf-l3vpn-ce-based-03.txt>

Jeremy De Clercq
Olivier Paridaens
Alcatel
Andrew Krywaniuk
Cliff Wang

December 2005
Expires June, 2006

An Architecture for Provider Provisioned CE-based Virtual Private Networks using IPsec

<draft-ietf-l3vpn-ce-based-03.txt>

Status of this Memo

Abstract

This informational document describes procedures for a Service Provider to offer Virtual Private Network Services to its customers by provisioning the CE devices on behalf of the customer. The IPsec technology is used to protect the customer traffic.

Table of Contents

1. Introduction

The L3VPN framework document [FRAMEWORK] identifies three basic provider provisioned VPN types : Provider Provisioned Network Based (also termed PE-based) Layer 3 VPNs, Provider Provisioned Layer 2 VPNs and Provider Provisioned CE-based VPNs.

This document describes a method enabling a Service Provider to offer IP VPN services to its customers by provisioning the CE devices on behalf of the customer (Provider Provisioned CE-based VPNs). This document describes which parameters need to be provisioned, but not which protocol to use for the provisioning. As such, this document is of informational nature and does not specify a protocol specification from which one can achieve interoperability.

For a CE-based VPN to be set up under the SP's control, the VPN customer informs the Service Provider of which sites (identified by a set of CE devices) should become part of the considered VPN and what the requested topology of the VPN should look like. The SP then configures and updates its VPN database, and then provisions and manages the Customer's VPN.

The model proposed in this document uses the IPsec protocol suite for the purpose of securely tunneling the customer VPN traffic and the inter-site reachability information distribution.

2. Reference Model

The reference model upon which the mechanisms and procedures described in this document are based, is taken from the CE-based VPN reference model described in [FRAMEWORK]. The most important aspects of that framework model and the restrictions that are relevant to this document are described in this section.

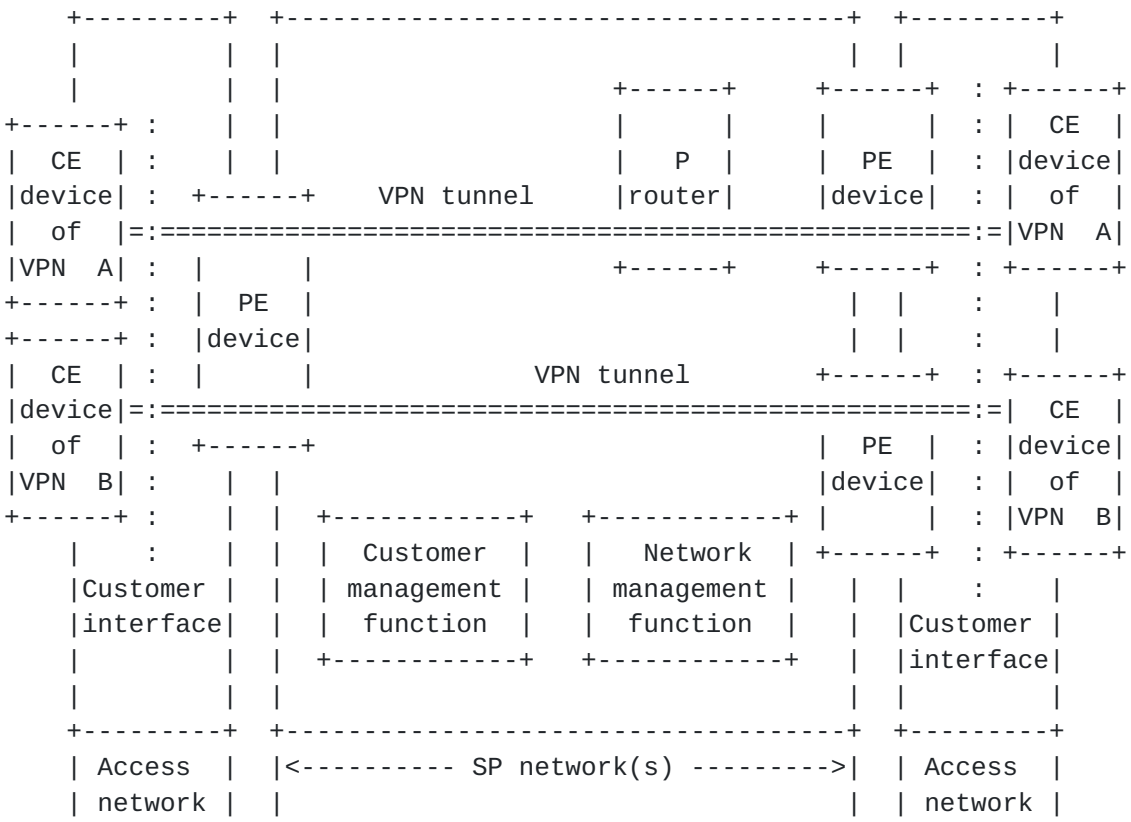

Figure 1: Reference model for provider provisioned CE-based VPNs

2.1 Entities in the reference model and Terminology

o Customer Edge (CE) device

In the context of this solution, a CE device is a router located at the edge of a customer site, that has IP connectivity with a SP's PE device (not necessarily Internet connectivity). A CE device maintains one or more VPN tunnel endpoints. The VPN-specific functions in the CE device are provisioned by the SP.

Note that other functions that are normally applied by the PE router may need to be performed by the CE device in this context (e.g. NAT functionality, QoS classification, etc.). These functions may be managed by the SP or alternatively be managed by the VPN customer, depending on the applicable service contract.

The CE device may also provide general (non VPN-oriented) Internet connectivity for the customer network. Such connectivity may be achieved via the SP's PE router that provides the VPN connectivity or some other router (of the same or another SP). In such a situation, the CE device must be able to distinguish between traffic to be sent through a VPN and traffic to be sent outside any VPN. Section 6 of this document discusses this in more details.

CE devices in a CE-based VPN model differ from CE devices in a PE-based VPN model in that they need to support VPN-specific functions. With CE-based PPVPNs, the VPN awareness is pushed even further towards the edges of the provider networks.

o Provider Edge (PE) router

In the context of Provider Provisioned CE-based VPNs, a PE router is a router, located at the edge of the Service Provider's network, that does not have any VPN-specific functionality. A PE router is attached via an access connection to one or more CE devices, and offers possibly limited or restricted IP connectivity, or possibly full Internet connectivity, over the access connections to these CE devices.

o SP network

A SP network is a network administrated by a single service provider. In the context of PP CE-based VPNs, the SP who owns the SP network can also be the VPN provider (managing the CE devices). This can lead to operational advantages (e.g. for offering QoS). Alternatively, the SP owning the SP network may be an ISP offering Internet connectivity, while another entity may provision the VPN service. This configuration allows for inter-SP and Internet-wide VPN scenarios.

o Access connection

An access connection represents a layer 2 connectivity between a CE device and a PE router. This includes dedicated physical circuits, logical circuits (such as Frame Relay and ATM), IP tunnels (e.g., using IPsec, L2TP) and shared medium access (such as Ethernet-based access). In the context of provider provisioned

CE-based VPNs, the CE device and the PE router have layer 3 connectivity over the Access Connection.

o VPN tunnel

A VPN tunnel is a logical link between two entities which is created by encapsulating packets within an encapsulating header for purpose of transmission between those two entities for support of VPNs. In the context of provider provisioned CE-based VPNs, a VPN tunnel is an IP tunnel (e.g., using IPsec [IPSEC], L2TP [L2TP], GRE [GRE], IP-in-IP [IPinIP]) between two CE devices over the SP's network. In the context of this document, a VPN tunnel is achieved using IPsec in tunnel mode or via an IP-in-IP tunnel protected by IPsec in transport mode between two CE devices.

o Security Association (SA)

Throughout this document, the acronym SA will be used to denote an IPsec Security Association.

2.2 IP connectivity between CE and PE devices

CE devices operating in a PP CE-based VPN will operate in two independent IP routing spaces.

The first routing space is the VPN routing space. Hosts and routers within the VPN will use IP addresses that belong to this VPN routing space. The CE router will participate in this VPN routing space, and will create VPN tunnels (virtual links) to be used as virtual interfaces by this VPN routing space.

The second routing space is the SP's routing space. Every CE device that belongs to a PP CE-based VPN is identified by an IP address that is routable in the SP's network. This IP address may be a global IP address or a private IP address. The CE device is reachable from the SP's core network via this IP address.

In order to easily differentiate between these two routing spaces, this document uses the following convention: IP addresses belonging to the VPN's routing realm will be followed by a 'v' between brackets: address (v); IP addresses belonging to the service provider's routable space will be followed by a 's' between brackets: address (s).

These two routing spaces may use overlapping address spaces and thus need to be kept separate in the CE devices. The way this is done is largely implementation dependent. This may be by using two separate sets of (virtual) routing and forwarding tables (figure 2). These

routing tables may then run independent routing protocols.

Only the CE's IP address (s) needs to be reachable in the provider's core network. This means that this approach requires only one IP address (s) per CE device to be injected in the core network. A CE device should not inject other routes into the SP's network (one exception is for Internet Access scenarios, which are discussed in section 6). In many cases, this CE's IP address (s) may be an IP address assigned by the SP who owns the core network. As such, aggregate routes can be distributed by the PE devices into the core network.

The CE device and the PE device may be routing protocol peers in the SP's routing space. Alternatively, a default route (s) (towards the PE) may be statically configured in the SP's routing space on the CE device, and the CE device's IP address (s) statically configured on the PE. The CE device should not inject SP's routes (s) towards the other routers within its VPN site (except in the context of the Internet Access scenarios described in section 6).

Note that, when the CE device is attached to only one PE device, via only one (sub-)interface, the CE's implementation can be fairly straightforward (see figure 3). With regards to the SP routing space, the CE device then acts as a host, having only one outgoing interface. The source IP address (s) (of the _outer_ IP header) of all packets leaving the CE device will always be the CE's identifier, and the IP next hop will always be the PE device to which it is attached. On the CE, no routing decisions need to be performed in the provider's routing space and only one forwarding action is possible.

The following figures give an overview of the routing spaces in the CE device. Note that this description is merely an example and is not meant to specify a particular implementation.

Section 5 describes the end-to-end processing of customer data-packets in more details.

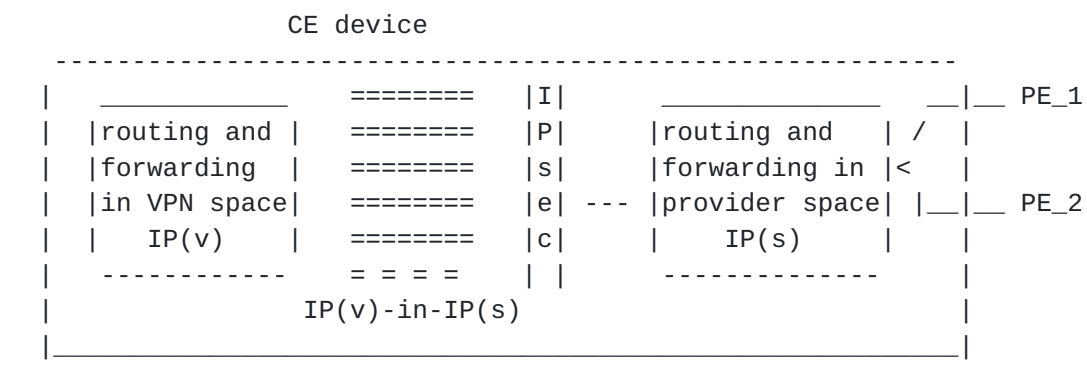

figure (2): routing spaces in the CE device
(''VPN space'' = customer ''private'' space)

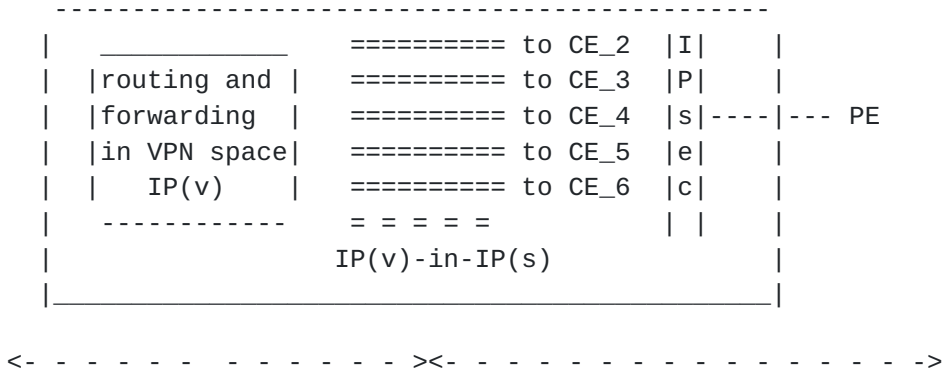

figure (3): the CE is connected to only one PE device

Note that there are no routing protocols operating in both routing spaces simultaneously. Packets can only go from one routing space to the other routing space via either (IP-in-IP) tunneling or after firewall and possibly NAT processing (as described in section 6).

This approach enables the CE devices to reach each other via tunnels over the SP's network, but does not prevent the interconnection of CE devices via so-called "backdoor routes". CE devices belonging to the same VPN can be interconnected via "backdoor routes". If "backdoor routes" are present in a certain VPN, the VPN's routing protocol metrics will dictate which routes will be used as the preferred routes for certain destinations.

2.3 Assumed Service Provider's infrastructure

The service provider maintains a secured VPN database (e.g. on a centralized server). One such VPN database may be used for the

provisioning of many VPNs. As the number of VPNs to be provisioned grows, other servers may be deployed. As such, the scalability of no single device is dependent on the total number of VPNs.

In order to provide a reliable service, the SP may choose to deploy backup VPN database servers that it keeps synchronized with the primary server.

The Service Provider's VPN management infrastructure needs to have a secure provisioning channel to every attached CE device. This secure provisioning channel will be used to exchange VPN-specific configuration information between the SP's VPN database and the CE devices.

Note that the management access to the CE devices may be in-band (i.e. using the same access connections as the VPN data traffic) or alternatively the management access may be out-of-band, for example using a dial-up connection.

Note that this document does not prescribe one particular protocol for this provisioning channel. Some examples are: SOAP/XML/HTTP/TLS, CLI/Telnet/SSH, an IPsec-protected remote configuration protocol, etc.

As the SP will be responsible for provisioning the secure tunnels between the CE devices, it needs to deploy a key management system.

3. Configuring the CE-based VPN

As was noted before, this document does not describe the protocol to use as a remote management protocol to provision CE devices. It does however describe with which information CE devices need to be pre-provisioned, and which parameters need to be configurable via this management protocol by the Service Provider.

3.1 Initializing the VPN database

As a first step in the VPN configuration process, the Service Provider configures its VPN database with a new VPN entry and with the IP addresses (s) or identifiers of the CE devices belonging to the VPN, and with a description of the VPN's topology.

For every CE device, the following information is configured and maintained in the VPN database:

- the security information that is necessary for the secure remote management protocol. This information should allow for mutual authentication between CE and SP's VPN server, and for encryption

of the management data. The details of this information will depend on the particular protocol (stack) used for remote management

- the security information that is necessary for the CE device to establish and maintain Security Associations with its peer CE devices belonging to the same VPN; section 3.3 defines which is the minimal set of information a CE device should be able to retrieve/receive from the SP's VPN management server.

3.2 Pre-configuration of the CE device

This document uses the term "pre-configuration" for the initial provisioning of a CE device. This pre-configuration happens before a CE is attached to a VPN (before the considered site actively belongs to the VPN). This pre-configuration can be performed by the SP before shipping the CE device to the customer's premises. Alternatively, some of the information can be auto-configured via for example DHCP or the SP can pre-provision the CE device manually at the customer's premises. Another possibility is for the SP to tell the customer how to pre-provision its CE device. Finally other scenarios such as remote management with for example secured SNMP are also possible.

Every CE device participating in a VPN needs to be pre-provisioned with the necessary configuration information that enables it to establish a secure communication path with the SP's VPN server.

The CE device must be configured with the IP address (s) of the Service Provider's VPN server or with a URL to the required CE's VPN information on the Service Provider's VPN database.

The CE device must be configured with the security information required by the SP's secure remote management protocol (stack).

And finally, the CE device must be provided with the CE's IP address (s) in the SP's space.

As mentioned before, the CE device is identified by an IP address (s) that belongs to the Service Provider's routing space. This IP address (s) may be an IP address assigned by the SP and manually configured on the CE device, together with the other (pre-) configuration information (this would require this IP address (s) to be configured as a static route on the attached PE too). Alternatively, the CE may dynamically obtain this IP address (s), using for example DHCP or IPCP over the CE-PE link. Yet another possibility is that the CE device has obtained a (global) IP address (s) from an ISP, and that the VPN customer communicates this IP address (s) to the VPN Service Provider. Note that the CE device needs to maintain this same IP

address (s) at least for the duration of its VPN membership.

Note that other information, such as timer-parameters etc. may be configurable by the SP. These parameters can be provisioned by the SP at pre-configuration time.

3.3 Fetching the VPN configuration information

The VPN service is initialized by the CE device by retrieving the VPN configuration information from the SP's VPN database using the appropriate secure remote configuration channel.

The CE device will retrieve from the SP's VPN server the information that is necessary to establish IPsec-secured tunnels with the other CE devices that belong to the same VPN (and to which it should establish a virtual VPN link - depending on the VPN topology). The SP may choose to let the CE devices authenticate the IKE negotiations between CE devices using (i) pre-shared keys or (ii) digital signatures and certificates. The IPsec implementation on the CE devices should support both modes of authentication.

(i) in case of pre-shared keys, the following information is to be retrieved from the SP's VPN server:

- a list of <peer CE IP address (s), pre-shared key, SA information, tunnel information> tuplets

  (SA information = the necessary information to negotiate a SA with the peer CE: security protocol, Diffie-Hellman group, IPsec transforms, etc. The (optional) presence of this information will overwrite possible default values in the CE)

  (tunnel information : traffic-driven tunnel or 'permanent' tunnel; tunnel mode IPsec or transport mode IPsec over an IP-in-IP encapsulation; dynamic routing trough the tunnel or not)

(ii) in case of digital signature authentication, the following information is to be retrieved from the SP's VPN server:

- a <private key, public key> pair

- a certificate for the public key

- a public key from the Certificate Authority

- a list of <peer CE IP address (s), SA information, tunnel information> tuplets

The above information is maintained on the SP's VPN server, and sent to the CE device when necessary.

3.4 Establishing the (secure) VPN tunnels/SAs

When one Site sends traffic to another Site belonging to the same VPN, these IP packets will be secured via IPsec. This means that an IPsec Security Association is needed between each pair of sites that directly exchange private VPN data with each other.

The Internet Key Exchange protocol (IKE, [IKE]) or its successor IKEv2 [IKEv2] will be used for the purpose of automatic setup of security associations between VPN sites within the same VPN. The CE devices will use the information that they have retrieved (or received) from the SP's VPN server to negotiate SAs with their peers, using IKE(v2).

The successful establishment of such a 'VPN' IPsec SA between two CEs will result in the auto-configuration of a new VPN tunnel (or virtual link) between the two considered CE devices.

As explained in section 5 of this memo, a 'VPN tunnel' is either an IP-in-IP tunnel protected by an IPsec transport mode SA or alternatively a tunnel mode IPsec SA. In both cases, the VPN tunnel is established once the protecting SA is established.

These dynamically established SAs can be set-up and maintained independently of the presence of actual inter-site user traffic, resulting in 'permanent' IPsec tunnels. These tunnels are then always available and not traffic-triggered. It is then required to frequently re-negotiate the SA (via IKE(v2)) before the IPsec timers of the connection time out. The set-up of a 'permanent' IPsec tunnel will be triggered by the configuration of a new peer CE device within the same VPN. An advantage of this method is that the IPsec tunnel is always available, and that eventual traffic does not encounter an extra delay due to the setup time of a new SA. The use of 'permanent' IPsec tunnels is recommended for CE-based site-to-site VPNs.

A CE device that first joins a VPN must retrieve the initial VPN configuration information from the SP's VPN server. Next, for 'permanent' IPsec tunnels, the considered CE subsequently establishes "VPN tunnel SAs" (using IKE) with every peer CE device listed in the VPN configuration information.

- o if the IKE negotiation is accepted and authentication succeeds, the SA is successfully established.

- o if the IKE negotiation is refused or the authentication fails,

the IKE negotiation will be stopped and the SA not be established; the CE device will then wait for a time interval larger than a certain minimum value (to be configured, depending on e.g. the responsiveness of the auto-discovery mechanism) and then try negotiating the SA with the considered peer again. After a new failure, the CE device will retry after a certain period of time (t1, to be configured). This process continues with exponential backoff of t1 until a certain limit (to be configured) upon which an alarm will trigger human interaction.

Provider provisioned CE-based IPsec VPNs as described by this document use 'permanent' IPsec Security Associations when dynamic routing through IPsec-secured tunnels is used.

Alternatively, the IPsec SA setup can be triggered by the effective (data) traffic flow going from one site to another. In this case, the arrival of data packets at the CE device, coming from within the VPN site and going to another VPN site, will be noticed by the CE's IPsec or routing database, and an IKE exchange will be initiated to set up an IPsec secured connection between both parties. Once the secure tunnel is set up, the data packets can flow from one site to the other in a secure way. When no traffic flows for a certain duration of time, the secure tunnel will be torn down again. An advantage of this method is that an IPsec tunnel is only to be maintained when there is effectively traffic flowing. A disadvantage is the extra delay introduced for the traffic during IKE signaling and the potentially large amount of data traffic that might need to be buffered or dropped during tunnel establishment for high-speed connections. Another disadvantage is the difficult interaction with the tunneled inter-site VPN routing information distribution.

Provider provisioned CE-based IPsec VPNs as described by this document could use traffic-driven IPsec SA establishment when static intra VPN inter-site routing is used (no dynamic routing through the IPsec tunnels), see section 4.3. Provider provisioned CE-based IPsec VPNs as described by this document don't use traffic-driven IPsec SA establishment when dynamic site-to-site routing through the IPsec-secured tunnels is used.

The CE configuration determines whether traffic-driven SA establishment is used or not, and whether dynamic routing through IPsec tunnels is used or not.

The procedures described in this memo can be used together with [IPSEC-DPD] that offers a mechanism to efficiently keep IPsec SAs alive.

Note that IPsec tunnels are unidirectional in nature, but that within

the application of this document, the set-up of one direction is accompanied by the set-up of the reverse direction IPsec tunnel.

This document describes two possible ways to use IPsec in CE-based VPN scenarios (see section 5): in 'transport mode' or in 'tunnel mode'. The CE configuration, IKE exchange and resulting SA's specify which mode will be used.

Note that the number of peer CE devices with which a specific CE device must have an IPsec connection to secure the data traffic, is dependent on the specific 'role' of the site in the considered VPN. A hub CE will for example have a larger number of tunnels to support than a spoke device.

3.5 Updating VPN configuration information

An important requirement for the scalability of L3VPNs is the availability of an 'auto-discovery' mechanism. Such an 'auto-discovery' mechanism should for example make sure that the addition/deletion of a VPN site to/from an existing VPN is possible by only configuring the 'new' CE device (and the SP's VPN database): the existing VPN sites should automatically 'discover' the new site in a reliable and secure manner.

The precise auto-discovery mechanism and related protocol actions will highly depend on the remote management protocol in use. As such this document does not describe a specific auto-discovery mechanism, and the principles of this document remain applicable with any auto-discovery mechanism.

The remote management protocol can operate in a 'push' model (when a new CE device is added to the VPN, the VPN server pushes the new VPN configuration information to all existing CE devices from that VPN), in a 'pull' model (CE devices periodically download their VPN configuration information from the SP's VPN server, or when receiving tunnel establishment requests from unknown CE devices), or in a combined mode (the SP's VPN server sends a 'notification' to the CE devices that tells them to update their VPN configuration information by downloading it from the VPN server). The different modes and the applied protocol dynamics will have different reliability characteristics.

3.6 Removing an existing VPN site

When the VPN customer wants to remove an existing site from a certain VPN, this customer first informs the VPN SP. The SP will then update the VPN database on the centralized server.

Different approaches can then be used. The SP can provision the considered CE device to delete its VPN information and to tear-down the IPsec SA's using IKE(v2). After completion of the IKE tear-down process, the peering CE devices should not attempt to re-establish the deleted SA. At this stage, the VPN tunnels are actually removed, and the routing protocols operating through the tunnels in the VPN's routing space will notice the topology change and react appropriately. The periodical retrieval of the VPN configuration information from the VPN database by the other CE devices will then make sure that the removed CE's information is no longer available. The discussed provisioning action can happen in the same way as the pre-provisioning action described in section 3.1, i.e. via manual configuration, via remote management or via interaction with the customer.

Alternatively, the SP will not provision the to-be-removed CE individually but the removal of the information relevant to the considered CE from the VPN database will ultimately automatically result in the removal of the CE from the VPN: peer CEs will notice the removal of the particular CE from their updated configuration file and will tear-down the appropriate SA using IKE(v2); the deletion of active SAs will effectively remove the VPN tunnels and the routing protocols running through the VPN tunnels will discover the topology changes and react accordingly. The to-be-removed CE will not be able to retrieve VPN information from the VPN database and will delete all its VPN information and try to tear-down the remaining SAs.

4. Exchanging and maintaining VPN routes

One of the requirements for PP CE-based VPNs is that dynamic routing is not only supported within individual VPN sites, but also between the different VPN sites of a specific VPN. This means that when a change in the routing information in a specific site occurs, the other sites that belong to the same VPN must be notified of that change.

This section deals with the exchange of routing information in the customer VPN's routing space (v). As depicted in figure 4, this exchange of routing information happens over the VPN tunnels and is as such transparent for the SP's network. CE devices don't leak VPN routes into the SP's network and don't leak routes from the SP's routing space into the VPN sites, unless explicitly configured to do so (as e.g. explained in section 6 of this document).

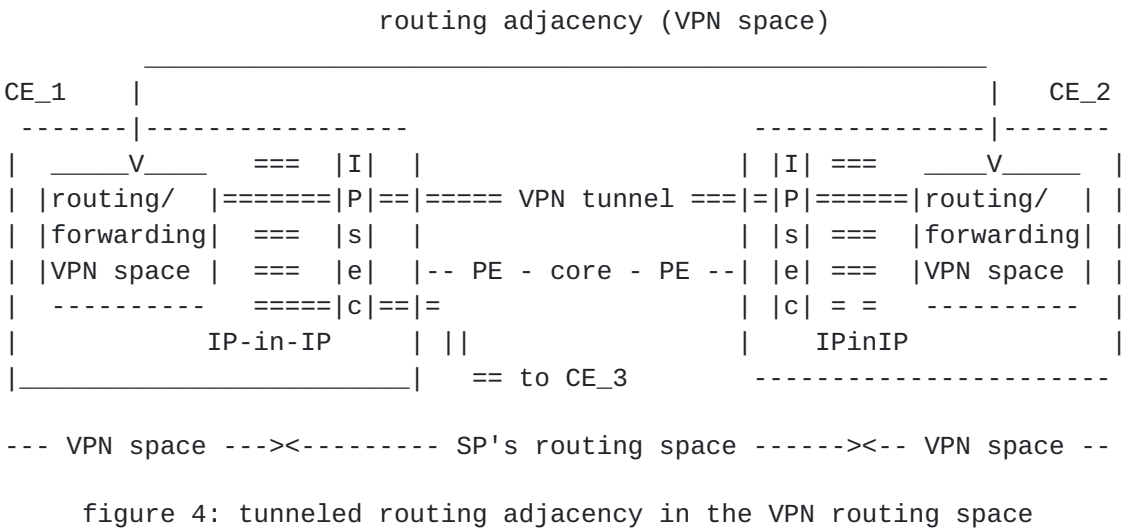

figure 4: tunneled routing adjacency in the VPN routing space

This document assumes that the routing within a VPN site is controlled by the VPN customer.

4.1 The CE device and VPN routing

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. And sometimes, "backdoor links" between routers of different sites of the same VPN exist.

In the VPN routing space (v), the CE is involved in (i) the intra-site routing, (ii) the VPN tunnel termination, and (iii) the inter-site VPN routing.

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 tunnel terminator resides on a firewall are also very common. For the sake of simplicity, we assume that the CE router is an integrated device that participates in the intra-site routing (e.g. via an IGP) and at the same time terminates VPN tunnels.

In the context of this document, the routing aspects within a VPN site (intra-site routing information distribution) are controlled by the VPN customer.

As was explained earlier, the SP's dynamic VPN discovery scheme and tunnel establishment mechanism provides the CE device with secure (virtual) links towards other CE devices in the same VPN. Whether the intra-VPN inter-site routing aspects that make use of these virtual links are managed by the customer or by the SP is dependent on the

service contract. In many situations, the SP will configure the necessary routing protocol information at pre-configuration time (see section 3.1), in close collaboration with the customer.

An important requirement for the routing protocol implementation that is configured to exchange reachability information through the inter-site tunnels, is that it must be able to autonomously deal with dynamically created new inter-site links.

4.2 IPsec and routing

IPsec is a layer 3 security protocol, which operates purely at the IP layer and which is defined by a number of IETF standards ([IPSEC], [RFC2402], [RFC2406], [RFC2407], [RFC2408], [IKE], [IKEv2], etc.). The interaction between IPsec and layer 3 routing is not always straightforward and has been described in [TOUCH]. Depending on individual implementations, difficulty may arise when an IPsec user wants to support robust routing across IPsec-interconnected VPNs sites.

4.3 Exchanging VPN routes between VPN sites

In the proposed mechanism to exchange VPN reachability information between VPN sites, routing protocol messages are tunneled through the IPsec-secured tunnels between peering sites. The CE-to-CE IPsec-secured tunnels between VPN sites are then being seen as point-to-point links by the customer networks and are interpreted as such by the routing protocol functions of the CE devices. This means that when a change in the reachability occurs in one particular site, a routing protocol (such as RIP, OSPF, etc.) will take care of the distribution of the new reachability information within the site, but also to all other sites, through the VPN tunnels that the considered CE is possibly maintaining.

As the described architecture allows for the dynamic creation of inter-site (IPsec-protected) VPN links, the routing protocol implementation(s) operating on the CE device must be able to support this.

Although very often it will be the SP's responsibility to configure the CE's routing information at pre-configuration time, the service agreement may specify that routing on the CE device falls under the customer's management.

The IPsec tunnels through which routing messages are exchanged may be implemented using IPsec tunnel mode or using IPsec transport mode (see section 5). Note that the same tunnels are used for exchanging intra-VPN inter-site routing messages as for exchanging VPN user data

traffic.

There are significant issues when a traffic-driven tunnel establishment mechanism is used at the same time as an approach whereby a routing protocol (with a keep-alive mechanism) runs on top of the VPN tunnel. In this case the delay introduced by the tunnel establishment phase could lead to a loss of routing updates and the routing protocol's keep-alive mechanism could interact with the tunnel establishment in an undesired way. For example the frequency with which dynamic routing protocols typically exchange Hello messages makes it undesirable to re-establish tunnels for each Hello packet. Therefore, when dynamic routing is used through IPsec-secured CE-to-CE tunnels, traffic-driven SA establishment should not be used.

5. Tunneling IP traffic (user data) among VPN sites

This section describes the processes that an IP packet that is sent from one VPN site to another will go through. This is depending on the way that IPsec is used. This document describes two possible ways to use IPsec in CE-based VPN implementations: IPsec in tunnel mode, and IPsec in transport mode.

An IP packet that is sent by an IP device in a certain site and destined for an IP device in another site belonging to the same VPN, will be forwarded as follows.

The device in the sending site sends an IP packet (possibly using a private address space) on its LAN network. The next hop for this destination IP address will (at some point in time) be the site's CE device (according to the routing/forwarding in the VPN site). The processing by the CE device now is dependent on the implemented mode for IPsec.

Note that the following description is not meant to specify an implementation strategy; any implementation procedure which produces the same results is acceptable.

o IPsec in transport mode (see also [TOUCH] for a detailed specification)

  When IPsec is used in transport mode in this context, the CE device first analyzes the private IP packets arriving from within the site and select the appropriate outgoing interface and required encapsulation, based on the VPN routing/forwarding information. For a destination located in another site, the outgoing interface will be a virtual interface (a VPN tunnel) and the required encapsulation will be IP-in-IP, using the considered CE's IP address (s) as the source address in the outer IP

encapsulation header and a peer CE's IP address (s) in the outer IP encapsulation header's destination address field. The CE device then processes this new IP packet to its IPsec driver.

The IPsec driver in the CE device then does the following:

- analyze the IP packets that have been IP-in-IP encapsulated and select the appropriate SA (based on the packet's outer header destination address (s)).

- authenticate and/or encrypt the private IP packet according to the (transport mode-specific) rules described in the SA and insert an appropriate IPsec header (according to IPsec in transport mode).

o IPsec in tunnel mode

When IPsec is used in tunnel mode in this context, the IPsec driver in the CE device does the following:

- analyze the private IP packets arriving from within the site and select/setup an appropriate SA with the appropriate destination CE device.

- authenticate and/or encrypt the private IP packet according to the (tunnel mode-specific) rules described in the SA, AND encapsulate the packet in an IPsec header AND encapsulate the packet in a new 'outer' IP header. This 'outer' IP header has the CE's non-private (i.e. routable in the SP's realm) IP address in the source IP address field and the destination CE's non-private (i.e. routable in the SP's realm) IP address in the destination IP address field.

The CE device then sends the IPsec packet to the PE device, and the IPsec packet will then be forwarded using 'normal' IP forwarding across the SP's network, based on the outer header's IP destination address (s), that is the destination CE's 'global' (i.e. routable in the SP's realm) IP address. The packet will be forwarded to the egress PE who will also only examine the outer IP header and send the IP(sec) packet to the destined CE device. The egress CE device will recognize itself as the destination node (the IP packet has the CE's IP address (s) in the outer IP destination address field) and process the IPsec packet to the IPsec driver that will then, based on the appropriate Security Association (identified by the packet's SPI field in the IPsec header), perform IPsec authentication and/or decryption. Dependent on whether tunnel mode or transport mode IPsec is used, the packet will be decapsulated by the IPsec driver itself

or sent to the IP-in-IP decapsulation function. The resulting (private) IP packet (v) will then be further processed in the CE's VPN IP forwarding table and send on the LAN network to the appropriate next hop router or destination IP device.

Note that IPsec tunnels might unintentionally terminate or break. For example, the CE device on one end point of an IPsec tunnel might fail, or one end point might become unreachable from the other end due to a failure of IP routing in the intermediate infrastructure. When dynamic routing is not supported through the inter-site VPN tunnels, this may have serious consequences if VPN membership and VPN routing information are not changed accordingly within the VPN. Indeed, where static routing is used the unnoticed termination of a VPN tunnel can result in the creation of black holes.

This means that a mechanism must exist to monitor the state of the VPN tunnels. When dynamic inter-site VPN routing is used, the routing protocol that runs on top of the IPsec VPN tunnels will serve that purpose. When dynamic inter-site routing is not used, alternatives are possible such as the use of an IPsec-specific keep-alive mechanism [IPSEC-DPD] or a SP-proprietary mechanism.

6. CE-based VPN and Internet

6.1 Allowing both VPN connectivity and Internet connectivity

In many VPNs, sites will need to both access the public Internet as well as to access other sites within the same VPN.

In order to achieve this, some sites within the VPN will obtain Internet Access by means of an "Internet Gateway" that is attached via one of its interfaces to an ISP's PE device. Such an Internet Gateway may for example be a firewall and may or may not need to implement network address translation functions. The ISP may be the same SP that offers the VPN service, or it may be a different SP. The PE to which the Internet Gateway is connected may be the same PE to which the CE is connected or it may be another PE.

The Internet Gateway may be a separate device, or alternatively the Internet Gateway functions may be integrated into the CE device. When the Internet Gateway functions are integrated into the CE device, the CE-PE interface used by the Internet Gateway functions may be the same or a different interface than the interface used by the VPN tunnels. In further discussions, we'll assume that the Internet Gateway is a separate device.

The service contract will define whether the Internet Gateway will be managed by the SP or by the VPN customer.

Note that when Internet Access is offered within a VPN, the address spaces used within the VPN must be non-overlapping. This means that the VPN either uses global addresses that have been assigned to the VPN customer, or private addressing in combination with NAT [NAT].

The sites that have Internet Access via an Internet Gateway will have a default route (v) pointing to their Internet Gateway and may be distributing a default route via their CE towards the other CEs of the same VPN through the VPN tunnels. This provides Internet Access for all the VPN sites. Note that other sites (that don't have their own Internet Gateway) must not distribute default routes in this scenario. A site that has distributed a default route to other sites for Internet Access should have either a default route to its Internet Gateway or Internet routes (leading to its Internet Gateway) in its forwarding table (of the VPN routing space).

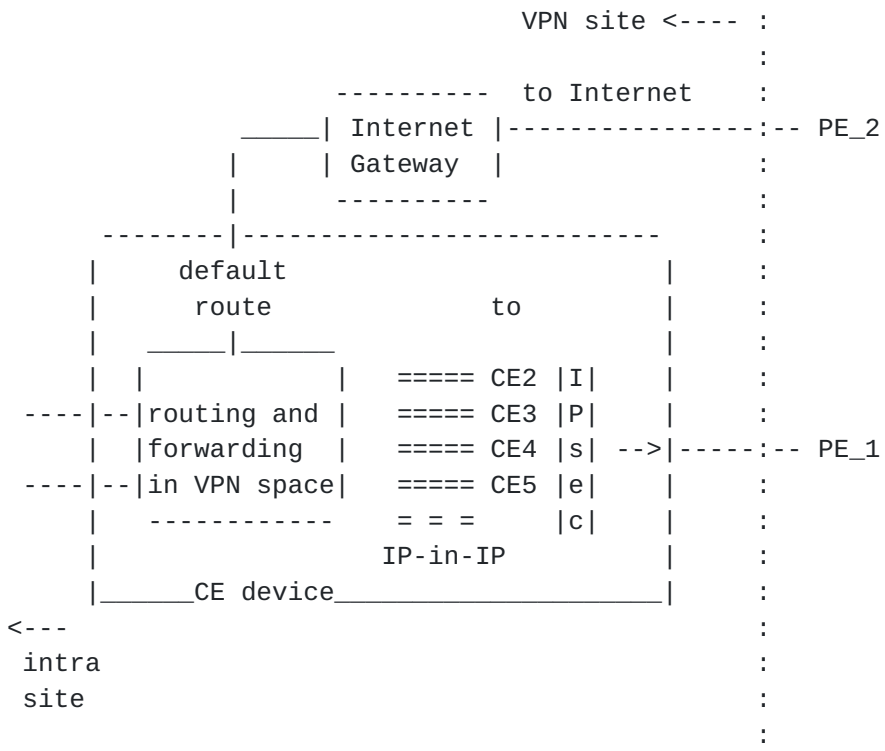

figure 5: Internet Access from within a VPN

The Internet Gateway will process (e.g. firewall) all traffic coming from within the VPN and, if accepted, send it to the PE with which it interfaces. As such the Internet Gateway effectively is the device that interfaces between the VPN routing space and the SP's/Internet routing space. Note that traffic that leaves a VPN via an Internet Gateway will not be IP-in-IP encapsulated and will not be IPsec processed. The traffic coming from the gateway will then be forwarded according to the PE's (default/Internet) forwarding table.

In order to allow for traffic in the reverse direction (from the Internet to the VPN sites), the ISP connected to the Internet Gateway must distribute, to the Internet, routes that lead to addresses that are within the VPN. NAT-like techniques are also sometimes used. As such there will be routes that will lead from the Internet to the site's Internet Gateway. The Internet Gateway will process traffic coming from the Internet and, if accepted (based on local policies), send it into the VPN site where intra-VPN routing and forwarding will lead the packets to their destination. This distribution of routes that lead to addresses within the VPN towards the Internet is independent of any intra-VPN route distribution as described elsewhere within this specification. Note also that normally the internal structure of the VPN will remain invisible to the outside world.

When the Internet Gateway functions are implemented in the CE device and the CE device is attached via only one (sub-)interface towards only one PE device, inspection of the packets coming from the PE will indicate whether the concerned traffic is intra-VPN traffic (when the packet is an IPsec packet with the CE device's own IP address (s) in the outer header's destination address field and the encapsulated payload is an IP-in-IP encapsulated private IP packet (v), and a matching SA is found), or control-plane traffic (IKE(v2) or VPN remote management traffic: when the inspected packets conform to the control plane's policies), or VPN <--> Internet traffic (then the Internet Gateway function will decide whether the considered packets will be accepted, (translated), and forwarded or not).

In the above discussed procedures, some sites will access the Internet via a VPN tunnel that leads to another site of the same VPN, because they don't have an own Internet Gateway, and will forward the traffic according to the default route. Ultimately though, Internet traffic will always go via an Internet Gateway before entering/leaving a VPN.

Further note that the PE to which the Internet Gateway is attached doesn't necessarily need to carry all the Internet routes; a default route to another Internet router suffices.

6.2 Prohibiting or restricting Internet connectivity from within a CE-based VPN

In the approach described in this document, the CE device sends IP packets (s) to the VPN-unaware PE device and receives IP packets from that PE device. The PE device forwards these packets based on the IP addresses (s) in the (outer) IP header. The packets received by the PE are as such either packets that are routable within the SP's private scope, or either in the public Internet's scope. This section

discusses the implications hereof with regards to security and access control.

o traffic that the CE sends to the PE

Following the procedures described in this document, three types of 'VPN' traffic can be sent by the CE device towards the PE device:

(i) customer VPN traffic: intra-VPN traffic sent from one VPN site to another VPN site; these packets will always have the sending CE's IP address (s) in the IP header's source IP address field, the IP address (s) of a peer CE device of the same VPN in the IP header's destination IP address field, and will always contain an IPsec header;

(ii) secure remote management traffic: this comprises both the traffic to establish the secure management channel (e.g. IPsec or secure TLS) and the traffic to download the VPN configuration file; these packets will always have the CE's IP address (s) in the IP header's source IP address field;

(iii) IKE(v2) traffic: the IP packets sent between CE devices in order to establish SAs; these packets will always have the CE's IP address (s) in the IP header's source IP address field.

o traffic that the CE receives from the PE

Following the procedures described in this document, the same three types of traffic can be received by the CE device from the PE device. As such, the CE device should perform the following actions:

+ for IP packets that have the CE's own IP address (s) in the outer IP header's destination address field and that have an IPsec header: process the packets through the CE router's IPsec daemon where conformance with an existing SA will be checked, and the packets further processed;

+ for IKE(v2) packets that have the CE's own IP address (s) in the outer IP header's destination address field: process according to the tunnel establishment procedures described in this specification;

+ for IP packets that have the CE's own IP address (s) in the outer IP header's destination address field and that correspond to secured management traffic: process according to the VPN secure remote management procedures, which will depend on the used

management protocols;

+ for CE devices that have an integrated Internet Gateway role: process all other packets to the Internet Gateway module;

+ for CE devices that don't have an integrated Internet Gateway role: drop all other IP packets, unless explicitly allowed by complementary procedures that are out of scope of this memo.

o SP's control over CE initiated traffic

Note that with this specification's concepts, the PE device that receives traffic from a CE device has no means to verify whether the received traffic is intra-VPN traffic, or traffic that is sent to for example another VPN or e.g. to the Internet.

From a VPN data privacy point of view, this has no implications, as the security is enforced at the CE devices themselves: traffic that doesn't conform to the security associations or other policy rules will be dropped at the CE.

One remaining issue is that customers might use CE devices (that have been granted VPN access) to access services they have not been granted access for, via the PE device. Although this would possibly compromise the security of the customer's own VPN, the SP may want to deploy measures to prevent this without bringing full VPN knowledge to the PE. One way of doing this would be by using specific IP address ranges for VPN purposes and to have specific access lists configured on the PE devices (this has inter-SP and Internet transparency issues though). Note that maintaining, at every PE, a list of <CE device IP address, VPN-ID> would add a considerable management burden and is as such not advised. Another strategy for the SP would be not to care about the particularities of the traffic and treat it at the PE level as it treats public Internet traffic (and as such to only control the total of the resources consumed by particular access connections).

Taking into consideration that in many cases, VPNs will also need to be able to access the public Internet, and that the above problem does not seem to be an important threat for the SP nor the VPN customer, this issue is not considered as a major drawback for the deployment of the discussed VPN approach.

7. Security Considerations

The security aspects of what is presented in this document are implicitly discussed in most of the sections. This draft is for a large part focusing on security aspects.

Note that the security of the mechanisms presented here is highly dependent on the following factors:

- the security of the 'management channel', used by the management protocol to configure the VPN CE devices.

- the security of the site and of the CE-device itself

- the security aspects of the credentials: the IPsec credential must be generated, provisioned, updated, and stored securely

- for a VPN with a complex topology, every tunnel must use the same grade of security strength, otherwise, a single weak link degrades the whole VPN

A more detailed analysis of the security aspects of CE-based PPVPNs is described in [RFC4111].

8. IANA Considerations

This document has no actions for IANA.

9. Acknowledgements

The authors would like to thank the following persons for their valuable contributions to this document: Lars Eggert, Brian Gleeson, Archana Khetan, Sankar Ramamoorthi, Eric Rosen, Michael Choung Shieh, Joe Touch, Eric Vyncke, S. Felix Wu, Yu-Shun Wang, Cliff Wang, Alex Zinin.

10. Informative References

[FRAMEWORK] Callon, R. et al., "A Framework for Provider Provisioned Virtual Private Networks", RFC 4110, July 2005

[GRE] Farinacci, D. et al., "Generic Route Encapsulation", March 2000, RFC 2784

[IKE] Harkins, D. and Carrel, D., "The Internet Key Exchange (IKE)", November 1998, RFC 2409

[IKEv2] Kaufman, C., "Internet Key Exchange (IKEv2) Protocol, draft-ietf-ipsec-ikev2, work in progress

[IPinIP] Perkins, C., "IP encapsulation within IP", October 1996, RFC 2003

[IPSEC] Kent, S., Atkinson, R., "Security Architecture for the Internet Protocol", November 1998, RFC 2401

[IPSEC-DPD] Huang, G., Beaulieu, S., Rochefort, D., "A Traffic-Based Method of Detecting Dead IKE Peers", February 2004, RFC 3706

[L2TP] Lau, J., et al., "Layer Two Tunneling Protocol (Version 3)", March 2005, RFC 3931

[NAT] Srisuresh, P., Egevang, K., "Traditional IP Network Address Translator (Traditional NAT)", January 2001, RFC 3022

[RFC2402] Kent, S., Atkinson, R., "IP Authentication Header", November 1998, RFC 2402

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

[RFC2407] Piper, D., "The Internet IP Security Domain of Interpretation for ISAKMP" November 1998, RFC 2407

[RFC2408] Maughan, D., et al., "Internet Security Association and Key Management Protocol (ISAKMP)", November 1998, RFC 2408

[TOUCH] Touch, J. and Eggert, L., "Use of IPSEC transport mode for Dynamic Routing", September 2004, RFC 3884

[RFC4111] Fang, L., "Security Framework for Provider-Provisioned Virtual Private Networks (PPVPNs)", July 2005, RFC 4111

11. Authors' Addresses

Jeremy De Clercq
Alcatel
Fr. Wellesplein 1, 2018 Antwerpen, Belgium
E-mail: jeremy.de_clercq@alcatel.be

Olivier Paridaens
Alcatel
Fr. Wellesplein 1, 2018 Antwerpen, Belgium
E-mail: olivier.paridaens@alcatel.be

Cliff Wang
E-mail: cliff.wang@us.army.mil

Intellectual Property Statement

Full Copyright Statement

Disclaimer of Validity