Internet Engineering Task Force
INTERNET-DRAFT
Expires May 24, 2001

M. Suzuki and J. Sumimoto (Ed.)
NTT
A. Malis
Vivace Networks, Inc.
K. Muthukrishnan
Lucent Technologies
November 24, 2000

A Framework for Network-based VPNs
<draft-suzuki-nbvpn-framework-02.txt>

Status of this Memo

Abstract

The objective of this draft is to clarify a framework for standardizing the mechanisms supporting interoperable network-based virtual private networks (NBVPNs). These are VPNs using IP facilities whose operating mechanisms are implemented within a network (or networks) and outsourced to one or more service providers. This draft first describes the assumed services of NBVPNs and clarifies the logical architecture model and reference model of an NBVPN. Considering the assumed services, this draft further clarifies the NBVPN requirements for interfaces and MIBs in the reference model. It also surveys and discusses current technologies supporting NBVPNs such as tunneling, VPN identifier, routing, and QoS/SLA. Additionally it will, in future, provide an outline of the

interface and MIB specifications and present criteria for achieving interoperability.

1. Objective and Scope of this Document

The objective of this document is to clarify a framework for standardizing the mechanisms supporting interoperable network-based virtual private networks (NBVPNs). Note that the document uses concepts and discussions in [RFC2764], but does not repeat the discussions therein.

This framework includes assumed services of NBVPN for which interoperable solutions need to be developed, a logical architecture model and reference model of NBVPN, requirements for interfaces and MIBs of the NBVPN reference model, an outline of the interface and MIB specifications, overview of related technologies, and criteria for achieving interoperability.

A VPN service is defined as a service that provides a network whose logical structure, such as addressing, reachability, and access control, is equivalent to part of or all of a conventional enterprise network using private facilities, it does not affect the logical structure in the rest of the enterprise network, and it is implemented with public network facilities.

In particular, a VPN service that uses facilities of the Internet is called an IP VPN service. Since IP VPN services are provided at lower costs and their service provisioning is more flexible than that of VPNs based on other technologies, various IP VPN implementations have been developed.

IP VPN implementations are further classified into "network-based VPNs (NBVPNs)" and "customer premises equipment (CPE)-based VPNs." The NBVPN is an IP VPN whose VPN operations mechanisms are implemented within a network (or networks) and outsourced to one or more service providers (SPs) [RFC2764]. Compared with a CPE-based VPN, in which the VPN operations mechanisms are implemented in CPE, the NBVPN has the advantage of reducing the customer's overhead for VPN operations, so it is attracting the attention of Internet users and SPs.

Looking at current implementations of NBVPNs, we see that a single technology cannot serve as the base technology, so various technologies such as MPLS [MPLS-ARC] [MPLS-FRAME] and IPsec [RFC2401] have been used. However, there has been no practical and commonly supported way of achieving interworking between an NBVPN of one technology and another NBVPN of another technology even though they have similar mechanisms. Thus, early provision of such a solution is

eagerly awaited by Internet users and SPs.

In order to support the standardization activity (responding to demands) to provide solutions for NBVPN interworking, this framework is created and serves as the basis for standardization in terms of the architecture and specifications of NBVPNs.

This standardization work aims to avoid applying excessive constraints on the mechanisms and specifications of base technologies (e.g., tunneling mechanisms) so that future advances in the base technologies for NBVPN can also be accommodated within this framework. This standardization work does not intend to modify any currently used mechanisms or specifications of the base technologies, either.

The NBVPNs targeted by this framework are:

o Virtual private routed networks, which are defined as an emulation of a multi-site wide area routed network using IP [RFC2764].

Excluded are:

o NBVPNs using VPN native (non-IP) protocols as their base technologies. However, this does not mean to exclude multi-protocol access to the NBVPN by users.

o Virtual leased lines, which provide a point-to-point link between two user sites [RFC2764].

o Virtual private dialup networks, which are defined as an emulation of on-demand isolated IP reachability from a remote user to a user site. The remote user is connected via a dial-up PSTN or ISDN link [RFC2764].

o Virtual private LAN segments, which are defined as an emulation of a LAN segment using Internet facilities [RFC2764].

This standardization is expected to lead to the following benefits.

o Benefits to SPs

  It will enable flexible NBVPN implementation over multi-vendor multi-mechanism subnetworks. It will remove the constraint that all user sites of an NBVPN are limited to a specific vendor or mechanism. It will also lead to lower costs than with the uniform NBVPN implementation.

o Benefits to customers

Customers will have more chance to construct wider area (e.g., international) NBVPNs as a result of the multi-SP multi-vendor environments provided by this technology. They will also get cheaper NBVPN services.

In this document, section 2 describes assumed services of NBVPNs, section 3 clarifies the logical architecture model and reference model for NBVPNs, section 4 clarifies requirements for interfaces and MIBs in the NBVPN reference model, and section 5 outlines interface and MIB specifications. Moreover, section 6 surveys current mechanisms and discusses their issues, section 7 discusses criteria for achieving interoperability, and section 8 summarizes security considerations.

2. Assumed Services of NBVPNs

This section describes assumed services of NBVPNs which are provided to user sites by the networks. The purpose of discussing assumed services is to extract the requirements for mechanisms to be standardized for interoperable NBVPNs. We do not intend to standardize these services for NBVPNs in any way.

2.1 Closed User Group (CUG)

A closed user group (CUG) service provides communications between various specific user sites through an NBVPN. Other user sites cannot reach them. This is the basic service of an NBVPN. Operation mechanisms are implemented within a network and the operations are performed by an SP. This service prevents packets from being injected into the network without authorization. It also prevents packets from being snooped on, modified in transit, or subjected to traffic analysis by unauthorized parties. Private IP addressing may be used in a CUG.

2.2 CUG Interconnection

A CUG interconnection service enables communications between specific CUGs or user sites belonging to other CUGs within the networks. Access control (including packet filtering and address translation) may be applied between CUGs according to policy. Interconnection of CUGs performed in user sites is outside the scope of this document.

2.3 QoS/SLAs

QoS/SLA services provide guaranteed and/or differentiated communications with NBVPN-specific SLAs covering loss rates, jitter, latency, and bandwidth etc. Various classes of QoS are provided, although they may depend on the supporting technologies, e.g.,

IntServ [RFC2211] [RFC2212], DiffServ [RFC2474] [RFC2475], or L2 traffic engineering capabilities [AF-TM-0121.000].

2.4 Dynamic Routing

A dynamic routing service enables the exchange of unicast routing information between user sites and an NBVPN using a routing protocol such as Open Shortest Path First (OSPF) [RFC2328] or Border Gateway Protocol 4 (BGP-4) [RFC1771]. Routing information about each user site can be distributed from one user site to another. This service is essential for multihomed user sites, in which the main purpose of multihoming is to improve reliability.

2.5 Multiprotocol Transport

A multiprotocol transport service supports traffic carried between user sites using various different protocols.

2.6 NBVPN over Multiple SPs

An NBVPN over multiple SPs service enables a single NBVPN to cover multiple SPs.

2.7 Multicast

A multicast service replicates multicast packets forwarded from user sites in the networks and forwards them to multiple user sites. Multicast routing information is exchanged between user sites and an NBVPN using a multicast routing protocol.

2.8 Note on Data Security Service

[RFC2764] discusses data security service which provides stronger security than that of the basic CUG service and which is supported by encryption and authentication. In this framework document, it is not assumed for a NBVPN service for the following reasons.

- o If a user requires stronger security than that of the NBVPN service, it should be provided by a CPE-based security mechanism. This is because a network-based solution cannot ensure the security of access links between user sites and a network.

- o If stronger security is provided by a network-based mechanism, it is located at the edge of the SP network providing the NBVPN service. Thus, security and NBVPNs service mechanisms are independent, because the security protocol layer is located on the protocol layer that provides NBVPN service. Therefore, this security mechanism is not discussed in this framework.

However, a similar security mechanism may be needed on the SP interworking interface of NBVPNs. See sections 4.4.1 and 8 for details.

3. Logical Architecture Model and Reference Model for NBVPN

This section describes the logical architecture model and reference model for NBVPN. These will be used in mapping the NBVPN service descriptions in section 2 to interfaces and MIBs requirements described in section 4.

3.1 Logical Architecture Model for NBVPN

The logical architecture model for NBVPN describes functions and their relationship for implementing NBVPN. Figure 3.1 shows the logical architecture model. The architecture is based on a real routed IP network.

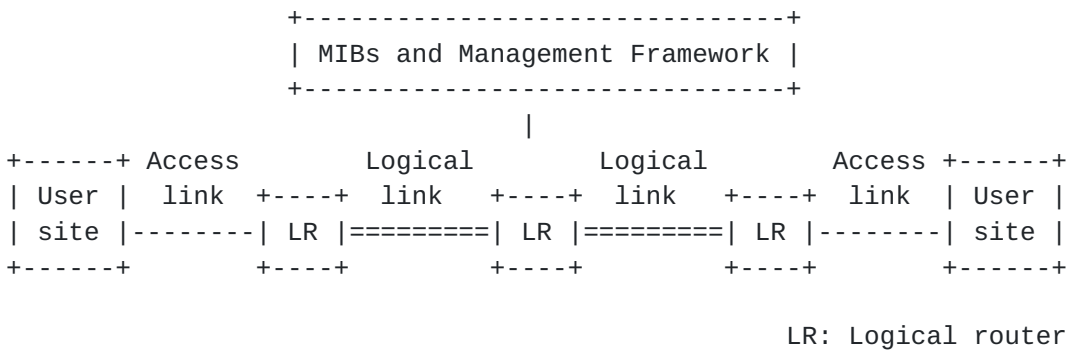

Figure 3.1: Logical architecture model.

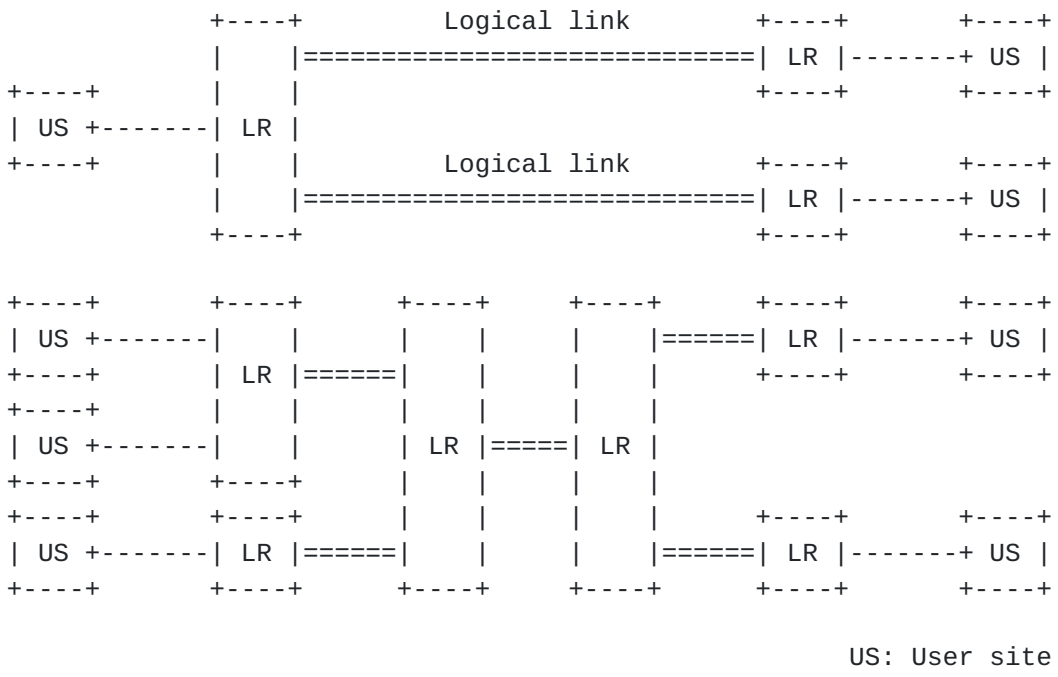

Figure 3.2: Example configurations
applying the logical architecture model.

Figure 3.1 shows a generalized model. It can represent various NBVPN configurations, as shown in Figure 3.2. The entities in the logical architecture model are described below.

o Logical router

A logical router supports router functions dedicated to a serving NBVPN. It has the following functions.

- Routing function: A logical router creates, modifies, and maintains entries in a routing table of the serving NBVPN using routing protocols.

- Forwarding function: A logical router forwards IP packets within the NBVPN by looking up entries in the routing table.

- Access control function: A logical router may control access (packet filtering and address translation) from other NBVPNs or from the Internet.

o User site

A user site is one or more subnetworks that are part of an NBVPN.

- Logical link

  A logical link is a connection (isolated from other NBVPNs and the Internet) between logical routers whose serving NBVPNs are identical. A logical link is terminated by logical routers.

- Access link

  An access link provides a user site with access to services associated with a specific NBVPN. Note that a physical facility may multiplex multiple access lines, but this is outside the scope of this model.

- MIBs and management framework

  These represent MIBs for managing the customer configuration associated with the concerned VPN, MIBs for managing logical routers, and other devices constructing the concerned NBVPN and associated managing functions.

3.2 Reference Model

In order to clarify possible mapping between the logical architecture model given in section 3.1 and implementation as well as to clarify the targets of the standardization work, this section describes a reference model illustrating the reference configuration of the NBVPN. Figure 3.3 shows the reference model.

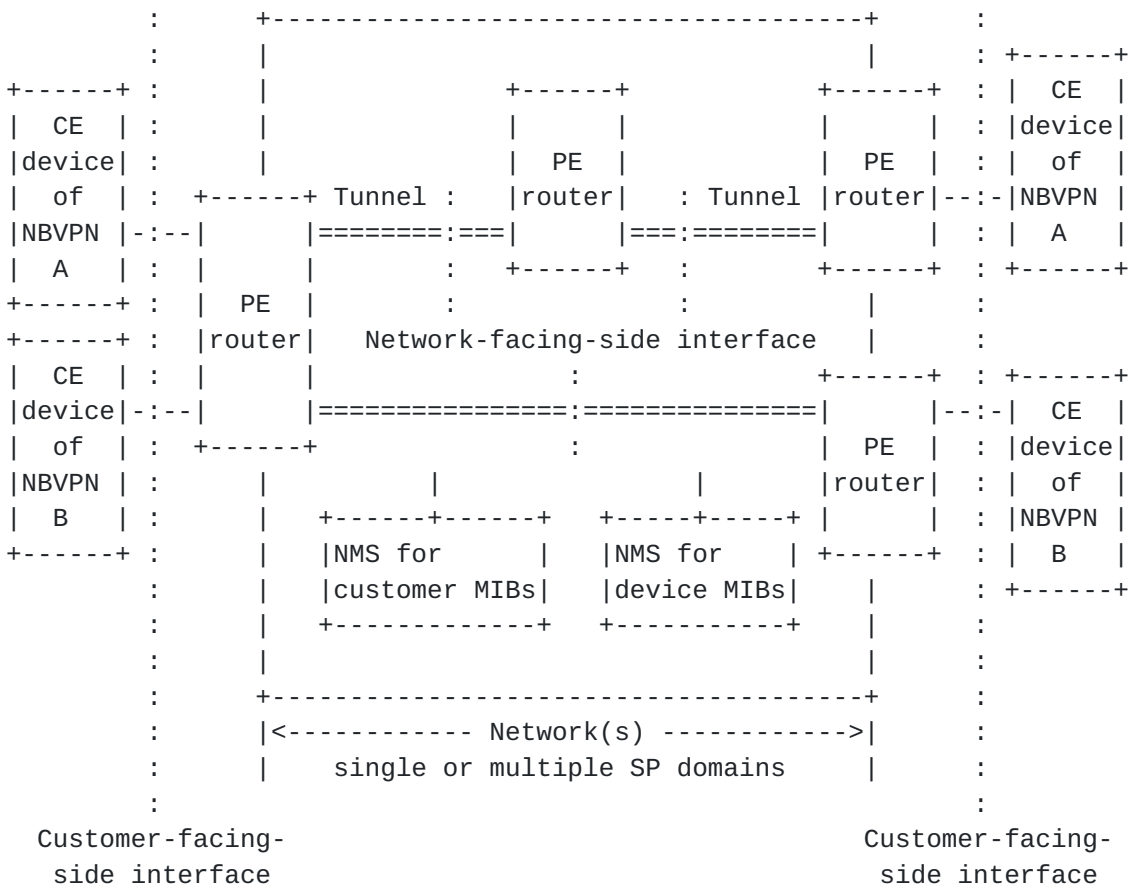

Figure 3.3: Reference model.

o Customer edge (CE) device

A CE device is usually a router located at the edge of a user site. It may also be a host or hosts belonging to a subnetwork. A CE device belongs to only one NBVPN, although it can reach other NBVPNs through a CUG interconnection service. It is usually accommodated by a single PE router. However, four types of double-homing arrangements, shown in Figure 3.4, must be supported.

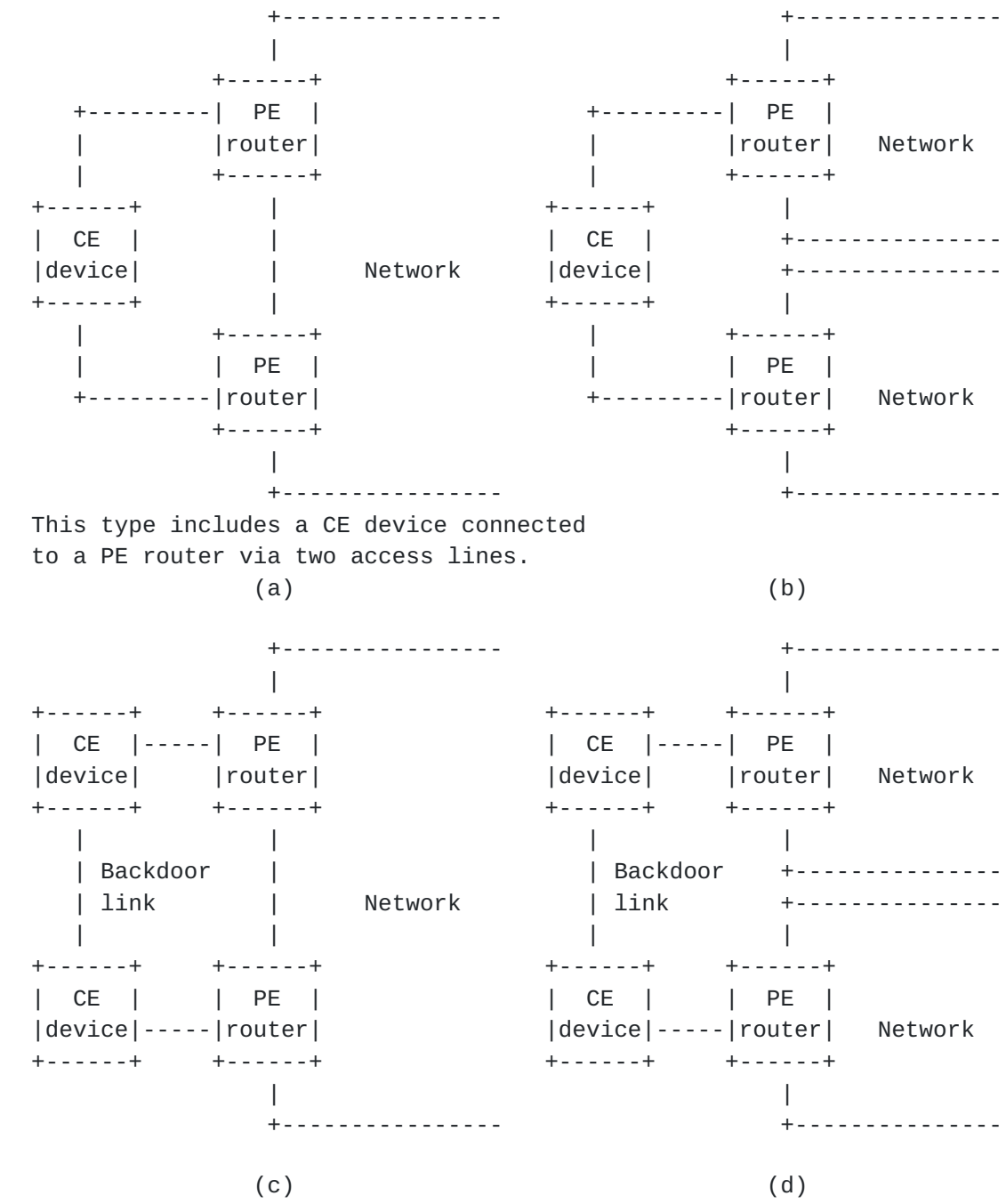

Figure 3.4: Four types of double-homing arrangements.

o Networks

NBVPN services are provided by one or more networks to CE devices as members of the concerned NBVPN. These networks support PE routers, tunnels, NMSs for customers and device MIBs. In this document, "a network" means a single domain of an SP. The NBVPN

operation in a network is outsourced to an SP, but the whole NBVPN operation may be spread over multiple SPs.

o Tunnel

A tunnel is a connection between PE routers. Multiple logical links defined in section 3.1 may be multiplexed into a single tunnel. A number of IP tunneling protocols have been proposed, but in this document, three different tunneling mechanisms--that is MPLS, GRE, and IPsec--are considered to support NBVPN. A single NBVPN may make use of a mixture of tunneling mechanisms.

When MPLS is used for the tunneling mechanism, LSPs implement tunnels and two multiplexing schemes are supported. The first scheme uses two-layer label stacking of the MPLS. In this scheme, the multiple logical links identified by second labels are multiplexed in the tunnel identified by the top label. The second scheme is applicable when the MPLS network is implemented by ATM, and it uses the CPCS user-to-user field in the AAL5 trailer or the VPN-ID field in the VPN encapsulation header [RFC2684]. In this scheme, the multiple logical links in the tunnel are identified by the CPCS-UU or VPN-ID field respectively.

When GRE is used for the tunneling mechanism and the key field extension is supported, the logical links are identified by the key field. Note that if the key field is not present, the tunnel supports only one logical link. When IPsec is used, they are identified by the SPI field.

Note that when the tunnel is provided by GRE or IPsec, it may pass through another tunneling mechanism (e.g., an IPsec tunnel over an MPLS network). In this document, a tunnel is identical to the tunnel that directly multiplexes logical links and does not include underlying tunneling mechanisms.

Figure 3.5 illustrates logical link multiplexing. Multiple logical links supporting connections for NBVPNs are multiplexed into a tunnel. This arrangement allows multiplexing of logical links of different NBVPNs.

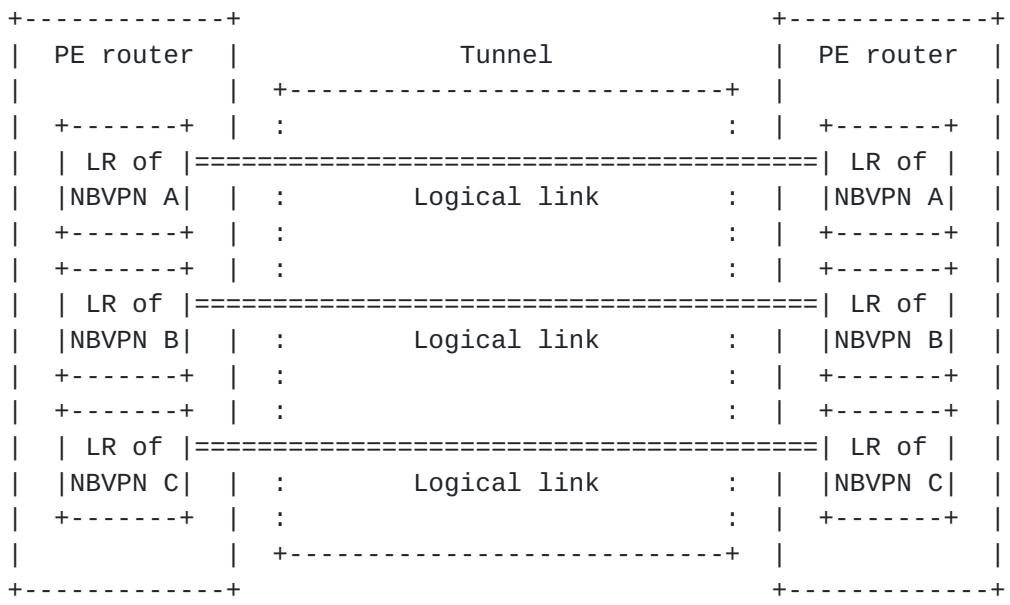

Figure 3.5: Logical link multiplexing.

o Provider edge (PE) router

A PE router implements one or more logical routers. It is usually located at the edge of an SP network. It may terminate access links. In this document, the virtual router (VR) [VPN-VR] and VPN routing and forwarding (VRF) tables [VPN-2547BIS] approaches are considered as methods of implementing logical routers in a PE router.

VR is a technology for implementing a router function in a PE router. A PE router may contain more than one VR and a VR supports only one NBVPN. A logical router can be implemented with a VR. A VR forwards user traffic from a CE device or another VR, which belonging to the same NBVPN, to another CE device or VR via an access or logical link respectively. For the dynamic routing service described in section 2.4, a VR also forwards route information inside user sites, which is received from a CE device or another VR, to another CE device or VR as user traffic.

The distinctive feature of this approach is that the current routing protocols are applicable between VRs or PE routers without any extensions or modifications. Thus, it can be implemented without difficulty and managed simply. However, an extension for a routing protocol between PE routers has been proposed to support auto-setup of tunnels and auto-discovery of PE router topology and NBVPN membership [VPN-BGP-VR].

A VRF table is a packet routing and forwarding table and a user site corresponds to a VRF table. In a PE router, each logical router can be implemented with an entity of routing protocol between PE routers whose processing is based on VRF tables. Based on the route information of a VRF table in a PE router, user traffic received from a CE device or another PE router is forwarded to another CE device or PE router via an access or logical link respectively. For the dynamic routing service, a PE router distributes route information inside user sites, which is received from a CE device or another PE router, to another CE device or PE router using routing protocol between PE routers. See [VPN-2547BIS] for detail.

This approach requires an extension of the route information format to distinguish the same IPv4 addresses belonging to different NBVPNs and an extension of the routing protocol between PE routers to distribute the extended route information. Currently, extensions for BGP-4 protocol have been proposed. Furthermore, for a dynamic routing service, when CE devices and PE routers in an NBVPN exchange route information inside user sites using OSPF, IS-IS, or RIP, and if different CE devices must belong to the same OSPF, IS-IS, or RIP domain, extensions which correspond to these protocols are required for the routing protocol between PE routers.

However, in this approach, the number of routing protocol entities in a PE router does not depend on the number of NBVPNs supported by the PE router, so it achieves high scalability. This approach assumes the use of LSP with two-layer label stacking as the tunneling mechanism, and basically, multiple logical links identified by second labels are multiplexed in the tunnel identified by the top label. Therefore, the tunnel enables high-speed packet forwarding, because the forwarding processing does not refer to the second label which reflects the number of NBVPNs supported by the PE router.

In this approach, a VRF table can support more than one NBVPNs, so, a user site is able to belong to multiple NBVPNs. However, the overlapping address space between NBVPNs can be allocated only when the NBVPNs have no common user sites. That is, if two NBVPNs have the common address space, a user site can belong to only one NBVPN. And if an NBVPN has a private address space and it is interconnected to the Internet via NAT, the user traffic must be forwarded to a CE device where the NAT is located. Therefore in this case, this approach does not optimize routing paths in the network(s) providing the NBVPN.

- NMS for customer MIB

An NMS that manages customer MIBs of an NBVPN.

o NMS for device MIB

An NMS that manages device MIBs of an NBVPN.

3.3 Classification of Network-facing-side Interface

In this section, the network-facing-side interface shown in Figure 3.3 is classified into three specific interfaces.

It is not necessary for a single SP's whole network to be constructed with a uniform technology. As shown in Figure 3.6, different subnetworks may be implemented with different technologies. In this case, a PE router must be placed at the edge of a subnetwork interconnecting with another subnetwork that is based on another technology. In this document, it is called a subnetwork edge (SE) router.

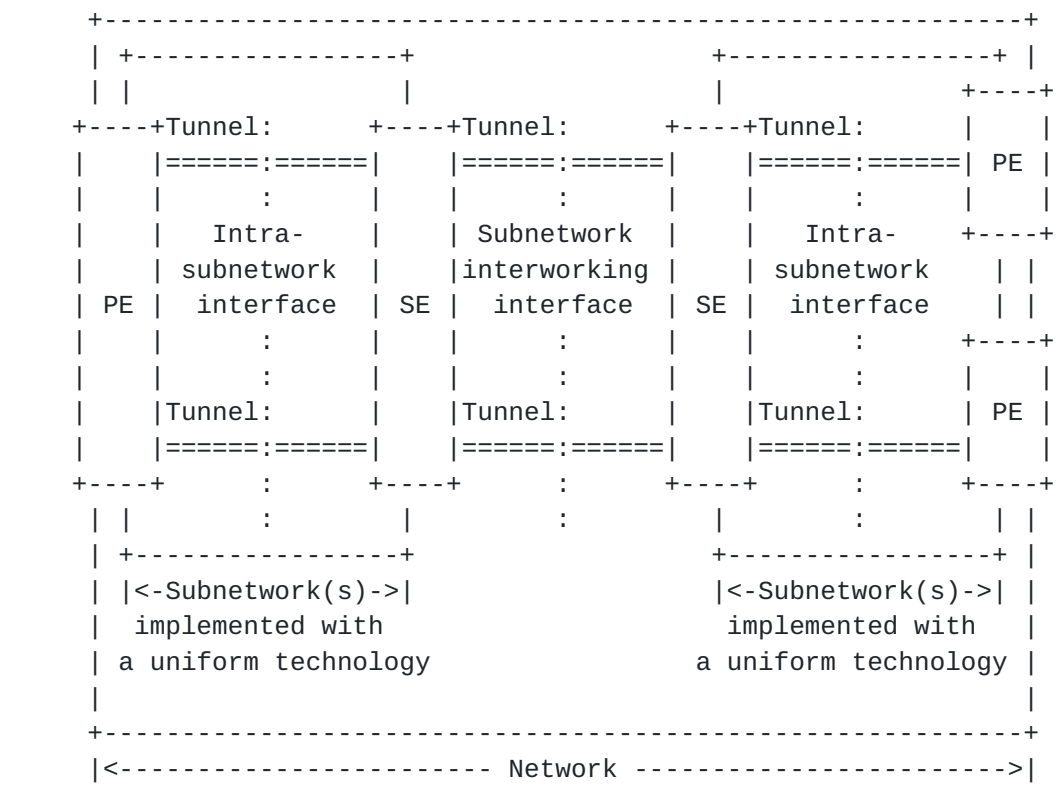

Figure 3.6: Intra-subnetwork interface and subnetwork interworking interface.

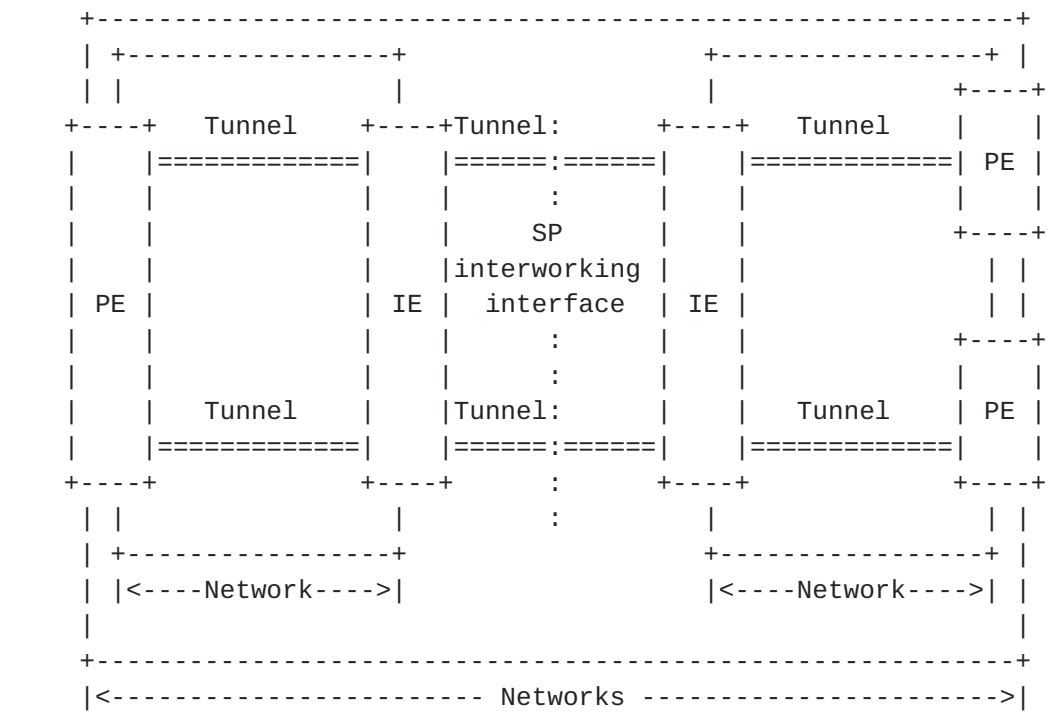

Figure 3.7: SP interworking interface.

Similarly, when a single NBVPN spans multiple SPs, PE routers should be placed at every SP interconnecting point as shown in Figure 3.7. In this document, they are called inter-provider edge (IE) routers.

In the rest of this document, SE and IE routers are also simply called "PE routers" unless they need to be differentiated.

The intra-subnetwork interface and subnetwork interworking interface are defined as shown in Figure 3.6. The former interface exists between a pair of PE routers and is restricted to one or more subnetworks implemented with a uniform technology. The latter interface exists between a pair of SE routers and connects two subnetworks implemented with different technologies. The SP interworking interface is defined as shown in Figure 3.7. It exists between a pair of IE routers and connects two SP networks.

3.4 Targets of the Standardization Work and Protocol Architecture

The targets of the standardization work are the following two interfaces and MIBs illustrated in the reference model given in Figures 3.3, 3.6, and 3.7.

o Customer-facing-side interface

  An interface between a CE device and a PE router.

- Network-facing-side interface

  An interface between PE routers. This interface is further classified into the following three interfaces.

  - Intra-subnetwork interface

  - Subnetwork interworking interface

  - SP interworking interface

- Customer MIBs

  MIBs of NBVPN customer attributes.

- Device MIBs

  MIBs of device attributes, covering PE routers and other devices constructing the concerned NBVPN.

To clarify the protocol architecture on the network-facing-side interface, protocols on the interface are classified into the u- and c-planes.

The u-plane provides forwarding of user traffic between CE devices belonging to the same NBVPN. For the dynamic routing service implemented with the VR approach, a VR forwards route information inside user sites to another VR as user traffic via a logical link. Therefore, this protocol is included in the u-plane. Tunneling protocols that connect PE routers belong to the u-plane protocols. However, tunnel setup protocols are included in the c-plane.

The c-plane provides auto-discovery of PE routers topology and NBVPN membership. It also provides auto-setup of tunnels based on the PE routers topology information. For the dynamic routing service implemented with the VRF approach, it provides distribution of route information inside user sites between PE routers. Routing protocols between PE routers and control protocols for MPLS and IPsec belong to the c-plane. Note that GRE is not equipped with standard ways to set up and maintain GRE tunnels.

4. Requirements for Interfaces and MIBs

4.1 General Requirements

The implementation providing an NBVPN must:

- be scalable

- be manageable
- enable a single NBVPN to span multiple subnetworks implemented with different technologies. For example, a single NBVPN must be able to span IPsec- and MPLS-based-subnetworks.
- enable a single NBVPN to span multiple SPs.

4.2 Requirements for Identifiers

This section clarifies the requirements for the identifiers to describe the requirements for the interfaces and MIBs. Several identifiers are defined, as illustrated in Figure 4.1.

Note that not all protocols and MIBs specified in section 3.4 need to support all identifiers described in this section. However, supported identifiers must be the same as, logically equivalent to, or inclusive of identifiers described in this section.

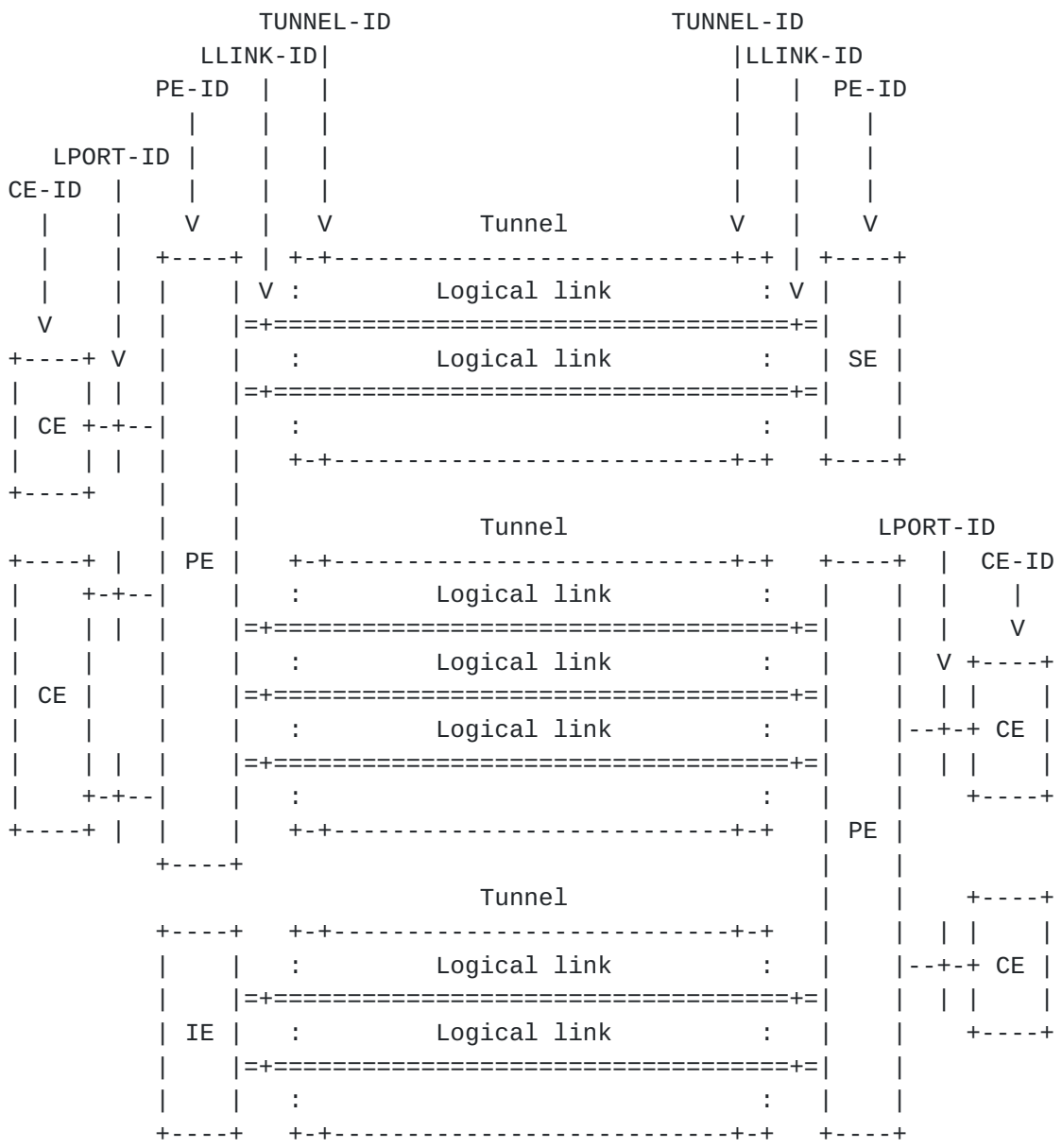

Figure 4.1: Identifiers.

- SP-ID, which identifies each SP, must be unique at least within all the interconnected networks of SPs. (In practice, it should be globally unique.) This is necessary when a single NBVPN spans multiple SPs.

- VPN-ID, which identifies each NBVPN, must be unique at least within each SP's network.

- CE-ID, which identifies each CE device, must be unique at least within each SP's network.

- PE-ID, which identifies each PE router, must be unique at least within each SP's network. The PE-ID of an IE must be unique at least within all the interconnected SP networks.

  Note: One of the IP addresses assigned to an edge device is usually used as PE-ID.

- LPORT-ID, which identifies a logical port, must be unique at least within each PE router containing the logical port. Here, a logical port represents a terminating point of an access link accommodating a user site.

- TUNNEL-ID, which identifies each tunnel, must be unique at least within each PE router supporting the tunnel.

- LLINK-ID, which identifies each logical link, must be unique at least within each tunnel supporting the logical link.

The scope of the identifiers is summarized in Figure 4.2. It shows that the right-side identifier must be unique at least within the scope of the left-side identifier for each arrow.

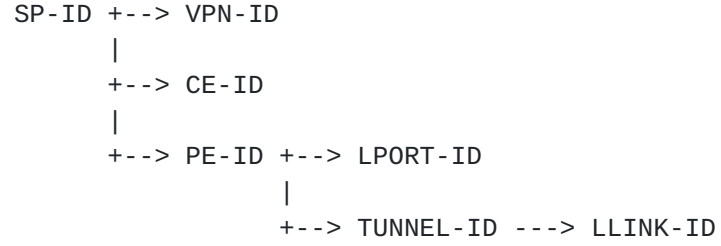

Figure 4.2: Scope of identifiers.

When a single NBVPN spans multiple SPs, their identifiers, except for SP-ID, must satisfy one of the following conditions: 1) their mappings are predefined, 2) their mappings are dynamically built by a protocol, or 3) they are linked together with the SP-ID.

The association among the identifiers must satisfy the following requirements.

- The CE-ID must be mapped to one or more pairs of PE-ID and LPORT-ID to configure the accommodation of CE devices. Note that it is not necessary for the mapping to be built in a one-to-one manner because a CE device may be connected to PE routers through multiple access links as shown in Figures 3.4(a) and (b). In this case, the CE-ID must be mapped to all the concerned pairs of PE-ID and LPORT-ID.

- The CE-ID must be uniquely mapped to the VPN-ID to distinguish the NBVPN associated with the CE device.

- A pair of PE-ID and LPORT-ID must be uniquely mapped to the VPN-ID to distinguish the NBVPN associated with the logical port.

- A set of PE-ID, TUNNEL-ID, and LLINK-ID must be uniquely mapped to the VPN-ID to support a logical link.

4.3 Requirements for Customer-facing-side Interface

This section describes the requirements for the customer-facing-side interface shown in Figure 3.3.

- Packet encapsulation

  Every packet must have the usual IP packet format without VPN-aware encapsulation, except in the case of providing multiprotocol transport service where every packet must have a protocol-specific packet format without VPN-aware encapsulation.

- QoS/SLA

  For QoS/SLA service, every access link connecting a CE device and a PE router must support the specified QoS/SLA.

- Dynamic routing

  For dynamic routing service, different routing protocols must be supported per access link connecting a CE device and a PE router.

4.4 Requirements for Network-facing-side Interface

This section describes the requirements for the three specific network-facing-side interfaces shown in Figures 3.3, 3.6, and 3.7.

4.4.1 Requirements for protocols on u-plane

- Packet encapsulation

  Every packet must be encapsulated with the LLINK-ID. Multiprotocol transport service requires multiprotocol-over-IP encapsulation.

- QoS/SLA

  For QoS/SLA service, every tunnel or logical link must support the specified QoS/SLA per NBVPN. Note that if QoS/SLA support is per-tunnel based, it can support only one logical link.

- Note on security

  If a tunnel on the SP interworking interface is not implemented with a direct circuit between IE routers and it passes through an unsecure SP, POP, NAP, or IX, then security mechanisms should be located at the edge routers. However, this security and NBVPN service mechanisms are independent, so the detailed specifications of the security mechanism depend on the implementation. See sections 2.8 and 8 for security discussions.

4.4.2 Requirements for protocols on c-plane

- Tunnel setup and maintenance

  To set up tunnels between PE routers, every PE router must support static configuration for tunneling and may support a tunnel setup protocol. When PE routers support the protocol, the information exchanged between them includes the VPN-ID, TUNNEL-ID, QoS/SLA information for QoS/SLA service, and multiprotocol-over-IP encapsulation information for multiprotocol transport service.

  A protocol for monitoring tunnel states must be supported.

  A protocol for tunnel restoration must be supported.

  For multicast service, multicast traffic must be forwarded through the created tunnels.

- Auto-discovery of PE routers topology and NBVPN membership

  For auto-discovery of PE routers topology and NBVPN membership, extensions for routing protocol between PE routers may be needed.

  Routing protocols on the SP interworking interface may support authentication.

  If policy routing is performed, routing protocols running between IE routers on the SP interworking interface may specify intermediate SPs by SP-ID in route distribution and then routing protocols running between IE routers on the intra-subnetwork and subnetwork interworking interface may also specify intermediate SPs by SP-ID in route distribution.

- Dynamic routing

  For dynamic routing service implemented with the VRF approach, routing protocols running between IE routers must support route control independently per NBVPN.

4.5 Requirements for Customer MIB

This section describes the requirements for the customer MIB shown in Figure 3.3.

- o Management information about CE devices and customer attributes of NBVPN must be configured and maintained. The information includes the CE-ID, PE-ID, LPORT-ID, VPN-ID, access control policy information for CUG interconnection service, routing protocols used for dynamic routing or multicast service, and QoS/SLA information for QoS/SLA service.

4.6 Requirements for Device MIB

This section describes the requirements for the device MIB shown in Figure 3.3.

- o The configuration and maintenance of PE routers must be supported. Their management information includes IP routing information and access control policy information for CUG interconnection service. For multiprotocol transport service, protocol-specific routing information must be managed instead of IP routing information.

- o The mappings between the LPORT-ID and VPN-ID must be configured and maintained. For QoS/SLA service, the mappings between LPORT-ID and QoS/SLA information must also be configured and maintained.

- o Tunnel information must be configured and maintained. It includes the TUNNEL-ID, LLINK-ID, tunnel states, and QoS/SLA information for QoS/SLA service.

- o Routing protocols running between PE routers and CE devices must be configured and maintained per NBVPN. For multicast service, multicast routing protocols must also be supported.

- o Routing protocols running between PE routers must be configured and maintained. For multicast service, multicast routing protocols must also be supported.

5. Outline of Interface and MIB Specifications

(To be written)

6. Survey of Available Technologies

The technologies surveyed in this section are relevant to NBVPNs. The framework, however, neither compels nor excludes their use.

6.1 Tunneling

Tunneling mechanisms provide isolated and secure communication between two CE devices. Available tunneling mechanisms include (but are not limited to): MPLS [MPLS-ARCH] [MPLS-FRAME] [MPLS-ATM], GRE [RFC2784] [RFC2890], and IPsec [RFC2401] [RFC2402]. In an NBVPN, a tunnel is a secure communication path within a network. A PE router encapsulates a data packet incoming from a CE device, and injects it into an appropriate tunnel. The data packet traverses the network, and reaches the PE router on the far side. In the course of traversal, the data packet may have transferred to other tunnels, if necessary. The PE router then retrieves the data packet from a tunnel, and passes it to the destination CE device.

6.1.1 MPLS [MPLS_ARCH] [MPLS_FRAME] [MPLS-ATM]

Multiprotocol Label Switching (MPLS) is a method for forwarding packets through a network. Routers at the edge of a network apply simple labels to packets. A label may be inserted between the data link and network headers, or may be carried in the data link header (e.g., the VPI/VCI field in an ATM header). Routers in the network switch packets according to the labels with minimal lookup overhead. A path, or a tunnel in the NBVPN, is called a "label switched path (LSP)."

- Multiplexing

  LSPs may be multiplexed into another LSP.

- Multiprotocol transport

  MPLS can carry data packets other than IP ones.

- QoS/SLA

  MPLS does not have intrinsic QoS or SLA management mechanisms. Some other techniques such as DiffServ may be used with it [DIFF-MPLS].

- Tunnel setup and maintenance

  LSPs are set up and maintained by LDP (Label Distribution Protocol) or RSVP (Resource Reservation Protocol) [LSP-RSVP].

- o Large MTUs, minimization of tunnel overhead, and frame sequencing

  MPLS does not restrict the MTU size. The overhead of label switching should be minimal. MPLS guarantees in-order delivery of packets.

6.1.2 GRE [RFC2784] [RFC2890]

Generic Routing Encapsulation (GRE) specifies a protocol for encapsulating an arbitrary payload protocol over an arbitrary delivery protocol [RFC2784]. In particular, it may encapsulate an IP payload packet over IP. An endpoint encapsulates and decapsulates GRE packets. A GRE tunnel is a communication path between two endpoints established by the use of GRE.

- o Multiplexing

  The GRE specification [RFC2784] does not support multiplexing. But the key field extension to GRE is specified in [RFC2890] and it may be used as a multiplexing field.

- o Multiprotocol transport

  GRE is assumed to support any type of payload protocol.

- o QoS/SLA

  These capabilities depend on the delivery protocol.

- o Tunnel setup and maintenance

  GRE is not equipped with standard ways for setting up and maintaining GRE tunnels.

- o Large MTUs, minimization of tunnel overhead, and frame sequencing

  These capabilities depend on the delivery protocol, but the GRE header overhead is designed to be minimal. The sequence field proposed in [RFC2890] may be used to achieve in-order delivery.

6.1.3 IPsec [RFC2401] [RFC2402] [RFC2406] [RFC2409]

IP Security (IPsec) provides security services at the IP layer [RFC2401]. It comprises authentication header (AH) protocol [RFC2402], encapsulating security payload (ESP) protocol [RFC2406], and Internet key exchange (IKE) protocol [RFC2409]. AH protocol provides data integrity, data origin authentication, and an anti-replay service. ESP protocol provides data confidentiality and

limited traffic flow confidentiality. It may also provide data integrity, data origin authentication, and an anti-replay service. AH and ESP may be used in combination.

IPsec may be employed in either transport or tunnel mode. In transport mode, either an AH or ESP header is inserted between the IPv4 header and the transport protocol header. In tunnel mode, an IP packet is encapsulated with an outer IP packet header. Either an AH or ESP header is inserted between them. AH and ESP establish a unidirectional secure communication path between two endpoints, which is called a security association. In tunnel mode, two security associations compose a tunnel between PE routers. IKE protocol is used to exchange encryption keys among IPsec endpoints.

- o Multiplexing

  The SPI field of AH and ESP is used to multiplex security associations (or tunnels) within a tunnel.

- o Multiprotocol transport

  IPsec needs extensions to carry packets other than IP ones. Alternatively, GRE may be used with it.

- o QoS/SLA

  IPsec itself does not have intrinsic QoS/SLA capabilities. Other mechanisms such as "RSVP Extensions for IPSEC Data Flows" [RFC2207] or DiffServ may be used with it.

- o Tunnel setup and maintenance

  IKE is used for the setup and maintenance of tunnels.

- o Large MTUs, minimization of tunnel overhead, and frame sequencing

  IPsec does not restrict the MTU size. IPsec may impose its own overhead. IPsec has a sequence number field that is used by a receiver to perform an anti-replay check, not to guarantee in-order delivery of packets.

Note: IPsec is applicable to a CPE-based VPN as well as to an NBVPN. This document deals with the aspects of IPsec that are relevant to an NBVPN.

6.2 VPN Identifiers

An NBVPN spanning multiple autonomous systems requires the use of a globally unique VPN identifier such as "a pair of an autonomous system-number and a VPN-index local to the autonomous system" and the "global VPN identifier" as specified in [RFC2685]. A globally unique VPN identifier may be included in an MIB for the VPN configuration. It may also be included in an encapsulation header of a data packet or may be exchanged as a parameter of signaling messages.

The global VPN identifier defined in [RFC2685] consists of a 3-byte VPN organizationally unique identifier that identifies a VPN administrative authority, and a 4-byte VPN index that identifies the VPN within the context of a given VPN administrator. The VPN encapsulation header defined in [RFC2684] supports the global VPN identifier. But it must be noted that the global VPN identifier, which is 56 bits long, does not fit into the 20-bit MPLS label or into the 32-bit IPsec SPI field.

6.3 Routing

Dynamic routing service as defined in section 2 requires the exchange of routing information between a network and user sites. A list of applicable technologies is given in section 6.3.1. The network may terminate a routing protocol, or it may transfer routing information between user sites transparently.

The network must maintain its routing configuration with integrity. The applicable technologies are listed in section 6.3.2.

6.3.1 Exchange of routing information between network and user sites

The following technologies are available for the exchange of routing information between a network and user sites.

o Static routing

  Routing tables may be configured through a management system.

o RIP (Routing Information Protocol) [RFC2453]

  RIP is an interior gateway protocol and is used within an autonomous system. It sends out routing updates at regular intervals and whenever the network topology changes. Routing information is then propagated by the adjacent routers to their neighbors and thus to the entire network. A route from a source to a destination is the path with the least number of routers. This number is called the "hop count" and its maximum value is 15. This

implies that RIP is suitable for a small- or medium-sized networks.

- OSPF (Open Shortest Path First) [RFC1583]

  OSPF is an interior gateway protocol and is applied to a single autonomous system. Each router distributes the state of its interfaces and neighboring routers as a link-state advertisement, and maintains a database describing the autonomous system's topology. A link-state is advertised every 30 minutes or when the topology is reconfigured.

  Each router maintains an identical topological database, from which it constructs a tree of shortest-paths with itself as the root. The algorithm is known as the Shortest Path First or SPF. The router generates a routing table from the tree of shortest-paths. OSPF supports a variable length subnet mask, which enables effective use of the IP address space.

  OSPF allows sets of networks to be grouped together into an area. Each area has its own topological database. The topology of the area is invisible from outside its area. The areas are interconnected via a "backbone" network. The backbone network distributes routing information between the areas. The area routing scheme can reduce the routing traffic and compute the shortest-path trees and is indispensable for larger-scale networks.

  Each multi-access network with multiple routers attached has a designated router. The designated router generates a link state advertisement for the multi-access network and synchronizes the topological database with other adjacent routers in the area. The concept of designated router can thus reduce the routing traffic and compute shortest-path trees. To achieve high availability, a backup designated router is used.

- IS-IS (intermediate system to intermediate system) [RFC1195]

  IS-IS is a routing protocol designed for the OSI (Open Systems Interconnection) protocol suites. Integrated IS-IS is derived from IS-IS in order to support the IP protocol. In the Internet community, IS-IS means integrated IS-IS. In this, a link-state is advertised over a connectionless network service. IS-IS has the same basic features as OSPF. They include: link-state advertisement and maintenance of a topological database within an area, calculation of a tree of shortest-paths, generation of a routing table from a tree of shortest-paths, the area routing scheme, a designated router, and a variable length subnet mask.

o BGP4 (Border Gateway Protocol version 4) [RFC1771]

BGP4 is an exterior gateway protocol and is applied to the routing of inter-autonomous systems. A BGP speaker establishes a session with other BGP speakers and advertises routing information to them. A session may be an External BGP (EBGP) that connects two BGP speakers within different autonomous systems, or an internal BGP (IBGP) that connects two BGP speakers within a single autonomous system. Routing information is qualified with path attributes, which differentiate routes for the purpose of selecting an appropriate one from possible routes. Also, routes are grouped by the community attribute [RFC1997] [BGP-COM].

The IBGP mesh size tends to increase dramatically with the number of BGP speakers in an autonomous system. BGP can reduce the number of IBGP sessions by dividing the autonomous system into smaller autonomous systems and grouping them into a single confederation [RFC1965]. Route reflection is another way to reduce the number of IBGP sessions [RFC1966]. BGP divides the autonomous system into clusters. Each cluster establishes the IBGP full-mesh within itself, and designates one or more BGP speakers as "route reflectors," which communicate with other clusters via their route reflectors. Route reflectors in each cluster maintain path and attribute information across the autonomous system. The autonomous system still functions like a fully meshed autonomous system. On the other hand, confederations provide finer control of routing within the autonomous system by allowing for policy changes across confederation boundaries, while route reflection requires the use of identical policies.

6.3.2 Exchange of routing information within a network

The following technologies can be used for exchanging routing information within a network.

o Static routing (see section 6.3.1)

o RIP (see section 6.3.1)

o OSPF (see section 6.3.1)

o BGP (see section 6.3.1)

o Multiprotocol BGP4 [RFC2858]

BGP4 has been extended to support IPv6, IPX, and others as well as IPv4 [RFC2283]. Multiprotocol BGP4 carries routes from multiple "address families."

o Extended BGP4 [VPN-2547BIS]

Extended BGP4 is a specific extension to Multiprotocol BGP4. The notion of "VPN-IPv4 address family" is introduced in [VPN-2547BIS]. A VPN-IPv4 address is 12 bytes long and consists of an 8-byte route distinguisher (RD) and a 4-byte IPv4 address. Overlapping of the IPv4 address space among multiple NBVPNs is resolved by using different RDs. Scalable configurations can be achieved by the use of route reflectors.

6.4 QoS/SLA

The following technologies for QoS/SLA are applicable to an NBVPN.

6.4.1 ATM [AF-TM-0121.000]

Asynchronous transfer mode (ATM) provides several service categories, such as CBR (constant bit rate) service, VBR (variable bit rate) service, and GFR (guaranteed frame rate) service. CBR service is used to guarantee a static amount of bandwidth. VBR service is designed for a wide range of applications, including real-time and non-real-time applications. GFR service is designed for applications that may require a minimum rate guarantee and can benefit from accessing additional bandwidth.

6.4.2 IntServ/RSVP [RFC2205] [RFC2208] [RFC2210] [RFC2746] [RSVP-LSP]

The integrated service, or IntServ for short, is a mechanism for providing QoS/SLA by admission control. RSVP is used to reserve network resources. The network needs to maintain a state for each reservation. The number of states in the network increases in proportion to the number of concurrent reservations.

6.4.3 DiffServ [RFC2474] [RFC2475]

The differentiated service, or DiffServ for short, is a mechanism for providing QoS/SLA by differentiating traffic. Traffic entering a network is classified into several behavior aggregates at the network-edge and each is assigned a corresponding DiffServ codepoint. Within the network, traffic is treated according to its DiffServ codepoint. Some behavior aggregates have already been defined. Expedited forwarding behavior [RFC2598] guarantees the QoS, whereas assured forwarding behavior [RFC2597] differentiates traffic packet precedence values.

7. Criteria for Achieving Interoperability

(To be written)

8. Security Considerations

As described in section 2.8, if a user requires stronger security than that of the basic CUG service, it should be provided by a CPE-based security mechanism. This is because a network-based solution cannot ensure the security of access links between user sites and network.

As described in section 4.4.1, if a tunnel on the SP interworking interface is not implemented with direct circuit between IE routers and it passes through an unsecure SP, POP, NAP, or IX, then security mechanisms should be located at the edge routers. However, detailed specifications of this security mechanism depend on the implementation, so it is not discussed in this framework.

Acknowledgments

VPNs are a huge technology and without the early work of RFC2764 "A Framework for IP Based Virtual Private Networks," it would have been impossible for us to complete this framework document. We would like to thank the authors of RFC2764, especially Bryan Gleeson of Nortel Networks.

We would also like to thank Joel Halpern of Longitude Systems, Eric Rosen of Cisco Systems, and Kazuo Kobayashi for their valuable comments and suggestions.

References

[RFC2764] Gleeson, B. et al., "A Framework for IP Based Virtual Private Networks," RFC2764, February 2000.

[RFC2547] Rosen, E. and Rekhter, Y., "BGP/MPLS VPN," RFC2547, March 1999.

[RFC2684] Grossman, D. and Heinanen, J., "Multiprotocol Encapsulation over ATM Adaptation Layer 5," RFC2684, September 1999.

[RFC2685] Fox B. and Gleeson, B., "Virtual Private Networks Identifier," RFC2685, September 1999.

[VPN-2547BIS] Rosen, E. et al., "BGP/MPLS VPNs," Internet-draft <draft-rosen-rfc2547bis-02.txt>, July 2000.

[VPN-BGP-OSPF] Rosen, E., "OSPF as the PE/CE Protocol in BGP/MPLS VPNs," Internet-draft <draft-rosen-vpns-ospf-bgp-mpls-00.txt>, July 2000.

[VPN-BGP-IPSEC] De Clercq, J. et al., "BGP/IPsec VPN," Internet-draft <draft-declercq-bgp-ipsec-vpn-00.txt>, July 2000.

[VPN-BGP-VR] Ould-Brahim, H. et al., "BGP/VPN: VPN Information Discovery for Network-based VPNs," Internet-draft <draft-ouldbrahim-bgp-vpn-00.txt>, July 2000.

[VPN-VR] Ould-Brahim H. et al., "Network based IP VPN Architecture Using Virtual Routers," Internet-draft <draft-ouldbrahim-vpn-vr-01.txt>, July 2000.

[VPN-IPSEC] Lordello, C. et al, "VPN-ID-Enhanced IPSec-VPN DOI for ISAKMP," Internet-draft <draft-lordello-ipsec-vpn-doi-00.txt>, August 2000.

[VPN-INTER] Sumimoto, J. et al., "MPLS VPN Interworking" Internet-Draft <draft-sumimoto-mpls-vpn-interworking-00.txt>," February 2000.

[RFC2917] Muthukrishnan, K. and Malis, A., "A Core MPLS IP VPN Architecture," RFC2917, September 2000.

[MPLS-ARCH] Rosen E. et al., "Multiprotocol Label Switching Architecture," Internet-draft <draft-ietf-mpls-arch-07.txt>, July 2000.

[MPLS-FRAME] Callon, R. et al., "A Framework for Multiprotocol Label Switching," <draft-ietf-mpls-framework-05.txt>, September 1999.

[MPLS-ATM] Davie, B. et al., "MPLS using LDP and ATM VC Switching," Internet-draft <draft-ietf-mpls-atm-04.txt>, June 2000.

[MPLS-DIFF] Le Faucheur, F. et al., "MPLS Support of Differentiated Services," Internet-draft <draft-ietf-mpls-diff-ext-07.txt>, August 2000.

[MPLS-GMNCL] GMN-CL home page:
http://www.gmncl.ecl.ntt.co.jp/top_e.html

[RFC2784] Farinacci, D. et al., "Generic Routing Encapsulation (GRE)," RFC2784, March 2000.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE," RFC2890, September 2000.

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

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

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

[RFC2409] Harkins, D. and Carrel, D., "The Internet Key Exchange (IKE)," RFC2409, November 1998.

[RFC2453] Malkin, G., "RIP Version 2," RFC2453, November 1994.

[RFC2328] Moy, J., "OSPF Version 2," RFC2328, April 1998.

[RFC1195] Callon, R., "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments," RFC1195, December 1990.

[RFC1771] Rekhter, Y. and Li, T., "A Border Gateway Protocol 4 (BGP-4)," RFC1771, March 1995.

[RFC1965] Traina, P., "Autonomous System Confederations for BGP," RFC1965, June 1996.

[RFC1966] Bates, T., "BGP Route Reflection: An alternative to full mesh IBGP," RFC 1966, June 1996.

[RFC1997] Chandra, R., Traina, P., and Li, T., "BGP Communities Attribute," RFC1997, August 1996.

[RFC2858] Bates, T., Chandra, R., Katz, D., and Rekhter, Y., "Multiprotocol Extensions for BGP-4," RFC2283, February 1998.

[BGP-COM] Ramachandra, S., "BGP Extended Communities Attribute," Internet-draft <draft-ramachandra-bgp-ext-communities-04.txt>, May 2000.

[AF-TM-0121.000] "Traffic Management Specification Version 4.1," ATM Forum, March 1999.

[RFC2205] Braden, R. et al., "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification," RFC2205, September 1997.

[RFC2208] Mankin, A. et al., "Resource ReSerVation Protocol (RSVP) Version 1 Applicability Statement Some Guidelines on Deployment," RFC2208, September 1997.

[RFC2210] Wroclawski, J., "The Use of RSVP with IETF Integrated Services," RFC2210, September 1997.

[RFC2211] Wroclawski, J., "Specification of the Controlled-Load Network Element Service," RFC2211, September 1997.

[RFC2212] Shenker, S., Partridge, C., and Guerin, R., "Specification of Guaranteed Quality of Service," RFC2212, September 1997.

[RFC2207] Berger, L. and O'Malley, T., "RSVP Extensions for IPSEC Data Flows," RFC2207, September 1997.

[RFC2746] Terzis, A. et al., "RSVP Operation Over IP Tunnels," RFC2746, January 2000.

[RSVP-LSP] Awduche, D. et al., "Extensions to RSVP for LSP Tunnels," Internet-draft <draft-ietf-mpls-rsvp-lsp-tunnel-07.txt>, August 2000.

[RFC2474] Nichols, K. et al., "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," RFC2474, December 1998.

[RFC2475] Blake S. et al., "An architecture for Differentiated Services," RFC2475, December 1998.

[RFC2597] Heinanen, J. et al., "Assured Forwarding PHB Group," RFC2597, June 1999.

[RFC2598] Jacobsen, V. et al., "An Expedited Forwarding PHB," RFC2598, June 1999.

[RFC2983] Black, D., "Differentiated Services and Tunnels," RFC2983, October 2000.

10. Authors' addresses

Muneyoshi Suzuki
NTT Information Sharing Platform Labs.
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: suzuki.muneyoshi@lab.ntt.co.jp

Junichi Sumimoto
NTT Information Sharing Platform Labs.
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: sumimoto.junichi@lab.ntt.co.jp

Andrew G. Malis
Vivace Networks, Inc.
2730 Orchard Parkway
San Jose, CA 95134, USA
Email: Andy.Malis@vivacenetworks.com

Karthik Muthukrishnan
Lucent Technologies
1 Robbins Road
Westford, MA 01886, USA
Email: mkarthik@lucent.com

Kosei Suzuki
NTT Information Sharing Platform Labs.
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: suzuki.kosei@lab.ntt.co.jp

Hiroshi Kurakami
NTT Information Sharing Platform Labs.
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: kurakami.hiroshi@lab.ntt.co.jp

Takafumi Hamano
NTT Information Sharing Platform Labs.
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: hamano.takafumi@lab.ntt.co.jp

Naoto Makinae
NTT Information Sharing Platform Labs.
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: makinae.naoto@lab.ntt.co.jp

Kenichi Kitami
NTT Information Sharing Laboratory Group
3-9-11, Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: kitami.kenichi@lab.ntt.co.jp