Independent Submission
Request for Comments: 5558
Category: Informational
ISSN: 2070-1721

F. Templin, Ed.
Boeing Research & Technology

February 2010

Virtual Enterprise Traversal (VET)

Abstract

Enterprise networks connect routers over various link types, and may also connect to provider networks and/or the global Internet. Enterprise network nodes require a means to automatically provision IP addresses/prefixes and support internetworking operation in a wide variety of use cases including Small Office, Home Office (SOHO) networks, Mobile Ad hoc Networks (MANETs), multi-organizational corporate networks and the interdomain core of the global Internet itself. This document specifies a Virtual Enterprise Traversal (VET) abstraction for autoconfiguration and operation of nodes in enterprise networks.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This is a contribution to the RFC Series, independently of any other RFC stream. The RFC Editor has chosen to publish this document at its discretion and makes no statement about its value for implementation or deployment. Documents approved for publication by the RFC Editor are not a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc5558.

IESG Note

This RFC is not a candidate for any level of Internet Standard. The IETF disclaims any knowledge of the fitness of this RFC for any purpose and in particular notes that the decision to publish is not based on IETF review for such things as security, congestion control, or inappropriate interaction with deployed protocols. The RFC Editor has chosen to publish this document at its discretion. Readers of this RFC should exercise caution in evaluating its value for implementation and deployment. See RFC 3932 for more information.

Note that the IETF AUTOCONF Working Group is working on a similar protocol solution that may become available in the future.

Table of Contents

1. Introduction

Enterprise networks [RFC4852] connect routers over various link types (see [RFC4861], Section 2.2). The term "enterprise network" in this context extends to a wide variety of use cases and deployment scenarios. For example, an "enterprise" can be as small as a SOHO network, as complex as a multi-organizational corporation, or as large as the global Internet itself. Mobile Ad hoc Networks (MANETs) [RFC2501] can also be considered as a challenging example of an enterprise network, in that their topologies may change dynamically over time and that they may employ little/no active management by a centralized network administrative authority. These specialized characteristics for MANETs require careful consideration, but the same principles apply equally to other enterprise network scenarios.

This document specifies a Virtual Enterprise Traversal (VET) abstraction for autoconfiguration and internetworking operation, where addresses of different scopes may be assigned on various types of interfaces with diverse properties. Both IPv4 [RFC0791] and IPv6 [RFC2460] are discussed within this context. The use of standard DHCP [RFC2131] [RFC3315] and neighbor discovery [RFC0826] [RFC1256] [RFC4861] mechanisms is assumed unless otherwise specified.

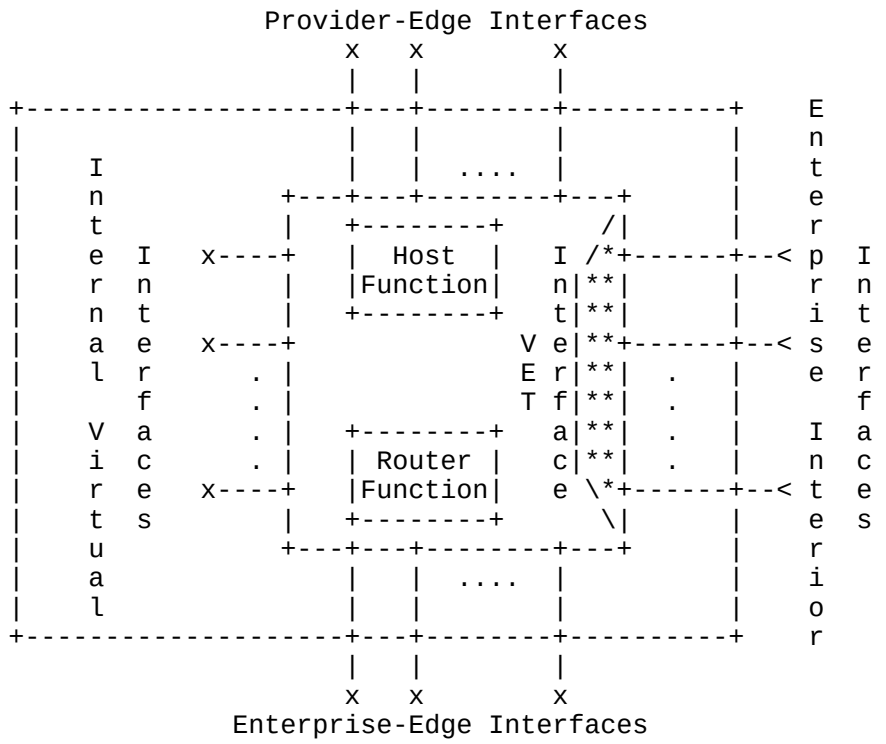

Figure 1: Enterprise Router (ER) Architecture

Figure 1 above depicts the architectural model for an Enterprise Router (ER). As shown in the figure, an ER may have a variety of interface types including enterprise-edge, enterprise-interior, provider-edge, internal-virtual, as well as VET interfaces used for IP-in-IP encapsulation. The different types of interfaces are defined, and the autoconfiguration mechanisms used for each type are specified. This architecture applies equally for MANET routers, in which enterprise-interior interfaces correspond to the wireless multihop radio interfaces typically associated with MANETs. Out of scope for this document is the autoconfiguration of provider interfaces, which must be coordinated in a manner specific to the service provider's network.

Enterprise networks must have a means for supporting both Provider-Independent (PI) and Provider-Aggregated (PA) IP prefixes. This is especially true for enterprise scenarios that involve mobility and multihoming. Also in scope are ingress filtering for multihomed sites, adaptation based on authenticated ICMP feedback from on-path routers, effective tunnel path MTU mitigations, and routing scaling suppression as required in many enterprise network scenarios.

Recognizing that one size does not fit all, the VET specification provides adaptable mechanisms that address these issues, and more, in a wide variety of enterprise network use cases.

VET represents a functional superset of 6over4 [RFC2529] and Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) [RFC5214], and it further supports additional encapsulations such as IPsec [RFC4301], Subnetwork Encapsulation and Adaptation Layer (SEAL) [RFC5320], etc. Together, these technologies serve as functional building blocks for a new Internetworking architecture known as Routing and Addressing in Networks with Global Enterprise Recursion [RFC5720][RANGERS].

The VET principles can be either directly or indirectly traced to the deliberations of the ROAD group in January 1992, and also to still earlier works including NIMROD [RFC1753], the Catenet model for internetworking [CATENET] [IEN48] [RFC2775], etc. [RFC1955] captures the high-level architectural aspects of the ROAD group deliberations in a "New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG".

VET is related to the present-day activities of the IETF AUTOCONF, DHC, IPv6, MANET, and v6OPS working groups, as well as the IRTF RRG working group.

2. Terminology

The mechanisms within this document build upon the fundamental principles of IP-in-IP encapsulation. The terms "inner" and "outer" are used to, respectively, refer to the innermost IP {address, protocol, header, packet, etc.} *before* encapsulation, and the outermost IP {address, protocol, header, packet, etc.} *after* encapsulation. VET also allows for inclusion of "mid-layer" encapsulations between the inner and outer layers, including IPsec [RFC4301], the Subnetwork Encapsulation and Adaptation Layer (SEAL) [RFC5320], etc.

The terminology in the normative references apply; the following terms are defined within the scope of this document:

subnetwork
: the same as defined in [RFC3819].

enterprise
: the same as defined in [RFC4852]. An enterprise is also understood to refer to a cooperative networked collective with a commonality of business, social, political, etc. interests.

Minimally, the only commonality of interest in some enterprise network scenarios may be the cooperative provisioning of connectivity itself.

site
a logical and/or physical grouping of interfaces that connect a topological area less than or equal to an enterprise in scope. A site within an enterprise can, in some sense, be considered as an enterprise unto itself.

Mobile Ad hoc Network (MANET)
a connected topology of mobile or fixed routers that maintain a routing structure among themselves over dynamic links, where a wide variety of MANETs share common properties with enterprise networks. The characteristics of MANETs are defined in [RFC2501], Section 3.

enterprise/site/MANET
throughout the remainder of this document, the term "enterprise" is used to collectively refer to any of enterprise/site/MANET, i.e., the VET mechanisms and operational principles can be applied to enterprises, sites, and MANETs of any size or shape.

Enterprise Router (ER)
As depicted in Figure 1, an Enterprise Router (ER) is a fixed or mobile router that comprises a router function, a host function, one or more enterprise-interior interfaces, and zero or more internal virtual, enterprise-edge, provider-edge, and VET interfaces. At a minimum, an ER forwards outer IP packets over one or more sets of enterprise-interior interfaces, where each set connects to a distinct enterprise.

Enterprise Border Router (EBR)
an ER that connects edge networks to the enterprise and/or connects multiple enterprises together. An EBR is a tunnel endpoint router, and it configures a separate VET interface over each set of enterprise-interior interfaces that connect the EBR to each distinct enterprise. In particular, an EBR may configure multiple VET interfaces -- one for each distinct enterprise. All EBRs are also ERs.

Enterprise Border Gateway (EBG)
an EBR that connects VET interfaces configured over child enterprises to a provider network -- either directly via a provider-edge interface or indirectly via another VET interface configured over a parent enterprise. EBRs may act as EBGs on some VET interfaces and as ordinary EBRs on other VET interfaces. All EBGs are also EBRs.

enterprise-interior interface
: an ER's attachment to a link within an enterprise. Packets sent over enterprise-interior interfaces may be forwarded over multiple additional enterprise-interior interfaces within the enterprise before they are forwarded via an enterprise-edge interface, provider-edge interface, or a VET interface configured over a different enterprise. Enterprise-interior interfaces connect laterally within the IP network hierarchy.

enterprise-edge interface
: an EBR's attachment to a link (e.g., an Ethernet, a wireless personal area network, etc.) on an arbitrarily complex edge network that the EBR connects to an enterprise and/or provider network. Enterprise-edge interfaces connect to lower levels within the IP network hierarchy.

provider-edge interface
: an EBR's attachment to the Internet or to a provider network outside of the enterprise via which the Internet can be reached. Provider-edge interfaces connect to higher levels within the IP network hierarchy.

internal-virtual interface
: an interface that is internal to an EBR and does not in itself directly attach to a tangible physical link, e.g., an Ethernet cable. Examples include a loopback interface, a virtual LAN interface, or some form of tunnel interface.

Virtual Enterprise Traversal (VET)
: an abstraction that uses IP-in-IP encapsulation to create an overlay that spans an enterprise in a single (inner) IP hop.

VET interface
: an EBR's tunnel virtual interface used for Virtual Enterprise Traversal. The EBR configures a VET interface over a set of underlying interfaces belonging to the same enterprise. When there are multiple distinct enterprises (each with their own distinct set of underlying interfaces), the EBR configures a separate VET interface over each set of underlying interfaces, i.e., the EBR configures multiple VET interfaces.

  The VET interface encapsulates each inner IP packet in any mid-layer headers plus an outer IP header, then it forwards it on an underlying interface such that the Time to Live (TTL) / Hop Limit in the inner header is not decremented as the packet traverses the enterprise. The VET interface therefore presents an automatic tunneling abstraction that represents the enterprise as a single IP hop.

VET interfaces in non-multicast environments are Non-Broadcast, Multiple Access (NBMA); VET interfaces in multicast environments are multicast capable.

VET host
: any node (host or router) that configures a VET interface for host operation only. Note that a single node may configure some of its VET interfaces as host interfaces and others as router interfaces.

VET node
: any node that configures and uses a VET interface.

Provider-Independent (PI) prefix
: an IPv6 or IPv4 prefix (e.g., 2001:DB8::/48, 192.0.2/24, etc.) that is either self-generated by an ER or delegated to an enterprise by a registry.

Provider Aggregated (PA) prefix
: an IPv6 or IPv4 prefix that is delegated to an enterprise by a provider network.

Routing Locator (RLOC)
: a non-link-local IPv4 or IPv6 address taken from a PI/PA prefix that can appear in enterprise-interior and/or interdomain routing tables. Global-scope RLOC prefixes are delegated to specific enterprises and are routable within both the enterprise-interior and interdomain routing regions. Enterprise-local-scope RLOC prefixes (e.g., IPv6 Unique Local Addresses [RFC4193], IPv4 privacy addresses [RFC1918], etc.) are self-generated by individual enterprises and routable only within the enterprise-interior routing region.

  ERs use RLOCs for operating the enterprise-interior routing protocol and for next-hop determination in forwarding packets addressed to other RLOCs. End systems use RLOCs as addresses for communications between endpoints within the same enterprise. VET interfaces treat RLOCs as *outer* IP addresses during IP-in-IP encapsulation.

Endpoint Interface iDentifier (EID)
: an IPv4 or IPv6 address taken from a PI/PA prefix that is routable within an enterprise-edge or VET overlay network scope, and may also appear in enterprise-interior and/or interdomain mapping tables. EID prefixes are typically separate and distinct from any RLOC prefix space.

Edge network routers use EIDs for operating the enterprise-edge or VET overlay network routing protocol and for next-hop determination in forwarding packets addressed to other EIDs. End systems use EIDs as addresses for communications between endpoints either within the same enterprise or within different enterprises. VET interfaces treat EIDs as *inner* IP addresses during IP-in-IP encapsulation.

The following additional acronyms are used throughout the document:

CGA	- Cryptographically Generated Address
DHCP(v4, v6)	- Dynamic Host Configuration Protocol
FIB	- Forwarding Information Base
ISATAP	- Intra-Site Automatic Tunnel Addressing Protocol
NBMA	- Non-Broadcast, Multiple Access
ND	- Neighbor Discovery
PIO	- Prefix Information Option
PRL	- Potential Router List
PRLNAME	- Identifying name for the PRL (default is "isatap")
RIO	- Route Information Option
RS/RA	- IPv6 ND Router Solicitation/Advertisement
SEAL	- Subnetwork Encapsulation and Adaptation Layer
SLAAC	- IPv6 StateLess Address AutoConfiguation

3. Enterprise Characteristics

Enterprises consist of links that are connected by Enterprise Routers (ERs) as depicted in Figure 1. ERs typically participate in a routing protocol over enterprise-interior interfaces to discover routes that may include multiple Layer 2 or Layer 3 forwarding hops. Enterprise Border Routers (EBRs) are ERs that connect edge networks to the enterprise and/or join multiple enterprises together. Enterprise Border Gateways (EBGs) are EBRs that either directly or indirectly connect enterprises to provider networks.

An enterprise may be as simple as a small collection of ERs and their attached edge networks; an enterprise may also contain other enterprises and/or be a subnetwork of a larger enterprise. An enterprise may further encompass a set of branch offices and/or nomadic hosts connected to a home office over one or several service providers, e.g., through Virtual Private Network (VPN) tunnels.

Enterprises that comprise link types with sufficiently similar properties (e.g., Layer 2 (L2) address formats, maximum transmission units (MTUs), etc.) can configure a sub-IP layer routing service such that IP sees the enterprise as an ordinary shared link the same as for a (bridged) campus LAN. In that case, a single IP hop is sufficient to traverse the enterprise without IP layer encapsulation.

Enterprises that comprise link types with diverse properties and/or configure multiple IP subnets must also provide a routing service that operates as an IP layer mechanism. In that case, multiple IP hops may be necessary to traverse the enterprise such that care must be taken to avoid multi-link subnet issues [RFC4903].

Conceptually, an ER embodies both a host function and router function. The host function supports Endpoint Interface iDentifier (EID)-based and/or Routing LOCator (RLOC)-based communications according to the weak end-system model [RFC1122]. The router function engages in the enterprise-interior routing protocol, connects any of the ER's edge networks to the enterprise, and may also connect the enterprise to provider networks (see Figure 1).

In addition to other interface types, VET nodes configure VET interfaces that view all other VET nodes in an enterprise as single-hop neighbors attached to a virtual link. VET nodes configure a separate VET interface for each distinct enterprise to which they connect, and discover other EBRs on each VET interface that can be used for forwarding packets to off-enterprise destinations.

For each distinct enterprise, an enterprise trust basis must be established and consistently applied. For example, in enterprises in which EBRs establish symmetric security associations, mechanisms such as IPsec [RFC4301] can be used to assure authentication and confidentiality. In other enterprise network scenarios, asymmetric securing mechanisms such as SEcure Neighbor Discovery (SEND) [RFC3971] may be necessary to authenticate exchanges based on trust anchors.

Finally, in enterprises with a centralized management structure (e.g., a corporate campus network), the enterprise name service and a synchronized set of EBGs can provide infrastructure support for virtual enterprise traversal. In that case, the EBGs can provide a "default mapper" [APT] service used for short-term packet forwarding until EBR neighbor relationships can be established. In enterprises with a distributed management structure (e.g., MANETs), peer-to-peer coordination between the EBRs themselves may be required. Recognizing that various use cases will entail a continuum between a fully distributed and fully centralized approach, the following sections present the mechanisms of Virtual Enterprise Traversal as they apply to a wide variety of scenarios.

4. Autoconfiguration

ERs, EBRs, EBGs, and VET hosts configure themselves for operation as specified in the following subsections.

4.1. Enterprise Router (ER) Autoconfiguration

ERs configure enterprise-interior interfaces and engage in any routing protocols over those interfaces.

When an ER joins an enterprise, it first configures a unique IPv6 link-local address on each enterprise-interior interface and configures an IPv4 link-local address on each enterprise-interior interface that requires an IPv4 link-local capability. IPv6 link-local address generation mechanisms that provide sufficient uniqueness include Cryptographically Generated Addresses (CGAs) [RFC3972], IPv6 Privacy Addresses [RFC4941], StateLess Address AutoConfiguration (SLAAC) using EUI-64 interface identifiers [RFC4291] [RFC4862], etc. The mechanisms specified in [RFC3927] provide an IPv4 link-local address generation capability.

Next, the ER configures an RLOC on each of its enterprise-interior interfaces and engages in any routing protocols on those interfaces. The ER can configure an RLOC via explicit management, DHCP autoconfiguration, pseudo-random self-generation from a suitably large address pool, or through an alternate autoconfiguration mechanism.

Alternatively (or in addition), the ER can request RLOC prefix delegations via an automated prefix delegation exchange over an enterprise-interior interface and can assign the prefix(es) on enterprise-edge interfaces. In that case, the ER can use an RLOC assigned to an enterprise-edge interface for enterprise-interior routing protocol operation and next-hop determination purposes. Note that in some cases, the same enterprise-edge interfaces may assign both RLOC and an EID addresses if there is a means for source address selection. In other cases (e.g., for separation of security domains), RLOCs and EIDs must be assigned on separate sets of enterprise-edge interfaces.

Self-generation of RLOCs for IPv6 can be from a large IPv6 local-use address range, e.g., IPv6 Unique Local Addresses [RFC4193]. Self-generation of RLOCs for IPv4 can be from a large IPv4 private address range (e.g., [RFC1918]). When self-generation is used alone, the ER must continuously monitor the RLOCs for uniqueness, e.g., by monitoring the routing protocol.

DHCP generation of RLOCs may require support from relays within the enterprise. For DHCPv6, relays that do not already know the RLOC of a server within the enterprise forward requests to the 'All_DHCP_Servers' site-scoped IPv6 multicast group [RFC3315]. For DHCPv4, relays that do not already know the RLOC of a server within the enterprise forward requests to the site-scoped IPv4 multicast

group address 'All_DHCPv4_Servers', which should be set to 239.255.2.1 unless an alternate multicast group for the site is known. DHCPv4 servers that delegate RLOCs should therefore join the 'All_DHCPv4_Servers' multicast group and service any DHCPv4 messages received for that group.

A combined approach using both DHCP and self-generation is also possible when the ER configures both a DHCP client and relay that are connected, e.g., via a pair of back-to-back connected Ethernet interfaces, a tun/tap interface, a loopback interface, inter-process communication, etc. The ER first self-generates a temporary RLOC used only for the purpose of procuring an actual RLOC taken from a disjoint addressing range. The ER then engages in the routing protocol and performs a DHCP client/relay exchange using the temporary RLOC as the address of the relay. When the DHCP server delegates an actual RLOC address/prefix, the ER abandons the temporary RLOC and re-engages in the routing protocol using an RLOC taken from the delegation.

In some enterprise use cases (e.g., MANETs), assignment of RLOCs on enterprise-interior interfaces as singleton addresses (i.e., as addresses with /32 prefix lengths for IPv4, and as addresses with /128 prefix lengths for IPv6) may be necessary to avoid multi-link subnet issues.

4.2. Enterprise Border Router (EBR) Autoconfiguration

EBRs are ERs that configure VET interfaces over distinct sets of underlying interfaces belonging to the same enterprise; an EBR can connect to multiple enterprises, in which case it would configure multiple VET interfaces. In addition to the ER autoconfiguration procedures specified in Section 4.1, EBRs perform the following autoconfiguration operations.

4.2.1. VET Interface Autoconfiguration

VET interface autoconfiguration entails:

1) interface initialization,
2) EBG discovery and enterprise identification, and
3) EID configuration.

These functions are specified in the following sections.

4.2.1.1. Interface Initialization

EBRs configure a VET interface over a set of underlying interfaces belonging to the same enterprise, where the VET interface presents a virtual-link abstraction in which all EBRs in the enterprise appear as single-hop neighbors through the use of IP-in-IP encapsulation. After the EBR configures a VET interface, it initializes the interface and assigns an IPv6 link-local address and an IPv4 link-local address if necessary.

When IPv6 and IPv4 are used as the inner/outer protocols (respectively), the EBR autoconfigures an ISATAP link-local address ([RFC5214], Section 6.2) on the VET interface to support packet forwarding and operation of the IPv6 neighbor discovery protocol. The ISATAP link-local address embeds an IPv4 RLOC, and need not be checked for uniqueness since the IPv4 RLOC itself is managed for uniqueness (see Section 4.1).

Link-local address configuration for other inner/outer IP protocol combinations is through administrative configuration or through an unspecified alternate method. Link-local address configuration for other inner/outer IP protocol combinations may not be necessary if an EID can be configured through other means (see Section 4.2.1.3).

After the EBR initializes a VET interface, it can communicate with other VET nodes as single-hop neighbors on the VET interface from the viewpoint of the inner IP protocol.

4.2.1.2. Enterprise Border Gateway Discovery and Enterprise Identification

The EBR next discovers a list of EBGs for each of its VET interfaces. The list can be discovered through information conveyed in the routing protocol, through the Potential Router List (PRL) discovery mechanisms outlined in Section 8.3.2 of [RFC5214], through DHCP options, etc. In multicast-capable enterprises, EBRs can also listen for advertisements on the 'rasadv' [RASADV] multicast group address.

In particular, whether or not routing information is available, the EBR can discover the list of EBGs by resolving an identifying name for the PRL ('PRLNAME') formed as 'hostname.domainname', where 'hostname' is an enterprise-specific name string and 'domainname' is an enterprise-specific DNS suffix. The EBR discovers 'PRLNAME' through manual configuration, a DHCP option, 'rasadv' protocol advertisements, link-layer information (e.g., an IEEE 802.11 Service Set Identifier (SSID)), or through some other means specific to the enterprise. In the absence of other information, the EBR sets the

'hostname' component of 'PRLNAME' to "isatap" and sets the 'domainname' component only if an enterprise-specific DNS suffix "example.com" is known (e.g., as "isatap.example.com").

The global Internet interdomain routing core represents a specific example of an enterprise network scenario, albeit on an enormous scale. The 'PRLNAME' assigned to the global Internet interdomain routing core is "isatap.net".

After discovering 'PRLNAME', the EBR can discover the list of EBGs by resolving 'PRLNAME' to a list of RLOC addresses through a name service lookup. For centrally managed enterprises, the EBR resolves 'PRLNAME' using an enterprise-local name service (e.g., the enterprise-local DNS). For enterprises with a distributed management structure, the EBR resolves 'PRLNAME' using Link-Local Multicast Name Resolution (LLMNR) [RFC4795] over the VET interface. In that case, all EBGs in the PRL respond to the LLMNR query, and the EBR accepts the union of all responses.

Each distinct enterprise must have a unique identity that EBRs can use to uniquely discern their enterprise affiliations. 'PRLNAME' as well as the RLOCs of EBGs and the IP prefixes they aggregate serve as an identifier for the enterprise.

4.2.1.3. EID Configuration

After EBG discovery, the EBR configures EIDs on its VET interfaces. When IPv6 and IPv4 are used as the inner/outer protocols (respectively), the EBR autoconfigures EIDs as specified in Section 5.4.1. In particular, the EBR acts as a host on its VET interfaces for router and prefix discovery purposes but acts as a router on its VET interfaces for routing protocol operation and packet forwarding purposes.

EID configuration for other inner/outer IP protocol combinations is through administrative configuration or through an unspecified alternate method; in some cases, such EID configuration can be performed independently of EBG discovery.

4.2.2. Provider-Aggregated (PA) EID Prefix Autoconfiguration

EBRs can acquire Provider-Aggregated (PA) EID prefixes through autoconfiguration exchanges with EBGs over VET interfaces, where each EBG may be configured as either a DHCP relay or DHCP server.

For IPv4 EIDs, the EBR acquires prefixes via an automated IPv4 prefix delegation exchange, explicit management, etc.

For IPv6 EIDs, the EBR acquires prefixes via DHCPv6 Prefix Delegation exchanges. In particular, the EBR (acting as a requesting router) can use DHCPv6 prefix delegation [RFC3633] over the VET interface to obtain IPv6 EID prefixes from the server (acting as a delegating router).

The EBR obtains prefixes using either a 2-message or 4-message DHCPv6 exchange [RFC3315]. For example, to perform the 2-message exchange, the EBR's DHCPv6 client forwards a Solicit message with an IA_PD option to its DHCPv6 relay, i.e., the EBR acts as a combined client/relay (see Section 4.1). The relay then forwards the message over the VET interface to an EBG, which either services the request or relays it further. The forwarded Solicit message will elicit a reply from the server containing PA IPv6 prefix delegations.

The EBR can propose a specific prefix to the DHCPv6 server per Section 7 of [RFC3633], e.g., if a prefix delegation hint is available. The server will check the proposed prefix for consistency and uniqueness, then return it in the reply to the EBR if it was able to perform the delegation.

After the EBR receives PA prefix delegations, it can provision the prefixes on enterprise-edge interfaces as well as on other VET interfaces for which it is configured as an EBG. It can also provision the prefixes on enterprise-interior interfaces as long as other nodes on those interfaces unambiguously associate the prefixes with the EBR.

4.2.3. Provider-Independent (PI) EID Prefix Autoconfiguration

Independent of any PA prefixes, EBRs can acquire and use Provider-Independent (PI) EID prefixes that are self-configured (e.g., using [RFC4193], etc.) and/or delegated by a registration authority (e.g., using [CENTRL-ULA], etc.). When an EBR acquires a PI prefix, it must also obtain credentials that it can use to prove prefix ownership when it registers the prefixes with EBGs within an enterprise (see Sections 5.4 and 5.5).

After the EBR receives PI prefix delegations, it can provision the prefixes on enterprise-edge interfaces as well as on other VET interfaces for which it is configured as an EBG. It can also provision the prefixes on enterprise-interior interfaces as long as other nodes on those interfaces can unambiguously associate the prefixes with the EBR.

The minimum-sized IPv6 PI prefix that an EBR may acquire is a /56.

The minimum-sized IPv4 PI prefix that an EBR may acquire is a /24.

4.3. Enterprise Border Gateway (EBG) Autoconfiguration

EBGs are EBRs that connect child enterprises to provider networks via provider-edge interfaces and/or via VET interfaces configured over parent enterprises. EBGs autoconfigure their provider-edge interfaces in a manner that is specific to the provider connections, and they autoconfigure their VET interfaces that were configured over parent enterprises, using the EBR autoconfiguration procedures specified in Section 4.2.

For each of its VET interfaces configured over a child enterprise, the EBG initializes the interface and configures an EID the same as for an ordinary EBR (see Section 4.2.1). It must then arrange to add one or more of its RLOCs associated with the child enterprise to the PRL, and it must maintain these resource records in accordance with [RFC5214], Section 9. In particular, for each VET interface configured over a child enterprise, the EBG adds the RLOCs to name-service resource records for 'PRLNAME'.

EBGs respond to LLMNR queries for 'PRLNAME' on VET interfaces configured over child enterprises with a distributed management structure.

EBGs configure a DHCP relay/server on VET interfaces configured over child enterprises that require DHCP services.

To avoid looping, EBGs must not configure a default route on a VET interface configured over a child interface.

4.4. VET Host Autoconfiguration

Nodes that cannot be attached via an EBR's enterprise-edge interface (e.g., nomadic laptops that connect to a home office via a Virtual Private Network (VPN)) can instead be configured for operation as a simple host connected to the VET interface. Such VET hosts perform the same VET interface autoconfiguration procedures as specified for EBRs in Section 4.2.1, but they configure their VET interfaces as host interfaces (and not router interfaces). VET hosts can then send packets to the EID addresses of other hosts on the VET interface, or to off-enterprise EID destinations via a next-hop EBR.

Note that a node may be configured as a host on some VET interfaces and as an EBR/EBG on other VET interfaces.

5. Internetworking Operation

Following the autoconfiguration procedures specified in Section 4, ERs, EBRs, EBGs, and VET hosts engage in normal internetworking operations as discussed in the following sections.

5.1. Routing Protocol Participation

Following autoconfiguration, ERs engage in any RLOC-based IP routing protocols and forward IP packets with RLOC addresses. EBRs can additionally engage in any EID-based IP routing protocols and forward IP packets with EID addresses. Note that the EID-based IP routing domains are separate and distinct from any RLOC-based IP routing domains.

5.2. RLOC-Based Communications

When permitted by policy and supported by routing, end systems can avoid VET interface encapsulation through communications that directly invoke the outer IP protocol using RLOC addresses instead of EID addresses. End systems can use source address selection rules to determine whether to use EID or RLOC addresses based on, e.g., name-service records.

5.3. EID-Based Communications

In many enterprise scenarios, the use of EID-based communications (i.e., instead of RLOC-based communications) may be necessary and/or beneficial to support address scaling, NAT avoidance, security domain separation, site multihoming, traffic engineering, etc.

The remainder of this section discusses internetworking operation for EID-based communications using the VET interface abstraction.

5.4. IPv6 Router Discovery and Prefix Registration

The following sections discuss router and prefix discovery considerations for the case of IPv6 as the inner IP protocol.

5.4.1. IPv6 Router and Prefix Discovery

EBGs follow the router and prefix discovery procedures specified in [RFC5214], Section 8.2. They send solicited RAs over VET interfaces for which they are configured as gateways with default router lifetimes, with PIOs that contain PA prefixes for SLAAC, and with any other required options/parameters. The RAs can also include PIOs with the 'L' bit set to 0 and with a prefix such as '2001: DB8::/48'

as a hint of an aggregated prefix from which the EBG is willing to delegate longer PA prefixes. When PIOs that contain PA prefixes for SLAAC are included, the 'M' flag in the RA should also be set to 0.

VET nodes follow the router and prefix discovery procedures specified in [RFC5214], Section 8.3. They discover EBGs within the enterprise as specified in Section 4.2.1.2, then perform RS/RA exchanges with the EBGs to establish and maintain default routes. In particular, the VET node sends unicast RS messages to EBGs over its VET interface(s) to receive RAs. Depending on the enterprise network trust basis, VET nodes may be required to use SEND to secure the RS/RA exchanges.

When the VET node receives an RA, it authenticates the message, then configures a default route based on the Router Lifetime. If the RA contains Prefix Information Options (PIOs) with the 'A' and 'L' bits set to 1, the VET node also autoconfigures IPv6 addresses from the advertised prefixes using SLAAC and assigns them to the VET interface. Thereafter, the VET node accepts packets that are forwarded by EBGs for which it has current default routing information (i.e., ingress filtering is based on the default router trust relationship rather than a prefix-specific ingress filter entry).

In enterprises in which DHCPv6 is preferred, DHCPv6 exchanges between EBRs and EBGs may be sufficient to convey default router and prefix information. In that case, RS/RA exchanges may not be necessary.

5.4.2. IPv6 PA Prefix Registration

After an EBR discovers default routes, it can use DHCP prefix delegation to obtain PA prefixes via an EBG as specified in Section 4.2.2. The DHCP server ensures that the delegations are unique and that the EBG's router function will forward IP packets over the VET interface to the correct EBR. In particular, the EBG must register and track the PA prefixes that are delegated to each EBR.

The PA prefix registrations remain active in the EBGs as long as the EBR continues to issue DHCP renewals over the VET interface before lease lifetimes expire. The lease lifetime also keeps the delegation state active even if communications between the EBR and DHCP server are disrupted for a period of time (e.g., due to an enterprise network partition) before being reestablished (e.g., due to an enterprise network merge).

5.4.3. IPv6 PI Prefix Registration

After an EBR discovers default routes, it must register its PI prefixes by sending RAs to a set of one or more EBGs with Route Information Options (RIOs) [RFC4191] that contain the EBR's PI prefixes. Each RA must include the RLOC of an EBG as the outer IP destination address and a link-local address assigned to the VET interface as the inner IP destination address. For enterprises that use SEND, the RAs also include a CGA link-local inner source address, SEND credentials, plus any certificates needed to prove ownership of the PI prefixes. The EBR additionally tracks the set of EBGs to which it sends RAs so that it can send subsequent RAs to the same set.

When the EBG receives the RA, it first authenticates the message; if the authentication fails, the EBG discards the RA. Otherwise, the EBG installs the PI prefixes with their respective lifetimes in its Forwarding Information Base (FIB) and configures them for both ingress filtering [RFC3704] and forwarding purposes. In particular, the EBG configures the FIB entries as ingress filter rules to accept packets received on the VET interface that have a source address taken from the PI prefixes. It also configures the FIB entries to forward packets received on other interfaces with a destination address taken from the PI prefixes to the EBR that registered the prefixes on the VET interface.

The EBG then publishes the PI prefixes in a distributed database (e.g., in a private instance of a routing protocol in which only EBGs participate, via an automated name-service update mechanism [RFC3007], etc.). For enterprises that are managed under a centralized administrative authority, the EBG also publishes the PI prefixes in the enterprise-local name-service (e.g., the enterprise-local DNS [RFC1035]).

In particular, the EBG publishes each /56 prefix taken from the PI prefixes as a separate Fully Qualified Domain Name (FQDN) that consists of a sequence of 14 nibbles in reverse order (i.e., the same as in [RFC3596], Section 2.5) followed by the string 'ip6' followed by the string 'PRLNAME'. For example, when 'PRLNAME' is "isatap.example.com", the EBG publishes the prefix '2001:DB8::/56' as:

'0.0.0.0.0.0.8.b.d.0.1.0.0.2.ip6.isatap.example.com'.

The EBG includes the outer RLOC source address of the RA (e.g., in a DNS A resource record) in each prefix publication. For enterprises that use SEND, the EBG also includes the inner IPv6 CGA source address (e.g., in a DNS AAAA record) in each prefix publication. If

the prefix was already installed in the distributed database, the EBG instead adds the outer RLOC source address (e.g., in an additional DNS A record) to the preexisting publication to support PI prefixes that are multihomed. For enterprises that use SEND, this latter provision requires all EBRs of a multihomed site that advertise the same PI prefixes in RAs to use the same CGA and the same SEND credentials.

After the EBG authenticates the RA and publishes the PI prefixes, it next acts as a Neighbor Discovery proxy (NDProxy) [RFC4389] on the VET interfaces configured over any of its parent enterprises, and it relays a proxied RA to the EBGs on those interfaces. (For enterprises that use SEND, the EBG additionally acts as a SEcure Neighbor Discovery Proxy (SENDProxy) [SEND-PROXY].) EBGs in parent enterprises that receive the proxied RAs in turn act as NDProxys/SENDProxys to relay the RAs to EBGs on their parent enterprises, etc. The RA proxying and PI prefix publication recurses in this fashion and ends when an EBR attached to an interdomain routing core is reached.

After the initial PI prefix registration, the EBR that owns the prefix(es) must periodically send additional RAs to its set of EBGs to refresh prefix lifetimes. Each such EBG tracks the set of EBGs in parent enterprises to which it relays the proxied RAs, and should relay subsequent RAs to the same set.

This procedure has a direct analogy in the Teredo method of maintaining state in network middleboxes through the periodic transmission of "bubbles" [RFC4380].

5.4.4. IPv6 Next-Hop EBR Discovery

VET nodes discover destination-specific next-hop EBRs within the enterprise by querying the name service for the /56 IPv6 PI prefix taken from a packet's destination address, by forwarding packets via a default route to an EBG, or by some other inner-IP-to-outer-IP address mapping mechanism. For example, for the IPv6 destination address '2001:DB8:1:2::1' and 'PRLNAME' "isatap.example.com" the VET node can lookup the domain name:

'0.0.1.0.0.0.8.b.d.0.1.0.0.2.ip6.isatap.example.com'.

If the name-service lookup succeeds, it will return RLOC addresses (e.g., in DNS A records) that correspond to next-hop EBRs to which the VET node can forward packets. (In enterprises that use SEND, it will also return an IPv6 CGA address, e.g., in a DNS AAAA record.)

Name-service lookups in enterprises with a centralized management structure use an infrastructure-based service, e.g., an enterprise-local DNS. Name-service lookups in enterprises with a distributed management structure and/or that lack an infrastructure-based name-service instead use LLMNR over the VET interface. When LLMNR is used, the EBR that performs the lookup sends an LLMNR query (with the /56 prefix taken from the IP destination address encoded in dotted-nibble format as shown above) and accepts the union of all replies it receives from other EBRs on the VET interface. When an EBR receives an LLMNR query, it responds to the query IFF it aggregates an IP prefix that covers the prefix in the query.

Alternatively, in enterprises with a stable and highly-available set of EBGs, the VET node can simply forward an initial packet via a default route to an EBG. The EBG will forward the packet to a next-hop EBR on the VET interface and return an ICMPv6 Redirect [RFC4861] (using SEND, if necessary). If the packet's source address is on-link on the VET interface, the EBG returns an ordinary "router-to-host" redirect with the source address of the packet as its destination. If the packet's source address is not on-link, the EBG instead returns a "router-to-router" redirect with the link-local ISATAP address of the previous-hop EBR as its destination. When IPv4 is used as the outer IP protocol, the EBG also includes in the redirect one or more IPv6 Link-Layer Address Options (LLAOs) that contain the IPv4 RLOCs of potential next-hop EBRs arranged in order from lowest to highest priority (i.e., the first LLAO contains the lowest priority RLOC and the final LLAO option contains the highest priority). These LLAOs are formatted using a modified version of the form specified in Section 5 of [RFC2529], as shown in Figure 2 (the LLAO format for IPv6 as the outer IP protocol is out of scope).

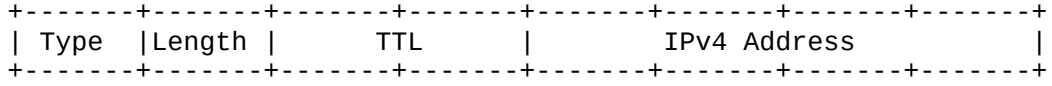

Figure 2: VET Link-Layer Address Option Format

For each such IPv6/IPv4 LLAO, the Type is set to 2 (for Target Link-Layer Address Option), Length is set to 1, and IPv4 Address is set to the IPv4 RLOC of the next-hop EBR. TTL is set to the time in seconds that the recipient may cache the RLOC, where the value 65535 represents infinity and the value 0 suspends forwarding through this RLOC.

When a VET host receives an ordinary "router-to-host" redirect, it processes the redirect exactly as specified in [RFC4861], Section 8. When an EBR receives a "router-to-router" redirect, it discovers the RLOC addresses of potential next-hop EBRs by examining the LLAOs

included in the redirect. The EBR then installs a FIB entry that contains the /56 prefix of the destination address encoded in the redirect and the list of RLOCs of potential next-hop EBRs. The EBR then enables the FIB entry for forwarding to next-hop EBRs but DOES NOT enable it for ingress filtering acceptance of packets from next-hop EBRs (i.e., the forwarding determination is unidirectional).

In enterprises in which spoofing is possible, after discovering potential next-hop EBRs (either through name-service lookup or ICMP redirect) the EBR must send authenticating credentials before forwarding packets via the next-hops. To do so, the EBR must send RAs over the VET interface (using SEND, if necessary) to one or more of the potential next-hop EBRs with an RLOC as the outer IP destination address. The RAs must include a Route Information Option (RIO) [RFC4191] that contains the /56 PI prefix of the original packet's source address. After sending the RAs, the EBR can either enable the new FIB entry for forwarding immediately or delay until it receives an explicit acknowledgement that a next-hop EBR received the RA (e.g., using the SEAL explicit acknowledgement mechanism -- see Section 5.7).

When a next-hop EBR receives the RA, it authenticates the message then it performs a name-service lookup on the prefix in the RIO if further authenticating evidence is required. If the name service returns resource records that are consistent with the inner and outer IP addresses of the RA, the next-hop EBR then installs the prefix in the RIO in its FIB and enables the FIB entry for ingress filtering but DOES NOT enable it for forwarding purposes. After an EBR sends initial RAs following a redirect, it should send periodic RAs to refresh the next-hop EBR's ingress filter prefix lifetimes as long as traffic is flowing.

EBRs retain the FIB entries created as a result of an ICMP redirect until all RLOC TTLs expire, or until no hints of forward progress through any of the associated RLOCs are received. In this way, RLOC liveness detection exactly parallels IPv6 Neighbor Unreachability Detection ([RFC4861], Section 3).

5.5. IPv4 Router Discovery and Prefix Registration

When IPv4 is used as the inner IP protocol, router discovery and prefix registration exactly parallel the mechanisms specified for IPv6 in Section 5.4. To support this, modifications to the ICMPv4 Router Advertisement [RFC1256] function to include SEND constructs and modifications to the ICMPv4 Redirect [RFC0792] function to support router-to-router redirects will be specified in a future

document. Additionally, publications for IPv4 prefixes will be in dotted-nibble format in the 'ip4.isatap.example.com' domain. For example, the IPv4 prefix 192.0.2/24 would be represented as:

'2.0.0.0.0.c.ip4.isatap.example.com'

5.6. VET Encapsulation

VET nodes forward packets by consulting the FIB to determine a specific EBR/EBG as the next-hop router on a VET interface. When multiple next-hop routers are available, VET nodes can use default router preferences, routing protocol information, traffic engineering configurations, etc. to select the best exit router. When there is no FIB information other than "default" available, VET nodes can discover the next-hop EBR/EBG through the mechanisms specified in Section 5.4 and Section 5.5.

VET interfaces encapsulate inner IP packets in any mid-layer headers followed by an outer IP header according to the specific encapsulation type (e.g., [RFC4301], [RFC5214], [RFC5320], etc.); they next submit the encapsulated packet to the outer IP forwarding engine for transmission on an underlying interface.

For forwarding to next-hop addresses over VET interfaces that use IPv6-in-IPv4 encapsulation, VET nodes determine the outer destination address (i.e., the IPv4 RLOC of the next-hop EBR) through static extraction of the IPv4 address embedded in the next-hop ISATAP address. For other IP-in-IP encapsulations, determination of the outer destination address is through administrative configuration or through an unspecified alternate method. When there are multiple candidate destination RLOCs available, the VET node should only select an RLOC for which there is current forwarding information in the outer IP protocol FIB.

5.7. SEAL Encapsulation

VET nodes should use SEAL encapsulation [RFC5320] over VET interfaces to accommodate path MTU diversity, to defeat source address spoofing, and to monitor next-hop EBR reachability. SEAL encapsulation maintains a unidirectional and monotonically incrementing per-packet identification value known as the 'SEAL_ID'. When a VET node that uses SEAL encapsulation sends a SEND-protected Router Advertisement (RA) or Router Solicitation (RS) message to another VET node, both nodes cache the new SEAL_ID as per-tunnel state used for maintaining a window of unacknowledged SEAL_IDs.

In terms of security, when a VET node receives an ICMP message, it can confirm that the packet-in-error within the ICMP message corresponds to one of its recently sent packets by examining the SEAL_ID along with source and destination addresses, etc. Additionally, a next-hop EBR can track the SEAL_ID in packets received from EBRs for which there is an ingress filter entry and discard packets that have SEAL_ID values outside of the current window.

In terms of next-hop reachability, an EBR can set the SEAL "Acknowledgement Requested" bit in messages to receive confirmation that a next-hop EBR is reachable. Setting the "Acknowledgement Requested" bit is also used as the method for maintaining the window of outstanding SEAL_IDs.

5.8. Generating Errors

When an EBR receives an IPv6 packet over a VET interface and there is no matching ingress filter entry, it drops the packet and returns an ICMPv6 [RFC4443] "Destination Unreachable; Source address failed ingress/egress policy" message to the previous-hop EBR subject to rate limiting.

When an EBR receives an IPv6 packet over a VET interface, and there is no longest-prefix-match FIB entry for the destination, it returns an ICMPv6 "Destination Unreachable; No route to destination" message to the previous hop EBR subject to rate limiting.

When an EBR receives an IPv6 packet over a VET interface and the longest-prefix-match FIB entry for the destination is via a next-hop configured over the same VET interface the packet arrived on, the EBR forwards the packet, then (if the FIB prefix is longer than ::/0) sends a router-to-router ICMPv6 Redirect message (using SEND, if necessary) to the previous-hop EBR as specified in Section 5.4.4.

Generation of other ICMP messages [RFC0792] [RFC4443] is the same as for any IP interface.

5.9. Processing Errors

When an EBR receives an ICMPv6 "Destination Unreachable; Source address failed ingress/egress policy" message from a next-hop EBR, and there is a longest-prefix-match FIB entry for the original packet’s destination that is more specific than ::/0, the EBR discards the message and marks the FIB entry for the destination as "forwarding suspended" for the RLOC taken from the source address of the ICMPv6 message. The EBR should then allow subsequent packets to flow through different RLOCs associated with the FIB entry until it

forwards a new RA to the suspended RLOC. If the EBR receives excessive ICMPv6 ingress/egress policy errors through multiple RLOCs associated with the same FIB entry, it should delete the FIB entry and allow subsequent packets to flow through an EBG if supported in the specific enterprise scenario.

When a VET node receives an ICMPv6 "Destination Unreachable; No route to destination" message from a next-hop EBR, it forwards the ICMPv6 message to the source of the original packet as normal. If the EBR has longest-prefix-match FIB entry for the original packet's destination that is more specific than ::/0, the EBR also deletes the FIB entry.

When an EBR receives an authentic ICMPv6 Redirect, it processes the packet as specified in Section 5.4.4.

When an EBG receives new mapping information for a specific destination prefix, it can propagate the update to other EBRs/EBGs by sending an ICMPv6 redirect message to the 'All Routers' link-local multicast address with an LLAO with the TTL for the unreachable LLAO set to zero, and with a NULL packet in error.

Additionally, a VET node may receive ICMP "Destination Unreachable; net / host unreachable" messages from an ER indicating that the path to a VET neighbor may be failing. The VET node should first check, e.g., the SEAL_ID, IPsec sequence number, source address of the original packet if available, etc. to obtain reasonable assurance that the ICMP message is authentic, then should mark the longest-prefix-match FIB entry for the destination as "forwarding suspended" for the RLOC destination address of the ICMP packet-in-error. If the VET node receives excessive ICMP unreachable errors through multiple RLOCs associated with the same FIB entry, it should delete the FIB entry and allow subsequent packets to flow through a different route.

5.10. Mobility and Multihoming Considerations

EBRs that travel between distinct enterprise networks must either abandon their PA prefixes that are relative to the "old" enterprise and obtain new ones relative to the "new" enterprise or somehow coordinate with a "home" enterprise to retain ownership of the prefixes. In the first instance, the EBR would be required to coordinate a network renumbering event using the new PA prefixes [RFC4192]. In the second instance, an ancillary mobility management mechanism must be used.

EBRs can retain their PI prefixes as they travel between distinct enterprise networks as long as they register the prefixes with new EBGs and (preferably) withdraw the prefixes from old EBGs prior to

departure. Prefix registration with new EBGs is coordinated exactly as specified in Section 5.4.3; prefix withdrawal from old EBGs is simply through re-announcing the PI prefixes with zero lifetimes.

Since EBRs can move about independently of one another, stale FIB entry state may be left in VET nodes when a neighboring EBR departs. Additionally, EBRs can lose state for various reasons, e.g., power failure, machine reboot, etc. For this reason, EBRs are advised to set relatively short PI prefix lifetimes in RIO options, and to send additional RAs to refresh lifetimes before they expire. (EBRs should place conservative limits on the RAs they send to reduce congestion, however.)

EBRs may register their PI prefixes with multiple EBGs for multihoming purposes. EBRs should only forward packets via EBGs with which it has registered its PI prefixes, since other EBGs may drop the packets and return ICMPv6 "Destination Unreachable; Source address failed ingress/egress policy" messages.

EBRs can also act as delegating routers to sub-delegate portions of their PI prefixes to requesting routers on their enterprise-edge interfaces and on VET interfaces for which they are configured as EBGs. In this sense, the sub-delegations of an EBR’s PI prefixes become the PA prefixes for downstream-dependent nodes. Downstream-dependent nodes that travel with a mobile provider EBR can continue to use addresses configured from PA prefixes; downstream-dependent nodes that move away from their provider EBR must perform address/prefix renumbering when they associate with a new provider.

The EBGs of a multihomed enterprise should participate in a private inner IP routing protocol instance between themselves (possibly over an alternate topology) to accommodate enterprise partitions/merges as well as intra-enterprise mobility events. These peer EBGs should accept packets from one another without respect to the destination (i.e., ingress filtering is based on the peering relationship rather than a prefix-specific ingress filter entry).

5.11. Multicast

In multicast-capable deployments, ERs provide an enterprise-wide multicasting service (e.g., Simplified Multicast Forwarding (SMF) [MANET-SMF], Protocol Independent Multicast (PIM) routing, Distance Vector Multicast Routing Protocol (DVMRP) routing, etc.) over their enterprise-interior interfaces such that outer IP multicast messages of site-scope or greater scope will be propagated across the enterprise. For such deployments, VET nodes can also provide an inner IP multicast/broadcast capability over their VET interfaces through mapping of the inner IP multicast address space to the outer

IP multicast address space. In that case, operation of link-scoped (or greater scoped) inner IP multicasting services (e.g., a link-scoped neighbor discovery protocol) over the VET interface is available, but link-scoped services should be used sparingly to minimize enterprise-wide flooding.

VET nodes encapsulate inner IP multicast messages sent over the VET interface in any mid-layer headers (e.g., IPsec, SEAL, etc.) plus an outer IP header with a site-scoped outer IP multicast address as the destination. For the case of IPv6 and IPv4 as the inner/outer protocols (respectively), [RFC2529] provides mappings from the IPv6 multicast address space to a site-scoped IPv4 multicast address space (for other IP-in-IP encapsulations, mappings are established through administrative configuration or through an unspecified alternate static mapping).

Multicast mapping for inner IP multicast groups over outer IP multicast groups can be accommodated, e.g., through VET interface snooping of inner multicast group membership and routing protocol control messages. To support inner-to-outer IP multicast mapping, the VET interface acts as a virtual outer IP multicast host connected to its underlying interfaces. When the VET interface detects that an inner IP multicast group joins or leaves, it forwards corresponding outer IP multicast group membership reports on an underlying interface over which the VET interface is configured. If the VET node is configured as an outer IP multicast router on the underlying interfaces, the VET interface forwards locally looped-back group membership reports to the outer IP multicast routing process. If the VET node is configured as a simple outer IP multicast host, the VET interface instead forwards actual group membership reports (e.g., IGMP messages) directly over an underlying interface.

Since inner IP multicast groups are mapped to site-scoped outer IP multicast groups, the VET node must ensure that the site-scope outer IP multicast messages received on the underlying interfaces for one VET interface do not "leak out" to the underlying interfaces of another VET interface. This is accommodated through normal site-scoped outer IP multicast group filtering at enterprise boundaries.

5.12. Service Discovery

VET nodes can perform enterprise-wide service discovery using a suitable name-to-address resolution service. Examples of flooding-based services include the use of LLMNR [RFC4795] over the VET

interface or multicast DNS [mDNS] over an underlying interface. More scalable and efficient service discovery mechanisms are for further study.

5.13. Enterprise Partitioning

EBGs can physically partition an enterprise by configuring multiple VET interfaces over multiple distinct sets of underlying interfaces. In that case, each partition (i.e., each VET interface) must configure its own distinct 'PRLNAME' (e.g., 'isatap.zone1.example.com', 'isatap.zone2.example.com', etc.).

EBGs can logically partition an enterprise using a single VET interface by sending RAs with PIOs containing different IPv6 PA prefixes to group nodes into different logical partitions. EBGs can identify partitions, e.g., by examining RLOC prefixes, observing the interfaces over which RSs are received, etc. In that case, a single 'PRLNAME' can cover all partitions.

5.14. EBG Prefix State Recovery

EBGs must retain explicit state that tracks the inner IP prefixes owned by EBRs within the enterprise, e.g., so that packets are delivered to the correct EBRs and not incorrectly "leaked out" of the enterprise via a default route. For PA prefixes, the state is maintained via an EBR's DHCP prefix delegation lease renewals, while for PI prefixes the state is maintained via an EBR's periodic prefix registration RAs.

When an EBG loses some or all of its state (e.g., due to a power failure), it must recover the state so that packets can be forwarded over correct routes. If the EBG aggregates PA prefixes from which the IP prefixes of all EBRs in the enterprise are sub-delegated, then the EBG can recover state through DHCP prefix delegation lease renewals, through bulk lease queries, or through on-demand name-service lookups based due to IP packet forwarding. If the EBG serves as an anchor for PI prefixes, however, care must be taken to avoid looping while state is recovered through prefix registration RAs from EBRs. In that case, when the EBG that is recovering state forwards an IP packet for which it has no explicit route other than ::/0, it must first perform an on-demand name-service lookup to refresh state.

6. Security Considerations

Security considerations for MANETs are found in [RFC2501].

Security considerations with tunneling that apply also to VET are found in [RFC2529] [RFC5214]. In particular, VET nodes must verify that the outer IP source address of a packet received on a VET interface is correct for the inner IP source address using the procedures specified in Section 7.3 of [RFC5214] in conjunction with the ingress filtering mechanisms specified in this document.

SEND [RFC3971], IPsec [RFC4301], and SEAL [RFC5320] provide additional securing mitigations to detect source address spoofing and bogus RA messages sent by rogue routers.

Rogue routers can send bogus RA messages with spoofed RLOC source addresses that can consume network resources and cause EBGs to perform extra work. Nonetheless, EBGs should not "blacklist" such RLOCs, as that may result in a denial of service to the RLOCs' legitimate owners.

7. Related Work

Brian Carpenter and Cyndi Jung introduced the concept of intra-site automatic tunneling in [RFC2529]; this concept was later called: "Virtual Ethernet" and investigated by Quang Nguyen under the guidance of Dr. Lixia Zhang. Subsequent works by these authors and their colleagues have motivated a number of foundational concepts on which this work is based.

Telcordia has proposed DHCP-related solutions for MANETs through the CECOM MOSAIC program.

The Naval Research Lab (NRL) Information Technology Division uses DHCP in their MANET research testbeds.

Security concerns pertaining to tunneling mechanisms are discussed in [TUNNEL-SEC].

Default router and prefix information options for DHCPv6 are discussed in [DEF-ROUTER].

An automated IPv4 prefix delegation mechanism is proposed in [SUBNET].

RLOC prefix delegation for enterprise-edge interfaces is discussed in [MANET-REC].

MANET link types are discussed in [LINKTYPE].

Various proposals within the IETF have suggested similar mechanisms.

8. Acknowledgements

The following individuals gave direct and/or indirect input that was essential to the work: Jari Arkko, Teco Boot, Emmanuel Bacelli, James Bound, Scott Brim, Brian Carpenter, Thomas Clausen, Claudiu Danilov, Ralph Droms, Dino Farinacci, Vince Fuller, Thomas Goff, Joel Halpern, Bob Hinden, Sapumal Jayatissa, Dan Jen, Darrel Lewis, Tony Li, Joe Macker, David Meyer, Thomas Narten, Pekka Nikander, Dave Oran, Alexandru Petrescu, John Spence, Jinmei Tatuya, Dave Thaler, Ole Troan, Michaela Vanderveen, Lixia Zhang, and others in the IETF AUTOCONF and MANET working groups. Many others have provided guidance over the course of many years.

9. Contributors

The following individuals have contributed to this document:

Eric Fleischman (eric.fleischman@boeing.com)
Thomas Henderson (thomas.r.henderson@boeing.com)
Steven Russert (steven.w.russert@boeing.com)
Seung Yi (seung.yi@boeing.com)

Ian Chakeres (ian.chakeres@gmail.com) contributed to earlier versions of the document.

Jim Bound’s foundational work on enterprise networks provided significant guidance for this effort. We mourn his loss and honor his contributions.

10. References

10.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, September 1981.

[RFC0826] Plummer, D., "Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, November 1982.

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC3007] Wellington, B., "Secure Domain Name System (DNS) Dynamic Update", RFC 3007, November 2000.

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3596] Thomson, S., Huitema, C., Ksinant, V., and M. Souissi, "DNS Extensions to Support IP Version 6", RFC 3596, October 2003.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.

[RFC3971] Arkko, J., Ed., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, March 2005.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, November 2005.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, March 2006.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.

[RFC5214] Templin, F., Gleeson, T., and D. Thaler, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)", RFC 5214, March 2008.

10.2. Informative References

[CATENET] Pouzin, L., "A Proposal for Interconnecting Packet Switching Networks", May 1974.

[mDNS] Cheshire, S. and M. Krochmal, "Multicast DNS", Work in Progress, September 2009.

[MANET-REC] Clausen, T. and U. Herberg, "MANET Router Configuration Recommendations", Work in Progress, February 2009.

[LINKTYPE] Clausen, T., "The MANET Link Type", Work in Progress, October 2008.

[DEF-ROUTER] Droms, R. and T. Narten, "Default Router and Prefix Advertisement Options for DHCPv6", Work in Progress, October 2009.

[SEND-PROXY] Krishnan, S., Laganier, J., and M. Bonola, "Secure Proxy ND Support for SEND", Work in progress, July 2009.

[SUBNET] Johnson, R., Kumarasamy, J., Kinnear, K., and M. Stapp, "Subnet Allocation Option", Work in Progress, October 2009.

[CENTRL-ULA] Hinden, R., Huston, G., and T. Narten, "Centrally Assigned Unique Local IPv6 Unicast Addresses", Work in Progress, June 2007.

[MANET-SMF] Macker, J., Ed. and SMF Design Team, "Simplified Multicast Forwarding for MANET", Work in Progress, July 2009.

[TUNNEL-SEC] Hoagland, J., Krishnan, S., and D. Thaler, "Security Concerns With IP Tunneling", Work in Progress, October 2008.

[APT] Jen, D., Meisel, M., Massey, D., Wang, L., Zhang, B., and L. Zhang, "APT: A Practical Transit Mapping Service", Work in Progress, November 2007.

[IEN48] Cerf, V., "The Catenet Model for Internetworking", IEN 48, July 1978.

[RASADV] Microsoft, "Remote Access Server Advertisement (RASADV) Protocol Specification", October 2008.

[RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, October 1989.

[RFC1256] Deering, S., Ed., "ICMP Router Discovery Messages", RFC 1256, September 1991.

[RFC1753] Chiappa, N., "IPng Technical Requirements Of the Nimrod Routing and Addressing Architecture", RFC 1753, December 1994.

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC1955] Hinden, R., "New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG", RFC 1955, June 1996.

[RFC2501] Corson, S. and J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, January 1999.

[RFC2529] Carpenter, B. and C. Jung, "Transmission of IPv6 over IPv4 Domains without Explicit Tunnels", RFC 2529, March 1999.

[RFC2775] Carpenter, B., "Internet Transparency", RFC 2775, February 2000.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, March 2004.

[RFC3819] Karn, P., Ed., Bormann, C., Fairhurst, G., Grossman, D., Ludwig, R., Mahdavi, J., Montenegro, G., Touch, J., and L. Wood, "Advice for Internet Subnetwork Designers", BCP 89, RFC 3819, July 2004.

[RFC3927] Cheshire, S., Aboba, B., and E. Guttman, "Dynamic Configuration of IPv4 Link-Local Addresses", RFC 3927, May 2005.

[RFC4192] Baker, F., Lear, E., and R. Droms, "Procedures for Renumbering an IPv6 Network without a Flag Day", RFC 4192, September 2005.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, February 2006.

[RFC4389] Thaler, D., Talwar, M., and C. Patel, "Neighbor Discovery Proxies (ND Proxy)", RFC 4389, April 2006.

[RFC4795] Aboba, B., Thaler, D., and L. Esibov, "Link-Local Multicast Name Resolution (LLMNR)", RFC 4795, January 2007.

[RFC4852] Bound, J., Pouffary, Y., Klynsma, S., Chown, T., and D. Green, "IPv6 Enterprise Network Analysis - IP Layer 3 Focus", RFC 4852, April 2007.

[RFC4903] Thaler, D., "Multi-Link Subnet Issues", RFC 4903, June 2007.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

[RFC5320] Templin, F., "The Subnetwork Encapsulation and Adaptation Layer (SEAL)", RFC 5320, February 2010.

[RFC5720] Templin, F., "Routing and Addressing in Networks with Global Enterprise Recursion (RANGER)", RFC 5720, February 2010.

[RANGERS] Russert, S., Ed., Fleischman, E., Ed., and F. Templin, Ed., "RANGER Scenarios", Work in Progress, September 2009.

Appendix A. Duplicate Address Detection (DAD) Considerations

A priori uniqueness determination (also known as "pre-service DAD") for an RLOC assigned on an enterprise-interior interface would require either flooding the entire enterprise or somehow discovering a link in the enterprise on which a node that configures a duplicate address is attached and performing a localized DAD exchange on that link. But, the control message overhead for such an enterprise-wide DAD would be substantial and prone to false-negatives due to packet loss and intermittent connectivity. An alternative to pre-service DAD is to autoconfigure pseudo-random RLOCs on enterprise-interior interfaces and employ a passive in-service DAD (e.g., one that monitors routing protocol messages for duplicate assignments).

Pseudo-random IPv6 RLOCs can be generated with mechanisms such as CGAs, IPv6 privacy addresses, etc. with very small probability of collision. Pseudo-random IPv4 RLOCs can be generated through random assignment from a suitably large IPv4 prefix space.

Consistent operational practices can assure uniqueness for EBG-aggregated addresses/prefixes, while statistical properties for pseudo-random address self-generation can assure uniqueness for the RLOCs assigned on an ER's enterprise-interior interfaces. Still, an RLOC delegation authority should be used when available, while a passive in-service DAD mechanism should be used to detect RLOC duplications when there is no RLOC delegation authority.

Author's Address

Fred L. Templin (editor)
Boeing Research & Technology
P.O. Box 3707 MC 7L-49
Seattle, WA 98124
USA

EMail: fltemplin@acm.org