Independent Submission
Request for Comments: 5720
Category: Informational
ISSN: 2070-1721

F. Templin
Boeing Phantom Works

February 2010

# Routing and Addressing in Networks with Global Enterprise Recursion (RANGER)

## Abstract

RANGER is an architectural framework for scalable routing and addressing in networks with global enterprise recursion. The term "enterprise network" within this context extends to a wide variety of use cases and deployment scenarios, where an "enterprise" can be as small as a Small Office, Home Office (SOHO) network, as dynamic as a Mobile Ad Hoc Network, as complex as a multi-organizational corporation, or as large as the global Internet itself. Such networks will require an architected solution for the coordination of routing and addressing plans with accommodations for scalability, provider-independence, mobility, multihoming, and security. These considerations are particularly true for existing deployments, but the same principles apply even for clean-slate approaches. The RANGER architecture addresses these requirements and provides a comprehensive framework for IPv6/IPv4 coexistence.

## Status of This Memo

Table of Contents

## 1. Introduction

RANGER is an architectural framework for scalable routing and addressing in networks with global enterprise recursion. The term "enterprise network" within this context extends to a wide variety of use cases and deployment scenarios, where an "enterprise" can be as small as a SOHO network, as dynamic as a Mobile Ad Hoc Network, as complex as a multi-organizational corporation, or as large as the global Internet itself. Such networks will require an architected solution for the coordination of routing and addressing plans with accommodations for scalability, provider-independence, mobility, multihoming, and security. These considerations are particularly true for existing deployments, but the same principles apply even for clean-slate approaches. The RANGER architecture addresses these requirements and also provides a comprehensive framework for IPv6/IPv4 coexistence [COEXIST].

RANGER provides a unifying architecture for enterprises that contain one or more distinct interior IP routing and addressing domains (or "Routing LOCator (RLOC) space"), with each distinct RLOC space containing one or more organizational groupings. Each RLOC space may coordinate their own internal addressing plans independently of one another, such that limited-scope addresses (e.g., [RFC1918] private-use IPv4 addresses) may be reused with impunity to provide unlimited scaling through spatial reuse. Each RLOC space therefore appears as an enterprise unto itself, where organizational partitioning of the enterprise into one or more "sub-enterprises" (or "sites") is also possible and beneficial in many scenarios. Without an architected approach, routing and addressing within such a framework would be fragmented due to address/prefix reuse between disjoint enterprises. With RANGER, however, multiple enterprises can be linked together to provide a multi-hop transit for nodes attached to enterprise edge networks that use Endpoint Interface iDentifier (EID) addresses taken from an IP addressing range that is distinct from any RLOC space.

RANGER is recursive in that multiple enterprises can be joined together in a nested "enterprise-within-enterprise" fashion, where each enterprise also connects edge networks with nodes that configure addresses taken from EID space to support edge/core separation. In this way, the same RANGER principles that apply in lower levels of recursion can extend upwards to parent enterprises and ultimately to the core of the global Internet itself. Furthermore, it is also worth considering whether today’s global Internet represents a limiting condition for recursion -- in particular, whether other internets could be manifested as "parallel universes" and joined together at still higher levels of recursion.

The RANGER architecture is manifested through composite technologies, including Virtual Enterprise Traversal (VET) [VET], the Subnetwork Encapsulation and Adaptation Layer (SEAL) [SEAL], and the Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) [RFC5214]. Other mechanisms such as IPsec [RFC4301] are also in scope for use within certain scenarios.

Noting that combinations with still other technologies are also possible, the issues addressed either in full or in part by RANGER include:

- scalable routing and addressing
- provider-independent addressing and its relation to provider-aggregated addressing
- site mobility and multihoming
- address and prefix autoconfiguration
- border router discovery
- router/host-to-router/host tunneling
- neighbor discovery over tunnels
- MTU determination for tunnels
- IPv6/IPv4 coexistence and transition

Note that while this document primarily uses the illustrative example of IPv6 [RFC2460] as a virtual overlay over IPv4 [RFC0791] networks, it is important to note that the same architectural principles apply to any combination of IPvX virtual overlays over IPvY networks.

## 2. Terminology

Routing Locator (RLOC)
: an IPv4 or IPv6 address assigned to an interface in an enterprise-interior routing region. Note that private-use IP addresses are local to each enterprise; hence, the same private-use addresses may appear within disjoint enterprises.

Endpoint Interface iDentifier (EID)
: an IPv4 or IPv6 address assigned to an edge network interface of an end system. Note that EID space must be separate and distinct from any RLOC space.

commons
: an enterprise-interior routing region that provides a subnetwork for cooperative peering between the border routers of diverse organizations that may have competing interests. A prime example of a commons is the Default-Free Zone (DFZ) of the global Internet. The enterprise-interior routing region within the commons uses an addressing plan taken from RLOC space.

enterprise
: the same as defined in [RFC4852], where the enterprise deploys a unified RLOC space addressing plan within the commons but may also contain partitions with disjoint RLOC spaces and/or organizational groupings that can be considered as enterprises unto themselves. An enterprise therefore need not be "one big happy family", but instead provides a commons for the cooperative interconnection of diverse organizations that may have competing interests (e.g., such as the case within the global Internet DFZ).

  Enterprise networks are typically associated with large corporations or academic campuses; however, the RANGER architectural principles apply to any network that has some degree of cooperative active management. This definition therefore extends to home networks, small office networks, ISP networks, a wide variety of Mobile Ad Hoc Networks (MANETs), and even to the global Internet itself.

site
: a logical and/or physical grouping of interfaces within an enterprise commons, where the topology of the site is a proper subset of the topology of the enterprise. A site may contain many interior sites, which may themselves contain many interior sites in a recursive fashion.

  Throughout the remainder of this document, the term "enterprise" refers to either enterprise or site, i.e., the RANGER principles apply equally to enterprises and sites of any size or shape. At the lowest level of recursive decomposition, a singleton Enterprise Border Router can be considered as an enterprise unto itself.

Enterprise Border Router (EBR)
: a router at the edge of an enterprise that is also configured as a tunnel endpoint in an overlay network. EBRs connect their directly attached networks to the overlay network, and connect to other networks via IP-in-IP tunneling across the commons to other EBRs. This definition is intended as an architectural equivalent of the functional term "EBR" defined in [VET].

Enterprise Border Gateway (EBG)
an EBR that also connects the enterprise to provider networks and/or to the global Internet. EBGs are typically configured as default routers in the overlay and provide forwarding services for accessing IP networks not reachable via an EBR within the commons. This definition is intended as an architectural equivalent of the functional term "EBG" defined in [VET], and is synonymous with the term "default mapper" used in other contexts (e.g., [JEN]).

Ingress Tunnel Endpoint (ITE)
a host or router interface that encapsulates inner IP packets within an outer IP header for transmission over an enterprise-interior routing region to the RLOC address of an Egress Tunnel Endpoint (ETE).

Egress Tunnel Endpoint (ETE)
a host or router interface that receives encapsulated packets sent to its RLOC address, decapsulates the inner IP packets, then delivers them to the EID address of the final destination.

overlay network
a virtual network manifested by routing and addressing over virtual links formed through automatic tunneling. An overlay network may span many underlying enterprises.

Provider-Independent (PI) prefix
an IPv6 or IPv4 EID prefix (e.g., 2001:DB8::/48, 192.0.2/24, etc.) that is routable within a limited scope and may also appear in enterprise mapping tables. PI prefixes that can appear in mapping tables are typically delegated to a Border Router (BR) by a registry but are not aggregated by a provider network. Local-use IPv6 and IPv4 prefixes (e.g., FD00::/8, 192.168/16, etc.) are another example of a PI prefix, but these typically do not appear in mapping tables.

Provider-Aggregated (PA) prefix
an IPv6 or IPv4 EID prefix that is either derived from a PI prefix or delegated directly to a provider network by a registry. Although not widely discussed, it bears specific mention that a prefix taken from a delegating router’s PI space becomes a PA prefix from the perspective of the requesting router.

Additionally, RANGER provides an informative consideration of functional specifications and operational practices outlined in other documents. These documents include:

6over4
: Transmission of IPv6 over IPv4 Domains without Explicit Tunnels [RFC2529]; functional specifications and operational practices for automatic tunneling of unicast/multicast IPv6 packets over multicast-capable IPv4 enterprises.

ISATAP
: Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) [RFC5214]; functional specifications and operational practices for automatic tunneling of unicast IPv6 packets over unicast-only IPv4 enterprises.

VET
: Virtual Enterprise Traversal (VET) [VET]; functional specifications and operational practices for automatic tunneling of both unicast and multicast IP packets with provisions for address/prefix autoconfiguration, provider-independent addressing, mobility, multihoming, and security. VET is descended from both 6over4 and ISATAP and is also known as "ISATAP version 2 (ISATAPv2)".

SEAL
: Subnetwork Encapsulation and Adaptation Layer (SEAL) [SEAL]; an encapsulation sublayer that provides an extended IP Identification field and mechanisms for link MTU adaptation over tunnels.

## 3. The RANGER Architecture

The RANGER architecture enables scalable routing and addressing in networks with global enterprise recursion while sustaining support for legacy networks and services. Key to this approach is a framework that accommodates interconnection of diverse organizations across a commons that have a mutual spirit of cooperation but also have the potential for competing interests. The following sections outline the RANGER architecture within the context of anticipated use cases:

### 3.1. Routing and Addressing

The Internet today is facing "growing pains", with indications that core Routing Information Base (RIB) scaling may not be sustainable over the long term and that the remaining space for IPv4 address allocations may be depleted in the near future. Therefore, there is an emerging need for scalable routing and addressing solutions. It must further be noted that the same solutions selected to address global Internet routing and addressing scaling can apply equally for large enterprises -- or for any enterprise that would benefit from a separation of core and edge addressing domains.

RANGER supports scalable routing through an approach that parallels the "New Scheme for Internet Routing and Addressing" described in [RFC1955]. This approach is also commonly known as "map-and-encaps". In this approach, an Ingress Tunnel Endpoint (ITE) that must forward an IP packet first consults a mapping system to discover a mapping for the destination Endpoint Interface iDentifier (EID) to a Routing LOCator (RLOC) assigned to an Egress Tunnel Endpoint (ETE). The mapping system is typically maintained as a per-enterprise distributed database that is synchronized among a limited set of mapping agents. Distributed database management alternatives include a routing protocol instance maintained by Enterprise Border Gateways (EBGs), a DNS reverse zone distributed among a restricted set of caching servers, etc. Mapping entries are inserted into the mapping system through administrative configuration, automated prefix registrations, etc.

RANGER allows for an ITE to either consult the mapping system itself (while delaying or dropping initial IP packets) or forward initial packets to an EBG acting as a "default mapper". In either case, the ITE receives a mapping reply that it can use to populate its Forwarding Information Base (FIB). The choice of mapping approaches must be considered with respect to the individual enterprise network scenario, e.g., forwarding to an EBG may be more appropriate in some scenarios while ITE self-discovery may be more appropriate in others. Use of other mapping mechanisms is also possible according to the specific enterprise scenario.

After discovering the mapping, the ITE encapsulates inner IP packets in an outer IP header for transmission across the commons to the RLOC address of an ETE. The ETE in turn decapsulates the packets and forwards them over the next hop toward the EID address of the final destination. Therefore, the Routing Information Base (RIB) within the commons only needs to maintain state regarding RLOCs and not EIDs, while the synchronized EID-to-RLOC mapping state is maintained by a smaller number of nodes and is not subject to oscillations due to link state changes within the commons. Finally, EIDs are routable only within a limited scope within edge networks (which may be as small as node-local scope in the limiting case).

RANGER supports scalable addressing by selecting a suitably large EID addressing range that is distinct and kept separate from any enterprise-interior RLOC addressing ranges. It should therefore come as no surprise that taking EID space from the IPv6 addressing architecture should lead to a viable, scalable addressing alternative, while taking EID space from the (already exhausted) IPv4 addressing architecture may not.

### 3.2. The Enterprise-within-Enterprise Framework

Enterprise networks traditionally distribute routing information via Interior Gateway Protocols (IGPs) such as Open Shortest Path First (OSPF), while large enterprises may even use an Exterior Gateway Protocol (EGP) internally in place of an IGP. Thus, it is becoming increasingly commonplace for large enterprises to use the Border Gateway Protocol (BGP) internally and independently from the BGP instance that maintains the RIB within the global Internet DFZ.

As such, large enterprises may run an internal instance of BGP across many internal Autonomous Systems (ASs). Such a large enterprise can therefore appear as an internet unto itself, albeit with default routes leading to the true global Internet. Additionally, each internal AS within such an enterprise may itself run BGP internally in place of an IGP, and can therefore also appear as an independent, lower-tier enterprise unto itself. This enterprise-within-enterprise framework can be extended in a recursive fashion as broadly and as deeply as desired to achieve scaling factors as well as organizational and/or functional compartmentalization, e.g., as shown in Figure 1.

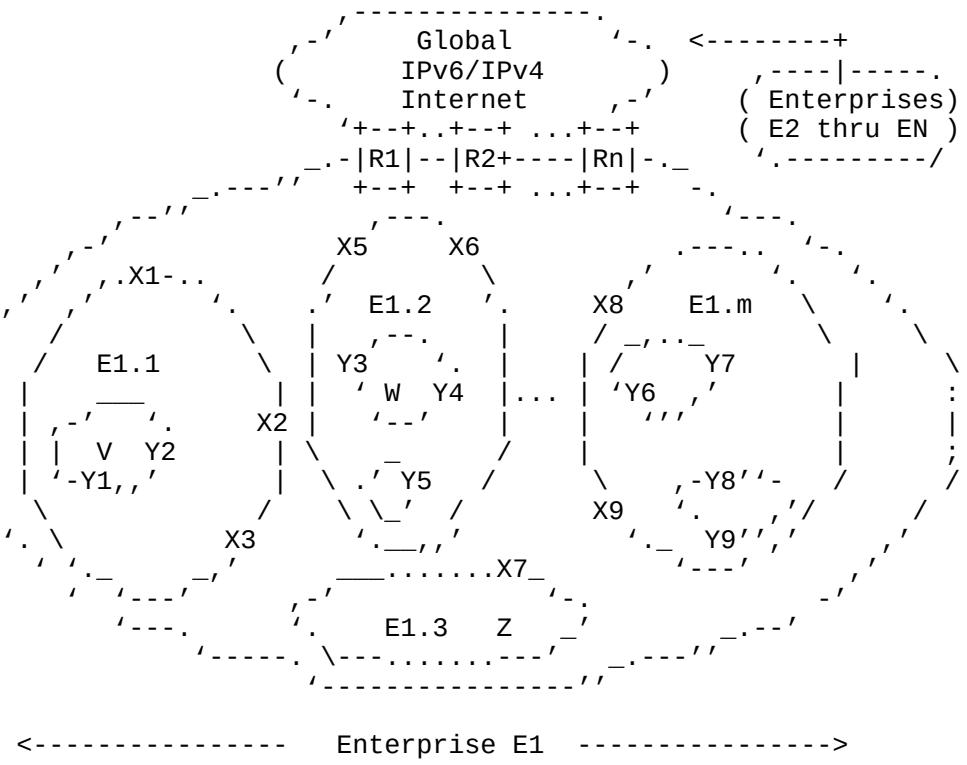


Figure 1: Enterprise-within-Enterprise Framework

Figure 1 depicts an enterprise 'E1' connected to the global IPv6/IPv4 Internet via routers 'R1' through 'Rn' and additional enterprises 'E2' through 'EN' that also connect to the global Internet. Within the 'E1' commons, there may be arbitrarily many hosts, routers, and networks (not shown in the diagram) that use addresses taken from RLOC space and over which both encapsulated and unencapsulated IP packets can be forwarded. There may also be many lower-tier enterprises, 'E1.1' through 'E1.m' (shown in the diagram), that interconnect within the 'E1' commons via Enterprise Border Routers (EBRs), depicted as 'X1' through 'X9' (where 'X1' through 'X9' see 'R1' through 'Rn' as EBGs). Within each 'E1.*' enterprise, there may also be arbitrarily many lower-tier enterprises that interconnect within the 'E1.*' commons via EBRs, depicted as 'Y1' through 'Y9' in the diagram (where 'Y1' through 'Y9' see 'X1' through 'X9' as EBGs). This recursive decomposition can be nested as deeply as desired and ultimately terminates at singleton nodes such as those depicted as 'V', 'W', and 'Z' in the diagram.

It is important to note that nodes such as 'V', 'W', and 'Z' may be simple hosts or they may be EBRs that attach arbitrarily complex edge networks with addresses taken from EID space. Such edge networks could be as simple as a home network behind a residential gateway or as complex as a major corporate/academic campus, a large service provider network, etc.

Again, this enterprise-within-enterprise framework can be recursively nested as broadly and deeply as desired. From the highest level of the recursion, consider now that the global Internet itself can be viewed as an "enterprise" that interconnects lower-tier enterprises E1 through EN such that all RANGER architectural principles apply equally within that context. Furthermore, the RANGER architecture recognizes that the global Internet need not represent a limiting condition for recursion, but rather allows that other internets could be manifested as "parallel universes" and joined together at still higher levels of recursion.

As a specific case in point, the future global Aeronautical Telecommunications Network (ATN), under consideration within the civil aviation industry [BAUER], will take the form of a large enterprise network that appears as an internet unto itself, i.e., exactly as depicted for 'E1' in Figure 1. Within the ATN, there will be many Air Communications Service Provider (ACSP) and Air Navigation Service Provider (ANSP) networks organized as autonomous systems internal to the ATN, i.e., exactly as depicted for 'E1.*' in the diagram. The ACSP/ANSP network EBGs will participate in a BGP instance internal to the ATN, and may themselves run independent BGP instances internally that are further sub-divided into lower-tier enterprises organized as regional, organizational, functional, etc. compartments. It is important to note that, while ACSPs/ANSPs within the ATN will share a common objective of safety-of-flight for civil aviation services, there may be competing business/social/political interests between them, such that the ATN is not necessarily "one big happy family". Therefore, the model parallels that of the global Internet itself.

Such an operational framework may indeed be the case for many next-generation enterprises. In particular, although the routing and addressing arrangements of all enterprises will require a mutual level of cooperative active management at a certain level, free market forces, business objectives, political alliances, etc. may drive internal competition.

### 3.3. Virtual Enterprise Traversal (VET)

Within the enterprise-within-enterprise framework outlined in Section 3.2, the RANGER architecture is based on overlay networks manifested through Virtual Enterprise Traversal (VET) ([VET], [RFC5214]). The VET approach uses automatic IP-in-IP tunneling in which ITEs encapsulate EID-based inner IP packets within RLOC-based, outer IP headers for transmission across the commons to ETEs.

For each enterprise they connect to, EBRs that use VET configure a Non-Broadcast, Multiple Access (NBMA) interface known as a "VET interface" that sees all other EBRs within the enterprise as potential single-hop neighbors from the perspective of the inner IP protocol. This means that, for many enterprise scenarios, standard neighbor discovery mechanisms (e.g., router advertisements, redirects, etc.) can be used between EBR pairs. This gives rise to a data-driven model in which neighbor relationships are formed based on traffic demand in the data plane, which in many cases can relax the requirement for dynamic routing exchanges across the overlay in the control plane.

When multiple VET interfaces are linked together, end-to-end traversal is seen as multiple VET hops from the perspective of the inner IP protocol. In that case, transition between VET interfaces entails a "re-encapsulation" approach in which a packet that exits VET interface 'i' is decapsulated then re-encapsulated before it is forwarded into VET interface 'i+1'. For example, if an end-to-end path between two EID-based peers crosses N distinct VET interfaces, a traceroute would show N inner IP forwarding hops corresponding to the portions of the path that traverse the VET interfaces.

VET and its related works specify necessary mechanisms and operational practices to manifest the RANGER architecture. The use of VET in conjunction with SEAL (see Section 3.4) is essential in certain deployments to avoid issues related to source address spoofing and black holing due to path Maximum Transmission Unit (MTU) limitations. (The use of VET in conjunction with IPsec [RFC4301] may also be necessary in some enterprise network scenarios.) The following sections discuss operational considerations and use cases within the VET approach.

#### 3.3.1. RANGER Organizational Principles

Figure 2 below depicts a vertical slice (albeit represented horizontally) from the enterprise-within-enterprise framework shown in Figure 1:

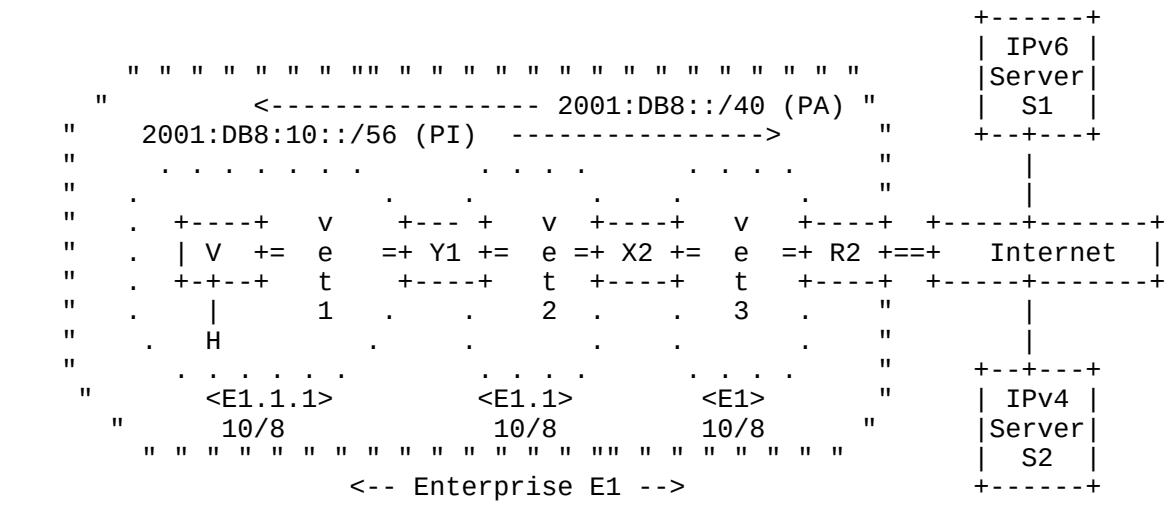


Figure 2: Virtual Enterprise Traversal

Within this vertical slice, each enterprise within the 'E1' recursive hierarchy is spanned by VET interfaces, represented as 'vet1' through 'vet3'. Each VET interface within this framework is a Non-Broadcast, Multiple Access (NBMA) interface that connects all EBRs within the same enterprise. Each enterprise within the 'E1' hierarchy may comprise a smaller topological region within a larger RLOC space, or they may configure an independent RLOC space from a common (but spatially reused) limited-scope prefix, e.g., depicted as multiple disjoint instances of '10/8' in the diagram.

In the RANGER approach, EBRs within lower-tier enterprises coordinate their EID prefixes with EBGs that connect to an upper-tier enterprise. EID prefixes could be either provider-independent (PI) prefixes owned by the EBR or provider-aggregated (PA) prefixes delegated by the EBG. In either case, EID prefixes must be coordinated with the enterprise routing/mapping systems.

When PA EID prefixes are used, the EBR for each 'E1*' enterprise receives an EID prefix delegation from a delegating EBG in a parent enterprise. In this example, when 'R2' is a delegating router for the prefix '2001:DB8::/40', it may delegate '2001:DB8::/48' to 'X2', which in turn delegates '2001:DB8::/52' to 'Y1', which in turn delegates '2001:DB8::/56' to 'V'. The preferred mechanism for this recursive PA prefix sub-delegation is DHCP Prefix Delegation [RFC3633], which also arranges for publication of the prefixes in the enterprise routing system.

When PI EID prefixes are used, individual EBRs (e.g., 'V') register their PI prefixes (e.g., '2001:DB1:10::/56') by sending Router Advertisement (RA) messages to EBGs within the enterprise to assert prefix ownership. When stronger authentication is necessary, the EBRs can digitally sign the messages using the mechanisms specified for SEcure Neighbor Discovery (SEND) [RFC3971]. EBGs that receive the RAs (e.g., 'Y1') first verify the sender's credentials, then register the prefixes in the enterprise mapping system. Next, they forward a proxied version of the RA to EBGs within their parent enterprises (e.g., 'X2'). This proxying process continues up the recursive hierarchy until a default-free commons is reached. (In this case, the proxying process ends at 'R2'). After the initial registration, the EBR that owns the PI prefixes must periodically send additional RAs to update prefix expiration timers.

### 3.3.2. RANGER End-to-End Addressing Example

In Figure 2, an IPv6 host 'H' that is deeply nested within Enterprise 'E1' connects to IPv6 server 'S1', located somewhere on the IPv6 Internet. 'H' is attached to a shared link with IPv6/IPv4 dual-stack router 'V', which advertises the IPv6 prefixes '2001:DB8:0:0::/64' and '2001:DB8:10:0::/64'. 'H' uses standard IPv6 neighbor discovery mechanisms to discover 'V' as a default IPv6 router that can forward its packets off the local link, and configures addresses from both of the advertised prefixes. 'V' in turn sees node 'Y1' as an EBG that is reachable via VET interface 'vet1' and that can forward packets toward IPv6 server 'S1'. Similarly, node 'Y1' is an EBR on the enterprise spanned by 'vet2' that sees 'X2' as an EBG, and node 'X2' is an EBR on 'vet3' that sees 'R2' as an EBG. Ultimately, 'R2' is an EBR that connects 'E1' to the global Internet.

### 3.3.3. Dynamic Routing and On-Demand Mapping

In the example shown in Figure 2, 'V', 'Y1', 'X2', and 'R2' configure separate VET interfaces for each enterprise they connect to in order to discover routes through a dynamic routing protocol and/or mapping database lookups. After tunnels 'vet1' through 'vet3' are established, the EBRs connected to a VET interface can run a dynamic routing protocol such as OSPVFv3 [RFC5340] and exchange topology information over the VET interface using the NBMA interface model. In this way, each EBR can discover other EBRs on the link via routing protocol control message exchanges.

In a second example, Figure 3 depicts an instance of on-demand discovery of more specific routes in which an IPv6 end system 'H' connects to a peer end system 'J', located in a different organizational entity within 'E1':

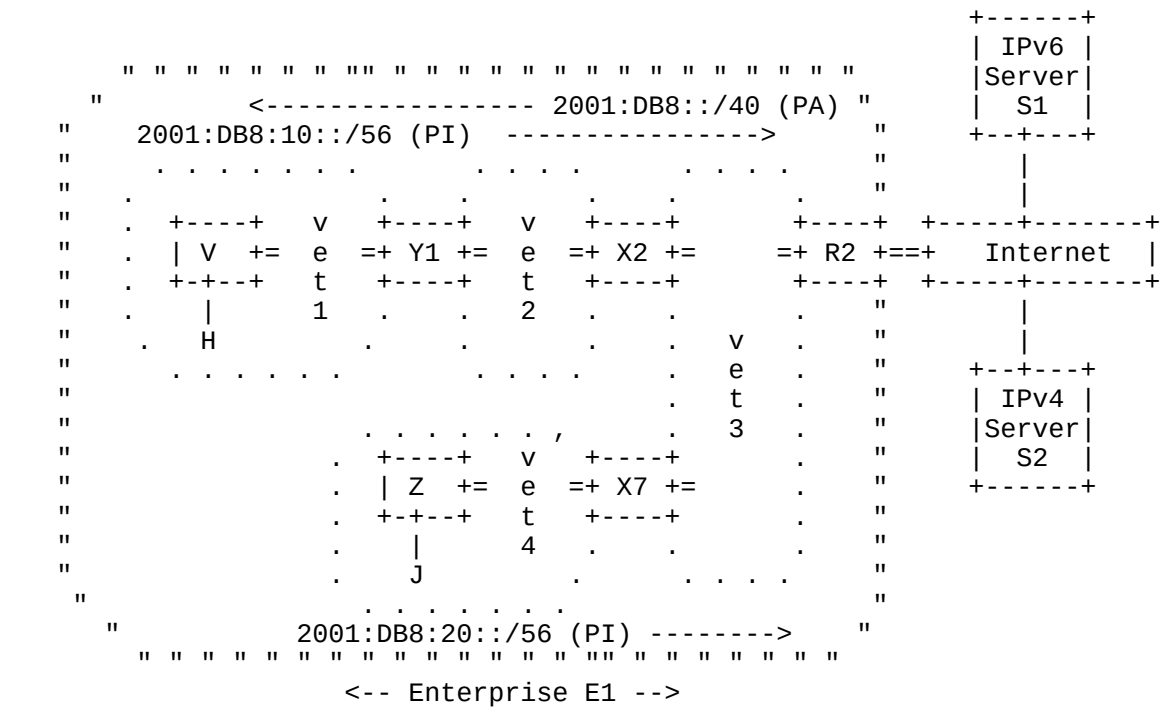


Figure 3: On-Demand Discovery

In this example, tunnel interfaces 'vet1' through 'vet4' as well as IPv6 PI prefix registrations have been established through VET enterprise autoconfiguration procedures. When IPv6 end system 'H' with IPv6 address '2001:DB8:10::1' sends packets to a peer end system 'J' with IPv6 address '2001:DB8:20::1', the packets will be conveyed through 'V', 'Y1', and finally to 'X2' via default routes. Then, unless 'X2' has an IPv6 FIB entry matching 'J', it must discover that 'J' can be reached using a more direct route via 'X7' as the next-hop across the 'E1' commons.

In particular, when 'X2' receives a packet on the 'vet2' interface with inner destination address 'J', it can perform an on-demand mapping lookup by consulting the enterprise mapping service, e.g., by consulting the DNS reverse zone. Alternatively, 'X2' can send the packet to a default router (e.g., 'R2'), which in turn can forward the packet to 'X7' and return an ICMPv6 redirect message. When 'X2' receives the redirect, it can send an RA message to 'X7' to prove that it is authorized to produce packets with a prefix that matches source address 'J'. 'X2' can then forward subsequent packets directly to 'X7' without involving 'R2'.

In some enterprise scenarios, dynamic routing and on-demand mapping can be combined as complementary functions. In other scenarios, it may be preferable to use either dynamic routing only or on-demand mapping only.

### 3.3.4. Support for Legacy RLOC-Based Services

Legacy hosts, routers, and networks that were already present in pre-RANGER deployments and have already numbered their interfaces with RLOC addresses must see continued support for RLOC-based services for the long term, even as EID-based services are rolled out in new deployments. For example, a legacy IPv4-only node behind an IPv4 Network Address Translator (NAT) must still be able to reach legacy IPv4-only Internet services (e.g., "http://example.com") long after the RANGER architecture and EID-based services are widely deployed.

Returning to the example diagrams, while virtual enterprise traversal across 'E1' provides a fully connected routing and addressing capability for EID-based services, legacy nodes will still require access to RLOC-based services within connected or disjoint RLOC spaces for an extended (and possibly indefinite) period. For example, Figure 4 below depicts the applicable RLOC-based IPv4 service-access scenarios for the RANGER architecture when VET interfaces are used to link recursively nested enterprises together:

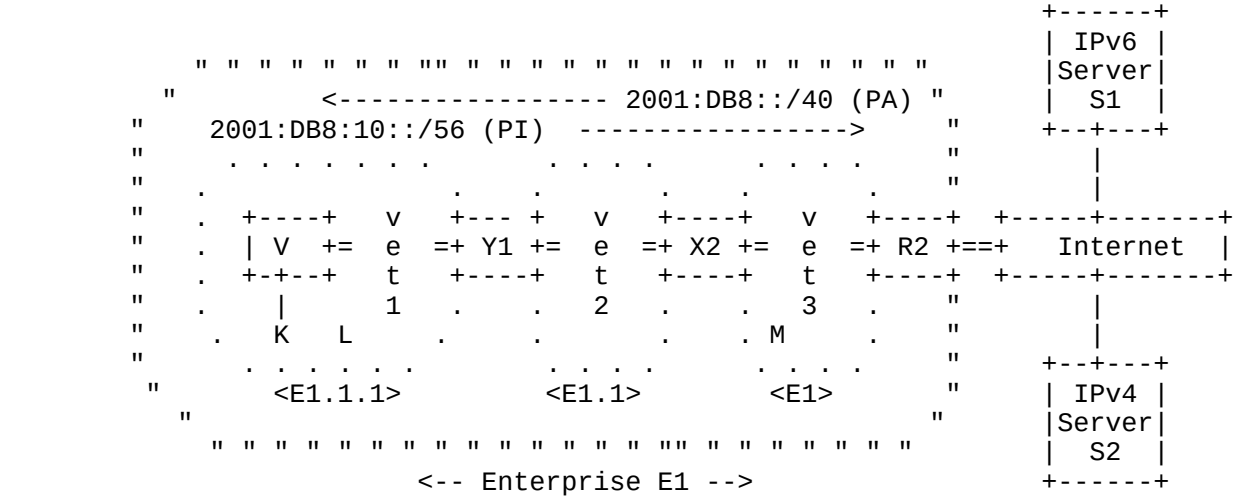


Figure 4: Support for Legacy RLOC-Based Services

In a first instance, a legacy RLOC-based IPv4 client 'K' within enterprise 'E1.1.1' can access RLOC-based IPv4 service 'L' within the same enterprise as normal and without the need for any encapsulation.

Instead, 'K' discovers a "mapping" for 'L' through a simple lookup within the 'E1.1.1' enterprise-local name service, and conveys packets to 'L' through unencapsulated RLOC-based IPv4 routing and addressing within the 'E1.1.1' commons. In many instances, this may indeed be the preferred service-access model, even when EID-based IPv6 services are widely deployed due to factors such as inability to replace legacy IPv4 applications, IPv6 header overhead avoidance, etc.

In a second instance, RLOC-based IPv4 client 'K' can access RLOC-based IPv4 server 'S2' on the legacy global IPv4 Internet in a number of ways, based on the way the recursively nested 'E1.*' enterprises are provisioned:

- if all of the recursively nested 'E1.*' enterprises are configured within the same IPv4 RLOC space, normal IPv4 forwarding will convey unencapsulated IPv4 packets from 'K' toward 'R2', which then acts as an IPv4 Network Address Translator (NAT) and/or an ordinary IPv4 Enterprise Border Router.

- if the recursively nested 'E1.*' enterprises are configured within disjoint RLOC spaces, all EBGs 'Y1', 'X2', and 'R2' can be configured to provide an IPv4 NAT capability (i.e., a recursive nesting of NATs within NATs). However, this approach places multiple instances of stateful NAT devices on the path, which can lead to an overall degree of brittleness and intolerance to routing changes. Instead, 'R2' can act as a "Carrier-Grade NAT (CGN)", and 'V' can convey packets from 'K' to the CGN using IPv4-in-IPv6-in-IPv4 tunneling. The CGN can then decapsulate the inner, RLOC-based IPv4 packets and translate the IPv4 source addresses into global IPv4 source addresses before sending the packets on to 'S2'.

- 'K' could be configured as an EID-based, IPv6-capable node and use standard IPv6 routing to reach an IPv6/IPv4 translator located at an EBR for the enterprise in which 'S2' resides. The translator would then use IPv6-to-IPv4 translation before sending packets onwards toward 'S2'. These and other alternatives are discussed in [WING].

In a final instance, RLOC-based IPv4 client 'K' can access an RLOC-based IPv4 server 'M' in a different enterprise within E1 as long as both enterprises are configured over the same IPv4 RLOC space. If the enterprises are configured over disjoint IPv4 RLOC spaces, however, 'K' would still be able to access 'M' by using EID-based IPv6 services, by using EID-based IPv4 services if an EID-based IPv4 overlay were deployed, or by using some form of RLOC-based IPv4 NAT traversal. 'K' could also access server 'M' if both 'V' and 'X2'

implemented an IPv6/IPv4 protocol translation capability. Finally, 'K' and/or 'M' could implement a bump-in-the-wire or bump-in-the-api IPv6/IPv4 protocol translation capability.

### 3.4. Subnetwork Encapsulation and Adaptation Layer (SEAL)

Tunnel endpoints that depend on ICMP feedback from routers within the enterprise commons may be susceptible to undetected black holes due to ICMP filtering gateways and/or off-path ICMP spoofing attacks from a node pretending to be a router. Furthermore, rogue nodes within enterprises that do not correctly implement ingress filtering can send spoofed packets of any kind, e.g., for the purpose of mounting denial-of-service and/or traffic amplification attacks targeting underprivileged links.

The Subnetwork Encapsulation and Adaptation Layer (SEAL) provisions each encapsulated packet with a monotonically incrementing, extended Identification field (i.e., the 32-bit SEAL_ID) that tunnel endpoints can use as a nonce to detect off-path spoofing. Moreover, tunnel endpoints that use SEAL can continue to operate correctly even if some/many ICMPs are lost. Finally, tunnel endpoints that use SEAL can adapt to subnetworks containing links with diverse MTUs properties.

### 3.5. Mobility Management

Enterprise mobility use cases must be considered along several different vectors:

- nomadic enterprises and end systems may be satisfied to incur address renumbering events as they move between new enterprise network attachment points.

- mobile enterprises with PI prefixes may be satisfied by dynamic updates to the mapping system as long as they do not impart unacceptable churn.

- mobile enterprises and end systems with PA addresses/prefixes may require additional supporting mechanisms that can accommodate address/prefix renumbering.

Nomadic enterprise mobility is already satisfied by currently deployed technologies. For example, transporting a laptop computer from a wireless-access hot spot to a home network LAN would allow the nomadic device to re-establish connectivity at the expense of address renumbering. Such renumbering may be acceptable, especially for

devices that do not require session persistence across mobility events and do not configure servers with addresses published in the global DNS.

Mobile enterprises with PI prefixes that use VET and SEAL can move between parent enterprise attachment points as long as they withdraw the prefixes from the mapping systems of departed enterprises and inject them into the mapping systems of new enterprises. When moving between the lower recursive tiers of a common parent enterprise, such a localized event mobility may result in no changes to the parent enterprise's mapping system. Hence, the organizational structure of a carefully arranged enterprise-within-enterprise framework may be able to dampen mobility-related churn. For enterprises that require in-the-network confidentiality, IKEv2 Mobility and Multihoming (MOBIKE) [RFC4555] may also be useful within this context.

Mobile enterprises and end systems that move quickly between disparate parent enterprise attachment points should not use PI prefixes if withdrawing and announcing the prefixes would impart unacceptable mapping/routing churn and packet loss. They should instead use PA addresses/prefixes that can be coordinated via a rendezvous service. Mobility management mechanisms such as Mobile IPv6 [RFC3775] and the Host Identity Protocol (HIP) [RFC4423] can be used to maintain a stable identifier for fast moving devices even as they move quickly between visited enterprise attachment points.

As a use case in point, consider an aircraft with a mobile router moving between ground station points of attachment. If the ground stations are located within the same enterprise, or within lower-tier sites of the same parent enterprise, it should suffice for the aircraft to announce its PI prefixes at its new point of attachment and withdraw them from the old. This would avoid excessive mapping system churn, since the prefixes need not be announced/withdrawn within the parent enterprise, i.e., the churn is isolated to lower layers of the recursive hierarchy. Note also that such movement would not entail an aircraft-wide readdressing event.

As a second example, consider a wireless handset moving between service coverage areas maintained by independent providers with peering arrangements. Since the coverage range of terrestrial cellular wireless technologies is limited, mobility events may occur on a much more aggressive timescale than some other examples. In this case, the handset may expect to incur a readdressing event for its access interface at least, and may be obliged to arrange for a rendezvous service linkage.

It should specifically be noted that the contingency of mobility management solutions is not necessarily mutually exclusive and must be considered in relation to specific use cases. The RANGER architecture is therefore naturally inclusive in this regard. In particular, RANGER could benefit from mechanisms that could support rapid dynamic updates of PI prefix mappings without causing excessive churn.

### 3.6. Multihoming

As with mobility management, multihoming use cases must be considered along multiple vectors. Within an enterprise, EBRs can discover multiple EBGs and use them in a fault-tolerant and load-balancing fashion as long as they register their PI prefixes with each such EBG, as described in Section 3.3.1. These registrations are created through the transmission of Router Advertisement messages that percolate up through the recursive enterprise-within-enterprise hierarchy.

As a first case in point, consider the enterprise network of a major corporation that obtains services from a number of ISPs. The corporation should be able to register its PI prefixes with all of its ISPs, and use any of the ISPs for its Internet access services.

As a second use case, consider an aircraft with a diverse set of wireless links (e.g., VHF, 802.16, directional, SATCOM, etc.). The aircraft should be able to select and utilize the most appropriate link(s) based on the phase of flight and to change seamlessly between links as necessary. Other examples include a nomadic laptop with both 802.11 and Ethernet links, a wireless handset with both CDMA wireless and 802.11, etc.

As with mobility management, the contingency of solutions is not necessarily mutually exclusive and can combine to suit use cases within the scope of the RANGER architecture.

### 3.7. Implications for the Internet

Selection of mapping alternatives may have significant implications for applications, server selection, route determination, security, etc. In particular, applications that expect all packets (including initial ones) to experience similar delays may be adversely affected by a scheme that imposes non-negligible delays when initial packets are queued while a look-aside mapping table is consulted. Still other applications may experience significant startup delays when its initial packets are dropped during a mapping lookup event. These

factors would seem to favor a scheme that is able to forward initial packets along a path with sub-optimal delay while a mapping lookup is performed in parallel, e.g., such as when a "default mapper" is used.

Generally speaking, proactive mapping-maintenance mechanisms may have scaling issues with the amount of updates they generate, while reactive mechanisms may involve effects to the delay of initial packets before the cached state is updated. Also to be considered are attacks against the mapping mechanism, which may result in denial of service of the mapping cache.

Encapsulation of packets in automatically created tunnels involves a number of issues as well. There are obvious interactions between encapsulation overhead and the effective tunnel MTU, which can be addressed by SEAL and (when necessary) careful operational link arrangements. Moreover, it is important to minimize the impact to the global routing table without at the same time impacting the ability of legacy Internet networks to connect to those employing RANGER. As long as other nodes in the Internet need to connect to networks implementing RANGER, EID routes need to appear both in the mapping system and the global BGP routing tables. This can be accommodated by keeping the number of prefixes aggregated by RANGER to the bare minimum through efficient aggregation (e.g., one or a few [PREF]::/4 IPv6 prefixes instead of millions of [PREF]::/32 prefixes).

## 4. Related Initiatives

The origins of the RANGER architectural principles can be traced to the "Catenet model for internetworking" ([CATENET], [IEN48], [RFC2775]) beginning as early as the mid-1970's. Subsequently, deliberations of the ROAD group [RFC1380] and related efforts such as NIMROD [RFC1753] provided a sustained evolution of the concepts. [RFC1955], "New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG", captures the high-level architectural aspects of the ROAD group deliberations.

These foundational works significantly influenced more recent developments, including the X-Bone initiative [XBONE], which explored virtual topologies manifested through tunneling. Various tunneling approaches including IP-in-IP ([RFC2003], [RFC4213]), 6over4 [RFC2529], and ISATAP [RFC5214] have evolved from the mid-1990's until the present day and are used in common, operational practice. Tunnel-mode IPsec [RFC4301] is also commonly used for separation of security domains within enterprises.

Currently, initiatives with similar properties to RANGER are under development within the IRTF Routing Research Group (RRG) and within IETF working groups such as LISP, SOFTWIRE, V6OPS, and others. Numerous proposals have been offered within the RRG, including the Locator-Identifier Split Protocol (LISP) [LISP], Six-One [VOGT], ILNP [ILNP], Internet vastly improved plumbing (Ivip) [WHITTLE], A Practical Transit-Mapping Service (APT) [JEN], and Virtual Aggregation [VA]. Still other similar initiatives almost certainly exist.

While RANGER shares many properties with these earlier works, it uniquely provides a top-to-bottom articulation of how the various pieces fit together within a recursively nested "enterprise-within-enterprise" (or "network-of-networks") framework. In this way, it bears striking resemblance to the network-of-networks model envisioned by CATENET. RANGER further provides a detailed consideration of challenging issues such as autoconfiguration, provider-independent addressing, border router discovery, tunnel MTU, multihoming, etc. that many other approaches have either overlooked or left for future work. A detailed analysis of RANGER applicability in various use case scenarios is provided in "RANGER Scenarios (RANGERS)" [RUSSERT].

## 5. Security Considerations

Communications between endpoints within different sites inside an enterprise are carried across a commons that joins organizational entities with a mutual spirit of cooperation, but between which there may be competing business/sociological/political interests. As a result, mechanisms that rely on feedback from routers on the path may become brittle or susceptible to spoofing attacks. This is due to the fact that IP packets can be lost due to congestion or packet-filtering gateways and that the source addresses of IP packets can be forged. Moreover, IP packets in general can be generated by anonymous attackers, e.g., from a rogue node within a third-party enterprise that has malicious intent toward a victim.

Path MTU Discovery is an example of a mechanism that relies on ICMP feedback from routers on the path and, as such, is susceptible to these issues. For IPv4, a common workaround is to disable Path MTU Discovery and let fragmentation occur in the network if necessary. For IPv6, lack of fragmentation support in the network precludes this option such that the mitigation typically recommended is to discard ICMP messages that contain insufficient information and/or to operate with the minimum IPv6 path MTU. This example extends also to other mechanisms that either rely on or are enhanced by feedback from network devices; however, attack vectors based on non-ICMP messages are also subject for concern.

The RANGER architecture supports effective mitigations for attacks such as distributed denial-of-service, traffic amplification, etc. In particular, when VET and SEAL are used, EBGs can use the 32-bit identification encoded in the SEAL header as well as ingress filtering to determine if a message has come from a topologically correct enterprise located across the commons. This allows enterprises to employ effective mitigations at their borders without the requirement for mutual cooperation from other enterprises. When source address spoofing by on-path attackers located within the commons is also subject for concern, additional securing mechanisms such as tunnel-mode IPsec between enterprise EBGs can also be used.

EBRs can obtain PI prefixes through arrangements with a prefix delegation authority. Thereafter, the EBR can announce and/or withdraw the prefixes within an enterprise by sending IPv6 Router Advertisements (RAs). In environments where additional authenticating mechanisms are necessary, the EBR can sign its RAs using SEcure Neighbor Discovery (SEND) [RFC3971].

While the RANGER architecture does not in itself address security considerations, it proposes an architectural framework for functional specifications that do. Security concerns with tunneling, along with recommendations that are compatible with the RANGER architecture, are found in [HOAGLAND].

## 6. Acknowledgements

This work was inspired through the encouragement of the Boeing DC&NT network technology team and through the communications of the IESG.

Many individuals have contributed to the architectural principles that form the basis for RANGER over the course of many years. The following individuals have given specific feedback on the RANGER document itself: Jari Arkko, Brian Carpenter, Eric Fleischman, Joel Halpern, Thomas Henderson, Steven Russert, Dallas Thomas, Robin Whittle.

## 7. References

### 7.1. Normative References

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

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

## 7.2. Informative References

[CATENET] Pouzin, L., "A Proposal for Interconnecting Packet Switching Networks", Proceedings of EUROCOMP, Bronel University, p. 1023-1036, May 1974.

[COEXIST] Arkko, J. and M. Townsley, "IPv4 Run-Out and IPv4-IPv6 Co-Existence Scenarios", Work in Progress, July 2009.

[BAUER] Bauer, C. and S. Ayaz, "ATN Topology Considerations for Aeronautical NEMO RO", Work in Progress, September 2009.

[LISP] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol (LISP)", Work in Progress, September 2009.

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

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

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

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

[VET] Templin, F., Ed., "Virtual Enterprise Traversal (VET)", RFC 5558, February 2010.

[WING] Wing, D., Ward, D., and A. Durand, "A Comparison of Proposals to Replace NAT-PT", Work in Progress, September 2008.

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

[ILNP] Atkinson, R., "ILNP Concept of Operations", Work in Progress, December 2008.

[RFC1380] Gross, P. and P. Almquist, "IESG Deliberations on Routing and Addressing", RFC 1380, November 1992.

[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.

[RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

[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.

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

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

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

[RFC4213] Nordmark, E. and R. Gilligan, "Basic Transition Mechanisms for IPv6 Hosts and Routers", RFC 4213, October 2005.

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

[RFC4423] Moskowitz, R. and P. Nikander, "Host Identity Protocol (HIP) Architecture", RFC 4423, May 2006.

[RFC4555] Eronen, P., "IKEv2 Mobility and Multihoming Protocol (MOBIKE)", RFC 4555, June 2006.

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

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

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

[VA] Francis, P., Xu, X., Ballani, H., Jen, D., Raszuk, R., and L. Zhang, "FIB Suppression with Virtual Aggregation", Work in Progress, October 2009.

[VOGT] Vogt, C., "Six/One: A Solution for Routing and Addressing in IPv6", Work in Progress, October 2009.

[WHITTLE] Whittle, R., "Ivip (Internet Vastly Improved Plumbing) Architecture", Work in Progress, August 2008.

[XBONE] Touch, J., "The X-Bone", March 1997, http://www.isi.edu/touch/pubs/ngi97/x-bone-ngi97.pdf

## Author's Address

Fred L. Templin
Boeing Phantom Works
P.O. Box 3707 MC 7L-49
Seattle, WA 98124
USA

EMail: fltemplin@acm.org