Routing and Addressing in Next-Generation Enterprises (RANGER)
draft-templin-ranger-07.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

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

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

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

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

This Internet-Draft will expire on August 10, 2009.

Copyright Notice

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

RANGER is an architectural framework for scalable routing and addressing in next generation enterprise networks. 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, multi-homing 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.

Table of Contents

1. Introduction .................................................. 3
2. Terminology ................................................... 4
3. The RANGER Architecture ...................................... 7
   3.1. Routing and Addressing .................................... 8
   3.2. The Enterprise-within-Enterprise Framework ................. 9
   3.3. Virtual Enterprise Traversal (VET) ....................... 11
      3.3.1. RANGER Organizational Principles .................... 12
      3.3.2. RANGER End-to-End Addressing Example ............... 14
      3.3.3. Dynamic Routing and On-Demand Mapping .............. 14
      3.3.4. Support for Legacy RLOC-Based Services ............. 16
   3.4. Subnetwork Encapsulation and Adaptation Layer (SEAL) .... 18
   3.5. Mobility Management .................................... 18
   3.6. Multihoming ............................................. 20
4. Related Initiatives ............................................ 21
   4.1. 6over4 and ISATAP ....................................... 21
   4.2. The Locator Identifier Split Protocol (LISP) ............ 21
5. IANA Considerations ........................................... 21
6. Security Considerations ....................................... 21
7. Acknowledgements ............................................. 23
8. References .................................................. 23
   8.1. Normative References .................................... 23
   8.2. Informative References .................................. 23
Author’s Address ................................................ 25
1. Introduction

RANGER is an architectural framework for scalable routing and addressing in next generation enterprise networks. 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, multi-homing 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 [I-D.arkko-townsley-coexistence].

RANGER provides a unifying architecture for enterprises that contain one or more distinct interior IP 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, or if 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) [I-D.templin-autoconf-dhcp], the Subnetwork Encapsulation and Adaptation Layer (SEAL) [I-D.templin-seal], and the Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) [RFC5214][I-D.templin-isatapv4]. Other mechanisms such as IPsec [RFC4301] are also in scope for use within certain deployments.

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 multi-homing
- address and prefix autoconfiguration
- border router discovery
- router-to-router tunneling
- neighbor discovery over tunnels
- MTU determination for tunnels
- IPv6/IPv4 coexistence and transition

The RANGER architectural principles can be traced to the deliberations of the ROAD group in January 1992 [RFC1380], and also to still earlier works including NIMROD [RFC1753], the Catenet model for internetworking [CATENET][IEN48] [RFC2775], and many others. [RFC1955] captures the high-level architectural aspects of the ROAD group deliberations in a "New Scheme for Internet Routing and Addressing [ENCAPS] for IPNG".

Note that while this document primarily uses the illustrative example of IPv6 as a virtual overlay over or IPv4 networks, it is important to note that the same architectural principles apply to any combination of IP* virtual overlays over IP* networks.

2. Terminology
Routing Locator (RLOC)  
an IPv4 or IPv6 address assigned to an interface in an enterprise-
interior routing region. Note that RLOC space is local to each
enterprise, hence the same RLOC space IP addresses may be reused
between disjoint enterprises.

Endpoint Interface iDentifier (EID)  
an IPv4 and 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  
a 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
default free zone).

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, 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 node 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 [I-D.templin-autoconf-dhcp], and is synonymous with the term "xTR" used in other contexts (e.g., [I-D.farinacci-lisp]).

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 [I-D.templin-autoconf-dhcp], and is synonymous with the term "default mapper" used in other contexts (e.g., [I-D.jen-apt]).

Ingress Tunnel Router (ITR)
a BR 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 Router (ETR).

Egress Tunnel Router (ETR)
a BR that receives encapsulated packets sent to its RLOC address, decapsulates the inner IP packets, then forwards them over an edge network 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.

6over4
ISATAP
Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)
[RFC5214]/[I-D.templin-isatapv4]; functional specifications and
operational practices for automatic tunneling over unicast-only
enterprises.

VET
Virtual Enterprise Traversal (VET) [I-D.templin-autoconf-dhcp];
functional specifications and operational practices that provide a
functional superset of 6over4 and ISATAP. In addition to both
unicast and multicast tunneling, VET also supports address/prefix
autoconfiguration as well as additional encapsulations such as
IPSec, SEAL, LISP/UDP, Teredo/UDP, etc.

SEAL
Subnetwork Encapsulation and Adaptation Layer (SEAL)
[I-D.templin-seal]; a functional specification for robust packet
identification and link MTU adaptation over tunnels. SEAL
supports effective ingress filtering and adapts to subnetworks
configured over links with diverse characteristics.

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 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.

3. The RANGER Architecture

The RANGER architecture enables scalable routing and addressing in
next-generation enterprise networks with sustaining support for
legacy networks and services. Key to this approach is a framework
that accommodates interconnection of diverse organizations across a
commons with a mutual spirit of cooperation, but with 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", as seen through evidence that core Routing Information Base (RIB) scaling will be unsustainable over the long term and that the remaining space for IPv4 address allocations will 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 known as map-and-encaps. In this approach, an Ingress Tunnel Router (ITR) 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 Router (ETR). The mapping system is maintained as a per-enterprise distributed database which is synchronized among a limited set of mapping agents. Distributed database management alternatives include a BGP instance maintained by 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 that an ITR either consult the mapping system itself (while delaying or dropping initial IP packets) or forward initial packets to an Enterprise Border Gateway (EBG) acting as a "default mapper". In either case, the ITR 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 ITR 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 ITR encapsulates inner IP packets in an outer IP header for transmission across the commons to the RLOC address of an ETR. The ETR in turn decapsulates the packets and forwards them over edge networks to the EID addresses of final destinations. The Routing Information Base (RIB) within the commons therefore 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 Default Free Zone (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 BRs depicted as ‘Y1’ through ‘Y9’ in the diagram (where ‘Y1’ through ‘Y9’ see ‘X1’ through ‘X9’ as BGs). 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 [I-D.bauer-mext-aero-topology] 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)
The VET approach uses automatic IP-in-IP tunneling in which ITRs encapsulate EID-based inner IP packets within RLOC-based outer IP headers for transmission across the commons to ETRs.

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] can also be beneficial 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:
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 digitally signed Router Advertisement (RA) messages to EBGs within the enterprise using the mechanisms specified for SEcure Neighbor Discovery (SEND) [RFC3971]. The signed messages must contain sufficient proof of the EBR’s authority to use the prefixes. 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 and to discover EBRs/EBGs through a dynamic routing protocol and/or mapping database lookups. After tunnels ‘vet1’ through ‘vet3’ are established and EBG’s discovered, the EBRs connected to a ‘vet∗’ interface can run a dynamic routing protocol such as OSPFv3 [RFC5340] and exchange topology information with one another over the ‘vet∗’ interface. It is important to note that EBR neighbor relationships can be formed on-demand and allowed to expire after idle periods such that a full mesh of neighbors need not be maintained. This allows an overlay network that spans ‘E1’ to dynamically adapt to changing conditions within the enterprise.

In a second example, Figure 3 depicts an instance of on-demand discovery of more-specific routes in which an IPv6 host ‘H’ connects
to an IPv6 server ‘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 host ‘H’ with IPv6 address ‘2001:DB8:10::1’ sends packets to server ‘J’ with IPv6 address ‘2001:DB8:20::1’, unless EBR ‘X2’ has an IPv6 FIB entry matching ‘J’, it must determine that ‘J’ can be reached using a more direct route via ‘X7’ as the next-hop across the ‘E1’ commons.

To determine the best next-hop, ‘X2’ can perform an on-demand mapping lookup by consulting the enterprise mapping service (e.g., an enterprise name service) while dropping or delaying initial packets. This can be done, e.g., by consulting the DNS for a FQDN that matches the EID prefix of the inner IP packet’s destination address. Alternatively, ‘X2’ can send the packet to a default router (e.g., EBG ‘R2’) which in turn can forward the packet to ‘X7’ and return an ICMPv6 redirect message to ‘X2’. When ‘X2’ receives the redirect, it can send a SEND-signed RA message to ‘X7’ then forward subsequent
packets directly via the route-optimized path to ‘X7’.

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 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:

```
+------+
| IPv6 |
|  S1  |
+------+

"<----------------- 2001:DB8::/40 (PA) " +------+
" 2001:DB8:10::/56 (PI) ---------------" +------+
" . . . . . . . . . . . . . . . . " +------+
" 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 901234567890 " +------+

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:

1. If all of the recursively nested ‘E1.∗’ enterprises are configured within the same IPv4 RLOC space, normal IPv4 forwarding will convey ‘K’s unencapsulated IPv4 packets toward ‘R2’ which then acts as an IPv4 Network Address Translator (NAT) and/or an ordinary IPv4 enterprise border router.

2. 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’ could act as a "Carrier-Grade NAT (CGN)", and ‘K’s IPv4 router (‘V’) could use the "dual-stack-lite" approach to convey ‘K’s packets to the CGN using IPv4-in-IPv6-in-IPv4 tunneling. The CGN then decapsulates the inner RLOC-based IPv4 packet and translates the IPv4 source address into a global IPv4 source address before sending the packets on to ‘S2’.

3. ‘K’ could be configured as an EID-based IPv6-only 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 [I-D.wing-nat-pt-replacement-comparison].

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 [I-D.templin-isatapv4], or by using some form of RLOC-based IPv4 NAT traversal.

3.4. Subnetwork Encapsulation and Adaptation Layer (SEAL)

Whenever the EBRs of disjoint enterprises are joined across a commons, mechanisms that rely on ICMP feedback from routers within the network may become brittle or susceptible to spoofing attacks. This is due to the fact that ICMP messages can be lost due to congestion or packet filtering gateways, and that network middleboxes are essentially "anonymous" and may include insufficient information in ICMPs that can be used to authenticate their source. Of even greater concern is the fact that a rogue node from a different enterprise could 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 Encapsulation Layer (SEAL) provides effective mitigations by only accepting packets from correspondent BRs that can be validated as topologically-correct routers within the commons (i.e., the subnetwork) using the VET Potential Router List (PRL) and ingress filtering [I-D.templin-autoconf-dhcp] in conjunction with the 32-bit SEAL_ID in the packet. Moreover, SEAL does not require reliable delivery of all ICMPs, but rather supports continued operation even if some/many ICMPs are lost. Finally, SEAL adapts to subnetworks that contain links with diverse MTUs properties, and can use probing to identify links in the path that configure marginal MTUs.

The advantages of using SEAL in conjunction with the RANGER enterprise-within-enterprise framework are tangible, and compare favorably with the alternative of deploying an all-IPv6 infrastructure even for clean-slate deployments. This is especially true within enterprises that provide a commons for joining organizational/political/functional entities with a spirit of mutual cooperation but with competing interests or objectives.

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 dampen mobility-related churn. For enterprises that require in-the-network confidentiality, 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 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. An analysis of other mobility approaches (e.g., Robin Whittle’s TTR proposal) will also be conducted.

3.6. Multihoming

As with mobility management, multi-homing 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 phase of flight, and 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.
4. Related Initiatives

4.1. 6over4 and ISATAP

Long before the RANGER architecture and VET/SEAL specifications were published, 6over4 [RFC2529] specified a means for automatic tunneling of unicast/multicast IPv6 packets over multicast-capable IPv4 enterprises, however it was unable to function in enterprises that did not support a full deployment of IPv4 multicast services.

To address these shortcomings, ISATAP [RFC5214][I-D.templin-isatapv4] was specified as a unicast-only variant of 6over4 and widely implemented among major software and equipment vendor products. ISATAP provides unicast-only neighbor discovery mechanisms and also adds a means for determining whether a node on an ISATAP interface is authorized to act as an IPv6 router; namely, the Potential Router List (PRL).

VET provides a functional superset of the 6over4 and ISATAP mechanisms; VET further combines with SEAL, IPSec, etc. to provide the functional elements of the RANGER architecture.

4.2. The Locator Identifier Split Protocol (LISP)

The Locator-Identifier Split Protocol (LISP) [I-D.farinacci-lisp] proposes a map-and-encaps architecture for scalable routing and addressing within the global Internet Default Free Zone (DFZ). Several companion documents (e.g., LISP-ALT, LISP-CONS, LISP-EMACS, LISP-NERD) propose mapping function solutions. A related mapping function solution proposal is found in [I-D.jen-apt].

LISP, and a number of related proposals being discussed in the Routing Research Group, share common properties with the solution proposed here. They may therefore be architecturally consistent with the RANGER architecture.

5. IANA Considerations

There are no IANA considerations for this document.

6. 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 must have a means of proving its ownership when it announces or withdraws the prefixes in enterprise routing systems. This can be accommodated through the use of SEcure Neighbor Discovery (SEND) [RFC3971] as well as a means for confirming prefix ownership, e.g., through name service lookup. The SEND mechanism is also useful for route optimization between lower-tier enterprises across a parent enterprise commons.

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 [I-D.ietf-v6ops-tunnel-security-concerns].
7. 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: Brian Carpenter, Eric Fleischman, Joe Halpern, Thomas Henderson, Steven Russert, Robin Whittle.

8. References

8.1. Normative References


8.2. Informative References


[I-D.jen-apt]

draft-jen-apt-01 (work in progress), November 2007.

[I-D.templin-autoconf-dhcp]
Templin, F., "Virtual Enterprise Traversal (VET)",
draft-templin-autoconf-dhcp-33 (work in progress),
February 2009.

[I-D.templin-isatapv4]
Templin, F., "Transmission of IPv4 Packets over ISATAP Interfaces",
draft-templin-isatapv4-00 (work in progress),
December 2008.

[I-D.templin-seal]
Templin, F., "The Subnetwork Encapsulation and Adaptation Layer (SEAL)",
draft-templin-seal-23 (work in progress),
August 2008.

[I-D.wing-nat-pt-replacement-comparison]
draft-wing-nat-pt-replacement-comparison-02 (work in progress),
September 2008.

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


[RFC1753]  Chiappa, J., "IPng Technical Requirements Of the Nimrod Routing and Addressing Architecture",

BCP 5, RFC 1918, February 1996.

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


[RFC2775]  Carpenter, B., "Internet Transparency",


Author’s Address

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

Email: fltemplin@acm.org