Internet Engineering Task Force (IETF)
Request for Comments: 6992
Category: Informational
ISSN: 2070-1721

D. Cheng
Huawei Technologies
M. Boucadair
France Telecom
A. Retana
Cisco Systems

July 2013

Routing for IPv4-Embedded IPv6 Packets

Abstract

This document describes a routing scenario where IPv4 packets are transported over an IPv6 network, based on the methods described in RFCs 6145 and 6052, along with a separate OSPFv3 routing table for IPv4-embedded IPv6 routes in the IPv6 network.

Status of This Memo

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

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6992.

Table of Contents

1. Introduction

This document describes a routing scenario where IPv4 packets are transported over an IPv6 network, based on [RFC6145] and [RFC6052], along with a separate OSPFv3 routing table for IPv4-embedded IPv6 routes in the IPv6 network. This document does not introduce any new IPv6 transition mechanism.

In this document, the following terminology is used:

- An IPv4-embedded IPv6 address denotes an IPv6 address that contains an embedded 32-bit IPv4 address constructed according to the rules defined in [RFC6052].

- IPv4-embedded IPv6 packets are packets of which destination addresses are IPv4-embedded IPv6 addresses.

- AFBR (Address Family Border Router) [RFC5565] refers to an edge router that supports both IPv4 and IPv6 address families, but the backbone network it connects to only supports either the IPv4 or IPv6 address family.

- AFXLBR (Address Family Translation Border Router) is defined in this document. It refers to a border router that supports both IPv4 and IPv6 address families located on the boundary of an IPv4-only network and an IPv6-only network and that is capable of performing IP header translation between IPv4 and IPv6 [RFC6145].

1.1. The Scenario

Due to exhaustion of public IPv4 addresses, there has been a continuing effort within the IETF to investigate and specify IPv6 transitional techniques. In the course of the transition, it is certain that networks based on IPv4 and IPv6 technologies, respectively, will coexist at least for some time. One such scenario is the interconnection of IPv4-only and IPv6-only networks, and in particular, when an IPv6-only network serves as an interconnection between several segregated IPv4-only networks. In this scenario, IPv4 packets are transported over the IPv6 network between IPv4 networks. In order to forward an IPv4 packet from a source IPv4 network to the destination IPv4 network, IPv4 reachability information must be exchanged between the IPv4 networks via some mechanism.

In general, running an IPv6-only network would reduce operational expenditures and optimize operations as compared to an IPv4-IPv6 dual-stack environment. Some proposed solutions allow the delivery of IPv4 services over an IPv6-only network. This document specifies an engineering technique that separates the routing table dedicated to IPv4-embedded IPv6 destinations from the routing table used for native IPv6 destinations.

OSPFv3 is designed to support multiple instances. Maintaining a separate routing table for IPv4-embedded IPv6 routes would simplify implementation, troubleshooting, and operation; it would also prevent overload of the native IPv6 routing table. A separate routing table can be generated from a separate routing instance.

1.2. Routing Solution per RFC 5565

The aforementioned scenario is described in [RFC5565], i.e., the IPv4-over-IPv6 scenario, where the network core is IPv6-only and the interconnected IPv4 networks are called IPv4 client networks. The P Routers (Provider Routers) in the core only support IPv6, but the AFBRs support IPv4 on interfaces facing IPv4 client networks and IPv6 on interfaces facing the core. The routing solution defined in [RFC5565] for this scenario is to run IBGP among AFBRs to exchange IPv4 routing information in the core, and the IPv4 packets are forwarded from one IPv4 client network to the other through a softwire using tunneling technology, such as MPLS, LSP, GRE, L2TPv3, etc.

1.3. An Alternative Routing Solution with OSPFv3

In this document, we propose an alternative routing solution for the scenario described in Section 1.1 where several segregated IPv4 networks, called IPv4 client networks, are interconnected by an IPv6 network. The IPv6 network and the interconnected IPv4 networks may or may not belong to the same Autonomous System (AS). We refer to the border node on the boundary of an IPv4 client network and the IPv6 network as an Address Family Translation Border Router (AFXLBR), which supports both the IPv4 and IPv6 address families and is capable of translating an IPv4 packet to an IPv6 packet, and vice versa, according to [RFC6145]. The described scenario is illustrated in Figure 1.

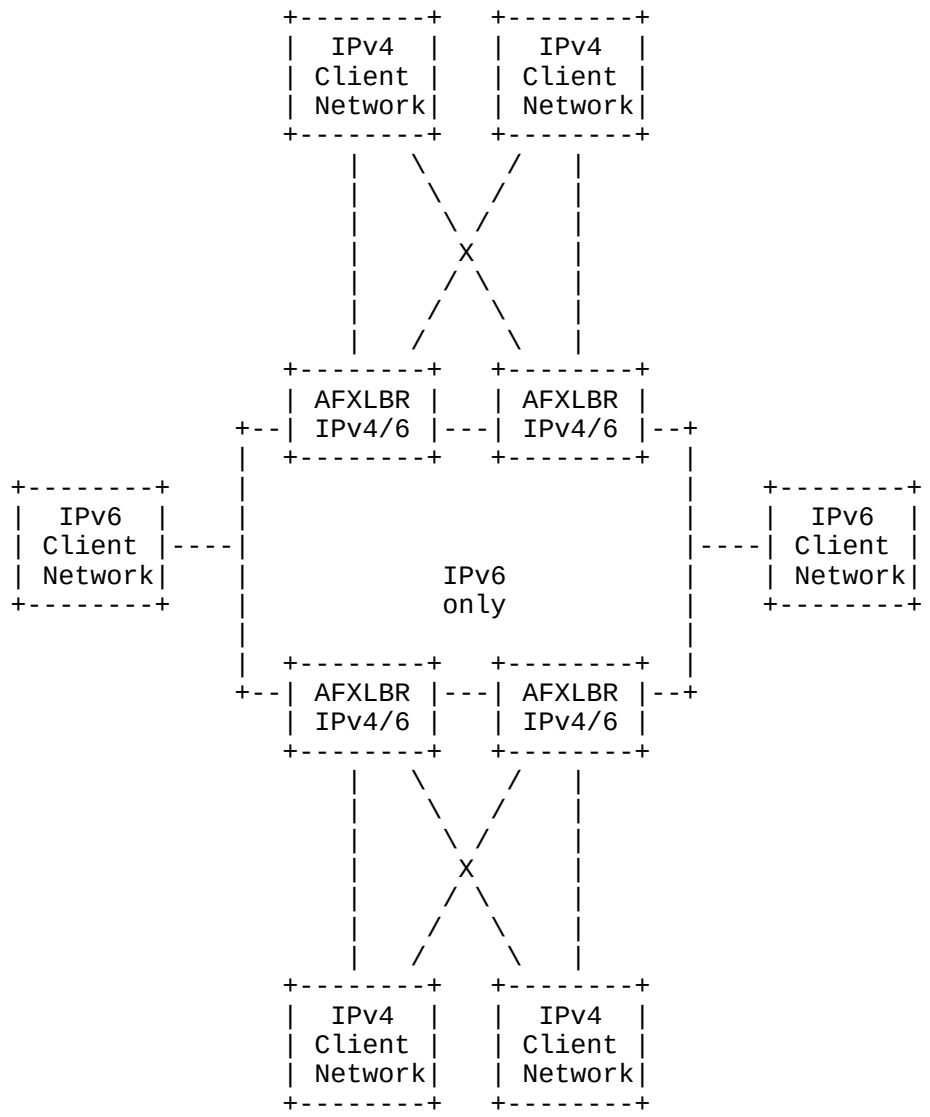

Figure 1: Segregated IPv4 Networks Interconnected by an IPv6 Network

Since the scenario occurs most commonly within an organization, an IPv6 prefix can be locally allocated and used by AFXLBRs to construct IPv4-embedded IPv6 addresses [RFC6052]. The embedded IPv4 address or prefix belongs to an IPv4 client network that is connected to the AFXLBR. An AFXLBR injects IPv4-embedded IPv6 addresses and prefixes into the IPv6 network using OSPFv3, and it also installs IPv4-embedded IPv6 routes advertised by other AFXLBRs.

When an AFXLBR receives an IPv4 packet from a locally connected IPv4 client network destined to a remote IPv4 client network, it translates the IPv4 header to the relevant IPv6 header [RFC6145], and in that process, the source and destination IPv4 addresses are translated into IPv4-embedded IPv6 addresses, respectively [RFC6052]. The resulting IPv6 packet is then forwarded to the AFXLBR that connects to the destination IPv4 client network. The remote AFXLBR derives the IPv4 source and destination addresses from the IPv4-embedded IPv6 addresses, respectively [RFC6052], and translates the header of the received IPv6 packet to the relevant IPv4 header [RFC6145]. The resulting IPv4 packet is then forwarded according to the IPv4 routing table maintained on the AFXLBR.

There are use cases where the proposed routing solution is useful. One case is that some border nodes do not participate in IBGP for the exchange of routes, or IBGP is not used at all. Another case is when tunnels are not deployed in the IPv6 network, or native IPv6 forwarding is preferred. Note that with this routing solution, the IPv4 and IPv6 header translation performed in both directions by the AFXLBR is stateless.

1.4. OSPFv3 Routing with a Specific Topology

In general, IPv4-embedded IPv6 packets can be forwarded just like native IPv6 packets with OSPFv3 running in the IPv6 network. However, this would require that IPv4-embedded IPv6 routes be flooded throughout the entire IPv6 network and stored on every router. This is not desirable from a scaling perspective. Moreover, since all IPv6 routes are stored in the same routing table, it would be inconvenient to manage the resource required for routing and forwarding based on traffic category, if so desired.

To improve the situation, a separate OSPFv3 routing table dedicated to the IPv4-embedded IPv6 topology can be constructed; that table would be solely used for routing IPv4-embedded IPv6 packets in the IPv6 network. The IPv4-embedded IPv6 topology includes all the participating AFXLBRs and a set of P Routers providing redundant connectivity with alternate routing paths.

To realize this, a separate OSPFv3 instance is configured in the IPv6 network [RFC5838]. This instance operates on all participating AFXLBRs and a set of P routers that interconnect them. As a result, there would be a dedicated IPv4-embedded IPv6 topology that is maintained on all these routers, along with a dedicated IPv4-embedded IPv6 routing table. This routing table in the IPv6 network is solely for forwarding IPv4-embedded IPv6 packets.

This document elaborates on how configuration is done with this method and on related routing issues.

This document only focuses on unicast routing for IPv4-embedded IPv6 packets using OSPFv3.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Provisioning

3.1. Deciding on the IPv4-Embedded IPv6 Topology

Before deploying configurations that use a separate OSPFv3 routing table for IPv4-embedded IPv6 addresses and prefixes, a decision must be made regarding the set of routers and their interfaces in the IPv6 network that should be part of the IPv4-embedded IPv6 topology.

For the purpose of this IPv4-embedded IPv6 topology, all AFXLBRs that connect to IPv4 client networks MUST be members of this topology. An AFXLBR MUST have at least one connection with a P Router in the IPv6 network or another AFXLBR.

The IPv4-embedded IPv6 topology is a subtopology of the entire IPv6 network, and if all routers (including AFXLBRs and P routers) and all their interfaces are included, the two topologies converge. Generally speaking, when this subtopology contains more interconnected P Routers, there would be more routing paths across the IPv6 network from one IPv4 client network to the other; however, this requires more routers in the IPv6 network to participate in IPv4-embedded IPv6 routing. In any case, the IPv4-embedded IPv6 topology MUST be continuous with no partitions.

3.2. Maintaining a Dedicated IPv4-Embedded IPv6 Routing Table

In an IPv6 network, in order to maintain a separate IPv6 routing table that contains routes for IPv4-embedded IPv6 destinations only, OSPFv3 needs to use the mechanism defined in [RFC5838].

It is assumed that the IPv6 network that is interconnected with IPv4 networks as described in this document is under one administration, and as such an OSPFv3 Instance ID (IID) is allocated locally and used for OSPFv3 operation dedicated to unicast IPv4-embedded IPv6 routing in an IPv6 network. This IID is configured on OSPFv3 router interfaces that participate in the IPv4-embedded IPv6 topology.

A locally configured OSPFv3 IID is allocated in the range 192 to 255, inclusive, in the "OSPFv3 Instance ID Address Family Values" registry; this range is reserved for "Private Use" [RFC6969]. This IID must be used to encode the "Instance ID" field in the packet header of OSPFv3 packets associated with the OSPFv3 instance.

In addition, the AF-bit in the OSPFv3 Option field MUST be set.

During Hello packet processing, an adjacency may only be established when the received Hello packet contains the same Instance ID as the Instance ID configured on the receiving OSPFv3 interface. This insures that only interfaces configured as part of the OSPFv3 unicast IPv4-embedded IPv6 topology are used for IPv4-embedded IPv6 unicast routing.

For more details, the reader is referred to [RFC5838].

4. Translation of IP Packets

When transporting IPv4 packets across an IPv6 network via the mechanism described above (Section 3.2), an IPv4 packet is translated to an IPv6 packet at the ingress AFXLBR, and the IPv6 packet is translated back to an IPv4 packet at the egress AFXLBR. IP packet header translation is accomplished in a stateless manner according to rules specified in [RFC6145]; the details of address translation are explained in the next subsection.

4.1. Address Translation

Prior to address translation, an IPv6 prefix is allocated by the operator, and it is used to form IPv4-embedded IPv6 addresses.

The IPv6 prefix can either be the IPv6 well-known prefix (WKP) 64:ff9b::/96 or a network-specific prefix that is unique to the organization; for the latter case, the IPv6 prefix length may be 32, 40, 48, 56, or 64. In either case, this IPv6 prefix is used during the address translation between an IPv4 address and an IPv4-embedded IPv6 address, as described in [RFC6052].

During translation from an IPv4 header to an IPv6 header at an ingress AFXLBR, the source IPv4 address and destination IPv4 address are translated into the corresponding source IPv6 address and destination IPv6 address, respectively. During translation from an IPv6 header to an IPv4 header at an egress AFXLBR, the source IPv6 address and destination IPv6 address are translated into the corresponding source IPv4 address and destination IPv4 address, respectively. Note that address translation is accomplished in a stateless manner.

When an IPv6 WKP is used, [RFC6052] allows only global IPv4 addresses to be embedded in the IPv6 address. An IPv6 address composed of a WKP and a non-global IPv4 address is hence invalid, and packets that contain such an address received by an AFXLBR are dropped.

In the case where both the IPv4 client networks and the IPv6 transit network belong to the same organization, non-global IPv4 addresses may be used with a network-specific prefix [RFC6052].

5. Advertising IPv4-Embedded IPv6 Routes

In order to forward IPv4 packets to the proper destination across an IPv6 network, IPv4 reachability information needs to be disseminated throughout the IPv6 network. This is performed by AFXLBRs that connect to IPv4 client networks using OSPFv3.

With the scenario described in this document, i.e., a set of AFXLBRs that interconnect multiple IPv4 client networks with an IPv6 network, the IPv4 networks and IPv6 networks belong to the same or separate Autonomous Systems (ASs), and as such, these AFXLBRs behave as AS Boundary Routers (ASBRs).

5.1. Advertising IPv4-Embedded IPv6 Routes through an IPv6 Transit Network

IPv4 addresses and prefixes in an IPv4 client network are translated into IPv4-embedded IPv6 addresses and prefixes, respectively, using the IPv6 prefix allocated by the operator and the method specified in [RFC6052]. These routes are then advertised by one or more attached ASBRs into the IPv6 transit network using AS-External-LSAs [RFC5340], i.e., with advertising scope comprising the entire Autonomous System.

5.1.1. Routing Metrics

By default, the metric in an AS-External-LSA that carries an IPv4-embedded IPv6 address or prefixes is a Type 1 external metric, which is comparable to the link state metric, and we assume that in most cases OSPFv2 is used in client IPv4 networks. This metric is added to the metric of the intra-AS path to the ASBR during the OSPFv3 route calculation. Through ASBR configuration, the metric can be set to a Type 2 external metric, which is considered much larger than the metric for any intra-AS path. Refer to the OSPFv3 specification [RFC5340] for more details. In either case, an external metric may take the same value as in an IPv4 network (using OSPFv2 or another routing protocol) but may also be specified based on some routing policy, the details of which are beyond the scope of this document.

5.1.2. Forwarding Address

If the "Forwarding Address" field of an OSPFv3 AS-External-LSA is used to carry an IPv6 address, that address must also be an IPv4-embedded IPv6 address where the embedded IPv4 address is the destination address in an IPv4 client network. However, since an AFXLBR sits on the border of an IPv4 network and an IPv6 network, it is RECOMMENDED that the "Forwarding Address" field not be used, so that the AFXLBR can make the forwarding decision based on its own IPv4 routing table.

5.2. Advertising IPv4 Addresses into Client Networks

IPv4-embedded IPv6 routes injected into the IPv6 network from one IPv4 client network MAY be advertised into another IPv4 client network after the associated destination addresses and prefixes are translated back to IPv4 addresses and prefixes, respectively. This operation is similar to normal OSPFv3 operation, wherein an AS-External-LSA can be advertised in a non-backbone area by default.

An IPv4 client network can limit which advertisements it receives through configuration.

For the purpose of this document, IPv4-embedded IPv6 routes MUST NOT be advertised into any IPv6 client networks that are also connected to the IPv6 transit network.

6. Aggregation on IPv4 Addresses and Prefixes

In order to reduce the amount of Link State Advertisements (LSAs) that are injected into the IPv6 network, an implementation should provide mechanisms to aggregate IPv4 addresses and prefixes at an AFXLBR prior to advertisement as IPv4-embedded IPv6 addresses and prefixes. In general, the aggregation practice should be based on routing policy, which is beyond the scope of this document.

7. Forwarding

There are three cases applicable to forwarding IP packets in the scenario described in this document:

1. On an AFXLBR, if an IPv4 packet is received on an interface connecting to an IPv4 segregated client network with a destination IPv4 address belonging to another IPv4 client network, the header of the packet is translated to the corresponding IPv6 header as described in Section 4, and the packet is then forwarded to the destination AFXLBR that advertised the IPv4-embedded IPv6 address into the IPv6 network.

2. On an AFXLBR, if an IPv4-embedded IPv6 packet is received and the embedded destination IPv4 address is in its IPv4 routing table, the header of the packet is translated to the corresponding IPv4 header as described in Section 4, and the packet is then forwarded accordingly.

3. On any router that is within the IPv4-embedded IPv6 topology subset of the IPv6 network, if an IPv4-embedded IPv6 packet is received and a route is found in the IPv4-embedded IPv6 routing table, the packet is forwarded to the IPv6 next hop, just like the handling for a normal IPv6 packet, without any translation.

The classification of an IPv4-embedded IPv6 packet is done according to the IPv6 prefix of the destination address, which is either the WKP (i.e., 64:ff9b::/96) or locally allocated as defined in [RFC6052].

8. Backdoor Connections

In some deployments, IPv4 client networks are interconnected across the IPv6 network but are also directly connected to each other. The direct connections between IPv4 client networks, sometimes called "backdoor" connections, can certainly be used to transport IPv4 packets between IPv4 client networks. In general, backdoor connections are preferred over the IPv6 network, since no address family translation is required.

9. Prevention of Loops

If an LSA sent from an AFXLBR into a client network could then be received by another AFXLBR, it would be possible for routing loops to occur. To prevent loops, an AFXLBR MUST set the DN bit [RFC4576] in any LSA that it sends to a client network. The AFXLBR MUST also ignore any LSA received from a client network that already has the DN bit set.

10. MTU Issues

In the IPv6 network, there are no new MTU issues introduced by this document. If a separate OSPFv3 instance (per [RFC5838]) is used for IPv4-embedded IPv6 routing, the MTU handling in the IPv6 network is the same as that of the default OSPFv3 instance.

However, the MTU in the IPv6 network may be different than that of IPv4 client networks. Since an IPv6 router will never fragment a packet, the packet size of any IPv4-embedded IPv6 packet entering the IPv6 network must be equal to or less than the MTU of the IPv6 network. In order to achieve this requirement, it is recommended

that AFXLBRs perform IPv6 path discovery among themselves. The resulting MTU, after taking into account the difference between the IPv4 header length and the IPv6 header length, must be "propagated" into IPv4 client networks, e.g., included in the OSPFv2 Database Description packet.

The details of passing the proper MTU into IPv4 client networks are beyond the scope of this document.

11. Security Considerations

There are several security aspects that require attention in the deployment practices described in this document.

In the OSPFv3 transit network, the security considerations for OSPFv3 are handled as usual, and in particular, authentication mechanisms described in [RFC6506] can be deployed.

When a separate OSPFv3 instance is used to support IPv4-embedded IPv6 routing, the same Security Association (SA) [RFC4552] MUST be used by the embedded IPv4 address instance as other instances utilizing the same link, as specified in [RFC5838].

Security considerations as documented in [RFC6052] must also be thought through and properly implemented, including the following:

- The IPv6 prefix that is used to carry an embedded IPv4 address (refer to Section 4.1) must be configured by the authorized operator on all participating AFXLBRs in a secure manner. This is to help prevent a malicious attack resulting in network disruption, denial of service, and possible information disclosure.

- Effective mechanisms (such as reverse path checking) must be implemented in the IPv6 transit network (including AFXLBRs) to prevent spoofing of embedded IPv4 addresses, which otherwise might be used as source addresses of malicious packets.

- If firewalls are used in IPv4 and/or IPv6 networks, configuration of the routers must be consistent, so that there are no holes in IPv4 address filtering.

The details of security handling are beyond the scope of this document.

12. Operational Considerations

This document puts together some mechanisms based on existing technologies developed by the IETF as an integrated solution to transport IPv4 packets over an IPv6 network using a separate OSPFv3 routing table. There are several aspects of these mechanisms that require attention for deployment and operation.

The tunnel-based solution documented in [RFC5565] and the solution proposed in this document are both used for transporting IPv4 packets over an IPv6 network, using different mechanisms. The two methods are not related to each other, and they can coexist in the same network if so deployed, without any conflict.

If one approach is to be deployed, the operator will decide which approach to use. Note that each approach has its own characteristics and requirements. For example, the tunnel-based solution requires a mesh of inter-AFBR softwires (tunnels) spanning the IPv6 network, as well as IBGP to exchange routes between AFBRs [RFC5565]; the approach in this document requires AFXLBRs that are capable of performing IPv4-IPv6 packet header translation per [RFC6145].

To deploy the solution as documented here, some configurations are required. An IPv6 prefix must first be chosen that is used to form all the IPv4-embedded IPv6 addresses and prefixes advertised by AFXLBRs in the IPv6 network; refer to Section 4.1 for details. The IPv4-embedded IPv6 routing table is created by using a separate OSPFv3 instance in the IPv6 network. As described in Section 3.2, this configuration is accomplished according to the mechanism described in [RFC5838].

Note that this document does not change any behavior of OSPFv3, and the existing or common practice should apply in the context of scalability. For example, the amount of routes that are advertised by OSPFv3 is one key concern. With the solution as described in this document, IPv4-embedded IPv6 addresses and prefixes will be injected by AFXLBRs into some part of the IPv6 network (see Section 3.1 for details), and a separate routing table will be used for IPv4-embedded IPv6 routing. Care must be taken during network design such that 1) aggregations are performed on IPv4 addresses and prefixes before being advertised in the IPv6 network as described in Section 6, and 2) estimates are made as to the amount of IPv4-embedded IPv6 routes that would be disseminated in the IPv6 network and to the size of the separate OSPFv3 routing table.

13. Acknowledgements

Many thanks to Acee Lindem, Dan Wing, Joel Halpern, Mike Shand, and Brian Carpenter for their comments.

14. References

14.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4576] Rosen, E., Psenak, P., and P. Pillay-Esnault, "Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4576, June 2006.

[RFC5565] Wu, J., Cui, Y., Metz, C., and E. Rosen, "Softwire Mesh Framework", RFC 5565, June 2009.

[RFC5838] Lindem, A., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC 5838, April 2010.

[RFC6145] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", RFC 6145, April 2011.

[RFC6969] Retana, A. and D. Cheng, "OSPFv3 Instance ID Registry Update", RFC 6969, July 2013.

14.2. Informative References

[RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, June 2006.

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

[RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.

[RFC6506] Bhatia, M., Manral, V., and A. Lindem, "Supporting Authentication Trailer for OSPFv3", RFC 6506, February 2012.

Authors' Addresses

Dean Cheng
Huawei Technologies
2330 Central Expressway
Santa Clara, California 95050
USA

EMail: dean.cheng@huawei.com

Mohamed Boucadair
France Telecom
Rennes, 35000
France

EMail: mohamed.boucadair@orange.com

Alvaro Retana
Cisco Systems
7025 Kit Creek Rd.
Research Triangle Park, North Carolina 27709
USA

EMail: aretana@cisco.com