Internet Engineering Task Force (IETF)
Request for Comments: 6535
Obsoletes: 2767, 3338
Category: Standards Track
ISSN: 2070-1721

B. Huang
H. Deng
China Mobile
T. Savolainen
Nokia

February 2012

Dual-Stack Hosts Using "Bump-in-the-Host" (BIH)

Abstract

Bump-in-the-Host (BIH) is a host-based IPv4 to IPv6 protocol translation mechanism that allows a class of IPv4-only applications that work through NATs to communicate with IPv6-only peers. The host on which applications are running may be connected to IPv6-only or dual-stack access networks. BIH hides IPv6 and makes the IPv4-only applications think they are talking with IPv4 peers by local synthesis of IPv4 addresses. This document obsoletes RFC 2767 and RFC 3338.

Status of This Memo

This is an Internet Standards Track document.

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). Further information on Internet Standards is available in 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/rfc6535.

Copyright Notice

Copyright (c) 2012 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. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction 4
1.1. Terminology 5
1.2. Acknowledgment of Previous Work 5
2. Components of the Bump-in-the-Host 6
2.1. Function Mapper 8
2.2. Protocol Translator 8
2.3. Extension Name Resolver 8
2.3.1. Special Exclusion Sets for A and AAAA Records 9
2.3.2. DNSSEC Support 10
2.3.3. Reverse DNS Lookup 10
2.3.4. DNS Caches and Synthetic IPv4 Addresses 10
2.4. Address Mapper 11
3. Behavior and Network Examples 11
4. Considerations 15
4.1. Socket API Conversion 15
4.2. Socket Bindings 15
4.3. ICMP Message Handling 15
4.4. IPv4 Address Pool and Mapping Table 15
4.5. Multi-Interface 17
4.6. Multicast 17
5. Application-Level Gateway Requirements Considerations 17
6. Security Considerations 17
6.1. Implications on End-to-End Security 18
6.2. Filtering 18
6.3. Attacks on BIH 18
6.4. DNS Considerations 19
7. Changes since RFC 2767 and RFC 3338 19
8. Acknowledgments 20
9. References 21
9.1. Normative References 21
9.2. Informative References 21
Appendix A. API List Intercepted by BIH 23

1. Introduction

This document describes Bump-in-the-Host (BIH), a successor and combination of the Bump-in-the-Stack (BIS)[RFC2767] and Bump-in-the-API (BIA) [RFC3338] technologies, which enable IPv4-only legacy applications to communicate with IPv6-only servers by synthesizing IPv4 addresses from AAAA records. Section 7 describes the reasons for making RFC 2767 and RFC 3338 obsolete.

The supported class of applications includes those that use DNS for IP address resolution and that do not embed IP address literals in application-protocol payloads. This includes legacy client-server applications using the DNS that are agnostic to the IP address family used by the destination and that are able to do NAT traversal. The synthetic IPv4 addresses shown to applications are taken from the private address pool of [RFC1918] in order to ensure that possible NAT traversal techniques will be initiated.

The IETF recommends using solutions based on dual stack or tunneling for IPv6 transition and specifically recommends against deployments utilizing double protocol translation. Use of BIH together with a NAT64 is NOT RECOMMENDED [RFC6180].

BIH includes two major implementation alternatives: a protocol translator between the IPv4 and the IPv6 stacks of a host or an API translator between the IPv4 socket API module and the TCP/IP module. Essentially, IPv4 is translated into IPv6 at the socket API layer or at the IP layer, the former of which is the recommended implementation alternative.

When BIH is implemented at the socket API layer, the translator intercepts IPv4 socket API function calls and invokes corresponding IPv6 socket API function calls to communicate with IPv6 hosts.

When BIH is implemented at the network layer, the IPv4 packets are intercepted and converted to IPv6 using the IP conversion mechanism defined in the Stateless IP/ICMP Translation Algorithm (SIIT) [RFC6145]. The protocol translation has the same benefits and drawbacks as SIIT.

The location of the BIH refers to the location of the protocol translation function. The location of the IPv4 address and DNS A record synthesis function is orthogonal to the location of the protocol translation and may or may not happen at the same location.

BIH can be used whenever an IPv4-only application needs to communicate with an IPv6-only server, independently of the address families supported by the access network. Hence, the access network can be IPv6-only or dual-stack capable.

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

This document uses terms defined in [RFC2460] and [RFC4213].

1.1. Terminology

DNS synthesis

The process of creating an A record containing a synthetic IPv4 address.

Real IPv4 address

An IPv4 address of a remote node a host has learned, for example, from DNS response to an A query.

Real IPv6 address

An IPv6 address of a remote node a host has learned, for example, from DNS response to a AAAA query.

Synthetic IPv4 address

An IPv4 address that has meaning only inside a host and that is used to provide IPv4 representation of remote node’s real IPv6 address.

1.2. Acknowledgment of Previous Work

This document is a direct derivative of [RFC2767], "Dual Stack Hosts using the "Bump-In-the-Stack" Technique (BIS)" by Kazuaki TSHUCHIYA, Hidemitsu HIGUCHI, and Yoshifumi ATARASHI and of [RFC3338], "Dual Stack Hosts Using "Bump-in-the-API" (BIA)" by Seungyun Lee, Myung-Ki Shin, Yong-Jin Kim, Alain Durand, and Erik Nordmark, which similarly provides IPv4-only applications on dual-stack hosts the means to operate over IPv6. Section 7 covers the changes since those documents.

2. Components of the Bump-in-the-Host

Figure 1 shows the architecture of a host in which BIH is implemented as a socket API-layer translator, i.e., as a "Bump-in-the-API".

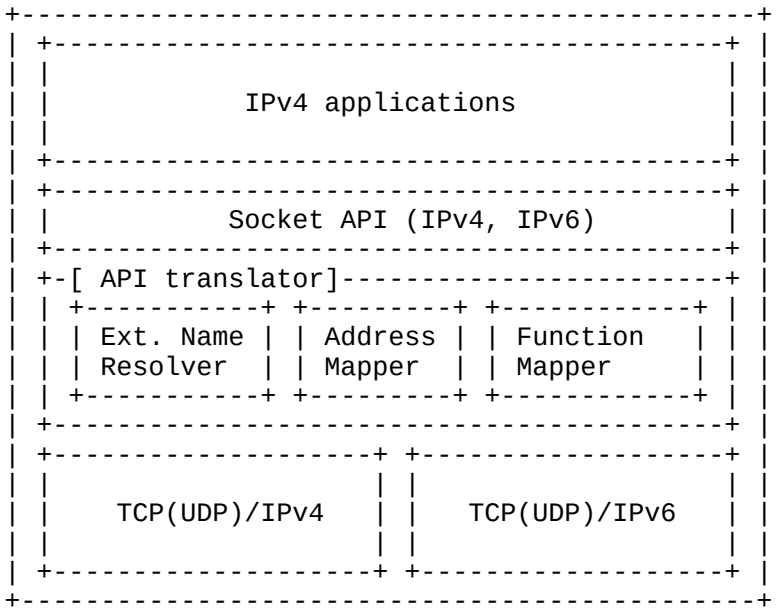


Figure 1: Architecture of a dual-stack host using protocol translation at the socket layer

Figure 2 shows the architecture of a host in which BIH is implemented as a network-layer translator, i.e., a "Bump-in-the-Stack".

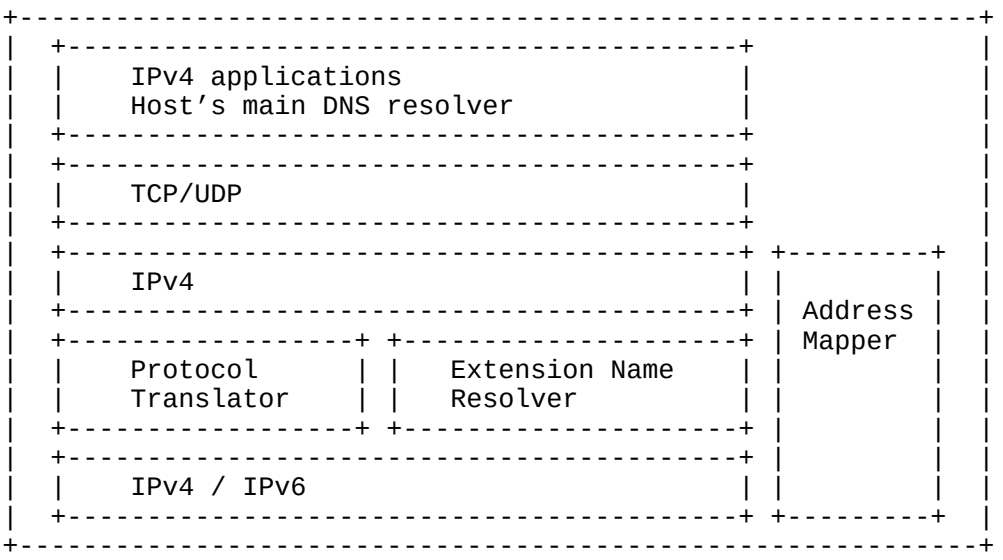


Figure 2: Architecture of a dual-stack host using protocol translation at the network layer

Dual-stack hosts, defined in [RFC4213], need applications, TCP/IP modules, and addresses for both IPv4 and IPv6. The proposed hosts in this document have an API or network-layer translator to allow legacy IPv4 applications to communicate with IPv6-only peers. The BIH architecture consists of an Extension Name Resolver, an address mapper, and depending on implementation either a function mapper or a protocol translator. It is worth noting that the Extension Name Resolver's placement is orthogonal to the placement of protocol translation. For example, the Extension Name Resolver may reside in the socket API while protocol translation takes place at the network layer.

The choice between the socket API- and network-layer architectures varies case by case. While the socket API architecture alternative is the recommended one, it may not always be possible to choose. This may be the case, for example, when the used operating system does not allow modifications to be done for API implementations, but does allow the addition of virtual network interfaces and related software modules. On the other hand, sometimes it may not be possible to introduce protocol translators inside the operating system, but it may be easy to modify implementations behind the API provided for applications. The choice of architecture also depends on who is creating implementation of BIH. For example, an

application framework provider, an operating system provider, and a device vendor may all choose different approaches due their different positions.

2.1. Function Mapper

The function mapper translates an IPv4 socket API function into an IPv6 socket API function.

When detecting IPv4 socket API function calls from IPv4 applications, the function mapper MUST intercept the function calls and invoke IPv6 socket API functions that correspond to the IPv4 socket API functions.

The function mapper MUST NOT perform function mapping when the application is initiating communications to the address range used by local synthesis and the address mapping table does not have an entry matching the address.

See Appendix A for an informational list of functions that would be appropriate to intercept by the function mapper.

2.2. Protocol Translator

The protocol translator translates IPv4 into IPv6, and vice versa, using the IP conversion mechanism defined in SIIT [RFC6145]. To avoid unnecessary fragmentation, the host's IPv4 module SHOULD be configured with a small enough MTU (MTU of the IPv6 enabled link - 20 bytes).

Protocol translation cannot be performed for IPv4 packets sent to the IPv4 address range used by local synthesis and for which a mapping table entry does not exist. The implementation SHOULD attempt to route such packets via IPv4 interfaces instead.

2.3. Extension Name Resolver

The Extension Name Resolver (ENR) returns an answer in response to the IPv4 application's name resolution request.

In the case of the socket API-layer implementation alternative, when an IPv4 application tries to do a forward lookup to resolve names via the resolver library (e.g., gethostbyname()), BIH intercepts the function call and instead calls the IPv6 equivalent functions (e.g., getaddrinfo()) that will resolve both A and AAAA records. This implementation alternative is name resolution protocol agnostic; hence, it supports techniques such as "hosts-file", NetBIOS, mDNS, and anything else the underlying operating system uses.

In the case of the network-layer implementation alternative, the ENR intercepts the A query and creates an additional AAAA query with similar content. The ENR will then collect replies to both A and AAAA queries and, depending on results, either return an A reply unmodified or synthesize a new A reply. If no reply for the A query is received after ENR-implementation-specific timeout, after reception of positive AAAA response, the ENR MAY choose to proceed as if there were only a AAAA record available for the destination.

The network-layer implementation alternative will only be able to catch applications' name resolution requests that result in actual DNS queries; hence, it is more limited when compared to the socket API-layer implementation alternative. Hence, the socket API-layer alternative is RECOMMENDED.

In either implementation alternative, if a DNS A record reply contains non-excluded real IPv4 addresses, the ENR MUST NOT synthesize IPv4 addresses.

The ENR asks the address mapper to assign a synthetic IPv4 address corresponding to each received IPv6 address if the A record query resulted in a negative response, all received real IPv4 addresses were excluded, or the A query timed out. The timeout value is implementation specific and may be short in order to provide a good user experience.

In the case of the API-layer implementation alternative, the ENR will simply make the API (e.g., gethostbyname) return the synthetic IPv4 address. In the case of the network-layer implementation alternative, the ENR synthesizes an A record for the assigned synthetic IPv4 address and delivers it up the stack. If the response contains a CNAME or a DNAME record, then the CNAME or DNAME chain is followed until the first terminating A or AAAA record is reached.

Application query	Network response	ENR behavior
IPv4 address(es)	IPv4 address(es)	return real IPv4 address(es)
IPv4 address(es)	IPv6 address(es)	synthesize IPv4 address(es)
IPv4 address(es)	IPv4/IPv6 address(es)	return real IPv4 address(es)

Figure 3: ENR Behavior Illustration

2.3.1. Special Exclusion Sets for A and AAAA Records

An ENR implementation SHOULD, by default, exclude certain real IPv4 and IPv6 addresses seen on received A and AAAA records. The addresses to be excluded by default MAY include addresses such as

those that should not appear in the DNS or on the wire (see Section 5.1.4 of [RFC6147] and [RFC5735]). Additional addresses MAY be excluded based on possibly configurable local policies.

2.3.2. DNSSEC Support

When the ENR is implemented at the network layer, the A record synthesis can cause similar issues as are described in [RFC6147] section 3. While running BIH, the main resolver of the host SHOULD NOT perform validation of A records, as synthetic A records created by ENR would fail in validation. While not running BIH, a host's resolver can use DNS Security (DNSSEC) in the same way that any other resolver can. The ENR MAY support DNSSEC, in which case the (stub) resolver on a host can be configured to trust validations done by the ENR located at the network layer. In some cases, the host's validating stub resolver can implement the ENR by itself.

When the ENR is implemented at the socket API level, there are no issues with DNSSEC use, as the ENR itself uses socket APIs for DNS resolution. This approach is RECOMMENDED.

2.3.3. Reverse DNS Lookup

When an application requests a reverse lookup (PTR query) for an IPv4 address, the ENR MUST check whether the queried IPv4 address can be found in the address mapper's mapping table and if it is a synthetic IPv4 address. If an entry is found and the queried IPv4 address is synthetic, the ENR MUST initiate a corresponding reverse lookup for the real IPv6 address. In the case where the application requested a reverse lookup for an address not part of the synthetic IPv4 address pool, e.g., a global address, the request MUST be passed on unmodified.

For example, when an application requests a reverse lookup for a synthetic IPv4 address, the ENR needs to intercept that query. The ENR asks the address mapper for the real IPv6 address that corresponds to the synthetic IPv4 address. The ENR shall perform a reverse lookup procedure for the destination's IPv6 address and return the name received as a response to the application that initiated the IPv4 query.

2.3.4. DNS Caches and Synthetic IPv4 Addresses

When BIH shuts down or address mapping table entries are cleared for any reason, DNS cache entries for synthetic IPv4 addresses MUST be flushed. There may be a DNS cache in the network-layer ENR itself and at the host's stub resolver.

2.4. Address Mapper

The address mapper maintains an IPv4 address pool that can be used for IPv4 address synthesis. The pool consists of the IPv4 addresses of [RFC1918] as per Section 4.4. Also, the address mapper maintains a table consisting of pairs of synthetic IPv4 addresses and destinations' real IPv6 addresses.

When the ENR, translator, or the function mapper requests the address mapper to assign a synthetic IPv4 address corresponding to an IPv6 address, the address mapper selects and returns an IPv4 address out of the local pool and registers a new entry into the table. The registration occurs in the following three cases:

1. When the ENR gets only IPv6 addresses for the target host name and there is no existing mapping entry for the IPv6 addresses. One or more synthetic IPv4 addresses will be returned to the application and mappings for synthetic IPv4 addresses to real IPv6 addresses are created.

2. When the ENR gets both real IPv4 and IPv6 addresses, but the real IPv4 addresses contain only excluded IPv4 addresses (e.g., 127.0.0.1). The behavior will follow case (1).

3. When the function mapper is triggered by a received IPv6 packet and there is no existing mapping entry for the IPv6 source address (for example, the client sent a UDP request to an anycast address, but a response was received from a unicast address).

Other possible combinations are outside of BIH.

3. Behavior and Network Examples

Figure 4 illustrates a very basic network scenario. An IPv4-only application is running on a host attached to the IPv6-only Internet and is talking to an IPv6-only server. Communication is made possible by Bump-in-the-Host.

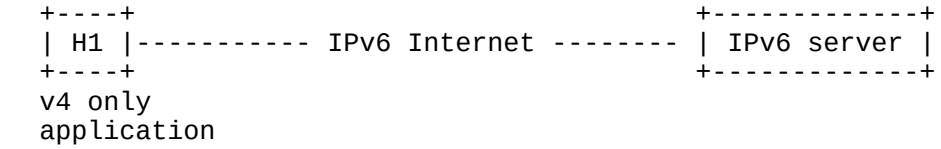


Figure 4: Network Scenario #1

Figure 5 illustrates a possible network scenario where an IPv4-only application is running on a host attached to a dual-stack network, but the destination server is running on a private site that is numbered with public IPv6 addresses and not globally reachable IPv4 addresses, such as the addresses of [RFC1918], without port forwarding set up on the NAT44. The only means to contact the server is to use IPv6.

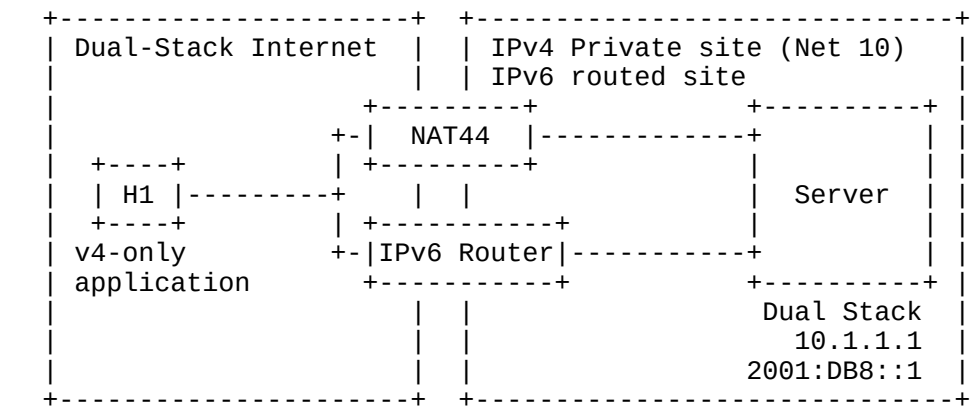
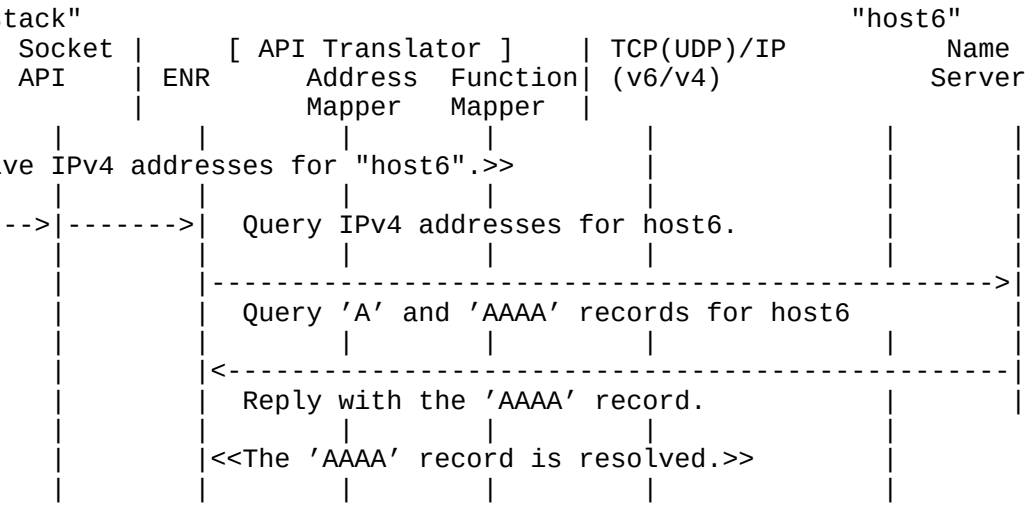


Figure 5: Network Scenario #2

Illustrations of host behavior in both implementation alternatives are given here. Figure 6 illustrates a setup where BIH (including the ENR) is implemented at the socket API layer, and Figure 7 illustrates a setup where BIH (including the ENR) is implemented at the network layer.



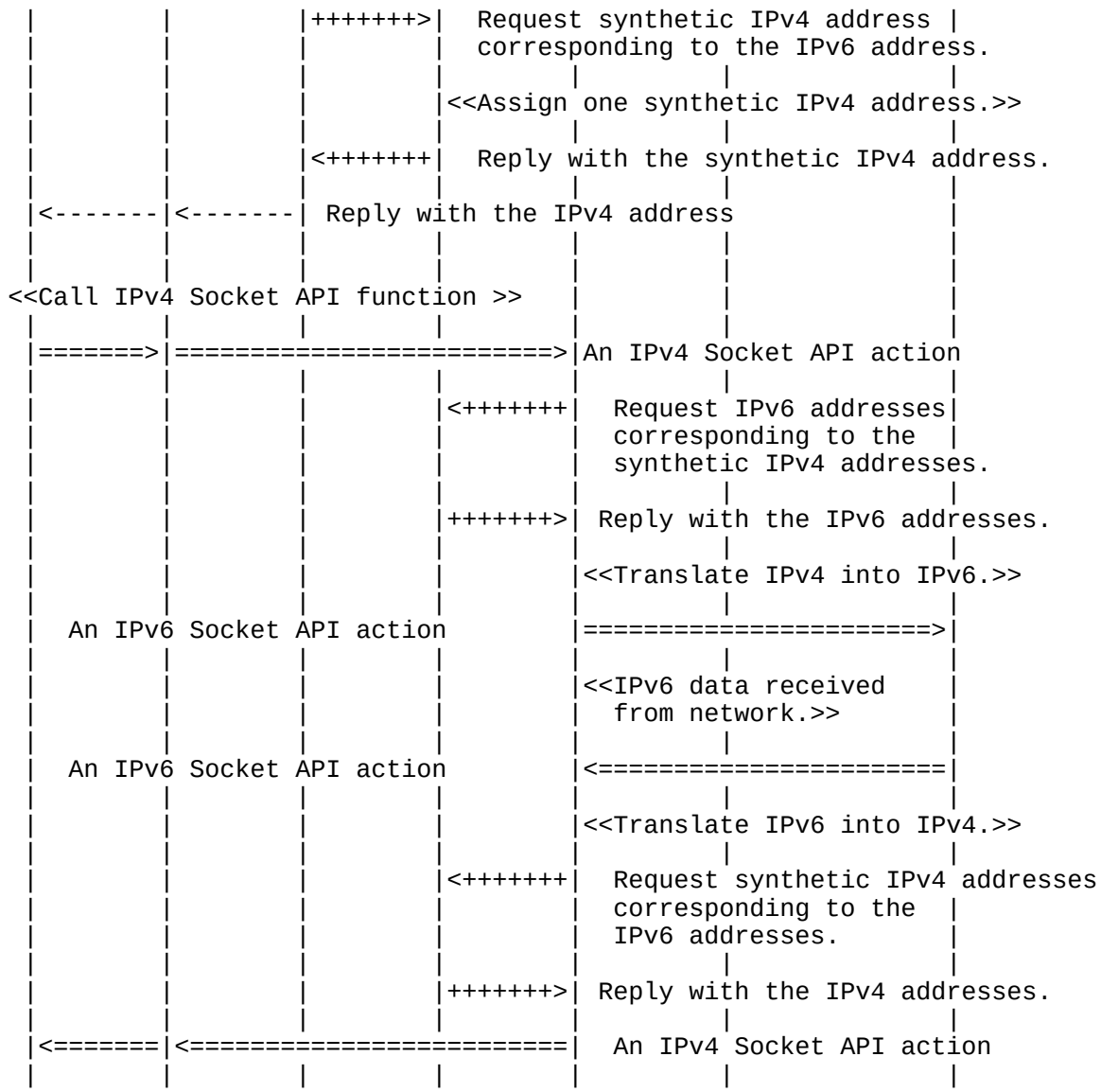
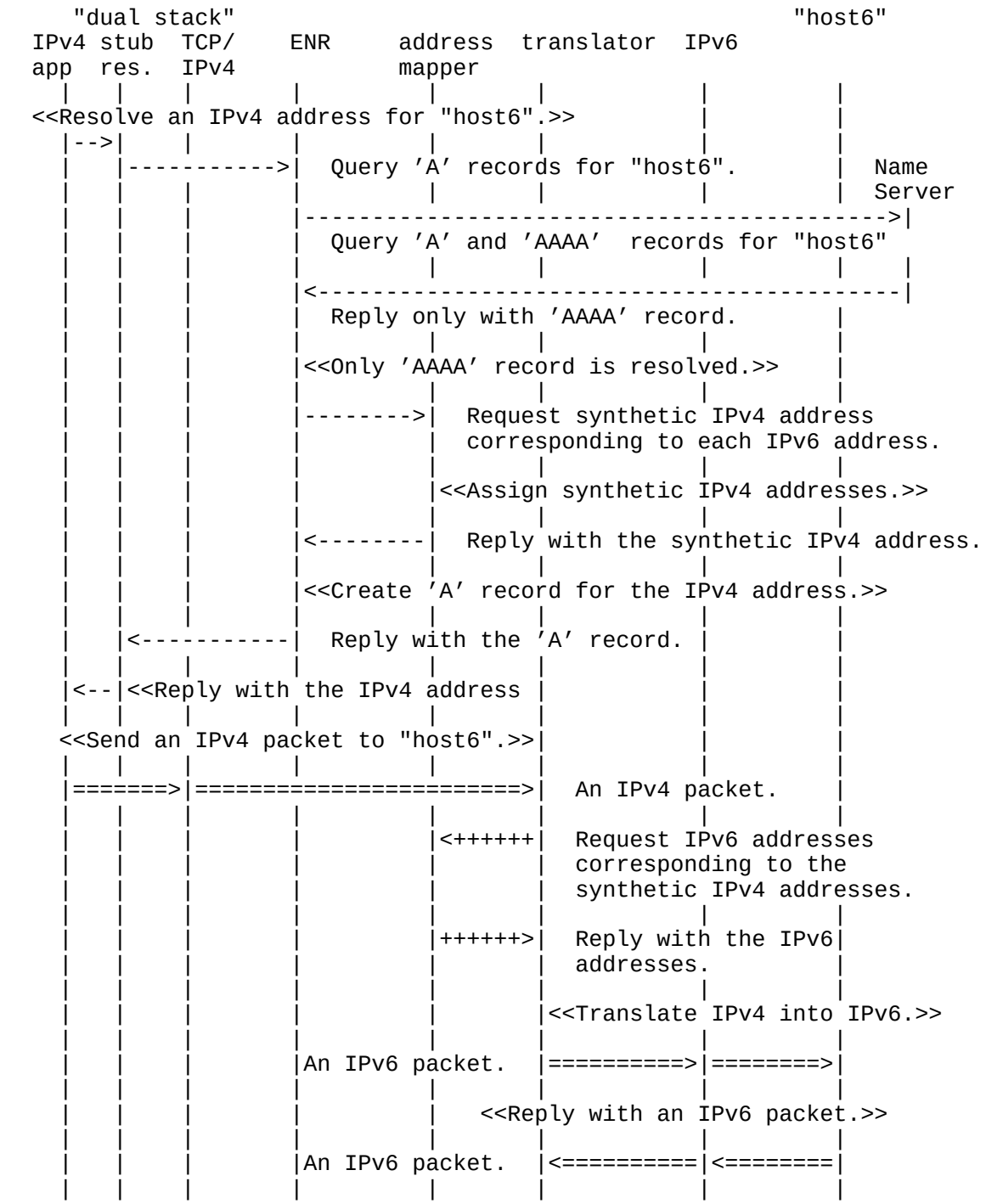


Figure 6: Example of BIH as API Addition


"dual stack"
"host6"
IPv4 stub app
res.
TCP/ IPv4
ENR
address mapper
translator
IPv6
<<Resolve an IPv4 address for "host6".>>
Query 'A' records for "host6".
Name Server
Query 'A' and 'AAAA' records for "host6"
Reply only with 'AAAA' record.
<<Only 'AAAA' record is resolved.>>
Request synthetic IPv4 address corresponding to each IPv6 address.
<<Assign synthetic IPv4 addresses.>>
Reply with the synthetic IPv4 address.
<<Create 'A' record for the IPv4 address.>>
Reply with the 'A' record.
<<Reply with the IPv4 address
<<Send an IPv4 packet to "host6".>>
An IPv4 packet.
Request IPv6 addresses corresponding to the synthetic IPv4 addresses.
Reply with the IPv6 addresses.
<<Translate IPv4 into IPv6.>>
An IPv6 packet.
<<Reply with an IPv6 packet.>>
An IPv6 packet.

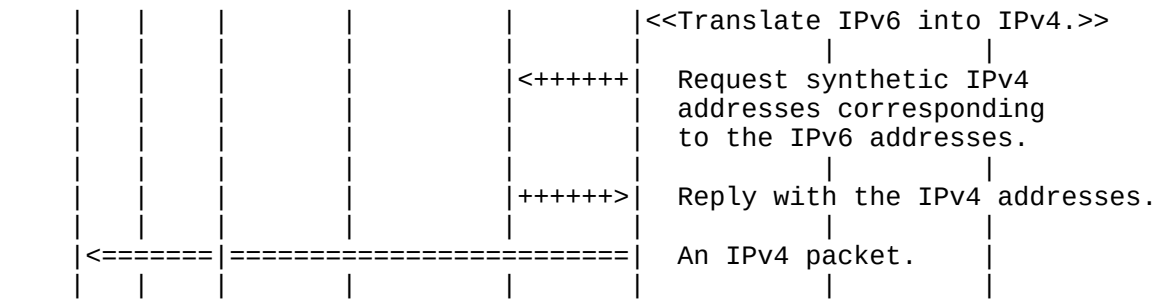


Figure 7: Example of BIH at the Network Layer

4. Considerations

4.1. Socket API Conversion

IPv4 socket API functions are translated into IPv6 socket API functions that are semantically as identical as possible, and vice versa. See Appendix A for the API list intercepted by BIH. However, some IPv4 socket API functions are not fully compatible with IPv6 since IPv4 supports features that are not present in IPv6, such as SO_BROADCAST.

4.2. Socket Bindings

BIH SHOULD select a source address for a socket from the recommended source address pool if a socket used for communications has not been explicitly bound to any IPv4 address.

The binding of an explicitly bound socket MUST NOT be changed by the BIH.

4.3. ICMP Message Handling

ICMPv4 and ICMPv6 messages MUST be translated as defined by SIIT [RFC6145]. In the network-layer implementation alternative, the protocol translator MUST translate ICMPv6 packets to ICMPv4 and vice versa, and in the socket API implementation alternative, the socket API MUST handle conversions in similar fashion.

4.4. IPv4 Address Pool and Mapping Table

The address pool consists of the private IPv4 addresses of [RFC1918]. This pool can be implemented at different granularities in the node, e.g., a single pool per node, or at some finer granularity such as per-user or per-process. In the case of a large number of IPv4 applications communicating with a large number of IPv6 servers, the

available address space may be exhausted if the granularity is not fine enough. This should be a rare event and chances will decrease as IPv6 support increases. The applications may use IPv4 addresses they learn for a much longer period than DNS time to live indicates. Therefore, the mapping table entries should be kept active for a long period of time. For example, a web browser may initiate one DNS query and then create multiple TCP sessions over time to the address it learns. When address mapping table clean-up is required, the BIH may utilize techniques used by network address translators, such as described in [RFC2663], [RFC5382], and [RFC5508].

The address space of RFC 1918 was chosen because legacy applications generally understand it as a private address space. A new dedicated address space would run the risk of not being understood by applications as private. 127/8 and 169.254/16 are rejected due to possible assumptions applications may make when seeing them.

The addresses of RFC 1918 used by the BIH have a risk of conflicting with addresses used in the host's possible IPv4 interfaces and corresponding local networks. The conflicts can be mitigated, but not fully avoided, by using less commonly used portions of the address space of RFC 1918. Addresses from 172.16/12 are thought to be less likely to be in conflict than addresses from 10/8 or 192.168/16 spaces. A source address can usually be selected in a non-conflicting manner, but a small possibility exists for synthesized destination addresses being in conflict with real addresses used in attached IPv4 networks.

The RECOMMENDED IPv4 addresses are following:

Primary source addresses: 172.21.112.0/20.

Source addresses have to be allocated because applications use getsockname() calls and, in the network-layer mode, an IP address of the IPv4 interface has to be shown (e.g., by 'ifconfig'). More than one address is allocated to allow implementation flexibility, e.g., for cases where a host has multiple IPv6 interfaces. The source addresses are from different subnets than destination addresses to ensure applications would not make on-link assumptions and would instead enable NAT traversal functions.

Secondary source addresses: 10.170.224.0/20.

These addresses are recommended if a host has a conflict with primary source addresses.

Primary destination addresses: 10.170.160.0/20.

The address mapper will select destination addresses primarily out of this pool.

Secondary destination addresses: 172.21.80.0/20.

The address mapper will select destination addresses out of this pool if the node has a dual-stack connection conflicting with primary destination addresses.

4.5. Multi-Interface

In the case of dual-stack destinations, BIH MUST NOT do protocol translation from IPv4 to IPv6 when the host has any IPv4 interfaces, native or tunneled, available for use.

It is possible that an IPv4 interface is activated during BIH operation, for example, if a node moves to a coverage area of an IPv4-enabled network. In such an event, BIH MUST stop initiating protocol translation sessions for new connections, and BIH MAY disconnect active sessions. The choice of disconnection is left for implementations, and it may depend on whether IPv4 address conflict occurs between addresses used by BIH and addresses used by the new IPv4 interface.

4.6. Multicast

Protocol translation for multicast is not supported.

5. Application-Level Gateway Requirements Considerations

No Application-Level Gateway (ALG) functionality is specified herein as ALG design is generally not encouraged for host-based translation and as BIH is intended for applications that do not include IP addresses in protocol payloads.

6. Security Considerations

The security considerations of BIH follows closely, but not completely, those of NAT64 [RFC6146] and DNS64 [RFC6147]. The following sections are copied from RFC 6146 and RFC 6147 and modified for BIH.

6.1. Implications on End-to-End Security

Any protocols that protect IP header information are essentially incompatible with BIH. This implies that end-to-end IPsec verification will fail when the Authentication Header (AH) is used (both transport and tunnel mode) and when ESP is used in transport mode. This is inherent in any network-layer translation mechanism. End-to-end IPsec protection can be restored, using UDP encapsulation as described in [RFC3948]. The actual extensions to support IPsec are out of the scope of this document.

6.2. Filtering

BIH creates binding state using packets flowing from the IPv4 side to the IPv6 side. In accordance with the procedures defined in this document, following the guidelines defined in [RFC4787], a BIH implementation MUST offer "Endpoint-Independent Mapping".

Implementations MAY also provide support for "Address-Dependent Mapping" following the guidelines defined in [RFC4787].

The security properties, however, are determined by which packets the BIH allows in and which it does not. The security properties are determined by the filtering behavior and by the possible filtering configuration in the filtering portions of the BIH, not by the address mapping behavior.

6.3. Attacks on BIH

The BIH implementation itself is a potential victim of different types of attacks. In particular, the BIH can be a victim of Denial-of-Service (DoS) attacks. The BIH implementation has a limited number of resources that can be consumed by attackers creating a DoS attack. The BIH has a limited number of IPv4 addresses that it uses to create the bindings. Even though the BIH performs address translation, it is possible for an attacker to consume the synthetic IPv4 address pool by triggering a host to issue DNS queries for names that cause ENR to synthesize A records. DoS attacks can also affect other limited resources available in the host running BIH such as memory or link capacity. For instance, it is possible for an attacker to launch a DoS attack on the memory of the BIH running device by sending fragments that the BIH will store for a given period. If the number of fragments is large enough, the memory of the host could be exhausted. BIH implementations MUST implement proper protection against such attacks, for instance, allocating a limited amount of memory for fragmented packet storage.

Another consideration related to BIH resource depletion is the preservation of binding state. Attackers may try to keep a binding state alive forever by sending periodic packets that refresh the state. In order to allow the BIH to defend against such attacks, the BIH implementation MAY choose not to extend the session entry lifetime for a specific entry upon the reception of packets for that entry through the external interface. However, such an action would not allow one-way communication sessions to stay alive.

6.4. DNS Considerations

BIH operates in combination with the DNS, and it is therefore subject to whatever security considerations are appropriate to the DNS mode in which the BIH is operating (i.e., recursive or stub-resolver mode).

BIH has the potential to interfere with the functioning of DNSSEC, because BIH modifies DNS answers, and DNSSEC is designed to detect such modifications and to treat modified answers as bogus.

7. Changes since RFC 2767 and RFC 3338

This document combines and obsoletes both [RFC2767] and [RFC3338].

The changes in this document mainly reflect the following:

1. Addresses of RFC 1918 used for synthesis

   RFC 3338 used unassigned IPv4 addresses (e.g., 0.0.0.1 - 0.0.0.255) for synthetic IPv4 addresses. Those addresses should not have been used and that may cause problems with applications. It is preferable to use addresses defined in RFC 1918 instead, as described in Section 4.4.

2. Support for reverse (PTR) DNS queries

   Neither RFC 2767 nor RFC 3338 included support for reverse (PTR) DNS queries. This document adds the support in Section 2.3.3.

3. DNSSEC support

   RFC 2767 did not include DNSSEC considerations, which are now included in Section 2.3.2

4. Architectural recommendation

   This document recommends the socket API-layer implementation option over network layer translation, i.e., it recommends the approach introduced in RFC 2767 over the approach of RFC 3338.

5. Standards-Track document

   RFC 2767 is classified as an Informational RFC and RFC 3338 as an Experimental RFC. It was discussed and decided in the IETF that this technology should be on the Standards Track.

6. Set of other extensions and improvements

   A set of lesser extensions, improvements, and clarifications have been introduced. These include but are not limited to IPv4 and IPv6 address exclusion sets at Section 2.3.1, host's DNS cache considerations, ENR behavior updates, updated security considerations, example updates, and deployment scenario updates.

8. Acknowledgments

The authors are grateful for discussion from Gang Chen, Dapeng Liu, Bo Zhou, Hong Liu, Tao Sun, Zhen Cao, and Feng Cao et al. in the development of this document.

The efforts of Mohamed Boucadair, Dean Cheng, Lorenzo Colitti, Paco Cortes, Ralph Droms, Stephen Farrell, Fernando Gont, Marnix Goossens, Wassim Haddad, Ala Hamarsheh, Dave Harrington, Ed Jankiewizh, Suresh Krishnan, Julien Laganier, Yiu L. Lee, Jan M. Melen, Qibo Niu, Pierrick Seite, Christian Vogt, Magnus Westerlund, Dan Wing, and James Woodyatt in reviewing this document are gratefully acknowledged.

Special acknowledgments go to Dave Thaler for his extensive review and support.

The authors of RFC 2767 acknowledged WIDE Project, Kazuhiko YAMAMOTO, Jun MURAI, Munechika SUMIKAWA, Ken WATANABE, and Takahisa MIYAMOTO. The authors of RFC 3338 acknowledged implementation contributions by Wanjik Lee (wjlee@arang.miryang.ac.kr) and i2soft Corporation (www.i2soft.net).

The authors of "Bump-in-the-Wire IPv4/IPv6 Translator" (a draft document submitted to the v6ops WG in October 2006), P. Moster, L. Chin, and D. Green, are acknowledged. Some ideas and clarifications from BIW have been adapted to this document.

9. References

9.1. Normative References

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

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

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

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

[RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

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

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.

[RFC6147] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", RFC 6147, April 2011.

9.2. Informative References

[RFC2663] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999.

[RFC2767] Tsuchiya, K., HIGUCHI, H., and Y. Atarashi, "Dual Stack Hosts using the "Bump-In-the-Stack" Technique (BIS)", RFC 2767, February 2000.

[RFC3338] Lee, S., Shin, M-K., Kim, Y-J., Nordmark, E., and A. Durand, "Dual Stack Hosts Using "Bump-in-the-API" (BIA)", RFC 3338, October 2002.

[RFC3493] Gilligan, R., Thomson, S., Bound, J., McCann, J., and W. Stevens, "Basic Socket Interface Extensions for IPv6", RFC 3493, February 2003.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, January 2005.

[RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.

[RFC5508] Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP", BCP 148, RFC 5508, April 2009.

[RFC5735] Cotton, M. and L. Vegoda, "Special Use IPv4 Addresses", BCP 153, RFC 5735, January 2010.

[RFC6180] Arkko, J. and F. Baker, "Guidelines for Using IPv6 Transition Mechanisms during IPv6 Deployment", RFC 6180, May 2011.

Appendix A. API List Intercepted by BIH

The following informational list includes some of the API functions that would be appropriate to intercept by BIH module when implemented at the socket API layer. Please note that this list is not fully exhaustive, as the function names and services that are available on different APIs vary significantly.

The functions that the application uses to pass addresses into the system are as follows:

- bind()
- connect()
- sendmsg()
- sendto()
- gethostbyaddr()
- getnameinfo()

The functions that return an address from the system to an application are as follows:

- accept()
- recvfrom()
- recvmsg()
- getpeername()
- getsockname()
- gethostbyname()
- getaddrinfo()

The functions that are related to socket options are as follows:

- getsocketopt()
- setsocketopt()

As well, raw sockets for IPv4 and IPv6 may be intercepted.

Most of the socket functions require a pointer to the socket address structure as an argument. Each IPv4 argument is mapped into corresponding an IPv6 argument, and vice versa.

According to [RFC3493], the following new IPv6 basic APIs and structures are required.

IPv4	new IPv6
AF_INET	AF_INET6
sockaddr_in	sockaddr_in6
gethostbyname()	getaddrinfo()
gethostbyaddr()	getnameinfo()
inet_ntoa()/inet_addr()	inet_pton()/inet_ntop()
INADDR_ANY	in6addr_any

Figure 8

BIH may intercept inet_ntoa() and inet_addr() and use the address mapper for those. Doing that enables BIH to support literal IP addresses. However, IPv4 address literals can only be used after a mapping entry between the IPv4 address and corresponding IPv6 address has been created.

The gethostbyname() and getaddrinfo() calls return a list of addresses. When the name resolver function invokes getaddrinfo(), and getaddrinfo() returns multiple IP addresses, whether IPv4 or IPv6, they should all be represented in the addresses returned by gethostbyname(). Thus, if getaddrinfo() returns multiple IPv6 addresses, this implies that multiple address mappings will be created: one for each IPv6 address.

Authors' Addresses

Bill Huang
China Mobile
No.32 Xuanwumen West Street
Xicheng District
Beijing 100053
China

EMail: bill.huang@chinamobile.com

Hui Deng
China Mobile
No.32 Xuanwumen West Street
Xicheng District
Beijing 100053
China

EMail: denghui@chinamobile.com

Teemu Savolainen
Nokia
Hermiankatu 12 D
FI-33720 TAMPERE
Finland

EMail: teemu.savolainen@nokia.com