

Internet Engineering Task Force
Internet-Draft
Intended status: BCP
Expires: January 6, 2013

I. Yamagata
S. Miyakawa
NTT Communications
A. Nakagawa
Japan Internet Exchange (JPIX)
J. Yamaguchi
Fiber 26 Network
H. Ashida
IS Consulting G.K.
July 5, 2012

ISP Shared Address
draft-shirasaki-isp-shared-addr-08

Abstract

This document defines IPv4 ISP Shared Address to be jointly used among Internet Service Providers (ISPs). This space is intended to be used in NAT444 model which is used during the transition period to IPv6.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on January 6, 2013.

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
2. ISP Shared Address	4
2.1. Definition	4
2.2. Details	4
3. Size of Address Space	5
4. Acknowledgements	5
5. IANA Considerations	5
6. Security Considerations	5
7. References	5
7.1. Normative References	5
7.2. Informative References	5
Authors' Addresses	6

1. Introduction

The only permanent solution of the IPv4 address exhaustion is to deploy IPv6. Now, just before the exhaustion, it's time to make a transition to IPv6.

NAT444 model [I-D.shirasaki-nat444] is one of the solutions for transition to IPv6.

This document defines ISP Shared Address to be used in NAT444 model [I-D.shirasaki-nat444-isp-shared-addr]. It is supposed to be used between Customer Premises Equipment (CPE) and Carrier Grade NAT (CGN) [I-D.ietf-behave-lsn-requirements].

ISP Shared Address is needed until the IPv4 Internet fades out.

2. ISP Shared Address

2.1. Definition

ISP Shared Address is intended to be assigned between CPE and CGN in a NAT444.

2.2. Details

- Each ISP can use ISP Shared Address without any coordination with IANA or Internet registries.
- ISP Shared Address can be used by many ISPs.
- ISP has to install CGN to use ISP Shared Address.
- ISP Shared Address must not be used at customers' site or Internet Exchanges.
- Routing information of ISP Shared Address must not be advertised to the Internet.
- Reverse DNS queries for this address space must not be sent to root DNS servers.
- Packets with this space as source address and/or destination address must be filtered out at the border of each ISP.
- Addresses within this address space should be unique within the ISP, or the set of ISPs which choose to cooperate over this space so they may directly communicate with each other in their networks.

3. Size of Address Space

Because the aggregation size of Tokyo area POP is around /10 in Japan, /10 should be the hard limit of minimum size ISP Shared Address. We understand this can be determined by further discussions.

4. Acknowledgements

Thanks for the input and review by Shirou Niinobe, Takeshi Tomochika, Tomohiro Fujisaki, Dai Nishino, JP address community members, AP address community members and JPNIC members.

5. IANA Considerations

IANA is to record the allocation of the IPv4 global unicast address as ISP Shared Address in the IPv4 address registry.

6. Security Considerations

ISP Shared Address is supposed to be used with CGN. The Global IPv4 address that is assigned outside CGN may be used as source address of 'Denial of Service' attack.

7. References

7.1. Normative References

[I-D.ietf-behave-lsn-requirements]
Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A.,
and H. Ashida, "Common requirements for Carrier Grade NATs
(CGNs)", draft-ietf-behave-lsn-requirements-07 (work in
progress), June 2012.

7.2. Informative References

[I-D.shirasaki-nat444-isp-shared-addr]
Yamaguchi, J., Shirasaki, Y., Miyakawa, S., Nakagawa, A.,
and H. Ashida, "NAT444 addressing models",
draft-shirasaki-nat444-isp-shared-addr-07 (work in
progress), January 2012.

[I-D.shirasaki-nat444]
Yamagata, I., Shirasaki, Y., Nakagawa, A., Yamaguchi, J.,

and H. Ashida, "NAT444", draft-shirasaki-nat444-05 (work in progress), January 2012.

[PROP58] Niinobe, S., Tomochika, T., Yamaguchi, J., Nishino, D., Ashida, H., Nakagawa, A., and T. Hosaka, "Proposal to create IPv4 shared use address space among LIRs", 2008, <<http://www.apnic.net/policy/proposals/prop-058-v001.html>>.

Authors' Addresses

Ikuhei Yamagata
NTT Communications Corporation
Gran Park Tower 17F, 3-4-1 Shibaura, Minato-ku
Tokyo 108-8118
Japan

Phone: +81 3 6700 8530
Email: ikuhei@nttv6.jp

Shin Miyakawa
NTT Communications Corporation
Gran Park Tower 17F, 3-4-1 Shibaura, Minato-ku
Tokyo 108-8118
Japan

Phone: +81 50 3812 4695
Email: miyakawa@nttv6.jp

Akira Nakagawa
Japan Internet Exchange Co., Ltd. (JPPIX)
Otemachi Building 21F, 1-8-1 Otemachi, Chiyoda-ku
Tokyo 100-0004
Japan

Phone: +81 90 9242 2717
Email: a-nakagawa@jpix.ad.jp

Jiro Yamaguchi
Fiber 26 Network Inc.
Haraguchi bldg., 5F, 3-11-4 Kanda Jinbo-cho, Chiyoda-ku
Tokyo 101-0051
Japan

Phone: +81 50 3463 6109
Email: jiro-y@f26n.jp

Hiroyuki Ashida
IS Consulting G.K.
12-17 Odenma-cho, Nihonbashi, Chuo-ku
Tokyo 103-0011
Japan

Email: assie@hir.jp

Network Working Group
Internet-Draft
Intended status: Informational
Expires: July 10, 2010

X. Li
C. Bao
M. Chen
H. Zhang
J. Wu
CERNET Center/Tsinghua University
January 6, 2010

The CERNET IVI Translation Design and Deployment for the IPv4/IPv6
Coexistence and Transition
draft-xli-behave-ivi-07

Abstract

This document presents the China Education and Research Network (CERNET)'s IVI translation design and deployment for the IPv4/IPv6 coexistence and transition.

The IVI is a prefix-specific and stateless address mapping mechanism for "an IPv6 network to the IPv4 Internet" and "the IPv4 Internet to an IPv6 network" scenarios. In the IVI design, subsets of the ISP's IPv4 addresses are embedded in the ISP's IPv6 addresses and the hosts using these IPv6 addresses can therefore communicate with the global IPv6 Internet directly and can communicate with the global IPv4 Internet via stateless translators, the communications can either be IPv6 initiated or IPv4 initiated. The IVI mechanism supports the end-to-end address transparency and incremental deployment. The IVI is an early design deployed in CERNET as a reference for the IETF standard documents on IPv4/IPv6 translation.

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/lid-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 July 10, 2010.

Copyright Notice

Copyright (c) 2010 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 BSD License.

Table of Contents

1. Introduction	4
1.1. Analysis of IPv4-IPv6 Translation Mechanisms	4
1.2. CERNET Translation Requirements	5
2. Terms and Abbreviations	7
3. The IVI Translation Algorithm	7
3.1. Address Format	9
3.2. Routing and Forwarding	9
3.3. Network-layer Header Translation	11
3.4. Transport-layer Header Translation	12
3.5. Fragmentation and MTU Handling	12
3.6. ICMP Handling	12
3.7. Application Layer Gateway	13
4. The IVI DNS Configuration	13
4.1. DNS Configuration for the IVI6(i) Addresses	13
4.2. DNS Service for the IVIG6(i) Addresses	13
5. The Advanced IVI Translation Functions	13
5.1. IVI Multicast	13
6. IVI Host Operation	14
6.1. IVI Address Assignment	14
6.2. IPv6 Source Address Selection	14
7. The IVI Implementation	15
7.1. Linux Implementation	15
7.2. Testing Environment	15
8. Security Considerations	15
9. IANA Considerations	16
10. Contributors	16
11. Acknowledgments	16
12. Appendix A. The IVI translator configuration example	17
13. Appendix B. The traceroute results	18
14. References	20
14.1. Normative References	20
14.2. Informative References	21
Authors' Addresses	22

1. Introduction

This document presents the CERNET IVI translation design and deployment for the IPv4/IPv6 coexistence and transition. In roman numerals, the IV stands for 4 and VI stands for 6, so IVI stands for the IPv4/IPv6 translation.

The experiences with IPv6 deployment in the past 10 years indicate that the ability to communicate between IPv4 and IPv6 address families would be beneficial. However, the current transition methods do not fully support this requirement [RFC4213]. For example, dual-stack hosts can communicate with both the IPv4 and IPv6 hosts, but single-stack hosts can only communicate with hosts in the same address family. While the dual-stack approach continues to work in many cases even in the face of IPv4 address depletion [COUNT], there are situations where it would be desirable to communicate with a device in another address family. Tunneling-based architectures can link the IPv6 islands across IPv4 networks, but they cannot provide communication between the two different address families [RFC3056] [RFC5214] [RFC4380]. Translation can relay communications for hosts located in IPv4 and IPv6 networks, but the current implementation of this kind of architecture is not scalable and it cannot maintain end-to-end address transparency [RFC2766] [RFC3142] [RFC4966] [RFC2775].

1.1. Analysis of IPv4-IPv6 Translation Mechanisms

Since IPv4 and IPv6 are different protocols with different addressing structures, a translation mechanism is necessary for communication between endpoints using different address families. There are several ways to implement the translation. One is the stateless IP/ICMP translation algorithm (SIIT) [RFC2765], which provides a mechanism for translation between IPv4 and IPv6 packet headers (including ICMP headers) without requiring any per-connection state. But, SIIT does not specify the address assignment and routing scheme [RFC2766]. For example, the SIIT uses IPv4 mapped IPv6 addresses [::FFFF:ipv4-addr/96] and IPv4 compatible IPv6 addresses [::ipv4-address/96] for the address mapping, but these addresses violate the aggregation principle of IPv6 routing [RFC4291]. The other translation mechanism is NAT-PT, which has serious technical and operational difficulties and IETF has reclassified it from proposed standard to historic status [RFC4966].

In order to solve the technical difficulties in NAT-PT, the issues and the possible workarounds are:

1. NAT-PT disrupts all protocols that embed IP addresses (and/or ports) in packet payloads. There is little that can be done

about this, other than using Application Layer Gateways (ALGs) or preferring protocols that transport DNS names instead of addresses.

2. Loss of end-to-end address transparency. End-to-end address transparency implies a global address space, ability to pass packets unaltered throughout the network, and the ability to use source and destination addresses as unique labels [RFC2775]. A reversible, algorithmic mapping can restore some of this transparency. However, it is still not possible to ensure that all nodes in the existing Internet support such reversible mappings.
3. The states maintained in the translator cause scalability, multihoming and load sharing problems. Hence, a stateless translation scheme is preferred.
4. Loss of information due to incompatible semantics between IPv4 and IPv6 versions of headers and protocols. A partial remedy to this is the proper attention to the details of the protocol translation, for example the error codes mapping between ICMP and ICMPv6. However, some semantic differences remain.
5. The DNS is tightly coupled with the translator and lack of address mapping persistence discussed in Section 3.3 of [RFC4966]. Hence, the DNS should be decoupled from the translator.
6. Support for referrals is difficult in NAT-PT, given that translated addresses may leak outside the network where these addresses have a meaning. Stateless translation, algorithmic address mappings, and the decoupling of DNS from the translation process can help the handling of referrals. Nevertheless, it is still possible that an address-based referral is passed to someone who cannot employ it. For instance, an IPv6-only node may pass a referral based on an IPv6 address to a node that only understands IPv4.

1.2. CERNET Translation Requirements

China Education and Research Network has two backbones using different address families. The CERNET is IPv4-only and CERNET2 is IPv6-only [CERNET] [CNGI-CERNET2], which fits in "an IPv6 network to the IPv4 Internet" and "the IPv4 Internet to an IPv6 network" scenarios in the IETF behave Working Group definition [BEHAVE] [I-D.ietf-behave-v6v4-framework]. In order to make CERNET2 communicate with the IPv4 Internet, we designed the IVI mechanism and installed IVI translators between CERNET and CERNET2.

The requirements of the IVI mechanism are:

1. It should support both IPv6 initiated and IPv4 initiated communications for the IPv6 clients/servers in "an IPv6 network".
2. It should follow current IPv4 and IPv6 routing practice without increasing the global routing table size in both address families.
3. It should be able to be deployed incrementally.
4. It should be able to use IPv4 addresses effectively due to the IPv4 address depletion problem.
5. It should be stateless to achieve scalability.
6. The DNS function should be decoupled from the translator.

The specific IVI design presented in this document can satisfy the above requirements with following notes.

1. It restricts the IPv6 hosts to use a subset of the addresses inside the ISP's IPv6 block. Therefore, IPv6 auto-configuration cannot be used for these IPv6 hosts. Manual configuration or autoconfiguration via stateful DHCPv6 is required.
2. It defines a one-to-one mapping between IPv4 addresses and IPv6 addresses, hence the IPv4 addresses cannot be used efficiently. We suggest using the IVI6 addresses for servers instead of clients.
3. An ALG is still required for any applications which embed address(es) in the payload.
4. Some issues with end-to-end transparency, address referrals, and incompatible semantics between protocol versions still remain, as discussed above.

The IVI is an early design deployed in CERNET for the stateless translation. The IETF standard IPv4-IPv6 stateless and stateful translation mechanisms are defined in [I-D.ietf-behave-v6v4-framework], [I-D.ietf-behave-address-format], [I-D.ietf-behave-v6v4-xlate], [I-D.ietf-behave-v6v4-xlate-stateful] and [I-D.ietf-behave-dns64], etc.

2. Terms and Abbreviations

The following terms and abbreviations are used in this document:

ISP(i): A specific Internet service provider "i".

IVIG4: The global IPv4 address space.

IPS4(i): A subset of IVIG4 allocated to ISP(i).

IVI4(i): A subset of IPS4(i), the addresses in this set will be mapped to IPv6 via IVI mapping mechanism and used by IPv6 hosts of ISP(i).

IPG6: The global IPv6 address space.

IPS6(i): A subset of IPG6 allocated to ISP(i).

IVIG6(i): A subset of IPS6(i), and an image of IVIG4 in IPv6 address family via IVI mapping mechanism. It is defined as the IPv4-converted address in [I-D.ietf-behave-v6v4-framework].

IVI6(i): A subset of IVIG6(i) and an image of IVI4(i) in IPv6 address family via IVI mapping mechanism. It is defined as the IPv4-translatable address in [I-D.ietf-behave-v6v4-framework].

IVI translator: The mapping and translation gateway between IPv4 and IPv6 based on IVI mechanism.

IVI DNS: Providing IVI Domain Name Service (DNS).

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

3. The IVI Translation Algorithm

The IVI is a prefix-specific and stateless address mapping scheme which can be carried out by individual ISPs. In the IVI design, subsets of the ISP's IPv4 addresses are embedded in ISP's IPv6 addresses and the hosts using these IPv6 addresses can therefore communicate with the global IPv6 Internet directly and can communicate with the global IPv4 Internet via stateless translators, the communications can either be IPv6 initiated or IPv4 initiated.

IVI mapping and translation mechanism is implemented in an IVI translator which connects between "an IPv6 network" and the IPv4

Internet via the ISP's IPv4 network as shown in the following figure.

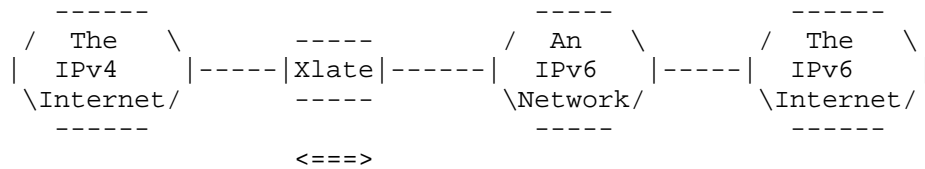


Figure 1: The scenarios: An IPv6 network to the IPv4 Internet and the IPv4 Internet to an IPv6 network

In order to perform the translation function between IPv4 and IPv6 addresses, the translator needs to represent the IPv4 addresses in IPv6 and the IPv6 addresses in IPv4.

To represent the IPv4 addresses in IPv6, a unique, prefix-specific and stateless mapping scheme is defined between IPv4 addresses and subsets of IPv6 addresses, so each provider-independent IPv6 address block (usually a /32) will have a small portion of IPv6 addresses (for example /40 defined by PREFIX), which is the image of the totality of the global IPv4 addresses, as shown in the following figure. The SUFFIX is all zeros.

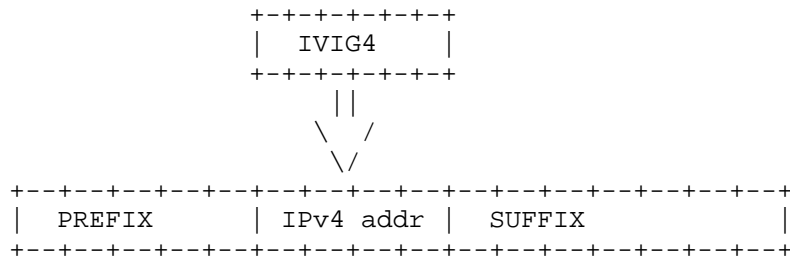


Figure 2: Representing the IPv4 addresses in IPv6

To represent the IPv6 addresses in IPv4, each provider can borrow a portion of its IPv4 addresses and map them into IPv6 based on the above mapping rule. These special IPv6 addresses will be physically used by IPv6 hosts. The original IPv4 form of the borrowed addresses is the image of these special IPv6 addresses and it can be accessed by the IPv4 Internet, as shown in the following figure. The SUFFIX can either be all zeros or some other value for future extensions.

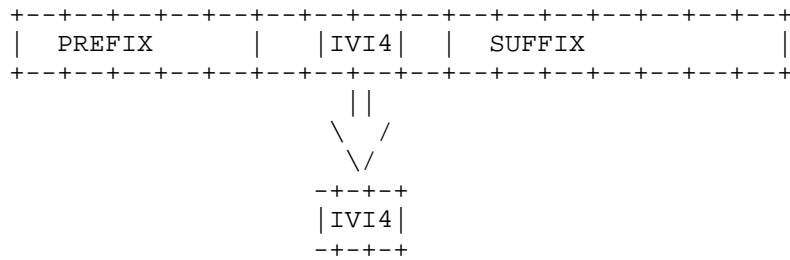


Figure 3: Representing the IPv6 addresses in IPv4

3.1. Address Format

The IVI address format is defined based on an individual ISP's IPv6 prefix as shown in the following figure.

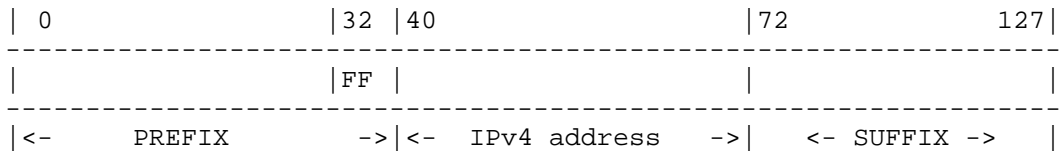


Figure 4: IVI Address Mapping

where bit 0 to bit 31 are the prefix of ISP(i)'s /32 (e.g. using document IPv6 address IPS6=2001:DB8::/32), in the CERNET implementation bit 32 to bit 39 are all one's as the identifier of the IVI addresses, bit 40 to bit 71 are embedded global IPv4 space (IVIG4) presented in hexadecimal format. (e.g. 2001:DB8:ff00::/40). Note that based on the IVI mapping mechanism, an IPv4 /24 is mapped to an IPv6 /64 and an IPv4 /32 is mapped to an IPv6 /72.

The IETF standard of the address format is defined in [I-D.ietf-behave-address-format].

3.2. Routing and Forwarding

Based on the IVI address mapping rule, routing is straightforward, as shown in the following figure.

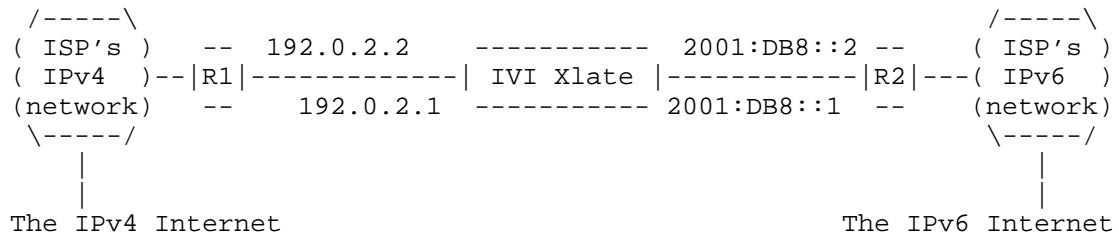


Figure 5: IVI Routing

where

1. IVI Xlate is a special dual-stack router, with two interfaces, one to the IPv4 network and the other to the IPv6 network (it is also possible to have a single interface configured with both IPv4 and IPv6 addresses). IVI Xlate can support dynamic routing protocols in IPv4 and IPv6 address families. In the above configuration, the static routing configuration can be used.
2. Router R1 has an IPv4 route for $IVI4(i)/k$ (k is the prefix length of $IVI4(i)$) with the next-hop equal to 192.0.2.1 and this route is distributed to the Internet with proper aggregation.
3. Router R2 has an IPv6 route for $IVI6(i)/40$ with the next-hop equal to 2001:DB8::1 and this route is distributed to the IPv6 Internet with proper aggregation.
4. The IVI translator has an IPv6 route for $IVI6(i)/(40+k)$ with next hop equal to 2001:DB8::2. The IVI translator also has IPv4 default route 0.0.0.0/0 with next hop equals to 192.0.2.2.

Note that the routes described above can be learned/inserted by dynamic routing protocols (IGP or BGP) in the IVI translator peering with R1 and R2.

Since both $IVI4(i)$ and $IVI6(i)$ are aggregated to $IPS4(i)$ and $IPS6(i)$ in $ISP(i)$'s border routers respectively, they will not affect the global IPv4 and IPv6 routing tables [RFC4632].

Since the IVI translation is stateless, it can support multi-homing when the same prefix is used for multiple translators.

Since the IVI translation can be implemented independently in each ISP's network, it can be incrementally deployed in the global Internet.

3.3. Network-layer Header Translation

IPv4 [RFC0791] and IPv6 [RFC2460] are different protocols with different network layer header formats; the translation of the IPv4 and IPv6 headers MUST be performed according to SIIT [RFC2765] except the source and destinations addresses in the header, as shown in the following figures.

IPv4 Field	Translated to IPv6
Version (0x4)	Version (0x6)
IHL	discarded
Type of Service	discarded
Total Length	Payload Length = Total Length - 20
Identification	discarded
Flags	discarded
Offset	discarded
Time to Live	Hop Limit
Protocol	Next Header
Header Checksum	discarded
Source Address	IVI address mapping
Destination Address	IVI address mapping
Options	discarded

Figure 6: IPv4 to IPv6 Header translation

IPv6 Field	Translated to IPv4 Header
Version (0x6)	Version (0x4)
Traffic Class	discarded
Flow Label	discarded
Payload Length	Total Length = Payload Length + 20
Next Header	Protocol
Hop Limit	TTL
Source Address	IVI address mapping
Destination Address	IVI address mapping
-	IHL = 5
-	Header Checksum recalculated

Figure 7: IPv6 to IPv4 Header translation

The IETF standard for IP/ICMP translation is defined in [I-D.ietf-behave-v6v4-xlate], which contains updated technical

specifications.

3.4. Transport-layer Header Translation

Since the TCP and UDP headers [RFC0793] [RFC0768] consist of check sums which include the IP header, the recalculation and updating of the transport-layer headers MUST be performed. Note that SIIT does not recalculate the transport-layer checksum, since checksum neutral IPv6 addresses are used in SIIT [RFC2765].

The IETF standard for Transport-layer Header Translation is defined in [I-D.ietf-behave-v6v4-xlate], which contains updated technical specifications.

3.5. Fragmentation and MTU Handling

When the packet is translated by the IVI translator, due to the different sizes of the IPv4 and IPv6 headers, the IVI6 packets will be at least 20 bytes larger than the IVI4 packets, which may exceed the MTU of the next link in the IPv6 network. Therefore, the MTU handling and translation between IPv6 fragmentation headers and fragmentation field in the IPv4 headers are necessary, which is performed in the IVI translator according to SIIT [RFC2765].

The IETF standard for Fragmentation and MTU Handling is defined in [I-D.ietf-behave-v6v4-xlate], which contains updated technical specifications.

3.6. ICMP Handling

For ICMP message translation between IPv4 and IPv6, IVI follows the ICMP/ICMPv6 message correspondence as defined in SIIT [RFC2765]. Note that the ICMP message may be generated by an intermediate router whose IPv6 address does not belong to IVIG6(i). Since ICMP translation is important to the path MTU discovery and trouble shooting, the IPv4 representation of the non-IVIG6 addresses in the ICMP packets is required. In the current IVI prototype, a small IPv4 address block is used to identify the non-IVIG6 addresses. This prevents translated ICMP messages from being discarded due to unknown or private IP source.

The IETF standard for IP/ICMP translation is defined in [I-D.ietf-behave-v6v4-xlate], which contains updated technical specifications.

3.7. Application Layer Gateway

Due to the features of 1-to-1 address mapping and stateless operation, IVI can support most of the existing applications, such as HTTP, SSH and Telnet. However, some applications are designed such that IP addresses are used to identify application-layer entities (e.g. FTP). In these cases, application layer gateway (ALG) is unavoidable, and it can be integrated into the IVI translator.

The discussion of the use of ALGs is in [I-D.ietf-behave-v6v4-framework].

4. The IVI DNS Configuration

The DNS [RFC1035] service is important for the IVI mechanism.

4.1. DNS Configuration for the IVI6(i) Addresses

For providing authoritative DNS service for IVI4(i) and IVI6(i), each host name will both have an A record and an AAAA record pointing to IVI4(i) and IVI6(i), respectively. Note that the same name always points to a unique host, which is an IVI6(i) host and it has IVI4(i) representation via the IVI translator.

4.2. DNS Service for the IVIG6(i) Addresses

For resolving the IPv6 form of the global IPv4 space (IVIG6(i)), each ISP must provide customized IVI DNS service for the IVI6(i) hosts. The IVI DNS server MUST be deployed in a dual stack environment. When the IVI6(i) host queries an AAAA record for an IPv4 only domain name, the IVI DNS will query the AAAA record first. If the AAAA record does not exist, the IVI DNS will query the A record and map it to IVIG6(i) and return an AAAA record to the IVI6(i) host. The technical specifications of this process are defined in [I-D.ietf-behave-dns64].

5. The Advanced IVI Translation Functions

5.1. IVI Multicast

The IVI mechanism can support IPv4/IPv6 communication of the protocol-independent specific-source sparse-mode multicast (PIM SSM) [RFC3171] [RFC3569] [RFC4607].

There will be 2^{24} group addresses for IPv4 SSM. The corresponding IPv6 SSM group addresses can be defined as shown in the following

figure.

IPv4 Group Address	IPv6 Group Address
232.0.0.0/8	ff3e:0:0:0:0:0:f000:0000/96
232.255.255.255/8	ff3e:0:0:0:0:0:f0ff:ffff/96

Figure 8: IVI Multicast Group Address Mapping

The source address in IPv6 MUST be IVI6(i) in order to perform reverse path forwarding (RPF) as required by PIM-SM.

The interoperation of PIM-SM for address families IPv4 and IPv6 can either be implemented via an application layer gateway or via static joins based on IGMPv3 and MLDv2 in IPv4 and IPv6, respectively.

6. IVI Host Operation

6.1. IVI Address Assignment

The IVI6 address has special format (for example IVI4=192.0.2.1/32 and IVI6=2001:db8:ffc0:2:100::/72), therefore, stateless IPv6 address auto-configuration cannot be used. However, the IVI6 can be assigned to the IPv6 end system via manual configuration or stateful auto-configuration via DHCPv6.

- o For the manual configuration, the host needs to configure the IVI6 address and the corresponding prefix length, as well as the default gateway address and the DNS resolver address.
- o For the DHCPv6 configuration, the DHCPv6 will assign the IVI6 address and the DNS resolver address to the host. The router in the subnet should enable router advertisement (RA), since the default gateway is learned from the router.

6.2. IPv6 Source Address Selection

Since each IPv6 host may have multiple addresses, it is important for the host to use an IVI6(i) address to reach the global IPv4 networks. The short-term work around is to use IVI6(i) as the default source IPv6 address of the host, defined as the policy table in [RFC3484]. The long-term solution requires that the application should be able to select the source addresses for different services.

7. The IVI Implementation

7.1. Linux Implementation

An implementation of IVI exists for the Linux operating system. The sources code can be downloaded from [LINUX]. An example of how to configure an IVI deployment is shown in Appendix A.

The IVI DNS source code for the IVID6(i) addresses presented in this document can be downloaded from [DNS].

7.2. Testing Environment

The IVI translator based on the Linux implementation has been deployed between [CERNET] (IPv4-only) and [CNGI-CERNET2] (IPv6-only) since March 2006. The pure IPv6 web servers using IVID6 addresses [2001:250:ffca:2672:100::] behind the IVI translator can be accessed by the IPv4 hosts [TEST4], and also by the global IPv6 hosts [TEST6]. The pure IPv6 clients using IVID6 addresses behind IVI translator can access IPv4 servers on the IPv4 Internet.

Two traceroute results are presented in Appendix B to show the address mapping of the IVI mechanism.

IVID6 manual configuration and DHCPv6 configuration of the IPv6 end system have also been tested with success.

8. Security Considerations

This document presents the prefix-specific and stateless address mapping mechanism (IVI) for the IPv4/IPv6 coexistence and transition. The IPv4 security and IPv6 security issues should be addressed by related documents of each address family and are not included in this document.

However, there are several issues that need special considerations, specifically (a) IPsec and its NAT traversal, (b) DNSSEC, and (c) firewall filter rules.

- o IPsec and its NAT traversal. Since the IVI scheme maintains end-to-end address transparency, IPsec could work without or with NAT traversal techniques.
- o DNSSEC verification will be terminated at the IVI DNS for the A record to AAAA record translation. It would be fine to have a translation in a local IVI DNS server that also verifies DNSSEC. Or in the host, if the host both translates the DNS entry and

again verifies DNSSEC validity. The DNSSEC discussion is in [I-D.ietf-behave-dns64].

- o Firewall filter rules. Since the IVI scheme maintains the end-to-end address transparency and there is a unique mapping between IPv4 and IPv6 addresses, therefore the firewall filter rule can be implemented for one address family or mapped to another address family and implemented in that address family. However, the current IPv6 routers may only support the access-list or uRPF (unicast Reverse Path Forwarding) for the prefix length shorter than /64, there may a practical constraint for the construction of such rules.

Except the issues discussed above, we have not found special security problems introduced by the IVI translation in our experiments.

9. IANA Considerations

This memo adds no new IANA considerations.

Note to RFC Editor: This section will have served its purpose if it correctly tells IANA that no new assignments or registries are required, or if those assignments or registries are created during the RFC publication process. From the author's perspective, it may therefore be removed upon publication as an RFC at the RFC Editor's discretion.

10. Contributors

The authors would like to acknowledge the following contributors in the different phases of the IVI development: Ang Li, Yuncheng Zhu, Junxiu Lu, Yu Zhai, Wentao Shang, Weifeng Jiang and Bizheng Fu.

The authors would like to acknowledge the following contributors who provided helpful inputs concerning the IVI concept: Bill Manning, David Ward, Elwyn Davies, Lixia Zhang, Jun Murai, Fred Baker, Jari Arkko, Ralph Droms, Tony Hain and Kevin Yin.

11. Acknowledgments

The authors thank to the funding supports of the CERNET, CNGI-CERNET2, CNGI Research and Development, China "863" and China "973" projects.

12. Appendix A. The IVI translator configuration example

IVI Configuration Example

```
#!/bin/bash
# open forwarding
echo 1 > /proc/sys/net/ipv6/conf/all/forwarding
echo 1 > /proc/sys/net/ipv4/conf/all/forwarding

# config route for IVI6 = 2001:db8:ffc0:2:0::/64,
#                      IVI4 = 192.0.2.0/24

# configure IPv6 route
route add -A inet6 2001:db8:ffc0:2:0::/64 \
gw 2001:da8:aaae::206 dev eth0

# config mapping for      source-PF = 2001:db8::/32
# config mapping for destination-PF = 2001:db8::/32

# for each mapping, a unique pseudo-address (10.0.0.x/8)
# should be configured.
# ip addr add 10.0.0.1/8 dev eth0

# IPv4-to-IPv6 mapping, multiple mappings can be done via multiple
# commands.
# mroute IVI4-network IVI4-mask pseudo-address interface \
# source-PF destination-PF
/root/mroute 192.0.2.0 255.255.255.0 10.0.0.1 \
eth0 2001:db8:: 2001:db8::

# IPv6-to-IPv4 mapping
# mroute6 destination-PF destination-PF-pref-len
/root/mroute6 2001:db8:ff00:: 40
```

Figure 9

13. Appendix B. The traceroute results

```
ivitraceroute
```

```
ivitraceroute 202.38.108.2
```

```
1  202.112.0.65 6 ms 2 ms 1 ms
2  202.112.53.73 4 ms 6 ms 12 ms
3  202.112.53.178 1 ms 1 ms 1 ms
4  202.112.61.242 1 ms 1 ms 1 ms
5  192.0.2.100 1 ms 1 ms 1 ms
6  192.0.2.102 1 ms 1 ms 1 ms
7  192.0.2.103 2 ms 2 ms 2 ms
8  192.0.2.104 2 ms 2 ms 2 ms
9  192.0.2.105 4 ms 4 ms 3 ms
10 202.38.108.2 2 ms 3 ms 3 ms
```

Figure 10

Note that the non-IVIG6 addresses are mapped to IPv4 document address 192.0.2.0/24.

```
ivitraceroute6
```

```
ivitraceroute6 www.mit.edu
```

```
src_ivi4=202.38.97.205 src_ivi6=2001:da8:ffca:2661:cd00::
dst_host=www.mit.edu
dst_ip4=18.7.22.83 dst_ivi6=2001:da8:ff12:716:5300::
```

```
traceroute to 2001:da8:ff12:716:5300:: (2001:da8:ff12:716:5300::),
30 hops max, 40 byte packets to not_ivi
```

```

1  2001:da8:ff0a:0:100::      0.304 ms 0.262 ms 0.190 ms
   10.0.0.1
2  2001:da8:ffca:7023:fe00::  0.589 ms * *
   202.112.35.254
3  2001:da8:ffca:7035:4900::  1.660 ms 1.538 ms 1.905 ms
   202.112.53.73
4  2001:da8:ffca:703d:9e00::  0.371 ms 0.530 ms 0.459 ms
   202.112.61.158
5  2001:da8:ffca:7035:1200::  0.776 ms 0.704 ms 0.690 ms
   202.112.53.18
6  2001:da8:ffcb:b5c2:7d00::  89.382 ms 89.076 ms 89.240 ms
   203.181.194.125
7  2001:da8:ffc0:cb74:9100::  204.623 ms 204.685 ms 204.494 ms
   192.203.116.145
8  2001:da8:ffcf:e7f0:8300::  249.842 ms 249.945 ms 250.329 ms
   207.231.240.131
9  2001:da8:ff40:391c:2d00::  249.891 ms 249.936 ms 250.090 ms
   64.57.28.45
10 2001:da8:ff40:391c:2a00::  259.030 ms 259.110 ms 259.086 ms
   64.57.28.42
11 2001:da8:ff40:391c:700::   264.247 ms 264.399 ms 264.364 ms
   64.57.28.7
12 2001:da8:ff40:391c:a00::   271.014 ms 269.572 ms 269.692 ms
   64.57.28.10
13 2001:da8:ffc0:559:dd00::   274.300 ms 274.483 ms 274.316 ms
   192.5.89.221
14 2001:da8:ffc0:559:ed00::   274.534 ms 274.367 ms 274.517 ms
   192.5.89.237
15 * * *
16 2001:da8:ff12:a800:1900::  276.032 ms 275.876 ms 276.090 ms
   18.168.0.25
17 2001:da8:ff12:716:5300::  276.285 ms 276.370 ms 276.214 ms
   18.7.22.83
```

Figure 11

Note that all of the IPv4 addresses can be mapped to prefix-specific IPv6 addresses (for example 18.7.22.83 is mapped to 2001:da8:ff12:716:5300::).

14. References

14.1. Normative References

- [I-D.ietf-behave-address-format]
Huitema, C., Bao, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", draft-ietf-behave-address-format-03 (work in progress), December 2009.
- [I-D.ietf-behave-dns64]
Bagnulo, M., Sullivan, A., Matthews, P., and I. Beijnum, "DNS64: DNS extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", draft-ietf-behave-dns64-05 (work in progress), December 2009.
- [I-D.ietf-behave-v6v4-framework]
Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", draft-ietf-behave-v6v4-framework-04 (work in progress), December 2009.
- [I-D.ietf-behave-v6v4-xlate]
Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", draft-ietf-behave-v6v4-xlate-05 (work in progress), December 2009.
- [I-D.ietf-behave-v6v4-xlate-stateful]
Bagnulo, M., Matthews, P., and I. Beijnum, "NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", draft-ietf-behave-v6v4-xlate-stateful-07 (work in progress), December 2009.
- [RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.
- [RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.
- [RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

- [RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.
- [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.
- [RFC2765] Nordmark, E., "Stateless IP/ICMP Translation Algorithm (SIIT)", RFC 2765, February 2000.
- [RFC2766] Tsirtsis, G. and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)", RFC 2766, February 2000.
- [RFC3056] Carpenter, B. and K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.
- [RFC3171] Albanna, Z., Almeroth, K., Meyer, D., and M. Schipper, "IANA Guidelines for IPv4 Multicast Address Assignments", BCP 51, RFC 3171, August 2001.
- [RFC4213] Nordmark, E. and R. Gilligan, "Basic Transition Mechanisms for IPv6 Hosts and Routers", RFC 4213, October 2005.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.
- [RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, February 2006.
- [RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, August 2006.
- [RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", BCP 122, RFC 4632, August 2006.
- [RFC5214] Templin, F., Gleeson, T., and D. Thaler, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)", RFC 5214, March 2008.

14.2. Informative References

- [BEHAVE] "The IETF Behave Working Group Charter:
<http://www.ietf.org/html.charters/behave-charter.html/>".

- [CERNET] "CERNET Homepage:
http://www.edu.cn/english_1369/index.shtml".
- [CNGI-CERNET2]
"CNGI-CERNET2 Homepage:
http://www.cernet2.edu.cn/index_en.htm".
- [COUNT] "IPv4 address count down: <http://penrose.uk6x.com/>".
- [DNS] "Source Code of the IVI DNS
<http://www.ivi2.org/IVI/src/ivids-0.1.tar.gz>".
- [LINUX] "Source Code of the IVI implementation for Linux:
<http://linux.ivi2.org/impl/>".
- [RFC2775] Carpenter, B., "Internet Transparency", RFC 2775,
February 2000.
- [RFC3142] Hagino, J. and K. Yamamoto, "An IPv6-to-IPv4 Transport
Relay Translator", RFC 3142, June 2001.
- [RFC3484] Draves, R., "Default Address Selection for Internet
Protocol version 6 (IPv6)", RFC 3484, February 2003.
- [RFC3569] Bhattacharyya, S., "An Overview of Source-Specific
Multicast (SSM)", RFC 3569, July 2003.
- [RFC4966] Aoun, C. and E. Davies, "Reasons to Move the Network
Address Translator - Protocol Translator (NAT-PT) to
Historic Status", RFC 4966, July 2007.
- [TEST4] "Test homepage for the IVI4(i): <http://202.38.114.1/>".
- [TEST6] "Test homepage for the IVI6(i):
[http://\[2001:250:ffca:2672:0100::0\]/](http://[2001:250:ffca:2672:0100::0]/)".

Authors' Addresses

Xing Li
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 62785983
Email: xing@cernet.edu.cn

Congxiao Bao
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 62785983
Email: congxiao@cernet.edu.cn

Maoke Chen
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 62785983
Email: mk@cernet.edu.cn

Hong Zhang
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 62785983
Email: neilzh@gmail.com

Jianping Wu
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University
Beijing 100084
CN

Phone: +86 62785983
Email: jianping@cernet.edu.cn

