

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: December 23, 2008

X. Li  
M. Chen  
C. Bao  
H. Zhang  
J. Wu  
CERNET Center/Tsinghua University  
June 21, 2008

Bi-directional Address Borrowing (IVI)  
draft-xli-ivi-00.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of 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 December 23, 2008.

## Abstract

This document presents the concept and practice of the bi-directional address borrowing scheme (IVI) for smooth and transparent transition from IPv4 to IPv6. Based on this scheme, special IPv6 address blocks are defined, which can communicate with the global IPv6 networks directly and with the global IPv4 networks via stateless (or almost stateless) gateways. The latter kind of communication also has the features of end-to-end address transparency, being globally routable, incremental deployable, having multi-homing support and being able to use the global IPv4 addresses effectively. The design considerations of the IVI scheme (the address mapping, the header processing, etc) are discussed in this document. The preliminary testing results of the IVI deployment on large scale, public IPv4 and IPv6 networks are presented as well. In addition, a smooth and transparent IPv4 to IPv6 transition scheme and the corresponding address allocation and assignment policies are discussed.

## Table of Contents

|                                                                 |    |
|-----------------------------------------------------------------|----|
| 1. Introduction . . . . .                                       | 4  |
| 2. Terms and Abbreviations . . . . .                            | 5  |
| 3. IVI Concept . . . . .                                        | 6  |
| 3.1. Address Mapping . . . . .                                  | 6  |
| 3.2. Routing and Forwarding . . . . .                           | 7  |
| 3.3. IVI Communication Scenarios . . . . .                      | 8  |
| 4. Design Considerations . . . . .                              | 10 |
| 4.1. Address Mapping . . . . .                                  | 10 |
| 4.2. Network-layer Header Translation . . . . .                 | 10 |
| 4.3. Transport-layer Header Translation . . . . .               | 11 |
| 4.4. Fragmentation and MTU Handling . . . . .                   | 11 |
| 4.5. ICMP Handling . . . . .                                    | 12 |
| 4.6. Application Layer Gateway . . . . .                        | 12 |
| 4.7. IPv6 Source Address Selection . . . . .                    | 12 |
| 4.8. IPv4 over IPv6 Support . . . . .                           | 12 |
| 5. DNS Configuration and Mapping . . . . .                      | 13 |
| 5.1. DNS Configuration for the IVI6(i) Block . . . . .          | 13 |
| 5.2. DNS Mapping for the IVIG46(i) Block . . . . .              | 13 |
| 6. Multiplexing of the Global IPv4 Addresses . . . . .          | 14 |
| 6.1. Temporal Multiplexing . . . . .                            | 14 |
| 6.2. Port Multiplexing . . . . .                                | 14 |
| 6.3. Spatial Multiplexing . . . . .                             | 14 |
| 7. IVI Multicast Support . . . . .                              | 16 |
| 8. IVI Implementation and Preliminary Testing Results . . . . . | 17 |
| 9. Features of IVI . . . . .                                    | 18 |
| 10. Address Policy and IVI Address Evolution . . . . .          | 20 |
| 10.1. IPv6 Address allocation and Assignment Policy . . . . .   | 20 |
| 10.2. IPv4 Address allocation and assignment Policy . . . . .   | 20 |
| 10.3. Evolution of the IVI Addresses . . . . .                  | 20 |
| 11. Security Considerations . . . . .                           | 22 |
| 12. IANA Considerations . . . . .                               | 23 |
| 13. Principal Authors . . . . .                                 | 24 |
| 14. Contributors . . . . .                                      | 25 |
| 15. Acknowledgments . . . . .                                   | 26 |
| 16. Appendix A. The IVI gateway configuration example . . . . . | 27 |
| 17. Appendix B. The traceroute results . . . . .                | 28 |
| 18. References . . . . .                                        | 31 |
| 18.1. Normative References . . . . .                            | 31 |
| 18.2. Informative References . . . . .                          | 33 |
| Authors' Addresses . . . . .                                    | 34 |
| Intellectual Property and Copyright Statements . . . . .        | 36 |

## 1. Introduction

This document presents the concept and practice of the bi-directional address borrowing scheme (IVI) for smooth and transparent transition from IPv4 to IPv6.

The experiences for the IPv6 deployment in the past 10 years strongly indicate that for a successful transition, the IPv6 hosts need to communicate with the global IPv4 networks [JJI07]. 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 the IPv4 address depletion problem makes dual stack approach not practical [COUNT]. The tunneled architectures can link the IPv6 islands cross IPv4 networks, but it does not help the communication between two address families [RFC3056] [RFC4214] [RFC4380]. The translation architecture can relay the communications for the hosts located in IPv4 and IPv6 networks, but it is not scalable and the end-to-end address transparency is lost [RFC2766] [RFC4966].

Since IPv4 and IPv6 are different protocols with different addressing structure, it is impossible to make seamless communication between IPv4 and IPv6 with arbitrary addresses without restrictions. For example, the Network Address Translation (NAT) approach lost the scalability and end-to-end address transparency [RFC3142] [RFC4966] [RFC2775]. The mapping methods (IPv4 mapped IPv6 address [::FF:0:0/96] and IPv4 compatible IPv6 address [::ipv4-address/96]) are not globally routable [RFC4291].

In this document, we allocate special IPv6 address blocks based on bi-directional borrowing scheme (IVI), the hosts which are using special IPv6 address block can communicate with the global IPv6 network directly and the global IPv4 network via stateless (or almost stateless) gateways. The latter kind of communication also has the features of end-to-end address transparency, being globally routable, incremental deployable, having multi-homing support and being able to use the global IPv4 addresses effectively.

## 2. Terms and Abbreviations

The following terms and abbreviations are used in this document:

IVI: IV means 4 and VI means 6 in Roman representation, so IVI means bi-directional borrowing, mapping and translation between IPv4 and IPv6.

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

IPG4: A set containing all IPv4 addresses, being used by pure IPv4 hosts.

IPS4(i): A subset of IPG4 allocated to ISP(i).

IVI4(i): A subset of IPS4(i), being physically borrowed by pure IPv6 hosts of ISP(i).

IPG6: A set containing all IPv6 addresses.

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

IVIG46(i): A subset of IPS6(i), being virtually borrowed by global IPv4 hosts.

IVI6(i): A subset of IVIG46(i), being used by pure IPv6 hosts. It is the IVI4(i) equivalent in IPv6.

IVI gateway: The mapping and forwarding gateway between IPv4 and IPv6 based on IVI scheme.

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

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. IVI Concept

The concept of IVI is a bi-directional borrowing scheme which can be carried out by individual ISPs.

(1) The ISP borrows a subset of ISP's global routable IPv4 address, maps it into IPv6 and physically uses it for ISP's pure IPv6 hosts;

(2) The ISP borrows a subset of ISP's global routable IPv6 address, maps it into IPv4 and virtually uses it for global IPv4 hosts.

#### 3.1. Address Mapping

The IVI address mapping rule is defined in the basis of individual ISP as shown in the following figure.

IVI Address Mapping

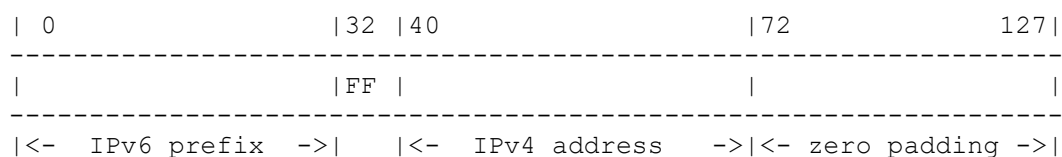


Figure 1

where bit 0 to bit 31 are the prefix of ISP(i)'s /32 (e.g. IPS6=2001:DB8::/32), bit 32 to bit 39 are all one's as the identifier of IVI, bit 40 to bit 71 are the global IPv4 space (IPG4) presented in hexadecimal format. (i.e. IVIG46(i)=2001:DB8:ff00::/40). Because this mapping is 1-to-1 defined by the IVI mapping rule, it is stateless and it has feature of end-to-end address transparency.

(1) The ISP(i) borrows a subset of ISP4(i) defined as IVI4(i), maps it into IPv6 as IVI6(i). Note that the IVI4(i) cannot be used by IPv4 hosts, while IVI6(i) is being physically used by pure IPv6 hosts inside ISP(i)'s IPv6 network. Therefore, IVI6(i) is the special IPv6 address block which can communicates with both address families.

(2) Based on the above mapping rule, the ISP(i) borrows a subset of ISP6(i), defined as IVIG46(i), maps it into IPv4 as IPG4. Note that the IVIG46(i) is being virtually used by global IPv4 hosts and it cannot be used by IPv6 hosts, except the portion of IVI6(i).

The relationships of different address sets are,

For IPv4 address family:  $IVI4(i) < IPS4(i) < IPG4$ .

For IPv6 address family:  $IVI6(i) < IVIG46(i) < IPS6(i) < IPG6$ .

where  $A < B$  means A is a subset of B.

$IVI6(i)$  is physically used by IPv6 hosts.  $IVI4(i)$  is a mirror image of  $IVI6(i)$ .

$IPG4$  is physically used by IPv4 hosts except the portion of  $IVI4(i)$ .  $IVIG46(i)$  is a mirror images of  $IPG4$ .

$IPG4$  has multiple mirror images for different ISP, i.e.  $IVIG46(i) \neq IVIG46(j)$  for  $i \neq j$ .

$IVI4(i)$  and  $IVI4(j)$  are not overlapped for  $i \neq j$ .

### 3.2. Routing and Forwarding

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

IVI Routing

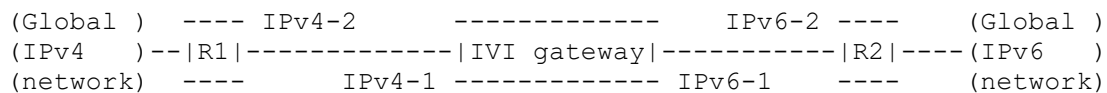


Figure 2

where

(1) Router R1 has IPv4 route of  $IVI4(i)/k$  ( $k$  is the prefix length of  $IVI4(i)$ ) with next-hop equals to IPv4-1 and this route is distributed to global IPv4 networks with proper aggregation.

(2) Router R2 has IPv6 route of  $IVIG46(i)/40$  with next-hop equals to IPv6-1 and this route is distributed to global IPv6 networks with proper aggregation.

(3) IVI gateway has IPv6 route of  $IVI6(i)/(40+k)$  with next hop equals to IPv6-2. IVI gateway also has IPv4 default route  $0.0.0.0/0$  with next hop equals to IPv4-2.

Note that the routes described above can be learned/inserted by dynamic routing protocols in the IVI gateway neighboring (IGP) or

peering (BGP) with R1 and R2.

The reachability matrix of the IPv4, IVI and IPv6 is shown in the following figure.

IVI reachability Matrix

|      |       | IPG4 | IVI | IPG6 |
|------|-------|------|-----|------|
|      | ----- |      |     |      |
| IPG4 |       | OK   | OK  | NO   |
| IVI  |       | OK   | OK  | OK   |
| IPG6 |       | NO   | OK  | OK   |

Figure 3

Since both IVI4(i) and IVI6(i) are aggregated to IPS4(i) and IPS6(i) in ISP(i)'s border routers respectively, there will be no affect to the global IPv4 and IPv6 routing tables [RFC1519].

If IVI4(i) and IVI6(i) has 1-to-1 mapping relationship, then IVI is stateless and it can support multi-homing.

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

### 3.3. IVI Communication Scenarios

Assume that there are IPv4 addresses A and IVI mapped IPv6 address A', IPv4 address B and IVI mapped IPv6 address B', IPv4 address D and IVI mapped IPv6 address D' as well as an arbitrary IPv6 addresses C'.

Scenario 1: If ISP(1) deployed IVI, then a pure IPv6 host A' can communicate with B, where B is a host with arbitrary IPv4 address. Note that A' is indeed communicating with B', a mirror image of B, and B is indeed communicates with A, a mirror image of A'. Since A' is an IPv6 address inside ISP6(1), it can also communicate with IPv6 address C'.

Scenario 2: If both ISP(1) and ISP(2) deployed IVI but ISP(1) and ISP(2) do not know the IVI deployment on the other end, then A' can still communicate with D. However, in this case, D is a mirror image of D', so the routing is not optimal (i.e. via two IVI gateways).

Scenario 3: If ISP(1) and ISP(2) deployed IVI and ISP(1) and ISP(2) know the IVI deployment on the other end, then A' can communicate

with D' directly and the routing is optimal (i.e. via none IVI gateways).

#### 4. Design Considerations

The components of IVI scheme includes: address mapping, network-layer header translation, transport-layer header translation, fragmentation/MTU handling, ICMP handling, application layer gateway, IPv6 source address selection and IPv4 over IPv6 support.

##### 4.1. Address Mapping

The address mapping rule is defined in Section 3.1.

In addition, depending on the implementation scope of the IVI gateway, IVIG46(i) block can also be defined as 2001:DB8:FFFF::/48, 2001:DB8:ABCD:FF00::/56 or 2001:DB8:ABCD:FFFF::/64.

##### 4.2. Network-layer Header Translation

IPv4 [RFC791] [RFC791] and IPv6 [RFC2460] are different protocols with different network layer header format, the translation of the IPv4 and IPv6 headers must be performed [MVB98] as shown in the following figures.

IVI IPv4 to IPv6 Header translation

| IPv4 Field        | Translated to IPv6                      |
|-------------------|-----------------------------------------|
| Version (0x4)     | Version (0x6)                           |
| IHL               | discarded                               |
| Type of Service   | discarded                               |
| Total Length      | Payload Length = Total Length - IHL * 4 |
| Identification    | discarded, cf. Subsection V-C           |
| Flags             | same as above                           |
| Offset            | same as above                           |
| Time to Live      | Hop Limit                               |
| Protocol          | Next Header                             |
| Header Checksum   | discarded                               |
| Source Address    | IVI address mapping                     |
| Destination Addr. | same as above                           |
| Options           | discarded                               |

Figure 4

## IVI IPv6 to IPv4 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 Addr. | same as above                      |
| -                 | IHL = 5                            |
| -                 | Header Checksum recalculated       |

Figure 5

## 4.3. Transport-layer Header Translation

Since the TCP and UDP headers [RFC793] [RFC768] consists of check sums which include the IP header, the recalculation and updating of the transport-layer headers must be performed.

## 4.4. Fragmentation and MTU Handling

When the packet is translated by the IVI gateway, due to the different sizes of IPv4 and IPv6 headers, the IVI6 packets will be at least 20 bytes larger than IVI4 packets, which may exceed the MTU of next link in the IPv6 network [CJ88]. However, the problem is not severe in reality. For TCP-based communications, the maximum segment size (MSS) is negotiated when every connection is initialized [RFC1191] [RFC1981]. For UDP-based applications, the applications restrict their messages size small enough to avoid fragmentation. For instance, DNS protocol requires truncation when message is larger than 512 bytes [RFC1035]; RTP uses self-defined format instead of relying on IP fragmentation [RFC2736]

Although the percentage of the fragmented packets is very small, to follow the standards, the IVI gateway performs the mapping between the fragmentation field in the IPv4 header and the IPv6 fragmentation header. [RFC2460]

#### 4.5. ICMP Handling

For ICMP message translation between IPv4 and IPv6, IVI follows the ICMP/ICMPv6 message correspondence defined in SIIT [RFC2765]. Note that ICMP message may be generated by an intermediate router which is not belonging to IVIG46(i). Since ICMP translation is important to path MTU discovery, the inverse mapping for unmapped addresses is defined in this document. In the current prototype, a pseudo IPv4 address is generated such that the first 16 bits are the IPv4 address of the IVI gateway, and the last 16 bits are the AS number of current domain. This prevents translated ICMP messages from being discarded due to unknown or private IP source.

#### 4.6. Application Layer Gateway

Due to the 1-to-1 address mapping and stateless features, IVI can support most of the existing applications.

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, but it can be integrated into IVI gateway.

A list of applications which support IVI scheme without ALG will be compiled.

#### 4.7. IPv6 Source Address Selection

Since IPv6 host may have multiple addresses, it is important for the host using IVI6(i) address to reach the global IPv4 networks. The short term work around is to use IVI6(i) as the default IPv6 address of the host. The long term solution is that the application should be able to select the source address for different services.

#### 4.8. IPv4 over IPv6 Support

The IVI scheme can support the IPv4 over IPv6 service, i.e. a stub IPv4 network can connected to an IVI gateway to reach IPv6 network and via another IVI gateway to reach the global IPv4 network [RFC4925]

A more interesting scenario is to integrate the functions of first IVI gateway into the end-host. In this case, the application software is IPv4 based and there is no need to have ALG support in the IVI gateway.

## 5. DNS Configuration and Mapping

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

### 5.1. DNS Configuration for the IVI6(i) Block

For providing authoritative DNS service for IVI4(i) and IVI6(i); each host name will have both 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 IVI gateway.

### 5.2. DNS Mapping for the IVIG46(i) Block

For resolving the IVI IPv6 mapped global IPv4 space (IVIG46(i)), each ISP must provide customized IVI DNS service for the IVI6(i) hosts. The IVI DNS server is in dual stack environment, when the IVI6(i) host queries an AAAA record for an IPv4 only domain name, the IVI DNS server will query the A record and mapped it to IVIG46(i) with ISP's IPv6 prefix and returns an AAAA record to the IVI6(i) host.

## 6. Multiplexing of the Global IPv4 Addresses

Since public-IPv4 address is a scarce resource, the effective use of the IPv4 address is important for IVI scheme. The multiplexing techniques are temporal multiplexing, port multiplexing and spatial multiplexing.

### 6.1. Temporal Multiplexing

The IVI6 can be temporally multiplexed inside the ISP(i)'s /32. This is to say that the ISP can dynamically assign IVI6(i) to an end system when it requests the IPv4 communication service. The corresponding fee can be charged, therefore, results in a more sustainable business model. For temporal multiplexing, the features of stateless in IVI gateway and end-to-end address transparency are maintained.

### 6.2. Port Multiplexing

To future increase the utilization ratio of the public IPv4 address; the port multiplexing inside the ISP(i)'s /32 can be deployed [RFC2766]. This is to say that a single borrowed IPv4 address (IVI4(i)) can be used for multiple IVI6(i) addresses. The mapping scheme is to use the least significant bits in the IVI6(i) to define the multiple mapping and combine the transport-layer port number to uniquely perform the mapping from IVI4(i) to IVI6(i). Based on this method, the mapping gain can be adjusted incrementally depending on the requirements. For example, zero bit means 1-to-1 mapping, and it is stateless. One bit means 1-to-1 mapping, and it has two states. Four bits means 1-to-16 mapping, etc. In the case of one-to-many mapping, when two IVI6(i) have the same port number, the IVI gateway will map one of the port number to another unused port number and maintain the mapping table (IVI4(i) plus port number). Since the one-to-many mapping lost the feature of stateless and may lost the end-to-end address transparency, the proper use of the one-to-many mapping is the balancing of tradeoffs [RFC4966].

The tradeoffs are: (1) The number of the port number ( $2^{16}$ ); (2) The gain of the IPv4 address utilization; (3) Half association or full association; (4) The number of states; (5) The average concurrent port used in an IPv6 host and; (6) The collision ratio of the port number.

### 6.3. Spatial Multiplexing

The spatial multiplexing means that for different operation modes of server and client, the different port multiplexing ratios can be applied. There are basically three cases.

(1) Server: we suggest have 1-to-1 mapping between IVI4(i) and IVI6(i). Because it provides the features of end-to-end address transparency, stateless, multi-homing support and it provides services via well-known ports.

(2) Client with self-initiated connection: we suggest have 1-to-N mapping between IVI4(i) and IVI6(i) (N is a positive integer greater than 1), i.e. one IVI4(i) can support several IVI6(i) users to access the IPv4 network. By adjusting N, The number of state can be controlled. In this case, the port number is randomly generated by the client operating system. The IVI gateway maintains the port mapping table to avoid collision. There is no need to modify the client operating system and/or client application.

(3) Client with peer initiated connection: we suggest have 1-to-M mapping between IVI4(i) and IVI6(i) (M is a positive integer greater than 1 and may be smaller than N), i.e. one IVI4(i) can support several IVI6(i) users as the peer-to-peer hosts for the IPv4 network. By adjusting M, The number of state can be controlled. In this case, we can define "pseudo-well-known port number", which is unique for IVI4(i) and known to the peers. However, modification for the client operating system and/or client application may be needed. By combining address and pseudo-well-known port number, the feature of end-to-end address transparency can still be maintained.

## 7. IVI Multicast Support

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

(1) IVI group address mapping rule: There will be  $2^{24}$  group addresses for IPv4 SSM. The corresponding IPv6 SSM group address can be defined as shown in the following figure.

### IVI Multicast Group Address Mapping

| 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 6

(2) IVI multicast source address selection: The source address in IPv6 has to be IVI6(i) to perform reverse path forwarding (RPF) required by PIM-SM.

(3) The protocol inter operation: The inter operation of PIM-SM for address families IPv4 and IPv6 can either be implemented via application layer gateway or via static join based on IGMPv3 and MLDv2 in IPv4 and IPv6, respectively.

The Any Source Multicast (ASM) cannot be supported in the cross domain environment, since IPv6 does not support MSDP [RFC4611], while IPv4 does not support embedded RP [RFC3956].

## 8. IVI Implementation and Preliminary Testing Results

The IVI scheme presented in this document is implemented in the Linux OS and the source code can be downloaded [LINUX]. The example of the configuration is shown in Appendix A.

The IVI gateway based on the Linux implementation has been deployed between CERNET (IPv4 and partially dual-stack) [CERNET]. and CNGI-CERNET2 (pure IPv6) [CERNET2]. since March 2006. The pure IPv6 web servers using IVI addresses behind CERNET's IVI gateway can be access by the global IPv4 [IVI4] and the global IPv6 [IVI6].

In addition, two traceroute results are presented in Appendix B to shown the address mapping of the IVI scheme.

Because CNGI-CERNET2 is a pure IPv6 network, the secure shell access (ssh) to the pure IPv6 hosts in CNGI-CERNET from the IPv4 Internet has been heavily rely on IVI scheme since March 2006.

## 9. Features of IVI

The basic features of IVI scheme are:

- (1) Special IPv6 addresses can communicate with global IPv6 network directly and can communicate with global IPv4 network via IVI gateways.
- (2) When the mapping is 1-to-1, the IVI gateway is stateless and support multi-homing. When mapping is 1-to-N ( $N \neq 1$ ), the IVI gateway is stateful, but the number of state can be controlled.
- (3) When the mapping is 1-to-1, IVI scheme has the features of end-to-end address transparency. When mapping is 1-to-M, by introducing pseudo-well-known ports, the feature of end-to-end address transparency can also be maintained.
- (4) The IVI addresses are globally routable.
- (5) The IVI scheme is incremental deployable.
- (6) Based on the multiplexing techniques, the global IPv4 addresses can be effectively used.

The IVI scheme can satisfy most of the basic and advanced requirements for the IPv4 to IPv6 transition specified by Internet Drafts [I-D.nat64].

For the basic requirements (MUST):

- (1) No need to change the end system (IPv4 and IPv6).
- (2) Support v4-initiated and v6-initiated short-lived local handle.
- (3) Support interaction with dual-stack hosts.
- (4) The standard IPv4 NAT can easily be integrated into the system.
- (5) Do not violate standard DNS semantics.
- (6) No affect to IPv6 routing.
- (7) Support TCP, UDP, ICMP.
- (8) No strange behavior
- (9) Can handle fragmentation.

(10) Do not introduce new network security vulnerabilities.

For the advanced requirements (SHOULD):

- (1) Support multicast (SSM).
- (2) Support operational flexibility.
- (3) Support central Management.
- (5) Support richer application behavior.

Other requirement specified by IETF RFC or IETF drafts will be studied.

## 10. Address Policy and IVI Address Evolution

Based on IVI scheme, a modified address allocation and assignment policy can be deployed for smooth and transparent IPv4 to IPv6 transition [RFC1744] [RFC2008].

### 10.1. IPv6 Address allocation and Assignment Policy

(1) Reserve 2001:DB8:ff00::/40 for each 2001:DB8::/32 (2001:DB8::/32 is the documentation address, which represents all /32s [RFC4291]).

(2) Encourage ISPs to deploy their IPv6 networks and to install their IVI gateways.

(3) Encourage ISPs to borrowing a subset (i.e.  $IVI4(i)$ ) of their own IPv4 address blocks and mapped into IPv6 via IVI scheme (i.e.  $IVI6(i)$ ) for their initial deployment of IPv6. For servers use 1-to-1 mapping, and for clients use 1-to-N mapping, therefore the scarce IPv4 addresses can be effectively used. This special IPv6 block can communicate with the global IPv6 networks directly and communicate with the global IPv4 networks via IVI gateway.

(4) Encourage ISPs to increase the size of  $IVI4(i)$ , when  $IVI4(i) = IPS4(i)$ , the IPv4 to IPv6 transition for  $ISP(i)$  will be accomplished.

### 10.2. IPv4 Address allocation and assignment Policy

(1) The remaining IPv4 address should be dedicated for the IVI transition use, i.e. using these blocks for the  $IVI4(i)$  deployment. The users use  $IVI6(i)$  can access the IPv6 networks directly and IPv4 networks via IVI gateway.

(2) Based on multiplexing techniques, the global IPv4 addresses can be used effectively. For example, with a reasonable port multiplexing ratio (say 16), one /8 can support 268M hosts. If 10 /8s can be allocated for the IVI use, it will be 2.6 billion addresses, possibly enough for the unwired population in the world. The 43.0.0.0/8 could be a good candidate for the trail [APNIC].

### 10.3. Evolution of the IVI Addresses

The IVI scheme can help to build a smooth and transparent IPv4 to IPv6 transition. The users who uses the  $IVI6(i)$  can communicate with the global IPv4 networks via IVI gateway and can communicate with IPv6 networks directly. In addition, they may enjoy new services provided by IPv6 networks. Therefore, IPv6 becomes an attractive service. When  $IVI4(i) = IPS4(i)$  for most of  $ISP(i)$ , the rest of the IPv6 addresses (non  $IVI6(i)$ ) can be used to the further development

of the global Internet.

## 11. Security Considerations

This document presents the bi-directional address borrowing scheme (IVI) for the IPv4 to IPv6 transition. The IPv4 security and IPv6 security issues should be addressed by related documents of each address family.

However, the specific security issues for the IVI gateway implementation should be studied and addressed during the development of the specific mechanisms.

## 12. IANA Considerations

The address allocation and assignment policies discussed in this document may have impact to IANA operation.

### 13. Principal Authors

#### 14. Contributors

The authors would like to acknowledge the following contributors in the different phases of the IVI development: Ang Li, Yuncheng Zhu and Yu Zhai.

The authors would like to acknowledge the following contributors who provided helpful inputs concerning the IVI concept: Bill Manning, David Ward, Lixia Zhang and Jun Murai.

## 15. Acknowledgments

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

## 16. Appendix A. The IVI gateway 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:da8:ffca:2661:cc00::/70,
# IVI4 = 202.38.97.204/30

# configure IPv6 route
route add -A inet6 2001:da8:ffca:2661:cc00::/70 \
gw 2001:da8:aaae::206 dev eth0

# config mapping for      source-PF = 2001:da8::/32
# config mapping for destination-PF = 2001:da8::/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 202.38.97.204 255.255.255.252 10.0.0.1 \
eth0 2001:da8:: 2001:da8::

# IPv6-to-IPv4 mapping
# mroute6 destination-PF destination-PF-pref-len
/root/mroute6 2001:da8:ff00:: 40
```

Figure 7

## 17. Appendix B. The traceroute results

```
ivitracroute
```

```
ivitracroute 202.38.108.2
```

```
1  202.112.0.65 6 ms 2 ms 1 ms
   202.112 AS65

2  202.112.53.73 4 ms 6 ms 12 ms
   202.112 AS13641

3  202.112.53.178 1 ms 1 ms 1 ms
   202.112 AS13746

4  202.112.61.242 1 ms 1 ms 1 ms
   202.112 AS15858

5  202.38.17.186 1 ms 1 ms 1 ms
   202.38 AS4538

6  202.38.17.186 1 ms 1 ms 1 ms
   202.38 AS4538

7  202.38.17.186 2 ms 2 ms 2 ms
   202.38 AS4538

8  202.38.17.186 2 ms 2 ms 2 ms
   202.38 AS4538

9  202.38.17.186 4 ms 4 ms 3 ms
   202.38 AS4538

10 202.38.108.2 2 ms 3 ms 3 ms
    202.38 AS27650
```

Figure 8

```
ivitracroute6
```

```
ivitracroute6 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_ivig=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\_ivl

```
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  i
   64.57.28.7

12 2001:da8:ff40:391c:a00::   271.014 ms 269.572 ms 269.692 ms  i
   64.57.28.10

13 2001:da8:ffc0:559:dd00::   274.300 ms 274.483 ms 274.316 ms  i
   192.5.89.221

14 2001:da8:ffc0:559:ed00::   274.534 ms 274.367 ms 274.517 ms  i
   192.5.89.237

15 * * *

16 2001:da8:ff12:a800:1900::  276.032 ms 275.876 ms 276.090 ms  i
```

18.168.0.25

17 2001:da8:ff12:716:5300:: 276.285 ms 276.370 ms 276.214 ms i  
18.7.22.83

Figure 9

## 18. References

### 18.1. Normative References

- [I-D.nat64]        Bagnulo, M., Baker, F., and I. van Beijnum, "IPv4/IPv6 Coexistence and Transition: Requirements for solutions", May 2008.
- [IVI4]            "Test homepage for the IVI4(i): <http://202.38.114.1/>".
- [IVI6]            "Test homepage for the IVI6(i): [http://\[2001:250:ffca:2672:0100::0\]/](http://[2001:250:ffca:2672:0100::0]/)".
- [LINUX]           "Source Code of the IVI implementation for Linux: <http://linux.ivi2.org/impl/>".
- [MVB98]           Fiuczynski, M., Lam, V., and B. Bershad, "The design and implementation of an ipv6/ipv4 network address and protocol translator", Proceedings of the USENIX Annual Technical Conference (NO 98), June 1998.
- [RFC1035]        Mockapetris, P., "Domain names - implementation and specification", RFC 1035, November 1987.
- [RFC1191]        Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.
- [RFC1519]        Fuller, V., Li, T., Yu, J., and K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", RFC 1519, September 1993.
- [RFC1744]        Huston, G., "Observations on the Management of the Internet Address Space", RFC 1744, December 1994.
- [RFC1981]        McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.
- [RFC2008]        Rekhter, Y. and T. Li, "Implications of Various Address Allocation Policies for Internet Routing", October 1996.
- [RFC2119]        Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.
- [RFC2460]        Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [RFC2736]        Handley, M. and C. Perkins, "Guidelines for Writers of RTP

Payload Format Specifications", RFC 2736, December 1999.

- [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, Feb. 2000.
- [RFC2775]     Carpenter, B., "Internet Transparency", RFC 2775, Feb. 2000.
- [RFC3056]     Carpenter, B. and K. Moore, "Connection of ipv6 domains via ipv4 clouds", RFC 3056, February 2001.
- [RFC3142]     Hagino, J. and K. Yamamoto, "An IPv6-to-IPv4 Transport Relay Translator", RFC 3142, June 2001.
- [RFC3171]     Albanna, Z., Almeroth, K., Meyer, D., and M. Schipper, "IANA Guidelines for IPv4 Multicast Address Assignments", RFC 3171, August 2001.
- [RFC3569]     Bhattacharyya, S., "An Overview of Source-Specific Multicast (SSM)", RFC 3569, July 2003.
- [RFC3956]     Savola , P. and B. Haberman , "Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address", RFC 3956, November 2004.
- [RFC4213]     Nordmark, E. and R. Gilligan, "Basic Transition Mechanisms for IPv6 Hosts and Routers", RFC 4213, October 2005.
- [RFC4214]     Templin, F., Gleeson, T., and D. Thaler, "Intra-site automatic tunnel addressing protocol (isatap)", RFC 4214, Coctober 2005.
- [RFC4291]     Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, Feb. 2006.
- [RFC4380]     Huitema , C., "Teredo: Tunneling ipv6 over udp through network address translations (nats)", RFC 4380, Feb. 2006.
- [RFC4607]     Holbrook , H. and B. Cain , "Source-Specific Multicast for IP", RFC 4607, August 2006.
- [RFC4611]     McBride, M., Meylor, J., and D. Meyer, "Multicast Source Discovery Protocol (MSDP) Deployment Scenarios", RFC 4611, August 2006.

- [RFC4925] Li, X., Dawkins, S., Ward, D., and A. Durand, "Softwire Problem Statement", RFC 4925, July 2007.
- [RFC4966] Aoun, C. and E. Davies, "Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status", RFC 4966, July 2007.
- [RFC768] Postel, J., "User Datagram Protocol", RFC 768, August 1981.
- [RFC791] Postel, J., "Internet Protocol", RFC 791, September 1981.
- [RFC793] Postel, J., "Transmission Control Protocol", RFC 793, September 1981.

#### 18.2. Informative References

- [APNIC] Ito, K., "Large IPv4 address space Usage trial for Future IPv6 Deployment", <http://www.apnic.net/meetings/25/program/policy/ito-large-ipv4-trial.pdf> .
- [CERNET] "CERNET Homepage:  
[http://www.edu.cn/english\\_1369/index.shtml](http://www.edu.cn/english_1369/index.shtml)".
- [CERNET2] "CNGI-CERNET2 Homepage:  
[http://www.cernet2.edu.cn/index\\_en.htm](http://www.cernet2.edu.cn/index_en.htm)".
- [CJ88] Kent, C. and J. Mogul, "Fragmentation considered harmful", SIGCOMM '87, 1988.
- [COUNT] "IPv4 address count down: <http://penrose.uk6x.com/>".
- [JJI07] Joseph, D., Chuang, J., and I. Stocia, "Modeling the Adoption of new Network Architectures", EECS Department, University of California, Berkeley Tech. Rep. UCB/EECS-2007-41, April 2007.

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

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

Phone: +86 62785983  
Email: mk@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

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](mailto:jianping@cernet.edu.cn)

#### Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).