behave
Internet-Draft
Intended status: Standards Track
Expires: July 5, 2013

C. Bao
X. Li
CERNET Center/Tsinghua University

January 1, 2013

# Extended IPv6 Addressing for Encoding Port Range
## draft-bcx-behave-address-fmt-extension-03

Abstract

This document discusses an extension of the algorithmic translation between IPv4 and IPv4-translatable IPv6 addresses. The extended address format contains transport-layer port set identification (PSID) which allows several IPv6 nodes to share a single IPv4 address with each node managing a different range of ports. This address format extension can be used for IPv4/IPv6 translation, as well as IPv4 over IPv6 tunneling.

Status of this Memo



Copyright Notice

Table of Contents

## 1. Introduction

This document discusses an extension of the address format defined in [RFC6052]. In Section 2.2, the IPv4-embedded IPv6 address format is defined which composed of a variable length prefix, the embedded IPv4 address, and a variable length suffix, as presented in the following diagram, in which PL designates the prefix length:

```
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|PL| 0-------------32--40--48--56--64--72--80--88--96--104-112-120-|
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|32|     prefix    |v4(32)         | u | suffix                    |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|40|     prefix        |v4(24)     | u |(8)| suffix                |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|48|     prefix            |v4(16) | u | (16)  | suffix            |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|56|     prefix                |(8)| u |  v4(24)   | suffix        |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|64|     prefix                    | u |   v4(32)      | suffix    |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|96|     prefix                                    |    v4(32)     |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

Figure 1: Address Format

In [RFC6052] Section 3.5, it states:

> "There have been proposals to complement stateless translation with a port-range feature. Instead of mapping an IPv4 address to exactly one IPv6 prefix, the options would allow several IPv6 nodes to share an IPv4 address, with each node managing a different range of ports. If a port range extension is needed, it could be defined later, using bits currently reserved as null in the suffix."

This document defines such a suffix encoding scheme and the corresponding port mapping algorithm.

### 1.1. Applicability Scope

The address format extension presented in this document is used for IPv4/IPv6 stateless translation and dual IPv4/IPv6 stateless translation without prefix delegation [I-D.xli-behave-divi]. The address format used for dual IPv4/IPv6 stateless translation and encapsulation with prefix delegation should refer to [I-D.ietf-softwire-map-t] [I-D.ietf-softwire-map].

### 1.2. Conventions

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

## 2. Port Mapping Algorithm

### 2.1. Mathematical representation of the Algorithm

There exist many port mapping algorithms and each one may have advantages and disadvantages, as well as has its best application scenario. Since different PSID MUST have non-overlapped port range, the two extreme cases are: (1) the port number is not continue for each PSID, but uniformly distributed cross the whole port range (0-65535); (2) the port number is continue in a single range for each PSID. The port mapping algorithm proposed here is called generalized modulus algorithm and it is flexible, meets these two cases and simple.

For given sharing ratio (R) and the maximum number of continue ports (M), the generalized modulus algorithm is defined as

1. The port number (P) of a given PSID (K) is composed of

   P = R*M*j + M*K + i

   Where
   o PSID: K=0 to R-1
   o Port range index: j = (1024/M)/R to ((65536/M)/R)-1, if the well-known port numbers (0-1023) are excluded.
   o Port continue index: i=0 to M-1

2. The PSID (K) of a given port number (P) is determined by

   K = (floor(P/M)) % R

   Where
   o % is modular operator
   o floor(arg) is a function returns the largest integer not greater than arg

3. The well-known port number (0-1023) can be used, if additional port mapping rule is defined.

## 2.2. Bit Representation of the Algorithm

Given sharing ratio (R=2^k), the maximum number of continue ports (M=2^m), for any PSID (K) available ports (P) can be represented as:

```
    0                          8                         15
    +---------------+----------+------+------------------+
    |                    P                               |
    ----------------+----------------+-------------------+
    |        A (j)  |   PSID (K)     |        M  (i)     |
    +---------------+----------+------+------------------+
    |<----a bits--->|<-----k bits---->|<------m bits----->|
                    |<---c bits--->|<-----(k+m-c) bits--->|
```

Figure 2: Bit representation

Where j and i are the same indexes defined in the port mapping algorithm.

For any port number, the PSID can be obtained by bit mask operation and therefore the generalized modulus algorithm does not introduce the computational complexity.

Note that in above figure there is a PSID prefix length (c). Based on this definition, PSID is also in CIDR style and more ports can be assigned to a single CE when PSID prefix length (c < k).

When m=0, the generalized modulus algorithm becomes modulus operation. When a=0, the generalized modulus algorithm becomes division operation.

## 2.3. Example of the Algorithm

### 2.3.1. PSID with fixed prefix length

For example, for R=128 (k=7), M=4 (m=2)

| | Port range-1 | Port rang-2 | Port |
|---|---|---|---|
| PSID=0 | 1024, 1025, 1026, 1027, | 1536, 1537, 1538, 1539, | 2048 |
| PSID=1 | 1028, 1029, 1030, 1031, | 1540, 1541, 1542, 1543, | .... |
| PSID=2 | 1032, 1033, 1034, 1035, | 1544, 1545, 1546, 1547, | .... |
| PSID=3 | 1036, 1037, 1038, 1039, | 1548, 1549, 1550, 1551, | .... |
| ... | | | |
| PSID=127 | 1532, 1533, 1534, 1535, | 2044, 2045, 2046, 2047, | .... |

Figure 3: Example 1

### 2.3.2. PSID with variable prefix length

For example, different PSIDs have different prefix length (c)

| Host | PSID prefix | Number of ports |
|---|---|---|
| Host0 | 000/2 | 2x8192 |
| Host1 | 010/3 | 1x8192 |
| Host2 | 011/3 | 1x8192 |
| Host3 | 100/1 | 4x8192 |

Figure 4: Example 2

## 2.4. Features of the Algorithm

The generalized modulus operation has the following features:

1. There is no waste of the port number, except the well-known ports.

2. The algorithm is flexible, the control parameters are sharing ratio (R), the continue port range (M) and PSID prefix length (c).

3. It does not introduce algorithm complexity.

4. It allows service providers to define their own address sharing ratio, the theoretical value is from 1:1 to 1:65536 and a more practical value is from 1:1 to 1:4096.

5. It supports deployments using differentiated port ranges.

6. It supports differentiated port ranges within a single shared IPv4 address.

7. It support excluding the well known ports 0-1023.

8. It supports assigning well known ports to a CE.

9. It supports legacy RTP/RTCP compatibility.

# 3. Extended IPv4-translatable IPv6 Address

### 3.1. Address Format

Based on the port mapping algorithm, the extended address format is shown in the following figure.

```
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|PL| 0-------------32--40--48--56--64--72--80--88--96--104-112-120-|
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|32|     prefix    |v4(32)         | u | PSID  | 0             | Q |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|40|     prefix        |v4(24)     | u |(8)| PSID | 0          | Q |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|48|     prefix            |v4(16) | u | (16)  | PSID | 0      | Q |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|56|     prefix                |(8)| u |  v4(24)   | PSID | 0  | Q |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|64|     prefix                    | u |   v4(32)      |PSID |0| Q |
+--+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

Figure 5: Address Format

Where PL designates the prefix length.

The PSID is placed right after the IPv4 address, since the combination of the IPv4 address and the PSID represents the more specifics in CIDR style which is sharing an IPv4 address with others.

The PSID prefix length (Q=c) is encoded in the last octet (bits 120-127) to indicate the number of ports can be used. When Q=0, the extended address format will become the address format defined in [RFC6052]. The relations between Q, the sharing ratio (R), the maximum continue port range (M) and the number of ports can be shown in the following figure.

| Q | Ratio | Maximum M | # of Ports |
|---|---|---|---|
| 0 | 1:1 | 65,536 | 65,536 |
| 1 | 1:2 | 32,786 | 32,786 |
| 2 | 1:4 | 16,384 | 16,384 |
| 3 | 1:8 | 8,192 | 8,192 |
| 4 | 1:16 | 4,096 | 4,096 |
| 5 | 1:32 | 2,048 | 2,048 |
| 6 | 1:64 | 1,024 | 1,024 |
| 7 | 1:128 | 512 | 512 |
| 8 | 1:256 | 256 | 256 |
| 9 | 1:512 | 128 | 128 |
| 10 | 1:1,024 | 64 | 64 |
| 11 | 1:2,048 | 32 | 32 |
| 12 | 1:4,096 | 16 | 16 |

Figure 6: Port range

Since newly defined IPv6 addresses with suffix are more specifics compared with the original address format defined in [RFC6052], the routing considerations in that document are also applied here. Furthermore, the port range is embedded in the extended IPv4-translatable IPv6 addresses and bound to the PSID therefore the packets containing extended IPv4-translatable IPv6 addresses as the destination can be routed to different IPv6 nodes.

### 3.2. Considerations of Using a Shorter Prefix length

Since IPv4 address plus variable length PSID represents the more specifics, the prefix length (PL) defined in [RFC6052] can be shorter. In these cases, the interface identifier (IID: second 64 bits) will not contains PSID and therefore can be used for regular prefix delegation, as shown in the following figure.

```
+--+---+---+----+----+-------+----+----+----+---+---+---+------+
|PL| 0-----24---28---32------56---60---64---72------96-----120-|
+--+---+---+----+----+-------+----+----+----+---+---+---+------+
|24| prefix|     v4(32)      |  PSID   | u  |              | Q |
+--+---+---+----+----+-------+----+----+----+---+---+---+------+
|28|   prefix   |     v4(32)      |PSID| u  |              | Q |
+--+---+---+----+----+-------+----+----+----+---+---+---+------+
|32|      prefix     |     v4(32)      | u  |              | Q |
+--+---+---+----+----+-------+----+----+----+---+---+---+------+
```

Figure 7: Shorter PL

Note that PL can take any value. For example,

o PL=24: Q=8, R=256
o PL=25: Q=7, R=128
o PL=26: Q=6, R=64
o PL=27: Q=5, R=32
o PL=28: Q=4, R=16
o PL=29: Q=3, R=8
o PL=30: Q=2, R=4
o PL=31: Q=1, R=2
o PL=32: Q=0, R=1

However, there will be a waste of the IPv6 address space in order to represent the IPv4-converted addresses.

### 3.3. Mapping Extended IPv4-translatable IPv6 Address to RFC1918 Space

Based on the algorithm defined in this document, a public IPv4 address and PSID can be mapped to extended IPv4-translatable IPv6 address and vise versa.

On the other hand, it is also possible to map the extended IPv4-translatable IPv6 address to [RFC1918] address space. In this case, one public IPv4 address can be mapped to several RFC1918 addresses and used by IPv4 or dual stack hosts.

For public IPv4 address a.b.c.d,

- o If R <= 256, the corresponding RFC1918 address is 10.c.d.PSID (PSID has 8 bits)

- o Otherwise, the corresponding RFC1918 address is 10.d.[PSID] (PSID has 16 bits)

## 4. DHCP Options Extensions

Based on the address format and the port mapping algorithm defined in this document, the IPv6 host needs to get the corresponding parameters via DHCPv6 [RFC3315][RFC3633] or others signaling scheme. These parameters are:

1. The IPv6 prefix

2. The IPv6 prefix length

3. The IPv4 prefix

4. The IPv4 prefix length

5. The sharing ratio (R)

6. The maximum number of continue ports (M)

7. The PSID (K)

8. The PSID length (c)

## 5. Comparisons with MAP

There are common parts and differences between this document and the address format defined in [I-D.ietf-softwire-map] [I-D.ietf-softwire-map-t].

1. The address format extension defined in this document is used for single and dual stateless translation without prefix delegation, while MAP is used for encapsulation and dual stateless translation with prefix delegation.

2. The address format extension defined in this document uses same IPv6 prefix for the source address from a CE to any destination (IPv4-translatable address) and the destination address from a CE to the outside IPv4 Internet (IPv4-converted address), while MAP uses different IPv6 prefixes, due to the requirements of prefix delegation.

3. The address format extension defined in this document uses same IPv6 prefix for all CEs, so there is no need to define prefix encoding scheme (e.g. CE index, or EA-bits), while MAP defines the prefix encoding scheme, due to the requirements of prefix delegation.

4. Due to the nature of using same IPv6 prefix for both IPv4-translatable address and IPv4-converted address, there is no referral problem and mesh scenarios can be supported without additional mapping rules, while MAP does require additional mapping rule for supporting mesh scenario.

5. The address format extension defined in this document and MAP share the same suffix coding scheme (IPv4 address + PSID).

6. The AFT and MAP share the same port mapping algorithm (generalized modulus algorithm).

## 6. IANA Considerations

This memo adds no new IANA considerations.

## 7. Security Considerations

There is no special security consideration.

## 8. Acknowledgements

Thanks to Wojciech Dec, Remi Despres, Ole Troan, Rajiv Asati, Chongfeng Xie, Qiong Sun, Fred Baker, Wing Dan, M. Boucadair and all the friends for the fruitful discussions during the unification of the port mapping algorithm of [I-D.mdt-softwire-mapping-address-and-port]. Also thanks to Wentao Shang, Guoliang Han, Yu Zhai, Haotao Zhang and Weicai Wang for their contributions.

## 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.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.

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

### 9.2. Informative References

[I-D.ietf-softwire-map]
Troan, O., Dec, W., Li, X., Bao, C., Matsushima, S., and

T. Murakami, "Mapping of Address and Port with Encapsulation (MAP)", draft-ietf-softwire-map-02 (work in progress), September 2012.

[I-D.ietf-softwire-map-t]
Li, X., Bao, C., Dec, W., Troan, O., Matsushima, S., and T. Murakami, "Mapping of Address and Port using Translation (MAP-T)", draft-ietf-softwire-map-t-00 (work in progress), October 2012.

[I-D.mdt-softwire-mapping-address-and-port]
Bao, C., Troan, O., Matsushima, S., Murakami, T., and X. Li, "Mapping of Address and Port (MAP)", draft-mdt-softwire-mapping-address-and-port-03 (work in progress), January 2012.

[I-D.xli-behave-divi]
Shang, W., Li, X., Zhai, Y., and C. Bao, "dIVI: Dual-Stateless IPv4/IPv6 Translation", draft-xli-behave-divi-04 (work in progress), October 2011.

[RFC6144] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", RFC 6144, April 2011.

[RFC6346] Bush, R., "The Address plus Port (A+P) Approach to the IPv4 Address Shortage", RFC 6346, August 2011.

## Authors' Addresses

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

Phone: +86 10-62785983
Email: congxiao@cernet.edu.cn

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

Phone: +86 10-62785983
Email: xing@cernet.edu.cn