Network Working Group   X. Li
Internet-Draft   C. Bao
Intended status: Experimental   CERNET Center/Tsinghua
Expires: January 12, 2014   University
                                                     W. Dec, Ed.
                                                       O. Troan
                                                   Cisco Systems
                                                  S. Matsushima
                                                 SoftBank Telecom
                                                    T. Murakami
                                                     IP Infusion
                                                   July 11, 2013

Mapping of Address and Port using Translation (MAP-T)
draft-ietf-softwire-map-t-03

Abstract

   This document specifies the "Mapping of Address and Port" double
   stateless NAT64 translation based solution (MAP-T) for providing
   shared or uniquely addressed IPv4 device connectivity to and across
   an IPv6 domain.

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 12, 2014.

Table of Contents

1. Introduction

   Experiences from IPv6 deployments in service provider networks such
   as [RFC6219] indicate that a successful transition to IPv6 can happen
   while allowing for continued support of IPv4 users, without incurring
   the burden of an end-end dual stack network. Due to IPv4 address
   exhaustion, this requires technology that supports shared IPv4
   address usage across an IPv6 network domain, while allowing the
   network operator to realize operational practices and deploy network
   equipment that are ideally the same for both IPv6 or IPv4 end-users.
   The use of double NAT64 translation based solutions is an optimal way
   to address these requirements, especially in combination with
   stateless translation techniques that seek to minimize operational
   challenges outlined in [I-D.ietf-softwire-stateless-4v6-motivation].

   The Mapping of Address and Port - Translation (MAP-T) solution
   specified in this document is a double NAT64 based solution, that
   builds on existing stateless NAT64 techniques specified in [RFC6145],
   along with a stateless algorithmic address & transport layer port
   mapping scheme, to allow the sharing of IPv4 addresses across an IPv6
   network. The MAP-T solution is closely related to MAP-E
   [I-D.ietf-softwire-map], with both utilizing the same address and
   port mapping & indexing method, but differing in their choice of IPv6
   domain transport, i.e. Translation [RFC6145] for MAP-T and
   encapsulation [RFC2473] for MAP-E. The translation mode is deemed
   valuable for environments where the encapsulation overhead, or IPv6
   oriented practices (e.g. use of IPv6 only servers, or IPv6 traffic
   classification) requirements, contribute to an encapsulation based
   solution being not feasable. These scenarios are presented in
   [I-D.maglione-softwire-map-t-scenarios]

   A companion document, applicable to both MAP-T and MAP-E, defines the
   DHCPv6 options for MAP provisioning [I-D.ietf-softwire-map-dhcp].

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 RFC 2119 [RFC2119].

3. Terminology

MAP domain: One or more MAP CEs and BRs connected by
                           means of an IPv6 network and sharing a common
                           set of MAP Rules. A service provider may
                           deploy a single MAP domain, or may utilize
                           multiple MAP domains.

MAP Rule: A set of parameters describing the mapping
                           between an IPv4 prefix, IPv4 address or
                           shared IPv4 address and an IPv6 prefix or
                           address. Each MAP domain uses a different
                           mapping rule set.

MAP Rule set   A Rule set is composed out of all thh MAP
                           Rules communicated to a device, that are
                           intended for determining traffic forwarding
                           operations. A set has at least one entry,
                           known as a default map rule. The Rule set is
                           interchangeably referred to in this document
                           as a Rule table.

MAP Rule table   See MAP Rule set.

MAP node: A device that implements MAP.

MAP Border Relay (BR):   A MAP enabled router managed by the service
                           provider at the edge of a MAP domain. A
                           Border Relay router has at least an IPv6-
                           enabled interface and an IPv4 interface
                           connected to the native IPv4 network. A MAP
                           BR may also be referred to simply as a "BR"
                           within the context of MAP.

MAP Customer Edge (CE): A device functioning as a Customer Edge
                           router in a MAP deployment. A typical MAP CE
                           adopting MAP rules will serve a residential
                           site with one WAN side interface, and one or
                           more LAN side interfaces. A MAP CE may also
                           be referred to simply as a "CE" within the
                           context of MAP.

Port-set: Each node has a separate part of the
                           transport layer port space; denoted as a
                           port-set.

Port-set ID (PSID):   Algorithmically identifies a set of ports
                           exclusively assigned to the CE.

   Shared IPv4 address: An IPv4 address that is shared among multiple
                           CEs. Only ports that belong to the assigned
                           port-set can be used for communication. Also
                           known as a Port-Restricted IPv4 address.

   End-user IPv6 prefix: The IPv6 prefix assigned to an End-user CE by
                           other means than MAP itself. E.g.
                           Provisioned using DHCPv6 PD [RFC3633],
                           assigned via SLAAC [RFC4862], or configured
                           manually. It is unique for each CE.

   MAP IPv6 address: The IPv6 address used to reach the MAP
                           function of a CE from other CEs and from BRs.

   Rule IPv6 prefix: An IPv6 prefix assigned by a Service Provider
                           for a MAP rule.

   Rule IPv4 prefix: An IPv4 prefix assigned by a Service Provider
                           for a MAP rule.

   Embedded Address (EA) bits: The IPv4 EA-bits in the IPv6 address
                           identify an IPv4 prefix/address (or part
                           thereof) or a shared IPv4 address (or part
                           thereof) and a port-set identifier.

4. Architecture

   Figure 1 depicts the overall MAP-T architecture, which sees any
   number of IPv4 users (N and M used as examples), connected by means
   of MAP-T CEs to an IPv6 network that is equipped with one or more
   MAP-T BR. The CEs and BRs form the MAP-T Domain, by means of
   configuration that they share.

   Functionally the MAP-T CE and BR utilize and extend some well
   established technical building blocks to allow the IPv4 users to
   correspond with nodes on the Public IPv4 network, or IPv6 network as
   follows:

   o A regular (NAT44) NAPT [RFC2663] function on a MAP CE is extended
      with support for restricting the allowable TCP/UDP ports for a
      given IPv4 address. The IPv4 address and port range used are
      determined by the MAP provisioning process and identical to MAP-E
      [I-D.ietf-softwire-map].

   o A standard stateless NAT64 function [RFC6145] is extended to allow
      stateless mapping of IPv4 and transport layer port ranges to IPv6
      address space. This algorithmic mapping is specified in section

5.

```
        User N
     Private IPv4
      |  Network
      |
  O--+--------------O
  |  | MAP-T CE      |
  | +-----+--------+ |
  | NAPT44|  MAP-T | |
  | +-----+       | |  -._     ,-------.                    .------.
  |      +--------+ |   ,-'           `-.              ,-'        `-.
  O----------------O  /               \  O--------O /   Public    \
                     /  IPv6 only    \ | MAP-T  |/     IPv4       \
                    (   Network    --+  Border +-   Network      )
                     \              /  | Relay  |\              /
  O----------------O  \            /   O--------O \            /
  |     MAP-T CE    |   ;".       ,-'             `-.       ,-'
  | +-----+--------+ | ,"    `----+--'              ------'
  | NAPT44|  MAP-T | | |,          |
  | +-----+       | | |     IPv6 node(s)
  |  |   +--------+ |      (w/ v4 mapped
  O---.-------------O        address)
      |
        User M
     Private IPv4
       Network
```

Figure 1: MAP-T Architecture

Each MAP-T CE is configured by means of MAP procedures with an IPv4
address and a port-range that is indexed by means of a Port Set
Identifier (PSID). Each CE is responsible for translating between a
given users' private IPv4 address space and the CE's MAP derived IPv4
address + port set, as well as adapting traffic between IPv4 and IPv6
using NAT64 procedures that are in accordance with the MAP Rules
applicable for a given domain. The MAP procedures can operate with
CE's using a shared IPv4 address, full IPv4 addresses or IPv4
prefixes, and place no assumption on the IPv6 addressing, other than
an IPv6 prefix of adequate size being allocated.

The MAP-T BR is responsible for connecting one or more MAP-T domains
to external IPv4 networks, using stateless NAT64 as extended by the
MAP rules in this document, to relay traffic between the two.

The intended role for NAT64 technology in the architecture is two
fold. Firstly, it is intended to allow the IPv6 network to focus on
IPv6 operational procedures with minimal consideration of IPv4-only

nodes attached to the domain. Secondly, it is intended to allow
IPv4-only nodes to correspond directly with IPv6-only nodes, provided
they have an IPv4 mapped IPv6 address belonging to the IPv6 prefix
assigned to the MAP-T domain (as per [RFC6052]).

The detailed operation of the above mechanism is governed by means of
MAP Rules and an address+port mapping algorithm covered in Section 5.
Section 7 describes how the mechanism is used for packet forwarding
operations.

5. Mapping Rules

A MAP node is provisioned with one or more mapping rules that govern
the IPv4 address and port-set are to a node in the IPv6 domain, as
well specific or default path forwarding behavior for the domain.
Three specific types of mapping rules are defined:

1. Basic Mapping Rule (BMR) - used for determining the CE's IPv4
    address and/or port set, as well as determining the MAP IPv6
    address that the CE is to use. For a given end-user IPv6 prefix
    there can be only one BMR. The BMR is defined out of the
    following parameters:

    * Rule IPv6 prefix (including prefix length)

    * Rule IPv4 prefix (including prefix length)

    * Rule EA-bits length (in bits)

    * Optional Rule Port Parameters

2. Forwarding Mapping Rule (FMR) - used for setting up forwarding
    between CEs in the MAP domain (a.k.a. Mesh mode). Each
    Forwarding Mapping Rule will result in a forwarding entry for the
    Rule IPv4 prefix + the given port range, i.e. Specific IPv4 +
    port routes.The FMR consists of the following parameters, which
    are shared with the BMR:

    * Rule IPv6 prefix (including prefix length)

    * Rule IPv4 prefix (including prefix length)

    * Rule EA-bits length (in bits)

    * Optional Rule Port Parameters

3. Default Mapping Rule (DMR) - used for mapping and forwarding to
    destinations outside the MAP domain, i.e. a default route for the
    MAP domain leading to the MAP BR. It consists of:

    * The IPv6 prefix (including prefix length) used to represent
       destinations outside the MAP domain. Typically a routed
       prefix to one or more BRs.

4. Optional Rule Port Parameters - used to represent additional
    configuration settings. Currently defined parameters are

    * Offset: Specifies the numeric value for the MAP algorithm's
       excluded port range/offset bits (A-bits). Unless explicitly
       defined this value MUST default to 6.

By default, every MAP node belonging to a MAP domain node, MUST be
provisioned with a Basic Mapping Rule (BMR). The rule is then used
for IPv4 prefix, address or shared address assignment.

A MAP IPv6 address is formed from the BMR Rule IPv6 prefix. This
address MUST be assigned to an interface of the MAP node and is used
to terminate all MAP traffic being sent or received to the node.

Port-aware IPv4 entries in the Rules table are installed for all the
Forwarding Mapping Rules and a default route to the MAP BR as per the
DMR (see section Section 5.3). A given domain can have only one DMR,
however be deployed for load balancing using multiple BRs, for
example by means of anycast addressing of the BRs.

Forwarding rules are used to allow direct communication between MAP
CEs, known as mesh mode. In hub and spoke mode, there are no
forwarding rules, and all traffic is forwarded from the CE to the BR
by means of the DMR.

The following subsections specify the MAP algorithm and its use of
Rules.

5.1. Basic mapping rule (BMR)

The Basic Mapping Rule is mandatoryfor a MAP CE, and is used by the
CE to derive its IPv4 prefix, IPv4 address or shared IPv4 address and
associated port-range in conjunction with the information in the end-
user IPv6 prefix. Recall from Section 5 that the BMR consists of the
following parameters:

o Rule IPv6 prefix, of a length n.

o Rule IPv4 prefix, of a length r.

o Rule EA-bits of length o.

o Optional Rule Port Parameters (a, k)

Figure 2 shows the structure of the complete MAP IPv6 address of a CE
as specified in this document, and its relation to the information
contained in the BMR and End-user IP6 prefix. The MAP CE IPv6
address is determined by concatenating the End-user IPv6 prefix with
the MAP subnet-id (if the End-user IPv6 prefix is shorter than 64
bits) and the MAP interface ID. The MAP interface-id is derived as
specified in Section 5.5. The MAP subnet ID is defined to be the
first subnet (all bits set to zero). For End-user IPv6 prefixes
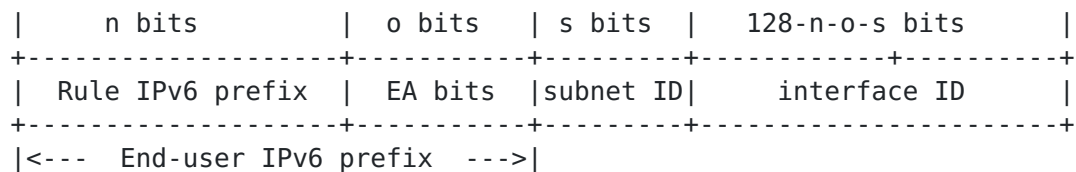longer than 64 bits, no MAP subnet id is used.


```
|      n bits        | o bits   | s bits |    128-n-o-s bits     |
+--------------------+----------+--------+-----------+----------+
|   Rule IPv6 prefix |  EA bits |subnet ID|    interface ID      |
+--------------------+----------+--------+----------------------+
|<---  End-user IPv6 prefix  --->|
```

Figure 2: IPv6 address format

The MAP CE's IPv4 address is determined by completing the r-bits of
the Rule IPv4 prefix with the remaining IPv4 suffix 32-r bits of
information (p), along with the optional k bits of the Port Set
Identifier (PSID). These remaining p + k bits of information
themselves come from the (o) Embedded-Address (EA) bits of the end-
user IPv6 prefix. The End-user IPv6 prefix is the IPv6 prefix
assigned to the CE and is unique per CE.

The n bit Rule IPv6 prefix, is the part of the End-user IPv6 prefix
that is common among all CEs using the same Basic Mapping Rule within
the MAP domain. Similarly, the Rule IPv4 prefix of length r is the
IPv4 prefix common among all CEs using the same BMR within the MAP
domain. An EA-bit length of 0 signifies that all relevant p and k
bits of addressing information are passed directly in the BMR, and
not derived from the EA bits of the End-user IPv6 prefix. Examples
of these and other cases are given in Appendix A.

For a given BMR, if o + r < 32 (length of the IPv4 address in bits),
then an IPv4 prefix is being intended for use by the BMR. This case
is shown in Figure 4.

```
              |   r bits   |      32-r bits       |
              +------------+----------------------+
              |  Rule IPv4 | IPv4 Address suffix |
              +------------+----------------------+
              |              < 32 bits            |
```
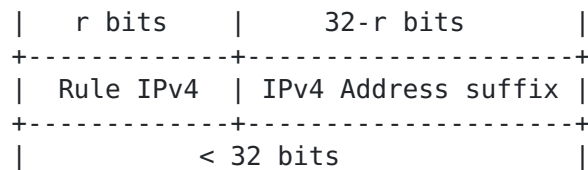
Figure 3: IPv4 prefix

If o + r is equal to 32, then a full IPv4 address is to be assigned.
The address is created by concatenating the Rule IPv4 prefix and the
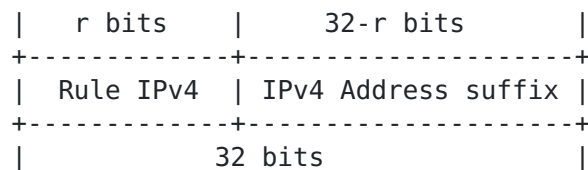EA-bits. This case is shown in Figure 5.

```
              |   r bits   |      32-r bits       |
              +------------+----------------------+
              |  Rule IPv4 | IPv4 Address suffix |
              +------------+----------------------+
              |              32 bits              |
```

Figure 4: Complete IPv4 address

If o + r is > 32, then a shared IPv4 address is to be assigned. The
number of IPv4 address suffix bits (p) in the EA bits is given by 32
- r. The PSID bits are used to create a port-set. The length of the
PSID bit field within EA bits is: k = o - 32 + r.

```
     |   r bits   |      32-r bits       |        |   k bits   |
     +------------+----------------------+        +------------+
     |  Rule IPv4 | IPv4 Address suffix |        |Port-Set ID |
     +------------+----------------------+        +------------+
     |              32 bits              |
```
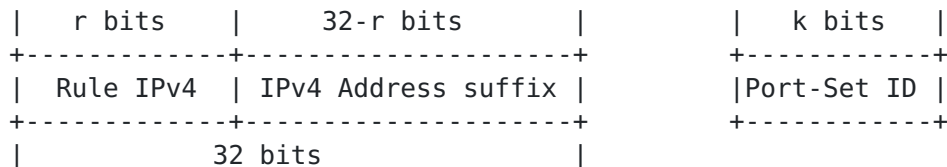
Figure 5: Shared IPv4 address

It should be noted that the length r MAY be zero, in which case the
complete IPv4 address or prefix is encoded in the EA bits. Similarly
the length of o MAY, in which case no part of the CE's IPv6 end-user
prefix is used to derive the CE's IPv4 address. To create a complete
IPv4 address (or prefix), the IPv4 address suffix (p = 32-r) from the
EA bits, is concatenated with the Rule IPv4 prefix (r bits).

The BMR is provisioned to the CE by means (e.g. a DHCPv6 option) not
specified in this document.

See [Appendix A](#) for an example of the Basic Mapping Rule.

5.2. Forwarding mapping rule (FMR)

The Forwarding Mapping Rule is an optional rule used in mesh mode to enable direct CE to CE connectivity.

The processing of an FMR rule results in a route entry being installed on the processing MAP device for the IPv4 Rule prefix and any associated port range. The "next hop" of such a route is the MAP transformation defined by the rule's key elements:

o The Rule IPv6 prefix, of a length n.

o The Rule IPv4 prefix, of a length r.

o The Rule EA-bits of length o.

o Optional Rule Port Parameters (a, k)

On forwarding an IPv4 packet, a best matching prefix look up is done and the closest matching FMR is chosen. The IPv6 destination address is derived from the destination IPv4 + port in combination with the rule's parameters as exemplified in Figure 6.
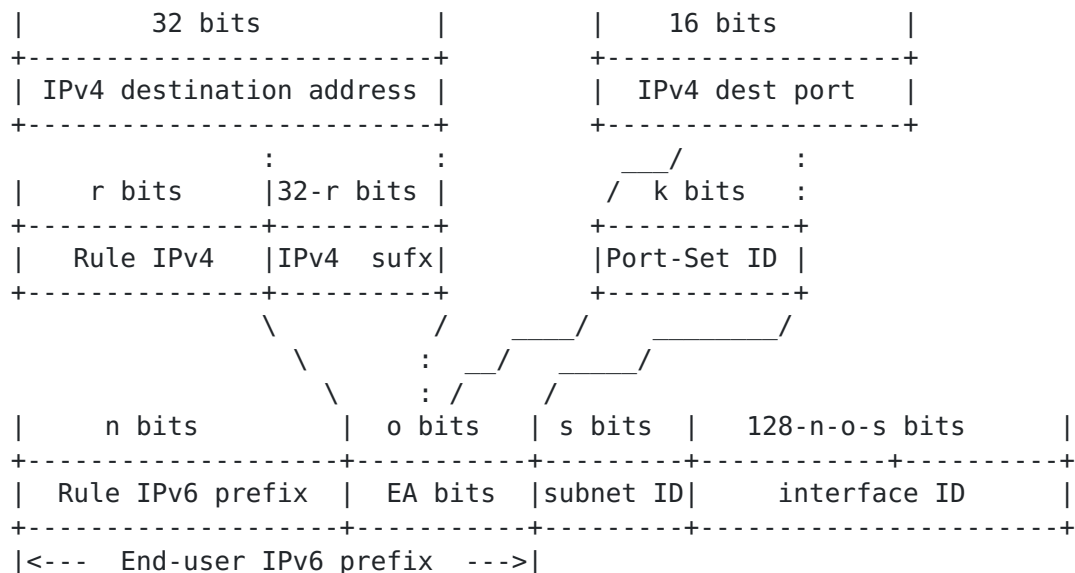
```
|        32 bits          |        |     16 bits        |
+-------------------------+        +-------------------+
| IPv4 destination address |        |  IPv4 dest port   |
+-------------------------+        +-------------------+
            :         :              ___/        :
|    r bits    |32-r bits |          /  k bits    :
+--------------+----------+        +------------+
|   Rule IPv4  |IPv4  sufx|        |Port-Set ID |
+--------------+----------+        +------------+
            \          /  ____/      _____/
              \      :  __/    _____/
               \    : /      /
|     n bits          | o bits  | s bits |   128-n-o-s bits      |
+--------------------+---------+---------+------------+----------+
|  Rule IPv6 prefix  | EA bits |subnet ID|    interface ID      |
+--------------------+---------+---------+----------------------+
|<---  End-user IPv6 prefix  --->|
```

Figure 6: Deriving of MAP IPv6 address

See [Appendix A](#) for an example of the Forwarding Mapping Rule.

[5.3](#). Port mapping algorithm

The port mapping algorithm is used in domains whose MAP Rules allows
IPv4 address sharing, and is intended to allow the a range of ports
to be represented by an algorithmically computable index, the Port
Set Identifier (PSID) that is unique for each CE.

The simplest way to represent a port range is using a notation
similar to CIDR [[RFC4632](#)]. For example the first 256 ports are
represented as port prefix 0.0/8. The last 256 ports as 255.0/8. In
hexadecimal, 0x0000/8 (PSID = 0) and 0xFF00/8 (PSID = 0xFF). Using
this technique, but wishing to avoid allocating the system ports
[[I-D.ietf-tsvwg-iana-ports](#)] to a give CE, one would have to exclude
the use of one or more PSIDs (e.g., PSIDs 0 to 3 in the example just
given).

As will be seen shortly, the PSID forms a portion of the End-user
IPv6 prefix, however it is desirable to minimize the dependencies
between the End-user IPv6 prefix and the assigned port set. This is
achieved by using an infix representation of the port value. Using
such a representation, the well-known ports are excluded by
restrictions on the value of the first A high-order bits of the
transport port space, known as the A-bit field, rather than the PSID
itself, whihc directly follows. For a given A-bit field value, and a
given PSID, the range of contiguous ports being represented are all
the combinations of the remaining m wildcard bits (i.e. 2^m
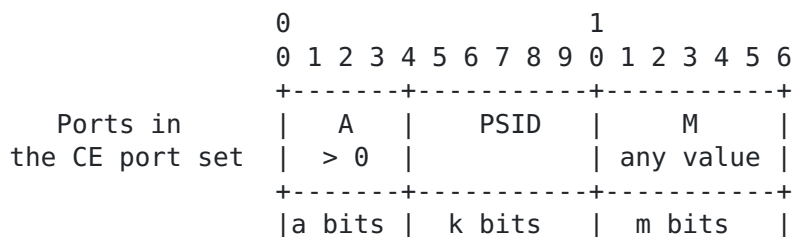combinations) out of the 16-bit field, as shown in the figure below.

```
                        0                   1
                        0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6
                        +-------+-----------+-----------+
         Ports in       |   A   |    PSID   |     M     |
     the CE port set    |  > 0  |           | any value |
                        +-------+-----------+-----------+
                        |a bits |  k bits   |  m bits   |
```

                        Figure 7: PSID

A Selects the range of the port number. For a > 0, A MUST be larger
   than 0. This ensures that the algorithm excludes the system
   ports. For this value of a, the system ports, but no others, are
   excluded by requiring that A be greater than 0. For smaller
   values of a, A still has to be greater than 0, but this excludes
   ports above 1023. For larger values of a, the minimum value of A
   has to be higher to exclude all the system ports. The interval

between successive contiguous ranges assigned to the same user is
2^a.

a-bits The number of offset bits. The default Offset bits (a) are:
6. To simplify the port mapping algorithm the defaults are chosen
so that the PSID field starts on a nibble boundary and the
excluded port range (0-1023) is extended to 0-4095.

PSID The Port Set Identifier. Different Port-Set Identifiers (PSID)
MUST have non-overlapping port-sets.

k-bits The length in bits of the PSID field. The sharing ratio is
k^2. The number of ports assigned to the user is 2^(16-k) - 2^m
(excluded ports)

M Selects the specific port within the particular range specified by
the concatenation of A and the PSID.

m bits The contiguous port size, i.e. the number of contiguous ports
allocated to a given PSID. The number of contiguous ports is
given by 2^m.

5.4. Default mapping rule (DMR)

IPv4 traffic between MAP-T nodes that are all within one MAP domain
is translated to IPv6, with the senders MAP IPv6 address as the IPv6
source address and the receiving MAP node's MAP IPv6 address as the
IPv6 destination address. To reach destinations outside the MAP-T
domain and/or for the case when the MAP domain is defined to be
composed out of a single CE and BR, the Default Mapping rule is used.
The DMR is specified in terms of the BR IPv6 prefix that MAP-T CEs
will use for mapping an IPv4 destination address.

 Default Mapping Rule:
     {2001:db8:0001::/Prefix-length (Rule IPv6 prefix),
       0.0.0.0/0 (Rule IPv4 prefix)}

                   Example: Default Mapping Rule

It is recommended that the BR prefix-length SHOULD be by default 64
bits long, and in any case MUST NOT exceed 96 bits. The mapping of
the IPv4 destination behind the IPv6 prefix will by default follow
the /64 rule as per [RFC6052]. Any trailing bits after the IPv4
address are set to 0x0.

5.5. The IPv6 Interface Identifier

The Interface identifier format of a MAP node is described below.

```
|         128-n-o-s bits          |
| 16 bits|    32 bits     | 16 bits|
+--------+----------------+--------+
|   0    |  IPv4 address  |  PSID  |
+--------+----------------+--------+
```
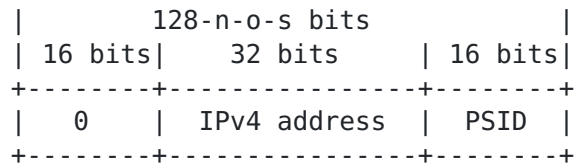
Figure 8

In the case of an IPv4 prefix, the IPv4 address field is right-padded
with zeros up to 32 bits. The k-bit PSID is zero left-padded to
create a 16 bit field. For an IPv4 prefix or a complete IPv4
address, the PSID field is zero.

If the End-user IPv6 prefix length is larger than 64, the most
significant parts of the interface identifier is overwritten by the
prefix.

6. MAP-T Configuration

For a given MAP domain, the BR and CE MUST be configured with the
following MAP elements. The configured values for these elements are
identical for all CEs and BRs within a given MAP domain.

o The Basic Mapping Rule and optionally the Forwarding Mapping
   Rules, including the Rule IPv6 prefix, Rule IPv4 prefix, and
   Length of EA bits

o Hub and spoke mode or Mesh mode. (If all traffic should be sent
   to the BR, or if direct CE to CE traffic should be supported).

o Use of Translation mode (MAP-T)

o The BR's IPv6 prefix used in the DMR

The MAP-T CE and BR configuration is the same as for MAP-E described
in Section 7 of [I-D.ietf-softwire-map] except for two differences:

o Translation mode is used instead of Encapsulation

   o Use of the BR's IPv6 prefix instead of address

6.1. MAP CE

   The MAP elements are set to values that are the same across all CEs
   within a MAP domain. The values may be configured in a variety of
   manners, including provisioning methods such as the Broadband Forum's
   "TR-69" Residential Gateway management interface, an XML-based object
   retrieved after IPv6 connectivity is established, DHCPv6, or manual
   configuration by an administrator. This document does not prescribe
   any of these methods, but recommends that a MAP CE SHOULD implement
   DHCPv6 options as per [I-D.ietf-softwire-map-dhcp]. Other
   configuration and management methods may use the format described by
   this option for consistency and convenience of implementation on CEs
   that support multiple configuration methods.

   The only remaining provisioning information the CE requires in order
   to calculate the MAP IPv4 address and enable IPv4 connectivity is the
   IPv6 prefix for the CE. The End-user IPv6 prefix is configured as
   part of obtaining IPv6 Internet access, and requires no special
   handling.

   The MAP provisioning parameters, and hence the IPv4 service itself,
   is tied to the End-user IPv6 prefix; thus, the MAP service is also
   tied to this in terms of authorization, accounting, etc. The MAP
   IPv4 address, prefix or shared IPv4 address and port set has the same
   lifetime as its associated End-user IPv6 prefix.

   A single MAP CE MAY be connected to more than one MAP domain, just as
   any router may have more than one IPv4-enabled service provider
   facing interface and more than one set of associated addresses
   assigned by DHCPv6. Each domain a given CE operates within would
   require its own set of MAP configuration elements and would generate
   its own IPv4 address. The MAP DHCPv6 option is specified in
   [I-D.ietf-softwire-map-dhcp]

6.2. MAP BR

   The MAP BR MUST be configured with the same MAP elements as the MAP
   CEs operating within the same domain.

   For increased reliability and load balancing, the BR IPv6 prefix MAY
   be shared across a given MAP domain. As MAP is stateless, any BR may
   be used at any time.

   Since MAP uses provider address space, no specific routes need to be
   advertised externally for MAP to operate, neither in IPv6 nor IPv4
   BGP. However, the BR prefix needs to be advertised in the service

provider's IGP.

7. MAP-T Packet Forwarding

The end-end packet flow in MAP-T involves an IPv4 or IPv6 packet
being forwarded across one or both of a CE and a BR, in one of two
directions in for each such case.

7.1. IPv4 to IPv6 at the CE

A MAP-T CE receiving IPv4 packets SHOULD perform NAPT NAT44 function,
and create any necessary NAPT44 bindings. The source address and
port of the packet obtained as a result of the NAPT44 process MUST
correspond to the source IPv4 address and source transport port
number derived to belong to the CE by means of the MAP Basic Mapping
Rule (BMR).

The resulting IPv4 packet is subject to a longest IPv4 address + port
match MAP rule selection, which then determines the parameters for
the subsequent NAT64 operation. By default, all traffic is matched
to the default mapping rule (DMR), and subject to the stateless NAT64
operation using the DMR parameters for the MAP algorithm and NAT64.
Packets matching destinations covered by any (optional) forward
mapping rules (FMRs) are subject to the stateless NAT64 operation
using the FMR parameters for the MAP algorithm and stateless NAT64.

A MAP-T CE MUST support a default mapping rule and SHOULD support one
or more forward mapping rules.

7.2. IPv6 to IPv4 at the CE

A MAP-T CE receiving an IPv6 packet performs its regular IPv6
operations (filtering, pre-routing, etc). Only packets that are
addressed to the CE's MAP-T addresses, and with source addresses
matching the IPv6 map-rule prefixes of a DMR or FMR, are processed by
the MAP-T CE. All other IPv6 traffic SHOULD be forwarded as per the
CE's IPv6 routing rules. The CE SHOULD check that MAP-T received
packets' destination transport-layer destination port number is in
the range allowed for by the CE's MAP BMR configuration. The CE
SHOULD drop any non conforming packet and respond with an ICMPv6
"Address Unreachable" (Type 1, Code 3). For packets whose source
address matches an FMR, the CE SHOULD perform a check of consistency
of the source against the allowed values from the source port-range.
If the packets' source port number is found to be outside the range
allowed, the CE MUST drop the packet and SHOULD respond with an
ICMPv6 "Destination Unreachable, Source address failed ingress/egress
policy" (Type 1, Code 5).

For each MAP-T processed packet, the CE's NAT64 function MUST derive
the IPv4 source and destination addresses. The IPv4 destination
address is derived by extracting relevant information from the IPv6
destination and the information stored in the BMR as per Section 5.1
of this document. The IPv4 source address is formed by classifying
the packet's source as matching a DMR or FMR rule prefix, and then
using that NAT64 rule-set, as per Section 5.4 or Section 5.2
respectively.

The resulting IPv4 packet is then forwarded to the CE's NAPT NAT44
function, where the destination IPv4 address and port number MUST be
mapped to their original value, before being forwarded according to
the CE's regular IPv4 rules. When the NAPT function is not enabled,
the traffic from the stateless NAT64 function is directly forwarded
according to the CE's IPv4 rules.

7.3. IPv6 to IPv4 at the BR

A MAP-T BR receiving IPv6 packets MUST select a matching MAP rule
based on a longest address match of the packets' source address
against the BR's configured MAP Rules. In combination with the port-
set-id contained in the packet's source IPv6 address, the selected
MAP rule allows the BR to verify that the CE is using its allowed
address and port range. Thus, the BR MUST perform a validation of
the consistency of the source against the allowed values from the
identified port-range. If the packets' source port number is found
to be outside the range allowed, the BR MUST drop the packet and
respond with an ICMPv6 "Destination Unreachable, Source address
failed ingress/egress policy" (Type 1, Code 5).

When constructing the IPv4 packet, the BR MUST derive the source and
destination IPv4 addresses as per Section 5 of this document and
translate the IPv6 to IPv4 headers as per [RFC6145]. The resulting
IPv4 packets are then passed to regular IPv4 forwarding.

7.4. IPv4 to IPv6 at the BR

A MAP-T BR receiving IPv4 packets uses a longest match IPv4 +
transport layer port lookup to identify the target MAP-T domain and
rule. The MAP-T BR MUST then compute the IPv6 destination addresses
from the IPv4 destination address and port as per Section 5.1 of this
document. The MAP-T BR MUST also compute the IPv6 source addresses
from the IPv4 source address as per Section 5.4 (i.e. It needs to
form an IPv6 mapped IPv4 address using the BR's DMR prefix).
Throughout the generic IPv4 to IPv6 header procedures following
[RFC6145] apply. The resulting IPv6 packets are then passed to
regular IPv6 forwarding.

Note that the operation of a BR when forwarding to MAP-T domains that
are defined without IPv4 address sharing is the same as stateless
NAT64 IPv4/IPv6 translation.

8. ICMP Handling

ICMP messages supported in the MAP-T domain need to take into
consideration also the NAPT44 component and best current practice
documented in [RFC5508] along with some additional specific
considerations.

MAP-T CEs and BRs MUST follow ICMP/ICMPv6 translation as per
[RFC6145], with the following extension to cover the address sharing/
port-range feature.

Unlike TCP and UDP, which provide two transport protocol port fields
to represent both source and destination, the ICMP/ICMPv6 [RFC0792],
[RFC4443] Query message header has only one ID field which needs to
be used to identify a sending IPv4 host.

When receiving IPv4 ICMP messages, the MAP-T CE MUST rewrite the ID
field to a port value derived from the CE's Port-set-id.

In the return path, when MAP-T BR receives an IPv4 ICMP packet
containing an ID field which is bound for a shared address in the
MAP-T domain, the MAP-T BR SHOULD use the ID value as a substitute
for the destination port in determining the IPv6 destination address.
In all other cases, the MAP-T BR MUST derive the destination IPv6
address by simply mapping the destination IPv4 address without
additional port info. Throughout the ICMP message MUST be translated
as per [RFC6145] with the the ID field preserved.

9. Fragmentation and Path MTU Discovery

Due to the different sizes of the IPv4 and IPv6 header, handling the
maximum packet size is relevant for the operation of any system
connecting the two address families. There are three mechanisms to
handle this issue: Path MTU discovery (PMTUD), fragmentation, and
transport-layer negotiation such as the TCP Maximum Segment Size
(MSS) option [RFC0897]. MAP uses all three mechanisms to deal with
different cases.

9.1. Fragmentation in the MAP domain

Translating an IPv4 packet to carry it across the MAP domain will
increase its size by 20 bytes respectively. It is strongly

recommended that the MTU in the MAP domain is well managed and that
the IPv6 MTU on the CE WAN side interface is set so that no
fragmentation occurs within the boundary of the MAP domain.

Fragmentation in MAP-T domain is to be handled as described in
section 4 and 5 of [RFC6145].

9.2. Receiving IPv4 Fragments on the MAP domain borders

Forwarding of an IPv4 packet received from the outside of the MAP
domain requires the IPv4 destination address and the transport
protocol destination port. The transport protocol information is
only available in the first fragment received. As described in
section 5.3.3 of [RFC6346] a MAP node receiving an IPv4 fragmented
packet from outside has to reassemble the packet before sending the
packet onto the MAP link. If the first packet received contains the
transport protocol information, it is possible to optimize this
behavior by using a cache and forwarding the fragments unchanged. A
description of this algorithm is outside the scope of this document.

9.3. Sending IPv4 fragments to the outside

If two IPv4 host behind two different MAP CE's with the same IPv4
address sends fragments to an IPv4 destination host outside the
domain. Those hosts may use the same IPv4 fragmentation identifier,
resulting in incorrect reassembly of the fragments at the destination
host. Given that the IPv4 fragmentation identifier is a 16 bit
field, it could be used similarly to port ranges. A MAP CE SHOULD
rewrite the IPv4 fragmentation identifier to be within its allocated
port set.

10. Usage Considerations

10.1. EA-bit length of 0

The MAP solution supports use and configuration of domains with a BMR
expressing an EA-bit length of 0. This results in independence
between the end-user IPv6 prefix assigned to the CE and the IPv4
address and/or port-range used by MAP. The k-bits of PSID
information may in this case be derived from the BMR.

The constraint imposed is that each such MAP domain be composed of
just 1 MAP CE which has a predetermined IPv6 prefix. The BR would be
configured with an FRM rule per CPE, where the FMR would uniquely
describe the IPv6 prefix of a given CE. Each CE would have a
distinct BMR, that would fully describe that CE's IPv4 address, and
PSID if any.

10.2. Mesh and Hub and spoke modes

The hub and spoke mode of communication, whereby all traffic sent by
a MAP-T CE is forwarded via a BR, and the mesh mode, whereby a CE is
directly able to forward traffic to another CE, are governed by the
activation of Forward Mapping Rule that cover the IPv4-prefix
destination, and port-index range. By default, a MAP CE configured
only with a BMR, as per this specification, will use it to configure
its IPv4 parameters and IPv6 MAP address without enabling mesh mode.

10.3. Communication with IPv6 servers in the MAP-T domain

By default, MAP-T allows communication between both IPv4-only and any
IPv6 enabled devices, as well as with native IPv6-only servers
provided that the servers are configured with an IPv4-mapped IPv6
address. This address could be part of the the IPv6 prefix used by
the DMR in the MAP-T domain. Such IPv6 servers (e.g. an HTTP server,
or a web content cache device) are thus able to serve both IPv6 users
as well as IPv4-only users users alike utilizing IPv6. Any such
IPv6-only servers SHOULD have both A and AAAA records in DNS. DNS64
[RFC6147] become required only when IPv6 servers in the MAP-T domain
are expected themselves to initiate communication to external IPv4-
only hosts.

10.4. Compatibility with other NAT64 solutions

A MAP-T CE is by default compatible with [RFC6146] stateful NAT64
devices that are placed to use/advertise the BR prefix. This in
effect allows the use of MAP-T CEs in environments that need to
perform statistical multiplexing of IPv4 addresses, while utilizing
stateful NAT64 devices, and can take the role of a CLAT as defined in
[RFC6877].

Furthermore, a MAP-T CE configured to operate without address sharing
(no PSID) is compatible with any stateless NAT64 devices positioned
as BRs.

11. IANA Considerations

This specification does not require any IANA actions.

12. Security Considerations

Spoofing attacks: With consistency checks between IPv4 and IPv6
    sources that are performed on IPv4/IPv6 packets received by MAP
    nodes, MAP does not introduce any new opportunity for spoofing
    attacks that would not already exist in IPv6.

Denial-of-service attacks: In MAP domains where IPv4 addresses are
    shared, the fact that IPv4 datagram reassembly may be necessary
    introduces an opportunity for DOS attacks. This is inherent to
    address sharing, and is common with other address sharing
    approaches such as DS-Lite and NAT64/DNS64. The best protection
    against such attacks is to accelerate IPv6 enablement in both
    clients and servers so that, where MAP is supported, it is less
    and less used.

Routing-loop attacks: This attack may exist in some automatic
    tunneling scenarios are documented in [RFC6324]. They cannot
    exist with MAP because each BRs checks that the IPv6 source
    address of a received IPv6 packet is a CE address based on
    Forwarding Mapping Rule.

Attacks facilitated by restricted port set: From hosts that are not
    subject to ingress filtering of [RFC2827], some attacks are
    possible by an attacker injecting spoofed packets during ongoing
    transport connections ([RFC4953], [RFC5961], [RFC6056]). The
    attacks depend on guessing which ports are currently used by
    target hosts, and using an unrestricted port set is preferable,
    i.e. Using native IPv6 connections that are not subject to MAP
    port range restrictions. To minimize this type of attacks when
    using a restricted port set, the MAP CE's NAT44 filtering behavior
    SHOULD be "Address-Dependent Filtering". Furthermore, the MAP CEs
    SHOULD use a DNS transport proxy function to handle DNS traffic,
    and source such traffic from IPv6 interfaces not assigned to
    MAP-T. Practicalities of these methods are discussed in Section
    5.9 of [I-D.dec-stateless-4v6].

ICMP Flood Given the necessity to process and translate ICMP and
    ICMPv6 messages by the BR and CE nodes, a foreseeable attack
    vector is that of a flood of such messages leading to a saturation
    of the nodes' compute resources. This attack vector is not
    specific to MAP, and its mitigation lies a combination of policing
    the rate of ICMP messages, policing the rate at which such
    messages can get processed by the MAP nodes, and of course
    identifying and blocking off the source(s) of such traffic.

[RFC6269] outlines general issues with IPv4 address sharing.

13. Contributors

The following individuals authored major contribution to this document:

Chongfeng Xie (China Telecom) Room 708, No.118, Xizhimennei Street Beijing 100035 CN Phone: +86-10-58552116 Email: xiechf@ctbri.com.cn

Qiong Sun (China Telecom) Room 708, No.118, Xizhimennei Street Beijing 100035 CN Phone: +86-10-58552936 Email: sunqiong@ctbri.com.cn

Rajiv Asati (Cisco Systems) 7025-6 Kit Creek Road Research Triangle Park NC 27709 USA Email: rajiva@cisco.com

Gang Chen (China Mobile) 53A,Xibianmennei Ave. Beijing 100053 P.R.China Email: chengang@chinamobile.com

Wentao Shang (CERNET Center/Tsinghua University) Room 225, Main Building, Tsinghua University Beijing 100084 CN Email: wentaoshang@gmail.com

Guoliang Han (CERNET Center/Tsinghua University) Room 225, Main Building, Tsinghua University Beijing 100084 CN Email: bupthgl@gmail.com

Yu Zhai CERNET Center/Tsinghua University Room 225, Main Building, Tsinghua University Beijing 100084 CN Email: jacky.zhai@gmail.com

14. Acknowledgements

This document is based on the ideas of many. In particular Remi Despres, who has tirelessly worked on generalized mechanisms for stateless address mapping.

The authors would like to thank Mohamed Boucadair, Guillaume Gottard, Dan Wing, Jan Zorz, Necj Scoberne, Tina Tsou, , Gang Chen, Maoke Chen, Xiaohong Deng, Jouni Korhonen, Tomasz Mrugalski, Jacni Qin, Chunfa Sun, Qiong Sun, Leaf Yeh, Andrew Yourtchenko for their review and comments.

15. References

15.1. Normative References

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

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

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

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

15.2. Informative References

   [I-D.dec-stateless-4v6]
              Dec, W., Asati, R., and H. Deng, "Stateless 4Via6 Address
              Sharing", draft-dec-stateless-4v6-04 (work in progress),
              October 2011.

   [I-D.ietf-softwire-map]
              Troan, O., Dec, W., Li, X., Bao, C., Matsushima, S.,
              Murakami, T., and T. Taylor, "Mapping of Address and Port
              with Encapsulation (MAP)", draft-ietf-softwire-map-07
              (work in progress), May 2013.

   [I-D.ietf-softwire-map-dhcp]
              Mrugalski, T., Troan, O., Dec, W., Bao, C.,
              leaf.yeh.sdo@gmail.com, l., and X. Deng, "DHCPv6 Options
              for Mapping of Address and Port",
              draft-ietf-softwire-map-dhcp-03 (work in progress),
              February 2013.

   [I-D.ietf-softwire-stateless-4v6-motivation]
              Boucadair, M., Matsushima, S., Lee, Y., Bonness, O.,
              Borges, I., and G. Chen, "Motivations for Carrier-side
              Stateless IPv4 over IPv6 Migration Solutions",
              draft-ietf-softwire-stateless-4v6-motivation-05 (work in
              progress), November 2012.

   [I-D.ietf-tsvwg-iana-ports]
              Cotton, M., Eggert, L., Touch, J., Westerlund, M., and S.
              Cheshire, "Internet Assigned Numbers Authority (IANA)
              Procedures for the Management of the Service Name and
              Transport Protocol Port Number Registry",
              draft-ietf-tsvwg-iana-ports-10 (work in progress),
              February 2011.

   [I-D.maglione-softwire-map-t-scenarios]
              Maglione, R., Dec, W., Kuarsingh, V., and E. Mallette,
              "Use cases for MAP-T",

draft-maglione-softwire-map-t-scenarios-02 (work in
               progress), June 2013.

   [I-D.xli-behave-divi]
               Bao, C., Li, X., Zhai, Y., and W. Shang, "dIVI: Dual-
               Stateless IPv4/IPv6 Translation", draft-xli-behave-divi-05
               (work in progress), June 2013.

   [RFC0792]   Postel, J., "Internet Control Message Protocol", STD 5,
               RFC 792, September 1981.

   [RFC0897]   Postel, J., "Domain name system implementation schedule",
               RFC 897, February 1984.

   [RFC2473]   Conta, A. and S. Deering, "Generic Packet Tunneling in
               IPv6 Specification", RFC 2473, December 1998.

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

   [RFC2827]   Ferguson, P. and D. Senie, "Network Ingress Filtering:
               Defeating Denial of Service Attacks which employ IP Source
               Address Spoofing", BCP 38, RFC 2827, May 2000.

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

   [RFC4443]   Conta, A., Deering, S., and M. Gupta, "Internet Control
               Message Protocol (ICMPv6) for the Internet Protocol
               Version 6 (IPv6) Specification", RFC 4443, March 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.

   [RFC4862]   Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless
               Address Autoconfiguration", RFC 4862, September 2007.

   [RFC4953]   Touch, J., "Defending TCP Against Spoofing Attacks",
               RFC 4953, July 2007.

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

   [RFC5961]   Ramaiah, A., Stewart, R., and M. Dalal, "Improving TCP's

                Robustness to Blind In-Window Attacks", RFC 5961,
                August 2010.

   [RFC6056]   Larsen, M. and F. Gont, "Recommendations for Transport-
                Protocol Port Randomization", BCP 156, RFC 6056,
                January 2011.

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

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

   [RFC6219]   Li, X., Bao, C., Chen, M., Zhang, H., and J. Wu, "The
                China Education and Research Network (CERNET) IVI
                Translation Design and Deployment for the IPv4/IPv6
                Coexistence and Transition", RFC 6219, May 2011.

   [RFC6269]   Ford, M., Boucadair, M., Durand, A., Levis, P., and P.
                Roberts, "Issues with IP Address Sharing", RFC 6269,
                June 2011.

   [RFC6324]   Nakibly, G. and F. Templin, "Routing Loop Attack Using
                IPv6 Automatic Tunnels: Problem Statement and Proposed
                Mitigations", RFC 6324, August 2011.

   [RFC6877]   Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT:
                Combination of Stateful and Stateless Translation",
                RFC 6877, April 2013.

Appendix A. Examples of MAP-T translation

Example 1 - Basic Mapping Rule:

Given the following MAP domain information and IPv6 end-user
prefix assigned to a MAP CE:

IPv6 prefix assigned to the end-user: 2001:db8:0012:3400::/56
Basic Mapping Rule: {2001:db8:0000::/40 (Rule IPv6 prefix),
    192.0.2.0/24 (Rule IPv4 prefix), 16 (Rule EA-bits length)}
PSID length: (16 - (32 - 24) = 8. (Sharing ratio of 256)
PSID offset: 6 (default)

A MAP node (CE or BR) can via the BMR, or equivalent FMR,
determine the IPv4 address and port-set as shown below:

EA bits offset: 40
IPv4 suffix bits (p) Length of IPv4 address (32) - IPv4 prefix
    length (24) = 8
IPv4 address 192.0.2.18 (0xc0000212)
PSID start: 40 + p = 40 + 8 = 48
PSID length (q): o - p = (End-user prefix len -
    rule IPv6 prefix len) - p = (56 - 40) - 8 = 8
PSID: 0x34

Available ports (63 ranges) : 1232-1235, 2256-2259, ... ,
                                    63696-63699, 64720-64723

The BMR information allows a MAP CE to determine (complete)
its IPv6 address within the indicated end-user IPv6 prefix.

IPv6 address of MAP CE: 2001:db8:0012:3400:0000:c000:0212:0034

   Example 2 - BR:

      Another example can be made of a hypothetical MAP-T BR,
      configured with the following FMR when receiving a packet
      with the following characteristics:

      IPv4 source address: 1.2.3.4 (0x01020304)
      TCP source port: 80
      IPv4 destination address: 192.0.2.18 (0xc0000212)
      TCP destination port: 1232

      Configured Forwarding Mapping Rule: {2001:db8::/40
         (Rule IPv6 prefix), 192.0.2.0/24 (Rule IPv4 prefix),
         16 (Rule EA-bits length)}

      MAP-T BR Prefix (DMR) 2001:db8:ffff::/64

      The above information allows the BR to derive as follows
      the mapped destination IPv6 address for the corresponding
      MAP-T CE, and also the source IPv6 address for
      the mapped IPv4 source address.

      IPv4 suffix bits (p) 32 - 24 = 8 (18 (0x12))
      PSID length: 8
      PSID: 0x34 (1232)

      The resulting IPv6 packet will have the following header fields:

      IPv6 source address 2001:db8:ffff:0:0001:0203:0400::
      IPv6 destination address: 2001:db8:0012:3400:0000:c000:0212:0034
      TCP source Port: 80
      TCP destination Port: 1232

Example 3- FMR:

    An IPv4 host behind a MAP-T CE (configured as per the previous
    examples) corresponding with an IPv4 host 1.2.3.4 will have its
    packets converted into IPv6 using the DMR configured on the MAP-T
    CE as follows:

    Default Mapping Rule used by MAP-T CE: {2001:db8:ffff::/64
    (Rule IPv6 prefix), 0.0.0.0/0 (Rule IPv4 prefix), null (BR IPv4
       address)}

    IPv4 source address (post NAT44 if present) 192.0.2.18
    IPv4 destination address: 1.2.3.4
    IPv4 source port (post NAT44 if present): 1232
    IPv4 destination port: 80
    IPv6 source address of MAP-T CE:
                        2001:db8:0012:3400:0000:c000:0212:0034
    IPv6 destination address: 2001:db8:ffff:0:0001:0203:0400::

  Example 4 - Rule with no embedded address bits and no address sharing

    End-user IPv6 prefix: 2001:db8:0012:3400::/56
    Basic Mapping Rule: {2001:db8:0012:3400::/56 (Rule IPv6 prefix),
       192.0.2.1/32 (Rule IPv4 prefix), 0 (Rule EA-bits length)}
    PSID length: 0 (Sharing ratio is 1)
    PSID offset: n/a

    A MAP node can via the BMR or equivalent FMR, determine
    the IPv4 address and port-set as shown below:

    EA bits offset: 0
    IPv4 suffix bits (p) Length of IPv4 address - IPv4 prefix
       length = 32 - 32 = 0
    IPv4 address 192.0.2.1 (0xc0000201)
    PSID start: 0
    PSID length: 0
    PSID: null

    The BMR information allows a MAP CE also to determine (complete)
    its full IPv6 address by combining the IPv6 prefix with the MAP
    interface identifier (that embeds the IPv4 address).

    IPv6 address of MAP CE: 2001:db8:0012:3400:0000:c000:0201:0000

   Example 5 - Rule with no embedded address bits and address sharing
   (sharing ratio 256)

      End-user IPv6 prefix: 2001:db8:0012:3400::/56
      Basic Mapping Rule: {2001:db8:0012:3400::/56 (Rule IPv6 prefix),
         192.0.2.1/32 (Rule IPv4 prefix), 0 (Rule EA-bits length)}
      PSID length: (16 - (32 - 24)) = 8. (Provisioned with DHCPv6.
                   Sharing ratio of 256.).
      PSID offset: 6 (default)
      PSID: 0x20 (Provisioned with DHCPv6)

      A MAP node can via the BMR determine the IPv4 address and port-set
      as shown below:

      EA bits offset: 0
      IPv4 suffix bits (p): Length of IPv4 address - IPv4 prefix
         length = 32 -32 = 0
      IPv4 address 192.0.2.1 (0xc0000201)
      PSID start: 0
      PSID length: 8
      PSID: 0x20

      Available ports (63 ranges) : 1536-1551, 2560-2575, ... ,
                                    64000-64015, 65024-65039

      The BMR information allows a MAP CE also to determine (complete)
      its full IPv6 address by combining the IPv6 prefix with the MAP
      interface identifier (that embeds the IPv4 address and PSID).

      IPv6 address of MAP CE: 2001:db8:0012:3400:0000:c000:0212:0034

      Note that the IPv4 address and PSID is not derived from the IPv6
      prefix assigned to the CE, but provisioned separately using for
      example MAP options in DHCPv6.

[Appendix B](#). Port mapping algorithm

   The driving principles and the mathematical expression of the mapping
   algorithm used by MAP can be found in [Appendix B](#) of
   [[I-D.ietf-softwire-map](#)]

Authors' Addresses

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

    Email: xing@cernet.edu.cn

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

    Email: congxiao@cernet.edu.cn

    Wojciech Dec (editor)
    Cisco Systems
    Haarlerbergpark Haarlerbergweg 13-19
    Amsterdam, NOORD-HOLLAND 1101 CH
    Netherlands

    Phone:
    Email: wdec@cisco.com

    Ole Troan
    Cisco Systems
    Oslo
    Norway

    Email: ot@cisco.com

    Satoru Matsushima
    SoftBank Telecom
    1-9-1 Higashi-Shinbashi, Munato-ku
    Tokyo
    Japan

    Email: satoru.matsushima@tm.softbank.co.jp

Tetsuya Murakami
IP Infusion
1188 East Arques Avenue
Sunnyvale
USA

Email: tetsuya@ipinfusion.com