Softwires Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 5, 2015

X. Li
C. Bao
CERNET Center/Tsinghua University
W. Dec, Ed.
O. Troan
Cisco Systems
S. Matsushima
SoftBank Telecom
T. Murakami
IP Infusion

December 2, 2014

Mapping of Address and Port using Translation (MAP-T)
draft-ietf-softwire-map-t-08

Abstract

This document specifies the "Mapping of Address and Port" stateless IPv6-IPv4 Network Address Translation (NAT64) based solution architecture for providing shared or non-shared IPv4 address connectivity to and across an IPv6 network.

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 June 5, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Conventions 3
3. Terminology 3
4. Architecture 5
5. Mapping Rules 7
5.1. Destinations outside the MAP domain 7
6. The IPv6 Interface Identifier 7
7. MAP-T Configuration 8
7.1. MAP CE 8
7.2. MAP BR 9
8. MAP-T Packet Forwarding 9
8.1. IPv4 to IPv6 at the CE 9
8.2. IPv6 to IPv4 at the CE 10
8.3. IPv6 to IPv4 at the BR 11
8.4. IPv4 to IPv6 at the BR 11
9. ICMP Handling 11
10. Fragmentation and Path MTU Discovery 12
10.1. Fragmentation in the MAP domain 12
10.2. Receiving IPv4 Fragments on the MAP domain borders . . . 12
10.3. Sending IPv4 fragments to the outside 13
11. NAT44 Considerations 13
12. Usage Considerations 13
12.1. EA-bit length 0 13
12.2. Mesh and Hub and spoke modes 13
12.3. Communication with IPv6 servers in the MAP-T domain . . 14
12.4. Compatibility with other NAT64 solutions 14
13. IANA Considerations 14
14. Security Considerations 14
15. Contributors 15
16. Acknowledgements 16
17. References 16
17.1. Normative References 16
17.2. Informative References 16
Appendix A. Examples of MAP-T translation 19
Appendix B. Port mapping algorithm 22
Authors' Addresses 22

1. Introduction

Experiences from initial service provider IPv6 network deployments, such as [RFC6219], indicate that successful transition to IPv6 can happen while supporting legacy IPv4 users without a full end-to-end dual IP stack deployment. However, due to public IPv4 address exhaustion this requires an IPv6 technology that supports IPv4 users utilizing shared IPv4 addressing, while also allowing the network operator to optimize their operations around IPv6 network practices. The use of double NAT64 translation based solutions is an optimal way to address these requirements, especially in combination with stateless translation techniques that minimize operational challenges outlined in [I-D.ietf-softwire-stateless-4v6-motivation].

The Mapping of Address and Port - Translation (MAP-T) architecture specified in this draft is such a double stateless NAT64 based solution. It builds on existing stateless NAT64 techniques specified in [RFC6145], along with the stateless algorithmic address & transport layer port mapping scheme defined in MAP-E [I-D.ietf-softwire-map]. The MAP-T solution differs from MAP-E in the use of IPv4-IPv6 translation, rather than encapsulation, as the form of IPv6 domain transport. The translation mode is considered advantageous in scenarios where the encapsulation overhead, or IPv6 operational practices (e.g. Use of IPv6 only servers, or reliance on IPv6 + protocol headers for traffic classification) rule out encapsulation. These scenarios are presented in [I-D.maglione-softwire-map-t-scenarios]

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-T
: Mapping of Address and Port by means of address Translation.

MAP Customer Edge (CE):
: A device functioning as a Customer Edge (CE) router in a MAP deployment. A typical MAP CE adopting MAP rules will serve a residential site with one WAN side IPv6 addressed interface, and one or more LAN side interfaces addressed using private IPv4 addressing.

MAP Border Relay (BR):	A MAP enabled router managed by the service provider at the edge of a MAP domain. A Border Relay (BR) 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 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 the MAP Rules communicated to a device, that are intended for determining the devices' IP+port mapping and forwarding operations. The MAP Rule set is interchangeably referred to in this document as a MAP Rule table or simply Rule table. Two specific types of rules, Basic Mapping Rule (BMR) and Forward Mapping Rule (FMR), are defined in Section 5 of [I-D.ietf-softwire-map]. The Default Mapping Rule (DMR) is defined in this document.
MAP Rule table:	See MAP Rule set.
MAP node:	A device that implements 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 privately addressed IPv4 users (N and M) connected by means of MAP-T CEs to an IPv6 network that is equipped with one or more MAP-T BR. CEs and BRs that share MAP configuration parameters, referred to as MAP rules, form a MAP-T Domain.

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

- A (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].

- A stateless NAT64 function [RFC6145] is extended to allow stateless mapping of IPv4 and transport layer port ranges to IPv6 address space.

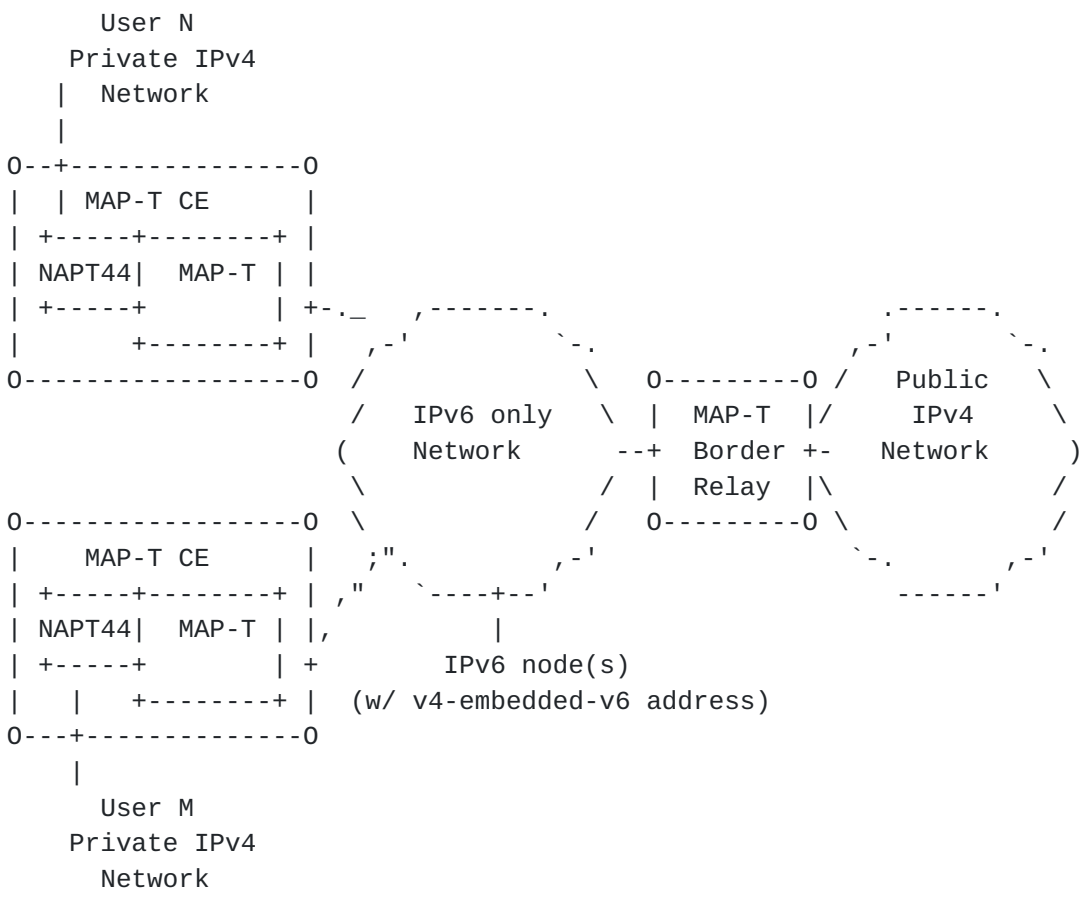


Figure 1: MAP-T Architecture

Each MAP-T CE is assigned with a regular IPv6 prefix from the operator's IPv6 network. This, in conjunction with MAP domain configuration settings and the use of the MAP procedures allows the computation of a MAP IPv6 address and a corresponding IPv4 address. To allow for IPv4 address sharing, the CE may also have be configured with a TCP/UDP port-range that is identified by means of a MAP Port Set Identifier (PSID) value. Each CE is responsible for forwarding traffic between a given user's private IPv4 address space and the MAP domain's IPv6 address space. The IPv4-IPv6 adaptation uses stateless NAT64, in conjunction with the MAP algorithm for address computation.

The MAP-T BR connects one or more MAP-T domains to external IPv4 networks using stateless NAT64 as extended by the MAP-T behaviour described in this document.

In contrast to MAP-E, NAT64 technology is used in the architecture for two purposes. Firstly, it is intended to diminish encapsulation overhead and allow IPv4 and IPv6 traffic to be treated as similarly as possible. Secondly, it is intended to allow IPv4-only nodes to

correspond directly with IPv6 nodes in the MAP-T domain that have IPv4 embedded IPv6 addresses as per [RFC6052]).

The MAP-T architecture is based on the following key properties i) algorithmic IPv4-IPv6 address mapping codified as MAP Rules covered in Section 5 ii) A MAP IPv6 address identifier, described in Section 6 iii) MAP-T IPv4-IPv6 forwarding behavior described in Section 8.

5. Mapping Rules

The MAP-T algorithmic mapping rules are identical to those in Section 5 of the MAP-E specification [I-D.ietf-softwire-map], with the following exception. The forwarding of traffic to and from IPv4 destinations outside a MAP-T domain is to be performed as described here under, instead of Section 5.4 of the MAP-E specification.

5.1. Destinations outside the MAP domain

IPv4 traffic sent by MAP nodes that are all within one MAP domain is translated to IPv6, with the sender's MAP IPv6 address, derived via the Basic Mapping Rule (BMR), as the IPv6 source address and the recipient's MAP IPv6 address, derived via the Forward Mapping Rule (FMR), as the IPv6 destination address.

IPv4 addressed destinations outside of the MAP domain are represented by means of IPv4-Embedded IPv6 address as per [RFC6052], using the BR's IPv6 prefix. For a CE sending traffic to any such destination, the source address of the IPv6 packet will be that of the CE's MAP IPv6 address, and the destination IPv6 address will be the destination IPv4-embedded-IPv6 address. This address mapping is termed as following the MAP-T Default Mapping Rule (DMR) and is defined in terms of the IPv6 prefix advertised by one or more BRs, which provide external connectivity. A typical MAP-T CE will install an IPv4 default route using this rule. A BR will use this rule when translating all outside IPv4 source addresses to the IPv6 MAP domain.

The DMR IPv6 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.

6. The IPv6 Interface Identifier

The Interface identifier format of a MAP-T node is the same as described in section 6 of [I-D.ietf-softwire-map]. For convenience this is cited below:

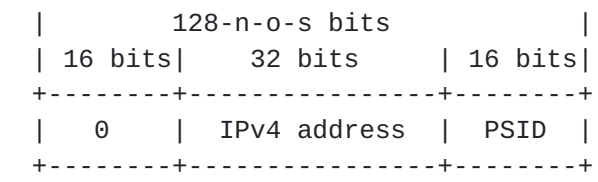


Figure 2

In the case of an IPv4 prefix, the IPv4 address field is right-padded with zeros up to 32 bits. The 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.

7. MAP-T Configuration

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

- The Basic Mapping Rule and optionally the Forwarding Mapping Rules, including the Rule IPv6 prefix, Rule IPv4 prefix, and Length of Embedded Address bits

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

- Use of IPv4-IPv6 Translation (MAP-T)

- The BR's IPv6 prefix used in the DMR

7.1. MAP CE

For a given MAP domain, the MAP configuration parameters are the same across all CEs within that domain. These values may be conveyed and configured on the CEs using a variety of methods, including; DHCPv6, Broadband Forum's "TR-69" Residential Gateway management interface, Netconf, or manual configuration. 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 data model described by this option for consistency and convenience of implementation on CEs that support multiple configuration methods.

Besides the MAP configuration parameters, a CE requires an IPv6 prefix to be assigned to the CE. This End-user IPv6 prefix is configured as part of obtaining IPv6 Internet access, and is acquired using standard IPv6 means applicable in the network where the CE is located.

The MAP provisioning parameters, and hence the IPv4 service itself, are tied to the End-user IPv6 prefix; thus, the MAP service is also tied to this in terms of authorization, accounting, etc.

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. Each MAP domain requires a distinct End-user IPv6 prefix.

7.2. 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 for forwarding to/from the domain at any time.

Since MAP uses provider address space, no specific IPv6 or IPv4 routes need to be advertised externally outside the service provider's network for MAP to operate. However, the BR prefix needs to be advertised in the service provider's IGP.

8. MAP-T Packet Forwarding

The end-to-end packet flow in MAP-T involves an IPv4 or IPv6 packet being forwarded by a CE of BR in one of two directions for each such case. This section presents a conceptual view of the operations involved in such forwarding.

8.1. IPv4 to IPv6 at the CE

A MAP-T CE receiving IPv4 packets SHOULD perform NAPT NAT44 processing, and create any necessary NAPT44 bindings. The source address and source port-range of packets resulting from the NAPT44 processing MUST correspond to the source IPv4 address and source transport port-range assigned to the CE by means of the MAP Basic Mapping Rule (BMR).

The IPv4 packet is subject to a longest IPv4 destination 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 NAT64 Section 5.1. Packets that are matched to (optional) Forward Mapping Rules (FMRs) are subject to the stateless NAT64 operation using the FMR parameters Section 5 for the MAP algorithm. In all cases the CE's MAP IPv6 address Section 6 is used as a source address.

A MAP-T CE MUST support a Default Mapping Rule and SHOULD support one or more Forward Mapping Rules.

8.2. 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 IPv6 addresses, and with source addresses matching the IPv6 map-rule prefixes of a DMR or FMR, are processed by the MAP-T CE, with the DMR or FMR being selected based on a longest match. The CE MUST check that each MAP-T received packet's destination transport-layer destination port number is in the range allowed for by the CE's MAP BMR configuration. The CE MUST silently drop any non conforming packet and an appropriate counter incremented. When receiving a packet whose source IP address longest matches an FMR prefix, the CE MUST perform a check of consistency of the source address against the allowed values as per the derived allocated source port-range. If the source port number of a packet is found to be outside the allocated range, 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 compute an IPv4 source and destination addresses. The IPv4 destination address is computed by extracting relevant information from the IPv6 destination and the information stored in the BMR as per Section 5. The IPv4 source address is formed by classifying a packet's source as longest matching a DMR or FMR rule prefix, and then using the respective rule parameters for the NAT64 operation.

The resulting IPv4 packet is then forwarded to the CE's NAPT NAPT44 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 NAPT44 function is not enabled, by virtue of MAP configuration, the traffic from the stateless NAT64 function is directly forwarded according to the CE's IPv4 rules.

8.3. IPv6 to IPv4 at the BR

A MAP-T BR receiving an IPv6 packet MUST select a matching MAP rule based on a longest address match of the packet's source address against the MAP Rules present on the BR. In combination with the Port-Set-Id derived from 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 packet's source port number is found to be outside the range allowed, the BR MUST drop the packet and increment a counter to indicate the event. The BR SHOULD also 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 packet is then passed to regular IPv4 forwarding.

8.4. 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 select the FMR and DMR rules. The MAP-T BR MUST then compute and apply the IPv6 destination addresses from the IPv4 destination address and port as per the selected FMR. The MAP-T BR MUST also compute and apply the IPv6 source addresses from the IPv4 source address as per Section 5.1 (i.e. Using the IPv4 source and the BR's IPv6 prefix it forms an IPv6 embedded IPv4 address). Throughout the generic IPv4 to IPv6 header translation 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/from MAP-T domains that are defined without IPv4 address sharing is the same as that of stateless NAT64 IPv4/IPv6 translation.

9. ICMP Handling

MAP-T CEs and BRs MUST follow ICMP/ICMPv6 translation as per [RFC6145], however additional behavior is also required due to the presence of NAPT44. 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.

A MAP-T BR receiving an IPv4 ICMP packet , which contains an ID field that is bound for a shared address in the MAP-T domain, 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.

10. 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 can use all three mechanisms to deal with different cases.

Note: The NAT64 [RFC6145]mechanism is not lossless. When IPv4 originated communication traverses across a double NAT64 function (a.k.a. NAT464), any IPv4 originated ICMP-independent PathMTU Discovery, as specified in [RFC 4821], ceases to be entirely reliable. This is because the[RFC4821] defined DF=1/MF=1 combination, following a double NAT64 translation, results in DF=0/ MF=1.

10.1. Fragmentation in the MAP domain

Translating an IPv4 packet to carry it across the MAP domain will increase its size typically by 20 bytes. The MTU in the MAP domain should be well managed and the IPv6 MTU on the CE WAN side interface SHOULD be configured so that no fragmentation occurs within the boundary of the MAP domain.

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

10.2. 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 SHOULD reassemble the packet before sending the packet onto the MAP domain. 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 such a caching algorithm is outside the scope of this document.

10.3. Sending IPv4 fragments to the outside

Two IPv4 hosts behind two different MAP CE's with the same IPv4 address sending fragments to an IPv4 destination host outside the domain may happen to 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 can be used similarly to port ranges. Thus, a MAP CE SHOULD rewrite the IPv4 fragmentation identifier to a value equivalent to a port of its allocated port-set.

11. NAT44 Considerations

The NAT44 implemented in the MAP CE SHOULD conform with the behavior and best current practice documented in [RFC4787], [RFC5508], and [RFC5382]. In MAP address sharing mode (determined by the MAP domain /rule configuration parameters) the operation of the NAT44 MUST be restricted to the available port numbers derived via the basic mapping rule.

12. Usage Considerations

12.1. EA-bit length 0

The MAP solution supports use and configuration of domains where a BMR expresses an EA-bit length of 0. This results in independence between the 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 end-user prefix. The BR would be configured with an FMR for each such CPE, where the rule would uniquely associate the IPv4 address + optional PSID and the IPv6 prefix of that given CE.

12.2. 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.

12.3. 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 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 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.

12.4. Compatibility with other NAT64 solutions

The MAP-T CEs NAT64 function is by default compatible for use with [RFC6146] stateful NAT64 devices that are placed in the operator's network. In such a case the MAP-T CE's DMR prefix is configured to correspond to the NAT64 device 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].

13. IANA Considerations

This specification does not require any IANA actions.

14. 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 support in both clients and servers.

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 Flooding 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 node's ICMP computing 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.

15. Contributors

The following individuals authored major contributions to this document, and made the document possible:

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.comc

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

16. 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 also like to thank Mohamed Boucadair, Guillaume Gottard, Dan Wing, Jan Zorz, Nejc Scoberne, Tina Tsou, Gang Chen, Maoke Chen, Xiaohong Deng, Jouni Korhonen, Tomasz Mrugalski, Jacni Qin, Chunfa Sun, Qiong Sun, Leaf Yeh, Andrew Yourtchenko, Roberta Maglione and Hongyu Chen for their review and comments.

17. References

17.1. Normative References

[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-12 (work in progress), November 2014.

[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.

17.2. 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-dhcp]
Mrugalski, T., Troan, O., Farrer, I., Perreault, S., Dec, W., Bao, C., leaf.yeh.sdo@gmail.com, l., and X. Deng, "DHCPv6 Options for configuration of Softwire Address and Port Mapped Clients", draft-ietf-softwire-map-dhcp-11 (work in progress), November 2014.

[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.maglione-softwire-map-t-scenarios]
Maglione, R., Dec, W., Leung, I., and E. Mallette, "Use cases for MAP-T", draft-maglione-softwire-map-t-scenarios-05 (work in progress), October 2014.

[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.

[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.

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

[RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", RFC 4821, March 2007.

[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.

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

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

[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:

End-user IPv6 prefix:	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 MAP-T BR,
configured with the following FMR when receiving a packet
with the following characteristics:

```
IPv4 source address:       10.2.3.4 (0x0a020304)
TCP source port:           80
IPv4 destination address:  192.0.2.18 (0xc0000212)
TCP destination port:      1232

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:  0               x34 (1232)
```

The resulting IPv6 packet will have the following header fields:

```
IPv6 source address:      2001:db8:ffff:0:000a: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 10.2.3.4 will have its packets converted into IPv6 using the DMR configured on the MAP-T CE as follows:

```
Default Mapping Rule:         {2001:db8:ffff::/64 (Rule IPv6 prefix),
                               0.0.0.0/0 (Rule IPv4 prefix)}

IPv4 source address:          192.0.2.18
IPv4 destination address:     10.2.3.4
IPv4 source port:             1232
IPv4 destination port:        80
MAP-T CE IPv6 source address: 2001:db8:0012:3400:0000:c000:0212:0034
IPv6 destination address:     2001:db8:ffff:0:000a: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:            192.0.2.18 (0xc0000212)
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),
                         192.0.2.18/32 (Rule IPv4 prefix),
                         0 (Rule EA-bits length)}
PSID length:             (16 - (32 - 24)) = 8. Sharing ratio of 256.
                         Provisioned with DHCPv6.
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             192.0.2.18 (0xc0000212)
PSID start:              0
PSID length:             8
PSID:                    0x34

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

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

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