Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 7, 2009

M. Boucadair, Ed.
P. Levis
J-L. Grimault
A. Villefranque
M. Kassi-Lahlou
France Telecom

March 6, 2009

Flexible IPv6 Migration Scenarios in the Context of IPv4 Address Shortage
draft-boucadair-behave-ipv6-portrange-01

Status of this Memo

Copyright Notice

Abstract

This memo presents a solution to solve IPv4 address shortage and ease IPv4-IPv6 interworking. The document presents a set of incremental steps for the deployment of IPv6 as a means to solve IPv4 address exhaustion. Stateless IPv4/IPv6 address mapping functions are introduced and IPv4-IPv6 interconnection scenarios presented. This memo advocates for a more proactive approach for the deployment of IPv6 into operational networks.

This document provides both the specification of the solution and deployment scenarios together with migrations paths.

Requirements Language

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].

Table of Contents

1. Introduction

1.1. IPv4 Address Exhaustion

It is commonly agreed by the Internet community that the exhaustion of public IPv4 addresses is an ineluctable fact. Regular alarms and reports have been emitted by the IETF particularly by the reports presented within the GROW working group (Global Routing Operations Working Group) meetings.

G. Huston introduced an extrapolation model to forecast the exhaustion date of IPv4 addresses managed by IANA. This effort indicates that if the current tendency of consumption continues at the same pace, IPv4 addresses exhaustion of IANA's pool would occur in 2011, while RIRs'pool would be exhausted in mid-2012. The state of the current consumption of public IPv4 addresses is daily updated and is available at this URL: http://www.potaroo.net/tools/ipv4/index.html.

1.2. To what extent does IPv6 solve the problem?

In this context, the community was mobilized in the past to adopt a promising solution (in particular with the definition of IPv6). IPv6 has been introduced for several years as the next version of the IP protocol. This new version offers an abundance of IP addresses as well as several enhancements compared to IPv4. IPv6 specifications are mature and current work within the IETF is related to operational aspects. Despite this effort, IPv6 is not globally activated by service providers for both financial and strategic reasons.

However, even if a service provider activates IPv6, it will be confronted with the problem to ensure a global connectivity towards nowadays Internet v4. Mechanisms such as NAT-PT (Network Address Translation Protocol Translation) were introduced to ensure the interconnection between two heterogeneous realms (i.e. IPv4/IPv6) and to ensure a continuity of IP communications (i.e. End-to-end). These solutions are statefull and are not suitable to interconnect an IPv6 domain with a dominant Internet which is IPv4-only. Further work should be undertaken with IETF to elaborate lightweight and hopefully stateless solution to ease IPv4-IPv6 interworking. Moreover, Service providers should adopt clear strategies so as to ease the adoption of IPv6 and to decrease the complexity related to IPv4-IPv6 interworking which is one of the critical issues to be taken into account when designing service platforms.

Last, it is worth to mention that migrating to IPv6 is a service provider issue and not the one of its customers. The ultimate requirement of the customers (mainly residential and mass market) is

to benefit from a global IP connectivity. How this connectivity is engineered and put into effect is of the business of the IP connectivity service providers. Of course, some corporate customers would specify the nature of their IP connectivity and reduce the interconnection engineering complexity of their interconnection nodes with the domain of their IP connectivity service provider(s). From this standpoint, service providers should be more proactive in order to avoid a crisis scenario where no IP addresses are available to be assigned to their customers.

1.3. Towards a proactive approach

The introduction of IPv6 into public networks becomes a reality. Several Internet providers have enabled IPv6 in their routers and launched therefore their IPv6 migration operations. The portion of the IPv6-enabled routers differs between SPs. The current trade is that operators offer dual connectivity to their customers, i.e. IPv4 and IPv6 access. IPv4 connectivity usage should be gradually decreased in favour of IPv6 one. This convergence phase towards a pure IPv6 connectivity will take several years depending on the policies adopted by service providers. For operators that adopt an aggressive position with regards to the activation of IPv6, this transition phase could be small compared to passive operators. Nevertheless the overall Internet IPv6 coverage will be long. This is due mainly to the significant number of involved actors to be convinced for a full migration towards IPv6, the significant number of existing ASes (more than 30000), etc. Moreover, customers do not have any reason to modify their local architecture (e.g. a given organisation does not have any motivations to migrate its FTP or HTTP servers towards IPv6). Operators must expect a long work of accompaniment for the migration towards IPv6. The final migration towards IPv6 would take several years (at least 10 years).

This migration to IPv6 should be incremental and not implemented in one shot. For these reasons, service providers should elaborate migration scenarios so as to achieve a transparent migration. This transparency is required because end-users should not be aware on the underlying technology used to deliver their subscribed services and the complexity related to service engineering should be hidden to end-users. Furthermore, service providers should use means to prioritise IPv6 traffic and the invocation of IPv6 transfer capabilities without relying on end-users behaviour. IPv6 transfer capabilities should be exploited and not considered as dormant ones. If no proactive means/procedures are adopted, the ratio of IPv6 traffic will depend on the behaviour of end-users and also on available IPv6 services.

Furthermore, in the perspective of IPv6 migration, the maintenance

and the operating of dual connectivity infrastructure would therefore be required for a long period. This option is not to be encouraged within service providers since it does not optimise both OPEX/CAPEX. Both technical skills should be maintained within each individual organisation. As an alternative, this document proposes a proactive and incremental deployment approach which consists at:

- Activation of IPv6 and port range IPv4 solution at the same time. Port-restricted devices are provisioned with an IPv6 prefix, a shared IPv4 public address and a port range.

- Activation of stateless functions and use of IPv6 to carry IPv4 traffic from/to port-restricted devices.

- Migration to IPv6-only core network.

- Maintain stateless IPv4-IPv6 interworking functions at interconnection segment to not alter intra-domain paths.

1.4. Contribution of this draft

This memo defines several solutions to solve the IPv4 address shortage problem and to migrate to IPv6 without requiring stateful nodes. The draft proposes also several migration paths. This target IPv6 deployment is a long term objective and can be reached incrementally through one or several intermediate steps. These intermediate steps perimeters differ from a service provider to another one depending on the service opportunities targeted by enabling IPv6-related capabilities and also on the level of risks on the already running services. Risks on existing services should be assessed. Careful or aggressive position may be adopted by each service provider. Service providers are free to deploy the step/the migration path suitable to their context and objectives, etc. This document only sketches scenarios and interconnection configurations. Voluntarily, no frozen architecture is described. Several options are supported.

This document provides:

o solution specification

o and deployment scenarios together with migrations paths.

1.5. Positioning this Draft

This draft proposes a solution to solve both IPv4 address exhaustion and IPv4-IPv6 interconnection. Unlike [I-D.ietf-softwire-dual-stack-lite], no additional NAT is required to

be deployed at the service provider's network. Both encapsulation and translation modes are presented in this memo.

A set of issues related to IPv4 Internet access in the context of public IPv4 address exhaustion are identified and described in [I-D.levis-behave-ipv4-shortage-framework]. To what extent the proposed solution handles those issues will be discussed in the next version of this document.

Alternative address mapping proposals may be found at [I-D.despres-sam].

2. Terminology

Within the context of this draft, the following terminology is used:

- Access segment: This segment encloses both IP access and backhaul network. Within this document, this access segment encompass an access PoP.

- Core network: Denotes a set of IP networking capabilities and resources which are between the interconnection and the access segments.

- Interconnection segment: Includes all nodes and resources which are deployed at the border of a given AS (Autonomous System) a la BGP (Border Gateway Protocol).

- PRR: Stands for Port Range Router. This function is responsible to handle a port-based routing. This function may retrieve the port value and use it to determine which routing action is to be executed or use it together with the destination IP address to build an IPv6 address.

- Interconnection PRR (i-PRR): A PRR which is deployed at interconnection segment.

- Access PRR (a-PRR): A PRR which is deployed at access segment.

- SMAP: Stands for Stateless A+P Mapping. This function is responsible to encapsulate (Resp. de-encapsulate), in a stateless scheme, IPv4 packets in (Resp. from) IPv6 ones. A SMAP function may be hosted in a PRR, end-user device, etc.

- Port-restricted device (PRD): A device which is able to constrain its source port number to be within a given Port Range. A port restricted device may be of several types such as:

- CPE (Customer Premise Equipment)/HGW (Home Gateway)
- PDA (Personal Digital Assistant)
- Mobile terminal

3. Reminder of the Port Range Solution

This section is a reminder of the solution presented in [I-D.boucadair-port-range]. For more details about the solution, the reader is invited to refer to [I-D.boucadair-port-range].

The main principle of the solution is to assign the same IP address (called Primary IP Address) to several end-users' devices and to constraint the source port numbers to be used by each device. In addition to the assigned IP address to access IP connectivity services, an additional parameter, called Port Range, is also assigned to the customer's device. For outbound communications, a given port restricted device proceeds to its classical operations except the constraint to control the source port number assignment so as to be within the Port Range. The traffic is then routed without any modification inside the Service Provider's domain and delivered to its final destination. For inbound communications, the traffic is trapped by a dedicated function called: Port Range Router (PRR). This function may be embedded in current routers or hosted by new nodes to be integrated in the IP infrastructure of these service providers. Appropriate routing tuning policies are enforced so as to drive the inbound traffic to cross a PRR. Each PRR correlate the Primary IP Address and information about the allowed port values with a specific identifier called: routing identifier (e.g. Private IPv4 address, IPv6 address, point-to-point link identifier, etc). This routing identifier is used to route the packets to the suitable device among all those having the same IP address.

Port-restricted devices are provisioned with port range to be used, especially the Port Mask to be applied when selecting a port value as a source port. A Port Mask defines a set of ports that all have in common a subset of pre-positioned bits. This set of ports is also called Port Range. Two port numbers are said to belong to the same Port Range if and only if, they have the same Port Mask. A Port Mask is composed of a Port Range Value and a Port Range Mask.

o The Port Range Value indicates the value of the significant bits of the Port Mask. The Port Range Value is coded as follows:

  - The significant bits may take a value of 0 or 1.

- All the other bits (non significant ones) are set to 0.

o The Port Range Mask indicates, by the bit(s) set to 1, the position of the significant bits of the Port Range Value.

```
    0                   1
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0| Port Range Mask
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     | | |
     | | | (3 significant bits)
     v v v
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0| Port Range Value
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |0 0 1 x x x x x x x x x x x x x| Usable ports (x may take a value of 0 or
1).
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Example of Port Range Mask and Port Range Value

An example of port range is provided in Figure 1. Ports belonging to this port range must have the 1st bit (Resp. the 2nd and 3rd), from the left, set to 0 (Resp. 0 and 1). The Port Mask is represented as: 001xxxxxxxxxxxxx.

4. IPv6-IPv4 Address Mapping Formalism

This section discusses issues related to the building of an IPv6 prefix or IPv6 address using IPv4-related information:

4.1. IPv6 Prefix: Another IPv4-mapped Prefix Scheme

[I-D.boucadair-port-range] proposes to assign the same public IPv4 address together with a port range to several devices. In order to discriminate those devices, an additional identifier called, routing identifier, is required. This identifier may be a secondary IPv4 address, PPP session identifier, etc. This document assumes that this identifier is an IPv6 address. This prefix is built using IPv4-related information as illustrated in Figure 2.

```
   +-------------------------+-----------+----------+
   |         WKIPv6          |  @IPv4    |   PRM    |
   +-------------------------+-----------+----------+
                                               Max.
   <-----------n bits------> < 32 bits> <16 bits >
   <-----------------64 bits max.---------------->
```

Figure 2: IPv6 prefix enclosing an IPv4 address and a port range

1. The length of this prefix is recommended to be less than 64 bits;

2. WKIPv6: is a sub-prefix belonging to the service provider or well-known prefix allocated by IANA for this service. The length of this field is variable (may be different from a service provider to another if not allocated by IANA). The WKIPv6 prefix used to build an IPv4-mapped IPv6 prefix may not be the same as the one used to execute the IPv4-to-IPv6 mapping function introduced in Section 4.2.

3. @IPv4 field encloses the shared IPv4 address. The length of this field is 32 bits;

4. PRM field includes the value of the significant bits of the Port Range. The maximum length of this field is 16 bits. But, in deployment scenarios this field may be 3, 4 or 5 bits. If n bits are used to build the PRM, the same IPv4 address may be shared between 2^n port-restricted devices.

For illustration purposes two examples are provided below.

Let suppose that a service provider dedicates the 2a01:c0a8::/29 to build an IPv4-inferred IPv6 prefix. In this example, we suppose that 8 port restricted devices share the same public address 193.51.145.206 owing to a port range mask with three significant bits (i.e. the three first bit are used to build the port mask. The remaining 13 bits may take a 0 or 1 value), yielding 8192 (2^16/2^3) possible ports per each port-restricted device. The corresponding IPv6Pref prefixes for these 8 port-restricted devices are thus the following ones:

- 1st port-restricted device (Port Mask: 000xxxxxxxxxxxxx):

```
     IPv6Pref = 2a01:c0a 1  11000001001100111001000111001110 000 ::
                            --------193.51.145.206---------- PRM

     IPv6Pref = 2a01:c0aE:099C:8E70::/64
```

- 2nd port-restricted device (Port Mask: 001xxxxxxxxxxxxx):

```
     IPv6Pref = 2a01:c0a 1  11000001001100111001000111001110 001 ::
                            --------193.51.145.206---------- PRM

     IPv6Pref = 2a01:c0aE:099C:8E71::/64
```

- ...

- 8th port-restricted device (Port Mask: 111xxxxxxxxxxxxx):

```
     IPv6Pref = 2a01:c0a 1  11000001001100111001000111001110 111 ::
                            --------193.51.145.206---------- PRM

     IPv6Pref = 2a01:c0aE:099C:8E77::/64
```

In this second example, let suppose that the service provider dedicates the 2a01:c::/20 prefix to build an IPv4-mapped IPv6 prefix. If we consider that 193.51.145.206 address is shared between 16 (2^4, 4 bits are used as the significant bits of the port range) port-restricted devices. The 16 port-restricted devices sharing that address have the following IPv6Pref prefixes (only the first prefix is represented below):

- 1st port-restricted device (Port Mask: 0000xxxxxxxxxxxx):

```
        IPv6Pref = 2a01:c 11000001001100111001000111001110 0000 ::
                          --------193.51.145.206---------- PRM-

        IPv6Pref = 2a01:cc13:391c:e0::/56
```

- ...

4.2. IPv4 to IPv6 Address Mapping Function

4.2.1. Overview

Within this memo, IPv4-to-IPv6 address mapping function denotes a function which uses IPv4-related information, as conveyed in a received IPv4 packet, to generate IPv6 one. This function generates an IPv6 address which builds as illustrated in Figure 3.

- WKIPv6 is configured by the service provider.

- Then, the next 32 bits are set to the value of the destination IPv4 address;

- The next 16 bits are set to the value of the destination port number;

- The remaining bits are then set to zeros.

```
+----------------------------------------------------------------------...+
|       WKIPv6             |Dest. IPv4    |Dest.  |          0:0:0:0      |
|                          |address       |port   |                       |
+----------------------------------------------------------------------...+
```

Figure 3: WKIPv6A Address Format

4.2.2. Example

Let suppose that a given device is provided with the WKIPv6 prefix equal to 2a01:c0a8::/29. Then the corresponding IPv6 address, using the IPv4-to-IPv6 address mapping function, to the IPv4 address equal to 193.51.145.206 and the port number equal to 19039 (0100101001011111) is the following:

```
IPv6PrefA=2a01:c0a 1 11000001001100111001000111001110 0100101001011111 ::
                     --------193.51.145.206---------  -----port-------

IPv6PrefA=2a01:c0aE:099C:8E72:52F8::/128
```

This IPv6 address falls in the IPv6 prefix of the second port-restricted device (port range 001) as listed in the previous section.

4.3. IPv6 to IPv4 Address Mapping Function

4.3.1. Overview

Unlike the previous function, the IPv6-to-IPv4 address mapping functions generates an IPv4 address together with a port number from the header and the transport part of a received IPv6 packet as follows:

- The destination IPv4 address corresponds to the 32 bits which follow a per-configured Provider prefix;

- Destination port number is equal to the one of the received IPv6 packet.

4.3.2. Example

Let suppose that the WKIPv6 prefix equal to 2a01:c0a8::/29 is used. Then the corresponding IPv4 address, resulting from the IPv6-to-IPv4 address mapping function applied to the address 2a01:c0aE:099C:8E71: A5F8::/128 is 193.51.145.206 since:

```
2a01:c0aE:099C:8E71:A5F8 =

2a01:c0a 1  11000001001100111001000111001110 0100101001011111 ::
            --------193.51.145.206---------  -----port-------
```

5. Stateless A+P Mapping Function

5.1. Stateless A+P Mapping gateway (SMAP) Function description

Stateless A+P Mapping gateway (SMAP) consists in two basic functions as described in Figure 4.

1. SMAP encapsulates an IPv4 packet, destined to a shared IPv4 address, in IPv6 one. The IPv6 source address is constructed (see Section 4.2) from the IPv4 source address and port number plus the IPv6 prefix which has been provisioned to the node performing the SMAP function. The destination IPv6 address is constructed using the shared IPv4 destination address and port number plus the IPv6 prefix which has been provisioned to the SMAP function and which is dedicated to IPv4 destination addresses.

2. SMAP extracts IPv4 incoming packets from IPv6 incoming ones which have IPv6 source addresses belonging to the prefix of the node performing the SMAP function. Extracted IPv4 packets are then

forwarded to the point identified by the IPv4 destination address and port number.

```
                           +-------------------+
                           |                   |----IPv6---\
               ----IPv4---\|                   |----IPv4---\\
               -----------/|                   |-----------//
                           |                   |-----------/
                           |       SMAP        |
                           |                   | /--IPv6-----
               /---IPv4----|                   |//---IPv4----
               \-----------|                   |\\-----------
                           |                   | \-----------
                           +-------------------+
```

Figure 4: Stateless A+P Mapping Gateway Function

A SMAP-enabled node will perform the stateless 6/4 mapping function for all public shared IPv4 addresses for which it was designated as a stateless 6/4 mapping gateway.

To perform stateless 6/4 mapping function a SMAP gateway must:

o be provided with an IPv6 prefix (i.e. WKIPv6). The SMAP gateway uses this prefix to construct IPv6 source addresses for all IPv4 shared addresses for which it was designated as a SMAP gateway. The IPv6 prefix may be provisioned statically or dynamically (e.g. DHCP)

o be able to know the IPv6 prefix of the node serving as another SMAP gateway for IPv4 destination addresses. This prefix may be known in various ways:

   * Default or Well known prefix which was provisioned statically or dynamically;

   * Retained at the reception of incoming IPv4-in-IPv6 encapsulated packets;

   * Discovered at the communication starting thanks to mechanisms as DNS resolution for example.

When the SMAP-enabled node receives IPv4 packets with IPv4 source addresses for which it was not designated as a SMAP gateway, it will not perform stateless 6/4 mapping function for those packets. Those packets will be handled in a classical way (i.e. forwarded, dropped

or locally processed).

When the SMAP-enabled node receives IPv6 packets with IPv6 addresses which do not match with its IPv6 prefix, it will not perform the stateless 6/4 mapping function for those packets. Those packets will be handled in a classical way (i.e. forwarded, dropped or locally processed).

5.2. Implementation modes

Stateless mapping function may be achieved in two main modes. Those modes consist in mapping the traffic only in one direction or in the two directions as described below.

5.2.1. SMAP to route incoming traffic destined to a shared IPv4 address

IPv4 traffic with shared IPv4 source addresses are forwarded by the node A without performing stateless mapping function. This traffic will reach its destination thanks to a classical routing. In the opposite direction, the traffic sent by the destination has to pass by the node B which performs the stateless mapping function (encapsulating in IPv6 packets) before forwarding to the node A. The node A performs the stateless mapping function (extract IP v4 packets) before forwarding IPv4 packets to the points identified by the IPv4 destination addresses and port number. In this case, both IPv4 and IPv6 traffic are routed in the network between the nodes A and B.

```
                  +----------+
         --IPv4---|----------|------------IPv4--------------------\
         ---------|----------|------------------------------------/
                  |          |
                  | +------+ |               +------+
                  | |      | | /----IPv6-----|      |
         /--IPv4--| | SMAP | |//---IPv4----  | SMAP |/---IPv4----
         \--------| |      | |\\-----------  |      |\-----------
                  | |      | | \-------------|      |
                  | +------+ |               +------+
                  |          |
                  +----------+
                     node A                    node B
```

Figure 5: First Configuration

5.2.2. No IPv4 capabilities are used anymore between two SMAP-enabled nodes

In this configuration, the node A performs the stateless mapping function on the received IPv4 traffic (encapsulated in IPv6 packets) before forwarding to the node B. The node B performs the stateless mapping function on the received IPv6 traffics (extracting IPv4 packets) before forwarding the IPv4 traffic to the destination identified by the IPv4 destination address and port number. In the opposite direction and as previously, the node B performs the stateless mapping function on the received IPv4 traffics (encapsulating in IPv6 packets) before forwarding to the node A. The node A performs the stateless mapping function on the received IPv6 traffic (extracting IPv4 packets) before forwarding the IPv4 traffic to the point identified by the IPv4 destination address and port number. In this case, only IPv6 traffic is managed in the network segment between the nodes A and B.

```
                       +------+             +------+
                       |      |----IPv6---\ |      |
           ----IPv4---\|      |----IPv4---\\|      |----IPv4---\
           -----------/|      |-----------//|      |-----------/
                       |      |-----------/ |      |
                       | SMAP |             | SMAP |
                       |      | /----IPv6---|      |
           /---IPv4----|      |//---IPv4----|      |/---IPv4----
           \-----------|      |\\-----------|      |\-----------
                       |      | \-----------|      |
                       +------+             +------+
                        node A               node B
```

Figure 6: Second Configuration

5.3. Deployment Scenarios

Several deployment scenarios of the SMAP function may be envisaged in the context of Port Range based solutions:

- A SMAP function is embedded in a port-restricted device. Other SMAP-enabled nodes are deployed in the boundaries between IPv6-enabled realms and IPv4 ones. This scenario may be particularly deployed for intra-domain communications so as to interconnect heterogeneous realms (i.e. IPv6/IPv4) within the same AS.

- A SMAP function is embedded in a port-restricted device. Other SMAP-enabled nodes are deployed in the interconnection segment

(with adjacent IPv4-only ones) of a given AS. This deployment scenario is more suitable for service providers targeting the deployment of IPv6 since it eases the migration to full IPv6. Core nodes are not required to activate anymore both IPv4 and IPv6 transfer capabilities.

Other considerations regarding the interconnection of SMAP-enabled domains should be elaborated. The following provides a non exhaustive list of interconnection schemes.

o The interconnection of two domains implementing the SMAP function may be deployed via IPv4 Internet (Figure 7): This means that IPv4 packets encapsulated in IPv6 one are transferred using IPv6 until reaching the first SMAP-node. Then these packets are de-encapsulated and are forwarded using IPv4 transfer capabilities. A remote SMAP-enabled node will receive those packets and proceeds to an IPv4-in-IPv6 encapsulation. These packets are then routed normally until reaching the port-restricted devices which de-encapsulates the packets.

```
+------+           +------+   +--------+   +------+           +------+
|      |--IPv6---\ |      |   |        |   |      |---IPv6--\ |      |
|      |--IPv4---\\|      |---|-IPv4---|--\|      |---IPv4--\\|      |
|      |---------//|      |---|--------|--/|      |---------//|      |
|      |---------/ |      |   |Internet|   |      |---------/ |      |
| SMAP |           | SMAP |   |  IPv4  |   | SMAP |           | SMAP |
|      | /---IPv6--|      |   |        |   |      | /---IPv6--|      |
|      |//---IPv4--|      |/--|-IPv4---|---|      |//--IPv4---|      |
|      |\\---------|      |\--|--------|---|      |\\---------|      |
|      | \---------|      |   |        |   |      | \---------|      |
+------+           +------+   +--------+   +------+           +------+
  Source            node A                  node B           Destination
```

Figure 7: Interconnection Scenario 1

o A second scheme is to interconnect two realms implementing the SMAP function using IPv6 (Figure 8). Two sub-scenarios are identified:

   * An IPv6 prefix (i.e. WKIPv6) assigned by IANA is used for this service. If appropriate routing configuration have been enforced, then the IPv6 encapsulated packets will be routed until the final destination.

   * If an IPv6 belonging a service provider prefix is used. This will be covered in the next versions of the document.

```
   +------+             +------------+              +------+
   |      |             |            |              |      |
   |      |----IPv6-----|----IPv6----|----IPv6----\ |      |
   |      |----IPv4-----|------------|----IPv4----\\|      |
   |      |-------------|------------|------------//|      |
   |      |-------------|------------|------------/ |      |
   | SMAP |             | Internet v6|              | SMAP |
   |      | /-----IPv6--|------------|-----IPv6-----|      |
   |      |//---IPv4----|------------|-------IPv4---|      |
   |      |\\-----------|------------|--------------|      |
   |      | \-----------|------------|--------------|      |
   |      |             |            |              |      |
   +------+             +------------+              +------+
    Source                                            Destination
```

Figure 8: Interconnection Scenario 2

5.4. SMAP and PRR

Within this draft, a PRR-enabled node implements both SMAP function and required functions to handle fragmentation and other portless protocols. More details about fragmentation may be found at [I-D.boucadair-port-range].

In the remaining part, the text refers only to PRR and not SMAP.

6. IPv6 Migration Scenarios

6.1. Overview

This section proposes a set of migration steps in the context of IPv4 address exhaustion and IPv6 deployment. Both objectives are taken into account.

The proposed steps are informational. An analysis of these steps and proposed IPv6 migration paths are discussed in Section 6.7.

The following figure (i.e. (Figure 9)) provides an overview of network segments and the localisation of PRR-enabled nodes. One or several PRR may be enabled. PRD1 and PRD2 are two port-restricted devices which have been provisioned with the same IP1 public address and two distinct port ranges (PR1 and PR2).

```
             +---------+  +---------------------+  +----------+  +---------+
   +----+    |         |  |                     |  | +-----+  |  |IPv4     |
   |PRD1|----|         |--|                     |--| |i-PRR|  |--|Internet |
   +----+    | +-----+ |  |                     |  | |     |  |  +---------+
   IP1, PR1  | |a-PRR| |  |    core network     |  | +-----+  |
             | +-----+ |  |                     |  |          |
   +----+    |         |  |                     |  |          |  +---------+
   |PRD2|----|         |  |                     |  |          |--|IPv6     |
   +----+    |         |  |                     |  |          |  |Internet |
   IP1, PR2  +---------+  +---------------------+  +----------+  +---------+
             access/                              interconnection
             backhaul                               segment
```

Figure 9: Reference Architecture

6.2. IPv6 Prefixes and Addresses

Different types of IPv6 prefixes and addresses are used in the scope of the solutions described in the document (i.e. Step_0 (Section 6.4), Step_1 (Section 6.5) and Step_2 (Section 6.6)). Theses prefixes and addresses are listed hereafter:

1. IPv6Pref: A prefix allocated to the port-restricted device. A packet sent to addresses belonging to this prefix are routed toward this port-restricted device. IPv6Pref prefix addresses may also be used to send and receive native IPv6 traffic. In stateless IPv6-IPv4 Address Mapping mode (as explained above), the IPv6Pref structure is related to the IPv4 address plus port range. In binding mode, IPv6Pref and IPv4 address plus port range are independent.

2. IPv6PrefA: An address belonging to IPv6Pref prefix used to send IPv4-in-IPv6 traffic.

3. WKIPv6: An IPv6 prefix (e.g. /21, /32) common to all of the IPv6 packets which must be routed to a PRR function. It is for further study to decide whether this prefix is to be:

   A. Service provider scope

   B. or common to all service providers (to be defined by IANA).

   Both alternatives are compatible with the proposed solutions.

4. WKIPv6A: An address belonging to the WKIPv6 prefix. A WKIPv6A address includes the WKIPv6 prefix on its left most part followed by the destination IPv4 address and destination port number, as

shown in Figure 3. When a binding table is implemented, a given a-PRR has to transform a destination address WKIPv6A to a destination address IPv6PrefA, it proceeds as illustrated in Figure 10.

```
+-------------------------------------+-------------------------------------+
|                  |                  |       |                             |
|      WKIPv6      |public IPv4       |Dest.  |          0:0:0:0            |
|                  |address           |port   |                             |
+---------------------------------------------------------------------------+
                   |                 ||       |
                                     ||
                                     vv
           +--------------------------------------------------+
           |              binding table                       |
           |                                                  |
           |(...)                                             |
           |   [public IPv4 address, Port Range] => IPv6PrefA |
           |(...)                                        ||   |
           +---------------------------------------------||---+
                                                         ||
                                     /====================/
                                     ||
                                     \/
+---------------------------------------------------------------------------+
|                                                                           |
|                                   IPv6PrefA                               |
|                                                                           |
+---------------------------------------------------------------------------+
```

Figure 10: Fetching IPv6PrefA from WKIPv6A

The following sections describe three migration steps and a set of proposed migration paths. The proposed solutions are stateless at interconnection segment. A binding table may be implemented to meet requirements of service providers which do not want to closely correlate their IPv6 address plan and the IPv4 one. More details are provided below.

6.3. On Stateless and Binding Table Modes

6.3.1. Stateless Mode

Complete stateless mapping implies that the IPv4 address and the significant bits coding the port range are reflected inside the IPv6 prefix assigned to the port-restricted device. Two alternatives are

offered when such a stateless mapping is to be enabled:

- either using the IPv6 prefix already used for native IPv6 traffic,

- or provide two prefixes to the port-restricted device: one for the native IPv6 traffic and one for the IPv4 traffic.

Note that:

- Providing two IPv6 prefixes has the advantages of allowing a /64 prefix for the port-restricted device along with another prefix (e.g. a /56 or /64) for native IPv6 traffic. This alternative spares the service provider to relate the native IPv6 traffic addressing plan to the IPv4 addressing plan. The drawback is the burden to allocate two prefixes to each port-restricted device and to route them. In addition, an address selection issue may be encountered.

- Providing one prefix for both needs (e.g. a /56 or a /64) spares the service provider to handle two types of IPv6 prefix for the port-restricted device and in routing tables. But the drawback is that it somewhat links strongly the IPv4 addressing plan to the allocated IPv6 prefixes.

6.3.2. Binding Table Mode

Another alternative is to assign a "normal" IPv6 prefix to the port-restricted device and to use a binding table, which can be hosted by a service node, to correlate the (shared IPv4 address, Port Range) with an IPv6 address part of the assigned IPv6 prefix. For scalability reasons, this table should be instantiated within PRR-enabled nodes which are close to the port-restricted devices. The number of required entries if hosted at interconnection segment would be equal to the amount of subscribed users (one per port-restricted device).

The stateless mode is recommended.

6.4. Step_0: IPv6 at Access Network

This step is described in [I-D.boucadair-port-range]. This step assumes that IPv6 is used to convey incoming traffic to its final destination. For this reason, an IPv6 address is used as the routing identifier. More information about this step is provided below.

6.4.1. Context

This step can be deployed at earlier stages of IPv6 deployment. The impact on routing (especially path optimisation) and also offered services is the same as for the Port Range solution described in [I-D.boucadair-port-range]. The service brokenness risk is optimized. IPv6 is used in this step as a means to convey incoming IPv4 traffic. Within this step, IPv6 is used in the access segment to deliver the received IPv4 traffic.

When this step is deployed, at least 50% of the handled traffic (incoming+outgoing), at the IP access segment, of a service provider's domain is achieved using IPv6 capabilities. IPv4 capabilities are used only for outgoing traffic.

Native IPv6 connectivity may also be offered to end-users.

6.4.2. Overall Procedure

This section discusses additional points related to IPv6 usage in the context of the Port Range solution described in [I-D.boucadair-port-range].

6.4.2.1. Provisioning Operations

This section lists the set of information required for a port-restricted device to access the connectivity service.

6.4.2.1.1. IP Connectivity Information

Each port-restricted device is assigned with:

1. A shared public IPv4 address. In addition to this address, a port range is also assigned to the device.

2. An IPv6 prefix: denominated in the rest as IPv6Pref. This prefix is allocated to the port-restricted device and it is used to discriminate (a given device) among all those having the same public IPv4 address. This address may be used also to send and receive native IPv6 traffic or a second prefix may be assigned specific for native IPv6 traffic.

6.4.2.1.2. Provisioning procedure

To convey IPv4 configuration information, one of these solutions may be implemented:

1. Activate DHCP and support port range options as described in [I-D.bajko-pripaddrassign];

2. Use PPP and support the IPCP Port Range Configuration Options as specified in [I-D.boucadair-pppext-portrange-option].

To convey IPv6 configuration information, DHCPv6 [RFC3315] may be activated.

6.4.2.2. Port Restricted Device's behaviour and supported functions

A port-restricted device may be a host, a CPE, etc.

6.4.2.2.1. Port Restriction Behaviour

The behaviour of the port-restricted device is as follows:

1. If the port-restricted device hosts a NAT function: For incoming traffic, the port-restricted device checks if the destination port number is within the Port Range, otherwise the packet is dropped. When the destination port number of a received packet (from outside the LAN) falls inside the Port Range, classical NAT operations are enforced and the packet is then routed to its final destination in the LAN.

2. Otherwise, the port-restricted device is an end-user host: the device restricts its source port numbers to be with the assigned Port Range. Received IPv4 packets with a destination port number outside the Port Range must be dropped.

6.4.2.2.2. Handling Outgoing traffic

The same procedure as described in [I-D.boucadair-port-range] applies. In addition to the normal NAT operations, the port-restricted device ensures that the source port number is within the allowed Port Range.

6.4.2.2.3. Handling Incoming traffic

For incoming traffic, two cases may be considered:

1. IPv6 encapsulated traffic: for encapsulated IPv6 packets, the port-restricted device de-encapsulates the packets and extracts the embedded IPv4 one. The original IPv4 packets is then treated and handled locally. If the destination port of that packet is within the Port Range of that port-restricted device, and depending on the local NAT implementation if any, the packet may be accepted and then proceed to classical NAT operation.

Otherwise, the packet is dropped.

2. IPv6 native traffic: No constraint is required. The traffic should be routed to it final destination, if the port-restricted device is a CPE.

6.4.2.3. PRR Behaviour

6.4.2.3.1. Supported functions

In addition to the functions listed in [I-D.boucadair-port-range], the PRR must support an IPv6 encapsulation function.

6.4.2.3.2. Localization

The PRR function is deployed under the same conditions as the ones discussed in [I-D.boucadair-port-range].

6.4.2.3.3. Behaviour

The PRR intervenes only for incoming traffic destined to a shared IPv4 address.

a. If a binding table is implemented: This binding table stores the required information to route the traffic destined to a shared IP address to the appropriate port-restricted device among all those sharing the same IP address. An IPv6 prefix may be used as routing identifier. In this case, the structure of the binding table is: (shared IPv4 address, Port Range) ==> IPv6 prefix. Instead of the IPv6 prefix itself (IPv6Pref), the binding table may contain a specific address under IPv6Pref (called in the rest IPv6PrefA). When a binding table is adopted, the IPv6 prefix assigned to a port-restricted device is not constrained. There is no need for the service provider to allocate two different IPv6 prefixes to the port-restricted devices (one for native IPv6 traffic and another one for the IPv4 encapsulated traffic). Only one may suffice for the two needs.

b. Stateless mapping: The service provider assigns an IPv4 shared address, a port range and an IPv6 prefix with IPv6 prefix containing explicitly the IPv4 address and the significant bits of the port range (i.e. Bits used to built the Port Range Mask, see [I-D.bajko-pripaddrassign]). When a stateless mapping is adopted, it is possible that the service provider has to cope with constraints when allocating the IPv6 prefix and the shared IPv4 address to the port-restricted devices. As a matter of fact the IPv6 prefix must reflect the shared IPv4 address. Alternatively the service provider may instead allocate two

different prefixes for the two needs (IPv6 native traffic and IPv4 encapsulated traffic).

In the remaining part of this section, "PRR retrieves the corresponding IPv6 prefix address" means that:

a. If a binding table is implemented: the PRR looks-up through this table and retrieves the IPv6Pref or IPv6PrefA corresponding to (IPv4 address, Port Range). If IPv6Pref is retrieved, the PRR builds an IPv6Pref address in complementing the IPv6Pref with a fixed bit sequence chosen by the service provider to be always the same complementary bit sequence for all port-restricted devices. .

b. Stateless mode: an IPv6 prefix address (i.e. IPv6PrefA) is built using the IPv4 address and the destination port number.

In both cases, when the PRR has got/built the corresponding IPv6PrefA , the PRR encapsulates the original packets in an IPv6 one with a destination IP address equal to the IPv6Pref address. The source IPv6 address is an address of the PRR. It may be an anycast IPv6 address or unicast one.

This packet is then routed according to instantiated IGP routes.

6.4.2.4. Routing considerations

The same IGP considerations as detailed in [I-D.boucadair-port-range] should be taken into account. In addition to these considerations, IPv6 routes should be installed to reach port-restricted devices from an IPv6-enabled PRR.

6.4.3. Focus on Communication Establishment

6.4.3.1. Outgoing IPv4 Communications

Outgoing IPv4 traffic is handled as described in [I-D.boucadair-port-range]. The traffic issued from a port-restricted device is routed to its final destination. The traffic is not altered and is transferred to its final destination.

6.4.3.2. Incoming IPv4 Communications

Owing to IGP configuration, the traffic destined to a shared IP address must cross a PRR. This latter encapsulates the received IPv4 packets in IPv6 ones as described in Section 6.4.2.3.3. The traffic is then routed using the IPv6PrefA as a destination address. The traffic is received by the device among those sharing the same IPv4

address (because this port-restricted device is allocated with this IPv6Pref). A de-encapsulation operation is then executed as described in Section 6.4.2.2. If the de-encapsulated IPv4 traffic is destined to a port within the assigned Port Range, the traffic is accepted, otherwise it is dropped.

6.4.3.3. Outgoing IPv6 Communications

Since the port-restricted device is IPv6-enabled, native IPv6 communications may be offered. This assumes that the service provider has deployed means for IPv6 transfer capability. The same prefix used to convey incoming IPv4 traffic (IPv6Pref) may be used also to send and receive native IPv6 traffic. Alternatively, a second IPv6 prefix may be assigned to that purpose.

6.4.3.4. Incoming IPv6 Communications

Native IPv6 communications are supported.

6.4.4. Typical Flow Example

In order to illustrate this step, let consider the example shown in the following figure (Figure 11). M1 is a machine behind a port-restricted device (called CPE as Customer Port restricted Equipment in the example and Figure 11).

M1 wants to establish an IPv4 communication with RM (Remote Machine). To do so, an IPv4 packet is issued by M1. This packet has as source IP address equal to Pri_IPv4. The packet is then received by CPE. This latter enforces its NAT operations. As a result, an IPv4 packet with a source IP address equal to Pub_IPv4 and a source port number within the Port Range of CPE is sent. The resulting packet is forwarded according to IPv4 transfer capabilities until reaching its final destination RM.

As a response, RM sends an IPv4 packet destined to Pub_IPv4 and a destination port number equal to the source port number of the received packet. This message is received by PRR. The PRR encapsulates the received IPv4 packet in an IPv6 one. The resulting IPv6 packet is then forwarded. The encapsulated packet is received by the appropriate CPE among those having the same IPv4 address. A de-encapsulation is enforced. The original IPv4 packet is extracted. Once classical NAT operations are executed, the CPE forwards the IPv4 packet to M1.

```
     +--+              +---+                                       +--+
     |M1|              |CPE|                                       |RM|
     +--+              +---+                                       +--+
       |                 |                                           |
       |==Pri_IPv4_Out==>|===============Pub_IPv4_Out===============>|
       |                 |                                           |
       |                 |                        +---+              |
       |                 |                        |PRR|              |
       |                 |                        +---+              |
       |                 |                          |                |
       |<==Pri_IPv4_In===|<==IPv6_Enc(Pub_IPv4_In)==|<==Pub_IPv4_In==|
       |                 |                          |                |
```

Figure 11: Flow Example (Step 0): Inter-domain

6.5. Step_1: IPv6 a Means to Transfer Incoming IPv4 Packets

6.5.1. Context

Step_1 is characterized by the activation of two levels of PRR functions in several segments of a given service provider's network. Some PRR-enabled nodes are deployed close to the port-restricted devices (e.g. In the access or backhaul network) whilst others are installed at the interconnection segment of the ISP network as shown in Figure 9.

The objective of this step is to maximize the invocation of IPv6 capabilities, particularly to convey incoming IPv4 traffic until delivery to final destination (e.g. port-restricted device).

6.5.2. Overall Procedure

Step_1 works exactly the same way as that the Step_0, apart what is specified in this section. In particular, there are none difference between Step_0 and Step_1 at the level of the port-restricted device.

6.5.2.1. Routing considerations

- a-PRR: As in Step_0, a-PRR announces the shared IPv4 prefixes it serves. In addition, in case of binding table mode, it announces in IGP the aggregates of all the WKIPv6A addresses of the port-restricted devices it serves so that IPv4-in-IPv6 packets reach this a-PRR.

- i-PRR: i-PRR announces in EGP (if embedded in an ASBR node) or in IGP (if deployed behind an ASBR), all (aggregated) IPv4 prefixes of all port-restricted devices it can route packets to (via a-PRR in case of binding table). Depending of the structure

of the service provider network, some i-PRR-enabled nodes may be positioned inside the service provider network to encapsulate more IPv4 traffic into IPv6. For example, from a region A to another region B of the service provider's network.

6.5.2.2. Behaviour of a-PPR and i-PRR

Figure 12 show the respective role of a-PRR and i-PRR-enabled nodes. The labels of the arrows are explained in further sub-sections.

- CPE1 (Customer Port restricted Equipment) is a port-restricted device 'served' by a-PRR (means that a-PRR announces in IGP the assigned shared IPv4 address to CPE1).

- CPE2 is a device of another customer connected to the network managed by the same service provider. It may be either another port-restricted device or a device with a plain IPv4 address.

- RM is a remote machine located outside the AS managed by the service provider.

```
     +----+                                 +-----+            +-----+
     |CPE1|                                 |i-PRR|            | RM  |
     +----+                                 +-----+            +-----+
       |                                       |                  |
       |<=====IPv6PrefA_Enc(Pub_IPv4_In)=======|<===Pub_IPv4_In===|
       |                                       |                  |
       |                                       |                  |
       |                 +-----+
       |                 |a-PRR|
       |                 +-----+            +----+
       |                     |              |CPE2|
       |                     |              +----+
       |                     |                |
       |<==IPv6PrefA_Enc     |<==Pub_IPv4_In==|
       |      (Pub_IPv4_In)==|                |
       |                     |                |
```

Figure 12: Step_1: Stateless mode

```
     +----+            +-----+              +-----+            +-----+
     |CPE1|            |a-PRR|              |i-PRR|            | RM  |
     +----+            +-----+              +-----+            +-----+
       |                  |                    |                  |
       |<==IPv6PrefA_Enc  |<==WKIPv6A_Enc      |<===Pub_IPv4_In===|
       |   (Pub_IPv4_In)==|      (Pub_IPv4_In)=|                  |
       |                  |                    |                  |
       |                  |
       |                  |
       |                  |             +----+
       |                  |             |CPE2|
       |                  |             +----+
       |                  |                |
       |<==IPv6PrefA_Enc  |<==Pub_IPv4_In==|
       |   (Pub_IPv4_In)==|                |
       |                  |                |
```

Figure 13: Step_1: Binding table mode

6.5.2.2.1. Localization

The PRR function is deployed under the same conditions as in Step_0 but as previously mentioned more PRR-enabled nodes are deployed within the ISP network.

6.5.2.2.2. Stateless Mapping Mode

In case the service provider assigns an IPv4 shared address, a port range and an IPv6 prefix containing explicitly the IPv4 address and the significant bits of the port range, i-PRR-nodes are able to build the IPv6Pref address of the port-restricted device using the IPv4 destination address and destination port bits of received IPv4 packets.

Hence, i-PPR behaviour is the same one as the PRR one described in Step_0, when stateless mapping is enforced. In that case, the IPv4-in-IPv6 packet does not pass through the a-PRR but it is routed directly to the port-restricted device (e.g. To CPE1 as depicted in Figure 12).

6.5.2.2.3. Binding Table Mode

This behaviour is illustrated in Figure 13.

If a binding table is implemented within a-PRR, i-PRR-enabled nodes can not hold all the binding table entries corresponding to all the port-restricted devices it may route traffic for. Consequently, it has to route the IPv4 traffic towards the a-PRR of which the port-

restricted device depends. More precisely, a given i-PRR encapsulates the incoming IPv4 traffic in IPv6 packets using the following addresses:

- The source IPv6 address is one of the global IPv6 addresses of the i-PRR.

- The destination IPv6 is built by i-PRR using an address under the WKIPv6 prefix conforming to the formalism defined in Section 4.2. This address, called WKIPv6A, includes the WKIPv6 prefix on its left most part and the destination IPv4 address and destination port number in this right most part.

Thus, an i-PRR-enabled node routes normally this IPv4-in-IPv6 packet (labeled WKIPv6A_Enc (Pub_IPv4_In) in Figure 13) using IPv6 transfer capabilities. The packet is routed towards the a-PRR serving the recipient port-restricted device (CPE1 is Figure 13) since appropriate routing configuration has been enforced (see Section 6.5.2.1).

Upon receipt of that packet, the a-PRR proceeds as follows:

- It retrieves the IPv4 shared address and port bits parts of the WKIPv6A address. Then, with these parts it looks-up through its binding table to fetch the IPv6PrefA corresponding to the couple (IPv4 address, Port Range).

- Once retrieved, the a-PRR positions the IPv6PrefA address as the destination address in place of the WKIPv6A address and forwards the packet. The IPv4-in-IPv6 packet is routed until reaching the port-restricted device (CPE1 in Figure 13).

As shown in Figure 13 (bottom part), when another machine (CPE2) within the same service provider's domain sends traffic to the port-restricted device CPE1, the working of a-PRR is the same one as that of the PRR is Step_0.

6.6. Step_2: Only IPv6 Is Used For Both Incoming and Outgoing Traffic

6.6.1. Context

This step is suitable for service providers wishing to migrate to full IPv6 and to offer a global connectivity using IPv6. This step provides a lightweight procedure to interconnect IPv6 with IPv4 realms. This procedure may be fully stateless or require a binding table. This table does not include session-based information.

Only IPv6 connectivity is used inside the service provider's domain;

IPv4 capabilities are deactivated. No parallel IPv4 and IPv6 operational tasks will be maintained anymore in the core segment.

Two implementation modes may be envisaged:

1. Encapsulation-based mode: This mode suggests using both inbound and outbound IPv6 encapsulation to carry IPv4 traffic (received from the remaining IPv4-only realms). This mode is almost similar to Step_1 for handling incoming IPv4 packets. Unlike Step_1, the required operations to build the outgoing encapsulated packets is also supported in this step.

2. Translation-based mode: Unlike the first mode, this one assumes that there is no need to maintain both IPv4 and IPv6 stacks in the CPE. It is recommended to implement this mode when mature IPv6 deployment has been observed and that IPv6 realms become more important than IPv4-only ones.

More information about these two modes is provided in the sub-sections hereafter.

6.6.2. The IPv6 Encapsulation-Based Mode

6.6.2.1. Provisioning Operations

6.6.2.1.1. IP Connectivity Information

In addition to the information required for Step_1, a WKIPv6 IPv6 prefix may e configured on the port-restricted device. This prefix is to be used when running the IPv4-to-IPv6 mapping function required to encapsulate IPv4 traffic in IPv6 one.

6.6.2.1.2. Provisioning Procedure

Idem as Step_1.

6.6.2.2. Routing Considerations

IPv4 IGP protocols are not anymore enabled in the core network. Only IPv6 routing table is maintained by involved routers.

Inter-domain IPv4 connectivity is maintained with IPv4-only realms. IPv4 network prefixes are mapped to IPv6 prefixes (using a WKIPv6 configured by the service provider) which are injected in the deployed IPv6 IGP protocol.

A given a-PRR MUST advertise in IGP the aggregated IPv6 prefixes it handles. Doing so, all intra-domain IPv6 packets will cross that

PRR.

6.6.2.3. Port Restricted Device's Behaviour and Supported Functions

6.6.2.3.1. Port Range Restriction

Idem as Step_1.

6.6.2.3.2. Handling Outgoing Traffic

Unlike Step_1, outgoing IPv4 traffic is encapsulated in IPv4-in-IPv6 packets. Concretely, the port-restricted device executes its port restricted NAT operations (if any). The resulting IPv4 packet is then encapsulated in an IPv6 packet. The port-restricted device selects an IPv6 address from its assigned IPv6 prefix (IPv6Pref). This address IPv6PrefA is used as the source IPv6 address of the encapsulated packet.

Two options may be considered to build the destination IPv6 address of the encapsulated packet as listed below:

1. The port-restricted device is provisioned to use an anycast IPv6 address. This anycast IPv6 address is configured on internal interfaces of all PRRs. This mode is may be implemented when the port-restricted device is not able to build a destination IPv6 address reflecting the IPv4 address and port of its correspondent (i.e. the port-restricted device does not support the mapping function defined in Section 4.2).

2. The port-restricted device is able to build an IPv6 address using a WKIPv6 prefix (which may be distinct than the one used to build mapped IPv6 prefixes by i-PRRs), the destination IPv4 address and the destination port number.

No constraints are to be followed for outgoing native IPv6 traffic.

6.6.2.3.3. Handling Incoming Traffic

Idem as Step_1.

6.6.2.4. PRR Behaviour

6.6.2.4.1. Supported Functions

In addition to the functions supported in Step_1, an IPv4-in-IPv6 de-encapsulation function must be supported by a-PRR.

6.6.2.4.2. Localization

Idem as Step_1.

6.6.2.4.3. Behaviour: Stateless Mode

The behaviour of both access and interconnection PRRs is elaborated below:

1. If an anycast IPv6 address is configured on interfaces of all a-PRRs:

   * An (access) PRR will receive the encapsulated packet issued from port-restricted devices. The packet is de-encapsulated and the original IPv4 one is retrieved. Then, the (access) PRR builds a destination IPv6 address using a WKIPv6 prefix, the destination IPv4 address and the destination port number. The original IPv4 packet is then encapsulated in IPv6 packet with a source IPv6 address of the (access) PRR and the destination IPv6 address equal to the newly built one. The packet is forwarded to next hop according to IPv6 routing table. If the correspondent is not a port-restricted device, the packet is intercepted by a-PRR or a i-PRR depending of where the correspondent is located. This a-PRR or i-PRR proceeds to the de-encapsulation operation. The extracted IPv4 packet is then forwarded to the IPv4 correspondent. This encapsulated packet is received by an (access/ interconnection) PRR which proceeds to a de-encapsulation operation. The extracted IPv4 packet is then forwarded to the next IPv4 hop.

   * Incoming IPv4 traffic is intercepted by an (interconnection) PRR. The PRR encapsulates the received IPv4 packet in an IPv6 one using the following information:

     + The source IPv6 address is one of the global IPv6 addresses of the PRR.

     + The destination IPv6 is built by the PRR using the formalism defined in Section 4.2. This address is build using a WKIPv6 prefix, the destination IPv4 address and the destination port number. The appropriate port-restricted device, among those having the same IPv4 address, will receive the packet. This is illustrated in Figure 14.

```
   +---+                    +-----+                     +-----+            +--+
   |CPE|                    |a-PRR|                     |i-PRR|            |RM|
   +---+                    +-----+                     +-----+            +--+
    |                          |                            |                |
    |=IPv6_Enc(Pub_IPv4_Out)==>|=WKIPv6A_Enc(Pub_IPv4_Out)=>|==Pub_IPv4_Out=>|
    |                          |                            |                |
    |                          |                            |                |
    |<=========IPv6PrefA_Enc(Pub_IPv4_In)===================|<==Pub_IPv4_In==|
    |                          |                            |                |
    |                          |                            |                |
```

LAN messages are not represented in the figure.

Figure 14: Encapsulation Mode: Anycast addresses are assigned to all PRR

2. Otherwise, all internal IPv4 traffic is encapsulated by the port restricted device in IPv4-in-IPv6 packets (the destination IPv6 address is the one built by the port-restricted device, see Section 4.2). As in the first bullet, the packet is forwarded to next hop according to IPv6 routing table. If the correspondent is not a port-restricted device, the packet is intercepted by a-PRR or a i-PRR depending of where the correspondent is located. This a-PRR or i-PRR proceeds to the de-encapsulation operation. The extracted IPv4 packet is then forwarded to the IPv4 correspondent. The same behaviour as for the first bullet applies for incoming IPv4 traffic. A flow is illustrated in Figure 15.

```
   +---+                                          +-----+             +--+
   |CPE|                                          |i-PRR|             |RM|
   +---+                                          +-----+             +--+
    |                                                |                  |
    |=======WKIPv6A_Enc(Pub_IPv4_Out)===============>|====Pub_IPv4_Out=>|
    |                                                |                  |
    |                                                |                  |
    |<======IPv6PrefA_Enc(Pub_IPv4_In)===============|<==Pub_IPv4_In====|
    |                                                |                  |
```

Figure 15: Encapsulation Mode: the port restricted device builds a destination IPv6 address

6.6.2.4.4. Behaviour: Binding Table Mode

The behaviour of both access and interconnection PRRs is elaborated below:

1. If an anycast IPv6 address is configured on interfaces of all a-PRRs:

   * An (access) PRR will receive the encapsulated packets issued from port-restricted devices. The packets are de-encapsulated and the original IPv4 is retrieved. Then, the (access) PRR builds a destination IPv6 address using a WKIPv6 prefix and the destination IPv4 address. The original IPv4 packets are then encapsulated in IPv6 packet with a source IPv6 address of the (access) PRR and the destination IPv6 address equal to the newly built one. The packet is forwarded to next hop according to IPv6 routing table. The packet is intercepted by a-PRR or a i-PRR depending of where the correspondent is located. This a-PRR or i-PRR proceeds to the de-encapsulation operation. The extracted IPv4 packet is then forwarded to the IPv4 correspondent.

   * Incoming IPv4 traffic is intercepted by an (interconnection) PRR. The PRR encapsulates the received IPv4 packet in an IPv6 one using the following information:

     + The source IPv6 address is one of the global IPv6 addresses of the PRR.

     + The destination IPv6 is built by the PRR using the formalism defined in Section 4.2. This address is build using a WKIPv6 prefix, the destination IPv4 address and the destination port number. The appropriate a-PRR managing the destination port-restricted device, among those having the same IPv4 address, will receive the packet. This is illustrated in Figure 16. The PRR de-encapsulates the packet and retrieves the original IPv4 packet. The access PRR retrieves the destination address IPv6PrefA stored in its binding table and forwards the packet to the port restricted device.

```
   +---+                    +-----+                    +-----+          +--+
   |CPE|                    |a-PRR|                    |i-PRR|          |RM|
   +---+                    +-----+                    +-----+          +--+
    |                          |                            |                |
    |==IPv6_Enc(Pub_IPv4_Out)=>|=WKIPv6A_Enc(Pub_IPv4_Out)=>|==Pub_IPv4_Out=>|
    |                          |                            |                |
    |<==IPv6_Enc(Pub_IPv4_In)==|<==WKIPv6A_Enc(Pub_IPv4_In)=|<==Pub_IPv4_In==|
    |                          |                            |                |
```

LAN messages are not represented in the figure.

Figure 16: Encapsulation Mode with a Binding table (Anycast)

2. Otherwise, all internal IPv4 traffic is encapsulated by the port restricted device in IPv4-in-IPv6 packets. As in the first bullet, the packet is forwarded to next hop according to its IPv6 routing table. The packet is intercepted by a-PRR or a i-PRR depending of where the correspondent is located. This a-PRR or i-PRR proceeds to the de-encapsulation operation. The extracted IPv4 packet is then forwarded to the IPv4 correspondent. The same behaviour as for the first bullet applies for incoming IPv4 traffic. A flow example is illustrated in Figure 17.

```
   +---+                                                     +-----+       +--+
   |CPE|                                                     |i-PRR|       |RM|
   +---+                                                     +-----+       +--+
    |                                                           |                |
    |================WKIPv6A_Enc(Pub_IPv4_Out)==============>|=Pub_IPv4_Out=>|
    |                                                           |                |
    |                          +-----+                          |                |
    |                          |a-PRR|                          |                |
    |                          +-----+                          |                |
    |                             |                             |                |
    |<=IPv6PrefA_Enc(Pub_IPv4_In)|<=WKIPv6A_Enc(Pub_IPv4_In)=|<=Pub_IPv4_In==|
    |                             |                             |                |
```

Figure 17: Encapsulation Mode with a Binding table

6.6.3. The IPv6 Translation-Based Mode

6.6.3.1. Context and Conditions

This mode assumes that IPv6-only terminals are deployed behind port-restricted devices. Particularly, all DNS resolution requests are AAAA ones [RFC3363] . Only IPv6 addresses are conveyed in DNS responses to requesting machines.

A dedicated ALG should be supported by the DNS infrastructure deployed by the service provider. The main function of this ALG is to form an IPv6 address based on a WKIPv6 prefix and a resolved IPv4 address, when no AAAA RR are available in the DNS system (following the formalism described in Section 4.2). The WKIPv6 prefix is configured by the service provider so as to identify that the resulting IPv6 address is not a native one. We refer to this IPv6 prefix as WKIPv6_v4. The procedure is illustrated in Figure 18.

```
            +--+                  +---+                 +---+
            |M1|                  |CPE|                 |DNS|
            +--+                  +---+                 +---+
              |                     |                     |
              |==(1)Request(AAAA)==>|==(2)Request(AAAA)==>|
              |                     |                     |
              |                     |             +------ALG------+
              |                     |             |No available   |
              |                     |             |AAAA Records,  |
              |                     |             +------| |------+
              |                     |             |      | |      |
              |                     |             +-------v-------+
              |                     |             | Return IPv4@  |
              |                     |             | Build IPv6@   |
              |                     |             +-------+-------+
              |                     |                     |
              |<=(4)Response(IPv6@)=|<=(3)Response(IPv6@)=|
              |                     |                     |
```

Figure 18: DNS ALG

In the remaining part of this section, it is assumed that M1 has retrieved an IPv6 address to contact.

6.6.3.2. Flow Example

Figure 19 shows a message exchange that occurs in the context of IPv6-IPv4 communications.

```
        +--+            +---+            +-----+               +--+
        |M1|            |CPE|            |i-PRR|               |RM|
        +--+            +---+            +-----+               +--+
          |                 |                 |                    |
          |==(1)IPv6_Out==>|==(2)IPv6_Out===>|==(3)Pub_IPv4_Out==>|
          |                 |                 |                    |
          |<==(6)IPv6_In===|<==(5)IPv6_In====|<===(4)Pub_IPv4_In==|
          |                 |                 |                    |
```

Figure 19: Translation Mode

Intra-domain communications are placed using IPv6 transfer capabilities. When the remote destination is an IPv4 (which is represented by an IPv6 address), the following exchanges are observed:

1. M1 issues an IPv6 message destined to RM.

2. Once received by the CPE, this latter checks if the destination address belongs to the WKIPv6_v4 prefix. If this is the case, a NAT66 operation is executed. As a result, a new IPv6 packet is generated.

3. This message is received by the interconnection PRR. It retrieves IPv4 information based on IPv6 one and translates the packet to a new IPv4 one.

4. This message is then routed using IPv4 capabilities of the connected IPv4-only realm.

5. Once received by RM, an answer may be issued. An IPv4 packet is then sent.

6. This IPv4 packet is received by i-PRR. It then proceeds to a stateless NAT46 operation. The newly built IPv6 packet is forwarded to the next hop.

7. Owing to underlying IGP configuration, the packet is received by the appropriate CPE which checks its NAT66 table.

8. Because a session has been instantiated, a NAT66 operation is executed. The resulting IPv6 packet is then received by M1.

6.6.3.3. Provisioning Operations

6.6.3.3.1. IP Connectivity Information

Unlike previous steps, no IPv4 connectivity is provided to customers. None IPv4 packets are sent, neither by the end-user's device within the LAN nor by the CPE itself.

IP connectivity is exclusively offered owing to IPv6 transfer capabilities. Thus, no IPv4 connectivity information is conveyed to end-user's device. In the meantime, an IPv6 prefix (IPv6Pref) is assigned to the end-user device (CPE or terminal). This assigned IPv6 prefix follows the constraints listed in Section 4.

As already mentioned in Section 6.3, the service provider may allocate to the customer's device a second prefix IPv6 prefix which is not IPv4-mapped.

6.6.3.3.2. Provisioning Procedure

In addition to what has been mentioned for Step_1 (IPv6 part), a specific policy should be installed so as to "guide" the behaviour of the NAT66 function introduced in "Handling Outgoing Traffic" section.

This specific policy needs to be aware of the port range allocated to the port-restricted device. It is for further study to defined how the port range can be allocated through IPv6 means (e.g. through a new DHCPv6 option).

6.6.3.4. Port Restricted Device's Behaviour and Supported Functions

6.6.3.4.1. Port Range Restriction

The port restriction is applied only if the destination IPv6 address belongs to the WKIPv6_v4 prefix. Otherwise, no port restriction is enforced, since it is assumed to be a native IPv6 communication.

A new NAT66 function should be supported by the CPE. This NAT66 is not required to be supported if a directly connected terminal is used. But then, its address selection process should follow the recommendations listed in "Handling Outgoing Traffic" sub-section.

6.6.3.4.2. Handling Outgoing Traffic

The following procedure is applied:

- o If the destination IPv6 address belongs to the WKIPv6_v4 prefix (this means the destination is not a native IPv6 host and an IPv6-IPv4 interconnection node will be crossed in the delivery path), the port-restricted device proceeds to NAT66 operations.

Concretely:

- A port number from the Port Range is selected, this port replaces the original source port number in the transport part of the received IPv6 packet;

- A source IPv6 address IPv6PrefA is selected under IPv6Pref (see Section 4) in such a way that the port value contained in the port part of this IPv6PrefA address is equal to the selected port number;

- The received IPv6 (from a machine in the LAN) packet is then translated to a new IPv6 one with the newly built IPv6PrefA address as source address and the newly selected source port number in transport part.

- Otherwise, the packet is forwarded to the next IPv6 hop.

6.6.3.4.3. Handling Incoming Traffic

The following procedure is applied:

- If the source IPv6 address belongs to the WKIPv6_v4 prefix, the port-restricted device proceeds to NAT66 operations according to its active NAT sessions.

- Otherwise, the packet is forwarded to the next IPv6 hop in the LAN.

6.6.3.5. PRR Behaviour

6.6.3.5.1. Supported Functions

Unlike previous steps, no encapsulation function is required to be supported by the PRR. Nevertheless, a stateless IPv6-IPv4 (and vice versa) translation must be supported.

6.6.3.5.2. Localization

No PRRs are required anymore to be maintained in the access segment. Only PRRs located in the interconnection segment should be deployed. These nodes should be near ASBRs used to interconnect with IPv4-only realms.

6.6.3.5.3. Behaviour

The behaviour of the PRR is as follows:

1. Traffic received from an IPv4-only realm: The PRR extracts destination IPv4 address, source IPv4 address, destination port number and source port number. A new IPv6 packet is generated following these rules:

   * The destination and source port numbers of the generated packet are the same as the original IPv4 one.

   * The destination IPv6 address follows the formalisms described in Section 4.2 using a WKIPv6 configured by the service provider.

   * The source IPv6 address follows the formalism described in Section 4.2 using a WKIPv6 provided by the service provider.

2. Traffic destined to an IPv4-only realm: A new IPv4 packet is generated according to these rules:

   * The destination and source port numbers of the generated packet are the same as the original IPv6 one.

   * The destination IPv4 address is extracted from the destination IPv6 address of the received IPv6 packet. See Section 4.3 for more information about the used IPv6-to-IPv4 address mapping function.

   * The source IPv4 address is extracted from the source IPv6 address of the received IPv6 packet using the IPv6-to-IPv4 address mapping function defined in Section 4.3.

6.6.3.6. Routing Considerations

IPv4 IGP protocols are not anymore enabled in the core network. Only IPv6 routing table is maintained by involved routers.

Inter-domain IPv4 connectivity is maintained with IPv4-only realms. IPv4 network prefixes are mapped to IPv6 prefixes (using a WKIPv6 prefix provided by the local service provider) which are injected in the IPv6 IGP deployed protocol.

6.7. Analysis and IPv6 Migration Scenarios

As aforementioned, the deployment of IPv6 is not a problem per se. The main issue is how to ensure a smooth interconnection between IPv4 and IPv6 realms. Interworking functions and procedures should be deployed. Currently proposed mechanisms rely on statefull interconnection nodes (e.g. CGN or DSLite server) or requires that dual-stack nodes (including end hosts and intermediary service nodes)

are deployed everywhere. The first category suffers from a performance issue and the second one is not realistic approach (since the adoption of IPv6 may require several years).

This document presents solutions which solve the problem of IPv4 address shortage and which prepare the migration to IPv6. As described in previous sections, three steps have been identified and specified. Table 1 gives an overview of the supported IP version per network segment and for each step. The proposed solution requires the migration of core segments to IPv6. Dual stacks would be maintained only at interconnection segments. Table 2 presents the ratio of IPv6 traffic when the solution is deployed. This table illustrates the invocation of IPv6 capabilities for the delivery of IP connectivity service. Owing to the deployment of the proposed solution, service providers have deterministic means to increase IPv6 traffic.

Step	Access	Core	Interconnection
Step_0	DS Network	IPv4 Network	IPv4 Network
Step_1	DS Network	DS Network	DS Network
Step_2	IPv6 Network	IPv6 Network	DS Network

Table 1: Supported IP version per network segment

Step	%IPv6 traffic	%IPv6 traffic core
	Access	
Step_0	at least 50%	variable
Step_1	at least 50%	at least 50%
Step_2 (Encapsulation)	100%	100%
Step_2 (Translation)	100%	100%

Table 2: % of IPv6

Table 3 lists the required function/node to be supported in order to ensure heterogeneous communication involving peers located in both IPv4 and IPv6 realms. Core network segment does not require the deployment of non IPv6 standards elements. Stateless functions (denoted as PRR in this document) are introduced first at access segment and then in the interconnection segment. Intra-domain communications are optimised from an IGP perspective. The presence of a-PRR allows a natural traffic distribution among deployed nodes. IPv4 traffic received from adjacent domains is handled by the i-PRR

and then routed to its final destination possibly through the a-PRR depending of the configuration (binding table with one line per port-restricted device or stateless mapping between shared IPv4 address and IPv6 prefixes).

Step	Access	Core	Interconnection
Step_0	a-PRR	None	None
Step_1	a-PRR	None	i-PRR
Step_2 (Encapsulation)	a-PRR/None	None	i-PRR
Step_2 (Translation)	None	None	i-PRR

Table 3: Required nodes per network segment

Table 4 lists the required functions to be enabled in the context of each step.

Step	port-restricted device	a/i-PRR
Step_0	Port Restricted IPv4, IPv4-in-IPv6 de-encapsulation	Stateless IPv4-in-IPv6 encapsulation
Step_1	Port Restricted IPv4, IPv4-in-IPv6 de-encapsulation	stateless IPv4-in-IPv6 encapsulation
Step_2 (Encapsulation)	Port Restricted IPv4, IPv4-in-IPv6 encapsulation, IPv4-in-IPv6 de-encapsulation	Stateless IPv4-in-IPv6 encapsulation, IPv4-in-IPv6 de-encapsulation
Step_2 (Translation)	Port Restricted NAT66	Stateless NAT46/NAT64

Table 4: Required Functions

Various migration paths may be adopted by service providers based on backward compatibility considerations and also to the service portfolio.

For service providers which offer already an IPv4-based connectivity service, several migration paths may be followed, depending on the service provider's objectives and profile, to adopt IPv6 without breaking global connectivity (i.e. Reach both IPv4 and IPv6 realms):

1. Deploy IPv6 with no major risks on currently offered services: a candidate migration path would be (ordered steps):

    A. Deploy first the procedure described in [I-D.boucadair-port-range]. Then, IPv6 may be activated in the access segment when deploying Step_0. Once IPv6 is deployed in core network, the service provider should activate Step_1 mainly by deploying PRR at interconnection segments. Once the connectivity service is stable, a final step would be to adopt Step_2 (encapsulation mode) and then Step_2 (translation mode).

    B. Deploy first Step_0, then adopt Step_1. Once the service is stable, move to Step_2 (encapsulation mode) and latter to Step_2 (translation mode).

2. Aggressive position with regards to IPv6 deployment: For this category of service providers, the migration path would be either

    A. either deploy Step_1 then Step_2 (encapsulation mode) and finally Step_2 (translation mode).

    B. or deploy Step_2 (encapsulation mode) and then Step_2 (translation mode).

For new service providers which do not have backward compatibility requirement, the following deployment path may be adopted to ensure a global IPv4/IPv6 connectivity service.

> Deploy Step_2 (encapsulation mode) and then migrate to Step_2 (translation mode).

7. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

8. Security Considerations

TBC

9. Acknowledgements

The authors would like to thank Pierrick MORAND for his review, support and suggestions.

10. References

10.1. Normative References

[I-D.bajko-pripaddrassign]
Bajko, G., Savolainen, T., Boucadair, M., and P. Levis, "Port Restricted IP Address Assignment", draft-bajko-pripaddrassign-00 (work in progress), February 2009.

[I-D.boucadair-pppext-portrange-option]
Boucadair, M., Levis, P., Grimault, J., and A. Villefranque, "Port Range Configuration Options for PPP IPCP", draft-boucadair-pppext-portrange-option-00 (work in progress), February 2009.

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

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

10.2. Informative References

[I-D.boucadair-port-range]
Boucadair, M., Levis, P., Bajko, G., and T. Savolainen, "IPv4 Connectivity Access in the Context of IPv4 Address Exhaustion", draft-boucadair-port-range-01 (work in progress), January 2009.

[I-D.despres-sam]
Despres, R., "Stateless Address Mappings (SAMs) IPv6 & extended IPv4 via local routing domains - possibly multihomed", draft-despres-sam-01 (work in progress), November 2008.

[I-D.ietf-softwire-dual-stack-lite]
Durand, A., Droms, R., Haberman, B., and J. Woodyatt, "Dual-stack lite broadband deployments post IPv4 exhaustion", draft-ietf-softwire-dual-stack-lite-00 (work in progress), March 2009.

[I-D.levis-behave-ipv4-shortage-framework]
Levis, P., Boucadair, M., Grimault, J., and A. Villefranque, "IPv4 Shortage: Needs and Open Issues", draft-levis-behave-ipv4-shortage-framework-01 (work in progress), March 2009.

[RFC3363] Bush, R., Durand, A., Fink, B., Gudmundsson, O., and T. Hain, "Representing Internet Protocol version 6 (IPv6) Addresses in the Domain Name System (DNS)", RFC 3363, August 2002.

Authors' Addresses

Mohamed Boucadair (editor)
France Telecom
42 rue des Coutures
BP 6243
Caen Cedex 4 14066
France

Email: mohamed.boucadair@orange-ftgroup.com

Pierre Levis
France Telecom

Email: pierre.levis@orange-ftgroup.com

Jean-Luc Grimault
France Telecom

Email: jeanluc.grimault@orange-ftgroup.com

Alain Villefranque
France Telecom

Email: alain.villefranque@orange-ftgroup.com

Mohamed Kassi-Lahlou
France Telecom

Email: mohamed.kassilahlou@orange-ftgroup.com