

Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 24, 2009

P. Levis, Ed.
M. Boucadair
JL. Grimault
A. Villefranke
France Telecom
June 22, 2009

IPv4 Address Shortage: Needs and Open Issues
draft-levis-behave-ipv4-shortage-framework-02.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on December 24, 2009.

Copyright Notice

Copyright (c) 2009 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 in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This document analyses the main issues related to IPv4 Internet

access in the context of public IPv4 address exhaustion.

Table of Contents

1.	Introduction	3
2.	Shared IPv4 Addresses	3
3.	Solution Space	4
3.1.	CGN-based Solutions	4
3.2.	A+P Solutions	5
3.3.	Common Architecture	6
4.	Address Space Multiplicative Factor	7
5.	Number of Current Sessions	8
6.	Service Management	10
7.	IPv6 Migration and IPv4-IPv6 Coexistence	10
8.	Network Addressing Capability	11
9.	Scarcity of Private Addressing	12
10.	Scalability	12
11.	Impact on Information System	14
12.	Impact on Services	14
13.	Port Space Boundaries	15
14.	Flow Discrimination	17
15.	Impact on Intra-Domain and Inter-Domain Routing	17
16.	Fragmentation	17
17.	QoS	17
17.1.	QoS performance	17
17.2.	QoS mechanisms	18
17.3.	Introduction of Single Point of Failure (Robustness)	18
18.	Support of Multicast	18
19.	Mobile-IP	18
20.	End-Users Facilities	19
21.	Management Tools	20
22.	Legal Obligations	20
22.1.	Traceability	20
22.2.	Interception	21
23.	Security	21
23.1.	Port Randomization	21
23.2.	Duplicate Effects	22
23.3.	IPsec	22
24.	Acknowledgements	22
25.	IANA Considerations	23
26.	Security Considerations	23
27.	Informative References	23

[1.](#) Introduction

Taking into consideration the IPv4 public address pool currently available at the Internet Assigned Numbers Authority (IANA), it is expected that the Regional Internet Registries (RIRs) will have no more public IPv4 addresses to allocate in the short term. At the time of writing, this foreseen date is mid-2012. See the IPv4 Address Report website (available at <http://www.potaroo.net/tools/ipv4/index.html>) for a thorough analysis of this issue, and an updated prediction.

This exhaustion phenomenon has been anticipated for a long time by the IETF, with the specification of the IPv6 protocol as the IPv4 successor. IPv6 increases the IPv4 addressing space by a power-of-four factor. IPv6 specifications are mature and current standardization effort is exclusively dedicated to operational aspects. Unfortunately, IPv6 was not devised as backwards compatible with IPv4; it is not possible to have IPv6-based applications directly exchange IP packets with their IPv4-based counterparts. The expected transition process was to assume hosts and routers will be Dual-Stack, that is, will support the two distinct protocol families IPv4 and IPv6, and to wait until the entire Internet supports Dual-Stack connectivity to deprecate IPv4. More than a decade later, IPv4-only communications are still the norm, and IPv6 is still rather marginal. ISPs will need to continue to provide IPv4 Internet access to their customers at the exhaustion date, if they do not want to see their business completely stalled or decreasing. They will wind up with public address pools that cannot grow. They will have to adapt to this situation, and enter an IPv4 address shortage management phase.

This document analyses the main issues related to IPv4 Internet access in the context of public IPv4 address exhaustion. Another complementary study can be found in [\[I-D.ford-shared-addressing-issues\]](#).

[2.](#) Shared IPv4 Addresses

So far, the current practice (particularly for fixed service offering) has been to give a unique IPv4 public address to each customer. A common design is to assign this global IPv4 address to the Customer Premises Equipment (CPE), and to assign IPv4 private addresses to the hosts connected behind the CPE. A private to public (and vice versa) translation occurs in the CPE owing to the activation of a Network Address and Port Translator (NAPT) function [[RFC3022](#)]. In this context, the public addresses that can be seen in any IP packets always refer to a unique customer. To cope with the

IPv4 address exhaustion, this kind of practices is no more affordable. Therefore, ISPs are bound to share the same IPv4 public address among several customers at the same time.

All IPv4 address shortage mechanisms extend the address space in adding port information. We call them by the generic name of IPv4 shortage solutions, or shared address solutions. IPv4 shortage solutions differ on the way they manage the port value. In this new context, a public IPv4 address seen in an IP packet can refer to several customers. The port information must be considered as well, in order to be able to unambiguously identify the customer pointed by that shared address. In particular, the port information along with the address information, must eventually be taken into account in order to correctly reach the intended destination.

[3.](#) Solution Space

[3.1.](#) CGN-based Solutions

These solutions propose the introduction of a NAPT function in the ISP network, denoted also as Carrier Grade NAT (CGN), or Large Scale NAT (LSN) [[I-D.nishitani-cgn](#)], or Provider NAT. The CGN is responsible for translating private addresses to publicly routable addresses. Private addresses are assigned to customers, a pool of public addresses is assigned to the CGN, the number of public addresses is much smaller than the number of customers. A public address of the CGN pool will therefore be shared by several customers at the same time.

Packet processing is as follows (for port-based communications):

- o Outgoing packets from an internal host to an Internet host: the internal host sends a packet with a private IPv4 source address. This packet goes up to the CGN, possibly after a first private to private address translation in the CPE, the CGN dynamically replaces the private source address by a public source address and modifies the source port value. The CGN creates an entry in its NAT mapping (or binding) table for each new mapping. Dynamic mappings can only be created by outgoing packets.
- o Incoming packets from an Internet host to an internal host: the remote host sends a packet with a destination address within the CGN public address pool. If, and only if, a mapping already exists for the (destination address, destination port, protocol) tuple, the CGN can reverse the mapping and forward the packet towards the internal side. However, even if this mapping exists, the CGN may discard the packet due to a particular filtering

policy. If no mapping exists for the (destination address, destination port, protocol), the CGN discards the packet.

CGN-based solutions can be deployed in different network configurations. One outstanding flavor is the DS-lite CGN [[I-D.ietf-softwire-dual-stack-lite](#)]: there is no NAT in the CPEs, IPv6 addresses (or prefixes) are assigned to CPEs, no IPv4 addresses are assigned to CPEs, traffic is tunnelled between CPEs and the DS-lite CGN into IPv6.

[3.2.](#) A+P Solutions

These solutions avoid the presence of a CGN function. They assign the same IP public address to several customers at the same time (shared address). They also assign a restricted port range to each customer so that two customers with the same IP address have two different port ranges that do not overlap. These solutions are called A+P (Address+Port) [[I-D.ymbk-aplusp](#)], or Port Range [[I-D.boucadair-port-range](#)], or SAM (Stateless Address Mapping) [[I-D.despres-sam](#)]. These solutions introduce a new function in the ISP network called Port Range Router (PRR).

Packet processing is as follows (for port-based communications):

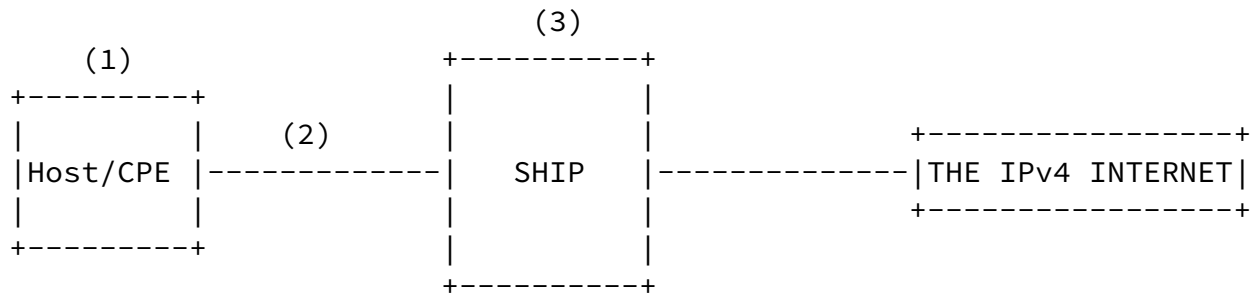
- o Outgoing packets from an internal host to an Internet host: the internal host sends a packet with its public shared address as the source address and a source port within his dedicated port range (the internal host may also send a packet with a private address, and then a NAPT in the CPE forces the source port to be within the port range and translates the private source address into the shared public address allocated). No specific handling is necessary in the ISP network, apart from the traditional IP routing process. In order to allow internal communications (within the same ISP realm), the outgoing packets have however to go up to the device that embeds the PRR function, but the packet will be processed by this PRR, if and only if, the PRR handles the destination subnet.
- o Incoming packets from an Internet host to an internal host: the remote host sends a packet with a shared destination public address. Owing to routing protocol advertisements, this packet is routed up to the PRR that handles this address subnet. The PRR is in charge to find a route towards the actual internal correspondent -for instance, CPEs have setup PPP sessions with the PRR, the PRR maintains a binding table between (IPv4 address, Port Range) couples and PPP session IDs-. The PRR that handles the shared address of a user must be in the data path of any packet intended to that user.

Port range solutions are sometimes described as a CGN function that would have been distributed among the CPEs. This is not false, but may be a rather misleading view, in the sense that it would assume CGN has not disappeared but simply melt into the CPEs, and that, to some extent, it is still CGN. Actually, the very purpose of A+P solutions is to completely get rid of any CGN function. One main interest of A+P solutions is that PRR devices are simple equipment which, unlike CGN devices, have no per-user session processing, and do not modify packet headers. Thus, PRR performance should not be an issue.

[3.3.](#) Common Architecture

To provide a common architecture, we introduce the SHared Ip Processor (SHIP) function embedded in a SHIP device. A SHIP device

can host either a CGN function or a PRR function, or both. In the latter case, some traffic is CGN-processed (outgoing and incoming packets), some is PRR-processed (incoming packets). There can be several SHIP devices in the same ISP network.



Shared Address Architecture

This architecture is composed of the following elements:

(1) The hosts are directly connected (e.g. mobile terminal), or connected through a CPE. For PRR-based solutions, packets are sent with source ports within their allocated Port Ranges. A Host/CPE must know a SHIP device to which it can send its traffic.

(2) The network infrastructure between CPEs and PRR:

- o Ensures outgoing packets are routed up to the SHIP device;
- o Ensures incoming packets are routed up to the intended Host/CPE.

(3) The SHIP device may:

- o Translate (NAPT) outgoing and incoming packets for CGN-processed traffic (CGN function only);

- o Find a route to incoming packets for PRR-processed traffic (PRR function only).

Many architectural issues are common to all IPv4 shortage solutions, whether they rely on the presence of a CGN or a PRR. Including, for instance, means to ensure proper routing between customers and SHIP devices, and SHIP devices location. Therefore, it does make sense to describe and specify all solutions on a concerted basis, and not in

completely separate ways.

4. Address Space Multiplicative Factor

The purpose of sharing public IPv4 addresses is to potentially increase the addressing space. A key parameter is the factor by which ISPs want or need to multiply their IPv4 public address space; and the consequence is the number of customers sharing the same public IPv4 address. This parameter is called the address space multiplicative factor, the inverse is called the compression ratio.

It is expected that IPv6 traffic will take an increasing part during the next years to come, at the expense of IPv4 traffic. We should reach a safety point in the future, where the number of IPv4 public addresses, in use at the same time, begins decreasing. From an ISP point of view, the multiplicative factor must be enough to survive until this event occurs for its own customers.

The multiplicative factor can only be applied to the part of customers that is eligible to shared address. The reasons a customer cannot have a shared address can be:

- o It would not be compatible with the service he has currently subscribed to (for example: business customer).
- o His CPE does not allow the solution selected by the ISP (for example it does not handle port restriction for PRR-based solutions or it does not allow IPv4 in IPv6 encapsulation for DS-lite solution), and its replacement is not easy.

Different ISPs may have very different needs. A long-lived ISP, whose number of customers is rather stable, will have an existing address pool that will only need a small extension to cope with the next years to come (small multiplicative factor, less than 10). A new comer will need a much bigger multiplicative factor (e.g. 1000). A mobile operator may see its address need explode if the IP mobile handset market dramatically grows, with Internet 3G access and Internet wireless (Wi-Fi/WiMAX) hotspot access.

The multiplicative factor is an important criterion in order to

select a solution. When the multiplicative factor is large, the average number of ports per customer is small; in that case, it is essential to manage port attribution the closest to the need as possible, and to avoid to dedicate ports to people who do not use them, when other people are running out of ports. Then, the larger the multiplicative factor is, the more dynamic the port assignment has to be.

CGN solutions provide the most dynamic port assignment scheme, hence the best possible ratio. However, CGNs can also be produced with functionalities that reduce their dynamicity; for instance, a port range can be attributed to each user, the CGN mapping is constrained into this range, in order, for example, to ease legal storage. As for A+P solutions, their dynamicity can vary a lot, depending on the Port Range assignment policy. Port Ranges can be assigned in a complete static way; one Port Range is assigned once and for all to each customer, no port can be added in, or removed from, the attributed Range. On the opposite, Port Range assignment can be fully dynamic; small Port Ranges (possibly up to one port) are dynamically assigned and removed.

Whatever the multiplicative factor selected, one must keep in mind that trying to deploy solutions that increase the IPv4 space by a big factor, might be seen as an incentive to postpone again and again IPv6 deployment.

5. Number of Current Sessions

In any kind of solutions, the number of current sessions per customer has, de facto, to be limited in some way. Therefore, the number of current sessions per customer is a limit to take into account in any architectural dimensioning. According to [\[RFC2663\]](#) terminology: "A session is defined as the set of traffic that is managed as a unit for translation. TCP/UDP sessions are uniquely identified by the tuple of (source IP address, source TCP/UDP port, target IP address, target TCP/UDP port)".

A CGN must sustain the traffic that will occur in its location point at the most loaded time, in terms of number of current sessions, and number of new sessions per second (i.e. when these parameter values are the highest). A Port Range solution should be less sensitive than a CGN to session rate and number, because a PRR is a per-user, and not a per-session, process.

UDP sessions and TCP sessions are separately handled in a CGN (a CGN mapping incorporates the protocol value); it is as if the same public

IPv4 address pool was allocated twice: once for the TCP communications and once for the UDP communications (of course other Transport protocols may be taken into account such as SCTP or DCCP). Therefore, the number of public addresses needed in a CGN will be driven by the type of communications -TCP or UDP- which consumes the highest number of public addresses. This consumption depends on the number of current sessions. Actually, for a Full Cone CGN (Endpoint-Independent Mapping and End point-Independent Filtering [[RFC4787](#)]), the address consumption should directly depend on the number of current couples (customer source address, customer source port) rather than on the number of current sessions.

For the same traffic conditions, a Port Range solution should need more addresses than a CGN. Let's assume, for the sake of simplicity, that port range allocation is a static one-size fits all. The range size should not only cover the average number of current sessions per user at peak time, but should also take into account the distribution of sessions, and should be significantly larger than the average if the standard deviation is large. So, where a CGN consumes no more ports than the number of current sessions, a Port Range solution, due to the attribution of ports by blocks, consumes more ports, and thus more addresses.

The IPv6 increasing deployment should have a decreasing effect on the number of current IPv4 sessions per customer. If the percentage of end-to-end IPv6 traffic significantly increases, so that the volume of IPv4 traffic begins decreasing, then the number of IPv4 sessions will be decreasing. The smaller the number of current IPv4 sessions per customer is, the higher the number of customers under the same IPv4 public address can be (better compression ratio), and consequently, the lower the number of IPv4 public addresses is needed. Hence, the pressure on IPv4 address shortage would be relaxed, because one IPv4 public address would be able to potentially serve more customers. However, this effect will only occur for customers who have both an IPv6 access and a shared IPv4 access. This advocates the strategy to systematically bound a shared IPv4 access to any IPv6 access. It is difficult to foresee to what extent the IPv6 traffic will decrease the number of current IPv4 sessions, but in any case, IPv6 activation should increase the potential IPv4 multiplicative factor. If it does not, that means IPv6 does not take over IPv4.

As for the current usage of ports, several hundreds as peak numbers per customer, seems current practice, although several thousands may be not unusual with some P2P applications (e.g. BitTorrent). The degree of fairness -balanced distribution of sessions between

customers-, traffic loss due to the limitation in number of sessions, may vary according to the traffic conditions, and the policy

enforced. The knowledge of the distribution of sessions among a set of users, makes it possible to know how many users might be adversely affected if/when system limits are reached. However, it is surely not that straightforward to really assess the degradation the customers should experience, from the knowledge of their session number limitations.

6. Service Management

At the time of IPv4 address exhaustion in the RIRs, ISPs will have to manage public address pools that cannot grow (at least from the RIRs). Concretely, they will have to decide to whom they allocate shared addresses and to whom they allocate unique public addresses, to the extent of the availability of addresses. Many policies can be envisaged, taking into account parameters such as: old vs. new customers, user profile, access type, geographic considerations, unique address as the privileged choice, shared address as the privileged choice, etc.

Care must be taken when considering the ratio that reflects the number of customers who will share a given global IPv4 address, not only to preserve some flexibility on the global address space that is left, but also to make sure that the ISP can adequately serve customer's requirements, without degrading the services they have subscribed to. ISPs can adjust the volume of IPv4 public addresses available playing on the balance between shared and unique allocations.

- o To increase the public IPv4 address pool: increase the number of customers with shared address; increase the ratio of customers per shared address.
- o To decrease the public IPv4 address pool: decrease the number of customers with shared address; decrease the ratio of customers per shared address.

7. IPv6 Migration and IPv4-IPv6 Coexistence

IPv6 is the only solution to solve the IPv4 public address exhaustion, that is, to actually get rid of the limitation on the number of IP addresses. IPv4 shared addresses are not candidate to IPv6 replacement. However, we should be very careful; whatever the network model deployed, applications and business will run on top of it. If we do not want to see IPv4 address shortage mechanisms postpone IPv6 deployment, all Internet actors must adopt a voluntary position towards IPv6.

Any IPv4 address shortage solution should make use, as much as possible, of the IPv6 transport capabilities available, in order to increase the IPv6 traffic and to move forward from an IPv4-enabled ISP network towards an almost IPv6-only ISP network. If it is not the case, the risk is to delay IPv6 operational deployments, in staying on a pure Dual-Stack attitude for ever, similar to the ships in the night routing approach, where the protocols independently live their own lives. IPv4 in IPv6 tunnels, and/or NAT64 and NAT46 translations should be favored. However, increasing the number of IPv6 packets does not automatically mean IPv6 is being generalized, if the main purpose of these packets is to carry IPv4 information. This is very similar to what occurred with ATM, especially in European countries, where ATM cells have heavily been used to convey IPv4 packets in the backhaul networks, but have never been used for end-to-end communications!

Some public IPv4 addresses will be required to connect IPv4 and IPv6 realms, through IPv4-IPv6 translators, for the sake of global reachability.

IPv4 shortage solutions may interfere with exiting IPv4 to IPv6 transition mechanisms, which were not designed with IPv4 shortage considerations. With A+P for instance, incoming 6to4 packets should be able to find their way from a 6to4 Relay up to the appropriate 6to4 CPE router, despite the lack of direct port range information (UDP/TCP initial source port did not pass through the CPE Port Range NAT translation process). One solution would be that a 6to4 IPv6 address embeds not only an IPv4 address but also a Port Range value.

[8.](#) Network Addressing Capability

The network addressing capability is the level of flexibility the network has to configure customers' devices, either with a unique public IPv4 address, or with a shared public IPv4 address. It may be assessed through the following considerations:

- o Is it possible to configure any customer's device with a shared address, regardless his location and his history?
- o Is it possible to configure any customer's device with a unique public address, regardless his location and his history?
- o Is it straightforward to switch, for any customer, from a shared address to a unique public address, and vice versa?

What is considered here is not the policy decision to allocate a unique or a shared address, but the network capability to enforce

such address management schemes.

9. Scarcity of Private Addressing

In several IPv4 shortage scenarios, private addresses, (as defined in [\[RFC1918\]](#)), can be assigned to CPEs, and to SHIP devices. This can be applied, for instance, to CGN double NAT architecture. In this case, if there are intermediate routers between CPEs and CGN, the outgoing packets are encapsulated into IPv4 packets addressed to a CGN private address (tunneling), and the incoming packets are natively routed towards the CPEs -the routing infrastructure must be able to route the CPEs' private addresses-. If there are no intermediate routers, no tunnel is necessary. However, private addresses are already in use in many ISPs' networks, they belong to a finite space which may rapidly raise overlapping issues as both the number of customers and the number of services that can be subscribed by these customers increase. As a consequence, some ISPs use Virtual Private Networks (VPNs) such as [\[RFC4364\]](#) to allow reusing the same private addresses several times with no routing overlaps. This brings a lot of complexity in network configuration and management.

It has been suggested to make the 240/4 block available for private addressing [\[I-D.wilson-class-e\]](#). This address block, formerly designated as "Class E", is still noted as being reserved in the IANA

IPv4 address registry. If it were reassigned for private addressing that would yield around 268 millions extra private addresses. However, many current implementations of the TCP/IP protocol stack do not allow the use of the 240/4 block. This is a severe blocking point for a lot of existing devices: CPE, NAT or routers. This issue will only be solved when the vendors' implementations accept the (240/4) addresses.

Another suggestion [[I-D.shirasaki-isp-shared-addr](#)] is to reserve some public blocks (typically three or four /8) only for internal usage. So far, there has been no consensus upon this proposal.

10. Scalability

The number of state entries in the ISP equipment (mainly the SHIP devices) can have a significant impact on performance, scalability, and solution cost. CGNs need to store many highly dynamic user session states. Basically, PRRs will keep information about Port Range attribution; they will not need to store a lot of states, the states dynamicity will depend on the assignment policy enforced by ISPs.

IPv4 shortage solutions must be able to adapt to the different ISP needs and be flexible enough to cover their evolutions. Customers under shared addresses can be geographically disseminated in very different ways: scattered vs. localised, sparse vs. dense. In term of architecture two types of responses are possible for the placement of the SHIP devices: close to the users (many devices that federate few customers) vs. close to the core (few devices that federate many customers).

A+P solutions should be rather flexible and scalable. A PRR treatment is a light process and should be straightforward to implement, a PRR function could then be implemented in almost any router devices. From this viewpoint, it will be rather comfortable for a network manager to activate the PRR functions he wants, when he needs and where he needs, with no big extra CAPEX (CAPital EXpenditure) increase. For the same number of customers federated, a CGN device should be more expensive than a PRR device. That could make CGN solutions less flexible and scalable than A+P solutions.

Another important scalability parameter is the impact on the CPEs. Some IPv4 shortage mechanisms require specific features in the CPEs. In that case, IPv4 shared addresses will not be able to spread to current CPEs that lack these appropriate features. Some CPEs are ISP branded which makes them, on the one hand, easier to control and to enhance, but which can be, on the other hand, very expensive for ISPs if a lot of legacy CPEs need to be replaced (if a software update is not enough).

A+P solutions require that CPEs receive specific Port Ranges (e.g. through DHCP), and that they control and restrict the source port of outgoing packets, according to the assigned Port Ranges. Supplementary CPE resources (hardware and software) to run these two features should be tiny (if not null) because, for example, in Linux OS the port range restriction is already part of the OS (Netfilter/Iptables) and the DHCP process is already in the CPE, and would only need to be complemented with port range negotiation as proposed in [\[I-D.bajko-pripaddrassign\]](#) for DHCPv4, and in [\[I-D.boucadair-dhcpv6-shared-address-option\]](#) for DHCPv6.

Some CGN solutions can be used with no specific features in the CPEs, just adding a second NAT level in the ISP network, to the extent that routing is correctly achieved between CPEs and CGN. The DS-lite CGN flavor requires CPEs to be NATless and to have a tunnel interface IPv4 into IPv6 (with several possible encapsulation types) to communicate with the DS-lite CGN device. A+P IPv6 variant also requires the same kind of IPv4 into IPv6 encapsulation.

11. Impact on Information System

IPv4 shortage solutions will have an impact on the Information System platforms and applications handling the administrative and technical information to control the activation of services (service ordering and service handling functions), and to manage the customer profiles. The possibility to give either a unique or a shared address, coupled or not with an IPv6 address, could yield several types of customers to deal with: IPv4 unique only, IPv4 shared only, IPv4 unique + IPv6, IPv4 shared + IPv6, IPv6 only.

Common practice used to rely upon the global IPv4 address assigned to a CPE device for customer identification purposes. The forthcoming address depletion therefore encourages ISPs to revisit their customer identification schemes since global IPv4 addresses will be shared amongst several customers. This clearly advocates for an IPv6-based customer identification scheme and thus impacts the way customer-specific management policies are enforced.

Additionally, "Service and Network Assurance" functions may be impacted by the introduction of SHIP function. Impacts should be assessed. Moreover, dedicated interface to access the CPE when no IPv4 address is assigned (e.g. DS-lite) should be defined and implemented (preferably such access should migrate towards IPv6).

12. Impact on Services

There is a potential danger for the following types of applications:

- o Applications that establish inbound communications;
- o Applications that carry address information in their payload;
- o Applications that carry port information in their payload;
- o Applications that use fixed ports (e.g. well known);
- o Applications that do not use any port (e.g. ICMP);
- o Applications that assume the uniqueness of customers' addresses (e.g. IP address as identifier);
- o Applications that explicitly prohibit twice the same address to access to a resource at the same time.

Current applications already implement some mechanisms in order to circumvent the presence of NATs (typically CPE NATs):

- o ALGs;
- o Port Forwarding;

- o UPnP IGD;
- o NAT-PMP;
- o NAT Traversal techniques: ICE, STUN, TURN, etc.

It should be considered to what extent these mechanisms can still be used with IPv4 shortage mechanisms.

Impact on existing services:

- o Will this service work as usual?
- o Will this service work but with a degradation?
- o What level of degradation?
- o Will this service not work at all?
- o What modifications are needed if any?

Impact on future services:

- o What new constraints are to be taken into account to devise new services?

CGN solutions should have more impact on applications than A+P solutions. Basically, from an end-to-end perspective, A+P solutions should not be perceived by applications, to the extent of the port range limitation, whereas a CGN should be perceived as a supplementary blocking mechanism that must be circumvented by specific techniques and protocols. This is particularly sensitive for P2P applications, even if Full Cone CGN behavior should alleviate some problems.

[13.](#) Port Space Boundaries

Most of the time the source port issued by a client application will be translated, apart from a direct knowledge of a Port Range restriction by the client's stack (A+P), either by a NAT in the user's device (A+P), or by a CPE NAT (A+P), or by a CPE NAT and/or a CGN NAT (CGN).

IANA has classified the whole port space in three categories (as defined in <http://www.iana.org/assignments/port-numbers>):

- o The Well Known Ports are those from 0 through 1023.
- o The Registered Ports are those from 1024 through 49151.
- o The Dynamic and/or Private Ports are those from 49152 through 65535.

[RFC4787] notices that current NATs have different policies with regard to this classification; some NATs restrict their translations to the use of dynamic ports, some also include registered ports, some preserve the port range if it is in the well-known range. [RFC4787] makes it clear that the use of port space [1024, 65535] is safe: "mapping a source port to a source port that is already registered is unlikely to have any bad effects". Therefore, for both A+P and CGN solutions, there is no reason to only consider a subset of the port space [1024, 65535] for outgoing source ports. In any case, limiting the number of ports available will limit the compression ratio.

The problem is trickier for Well Known Ports. For outgoing source ports, [RFC4787] points out that "certain applications expect the source UDP port to be in the well-known range", hence, it can be safe to keep the source port in the Well Known Ports in that case. This is possible for a CGN, to the extent that this port is still available; this is not possible for A+P if the well known port value is not in the attributed Port Range. Apart from port preservation, the use of Well Known Ports should be prohibited for outgoing source ports, because some applications will not work (for instance MSN messenger cannot sign in).

For inbound communications, it is interesting to be able to use a destination port within the Well Known Ports, in order to reach a server hosted by a customer (however, this constraint can be alleviated with the use of SRV records [RFC2782]). This possibility is limited. It is unlikely, for example, to satisfy all customers who may request port 80. For a CGN, a given customer will get port 80 if there is still one address of the CGN public pool where port 80 is currently not used, with the restriction that if this customer already uses an external public address, it can be harmful to give him a second current external public address (this is referenced in behave WG documents as "IP address pooling behavior" of "arbitrary" vs. "paired"). For A+P, a given customer will only get port 80 if port 80 is within his Port Range.

[14.](#) Flow Discrimination

ISPs can offer walled garden services along with Internet services. ISPs may want these flows not to traverse the IPv4 shortage facilities put in place. For instance, all the IPv4 traffic does not have to be processed by a CGN facility, for various reasons that are mostly ISP policy-specific and which include -but are not necessarily limited to- performance considerations, service-specific forwarding policies. It should be clear how these IPv4 flows can bypass the IPv4 shortage facilities and how they can be handled by the corresponding service platform/gateway. However, the best practice seems to rapidly migrate these services from IPv4 to IPv6.

[15.](#) Impact on Intra-Domain and Inter-Domain Routing

The introduction of port consideration to route packets to their final destinations may have an impact on the current routing infrastructure: on the architecture, the IGP and EGP configuration, the addressing configuration, and on routers performances.

The introduction of new nodes that cannot be circumvent could also yield non optimized paths, especially for communications between customers attached to the same ISP realm.

Centralized SHIP devices could also strongly modify the current flow distribution scheme among the different links and nodes, and then lead to non optimal paths.

[16.](#) Fragmentation

When a packet is fragmented, the port information (either UDP or TCP) will only be present in the first fragment. The other fragments will not bear the port information which is necessary to a correct treatment up to the destination. Dedicated means to ease fragmented packets routing should be activated.

[17.](#) QoS

[17.1.](#) QoS performance

The possible degradation of end-to-end performances (e.g. delay) experienced in the context of IPv4 shortage solutions should be evaluated.

Levis, et al.

Expires December 24, 2009

[Page 17]

Internet-Draft

IPv4 Address Shortage

June 2009

[17.2.](#) QoS mechanisms

The impact on QoS mechanisms should be investigated. In particular the ability to classify traffic in order to apply differentiated treatments could be hindered by the fact that an IPv4 address is shared among several users, possibly in a dynamic way. In particular, transparent DSCP handling should be supported by SHIP devices.

[17.3.](#) Introduction of Single Point of Failure (Robustness)

The introduction of new nodes/functions, specifically where the port information is managed, can create single points of failure. Any IPv4 shortage solution should consider the opportunity to add redundancy features in order to alleviate the impact on the robustness of the offered IP connectivity service.

Additionally, load balancing and load sharing means should be evaluated. The ability of the solution to allow hot swapping from a machine to another, in minimizing the perturbations, should be considered.

For CGN solutions, the ability to switch from a CGN node to another one without losing active sessions, might be rather complex to achieve due to the need to keep the two devices synchronized with the same active session states. For the A+P solutions such hot swapping is clearly achievable because an A+P node in the network does not maintain any session-based states; thus, fail-over means would be lightweight.

[18.](#) Support of Multicast

It should be assessed if a customer with a shared address can receive multicast packets and source multicast packets.

Particularly, impact on IGMP should be identified and solutions proposed. Because of the presence of several end-user devices with the same IP address, membership to multicast groups should be evaluated and enhancement should be proposed if required. Besides the membership issue, building multicast trees may be impacted. This impact should be assessed and alternatives proposed.

[19.](#) Mobile-IP

Owing to the deployment of a Mobile-IP architecture, a mobile terminal continues to access its connectivity service when visiting a

Levis, et al.

Expires December 24, 2009

[Page 18]

Internet-Draft

IPv4 Address Shortage

June 2009

Foreign Network. In order to avoid traffic loss, it is recommended to use the home address (HoA), and not the care-of address (CoA), to reach that mobile terminal. A dedicated entity called HA (Home Agent) is responsible for routing the traffic according to the binding table it maintains. This table includes in particular the association between the HoA and CoA. A Foreign Agent (FA) can optionally be deployed in the visited network. If an IP address is shared (in the home network or/and in the visited network), HA or FA must be updated so as to take into account the port information to achieve its operations (i.e. relay traffic destined to HoA to the current CoA).

[20.](#) End-Users Facilities

In the current deployments, end-users are used to configuring their CPEs in order to control the traffic entering/exiting their home LAN.

One important and very used feature is the ability to open ports (port forwarding) either manually, or with a protocol such as UPnP IGD. If a CGN is present, the port must also be open in the CGN. However, ISPs may not be very incline to see their customers configure some specific parameters in a device inside their networks. Furthermore, any supplementary treatment in the NAT process is also prone to decrease the overall CGN performances. The situation may be alleviated if the CGN architecture is composed of only one NAT level

(no NAT in the CPE) as for DS-lite. If all NATs are Full Cone the need for port forwarding may be less critical, especially to allow P2P communications. For A+P the port forwarding capability may still be used at CPE level, with the limitation that the port open must be within the Port Range (for instance no possibility to open port 80, if port 80 is not in the allocated Port Range). UPnP IGD version 2 should, on this matter, facilitate the Port Range working, in allowing CPEs to allocate another port number than the one first requested by the terminals.

The use of the DNS SRV records [[RFC2782](#)] could be the solution to host servers in customers' premises under shared addresses. SRV records give the possibility to specify a port value related to a service, and then allow services to be accessible on ports which are not Well Known ports (e.g. a web server accessible on a port different from 80).

Another widely used feature is the ability to store specific ALGs on the CPE to allow applications to correctly behave despite the presence of the NAT CPE (even if it is harmful and should be avoided, to the extent possible, according to behave WG recommendations). When the CPE belongs to the customer, the customer has all

flexibility to tailor the device to his needs, if the CPE belongs to the ISP, the customer depends on the ISP good will to satisfy his request. For CGNs, it would be difficult to customize the CGN with specific ALGs coming from specific customers' requirements, the customers should wind up with only a limited set of possibilities if any. For A+P, ALGs should work as usual, and have the same possibilities as today, possibly taking into account the limited port choice. Like current deployment, and in the context of A+P, the resources required to run ALGs concern the CPE and no network nodes.

[21.](#) Management Tools

ISPs deploy a set of tools and applications for the management of their infrastructure, especially for supervision purposes. Impact on these tools should be evaluated and solutions proposed when required. Particularly, means to assign IP connectivity information, means to monitor the overall network, to assess the reachability of devices should be specified. In this context, impact on tools (e.g. ICMP-

based) to check the reachability of network nodes should be evaluated.

22. Legal Obligations

ISPs can be required by governmental and/or regulation authorities to provide customer-specific information upon request.

22.1. Traceability

Legal obligations require an ISP to provide the identity of a customer upon request of the authorities. Because one public IPv4 address may be shared between several customers, the knowledge of the IP address is not enough to have a chance to find the appropriate customer. The legal request should include the information: [IP address - Port - Protocol- Begin_Timestamp - End_Timestamp].

A CGN must record and store all mappings (typically during 6 months to one year, depending on the legislation) that it creates. The piece of information to store by mapping is two-part: Index, and Customer_Discriminator.

- o Index: is the information that will match legal request input. It is composed of: [Public IP address - Public port - Protocol - Date of mapping creation - Date of mapping deletion].
- o Customer_Discriminator: is the Information that will identify the customer for a given index value. The customer discriminator

depends on the kind of CGN solution put in place. For a DS-lite CGN the customer discriminator is his IPv6 prefix, for other CGN architectures it can be, for instance and among other possibilities, an IPv4 private address, a PPP session ID.

The complete DS-lite storage tuple is: [IPv6 prefix - Public IP address - Public port - Protocol - Date of mapping creation - Date of mapping deletion] per mapping. It should not be necessary to store the private address and private port. The ISP should be responsible for resolving the customer identity: who the ISP has an agreement with, but not the identity of who in the customer's house was actually connected (which is of the customer responsibility).

If we consider one mapping per session (in fact it should be less for a Full Cone, one mapping per couple as seen in [Section 5](#)) an ISP should record and retain traces of all sessions created by all customers during one year (if the legal storage duration is one year). This can be a problem not only as a big volume of data to store, but also as a big volume of data to constantly transfer from the CGN to the storage place.

A PRR has only to keep one entry [Public IP address - Port Range -Date of beginning of Port Range assignment-Date of end of Port Range assignment] per customer. Legal storage should not be a main issue for A+P solutions, especially if Port Range assignment remains rather static.

[22.2.](#) Interception

This process is proactive, a given group of communications is replicated in real time towards a law enforcement agency. Typically, the point of replication is the first IP hop in the ISP network. Wiretapping techniques need to be transparent to the customer, so that the targeted customer cannot be aware of the interception, owing to tools such as traceroute that would make appear modification on the path. They even need to be transparent to network administration team, and need special login with special privileges only accessible to authorized members.

[23.](#) Security

[23.1.](#) Port Randomization

A kind of blind attacks that can be performed against TCP relies on the attacker's ability to guess the 5-tuple (Protocol, Source Address, Destination Address, Source Port, Destination Port) that identifies the transport protocol instance to be attacked. Document

[I-D.ietf-tsvwg-port-randomization] describes a number of methods for the random selection of the client port number, such that the possibility of an attacker guessing the exact value is reduced. With shared IPv4 addresses, the port selection space is reduced. This issue seems more severe for A+P solutions than for CGN; Intuitively,

assuming the port range is known, the smaller the port range is, the more predictable the port choice is.

It should be noted that to guess the port information may not be sufficient to carry out a successful blind attack. The exact TCP Sequence Number (SN) should also be known, a TCP segment is processed only if all previous segments have been received, except for some Reset Segment implementations which immediately process the Reset as long as it is within the Window. If SN is randomly chosen it will be difficult to guess it (SN is 32 bits long); port randomization is one protection among others against blind attacks.

[23.2.](#) Duplicate Effects

Some types of attacks that have an impact on a targeted IPv4 public address, could see their effects increased by the number of customers who share this address. For example, if a given address that has, deliberately or not misbehaved, is consequently forbidden to access some resources, the whole set of customers who share this address will be impacted. Application developers should find alternative information to uniquely identify connected users.

[23.3.](#) IPsec

Even if IPsec is not deployed for mass market (e.g. residential), impacts of solutions based on shared IP addresses should be evaluated and assessed. [[RFC3947](#)] proposes a solution to solve issues documented in [[RFC3715](#)]. The applicability of [[RFC3947](#)] in the context of shared IP address should be evaluated.

[24.](#) Acknowledgements

We are grateful to Christian Jacquenet, Iain Calder, and Marcelo Bagnulo, for their helpful comments and suggestions for improving this document.

This document has also been deeply improved by the fruitful exchanges with all people who have actively participated in the proposal of IPv4 shortage solutions.

[25.](#) IANA Considerations

There are no IANA considerations.

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

[26.](#) Security Considerations

Security considerations are discussed in the Security section

[27.](#) Informative References

[I-D.bajko-pripaddrassign]

Bajko, G., Savolainen, T., Boucadair, M., and P. Levis, "Port Restricted IP Address Assignment", [draft-bajko-pripaddrassign-01](#) (work in progress), March 2009.

[I-D.boucadair-dhcpv6-shared-address-option]

Boucadair, M., Levis, P., Grimault, J., Savolainen, T., and G. Bajko, "Dynamic Host Configuration Protocol (DHCPv6) Options for Shared IP Addresses Solutions", [draft-boucadair-dhcpv6-shared-address-option-00](#) (work in progress), May 2009.

[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 Mapping (SAM) Avoiding NATs and restoring the end-to-end model in IPv6", [draft-despres-sam-02](#) (work in progress), March 2009.

[I-D.ford-shared-addressing-issues]

Durand, A., Ford, M., and P. Roberts, "Issues with ISP Responses to IPv4 Address Exhaustion", [draft-ford-shared-addressing-issues-00](#) (work in progress), March 2009.

[I-D.ietf-software-dual-stack-lite]

Durand, A., Droms, R., Haberman, B., and J. Woodyatt, "Dual-stack lite broadband deployments post IPv4

Internet-Draft

IPv4 Address Shortage

June 2009

exhaustion", [draft-ietf-softwire-dual-stack-lite-00](#) (work in progress), March 2009.

[I-D.ietf-tsvwg-port-randomization]

Larsen, M. and F. Gont, "Port Randomization", [draft-ietf-tsvwg-port-randomization-03](#) (work in progress), March 2009.

[I-D.nishitani-cgn]

Nishitani, T., Miyakawa, S., Nakagawa, A., and H. Ashida, "Common Functions of Large Scale NAT (LSN)", [draft-nishitani-cgn-02](#) (work in progress), June 2009.

[I-D.shirasaki-isp-shared-addr]

Shirasaki, Y., Miyakawa, S., Nakagawa, A., Yamaguchi, J., and H. Ashida, "ISP Shared Address", [draft-shirasaki-isp-shared-addr-02](#) (work in progress), March 2009.

[I-D.wilson-class-e]

Wilson, P., Michaelson, G., and G. Huston, "Redesignation of 240/4 from "Future Use" to "Private Use"", [draft-wilson-class-e-02](#) (work in progress), September 2008.

[I-D.ymbk-aplusp]

Bush, R., Maennel, O., Zorz, J., Bellovin, S., and L. Cittadini, "The A+P Approach to the IPv4 Address Shortage", [draft-ymbk-aplusp-03](#) (work in progress), March 2009.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", [BCP 5](#), [RFC 1918](#), February 1996.

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

[RFC2782] Gulbrandsen, A., Vixie, P., and L. Esibov, "A DNS RR for specifying the location of services (DNS SRV)", [RFC 2782](#), February 2000.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", [RFC 3022](#), January 2001.

[RFC3715] Aboba, B. and W. Dixon, "IPsec-Network Address Translation

Levis, et al.

Expires December 24, 2009

[Page 24]

Internet-Draft

IPv4 Address Shortage

June 2009

(NAT) Compatibility Requirements", [RFC 3715](#), March 2004.

[RFC3947] Kivinen, T., Swander, B., Huttunen, A., and V. Volpe, "Negotiation of NAT-Traversal in the IKE", [RFC 3947](#), January 2005.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), February 2006.

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

Authors' Addresses

Pierre Levis (editor)
France Telecom
42 rue des Coutures
BP 6243
Caen Cedex 4 14066
France

Email: pierre.levis@orange-ftgroup.com

Mohamed Boucadair
France Telecom

Email: mohamed.boucadair@orange-ftgroup.com

Jean-Luc Grimault
France Telecom

Email: jeanluc.grimault@orange-ftgroup.com

Alain Villefranque
France Telecom

Email: alain.villefranque@orange-ftgroup.com

Levis, et al.

Expires December 24, 2009

[Page 25]