Issues with IP Address Sharing
draft-ietf-intarea-shared-addressing-issues-05

Abstract

The completion of IPv4 address allocations from IANA and the RIRs is
causing service providers around the world to question how they will
continue providing IPv4 connectivity service to their subscribers when
there are no longer sufficient IPv4 addresses to allocate them one per
subscriber. Several possible solutions to this problem are now emerging
based around the idea of shared IPv4 addressing. These solutions give
rise to a number of issues and this memo identifies those common to all
such address sharing approaches. Such issues include application
failures, additional service monitoring complexity, new security
vulnerabilities and so on. Solution-specific discussions are out of
scope.
Deploying IPv6 is the only perennial way to ease pressure on the public
IPv4 address pool without the need for address sharing mechanisms that
give rise to the issues identified herein.

Status of this Memo

Table of Contents

1. Introduction

Allocations of IPv4 addresses from the Internet Assigned Numbers
Authority (IANA) were completed on February 3, 2011 [IPv4_Pool].
Allocations from Regional Internet Registries (RIRs) are anticipated to
be complete around a year later, although the exact date will vary from
registry to registry. This is causing service providers around the
world to start to question how they will continue providing IPv4
connectivity service to their subscribers when there are no longer
sufficient IPv4 addresses to allocate them one per subscriber. Several
possible solutions to this problem are now emerging based around the
idea of shared IPv4 addressing. These solutions give rise to a number
of issues and this memo identifies those common to all such address
sharing approaches and collects them in a single document.
Deploying IPv6 is the only perennial way to ease pressure on the public
IPv4 address pool without the need for address sharing mechanisms that
give rise to the issues identified herein. In the short term,
maintaining growth of IPv4 services in the presence of IPv4 address
depletion will require address sharing. Address sharing will cause
issues for end-users, service providers and third parties such as law
enforcement agencies and content providers. This memo is intended to
highlight and briefly discuss these issues.
Increased IPv6 deployment should reduce the burden being placed on an
address sharing solution, and should reduce the costs of operating that
solution. Increasing IPv6 deployment should cause a reduction in the
number of concurrent IPv4 sessions per subscriber. 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
decrease. The smaller the number of concurrent IPv4 sessions per
subscriber, the higher the number of subscribers able to share the same
IPv4 public address, and consequently, the lower the number of IPv4
public addresses required. However, this effect will only occur for
subscribers who have both an IPv6 access and a shared IPv4 access. This
motivates a strategy to systematically bind a shared IPv4 access to an
IPv6 access. It is difficult to foresee to what extent growing IPv6
traffic will reduce the number of concurrent IPv4 sessions, but in any
event, IPv6 deployment and use should reduce the pressure on the
available public IPv4 address pool.

2. Shared Addressing Solutions

In many networks today a subscriber is provided with a single public
IPv4 address at their home or small business. For instance, in fixed
broadband access, an IPv4 public address is assigned to each CPE
(Customer Premises Equipment). CPEs embed a NAT function which is

responsible for translating private IPv4 addresses ([RFC1918] addresses) assigned to hosts within the local network, to the public IPv4 address assigned by the service provider (and vice versa). Therefore, devices located with the LAN share the single public IPv4 address and they are all associated with a single subscriber account and a single network operator.

A number of proposals currently under consideration in the IETF rely upon the mechanism of multiplexing multiple subscribers' connections over a smaller number of shared IPv4 addresses. This is a significant change. These proposals include Carrier Grade NAT (CGN. a.k.a., LSN for Large Scale NAT) [I-D.ietf-behave-lsn-requirements], Dual-Stack Lite [I-D.ietf-softwire-dual-stack-lite], NAT64 [I-D.ietf-behave-v6v4-xlate-stateful] [I-D.ietf-behave-v6v4-xlate], Address+Port (A+P) proposals [I-D.ymbk-aplusp], [I-D.boucadair-port-range] and SAM [I-D.despres-sam]. Section 26 provides a classification of these different types of solutions and discusses some of the design considerations to be borne in mind when deploying large-scale address sharing. Whether we're talking about DS-Lite, A+P, NAT64 or CGN isn't especially important - it's the view from the outside that matters, and given that, most of the issues identified below apply regardless of the specific address sharing scenario in question.

In these new proposals, a single public IPv4 address would be shared by multiple homes or small businesses (i.e., multiple subscribers) so the operational paradigm described above would no longer apply. In this document we refer to this new paradigm as large-scale address sharing. All these proposals extend the address space by adding port information, they differ in the way they manage the port value.

Security issues associated with NAT have long been documented (see [RFC2663] and [RFC2993]). However, sharing IPv4 addresses across multiple subscribers by any means, either moving the NAT functionality from the home gateway to the core of the service provider network, or restricting the port choice in the subscriber's NAT, creates additional issues for subscribers, content providers and network operators. Many of these issues are created today by public Wi-Fi hotspot deployments. As such large-scale address sharing solutions become more widespread in the face of IPv4 address depletion, these issues will become both more severe and more commonly felt. NAT issues in the past typically only applied to a single legal entity; as large-scale address sharing becomes more prevalent these issues will increasingly span across multiple legal entities simultaneously.

All large-scale address sharing proposals share technical and operational issues and these are addressed in the sections that follow. These issues are common to any service-provider NAT, enterprise NAT, and also non-NAT solutions that share individual IPv4 addresses across multiple subscribers. This document is intended to bring all of these issues together in one place.

3. Analysis of Issues as they Relate to First and Third Parties

In this section we present an analysis of whether the issues identified in the remainder of this document are applicable to the organization deploying the shared addressing mechanism (and by extension their subscribers) and/or whether these issues impact third parties (e.g., content providers, law enforcement agencies, etc.). In this analysis, issues that affect end-users are deemed to affect 1st parties, as end-users can be expected to complain to their service provider when problems arise. Where issues can expect to be foreseen and addressed by the party deploying the shared addressing solution, they are not attributed.
In Figure 1 we have also tried to indicate (with 'xx') where issues are newly created in addition to what could be expected from the introduction of a traditional NAT device. Issues marked with a single 'x' are already present today in the case of typical CPE NAT, however they can be expected to be more severe and widespread in the case of large-scale address sharing.

```
+------------------------------------------------+--------+---------+
|                    Issue                       |  1st   |   3rd   |
|                                                | party  | parties |
+------------------------------------------------+--------+---------+
| Restricted allocations of outgoing             |   x    |         |
| ports will impact performance for end users    |        |         |
|                                                |        |         |
| Incoming port negotiation mechanisms may fail  |   xx   |         |
|                                                |        |         |
| Incoming connections to Assigned Ports will    |   x    |         |
| not work                                       |        |         |
|                                                |        |         |
| Port discovery mechanisms will not work        |   x    |         |
|                                                |        |         |
| Some applications will fail to operate         |   x    |    x    |
|                                                |        |         |
| Assumptions about parallel/serial connections  |   x    |    x    |
| may fail                                       |        |         |
|                                                |        |         |
+------------------------------------------------+--------+---------+


+------------------------------------------------+--------+---------+
|                    Issue                       |  1st   |   3rd   |
|                                                | party  | parties |
+------------------------------------------------+--------+---------+
| TCP control block sharing will be affected     |   x    |    x    |
|                                                |        |         |
| Reverse DNS will be affected                   |   x    |    x    |
|                                                |        |         |
| Inbound ICMP will fail in many cases           |   x    |    x    |
|                                                |        |         |
| Amplification of security issues               |   xx   |   xx    |
|                                                |        |         |
| Fragmentation will require special handling    |   x    |         |
|                                                |        |         |
| Single points of failure and increased         |   x    |         |
| network instability                            |        |         |
|                                                |        |         |
| Port randomization will be affected            |   x    |         |
|                                                |        |         |
| Service usage monitoring and abuse logging     |   xx   |   xx    |
| will be impacted for all elements in the chain |        |         |
| between service provider and content provider  |        |         |
|                                                |        |         |
| Penalty boxes will no longer work              |   xx   |   xx    |
|                                                |        |         |
| Spam blacklisting will be affected             |   xx   |   xx    |
|                                                |        |         |
| Geo-location services will be impacted         |   xx   |   xx    |
```

```
|                                                   |        |         |
| Geo-proximity mechanisms will be impacted         |   xx   |   xx    |
|                                                   |        |         |
| Load balancing algorithms may be impacted         |        |   xx    |
|                                                   |        |         |
| Authentication mechanisms may be impacted         |        |    x    |
|                                                   |        |         |
| Traceability of network usage and abusage will    |        |   xx    |
| be affected                                       |        |         |
|                                                   |        |         |
| IPv6 transition mechanisms will be affected       |   xx   |         |
|                                                   |        |         |
| Frequent keep-alives reduce battery life          |    x   |         |
|                                                   |        |         |
+---------------------------------------------------+--------+---------+
```

As can be seen from Figure 1, deployment of large-scale address sharing will create almost as many problems for third parties as it does for the service provider deploying the address sharing mechanism. Several of these issues are specific to the introduction of large-scale address sharing as well. All of these issues are discussed in further detail below.

4. Content Provider Example

Taking a content provider as an example of a third-party, and focusing on the issues that are created specifically by the presence of large-scale address sharing, we identify the following issues as being of particular concern:

   *Degraded geolocation for targeted advertising and licensed
    content restrictions (see Section 7).

   *Additional latency due to indirect routing and degraded
    geoproximity (see Section 7).

   *Exposure to new amplification attacks (see Section 13).

   *Service usage monitoring is made more complicated (see Section
    8).

   *Incoming port negotiation mechanisms may fail (see Section
    5.2.1).

   *IP blocking for abuse/spam will cause collateral damage (see
    Section 13).

   *Load balancing algorithms may be impacted (see Section 16).

*Traceability of network usage and abuse will be impacted (see
   [Section 12](#)).

[5.](#) Port Allocation

When we talk about port numbers we need to make a distinction between
outgoing connections and incoming connections. For outgoing
connections, the actual source port number used is usually irrelevant.
(While this is true today, in a port-range solution it is necessary for
the source port to be within the allocated range). But for incoming
connections, the specific port numbers allocated to subscribers matter
because they are part of external referrals (used by third parties to
contact services run by the subscribers).
The total number of subscribers able to share a single IPv4 address
will depend upon assumptions about the average number of ports required
per active subscriber, and the average number of simultaneously active
subscribers. It is important to realize that the TCP design makes it
undesirable for clients to re-use ports while they remain in the TIME-
WAIT state (typically 4 minutes after the connection has concluded).
TIME-WAIT state removes the hazard of old duplicates for "fast" or
"long" connections, in which clock-driven Initial Sequence Number
selection is unable to prevent overlap of the old and new sequence
spaces. The TIME-WAIT delay allows all old duplicate segments time
enough to die in the Internet before the connection is reopened
[RFC1337]. Therefore ports in this state must be included in
calculations concerning port usage per subscriber.
Most of the time the source port selected by a client application will
be translated (unless there is direct knowledge of a port-range
restriction in the client's stack), either by a NAT in the subscriber's
device, or by a CPE NAT, or by a CPE NAT and a CGN.
[RFC1700] defines the Assigned Ports (both System and User). IANA has
further classified the whole port space into three categories, as
defined in [IANA_Ports]

    *The Well-Known Ports are those from 0 through 1023.

    *The Registered Ports are those from 1024 through 49151.

    *The Dynamic and/or Private Ports are those from 49152 through
     65535.

[RFC4787] notes 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 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 all address sharing solutions, there is no

reason to only consider a subset of the port space (1024-65535) for outgoing source ports.

5.1. Outgoing Ports

According to measurements the average number of outgoing ports consumed per active subscriber is much, much smaller than the maximum number of ports a subscriber can use at any given time. However, the distribution is heavy-tailed, so there are typically a small number of subscribers who use a very high number of ports [CGN_Viability]. This means that an algorithm that dynamically allocates outgoing port numbers from a central pool will typically allow more subscribers to share a single IPv4 address than algorithms that statically divide the resource by pre-allocating a fixed number of ports to each subscriber. Similarly, such an algorithm should be more able to accommodate subscribers wishing to use a relatively high number of ports.
It is important to note here that the desire to dynamically allocate outgoing port numbers will make a service provider's job of maintaining records of subscriber port number allocations considerably more onerous (see Section 12). The number of records per subscriber will increase from 1 in a scheme where ports are statically allocated, to a much larger number equivalent to the total number of outgoing ports consumed by that subscriber during the time period for which detailed logs must be kept.
A potential problem with dynamic allocation occurs when one of the subscriber devices behind such a port-shared IPv4 address becomes infected with a worm, which then quickly sets about opening many outbound connections in order to propagate itself. Such an infection could rapidly exhaust the shared resource of the single IPv4 address for all connected subscribers. It is therefore necessary to impose limits on the total number of ports available to an individual subscriber to ensure that the shared resource (the IPv4 address) remains available in some capacity to all the subscribers using it. However, static schemes for ports assignment may introduce security issues [RFC6056] when small contiguous port ranges are statically assigned to subscribers. Another way to mitigate this issue is to kill off (randomly) established connections when the port space runs out. A user with many connections will be proportionally more likely to get impacted.
Session failure due to NAT state overflow or timeout (when the NAT discards session state because it's run out of resource) can be experienced when the configured quota per user is reached or if the NAT is out of recourses.

5.2. Incoming Ports

It is desirable to ensure that incoming ports remain stable over time. This is challenging as the network doesn't know anything in particular about the applications that it is supporting and therefore has no real

notion of how long an application/service session is still ongoing and
therefore requiring port stability.
Early measurements [CGN Viability] also seem to indicate that, on
average, only very few ports are used by subscribers for incoming
connections. However, a majority of subscribers accept at least one
inbound connection.
This means that it is not necessary to pre-allocate a large number of
incoming ports to each subscriber. It is possible to either pre-
allocate a small number of ports for incoming connections or do port
allocation on demand when the application wishing to receive a
connection is initiated. The bulk of incoming ports can be reserved as
a centralized resource shared by all subscribers using a given public
IPv4 address.

5.2.1. Port Negotiation

In current deployments, one important and widely used feature of many
CPE devices is the ability to open incoming ports (port forwarding)
either manually, or with a protocol such as Universal Plug and Play
Internet Gateway Device (UPnP IGD) [UPnP-IGD]. If a CGN is present, the
port must also be opened in the CGN. CGN makes subscribers dependent on
their service provider for this functionality.
If the CPE and the CGN are required to co-operate in order for port
forwarding functionality to work, protocol development will be required
to provide an automated solution. If the CGN architecture is composed
of only one NAT level (no NAT in the CPE) as for DS-Lite, the service
provider will still be required to offer some means for configuring
incoming ports by their subscribers. This may be either via a PCP [I-
D.ietf-pcp-base], UPnP or NAT-PMP proxy over a tunneled direct
connection between the CPE and CGN, or a web interface to configure the
incoming port mapping on the CGN. Note, that such an interface
effectively makes public what was previously a private management
interface and this raises security concerns that must be addressed.
For port-range solutions, port forwarding capabilities may still be
present at the CPE, with the limitation that the open incoming port
must be within the allocated port-range (for instance it is not
possible to open port 5002 for incoming connections if port 5002 is not
within the allocated port-range).

5.2.1.1. Universal Plug and Play (UPnP)

Using the UPnP semantic, an application asks "I want to use port number
X, is that OK?" and the answer is yes or no. If the answer is no, the
application will typically try the next port in sequence, until it
either finds one that works or gives up after a limited number of
attempts. UPnP IGD 1.0 has no way to redirect the application to use
another port number. UPnP IGD 2.0 improves this situation and allows
for allocation of any available port.

5.2.1.2. NAT Port Mapping Protocol (NAT-PMP)

NAT-PMP enables the NAT to redirect the requesting application to a port deemed to be available for use by the NAT state mapping table.

5.2.2. Connection to a Well-Known Port Number

Once an IPv4 address sharing mechanism is in place, inbound connections to well-known port numbers will not work in the general case. Any application that is not port-agile cannot be expected to work. Some workaround (e.g., redirects to a port-specific URI) could be deployed given sufficient incentives. There exist several proposals for 'application service location' protocols which would provide a means of addressing this problem, but historically these proposals have not gained much deployment traction.
For example, the use of DNS SRV records [RFC2782] provides a potential solution for subscribers wishing to host services in the presence of a shared-addressing scheme. SRV records make it possible to specify a port value related to a service, thereby making services accessible on ports other than the Well-Known ports. It is worth noting that this mechanism is not applicable to HTTP and many other protocols.

5.2.3. Port Discovery Mechanisms

Port discovery using a UDP port to discover a service available on a corresponding TCP port, either through broadcast, multicast or unicast, is a commonly deployed mechanism. Unsolicited inbound UDP will be dropped by address sharing mechanisms as they have no live mapping to enable them to forward the packet to the appropriate host. Address sharing thereby breaks this service discovery technique.

6. Impact on Applications

Address sharing solutions will have an impact on the following types of applications:

> *Applications that establish inbound communications - these
> applications will have to ensure that ports selected for inbound
> communications are either within the allocated range (for port-
> range solutions) or are forwarded appropriately by the CGN (for
> CGN-based solutions). See Section 5.2 for more discussion of
> this;

> *Applications that carry address and/or port information in their
> payload - where translation of port and/or address information is
> performed at the IP and transport layers by the address sharing
> solution, an ALG will also be required to ensure application
> layer data is appropriately modified. Note that ALGs are required
> in some cases, and in many other cases end-to-end protocol
> mechanisms have developed to work around a lack of ALGs in

address translators, to the point of it being preferable to avoid
any support in the NAT;

   *Applications that use fixed ports - see Section 5.2.2 for more
    discussion of this;

   *Applications that do not use any port (e.g., ICMP echo) - will
    require special handling - see Section 9 for more discussion of
    this;

   *Applications that assume the uniqueness of source addresses
    (e.g., IP address as identifier) - such applications will fail to
    operate correctly in the presence of multiple, discrete,
    simultaneous connections from the same source IP address;

   *Applications that explicitly prohibit concurrent connections from
    the same address - such applications will fail when multiple
    subscribers sharing an IP address attempt to use them
    simultaneously.

   *Applications that do not use TCP or UDP for transport - All IP
    address sharing mechanisms proposed to date are limited to TCP,
    UDP, and ICMP, thereby preventing end users from fully utilizing
    the Internet (e.g., SCTP, DCCP, RSVP, protocol 41 (IPv6-over-
    IPv4), protocol 50 (IPsec ESP)).

Applications already frequently implement mechanisms in order to
circumvent the presence of NATs (typically CPE NATs):

   *Application Layer Gateways (ALGs): Many CPE devices today embed
    ALGs that allow applications to behave correctly despite the
    presence of NAT on the CPE. When the NAT belongs to the
    subscriber, the subscriber has flexibility to tailor the device
    to his or her needs. For CGNs, subscribers will be dependent on
    the set of ALGs that their service provider makes available. For
    port-range solutions, ALGs will require modification to deal with
    the port-range restriction, but will otherwise have the same
    capabilities as today. Note that ALGs are required in some cases,
    and in many other cases end-to-end protocol mechanisms have
    developed to work around lack of ALGs, to the point of it being
    preferable to avoid any support in the NAT.

   *NAT Traversal Techniques: There are several commonly-deployed
    mechanisms that support operating servers behind a NAT by
    forwarding specific TCP or UDP ports to specific internal hosts
    ([UPnP-IGD], [I-D.cheshire-nat-pmp], and manual configuration).
    All of these mechanisms assume the NAT's WAN address is a
    publicly-routable IP address, and fail to work normally when that
    assumption is wrong. There have been attempts to avoid that
    problem by automatically disabling the NAT function and bridging

traffic instead upon assignment of a private IP address to the
        WAN interface (as is required for [Windows-Logo] certification).
        Bridging (rather than NATting) has other side effects (DHCP
        requests are served by an upstream DHCP server which can increase
        complexity of in-home networking).

7. Geo-location and Geo-proximity

IP addresses are frequently used to indicate, with some level of
granularity and some level of confidence, where a host is physically
located. Using IP addresses in this fashion is a heuristic at best, and
is already challenged today by other deployed capabilities, e.g.,
tunnels. Geo-location services are used by content providers to allow
them to conform with regional content licensing restrictions, to target
advertising at specific geographic areas, or to provide customized
content. Geo-location services are also necessary for emergency
services provision. In some deployment contexts (e.g., centralized
CGN), shared addressing will reduce the level of confidence and level
of location granularity that IP-based geo-location services can
provide. Viewed from the content provider, a subscriber behind a CGN
geolocates to wherever the prefix of the CGN appears to be; very often
that will be in a different city than the subscriber.
IP addresses are also used as input to geolocation services that
resolve an IP address to a physical location using information from the
network infrastructure. Current systems rely on resources such as
RADIUS databases and DHCP lease tables. The use of address sharing will
prevent these systems from resolving the location of a host based on IP
address alone. It will be necessary for users of such systems to
provide more information (e.g., TCP or UDP port numbers), and for the
systems to use this information to query additional network resources
(e.g., NAT-PT binding tables). Since these new data elements tend to be
more ephemeral than those currently used for geolocation, their use by
geolocation systems may require them to be cached for some period of
time.
Other forms of geo-location will still work as usual.
A slightly different use of an IP address is to calculate the proximity
of a connecting host to a particular service delivery point. This use
of IP address information impacts the efficient delivery of content to
an end-user. If a CGN is introduced in communications and it is far
from an end-user connected to it, application performance may be
degraded insofar as IP-based geo-proximity is a factor.

8. Tracking Service Usage

As large-scale address sharing becomes more commonplace, monitoring the
number of unique users of a service will become more complex than
simply counting the number of connections from unique IP addresses.
While this is a somewhat inexact methodology today due to the
widespread use of CPE NAT, it will become a much less useful approach

in the presence of widespread large-scale address sharing solutions. In general, all elements that monitor usage or abusage in the chain between a service provider that has deployed address sharing and a content provider will need to be upgraded to take account of the port value in addition to IP addresses.

9. ICMP

ICMP does not include a port field and is consequently problematic for address sharing mechanisms. Some ICMP message types include a fragment of the datagram that triggered the signal to be sent, which is assumed to include port numbers. For some ICMP message types, the Identifier field has to be used as a de-multiplexing token. Sourcing ICMP echo messages from hosts behind an address sharing solution does not pose problems, although responses to outgoing ICMP echo messages will require special handling, such as making use of the ICMP identifier value to route the response appropriately.
For inbound ICMP there are two cases. The first case is that of ICMP sourced from outside the network of the address sharing solution provider. Where ICMP messages include a fragment of an outgoing packet including port numbers it may be possible to forward the packet appropriately. In addition to these network signaling messages, several applications (e.g., P2P applications) make use of ICMP echo messages which include no hints that could be used to route the packet correctly. Measurements derived by such applications in the presence of an address sharing solution will be erroneous or fail altogether. The second case is that of ICMP sourced from within the network of the address sharing solution provider (e.g., for network management, signaling and diagnostic purposes). In this case ICMP can be routed normally for CGN-based solutions owing to the presence of locally unique private IP addresses for each CPE device. For port-range solutions, ICMP echo messages will not be routable without special handling, e.g., placing a port number in the ICMP identifier field, and having port-range routers make routing decisions based upon that field. Considerations related to ICMP message handling in NAT-based environments are specified in [RFC5508].

10. MTU

In applications where the end hosts are attempting to use path MTU Discovery [RFC1191] to optimize transmitted packet size with underlying network MTU, shared addressing has a number of items which must be considered. As covered in Section 9, ICMP "Packet Too Big" messages must be properly translated through the address sharing solution in both directions. However, even when this is done correctly, MTU can be a concern. Many end hosts cache PMTUd information for a certain period of time. If the MTU behind the address sharing solution is inconsistent, the public end host may have the incorrect MTU value cached. This may cause it to send packets that are too large, causing

them to be dropped if the DF (Don't Fragment) bit is set, or causing
them to be fragmented by the network, increasing load and overhead.
Because the host eventually will reduce MTU to the lowest common value
for all hosts behind a given public address, it may also send packets
that are below optimal size for the specific connection, increasing
overhead and reducing throughput.
This issue also generates a potential attack vector, that a malevolent
user could send an ICMP "Packet Too Big" (Type 3, Code 4) message
indicating a Next-Hop MTU of anything down to 68 octets. This value
will be cached by the off-net server for all subscribers sharing the
address of the malevolent user. This could lead to a DoS against both
the remote server and the large-scale NAT device itself (as they will
both have to handle many more packets per second).

11. Fragmentation

When a packet is fragmented, transport-layer port information (either
UDP or TCP) is only present in the first fragment. Subsequent fragments
will not carry the port information and so will require special
handling. In addition, the IP Identifier may no longer be unique as
required by the receiver to aid in assembling the fragments of a
datagram.

12. Traceability

In many jurisdictions, service providers are legally obliged to provide
the identity of a subscriber upon request to the appropriate
authorities. Such legal requests have traditionally included the source
IPv4 address and date (and usually the time), which is sufficient
information when subscribers are assigned IPv4 addresses for a long
duration.
However, where one public IPv4 address is shared between several
subscribers, the IPv4 address no longer uniquely identifies a
subscriber. There are two solutions to this problem:

   *The first solution is for servers to additionally log the source
    port of incoming connections and for the legal request to include
    the source port. The legal request should include the
    information: [Source IP address, Source Port, Timestamp] (and
    possibly other information). Accurate time-keeping (e.g., use of
    NTP or SNTP) is vital because port assignments are dynamic. A
    densely populated CGN could mean even very small amounts of clock
    skew between a third party's server and the CGN operator will
    result in ambiguity about which customer was using a specific
    port at a given time.

   *The second solution considers it is unrealistic to expect all
    servers to log the source port number of incoming connections. To

deal with this, service providers using IPv4 address sharing may
need to log IP destination addresses.

Destination logging is imperfect if multiple subscribers are accessing
the same (popular) server at nearly the same time, it can be impossible
to disambiguate which subscriber accessed the server, especially with
protocols that involve several connections (e.g., HTTP). Thus, logging
the destination address on the NAT is inferior to logging the source
port at the server.
If neither solution is used (that is, the server is not logging source
port numbers and the NAT is not logging destination IP addresses), the
service provider cannot trace a particular activity to a specific
subscriber. In this circumstance, the service provider would need to
disclose the identity of all subscribers who had active sessions on the
NAT during the time period in question. This may be a large number of
subscribers.
Address sharing solutions must record and store all mappings (typically
during 6-12 months, depending on the local jurisdiction) that they
create. If we consider one mapping per session, a service provider
should record and retain traces of all sessions created by all
subscribers during one year (if the legal storage duration is one
year). This may be challenging due to the volume of data requiring
storage, the volume of data to repeatedly transfer to the storage
location, and the volume of data to search in response to a query.
Address sharing solutions may mitigate these issues to some extent by
pre-allocating groups of ports. Then only the allocation of the group
needs to be recorded, and not the creation of every session binding
within that group. There are trade-offs to be made between the sizes of
these port groups, the ratio of public addresses to subscribers,
whether or not these groups timeout, the impact on logging requirements
and port randomization security [RFC6056].

13. Security

Before noting some specific security-related issues caused by large-
scale address sharing, it is perhaps worth noting that, in general,
address sharing creates a vector for attack amplification in numerous
ways. See Section 10 for one example.

13.1. Abuse Logging and Penalty Boxes

When an abuse is reported today, it is usually done in the form: IPv4
address X has done something bad at time T0. This is not enough
information to uniquely identify the subscriber responsible for the
abuse when that IPv4 address is shared by more than one subscriber. Law
enforcement authorities may be particularly impacted because of this.
This particular issue can be fixed by logging port numbers, although
this will increase logging data storage requirements.

A number of services on the network today log the IPv4 source addresses used in connection attempts to protect themselves from certain attacks. For example, if a server sees too many requests from the same IPv4 address in a short period of time, it may decide to put that address in a penalty box for a certain time during which requests are denied, or it may require completion of a CAPTCHA for future requests. If an IPv4 address is shared by multiple subscribers, this would have unintended consequences in a couple of ways. First it may become the natural behavior to see many login attempts from the same address because it is now shared across a potentially large number of subscribers. Second and more likely is that one user who fails a number of login attempts may block out other users who have not made any previous attempts but who will now fail on their first attempt. In the presence of widespread large-scale address sharing, penalty box solutions to service abuse simply will not work.
In addition, there are web tie-ins into different blacklists that web administrators subscribe to redirect users with infected machines (e.g., detect the presence of a worm) to a URL that says "Hey, your machine is infected!". With address sharing, someone else's worm can interfere with the ability to access the service for other subscribers sharing the same IP address.

13.2. Authentication

Simple address-based identification mechanisms that are used to populate access control lists will fail when an IP address is no longer sufficient to identify a particular subscriber. Including port numbers in access control list definitions may be possible at the cost of extra complexity, and may also require the service provider to make static port assignments, which conflicts with the requirement for dynamic assignments discussed in Section 5.1.
Address or DNS-name based signatures (e.g., some X.509 signatures) may also be affected by address sharing as the address itself is now a shared token, and the name to address mapping may not be current.

13.3. SPAM

Another case of identifying abusers has to do with SPAM blacklisting. When a spammer is behind a CGN or using a port-shared address, blacklisting of their IP address will result in all other subscribers sharing that address having their ability to source SMTP packets restricted to some extent.

13.4. Port Randomisation

A blind attack 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. [RFC6056] describes a

number of methods for the random selection of the source port number, such that the ability of an attacker to correctly guess the 5-tuple is reduced. With shared IPv4 addresses, the port selection space is reduced. Preserving port randomization is important and may be more or less difficult depending on the address sharing solution and the size of the port space that is being manipulated. Allocation of non-contiguous port ranges could help to mitigate this issue.
It should be noted that guessing the port information may not be sufficient to carry out a successful blind attack. An in-window TCP Sequence Number (SN) should also be known or guessed. 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. There is more detailed discussion of improving TCP's robustness to Blind In-Window Attacks in [RFC5961].

13.5. IPsec

The impact of large-scale IP address sharing for IPsec operation should be evaluated and assessed. [RFC3947] proposes a solution to solve issues documented in [RFC3715]. [RFC5996] specifies Internet Key Exchange (IKE) Protocol Version 2 which includes NAT traversal mechanisms that are now widely used to enable IPsec to work in the presence of NATs in many cases. Nevertheless, service providers may wish to ensure that CGN deployments do not inadvertently block NAT traversal for security protocols such as IKE (refer to [I-D.gont-behave-nat-security] for more information).

13.6. Policing Forwarding Behaviour

[RFC2827] motivates and discusses a simple, effective, and straightforward method for using ingress traffic filtering to prohibit Denial-of-Service (DoS) attacks which use forged IP addresses.
Following this recommendation, service providers operating shared-addressing mechanisms should ensure that source addresses, or source ports in the case of port-range schemes, are set correctly in outgoing packets from their subscribers or they should drop the packets.
If some form of IPv6 ingress filtering is deployed in the broadband network and DS-Lite service is restricted to those subscribers, then tunnels terminating at the CGN and coming from registered subscriber IPv6 addresses cannot be spoofed. Thus a simple access control list on the tunnel transport source address is all that is required to accept traffic on the internal interface of a CGN.

14. **Transport Issues**

14.1. **Parallel connections**

Systems that assume that multiple simultaneous connections to a single IP address implies connectivity to a single host - such systems may experience unexpected results.

14.2. **Serial connections**

Systems that assume that returning to a given IP address means returning to the same physical host, as with stateful transactions. This may also affect cookie-based systems.

14.3. **TCP Control Block Sharing**

[RFC2140] defines a performance optimization for TCP based on sharing state between TCP control blocks that pertain to connections to the same host, as opposed to maintaining state for each discrete connection. This optimization assumes that an address says something about the properties of the path between two hosts, which is clearly not the case if the address in question is shared by multiple hosts at different physical network locations. While CPE NAT today causes problems for sharing TCP control block state across multiple connections to a given IP address, large-scale address sharing will make these issues more severe and more widespread.

15. **Reverse DNS**

Many service providers populate forward and reverse DNS zones for the public IPv4 addresses that they allocate to their subscribers. In the case where public addresses are shared across multiple subscribers, such strings are, by definition, no longer sufficient to identify an individual subscriber without additional information.

16. **Load Balancing**

Algorithms used to balance traffic load for popular destinations may be affected by the introduction of address sharing. Where balancing is achieved by deterministically routing traffic from specific source IP addresses to specific servers, imbalances in load may be experienced as address sharing is enabled for some of those source IP addresses. This will require re-evaluation of the algorithms used in the load-balancing design. In general, as the scale of address sharing grows, load-balancing designs will need to be re-evaluated and any assumptions about average load per source IP address revisited.

[17.](#) IPv6 Transition Issues

IPv4 address sharing solutions may interfere with existing IPv4 to IPv6
transition mechanisms, which were not designed with IPv4 shortage
considerations in mind. With port-range solutions for instance,
incoming 6to4 packets should be able to find their way from a 6to4
relay 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 translation process). One solution would be
for a 6to4 IPv6 address to embed not only an IPv4 address but also a
port range value.
Subscribers allocated with private addresses will not be able to
utilize 6to4 [RFC3056] to access IPv6, but may be able to utilize
Teredo [RFC4380].
Some routers enable 6to4 on their WAN link. 6to4 requires a publicly-
routable IPv4 address. Enabling 6to4 when the apparently public IPv4
WAN address is in fact behind a NAT creates a disconnected IPv6 island.

[18.](#) Introduction of Single Points of Failure

In common with all deployments of new network functionality, the
introduction of new nodes or functions to handle the multiplexing of
multiple subscribers across shared IPv4 addresses could create single
points of failure in the network. Any IP address sharing solution
should consider the opportunity to add redundancy features in order to
alleviate the impact on the robustness of the offered IP connectivity
service. The ability of the solution to allow hot swapping from one
machine to another should be considered. This is especially important
where the address sharing solution in question requires the creation of
per-flow state in the network.

[19.](#) State Maintenance Reduces Battery Life

In order for hosts to maintain network state in the presence of NAT,
keep-alive messages have to be sent at frequent intervals. For battery-
powered devices, sending these keep-alive messages can result in
significantly reduced battery performance than would otherwise be the
case [Mobile Energy Consumption].

[20.](#) Support of Multicast

[RFC5135] specifies requirements for a NAT that supports Any Source IP
Multicast or Source-Specific IP Multicast. Port-range routers that form
part of port-range solutions will need to support similar requirements
if multicast support is required.

[21.](#) Support of Mobile-IP

IP address sharing within the context of Mobile-IP deployments (in the
home network and/or in the visited network), will require Home Agents

and/or Foreign Agents to be updated so as to take into account the relevant port information. There may also be issues raised when an additional layer of encapsulation is required thereby causing, or increasing the need for, fragmentation and reassembly.
Issues for Mobile-IP in the presence of NAT are discussed in [I-D.haddad-mext-nat64-mobility-harmful]

22. IANA Considerations

This memo includes no request to IANA.

23. Security Considerations

This memo does not define any protocol and therefore creates no new security issues. Section 13 discusses some of the security and identity-related implications of IP address sharing.

24. Contributors

This document is based on sources co-authored by J.L. Grimault and A. Villefranque of France Telecom.

25. Acknowledgments

This memo was partly inspired by conversations that took place as part of Internet Society (ISOC) hosted roundtable events for operators and content providers deploying IPv6. Participants in those discussions included John Brzozowski, Leslie Daigle, Tom Klieber, Yiu Lee, Kurtis Lindqvist, Wes George, Lorenzo Colliti, Erik Kline, Igor Gashinsky, Jason Fesler, Rick Reed, Adam Bechtel, Larry Campbell, Tom Coffeen, David Temkin, Pete Gelbman, Mark Winter, Will Charnock, Martin Levy, Greg Wood and Christian Jacquenet.
The authors are also grateful to Christian Jacquenet, Iain Calder, Joel Halpern, Brian Carpenter, Gregory Lebovitz, Bob Briscoe, Marcelo Bagnulo, Dan Wing and Wes George for their helpful comments and suggestions for improving the document.
This memo was created using the xml2rfc tool.

26. Annex

26.1. Classes of Address Sharing Solution

IP address sharing solutions fall into two classes. Either a service-provider operated NAT function is introduced and subscribers are allocated addresses from [RFC1918] space, or public IPv4 addresses are shared across multiple subscribers by restricting the range of ports

available to each subscriber. These classes of solution are described
in a bit more detail below.

> *CGN-based solutions: These solutions propose the introduction of
> a NAPT function in the service provider's network, denoted also
> as Carrier Grade NAT (CGN), or Large Scale NAT (LSN) [I-D.ietf-
> behave-lsn-requirements], or Provider NAT. The CGN is responsible
> for translating private addresses to publicly routable addresses.
> Private addresses are assigned to subscribers, a pool of public
> addresses is assigned to the CGN, and the number of public
> addresses is smaller than the number of subscribers. A public
> IPv4 address in the CGN pool is shared by several subscribers at
> the same time. Solutions making use of a service provider-based
> NAT include [I-D.shirasaki-nat444] (two layers of NAT) and [I-
> D.ietf-softwire-dual-stack-lite] (a single layer of NAT).

> *Port-range solutions: These solutions avoid the presence of a CGN
> function. A single public IPv4 address is assigned to several
> subscribers at the same time. A restricted port range is also
> assigned to each subscriber so that two subscribers with the same
> IPv4 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].

26.2. Address Space Multiplicative Factor

The purpose of sharing public IPv4 addresses is to increase the
addressing space. A key parameter is the factor by which service
providers want or need to multiply their IPv4 public address space; and
the consequence is the number of subscribers sharing the same public
IPv4 address. We refer to this parameter as the address space
multiplicative factor, the inverse is called the compression ratio.
The multiplicative factor can only be applied to the subset of
subscribers that are eligible for a shared address. The reasons a
subscriber cannot have a shared address can be:

> *It would not be compatible with the service they are currently
> subscribed to (for example: business subscriber).

> *Subscriber CPE is not compatible with the address sharing
> solution selected by the service provider (for example it does
> not handle port restriction for port-range solutions or it does
> not allow IPv4 in IPv6 encapsulation for the DS-Lite solution),
> and its replacement is not easy.

Different service providers may have very different needs. A long-lived
service provider, whose number of subscribers is rather stable, may
have an existing address pool that will only need a small extension to

cope with the next few years, assuming that this address pool can be re-purposed for an address sharing solution (small multiplicative factor, less than 10). A new entrant or a new line of business will need a much bigger multiplicative factor (e.g., 1000). A mobile operator may see its addressing needs grow dramatically as the IP-enabled mobile handset market grows.

When the multiplicative factor is large, the average number of ports per subscriber is small. Given the large measured disparity between average and peak port consumption [CGN Viability], this will create service problems in the event that ports are allocated statically. In this case, it is essential for port allocation to map to need as closely as possible, and to avoid allocating ports for longer than necessary. Therefore, the larger the multiplicative factor, the more dynamic the port assignment has to be.

27. References

[RFC1337]	Braden, B., "TIME-WAIT Assassination Hazards in TCP", RFC 1337, May 1992.
[RFC2140]	Touch, J., "TCP Control Block Interdependence", RFC 2140, April 1997.
[RFC2663]	Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999.
[RFC2993]	Hain, T., "Architectural Implications of NAT", RFC 2993, November 2000.
[RFC2782]	Gulbrandsen, A., Vixie, P. and L. Esibov, "A DNS RR for specifying the location of services (DNS SRV)", RFC 2782, February 2000.
[RFC5135]	Wing, D. and T. Eckert, "IP Multicast Requirements for a Network Address Translator (NAT) and a Network Address Port Translator (NAPT)", BCP 135, RFC 5135, February 2008.
[RFC2827]	Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.
[RFC3947]	Kivinen, T., Swander, B., Huttunen, A. and V. Volpe, "Negotiation of NAT-Traversal in the IKE", RFC 3947, January 2005.

[RFC1918]	Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G. and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.
[RFC3715]	Aboba, B. and W. Dixon, "IPsec-Network Address Translation (NAT) Compatibility Requirements", RFC 3715, March 2004.
[RFC4787]	Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.
[RFC3056]	Carpenter, B. and K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.
[RFC1191]	Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.
[I-D.gont-behave-nat-security]	Gont, F and P Srisuresh, "Security implications of Network Address Translators (NATs)", Internet-Draft draft-gont-behave-nat-security-03, October 2009.
[I-D.ietf-behave-lsn-requirements]	Perreault, S, Yamagata, I, Miyakawa, S, Nakagawa, A and H Ashida, "Common requirements for Carrier Grade NAT (CGN)", Internet-Draft draft-ietf-behave-lsn-requirements-04, October 2011.
[I-D.ietf-softwire-dual-stack-lite]	Durand, A, Droms, R, Woodyatt, J and Y Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", Internet-Draft draft-ietf-softwire-dual-stack-lite-11, May 2011.
[I-D.ymbk-aplusp]	Bush, R, "The A+P Approach to the IPv4 Address Shortage", Internet-Draft draft-ymbk-aplusp-10, May 2011.
[I-D.ietf-behave-v6v4-xlate-stateful]	Bagnulo, M, Matthews, P and I Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", Internet-Draft draft-ietf-behave-v6v4-xlate-stateful-12, July 2010.
[I-D.despres-sam]	Despres, R, "Scalable Multihoming across IPv6 Local-Address Routing Zones Global-Prefix/Local-Address

	Stateless Address Mapping (SAM)", Internet-Draft draft-despres-sam-03, July 2009.
[I-D.ietf-behave-v6v4-xlate]	Li, X, Bao, C and F Baker, "IP/ICMP Translation Algorithm", Internet-Draft draft-ietf-behave-v6v4-xlate-23, September 2010.
[I-D.shirasaki-nat444]	Yamagata, I, Shirasaki, Y, Nakagawa, A, Yamaguchi, J and H Ashida, "NAT444", Internet-Draft draft-shirasaki-nat444-04, July 2011.
[I-D.boucadair-port-range]	Boucadair, M, Levis, P, Bajko, G and T Savolainen, "IPv4 Connectivity Access in the Context of IPv4 Address Exhaustion: Port Range based IP Architecture", Internet-Draft draft-boucadair-port-range-02, July 2009.
[I-D.haddad-mext-nat64-mobility-harmful]	Haddad, W and C Perkins, "A Note on NAT64 Interaction with Mobile IPv6", Internet-Draft draft-haddad-mext-nat64-mobility-harmful-02, March 2011.
[I-D.cheshire-nat-pmp]	Cheshire, S, "NAT Port Mapping Protocol (NAT-PMP)", Internet-Draft draft-cheshire-nat-pmp-03, April 2008.
[IANA_Ports]	, , "IANA Port Number Assignments", February 2011.
[IPv4_Pool]	, , "Available Pool of Unallocated IPv4 Internet Addresses Now Completely Emptied", 2011.
[CGN_Viability]	Alcock, S, "Research into the Viability of Service-Provider NAT", 2008.
[Mobile_Energy_Consumption]	Haverinen, H, Siren, J and P Eronen, "Energy Consumption of Always-On Applications in WCDMA Networks", 2007.
[UPnP-IGD]	UPnP Forum, , "Universal Plug and Play (UPnP) Internet Gateway Device (IGD) V 2.0", December 2010.
[Windows-Logo]	Microsoft, , "Windows Logo Program Device Requirements", 2006.
[RFC5508]	Srisuresh, P., Ford, B., Sivakumar, S. and S. Guha, "NAT Behavioral Requirements for ICMP", BCP 148, RFC 5508, April 2009.

[RFC4380]	Huitema, C., "[Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)](#)", RFC 4380, February 2006.
[I-D.ietf-pcp-base]	Wing, D, Cheshire, S, Boucadair, M, Penno, R and P Selkirk, "[Port Control Protocol (PCP)](#)", Internet-Draft draft-ietf-pcp-base-17, October 2011.
[RFC1700]	[Reynolds, J.](#) and [J. Postel](#), "[Assigned Numbers](#)", RFC 1700, October 1994.
[RFC6056]	Larsen, M. and F. Gont, "[Recommendations for Transport-Protocol Port Randomization](#)", BCP 156, RFC 6056, January 2011.
[RFC5961]	Ramaiah, A., Stewart, R. and M. Dalal, "[Improving TCP's Robustness to Blind In-Window Attacks](#)", RFC 5961, August 2010.
[RFC5996]	Kaufman, C., Hoffman, P., Nir, Y. and P. Eronen, "[Internet Key Exchange Protocol Version 2 (IKEv2)](#)", RFC 5996, September 2010.

Authors' Addresses

Mat Ford editor Ford Internet Society Geneva, Switzerland EMail: ford@isoc.org

Mohamed Boucadair Boucadair France Telecom Rennes, 35000 France EMail: mohamed.boucadair@orange-ftgroup.com

Alain Durand Durand Juniper Networks EMail: adurand@juniper.net

Pierre Levis Levis France Telecom 42 rue des Coutures BP 6243 Caen Cedex 4, 14066 France EMail: pierre.levis@orange-ftgroup.com

Phil Roberts Roberts Internet Society Reston, VA, USA EMail: roberts@isoc.org