6MAN   B. Carpenter, Ed.
Internet-Draft   Univ. of Auckland
Intended status: Informational   T. Chown
Expires: April 5, 2015   Univ. of Southampton
                                                               F. Gont
                                                  SI6 Networks / UTN-FRH
                                                               S. Jiang
                                           Huawei Technologies Co., Ltd
                                                            A. Petrescu
                                                             CEA, LIST
                                                        A. Yourtchenko
                                                                  cisco
                                                       October 2, 2014

            Analysis of the 64-bit Boundary in IPv6 Addressing
                       draft-ietf-6man-why64-06

Abstract

   The IPv6 unicast addressing format includes a separation between the
   prefix used to route packets to a subnet and the interface identifier
   used to specify a given interface connected to that subnet.
   Currently the interface identifier is defined as 64 bits long for
   almost every case, leaving 64 bits for the subnet prefix. This
   document describes the advantages of this fixed boundary and analyses
   the issues that would be involved in treating it as a variable
   boundary.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   Rather than simply overcoming the IPv4 address shortage by doubling
   the address size to 64 bits, IPv6 addresses were originally chosen to
   be 128 bits long to provide flexibility and new possibilities. In
   particular, the notion of a well-defined interface identifier was

added to the IP addressing model. The IPv6 addressing architecture
[RFC4291] specifies that a unicast address is divided into n bits of
subnet prefix followed by (128-n) bits of interface identifier (IID).
The bits in the IID have no meaning and the entire identifier should
be treated as an opaque value [RFC7136]. Also, since IPv6 routing is
entirely based on variable length prefixes (also known as variable
length subnet masks), there is no basic architectural assumption that
n has any particular fixed value. All IPv6 routing protocols support
prefixes of any length up to /128.

The IID is of basic importance in the IPv6 stateless address
autoconfiguration (SLAAC) process [RFC4862]. However, it is
important to understand that its length is a parameter in the SLAAC
process, and it is determined in a separate link-type specific
document (see Section 2 of RFC 4862). The SLAAC protocol does not
define its length or assume any particular length. Similarly, DHCPv6
[RFC3315] does not include a prefix length in its address assignment.

The notion of a /64 boundary in the address was introduced after the
initial design of IPv6, following a period when it was expected to be
at /80. There were two motivations for setting it at /64. One was
the original "8+8" proposal [DRAFT-odell] that eventually led to ILNP
[RFC6741], which required a fixed point for the split between local
and wide-area parts of the address. The other was the expectation
that EUI-64 MAC addresses would become widespread in place of 48-bit
addresses, coupled with the plan at that time that auto-configured
addresses would normally be based on interface identifiers derived
from MAC addresses.

As a result, RFC 4291 describes a method of forming interface
identifiers from IEEE EUI-64 hardware addresses [IEEE802] and this
specifies that such interface identifiers are 64 bits long. Various
other methods of forming interface identifiers also specify a length
of 64 bits. The addressing architecture, as modified by [RFC7136],
states that "For all unicast addresses, except those that start with
the binary value 000, Interface IDs are required to be 64 bits long.
If derived from an IEEE MAC-layer address, they must be constructed
in Modified EUI-64 format." The de facto length of almost all IPv6
interface identifiers is therefore 64 bits. The only documented
exception is in [RFC6164], which standardises 127-bit prefixes for
point-to-point links between routers, among other things to avoid a
loop condition known as the ping-pong problem.

With that exception, and despite the comments above about the routing
architecture and the design of SLAAC, using an IID shorter than 64
bits and a subnet prefix longer than 64 bits is outside the current
IPv6 specifications, so results may vary.

The question is often asked why the subnet prefix boundary is set
rigidly at /64. The first purpose of this document is to explain the
advantages of the fixed IID length. Its second purpose is to analyse
in some detail the effects of hypothetically varying the IID length.
The fixed length limits the practical length of a routing prefix to
64 bits, whereas architecturally, and from the point of view of
routing protocols, it could be any value up to /128, as for host
routes. Whatever the length of the IID, the longest match is done on
the concatenation of prefix and IID. Here, we mainly discuss the
question of a shorter IID, which would allow a longer subnet prefix.
The document makes no proposal for a change to the IID length.

The following three sections describe in turn the advantages of the
fixed length IID, some arguments for shorter lengths, and the
expected effects of varying the length.

2. Advantages of a fixed identifier length

As mentioned in Section 1, the existence of an IID of a given length
is a necessary part of IPv6 stateless address autoconfiguration
(SLAAC) [RFC4862]. This length is normally the same for all nodes on
a given link that is running SLAAC. Even though this length is a
parameter for SLAAC, determined separately for the link layer media
type of each interface, a globally fixed IID length for all link
layer media is the simplest solution, and is consistent with the
principles of Internet host configuration described in [RFC5505].

An interface identifier of significant length, clearly separated from
the subnet prefix, makes it possible to limit the traceability of a
host computer by varying the identifier. This is discussed further
in Section 4.5.

An interface identifier of significant length guarantees that there
are always enough addresses in any subnet to add one or more real or
virtual interfaces. There might be other limits, but IP addressing
will never get in the way.

The addressing architecture [RFC4291] [RFC7136] sets the IID length
at 64 bits for all unicast addresses, and therefore for all media
supporting SLAAC. An immediate effect of fixing the IID length at 64
bits is, of course, that it fixes the subnet prefix length also at 64
bits, regardless of the aggregate prefix assigned to the site
concerned, which in accordance with [RFC6177] should be /56 or
shorter. This situation has various specific advantages:

o Everything is the same. Compared to IPv4, there is no more
   calculating leaf subnet sizes, no more juggling between subnets,
   and fewer consequent errors. Network design is therefore simpler

and much more straightforward. This is of importance for all
types of networks - enterprise, campus, small office, or home
networks - and for all types of operator, from professional to
consumer.

o Adding a subnet is easy - just take another /64 from the pool. No
   estimates, calculations, consideration or judgement is needed.

o Router configurations are homogeneous and easier to understand.

o Documentation is easier to write and easier to read; training is
   easier.

The remainder of this document describes arguments that have been
made against the current fixed IID length and analyses the effects of
a possible change. However, the consensus of the IETF is that the
benefits of keeping the length fixed at 64 bits, and the practical
difficulties of changing it, outweigh the arguments for change.

3. Arguments for shorter identifier lengths

In this section we describe arguments for scenarios where shorter
IIDs, implying prefixes longer than /64, have been used or proposed.

3.1. Insufficient address space delegated

A site may not be delegated a sufficiently generous prefix from which
to allocate a /64 prefix to all of its internal subnets. In this
case the site may either determine that it does not have enough
address space to number all its network elements and thus, at the
very best, be only partially operational, or it may choose to use
internal prefixes longer than /64 to allow multiple subnets and the
hosts within them to be configured with addresses.

In this case, the site might choose, for example, to use a /80 per
subnet, in combination with hosts using either manually configured
addressing or DHCPv6 [RFC3315].

Scenarios that have been suggested where an insufficient prefix might
be delegated include home or small office networks, vehicles,
building services and transportation services (road signs, etc.). It
should be noted that the homenet architecture text
[I-D.ietf-homenet-arch] states that a CPE should consider the lack of
sufficient address space to be an error condition, rather than using
prefixes longer than /64 internally.

Another scenario occasionally suggested is one where the Internet
address registries actually begin to run out of IPv6 prefix space,

such that operators can no longer assign reasonable prefixes to users
in accordance with [RFC6177]. It is sometimes suggested that
assigning a prefix such as /48 or /56 to every user site (including
the smallest) as recommended by [RFC6177] is wasteful. In fact, the
currently released unicast address space, 2000::/3, contains 35
trillion /48 prefixes ((2**45 = 35,184,372,088,832), of which only a
small fraction have been allocated. Allowing for a conservative
estimate of allocation efficiency, i.e., an HD-ratio of 0.94
[RFC4692], approximately 5 trillion /48 prefixes can be allocated.
Even with a relaxed HD-ratio of 0.89, approximately one trillion /48
prefixes can be allocated. Furthermore, with only 2000::/3 currently
committed for unicast addressing, we still have approximately 85% of
the address space in reserve. Thus there is no objective risk of
prefix depletion by assigning /48 or /56 prefixes even to the
smallest sites.

3.2. Hierarchical addressing

Some operators have argued that more prefix bits are needed to allow
an aggregated hierarchical addressing scheme within a campus or
corporate network. However, if a campus or enterprise gets a /48
prefix (or shorter), then that already provides 16 bits for
hierarchical allocation. In any case, flat IGP routing is widely and
successfully used within rather large networks, with hundreds of
routers and thousands of end systems. Therefore there is no
objective need for additional prefix bits to support hierarchy and
aggregation within enterprises.

3.3. Audit requirement

Some network operators wish to know and audit which nodes are active
on a network, especially those that are allowed to communicate off
link or off site. They may also wish to limit the total number of
active addresses and sessions that can be sourced from a particular
host, LAN or site, in order to prevent potential resource depletion
attacks or other problems spreading beyond a certain scope of
control. It has been argued that this type of control would be
easier if only long network prefixes with relatively small numbers of
possible hosts per network were used, reducing the discovery problem.
However, such sites most typically operate using DHCPv6, which means
that all legitimate hosts are automatically known to the DHCPv6
servers, which is sufficient for audit purposes. Such hosts could,
if desired, be limited to a small range of IID values without
changing the /64 subnet length. Any hosts inadvertently obtaining
addresses via SLAAC can be audited through Neighbor Discovery logs.

3.4. Concerns over ND cache exhaustion

   A site may be concerned that it is open to neighbour discovery (ND)
   cache exhaustion attacks [RFC3756], whereby an attacker sends a large
   number of messages in rapid succession to a series of (most likely
   inactive) host addresses within a specific subnet. Such an attack
   attempts to fill a router's ND cache with ND requests pending
   completion, in so doing denying correct operation to active devices
   on the network.

   One potential way to mitigate this attack would be to consider using
   a /120 prefix, thus limiting the number of addresses in the subnet to
   be similar to an IPv4 /24 prefix, which should not cause any concerns
   for ND cache exhaustion. Note that the prefix does need to be quite
   long for this scenario to be valid. The number of theoretically
   possible ND cache slots on the segment needs to be of the same order
   of magnitude as the actual number of hosts. Thus small increases
   from the /64 prefix length do not have a noticeable impact: even 2^32
   potential entries, a factor of two billion decrease compared to 2^64,
   is still more than enough to exhaust the memory on current routers.
   Given that most link layer mappings cause SLAAC to assume a 64 bit
   network boundary, in such an approach hosts would likely need to use
   DHCPv6, or be manually configured with addresses.

   It should be noted that several other mitigations of the ND cache
   attack are described in [RFC6583], and that limiting the size of the
   cache and the number of incomplete entries allowed would also defeat
   the attack. For the specific case of a point-to-point link between
   routers, this attack is indeed mitigated by a /127 prefix [RFC6164].

4. Effects of varying the interface identifier length

   This section of the document analyses the impact and effects of
   varying the length of an IPv6 unicast IID by reducing it to less than
   64 bits.

4.1. Interaction with IPv6 specifications

   The precise 64-bit length of the Interface ID is widely mentioned in
   numerous RFCs describing various aspects of IPv6. It is not
   straightforward to distinguish cases where this has normative impact
   or affects interoperability. This section aims to identify
   specifications that contain an explicit reference to the 64-bit
   length. Regardless of implementation issues, the RFCs themselves
   would all need to be updated if the 64-bit rule was changed, even if
   the updates were small, which would involve considerable time and
   effort.

First and foremost, the RFCs describing the architectural aspects of
IPv6 addressing explicitly state, refer and repeat this apparently
immutable value: Addressing Architecture [RFC4291], IPv6 Address
Assignment to End Sites [RFC6177], Reserved Interface Identifiers
[RFC5453], ILNP Node Identifiers [RFC6741]. Customer Edge routers
impose /64 for their interfaces [RFC7084]. The IPv6 Subnet Model
[RFC5942] points out that the assumption of a /64 prefix length is a
potential implementation error.

Numerous IPv6-over-foo documents make mandatory statements with
respect to the 64-bit length of the Interface ID to be used during
the Stateless Autoconfiguration. These documents include [RFC2464]
(Ethernet), [RFC2467] (FDDI), [RFC2470] (Token Ring), [RFC2492]
(ATM), [RFC2497] (ARCnet), [RFC2590] (Frame Relay), [RFC3146] (IEEE
1394), [RFC4338] (Fibre Channel), [RFC4944] (IEEE 802.15.4),
[RFC5072] (PPP), [RFC5121] [RFC5692] (IEEE 802.16), [RFC2529]
(6over4), [RFC5214] (ISATAP), [I-D.templin-aerolink] (AERO),
[I-D.ietf-6lowpan-btle], [I-D.ietf-6man-6lobac],
[I-D.brandt-6man-lowpanz].

To a lesser extent, the address configuration RFCs themselves may in
some ways assume the 64-bit length of an Interface ID (e.g, RFC 4862
for the link-local addresses, DHCPv6 for the potentially assigned
EUI-64-based IP addresses, Optimistic Duplicate Address Detection
[RFC4429] which computes 64-bit-based collision probabilities).

The MLDv1 [RFC2710] and MLDv2 [RFC3810] protocols mandate that all
queries be sent with a link-local source address, with the exception
of MLD messages sent using the unspecified address when the link-
local address is tentative [RFC3590]. At the time of publication of
RFC 2710, the IPv6 addressing architecture specified link-local
addresses with 64-bit interface identifiers. MLDv2 explicitly
specifies the use of the fe80::/64 link-local prefix, and bases the
querier election algorithm on the link-local subnet prefix of length
/64.

The IPv6 Flow Label Specification [RFC6437] gives an example of a
20-bit hash function generation which relies on splitting an IPv6
address in two equally-sized 64bit-length parts.

The basic transition mechanisms [RFC4213] refer to IIDs of length 64
for link-local addresses, and other transition mechanisms such as
Teredo [RFC4380] assume the use of IIDs of length 64. Similar
assumptions are found in 6to4 [RFC3056] and 6rd [RFC5969].
Translation-based transition mechanisms such as NAT64 and NPTv6 have
some dependency on prefix length, discussed below.

The proposed method [RFC7278] of extending an assigned /64 prefix
from a smartphone's cellular interface to its WiFi link relies on
prefix length, and implicitly on the length of the Interface ID, to
be valued at 64.

The CGA and HBA specifications rely on the 64-bit identifier length
(see below), as do the Privacy extensions [RFC4941] and some examples
in IKEv2bis [RFC5996].

464XLAT [RFC6877] explicitly mentions acquiring /64 prefixes.
However, it also discusses the possibility of using the interface
address on the device as the endpoint for the traffic, thus
potentially removing this dependency.

[RFC2526] reserves a number of subnet anycast addresses by reserving
some anycast IIDs. An anycast IID so reserved cannot be less than 7
bits long. This means that a subnet prefix length longer than /121
is not possible, and a subnet of exactly /121 would be useless since
all its identifiers are reserved. It also means that half of a /120
is reserved for anycast. This could of course be fixed in the way
described for /127 in [RFC6164], i.e., avoiding the use of anycast
within a /120 subnet. Note that support for "on-link anycast" is a
standard IPv6 neighbor discovery capability [RFC4861][RFC7094], and
therefore applications and their developers would expect it to be
available.

The Mobile IP home network models [RFC4887] rely heavily on the /64
subnet length and assume a 64-bit IID.

While preparing this document, it was noted that many other IPv6
specifications refer to mandatory alignment on 64-bit boundaries,
64-bit data structures, 64-bit counters in MIBs, 64-bit sequence
numbers and cookies in security, etc. Finally, the number "64" may
be considered "magic" in some RFCs, e.g., 64k limits in DNS and
Base64 encodings in MIME. None of this has any influence on the
length of the IID, but might confuse a careless reader.

4.2. Possible failure modes

This section discusses several specific aspects of IPv6 where we can
expect operational failures with subnet prefixes other than /64.

o Router implementations: Router implementors might interpret IETF
   standards such as [RFC6164] and [RFC7136] to indicate that
   prefixes between /65 and /126 inclusive for unicast packets on-
   the-wire are invalid, and operational practices that utilize
   prefix lengths in this range may fail on some devices, as
   discussed in Section 4.3.2.

o Multicast: [RFC3306] defines a method for generating IPv6
   multicast group addresses based on unicast prefixes. This method
   assumes a longest prefix of 64 bits. If a longer prefix is used,
   there is no way to generate a specific multicast group address
   using this method. In such cases the administrator would need to
   use an "artificial" prefix from within their allocation (a /64 or
   shorter) from which to generate the group address. This prefix
   would not correspond to a real subnet.

   Similarly [RFC3956], which specifies Embedded-RP, allowing IPv6
   multicast rendezvous point addresses to be embedded in the
   multicast group address, would also fail, as the scheme assumes a
   maximum prefix length of 64 bits.

o CGA: The Cryptographically Generated Address format (CGA,
   [RFC3972]) is heavily based on a /64 interface identifier.
   [RFC3972] has defined a detailed algorithm showing how to generate
   a 64-bit interface identifier from a public key and a 64-bit
   subnet prefix. Changing the /64 boundary would certainly
   invalidate the current CGA definition. However, CGA might benefit
   in a redefined version if more bits are used for interface
   identifier (which means shorter prefix length). For now, 59 bits
   are used for cryptographic purposes. The more bits are available,
   the stronger CGA could be. Conversely, longer prefixes would
   weaken CGA.

o NAT64: Both stateless [RFC6052] NAT64 and stateful NAT64 [RFC6146]
   are flexible for the prefix length. [RFC6052] has defined
   multiple address formats for NAT64. In Section 2 "IPv4-Embedded
   IPv6 Prefix and Format" of [RFC6052], the network-specific prefix
   could be one of /32, /40, /48, /56, /64 and /96. The remaining
   part of the IPv6 address is constructed by a 32-bit IPv4 address,
   a 8-bit u byte and a variable length suffix (there is no u byte
   and suffix in the case of 96-bit Well-Known Prefix). NAT64 is
   therefore OK with a subnet boundary out to /96, but not longer.

o NPTv6: IPv6-to-IPv6 Network Prefix Translation [RFC6296] is also
   bound to /64 boundary. NPTv6 maps a /64 prefix to another /64
   prefix. When the NPTv6 Translator is configured with a /48 or
   shorter prefix, the 64-bit interface identifier is kept unmodified
   during translation. However, the /64 boundary might be changed as
   long as the "inside" and "outside" prefixes have the same length.

o ILNP: Identifier-Locator Network Protocol (ILNP) [RFC6741] is
   designed around the /64 boundary, since it relies on locally
   unique 64-bit node identifiers (in the interface identifier
   field). While a re-design to use longer prefixes is not

inconceivable, this would need major changes to the existing
specification for the IPv6 version of ILNP.

o shim6: The Multihoming Shim Protocol for IPv6 (shim6) [RFC5533] in
   its insecure form treats IPv6 address as opaque 128-bit objects.
   However, to secure the protocol against spoofing, it is essential
   to either use CGAs (see above) or Hash-Based Addresses (HBA)
   [RFC5535]. Like CGAs, HBAs are generated using a procedure that
   assumes a 64-bit identifier. Therefore, in effect, secure shim6
   is affected by the /64 boundary exactly like CGAs.

o Duplicate address risk: If SLAAC was modified to work with shorter
   IIDs, the statistical risk of hosts choosing the same pseudo-
   random identifier [RFC7217] would increase correspondingly. The
   practical impact of this would range from slight to dramatic,
   depending on how much the IID length was reduced. In particular,
   a /120 prefix would imply an 8 bit IID and address collisions
   would be highly probable.

o The link-local prefix: While RFC 4862 is careful not to define any
   specific length of link-local prefix within fe80::/10, the
   addressing architecture [RFC4291] does define the link-local IID
   length to be 64 bits. If different hosts on a link used IIDs of
   different lengths to form a link-local address, there is potential
   for confusion and unpredictable results. Typically today the
   choice of 64 bits for the link-local IID length is hard-coded per
   interface, in accordance with the relevant IPv6-over-foo
   specification, and systems behave as if the link local prefix was
   actually fe80::/64. There might be no way to change this except
   conceivably by manual configuration, which will be impossible if
   the host concerned has no local user interface.

It goes without saying that if prefixes longer than /64 are to be
used, all hosts must be capable of generating IIDs shorter than 64
bits, in order to follow the auto-configuration procedure correctly
[RFC4862].

4.3. Experimental observations

4.3.1. Survey of the processing of Neighbor Discovery options with prefixes other than /64

This section provides a survey of the processing of Neighbor
Discovery options which include prefixes that are different than /64.

The behavior of nodes was assessed with respect to the following
options:

   o PIO-A: Prefix Information Option (PIO) [RFC4861] with the A bit
      set.

   o PIO-L: Prefix Information Option (PIO) [RFC4861] with the L bit
      set.

   o PIO-AL: Prefix Information Option (PIO) [RFC4861] with both the A
      and L bits set.

   o RIO: Route Information Option (RIO) [RFC4191].

   In the tables below, the following notation is used:

   NOT-SUP:
      This option is not supported (i.e., it is ignored no matter the
      prefix length used).

   LOCAL:
      The corresponding prefix is considered "on-link".

   ROUTE
      The corresponding route is added to the IPv6 routing table.

   IGNORE:
      The Option is ignored as an error.

      +--------------------+--------+-------+--------+---------+
      | Operating System | PIO-A | PIO-L | PIO-AL | RIO |
      +--------------------+--------+-------+--------+---------+
      | FreeBSD 9.0 | IGNORE | LOCAL | LOCAL | NOT-SUP |
      +--------------------+--------+-------+--------+---------+
      | Linux 3.0.0-15 | IGNORE | LOCAL | LOCAL | NOT-SUP |
      +--------------------+--------+-------+--------+---------+
      | Linux-current | IGNORE | LOCAL | LOCAL | NOT-SUP |
      +--------------------+--------+-------+--------+---------+
      | NetBSD 5.1 | IGNORE | LOCAL | LOCAL | NOT-SUP |
      +--------------------+--------+-------+--------+---------+
      | OpenBSD-current | IGNORE | LOCAL | LOCAL | NOT-SUP |
      +--------------------+--------+-------+--------+---------+
      | Win XP SP2 | IGNORE | LOCAL | LOCAL | ROUTE |
      +--------------------+--------+-------+--------+---------+
      | Win 7 Home Premium | IGNORE | LOCAL | LOCAL | ROUTE |
      +--------------------+--------+-------+--------+---------+

      Table 1: Processing of ND options with prefixes longer than /64

```
   +--------------------+--------+-------+--------+---------+
   |  Operating System  | PIO-A  | PIO-L | PIO-AL |   RIO   |
   +--------------------+--------+-------+--------+---------+
   |    FreeBSD 9.0      | IGNORE | LOCAL | LOCAL  | NOT-SUP |
   +--------------------+--------+-------+--------+---------+
   |   Linux 3.0.0-15   | IGNORE | LOCAL | LOCAL  | NOT-SUP |
   +--------------------+--------+-------+--------+---------+
   |   Linux-current    | IGNORE | LOCAL | LOCAL  | NOT-SUP |
   +--------------------+--------+-------+--------+---------+
   |     NetBSD 5.1      | IGNORE | LOCAL | LOCAL  | NOT-SUP |
   +--------------------+--------+-------+--------+---------+
   |  OpenBSD-current   | IGNORE | LOCAL | LOCAL  | NOT-SUP |
   +--------------------+--------+-------+--------+---------+
   |    Win XP SP2       | IGNORE | LOCAL | LOCAL  |  ROUTE  |
   +--------------------+--------+-------+--------+---------+
   | Win 7 Home Premium | IGNORE | LOCAL | LOCAL  |  ROUTE  |
   +--------------------+--------+-------+--------+---------+
```

   Table 2: Processing of ND options with prefixes shorter than /64

   The results obtained can be summarized as follows:

   o   the "A" bit in the Prefix Information Options is honored only if
       the prefix length is 64. At least for the case where the IID
       length is defined to be 64 bits in the corresponding link-type-
       specific document, which is the case for all currently published
       such documents, this is consistent with [RFC4862], which defines
       the case where the sum of the advertised prefix length and the IID
       length does not equal 128 as an error condition.

   o   the "L" bit in the Prefix Information Options is honored for any
       arbitrary prefix length (whether shorter or longer than /64).

   o   nodes that support the Route Information Option, allow such routes
       to be specified with prefixes of any arbitrary length (whether
       shorter or longer than /64)

4.3.2. Other Observations

   Participants in the V6OPS working group have indicated that some
   forwarding devices have been shown to work correctly with long
   prefixes such as /80 or /96. Indeed, it is to be expected that
   longest prefix match based forwarding will work for any prefix
   length, and no reports of this completely failing have been noted.
   Also, DHCPv6 is in widespread use without any dependency on the /64
   boundary. Reportedly, there are deployments of /120 subnets
   configured using DHCPv6.

There have been definite reports that some routers have a performance
drop-off or even resource exhaustion for prefixes longer than /64,
due to design issues. In particular, some routing chip designs
allocate much less space for longer prefixes than for prefixes up to
/64, for the sake of savings in memory, power and lookup latency.
Some devices need special-case code to handle point-to-point links
according to [RFC6164].

It has been reported that at least one type of switch has a content-
addressable memory limited to 144 bits, which is indeed a typical
value for commodity components [TCAM]. This means that packet
filters or access control lists cannot be defined based on 128-bit
addresses and two 16-bit port numbers; the longest prefix that could
be used in such a filter is a /112.

4.4. Implementation and deployment issues

From an early stage, implementations and deployments of IPv6 assumed
the /64 subnet length, even though routing was based on prefixes of
any length. As shown above, this became anchored in many
specifications (Section 4.1) and in important aspects of
implementations commonly used in local area networks (Section 4.3).
In fact, a programmer might be lulled into assuming a comfortable
rule of thumb that subnet prefixes are always /64 and an IID is
always of length 64. Apart from the limited evidence in
Section 4.3.1, we cannot tell without code inspections or tests
whether existing stacks are able to handle a flexible IID length, or
whether they would require modification to do so. A conforming
implementation of an IPv6-over-foo that specifies a 64 bit IID for
foo links will of course only support 64. But in a well designed
stack, the IP layer itself will treat that 64 as a parameter, so
changing the IID length in the IPv6-over-foo code should be all that
is necessary.

The main practical consequence of the existing specifications is that
deployments in which longer subnet prefixes are used cannot make use
of SLAAC-configured addresses, and require either manually configured
addresses or DHCPv6. To reverse this argument, if it was considered
desirable to allow auto-configured addresses with subnet prefixes
longer than /64, all of the specifications identified above as
depending on /64 would have to be modified, with due regard to
interoperability with unmodified stacks. In fact [RFC7217] allows
for this possibility. Then modified stacks would have to be
developed and deployed. It might be the case that some stacks
contain dependencies on the /64 boundary which are not directly
implied by the specifications, and any such hidden dependencies would
also need to be found and removed.

At least one DHCPv6 client unconditionally installs a /64 prefix as
on-link when it configures an interface with an address, although
some specific operating system vendors seem to change this default
behavior by tweaking a client-side script. This is in clear
violation of the IPv6 subnet model [RFC5942]. The motivation for
this choice is that if there is no router on the link, the hosts
would fail to communicate with each other using the configured
addresses because the "on-link assumption" was removed in [RFC4861].
This is not really about the magic number of 64, but an
implementation may sometimes pick an arbitrary value of prefix length
due to the removal of the on-link assumption, and the value chosen
will most likely be 64.

Typical IP Address Management (IPAM) tools treat /64 as the default
subnet length, but allow users to specify longer subnet prefixes if
desired. Clearly, all IPAM tools and network management systems
would need to be checked in detail.

Finally, IPv6 is already deployed at many sites, with a large number
of staff trained on the basis of the existing standards, supported by
documentation and tools based on those standards. Numerous existing
middlebox devices are also based on those standards. These people,
documents, tools and devices represent a very large investment that
would be seriously impacted by a change in the /64 boundary.

4.5. Privacy issues

The length of the interface identifier has implications for privacy
[I-D.ietf-6man-ipv6-address-generation-privacy]. In any case in
which the value of the identifier is intended to be hard to guess,
whether or not it is cryptographically generated, it is apparent that
more bits are better. For example, if there are only 20 bits to be
guessed, at most just over a million guesses are needed, today well
within the capacity of a low cost attack mechanism. It is hard to
state in general how many bits are enough to protect privacy, since
this depends on the resources available to the attacker, but it seems
clear that a privacy solution needs to resist an attack requiring
billions rather than millions of guesses. Trillions would be better,
suggesting that at least 40 bits should be available. Thus we can
argue that subnet prefixes longer than say /80 might raise privacy
concerns by making the IID guessable.

A prefix long enough to limit the number of addresses comparably to
an IPv4 subnet, such as /120, would create exactly the same situation
for privacy as IPv4 except for the absence of NAT. In particular, a
host would be forced to pick a new IID when roaming to a new network,
to avoid collisions. As mentioned earlier, it is likely that SLAAC
will not be used on such a subnet.

5. Security Considerations

   In addition to the privacy issues mentioned in Section 4.5, and the
   issues mentioned with CGAs and HBAs in Section 4.2, the length of the
   subnet prefix affects the matter of defence against scanning attacks
   [I-D.ietf-opsec-ipv6-host-scanning]. Assuming the attacker has
   discovered or guessed the prefix length, a longer prefix reduces the
   space that the attacker needs to scan, e.g., to only 256 addresses if
   the prefix is /120. On the other hand, if the attacker has not
   discovered the prefix length and assumes it to be /64, routers can
   trivially discard attack packets that do not fall within an actual
   subnet.

   However, assume that an attacker finds one valid address A and
   assumes that it is within a long prefix such as a /120. The attacker
   then starts a scanning attack by scanning "outwards" from A, by
   trying A+1, A-1, A+2, A-2, etc. This attacker will easily find all
   hosts in any subnet with a long prefix, because they will have
   addresses close to A. We therefore conclude that any prefix
   containing densely packed valid addresses is vulnerable to a scanning
   attack, without the attacker needing to guess the prefix length.
   Therefore, to preserve IPv6's advantage over IPv4 in resisting
   scanning attacks, it is important that subnet prefixes are short
   enough to allow sparse allocation of identifiers within each subnet.
   The considerations are similar to those for privacy, and we can again
   argue that prefixes longer than say /80 might significantly increase
   vulnerability. Ironically, this argument is exactly converse to the
   argument for longer prefixes to resist an ND cache attack, as
   described in Section 3.4.

   Denial of service attacks related to Neighbor Discovery are discussed
   in Section 3.4 and in [RFC6583]. One of the mitigations suggested by
   that document is "sizing subnets to reflect the number of addresses
   actually in use", but the fact that this greatly simplifies scanning
   attacks is not noted. For further discussion of scanning attacks,
   see [I-D.ietf-opsec-ipv6-host-scanning].

   Note that, although not known at the time of writing, there might be
   other resource exhaustion attacks available, similar in nature to the
   ND cache attack. We cannot exclude that such attacks might be
   exacerbated by sparsely populated subnets such as a /64. It should
   also be noted that this analysis assumes a conventional deployment
   model with a significant number of end-systems located in a single
   LAN broadcast domain. Other deployment models might lead to
   different conclusions.

6. IANA Considerations

   This document requests no action by IANA.

7. Acknowledgements

   This document was inspired by a vigorous discussion on the V6OPS
   working group mailing list with at least 20 participants. Later,
   valuable comments were received from Ran Atkinson, Fred Baker, Steven
   Blake, Lorenzo Colitti, David Farmer, Bill Fenner, Ray Hunter,
   Paraskevi Iliadou, Jen Linkova, Philip Matthews, Matthew Petach,
   Scott Schmit, Tatuya Jinmei, Fred Templin, Ole Troan, Stig Venaas,
   and numerous other participants in the 6MAN working group. An
   extremely detailed review by Mark Smith was especially helpful.

   This document was produced using the xml2rfc tool [RFC2629].

8. Change log [RFC Editor: Please remove]

   draft-ietf-6man-why64-06: minor IETF Last Call comments, 2014-10-02.

   draft-ietf-6man-why64-05: Area Director review comments, 2014-09-16.

   draft-ietf-6man-why64-04: fixed reference error, 2014-09-10.

   draft-ietf-6man-why64-03: fixed nits, 2014-08-27.

   draft-ietf-6man-why64-02: responded to WGLC reviews and comments,
   2014-08-16.

   draft-ietf-6man-why64-01: language improvements, added TCAM
   reference, 2014-05-07.

   draft-ietf-6man-why64-00: WG adoption, WG comments, including major
   text reorganisation: 3 main sections describe advantages of fixed
   length IID, arguments for shorter lengths, and expected effects of
   varying the length, 2014-04-11.

   draft-carpenter-6man-why64-01: WG comments, added experimental
   results, implementation/deployment text, 2014-02-06.

   draft-carpenter-6man-why64-00: original version, 2014-01-06.

9. References

9.1. Normative References

   [RFC2464] Crawford, M., "Transmission of IPv6 Packets over Ethernet
              Networks", RFC 2464, December 1998.

   [RFC2467] Crawford, M., "Transmission of IPv6 Packets over FDDI
              Networks", RFC 2467, December 1998.

   [RFC2470] Crawford, M., Narten, T., and S. Thomas, "Transmission of
              IPv6 Packets over Token Ring Networks", RFC 2470, December
              1998.

   [RFC2492] Armitage, G., Schulter, P., and M. Jork, "IPv6 over ATM
              Networks", RFC 2492, January 1999.

   [RFC2497] Souvatzis, I., "Transmission of IPv6 Packets over ARCnet
              Networks", RFC 2497, January 1999.

   [RFC2526] Johnson, D. and S. Deering, "Reserved IPv6 Subnet Anycast
              Addresses", RFC 2526, March 1999.

   [RFC2529] Carpenter, B. and C. Jung, "Transmission of IPv6 over IPv4
              Domains without Explicit Tunnels", RFC 2529, March 1999.

   [RFC2590] Conta, A., Malis, A., and M. Mueller, "Transmission of
              IPv6 Packets over Frame Relay Networks Specification", RFC
              2590, May 1999.

   [RFC2710] Deering, S., Fenner, W., and B. Haberman, "Multicast
              Listener Discovery (MLD) for IPv6", RFC 2710, October
              1999.

   [RFC3056] Carpenter, B. and K. Moore, "Connection of IPv6 Domains
              via IPv4 Clouds", RFC 3056, February 2001.

   [RFC3146] Fujisawa, K. and A. Onoe, "Transmission of IPv6 Packets
              over IEEE 1394 Networks", RFC 3146, October 2001.

   [RFC3306] Haberman, B. and D. Thaler, "Unicast-Prefix-based IPv6
              Multicast Addresses", RFC 3306, August 2002.

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

   [RFC3590] Haberman, B., "Source Address Selection for the Multicast
              Listener Discovery (MLD) Protocol", RFC 3590, September
              2003.

   [RFC3810] Vida, R. and L. Costa, "Multicast Listener Discovery
              Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

   [RFC3956] Savola, P. and B. Haberman, "Embedding the Rendezvous
              Point (RP) Address in an IPv6 Multicast Address", RFC
              3956, November 2004.

   [RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)",
              RFC 3972, March 2005.

   [RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and
              More-Specific Routes", RFC 4191, November 2005.

   [RFC4213] Nordmark, E. and R. Gilligan, "Basic Transition Mechanisms
              for IPv6 Hosts and Routers", RFC 4213, October 2005.

   [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing
              Architecture", RFC 4291, February 2006.

   [RFC4338] DeSanti, C., Carlson, C., and R. Nixon, "Transmission of
              IPv6, IPv4, and Address Resolution Protocol (ARP) Packets
              over Fibre Channel", RFC 4338, January 2006.

   [RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through
              Network Address Translations (NATs)", RFC 4380, February
              2006.

   [RFC4429] Moore, N., "Optimistic Duplicate Address Detection (DAD)
              for IPv6", RFC 4429, April 2006.

   [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman,
              "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861,
              September 2007.

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

   [RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy
              Extensions for Stateless Address Autoconfiguration in
              IPv6", RFC 4941, September 2007.

   [RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler,
              "Transmission of IPv6 Packets over IEEE 802.15.4
              Networks", RFC 4944, September 2007.

   [RFC5072] Varada, S., Haskins, D., and E. Allen, "IP Version 6 over
              PPP", RFC 5072, September 2007.

   [RFC5121] Patil, B., Xia, F., Sarikaya, B., Choi, JH., and S.
              Madanapalli, "Transmission of IPv6 via the IPv6
              Convergence Sublayer over IEEE 802.16 Networks", RFC 5121,
              February 2008.

   [RFC5214] Templin, F., Gleeson, T., and D. Thaler, "Intra-Site
              Automatic Tunnel Addressing Protocol (ISATAP)", RFC 5214,
              March 2008.

   [RFC5453] Krishnan, S., "Reserved IPv6 Interface Identifiers", RFC
              5453, February 2009.

   [RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming
              Shim Protocol for IPv6", RFC 5533, June 2009.

   [RFC5535] Bagnulo, M., "Hash-Based Addresses (HBA)", RFC 5535, June
              2009.

   [RFC5692] Jeon, H., Jeong, S., and M. Riegel, "Transmission of IP
              over Ethernet over IEEE 802.16 Networks", RFC 5692,
              October 2009.

   [RFC5942] Singh, H., Beebee, W., and E. Nordmark, "IPv6 Subnet
              Model: The Relationship between Links and Subnet
              Prefixes", RFC 5942, July 2010.

   [RFC5969] Townsley, W. and O. Troan, "IPv6 Rapid Deployment on IPv4
              Infrastructures (6rd) -- Protocol Specification", RFC
              5969, August 2010.

   [RFC5996] Kaufman, C., Hoffman, P., Nir, Y., and P. Eronen,
              "Internet Key Exchange Protocol Version 2 (IKEv2)", RFC
              5996, September 2010.

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

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

   [RFC6164] Kohno, M., Nitzan, B., Bush, R., Matsuzaki, Y., Colitti,
              L., and T. Narten, "Using 127-Bit IPv6 Prefixes on Inter-
              Router Links", RFC 6164, April 2011.

   [RFC6177] Narten, T., Huston, G., and L. Roberts, "IPv6 Address
              Assignment to End Sites", BCP 157, RFC 6177, March 2011.

   [RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix
              Translation", RFC 6296, June 2011.

   [RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme,
              "IPv6 Flow Label Specification", RFC 6437, November 2011.

   [RFC7084] Singh, H., Beebee, W., Donley, C., and B. Stark, "Basic
              Requirements for IPv6 Customer Edge Routers", RFC 7084,
              November 2013.

   [RFC7136] Carpenter, B. and S. Jiang, "Significance of IPv6
              Interface Identifiers", RFC 7136, February 2014.

9.2. Informative References

   [DRAFT-odell]
              O'Dell, M., "8+8 - An Alternate Addressing Architecture
              for IPv6", draft-odell-8+8.00 (work in progress), October
              1996.

   [I-D.brandt-6man-lowpanz]
              Brandt, A. and J. Buron, "Transmission of IPv6 packets
              over ITU-T G.9959 Networks", draft-brandt-6man-lowpanz-02
              (work in progress), June 2013.

   [I-D.ietf-6lowpan-btle]
              Nieminen, J., Savolainen, T., Isomaki, M., Patil, B.,
              Shelby, Z., and C. Gomez, "Transmission of IPv6 Packets
              over BLUETOOTH Low Energy", draft-ietf-6lowpan-btle-12
              (work in progress), February 2013.

   [I-D.ietf-6man-6lobac]
              Lynn, K., Martocci, J., Neilson, C., and S. Donaldson,
              "Transmission of IPv6 over MS/TP Networks", draft-ietf-
              6man-6lobac-01 (work in progress), March 2012.

   [I-D.ietf-6man-ipv6-address-generation-privacy]
              Cooper, A., Gont, F., and D. Thaler, "Privacy
              Considerations for IPv6 Address Generation Mechanisms",
              draft-ietf-6man-ipv6-address-generation-privacy-01 (work
              in progress), February 2014.

   [I-D.ietf-homenet-arch]
              Chown, T., Arkko, J., Brandt, A., Troan, O., and J. Weil,
              "IPv6 Home Networking Architecture Principles", draft-
              ietf-homenet-arch-17 (work in progress), July 2014.

   [I-D.ietf-opsec-ipv6-host-scanning]
             Gont, F. and T. Chown, "Network Reconnaissance in IPv6
             Networks", draft-ietf-opsec-ipv6-host-scanning-04 (work in
             progress), June 2014.

   [I-D.templin-aerolink]
             Templin, F., "Transmission of IP Packets over AERO Links",
             draft-templin-aerolink-40 (work in progress), September
             2014.

   [IEEE802] IEEE, "IEEE Standard for Local and Metropolitan Area
             Networks: Overview and Architecture", IEEE Std 802-2001
             (R2007), 2007.

   [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629,
             June 1999.

   [RFC3756] Nikander, P., Kempf, J., and E. Nordmark, "IPv6 Neighbor
             Discovery (ND) Trust Models and Threats", RFC 3756, May
             2004.

   [RFC4692] Huston, G., "Considerations on the IPv6 Host Density
             Metric", RFC 4692, October 2006.

   [RFC4887] Thubert, P., Wakikawa, R., and V. Devarapalli, "Network
             Mobility Home Network Models", RFC 4887, July 2007.

   [RFC5505] Aboba, B., Thaler, D., Andersson, L., and S. Cheshire,
             "Principles of Internet Host Configuration", RFC 5505, May
             2009.

   [RFC6583] Gashinsky, I., Jaeggli, J., and W. Kumari, "Operational
             Neighbor Discovery Problems", RFC 6583, March 2012.

   [RFC6741] Atkinson,, RJ., "Identifier-Locator Network Protocol
             (ILNP) Engineering Considerations", RFC 6741, November
             2012.

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

   [RFC7094] McPherson, D., Oran, D., Thaler, D., and E. Osterweil,
             "Architectural Considerations of IP Anycast", RFC 7094,
             January 2014.

   [RFC7217]   Gont, F., "A Method for Generating Semantically Opaque
               Interface Identifiers with IPv6 Stateless Address
               Autoconfiguration (SLAAC)", RFC 7217, April 2014.

   [RFC7278]   Byrne, C., Drown, D., and A. Vizdal, "Extending an IPv6
               /64 Prefix from a Third Generation Partnership Project
               (3GPP) Mobile Interface to a LAN Link", RFC 7278, June
               2014.

   [TCAM]   Meiners, C., Liu, A., and E. Torng, "Algorithmic
               Approaches to Redesigning TCAM-Based Systems", ACM
               SIGMETRICS'08 467-468, 2008.

Authors' Addresses

   Brian Carpenter (editor)
   Department of Computer Science
   University of Auckland
   PB 92019
   Auckland 1142
   New Zealand

   Email: brian.e.carpenter@gmail.com

   Tim Chown
   University of Southampton
   Southampton, Hampshire SO17 1BJ
   United Kingdom

   Email: tjc@ecs.soton.ac.uk

   Fernando Gont
   SI6 Networks / UTN-FRH
   Evaristo Carriego 2644
   Haedo, Provincia de Buenos Aires 1706
   Argentina

   Email: fgont@si6networks.com

   Sheng Jiang
   Huawei Technologies Co., Ltd
   Q14, Huawei Campus
   No.156 Beiqing Road
   Hai-Dian District, Beijing 100095
   P.R. China

   Email: jiangsheng@huawei.com

   Alexandru Petrescu
   CEA, LIST
   CEA Saclay
   Gif-sur-Yvette, Ile-de-France 91190
   France

   Email: Alexandru.Petrescu@cea.fr

   Andrew Yourtchenko
   cisco
   7a de Kleetlaan
   Diegem 1830
   Belgium

   Email: ayourtch@cisco.com