

Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 3, 2012

J. Arkko
A. Lindem
Ericsson
October 31, 2011

Prefix Assignment in a Home Network
draft-arkko-homenet-prefix-assignment-01

Abstract

This memo describes a prefix assignment mechanism for home networks. It is expected that home gateway routers are assigned an IPv6 prefix through DHCPv6 Prefix Delegation (PD). This prefix needs to be divided among the multiple subnets in a home network. This memo describes a mechanism for such division via OSPFv3. This is an alternative design to using DHCPv6 PD also for the prefix assignment. The memo is input to the working group so that it can make a decision on which type of design to pursue. It is expected that a routing-protocol based assignment uses a minimal amount of prefixes.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Requirements language	3
3. Role of Prefix Assignment	3
4. Router Behavior	5
5. Prefix Assignment in OSPFv3	7
5.1. Usable Prefix TLV	7
5.2. Assigned Prefix TLV	8
5.3. OSPFv3 Prefix Assignment	9
6. Manageability Considerations	11
7. Security Considerations	11
8. IANA Considerations	12
9. Analysis	12
10. References	12
10.1. Normative References	12
10.2. Informative References	13
Appendix A. Acknowledgments	13
Authors' Addresses	13

1. Introduction

This memo describes a prefix assignment mechanism for home networks. It is expected that home gateway routers are assigned an IPv6 prefix through DHCPv6 Prefix Delegation (PD) [RFC3633], or in some cases manually configured. This prefix needs to be divided among the multiple subnets in a home network. This memo describes a mechanism for such division via OSPFv3 [RFC5340].

The OSPFv3-based mechanism is an alternative design to using DHCPv6 PD also for the prefix assignment in the internal network. This memo has been written so that the working group can make a decision on which type of design to pursue. The main benefit of using a routing protocol to handle the prefix assignment is that it can provide a more efficient allocation mechanism than hierarchical assignment through DHCPv6 PD. This may be important for home networks that get only a /60 allocation from their ISPs.

The rest of this memo is organized as follows. Section 2 defines the usual keywords, Section 3 explains the main requirements for prefix assignments, Section 4 describes how a home gateway router makes assignments when it itself has multiple subnets, and Section 5 describes how the assignment can be performed in a distributed manner via OSPFv3 in the entire home network. Finally, Section 6 explains what administrative interfaces are useful for advanced users that wish to manually interact with the mechanisms, Section 7 discusses the security aspects of the design, and Section 8 explains the necessary IANA actions.

2. Requirements language

In this document, the key words "MAY", "MUST", "MUST NOT", "OPTIONAL", "RECOMMENDED", "SHOULD", and "SHOULD NOT", are to be interpreted as described in [RFC2119].

3. Role of Prefix Assignment

Given a prefix shorter than /64 for the entire home network, this prefix needs to be subdivided so that every subnet is given its own /64 prefix. In many cases there will be just one subnet, the internal network interface attached to the router. But it is also common to have two or more internal network interfaces with intentionally separate networks. For instance, "private" and "guest" SSIDs are automatically configured in many current home network routers. When all the network interfaces that require a prefix are attached to the same router, the prefix assignment problem is simple,

and procedures outlined in Section 4 can be employed.

In a more complex setting there are multiple routers in the internal network. There are various possible reasons why this might be necessary [I-D.chown-homenet-arch]. For instance, networks that cannot be bridged together should be routed, speed differences between wired and wireless interfaces make the use of the same broadcast domain undesirable, or simply that router devices keep being added. In any case, it then becomes necessary to assign prefixes across the entire network, and this assignment can no longer be done on a local basis as proposed in Section 4. A distributed mechanism and a protocol is required.

The key requirements for this distributed mechanism are as follows.

- o The short prefix assigned to the home gateway router must be used to assign /64 prefixes on each subnet that requires one.
- o The assignment mechanism should provide reasonable efficiency. As a practical benchmark, some ISPs may employ /60 assignments to individual subscribers. As a result, the assignment mechanism should avoid wasting too many prefixes so that this set of 16 /64 prefixes does not run out in the foreseeable future for commonly occurring network configurations.
- o In particular, the assignment of multiple prefixes to the same network from the same top-level prefix must be avoided.

Example: When a home network consists of a home gateway router connected to another router which in turn is connected to hosts, a minimum of two /64 prefixes are required in the internal network: one between the two routers, and another one for the host-side interface on the second router. But an ineffective assignment mechanism in the two routers might have both of them asking for an assignment for this shared interface.

- o The assignments must be stable across reboots, power cycling, router software updates, and preferably, should be stable across simple network changes. Simple network changes are in this case defined as those that could be resolved through either deletion or addition of a prefix assignment. For instance, the addition of a new router without changing connections between existing routers requires just the assignment of new prefixes for the new networks that the router introduces. There are no stability requirements across more complex types of network reconfiguration events. For instance, if a network is separated into two networks connected by a newly inserted router, this may lead into renumbering all

networks within the home.

In an even more complex setting there may be multiple home gateway routers and multiple connections to ISP(s). These cases are analogous to the case of a single gateway router. Each gateway will simply distribute the prefix it has, and participating routers throughout the network may assign themselves prefixes from several gateways.

Similarly, it is also possible that it is necessary to assign both a global prefix delegated from the ISP and a local, Unique Local Address (ULA) prefix [RFC4193]. The mechanisms in this memo are applicable to both types of prefixes. For ULA-based prefixes, it is necessary to elect one or more router as the generator of such prefixes, and have it perform the generation and employ the prefixes for local interfaces and the entire router network. The generation of ULAs in this manner -- and indeed even the question of whether ULAs are needed -- is outside the scope of this memo, however. We only note that if ULA prefixes are generated, then the mechanisms in this memo can be used to subdivide that prefix for the rest of the network.

Finally, the mechanisms in this memory can also be used in standalone or ad hoc networks where no global prefixes or Internet connectivity are available, by distributing ULA prefixes within the network.

4. Router Behavior

This section describes how a router assigns prefixes to its directly connected interfaces. We assume that the router has prefix(es) that it can use for this allocation. These prefix(es) can be manually configured, acquired through DHCPv6 PD from the ISP, or learned through the distributed prefix assignment protocols described in Section 5. Each such prefix is called a usable prefix. Parts of the usable prefix may already be assigned for some purpose; a coordinated allocation from the prefix is necessary before it can actually be assigned to an interface.

Even if the assignment process is local, it still needs to follow the requirements from Section 3. This is achieved through the following rules:

- o The router MUST maintain a list of assigned prefixes on a per-interface basis. The contents of this list consists of prefixes that the router itself has assigned to the interface, as well as prefixes assigned to the interface by other routers. The latter are learned through the mechanisms described in Section 5, when

they are used.

- o Whenever the router finds a combination of an interface and usable prefix that is not used on the interface, it SHOULD make a new assignment. That is, the router checks to see if there exists an interface and usable prefix such that there are no assigned prefixes within that interface that are more specific than the usable prefix. In this situation the router makes an allocation from the usable prefix (if possible) and adds the allocation to the list of assigned prefixes on that interface.
- o An allocation from a usable prefix MUST check for other allocations from the same usable prefix. Allocations are made for individual /64 prefixes. The choice of a /64 among multiple free ones MUST be made randomly or based on an algorithm that takes unique hardware characteristics of the router and the interface into account. This helps avoid collisions when simultaneous allocations are made within a network.
- o In order to provide a stable assignment, the router MUST store assignments affecting directly connected interfaces in non-volatile memory and attempt to re-use them in the future when possible. At least the 5 most recent assignments SHOULD be stored. Note that this applies to both its own assignments as well as assignments made by others. This ensures that the same prefix assignments are made regardless of the order that different devices are brought up. To avoid attacks on flash memory write cycles, assignments made by others SHOULD be recorded only after 10 minutes have passed and the assignment is still valid.
- o Re-using a memorized assignment is possible when there exists a usable prefix that is less specific than the prefix in the assignment (or it is the prefix itself in the assignment), and the prefix in the assignment can be allocated for that purpose.

Once the router has assigned a prefix to an interface, it MUST act as a router as defined in [RFC4861] and advertise the prefix in Router Advertisements. The lifetime of the prefix SHOULD be advertised as a reasonably long period, at least 48 hours or the lifetime of the assigned prefixes, whichever is smaller. To support a variety of IPv6-only hosts in these networks, the router needs to ensure that sufficient DNS discovery mechanisms are enabled. It is RECOMMENDED that both stateless DHCPv6 [RFC3736] and Router Advertisement options [RFC6106] are supported and turned on by default. This requires, however, that a working DNS server is known and addressable via IPv6. The mechanism in [RFC3736] and [RFC3646] can be used for this.

5. Prefix Assignment in OSPFv3

This section describes how prefix assignment in a home network can be performed in a distributed manner via OSPFv3. It is expected that the router already support the auto-configuration extensions defined in [I-D.acee-ospf-ospfv3-autoconfig].

An overview to OSPFv3-based prefix assignment is as follows. OSPFv3 routers that are capable of auto-configuration advertise OSPFv3 Auto-Configuration (AC) LSA [I-D.acee-ospf-ospfv3-autoconfig] with suitable TLVs. For prefix assignment, two TLVs are used. The Usable Prefix TLV (Section 5.1) advertises a usable prefix, usually the prefix that has been delegated to the home gateway router from the ISP through DHCPv6 PD. These usable prefixes are necessary for running the algorithm in Section 4 for determining whether prefix assignments can and should be made.

The Assigned Prefix TLV (Section 5.2) is used to communicate assignments that routers make out of the usable prefixes.

An assignment can be made when the algorithm in Section 4 indicates that it can be made and no other router has claimed the same assignment. The router emits an OSPFv3 advertisement with Assigned Prefix TLV included to let other devices know that the prefix is now in use. Unfortunately, collisions are still possible, when the algorithms on different routers happen to choose the same free /64 prefix or when more /64 prefixes are needed than there are available. This situation is detected through an advertisement where a different router claims the allocation of the same prefix. In this situation the router with numerically lower OSPFv3 Router ID has to select another prefix. See also [I-D.acee-ospf-ospfv3-autoconfig] Section 5.2.

5.1. Usable Prefix TLV

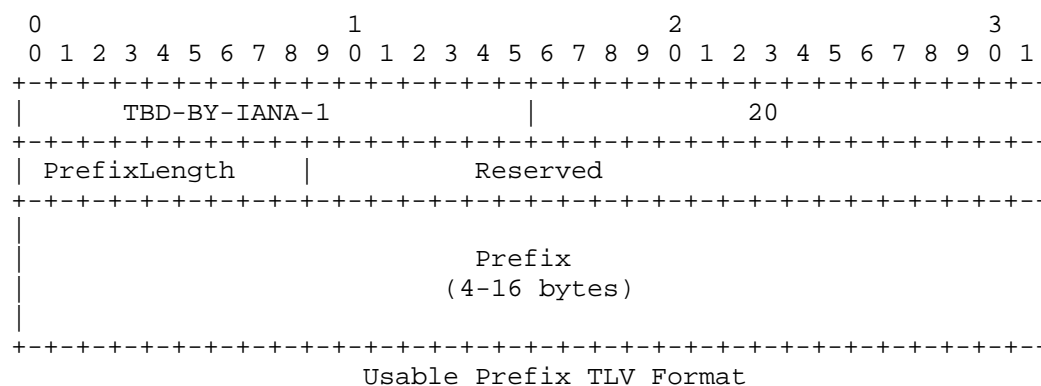
The Usable Prefix TLV is defined for the OSPFv3 Auto-Configuration (AC) LSA [I-D.acee-ospf-ospfv3-autoconfig]. It will have type TBD-BY-IANA-1 and MUST be advertised in the LSID OSPFv3 AC LSA with an LSID of 0. It MAY occur once or multiple times and the information from all TLV instances is retained. The length of the TLV is variable.

The contents of the TLV include a usable prefix (Prefix) and prefix length (PrefixLength). PrefixLength is the length in bits of the prefix and is an 8-bit field. The PrefixLength MUST be greater than or equal to 8 and less than or equal to 64. The prefix describes an allocation of a global or ULA prefix for the entire auto-configured home network. The Prefix is an encoding of the prefix itself as an

even multiple of 32-bit words, padding with zero bits as necessary. This encoding consumes $(\text{PrefixLength} + 31) / 32$ 32-bit words and is consistent with [RFC5340]. It MUST NOT be directly assigned to any interface before following through the procedures defined above.

This TLV SHOULD be emitted by every home gateway router that has either a manual or DHCPv6 PD based prefix that is shorter than /64.

This TLV MUST appear inside an OSPFv3 Router Auto-Configuration LSA, and only in combination with the Router-Hardware-Fingerprint TLV [I-D.acee-ospf-ospfv3-autoconfig] Section 5.2.2 in the same LSA.



5.2. Assigned Prefix TLV

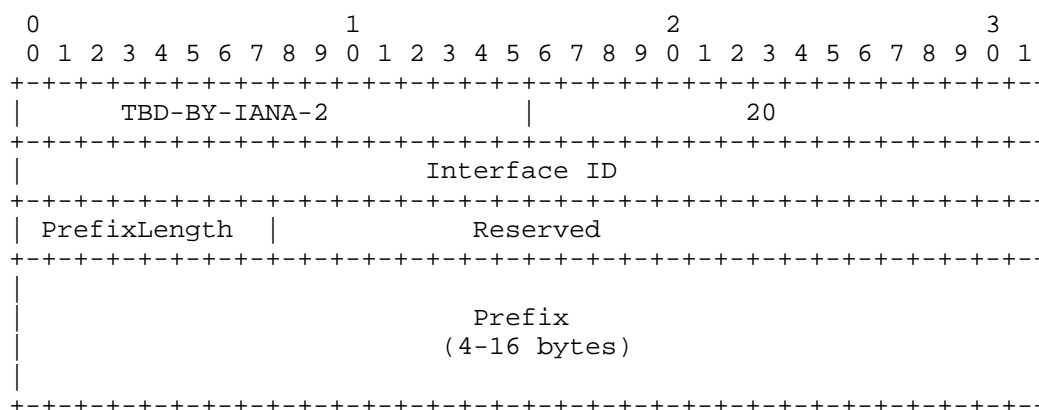
The Assigned Prefix TLV is defined for the OSPFv3 Auto-Configuration (AC) LSA [I-D.acee-ospf-ospfv3-autoconfig]. It will have type TBD-BY-IANA-2 and MUST be advertised in the LSID OSPFv3 AC LSA with an LSID of 0. It MAY occur once or multiple times and the information from all TLV instances is retained. The length of the TLV is variable.

The contents of the TLV include an Interface ID, assigned prefix (Prefix), and prefix length (PrefixLength). The Interface ID is the same OSPFv3 Interface ID that is described in section 4.2.1 or [RFC5340]. PrefixLength is the length in bits of the prefix and is an 8-bit field. The PrefixLength value MUST be 64 in this version of the specification. The prefix describes an assignment of a global or ULA prefix for a directly connected interface in the advertising router. The Prefix is an encoding of the prefix itself as an even multiple of 32-bit words, padding with zero bits as necessary. This encoding consumes $(\text{PrefixLength} + 31) / 32$ 32-bit words and is consistent with xref target="RFC5340"/>.

This TLV MUST be emitted by every home router that has made

assignment from a usable prefix per Section 4.

This TLV MUST appear inside an OSPFv3 Router Auto-Configuration LSA, and only in combination with the Router-Hardware-Fingerprint TLV [I-D.acee-ospf-ospfv3-autoconfig] Section 5.2.2 in the same LSA.



Assigned Prefix TLV Format

5.3. OSPFv3 Prefix Assignment

OSPFv3 Routers supporting the mechanisms in the memo will learn or assign a global /64 IPv6 prefix for each IPv6 interface. Since the mechanisms described herein are based on OSPFv3, Router ID assignment as described in [I-D.acee-ospf-ospfv3-autoconfig] MUST have completed successfully.

When an OSPFv3 Router receives a global prefix through DHCPv6 prefix delegation, manual configuration, or other means, it will advertise this prefix by including the Usable Prefix TLV in its OSPFv3 AC LSA. This will trigger prefix assignment for auto-configured OSPFv3 routers within the routing domain including the originating OSPFv3 router.

When an OSPFv3 Router receives an AC LSA containing a Usable Prefix TLV, it will determine whether or not a new prefix needs to be assigned for each of its attached IPv6 interfaces. For the purposes of this discussion, the received prefix will be referred to as the Current Usable Prefix. The following steps will be performed for each IPv6 interface:

1. The OSPFv3 Router will determine whether there are any other OSPFv3 Routers connected to the same link by examining its list

of neighbors.

2. If no OSPFv3 neighbors have been discovered, the router will wait TBD seconds before allocating a unique /64 IPv6 prefix for the link as described in step 5.
3. If OSPFv3 neighbors are present on the link, the router needs to determine whether any of them have already assigned an IPv6 prefix. This is done by examining the AC LSAs for neighbors on the link and looking for any that include an Assigned Prefix TLV with the same OSPFv3 Interface ID as the neighbor. If one is found and is it a subnet of the IPv6 prefix advertised in the Usable Prefix TLV, this global IPv6 prefix has been already been assigned to the link. If more than one neighbor's Assigned Prefix TLV is found with an IPv6 prefix matching the criteria above, the Assigned Prefix advertised by the OSPFv3 router with the numerically highest OSPFv3 Router ID takes precedence.
4. If there are OSPFv3 neighbors on the link but no IPv6 Prefix is found, the task of prefix allocation is delegated to the OSPFv3 Router with the numerically highest OSPFv3 Router ID. Note that this is different from OSPFv3 Designated Router (DR) election, as described in [RFC5340], in that the router priority is not taken into consideration and that the election will work for networks types where no DR is elected, e.g., point-to-point links.
5. If it is determined that the OSPFv3 Router is responsible for prefix assignment on the link, it will:
 - * Examine all the AC LSA including Assigned Prefix TLVs that are subnets of the Current Usable Prefix to determine which /64s prefixes are already assigned.
 - * Examine former prefix assignments stored in non-volatile storage for interface. Starting with the most recent assignment, if the prefix is both a subnet of the Current Usable Prefix and is currently unassigned, reuse the assignment for the interface.
 - * If no unused former prefix allocation is found, allocate a new one from the subnets of the Current Usable Prefix which are unallocated.
 - * Once a global IPv6 prefix is assigned, a new instance of the AC LSA will be re-originated including the Assigned Prefix TLV.

- * In the rare event that no global /64 IPv6 prefixes are available within the Current Usable Prefix, no IPv6 prefix is assigned and an error condition must be raised.

There are two types of conflicts that may be detected:

1. Two or more OSPFv3 routers have assigned the same IPv6 prefix for different networks.
2. Two of more OSPFv3 routers have assigned different IPv6 prefixes for the same network.

In the case of the former, the OSPFv3 Router with the numerically lower OSPFv3 Router ID must select a new prefix and advertise a new instance of its AC LSA with an updated Assign Prefix TLV for the link. In the latter case, the OSPFv3 Router with the numerically lower OSPFv3 Router ID should accept the global IPv6 prefix from the neighbor with the highest OSPFv3 Router ID and originate a new AC LSA excluding the Assigned Prefix TLV for the link.

6. Manageability Considerations

Advanced users may wish to manage their networks without automation, and there may also be situations where manual intervention may be needed. For these purposes there MUST be a configuration mechanism that allows users to turn off the automatic prefix assignment on a given interface. This setting can be a part of disabling the entire routing auto-configuration [I-D.acee-ospf-ospfv3-autoconfig].

In addition, there SHOULD be a configuration mechanism that allows users to specify the prefix that they would like the router to request for a given interface. This can be useful, for instance, when a router is replaced and there is a desire for the new router to be configured to ask for the same prefix as the old one, in order to avoid renumbering other devices on this network.

Finally, there SHOULD be mechanisms to display what prefixes the router has been assigned, and where they came from (manual configuration, DHCPv6 PD, OSPFv3).

7. Security Considerations

Security can be always added later.

8. IANA Considerations

This memo makes two allocations out of the OSPFv3 Auto- Configuration (AC) LSA TLV namespace [I-D.acee-ospf-ospfv3-autoconfig]:

- o The Usable Prefix TLV in Section 5.1 takes the value TBD-BY-IANA-1 (suggested value is 2).
- o The Assigned Prefix TLV in Section 5.2 takes the value TBD-BY-IANA-2 (suggested value is 3).

9. Analysis

An analysis of a mechanism reminiscent of the one described in this specification has been published in the SIGCOMM IPv6 Workshop [SIGCOMM.IPV6]. Further analysis is encouraged.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3646] Droms, R., "DNS Configuration options for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3646, December 2003.
- [RFC3736] Droms, R., "Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736, April 2004.
- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.
- [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [RFC6106] Jeong, J., Park, S., Beloeil, L., and S. Madanapalli, "IPv6 Router Advertisement Options for DNS Configuration", RFC 6106, November 2010.
- [I-D.acee-ospf-ospfv3-autoconfig]

Lindem, A. and J. Arkko, "OSPFv3 Auto-Configuration", draft-acee-ospf-ospv3-autoconfig-00 (work in progress), October 2011.

10.2. Informative References

- [RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.
- [I-D.chown-homenet-arch]
Arkko, J., Chown, T., Weil, J., and O. Troan, "Home Networking Architecture for IPv6", draft-chown-homenet-arch-00 (work in progress), September 2011.
- [I-D.chelius-router-autoconf]
Chelius, G., Fleury, E., and L. Toutain, "Using OSPFv3 for IPv6 router autoconfiguration", draft-chelius-router-autoconf-00 (work in progress), June 2002.
- [I-D.dimitri-zospf]
Dimitrelis, A. and A. Williams, "Autoconfiguration of routers using a link state routing protocol", draft-dimitri-zospf-00 (work in progress), October 2002.
- [SIGCOMM.IPV6]
Chelius, G., Fleury, E., Sericola, B., Toutain, L., and D. Binet, "An evaluation of the NAP protocol for IPv6 router auto-configuration", ACM SIGCOMM IPv6 Workshop, Kyoto, Japan, 2007.

Appendix A. Acknowledgments

The authors would like to thank to Tim Chown, Fred Baker, Mark Townsley, Lorenzo Colitti, Ole Troan, Ray Bellis, Wassim Haddad, Joel Halpern, Samita Chakrabarti, Michael Richardson, Anders Brandt, Erik Nordmark, Laurent Toutain, and Ralph Droms for interesting discussions in this problem space. The authors would also like to point out some past work in this space, such as those in [I-D.chelius-router-autoconf] or [I-D.dimitri-zospf].

Authors' Addresses

Jari Arkko
Ericsson
Jorvas 02420
Finland

Email: jari.arkko@piuha.net

Acee Lindem
Ericsson
Cary, NC 27519
USA

Email: acee.lindem@ericsson.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 6, 2012

F. Baker
R. Droms
Cisco Systems
October 4, 2011

IPv6 Prefix Assignment in Small Networks
draft-baker-homenet-prefix-assignment-00

Abstract

It is necessary to allocate prefixes in small networks, which include residential and Small Office/Home Office (SOHO) networks in a manner that minimizes or eliminates manual configuration. This note suggests an approach.

Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 6, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

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

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

Table of Contents

1. Introduction	3
2. Scope of this Document	4
3. Simple Tree Network Case	4
3.1. Assignment of prefixes in a simple network	4
3.1.1. CPE Router Behavior	5
3.1.2. Interior Router Behavior	5
4. Issues in a simple cascade procedure	8
4.1. Sequence of subnet number allocation	8
4.2. Multihoming Issues	8
4.3. Race Conditions	8
4.4. Scaling Issues	9
4.5. Prefix Stability	9
4.6. When you run out of prefixes	9
5. Router Advertisement Allocator Information Element	10
6. IANA Considerations	10
7. Security Considerations	10
7.1. Privacy Considerations	10
8. Change Log	10
9. References	10
9.1. Normative References	10
9.2. Informative References	11
Authors' Addresses	12

1. Introduction

One of the objectives of the design of IPv6 [RFC2460] has been to reduce or minimize the need for manual configuration in networks. IPv4 [RFC0791] networks, when it became widely deployed in the 1980's, required manual configuration, and the scaling limits of the approach quickly became apparent. One of the outcomes of that was the Dynamic Host Configuration Protocol [RFC2131] (DHCP), which facilitated central administration of desktop computers. In practice, DHCP itself has been of limited utility in the administration of network equipment; while it is conceptually possible to use it for any kind of configuration, more flexible protocols such as the Network Configuration Protocol [RFC6241][RFC6242] have been preferred.

Allocation of prefixes in small networks calls for an approach that can be completely automated. This note documents a procedure that has been suggested by several. It builds on a few basic assumptions:

- o IPv6 prefixes are allocated to a small network by one or more upstream service providers using [RFC3315] and [RFC3363].
- o IPv6 prefixes may allocated to LAN within a small network by the CPE Router using [RFC3315] and [RFC3363].
- o Occasional inefficiencies such as allocating two /64s to a LAN from a given upstream prefix are acceptable, especially if short-lived.
- o Small networks, such as described in Home Networking Architecture for IPv6 [I-D.chown-homenet-arch], are simple enough in structure that the mechanism described in this note is adequate.

These assumptions bear analysis. The first two, that prefixes can and may be allocated using mechanisms designed for the purpose, seems self-evident. The third builds on the IPv6 premise that a host may have more than one prefix on an interface and one or more addresses in each prefix; in such a case, while it may be suboptimal to allocate more than one /64 from the same upstream prefix, the hosts will not complain and the routing protocols will route them. The fourth may be considered the limit of applicability; if a network requires a prefix aggregation design or is otherwise too complex for this procedure to be effective, other procedures are more appropriate.

2. Scope of this Document

This document describes a procedure for prefix delegation and assignment. It results in the assignment of a series of /64 prefixes on the links in a small home network.

While this document describes the use of DHCPv6 for prefix delegation, specification of the use of DHCPv6 for address assignment and other purposes is out of scope.

If a network includes interior routers and the CPE router is not directly to all of the links in the network, the routers in the network will need routing information to forward traffic in the network and between the network and the service provider network. The specification of a routing protocol or other mechanism to provide that routing information to the routers is beyond the scope of this document.

3. Simple Tree Network Case

The first case to describe is that of a network with a simple tree topology. In this network, there is a single CPE router attached to a single SP network. The interior of the network is organized as a tree. Each interior router has one "upstream" interface and one or more "downstream" interfaces. Each link in the network has a single interior router with a downstream interface attached and zero or more interior routers with an upstream interface attached.

The fundamental procedure for prefix allocation takes three phases:

- o Allocating a prefix from the upstream network,
- o Prefix allocation by the CPE Router, and
- o Prefix allocation by a subsequent router.

3.1. Assignment of prefixes in a simple network

This section describes the assignment of prefixes in a simple network. The network is assumed to be tree-structured, including one CPE router that is connected to a SP network and one or more interior routers. The interior routers each have a single "upstream" interface and one or more "downstream" interfaces. The upstream interface of each interior router is connected to a link in the network to which a downstream interface of a router closer to the CPE router is already connected.

The CPE router obtains a delegated prefix for the entire home network, and manages prefix allocations for all of the interior routers. Each interior router uses DHCPv6 on its upstream interface to obtain delegated prefixes from the CPE router for each of the interior routers downstream interfaces.

3.1.1. CPE Router Behavior

The CPE router obtains a delegated prefix from the SP provisioning system using [RFC3315] and [RFC3363] and other appropriate provisioning systems. The prefix delegated from the service provider includes a preferred and valid lifetime for the prefix.

Once the CPE router has received a delegated prefix, it assigns a /64 subprefix to each of the links to which the router is attached. The CPE router configures an address to each of its interfaces from the prefix assigned to the link to which the interface is attached. After assigning the interface addresses, the CPE router begins sending Router Advertisement (RA) messages [RFC4861] advertising the appropriate prefix on each attached link.

The CPE router includes a Router Advertisement Allocator Information (RAAI) option, identifying itself as the allocating server for prefixes related to the prefix announced in the RA. The RAs include preferred and valid lifetimes derived from the lifetimes associated with the delegated prefix from the service provider. The RA also advertises the CPE router as the default router for the link. Other fields in the RAs are set as appropriate.

At this point, the links to which the CPE router is attached is now provisioned with prefixes taken from the prefix obtained from the service provider. The CPE router uses ongoing DHCPv6 messages exchanges according to [RFC3315] and [RFC3363] to maintain and update its delegated prefix.

The CPE router uses a DHCPv6 server for prefix subdelegation throughout the rest of the network. In preparation for assigning prefixes to links in the rest of the network, the CPE router makes all of the remaining prefixes from the network prefix available for subdelegation through a DHCPv6 server. The CPE router configures the preferred and valid lifetimes for the subdelegated prefixes from the values received from the service provider.

3.1.2. Interior Router Behavior

When an interior router is connected to the home network, its upstream interface is attached to a link in the home network, and its downstream interfaces are connected to other links to be added to the

home network.

3.1.2.1. Network with a Tree Topology

After the upstream interface is attached to a link, the interior router listens for RAs on the upstream interface and configures the upstream interface according to the information contained in the received RAs.

When the interior router receives an RA with an RAAI option, the router initiates a DHCPv6 message exchange to obtain prefixes from the prefix managed by the allocating router. The interior router requests the delegation of a separate /64 prefix for each of its downstream interfaces. The DHCPv6 service in the home network delivers the DHCPv6 traffic between the interior router and the CPE router.

Discussion: The interior router conducts the DHCPv6 message exchange directly with the allocating DHCPv6 server using IPv6 unicast. This technique assumes that the interior router has already obtained an address of sufficient scope through SLAAC or an earlier DHCPv6 address assignment. This technique also breaks the rule in RFC 3315 requiring the use of multicast and the DHCPv6 client's link-local address.

The requirements regarding DHCPv6 message addressing in RFC 3315 are based primarily on the need for some sort of address on the DHCPv6 client before address assignment is completed and the desire to forward all DHCPv6 traffic through a relay agent to allow for relay agent processing. The procedures in this specification require that the interior router (DHCPv6 client) already has an IPv6 address of sufficient scope before initiating any DHCPv6 message exchanges for prefix delegation. There is no need, in this specification, for relay agent processing, so direct communication between the interior router and the allocating DHCPv6 server is allowed.

The primary advantage to allowing direct DHCPv6 message exchanges in this specification is the avoiding the need for a relay agent infrastructure throughout the network. Otherwise, each interior router would have to act as a relay agent for potentially several DHCPv6 servers delegating prefixes for the network.

The CPE router delegates the requested prefixes from the prefix delegated to the network. The interior router then assigns a prefix to each link attached to which a downstream interface is attached, configures those downstream interfaces with addresses from the assigned prefixes and begins sending RAs on the downstream

interfaces. The interior router includes an RAAI option in the RAs, indentifying the CPE router as the allocating DHCPv6 server. The preferred and valid lifetimes for the advertised prefix are derived from the lifetimes in the DHCPv6 delegation, and the RAs advertise the interior router as the default router for the link.

3.1.2.2. Non-tree Topologies

It is quite likely that real world deployments will violate the assumption in the previous section that only one downstream interface will be attached to each link in the home network. In this situation, it is desirable that the link only be assigned one prefix and, therefore, only one of the interior routers with a downstream interface on the link be responsible for assigning a prefix and sending RAs on the link.

To avoid duplicate address assignment, a router first listens for RAs on the link attached to its downstream interface. If the router does not receive an RA after listening for INTERVAL1 microfortnights, the router assumes it is responsible for assigning a prefix to that link and initiates the DHCPv6 process for obtaining a delegated prefix.

After the router determines it is responsible for the link attached to its downstream interface, it continues to listen for RAs from other routers on the link. If it receives an RA from another router, it deassigns its delegated prefix from the link, unconfigures any addresses assigned from that prefix and releases the delegated prefix to the CPE router using DHCPv6.

If a router hears an RA such as described in Section 3.1.2, it uses IPv6 Stateless Address Autoconfiguration [RFC4862][RFC4941] or a DHCPv6 [RFC3315] request to each announced allocator to generate an address within the prefix for use in that subnet.

After the router determines that some other router is responsible for the link attached to its downstream interface, it continues to listen on the interface for RAs. If the router receives no RA on the interface for INTERVAL2 microfortnights, the router takes responsibility for the link and initiates the process described above to obtain and assign a prefix to the link.

3.1.2.3. Multi-homed Network

If a network has multiple service provider networks, it will have multiple prefixes. This situation is easiest to describe if the network is connected to each service provider through a separate CPE router.

Each CPE router obtains a delegated prefix from its service provider and then manages the prefix according to the

First layer of interior router get multiple direct DHCPv6 prefixes. Assigns each prefix in parallel. Sets up DHCPv6 relay agent to point to each of the CPE routers.

Next layer receives DHCPv6 transaction from each CPE router because upstream router forwards DHCPv6 messages to each of the CPE routers.

4. Issues in a simple cascade procedure

There are a number of potential issues in this procedure.

4.1. Sequence of subnet number allocation

Apart from cases in which the administration has chosen to fix a given subnet to a given LAN, such as to support server deployment in DNS, it is generally advised that subnet numbers be randomized. This is to make certain network attacks a little more difficult.

4.2. Multihoming Issues

One issue is "what happens if one has multiple upstream networks with multiple CPE Routers and therefore multiple allocators?" The design of the RA information element announcing the allocator is intended to simplify that by announcing an allocator.

4.3. Race Conditions

In the simplest case, there are no race conditions; the home has exactly one router, it obtains a prefix from its upstream network, and sub-allocates to its interfaces. If there are additional routers in the home, however, either there are one or more links that are not attached to the CPE Router or there are zero; in the event that there are one or more such links, they may be connected by one router or by multiple routers.

One race condition is when two interior routers are attached to the same LANs as the CPE. For example, one might have a wireless router in the home that connects both to the wired and the wireless network that the CPE Router is on. In such a case, it will hear and interpret one of the CPE Router's RAs first, and then the other some amount of time later. The purpose of the INTERVAL1 delay in Section 3.1.2 is to allow this race condition to stabilize before the router acts on this information it has.

A second race condition occurs when two "subsequent" routers are on the same LAN but it is not serviced by the CPE Router. These routers will both use the procedure of Section 3.1.2 to attempt to allocate a prefix to the LAN and so create a subnet. It is RECOMMENDED that the allocator allocate at most one prefix per INTERVAL2, ignoring all other requests, in order to allow the "subsequent" routers to sort out this class of race condition. If needed, ignored routers will re-request the allocation.

Due to the possibility of packet loss in the network, it is possible that these race conditions may result in a given LAN developing multiple subnets. While suboptimal, this is not a violation of the architecture and should cause no issues. However, in the event that two routers observe that they are announcing different subnets in the same upstream prefix on the same LAN, the one with the numerically least subnet number SHOULD NOT allow its prefix to expire, but any others SHOULD allow their prefixes to expire.

4.4. Scaling Issues

Obviously, use of this procedure in a complex network results in a serialization of prefix allocation that may take more time to settle than is operationally desirable (number of LANs times INTERVAL2). In such cases, the administration will have to decide how it wants to handle the issue. One approach would be to divide the network into easily-aggregated sections and use the procedure within each section; another would be to use a different procedure.

In such networks, the routers requesting prefixes can also act as a denial of service attack, by flooding the CPE Router with requests. Given that the procedure eventually terminates, this is undesirable but of limited duration.

4.5. Prefix Stability

In networks that contain servers or names that are announced in DNS, it is often valuable to have the same LAN always have the same subnet number applied to it. The procedure as described could accomplish that if the CPE Router maintains memory of what router it has allocated a given prefix to recently, or would fail to provide that if it does not. The distinction is essentially a marketing requirement that the implementation will need to decide for itself.

4.6. When you run out of prefixes

If a network runs out of subnet numbers and therefore subnet prefixes, this is considered a provisioning failure. It can result when multiple prefixes are allocated to the same LAN, which should be

unusual and will end when one of the routers releases its prefix. It can also result when the upstream network allocates a prefix that is too long and as a result contains too few potential prefixes. In that case, the administration is forced to either reorganize its network or negotiate for a shorter prefix.

5. Router Advertisement Allocator Information Element

On a Neighbor Discovery RA, Section 3.1.2 and Section 3.1.2 call for the RA to identify the allocator that a "subsequent" router may use to request a related prefix for use on a different interface. This information element contains a list of the IPv6 addresses of one or more allocators, and an element length option to permit parsing of the information element.

6. IANA Considerations

In Section 5, this note specifies an information element to be carried in the Router Advertisement message specified in Neighbor Discovery.

7. Security Considerations

<TBD>

7.1. Privacy Considerations

<TBD>

8. Change Log

Initial Version: 4 Octoboer 2011

9. References

9.1. Normative References

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

9.2. Informative References

- [I-D.chown-homenet-arch]
Arkko, J., Chown, T., Weil, J., and O. Troan, "Home Networking Architecture for IPv6", draft-chown-homenet-arch-00 (work in progress), September 2011.
- [RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.
- [RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [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.
- [RFC3363] Bush, R., Durand, A., Fink, B., Gudmundsson, O., and T. Hain, "Representing Internet Protocol version 6 (IPv6) Addresses in the Domain Name System (DNS)", RFC 3363, August 2002.
- [RFC3971] Arkko, J., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.
- [RFC4389] Thaler, D., Talwar, M., and C. Patel, "Neighbor Discovery Proxies (ND Proxy)", RFC 4389, 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.
- [RFC6241] Enns, R., Bjorklund, M., Schoenwaelder, J., and A. Bierman, "Network Configuration Protocol (NETCONF)", RFC 6241, June 2011.
- [RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure

Shell (SSH)", RFC 6242, June 2011.

Authors' Addresses

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

Email: fred@cisco.com

Ralph Droms
Cisco Systems
1414 Massachusetts Avenue
Boxborough, Massachusetts 01719
USA

Email: rdroms@cisco.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 3, 2012

J. Arkko
Ericsson
T. Chown
University of Southampton
J. Weil
Time Warner Cable
O. Troan
Cisco Systems, Inc.
October 31, 2011

Home Networking Architecture for IPv6
draft-chown-homenet-arch-01

Abstract

This text describes evolving networking technology within small "residential home" networks. The goal of this memo is to define the architecture for IPv6-based home networking and the associated principles and considerations. The text highlights the impact of IPv6 on home networking, illustrates topology scenarios, and shows how standard IPv6 mechanisms and addressing can be employed in home networking. The architecture describes the need for specific protocol extensions for certain additional functionality. It is assumed that the IPv6 home network runs as an IPv6-only or dual-stack network, but there are no recommendations in this memo for the IPv4 part of the network.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the

document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Effects of IPv6 on Home Networking	3
3. Architecture	7
3.1. Network Models	8
3.2. Requirements	12
3.3. Considerations	13
3.4. Principles	15
3.5. Summary of Homenet Architecture Recommendations	21
3.6. Implementing the Architecture on IPv6	22
4. References	22
4.1. Normative References	22
4.2. Informative References	23
Appendix A. Acknowledgments	25
Authors' Addresses	26

1. Introduction

This memo focuses on evolving networking technology within small "residential home" networks and the associated challenges. For example, a trend in home networking is the proliferation of networking technology in an increasingly broad range of devices and media. This evolution in scale and diversity sets requirements on IETF protocols. Some of these requirements relate to the need for multiple subnets, for example for private and guest networks, the introduction of IPv6, and the introduction of specialized networks for home automation and sensors.

While advanced home networks have been built, most operate based on IPv4, employ solutions that we would like to avoid such as (cascaded) network address translation (NAT), or require expert assistance to set up. The architectural constructs in this document are focused on the problems to be solved when introducing IPv6 with a eye towards a better result than what we have today with IPv4, as well as a better result than if the IETF had not given this specific guidance.

This architecture document aims to provide the basis and guiding principles for how standard IPv6 mechanisms and addressing [RFC2460] [RFC4291] can be employed in home networking, while coexisting with existing IPv4 mechanisms. In emerging dual-stack home networks it is vital that introducing IPv6 does not adversely affect IPv4 operation. Future deployments, or specific subnets within an otherwise dual-stack home network, may be IPv6-only.

[RFC6204] defines basic requirements for customer edge routers (CPEs). The scope of this text is the homenet, and thus the internal facing interface described that RFC as well as other components within the home network. While the network may be dual-stack or IPv6-only, specific transition tools on the CPE are out of scope of this text, as is any advice regarding architecture of the IPv4 part of the network. We assume that IPv4 network architecture in home networks is what it is, and can not be affected by new recommendations.

2. Effects of IPv6 on Home Networking

Service providers are deploying IPv6, content is becoming available on IPv6, and support for IPv6 is increasingly available in devices and software used in the home. While IPv6 resembles IPv4 in many ways, it changes address allocation principles, makes multi-addressing the norm, and allows direct IP addressability and routing to devices in the home from the Internet. This section presents an overview of some of the key areas impacted by the implementation of

IPv6 into the home network that are both promising and problematic:

Multiple segments and routers

Simple layer 3 topologies involving as few subnets as possible are preferred in home networks for a variety of reasons including simpler management and service discovery. However, the incorporation of dedicated (routed) segments remains necessary for a variety of reasons.

For instance, a common feature in modern home routers is the ability to support both guest and private network segments. Also, link layer networking technology is poised to become more heterogeneous, as networks begin to employ both traditional Ethernet technology and link layers designed for low-powered and lossy networks (LLNs) such as those used for certain types of sensor devices. Similar needs for segmentation may occur in other cases, such as separating building control or corporate extensions from the Internet access network. Also, different segments may be associated with subnets that have different routing and security policies.

Documents that provide some more specific background and depth on this topic include: [I-D.herbst-v6ops-cpeneenhancements], [I-D.baker-fun-multi-router], and [I-D.baker-fun-routing-class].

In addition to routing, rather than NATing, between subnets, there are issues of when and how to extend mechanisms such as service discovery which currently rely on link-local addressing to limit scope.

The presence of a multiple segment, multi-router network implies that there is some kind of automatic routing mechanism in place. In advanced configurations similar to those used in multihomed corporate networks, there may also be a need to discover border router(s) by an appropriate mechanism.

Multi-Addressing of devices

In an IPv6 network, devices may acquire multiple addresses, typically at least a link-local address and a globally unique address. Thus it should be considered the norm for devices on IPv6 home networks to be multi-addressed, and to also have an IPv4 address where the network is dual-stack. Default address selection mechanisms [I-D.ietf-6man-rfc3484-revise] allow a node to select appropriate src/dst address pairs for communications, though such selection may face problems in the event of multihoming, where nodes will be configured with one address from

each upstream ISP prefix, and the presence of upstream ingress filtering thus requires multi-addressed nodes to select the right source address to be used for the corresponding uplink.

Unique Local Addresses (ULAs)

[RFC4193] defines Unique Local Addresses (ULAs) for IPv6 that may be used to address devices within the scope of a single site. Support for ULAs for IPv6 CPEs is described in [RFC6204]. A home network running IPv6 may deploy ULAs for communication between devices within the network. ULAs have the potential to be used for stable addressing in a home network where the externally allocated global prefix changes over time or where external connectivity is temporarily unavailable. However, it is undesirable to aggressively deprecate global prefixes for temporary loss of connectivity, so for this to matter there would have to be a connection breakage longer than the lease period, and even then, deprecating prefixes when there is no connectivity may not be advisable. However, while setting a network up there may be a period with no connectivity.

Another possible reason for using ULAs would be to provide an indication to applications that the traffic is local. This could then be used with security settings to designate where a particular application is allowed to connect to.

Address selection mechanisms should ensure a ULA source address is used to communicate with ULA destination addresses. The use of ULAs does not imply IPv6 NAT, rather that external communications should use a node's global IPv6 source address.

Security, Borders, and the elimination of NAT

Current IPv4 home networks typically receive a single global IPv4 address from their ISP and use NAT with private [RFC1918]. addressing for devices within the network. An IPv6 home network removes the need to use NAT given the ISP offers a sufficiently large IPv6 prefix to the homenet, allowing every device on every link to be assigned a globally unique IPv6 address.

The end-to-end communication that is potentially enabled with IPv6 is both an incredible opportunity for innovation and simpler network operation, but it is also a concern as it exposes nodes in the internal networks to receipt of otherwise unwanted traffic from the Internet.

In IPv4 NAT networks, the NAT provides an implicit firewall function. [RFC4864] suggests that IPv6 networks with global addresses utilise "Simple Security" in border firewalls to restrict incoming connections through a default deny policy. Applications or hosts wanting to accept inbound connections then need to signal that desire through a protocol such as uPNP or PCP [I-D.ietf-pcp-base].

Such an approach would reduce the efficacy of end-to-end connectivity that IPv6 has the potential to restore, since the need for IPv4 NAT traversal is replaced by a need to use a signalling protocol to request a firewall hole be opened. [RFC6092] provides recommendations for an IPv6 firewall that applies "limitations on end-to-end transparency where security considerations are deemed important to promote local and Internet security." The firewall operation is "simple" in that there is an assumption that traffic which is to be blocked by default is defined in the RFC and not expected to be updated by the user or otherwise. The RFC does however state that CPEs should have an option to be put into a "transparent mode" of operation.

It is important to distinguish between addressability and reachability; i.e. IPv6 through use of globally unique addressing in the home makes all devices potentially reachable from anywhere. Whether they are or not should depend on firewall or filtering configuration, and not the presence or use of NAT.

Advanced Security for IPv6 CPE [I-D.vyncke-advanced-ipv6-security] takes the approach that in order to provide the greatest end-to-end transparency as well as security, security policies must be updated by a trusted party which can provide intrusion signatures and other "active" information on security threats. This is much like a virus-scanning tool which must receive updates in order to detect and/or neutralize the latest attacks as they arrive. As the name implies "advanced" security requires significantly more resources and infrastructure (including a source for attack signatures) in comparison to "simple" security.

In addition to establishing the security mechanisms themselves, it is important to know where to enable them. If there is some indication as to which router is connected to the "outside" of the home network, this is feasible. Otherwise, it can be difficult to know which security policies to apply where. Further, security policies may be different for various address ranges if ULA addressing is setup to only operate within the homenet itself and not be routed to the Internet at large. Finally, such policies must be able to be applied by typical home users, e.g. to give a visitor in a "guest" network access to media services in the home.

It may be useful to classify the border of the home network as a unique logical interface separating the home network from service provider network/s. This border interface may be a single physical interface to a single service provider, multiple layer 2 sub-interfaces to a single service provider, or multiple connections to a single or multiple providers. This border is useful for describing edge operations and interface requirements across multiple functional areas including security, routing, service discovery, and router discovery.

Naming, and manual configuration of IP addresses

In IPv4, a single subnet NATed home network environment is currently the norm. As a result, it is for example common practice for users to be able to connect to a router for configuration via a literal address such as 192.168.1.1 or some other commonly used RFC 1918 address. In IPv6, while ULAs exist and could potentially be used to address internally-reachable services, little deployment experience exists to date. Given a true ULA prefix is effectively a random 48-bit prefix, it is not reasonable to expect users to manually enter such address literals for configuration or other purposes. As such, even for the simplest of functions, naming and the associated discovery of services is imperative for an easy to administer homenet.

In a multi-subnet homenet, naming and service discovery should be expected to operate across the scope of the entire home network, and thus be able to cross subnet boundaries. It should be noted that in IPv4, such services do not generally function across home router NAT boundaries, so this is one area where there is scope for an improvement in IPv6.

3. Architecture

An architecture outlines how to construct home networks involving multiple routers and subnets. In this section, we present a set of typical home network topology models/scenarios, followed by a list of topics that may influence the architecture discussions, and a set of architectural principles that govern how the various nodes should work together. Finally, some guidelines are given for realizing the architecture with the IPv6 addressing, prefix delegation, global and ULA addresses, source address selection rules and other existing components of the IPv6 architecture. The architecture also drives what protocol extensions are necessary, as will be discussed in Section 3.6.

3.1. Network Models

Figure 1 shows the simplest possible home network topology, involving just one router, a local area network, and a set of hosts. Setting up such networks is in principle well understood today [RFC6204].

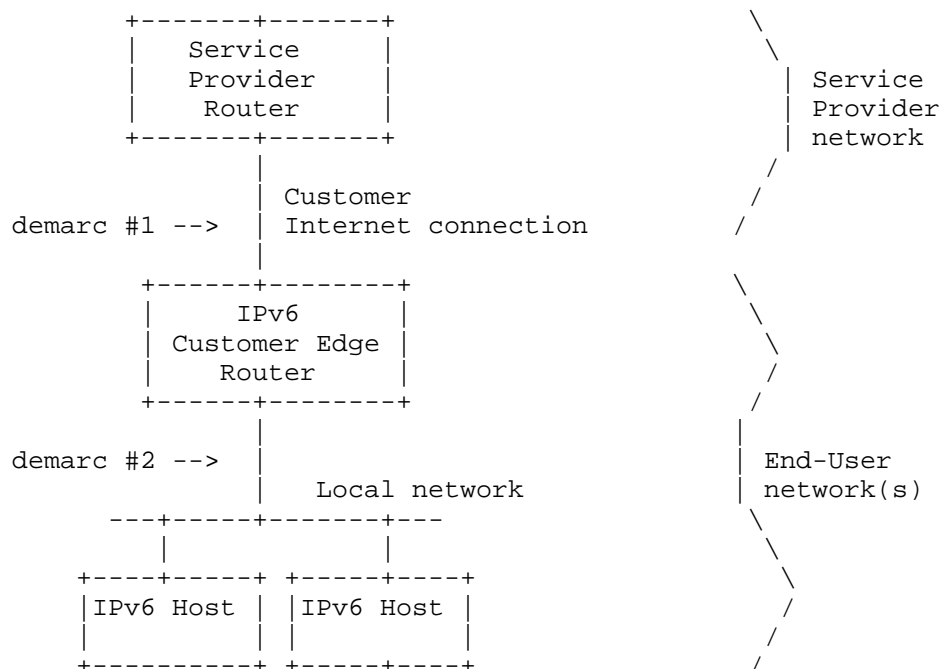


Figure 1

Two possible demarcation points are illustrated in Figure 1, which indicate which party is responsible for configuration or autoconfiguration. Demarcation #1 makes the Customer Edge Router the responsibility of the customer. This is only practical if the Customer Edge Router can function with factory defaults installed. The Customer Edge Router may be pre-configured by the ISP, or by some suitably simple method by the home customer. Demarcation #2 makes the Customer Edge Router the responsibility of the provider. Both models of operation must be supported in the homenet architecture, including the scenarios below with multiple ISPs and demarcation points.

Figure 2 shows another network that now introduces multiple local area networks. These may be needed for reasons relating to different link layer technologies in use or for policy reasons. Note that a

common arrangement is to have different link types supported on the same router, bridged together.

This topology is also relatively well understood today [RFC6204], though it certainly presents additional demands with regards suitable firewall policies and limits the operation of certain applications and discovery mechanisms (which may typically today only succeed within a single subnet).

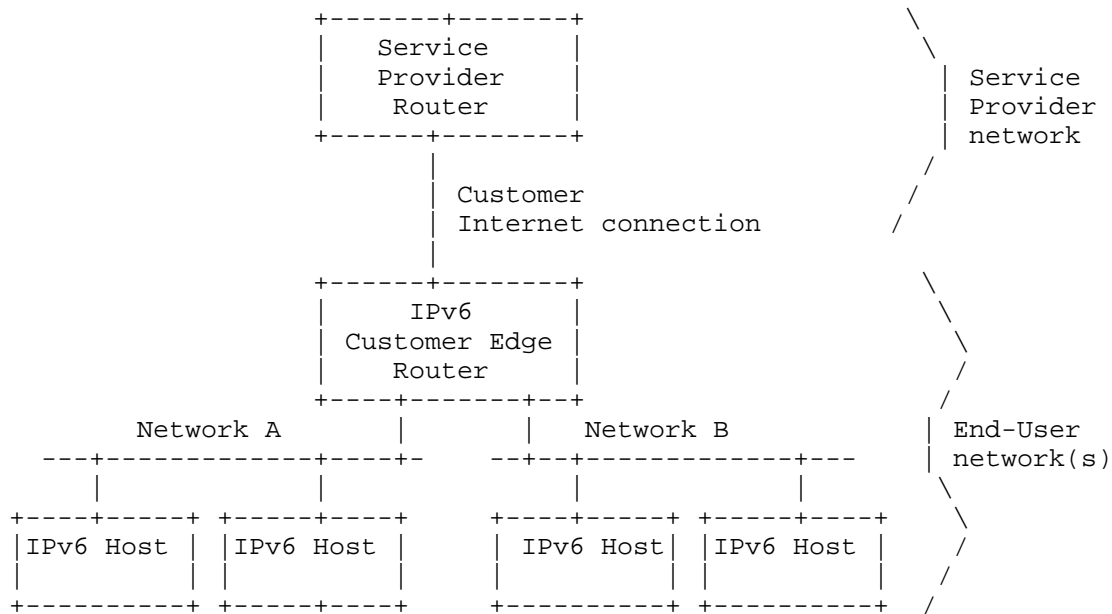


Figure 2

Figure 3 shows a little bit more complex network with two routers and eight devices connected to one ISP. This network is similar to the one discussed in [I-D.ietf-v6ops-ipv6-cpe-router-bis]. The main complication in this topology compared to the ones described earlier is that there is no longer a single router that a priori understands the entire topology. The topology itself may also be complex. It may not be possible to assume a pure tree form, for instance. This would be a consideration if there was an assumption that home users may plug routers together to form arbitrary topologies.

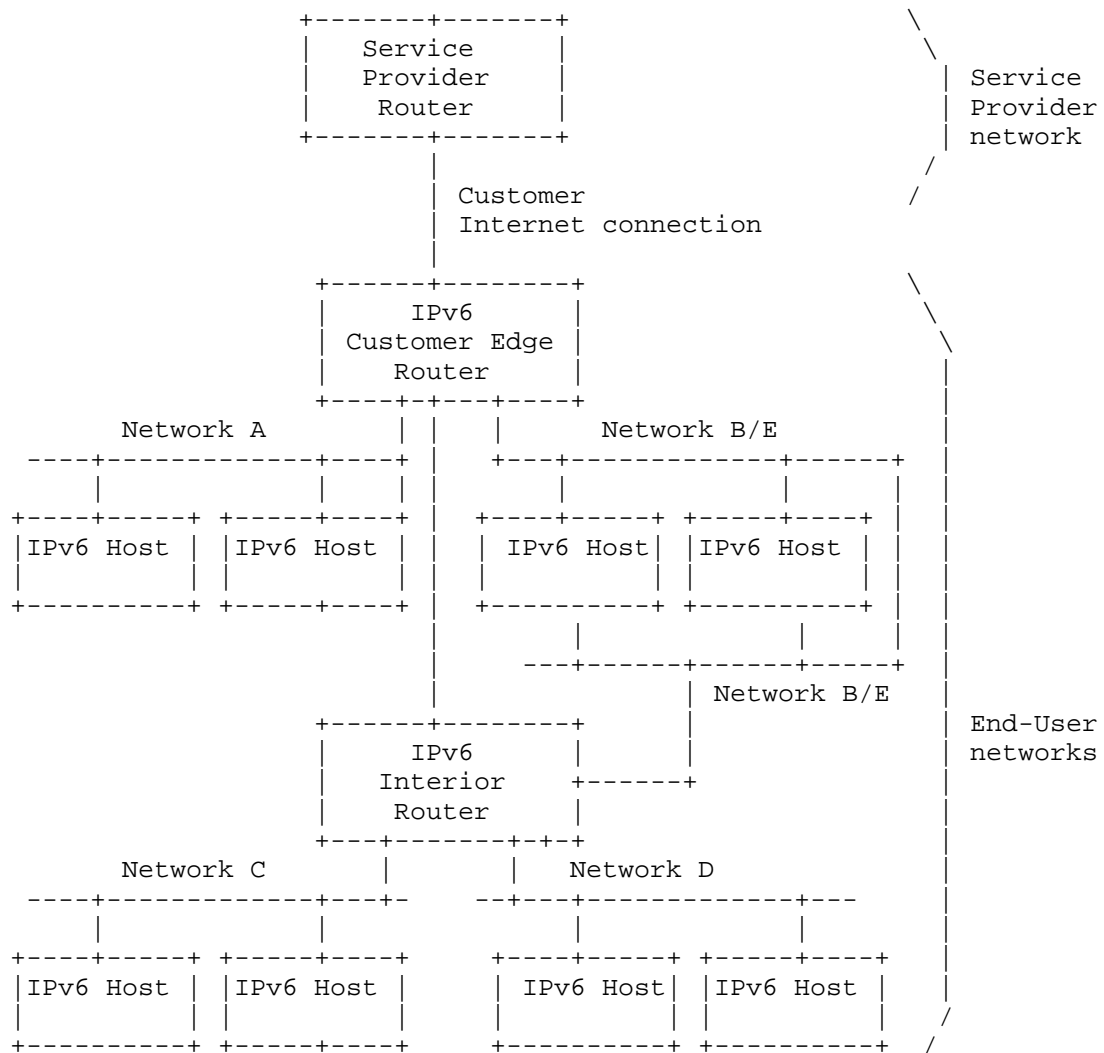


Figure 3

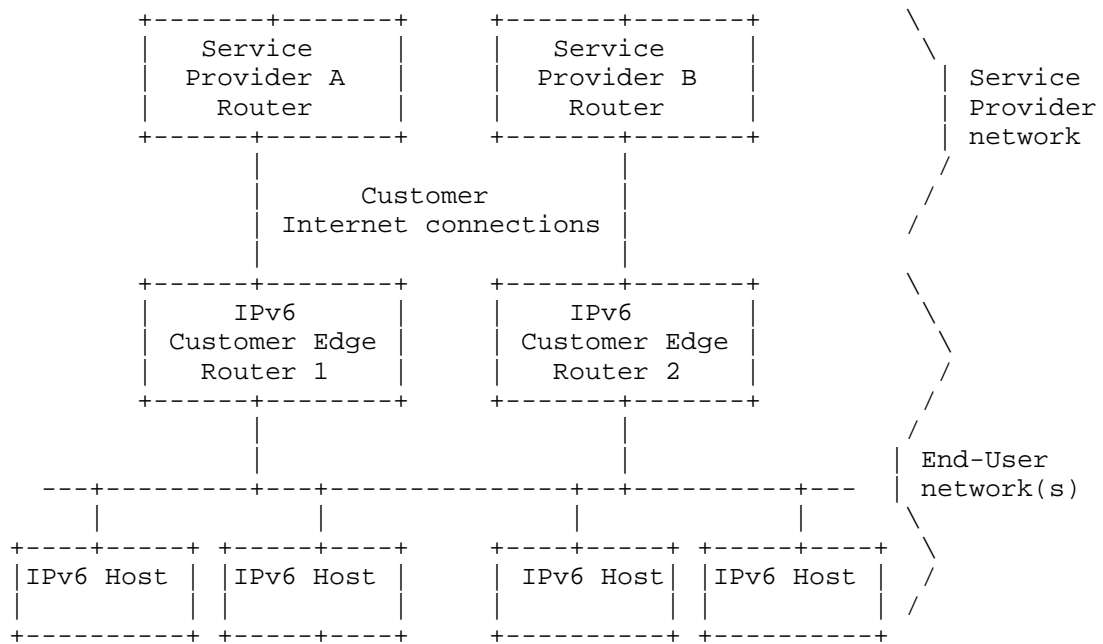


Figure 4

Figure 4 illustrates a multihomed home network model, where the customer has connectivity via CPE1 to ISP A and via CPE2 to ISP B. This example shows one shared subnet where IPv6 nodes would potentially be multihomed and receive multiple IPv6 global addresses, one per ISP. This model may also be combined with that shown in Figure 3 for example to create a more complex scenario.

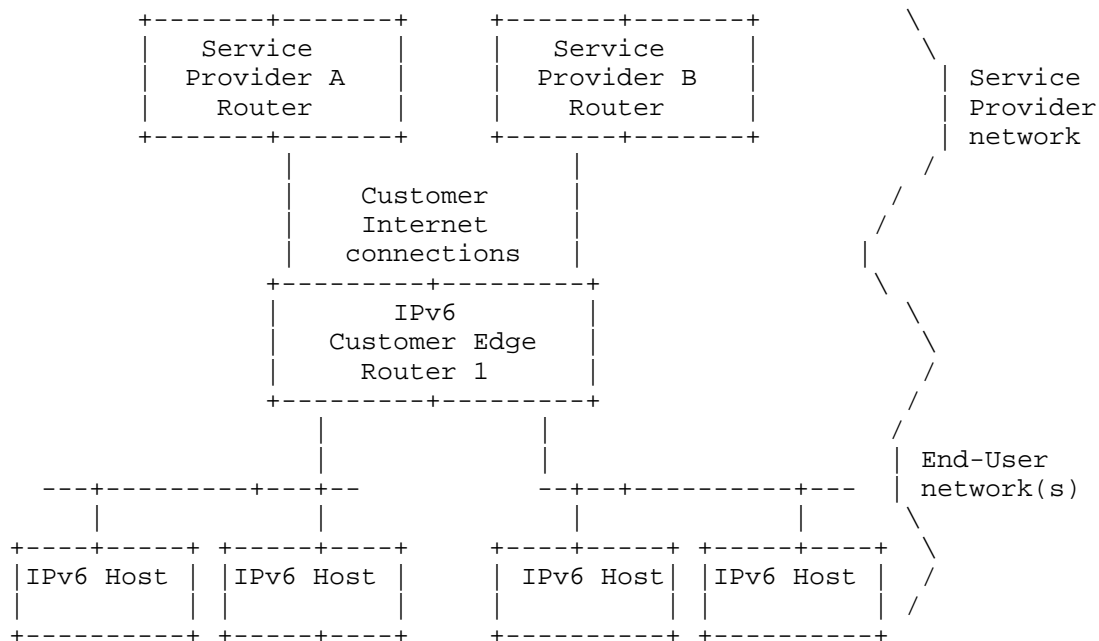


Figure 5

Figure 5 illustrates a model where a home network may have multiple connections to multiple providers or multiple logical connections to the same provider, but the associated subnet(s) are isolated. Some deployment scenarios may require this model.

3.2. Requirements

[RFC6204] defines "basic" requirements for IPv6 Customer Edge Routers, while [I-D.ietf-v6ops-ipv6-cpe-router-bis] describes "advanced" features. In general, home network equipment needs to cope with the different types of network topologies discussed above. Manual configuration is rarely, if at all, possible, given the knowledge lying with typical home users. The equipment needs to be prepared to handle at least

- o Prefix configuration for routers
- o Managing routing
- o Name resolution
- o Service discovery

- o Network security

3.3. Considerations

This section lists some considerations for home networking that may affect the architecture and associated requirements.

Multihoming

A homenet may be multihomed to multiple providers. This may either take a form where there are multiple isolated networks within the home or a more integrated network where the connectivity selection is dynamic. Current practice is typically of the former kind, but the latter is expected to become more commonplace.

In an integrated network, specific appliances or applications may use their own external connectivity, or the entire network may change its connectivity based on the status of the different upstream connections. Many general solutions for IPv6 multihoming have been worked on for years in the IETF, though to date there is little deployment of these mechanisms. While an argument can be made that home networking standards should not make another attempt at this, the obvious counter-argument is that multihoming support will be necessary for many deployment situations.

One such approach is the use of NPTv6 [RFC6296], which is a prefix translation-based mechanism. An alternative is presented in [I-D.v6ops-multihoming-without-ipv6nat]. Host-based methods such as Shim6 [RFC5533] have also been defined.

In any case, if multihoming is supported additional requirements are necessary. The general multihoming problem is broad, and solutions may include complex architectures for monitoring connectivity, traffic engineering, identifier-locator separation, connection survivability across multihoming events, and so on. However, there is a general agreement that for the home case, if there is any support for multihoming it should be limited to a very small subset of the overall problem. Specifically, multi-addressed hosts selecting the right source address to avoid falling foul of ingress filtering on upstream ISP connections [I-D.baker-fun-multi-router]. A solution to this particular problem is desirable.

Some similar multihoming issues have already been teased out in the work described in [I-D.ietf-mif-dns-server-selection], which has led to the definition of a DHCPv6 route option [I-D.ietf-mif-dhcpv6-route-option].

One could also argue that a "happy eyeballs" approach, not too dissimilar to that proposed for multiple interface (mif) scenarios, is also acceptable if such support becomes commonplace in hosts and applications.

A further consideration and complexity here is that at least one upstream may be a "walled garden", and thus only appropriate to be used for connectivity to the services of that provider.

Quality of Service in multi-service home networks

Support for QoS in a multi-service homenet may be a requirement, e.g. for a critical system (perhaps healthcare related), or for differentiation between different types of traffic (file sharing, cloud storage, live streaming, VoIP, etc). Different media types may have different QoS properties or capabilities.

However, homenet scenarios should require no new QoS protocols. A DiffServ [RFC2475] approach with a small number of predefined traffic classes should generally be sufficient, though at present there is little experience of QoS deployment in home networks. There may also be complementary mechanisms that could be beneficial in the homenet domain, such as ensuring proper buffering algorithms are used as described in [Gettys11].

DNS services

A desirable target may be a fully functional self-configuring secure local DNS service so that all devices are referred to by name, and these FQDNs are resolved locally. This will make clean use of ULAs and multiple ISP-provided prefixes much easier. The local DNS service should be (by default) authoritative for the local name space in both IPv4 and IPv6. A dual-stack residential gateway should include a dual-stack DNS server.

Consideration will also need to be given for existing protocols that may be used within a network, e.g. mDNS, and how these interact with unicast-based DNS services.

With the introduction of new top level domains, there is potential for ambiguity between for example a local host called apple and (if it is registered) an apple gTLD, so some local name space is probably required, which should also be configurable to something else by a home user if desired.

Privacy considerations

There are no specific privacy concerns for this text. It should be noted that most ISPs are expected to offer static IPv6 prefixes to customers, and thus the addresses they use would not generally change over time.

3.4. Principles

There is little that the Internet standards community can do about the physical topologies or the need for some networks to be separated at the network layer for policy or link layer compatibility reasons. However, there is a lot of flexibility in using IP addressing and inter-networking mechanisms. In this section we provide some guidance on how this flexibility should be used to provide the best user experience and ensure that the network can evolve with new applications in the future.

The following principles should be used as a guide in designing these networks in the correct manner. There is no implied priority by the order in which the principles are listed.

Reuse existing protocols

It is desirable to reuse existing protocols where possible, but at the same time to avoid consciously precluding the introduction of new or emerging protocols. For example, [I-D.baker-fun-routing-class] suggests introducing a routing protocol that may route on both source and destination addresses.

A generally conservative approach, giving weight to running code, is preferable. Where new protocols are required, evidence of commitment to implementation by appropriate vendors or development communities is highly desirable. Protocols used should be backwardly compatible.

Where possible, changes to hosts should be minimised. Some changes may be unavoidable however, e.g. signalling protocols to punch holes in firewalls where "Simple Security" is deployed in a CPE.

Liaisons with other appropriate standards groups and related organisations is desirable, e.g. the IEEE and Wi-Fi Alliance.

Dual-stack Operation

The homenet architecture targets both IPv6-only and dual-stack networks. While the CPE requirements in RFC 6204 are targeted at IPv6-only networks, it is likely that dual-stack homenets will be the norm for some period of time. IPv6-only networking may first be deployed in home networks in "greenfield" scenarios, or perhaps as one element of an otherwise dual-stack network. The homenet architecture must operate in the absence of IPv4, and IPv6 must work in the same scenarios as IPv4 today. Running IPv6-only may require documentation of additional considerations such as:

Ensuring there is a way to access content in the IPv4 Internet. This can be arranged through incorporating NAT64 [RFC6144] functionality in the home gateway router, for instance.

DNS discovery mechanisms are enabled even for IPv6. Both stateless DHCPv6 [RFC3736] [RFC3646] and Router Advertisement options [RFC6106] may have to be supported and turned on by default to ensure maximum compatibility with all types of hosts in the network. This requires, however, that a working DNS server is known and addressable via IPv6.

All nodes in the home network support operations in IPv6-only mode. Some current devices work well with dual-stack but fail to recognize connectivity when IPv4 DHCP fails, for instance.

In dual-stack networks, solutions for IPv6 must not adversely affect IPv4 operation. It is likely that topologies of IPv4 and IPv6 networks would be as congruent as possible.

Note that specific transition tools, particularly those running on the border CPE, are out of scope. The homenet architecture focuses on the internal home network.

Largest Possible Subnets

Today's IPv4 home networks generally have a single subnet, and early dual-stack deployments have a single congruent IPv6 subnet, possibly with some bridging functionality.

Future home networks are highly likely to need multiple subnets, for the reasons described earlier. As part of the self-organisation of the network, the network should subdivide itself to the largest possible subnets that can be constructed within the constraints of link layer mechanisms, bridging, physical connectivity, and policy. For instance, separate subnetworks are necessary where two different links cannot be bridged, or when a

policy requires the separation of a private and visitor parts of the network.

While it may be desirable to maximise the chance of link-local protocols succeeding, multiple subnet home networks are inevitable, so their support must be included. A general recommendation is to follow the same topology for IPv6 as is used for IPv4, but not to use NAT. Thus there should be routed IPv6 where an IPv4 NAT is used, and where there is no NAT there should be bridging.

In some cases IPv4 NAT home networks may feature cascaded NATs, e.g. where NAT routers are included within VMs or Internet connection services are used. IPv6 routed versions of such tools will be required.

Transparent End-to-End Communications

An IPv6-based home network architecture should naturally offer a transparent end-to-end communications model. Each device should be addressable by a unique address. Security perimeters can of course restrict the end-to-end communications, but it is simpler given the availability of globally unique addresses to block certain nodes from communicating by use of an appropriate filtering device than to configure the address translation device to enable appropriate address/port forwarding in the presence of a NAT.

As discussed previously, it is important to note the difference between hosts being addressable and reachable. Thus filtering is to be expected, while IPv6 NAT is not. End-to-end communications are important for their robustness to failure of intermediate systems, where in contrast NAT is dependent on state machines which are not self-healing.

When configuring filters, protocols for securely associating devices are desirable. In the presence of "Simple Security" the use of signalling protocols such as uPnP or PCP may be expected to punch holes in the firewall. Alternatively, RFC 6092 supports the option for a border CPE to run in "transparent mode", in which case a protocol like PCP is not required, but the security model is more open.

IP Connectivity between All Nodes

A logical consequence of the end-to-end communications model is that the network should by default attempt to provide IP-layer connectivity between all internal parts as well as between the

internal parts and the Internet. This connectivity should be established at the link layer, if possible, and using routing at the IP layer otherwise.

Local addressing (ULAs) may be used within the scope of a home network. It would be expected that ULAs may be used alongside one or more globally unique ISP-provided addresses/prefixes in a homenet. ULAs may be used for all devices, not just those intended to have internal connectivity only. ULAs may then be used for stable internal communications should the ISP-provided prefix change, or external connectivity be temporarily lost. The use of ULAs should be restricted to the homenet scope through filtering at the border(s) of the homenet; thus "end-to-end" for ULAs is limited to the homenet.

In some cases full internal connectivity may not be desirable, e.g. in certain utility networking scenarios, or where filtering is required for policy reasons against guest network subnet(s). Note that certain scenarios may require co-existence of ISP connectivity providing a general Internet service with provider connectivity to a private "walled garden" network.

Some home networking scenarios/models may involve isolated subnet(s) with their own CPEs. In such cases connectivity would only be expected within each isolated network (though traffic may potentially pass between them via external providers).

Routing functionality

Routing functionality is required when multiple subnets are in use. This functionality could be as simple as the current "default route is up" model of IPv4 NAT, or it could involve running an appropriate routing protocol.

The homenet routing environment may include traditional IP networking where existing link-state or distance-vector protocols may be used, but also new LLN or other "constrained" networks where other protocols may be more appropriate. IPv6 VM solutions may also add additional routing requirements. Current home deployments use largely different mechanisms in sensor and basic Internet connectivity networks. In general, LLN or other networks should be able to attach and participate the same way or map/be gatewayed to the main homenet.

It is desirable that the routing protocol has knowledge of the homenet topology, which implies a link-state protocol may be preferable. If so, it is also desirable that the announcements and use of LSAs and RAs are appropriately coordinated.

The routing environment should be self-configuring, as discussed in the next subsection. An example of how OSPFv3 can be self-configuring in a homenet is described in [I-D.acee-ospf-ospfv3-autoconfig]. It is important that self-configuration with "unintended" devices is avoided.

To support multihoming within a homenet, a routing protocol that can make routing decisions based on source and destination addresses is desirable, to avoid upstream ISP ingress filtering problems. In general the routing protocol should support multiple ISP uplinks and prefixes in concurrent use.

Self-Organisation

A home network architecture should be naturally self-organising and self-configuring under different circumstances relating to the connectivity status to the Internet, number of devices, and physical topology.

The most important function in this respect is prefix delegation and management. Delegation should be autonomous, and not assume a flat or hierarchical model. From the homenet perspective, a single prefix should be received on the border CPE from the upstream ISP, via [RFC3363]. The ISP should only see that aggregate, and not single /64 prefixes allocated within the homenet.

Each link in the homenet should receive a prefix from within the ISP-provided prefix. Delegation within the homenet should give each link a prefix that is persistent across reboots, power outages and similar short-term outages. Addition of a new routing device should not affect existing persistent prefixes, but persistence may not be expected in the face of significant "replumbing" of the homenet. Persistence should not depend on router boot order. Persistent prefixes may imply the need for stable storage on routing devices, and also a method for a home user to "reset" the stored prefix should a significant reconfiguration be required.

The assignment mechanism should provide reasonable efficiency, so that typical home network prefix allocation sizes can accommodate all the necessary /64 allocations in most cases. For instance, duplicate assignment of multiple /64s to the same network should be avoided.

Several proposals have been made for prefix delegation within a homenet. One group of proposals is based on DHCPv6 PD, as described in [I-D.baker-homenet-prefix-assignment], [I-D.chakrabarti-homenet-prefix-alloc], [RFC3315] and [RFC3363]. The other uses OSPFv3, as described in [I-D.arkko-homenet-prefix-assignment].

While the homenet should be self-organising, it should be possible to manually adjust (override) the current configuration. The network should also cope gracefully in the event of prefix exhaustion.

The network elements will need to be integrated in a way that takes account of the various lifetimes on timers that are used, e.g. DHCPv6 PD, router, valid prefix and preferred prefix timers.

The homenet will have one or more borders, with external connectivity providers and potentially parts of the internal network (e.g. for policy-based reasons). It should be possible to automatically perform border discovery at least for the ISP borders. Such borders determine for example the scope of ULAs, site scope multicast boundaries and where firewall policies may be applied.

The network cannot be expected to be completely self-organising, e.g. some security parameters are likely to need manual configuration, e.g. WPA2 configuration for wireless access control.

Fewest Topology Assumptions

There should be ideally no built-in assumptions about the topology in home networks, as users are capable of connecting their devices in ingenious ways. Thus arbitrary topologies will need to be supported.

It is important not to introduce new IPv6 scenarios that would break with IPv4+NAT, given dual-stack homenets will be commonplace for some time. There may be IPv6-only topologies that work where IPv4 is not used or required.

Naming and Service Discovery

The most natural way to think about naming and service discovery within a home is to enable it to work across the entire residence, disregarding technical borders such as subnets but respecting policy borders such as those between visitor and internal networks.

This may imply support is required for IPv6 multicast across the scope of the home network, and thus at least all routing devices in the network.

Homenet naming systems will be required that work internally or externally, though the domains used may be different in each case.

Proxy or Extend?

Related to the above, we believe that general existing discovery protocols that are designed to only work within a subnet are modified/extended to work across subnets, rather than defining proxy capabilities for those functions.

We may need to do more analysis (a survey?) on which functions/protocols assume subnet-only operation, in the context of existing home networks. Some experience from enterprises may be relevant here.

Adapt to ISP constraints

The home network may receive an arbitrary length IPv6 prefix from its provider, e.g. /60 or /56. The offered prefix may be static or dynamic. The home network needs to be adaptable to such ISP policies, e.g. on constraints placed by the size of prefix offered by the ISP. The ISP may use [I-D.ietf-dhc-pd-exclude] for example.

The internal operation of the home network should not also depend on the availability of the ISP network at any given time, other than for connectivity to services or systems off the home network. This implies the use of ULAs as supported in RFC6204. If used, ULA addresses should be stable so that they can always be used internally, independent of the link to the ISP.

It is expected that ISPs will deliver a static home prefix to customers. However, it is possible, however unlikely, that an ISP may need to restructure and in doing so renumber its customer homenets. In such cases "flash" renumbering may be imposed. Thus it's desirable that homenet protocols or operational processes don't add unnecessary complexity for renumbering.

3.5. Summary of Homenet Architecture Recommendations

In this section we present a summary of the homenet architecture recommendations that were discussed in more detail in the previous sections.

(Bullet points to be added in next version)

3.6. Implementing the Architecture on IPv6

The necessary mechanisms are largely already part of the IPv6 protocol set and common implementations, though there are some exceptions. For automatic routing, it is expected that existing routing protocols can be used as is. However, a new mechanism may be needed in order to turn a selected protocol on by default. Support for multiple exit routers and multi-homing would also require extensions, even if focused on the problem of multi-addressed hosts selecting the right source address to avoid falling foul of ingress filtering on upstream ISP connections.

For name resolution and service discovery, extensions to existing multicast-based name resolution protocols are needed to enable them to work across subnets, within the scope of the home network.

The hardest problems in developing solutions for home networking IPv6 architectures include discovering the right borders where the domain "home" ends and the service provider domain begins, deciding whether some of necessary discovery mechanism extensions should affect only the network infrastructure or also hosts, and the ability to turn on routing, prefix delegation and other functions in a backwards compatible manner.

4. References

4.1. Normative References

- [RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.
- [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.
- [RFC3363] Bush, R., Durand, A., Fink, B., Gudmundsson, O., and T. Hain, "Representing Internet Protocol version 6 (IPv6)

Addresses in the Domain Name System (DNS)", RFC 3363, August 2002.

- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.
- [RFC4864] Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6", RFC 4864, May 2007.
- [RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, June 2009.
- [RFC6092] Woodyatt, J., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, January 2011.
- [RFC6204] Singh, H., Beebe, W., Donley, C., Stark, B., and O. Troan, "Basic Requirements for IPv6 Customer Edge Routers", RFC 6204, April 2011.
- [RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, June 2011.

4.2. Informative References

- [RFC3646] Droms, R., "DNS Configuration options for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3646, December 2003.
- [RFC3736] Droms, R., "Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736, April 2004.
- [RFC6106] Jeong, J., Park, S., Beloeil, L., and S. Madanapalli, "IPv6 Router Advertisement Options for DNS Configuration", RFC 6106, November 2010.
- [RFC6144] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", RFC 6144, April 2011.
- [I-D.baker-fun-multi-router]
Baker, F., "Exploring the multi-router SOHO network", draft-baker-fun-multi-router-00 (work in progress), July 2011.

- [I-D.baker-fun-routing-class]
Baker, F., "Routing a Traffic Class",
draft-baker-fun-routing-class-00 (work in progress),
July 2011.
- [I-D.herbst-v6ops-cpeenancements]
Herbst, T. and D. Sturek, "CPE Considerations in IPv6
Deployments", draft-herbst-v6ops-cpeenancements-00 (work
in progress), October 2010.
- [I-D.vyncke-advanced-ipv6-security]
Vyncke, E., Yourtchenko, A., and M. Townsley, "Advanced
Security for IPv6 CPE",
draft-vyncke-advanced-ipv6-security-02 (work in progress),
July 2011.
- [I-D.ietf-v6ops-ipv6-cpe-router-bis]
Singh, H., Beebee, W., Donley, C., Stark, B., and O.
Troan, "Advanced Requirements for IPv6 Customer Edge
Routers", draft-ietf-v6ops-ipv6-cpe-router-bis-01 (work in
progress), July 2011.
- [I-D.ietf-6man-rfc3484-revise]
Matsumoto, A., Kato, J., Fujisaki, T., and T. Chown,
"Update to RFC 3484 Default Address Selection for IPv6",
draft-ietf-6man-rfc3484-revise-04 (work in progress),
July 2011.
- [I-D.ietf-dhc-pd-exclude]
Korhonen, J., Savolainen, T., Krishnan, S., and O. Troan,
"Prefix Exclude Option for DHCPv6-based Prefix
Delegation", draft-ietf-dhc-pd-exclude-03 (work in
progress), August 2011.
- [I-D.v6ops-multihoming-without-ipv6nat]
Troan, O., Miles, D., Matsushima, S., Okimoto, T., and D.
Wing, "IPv6 Multihoming without Network Address
Translation", draft-v6ops-multihoming-without-ipv6nat-00
(work in progress), March 2011.
- [I-D.ietf-mif-dns-server-selection]
Savolainen, T., Kato, J., and T. Lemon, "Improved DNS
Server Selection for Multi-Interfaced Nodes",
draft-ietf-mif-dns-server-selection-07 (work in progress),
October 2011.
- [I-D.ietf-mif-dhcpv6-route-option]
Dec, W., Mrugalski, T., Sun, T., and B. Sarikaya, "DHCPv6

Route Options", draft-ietf-mif-dhcpv6-route-option-03 (work in progress), September 2011.

[I-D.baker-homenet-prefix-assignment]
Baker, F. and R. Droms, "IPv6 Prefix Assignment in Small Networks", draft-baker-homenet-prefix-assignment-00 (work in progress), October 2011.

[I-D.arkko-homenet-prefix-assignment]
Arkko, J. and A. Lindem, "Prefix Assignment in a Home Network", draft-arkko-homenet-prefix-assignment-00 (work in progress), October 2011.

[I-D.acee-ospf-ospfv3-autoconfig]
Lindem, A. and J. Arkko, "OSPFv3 Auto-Configuration", draft-acee-ospf-ospfv3-autoconfig-00 (work in progress), October 2011.

[I-D.ietf-pcp-base]
Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", draft-ietf-pcp-base-16 (work in progress), October 2011.

[I-D.chakrabarti-homenet-prefix-alloc]
Nordmark, E., Chakrabarti, S., Krishnan, S., and W. Haddad, "Simple Approach to Prefix Distribution in Basic Home Networks", draft-chakrabarti-homenet-prefix-alloc-01 (work in progress), October 2011.

[Gettys11]
Gettys, J., "Bufferbloat: Dark Buffers in the Internet", March 2011,
<<http://www.ietf.org/proceedings/80/slides/tsvarea-1.pdf>>.

Appendix A. Acknowledgments

The authors would like to thank Brian Carpenter, Mark Andrews, Fred Baker, Ray Bellis, Cameron Byrne, Stuart Cheshire, Lorenzo Colitti, Ralph Droms, Lars Eggert, Jim Gettys, Wassim Haddad, Joel M. Halpern, David Harrington, Lee Howard, Ray Hunter, Joel Jaeggli, Heather Kirksey, Ted Lemon, Erik Nordmark, Michael Richardson, Barbara Stark, Sander Steffann, Dave Thaler, JP Vasseur, Curtis Villamizar, Russ White, and James Woodyatt for their contributions within homenet WG meetings and the mailing list, and Mark Townsley for being an initial editor/author of this text before taking his position as homenet WG co-chair.

Authors' Addresses

Jari Arkko
Ericsson
Jorvas 02420
Finland

Email: jari.arkko@piuha.net

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

Email: tjc@ecs.soton.ac.uk

Jason Weil
Time Warner Cable
13820 Sunrise Valley Drive
Herndon, VA 20171
USA

Email: jason.weil@twcable.com

Ole Troan
Cisco Systems, Inc.
Drammensveien 145A
Oslo N-0212
Norway

Email: ot@cisco.com

Home Networking
Internet-Draft
Intended status: Standards Track
Expires: April 26, 2012

W. Haddad
J. Halpern
Ericsson
October 24, 2011

Ensuring Home Network Visibility to Home Gateway
draft-haddad-homenet-gateway-visibility-00

Abstract

This memo describes a mechanism designed to increase the home gateway visibility on the home network that it is serving. This includes knowledge of all IPv6 addresses configured using prefixes assigned by the home gateway and advertised by router(s) attached to it.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on April 26, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Conventions used in this document	4
3. Motivation	5
4. Proposal	6
5. New Messages Structures and Options Format	8
6. Security Considerations	9
7. IANA Considerations	10
8. Normative References	11
Authors' Addresses	12

1. Introduction

With the expected proliferation of "smart home" networks, enabling multiple features and capabilities may require installing additional routers within the home that will connect to one or multiple home gateway (HGW(s)). In such scenario, it can be useful for the HGW(s) to keep track of all IPv6 addresses configured by different types of end devices that get attached to the home network via router(s) connected to the HGW(s).

This memo describes a mechanism designed to address this scenario by increasing the HGW visibility on the home network that it is serving, without incurring any change on the end devices.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Motivation

Future smart home networks are all about deploying new services within homes and enabling average users (i.e., the vast majority of Internet users) to easily interact with them. For this purpose, enabling automatic services/features discovery as well as associated home device(s) configuration (i.e., specifically for end devices that are not directly connected to the HGW) is a useful feature to provide. In fact, such feature would help assisting average user to seamlessly manage and configure home devices.

4. Proposal

For simplicity and better clarity, we consider in the following a home network composed of one HGW, two additional routers (R1) and (R2) and a set of home devices that are spread around the three network entities, i.e., both (R1) and (R2) are connecting a subset of home devices while the remaining ones are directly connected to the HGW. Such topology is shown in figure 1.

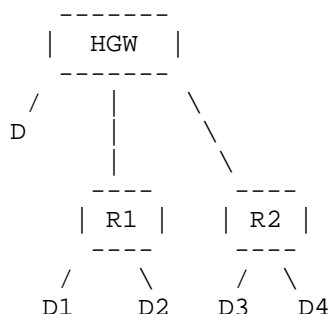


Figure 1

In this topology, one home device (D) is attached to the HGW WLAN interface in addition to (R1) and (R2). Two home devices {(D1), (D2)} are connected to (R1) and a second pair {(D3), (D4)} is connected to (R2). Finally, we assume that the HGW is able to delegate prefixes to both routers, and home devices are using stateless address autoconfiguration (described in [RFC4862]), in order to generate their IPv6 addresses.

Our goal is to keep the HGW fully aware of the four IPv6 addresses configured by the set of devices {D1, D2, D3, D4} despite not being directly connected to the HGW.

Our suggested proposal is described in the following steps:

- a. when delegating prefixes to (R1) and (R2) as described in [RFC3633], the HGW issues an explicit request to get notified about IPv6 addresses that appears on each router link, i.e., IPv6 addresses configured using the corresponding delegated prefix. Such request can be sent to the requesting routers (i.e., (R1) and/or (R2)) by inserting, for example, a new IA_PD option in the DHCP (Reply) message sent by the delegating router (i.e., HGW).

- b. upon receiving a request to convey IPv6 addresses that are (auto)-configured using the delegated (and advertised) prefix, (R1) proceeds to collect and store all IPv6 addresses which pass the duplicate address detection (DAD) procedure performed on its link. In our example, (R1) should convey to HGW all IPv6 addresses that are configured by (D1) and (D2) while (R2) should convey the addresses that are configured by (D3) and (D4).
- c. (R1) sends the collected IPv6 addresses to HGW using one (or multiple) new ICMP unicast message called "ICMP Notify (ICMP_NTY)". Such message may be sent whenever a new IPv6 address is successfully tested on the link or may be used to carry a set of IPv6 addresses. Other parameter(s) that are specific to the end device may also be sent in the ICMP_NTY message, along with the device's IPv6 address(es) (e.g., MAC address).
- d. After receiving a valid ICMP_NTY message, the HGW SHOULD send an acknowledgment to the sending router. For this purpose, we introduce another ICMP message called "ICMP Notify Acknowledgment (ICMP_NTA)". It follows that ICMP_NTA message MUST be sent only by the delegating router.

Note that the pair of new ICMP messages is also used to convey IPv6 addresses to the HGW when end devices configure their IPv6 addresses using DHCPv6 mechanism (described in [RFC3315]).

In more complicated scenarios, (R1) can be directly connected to one or multiple routers on the downstream path in which case, the prefix delegation functionality is not limited to the HGW. In such case, the suggested IPv6 address notification procedure requires the requesting router to send the ICMP_NTY messages directly to the HGW. For this purpose, the HGW address is sent by the delegating router in the DHCP Reply message (e.g., using another IA_PD option).

5. New Messages Structures and Options Format

TBD

6. Security Considerations

TBD

7. IANA Considerations

TBD

8. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.
- [RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.
- [RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.

Authors' Addresses

Wassim Michel Haddad
Ericsson
300 Holger Dr
San Jose, CA 95134
US

Phone: +1 646 256 2030
Email: Wassim.Haddad@ericsson.com

Joel Halpern
Ericsson
PO Box 6049
Leesburg, VA 20178
US

Phone: +1 703 371 3043
Email: Joel.Halpern@ericsson.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 20, 2012

L. Howard
Time Warner Cable
November 17, 2011

The UP PIO Field: Finding Up in an Unmanaged Network
draft-howard-up-pio-00

Abstract

It is difficult to find a path through an unmanaged network with multiple routers. This document describes a new Prefix Information Option field which can provide information to routers to find a path.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on May 20, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF

Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Format	3
3. Use Cases	4
3.1. Default Route	4
3.2. Walled Garden	4
3.3. ULA	4
3.4. Delegated Prefix	5
4. Implementation	5
4.1. Host Behavior	5
4.2. Tie Breaking	5
4.3. Multiple Paths	5
4.4. Route Withdrawal	6
5. Examples	6
6. Evaluation	8
7. Additional Work Required	10
8. Alternatives	10
9. Security Considerations	10
10. IANA Considerations	10
11. References	11
11.1. Normative References	11
11.2. Informative References	11
Author's Address	11

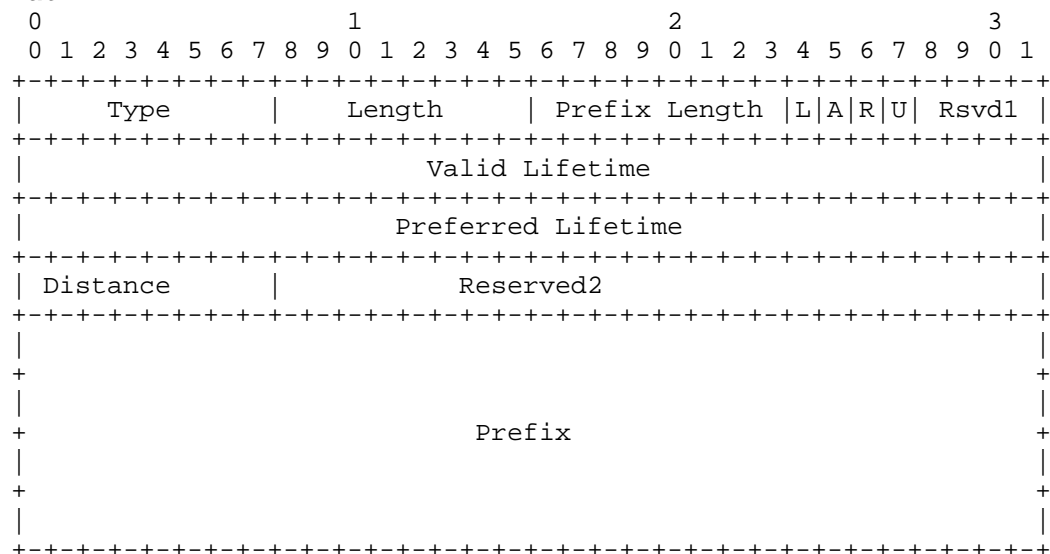
1. Introduction

This document describes a new Prefix Information Option field to be used in unmanaged networks (such as home networks) to find a path to a given prefix. This PIO field is not intended to replace dynamic routing protocols, and will not find the best path to a given destination, though it can provide useful information to routers.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119. [RFC2119].

2. Format



The use of L and A bits are specified in rfc4861. The use of the R bit is specified in rfc3775.

This format represents the following changes over those RFCs:

Upstream Prefix (U) 1-bit router address flag. When this UP bit is set to 1, indicates that the prefix was learned, rather than configured.

Rsvd1 Reduced from a 5-bit field to a 4-bit field to account for the addition of the above bit.

Distance An 8-bit metric used to determine distance and prevent loops.

The UP bit is used to find a path through an unmanaged network. When a router learns prefix information from a Router Advertisement with the UP bit set, the router SHOULD add that prefix to its own RAs. When sending RAs containing the learned prefix, it MUST increment the Distance value by one.

3. Use Cases

3.1. Default Route

The most common implementation of this field would be the advertisement of a default route, where Prefix Length = 0 and Prefix = 0. An Internet access provider could use Router Advertisements to customer gateways with the UP bit set, and a distance of 0, to indicate the border of the administrative domain and the default gateway for the customer access router. That customer access router, having learned the default gateway, SHOULD add the prefix to its routing table, then SHOULD include this information in its own RAs. When the router sends RAs including this prefix, it MUST increment the Distance in those RAs to indicate that it is one hop further than the origin of the prefix.

3.2. Walled Garden

For a network provider who does not provide a default route, the UP option can be used to indicate that a prefix is available. For instance, an operator who only wanted traffic for its hosted services on 2001:db8::/32, would send an RA for that prefix to the access gateway, with Distance=0. That gateway SHOULD add the prefix to its routing table, then SHOULD include the prefix in its own RAs, and MUST increment the Distance in those RAs to indicate that it is one hop away from the border.

3.3. ULA

Unique Local Addresses may be used for a variety of reasons [RFC6204]. When a router generates a ULA prefix, it MUST include that prefix in RAs. It SHOULD include the UP option field, with a Distance = 0. When another router learns that prefix, it SHOULD add the prefix to its routing table, then SHOULD include the prefix in its own RAs, and MUST increment the Distance in those RAs to indicate that it is one hop away from the border.

3.4. Delegated Prefix

In home network scenarios, routers are often also DHCPv6 servers. When a device is a DHCPv6-PD server, and receives a prefix to be used for host address assignments (regardless of setting of M-bit), if that device is also the router for that prefix, that router becomes authoritative for the prefix. "Authoritative for the prefix" is analogous to the notion of the "delegating router" responding to a request from the "requesting router" per [RFC3633]. In other words, when a router receives Prefix Delegation, it SHOULD include that prefix in its RAs, and SHOULD set the Distance to 0. Note that other values are possible, but reduce the possible diameter of the network.

Note that in this way, more specific routes may be propagated through the network via Router Advertisements. The longest match rule applies, and establishes the route preference. See Examples section.

4. Implementation

4.1. Host Behavior

Hosts use RAs and the PIO to find their next hop, and for address autoconfiguration. Nothing in the use of the UP bit changes these behaviors, though it is possible that hosts will learn multiple prefixes, and might have multiple paths to the same prefix. It is expected that these are harmless. Details of path selection are left to implementers.

A host MAY use the Distance metric to select a better path for a prefix. A host might learn multiple prefixes from which SLAAC may be used. This is not a new function introduced with the UP bit, and host behavior is expected to be unchanged.

4.2. Tie Breaking

When a router learns the same prefix from two other routers, and both RAs have the same Distance, a tie-breaker mechanism is required. The tie-breaker could be arbitrary, such as the time the RA was received, or it could be based on (e.g.) the Preferred Lifetime value, the layer 2 link type, or other suitable information. Implementers MUST have a tie-breaker rule or rules to resolve all ties.

4.3. Multiple Paths

A router receiving multiple RAs for the same prefix may choose to discard the path not chosen, or may add the route to its routing table with a higher Administrative Distance.

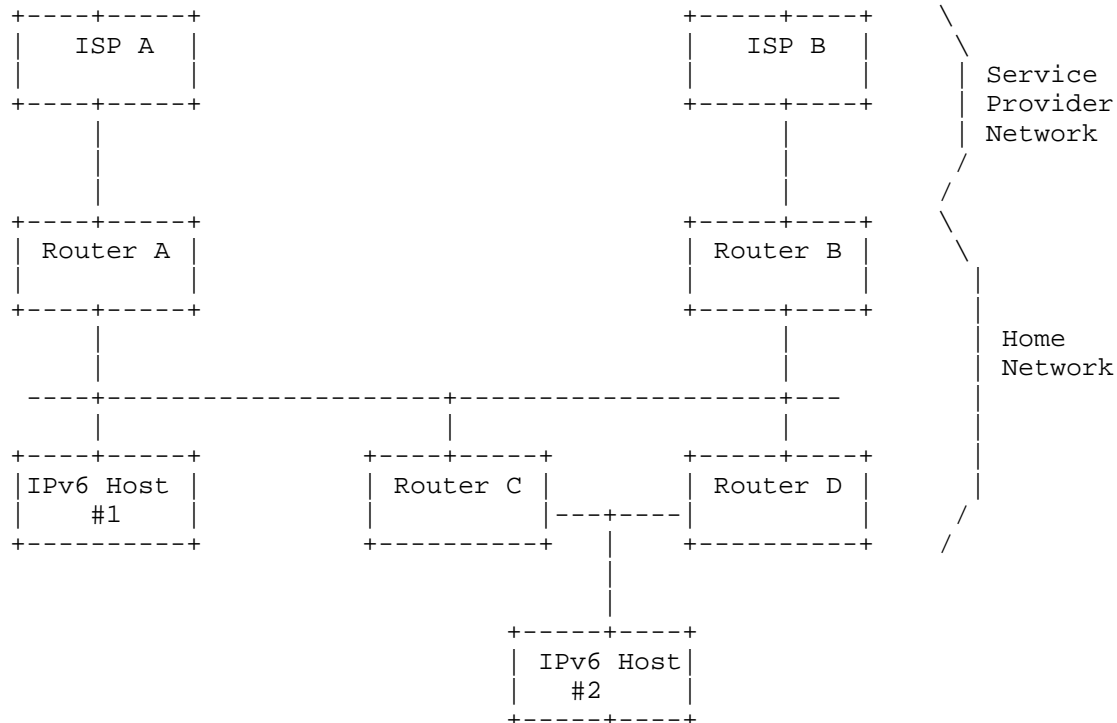
4.4. Route Withdrawal

As with other RAs, when an RA is received from a router with no information for a prefix, even if that router had previously provided prefix information, the receiver SHOULD remove references to that prefix from its routing table. Similarly, if no RAs are received from a router, the prefix SHOULD be removed. Further, it SHOULD send RAs without that prefix.

When the Valid Lifetime has passed, the route MUST be removed. When the Preferred Lifetime has passed, the route MAY be lowered in preference.

5. Examples

Consider the example in Figure 1, in which a network has two upstream ISPs, ISP A and ISP B. A different router connects to each ISP. Routers A and B connect to their respective upstream ISPs, and a Router C and Router D connect to the same link. Routers C and D share another link, causing a potential loop situation. A Host #1 connects to the shared link between Routers A, B, C, and D. A Host #2 connects to the shared link between Routers C and D.



Multi-homed Topology

Suppose ISP A and ISP B both provide a default route via Router Advertisements. Further suppose that ISP A delegates (via DHCPv6) the prefix 2001:db8:000a::/48, and ISP B delegates the prefix 2001:db8:0000:00b0::/56.

Router A sends 0::/0 with UP bit and Distance=1, and sends 2001:db8:a::/48 with UP bit and Distance=0.

Router B receives the RA. It prefers the default route from ISP B, because its Distance=0. It learns prefix 2001:db8:a::/48 and adds it to its routing table.

Router C and Router D receive the RA. They each install the default route and the /48 prefix in their routing tables.

Host #1 might configure an address from the prefix via DHCP or SLAAC.

Router B sends 0::/0 with UP bit and Distance=1, and sends 2001:db8:0:b0::/56 with UP bit and Distance=0.

Router A receives the RA. It prefers the default route from ISP A, because its Distance=0. It learns prefix 2001:db8:b0::/56 and adds it to its routing table.

Router C and Router D receive the RA. They compare the default route to the existing default route. Each applies its tie-breaking rule, and installs the winning route. Suppose they have different methods, and Router C uses the default to Router A, and Router D uses the default to Router B. They each add the /56 prefix to their routing tables.

Host #1 might configure an additional address from the prefix via DHCP or SLAAC. Whether it does is beyond the scope of this document.

Router C sends 0::/0 with UP bit and Distance=2, and sends 2001:db8:a::/48 with UP bit and Distance=2, and sends 2001:db8:0:b0::/56 with UP bit and Distance=2.

Routers A and B receive the RAs, but ignore the duplicate prefixes within them because they already have those prefixes with a shorter Distance.

Router D receives the RA for 0::/0, but already has a route with a lower Distance. Router D receives the RA for 2001:db8:a::/48, but

already has a route with a lower Distance (from Router A, where Distance=0). Router D receives the RA for 2001:db8:0:b::/56, but already has a route with a lower Distance (from Router B, where Distance=0).

Router C then receives delegated prefix 2001:db8:a:c::/64. Host #2 gets an address for this prefix, either via SLAAC or DHCP. Router C sends RAs for 2001:db8:a:c::/64 with UP bit and Distance=0. The host also receives RAs for the /64 prefix.

Router D receives the RA for 2001:db8:a:c::/64. It is a longer (more specific prefix) than its previous 2001:db8:a::/48 route, so it adds the more specific to its routing table.

Routers A and B receive the RA for 2001:db8:a:c::/64. It is a longer (more specific prefix) than their previous 2001:db8:a::/48 route, so they add the more specific to its routing table. They will update their own RAs, but Routers C and D already have routes with lower Distance.

If Router A or Router B sends the more-specific prefix to the ISP, the ISP MAY choose to add the route or ignore it if a route preferred by policy exists (for the delegated prefix).

6. Evaluation

Here follows an evaluation of whether this solution meets all of the Homenet routing requirements.

Can Host #1 reach Host #2? Yes, via a path through either router.

Is the network border clear? Yes, as Distance=0 for the shortest prefix.

Can it handle a router being added? Yes, the new router will learn prefixes via RAs quickly.

Can it handle a router being moved? Yes, with some delay in relearning the routes. If Host #2 were a Router E, and it was moved to the shared ABCD link without the router or interface going down long enough for RAs to be missed, it might propagate RAs learned from Routers C and D. It would have a higher Distance for the shorter prefixes (the /0, /48, /56), so neighboring routers would ignore its RAs. If it were two layers deeper in the network, such that it had a Distance=2 for a more-specific, which was better than the Distance previously seen by Routers A and B, neighbors would prefer its RAs. However, it would stop sending

those RAs once it stopped receiving them. The router would continue sending RAs for its delegated prefix for as long as it had that prefix; this problem is related to addressing, not routing.

Can it handle a router being removed? Yes. Some routers may install multiple routes; as soon as they miss seeing an RA for a prefix, they will have a back route. Other routers may have to wait until they see a new RA before having a backup path to the prefix.

Will it work in a no-configuration environment? Yes.

Can it support multiple upstream networks? Yes. It does not solve address configuration or source selection issues, but those are not routing problems per se.

Does it work if prefix delegation is not hierarchical? Yes. Each prefix has its own Prefix Information Option, each with its own Distance.

Can another path be found in case of failure? Yes, within a couple of RA intervals.

Does it prevent loops? Yes.

Does it allow for stable prefixes through reboots? Yes, though paths will have to be rediscovered, through the period configured for Router Advertisement transmissions.

Is it lightweight/cheap? Yes. A simple addition to the PIO, which already exists. Some additional routing decision-tree logic is required for comparing best paths, tie-breaking among equal paths, etc.

Can it handle multiple-dwelling units, or other potentially dense networks? Each upstream router will increment Distance, so each router should choose the shortest upstream path. However, a host with no router could end up choosing a neighbor's router instead of their ISP. The solution is no chattier than existing RAs.

Can it stand up to wireless networks, and networks with wired and wireless segments? Yes, with consideration for multiple-dwelling units.

Can it stand up to unintentional connections of networks? Yes. Although the Distance would seem to limit the diameter of the network to 255 segments, when each router is given a subnet

prefix, it resets the Distance to 0 for its prefix. Thus, the Diameter is as wide as prefix topology allows.

7. Additional Work Required

This option essentially overloads the PIO to be a lightweight distance-vector routing protocol of sorts. As such, it needs to avoid the problems of distance-vector protocols, such as the count-to-infinity problem. Several possibilities exist to mitigate this problem:

- Use a smaller possible infinity (i.e., change the Distance field to be a 4-bit field)

- Shorten the RA transmission intervals

- Split horizon (do not advertise a prefix out the interface it was learned) with poison reverse (when a neighbor is lost, send an RA with Distance=255) could be used

8. Alternatives

An extension of rfc4191, Default Router Preferences and More-Specific Routes, with its definition of a Route Information Option, might be a closer approximation to the distance-vector protocol described here.

A link-state protocol would solve standard problems with distance-vector protocols. However, most link-state protocols are much heavier implementations.

9. Security Considerations

By using unsecured Router Advertisements, attacks that compromise RAs would have an extended effect. Use of SeND should mitigate these attacks.

10. IANA Considerations

There are no IANA considerations or implications that arise from this document.

11. References

11.1. Normative References

- [RFC2119] "Key words for use in RFCs to Indicate Requirement Levels".
- [RFC2461] "Neighbor Discovery for IP Version 6 (IPv6)".
- [RFC3775] "Mobility Support in IPv6".
- [RFC4861] "Neighbor Discovery for IP version 6 (IPv6)".

11.2. Informative References

- [RFC3633] "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6".
- [RFC6204] Cisco Systems, Inc., Cisco Systems, Inc., CableLabs, AT&T, and Cisco Systems, Inc., "Basic Requirements for IPv6 Customer Edge Routers".

Author's Address

Lee Howard
Time Warner Cable
13820 Sunrise Valley Drive
Herndon, VA 20171
US

Phone: +1 703 345 3513
Email: lee.howard@twcable.com

Internet Draft
<draft-kitamura-ipv6-auto-name-00.txt>

H. Kitamura
NEC Corporation
S. Ata
Osaka City University
October 24, 2011

Expires April 2012

Corresponding Auto Names for IPv6 Addresses
<draft-kitamura-ipv6-auto-name-00.txt>

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on September 2011.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This document discusses notion and actual mechanisms of "Corresponding Auto Names" for IPv6 Addresses. With this mechanism, all IPv6 addresses (even if they are link-local scoped addresses) can obtain their own Names, and it will be able to use Names anywhere instead of IPv6 Addresses.

IPv6 address is too long and complicated to remember, and it is very nuisance thing to type a literal IPv6 address manually as an argument of applications. Also, it is very difficult for human beings to tell which IPv6 address is set to which actual IPv6 node. In this sense, literal IPv6 address information can be called meaningless information for human beings.

In order to solve above problems and to make the information meaningful, mechanisms called Corresponding Auto Names for IPv6 addresses is introduced. They will become gentle information for human beings. By applying a simple naming rule to the Auto Names (e.g., use the same name-prefix for IPv6 addresses that are set to the same interface (node)), this will contribute to help people to understand which IPv6 address / Name indicates which actual IPv6 node.

1. Introduction

This document discusses notion and actual mechanisms of "Corresponding Auto Names" for IPv6 Addresses.

IPv6 address is too long and complicated to remember, and it is very nuisance thing to type a literal IPv6 address manually as an argument of applications.

Furthermore, it is very normal and popular cases to set multiple IPv6 addresses to one node. One IPv6 node owns more than two IPv6 addresses (typically: one is link-local scoped address. the other is global scoped address) at least. Some IPv6 addresses (such as link-local scoped stateless auto-configuration addresses and temporary addresses) may become users' conscious-less address, because they are automatically set to the IPv6 node.

It is too difficult for human beings to tell which IPv6 address is set to which IPv6 node. In other words, when an IPv6 address is shown to a person, he almost can not tell that the shown IPv6 address indicates which IPv6 node. In this sense, literal IPv6 address information can be called useless or meaningless information for human beings.

So, there are strong desires to use Name information (that is gentle for human beings) instead of literal IPv6 Address information and to use meaningful information that can easily show which IPv6 address / name indicates which actual IPv6 node.

The Corresponding Auto Names for IPv6 Addresses is introduced to solve above problems and to satisfy the above desires.

2. Goals (What can be achieved)

In this section, goals of the mechanisms of the Corresponding Auto Names for IPv6 Addresses and what can be achieved are shown by using examples.

2.1 Assumed typical IPv6 communication environment:

Two IPv6 nodes (Node A and Node B) are located on the same link. Their IPv6 Addresses are shown below.

Node A:	Literal Address

MAC Address:	00:0d:5e:b8:80:7b

LL-Address:	fe80::20d:5eff:feb8:807b%fxp0
ULA:	fd01:2345:6789::20d:5eff:feb8:807b
	fd01:2345:6789::1234
Global Addr:	2001:DB8::20d:5eff:feb8:807b
	2001:DB8::1234
Node B:	Literal Address

MAC Address:	00:0c:76:d9:14:e3

LL-Address:	fe80::20c:76ff:fed9:14e3%em0
ULA:	fd01:2345:6789::20c:76ff:fed9:14e3
	fd01:2345:6789::5678
Global Addr:	2001:DB8::20c:76ff:fed9:14e3
	2001:DB8::5678

They own altogether 5 IPv6 addresses respectively;

- 1 Link-Local scoped Address
- 2 Unique Local Addresses (SLLAC address and manual set address)
- 2 Global scoped Addresses (SLLAC address and manual set address)

They communicate each other.

2.2 Auto Names examples

For all addresses, respective Corresponding Auto Names are prepared and registered to a name resolving DB and its service (such as the DNS) automatically by the mechanism that detects these addresses (that is explained after in this document). Prepared Auto Names are shown below.

Node A:	Literal Address	Auto Name
	-----	-----
MAC Address:	00:0d:5e:b8:80:7b	

LL-Address:	fe80::20d:5eff:feb8:807b%fxp0	-> n7bz-l1%fxp0
ULA:	fd01:2345:6789::20d:5eff:feb8:807b	-> n7bz-ul
	fd01:2345:6789::1234	-> n7bz-u2
Global Addr:	2001:DB8::20d:5eff:feb8:807b	-> n7bz-gl
	2001:DB8::1234	-> n7bz-g2
Node B:	Literal Address	Auto Name
	-----	-----
MAC Address:	00:0c:76:d9:14:e3	

LL-Address:	fe80::20c:76ff:fed9:14e3%em0	-> n3ez-l1%em0
ULA:	fd01:2345:6789::20c:76ff:fed9:14e3	-> n3ez-ul
	fd01:2345:6789::5678	-> n3ez-u2
Global Addr:	2001:DB8::20c:76ff:fed9:14e3	-> n3ez-gl
	2001:DB8::5678	-> n3ez-g2

2.3 Auto Name Prefix for Grouped Addresses

In order to make Auto Names meaningful, IPv6 addresses are grouped and Auto Name Prefix is used to show grouped addresses.

For IPv6 addresses that are set to the same interface (node), the same Auto Name-Prefix that stands for the Group ID is used for their Auto Names.

As shown above:

'n7bz-' is used for Auto Name Prefix (Group ID) for Node A and
'n3ez-' is used for Auto Name Prefix (Group ID) for Node B.

In order to make easier to identify and remember the Auto Name Prefixes, their naming rule is based on inheriting the last octet of the node's MAC address in this example.

2.4 Contribution in Regular Resolving (Name -> Address)

In order to communicate with the specific IPv6 address of the destination node, the following procedure to type literal IPv6 address is required in the current environment. They are very stressful and nuisance procedures for human beings.

When 'ping6' or 'telnet' to the specific IPv6 address of Node B from Node A is executed, the following commands are typed.

```
>ping6 fe80::20c:76ff:fed9:14e3%fxp0
>telnet fd01:2345:6789::20c:76ff:fed9:14e3
```

Especially for link-local scoped addresses or temporary addresses, there are no way to type Names instead of literal IPv6 addresses, because they are generally not registered to name resolving services.

By introducing the Corresponding Auto Names, above typed commands are changed and replaced with the following easy and rememberable name typing procedures.

```
>ping6 n3ez-11%fxp0
>telnet n3ez-ul
```

2.5 Contribution in Reverse Resolving (Address -> Name)

Communication related status information is shown to human beings in literal IPv6 address format in the current environment.

'netstat -a' (on Node A) shows connection status as followed:

Local Address	Foreign Address	(state)
fe80::20d:5eff:feb8:807b.8722	fe80:3::20c:76ff:fed9:14e3.23	ESTABLISH
fd01:2345:6789::1234.16258	fd01:2345:6789::5678.23	TIME_WAIT

'ndp -a' (on Node A) shows neighbor cache status as followed:

Neighbor	Linklayer Addr.	Netif	Expire	S
fe80::20d:5eff:feb8:807b%fxp0	0:0d:5e:b8:80:7b	fxp0	permanent	R
fd01:2345:6789::20d:5eff:feb8:807b	0:0d:5e:b8:80:7b	fxp0	permanent	R
fd01:2345:6789::1234	0:0d:5e:b8:80:7b	fxp0	permanent	R
2001:DB8::20d:5eff:feb8:807b	0:0d:5e:b8:80:7b	fxp0	permanent	R
2001:DB8::1234	0:0d:5e:b8:80:7b	fxp0	permanent	R
fe80::221:85ff:fea7:82ff%fxp0	0:21:85:a7:82:ff	fxp0	23h50m51s	S
fe80::20c:76ff:fed9:14e3%fxp0	0:0c:76:d9:14:e3	fxp0	23h51m56s	S
fd01:2345:6789::20c:76ff:fed9:14e3	0:0c:76:d9:14:e3	fxp0	23h52m50s	S
fd01:2345:6789::5678	0:0c:76:d9:14:e3	fxp0	23h53m51s	S
2001:DB8::20c:76ff:fed9:14e3	0:0c:76:d9:14:e3	fxp0	23h54m53s	S
2001:DB8::5678	0:0c:76:d9:14:e3	fxp0	23h55m54s	S

People almost can not tell which shown literal IPv6 address indicates which IPv6 node. In this sense, shown information is meaningless and useless.

By introducing the Corresponding Auto Names, above complicated information is converted into simple and meaningful information and shown as followed.

'netstat -a' (on Node A) shows connection status as followed:

Local Address	Foreign Address	(state)
n7bz-l1.8722	ne3z-l1.23	ESTABLISH
n7bz-ul.16258	ne3z-ul..23	TIME_WAIT

'ndp -a' (on Node A) shows neighbor cache status as followed:

Neighbor	Linklayer Addr.	Netif	Expire	S
n7bz-l1%fxp0	0:0d:5e:b8:80:7b	fxp0	permanent	R
n7bz-ul	0:0d:5e:b8:80:7b	fxp0	permanent	R
n7bz-u2	0:0d:5e:b8:80:7b	fxp0	permanent	R
n7bz-g1	0:0d:5e:b8:80:7b	fxp0	permanent	R
n7bz-g2	0:0d:5e:b8:80:7b	fxp0	permanent	R
nffz-l1%fxp0	0:21:85:a7:82:ff	fxp0	23h50m51s	S
n3ez-l1%fxp0	0:0c:76:d9:14:e3	fxp0	23h51m56s	S
n3ez-l1	0:0c:76:d9:14:e3	fxp0	23h52m50s	S
n3ez-l2	0:0c:76:d9:14:e3	fxp0	23h53m51s	S
n3ez-g1	0:0c:76:d9:14:e3	fxp0	23h54m53s	S
n3ez-g2	0:0c:76:d9:14:e3	fxp0	23h55m54s	S

Other examples where the Auto Name technique can contribute:

In log files of a server application, accesses from clients are recorded into them in literal IPv6 address format. It is almost impossible to read and understand the log files effectively without this Auto Name technique.

Also, in packet dumping applications, address information is shown in literal IPv6 address format. This Auto Name technique can significantly help for human beings to analyze and understand dumped packets.

Shown communication related status information in Auto Name format is simple and easy enough for human beings to understand. As shown above, troublesome IPv6 literal Address information can be converted into meaningful information by using the Corresponding Auto Names technique, and we can achieve our goals.

3 Deployed Notions and Functions that are used in Auto Names

3.1. Stateless Name

We know that we can categorize Addresses into two types. One is "stateful" address type, and the other is 'stateless' address type.

On the other, we have not been applied the same categorization to domain Names or host Names clearly. It has been assumed that existing all Names are categorized into stateful type and there is no stateless name type. Authors think that it is a time to change this preconception.

We can grasp that the introduced Corresponding Auto Name is realization of "stateless" name type, and we have deployed a notion Stateless Name clearly here.

3.2 Scoped Name

We also know that a notion called "scope" (such as link-local scope, global-scope) is introduced when we deal with addresses. Every address has its own scope.

In domain Names or host Names cases, the "scope" notion have not clearly introduced. It is assumed that all names are global information and "scope" notion does not exist.

The Corresponding Auto Name is achieved by introducing Scoped Name obviously.

Scope of Auto Name for IPv6 address is the same to the scope of its IPv6 address. For example, scope of the Auto Name for the link-local IPv6 address is link-local. They are only effective within the link-local scope.

3.3 Target IPv6 Addresses

One of the goals of the Auto Name technique is to provide and set Names to all IPv6 addresses.

If an address has its own name and it is registered into name resolving services (such as the DNS) already, it is basically not necessary to provide Auto Name to such addresses.

We can assume that IPv6 addresses whose names are registered into the name resolving services are well managed. So, they will not become targets of the Auto Name technique. However, we can provide Auto Names to such addresses, because one-to-multiple mapping is allowed in name resolving services.

4. Design of Auto Names

4.1 Conceptual Design on Naming Rules

Auto Names are composed of "<NGI>-<P><I>" format:

<NGI>: stands for Node (Interface) Group ID

4 characters (starting from 'n')
(e.g., 'n7bz', 'n3ez')

<P>: stands for Prefix of Address

1 character: (e.g., 'l', 'u', 'g')

<I>: stands for Interface ID of Address

1 character: (e.g., '1', '2')

Above discussed Auto Name examples satisfy <NGI>-<P><I> format.

on Node A: n7bz-l1, n7bz-u1, n7bz-u2, n7bz-g1, n7bz-g2
on Node B: n3ez-l1, n3ez-u1, n3ez-u2, n3ez-g1, n3ez-g2

4.1.1 <NGI> Value:

<NGI> value is also called Auto Name-Prefix.

In order to make IPv6 addresses meaningful, IPv6 addresses are grouped. It is very natural to group IPv6 addresses by which node (interface) they are set. So, IPv6 addresses that are set to the same node (interface) are grouped into the same group.

<NGI> value is shown as 'nXYZ' format:

- 'n' : (1st char) fixed and not changed
- 'XY': (2nd, 3rd chars) are inherited from
the last octet (2 characters) of the node's MAC address
- 'Z' : (4th char) suffix char to avoid a collision of 'XY'
starting from "z"
if 'XY' is collided, 'Z' is changed into "y", "x" , , ,

By using the birthday paradox theorem, collision probability of 256 states (1 octet) is calculated. If there are 19 nodes (interfaces), collision is happened with 50% probability.

Collision check procedure of the last octet of MAC addresses is necessary.

4.1.2 <P> Value:

<P> value stands for Prefix (Scope) of Address as 1 character format.

Auto Names of IPv6 addresses whose prefixes are same use the same <P> value.

Typically, following characters are used:

- "l": used for link-local scoped addresses.
- "u": used for ULA
- "g": used for global scoped address

If multiple prefixes for the same scope are used, other character (such as "h", "i", , , ,) can be used depending on the circumstances.

4.1.3 <I> Value:

<I> value stands for Interface ID of Address as 1 character format.

<I> value is starting from "1". If multiple IPv6 addresses whose <NGI> and <P> values are same are found, other <I> value (such as "2", "3", , , ,) is used.

4.2 IPv6 Address Appearance Detection and Auto Name Registration

In order to generate and register Auto Names automatically, two types of mechanisms are needed. One is a mechanism that detects IPv6 address appearance. The other is a mechanism that checks the detected addresses and generates Auto Names and registers them to name service.

Two functions ("Detector" and "Registrar") are introduced. "Detector" function takes in charge of the former mechanism, and "Registrar" takes in charge the latter mechanism.

4.2.1 IPv6 Address Appearance Detection mechanism

In order to detect newly appeared IPv6 address, DAD message (NS for DAD) is effectively used.

DAD message has the following good capabilities:

- issued only when node would like to set new IPv6 address
- issued for All types (link-local, global, temporary,...)
- L2 broadcast and easy to capture (without using mirror port)
- distinguishable from other NS messages, because source address of the message is unspecified ("::") and different from others
- Captured DAD message includes all necessary information (such as, IPv6 address and MAC address)

Detector captures DAD messages and detects newly appeared IPv6 addresses. Detected information is sent to Registrar.

4.2.2 Auto Names Generation and Registration mechanism

At first, Registrar checks the Detected address information that is sent from Detector(s). By using the reverse resolving (Address -> Name), it is checked whether the Detected address information is first appearance or not. If an entry for the address does NOT exist, it is confirmed that the address is first appearance and it should be registered to the name server.

After Name for the address is prepared, duplication of the Name can be checked by using the regular resolving (Name -> Address). If an entry for the Name exist, it is confirmed that Name is duplicated (collided). Another Name is prepared and checked again until the Name is not duplicated.

Finally, Registrar registers both Regular and Reverse resolving entries for the address and prepared Auto Name are registered to the name server.

4.2.3 Placement of Detector and Registrar

Placement of Detector and Registrar is designed to make the mechanisms flexible and to make it to be applied to various environments (office networks, home networks, etc.)

Figure 1 and 2 show typical examples that indicate locations where Detector and Registrar functions are placed on the IPv6 network. Figure 1 shows a case for a single link, and Figure 2 shows a case for multiple links.

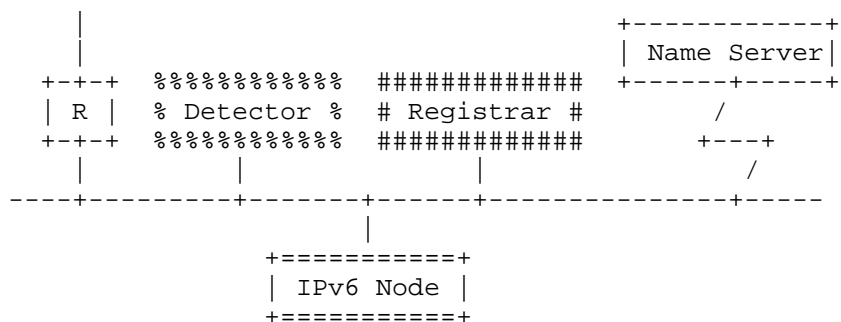


Fig. 1 Single-Link Case Example

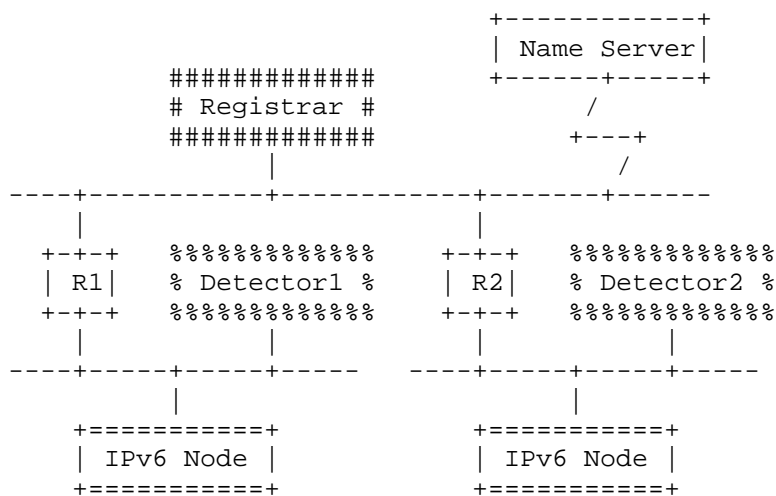
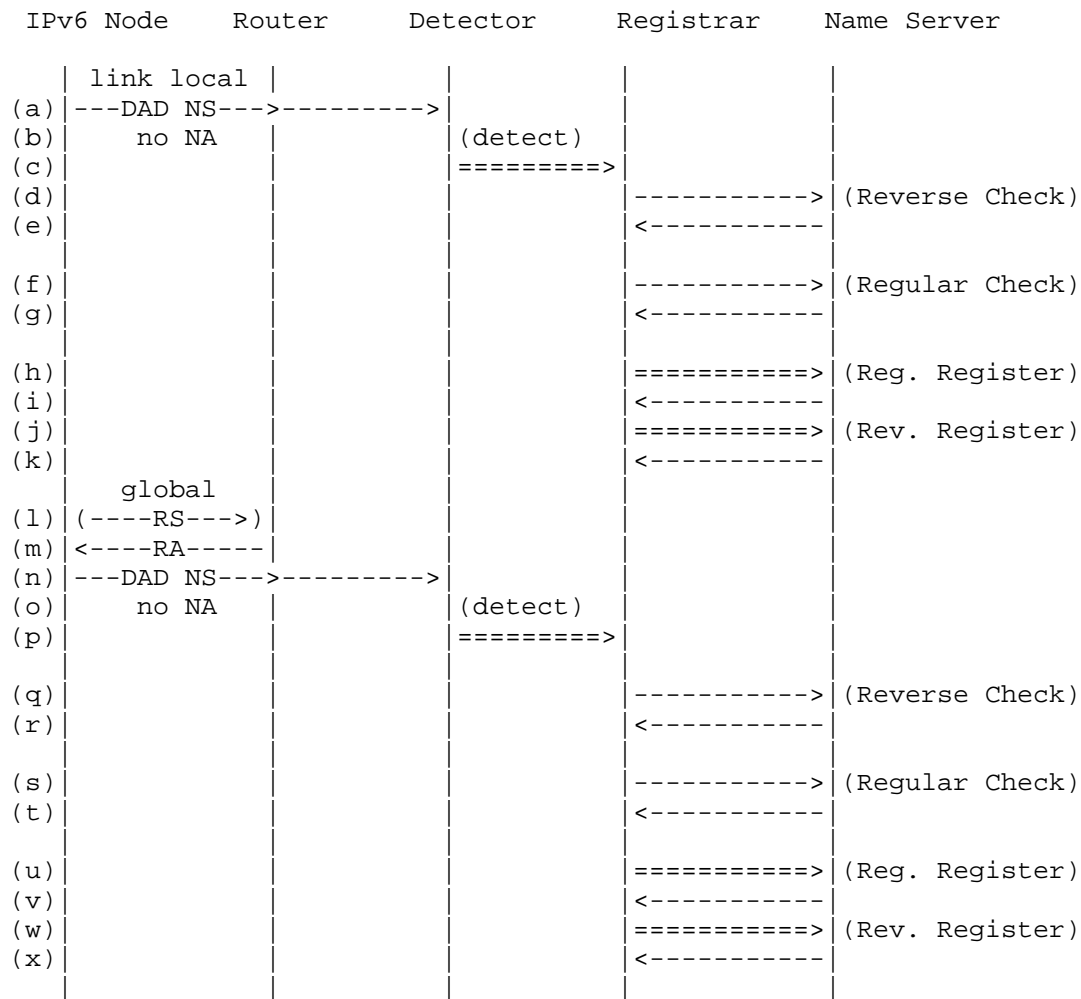


Fig. 2 Multiple-Link Case Example

4.2.4 Detection and Registration Procedures

Figure 3 shows an example of typical detection and registration procedures at IPv6 links where DAD packets are issued. DAD message packets are used for the appearance detection.



5. Security Considerations

Auto Names are generated and registered to the name service in this document. In order to register correct Auto Names information, communication between Detector and Registrar and communication between Registrar and Name Server should be protected and be secured.

In general usage, scope of Auto Names will be local (not global). Auto Names are usually local scoped names. So, we do not have to be too sensitive on the correctness of Auto Names.

6. IANA Considerations

This document does not require any resource assignments to IANA.

References

Normative References

- [RFC4291] R. Hinden and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006
- [RFC4861] T. Narten, E. Nordmark, W. Simpson and H. Soliman, "Neighbor Discovery for IP Version 6 (IPv6)," RFC 4861, September 2007
- [RFC4862] S. Thomson, T. Narten and T. Jinmei "IPv6 Stateless Address Autoconfiguration," RFC4862, September 2007
- [RFC4941] T. Narten, R. Draves and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6," RFC4941, September 2007
- [RFC1034] P. Mockapetris, "Domain names - concepts and facilities ", RFC 1034, November 1987
- [RFC1035] P. Mockapetris, "Domain names - implementation and specification", RFC 1035, November 1987
- [RFC2136] P. Vixie, S. Thomson, Y. Rekhter, and J. Bound, "Dynamic Updates in the Domain Name System," RFC 2136, April 1997
- [RFC4795] B. Aboba, D. Thaler, and L. Esibov, "Link-Local Multicast Name Resolution (LLMNR)," RFC4795, January 2007

Informative References

[RFC4620] M. Crawford and B. Haberman, "IPv6 Node Information Queries," RFC4620, August 2006

[mDNS] S. Cheshire and M. Krochmal, "Multicast DNS" <draft-cheshire-dnsext-multicastdns-14.txt> work in progress, February 2011

[RFC3849] G. Huston, A. Lord and P. Smith, "IPv6 Address Prefix Reserved for Documentation," RFC3849, July 2004

Authors' Addresses

Hiroshi Kitamura
Service Platform Research Laboratories, NEC Corporation
(SC building 12F)1753, Shimonumabe, Nakahara-Ku, Kawasaki,
Kanagawa 211-8666, JAPAN
Graduate School of Information Systems,
University of Electro-Communications
5-1 Chofugaoka 1-Chome, Chofu-shi, Tokyo 182-8585, JAPAN
Phone: +81 44 431 7686
Fax: +81 44 431 7680
Email: kitamura@da.jp.nec.com

Shingo Ata
Graduate School of Engineering, Osaka City University
3-3-138, Sugimoto, Sumiyoshi-Ku, Osaka 558-8585, JAPAN
Phone: +81 6 6605 2191
Fax: +81 6 6605 2191
Email: ata@info.eng.osaka-cu.ac.jp

Homenet
Internet-Draft
Intended status: Informational
Expires: May 3, 2012

E. Vyncke
A. Yourtchenko
M. Townsley
Cisco Systems
October 31, 2011

Advanced Security for IPv6 CPE
draft-vyncke-advanced-ipv6-security-03.txt

Abstract

This document describes how an IPv6 residential Customer Premise Equipment (CPE) can leverage modern security techniques to have strong security, while retaining as much of the end-to-end reachability of IPv6 as possible.

It is a re-submission in the framework of the HOMENET working group. The reputation part of this document should leverage the work done in the REPUTE working group of the Application are.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Threats	3
3. Overview	4
3.1. Rules for Security Policy	5
3.2. Security Analysis	6
4. IANA Considerations	7
5. Security Considerations	7
6. Acknowledgements	8
7. References	8
7.1. Normative References	8
7.2. Informative References	8
Authors' Addresses	8

1. Introduction

Internet access in residential IPv4 deployments generally consist of a single IPv4 address provided by the service provider for each home. Residential CPE then translates the single address into multiple private addresses allowing more than one device in the home, but at the cost of losing end-to-end reachability. IPv6 allows all devices to have a unique, global, IP address, restoring end-to-end reachability directly between any device. Such reachability is very powerful for ubiquitous global connectivity, and is often heralded as one of the significant advantages to IPv6 over IPv4. Despite this, concern about exposure to inbound packets from the IPv6 Internet (which would otherwise be dropped by the address translation function if they had been sent from the IPv4 Internet) remain. This document describes firewall functionality for an IPv6 CPE which departs from the "simple security" model described in [RFC6092]. The intention is to provide an example of a security model which allows most traffic, including incoming unsolicited packets and connections, to traverse the CPE unless the CPE identifies the traffic as potentially harmful based on a set of signatures (and other correlation data and heuristics) that are kept up to date on a regular basis. The computational resources necessary to support some, not all, functionalities of this model are likely more intensive than those described in [RFC6092], but are easily within the realm of what is commonly available in 2011 on medium to high-end network based firewall systems for small and medium businesses, or host-based commercial firewalls that run on laptop and desktop PCs. This set of techniques is also known as Universal Threat Mitigation (UTM).

2. Threats

For a typical residential network connected to the Internet over a broadband connection, the threats can be classified into:

- o denial of service by packet flooding: overwhelming either the access bandwidth or the bandwidth of a slower link in the residential network (like a slow home automation network) or the CPU power of a slow IPv6 host (like networked thermostat or any other sensor type nodes)
- o denial of service by service requests: like sending print jobs from the Internet to an ink jet printer until the ink cartridge is empty or like filing some file server with junk data
- o unauthorized use of services: like accessing a webcam or a file server which are open to anonymous access within the residential network but should not be accessed freely and anonymously from

outside of the home network

- o exploiting a vulnerability in the host in order to get access to data or to execute some arbitrary code in the attacked host. Exploitation can be further divided in two classes:
 1. day-0 attack when this attack has never been seen before (hence nothing can really detect it) and
 2. day+n attack where this attack is known and can be detected by the use of an attack signature
- o trojanized host (belonging to a Botnet) can communicate via a covert channel to its master and launch attacks to Internet targets.

3. Overview

The basic goal is to provide an adaptive security policy which aims to block known harmful traffic and allow the rest, restoring as much of end-to-end communication as possible. In addition, new protocols may evolve and be deployed over time; only if they become a threat vector does the CPE firewall receive a signature update (including dynamic correlation data) to classify and block them. This is in direct contrast to [RFC6092], which requires built-in knowledge of a number of protocols, or requires Internet communication to be limited to a handful of protocols that the CPE understands how to process.

- o Intrusion Prevention System (IPS) is a signature-based technology which inspects a pre-defined set of protocols at all layers (from layer-3 to layer-7) and uses a vast set of heuristics to detect attacks within one or several flow. Upon detection, the flow is terminated and an event is logged for further optional auditing. As exploits are added every day, the signature database must be updated daily and is usually quite large (more than 100 MB). This requires both large local storage (large flash or even a hard disk) and a subscription to an update service.
- o Reputation database is a centralized database which gives a reputation score to any IPv6 address (or prefix). The score varies from untrusted to trusted. Untrusted IPv6 addresses are typically addresses of a well-known attacker or from a Botnet member or from an ISP with a poor track of security... Protocols exist to dynamically request a reputation (based on DNS or HTTP). This usually requires a subscription. Note: in IPv6 the reputation database concept is still in its infancy, for example, little experience exists on the scope of the reputation: a host

/128, a LAN prefix /64 or a delegated prefix size of /56 or /48...

- o Local correlation uses another set of heuristics (like TCP distribution of Initial Sequence Number or used TCP ports or protocol handshake banners) to assert the variety of local hosts (namely operating system (OS) version and set of application) and raise or decrease the importance of a specific attack signature. For example, if the OS of host A is OS-A, then there is no point to inspect traffic to or from host A for attacks which are only relevant to OS-B.
- o Global correlation leverage all IPS distributed on the Internet to build the reputation database as well as changing the relevance of an IPS signature (for example, a propagating worm will trigger a lot of identical signatures on several IPS, this should raise the relevance of a specific signature up to the point of blocking all inbound/outbound connections on a specific layer-4 port).

The above techniques are common in the large network where budget is enough to buy firewalls, IPS and subscribe to signature or reputation source. The authors of this document believes that competition and Moore's law will make the set of those techniques (commonly referred to as 'Universal Threat Mitigation') affordable for consumer space.

3.1. Rules for Security Policy

These are an example set of rules to be applied. Each would normally be configurable, either by the user directly or on behalf of the user by a subscription service. The default preferred state hasn't been listed, though it is expected that all rules would be on by default.

If we named all hosts on the residential side of the CPE as 'inside' and all hosts on the Internet as 'outside', then the behavior of the CPE is described by a small set of rules:

1. Rule RejectBogon: apply unicast reverse path forwarding (RPF) checks (anti-spoofing) for all inbound and outbound traffic (implicitly blocking link-local and ULA in the same shot)
2. Rule BlockBadReputation: block all inbound and outbound packets whose outside IPv6 address has a bad reputation score
3. Rule AllowReturn: inspect all outbound traffic and allow the return traffic matching the states (5-tuple + TCP sequence number or any layer-4 state), apply IPS on the outbound (to block Botnet) and inbound (to block malicious/cracked servers which could inject malware) with IPS. If the protocol is not supported/recognized by the IPS, accept it anyway.

4. Rule AllowToPublicDnsHost: allow all inbound traffic to any inside address which is listed in the public DNS with a AAAA record (this requires that the CPE/RG can do a zone transfer, i.e., that the CPE/RG appears like a secondary name server), all inbound traffic is also inspected with IPS. If the protocol is not supported/recognized by the IPS, accept it anyway.
5. Rule ProtectLocalOnly: block all inbound traffic to any inside address as long as the inside address has never sent a packet to the outside. The intent is to protect local-only devices like thermostat or printers. Most (if not all) hosts expecting inbound connections have to send a couple of outbound packets to the outside (registration, DNS request, ...). This is the usual IPv4 firewall behavior augmented with IPS and reputation
6. Rule CryptoIntercept: at the exception of IPsec, all inbound connections that are encrypted (notably TLS [RFC5246]) must be intercepted (this is terminated by the CPE that will present its own self-signed certificate to the remote party which should have installed the CPE self-signed certificate in a secure way in its trust anchors store) in order to allow for further inspection. The decrypted flow is then passed again through those rules and encrypted again before being forwarded to the local host. This is actually a Man-in-the-Middle attack done for a good reason: protect the naive residential user. Of course, documentation and GUI MUST be provided to educate the user and help him/her to understand how to do it in a secure way. Note: this technique is also used nowadays by large enterprise web proxies with the self-signed certificate being securely distributed to all clients.
7. Rule ParanoidOpeness: allow all unsolicited inbound connections rate limited to protect against port and address scanning attacks or overloading devices or slow links within the home. The connection MUST be inspected by the IPS engine. If the connection is anonymous or using a default password (like connecting to a webcam as a guest), then the flow SHOULD be dropped. If the IPS detects an attack, then the flow MUST be closed. If the protocol is not recognized as supported by the IPS, the flow MAY be allowed.

3.2. Security Analysis

This proposal of 'paranoid openness' stops the following attacks:

- o unauthorized use of services/denial of service: because all anonymous access to inside servers are blocked.

- o Denial of services on low bandwidth or low CPU inside hosts IFF those hosts never access the Internet
- o Exploiting of a day+1 attack, those attacks are blocked with the IPS signature and address reputation database

The CryptoIntercept part can also be leveraged as a small Certification Authority (CA) that could generate RSA key pairs and X.509 certificates at the CPE/RG owner's request. Those key pairs and certificates can then be given to trusted devices or users (like the owner's laptop so that he/she could easily and safely connect from the outside).

This proposal cannot help with the following attacks:

- o flooding the access link to the Internet, this is exactly the same as with the old layers-3/4 firewall approach as only the ISP can effectively stop the flooding of the CE-PE link;
- o weak password on inside services, of course the IPS component will detect multiple failed attempts (dictionary attack) and report the offender to the Global Correlation system;
- o exploiting of day-0 attack: until now, these day-0 attacks are caused either by rapidly propagating worms (then the global correlation of unusual traffic pattern will raise an alert and block the traffic after a couple of hundred's of successful attacks) or by targeted attacks against high-profile targets (like Government or banks or ;..) which should be protected by conventional less open security policies;
- o exploiting a vulnerability in a rare or new protocol (not yet supported by the IPS), this case will probably never occur on a wide scale in a residential use of Internet.

4. IANA Considerations

There are no extra IANA consideration for this document.

5. Security Considerations

All security considerations have been done in the Security Analysis Section 3.2.

It is also advisable that the inbound rate limiter system could be added to the [RFC6092] as it is light and does not depend on a

centralized policy server.

6. Acknowledgements

Many thanks to Ole Troan, Stuart Cheshire, Dave Oran and Eliot Lear for the review of the -00 version and to Ron Bonica, Sam Hartmans, Lee Howard, Greg Lebovitz, Jordi Palet, Tina Tsou and others for their comments during and after the first presentation at the Hiroshima IETF meeting in November 2009.

A previous IETF work has similar ideas
[I-D.palet-v6ops-ipv6security].

7. References

7.1. Normative References

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.

7.2. Informative References

- [I-D.palet-v6ops-ipv6security]
Palet, J., Vives, A., Martinez, G., and A. Gomez, "IPv6 distributed security requirements",
draft-palet-v6ops-ipv6security-02 (work in progress),
February 2005.
- [RFC2993] Hain, T., "Architectural Implications of NAT", RFC 2993,
November 2000.
- [RFC6092] Woodyatt, J., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092,
January 2011.

Authors' Addresses

Eric Vyncke
Cisco Systems
De Kleetlaan 6a
Diegem 1831
Belgium

Phone: +32 2 778 4677
Email: evyncke@cisco.com

Andrew Yourtchenko
Cisco Systems
De Kleetlaan 6a
Diegem 1831
Belgium

Phone: +32 2 704 5494
Email: ayourtch@cisco.com

Mark Townsley
Cisco Systems
11, Rue Camille Desmoulins
Issy Les Moulineaux 92782
France

Phone: +33 15 804 3483
Email: townsley@cisco.com

