Abstract

This memo focuses on the evolving networking technology within and among relatively small "residential home" networks. The goal of this memo is to define the architecture for IPv6-based home networking that supports the demands placed on it. This architecture shows how standard IPv6 mechanisms and addressing can be employed in home networking, and outlines the need for specific protocol extensions for certain additional functionality.

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 January 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. Effects of IPv6 on Home Networking .................. 3
3. Architecture ......................................... 5
   3.1. Requirements ..................................... 8
   3.2. Principles ....................................... 9
   3.3. Implementing the Architecture on IPv6 .......... 10
4. References ........................................... 11
   4.1. Normative References ............................. 11
   4.2. Informative References ........................... 11
Appendix A. Acknowledgments ............................ 12
Authors’ Addresses ..................................... 12
1. Introduction

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

While many advanced home networks have been built, most operate based on IPv4, employ solutions that we would like to avoid such as network address translation (NAT), or require an 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 for how standard IPv6 mechanisms and addressing [RFC2460] [RFC4291] can be employed in home networking, while coexisting with existing IPv4 mechanisms that are widely deployed.

2. Effects of IPv6 on Home Networking

Service providers are deploying IPv6, widely accessed 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 and allows direct IP addressability and routing to devices in the home from the Internet. Following is an overview of some of the areas of that are both promising and problematic:

Multiple segments

While less complex L3-topologies involving as few subnets as possible are preferred in home networks for a variety of reasons including simpler management and service discovery, incorporation of dedicated segments remain necessary for some cases. For instance, a common feature in modern home routers in 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 sensor networks. Finally, similar needs for segmentation may
occurs in other cases, such as separating building control or corporate extensions from the Internet access network. 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-cpeenhancements], r [I-D.baker-fun-multi-router], and [I-D.baker-fun-routing-class].

In addition to routing, rather than natting, 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.

Security, Borders, and the elimination of NAT

The End-to-end communication that is promised with IPv6 is both an incredible opportunity for innovation and ease of 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. Firewalls that restrict incoming connections may be used to prevent exposure, however, this reduces the efficacy of end-to-end connectivity that IPv6 has the potential to restore.

[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. 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 an 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) vs. "Simple" security.

In addition to 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
Naming, and manual configuration of IP addresses

In IPv4, it is common practice to reach a router for configuration, DNS resolver functions, or otherwise via 192.168.1.1 or some other well-known RFC 1918 address. In IPv6, there is no such address space available, and generally IPv6 addresses are more cumbersome for humans to manually configure. As such, even for the simplest of functions, naming and the associated discovery of service is imperative for an easy to administer homenet.

3. Architecture

An architecture outlines how to construct home networks involving multiple routers and subnets. In the following this memo presents a few typical home network topology models, followed by architectural principles that govern how the various nodes should work together. Finally, some guidelines are given for realizing the architecture with the IPv6 addressing architecture, prefix delegation, global and ULA addresses, source address selection rules and other existing components of the IPv6 architecture. The architecture also drives what protocols extensions are necessary, as will be discussed in Section 3.3.

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 well understood today [RFC6204].
Figure 2 shows another network that now introduces multiple local area networks. These may be needed for reasons relating to different link layer technology or for policy reasons. Note that a common arrangement is to have different link types supported on the same router, bridged together. For the purposes of this memo and IP layer operation this arrangement is considered equivalent to the topology in Figure 1. 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.
Figure 2 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 understand the entire topology. The topology itself may also be complex, it may not be possible to assume a pure tree form, for instance.
3.1. 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 different types of network topologies discussed above. Manual configuration is rarely, if at all, possible. The equipment needs to be prepared to handle at least
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 internetworking mechanisms. It would be desirable to 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 authors believe that the following principles guide us in designing these networks in the correct manner:

**Largest Possible Subnets**

As part of the self-organization of the network, the network should subdivide itself to the largest possible subnets that can be constructed with 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.

**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 much easier to block certain nodes from communicating than it is to re-enable nodes to communicate if they have been hidden behind local addressing domains and address translation.
IP Connectivity between All Nodes

A logical consequence of the end-to-end communications model is that the network should 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.

Self-Organization

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

Least 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.

Discovery

The most natural way to think about name 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.

Intelligent Policy

As the Internet continues to evolve, no part of the architecture or security design should depend on hardcoding acceptable or unacceptable traffic patterns into the devices. Rather, these traffic patterns should be driven off up-to-date databases in the Internet.

3.3. Implementing the Architecture on IPv6

The necessary mechanisms are largely already part of the IPv6 protocol set and common implementations. The few known counter-examples are discussed in the following. For prefix configuration, existing protocols are likely sufficient, but may at worst may need some small enhancements, such as new options. 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. For name resolution and service discovery, extensions to existing multicast-based name resolution protocols are needed to enable them to work across subnets.

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


4.2. Informative References


[I-D.herbst-v6ops-cpeenhancements] Herbst, T. and D. Sturek, "CPE Considerations in IPv6 Deployments", draft-herbst-v6ops-cpeenhancements-00 (work
Appendix A. Acknowledgments

The authors would like to thank to Stuart Cheshire, James Woodyatt, Ole Troan, Lars Eggert, Ray Bellis, David Harrington, Wassim Haddad, Heather Kirksey, Dave Thaler, Fred Baker, and Ralph Droms for interesting discussions in this problem space.

Authors’ Addresses

Jari Arkko
Ericsson
Jorvas 02420
Finland

Email: jari.arkko@piuha.net

Mark Townsley
Cisco
Paris 75006
France

Email: townsley@cisco.com
Exploring the multi-router SOHO network
draft-baker-fun-multi-router-00

Abstract

This note explores the ramifications of a multi-router or multihomed small network, such as a residential or SOHO network.

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].

For clarity, in this document the word "may" is distinguished from "MAY". Consistent with [RFC2119], "MAY" refers to permission - something MAY or MAY NOT be done within a context. The word "may" refers to possibility; it is possible and correct for something to happen.

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 January 2, 2012.

Copyright Notice

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

1. Introduction ............................................. 3
   1.1. Definitions ........................................... 3
   1.2. Use Cases ........................................... 4
2. Issues .................................................... 7
   2.1. Routing in a small network .......................... 7
   2.2. Assigning Subnet Numbers ........................... 8
3. Possible Requirements ................................. 9
   3.1. Source/Destination Routing .......................... 9
   3.2. Subnet assignment .................................. 9
   3.3. Recommended upstream route ......................... 10
4. IANA Considerations ..................................... 10
5. Security Considerations ............................... 10
   5.1. Privacy Considerations ............................ 10
6. Acknowledgements ....................................... 11
7. Change Log .............................................. 11
8. References ............................................... 11
   8.1. Normative References ............................... 11
   8.2. Informative References ............................. 11
Author’s Address ....................................... 12
1. Introduction

This note explores the ramifications of a multi-router or multihomed small network, such as a residential or SOHO network. It has relevance to the IETF CPE Router discussion in the IPv6 Operations Working Group, numbering and renumbering in the "renum" effort, and scoping of the "fun" effort.

Much of the commentary in this draft applies equally to IPv4 [RFC0791] or IPv6 [RFC2460]. As such, the protocol will simply be called "IP" unless there is a reason to distinguish. References will be IPv6-related unless there is a reason for an IPv4-specific reference.

1.1. Definitions

I don’t think I’m introducing new terms, but let me state the definitions of the terms I’m using for clarity.

LAN: A Local Area Network (LAN) is a network of link layer interfaces that can directly communicate without the use of a router. Example of LANs include IEEE 802.3 Ethernet, IEEE 802.11 WiFi, and IEEE 802.15.1 and IEEE 802.15.4 WPANs.

Subnet: A Subnet is a set of IP interfaces connected by a LAN. It by definition has a prefix, which serves as a locator in routing. Multiple subnets may use the same LAN, and the membership of those subnets may or may not be congruent.

LAN Switch: A LAN switch connects different physical media (wired or wireless) in a LAN, switching packets to make them appear to be a single LAN from the perspective of the systems attached to it.

Host: With respect to a given subnet, a host is a system that has an address in it and may originate traffic using that address, but does not switch packets between it and other subnets. See [RFC2460].

Router: With respect to a given subnet, a router is a system that has an address in it, may originate traffic using that address, and additionally switches packets between it and other subnets. See [RFC2460].

CPE Router: The Customer Premises Equipment (CPE) Router is the router on the customer premises that connects it to another network. The term is often used in the context of a Managed Service, which is a service in which an upstream network own and operates the router. In residential broadband networks, it is
more common for the customer to own the router. For the purposes of this document, the term "CPE" simply means that it is on the customer’s premises.

1.2. Use Cases

The Basic Requirements for IPv6 Customer Edge Routers [RFC6204] postulate a very simple network: a single router connecting a single subnet to a single upstream ISP. In general, one would expect that router to also implement a simple firewall implementing the Recommended Simple Security Capabilities [RFC6092].

However, it is common, and in the future perhaps normal, for residential and SOHO networks to be more complex, with separate domains for

- domain-wide wired and/or wireless LANs with IP subnets,
- offices with differing corporate security requirements for the residents,
- networks with different PHY/MAC layers supporting the Smart Grid HAN or medical telemetry, and
- networks for entertainment such as home-wide audio systems or high definition TV.

In addition, there is evidence that individual residences are likely to be multihomed, in the sense of having multiple upstream networks. There are at least three obvious cases:

- One obvious case is a home using traditional broadband (DSL or Cable Modem) and additionally using 3GPP or LTE as an upstream.

- Another is a home equipped with a Smart Grid Energy Services Interface (ESI); such a home could be assigned a prefix by the utility for use in communications through the Advanced Metering Infrastructure (AMI). This is not an "ISP" in the usual sense of the term, as it provides limited services and only communicates to the relevant utility. It is, however, an "upstream" network in the sense that it allocates a prefix to the home and provides services for hire.

- A third, which has been proposed in Japan, is a Content Delivery Network (CDN) that delivers entertainment via IP, and allocates a prefix to the home for communication with it. Again, this is not an "ISP" in the usual sense of the term, as it provides limited services and only communicates to the CDN servers. It is,
however, an "upstream" network in the sense that it allocates a prefix to the home and provides services for hire.

As such services are deployed, it is reasonable to expect that the typical residence or SOHO will be multihomed.

If one takes [RFC6204]’s view of such a network, one gets a picture something like Figure 1 (in that picture, consider "ISP" to be a generalized upstream network, not specifically one that delivers Internet access as a service). From the perspective of some, this would be a wireless LAN, or at most a wired LAN and a wireless LAN bridged together.

```
---.       +--------+     | HAN Sensors, ESI,.
 \
 ISP1 )-----+        |     | Medical Sensors, .
 /---'        | CPE     | Office Equipment
---'         +-----+    A/V Equipment
 /---.       |        |     | home-wide
 ISP2 )-----+        |     |
 /---'       +--------+     |
 | Router +-----+
 /---.       |        |     |
 ISP3 )-----+        |     |
 /---'       +--------+     |
 |          +-----+
 |          ++-----+   home-wide LAN
 |          +++-----+
 | HAN Sensors, ESI,. |
```

Figure 1: Single router multihomed residence

The model in Figure 1 has some obvious problems, however. Cisco Systems, for example, requires that telecommuter’s offices have a security guard (firewall or separate Internet access) between the office and the home. There are known cases in which husband and wife or roommates each work for different companies and each company has such an information security guideline. In addition, especially with a single SSID 802.11 implementation, one could readily imagine HD TV crowding out other uses, or BitTorrent crowding out the A/V uses. Separating the uses into separate LANs for manageability and service isolation, and especially with the Smart Grid’s Home Area Network having a different physical interface (IEEE 802.15.4 being a prominent option), such a network begins to look like Figure 2.
Figure 2: Single router multihomed SOHO

Modulo the 802.15.4 interface, such routers exist today, providing multiple 802.3 and 802.11 LANs and routing between their subnets. However, such a router begins to look like the IT counterpart to a Swiss Army knife, and networks are constrained by their interface count and type. An alternative would be to build the network out of two-port routers, as in Figure 3.
Figure 3: Multiple router multihomed SOHO

Reality is probably somewhere in the middle - some multiport routers and some two-port routers, depending on the application.

2. Issues

Three obvious questions arise in such networks:

- How does routing, including routing between interior routers and to exit routers, actually work? What features are needed?
- How does the network identify and number its subnets?

2.1. Routing in a small network

A brief analysis of the advertised features of commercial residential routers - products designed for use by the uninitiated in their homes - found that almost without exception, they support RIP Version 2 [RFC2453]. At least one was found that supports RIP Version 1 [RFC1058], and one that supports OSPF Version 2 [RFC2328]. By analogy, it seems rational to expect residential and SOHO routers for IPv6 to support RIPng [RFC2080], and possibly OSPF [RFC5340].

The issues in distance vector routing, which are discussed in some detail in [RFC1058], primarily relate to bogus information that has...
not been removed from the routing system, especially during a "count to infinity" event. Such events happen in networks that have parallel connectivity, which is usually implemented for robustness. The network in Figure 3 does not have parallel paths, and so would be unlikely to have that issue. More generally, an outage in a small network would likely result in the network administrator resetting the router in question. So RIPng should be adequate for the purpose.

Another issue that arises, however, has to do with upstream Ingress Filtering [RFC2827]. In a network with a router per upstream network, one would really like to direct traffic intended for a specific upstream network to the correct router. If hosts select the correct source address using [I-D.ietf-6man-rfc3484-revise], [RFC3704] addresses that in part by suggesting that such routers redirect traffic to each other; a better approach would be to have a routing protocol that looks at {source, destination} address pairs and routes traffic to the appropriate exit.

2.2. Assigning Subnet Numbers

In order for an upstream network such as an ISP or utility to assign a prefix to a small network, the CPE router must support DHCPv6 [RFC3315] and its IPv6 Prefix Options [RFC3633]. This enables the CPE Router to obtain a prefix from its upstream network, be it an ISP, a content delivery service, or a utility, and begin to use it.

If, for example, an ISP allocated a global prefix to the CPE, one would expect the CPE to allocate a default unicast route (in IPv6, a route to 2000::/3, which is to say "all unicast addresses", as opposed to ::/0, which would include link-local addresses, ULAs, and multicast traffic) toward the ISP. In the more limited cases of a CDN or utility, it may be appropriate for the upstream prefix to be more limited — it might recommend an upstream route to exactly the CDN service, or the address of a single anycast server/service.

Within the small network, one would also hope for a way to assign subnet numbers; as others have suggested, this could build on the same capability as in [RFC3633]. If the CPE router has a prefix shorter than /64, other routers within the domain could ask it for /64s and have them assigned by the same mechanism. A hand-wave description would have any small-network router

1. Come up as a host on each interface, emitting an RS and awaiting an RA from another router. On any interface that calculates its address using SLAAC, one would expect the router to use the same prefix(es) that it learned.
2. If no address is allocated on an interface via SLAAC within some interval, perhaps sixty microfortnights, the router could request a prefix using the mechanisms in [RFC3315] and [RFC3633]. For the ISP-facing CPE Router, this would result in a prefix shorter than /64; within the domain, it should result in a /64. In the ISP-facing case, one would expect the router to allocate a /64 to each of its non-ISP-facing interfaces and immediately emitting an RA; routers in those subnets would then create addresses using SLAAC.

3. If an additional interval of (perhaps) sixty microfortnights elapses, and the router has an IPv6 address on one or more of its interfaces, one could imagine the router requesting a new /64 on one of its addressed interfaces and assigning it to an un-addressed interface.

This procedure has an obvious race condition: if there are two routers on the same LAN, they could both request a prefix and simultaneously apply it. While not incorrect (IPv6 allows for multiple subnets on a LAN), it is inefficient. Two obvious mechanisms exist to counter this. If an SPF-based protocol such as OSPF or IS-IS is in use, and only the designated router requests a prefix, there will be a minimum number of subnets on the LAN. Alternatively, if the DHCPv6 server allocates prefixes with some non-trivial inter-assignment interval, the LANs should similarly have a minimum number of subnets.

3. Possible Requirements

As the document is written, various possible requirements have popped up. These include at least the following.

3.1. Source/Destination Routing

Section 2.1 notes that it would be nice to have a routing protocol that steered traffic toward an appropriate exit. Stated generally, it would be nice to have a routing protocol that could generate routes *from* a source *to* a destination, as opposed to being simply *to* a destination.

3.2. Subnet assignment

A subnet assignment procedure such as described in Section 2.2 is needed. That section shows a "hand-wavy" mechanism, but the mechanism needs to be worked out in detail.
3.3. Recommended upstream route

Section 2.2 notes that it would be nice to have a way for the upstream network that provides a prefix to a customer also be able to give it a recommended upstream route. One obvious solution would be a DHCPv6 option that indicated some number of tuples, each consisting of

- A source prefix (which could be ::/0, the prefix assigned to the small network, or something else)
- A destination prefix (which could be ::/0, 2000::/3, or a more specific appropriate to the service such as an anycast address or a data center prefix for a specific service offered by the ISP)
- A set of DSCP values, which could be "any" or any more specific subset
- One or more next hop router addresses

If source/destination routing is implemented as described in Section 3.1, it might want to be able to specify that such datagrams must come *from* the prefix it assigned to the network. This could be implemented using a routing protocol, but that is a big change to the way residential broadband networks usually work; a more acceptable approach may be a DHCPv6 option.

4. IANA Considerations

This memo asks the IANA for no new parameters.

Note to RFC Editor: This section will have served its purpose if it correctly tells IANA that no new assignments or registries are required, or if those assignments or registries are created during the RFC publication process. From the author's perspective, it may therefore be removed upon publication as an RFC at the RFC Editor's discretion.

5. Security Considerations

5.1. Privacy Considerations
6. Acknowledgements

7. Change Log

Initial Version: 17 June 2011

8. References

8.1. Normative References


8.2. Informative References


Internet-Draft   Exploring the multi-router SOHO network       July 2011

[ RFC 3633 ] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic

[ RFC 3704 ] Baker, F. and P. Savola, "Ingress Filtering for Multihomed

[ RFC  5340 ] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF

[ RFC 6092 ] Woodyatt, J., "Recommended Simple Security Capabilities in
Customer Premises Equipment (CPE) for Providing

[ RFC 6204 ] Singh, H., Beebee, W., Donley, C., Stark, B., and O.
Troan, "Basic Requirements for IPv6 Customer Edge
Routers", RFC 6204, April 2011.

Author's Address

Fred Baker
Cisco Systems
Santa Barbara, California  93117
USA

Email: fred@cisco.com
Routing a Traffic Class
draft-baker-fun-routing-class-00

Abstract

This note addresses the concept of routing a traffic class. This has many possible implementations, IGP and BGP, and link state as well as distance vector. The fundamental impetus is the question raised in RFC 3704 and shim6 of exit routing, the question raised by Mike O’Dell of source/destination routing, and the "fish" problem, raised in many networks, in which distinct traffic classes that could conceivably use the same route predictably use different routes. Instead of handling these as "destination routing with a twist", the paper looks at the matter systemically.

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 January 2, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.
1. Introduction .............................................. 3
  1.1. Scope ............................................... 3
  1.2. Structure of the paper .............................. 3
2. The fundamental concept of routing a class ............ 3
  2.1. Define: "traffic class" ............................ 4
  2.2. Define: "metric" .................................. 5
  2.3. Define: "route announcement" ....................... 5
  2.4. Route Precedence ................................ 7
3. Sketch of a distance vector implementation ........... 8
4. Sketch of a link state implementation .................. 10
5. IANA Considerations ..................................... 12
6. Security Considerations ................................ 12
7. Acknowledgements ....................................... 12
8. Change Log ............................................. 13
9. References .............................................. 13
  9.1. Normative References ............................... 13
  9.2. Informative References ............................ 13
Author's Address ........................................... 14
1. Introduction

This note addresses the concept of routing a traffic class. This has many possible implementations, IGP and BGP, and link state as well as distance vector. The fundamental impetus is the question raised in [RFC3704] and shim6 of exit routing, the question raised by Mike O’Dell of source/destination routing, and the "fish" problem, raised in many networks, in which distinct traffic classes that could conceivably use the same route predictably use different routes. Instead of handling these as "destination routing with a twist", the paper looks at the matter systematically.

1.1. Scope

The question of implementation of IPv4 or IPv6 routes is moot; the algorithms discussed here can be used for either. Examples, however, will be drawn from IPv6 [RFC2460] and will presume its addressing architecture [RFC4291].

1.2. Structure of the paper

The paper looks first at the fundamental concept in Section 2. It then goes on to imagine an IS-IS-like [RFC5308] [RFC6119] implementation. Unfortunately, this really can’t be implemented in IS-IS by adding a TLV, which is our usual approach, due to the changed concept of a metric. However, given the changed concepts of a metric and a TLV, it shows how the same exchange and calculation algorithms can be used to build such a routing protocol. It also looks at a RIP-like [RFC2080] algorithm that includes a sequence number (derived from AODV [RFC3561]) to simplify count-to-infinity-related problems. It stops short of a BGP implementation, although one modeled on the distance vector model would make sense.

2. The fundamental concept of routing a class

This section introduces the fundamental concepts involved in routing a traffic class. These include the definitions of

- a "traffic class", which is a set-theoretic definition - all traffic matching a constraint,
- a "metric", which is an administrative number applied to a class of traffic crossing an interface, and
- a "route announcement", which is the accumulation of information regarding the routing of a traffic class as modified by various metrics en route.
In addition, it describes the fundamental algorithm applied in modifying a route announced by a predecessor node using a metric for announcement on the interface implied.

2.1. Define: "traffic class"

A "class" of traffic, in any routing protocol, is the selector that is used to identify a traffic stream for the purpose of routing. For traditional internet routing protocols like RIP, OSPF, IS-IS, EIGRP, or BGP, a "traffic class" is "the traffic destined to a stated prefix". In the context of this design, a traffic class is the traffic that goes

- to a destination prefix, which may be a default route (::/0), a host route (any /128 prefix), or anything in between,
- from a source prefix, which may be a default route (::/0), a host route (any /128 prefix), or anything in between, and
- using one of a set of DSCP [RFC2474] values.

The set of DSCP values includes and "any DSCP". "Any DSCP" has the obvious meaning: we are not really looking at the DSCP in traffic classification.

A protocol could also identify a route as a "null route"; this is like any other route, but specifically directs that traffic matching the traffic class is to be dropped.

A traffic class is represented in this document as

```
{destination, source, {list of DSCPs}|any|none}
```

Examples of common traffic classes include:

- Standard Default Route: {::/0, ::/0, any}
- Default Route using a source prefix {::/0, source, any}
- Default route used only for a QoS Traffic Class: {::/0, ::/0, {list of DSCPs}}

Examples of QoS traffic classes are found in the Configuration Guidelines for DiffServ Service Classes [RFC4594] and related documents [RFC5127] and [RFC5865].

2.2. Define: "metric"

A "metric" is an attribute of an interface, and associates a traffic class with an administrative value used in comparison of routes. In this document, it is represented as

\{(destination, source, (DSCP)|any|none), administrative value\}

and should be understood as "the metric for traffic from source to destination using DSCP".

For example, if an interface is available to all VoIP traffic, one of the metrics on an interface might be

\{::/0, ::/0, {EF}, administrative value\}

2.3. Define: "route announcement"

A route announcement is identical to a metric in structure, but is semantically different. Where a metric is an attribute of an interface, a route is an entry in the route table calculated by the routing protocol, and is used in making forwarding decisions.

The calculation of a route follows the following logic:

1. Inputs to the route calculation are a route announcement (which may be derived from an interface metric in the case of local routes, or received from a neighboring route for remote routes), and a metric associated with an interface.

2. The source, destination, and set of DSCPs of the route announcement and the metric are intersected to generate the traffic class for the resulting route.

3. The resulting route’s administrative value is the sum of the original route’s administrative value and the interface’s administrative value.

For example, in Figure 1, if the prefix allocated by ISP-1 to a network is 2001:db8:1::/48 and the prefix allocated by ISP-2 to the network is 2001:db8:2::/48, the exit routers for the network might advertise into the network the default routes:

- {{2000::/3, 2001:db8:1::/48, any}, 1} ("unicast traffic using source addresses in 2001:db8:1::/48 for any application"), and
- {{::/0, 2001:db8:2::/48, any}, 2} ("all traffic using source addresses in 2001:db8:2::/48 for any application").
One edge case is the case in which a route intersected with the available metrics yields the null set. In such a case, the resulting route cannot be used as no traffic matches it.

If an interior distance vector router in the network receives

\{(2000::/3, 2001:db8:1::/48, any), 1\}
\{(::/0, 2001:db8:2::/48, any), 2\}

from its upstream router, and on a given interface has a metric

\{(::/0, ::/0, {EF}), 5\}
("any VoIP traffic")

it would validly infer the routes

\{(2000::/3, 2001:db8:1::/48, {EF}), 6\}
("unicast VoIP traffic using source addresses in 2001:db8:1::/48")

\{(::/0, 2001:db8:2::/48, {EF}), 7\}
("all VoIP traffic using source addresses in 2001:db8:2::/48")

It would install the routes it received into its own route table, and advertise the calculated routes to neighboring routers.

Note that there is no question of a unicast RPF or other forms of Ingress Filtering [RFC2827] required in such a network. If a host spoofs a source address in another routing domain and sends a datagram upstream, there is no route in the network "from" that routing domain, and the router has no idea what to do with it. In
such cases the router would silently drop the traffic (it might be nice to respond with an ICMP, but to whom?); it should of course maintain appropriate counters and/or logs. Similarly, since traffic is routed according to both its source and destination addresses, traffic using a source address in 2001:db8:2::/48 would have no chance of existing to ISP-1. It is still possible for traffic from outside to come in, however, as such routes would have a source prefix of ::/0 or 2000::/3.

2.4. Route Precedence

Precedence in selection of routes is based on intersection, and follows the rule "most specific first". This is a generalization of the "longest match first" rule used in destination routing. That may require, in generating the Forwarding Information Base (FIB) from the Routing Information Base (RIB), that we calculate the least intersection of two route announcements.

In calculating routes, either one route’s traffic class is a subset of another’s, or they mutually incompatible. If one is a subset of the other, we consider the superset to be "less specific" and the subset to be "more specific"; the most specific matching traffic class rules. If the most specific option is in fact multiple traffic classes none of which are subsets of the others, we calculate the intersections of those routes for the purpose of comparison, and apply the rule to the resulting route announcements.

For example, consider the two routes

$$\{(2000::/3, ::/0, \text{any}), 1\} \text{ and } \{(:/0, 2000::/3, \text{any}), 5\}.$$

They are not comparable because neither is a subset of the other. We therefore calculate the intersection, which is a choice between

$$\{(2000::/3, 2000::/3, \text{any}), 1\} \text{ and } \{2000::/3, 2000::/3, \text{any}), 5\}.$$

Given that choice, the metric clearly selects $$\{(2000::/3, 2000::/3, \text{any}), 1\}$$. So we install in the RIB the three routes

$$\{(2000::/3, 2000::/3, \text{any}), 1\}, \{(2000::/3, ::/0, \text{any}), 1\}, \text{ and } \{(2000::/3, ::/0, \text{any}, 1\}.$$
((::/0, 2000::/3, any), 5).

The first is used by unicast traffic, as it is the most specific that matches traffic from and to addresses in 2000::/3. The second is used, for example, with traffic that is from addresses whose most significant three bits are not 001 and to an address in 2000::/3. The third is used for traffic that is to addresses whose most significant three bits are not 001 but are from addresses in 2000::/3.

3. Sketch of a distance vector implementation

A distance vector implementation would operate much as described in Section 2. In addition, however, we might take advantage of an algorithm used in AODV [RFC3561]) to simplify count-to-infinity-related problems.

To this end, we use the mechanisms and algorithms of [RFC2080] with certain modifications.

A Route Table Entry (RTE), in this model, might be structured as in Figure 2.
Figure 2: Route Table Entry

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>route tag (2)</th>
<th>A</th>
<th>N</th>
<th>flags</th>
<th>metric (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Announcement sequence number (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPv6 originator address (16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dest prefix len</td>
<td>source prf len</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPv6 destination prefix (up to 16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPv6 source prefix (up to 16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPv6 Traffic Class DSCP (64 bits)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The fields are:

Route Tag: The Route Tag is as defined in [RFC2080].

A: If 1, any DSCP applies, and the IPv6 Traffic Class DSCP mask is absent. If 0, the IPv6 Traffic Class DSCP is present.

N: if 1, a null route; traffic in this traffic class and not in more explicit match SHOULD be dropped.

flags: unspecified in this version of the specification. Set to 0 on transmission and ignored on receipt. Future versions may specify additional flags.

Metric: The Metric is as defined in [RFC2080].

Announcement sequence number: 0..FFFFFFFF, follows the rules for a sequence number specified in [RFC3561].
IPv6 originator address: One of the global addresses of the router originating this RTE, presumably the one used on the relevant interface, which is in control of the sequence number. See [RFC3561] for considerations and algorithms.

Destination Prefix Length: 0..128

Source Prefix Length: 0..128

IPv6 destination prefix: The significant bits of the prefix in question. Occupies ceiling(destination prefix length/8) bytes.

IPv6 source prefix: The significant bits of the prefix in question. Occupies ceiling(source prefix length/8) bytes.

IPv6 Traffic Class DSCP: if A=0, not present; if A=0, the most significant bit is numbered 0, and indicates that the traffic class includes traffic with a DSCP of 000000; if A=1, not present.

Distribution and calculation of routes follows [RFC2080], including split horizon (an announcement received from a router on the interface should not be reannounced to the same router) and poison reverse (which should be used on triggered updates but not regular announcements).

Elimination of stale routing information of routes follows the algorithm with the sequence number specified in [RFC3561].

Intersection of route announcements with interface metric data is as described in Section 2.

4. Sketch of a link state implementation

A link state (SPF) implementation would also operate much as described in Section 2, although the algorithm operates in memory as opposed to being distributed.

To this end, we use the mechanisms and algorithms of [RFC5308] with certain modifications.

IS-IS structures its network as a lattice of routers that lead one to sets of hosts identified by "TLVs". A TLV, in this model, might be structured as in Figure 3.
The fields are:

U: up/down bit

X: external original bit

S: sub-TLV present bit

A: If 1, any DSCP applies, and the IPv6 Traffic Class DSCP mask is absent. If 0, the IPv6 Traffic Class DSCP is present.

N: if 1, a null route; traffic in this traffic class and not in more explicit match SHOULD be dropped.
flags: unspecified in this version of the specification. Set to 0 on transmission and ignored on receipt. Future versions may specify additional flags.

Destination Prefix Length: 0..128

Source Prefix Length: 0..128

IPv6 destination prefix: The significant bits of the prefix in question. Occupies ceiling(destination prefix length/8) bytes.

IPv6 source prefix: The significant bits of the prefix in question. Occupies ceiling(source prefix length/8) bytes.

IPv6 Traffic Class DSCP: if A=0, not present; if A=0, the most significant bit is numbered 0, and indicates that the traffic class includes traffic with a DSCP of 000000; if A=1, not present.

Distribution and calculation of routes follows [RFC5308].

Intersection of route announcements with interface metric data is as described in Section 2.

5. IANA Considerations

At this point, this memo asks the IANA for no new parameters and gives the IANA no instructions. As development progresses, that might change.

6. Security Considerations

Security issues in routing protocols such as these are the same as in other distance vector and link state routing protocols, and need to be mitigated in the same ways. Since this paper looks primarily at the algorithms for route calculation, those issues are largely ignored. If a protocol such as described in Section 3 or Section 4 is implemented, however, care must be taken to ensure the integrity of communications between routers and their mutual authentication.

7. Acknowledgements

Dana Blair reviewed the initial version of the draft.
8. Change Log


9. References

9.1. Normative References


9.2. Informative References


Author’s Address

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

Email: fred@cisco.com
CPE Considerations in IPv6 Deployments
draft-herbst-v6ops-cpeenhancements-00

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/1id-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 April 17, 2011.

Copyright Notice

Copyright (c) 2010 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 BSD License.

This Internet-Draft will expire on April 17, 2011.
Abstract

Smart metering deployments in residential settings introduce the prospects of ad-hoc deployment of internetworked IPv6 customer premise equipment (CPE). WiFi access points, cable boxes and other home devices with internet access could all be internetworked with smart metering devices by customers with no data networking expertise resulting in a complex multi-segment network with differing prefixes, routing support and service discovery needs.

Table of Contents

1. Introduction ....................................................... 2
2. Description ......................................................... 2
   2.1. Unique Local Addresses (ULAs) for Site Local Multicast. .... 3
   2.2. ULA Delegation When Combining Network Segments. ......... 4
   2.3. Intra-network routing with multiple internet connected CPEs 4
3. Future Work ......................................................... 4
4. Conclusions ......................................................... 4
5. Security Considerations ............................................ 4
6. IANA Considerations ................................................. 4
7. Acknowledgments ..................................................... 4
8. References .......................................................... 4
   8.1. Normative References ........................................... 5
   9.2. Informative References ......................................... 5
Authors’ Address ....................................................... 5
1. Introduction

The availability of energy usage information within the Home Area Network, enabled through smart meter deployment, adds a popular interconnection target for electricity customers, service providers and third party suppliers. These opportunities for energy usage management are all assuming a home owner with no data networking expertise can link together a collection of standalone networks and enable a consistent set of services and device addressing modes. This draft starts a discussion on needed standards work to make the deployment of these services a reality in an IPv6 environment.

2. Description

In a regulated utility environment, utilities must deploy energy savings programs accessible to all customers. Broadband internet access cannot be assumed since around 40% of customers don’t have broadband. The smart meter is then architected as a standalone border gateway with a unique prefix supplied by the smart meter.

To fully enable deployment of energy savings applications onto a variety of devices, ad hoc internetworking of smart meters with HAN devices and existing networks employing diverse data links such as IEEE 802.15.4, IEEE P1901, and WiFi must be supported. The set of issues to be addressed include:

- Assignment of /64 prefixes from Globally Unique Address (GUA) and Unique Local Address (ULA) [RFC4193] prefixes available to the residential network
- Introduction of ULAs for a residence
- ULA Delegation and Reassignment when network segments with differing ULAs are combined
- Enablement of intra-network routing when CPEs within the residence are interconnected
- Extensions to multicast DNS to extend local name resolution across a multi-link residential network
2.1 Assignment of /64 prefixes from GUA and ULA prefixes

The residential network that includes multiple links will need a mechanism for assigning /64 prefixes for each link from one or more shorter prefixes assigned to the network. For example, DOCSIS 3.0 uses DHCPv6-PD [RFC3633] to delegate a prefix to the residential gateway. /64 prefixes from this delegated prefix must be assigned to every link within the residential network.

Similarly, the residential network may have a ULA prefix for local traffic if the residential network does not have any GUA prefixes (see section 2.2). /64 prefixes from the ULA must be assigned to the links in the residential network.

2.2 Unique Local Addresses (ULAs)

IPv6 offers three types of addressing prefixes: GUA, ULA and link-local. ULA prefixes are useful in the residential network scenario for local communication when no GUA prefixes are available; e.g., when the external link to the ISP is unavailable and no delegated prefixes are available.

The first requirement is that gateways in the residential network create a ULA for use within the network rooted at the gateway and no other ULA prefix is available.

The second requirement is that when multiple networks are created, then interconnected in a home, multiple ULAs may be present. When these networked are interconnected (by a homeowner without networking skills), the ULAs for these network segments should be harmonized without user interaction into a single set of ULAs and notification made to hosts holding references to the previous ULAs.

2.3 Intra-network routing with multiple internet connected CPEs

As network segments are interconnected, and CPE devices become border gateways for new bordering network segments, a routing protocol like RIPng needs to be supported. As noted for ULA delegation, the CPE needs to automatically detect the need in support for inter-segment routing and provide support automatically.

2.4 Extensions to multicast DNS for sitewide name resolution

For service discovery, two alternatives exist: user agent based
discovery and directory agent based discovery described as follows:

- **User agent**: Devices hold service discovery information themselves and respond to discovery requests based on matching criteria in the request. DNS Service Discovery [DNS-SD] resolved over Multicast DNS [mDNS] is an example of this type of solution.
- **Directory agent**: Devices register service discovery information with a central repository. A well known example of this type of solution includes uPnP [uPnP] which uses the Simple Service Discovery Protocol (SSDP) [SSDP]. Note that uPnP supports both user agent and directory agent service discovery methods.

mDNS only provides link-local name resolution. Use of mDNS in the residential network requires extensions so that mDNS can use site-local multicast that spans multiple hops using IP forwarding for sitewide local name resolution.

### 3. Future Work

The following work items are proposed:

- Create extensions to DHCPv6-PD to delegate prefixes across multiple links
- Define procedures for gateways to generate a ULA if required
- Create procedures for HAN devices to join the ULA and procedures to combine network segments with different ULAs into a single ULA.
- Define mechanisms for automated provisioning and operation of routing across multiple links in a residential network
- Create extensions to multicast DNS to support sitewide local name resolution across multiple links

### 4. Conclusions

To realize deployment requirements of self installed, ad hoc networking where different segments can be installed and provisioned at different times and where various link technologies may be used, additional features are needed in CPE.

### 5. Security Considerations

This requirements document introduces no security considerations.

### 6. IANA Considerations

This requirements document introduces no IANA considerations.

### 7. Acknowledgments
8. References

8.1. Normative References

8.2. Informative References


Authors’ Addresses

Tom Herbst
Silver Spring Networks
Redwood City, CA
USA

Phone: +1 650-542-4782
    Email: therbst@silverspringnet.com

Don Sturek
Pacific Gas & Electric
77 Beale Street
San Francisco, CA
USA

Phone: +1-619-504-3615
    Email: d.sturek@att.net
Advanced Requirements for IPv6 Customer Edge Routers
draft-ietf-v6ops-ipv6-cpe-router-bis-01

Abstract

This document continues the work undertaken by the IPv6 CE Router Phase I work in the IETF v6ops Working Group. Advanced requirements or Phase II work is covered in this document.

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 January 12, 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
Table of Contents

1. Introduction .............................................. 3
   1.1. Requirements Language ................................. 3
2. Terminology ................................................. 3
3. Conceptual Configuration Variables .......................... 4
4. Architecture .................................................. 4
5. Advanced Features and Feature Requirements ................. 6
   5.1. DNS ......................................................... 6
   5.2. Multicast Behavior ....................................... 6
   5.3. Routed network behavior ............................... 7
   5.4. Transition Technologies Support ....................... 7
      5.4.1. Dual-Stack(DS)-Lite ............................... 7
      5.4.2. 6rd ................................................ 9
      5.4.3. Transition Technologies Coexistence ............... 9
   5.5. Quality Of Service ..................................... 10
   5.6. Unicast Data Forwarding ............................... 10
   5.7. Additional DHCPv6 WAN Requirement .................... 10
6. Security Considerations ..................................... 10
7. Acknowledgements ........................................... 10
8. Contributors ................................................. 11
9. IANA Considerations ........................................ 11
10. References ................................................ 11
    10.1. Normative References ................................. 11
    10.2. Informative References ............................... 14
Authors’ Addresses ............................................. 14
1. Introduction

This document defines Advanced IPv6 features for a residential or small office router referred to as an IPv6 CE router. Typically these routers also support IPv4. The IPv6 End-user Network Architecture for such a router is described in [RFC6204]. This version of the document includes the requirements for Advanced features.

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. Terminology

End-user Network one or more links attached to the IPv6 CE router that connect IPv6 hosts.

IPv6 Customer Edge router a node intended for home or small office use which forwards IPv6 packets not explicitly addressed to itself. The IPv6 CE router connects the end-user network to a service provider network.

IPv6 host any device implementing an IPv6 stack receiving IPv6 connectivity through the IPv6 CE router.

LAN interface an IPv6 CE router’s attachment to a link in the end-user network. Examples are Ethernets (simple or bridged), 802.11 wireless or other LAN technologies. An IPv6 CE router may have one or more network layer LAN interfaces.

Service Provider an entity that provides access to the Internet. In this document, a Service Provider specifically offers Internet access using IPv6, and may also offer IPv4 Internet access. The Service Provider can provide such access over a variety of different transport methods such as DSL, cable, wireless, and others.
WAN interface an IPv6 CE router’s attachment to a link used to provide connectivity to the Service Provider network; example link technologies include Ethernets (simple or bridged), PPP links, Frame Relay, or ATM networks as well as Internet-layer (or higher-layer) "tunnels", such as tunnels over IPv4 or IPv6 itself.

3. Conceptual Configuration Variables

The CE Router maintains such a list of conceptual optional configuration variables.

1. Enable an IGP on the LAN.

2. Configure 6rd configuration.

3. Configure IPv6 for 6rd to have IPv6 traffic go to the 6rd Border Relay vs. directly to peers.

4. Architecture

This document extends the architecture described in [RFC6204] to cover a strictly larger set of operational scenarios. In particular, QoS, multicast, DNS, routed network in the home, transition technologies, and conceptual configuration variables. This document also extends the model described in [RFC6204] to a two router topology where the two routers are connected back-to-back (the LAN of one router is connected to the WAN of the other router). This topology is depicted below:

For DNS, the operational expectation is that the end-user would be able to access home hosts from the home using DNS names instead of more cumbersome IPv6 addresses. Note that this is distinct from the requirement to access home hosts from outside the home.

End-users are expected to be able to receive multicast video in the home without requiring the CE router to include the cost of supporting full multicast routing protocols.
5. Advanced Features and Feature Requirements

The IPv6 CE router will need to support connectivity to one or more access network architectures. This document describes an IPv6 CE router that is not specific to any particular architecture or Service Provider, and supports all commonly used architectures.

5.1. DNS

D-1: The CE Router MAY include a DNS server authoritative for .local to handle local queries. If the service provider specifies one or more DNS resolvers in DHCP configuration options, the CE router SHOULD forward all non-local DNS queries unchanged to those servers. The CE Router MAY also include DNS64 functionality which is specified in [RFC6147].

5.2. Multicast Behavior

This section is only applicable to a CE Router with at least one LAN interface. A host in the home is expected to receive multicast video. Note the CE Router resides at edge of the home and the Service Provider, and the CE Router has at least one WAN connection for multiple LAN connections. In such a multiple LAN to a WAN topology at the CE Router edge, it is not necessary to run a multicast routing protocol and thus MLD Proxy as specified in [RFC4605] can be used. The CE Router discovers the hosts via a MLDv2 Router implementation on a LAN interface. A WAN interface of the CE Router interacts with the Service Provider router by sending MLD Reports and replying to MLD queries for multicast Group memberships for hosts in the home.

The CE router SHOULD implement MLD Proxy as specified in [RFC4605]. For the routed topology shown in Figure 1, each router implements a MLD Proxy. If the CE router implements MLD Proxy, the requirements on the CE Router for MLD Proxy are listed below.

WAN requirements, MLD Proxy:

WMLD-1: Consistent with [RFC4605], the CE router MUST NOT implement the router portion of MLDv2 for the WAN interface.

LAN requirements, MLD Proxy:

LMMLD-1: The CPE Router MUST follow the model described for MLD Proxy in [RFC4605] to implement multicast.
LMMLD-2: Consistent with [RFC4605], the LAN interfaces on the CPE router MUST NOT implement an MLDv2 Multicast Listener.

LAN requirements:

LM-1: If the CE Router has bridging configured between the LAN interfaces, then the LAN interfaces MUST support snooping of MLD [RFC3810] messages as per [RFC4541].

5.3. Routed network behavior

CPE Router Behavior in a routed network:

R-1: One example of the CPE Router use in the home is shown below. The home has a broadband modem combined with a CPE Router, all in one device. The LAN interface of the device is connected to another standalone CPE Router that supports a wireless access point. To support such a network, this document recommends using prefix delegation of the prefix obtained either via IA_PD from WAN interface or a ULA from the LAN interface. The network interface of the downstream router MAY obtain an IA_PD via stateful DHCPv6. If the CPE router supports the routed network through a vendor specific automatic prefix delegation, the CPE router MUST support a DHCPv6 server or DHCPv6 relay agent. Further, if an IA_PD is used, the Service Provider or user MUST allocate an IA_PD or ULA prefix short enough to be delegated and subsequently used for SLAAC. Therefore, a prefix length shorter than /64 is needed. The CPE Router MAY support and IGP in the home network.

/-------+------------\    /------------+-----\ 
SP <--+ Modem | CPE Router    +--+ CPE Router | WAP + --> PC 
\-------+------------/    \------------+-----/ 

WAP = Wireless Access Point

Figure 2.

5.4. Transition Technologies Support

5.4.1. Dual-Stack(DS)-Lite

Even as users migrate from IPv4 to IPv6 addressing, a significant percentage of Internet resources and content will remain accessible.
only through IPv4. Also, many end-user devices will only support IPv4. As a consequence, Service Providers require mechanisms to allow customers to continue to access content and resources using IPv4 even after the last IPv4 allocations have been fully depleted. One technology that can be used for IPv4 address extension is DS-Lite.

DS-Lite enables a Service Provider to share IPv4 addresses among multiple customers by combining two well-known technologies: IP in IP (IPv4-in-IPv6) tunneling and Carrier Grade NAT. More specifically, Dual-Stack-Lite encapsulates IPv4 traffic inside an IPv6 tunnel at the IPv6 CE Router and sends it to a Service Provider Address Family Translation Router (AFTR). Configuration of the IPv6 CE Router to support IPv4 LAN traffic is outside the scope of this document.

The IPv6 CE Router SHOULD implement DS-Lite functionality as specified in [I-D.ietf-softwire-dual-stack-lite].

WAN requirements:

DLW-1: To facilitate IPv4 extension over an IPv6 network, if the CE Router supports DS-Lite functionality, the CE Router WAN interface MUST implement a B4 Interface as specified in [I-D.ietf-softwire-dual-stack-lite].

DLW-2: If the IPv6 CE Router implements DS-Lite functionality, the CE Router MUST support using a DS-Lite DHCPv6 option [I-D.ietf-softwire-ds-lite-tunnel-option] to configure the DS-Lite tunnel. The IPv6 CE Router MAY use other mechanisms to configure DS-Lite parameters. Such mechanisms are outside the scope of this document.

DLW-3: IPv6 CE Router MUST NOT perform IPv4 Network Address Translation (NAT) on IPv4 traffic encapsulated using DS-Lite.

DLW-4: If the IPv6 CE Router is configured with a public IPv4 address on its WAN interface, where public IPv4 address is defined as any address which is not in the private IP address space specified in [RFC1918] and also not in the reserved IP address space specified in [I-D.ietf-softwire-dual-stack-lite], then the IPv6 CE Router MUST disable the DS-Lite B4 element.

DLW-5: If DS-Lite is operational on the IPv6 CE Router, multicast data MUST NOT be sent on any DS-Lite tunnel.
5.4.2. 6rd

The IPv6 CE Router can be used to offer IPv6 service to a LAN, even when the WAN access network only supports IPv4. One technology that supports IPv6 service over an IPv4 network is IPv6 Rapid Deployment (6rd). 6rd encapsulates IPv6 traffic from the end user LAN inside IPv4 at the IPv6 CE Router and sends it to a Service Provider Border Relay (BR). The IPv6 CE Router calculates a 6rd delegated IPv6 prefix during 6rd configuration, and sub-delegates the 6rd delegated prefix to devices in the LAN.

The IPv6 CE Router SHOULD implement 6rd functionality as specified in [RFC5969].

6rd requirements:

6RD-1: If the IPv6 CE Router implements 6rd functionality, the CE Router WAN interface MUST support at least one 6rd Virtual Interface and 6rd CE functionality as specified in [RFC5969].

6RD-2: If the IPv6 CE Router implements 6rd CE functionality, it MUST support user-entered configuration and using the 6rd DHCPv4 Option (212) for 6rd configuration. The IPv6 CE Router MAY use other mechanisms to configure 6rd parameters. Such mechanisms are outside the scope of this document.

6RD-3: If the CE router implements 6rd functionality, it MUST allow the user to specify whether all IPv6 traffic goes to the 6rd Border Relay, or whether other destinations within the same 6rd domain are routed directly to those destinations. The CE router MAY use other mechanisms to configure this. Such mechanisms are outside the scope of this document.

6RD-4: If 6rd is operational on the IPv6 CE Router, multicast data MUST NOT be sent on any 6rd tunnel.

5.4.3. Transition Technologies Coexistence

Run the following four in parallel to provision CPE router connectivity to the Service Provider:

1. Initiate IPv4 address acquisition.
2. Initiate IPv6 address acquisition as specified by [RFC6204].
3. If 6rd is provisioned, initiate 6rd.
4. If DS-Lite is provisioned, initiate DS-Lite.

The default route for IPv6 through the native physical interface should have preference over the 6rd tunnel interface. The default route for IPv4 through the native physical interface should have preference over the DS-Lite tunnel interface.

5.5. Quality Of Service

Q-1: The CPE router MAY support differentiated services [RFC2474].

5.6. Unicast Data Forwarding

The null route introduced by the WPD-6 requirement in [RFC6204] has lower precedence than other routes except for the default route.

5.7. Additional DHCPv6 WAN Requirement

When the WAN interface sends a DHCPv6 SOLICIT message, the CE router SHOULD request all mandatory information (IA_NA and IA_PD options) in the SOLICIT regardless of whether any partial information was received in response to previous SOLICITs.

6. Security Considerations

None.

7. Acknowledgements

Thanks to the following people (in alphabetical order) for their guidance and feedback:

Mikael Abrahamsson, Merete Asak, Scott Beuker, Mohamed Boucadair, Rex Bullinger, Brian Carpenter, Remi Denis-Courmont, Gert Doering, Alain Durand, Katsunori Fukuoka, Tony Hain, Thomas Herbst, Kevin Johns, Stephen Kramer, Victor Kuarsingh, Francois-Xavier Le Bail, Chad Mikkelsen, David Miles, Shin Miyakawa, Jean-Francois Mule, Michael Newbery, Carlos Pignataro, John Pomeroy, Antonio Querubin, Teemu Savolainen, Matt Schmitt, Hiroki Sato, Mark Townsley, Bernie Volz, James Woodyatt, Dan Wing and Cor Zwart

This draft is based in part on CableLabs’ eRouter specification. The authors wish to acknowledge the additional contributors from the eRouter team:

Ben Bekele, Amol Bhagwat, Ralph Brown, Eduardo Cardona, Margo Dolas,
8. Contributors

The following people have participated as co-authors or provided substantial contributions to this document: Ralph Droms, Kirk Erichsen, Fred Baker, Jason Weil, Lee Howard, Jean-Francois Tremblay, Yiu Lee, John Jason Brzozowski and Heather Kirksey.

9. IANA Considerations

This memo includes no request to IANA.

10. References

10.1. Normative References

[I-D.ietf-softwire-ds-lite-tunnel-option]

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

[I-D.vyncke-advanced-ipv6-security]


[RFC2080] Malkin, G. and R. Minnear, "RIPng for IPv6", RFC 2080,


[RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", BCP 122, RFC 4632, August 2006.


10.2. Informative References


Authors’ Addresses

Hemant Singh
Cisco Systems, Inc.
1414 Massachusetts Ave.
Boxborough, MA  01719
USA

Phone: +1 978 936 1622
Email: shemant@cisco.com
URI:   http://www.cisco.com/

Wes Beebee
Cisco Systems, Inc.
1414 Massachusetts Ave.
Boxborough, MA  01719
USA

Phone: +1 978 936 2030
Email: wbeebee@cisco.com
URI:   http://www.cisco.com/
Chris Donley  
CableLabs  
858 Coal Creek Circle  
Louisville, CO  80027  
USA  
Email: c.donley@cablelabs.com  

Barbara Stark  
ATT  
725 W Peachtree St  
Atlanta, GA  30308  
USA  
Email: barbara.stark@att.com  

Ole Troan (editor)  
Cisco Systems, Inc.  
Veversmauet 8  
N-5017 BERGEN,  
Norway  
Email: ot@cisco.com