Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 4, 2011

T. Boot
Infinity Networks B.V.
A. Holtzer
TNO ICT

January 31, 2011

BRDP Framework
draft-boot-brdp-framework-00.txt

Abstract

This document describes the Border Router Discovery Protocol (BRDP) framework. This framework enables multi-homing for small to medium sites, using Provider Aggregatable IPv6 addresses. It describes a mechanism for automated IP address configuration and renumbering, a mechanism for optimized source address selection and a new paradigm for packet forwarding. The BRDP framework prevents ingress filtering problems with multi-homed sites and supports load-balancing for multi-path transport protocols. This work also prevents routing scalability problems in the provider network and Internet Default Free Zone because small to medium multi-homed size sites would not need to request Provider Independent address blocks.

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 August 4, 2011.

Table of Contents

1. Introduction

IPv6 provides basic functionality for multi-homing, since nodes can have multiple addresses configured on their interfaces. However, it is difficult to utilize the advantages of this, as there is a strong tendency among network administrators for shielding the network topology from hosts. As a result, it is difficult or impossible for a host to utilize available facilities of the network, such as multi-path. Also scalability of the Internet routing system is getting a problem due to a high demand of Provider Independent (PI) addresses.

The Border Router Discovery Protocol (BRDP) enhances the IPv6 model by enabling automated renumbering in dynamically changing multi-homed environments, such that routers and hosts cooperate on address configuration and path selection. BRDP utilizes Provider Aggregatable (PA) addresses and uses them as locator. Mapping identifiers to locators is out of scope of the BRDP framework, also because many solutions exists or are being worked on. All these solutions work fine with BRDP, as long as they don’t break IPv6.

The BRDP framework can be applied to edge networks. These networks can be fixed, for example enterprise networks, small offices / home offices (SOHO) or home sites. BRDP also can be used in wireless access networks, for example wireless access networks such as 3G or 4G, wireless LANs or mobile ad hoc networks (MANETs). A nice attribute of BRDP is that it supports multi-homing in heterogeneous networks, meaning that e.g. a SOHO network can have multiple wired broadband and 3G connections or a mixture of wireless access networks and MANET.

In a multi-homed network, nodes are connected to the Internet via multiple exit points, possibly via multiple providers. [RFC5887] argues that if a site is multi-homed using multiple PA routing prefixes, then the interior routers need a mechanism to learn which upstream providers and corresponding PA prefixes are currently reachable and valid. Next to that, these upstream providers or PA prefixes may change over time. This requires a dynamic renumbering mechanism that can handle planned or unplanned changes in the prefixes used. BRDP proposes a mechanism for automated renumbering in larger networks that goes beyond hosts in a single subnet.

The BRDP framework uses the following key elements:

- Propagation of available Border Routers and corresponding prefixes;
- Address autoconfiguration and prefix delegation, using BRDP provided hints;

- Source address selection, using BRDP provided hints;
- Packet forwarding to the Border Router that corresponds with the source address prefix, in case the destination address is not found in the routing domain.

The propagation of available Border Routers and corresponding prefixes is implemented as an extension on the Neighbor Discovery Protocol [RFC4861]. Border Router Information Options (BRIOs) are sent with Router Advertisements, and contain information about the Border Routers, such as the Border Router address, the prefix that corresponds with that Border Router, and the costs of the path via that Border Router to the Internet Default Free Zone (DFZ).

BRIOs are disseminated downstream through the network and all nodes store the information from BRIOs they receive in a BRIO cache. When a node is multi-homed it will receive multiple prefix information, from multiple upstream Border Routers. BRIOs contain a Border Router prefix and routers can generate an IPv6 address based on this prefix [I-D.boot-autoconf-brdp], just like in regular SLAAC [RFC4862]. Routers can set up reachability to this address automatically, by adding the generated address in the routing protocol.

Routers automatically learn Border Routers that act as DHCP server or relay agent [RFC3633]. When routers detect an alternate path to the DFZ, a new prefix is requested from this newly learned Border Router. Prefixes, of which the path to the DFZ is no longer available, are put ’out of service’ by routers, meaning they are not used for address assignments anymore. Optionally, if the cost to the DFZ through a Border Router is far higher than via other available paths, a router can put the corresponding prefix out of service. Prefixes that are out of service are released.

Prefixes that are in service are configured on interfaces with a 64-bit prefix length and advertised with a Prefix Information Option in Router Advertisements. The Prefix Information lifetime is copied from lifetime information in the BRIO cache.

Hosts can use the BRDP provided information together with the Prefix Information to autoconfigure addresses, based on IPv6 Stateless Address Autoconfiguration [RFC4862]. A host may also use DHCPv6 to get addresses or "Other configuration".

Nodes with multiple configured addresses need to select a source address for outgoing connections. Default Address Selection for IPv6 [RFC3484] defines a mechanism, used as default behavior. It is open to more advanced mechanisms or site policies. BRDP provided information can be used for a more advanced mechanism, where the

hosts select automatically a source address that corresponds with a path with the lowest cost to the DFZ. When multiple Border Routers are available, automatic load distribution and multi-path transport becomes available.

Hosts can use information in the BRIO cache to select a default router. For selecting the best paths, hosts may use next hop selection based on source address and path costs to the corresponding Border Router, if such information is available to the host.

Network Ingress Filtering [RFC2827] describes the need for ingress filtering, to limit the impact of distributed denial of service attacks, by denying traffic with spoofed source addresses access. It also helps ensure that traffic is traceable to its correct source network. Ingress Filtering for Multihomed Networks [RFC3704] provides solutions for multi-homed sites. However, the proposal applicable for PA addresses requires careful planning and configuration. It suggests routing based on source address, and a path on each Border Router to all ISPs in use, either with a direct connection or with tunnels between all Border Routers. It is hard to make such mechanisms work in an automated fashion, or mechanisms are not supported on Border Routers used today. As an evolutionary approach, BRDP provided information is to be used to forward packets to their destination without ingress filtering problems. The BRIO cache contains a mapping between Border Routers and the addresses that do pass ingress filtering. So the packet forwarding heuristic can be straightforward: send packets, where the destination is not in the routing domain itself to the Border Router that owns the prefix of the source address.

Enabling BRDP in an existing network is straightforward. First, all routers have to be updated for BRDP support. At this step, Border Router information is propagated in the network enabling BRDP assisted address autoconfiguration and prefix delegation and BRDP assisted source address selection. The second step is updating all routers with the BRDP based routing mechanism. To enable this mechanism the default gateway is removed from the routing table. This second step is a flag day operation. Rolling back is easy, by just re-inserting the default gateway.

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. Reference Scenarios

This section describes the use of BRDP in four different scenarios: a small multi-homed site or DMZ, a medium multi-homed site, a medium multi-homed site with ULA with DHCP server and a MANET site.

2.1. Small multi-homed site or DMZ

This scenario discusses BRDP operation for multi-homed Small Office - Home Office (SOHO) networks and De-Militarized Zones (DMZ). The scenario is shown in Figure 1. Each provider assigns a PA /48 prefix to its customers. All addresses and prefixes are configured completely automatically. The feature of BRDP that adds value in this scenario is BRDP based Border Router selection for multi-homed hosts. This is enabled by using BRDP based forwarding.

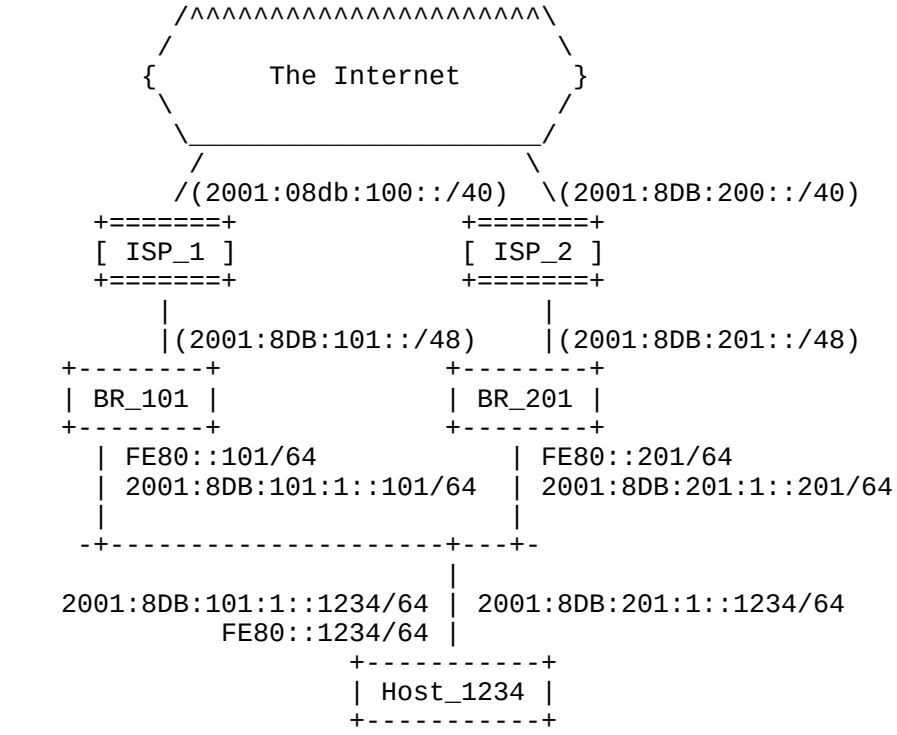

Figure 1: Scenario 1: multi-homed Small Office - Home Office (SOHO) network or DMZ

In this scenario, Host_1234 has configured two addresses using SLAAC [RFC4862], one with prefix 2001:8DB:101:1::/64 from Border Router BR_101 and one with prefix 2001:8DB:201:1::/64 from Border Router

BR_201. Host_1234 has learned these prefixes from Prefix Information Options sent by both Border Routers according to [RFC4861]. The host has learned via BRIOs that these prefixes belong to Border Routers. The host can use optimal paths by selecting BR_101 as default router for all packets with a source address with prefix 2001:8DB:101:1::/64 and default gateway BR_201 for all packets with a source address with prefix 2001:8DB:201:1::/64. Non-optimal default router selection on hosts is handled by the routers, "misdirected" packets are forwarded to the correct Border Router.

BRDP enables routers to deliver non-optimal directed packets from attached hosts towards a Border Router that owns the prefix of the source address, if such a Border Router exists. In the above scenario, a packet sent from Host_1234 with source address 2001:8DB: 201:1::1234 to default router BR_101 would be dropped due to on an ingress filter, when no mechanism is in place to redirect the packet. BRDP based forwarding provides such a mechanism automatically. Instead of dropping the packet, BR_101 forwards it to BR_201.

2.2. Medium multi-homed site

This scenario discusses BRDP operation for medium sized multi-homed networks. The difference with the previous scenario is that the network paths between hosts and the Border Routers have intermediate routers. The scenario is shown in Figure 2. The added value of BRDP in this scenario is the discovery of Border Routers for hosts and routers beyond the first hop as well as Border Router Selection for hosts and routers.

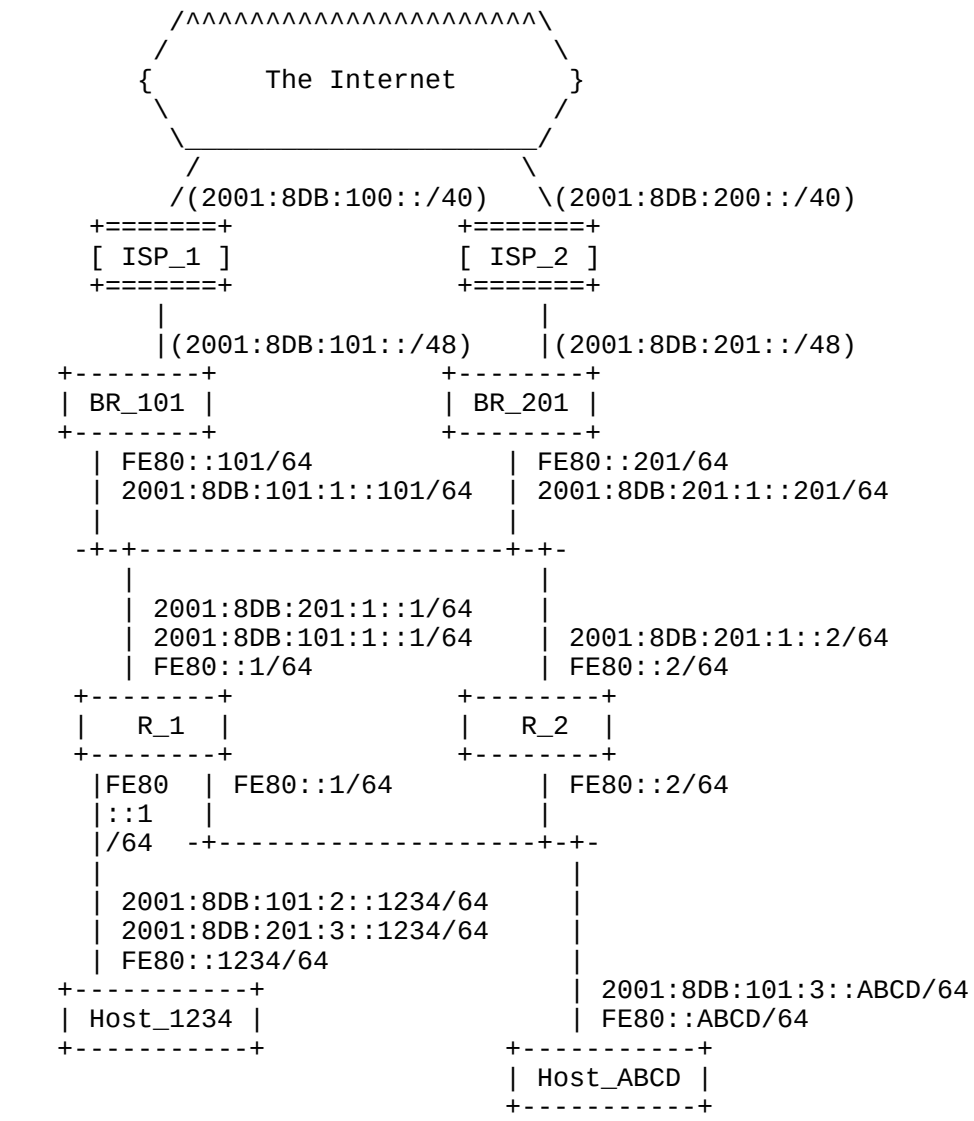

Figure 2: Scenario 2: medium sized multi-homed network

Routers can learn advertised on-link prefixes automatically via the Prefix Information Option in IPv6 ND RAs. In this scenario, routers R_1 and R_2 learn prefix 2001:8DB:101:1::/64 from BR_101 and prefix 2001:8DB:201:1::/64 from BR_201. Routers may autoconfigure addresses on their interfaces. In this example, R_1 has configured addresses from both providers on its upstream interface, R_2 only configured an address based on the prefix of BR_201. If the routers run a routing

protocol, the learned prefixes are made reachable in the network. In the next steps of the autoconfiguration proces, the prefixes and addresses on the other links are automatically configured, but first we discuss the BRDP messages that are disseminated through the network.

Routers automatically learn Border Routers and mapping between prefixes and Border Routers using BRDP. The diagram in Figure 3 depicts BRIO message dissemination in scenario 2. The two Border Routers advertise their own address and corresponding prefix with an address prefix. Nothing prevents them from forwarding each other's BRIO message, although resending BRIO information on non-MANET interfaces is not useful. Both routers R_1 and R_2 forward both Border Router address prefixes, using separate BRIOs in RAs, on downstream interfaces. In this way all routers and hosts in the network are aware of all reachable Border Routers and corresponding prefixes.

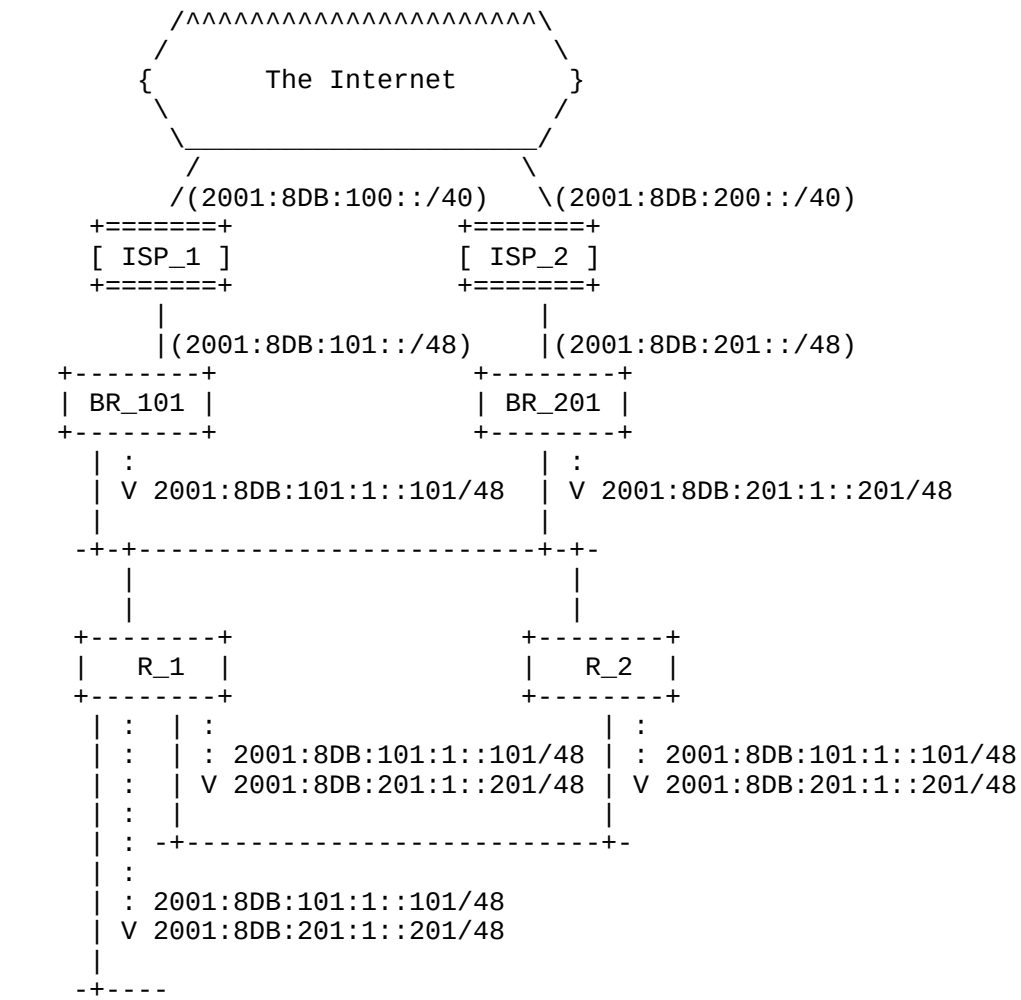

Figure 3: BRIO dissemination in Scenario 2

Routers are not required to configure global addresses on each interface. In the example, only the interface pointing to the Internet has configured global addresses. Routers may also use a (logical) management interface for global reachability.

So, the one-hop neighbours of BR_101 and BR_201, being R_1 and R_2, have learned the prefixes and configured addresses on their upstream interfaces. And all nodes in the network have learned the Border Router prefixes. The next step is to get configured addresses on the hosts in Figure 2. This is done by using DHCP Prefix Delegation. R_1 and R_2 request a prefix from either or both BR_101 and BR_201

for binding as on-link prefix on the links, and advertise those using Prefix Information Options to the hosts. This will result in a maximum of four prefixes that are advertised on the downlink of R_1 and R_2. Having multiple prefixes from the same ISP bound on a link is not useful. So a router requests a prefix from a Border Router only if no other prefix of that Border Router is advertised already by another router on this network segment.

In this example, R_1 has been delegated two prefixes by DHCP PD for the link with host Host_1234; 2001:8DB:101:2::/64 and 2001:8DB:201:3::/64. No other router is on this link. R_1 or R_2 has also been delegated a prefix on the link to host Host_ABCD; 2001:8DB:101:3::/64. It cannot be seen in Figure 2 which router has been delegated the prefix, nor if another prefix for this link has been delegated. No redundant prefix is delegated, as the routers learned with RA PIO already delegated prefixes for known Border Routers.

Now, Host_1234 and Host_ABCD can autoconfigure addresses for their interfaces. Host_1234 configures two addresses, one for each Border Router. Host_ABCD chooses not to use ISP_2.

Nodes R_1 and Host_1234 can use both providers, by using two configured global addresses. Any multi-path facility can be used, either on an application layer or with a multi-path transport protocol.

Host_ABCD may forward packets to the Internet via router R_1 or R_2. If R_2 is selected as default router, R_2 forwards the packets to BR_101 as this Border Router corresponds to the prefix of the source address 2001:8DB:101:3::ABCD. This works well, even in this case where R_2 hasn’t configured an address with a BR_101 prefix for itself, and selected a global address from the BR_201 prefix only.

2.3. Medium multi-homed site with ULAs and DHCP server

In this example, the scenario 2 is extended by adding Unique Local Addresses (ULA) for communication within the site itself. For simplicity there is only one ISP present. The ULA IP configuration, with prefix FD00:8DB::/48, is managed by DHCPv6 server DHCP_201. The scenario is shown in Figure 4.

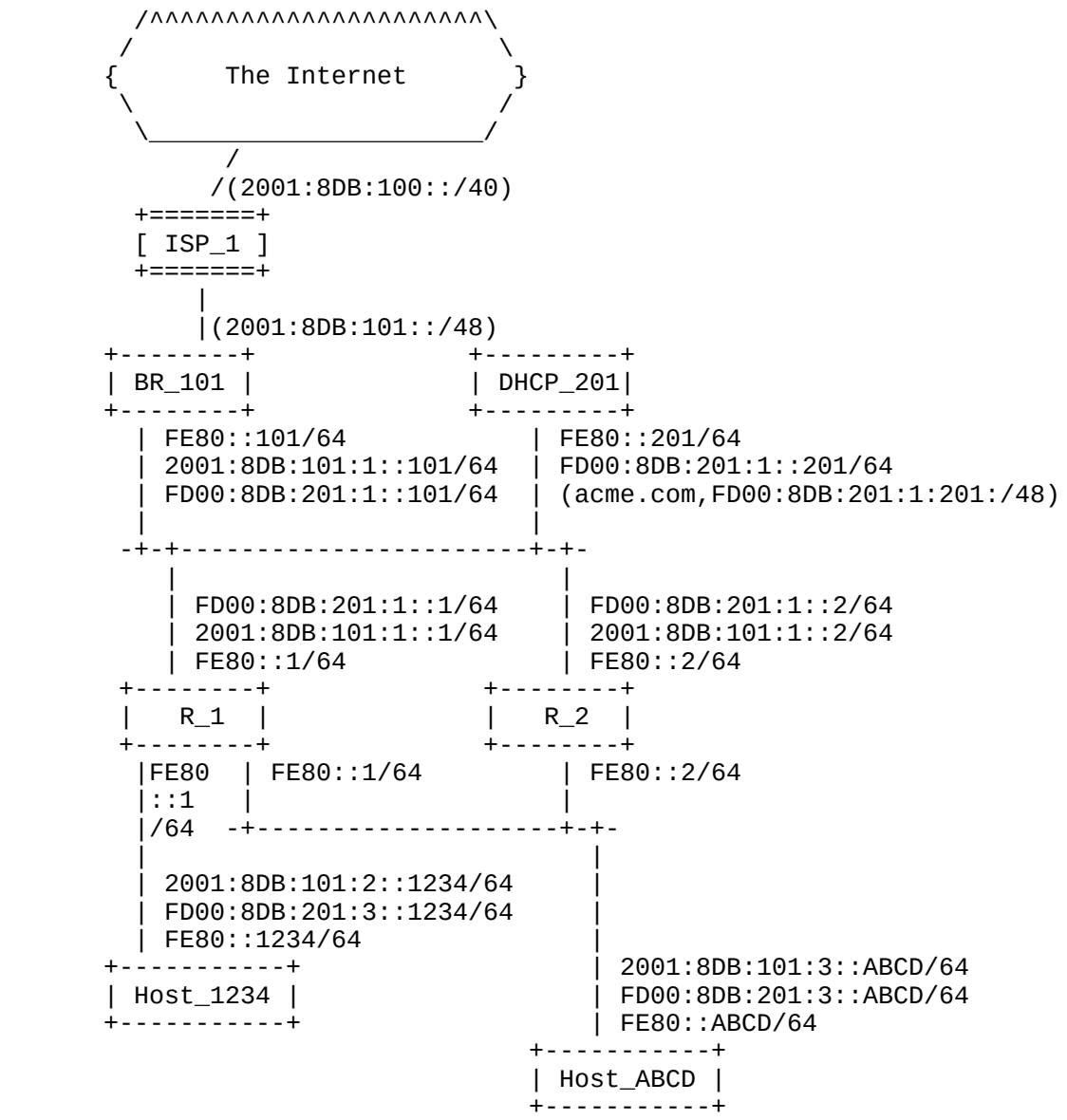

Figure 4: Scenario 3: a medium sized multi-homed site with ULAs and DHCP server.

In this scenario, all nodes have configured a ULA and a Global Unicast Address using prefix delegation in the way that was described in Section 2.2. ULA prefix delegation is automated just like PA addresses. The DHCP server is therefore implemented on a router, in

this case DHCP_201. This router advertises the ULA prefix with BRDP, here FD00:8DB:201::/48.

Although BRDP provides automatic prefix and address configuration for ULA, a network administrator is free to configure it manually, along using BRDP for global addresses.

BRDP based ULA configuration with BRDP based routing would result in routing packets with ULA destinations outside the site to the originator of the ULA prefix, in this case router DHCP_201. DHCP_201 is not connected to the Internet or another site owning the ULA, so packets to non-existing destinations are dropped. DHCP_201 indicates such with the BRIO F-bit set, meaning the Border Router is floating.

This scenario, it is demonstrated that BRDP and DHCPv6 cooperate in address configuration. BRDP provides announcements of Border Routers and DHCP servers. Routers request prefixes with DHCP, and can request other parameters also. Such parameters are disseminated to other nodes, either with router advertisements or acting as DHCP server itself. Routers may also act as DHCP relay, redirecting address requests to the Border Router(s). The Router Advertisement M-bit and O-bit indicates availability of DHCPv6 services to attached nodes.

Difficulties may arise when both ULA and global addresses are used for Internet connectivity, e.g. when address translation is used. To distinguish, the Border Router not providing Internet connectivity informs nodes in the network using Service Selection suboption, similar to "Service Selection for Mobile IPv6" [RFC5149]. This procedure helps also for extranet connectivity. In this scenario, the ULA is used within the ACME Corporation, nodes are made aware by adding "acme.com" in the BRIO Service Selection Option.

It is for the reader to work out extensions for this scenario, where the ULA prefix originator is a Border Router to another site, e.g. a link from a branch office to a head quarter, or a ULA-only side connected to the Internet with NAT66.

2.4. MANET site

BRDP was developed for address autoconfiguration in MANETs. This scenario, see Figure 5 demonstrates the powerful multi-homing facilities provided to the MANET nodes.

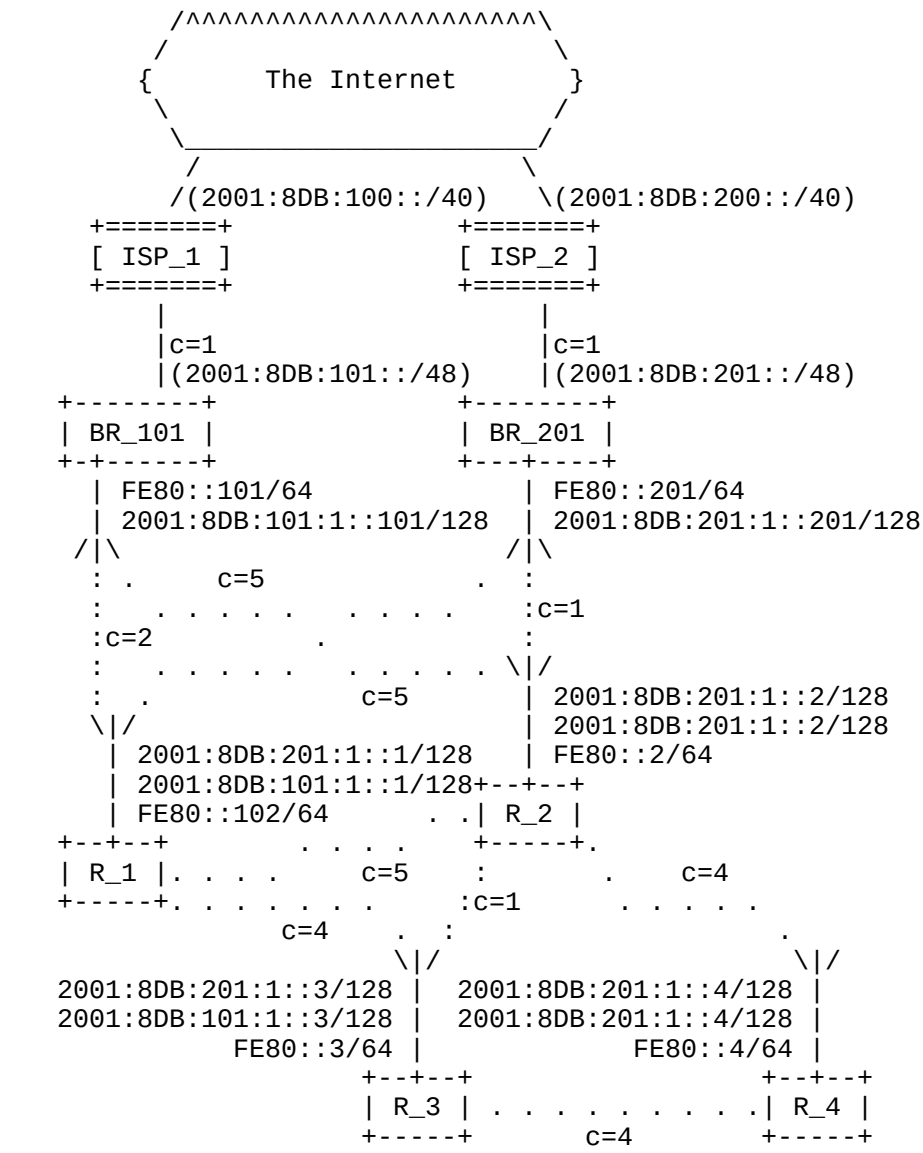

Figure 5: Scenario 4: a MANET site

On the MANET interfaces, addresses are configured using a 64-bit prefix provided by BRDP, appending it with a 64-bit Interface Identifier according to BRDP based address autoconfiguration. This creates a 128-bit prefix length as recommended in IP Addressing Model in Ad Hoc Networks [RFC5889]. Each MANET node has configured two global addresses, one for each ISP. With BRDP, the nodes are aware

of the cost of the path to the DFZ, defined as dimensionless metric for both directions of the patch. This enables optimized source address selection, and as an implicit result a Border Router and ISP selection. In the scenario, R_1 is near to BR_101 and the cost via this Border Router is lower than via BR_201. The table below shows costs to the DFZ for all nodes, via both ISPs. Paths with lowest costs are marked with *.

Costs to DFZ	Via ISP_1	Via ISP_2
BR_101	1*	7
BR_201	7	1*
R_1	3*	6
R_2	6	2*
R_3	7	3*
R_4	10	6*

The optimized source address selection facility is also of utility in the other scenarios. For example, the cost of the link to the ISP could be set depending of bandwidth and optionally on utilization. Nodes would use a near uplink to an ISP, and as a result some form of load distribution is enabled. Note that nodes still can use the alternative addresses, in fact it is recommended to use multi-path transport protocols for better load balancing and improved robustness.

For isolated MANETs, a DHCP server election mechanism can be used. Nodes may initiate to advertise a self-generated ULA. In such cases, it is recommended that a prefix is used with a 56-bit random ULA identifier (including random 16-bit Subnet ID) and 64-bit prefix length. Other nodes join this prefix, although some may wish to start or continue using their own prefix. The latter would occur in cases of a merge of previous isolated MANETs.

3. Border Router Discovery Protocol (BRDP)

BRDP is an extension to the IPv6 ND mechanism [RFC4861] that provides information about the reachability, availability, prefix information, quality and cost of upstream providers, and enables automated (re)numbering of possibly multi-homed routers and hosts.

BRDP adds the Border Router Information Option (BRIO) to the Router Advertisement (RA) of IPv6 ND. A BRIO contains all relevant information of an upstream Border Router and the corresponding

provider.

Border Routers initiate sending BRIO messages, other routers in the network disseminate the messages downstream through the network. Nodes store the information from received BRIOs in a BRIO cache, to be used for address generation, DHCP server discovery, address selection or packet forwarding.

A BRIO cache entry records reception of a BRIO for a single advertised prefix, received via a neighbor router. Border Routers that need to advertise multiple prefixes simply use multiple BRIOs, each with its own address prefix. For further processing of BRIO entries, only the entry with the lowest cost to a Border Router is used.

When a node is multi-homed, it will receive BRIOs from multiple upstream Border Routers. BRIO may have options to indicate connectivity to networks other than the Internet, or indications that usage of the Border Router needs authentication and authorization.

A BRIO message contains informational elements listed below. A more detailed description is provided in BRDP based Address Autoconfiguration [I-D.boot-autoconf-brdp].

Border Router Address:

128-bit address of the Border Router. The Border Router should make this address reachable in the IGP, if a site use an IGP.

Prefix Length:

8-bit unsigned integer. The number of leading bits in the Border Router Address, that indicates the assigned prefix for that Border Router. The Prefix Length is used for BRDP Based Routing [I-D.boot-brdp-based-routing]. The Border Router address prefix specifies the source address ingress filter, if ingress filtering is implemented on this Border Router.

Uniform Path Metric (UPM):

A dimensionless cost measure for the quality of the bi-directional path between the upstream router to the Border Router and the DFZ of the Internet, or to the Border Router itself in case it is floating. UPM is set to some initial value by the Border Router and is incremented by each Router that propagates the BRIO.

Floating(F) flag:

When the F-flag is set, the Border Router does not provide access to the Internet. BRIOs with an ULA prefix SHOULD have the F-bit set.

DHCP (D) flag:

When the D-flag is set, the Border Router is acting as a DHCP server or DHCP relay agent [RFC3315].

Service Selection Identifier:

An option for a variable length UTF-8 encoded Service Selection Identifier string used to identify the Border Routers' type of service. A valid example is 'acme.com'.

A Border Router MAY offer multiple services using multiple BRIOs. However, each of those BRIOs MUST use a unique Border Router address.

A detailed description of BRDP message processing is found in [I-D.boot-autoconf-brdp].

4. BRDP based Address Configuration and Prefix Delegation

BRDP supports stateless address autoconfiguration [RFC4862], DHCP managed IP configuration [RFC3315] and DHCP prefix delegation [RFC3633]. MANET routers can also use a variant of stateless address autoconfiguration, where BRDP provided information is used to configure off-link addresses, used in ad hoc networks [RFC5889].

BRDP adds topology awareness in address configuration. Nodes can configure multiple addresses, each to support a different facility. ULAs can be used for site internal traffic or for Extranets. Global addresses are mandatory for access to the Internet, assuming address translation is not used.

A node that is offered multiple prefixes for stateless address autoconfiguration or multiple addresses by DHCP chooses to configure one or more addresses. BRDP provides information for the candidate addresses. An important criterion is the costs of the path to the Internet DFZ. A node would select an address with the lowest costs. Another criterion is the Service Selection Identifier, to be used for access to private networks.

The BRDP framework does not modify stateless address autoconfiguration and DHCP protocols, except that in a MANET, MANET

routers perform stateless address autoconfiguration from the Border Router Information Option (BRIO) instead of the Prefix Information Option (PIO) [I-D.boot-autoconf-brdp]. This enables MANET-wide address configuration, because BRIOs are disseminated over multiple hops in the MANET, while PIOs are link local messages only.

When a BRIO is stored in the BRIO cache table, the node checks if a corresponding address already exists for the Border Router from which this BRIO originates. If not, and a corresponding address for that Border Router is beneficial, address generation for that Border Router is triggered.

5. BRDP based Source Address Selection

As a next step, multi-homed nodes perform source address selection for new, self-initiated connections. The algorithm described in Default Address Selection for IPv6 [RFC3484] uses the concept of a "candidate set" of potential source addresses. Rule 8 of source address selection is "Uses longest match prefix". The goal of this rule is to select the address with good communications performance. If other means of choosing among source addresses for better performance is available, that should be used.

BRDP provides attributes for prefix, such as a cost metric to the Internet or a Service Selection Identifier. This information van be used to select the "best" source address. For multi-path transport protocols, it is also important to have a mechanism to select alternative addresses. For example, rule 4 gives preference to a Home Address. Alternate addresses can be used for route optimization and to avoid overhead of the Mobile IP tunnel.

BRDP provided information can also be utilized by address lookup protocols such as DNS. A node can register its addresses dynamically, with support of preference and load balancing if the mechanism used support such.

6. BRDP based Routing

The BRDP framework introduces a new paradigm for packet forwarding for multi-homed sites, where forwarding to a default gateway is replaced by source address based forwarding towards a corresponding Border Router [I-D.boot-brdp-based-routing]. This enforces that packets will be sent via the selected upstream provider, without the need of tunneling. As such, it prevents problems with ingress filters in multi-homed edge networks [RFC3704].

The BRDP Based Routing mechanism provides basic support for load distribution over multiple Border Routers. BRDP Based Routing forwards the packets to the Border Router that corresponds with the source address. As a result, nodes can utilize multiple paths, if available. Standardization of this load balancing functionality is work in progress in the IETF MPTCP working group.

When a router forwards a packet to a next-hop node, via the interface where this packet was received, and the next-hop address was selected using BRDP based routing, then the router should not send an ICMP redirect message to the upstream host. This is because the upstream node would cache the redirect for the destination address, while the forwarding decision was based on the source address.

7. BRDP and IRTF RRG goals

The IRTF Routing Research Group (RRG) was chartered to explore solutions for problems on routing and addressing, when the Internet continues to evolve. It has explored a number of proposed solutions, but did not reach consensus on a single, best approach [I-D.irtf-rrg-recommendation]. In fulfillment of the routing research group's charter, the co-chairs recommend that the IETF pursue work in three areas, "Evolution" [I-D.zhang-evolution], "Identifier/Locator Network Protocol (ILNP) [I-D.rja-ilnp-intro] and "Renumbering" [RFC5887] . BRDP fits in all three approaches.

BRDP is an evolution in IPv6 address configuration and address selection, as well as forwarding to destinations outside the routing domain. As a result, it removes a demand for Provider Independent addresses for (small) multi-homed edge networks. BRDP enables sites to use multiple Provider Aggregatable address blocks, while being able to utilize multi-homing for improved redundancy of communications and enlarged capacity. Each site that continues to use Provider Aggregatable addresses when getting multi-homed, instead of using its own Provider Independent address space, reduces the growth of the routing tables in the Default Free Zone.

BRDP can cooperate or live next many other solutions. ILNP is a good example for cooperation, BRDP provides multi-path transport capabilities to ILNP nodes. This multi-path transport capability applies to many other approaches also, such as map&encap and nat66.

Because BRDP provides automatic address and prefix configuration, Renumbering is far less problematic. That said, legacy (IPv4) hosts, applications and network equipment is not BRDP enabled and manual address configuration will be used for many years to come.

In Design Goals for Scalable Internet Routing [I-D.irtf-rrg-design-goals], a number of design goals are defined. The role BRDP can play for these goals are briefly described in the next sections.

7.1. Scalability

Because BRDP is implemented in edge networks, and not in the core, scalability of BRDP is less an issue. BRDP solves the Internet routing problem at the source, by reducing the demand for PI addresses.

7.2. Traffic engineering

BRDP provides traffic engineering options to end-nodes. End-nodes can configure multiple addresses and use these for utilizing multi-path capabilities of the network. Using multi-path is being worked on by the IETF MPTCP working group.

7.3. Multi-homing

The core function of BRDP is providing support for IPv6 multi-homing, without any problems caused by ingress filtering [RFC3704].

7.4. Loc/id separation

BRDP does not mandate any approach for location / identification. For packet forwarding, addresses are used as locator. If addresses are used as identifiers also, for example in Mobile IP, BRDP supports route optimization where traffic uses the Home Address as identifier and care-of addresses as locator. MPTCP provides the route optimization capability.

7.5. Mobility

BRDP was defined as a solution for address autoconfiguration for ad hoc networks. With BRDP, nodes can easily configure topology correct addresses in a multi-homes ad hoc network. BRDP does not provide session continuity functions. Mobility solutions are already in place or new approaches are proposed. All approaches should work well with BRDP, as BRDP does not modify the IPv6 protocol.

7.6. Simplified renumbering

BRDP makes site renumbering fully automatic. This applies to node address configuration on the IPv6 stack and prefix delegation and configuration on routers. IP addresses could be configured on many other places, either manually or using specific protocols for such

purpose. Complete automatic numbering is possible if all mechanisms in use support dynamic addresses. There is definitely more work to do [RFC5887].

7.7. Modularity

BRDP is a small, but important piece of the puzzle. It applies to edge networks only. It helps other mechanisms to work well in a multi-homed network using PA addresses, but also provides multi-path capabilities in multi-homed networks with PI addresses or multi-homing with connections to Extranets.

7.8. Routing quality

BRDP is not a routing protocol, so it has no influence on routing quality. But the functionality of routing to a default gateway is changed. BRDP based routing supports paths to multiple Border Routers, where hosts can select which Border Router to use. In such scheme, nodes can select the route to use, based on quality of available routes. MPTCP provides this route selection functionality.

7.9. Routing security

BRDP doesn’t update any routing protocols. BRDBP based routing modifies the default gateway heuristic, the route to prefix ::/0 is replaced by a route to a Border Router, which corresponds with the source address of a packet. As a result, ingress filtering is distributed over all routers in the edge network and invalid packets are dropped as near to the source as possible.

The BRDP protocol runs on IPv6 NDP and inherits all security aspects. BRDP messages are disseminated in the edge network, which may enlarge the needs for protection. Implementing SeND [RFC3971] is recommended.

7.10. Deployability

BRDP deployment takes place edge network by edge network. Each network that migrates to BRDP, instead of getting a PI address bock, reduces the load on the Internet routing infrastructure.

For implementing BRDP on an edge network, all routers in the network must support BRDP. BRDP support for hosts is optional. Enterprise networks can migrate site by site.

8. Currently unaddressed issues

BRDP based routing may have impact on multicast routing, e.g. selecting the route to a RP.

It is not fully understood how BRDP may influence host behavior on RA M and O bits, and may bypass a 1-hop router DHCP relay server for getting information for a BRDP-learned DHCP server.

Currently unaddressed issues are addressed in a next version of this document.

9. Acknowledgements

TBD

10. IANA Considerations

This memo includes no request to IANA.

11. Security Considerations

No new security considerations arise.

12. References

12.1. Normative References

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

[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.

12.2. Informative References

[I-D.boot-autoconf-brdp] Boot, T. and A. Holtzer, "Border Router Discovery Protocol (BRDP) based Address Autoconfiguration", draft-boot-autoconf-brdp-02 (work in progress), July 2009.

[I-D.boot-brdp-based-routing]
Boot, T., "Border Router Discovery Protocol (BRDP) Based Routing", draft-boot-brdp-based-routing-00 (work in progress), November 2008.

[I-D.irtf-rrg-design-goals]
Li, T., "Design Goals for Scalable Internet Routing", draft-irtf-rrg-design-goals-06 (work in progress), January 2011.

[I-D.irtf-rrg-recommendation]
Li, T., "Recommendation for a Routing Architecture", draft-irtf-rrg-recommendation-16 (work in progress), November 2010.

[I-D.rja-ilnp-intro]
Atkinson, R., "ILNP Concept of Operations", draft-rja-ilnp-intro-09 (work in progress), January 2011.

[I-D.zhang-evolution]
Zhang, B. and L. Zhang, "Evolution Towards Global Routing Scalability", draft-zhang-evolution-02 (work in progress), October 2009.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[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.

[RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, February 2003.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, March 2004.

[RFC3971] Arkko, J., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.

[RFC5149] Korhonen, J., Nilsson, U., and V. Devarapalli, "Service Selection for Mobile IPv6", RFC 5149, February 2008.

[RFC5887] Carpenter, B., Atkinson, R., and H. Flinck, "Renumbering Still Needs Work", RFC 5887, May 2010.

[RFC5889] Baccelli, E. and M. Townsley, "IP Addressing Model in Ad Hoc Networks", RFC 5889, September 2010.

Authors' Addresses

Teco Boot
Infinity Networks B.V.

Email: teco@inf-net.nl

Arjen Holtzer
TNO Information and Communication Technology

Email: arjen.holtzer@tno.nl

Network Working Group
Internet Draft
Intended status: Best Current Practice
Expires: August 5, 2011

S. Jiang
B. Liu
Huawei Technologies Co., Ltd

January 26, 2011

IPv6 Site Renumbering Guidelines and Further Works
draft-jiang-ipv6-site-renum-guideline-00.txt

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 August 5, 2011.

Copyright Notice

Abstract

This document analyzes the existing issues for IPv6 site renumbering. It also analyzes the possible directions to solve these issues and gives recommendations. This document only takes the perspective of network and network protocols. Renumbering in IPv4 networks, in the

dual-stack network or in the IPv4/IPv6 transition networks are out of scope.

This document only takes the perspective of network and network protocols. According to the different stages, these issues are described in three categories: considerations during network design, considerations for routine network management, and considerations during renumbering operation. Recommended solutions or strategies are also described. Issues that still remain unsolvable are listed as the fourth category.

Although we list a few non-network issues in this document, we consider them as issues that ISPs or network providers cannot affect. So, these issues are considered to be unsolvable and not explore further in these document, though they may be solved by OS implementations or application implementations.

We summary the requests that need to extend current protocols as further works at the end of this document.

Table of Contents

1. Introduction

[RFC5887] has reviewed the existing mechanisms for site renumbering for both IPv4 and IPv6, and identified operational issues with those mechanisms. However, the discussion and analysis were too wide. It exposure many issues, but not give enough analysis on whether these issues are solvable and how. Even the document itself indicates more fractionized and detailed analysis is needed. On another side, the mechanisms analyzed in the document are still not in used.

This document focuses on IPv6 network renumbering only, by leaving IPv4 out of scope. Renumbering in IPv4 networks, in the dual-stack network or in the IPv4/IPv6 transition networks are out of scope. This is also consistent preference from IETF renumber mail list by the time of writing up.

This document is mainly concerned with issues affecting medium to large sites, which is taken as the conclusion from [RFC5887]. It takes the analysis conclusions from [RFC5887] and other relevant documents as the primary input.

This document only takes the perspective of network and network protocols. According to the different stages, these issues are described in three categories: considerations during network design, considerations for routine network management, and considerations during renumbering operation. Recommended solutions or strategies are also described. Issues that still remain unsolvable are listed as the fourth category.

Issues that need further analysis are temporarily listed for now. They should all be relocated into abovementioned four categories.

Although we list a few non-network issues in this document, we consider them as issues that ISPs or network providers cannot affect. So, these issues are considered to be unsolvable and not explore further in these document, though they may be solved by OS implementations or application implementations.

At the end of this document, we summary the requests that need to extend current protocols.

2. Network Renumbering Considerations and Solutions/Strategies

The purpose of this section is not to describe the renumbering operation or event completely or entirely, but to expose the existing issues and give the recommended solutions or strategies.

2.1. Considerations/issues during network design

This section describes the renumbering relevant considerations or issues that a network architect should carefully plan when he builds or designs a new network.

- Address configuration models

In IPv6 networks, there are two auto-configuration models for address assignment: the Stateless Address Auto-Configuration (SLAAC) by Neighbor Discovery (ND, [RFC4861, RFC4862]) and the stateful address configuration by Dynamic Host Configuration Protocol for IPv6 (DHCPv6, [RFC3315]). (Manual address configuration is not scalable in medium to large sites, hence be out of scope.)

SLAAC is considered easier to renumbering by broadcasting a Router Advertisement message with a new prefix. DHCPv6 can also trigger the renumbering process by sending unicast RECONFIGURE messages though it may cause large number of interactions between hosts and DHCPv6 server. However, DHCPv6 reconfiguration "doesn’t appear to be widely used for bulk renumbering purposes" [RFC5887].

In principle, a network should choice only one address configuration model and employs either ND or DHCPv6. However, since DHCPv6 is also used to configure many other network parameters, there are ND and DHCPv6 co-existing scenarios. The current protocols do not effectively prevent that both SLAAC and DHCPv6 address assignment are used in the same network (see M bit analysis in section 5.1.1 [RFC5887]. It is network architects’ job to make sure only one configuration model is employed. Even in a large network that contains several subnet works, it is recommended not to mixture the two address configuration models though isolately using them in different subnet works may reduce the risk partly.

On another side, new protocol extension may help to diagnose the fault situation. This diagnose function could be particularly useful in the scenario where a multihomed network uses SLAAC for one address prefix and DHCPv6 for another.

- DNS

It is recommended that the site have an automatic and systematic procedure for updating/synchronising its DNS records, including both forward and reverse mapping. Manually on-demand updating

model is considered as a harmful problem creator in renumbering event.

In order to simplify the operation procedure, the network architect should combine the forward and reverse DNS updates in a single procedure.

If a small site depends on its ISP's DNS system rather than maintains its own one. When renumbering, it requires administrative coordination between the site and its ISP. Alternatively, the DNS synchronizing may be completed through the Secure Dynamic DNS Update.

- Security

Any automatic renumbering scheme has a potential exposure to hijacking at the moment that a new address is announced. Proper network security mechanisms should be employed. Secure Neighbour Discovery (SEND, [RFC3971]), which does not widely deployed, is recommended to replace ND. Alternatively, certain lightweight renumbering specific security mechanism may be developed in the future. DHCPv6 build-in secure mechanisms, like Secure DHCPv6 [I-D.ietf-dhc-secure-dhcpv6] or authentication of DHCPv6 messages [RFC3315] are recommended.

- Miscellaneous

Addresses should not be used to configure network connectivity, such as tunnels. A site or network should also avoid to embed addresses from other sites or networks in its own configuration data. Instead, the Fully-Qualified Domain Names should be used. Thusness, these connectivities can survive after renumbering events. This also applies to host-based connectivities.

Service Location Protocol and multicast DNS with SRV records for service discovery can reduce the number of places that IP addresses need to be configured.

2.2. Considerations/issues for the routine network management

This session describes several recommendations for the routine network management. To adopt these recommendations, a site could be renumbered easier. However, these recommendations are not cost free. They are possible to increase the daily burden of networks. Therefore, only these networks that are expected to be renumbered soon or very frequent should adopt these recommendations with the balance consideration between daily cost and renumbering cost.

- Reduce the address preferred time or valid time or both.

Long-lifetime addresses may cause issues for renumbering events. Particularly, some offline hosts may reconnect back using these addresses after renumbering events. Shorter preferred lifetime with relevant long valid lifetime may get short transition period for renumbering event and avoid address renew too frequent.

- Reduce the DNS record TTL.

The DNS record TTL on the local DNS server should be manipulated to ensure that stale addresses are not cached.

- Reduce the DNS configuration lifetime on the hosts.

Since the DNS server could be renumbered as well, the DNS configuration lifetime on the hosts should also be reduced if renumbering events are expected. The DNS configuration can be done through either ND [RFC6106] or DHCPv6 [RFC3646]. However, DHCPv6 DNS option does not include associated lifetime. It should be updated.

- Reduce the NAT mapping session keepalive time.

Idle NAT mapping session may be keep alive for a long period if the external network addresses space is plenteous and the internal network address architecture is stable. However, renumbering events mean to restructure the internal network address architecture fully or partly. Reducing the NAT mapping session keepalive time may help to tear down the idle TCP connectivities. This will reduce the TCP surviving issue during the renumbering event.

2.3. Considerations/issues during renumbering operation

Renumbering events are not instantaneous events. Normally, there is a transition period, in which both the old prefix and the new prefix are used in the site. Better network design and management, better pre-preparation and longer transition period are helpful to reduce the issues during renumbering operation.

- Transition period

If renumbering transition period is longer than all addresses life, after which the addresses lease expire, each host will automatically pick up its new IP address. In this case, it would

be the DHCP server or Router Advertisement itself that automatically accomplishes client renumbering.

- Network initiative enforced renumbering

If the network has to enforce renumbering before addresses lease expire, the network should initiate enforcement messages, either in Router Advertisement messages or DHCPv6 RECONFIGURE messages.

- DNS record update and DNS configuration on hosts

DNS records should be updated if hosts are renumbered. If the site depends on ISP's DNS system, it should report the new host's DNS records to its ISP. During the transition period, both old and new DNS records are valid. If the TTL of DNS records is shorter than the transition period, administrative operation may not be necessary.

DNS configuration on hosts should be updated if local recursive DNS servers are renumbered. During the transition period, both old and new DNS addresses may co-exist on the hosts. If the lifetime of DNS configuration is shorter than the transition period, name resolving failure may not be reduced to minimum. A notification mechanism may be needed to indicate the hosts that a renumbering event of local re DNS happen or is going to take place recursive.

- Router awareness

In a site with multiple border routers, portion renumbering should be aware by all border routers in order to correctly handle inbound packets. Internal forwarding tables need to be updated.

- Border filtering

In a multihomed site, an egress router to ISP A could normally filter packets with source addresses from other ISPs. The egress router connecting to ISP A should be notified if the egress router connecting to ISP B initiates a renumbering event in order to properly act filter function.

- NAT or tunnel concentrator renumbering

NAT or tunnel concentrator itself might be renumbered. This change should be reconfigured to relevant hosts or router.

2.4. Issues that still remain unsolvable

This section lists a few issues that still remain unsolvable. Some of them may be inherently unsolvable.

- It is not possible to reduce a prefix’s lifetime to below two hours. So, renumbering should not be an unplanned sudden event. This issue could only be avoided by early planning.

- Manual or script-driven procedures will break the completely automatic host renumbering.

- Some environments like embedded systems might not use DHCP or SLAAC and even configuration scripts might not be an option. This creates special problems that no general-purpose solution is likely to address.

- TCP and UDP flows can’t survive at renumbering event at either end.

- Some address configuration data might be widely dispersed and much harder to find, even will inevitably be found only after the renumbering event.

- The embedding of IPv6 unicast addresses into multicast addresses and the embedded-RP (Rendezvous Point) will cause issues when renumbering.

- Changing the unicast source address of a multicast sender might also be an issue for receivers.

- When a renumbering event takes place, entries in the state table of NAT or tunnel concentrator that happen to contain the affected addresses will become invalid and will eventually time out. However, this can be considered as harmless though it takes sources on these devices for a while.

- A site that is listed in a black list can escape that list by renumbering itself. The site itself of course will not initiatively to report its renumbering and the black list may not be able to monitor or discover the renumbering event.

Some of these issues can be considered as harmless or have minimum impacts.

2.5. Issues that need further analysis

This section lists a few issues that still need further analysis. Some of them may be addressed in later version of this document and relocated into other sections. Some of them may be worthy separated document. (Editor note: if all issues addressed, this section should be removed.)

- "Some routers cache IP addresses in some situations. So routers might need to be restarted as a result of site renumbering" [RFC2072]. It seems this caused by individual implementation and only happen on the old type of routers. (Author note: to be removed, if confirmed)

- Multihomed site, using SLAAC for one address prefix and DHCPv6 for another, would clearly create a risk of inconsistent host behaviour and operational confusion.

- It seems so far the renumbering studies only focusing on the individual network using a single prefix. In a large network, a short prefix may be used. The prefix is assigned to be longer and prefixes and delegated to several sub-networks. To make the scenario even more complicated, it may be some sub-networks employ SLAAS while some others are managed by DHCPv6. How to coordinate among these sub-networks to be renumbered together may be worth of analyzing.

- The impact of portion renumbering may need to be analyzed further.

3. Non-network issues

Although we focus on network side, in this section, we also list a few non-network issues. They are out of network providers/operators reach. Therefore, from network perspective, these issues are considered to be unsolvable though they may be solved by OS implementations or application/service implementations. It is out of scope to explore these issues further.

- Any implementation is at risk from renumbering if it does not check that an address is valid each time it opens a new communications session

- Socket API encourages applications to be aware of and to store IP address. And the API relative functions do not return an address lifetime so that applications have no way to know the address is no longer valid.

- "DNS Pining": limits acceptance of server IP address changes for JavaScript security considerations and it may directly damage the ability of applications to deal with renumbering.

- Server applications might need to be restarted when the host they contain is renumbered. In an IPv6 multi-addressed host, server applications need to be able to listen on more than one address simultaneously. Name-based APIs or implementations are recommended.

- When a nameserver is renumbered, the host may not be aware or notified immediately; or even the host is notified, but it still considers the old nameserver is available. The host will at some point find it unavailable. This will cause name resolving failure though these failure may be recoverable.

- Renumbering may cause issues for ACLs or group login services.

4. Requests to extend current protocol

As mentioned in section 2, the following request to extend the current protocols.

- A diagnose function to detect and report the confliction of SLAAC and DHCPv6 address assignment.

- The current protocol needs to be extended if it does not support to combine the forward and reverse DNS updates in a single procedure. (Author note: it seems possible. If so, remove this item.)

- DHCPv6 should be extended to indicate hosts the associated DNS lifetimes when making DNS configuration.

- A lightweight renumbering specific security mechanism may be developed if SEND is too weight to be widely deployed.

- If the issues of coordination among these sub-networks to be renumbered together are confirmed, new interaction may need to be defined to achieve the cooperation.

- A notification mechanism may be needed to indicate the hosts that a renumbering event of local recursive DNS happen or is going to take place recursive.

- NAT or tunnel concentrator configuration procedure may need to be extended to be able to notify the host the renumbering of NAT or tunnel concentrator.

5. Security Considerations

A site that is listed in a black list can escape that list by renumbering itself.

Any automatic renumbering scheme has a potential exposure to hijacking at the moment that a new address is announced. Proper network security mechanisms should be employed. SEND is recommended to replace ND. Alternatively, certain lightweight renumbering specific security mechanism may be developed in the future. DHCPv6 build-in secure mechanisms, like Secure DHCPv6 [I-D.ietf-dhc-secure-dhcpv6] or authentication of DHCPv6 messages [RFC3315] are recommended.

The security updates will need to be made in two stages (immediately before and immediately after the event).

[Editor note: this section needs further work.]

6. IANA Considerations

This draft does not request any IANA action.

7. Acknowledgements

This work is illumined by RFC5887, so thank for RFC 5887 authors, Brian Carpeter, Randall Atkinson and Hannu Flinck. Useful ideas were also illumined by documents from Tim Chown and Fred Bark.

8. Change Log [RFC Editor please remove]

draft-jiang-ipv6-site-renumbering-ps-00, original version, 2011-01-28

9. References

9.1. Normative References

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

[RFC3646] R. Droms, "DNS Configuration options for Dynamic Host Configuration Protocol for IPv6 (DHCPv6) ", RFC3646, December 2003.

[RFC3736] R. Droms, "Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736, April 2004.

[RFC3971] J. Arkko, Ed., J. Kempf, B. Zill, and P. Nikander "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005

[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.

[RFC6106] J. Jeong, Ed., S. Park, L. Beloeil, and S. Madanapalli "IPv6 Router Advertisement Option for DNS Configuration", RFC 6106, November 2011.

9.2. Informative References

[RFC4076] Chown, T., Venaas, S., and A. Vijayabhaskar, "Renumbering Requirements for Stateless Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 4076, May 2005.

[RFC4192] Baker, F., Lear, E., and R. Droms, "Procedures for Renumbering an IPv6 Network without a Flag Day", RFC 4192, September 2005.

[RFC5887] Carpenter, B., Atkinson, R., and H. Flinck, "Renumbering Still Needs Work", RFC 5887, May 2010.

[I-D.ietf-dhc-secure-dhcpv6]
Jiang, S., and Shen S., "Secure DHCPv6 Using CGAs", working in progress.

Author's Addresses

Sheng Jiang
Huawei Technologies Co., Ltd
Huawei Building, No.3 Xinxi Rd.,
Shang-Di Information Industry Base, Hai-Dian District, Beijing 100085
P.R. China
Email: shengjiang@huawei.com

Bing Liu
Huawei Technologies Co., Ltd
Huawei Building, No.3 Xinxi Rd.,
Shang-Di Information Industry Base, Hai-Dian District, Beijing 100085
P.R. China
Email: leo.liubing@huawei.com

INTERNET-DRAFT
Intended Status: Standard Track
Expires: September 8, 2011

B.Liu
S.Jiang
Huawei Technologies Co., Ltd

March 7, 2011

SLAAC/DHCPv6 conflicts diagnostic during site renumbering
draft-liu-ipv6-renum-conflicts-00.txt

Status of this Memo

Copyright and License Notice

Abstract

While an IPv6 site is being renumbered, both DHCPv6 and ND may be used to reconfigure the host addresses. This may cause potential address configuration conflicts during renumbering procedure. The issue mainly include two situations: a) Address configuration method conflict, which means a host receives a new prefix comes from another address configuration protocol. b) Address prefix conflict, a host receives both DHCPv6 and ND address configuration messages which assign different address prefixes. This documents analyzes the conflict issues and proposes a conflict report mechanism for hosts to report the conflicts to DHCPv6 servers.

Table of Contents

1 Introduction

1.1 IPv6 renumbering

"Renumbering" is an event of changing in IP addressing information associated with a host or subnet [RFC1900]. [RFC5887] and [I-D.chown-v6ops-renumber-thinkabout] described numerous reasons why such sites might need to renumber in a planned fashion, for example, change of site topology, change of service provider etc.

[RFC4192] provided a general procedure of renumbering in an IPv6 network, which is achieved by changing address prefix. Before the old prefix is withdrawn, the hosts are assigned a new prefix. Both the old and the new prefixes may be usable till the new prefix is stable in the site systems, such as DNS, ACL .etc. Then the old prefix will be withdrawn. The transition periods are variable according to different network management settings.

[RFC4192] also mentioned several methods to reconfigure addresses while renumbering:

- Stateful address configuration through Dynamic Host Configuration Protocol for IPv6 (DHCPv6) protocol [RFC3315]
- Stateless address configuration through Neighbor Discovery Protocol[RFC4861] (SLAAC) [RFC4862]
- Manual configuration

This document focuses on the address reconfiguration problem and there is a specific issue of IPv6 site renumbering described as the following.

1.2 DHCPv6/ND address configuration conflict

Both of the DHCPv6 and ND protocols have IP address configuration function. They are suitable for different scenarios respectively. During renumbering, the SLAAC-configured hosts can reconfigure IP addresses by receiving ND Router Advertisement (RA) messages containing new prefix information. The DHCPv6-configured hosts can reconfigure addresses by initialing RENEW sessions when the current addresses' lease time is expired or receiving the reconfiguration messages initialed by the DHCPv6 servers.

But DHCPv6 and ND address configuration may overlap and cause conflict on a host. The issue includes two situations:

- A DHCPv6-configured host receives RA messages containing new

prefix

Ideally, hosts in a DHCPv6-managed network should not receive any ND address configuration messages to avoid potential confusion. But some factors may cause this happen. For example, a sub-net of a DHCPv6-managed network may be mis-configured to use SLAAC for address configuration; or a DHCPv6-managed network contains hosts (some specific types of Apple Mac computers, e.g.) that don't support DHCPv6 as the default so that SLAAC will be used along with DHCPv6 as a necessary supplement.

There are no standards specifying what approach should be taken by a DHCPv6-configured host when it receives RA messages containing new prefix. It depends on the operation system of the host and cannot be predicted or controlled by the network. If the host accepts the new prefix in RA, it may violet the DHCPv6-managed policies. But if it ignores the RA messages and there are no DHCPv6 reconfiguration messages received either, the renumbering would fail. What is worse, the host may even receive both the RA messages and DHCP-v6 reconfiguration messages and finds the prefixes in the two protocols are different. This means serious network configuration error occurring.

- A SLAAC-configured finds DHCPv6 is in use

[RFC5887] and [I-D.jiang-ipv6-site-renum-guideline] mentioned RA message of ND protocol contains a "Managed Configuration" flag to indicate DHCPv6 is in use. But it is unspecified what behavior should be taken when the host receives RA messages with "M" set to 1. The gap of standard will cause ambiguous host behavior because it depends on the operation system of the host.

The host may start a DHCPv6 session and receives the DHCPv6 address configuration. It is also possible that the host finds the DHCPv6 assigned prefix is different from the prefix in the RA messages, which means there is a serious network configuration error.

Another possibility is that the host may receive no response from any DHCPv6 servers, which means the DHCPv6 service is not available and the "Managed Configuration" flag was mis-configured.

These potential conflicts described above need to be addressed. This document proposes a report mechanism for hosts to report the conflicts to DHCPv6 servers. (The mis-configured "Managed Configuration" flag issue described above is not addressed in this document because there are no DHCPv6 servers available in that circumstance. Whether it needs to report the conflict to routers or some other servers such as network management systems needs further

study.)

2 Terminolog

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].

3 Conflict Report Mechanism

As analyzed above, while renumbering, hosts received address configuration messages (either ND or DHCP protocol messages), if the messages conflict with existing/previous configuration mechanism, hosts report address configuration policy conflict information to the network. And then they accept the address configuration indication from the network.

The conflicts can be outlined as two types:

- Prefix conflict, which means prefixes in DHCPv6 and ND messages are different.

- Address configuration method conflict, which means a host receives a new prefix comes from another address configuration protocol.

There are several approaches of the mechanism as the following clauses.

3.1 Host behavior

For the DHCPv6-configured hosts, it assumes that they will monitor the RA messages after being configured by DHCPv6.

For the SLAAC-configured hosts, it assumes that the hosts initially explored the DHCPv6 servers based on having already chosen DHCPv6 as high priority of address configuration protocol when it finds the "Managed Configuration" flag is set.

3.2 Conflict Report Trigger

Rules for the hosts to trigger conflict reports are as the following:

- Prefix conflict trigger, a host will trigger the report when it finds the prefixes are different in DHCPv6 and ND messages.

- Address configuration method conflict trigger, a DHCPv6-managed host receives a new prefix comes from RA messages.

3.3 DHCPv6 Reconfiguration Conflict Options

New DHCPv6 options could be defined respectively for the clients to report conflicts to servers and for servers to response according to analysis of the reported conflict details.

- Option_ReconfigConflict_Report

It is possible to include this option into the renew message. The content of the option could be:

a) New available address prefix (prefix in RA e.g.).
b) Serious error of prefix conflict.

- Option_ReconfigConflict_Response

DHCPv6 server could response as the following possibilities according to information got from the option:

a) Directly assigns new addresses to the hosts who report conflicts.
b) Indicates hosts to make SLAAC according to RA messages received.
c) Report the prefixes conflict to network management system.

3.4 Report processing by DHCPv6 server

When a DHCPv6 server receives the conflict reports, it should analyze the report and decides whether to forward the report to relative network management systems or indicate what approach should be taken to the hosts through DHCPv6 messages defined in section 3.3, however, the analysis processing of DHCPv6 servers is not in the scope of this memo.

4 Security Considerations

This document doesn’t provide additional security considerations for IPv6 site renumbering more than [RFC5887], [RFC4192] and other relative documents.

For the conflict report mechanism, there is a potential threat that any malicious host can fake conflict reports to DHCPv6 servers which may disturb the network manager when the report is prefix conflict, however, it cannot directly break the availability of the network.

5 IANA Considerations

This document requests IANA to assign new DHCPv6 option number.(TBD)

6 Acknowledgements

This document inherits various previous work. Thanks for Brian Carpenter, Randall Atkinson, Hannu Flinck, Fred Baker, Eliot Lear, Ralph Droms, Tim J. Chown, Mark K. Thompson, Alan Ford, and Stig Venaas.

7 References

7.1 Normative References

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

[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.

7.2 Informative References

[RFC5887] Carpenter, B., Atkinson, R., and H. Flinck,"Renumbering Still Needs Work", RFC 5887, May 2010.

[RFC4192] Baker, F., Lear, E., and R. Droms, "Procedures for Renumbering an IPv6 Network without a Flag Day", RFC 4192, September 2005.

[I-D.jiang-ipv6-site-renum-guideline] Jiang, S., and Liu B., "IPv6 Site Renumbering Guidelines

and Further Works", working in progress.

[I-D.chown-v6ops-renumber-thinkabout]
Chown, T., and Thompson, M., "Things to think about when Renumbering an IPv6 network", September 2006.

Author's Addresses

Bing Liu
Huawei Technologies Co., Ltd
Huawei Building, No.3 Xinxi Rd.,
Shang-Di Information Industry Base, Hai-Dian District, Beijing 100085
P.R. China
Email: leo.liubing@huawei.com

Sheng Jiang
Huawei Technologies Co., Ltd
Huawei Building, No.3 Xinxi Rd.,
Shang-Di Information Industry Base, Hai-Dian District, Beijing 100085
P.R. China
Email: jiangsheng@huawei.com