INTERNET DRAFT
<draft-huitema-ipngwg-dialondemand-00.txt>
Expires May 14, 2002

C. Huitema
Microsoft
November 14, 2001

IPv6 Dial-up on Demand

Status of this memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

This document is an Internet-Draft. 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.

Abstract

In an IPv6 home network, there is a possibility that the residential gateway, learns a different IPv6 prefix after each dial-up; attempts to communicate using old address would thus fail. This implies that we should not attempt to trigger dial-up by simply sending the first packet of a connection. Our solution to solve this issue has two parts: first, a convention that dial-up routers should immediately assign a "preferred lifetime" of zero to the global prefix learnt over a dial-up connection, after they hang up the connection; second, a rule that nodes who want to start a connection to the outside world (global scope address) notice that there is no available prefix with a non-zero preferred lifetime, and send a "connectivity test" ICMP message to the destination; the payload will contain information such as TCP ports needed for the gateway to evaluate whether to dial-up or not.

1 Introduction

Dial-up on demand is a classic feature of many networking configurations, in which a node connected to the Internet by a modem only dials-up the connection when there is actual traffic to send, e.g. when an applications attempts to initiate a TCP connection. In a classic implementation, the dialup results in the establishment of a phone or ISDN connection to a network access point, then the

synchronization of modems, the establishment of a PPP connection and of an IPv6 context over the PPP connection. At this stage, the node

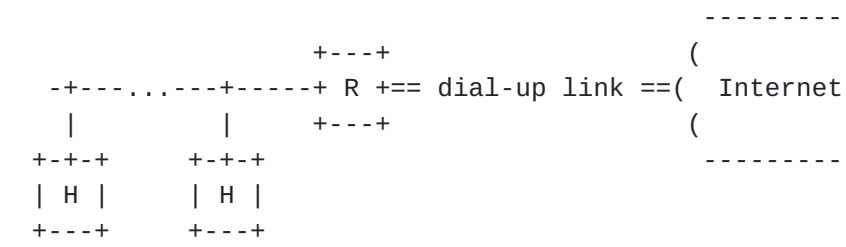
will normally send a "router solicit" to the access point, and receive a "router advertisement"; the node will then configure at least one IPv6 address based on the advertised prefix for use in the TCP connection. Once the address has been allocated, the TCP connection will proceed. The node will monitor usage of the connection, and disconnect it when some heuristic determines that it is not needed. This feature is widely used when connections are paid "by the minute", which is common in several countries.

We are concerned with a variation of this scenario, when the node establishing the dial-up on demand connection is in fact a router. In this variation, the router will trigger dial-up when it somehow detects that a node on the local network wants to communicate; it will then acquire a prefix by an RS/RA exchange with the access point, and publish a derived prefix on the local network. The local node will at that point configure an IPv6 address based on the prefix, and will be able to send traffic. The big issue with this scenario is that it requires some synchronization between dial-up on the router and initiation of a TCP connection on the host.

2 Notation

2.1 Single link domain

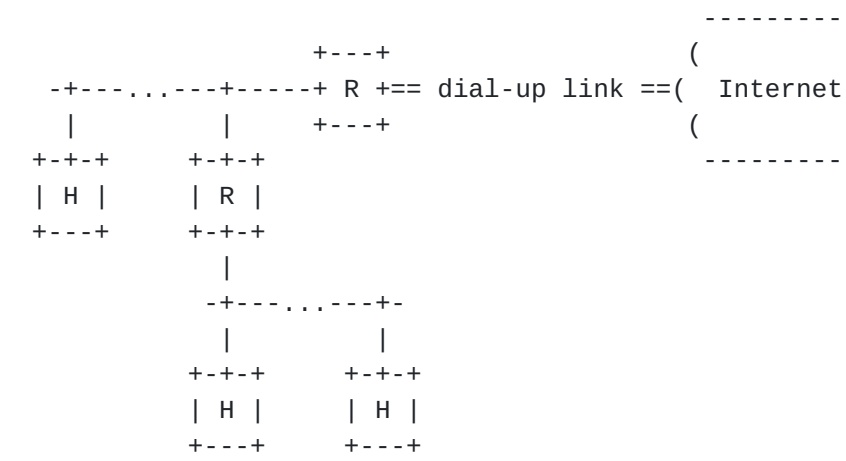
We designate as "single link domain" a configuration in which a domain is composed of a single subnet, connected by a router to the Internet:



The subnet may be composed of one or several segments, using technologies such as Ethernet bridging or Ethernet Switching.

2.2 Multiple link domain

We designate as "multiple link domain" a configuration in which a domain is composed multiple links connected by routers, one of which manages the connection to the Internet:



The domain will be composed of several logical links, using technologies that cannot be easily bridged, e.g. 1394 and Ethernet. In such domains, we expect that routers will be coordinated, using for example the router renumbering [RFC2894].

2.3 Dialup router

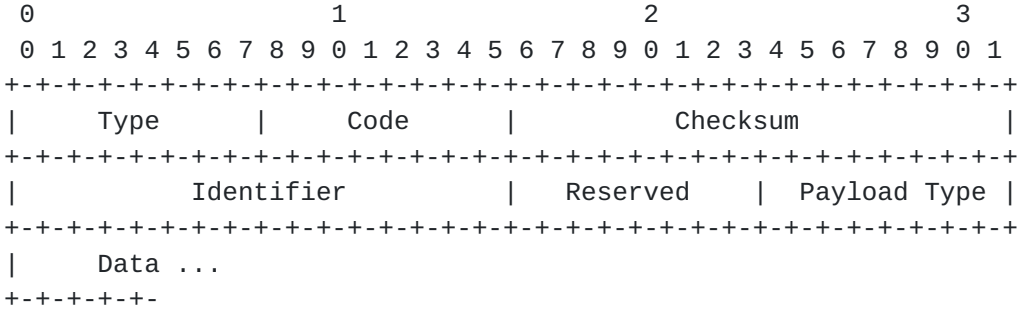
In a single link or multiple link domain, the dialup router is the router connecting the domain to the Internet via a dialup link.

3 Description of the solution

3.1 New ICMP messages

The dial-up on demand solution uses two new ICMP messages, the Connectivity Test and the Connectivity Reply. These two messages are variation of the informational ICMP messages "Echo Request" and "Echo Response."

3.1.1 Connectivity Test



IPv6 Fields:
Destination Address: the address towards which connectivity is being tested.

ICMPv6 Fields:

Type 128

Code 1

Identifier An identifier to aid in matching Connectivity Replies to this Connectivity Test. May be zero.

Payload Type

The payload type that the application intend to use to communicate with the destination. May be zero if unspecified.

Data Zero or more octets of arbitrary data, normally set to mimic the payload that the application intend to send, for example a TCP or UDP header if the payload type is set to indicate TCP or UDP.

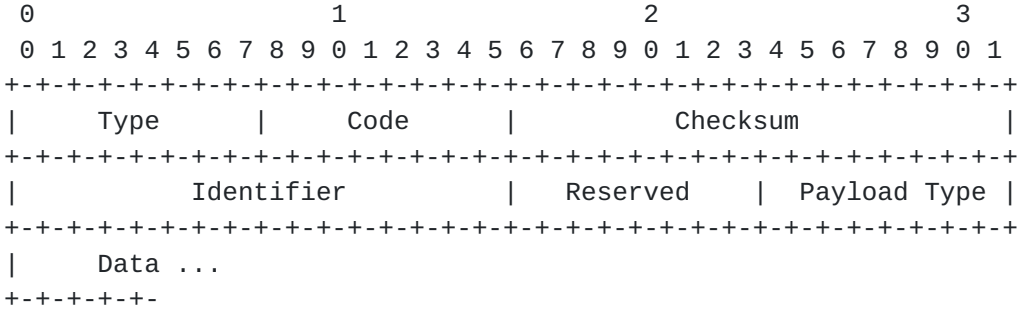
Description

Every node SHOULD implement an ICMPv6 Connectivity responder function that receives Connectivity Tests and sends corresponding Connectivity Replies. A node SHOULD also implement an application-layer interface for sending Connectivity Tests and receiving Connectivity Replies, for dial-up on demand purposes.

Upper layer notification

Connectivity Test messages MAY be passed to processes receiving ICMP messages.

3.1.2 Connectivity Reply Message



IPv6 Fields:

Destination Address

Copied from the Source Address field of the invoking Connectivity Test packet.

ICMPv6 Fields:

Type 129

Code 1

Identifier The identifier from the invoking Connectivity Test message.

Payload Type The payload type from the invoking Connectivity Test message.

Data The data from the invoking Connectivity Test message.

Description

Every node SHOULD implement an ICMPv6 Connectivity responder function that receives Connectivity Tests and sends corresponding Connectivity Replies. A node SHOULD also implement an application-layer interface for sending Connectivity Tests and receiving Connectivity Replies, for dial-up on demand purposes.

The source address of an Connectivity Reply sent in response to a unicast Connectivity Test message MUST be the same as the destination address of that Connectivity Test message.

A Connectivity Reply SHOULD NOT be sent in response to a Connectivity Test message sent to an IPv6 multicast address.

The data received in the ICMPv6 Connectivity Test message MUST be returned entirely and unmodified in the ICMPv6 Connectivity Test message.

Upper layer notification

Connectivity Reply messages MUST be passed to the process that originated a Connectivity Test message. It may be passed to processes that did not originate the Connectivity Test message.

3.2 Host behavior

Hosts located behind a dialup router that handles static prefixes will not perceive any difference compared to a regular IPv6 set up. Hosts located behind a dialup router that handle dynamic prefixes may experience the effect of a renumbering event.

Hosts will notice that the configuration is not standard when they attempt to perform "source address selection" and observe that there is no available address in the right scope, or if addresses exist that these addresses have a null preferred lifetime. In this situation, the host should send a "Connectivity Test" message, configured as follow:

IPv6 Fields:

Destination Address

The destination address that the application wishes to use.

Source Address
The highest scope address that the host can use on the selected interface, typically a site local address, possibly a link local address.

ICMPv6 Fields:

Payload Type The payload type that the application intends to use, or zero if it unknown.

Data A typical payload that the application intend to send, for example a TCP or UDP header if the payload type is set to indicate TCP or UDP. No data is sent if the application characteristics are unknown.

The application will then wait for a response from the destination, or for an ICMP error message from a router, or for a timer.

If the destination is reachable, it will respond with a connectivity reply, and the application MUST proceed with the source address used in the ICMP message. If the destination is reachable but does not understand the new ICMP message code, it will normally reply with an error message; if the error message specifies a code other than "destination unreachable", the application SHOULD proceed with the source address used in the ICMP message.

If the host does not receive any reply after a timer, it MAY repeat the Connectivity Test a limited number of times, after which the application SHOULD proceed with the source address used in the ICMP message.

If the destination is not reachable, for example because it is on the other side of a site boundary and packets sourced from a site local address cannot be forwarded, the host will normally receive an "ICMP error: destination unreachable" message. The host should then wait for the arrival of a router advertisement before configuring an address and completing the source address selection. If the host fails to receive a router advertisement in a delay compatible with the establishment of a dial-up link, it SHOULD send a router solicitation to the all routers multicast address.

Once source address selection is completed, the host communication will proceed normally, e.g. with a TCP SYN, or a UDP packet.

3.3 Advertising of dynamic prefixes by the dialup router

This section describes the behavior of routers that manage a dynamic

prefix. When these routers establish a connection, they acquire a prefix by means of an RS/RA exchange with a network access point.

The prefix or prefixes learned after dial-up may or may not be the same than the prefixes learned in the previous connection; when they are not the same, we have a renumbering event.

3.3.1 Dial-up behavior

The router managing a dynamic prefix dialup link will trigger dial-up based on heuristics such as the presence of packets queued for transmission on the dial-up interface. The router MAY implement a dial-up policy based on the type of packets queued in front of the interface.

The router may receive connectivity test messages bound to destination that are "on the other side" of a dial-up link; the router will immediately respond to such packets by an ICMP error indicating that the destination is unreachable. The router receiving such messages MAY implement a dial-up policy based on parameters such as the payload type indicated in the ICMP header, or the port numbers present in the first bytes of the ICMP data when the payload type is set to TCP or UDP. If the policy conditions are met, the router should trigger the dial-up of the link.

When a router successfully establishes a dial-up connection, it should send a router advertisement to all nodes on the link, advertising the dynamic prefix corresponding to the link. The preferred life time of the prefix should have a non null value.

In case of a renumbering event, the advertisement should either omit the prefixes obtained in previous connections, or advertise a null valid lifetime for these prefixes.

After establishing the connection, the router should perform ingress filtering, i.e. check that the source address in any queued packet is consistent with the address prefixes advertised for the connection. A packet that doesn't have a valid address prefix should be discarded; the router may send an ICMP message indicating that the packet has been rejected for an administrative reason.

3.3.2 Preferred lifetime of dynamic prefixes and hang-up

A router should not hang-up the connection before a delay corresponding to the preferred lifetime has elapsed since the last router advertisement announcing a dynamic prefix corresponding to that connection.

3.3.3 Deprecating dynamic prefixes on hang-up

When a router hangs up a dial-up connection, it should send a router advertisement to all nodes on the link, which specifies that the preferred lifetime of the dynamic prefixes learned from the Internet

access point is now null.

3.3.4 Valid lifetime of dynamic prefixes

When a router advertises a prefix associated to a connection, it should associate to that prefix a valid lifetime. This lifetime will be set to a conventional infinite value if there is a static prefix assignment associated to the connection. It should be set to a short value otherwise.

3.3.5 Responding to router solicitations

When a router receives a router solicitation and a connection is currently established, the router should advertise the dynamic prefix corresponding to the connection with a non null preferred lifetime.

When a router receives a router solicitation and a previously established connection has been hang-up, the router may advertise the dynamic prefix corresponding to the connection with a null preferred lifetime.

If no connection has been previously established, or if a delay larger or equal to the valid lifetime of the dynamic prefix has elapsed since the last advertisement of the dynamic prefix acquired on a previously established connection, the router should not advertise a prefix attached to the connection.

3.3.6 Multiple link configurations

In multiple link configurations, the router should use the "router renumbering" mechanism to inform other routers in the domain of any new prefix, or any change in the preferred lifetime or valid lifetime of current prefixes.

3.4 Handling of static prefixes by the dial-up router

This section describes the behavior of routers when a static prefix has been assigned to the local domain.

3.4.1 Advertising the local prefix

When a router has been configured to advertise a static prefix, it advertises this prefix in router advertisements sent on the local link. The preferred life time of the prefix should have a non null value, irrespective of the status of the local connection.

3.4.2 Handling of the dial-up connection

The router which manages a statically allocated prefix will dial-up or hang-up the line based on local heuristics, such as the number of packets queuing in front of the connection, or the time elapsed

since the last packet was sent...

3.4.3 Responding to router solicitations

When a router that manages a static prefix receives a router solicitation, the router should advertise the dynamic prefix corresponding to the connection with a non null preferred lifetime, regardless of the dial-up status of the connection.

4 Discussion of the solution

4.1 Dialup policies & triggering packet

In single node operation, dial-up is normally triggered when an application attempts to set up a TCP connection to a remote destination, or more generally when an application attempts to send a packet towards a remote destination. It is fairly common to implement dialup control policies that specify which applications can actually trigger dialup, e.g. make it automatic when looking for a web page, but not allow it for a network news download. The policies generally operate in terms of address ranges and port numbers.

In order to implement the same kind of policies in the dial-up router, we must be able to forward a packet that contains all the fields used in the policy decision: destination address, protocol type, port numbers, etc. On the other hand, we don't want to send the actual packets, since there will cases in which the destination is actually reachable, despite the mismatch in address scoped; sending the actual packet would result in confusion. The simplest way is to send as triggering packet an ICMP message that is sent to the actual destination, and contains sufficient parameters to enable policy decision, i.e. a payload type and a carbon copy of a regular packet payload.

4.2 Why sourcing the triggering packets from a local scope address?

Triggering packets can meet one of three fates: they may be dropped because the dial-up procedure took too long; they may be forwarded if the prefix learned after the dial-up happens to match the prefix used by the source; or they may be dropped because the dial-up procedure results in a "prefix renumbering". In order to minimize uncertainties and obtain a robust solution, we use as source address a site local address whenever one is available, a link local address when this is not possible. This ensures that in most cases the triggering packet will actually be dropped by the dialup router, since site local packets should not be forwarded on the global Internet, thus eliminating the risk of duplicate packets creating unwanted state at the correspondent node.

We should note that there is a particular case in which a site scopes include two domains linked by a dialup link, such as for

example a small office linked to a main campus by an ISDN

connection. In such cases, however, we can assume that the address assignment are permanent; we thus fall back to the case of "static address operation."

4.3 Why not use UPNP?

A possible solution would be to use UPNP as control protocol in order to trigger dial-up. However, this solution would entail using the user mode UPNP service, which is a source of delays and instabilities.

4.4 Why not just send a SYN?

A simpler solution would be to not bother the hosts at al, and let them just send a natural packet, e.g. a SYN. This is indeed what unsophisticated hosts will do, no matter what we specify. The problem with this procedure is that the dial-up may trigger a renumbering event. In that case, not only will the initial SYN be dropped, but also the successive repetitions; the application will be notified of the failure after a rather long delay, and may or may not decide to restart the connection.

Waiting for a router advertisement after dialup only creates a minimal additional delay: a round trip on a local network, compared to a dialup delay. It is worthwhile to wait this delay and have a robust execution.

5 Security Considerations

It could be argued that dial-up on demand creates some security issues. However, the particular procedures detailed here have a minimal security impact.

6 IANA Considerations

This document does not call for an IANA action.

7 Copyright

The following copyright notice is copied from RFC 2026 [Bradner, 1996], Section 10.4, and describes the applicable copyright for this document.

8 Intellectual Property

The following notice is copied from RFC 2026 [Bradner, 1996], Section 10.4, and describes the position of the IETF concerning intellectual property claims made against this document.

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use other technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

9 Acknowledgements

Several of my colleagues contributed to this document.

10 References

[RFC2894] M. Crawford, "Router Renumbering for IPv6", RFC 2894, August 2000.

11 Author's Addresses

Christian Huitema
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-6399

Email: huitema@microsoft.com