DHC Working Group
INTERNET-DRAFT
Category: Proposed Standard
<draft-ietf-dhc-dna-ipv4-18.txt>
1 December 2005

Bernard Aboba
Microsoft Corporation
James Carlson
Sun Microsystems
Stuart Cheshire
Apple Computer

Detecting Network Attachment in IPv4 (DNAv4)

This Internet-Draft will expire on June 10, 2006.

Copyright Notice

Abstract

The time required to detect movement between networks, and to obtain (or continue to use) an IPv4 configuration may be significant as a fraction of the total handover latency in moving between points of attachment. This document synthesizes from experience in the deployment of hosts supporting ARP, DHCP, and IPv4 Link-Local addresses a set of steps known as Detecting Network Attachment for IPv4 (DNAv4), in order to decrease the handover latency in moving between points of attachment.

Table of Contents

1. Introduction

The time required to detect movement between networks and to obtain (or continue to use) an operable IPv4 configuration may be significant as a fraction of the total handover latency in moving between points of attachment.

This document synthesizes from experience in the deployment of hosts supporting ARP [RFC826], DHCP [RFC2131], and IPv4 Link-Local addresses [RFC3927] a set of steps known as Detecting Network Attachment for IPv4 (DNAv4). DNAv4 optimizes the (common) case of re-attachment to a network that one has been connected to previously by attempting to re-use a previous (but still valid) configuration, reducing the re-attachment time to a few milliseconds on LANs. Since this procedure is dependent on the ARP protocol, it is not suitable for use on media that do not support ARP.

1.1. Applicability

DHCP is an effective and widely adopted mechanism for a host to obtain an IP address for use on a particular network link, or to re-validate a previously obtained address via DHCP's INIT-REBOOT mechanism [RFC2131].

When obtaining a new address, DHCP specifies that the client SHOULD use ARP to verify that the offered address is not already in use. The process of address conflict detection [ACD] can take as much as seven seconds. In principle this time interval could be shortened, with the obvious trade-off: the less time a host spends waiting to see if another host is already using its intended address, the greater the risk of inadvertent address conflicts.

Where the client successfully re-validates a previously obtained address using the INIT-REBOOT mechanism, the DHCP specification does not require the client to perform address conflict detection, so this seven-second delay does not apply. However, the DHCP server may be slow to respond, or may be down and not responding at all, so hosts could benefit from having an alternative way to quickly determine that a previously obtained address is valid for use on this particular link.

When the client moves between networks, the address re-validation attempt may be unsuccessful; a DHCPNAK may be received in response to a DHCPREQUEST, causing the client to restart the configuration process by moving to the INIT state. If an address previously obtained on the new network is still operable, DNAv4 enables the host to quickly confirm the new configuration, bypassing restart of the configuration process as well as conflict detection.

The alternative mechanism specified by this document applies when a host has a previously-allocated DHCP address, which was not returned to the DHCP server via a DHCPRELEASE message, and which still has time remaining on its lease. In this case, the host may determine whether it has re-attached to the logical link where this address is valid for use, by sending a unicast ARP Request packet to a router previously known for that link (or in the case of a link with more than one router, by sending one or more unicast ARP Request packets to one or more of those routers).

The use of unicast ARP has a number of benefits. One benefit is that unicast packets impose less burden on the network than broadcast packets, particularly on 802.11 networks where broadcast packets may be sent at rates as low as 1 Mb/sec. Another benefit is that if the host is not on the link it hoped to find itself on, a broadcast ARP Request could pollute the ARP caches of peers on that link. When using private addresses [RFC1918] another device could be legitimately using the same address, and a broadcast ARP Request could disrupt its communications, causing TCP connections to be broken, and similar problems. By using a unicast ARP packet instead, addressed to the MAC address of the router the host is expecting to find, if the host is not on the expected link, then there will be no device with that MAC address, and the ARP packet will harmlessly disappear into the void without doing any damage.

These issues that define the applicability of DNAv4 lead us to a number of conclusions:

- DNAv4 is a performance optimization. Its purpose is to speed up a process that may require as much as a few hundred milliseconds (DHCP INIT-REBOOT), as well as to reduce multi-second conflict detection delays when a host changes networks.

- As a performance optimization, it must not sacrifice correctness. In other words, false positives are not acceptable. DNAv4 must not conclude that a host that has returned to a previously-visited link where it has an operable IP address if this is not in fact the case.

- As a performance optimization, false negatives are acceptable. It is not an absolute requirement that this optimization correctly recognize a previously-visited link in all possible cases. For example, if a router ignores unicast ARP Requests then DNAv4 will not be able to detect that it has returned to the same link in future. This is acceptable because the host still operates correctly as it did without DNAv4, just without the performance benefit. Users and network operators who

desire the performance improvement offered by DNAv4 can upgrade their routers to support DNAv4.

- As a performance optimization, where DNAv4 fails to provide a benefit, it should add little or no delay compared to today's DHCP processing. In practice, this implies that DHCP processing needs to proceed in parallel. Waiting for DNAv4 to fail before beginning DHCP processing can greatly increase total processing time, the opposite of the desired effect.

- Trials are inexpensive. DNAv4 performs its checks using small unicast packets. An IPv4 ARP packet on Ethernet is just 42 octets, including the Ethernet header. This means that the cost of an unsuccessful attempt is small, whereas the cost of a missed opportunity (having the right address available as a candidate and choosing not to try it for some reason) is large. As a result, the best strategy is often to try all available candidate configurations, rather than trying to determine which candidates, if any, may be correct for this link, based on heuristics or hints. For a heuristic to usefully eliminate a configuration from the candidate list, it has to (a) be fast and inexpensive to compute, compared to sending a 42-octet unicast packet, and (b) have high probability of not falsely eliminating a candidate configuration that could be found to be the correct one.

- Time is limited. If DNAv4 is to be effective in enabling low latency handoffs, it needs to complete in less than 10 ms. This implies that any heuristic used to eliminate candidate configurations needs to take at most a few milliseconds to compute. This does not leave much room for heuristics based on observation of link layer or Internet layer traffic.

1.2. Requirements

In this document, several words are used to signify the requirements of the specification. 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 "Key words for use in RFCs to Indicate Requirement Levels" [RFC2119].

1.3. Terminology

This document uses the following terms:

ar$sha
: ARP packet field: Sender Hardware Address [RFC826]. The hardware (MAC) address of the originator of an ARP packet.

ar$spa
: ARP packet field: Sender Protocol Address [RFC826]. For IP Address Resolution this is the IPv4 address of the sender of the ARP packet.

ar$tha
: ARP packet field: Target Hardware Address [RFC826]. The hardware (MAC) address of the target of an ARP packet.

ar$tpa
: ARP packet field: Target Protocol Address [RFC826]. For IPv4 Address Resolution, the IPv4 address for which one desires to know the hardware address.

DHCP client
: A DHCP client or "client" is an Internet host using the Dynamic Host Configuration protocol (DHCP) [RFC2131] to obtain configuration parameters such as a network address.

DHCP server
: A DHCP server or "server" is an Internet host that returns configuration parameters to DHCP clients.

Link
: A communication facility or medium over which network nodes can communicate. Each link is associated with a minimum of two endpoints. Each link endpoint has a unique link-layer identifier.

Link Down
: An event provided by the link layer that signifies a state change associated with the interface no longer being capable of communicating data frames; transient periods of high frame loss are not sufficient. DNAv4 does not utilize "Link Down" indications.

Link Layer
: Conceptual layer of control or processing logic that is responsible for maintaining control of the data link. The data link layer functions provide an interface between the higher-layer logic and the data link. The link layer is the layer immediately below IP.

Link Up
: An event provided by the link layer that signifies a state change associated with the interface becoming capable of communicating data frames.

Point of Attachment
: The link endpoint on the link to which the host is currently connected.

Routable address
In this specification, the term "routable address" refers to any IPv4 address other than an IPv4 Link-Local address. This includes private addresses as specified in "Address Allocation for Private Internets" [RFC1918].

Operable address
In this specification, the term "operable address" refers to either a static IPv4 address, or an address assigned via DHCPv4 which has not been returned to the DHCP server via a DHCP RELEASE message, and whose lease has not yet expired.

2. Overview

On connecting to a new point of attachment, the host responds to a "Link Up" indication from the link layer by carrying out the DNAv4 procedure.

For each network that it connects to, it is assumed that the host saves to stable storage the following parameters:

[1] The IPv4 and MAC address of one or more test nodes on the network

[2] The IPv4 configuration parameters, including the DHCP client identifier, assigned address and lease expiration time

From the set of networks which have operable IPv4 address(es) associated with them, the host selects a subset, and attempts to confirm the configuration for each network, using the reachability test described in Section 2.1.

For a particular network, the host SHOULD use the addresses of local routers (preferably default gateways) as its test nodes. If more than one address is known, those addresses may be tested in parallel. In order to ensure configuration validity, the host SHOULD only configure routes for which the next hop address passes the reachability test. Other routes SHOULD be re-learned.

DNAv4 does not significantly increase the likelihood of an address conflict. The reachability test is only carried out for a network when the host has previously completed conflict detection as recommended in Section 2.2 of the DHCP specification [RFC2131], and obtained an operable IPv4 configuration on that network. Restrictions on sending ARP Requests and Responses are described in Section 2.1.1.

One case where DNAv4 does increase the likelihood of an address conflict is when:

- o a DHCP server hands out an address lease
- o the host with that lease leaves the network
- o the DHCP server is power-cycled, or crashes and is rebooted
- o the DHCP server, having failed to save leases to stable storage, assigns that same address to another host
- o the first host returns, and having a still-valid lease with time remaining, proceeds to use its assigned address, conflicting with the new host that is now using that same address

While Section 4 of the DHCP specification [RFC2131] assumes that DHCP servers save their leases in persistent storage, almost no consumer-grade NAT gateway does so. Short DHCP lease lifetimes can mitigate this risk, though this also limits the operable candidate configurations available for DNAv4 to try.

2.1. Reachability Test

The host skips the reachability test for a network if any of the following conditions are true:

[a] The host does not have an operable routable IPv4 address on that network. In this case, the reachability test cannot confirm that the host has an operable routable IPv4 address, so completing the reachability test would serve no purpose.

[b] The host does not know the addresses of any test nodes on that network. In this case, insufficient information is available to carry out the reachability test.

[c] If DHCP authentication [RFC3118] is configured. The reachability test utilizes ARP which is insecure. Hosts that have been configured to attempt DHCP authentication SHOULD NOT utilize the reachability test. Security issues are discussed in Section 4.

[d] The contents of the DHCP Client Identifier option the client used to obtain the candidate configuration is different from the DHCP Client Identifier option the client intends to present on the interface in question. In this case, it is anticipated that a DHCP server would NAK any request made by the client to acquire or extend the candidate configuration, since the two interfaces are presenting differing identities.

If the reachability test is successful, the host SHOULD continue to use the operable routable IPv4 address associated with the confirmed network, without needing to re-acquire it. Once a valid reachability test response is received, validation is complete, and additional responses should be discarded.

If a DHCPv4 client is operational, it is RECOMMENDED that the host attempt to obtain IPv4 configuration via DHCPv4 in parallel with the reachability tests, with the host using the first answer returned. This ensures that the DNAv4 procedure will not result in additional delay in the case where reachability test(s) fail, or where sending a DHCPREQUEST from the INIT-REBOOT state, as described in Section 3.2 and 4.3.2 of the DHCP specification [RFC2131], completes more quickly than the reachability test(s).

In situations where both DNAv4 and DHCP are used on the same link, it is possible that the reachability test will complete successfully, and then DHCP will complete later with a different result. If this happens, the implementation SHOULD abandon the reachability test results and use the DHCP result instead, unless the address confirmed via the reachability test has been manually assigned (see Section 2.4).

Where the reachability test does not return an answer, this is typically because the host is not attached to the network whose configuration is being tested. In such circumstances, there is typically little value in aggressively retransmitting reachability tests that do not elicit a response.

Where DNAv4 and DHCP are tried in parallel, one strategy is to forsake reachability test retransmissions, allowing only the DHCP client to retransmit. In order to reduce competition between DNAv4 and DHCP retransmissions, a DNAv4 implementation that retransmits may utilize the retransmission strategy described in Section 4.1 of the DHCP specification [RFC2131], scheduling DNAv4 retransmissions between DHCP retransmissions.

If a response is received to any reachability test or DHCP message, pending retransmissions are canceled. It is RECOMMENDED that a DNAv4 implementation retransmit no more than twice. To provide damping in the case of spurious "Link Up" indications, it is RECOMMENDED that the DNAv4 procedure be carried out no more than once a second.

2.1.1. Packet Format

The reachability test is performed by sending a unicast ARP Request. The host MUST set the target protocol address (ar$tpa) to the IPv4 address of the node being tested, and the sender protocol address

field (ar$spa) to its own candidate IPv4 address. The ARP Request MUST use the host MAC address as the source, and the test node MAC address as the destination. The host includes its MAC address in the sender hardware address field (ar$sha), and sets the target hardware address field (ar$tha) to 0.

If a valid ARP Reply is received, the MAC address in the sender hardware address field (ar$sha) in the ARP Reply is matched against the target hardware address field (ar$tpa) in the ARP Request, and the IPv4 address in the sender protocol address field (ar$spa) of the ARP Reply is matched against the target protocol address field (ar$tpa) in the ARP Request. If a match is found, then the host continues to use that IPv4 address, subject to the lease re-acquisition and expiration behavior described in Section 4.4.5 of the DHCP specification [RFC2131].

The risk of an address conflict is greatest when the host moves between private networks, since in this case the completion of conflict detection on the former network does not provide assurance against an address conflict on the new network. Until a host has confirmed the operability of its IPv4 configuration by receipt of a response to the reachability test, it SHOULD NOT respond to ARP Requests and SHOULD NOT broadcast ARP Requests containing its address within the sender protocol address field (ar$spa).

Sending an ICMP Echo Request [RFC792] would not be an acceptable way of testing a candidate configuration since sending any IP packet generally requires an ARP Request/Reply exchange, and as explained above, ARP packets may not be broadcast safely until after the candidate configuration has been confirmed. Also where a host moves from one private network to another, an ICMP Echo Request can result in an ICMP Echo Response even when the MAC address has changed, as long as the IPv4 address remains the same. This can occur, for example, where a host moves from one home network using prefix 192.168/16 to another one. In addition, if the ping is sent with TTL > 1, then an ICMP Echo Response can be received from an off-link router. As a result, if the MAC address of the test node is not checked, the host can mistakenly confirm attachment, potentially resulting in an address conflict. As a result, sending of an ICMP Echo Request SHOULD NOT be used as a substitute for the reachability test.

2.2. IPv4 Address Acquisition

If the host has an operable routable IPv4 address on one or more networks, and DHCPv4 is enabled on the interface, the host SHOULD attempt to acquire an IPv4 configuration using DHCPv4, in parallel with one or more reachability tests. This is accomplished by

entering the INIT-REBOOT state, and sending a DHCPREQUEST to the broadcast address as specified in Section 4.4.2 of the DHCP specification [RFC2131].

If the host does not have an operable routable IPv4 address on any network, the host enters the INIT state and sends a DHCPDISCOVER packet to the broadcast address, as described in Section 4.4.1 of the DHCP specification [RFC2131]. If the host supports the Rapid Commit Option [RFC4039], it is possible that the exchange can be shortened from a 4-message exchange to a 2-message exchange.

If the host does not receive a response to a DHCPREQUEST or DHCPDISCOVER, then it retransmits as specified in Section 4.1 of the DHCP specification [RFC2131].

As discussed in Section 4.4.4 of the DHCP specification [RFC2131], a host in INIT or REBOOTING state that knows the address of a DHCP server may use that address in the DHCPDISCOVER or DHCPREQUEST rather than the IPv4 broadcast address. In the INIT-REBOOT state a DHCPREQUEST is sent to the broadcast address so that the host will receive a response regardless of whether the previously configured IPv4 address is correct for the network to which it has connected.

Sending a DHCPREQUEST to the unicast address in INIT-REBOOT state is not appropriate, since if the DHCP client has moved to another subnet, a DHCP server response cannot be routed back to the client since the DHCPREQUEST will bypass the DHCP relay and will contain an invalid source address.

2.3. IPv4 Link-Local Addresses

DNAv4 applies only to previously-configured addresses that had some lease lifetime associated with them, during which lifetime the address may be legitimately regarded as being reserved for exclusive use by the assigned host. DHCP-assigned addresses fit this description, but IPv4 Link-Local address [RFC3927] do not, since IPv4 Link-Local addresses are not handed out by an authoritative server, and do not come with any guaranteed usable lifetime.

A host's claim on an IPv4 Link-Local address is valid only as long as that host remains connected to the link, actively defending against probes for its chosen address. As soon as a host shuts down, sleeps, or otherwise disconnects from a link, it immediately relinquishes any claim it may have had on any IPv4 Link-Local address on that link. A host wishing to reclaim a previously-used IPv4 Link-Local address MUST perform the full probing and announcement process required by "Dynamic Configuration of IPv4 Link-Local Addresses" [RFC3927], and MUST NOT attempt to use DNAv4 as a shortcut to bypass that process.

Where the host does not have an operable routable IPv4 address on any network, the host MAY configure an IPv4 Link-Local address prior to entering the INIT state and sending a DHCPDISCOVER packet, as described in Section 2.3 of the DHCP specification [RFC2131]. Where a host can confirm that it remains connected to a network on which it possesses an operable routable IPv4 address, that address should be used and the IPv4 Link-Local address is deprecated, as noted in Section 1.9 of the IPv4 Link-Local specification [RFC3927].

Where a host has an operable routable IPv4 address on one or more networks, but the reachability test cannot confirm the configuration and the DHCPv4 client does not receive a response after employing the retransmission algorithm, Section 3.2 of the DHCP specification [RFC2131] states that the client MAY choose to use the previously allocated network address and configuration parameters for the remainder of the unexpired lease.

2.4. Manually Assigned Addresses

An implementation may use DNAv4 to confirm the configuration of manually assigned addresses. However, special consideration is required for this to produce reliable results, so that it SHOULD NOT be enabled by default.

For the purposes of DNAv4, manually assigned addresses may be treated as equivalent to DHCP-assigned addresses with an infinite lifetime. This does not significantly increase the probability of an address conflict as long as the manually assigned address is reserved by the DHCP server or is outside the scope of addresses assigned by a DHCP server. However, where the manually assigned address is within an address scope utilized by a DHCP server, it is possible that the host will be unavailable when the DHCP server checks for a conflict prior to assigning the conflicting address. In this case, a host utilizing DNAv4 could confirm an address that had been assigned to another host.

Typically an address is manually assigned on a network because a dynamically assigned address was not suitable for some reason. Therefore where both DNAv4 and DHCP are run in parallel and DNAv4 confirms a manual configuration, it may be undesirable to allow this configuration to be overridden by DHCP, as described in Section 2.1. However, packet loss may cause the reachability test to fail while DHCP completes successfully, resulting in the host obtaining a dynamic address where a static address is desired. In order to provide for reliable reconfirmation of manually assigned addresses, reachability tests for manual configurations require a more aggressive retransmission strategy than that detailed in Section 4.1 of the DHCP specification [RFC2131]. For example, shorter

retransmission intervals and more persistent retransmissions may be required.

3. IANA Considerations

This specification does not request the creation of any new parameter registries, nor does it require any other IANA assignments.

4. Security Considerations

Detecting Network Attachment for IPv4 (DNAv4) is based on ARP and DHCP and inherits the security vulnerabilities of these two protocols.

ARP [RFC826] traffic is not secured, so that an attacker gaining access to the network can spoof a response to the reachability test described in Section 2.1, leading the querier to falsely conclude that it is attached to a network that it is not connected to.

Similarly, where DHCPv4 traffic is not secured, an attacker could masquerade as a DHCPv4 server, in order to convince the host that it was attached to a particular network. This and other threats relating to DHCPv4 are described in "Authentication for DHCP Messages" [RFC3118].

The effect of these attacks will typically be limited to denial of service, unless the host utilizes its IP configuration for other purposes, such as security configuration or location determination. For example, a host that disables its personal firewall based on evidence that it had attached to a home network could be compromised by spoofing of the DNAv4 reachability test. In general, adjustment of the security configuration based on DNAv4 is NOT RECOMMENDED.

Hosts that depend on secure IP configuration SHOULD NOT use DNAv4, but SHOULD instead utilize DHCP authentication [RFC3118], possibly in combination with the Rapid Commit Option [RFC4039].

5. References

5.1. Normative References

[RFC826] D. Plummer, "An Ethernet Address Resolution Protocol -or- Converting Network Addresses to 48-bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, November 1982.

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

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.

5.2. Informative References

[ACD] Cheshire, S., "IPv4 Address Conflict Detection", Internet draft (work in progress), draft-cheshire-ipv4-acd-04.txt, July 2005.

[RFC792] Postel, J., "Internet Control Message Protocol", RFC 792, USC/Information Sciences Institute, September 1981.

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G. and E. Lear, "Address Allocation for Private Internets", RFC 1918, February 1996.

[RFC3118] Droms, R. and W. Arbaugh, "Authentication for DHCP Messages", RFC 3118, June 2001.

[RFC3927] Cheshire, S., Aboba, B. and E. Guttman, "Dynamic Configuration of IPv4 Link-Local Addresses", RFC 3927, May 2005.

[RFC4039] Park, S., Kim, P., and B. Volz, "Rapid Commit Option for the Dynamic Host Configuration Protocol version 4 (DHCPv4)", RFC 4039, March 2005.

Acknowledgments

The authors would like to acknowledge Greg Daley of Monash University, Erik Guttman and Erik Nordmark of Sun Microsystems, Ralph Droms of Cisco Systems, Ted Lemon of Nominum, John Loughney of Nokia, Thomas Narten of IBM and David Hankins of ISC for contributions to this document.

Authors' Addresses

Bernard Aboba
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

EMail: bernarda@microsoft.com
Phone: +1 425 818 4011
Fax: +1 425 936 7329

James Carlson
Sun Microsystems, Inc
1 Network Drive

Burlington, MA 01803-2757
USA

Phone: +1 781 442 2084
Fax: +1 781 442 1677
EMail: james.d.carlson@sun.com

Stuart Cheshire
Apple Computer, Inc.
1 Infinite Loop
Cupertino, California 95014, USA

Phone: +1 408 974 3207
EMail: rfc@stuartcheshire.org

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

Open issues

Open issues relating to this specification are tracked on the following web site:

http://www.drizzle.com/~aboba/DNA/