Network Working Group
Request for Comments: 5071
Category: Informational

D. Hankins
ISC

December 2007

Dynamic Host Configuration Protocol Options Used by PXELINUX

Status of This Memo

Abstract

This document describes the use by PXELINUX of some DHCP Option Codes numbering from 208-211.

Table of Contents

1. Introduction

PXE, the Preboot eXecution Environment, is a first-stage network bootstrap agent. PXE is loaded out of firmware on the client host, and performs DHCP [3] queries to obtain an IP address.

Once on the network, it loads a second-stage bootstrap agent as configured by DHCP header and option contents.

PXELINUX is one such second-stage bootstrap agent. Once PXE has passed execution to it, PXELINUX seeks its configuration from a cache of DHCP options supplied to the PXE first-stage agent, and then takes action based upon those options.

Most frequently, this implies loading via Trivial File Transfer Protocol (TFTP) [6] one or more images that are decompressed into memory, then executed to pass execution to the final Host Operating System.

PXELINUX uses DHCP options 208-211 to govern parts of this bootstrap process, but these options are not requested by the PXE DHCP client at the time it acquires its lease. At that time, the PXE bootloader has no knowledge that PXELINUX is going to be in use, and even so, would have no way to know what option(s) PXELINUX might digest. Local installations that serve this PXELINUX image to its clients must also configure their DHCP servers to provide these options even though they are not on the DHCP Parameter Request List [4].

These options are:

- "MAGIC" - 208 - An option whose presence and content verifies to the PXELINUX bootloader that the options numbered 209-211 are for the purpose as described herein.

- "ConfigFile" - 209 - Configures the path/filename component of the configuration file’s location, which this bootloader should use to configure itself.

- "PathPrefix" - 210 - Configures a value to be prepended to the ConfigFile to discern the directory location of the file.

- "RebootTime" - 211 - Configures a timeout after which the bootstrap program will reboot the system (most likely returning it to PXE).

Historically, these option codes numbering from 208-211 were designated ’Site Local’, but after publication of RFC3942 [8], they were made available for allocation as new standard DHCP options.

This document marks these codes as assigned.

This direct assignment of option code values in the option definitions below is unusual as it is not mentioned in DHCP Option Code assignment guidelines [5]. This document's Option Code assignments are done within RFC 3942's provisions for documenting prior use of option codes within the new range (128-223 inclusive).

2. Terminology

- "first-stage bootloader" - Although a given bootloading order may have many stages, such as where a BIOS boots a DOS Boot Disk, which then loads a PXE executable, it is, in this example, only the PXE executable that this document describes as the "first-stage bootloader" -- in essence, this is the first stage of booting at which DHCP is involved.

- "second-stage bootloader" - This describes a program loaded by the first-stage bootloader at the behest of the DHCP server.

- "bootloader" and "network bootstrap agent" - These are synonyms, excepting that "bootloader" is intentionally vague in that its next form of bootstrapping may not in fact involve network resources.

The key words "MAY", "MUST", "MUST NOT", "SHOULD", and "SHOULD NOT" in this document are to be interpreted as described in RFC 2119 [2].

3. MAGIC Option

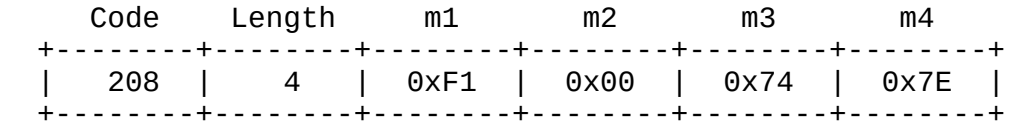
3.1. Description

If this option is provided to the PXE bootloader, then the value is checked by PXELINUX to match the octet string f1:00:74:7e. If this matches, then PXELINUX bootloaders will also consume options 209-211, as described below. Otherwise, they are ignored.

This measure was intended to ensure that, as the 'Site Local' option space is not allocated from a central authority, no conflict would result in a PXELINUX bootloader improperly digesting options intended for another purpose.

3.2. Packet Format

The MAGIC Option format is as follows:



The code for this option is 208. The length is always four.

3.3. Applicability

This option is absolutely inapplicable to any other purpose.

3.4. Response to RFC 3942

The option code 208 will be adopted for this purpose and immediately deprecated. Future standards action may return this option to an available status should it be necessary.

A collision of the use of this option is harmless (at least from PXELINUX' point of view) by design: if it does not match the aforementioned magic value, the PXELINUX bootloader will take no special action.

The PXELINUX project will deprecate the use of this option; future versions of the software will not evaluate its contents.

It is reasonable to utilize this option code for another purpose, but it is recommended to do this at a later time, given the desire to avoid potential collisions in legacy user bases.

4. Configuration File Option

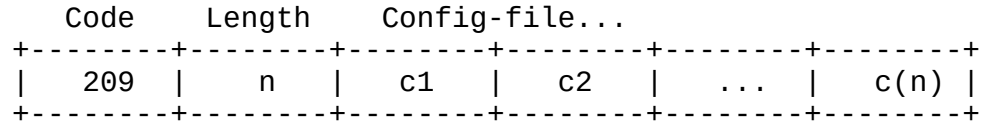
4.1. Description

Once the PXELINUX executable has been entered from the PXE bootloader, it evaluates this option and loads a file of that name via TFTP. The contents of this file serve to configure PXELINUX in its next stage of bootloading (specifying boot image names, locations, boot-time flags, text to present the user in menu selections, etc).

In the absence of this option, the PXELINUX agent will search the TFTP server (as determined by PXE prior to this stage) for a config file of several default names.

4.2. Packet Format

The Configuration File Option format is as follows:



The code for this option is 209. The Config-file (c1..c(n)) is an NVT-ASCII [1] printable string; it is not terminated by a zero or any other value.

4.3. Applicability

Any bootloader, PXE or otherwise, that makes use of a separate configuration file rather than containing all configurations within DHCP options (which may be impossible due to the limited space available for DHCP options) may conceivably make use of this option.

4.4. Response to RFC 3942

The code 209 will be adopted for this purpose.

4.5. Client and Server Behaviour

The Config File Option MUST be supplied by the DHCP server if it appears on the Parameter Request List, but MUST also be supplied if the server administrator believed it would later be useful to the client (such as because the server is configured to offer a second-stage boot image, which they know will make use of it). The option MUST NOT be supplied if no value has been configured for it, or if a value of zero length has been configured.

The DHCP client MUST only cache this option in a location the second-stage bootloader may access.

The second-stage bootloader MUST, in concert with other DHCP options and fields, use this option’s value as a filename to be loaded via TFTP and read for further second-stage-loader-specific configuration parameters. The format and content of such a file is specific to the second-stage bootloader, and as such, is out of scope of this document.

5. Path Prefix Option

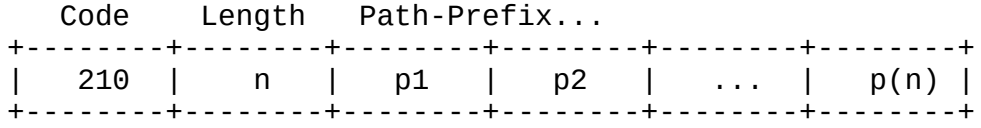
5.1. Description

In PXELINUX' case, it is often the case that several different environments would have the same TFTP path prefix, but would have different filenames (for example: hosts' bootloader images and config files may be kept in a directory structure derived from their Media Access Control (MAC) address). Consequently, it was deemed worthwhile to deliver a TFTP path prefix configuration option, so that these two things could be configured separately in a DHCP Server configuration: the prefix and the possibly host-specific file location.

The actual filename that PXELINUX requests from its TFTP server is derived by prepending this value to the Config File Option above. Once this config file is loaded and during processing, any TFTP file paths specified within it are similarly processed -- prepending the contents of this option.

5.2. Packet Format

The Path Prefix Option format is as follows:



The code for this option is 210. The Path Prefix is an NVT-ASCII printable string; it is not terminated by zero or any other value.

5.3. Applicability

This option came into existence because server administrators found it useful to configure the prefix and suffix of the config file path separately. A group of different PXE booting clients may use the same path prefix, but different filenames, or vice versa.

The 'shortcut' this represents is worthwhile, but it is questionable whether that needs to manifest itself on the protocol wire.

It only becomes interesting from a protocol standpoint if other options are adopted that prefix this value as well -- performing a kind of string compression is highly beneficial to the limited available DHCP option space.

But it's clearly inapplicable to any current use of, e.g., the FILENAME header contents or the DHCP Boot File Name option (#67). Use of these fields is encoded on firmware of thousands of devices that can't or are not likely to be upgraded. Altering any behaviour here is likely to cause severe compatibility problems.

Although compression of the TFTP-loaded configuration file contents is not a compelling factor, contrived configurations using these values may also exist: where each of a large variety of different clients load the same configuration file, with the same contents, but due to a differently configured path prefix actually load different images. Whether this sort of use is truly needed remains unproven.

5.4. Response to RFC 3942

The code 210 will be adopted for this purpose.

5.5. Client and Server Behaviour

The Path Prefix option MUST be supplied by the DHCP server if it appears on the Parameter Request List, but MUST also be supplied if the server administrator believed it would later be useful to the client (such as because the server is configured to offer a second-stage boot image that they know will make use of it). The option MUST NOT be supplied if no value has been configured for it, or if a value of zero length has been configured.

The DHCP client MUST only cache this option in a location where the second-stage bootloader may access it.

The second-stage bootloader MUST prepend this option's value, if any, to the contents of the ConfigFile option prior to obtaining the resulting value via TFTP, or the default 'Config File Search Path', which the second-stage bootloader iterates in the absence of a Config File Option. The client MAY prepend the value to other configuration directives within that file once it has been loaded. The client MUST NOT prepend this option's value to any other DHCP option contents or field, unless explicitly stated in a document describing that option or field.

6. Reboot Time Option

6.1. Description

Should PXELINUX be executed, and then for some reason, be unable to reach its TFTP server to continue bootstrapping, the client will, by default, reboot itself after 300 seconds have passed. This may be too long, too short, or inappropriate behaviour entirely, depending on the environment.

By configuring a non-zero value in this option, admins can inform PXELINUX of which specific timeout is desired. The client will reboot itself if it fails to achieve its configured network resources within the specified number of seconds.

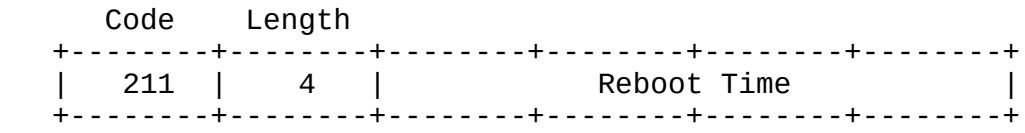
This reboot will run through the system's normal boot-time execution path, most likely leading it back to PXE and therefore PXELINUX. So, in the general case, this is akin to returning the client to the DHCP INIT state.

By configuring zero, the feature is disabled, and instead the client chooses to remove itself from the network and wait indefinitely for operator intervention.

It should be stressed that this is in no way related to configuring a lease time. The perceived transition to INIT state is due to client running state -- reinitializing itself -- not due to lease timer activity. That is, it is not safe to assume that a PXELINUX client will abandon its lease when this timer expires.

6.2. Packet Format

The Reboot Time Option format is as follows:



The code for this option is 211. The length is always four. The Reboot Time is a 32-bit (4 byte) integer in network byte order.

6.3. Applicability

Any network bootstrap program in any sufficiently complex networking environment could conceivably enter into such a similar condition, either due to having its IP address stolen out from under it by a rogue client on the network, by being moved between networks where its PXE-derived DHCP lease is no longer valid, or any similar means.

It seems desirable for any network bootstrap agent to implement an ultimate timeout for it to start over.

The client may, for example, get different working configuration parameters from a different DHCP server upon restarting.

6.4. Response to RFC 3942

The code 211 will be adopted for this purpose.

6.5. Client and Server Behaviour

The Reboot Time Option MUST be supplied by the DHCP server if it appears on the Parameter Request List, but MUST also be supplied if the server administrator believed it would later be useful to the client (such as because the server is configured to offer a second-stage boot image that they know will make use of it). The option MUST NOT be supplied if no value has been configured for it, or if it contains a value of zero length.

The DHCP client MUST only cache this option in a location the second-stage bootloader may access.

If the value of this option is nonzero, the second-stage bootloader MUST schedule a timeout: after a number of seconds equal to this option’s value have passed, the second-stage bootloader MUST reboot the system, ultimately returning the path of execution back to the first-stage bootloader. It MUST NOT reboot the system once the thread of execution has been passed to the host operating system (at which point, this timeout is effectively obviated).

If the value of this option is zero, the second-stage bootloader MUST NOT schedule such a timeout at all. Any second-stage bootloader that finds it has encountered excessive timeouts attempting to obtain its host operating system SHOULD disconnect itself from the network to wait for operator intervention, but MAY continue to attempt to acquire the host operating system indefinitely.

7. Specification Conformance

To conform to this specification, clients and servers MUST implement the Configuration File, Path Prefix, and Reboot Time options as directed.

The MAGIC option MAY NOT be implemented, as it has been deprecated.

8. Security Considerations

PXE and PXELINUX allow any entity acting as a DHCP server to execute arbitrary code upon a system. At present, no PXE implementation is known to implement authentication mechanisms [7] so that PXE clients can be sure they are receiving configuration information from the correct, authoritative DHCP server.

The use of TFTP by PXE and PXELINUX also lacks any form of cryptographic signature -- so a 'Man in the Middle' attack may lead to an attacker's code being executed on the client system. Since this is not an encrypted channel, any of the TFTP loaded data may also be exposed (such as in loading a "RAMDISK" image, which contains /etc/passwd or similar information).

The use of the Ethernet MAC Address as the client's unique identity may allow an attacker who takes on that identity to gain inappropriate access to a client system's network resources by being given by the DHCP server whatever 'keys' are required, in fact, to be the target system (to boot up as though it were the target).

Great care should be taken to secure PXE and PXELINUX installations, such as by using IP firewalls, to reduce or eliminate these concerns.

A nearby attacker might feed a "Reboot Time" option value of 1 second to a mass of unsuspecting clients, to effect a Denial Of Service (DoS) upon the DHCP server, but then again it may just as easily supply these clients with rogue second-stage bootloaders that simply transmit a flood of packets.

This document in and by itself provides no security, nor does it impact existing DCHP security as described in RFC 2131 [3].

9. IANA Considerations

IANA has done the following:

1. Moved DHCPv4 Option code 208 from 'Tentatively Assigned' to 'Assigned', referencing this document. IANA has marked this same option code, 208, as Deprecated.

2. Moved DHCPv4 Option code 209 from 'Tentatively Assigned' to 'Assigned', referencing this document.

3. Moved DHCPv4 Option code 210 from 'Tentatively Assigned' to 'Assigned', referencing this document.

4. Moved DHCPv4 Option code 211 from 'Tentatively Assigned' to 'Assigned', referencing this document.

10. Acknowledgements

These options were designed and implemented for the PXELINUX project by H. Peter Anvin, and he was instrumental in producing this document. Shane Kerr has also provided feedback that has improved this document.

11. References

11.1. Normative References

[1] Postel, J. and J. Reynolds, "Telnet Protocol Specification", STD 8, RFC 854, May 1983.

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

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

[4] Alexander, S. and R. Droms, "DHCP Options and BOOTP Vendor Extensions", RFC 2132, March 1997.

[5] Droms, R., "Procedures and IANA Guidelines for Definition of New DHCP Options and Message Types", BCP 43, RFC 2939, September 2000.

11.2. Informative References

[6] Sollins, K., "The TFTP Protocol (Revision 2)", STD 33, RFC 1350, July 1992.

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

[8] Volz, B., "Reclassifying Dynamic Host Configuration Protocol version 4 (DHCPv4) Options", RFC 3942, November 2004.

Author’s Address

David W. Hankins
Internet Systems Consortium, Inc.
950 Charter Street
Redwood City, CA 94063
US

Phone: +1 650 423 1307
EMail: David_Hankins@isc.org

Full Copyright Statement

Intellectual Property