IPng Working Group
Internet Draft

Matt Crawford
Fermilab
July 14, 2000

IPv6 Node Information Queries
<draft-ietf-ipngwg-icmp-name-lookups-06.txt>

Status of this Memo

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

1. Abstract

This document describes an experimental protocol for asking an IPv6 node to supply certain network information, such as its fully-qualified domain name. IPv6 implementation experience has shown that direct queries for a DNS name are useful, and a direct query mechanism for other information has been requested.

2. Terminology

A "Node Information (or NI) Query" message is sent by a "Querier" node to a "Responder" node in an ICMPv6 packet addressed to the "Queried Address." The Query concerns a "Subject Address" which may differ from the Queried Address, or a "Subject Name". The Responder sends a "Node Information Reply" to the Querier, containing information associated with the node at the Queries address. A node receiving a NI Query will be termed a Responder even if it does not send a Reply.

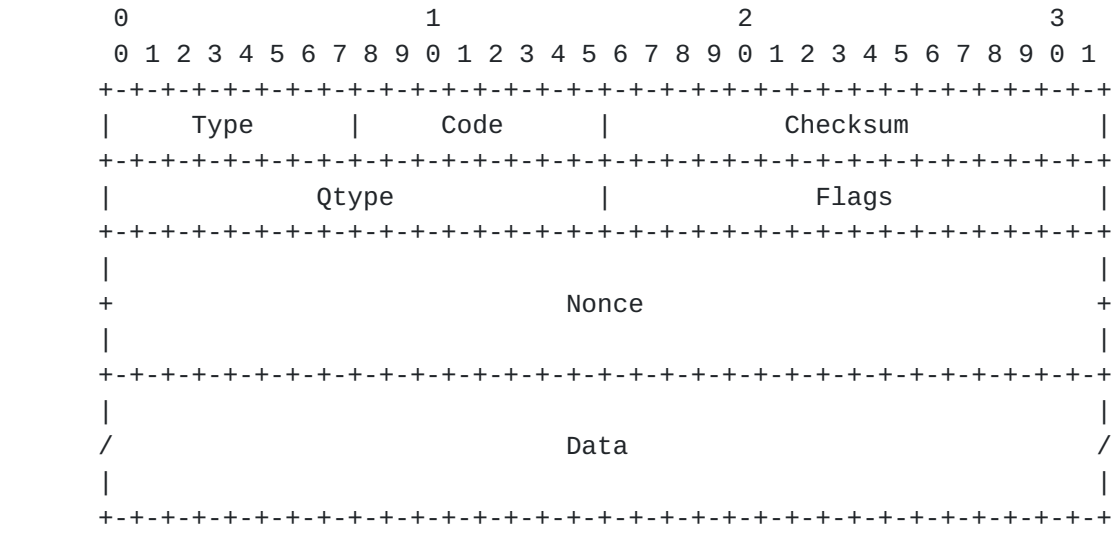
The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this

document are to be interpreted as described in [2119].

Packet fields marked "unused" must be zero on transmission and, aside from inclusion in checksums or message integrity checks, ignored on reception.

3. Node Information Messages

Two types of Node Information messages, the NI Query and the NI Reply, are carried in ICMPv6 [2463] packets. They have the same format.



Fields:

Type 139 - NI Query.
140 - NI Reply.

Code For NI Query:

0 Indicates that the Data field contains an IPv6 address which is the Subject of this Query.

1 Indicates that the Data field contains a domain name which is the Subject of this Query, or is empty, as in the case of a NOOP or Supported Qtypes query.

2 Indicates that the Data field contains an IPv4 address which is the Subject of this Query.

For NI Reply:

0 Indicates a successful reply. The Reply Data field may or may not be empty.

1 Indicates that the Responder refuses to supply the answer. The Reply Data field will be empty.

2 Indicates that the Qtype of the Query is unknown to the Responder. The Reply Data field will be empty.

Checksum The ICMPv6 checksum.

Qtype A 16-bit field which designates the type if information requested in a Query or supplied in a Reply. Its value in a Reply is always copied from the corresponding Query by the Responder. Five values of Qtype are specified in this document.

Flags Qtype-specific flags which may be defined for certain Query types and their Replies. Flags not defined for a given Qtype must be zero on transmission and ignored on reception, and must not be copied from a Query to a Reply unless so specified in the definition of the Qtype.

Nonce An opaque 64-bit field to help avoid spoofing and/or to aid in matching Replies with Queries. Its value in a Query is chosen by the Querier. Its value in a Reply is always copied from the corresponding Request by the Responder.

Data In a Query, the Subject Address or Name. In a Reply, Qtype-specific data present only when the ICMPv6 Type field is zero. The length of the Data may be inferred from the IPv6 header's Payload Length field [2460], the length of the fixed portion of the NI packet and the lengths of the ICMPv6 header and intervening extension headers.

Note that the type of information present in the Data field of a Query is inferred from the ICMP Code, while the type of information, if any, in the Data field of a Reply is inferred from the Qtype.

When the Subject of a Query is a name, the name MUST be in DNS wire format [1035]. The name may be either a fully-qualified domain name, including the terminating zero-length label, or a single DNS label followed by two zero-length labels. Since a Query contains at most one DNS name, DNS compression will not be used.

4. Message Processing

The Querier constructs an ICMP NI Query and sends it to the address from which information is wanted. When the Subject of the Query is an IPv6 address, that address will normally be used as the IPv6 destination address of the Query, but need not be if the Querier has useful a priori information about the addresses of the target node. An NI Query may also be sent to a multicast address of link-local scope [2373].

When the Subject is a domain name, either fully-qualified or single-component, and the Querier does not have a unicast address for the target node, the query MUST be sent to a link-scope multicast address formed in the following way. The Subject Name is converted to canonical form, as defined by DNS Security [2065], which is uncompressed with all alphabetic characters in lower case. (If additional DNS label types for host names are defined, the rules for canonicalizing those labels will be found in the defining specification.) Compute the MD5 hash [1321] of the first label of the Subject Name -- the portion beginning with the first one-octet length field and up to, but excluding, any subsequent length field. Append the first 32 bits of that 128-bit hash to the prefix FF02:0:0:0:0:2::/96. The resulting multicast address will be termed the "NI Group Address" for the name.

The Nonce should be a random or good pseudo-random value to foil spoofed replies. An implementation which allows multiple independent processes to send NI queries MAY use the Nonce value to deliver Replies to the correct process. Nonetheless, such processes MUST check the received Nonce and ignore extraneous Replies.

If true communication security is required, IPsec [2401] must be used.

Upon receiving a NI Query, the Responder must check the Query's IPv6 destination address and discard the Query without further processing unless it is one of the Responder's unicast or anycast addresses, or a link-local scope multicast address which the Responder has joined. Typically the latter will be a NI Group Address for a name belonging to the Responder or a NI Group Address for a name for which the Responder is providing proxy service. A node MAY be configurable to discard NI Queries to multicast addresses other than its NI Group Address(es) but if so, the default configuration MUST be not to discard them.

A Responder must also silently discard a Query whose Subject Address or Name (in the Data field) does not belong to that node, unless it is providing proxy service for that Subject. A single-component

Subject Name matches any fully-qualified name whose first label matches the Subject. All name matching is done in a case-independent manner consistent with DNSSEC name canonicalization [2065].

Next, if Qtype is unknown to the Responder, it must return a NI Reply with ICMPv6 Type = 2 and no Reply Data. The Responder should rate-limit such replies as it would ICMPv6 error replies [2463, 2.4(f)].

Next, the Responder should decide whether to refuse an answer, based on local policy not addressed in this document. If an answer is refused, the Responder may send a NI Reply with ICMPv6 Type = 1 and no Reply Data. Again, the Responder should rate-limit such replies as it would ICMPv6 error replies [2463, 2.4(f)].

Finally, if the Qtype is known and the response is allowed by local policy, the Responder must fill in the Flags and Reply Data of the NI Reply in accordance with the definition of the Qtype and transmit the NI Reply with an ICMPv6 source address equal to the Queried Address, unless that address was an anycast address. If the Queried Address was anycast, the source address for the Reply SHOULD be one belonging to the interface on which the Query was received.

If the Query was sent to an anycast or multicast address, transmission of the Reply MUST be delayed by a random interval between zero and MAX_ANYCAST_DELAY_TIME, as defined by IPv6 Neighbor Discovery [2461].

5. Defined Qtypes

The following five Qtypes are defined. The first four (number 0 to 3) MUST be supported by any implementation of this protocol. The last one SHOULD be supported by any implementation on an IPv4/IPv6 dual-stack node and MAY be supported on an IPv6-only node.

0 NOOP.

1 Supported Qtypes.

2 DNS Name.

3 Node Addresses.

4 IPv4 Addresses.

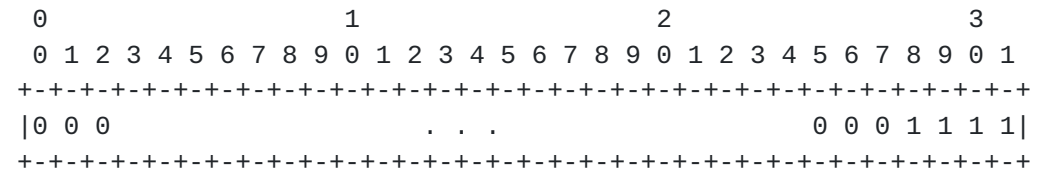
5.1. NOOP

This NI type has no defined flags and never has a Data field. A Reply to a NI NOOP Query tells the Querier that a node with the Queried Address is up and reachable, implements the Node Information protocol, and incidentally happens to reveal whether the Queried Address was an anycast address. On transmission, the ICMPv6 Code in a NOOP Query must be set to 1 and the Code in a NOOP Reply must be 0. On reception of a NOOP Query or Reply, the Code must be ignored.

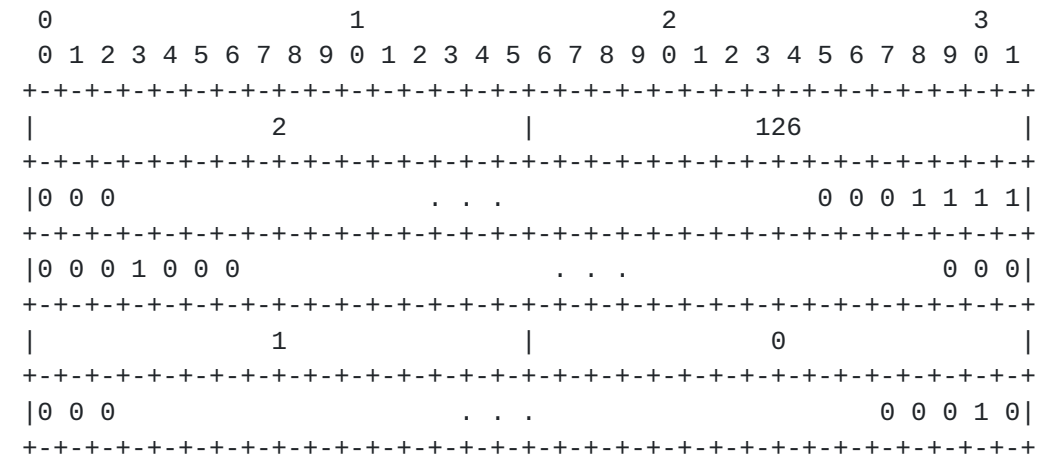
5.2. Supported Qtypes

This Query contains no Data field. The Reply Data is a bit-vector showing which Qtypes are supported by the Responder. The Reply Data has two variant forms: uncompressed and compressed. The uncompressed Data format is one or more complete 32-bit words, each word a bitmask with the low-order bit in each word corresponding to the lowest numbered Qtype in a group of 32. The first word describes the Responder's support for Qtypes 0 to 31, the second word 32 to 63, and so on.

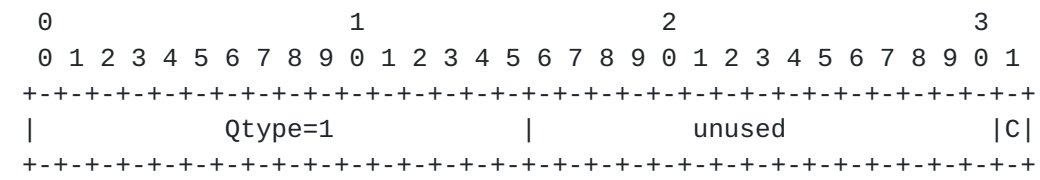
A 1-valued bit indicates support for the corresponding Qtype. The lowest-order four bits in the first 32-bit word must be set to 1, showing support for the four mandatory Qtypes defined in this specification. Thus the Data field of a NI Supported Qtypes Reply from a Responder implementing only the mandatory Qtypes will contain 32 bits in the following form:



The compressed form of the Reply Data consists of a sequence of blocks, each block consisting of two 16-bit unsigned integers, nWord and nSkip, followed by nWord 32-bit bitmasks describing the Responder's support for 32 consecutive Qtypes. nSkip is a count of 32-bit words following the included words which would have been all-zero and have been suppressed. The last block MUST have nSkip = 0. As an example, a Responder supporting Qtypes 0, 1, 2, 3, 60, and 4097 could express that information with the following Reply Data (nWord and nSkip fields are written in decimal for easier reading):



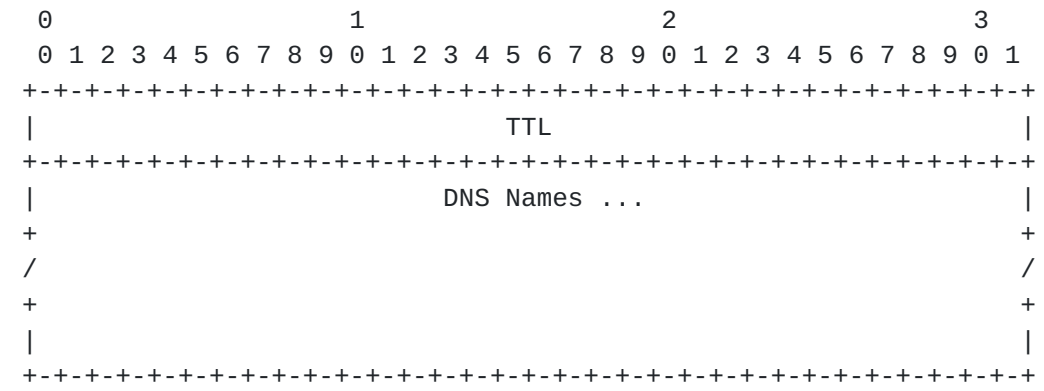
One flag bit is defined.



In a Query, a C-flag set to 1 indicates that the Querier will accept the compressed form of the Reply Data. In a Reply, a C-flag set to 1 indicates that the Reply Data is compressed. The compressed form MAY be used in a Reply only if the Query had the C-flag set. Implementations of this specification SHOULD support the compressed form and if they do, SHOULD set the C-flag in all Supported Qtypes Queries and SHOULD use the compressed form in Supported Qtypes Replies (when allowed by the C-flag in the query) if doing so would avoid fragmentation or save significant space in the Reply.

5.3. DNS Name

The NI DNS Name Query requests the fully-qualified or single-component name corresponding to the Subject Address or Name. The Reply Data has the following format.



TTL The number of seconds that the name may be cached. For compatibility with DNS [1035], this is a 32-bit signed, 2's-complement number, which must not be negative.

DNS Names The fully-qualified or single-component name or names of the Responder which correspond(s) to the Subject Address or Name, in DNS wire format [1035]. Each name MUST be fully-qualified if the responder knows the domain suffix, and otherwise be a single DNS label followed by two zero-length labels.

When multiple DNS names are returned, DNS name compression [1035] SHOULD be used, and the offsets are counted from the first octet of the Data field. An offset of 4, for example, will point to the beginning of the first name.

The Responder must fill in the TTL field of the Reply with a meaningful value if possible. That value should be one of the following.

The remaining lifetime of a DHCP lease on the Subject Address;

The remaining Valid Lifetime of a prefix from which the Subject Address was derived through Stateless Autoconfiguration [2461, 2462];

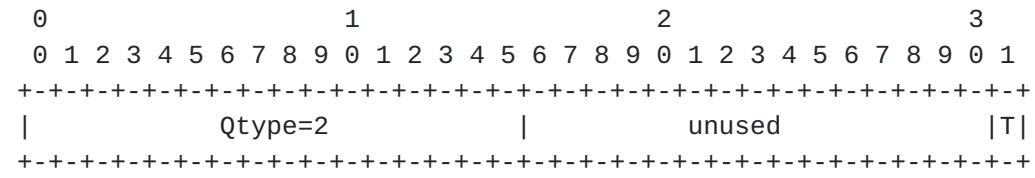
The TTL of an existing AAAA or A6 record which associates the Subject Address with the DNS Name being returned.

If the Responder returns multiple names but considers one name to be official or canonical, that name MUST be placed immediately after the TTL.

Only one TTL is included in the reply. If the Responder considers different names to be cacheable for different times, the TTL field must be set no larger than the minimum of those times.

If the Responder does not know its name at all it MUST send a Reply with TTL=0 and no DNS Names. The Querier will be able to determine from the packet length that the Data field contains only a TTL.

One Flag bit is defined, in the Reply only.



A T-flag set to 1 in a NI DNS Name Reply indicates that the TTL field contains a meaningful value. If the T-flag is 0, the TTL SHOULD be set to zero by the Responder and MUST be ignored by the Querier.

If a name rather than an address was the Subject of the Query, the T-flag MUST be zero and the TTL SHOULD be zero.

The information in a NI DNS Name Reply with T-flag 1 may be cached and used for the period indicated by that TTL. If a Reply has no TTL (T-flag 0), the information in that Reply must not be used more than once. If the Query was sent by a DNS server on behalf of a DNS client, the result may be returned to that client as a DNS response with TTL zero. However, if the server has the matching AAAA record, either in cache or in an authoritative zone, then the TTL of that record may be used as the missing TTL of the NI DNS Name Reply and the information in the reply may be cached and used for that period.

It would be an implementation choice for a server to perform a DNS query for the AAAA or A6 record that matches a received NI DNS Name Reply. This might be done to obtain a TTL to make the Reply cacheable or in anticipation of such a AAAA query from the client that caused the DNS Name Query.

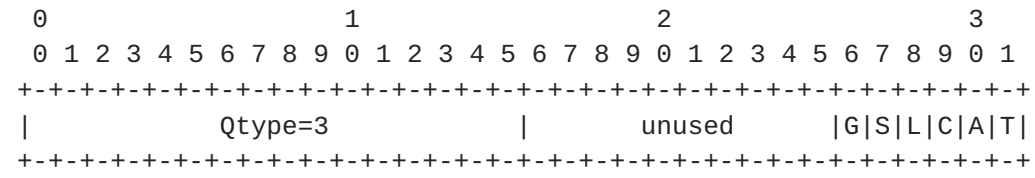
5.3.1. Discussion

Because a node can only answer a DNS Name Request when it is up and reachable, it may be useful to create a proxy responder for a group of nodes, for example a subnet or a site. Such a mechanism is not addressed here.

IPsec can be applied to NI DNS Name messages to achieve greater trust in the information obtained, but such a need may be obviated by applying IPsec directly to some other communication which is going on (or contemplated) between the Querier and Responder.

5.4. Node Addresses

The NI Node Addresses Query requests some set of the Responder's IPv6 unicast addresses. The Reply Data is a sequence of 128-bit IPv6 addresses, each address preceded by separate a 32-bit TTL value, with Preferred addresses listed before Deprecated addresses [2461], but otherwise in no special order. Five flag bits are defined in the Query, and six in the Reply.



G If set to 1, Global-scope addresses [2374] are requested.

S If set to 1, Site-local addresses [2374] are requested.

L If set to 1, Link-local addresses [2374] are requested.

C If set to 1, IPv4-compatible and IPv4-mapped addresses [2373] are requested.

A If set to 1, all the Responder's unicast addresses (of the specified scope(s)) are requested. If 0, only those addresses are requested which belong to the interface (or any one interface) which has the Subject Address, or which are associated with the Subject Name.

T Defined in a Reply only, indicates that the set of addresses is incomplete for space reasons.

Flags G, S, L, C and A are copied from a Query to the corresponding Reply.

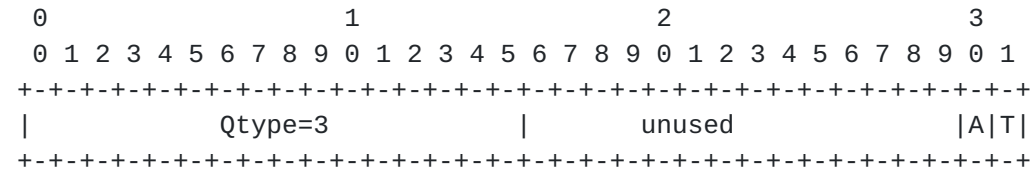
The TTL associated with each address are to be determined by the rules in section 5.3, applied to the returned address rather than the Subject. If no meaningful caching time can be given for an address, the corresponding TTL field MUST be zero.

Each address with non-zero TTL in a NI Node Address Reply may be cached and used for the period indicated by that TTL. If the TTL is zero, the corresponding address must not be used more than once. If the Query was sent by a DNS server on behalf of a DNS client, the result may be returned to that client as a DNS response with TTL zero.

IPv4-mapped addresses can only be returned by a Node Information proxy, since they represent addresses of IPv4-only nodes, which perforce do not implement this protocol.

5.5. IPv4 Addresses

The NI IPv4 Addresses Query requests some set of the Responder's IPv4 unicast addresses. The Reply Data is a sequence of 32-bit IPv4 addresses, each address preceded by a 32-bit TTL value. One flag bit is defined in the Query, and two in the Reply.



A If set to 1, all the Responder's unicast addresses are requested. If 0, only those addresses are requested which belong to the interface (or any one interface) which has the Subject Address.

T Defined in a Reply only, indicates that the set of addresses is incomplete for space reasons.

Flag A is copied from a Query to the corresponding Reply.

The TTL associated with each address are to be determined by the rules in section 5.3, applied to the returned address rather than the Subject, excluding the autoconfiguration Valid Lifetime. If no meaningful caching time can be given for an address, the corresponding TTL field MUST be zero.

Each address with non-zero TTL in a NI IPv4 Address Reply may be cached and used for the period indicated by that TTL. If the TTL is zero, the corresponding address must not be used more than once. If the Query was sent by a DNS server on behalf of a DNS client, the result may be returned to that client as a DNS response with TTL zero.

5.5.1. Discussion

It is possible that a node may treat IPv4 interfaces and IPv6 interfaces as distinct, even though they are associated with the same hardware. When such a node is responding to a NI Query having a Subject Address of one type requesting the other type, and the Query has the A flag set to 0, it SHOULD consider IP interfaces, other than tunnels, associated with the same hardware as being the same interface.

6. IANA Considerations

ICMPv6 type values 139 and 140 have been assigned by IANA for this protocol.

This document defines five values of Qtype, numbers 0 through 4. Following the policies outlined in [2434], new values, and their associated Flags and Reply Data, may be defined as follows.

Qtypes 5 through 255, by IETF Consensus.

Qtypes 256 through 1023, Specification Required.

Qtypes 1024 through 4095, First Come First Served.

Qtypes 4096 through 65535, Private Use.

Users of Private Use values should note that values above 8000 to 9000 are likely to lead to fragmentation of "Supported Qtypes" Replies unless the compressed for of the Reply Data is used.

Assignment of the multicast address prefix FF02:0:0:0:0:2::/96 used in section 4 as a destination for IPv6 Node Information Queries is requested.

7. Security Considerations

The anti-spoofing Nonce does not give any protection from spoofers who can snoop the Query or the Reply.

In a large Internet with relatively frequent renumbering, the maintenance of of KEY and SIG records [2065] in the zones used for address-to-name translations will be no easier than the maintenance of the NS, SOA and PTR records themselves, which already appears to be difficult in many cases. The author expects, therefore, that address-to-name mappings, either through the original DNS mechanism

or through this new mechanism, will generally be used as only a hint to find more trustworthy information using the returned name as an index.

8. Acknowledgments

Alain Durand contributed to this specification and valuable feedback and implementation experience was provided by Jun-Ichiro Hagino and Tatuya Jinmei. This document is not the first proposal of a direct query mechanism for address-to-name translation. The idea had been discussed briefly in the IPng working group and an experimental RFC [1788] describes such a mechanism for IPv4.

9. References

[1035] P. Mockapetris, "Domain Names - Implementation and Specification", RFC 1035, STD 13, November 1987.

[1321] R. Rivest, "The MD5 Message-Digest Algorithm", RFC 1321, April 1992.

[1788] W. Simpson, "ICMP Domain Name Messages", RFC 1788, April 1995.

[2065] Eastlake 3rd, D. and C. Kaufman, "Domain Name System Security Extensions", RFC 2065, January 1997.

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

[2373] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[2401] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[2434] Narten, T. and H. T. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 2434, October 1998.

[2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[2461] Narten, T., Nordmark, E. and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[2462] Thomson, S. and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, December 1998.

[2463] Conta, A. and S. Deering, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 2463, December 1998.

10. Author's Address

Matt Crawford
Fermilab MS 368
PO Box 500
Batavia, IL 60510
USA

Phone: +1 630 840 3461

Email: crawdad@fnal.gov