Internet Engineering Task Force (IETF)
Request for Comments: 8427
Category: Informational
ISSN: 2070-1721

P. Hoffman
ICANN

July 2018

Representing DNS Messages in JSON

Abstract

Some applications use DNS messages, or parts of DNS messages, as data. For example, a system that captures DNS queries and responses might want to be able to easily search them without having to decode the messages each time. Another example is a system that puts together DNS queries and responses from message parts. This document describes a general format for DNS message data in JSON. Specific profiles of the format in this document can be described in other documents for specific applications and usage scenarios.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are candidates for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8427.

Table of Contents

1. Introduction

The DNS message format is defined in [RFC1035]. DNS queries and DNS responses have exactly the same structure. Many of the field names and data type names given in [RFC1035] are commonly used in discussions of DNS. For example, it is common to hear things like "the query had a QNAME of 'example.com'" or "the RDATA has a simple structure".

There are hundreds of data-interchange formats for serializing structured data. Currently, JSON [RFC8259] is quite popular for many types of data, particularly data that has named subfields and optional parts.

This document uses JSON to describe DNS messages. It also defines how to describe a paired DNS query and response and how to stream DNS objects.

1.1. Design of the Format

There are many ways to design a data format. This document uses a specific design methodology based on the DNS format.

- The format is based on JSON objects in order to allow a writer to include or exclude parts of the format at will. No object members are ever required.

- This format is purposely overly general. A protocol or application that uses this format is expected to use only a subset of the items defined here; it is expected to define its own profile from this format.

- The format allows transformation through JSON that would permit re-creation of the wire content of the message.

- All members whose values are always 16 bits or shorter are represented by JSON numbers with no minus sign, no fractional part (except in fields that are specifically noted below), and no exponent part. One-bit values are represented as JSON numbers whose values are either 0 or 1. See Section 6 of [RFC8259] for more detail on JSON numbers.

- The JSON representation of the objects described in this document is limited to the UTF-8 codepoints from U+0000 to U+007F. This is done to prevent an attempt to use a different encoding such as UTF-8 for octets in names or data.

- Items that have string values can have "HEX" appended to their names to indicate a non-ASCII encoding of the value. Names that end in "HEX" have values stored in base16 encoding (hex with uppercase letters) defined in [RFC4648]. This is particularly useful for RDATA that is binary.

- All field names in this format are used as in [RFC1035], including their capitalization. Names not defined in [RFC1035] generally use "camel case".

- The same data may be represented in multiple object members multiple times. For example, there is a member for the octets of the DNS message header, and there are members for each named part of the header. A message object can thus inadvertently have inconsistent data, such as a header member whose value does not match the value of the first bits in the entire message member.

- It is acceptable that there are multiple ways to represent the same data. This is done so that application designers can choose what fields are best for them and even so that they are able to allow multiple representations. That is, there is no "better" way to represent DNS data, so this design doesn’t prefer specific representations.

- The design explicitly allows for the description of malformed DNS messages. This is important for systems that are logging messages seen on the wire, particularly messages that might be used as part of an attack. A few examples of malformed DNS messages include:

  * a resource record (RR) that has an RDLENGTH of 4 but an RDATA whose length is longer than 4 (if it is the last RR in a message)

  * a DNS message whose QDCOUNT is 0

  * a DNS message whose ANCOUNT is large but there are insufficient bytes after the header

  * a DNS message whose length is less than 12 octets, meaning it doesn’t even have a full header

- An object in this format can have zero or more of the members defined here; that is, no members are required by the format itself. Instead, profiles that use this format might have requirements for mandatory members, optional members, and prohibited members from the format. Also, this format does not prohibit members that are not defined in this format; profiles of the format are free to add new members in the profile.

- This document defines DNS messages, not the zone files described in [RFC1035]. A different specification could be written to extend it to represent zone files. Note that DNS zone files allow escaping of octet values using "\DDD" notation, but this specification does not allow that; when encoding from a zone file to this JSON format, you need to do a conversion for many types of values.

1.2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. JSON Format for DNS Messages

The following gives all of the members defined for a DNS message. It is organized approximately by levels of the DNS message.

2.1. Message Object Members

- ID - Integer whose value is 0 to 65535
- QR - Boolean
- Opcode - Integer whose value is 0 to 15
- AA - Boolean
- TC - Boolean
- RD - Boolean
- RA - Boolean
- AD - Boolean
- CD - Boolean
- RCODE - Integer whose value is 0 to 15
- QDCOUNT - Integer whose value is 0 to 65535
- ANCOUNT - Integer whose value is 0 to 65535
- NSCOUNT - Integer whose value is 0 to 65535

- ARCOUNT - Integer whose value is 0 to 65535
- QNAME - String of the name of the first Question section of the message; see Section 2.6 for a description of the contents
- compressedQNAME - Object that describes the name with two optional values: "isCompressed" (with a value of 0 for no and 1 for yes) and "length" (with an integer giving the length in the message)
- QTYPE - Integer whose value is 0 to 65535, of the QTYPE of the first Question section of the message
- QTYPEname - String whose value is from the IANA "Resource Record (RR) TYPEs" registry or has the format in [RFC3597]; this is case sensitive, so "AAAA", not "aaaa"
- QCLASS - Integer whose value is 0 to 65535, of the QCLASS of the first Question section of the message
- QCLASSname - String whose value is "IN", "CH", or "HS" or that has the format in [RFC3597]
- questionRRs - Array of zero or more resource records or rrSet objects in the Question section
- answerRRs - Array of zero or more resource records or rrSet objects in the Answer section
- authorityRRs - Array of zero or more resource records or rrSet objects in the Authority section
- additionalRRs - Array of zero or more resource records or rrSet objects in the Additional section

2.2. Resource Record Object Members

A resource record is represented as an object with the following members.

- NAME - String of the NAME field of the resource record; see Section 2.6 for a description of the contents
- compressedNAME - Object that describes the name with two optional values: "isCompressed" (with a value of 0 for no and 1 for yes) and "length" (with an integer giving the length in the message)
- TYPE - Integer whose value is 0 to 65535

- TYPEname - String whose value is from the IANA "Resource Record (RR) TYPEs" registry or has the format in [RFC3597]; this is case sensitive, so "AAAA", not "aaaa"

- CLASS - Integer whose value is 0 to 65535

- CLASSname - String whose value is "IN", "CH", or "HS" or has the format in [RFC3597]

- TTL - Integer whose value is -2147483648 to 2147483647 (it will only be 0 to 2147483647 in normal circumstances)

- RDLENGTH - Integer whose value is 0 to 65535. Applications using this format are unlikely to use this value directly, and instead calculate the value from the RDATA.

- RDATAHEX - Hex-encoded string (base16 encoding, described in [RFC4648]) of the octets of the RDATA field of the resource record. The data in some common RDATA fields are also described in their own members; see Section 2.3

- rrSet - List of objects that have RDLENGTH and RDATA members

A Question section can be expressed as a resource record. When doing so, the TTL, RDLENGTH, and RDATA members make no sense.

2.3. Specific RDATA Field Members

The following are common RDATA types and how to specify them as JSON members. The name of the member contains the name of the RDATA type. The data type for each of these members is a string. Each name is prefaced with "rdata" to prevent a name collision with fields that might later be defined that have the same name as the raw type name.

- rdataA - IPv4 address, such as "192.168.33.44"

- rdataAAAA - IPv6 address, such as "fe80::a65e:60ff:fed6:8aaf", as defined in [RFC5952]

- rdataCNAME - A domain name

- rdataDNAME - A domain name

- rdataNS - A domain name

- rdataPTR - A domain name

- rdataTXT - A text value

In addition, each of the following members has a value that is a space-separated string that matches the display format definition in the RFC that defines that RDATA type. It is not expected that every receiving application will know how to parse these values.

rdataCDNSKEY, rdataCDS, rdataCSYNC, rdataDNSKEY, rdataHIP, rdataIPSECKEY, rdataKEY, rdataMX, rdataNSEC, rdataNSEC3, rdataNSEC3PARAM, rdataOPENPGPKEY, rdataRRSIG, rdataSMIMEA, rdataSPF, rdataSRV, rdataSSHFP, rdataTLSA

2.4. The Message and Its Parts as Octets

The following can be members of a message object. These members are all encoded in base16 encoding, described in [RFC4648]. All these items are strings.

- messageOctetsHEX - The octets of the message
- headerOctetsHEX - The first 12 octets of the message (or fewer, if the message is truncated)
- questionOctetsHEX - The octets of the Question section
- answerOctetsHEX - The octets of the Answer section
- authorityOctetsHEX - The octets of the Authority section
- additionalOctetsHEX - The octets of the Additional section

The following can be a member of a resource record object.

- rrOctetsHEX - The octets of a particular resource record

The items in this section are useful in applications to canonically reproduce what appeared on the wire. For example, an application that is converting wire-format requests and responses might do decompression of names, but the system reading the converted data may want to be sure the decompression was done correctly. Such a system would need to see the part of the message where the decompressed labels resided, such as in one of the items in this section.

2.5. Additional Message Object Members

The following are members that might appear in a message object:

- dateString - The date that the message was sent or received, given as a string in the standard format described in [RFC3339] and refined by Section 3.3 of [RFC4287].

- dateSeconds - The date that the message was sent or received, given as a JSON number that is the number of seconds since 1970-01-01T00:00Z in UTC time; this number can be fractional. This number must have no minus sign, can have an optional fractional part, and can have no exponent part.

- comment - An unstructured comment as a string.

2.6. Name Fields

Names are represented by JSON strings. The rules for how names are encoded are described in Section 1.1. (To recap: it is limited to the UTF-8 codepoints from U+0000 to U+007F.) The contents of these fields are always uncompressed; that is, after [RFC1035], name compression has been removed.

There are two encodings for names:

- If the member name does not end in "HEX", the value is a domain name encoded as DNS labels consisting of UTF-8 codepoints from U+0000 to U+007F. Within a label, codepoints above U+007F and the codepoint U+002E (ASCII period) MUST be expressed using JSON's escaping rules within this set of codepoints. Separation between labels is indicated with a period (codepoint U+002E). Internationalized Domain Name (IDN) labels are always expressed in their A-label form, as described in [RFC5890].

- If the member name ends in "HEX", the value is the wire format for an entire domain name stored in base16 encoding, which is described in [RFC4648].

3. JSON Format for a Paired DNS Query and Response

A paired DNS query and response is represented as an object. Two optional members of this object are named "queryMessage" and "responseMessage", and each has a value that is a message object. This design was chosen (as compared to the more obvious array of two values) so that a paired DNS query and response could be differentiated from a stream of DNS messages whose length happens to be two.

4. Streaming DNS Objects

Streaming of DNS objects is performed using JSON text sequences [RFC7464].

5. Examples

5.1. Example of the Format of a DNS Query

The following is an example of a query for the A record of example.com.

```
{ "ID": 19678, "QR": 0, "Opcode": 0,
  "AA": 0, "TC": 0, "RD": 0, "RA": 0, "AD": 0, "CD": 0, "RCODE": 0,
  "QDCOUNT": 1, "ANCOUNT": 0, "NSCOUNT": 0, "ARCOUNT": 0,
  "QNAME": "example.com", "QTYPE": 1, "QCLASS": 1
}
```

As stated earlier, all members of an object are optional. This example object could have one or more of the following members as well:

```
"answerRRs": []
"authorityOctetsHEX": ""
"comment": "Something pithy goes here"
"dateSeconds": 1408504748.657783
"headerOctetsHEX": "4CDE00000001000000000000"
"QNAMEHEX": "076578616D706C6503636F6D00",
"compressedQNAME": { "isCompressed": 0 },
"messageOctetsHEX":
     "4CDE00000001000000000000076578616D706C6503636F6D0000010001"
"questionOctetsHEX": "076578616D706C6503636F6D0000010001"
"questionRRs": [ { "NAMEHEX": "076578616D706C6503636F6D00",
                "TYPE": 1, "CLASS": 1, "hostNAME" : "example.com." } ]
"questionRRs": [ { "NAME": "example.com.", "TYPE": 1,
                "CLASS": 1, } ]
```

(Note that this is an incomplete list of what else could be in the object.)

5.2. Example of the Format of a Paired DNS Query and Response

The following is a paired DNS query and response for a query for the A record of example.com.

```
{
  "queryMessage": { "ID": 32784, "QR": 0, "Opcode": 0, "AA": 0,
                    "TC": 0, "RD": 0, "RA": 0, "AD": 0, "CD": 0,
                    "RCODE": 0, "QDCOUNT": 1, "ANCOUNT": 0,
                    "NSCOUNT": 0, "ARCOUNT": 0,
                    "QNAME": "example.com.",
                    "QTYPE": 1, "QCLASS": 1 },
  "responseMessage": { "ID": 32784, "QR": 1, "AA": 1, "RCODE": 0,
                       "QDCOUNT": 1, "ANCOUNT": 1, "NSCOUNT": 1,
                       "ARCOUNT": 0,
                       "answerRRs": [ { "NAME": "example.com.",
                                        "TYPE": 1, "CLASS": 1,
                                        "TTL": 3600,
                                        "RDATAHEX": "C0000201" },
                                      { "NAME": "example.com.",
                                        "TYPE": 1, "CLASS": 1,
                                        "TTL": 3600,
                                        "RDATAHEX": "C000AA01" } ],
                        "authorityRRs": [ { "NAME": "ns.example.com.",
                                            "TYPE": 1, "CLASS": 1,
                                            "TTL": 28800,
                                            "RDATAHEX": "CB007181" } ]
                     }
}
```

The Answer section could instead be given with an rrSet:

```
"answerRRs": [ { "NAME": "example.com.",
                 "TYPE": 1, "CLASS": 1,
                 "TTL": 3600,
                 "rrSet": [ { "RDATAHEX": "C0000201" },
                            { "RDATAHEX": "C000AA01" } ] } ],
```

(Note that this is an incomplete list of what else could be in the Answer section.)

6. Local Format Policy

Systems using the format in this document will likely have policy about what must be in the objects. Those policies are outside the scope of this document.

For example, passive DNS systems such as those described in [PASSIVE-DNS] cover just DNS responses. Such a system might have a policy that makes QNAME, QTYPE, and answerRRs mandatory. That document also describes two mandatory times that are not in this format, so the policy would possibly also define those members and make them mandatory. The policy could also define additional members that might appear in a record.

As another example, a program that uses this format for configuring what a test client sends on the wire might have a policy of "each record object can have as few members as it wants; all unstated members are filled in from previous records".

7. IANA Considerations

7.1. Media Type Registration of application/dns+json

Type name: application

Subtype name: dns+json

Required parameters: N/A

Optional parameters: N/A

Encoding considerations: Encoding considerations are identical to those specified for the "application/json" media type.

Security considerations: This document specifies the security considerations for the format.

Interoperability considerations: This document specifies format of conforming messages and the interpretation thereof.

Published specification: This document

Applications that use this media type: Systems that want to exchange DNS messages

Fragment identifier considerations: None

Additional information:

Deprecated alias names for this type: N/A

Magic number(s): N/A

File extension(s): This document uses the media type to refer to protocol messages and thus does not require a file extension.

Macintosh file type code(s): N/A

Person & email address to contact for further information:
Paul Hoffman, paul.hoffman@icann.org

Intended usage: COMMON

Restrictions on usage: N/A

Author: Paul Hoffman, paul.hoffman@icann.org

Change controller: Paul Hoffman, paul.hoffman@icann.org

8. Security Considerations

As described in Section 1.1, a message object can have inconsistent data, such as a message with an ANCOUNT of 1 but that has either an empty answerRRs array or an answerRRs array that has two or more RRs. Other examples of inconsistent data would be resource records whose RDLENGTH does not match the length of the decoded value in the RDATAHEX member, a record whose various header fields do not match the value in headerOctetsHEX, and so on. A reader of this format must never assume that all of the data in an object are all consistent with each other.

This document describes a format, not a profile of that format. Lack of profile can lead to security issues. For example, if a system has a filter for JSON representations of DNS packets, that filter needs to have the same semantics for the output JSON as the consumer has. Unless the profile is quite tight, this can lead to the producer being able to create fields with different contents (using the HEX and regular formats), fields with malformed lengths, and so on.

Numbers in JSON do not have any bounds checking. Thus, integer values in a record might have invalid values, such as an ID field whose value is greater than or equal to 2^16, a QR field that has a value of 2, and so on.

9. Privacy Considerations

The values that can be contained in this format may contain privacy-sensitive information. For example, a profile of this format that is used for logging queries sent to recursive resolvers might have source IP addresses that could identify the location of the person who sent the DNS query.

10. References

10.1. Normative References

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <https://www.rfc-editor.org/info/rfc1035>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3339] Klyne, G. and C. Newman, "Date and Time on the Internet: Timestamps", RFC 3339, DOI 10.17487/RFC3339, July 2002, <https://www.rfc-editor.org/info/rfc3339>.

[RFC3597] Gustafsson, A., "Handling of Unknown DNS Resource Record (RR) Types", RFC 3597, DOI 10.17487/RFC3597, September 2003, <https://www.rfc-editor.org/info/rfc3597>.

[RFC4287] Nottingham, M., Ed. and R. Sayre, Ed., "The Atom Syndication Format", RFC 4287, DOI 10.17487/RFC4287, December 2005, <https://www.rfc-editor.org/info/rfc4287>.

[RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, DOI 10.17487/RFC4648, October 2006, <https://www.rfc-editor.org/info/rfc4648>.

[RFC5890] Klensin, J., "Internationalized Domain Names for Applications (IDNA): Definitions and Document Framework", RFC 5890, DOI 10.17487/RFC5890, August 2010, <https://www.rfc-editor.org/info/rfc5890>.

[RFC5952] Kawamura, S. and M. Kawashima, "A Recommendation for IPv6 Address Text Representation", RFC 5952, DOI 10.17487/RFC5952, August 2010, <https://www.rfc-editor.org/info/rfc5952>.

[RFC7464] Williams, N., "JavaScript Object Notation (JSON) Text Sequences", RFC 7464, DOI 10.17487/RFC7464, February 2015, <https://www.rfc-editor.org/info/rfc7464>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8259] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", STD 90, RFC 8259, DOI 10.17487/RFC8259, December 2017, <https://www.rfc-editor.org/info/rfc8259>.

10.2. Informative References

[DNS-QUERY] Parthasarathy, M. and P. Vixie, "Representing DNS messages using XML", Work in Progress, draft-mohan-dns-query-xml-00, September 2011.

[DNSXML] Daley, J., Morris, S., and J. Dickinson, "dnsxml - A standard XML representation of DNS data", Work in Progress, draft-daley-dnsxml-00, July 2013.

[PASSIVE-DNS] Dulaunoy, A., Kaplan, A., Vixie, P., and H. Stern, "Passive DNS - Common Output Format", Work in Progress, draft-dulaunoy-dnsop-passive-dns-cof-04, June 2018.

Acknowledgements

Some of the ideas in this document were inspired by earlier, abandoned work such as [DNSXML], [DNS-QUERY], and [PASSIVE-DNS]. The document was also inspired by early ideas from Stephane Bortzmeyer. Many people in the Domain Name System Operations (DNSOP) and DNS Over HTTPS (DOH) working groups contributed very useful ideas (even though this was not a WG work item).

Author’s Address

Paul Hoffman
ICANN

Email: paul.hoffman@icann.org