CoRE Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 27, 2021

M. Boucadair
Orange
J. Shallow

May 26, 2021

Constrained Application Protocol (CoAP) Block-Wise Transfer Options Supporting Robust Transmission draft-ietf-core-new-block-14

Abstract

This document specifies alternative Constrained Application Protocol (CoAP) Block-Wise transfer options: Q-Block1 and Q-Block2 Options.

These options are similar to, but distinct from, the CoAP Block1 and Block2 Options defined in RFC 7959. Q-Block1 and Q-Block2 Options are not intended to replace Block1 and Block2 Options, but rather have the goal of supporting Non-confirmable messages for large amounts of data with fewer packet interchanges. Also, the Q-Block1 and Q-Block2 Options support faster recovery should any of the blocks get lost in transmission.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on November 27, 2021.

Table of Contents

1. Introduction

The Constrained Application Protocol (CoAP) [RFC7252], although inspired by HTTP, was designed to use UDP instead of TCP. The message layer of CoAP over UDP includes support for reliable delivery, simple congestion control, and flow control. CoAP supports two message types (Section 1.2 of [RFC7252]): Confirmable (CON) and Non-confirmable (NON) messages. Unlike NON messages, every CON message will elicit an acknowledgement or a reset.

The CoAP specification recommends that a CoAP message should fit within a single IP packet (i.e., avoid IP fragmentation). To handle data records that cannot fit in a single IP packet, [RFC7959] introduced the concept of block-wise transfer and the companion CoAP Block1 and Block2 Options. However, this concept is designed to work exclusively with Confirmable messages (Section 1 of [RFC7959]). Note that the block-wise transfer was further updated by [RFC8323] for use over TCP, TLS, and WebSockets.

The CoAP Block1 and Block2 Options work well in environments where there are no, or minimal, packet losses. These options operate synchronously, i.e., each individual block has to be requested. A CoAP endpoint can only ask for (or send) the next block when the transfer of the previous block has completed. Packet transmission rate, and hence block transmission rate, is controlled by Round Trip Times (RTTs).

There is a requirement for blocks of data larger than a single IP datagram to be transmitted under network conditions where there may be asymmetrical transient packet loss (e.g., acknowledgment responses may get dropped). An example is when a network is subject to a Distributed Denial of Service (DDoS) attack and there is a need for DDoS mitigation agents relying upon CoAP to communicate with each other (e.g., [RFC8782][I-D.ietf-dots-telemetry]). As a reminder,

[RFC7959] recommends the use of CON responses to handle potential packet loss. However, such a recommendation does not work with a flooded pipe DDoS situation (e.g., [RFC8782]).

This document introduces the CoAP Q-Block1 and Q-Block2 Options which allow block-wise transfer to work with series of Non-confirmable messages, instead of lock-stepping using Confirmable messages (Section 3). In other words, this document provides a missing piece of [RFC7959], namely the support of block-wise transfer using Non-confirmable where an entire body of data can be transmitted without the requirement that intermediate acknowledgments be received from the peer (but recovery is available should it be needed).

Similar to [RFC7959], this specification does not remove any of the constraints posed by the base CoAP specification [RFC7252] it is strictly layered on top of.

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.

Readers should be familiar with the terms and concepts defined in [RFC7252], [RFC7959], and [RFC8132]. Particularly, the document uses the following key concepts:

Token: is used to match responses to requests independently from the underlying messages (Section 5.3.1 of [RFC7252]).

ETag: is used as a resource-local identifier for differentiating between representations of the same resource that vary over time (Section 5.10.6 of [RFC7252]).

The terms "payload" and "body" are defined in [RFC7959]. The term "payload" is thus used for the content of a single CoAP message (i.e., a single block being transferred), while the term "body" is used for the entire resource representation that is being transferred in a block-wise fashion.

Request-Tag refers to an option that allows a CoAP server to match message fragments belonging to the same request [I-D.ietf-core-echo-request-tag].

MAX_PAYLOADS is the maximum number of payloads that can be transmitted at any one time.

MAX_PAYLOADS_SET is the set of blocks identified by block numbers that, when divided by MAX_PAYLOADS, have the same numeric result. For example, if MAX_PAYLOADS is set to '10', a MAX_PAYLOADS_SET could be blocks #0 to #9, #10 to #19, etc. Depending on the overall data size, there could be fewer than MAX_PAYLOADS blocks in the final MAX_PAYLOADS_SET.

3. Alternative CoAP Block-Wise Transfer Options

This document introduces the CoAP Q-Block1 and Q-Block2 Options. These options are designed to work in particular with NON requests and responses.

Using NON messages, faster transmissions can occur as all the blocks can be transmitted serially (akin to fragmented IP packets) without having to wait for a response or next request from the remote CoAP peer. Recovery of missing blocks is faster in that multiple missing blocks can be requested in a single CoAP message. Even if there is asymmetrical packet loss, a body can still be sent and received by the peer whether the body comprises a single or multiple payloads, assuming no recovery is required.

A CoAP endpoint can acknowledge all or a subset of the blocks. Concretely, the receiving CoAP endpoint either informs the CoAP sender endpoint of successful reception or reports on all blocks in the body that have not yet been received. The CoAP sender endpoint will then retransmit only the blocks that have been lost in transmission.

Note that similar transmission rate benefits can be applied to Confirmable messages if the value of NSTART is increased from 1 (Section 4.7 of [RFC7252]). However, the use of Confirmable messages will not work effectively if there is asymmetrical packet loss. Some examples with Confirmable messages are provided in Appendix A.

There is little, if any, benefit of using these options with CoAP running over a reliable connection [RFC8323]. In this case, there is no differentiation between CON and NON as they are not used. Some examples using a reliable transport are provided in Appendix B.

Q-Block1 and Q-Block2 Options are similar in operation to the CoAP Block1 and Block2 Options, respectively. They are not a replacement for them, but have the following benefits:

- They can operate in environments where packet loss is highly asymmetrical.

- They enable faster transmissions of sets of blocks of data with fewer packet interchanges.
- They support faster recovery should any of the blocks get lost in transmission.
- They support sending an entire body using NON messages without requiring that an intermediate response be received from the peer.

There are the following disadvantages over using CoAP Block1 and Block2 Options:

- Loss of lock-stepping so payloads are not always received in the correct (block ascending) order.
- Additional congestion control measures need to be put in place for NON messages (Section 7.2).
- To reduce the transmission times for CON transmission of large bodies, NSTART needs to be increased from 1, but this affects congestion control and incurs a requirement to re-tune other parameters (Section 4.7 of [RFC7252]). Such tuning is out of scope of this document.
- Mixing of NON and CON during requests/responses using Q-Block is not supported.
- The Q-Block Options do not support stateless operation/random access.
- Proxying of Q-Block is limited to caching full representations.
- There is no multicast support.

Q-Block1 and Q-Block2 Options can be used instead of Block1 and Block2 Options when the different transmission properties are required. If the new options are not supported by a peer, then transmissions can fall back to using Block1 and Block2 Options (Section 4.1).

The deviations from Block1 and Block2 Options are specified in Section 4. Pointers to appropriate [RFC7959] sections are provided.

The specification refers to the base CoAP methods defined in Section 5.8 of [RFC7252] and the new CoAP methods, FETCH, PATCH, and iPATCH introduced in [RFC8132].

The No-Response Option [RFC7967] was considered but was abandoned as it does not apply to Q-Block2 responses. A unified solution is defined in the document.

3.1. CoAP Response Code (4.08) Usage

This document adds a media type for the 4.08 (Request Entity Incomplete) response defining an additional message format for reporting on payloads using the Q-Block1 Option that are not received by the server.

See Section 5 for more details.

3.2. Applicability Scope

The block-wise transfer specified in [RFC7959] covers the general case using Confirmable messages, but falls short in situations where packet loss is highly asymmetrical or there is no need for an acknowledgement. In other words, there is a need for Non-confirmable support.

The mechanism specified in this document provides roughly similar features to the Block1/Block2 Options. It provides additional properties that are tailored towards the intended use case of Non-confirmable transmission. Concretely, this mechanism primarily targets applications such as DDoS Open Threat Signaling (DOTS) that cannot use CON requests/responses because of potential packet loss and that support application-specific mechanisms to assess whether the remote peer is not overloaded and thus is able to process the messages sent by a CoAP endpoint (e.g., DOTS heartbeats in Section 4.7 of [RFC8782]). Other use cases are when an application sends data but has no need for an acknowledgement of receipt and, any data transmission loss is not critical.

The mechanism includes guards to prevent a CoAP agent from overloading the network by adopting an aggressive sending rate. These guards MUST be followed in addition to the existing CoAP congestion control as specified in Section 4.7 of [RFC7252]. See Section 7 for more details.

Any usage outside the primary use case of Non-confirmable with block transfers should be carefully weighed against the potential loss of interoperability with generic CoAP applications (See the disadvantages listed in Section 3). It is hoped that the experience gained with this mechanism can feed future extensions of the block-wise mechanism that will both be generally applicable and serve this particular use case.

It is not recommended that these options are used in a NoSec security mode (Section 9 of [RFC7252]) as the source endpoint needs to be trusted. Using OSCORE [RFC8613] does provide a security context and, hence, a trust of the source endpoint that prepared the inner OSCORE content. However, even with OSCORE, using a NoSec security mode with these options may still be inadequate, for reasons discussed in Section 11.

4. The Q-Block1 and Q-Block2 Options

4.1. Properties of Q-Block1 and Q-Block2 Options

The properties of the Q-Block1 and Q-Block2 Options are shown in Table 1. The formatting of this table follows the one used in Table 4 of [RFC7252] (Section 5.10). The C, U, N, and R columns indicate the properties Critical, UnSafe, NoCacheKey, and Repeatable defined in Section 5.4 of [RFC7252]. Only Critical and UnSafe columns are marked for the Q-Block1 Option. Critical, UnSafe, and Repeatable columns are marked for the Q-Block2 Option. As these options are UnSafe, NoCacheKey has no meaning and so is marked with a dash.

Number	C	U	N	R	Name	Format	Length	Default
TBA1	x	x	-		Q-Block1	uint	0-3	(none)
TBA2	x	x	-	x	Q-Block2	uint	0-3	(none)

Table 1: CoAP Q-Block1 and Q-Block2 Option Properties

The Q-Block1 and Q-Block2 Options can be present in both the request and response messages. The Q-Block1 Option pertains to the request payload and the Q-Block2 Option pertains to the response payload. When the Content-Format Option is present together with the Q-Block1 or Q-Block2 Option, the option applies to the body not to the payload (i.e., it must be the same for all payloads of the same body).

The Q-Block1 Option is useful with the payload-bearing, e.g., POST, PUT, FETCH, PATCH, and iPATCH requests and their responses.

The Q-Block2 Option is useful, e.g., with GET, POST, PUT, FETCH, PATCH, and iPATCH requests and their payload-bearing responses (response codes 2.01, 2.02, 2.04, and 2.05) (Section 5.5 of [RFC7252]).

A CoAP endpoint (or proxy) MUST support either both or neither of the Q-Block1 and Q-Block2 Options.

If the Q-Block1 Option is present in a request or the Q-Block2 Option is returned in a response, this indicates a block-wise transfer and describes how this specific block-wise payload forms part of the entire body being transferred. If it is present in the opposite direction, it provides additional control on how that payload will be formed or was processed.

To indicate support for Q-Block2 responses, the CoAP client MUST include the Q-Block2 Option in a GET or similar request (FETCH, for example), the Q-Block2 Option in a PUT or similar request (POST, for example), or the Q-Block1 Option in a PUT or similar request so that the server knows that the client supports this Q-Block functionality should it need to send back a body that spans multiple payloads. Otherwise, the server would use the Block2 Option (if supported) to send back a message body that is too large to fit into a single IP packet [RFC7959].

How a client decides whether it needs to include a Q-Block1 or Q-Block2 Option can be driven by a local configuration parameter, triggered by an application (DOTS, for example), etc. Such considerations are out of the scope of the document.

Implementation of the Q-Block1 and Q-Block2 Options is intended to be optional. However, when it is present in a CoAP message, it MUST be processed (or the message rejected). Therefore, Q-Block1 and Q-Block2 Options are identified as Critical options.

With CoAP over UDP, the way a request message is rejected for critical options depends on the message type. A Confirmable message with an unrecognized critical option is rejected with a 4.02 (Bad Option) response (Section 5.4.1 of [RFC7252]). A Non-confirmable message with an unrecognized critical option is either rejected with a Reset message or just silently ignored (Sections 5.4.1 and 4.3 of [RFC7252]). To reliably get a rejection message, it is therefore REQUIRED that clients use a Confirmable message for determining support for Q-Block1 and Q-Block2 Options. This CON message can be sent under the base CoAP congestion control setup specified in Section 4.7 of [RFC7252] (that is, NSTART does not need to be increased (Section 7.1)).

The Q-Block1 and Q-Block2 Options are unsafe to forward. That is, a CoAP proxy that does not understand the Q-Block1 (or Q-Block2) Option must reject the request or response that uses either option (See Section 5.7.1 of [RFC7252]).

The Q-Block2 Option is repeatable when requesting retransmission of missing blocks, but not otherwise. Except that case, any request

carrying multiple Q-Block1 (or Q-Block2) Options MUST be handled following the procedure specified in Section 5.4.5 of [RFC7252].

The Q-Block1 and Q-Block2 Options, like the Block1 and Block2 Options, are of both class E and class U for OSCORE processing (Table 2). The Q-Block1 (or Q-Block2) Option MAY be an Inner or Outer option (Section 4.1 of [RFC8613]). The Inner and Outer values are therefore independent of each other. The Inner option is encrypted and integrity protected between clients and servers, and provides message body identification in case of end-to-end fragmentation of requests. The Outer option is visible to proxies and labels message bodies in case of hop-by-hop fragmentation of requests.

Number	Name	E	U
TBA1	Q-Block1	x	x
TBA2	Q-Block2	x	x

Table 2: OSCORE Protection of Q-Block1 and Q-Block2 Options

Note that if Q-Block1 or Q-Block2 Options are included in a packet as Inner options, Block1 or Block2 Options MUST NOT be included as Inner options. Similarly, there MUST NOT be a mix of Q-Block and Block for the Outer options. Messages that do not adhere with this behavior MUST be rejected with 4.02 (Bad Option). Q-Block and Block Options can be mixed across Inner and Outer options as these are handled independently of each other. For clarity, if OSCORE is not being used, there MUST NOT be a mix of Q-Block and Block Options in the same packet.

4.2. Structure of Q-Block1 and Q-Block2 Options

The structure of Q-Block1 and Q-Block2 Options follows the structure defined in Section 2.2 of [RFC7959].

There is no default value for the Q-Block1 and Q-Block2 Options. Absence of one of these options is equivalent to an option value of 0 with respect to the value of block number (NUM) and more bit (M) that could be given in the option, i.e., it indicates that the current block is the first and only block of the transfer (block number is set to 0, M is unset). However, in contrast to the explicit value 0, which would indicate a size of the block (SZX) of 0, and thus a size value of 16 bytes, there is no specific explicit size implied by the absence of the option -- the size is left unspecified. (As for any uint, the explicit value 0 is efficiently indicated by a zero-length

option; this, therefore, is different in semantics from the absence of the option).

4.3. Using the Q-Block1 Option

The Q-Block1 Option is used when the client wants to send a large amount of data to the server using the POST, PUT, FETCH, PATCH, or iPATCH methods where the data and headers do not fit into a single packet.

When Q-Block1 Option is used, the client MUST include a Request-Tag Option [I-D.ietf-core-echo-request-tag]. The Request-Tag value MUST be the same for all of the requests for the body of data that is being transferred. The Request-Tag is opaque, but the client MUST ensure that it is unique for every different body of transmitted data.

> Implementation Note: It is suggested that the client treats the Request-Tag as an unsigned integer of 8 bytes in length. An implementation may want to consider limiting this to 4 bytes to reduce packet overhead size. The initial Request-Tag value should be randomly generated and then subsequently incremented by the client whenever a new body of data is being transmitted between peers.

Section 4.6 discusses the use of Size1 Option.

For Confirmable transmission, the server continues to acknowledge each packet, but a response is not required (whether separate or piggybacked) until successful receipt of the body by the server. For Non-confirmable transmission, no response is required until either the successful receipt of the body by the server or a timer expires with some of the payloads having not yet arrived. In the latter case, a "retransmit missing payloads" response is needed. For reliable transports (e.g., [RFC8323]), a response is not required until successful receipt of the body by the server.

Each individual message that carries a block of the body is treated as a new request (Section 6).

The client MUST send the payloads in order of increasing block number, starting from zero, until the body is complete (subject to any congestion control (Section 7)). Any missing payloads requested by the server must in addition be separately transmitted with increasing block numbers.

The following Response Codes are used:

2.01 (Created)

This Response Code indicates successful receipt of the entire body and that the resource was created. The token to use MUST be one of the tokens that were received in a request for this block-wise exchange. However, it is desirable to provide the one used in the last received request, since that will aid any troubleshooting. The client should then release all of the tokens used for this body. Note that the last received payload might not be the one with the highest block number.

2.02 (Deleted)

This Response Code indicates successful receipt of the entire body and that the resource was deleted when using POST (Section 5.8.2 [RFC7252]). The token to use MUST be one of the tokens that were received in a request for this block-wise exchange. However, it is desirable to provide the one used in the last received request. The client should then release all of the tokens used for this body.

2.04 (Changed)

This Response Code indicates successful receipt of the entire body and that the resource was updated. The token to use MUST be one of the tokens that were received in a request for this block-wise exchange. However, it is desirable to provide the one used in the last received request. The client should then release all of the tokens used for this body.

2.05 (Content)

This Response Code indicates successful receipt of the entire FETCH request body (Section 2 of [RFC8132]) and that the appropriate representation of the resource is being returned. The token to use MUST be one of the tokens that were received in a request for this block-wise exchange. However, it is desirable to provide the one used in the last received request.

If the FETCH request includes the Observe Option, then the server MUST use the same token as used for the 2.05 (Content) response for returning any Observe triggered responses so that the client can match them up.

The client should then release all of the tokens used for this body apart from the one used for tracking an observed resource.

2.31 (Continue)

This Response Code can be used to indicate that all of the blocks up to and including the Q-Block1 Option block NUM (all having the M bit set) have been successfully received. The token to use MUST be one of the tokens that were received in a request for this latest MAX_PAYLOADS_SET block-wise exchange. However, it is desirable to provide the one used in the last received request.

The client should then release all of the tokens used for this MAX_PAYLOADS_SET.

A response using this Response Code MUST NOT be generated for every received Q-Block1 Option request. It SHOULD only be generated when all the payload requests are Non-confirmable and a MAX_PAYLOADS_SET has been received by the server. More details about the motivations for this optimization are discussed in Section 7.2.

This Response Code SHOULD NOT be generated for CON as this may cause duplicated payloads to unnecessarily be sent.

4.00 (Bad Request)

This Response Code MUST be returned if the request does not include a Request-Tag Option or a Size1 Option but does include a Q-Block1 option.

4.02 (Bad Option)

This Response Code MUST be returned for a Confirmable request if the server does not support the Q-Block Options. Note that a reset message may be sent in case of Non-confirmable request.

4.08 (Request Entity Incomplete)

As a reminder, this Response Code returned without Content-Type "application/missing-blocks+cbor-seq" (Section 12.3) is handled as in Section 2.9.2 [RFC7959].

This Response Code returned with Content-Type "application/missing-blocks+cbor-seq" indicates that some of the payloads are missing and need to be resent. The client then retransmits the individual missing payloads using the same Request-Tag, Size1, and, Q-Block1 Option to specify the same NUM, SZX, and M bit as sent initially in the original, but not received, packet.

The Request-Tag value to use is determined by taking the token in the 4.08 (Request Entity Incomplete) response, locating the matching client request, and then using its Request-Tag.

The token to use in the 4.08 (Request Entity Incomplete) response MUST be one of the tokens that were received in a request for this block-wise body exchange. However, it is desirable to provide the one used in the last received request. See Section 5 for further information.

If the server has not received all the blocks of a body, but one or more NON payloads have been received, it SHOULD wait for NON_RECEIVE_TIMEOUT (Section 7.2) before sending a 4.08 (Request Entity Incomplete) response.

4.13 (Request Entity Too Large)

This Response Code can be returned under similar conditions to those discussed in Section 2.9.3 of [RFC7959].

This Response Code can be returned if there is insufficient space to create a response PDU with a block size of 16 bytes (SZX = 0) to send back all the response options as appropriate. In this case, the Size1 Option is not included in the response.

Further considerations related to the transmission timings of 4.08 (Request Entity Incomplete) and 2.31 (Continue) Response Codes are discussed in Section 7.2.

If a server receives payloads with different Request-Tags for the same resource, it should continue to process all the bodies as it has no way of determining which is the latest version, or which body, if any, the client is terminating the transmission for.

If the client elects to stop the transmission of a complete body, and absent any local policy, the client MUST "forget" all tracked tokens associated with the body's Request-Tag so that a reset message is generated for the invalid token in the 4.08 (Request Entity Incomplete) response. The server on receipt of the reset message SHOULD delete the partial body.

If the server receives a duplicate block with the same Request-Tag, it MUST ignore the payload of the packet, but MUST still respond as if the block was received for the first time.

A server SHOULD maintain a partial body (missing payloads) for NON_PARTIAL_TIMEOUT (Section 7.2).

4.4. Using the Q-Block2 Option

In a request for any block number, the M bit unset indicates the request is just for that block. If the M bit is set, this has different meanings based on the NUM value:

NUM is zero: This is a request for the entire body.

'NUM modulo MAX_PAYLOADS' is zero, while NUM is not zero: This is used to confirm that the current MAX_PAYLOADS_SET (the latest block having block number NUM-1) has been successfully received and that, upon receipt of this request, the server can continue to send the next MAX_PAYLOADS_SET (the first block having block number NUM). This is the 'Continue' Q-Block-2 and conceptually has the same usage (i.e., continue sending the next set of data) as the use of 2.31 (Continue) for Q-Block1.

Any other value of NUM: This is a request for that block and for all of the remaining blocks in the current MAX_PAYLOADS_SET.

If the request includes multiple Q-Block2 Options and these options overlap (e.g., combination of M being set (this and later blocks) and being unset (this individual block)) resulting in an individual block being requested multiple times, the server MUST only send back one instance of that block. This behavior is meant to prevent amplification attacks.

The payloads sent back from the server as a response MUST all have the same ETag (Section 5.10.6 of [RFC7252]) for the same body. The server MUST NOT use the same ETag value for different representations of a resource.

The ETag is opaque, but the server MUST ensure that it is unique for every different body of transmitted data.

> Implementation Note: It is suggested that the server treats the ETag as an unsigned integer of 8 bytes in length. An implementation may want to consider limiting this to 4 bytes to reduce packet overhead size. The initial ETag value should be randomly generated and then subsequently incremented by the server whenever a new body of data is being transmitted between peers.

Section 4.6 discusses the use of Size2 Option.

The client may elect to request any detected missing blocks or just ignore the partial body. This decision is implementation specific.

For NON payloads, the client SHOULD wait NON_RECEIVE_TIMEOUT (Section 7.2) after the last received payload before requesting retransmission of any missing blocks. Retransmission is requested by issuing a GET, POST, PUT, FETCH, PATCH, or iPATCH request that contains one or more Q-Block2 Options that define the missing block(s). Generally the M bit on the Q-Block2 Option(s) SHOULD be unset, although the M bit MAY be set to request this and later blocks from this MAX_PAYLOADS_SET, see Section 10.2.4 for an example of this in operation. Further considerations related to the transmission timing for missing requests are discussed in Section 7.2.

The missing block numbers requested by the client MUST have an increasing block number in each additional Q-Block2 Option with no duplicates. The server SHOULD respond with a 4.00 (Bad Request) to requests not adhering to this behavior. Note that the ordering constraint is meant to force the client to check for duplicates and remove them. This also helps with troubleshooting.

If the client receives a duplicate block with the same ETag, it MUST silently ignore the payload.

A client SHOULD maintain a partial body (missing payloads) for NON_PARTIAL_TIMEOUT (Section 7.2) or as defined by the Max-Age Option (or its default of 60 seconds (Section 5.6.1 of [RFC7252])), whichever is the less. On release of the partial body, the client should then release all of the tokens used for this body apart from the token that is used to track a resource that is being observed.

The ETag Option should not be used in the request for missing blocks as the server could respond with a 2.03 (Valid) response with no payload. It can be used in the request if the client wants to check the freshness of the locally cached body response.

The server SHOULD maintain a cached copy of the body when using the Q-Block2 Option to facilitate retransmission of any missing payloads.

If the server detects part way through a body transfer that the resource data has changed and the server is not maintaining a cached copy of the old data, then the transmission is terminated. Any subsequent missing block requests MUST be responded to using the latest ETag and Size2 Option values with the updated data.

If the server responds during a body update with a different ETag Option value (as the resource representation has changed), then the client should treat the partial body with the old ETag as no longer being fresh. The client may then request all of the new data by specifying Q-Block2 with block number '0' and the M bit set.

If the server transmits a new body of data (e.g., a triggered Observe notification) with a new ETag to the same client as an additional response, the client should remove any partially received body held for a previous ETag for that resource as it is unlikely the missing blocks can be retrieved.

If there is insufficient space to create a response PDU with a block size of 16 bytes (SZX = 0) to send back all the response options as appropriate, a 4.13 (Request Entity Too Large) is returned without the Size1 Option.

For Confirmable traffic, the server typically acknowledges the initial request using an ACK with a piggybacked payload, and then sends the subsequent payloads of the MAX_PAYLOADS_SET as CON responses. These CON responses are individually ACKed by the client. The server will detect failure to send a packet and SHOULD terminate the body transfer, but the client can issue, after a MAX_TRANSMIT_SPAN delay, a separate GET, POST, PUT, FETCH, PATCH, or iPATCH for any missing blocks as needed.

4.5. Using Observe Option

For a request that uses Q-Block1, the Observe value [RFC7641] MUST be the same for all the payloads of the same body. This includes any missing payloads that are retransmitted.

For a response that uses Q-Block2, the Observe value MUST be the same for all the payloads of the same body. This is different from Block2 usage where the Observe value is only present in the first block (Section 3.4 of [RFC7959]). This includes payloads transmitted following receipt of the 'Continue' Q-Block2 Option (Section 4.4) by the server. If a missing payload is requested by a client, then both the request and response MUST NOT include the Observe Option.

4.6. Using Size1 and Size2 Options

Section 4 of [RFC7959] defines two CoAP options: Size1 for indicating the size of the representation transferred in requests and Size2 for indicating the size of the representation transferred in responses.

For Q-Block1 and Q-Block2 Options, the Size1 or Size2 Option values MUST exactly represent the size of the data on the body so that any missing data can easily be determined.

The Size1 Option MUST be used with the Q-Block1 Option when used in a request and MUST be present in all payloads of the request, preserving the same value. The Size2 Option MUST be used with the

Q-Block2 Option when used in a response and MUST be present in all payloads of the response, preserving the same value.

4.7. Using Q-Block1 and Q-Block2 Options Together

The behavior is similar to the one defined in Section 3.3 of [RFC7959] with Q-Block1 substituted for Block1 and Q-Block2 for Block2.

4.8. Using Q-Block2 Option With Multicast

Servers MUST ignore multicast requests that contain the Q-Block2 Option. As a reminder, Block2 Option can be used as stated in Section 2.8 of [RFC7959].

5. The Use of 4.08 (Request Entity Incomplete) Response Code

4.08 (Request Entity Incomplete) Response Code has a new Content-Type "application/missing-blocks+cbor-seq" used to indicate that the server has not received all of the blocks of the request body that it needs to proceed. Such messages must not be treated by the client as a fatal error.

Likely causes are the client has not sent all blocks, some blocks were dropped during transmission, or the client has sent them sufficiently long ago that the server has already discarded them.

The new data payload of the 4.08 (Request Entity Incomplete) response with Content-Type set to "application/missing-blocks+cbor-seq" is encoded as a CBOR Sequence [RFC8742]. It comprises one or more missing block numbers encoded as CBOR unsigned integers [RFC8949]. The missing block numbers MUST be unique in each 4.08 (Request Entity Incomplete) response when created by the server; the client MUST ignore any duplicates in the same 4.08 (Request Entity Incomplete) response.

The Content-Format Option (Section 5.10.3 of [RFC7252]) MUST be used in the 4.08 (Request Entity Incomplete) response. It MUST be set to "application/missing-blocks+cbor-seq" (Section 12.3).

The Concise Data Definition Language [RFC8610] (and see Section 4.1 [RFC8742]) for the data describing these missing blocks is as follows:

```
; This defines an array, the elements of which are to be used
; in a CBOR Sequence:
payload = [+ missing-block-number]
; A unique block number not received:
missing-block-number = uint
```

Figure 1: Structure of the Missing Blocks Payload

This CDDL syntax MUST be followed.

It is desirable that the token to use for the response is the token that was used in the last block number received so far with the same Request-Tag value. Note that the use of any received token with the same Request-Tag would be acceptable, but providing the one used in the last received payload will aid any troubleshooting. The client will use the token to determine what was the previously sent request to obtain the Request-Tag value that was used.

If the size of the 4.08 (Request Entity Incomplete) response packet is larger than that defined by Section 4.6 [RFC7252], then the number of reported missing blocks MUST be limited so that the response can fit into a single packet. If this is the case, then the server can send subsequent 4.08 (Request Entity Incomplete) responses containing the missing other blocks on receipt of a new request providing a missing payload with the same Request-Tag.

The missing blocks MUST be reported in ascending order without any duplicates. The client SHOULD silently drop 4.08 (Request Entity Incomplete) responses not adhering with this behavior.

Implementation Note: Consider limiting the number of missing payloads to MAX_PAYLOADS to minimize congestion control being needed. The CBOR sequence does not include any array wrapper.

The 4.08 (Request Entity Incomplete) with Content-Type "application/missing-blocks+cbor-seq" SHOULD NOT be used when using Confirmable requests or a reliable connection [RFC8323] as the client will be able to determine that there is a transmission failure of a particular payload and hence that the server is missing that payload.

6. The Use of Tokens

Each new request generally uses a new Token (and sometimes must, see Section 4 of [I-D.ietf-core-echo-request-tag]). Additional responses to a request all use the token of the request they respond to.

Implementation Note: By using 8-byte tokens, it is possible to easily minimize the number of tokens that have to be tracked by

clients, by keeping the bottom 32 bits the same for the same body and the upper 32 bits containing the current body's request number (incrementing every request, including every re-transmit). This allows the client to be alleviated from keeping all the per-request-state, e.g., in Section 3 of [RFC8974]. However, if using NoSec, Section 5.2 of [RFC8974] needs to be considered for security implications.

7. Congestion Control for Unreliable Transports

The transmission of all the blocks of a single body over an unreliable transport MUST either all be Confirmable or all be Non-confirmable. This is meant to simplify the congestion control procedure.

As a reminder, there is no need for CoAP-specific congestion control for reliable transports [RFC8323].

7.1. Confirmable (CON)

Congestion control for CON requests and responses is specified in Section 4.7 of [RFC7252]. In order to benefit from faster transmission rates, NSTART will need to be increased from 1. However, the other CON congestion control parameters will need to be tuned to cover this change. This tuning is not specified in this document, given that the applicability scope of the current specification assumes that all requests and responses using Q-Block1 and Q-Block2 will be Non-confirmable (Section 3.2) apart from the initial Q-Block functionality negotiation.

Following the failure to transmit a packet due to packet loss after MAX_TRANSMIT_SPAN time (Section 4.8.2 of [RFC7252]), it is implementation specific as to whether there should be any further requests for missing data.

7.2. Non-confirmable (NON)

This document introduces new parameters MAX_PAYLOADS, NON_TIMEOUT, NON_TIMEOUT_RANDOM, NON_RECEIVE_TIMEOUT, NON_MAX_RETRANSMIT, NON_PROBING_WAIT, and NON_PARTIAL_TIMEOUT primarily for use with NON (Table 3).

Note: Randomness may naturally be provided based on the traffic profile, how PROBING_RATE is computed (as this is an average), and when the peer responds. Randomness is explicitly added for some of the congestion control parameters to handle situations where every thing is in sync when retrying.

MAX_PAYLOADS should be configurable with a default value of 10. Both CoAP endpoints MUST have the same value (otherwise there will be transmission delays in one direction) and the value MAY be negotiated between the endpoints to a common value by using a higher level protocol (out of scope of this document). This is the maximum number of payloads that can be transmitted at any one time.

> Note: The default value of 10 is chosen for reasons similar to those discussed in Section 5 of [RFC6928], especially given the target application discussed in Section 3.2.

NON_TIMEOUT is used to compute the delay between sending MAX_PAYLOADS_SET for the same body. By default, NON_TIMEOUT has the same value as ACK_TIMEOUT (Section 4.8 of [RFC7252]).

NON_TIMEOUT_RANDOM is the initial actual delay between sending the first two MAX_PAYLOADS_SETs of the same body. The same delay is then used between the subsequent MAX_PAYLOADS_SETs. It is a random duration (not an integral number of seconds) between NON_TIMEOUT and (NON_TIMEOUT * ACK_RANDOM_FACTOR). ACK_RANDOM_FACTOR is set to 1.5 as discussed in Section 4.8 of [RFC7252].

NON_RECEIVE_TIMEOUT is the initial time to wait for a missing payload before requesting retransmission for the first time. Every time the missing payload is re-requested, the time to wait value doubles. The time to wait is calculated as:

```
Time-to-Wait = NON_RECEIVE_TIMEOUT * (2 ** (Re-Request-Count - 1))
```

NON_RECEIVE_TIMEOUT has a default value of twice NON_TIMEOUT. NON_RECEIVE_TIMEOUT MUST always be greater than NON_TIMEOUT_RANDOM by at least one second so that the sender of the payloads has the opportunity to start sending the next MAX_PAYLOADS_SET before the receiver times out.

NON_MAX_RETRANSMIT is the maximum number of times a request for the retransmission of missing payloads can occur without a response from the remote peer. After this occurs, the local endpoint SHOULD consider the body stale, remove any body, and release Tokens and Request-Tag on the client (or the ETag on the server). By default, NON_MAX_RETRANSMIT has the same value as MAX_RETRANSMIT (Section 4.8 of [RFC7252]).

NON_PROBING_WAIT is used to limit the potential wait needed when using PROBING_RATE. By default, NON_PROBING_WAIT is computed in a similar way as EXCHANGE_LIFETIME (Section 4.8.2 of [RFC7252]) but with ACK_TIMEOUT, MAX_RETRANSMIT, and PROCESSING_DELAY substituted

with NON_TIMEOUT, NON_MAX_RETRANSMIT, and NON_TIMEOUT_RANDOM, respectively:

```
NON_PROBING_WAIT = NON_TIMEOUT * ((2 ** NON_MAX_RETRANSMIT) - 1) *
ACK_RANDOM_FACTOR + (2 * MAX_LATENCY) + NON_TIMEOUT_RANDOM
```

NON_PARTIAL_TIMEOUT is used for expiring partially received bodies. By default, NON_PARTIAL_TIMEOUT is computed in the same way as EXCHANGE_LIFETIME (Section 4.8.2 of [RFC7252]) but with ACK_TIMEOUT and MAX_RETRANSMIT substituted with NON_TIMEOUT and NON_MAX_RETRANSMIT, respectively:

```
NON_PARTIAL_TIMEOUT = NON_TIMEOUT * ((2 ** NON_MAX_RETRANSMIT) -
1) * ACK_RANDOM_FACTOR + (2 * MAX_LATENCY) + NON_TIMEOUT
```

Parameter Name	Default Value
MAX_PAYLOADS	10
NON_MAX_RETRANSMIT	4
NON_TIMEOUT	2 s
NON_TIMEOUT_RANDOM	between 2-3 s
NON_RECEIVE_TIMEOUT	4 s
NON_PROBING_WAIT	between 247-248 s
NON_PARTIAL_TIMEOUT	247 s

Table 3: Congestion Control Parameters

The PROBING_RATE parameter in CoAP indicates the average data rate that must not be exceeded by a CoAP endpoint in sending to a peer endpoint that does not respond. A single body will be subjected to PROBING_RATE (Section 4.7 of [RFC7252]), not the individual packets. If the wait time between sending bodies that are not being responded to based on PROBING_RATE exceeds NON_PROBING_WAIT, then the wait time is limited to NON_PROBING_WAIT.

Note: For the particular DOTS application, PROBING_RATE and other transmission parameters are negotiated between peers. Even when not negotiated, the DOTS application uses customized defaults as discussed in Section 4.5.2 of [RFC8782]. Note that MAX_PAYLOADS, NON_MAX_RETRANSMIT, NON_TIMEOUT, NON_PROBING_WAIT, and NON_PARTIAL_TIMEOUT can be negotiated between DOTS peers, e.g., as per [I-D.bosh-dots-quick-blocks]. When explicit values are configured for NON_PROBING_WAIT and NON_PARTIAL_TIMEOUT, these values are used without applying any jitter.

Each NON 4.08 (Request Entity Incomplete) response is subject to PROBING_RATE.

Each NON GET or FETCH request using a Q-Block2 Option is subject to PROBING_RATE.

As the sending of many payloads of a single body may itself cause congestion, after transmission of every MAX_PAYLOADS_SET of a single body, a delay MUST be introduced of NON_TIMEOUT_RANDOM before sending the next MAX_PAYLOADS_SET unless a 'Continue' is received from the peer for the current MAX_PAYLOADS_SET, in which case the next MAX_PAYLOADS_SET MAY start transmission immediately.

> Note: Assuming 1500-byte packets and the MAX_PAYLOADS_SET having 10 payloads, this corresponds to 1500 * 10 * 8 = 120 Kbits. With a delay of 2 seconds between MAX_PAYLOADS_SET, this indicates an average speed requirement of 60 Kbps for a single body should there be no responses. This transmission rate is further reduced by being subject to PROBING_RATE.

The sending of a set of missing blocks of a body is restricted to those in a MAX_PAYLOADS_SET at a time. In other words, a NON_TIMEOUT_RANDOM delay is still observed between each MAX_PAYLOAD_SET.

For Q-Block1 Option, if the server responds with a 2.31 (Continue) Response Code for the latest payload sent, then the client can continue to send the next MAX_PAYLOADS_SET without any further delay. If the server responds with a 4.08 (Request Entity Incomplete) Response Code, then the missing payloads SHOULD be retransmitted before going into another NON_TIMEOUT_RANDOM delay prior to sending the next set of payloads.

For the server receiving NON Q-Block1 requests, it SHOULD send back a 2.31 (Continue) Response Code on receipt of all of the MAX_PAYLOADS_SET to prevent the client unnecessarily delaying. If not all of the MAX_PAYLOADS_SET were received, the server SHOULD delay for NON_RECEIVE_TIMEOUT (exponentially scaled based on the repeat request count for a payload) before sending the 4.08 (Request Entity Incomplete) Response Code for the missing payload(s). If all of the MAX_PAYLOADS_SET were received and a 2.31 (Continue) had been sent, but no more payloads were received for NON_RECEIVE_TIMEOUT (exponentially scaled), the server SHOULD send a 4.08 (Request Entity Incomplete) response detailing the missing payloads after the block number that was indicated in the sent 2.31 (Continue). If the repeated response count of the 4.08 (Request Entity Incomplete) exceeds NON_MAX_RETRANSMIT, the server SHOULD discard the partial body and stop requesting the missing payloads.

It is likely that the client will start transmitting the next MAX_PAYLOADS_SET before the server times out on waiting for the last of the previous MAX_PAYLOADS_SET. On receipt of a payload from the next MAX_PAYLOADS_SET, the server SHOULD send a 4.08 (Request Entity Incomplete) Response Code indicating any missing payloads from any previous MAX_PAYLOADS_SET. Upon receipt of the 4.08 (Request Entity Incomplete) Response Code, the client SHOULD send the missing payloads before continuing to send the remainder of the MAX_PAYLOADS_SET and then go into another NON_TIMEOUT_RANDOM delay prior to sending the next MAX_PAYLOADS_SET.

For the client receiving NON Q-Block2 responses, it SHOULD send a 'Continue' Q-Block2 request (Section 4.4) for the next MAX_PAYLOADS_SET on receipt of all of the MAX_PAYLOADS_SET, to prevent the server unnecessarily delaying. Otherwise the client SHOULD delay for NON_RECEIVE_TIMEOUT (exponentially scaled based on the repeat request count for a payload), before sending the request for the missing payload(s). If the repeat request count for a missing payload exceeds NON_MAX_RETRANSMIT, the client SHOULD discard the partial body and stop requesting the missing payloads.

The server SHOULD recognize the 'Continue' Q-Block2 request as a continue request and just continue the transmission of the body (including Observe Option, if appropriate for an unsolicited response) rather than as a request for the remaining missing blocks.

It is likely that the server will start transmitting the next MAX_PAYLOADS_SET before the client times out on waiting for the last of the previous MAX_PAYLOADS_SET. Upon receipt of a payload from the new MAX_PAYLOADS_SET, the client SHOULD send a request indicating any missing payloads from any previous MAX_PAYLOADS_SET. Upon receipt of such request, the server SHOULD send the missing payloads before continuing to send the remainder of the MAX_PAYLOADS_SET and then go into another NON_TIMEOUT_RANDOM delay prior to sending the next MAX_PAYLOADS_SET.

The client does not need to acknowledge the receipt of the entire body.

> Note: If there is asymmetric traffic loss causing responses to never get received, a delay of NON_TIMEOUT_RANDOM after every transmission of MAX_PAYLOADS_SET will be observed. The endpoint receiving the body is still likely to receive the entire body.

8. Caching Considerations

Caching block based information is not straight forward in a proxy. For Q-Block1 and Q-Block2 Options, for simplicity it is expected that the proxy will reassemble the body (using any appropriate recovery options for packet loss) before passing on the body to the appropriate CoAP endpoint. This does not preclude an implementation doing a more complex per payload caching, but how to do this is out of the scope of this document. The onward transmission of the body does not require the use of the Q-Block1 or Q-Block2 Options as these options may not be supported in that link. This means that the proxy must fully support the Q-Block1 and Q-Block2 Options.

How the body is cached in the CoAP client (for Q-Block1 transmissions) or the CoAP server (for Q-Block2 transmissions) is implementation specific.

As the entire body is being cached in the proxy, the Q-Block1 and Q-Block2 Options are removed as part of the block assembly and thus do not reach the cache.

For Q-Block2 responses, the ETag Option value is associated with the data (and onward transmitted to the CoAP client), but is not part of the cache key.

For requests with Q-Block1 Option, the Request-Tag Option is associated with the build up of the body from successive payloads, but is not part of the cache key. For the onward transmission of the body using CoAP, a new Request-Tag SHOULD be generated and used. Ideally this new Request-Tag should replace the client's request Request-Tag.

It is possible that two or more CoAP clients are concurrently updating the same resource through a common proxy to the same CoAP server using Q-Block1 (or Block1) Option. If this is the case, the first client to complete building the body causes that body to start transmitting to the CoAP server with an appropriate Request-Tag value. When the next client completes building the body, any existing partial body transmission to the CoAP server is terminated and the new body representation transmission starts with a new Request-Tag value. Note that it cannot be assumed that the proxy will always receive a complete body from a client.

A proxy that supports Q-Block2 Option MUST be prepared to receive a GET or similar request indicating one or more missing blocks. The proxy will serve from its cache the missing blocks that are available in its cache in the same way a server would send all the appropriate Q-Block2 responses. If a body matching the cache key is not

available in the cache, the proxy MUST request the entire body from the CoAP server using the information in the cache key.

How long a CoAP endpoint (or proxy) keeps the body in its cache is implementation specific (e.g., it may be based on Max-Age).

9. HTTP-Mapping Considerations

As a reminder, the basic normative requirements on HTTP/CoAP mappings are defined in Section 10 of [RFC7252]. The implementation guidelines for HTTP/CoAP mappings are elaborated in [RFC8075].

The rules defined in Section 5 of [RFC7959] are to be followed.

10. Examples with Non-confirmable Messages

This section provides some sample flows to illustrate the use of Q-Block1 and Q-Block2 Options with NON. Examples with CON are provided in Appendix A.

The examples in the following subsections assume MAX_PAYLOADS is set to 10 and NON_MAX_RETRANSMIT is set to 4.

Figure 2 lists the conventions that are used in the following subsections.

```
           T: Token value
           O: Observe Option value
           M: Message ID
          RT: Request-Tag
          ET: ETag
         QB1: Q-Block1 Option values NUM/More/Size
         QB2: Q-Block2 Option values NUM/More/Size
        Size: Actual block size encoded in SZX
          \: Trimming long lines
        [[]]: Comments
        -->X: Message loss (request)
        X<--: Message loss (response)
         ...: Passage of time
  Payload N: Corresponds to the CoAP message that conveys
             a block number (N-1) of a given block-wise exchange.
```

Figure 2: Notations Used in the Figures

10.1. Q-Block1 Option

10.1.1. A Simple Example

Figure 3 depicts an example of a NON PUT request conveying Q-Block1 Option. All the blocks are received by the server.

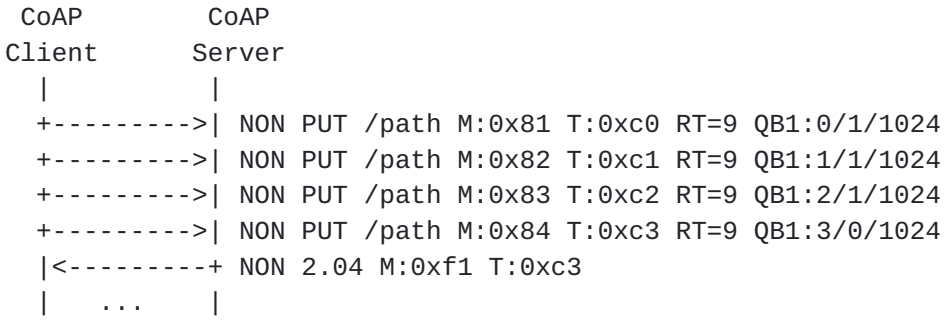

Figure 3: Example of NON Request with Q-Block1 Option (Without Loss)

10.1.2. Handling MAX_PAYLOADS Limits

Figure 4 depicts an example of a NON PUT request conveying Q-Block1 Option. The number of payloads exceeds MAX_PAYLOADS. All the blocks are received by the server.

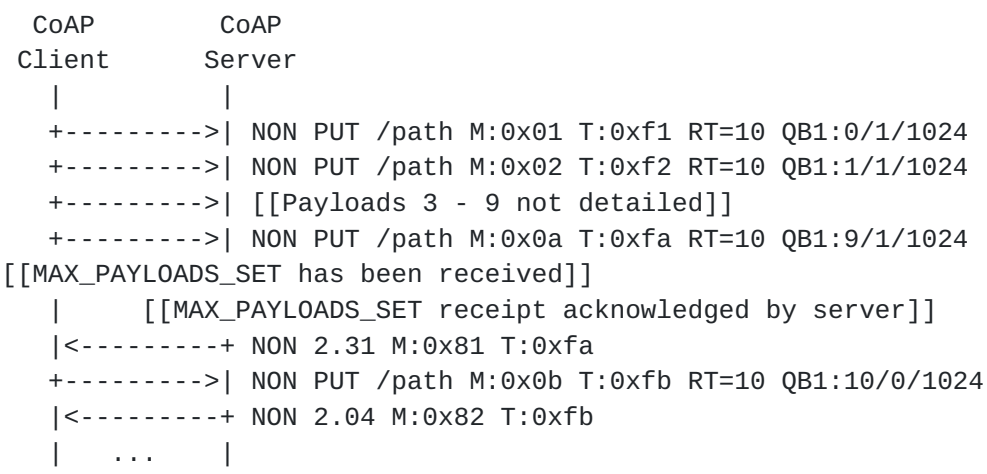

Figure 4: Example of MAX_PAYLOADS NON Request with Q-Block1 Option (Without Loss)

10.1.3. Handling MAX_PAYLOADS with Recovery

Consider now a scenario where a new body of data is to be sent by the client, but some blocks are dropped in transmission as illustrated in Figure 5.

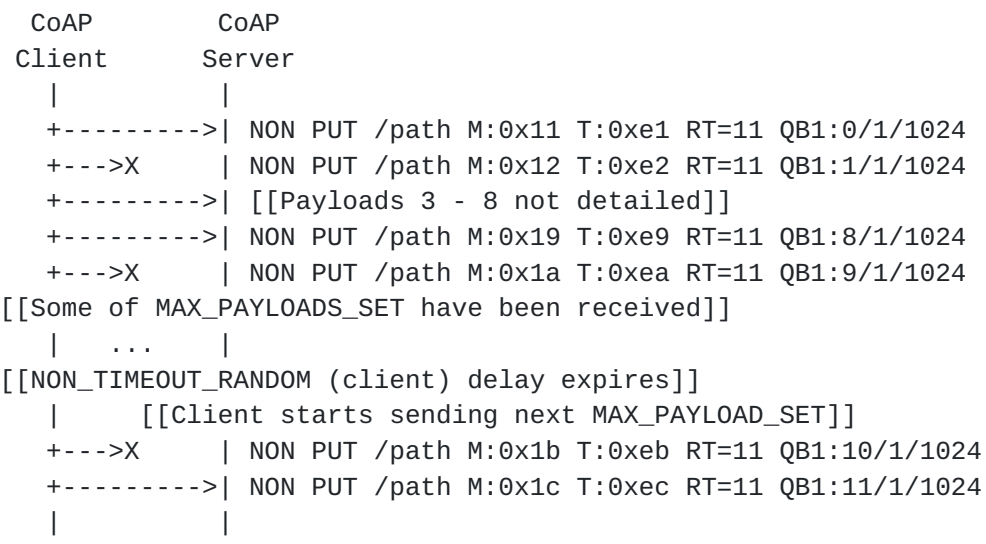

Figure 5: Example of MAX_PAYLOADS NON Request with Q-Block1 Option (With Loss)

On seeing a payload from the next MAX_PAYLOAD_SET, the server realizes that some blocks are missing from the previous MAX_PAYLOADS_SET and asks for the missing blocks in one go (Figure 6). It does so by indicating which blocks from the previous MAX_PAYLOADS_SET have not been received in the data portion of the response (Section 5). The token used in the response should be the token that was used in the last received payload. The client can then derive the Request-Tag by matching the token with the sent request.

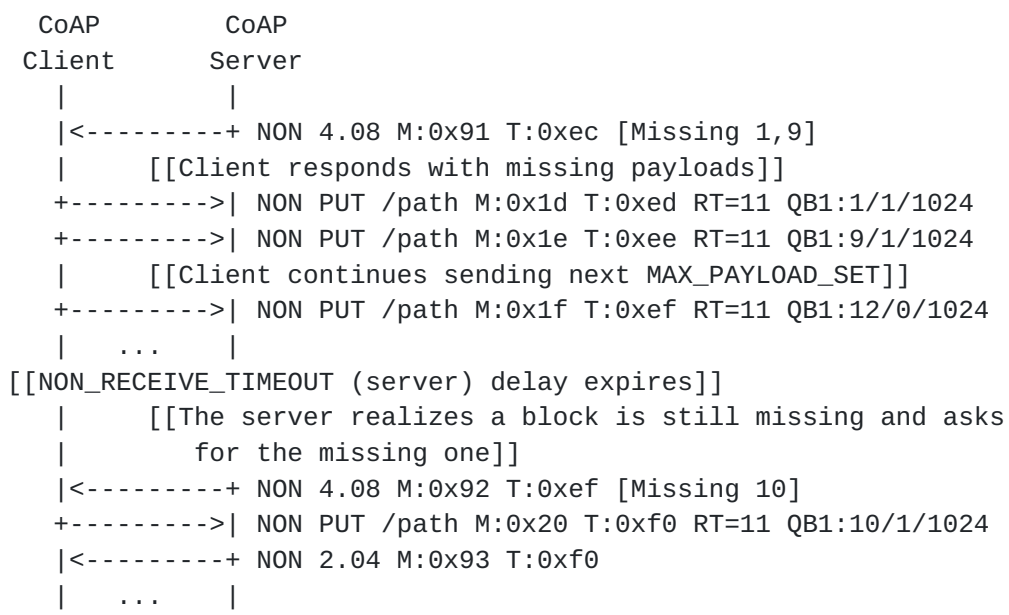

Figure 6: Example of NON Request with Q-Block1 Option (Blocks Recovery)

10.1.4. Handling Recovery with Failure

Figure 7 depicts an example of a NON PUT request conveying Q-Block1 Option where recovery takes place, but eventually fails.

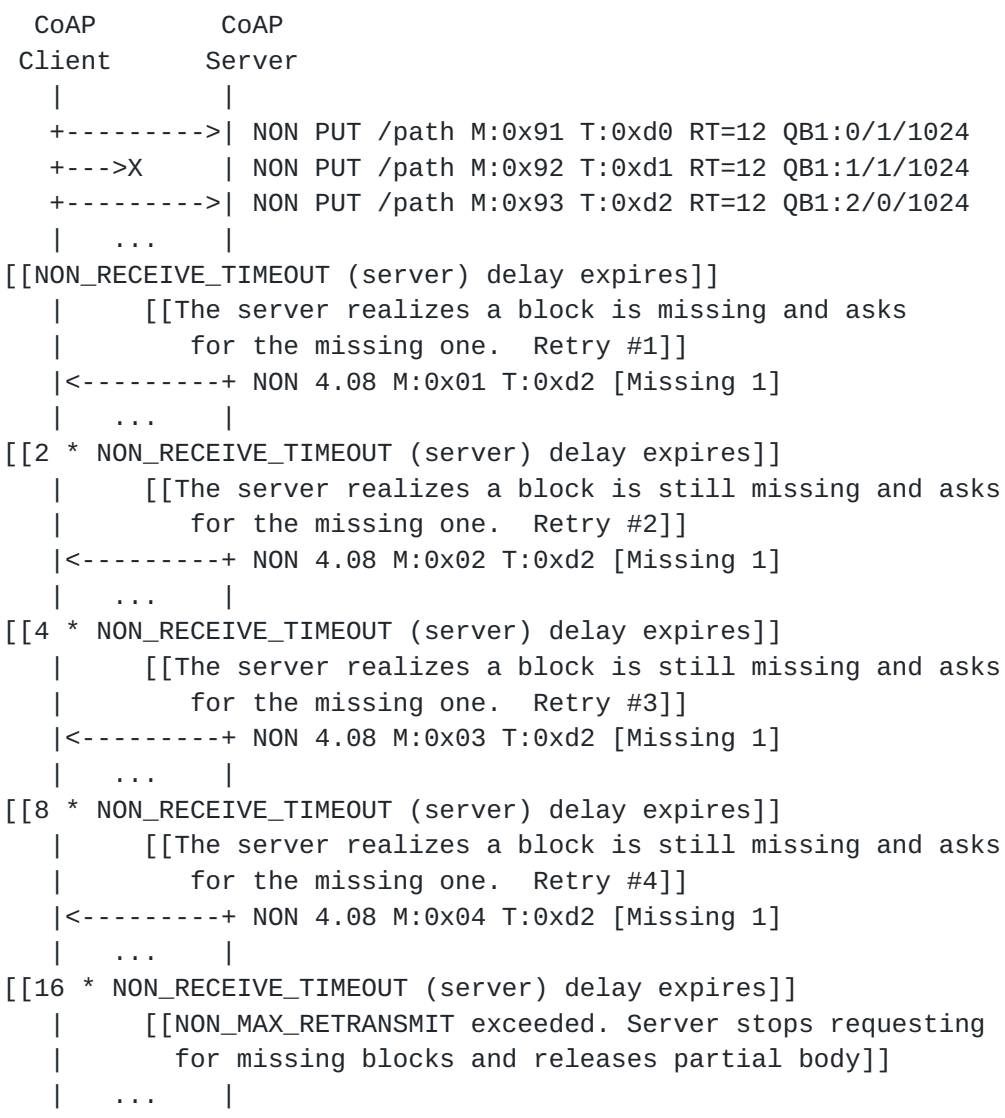

Figure 7: Example of NON Request with Q-Block1 Option (With Eventual Failure)

10.2. Q-Block2 Option

These examples include the Observe Option to demonstrate how that option is used. Note that the Observe Option is not required for Q-Block2; the observe detail can thus be ignored.

10.2.1. A Simple Example

Figure 8 illustrates the example of Q-Block2 Option. The client sends a NON GET carrying Observe and Q-Block2 Options. The Q-Block2 Option indicates a block size hint (1024 bytes). This request is replied to by the server using four (4) blocks that are transmitted to the client without any loss. Each of these blocks carries a

Q-Block2 Option. The same process is repeated when an Observe is triggered, but no loss is experienced by any of the notification blocks.

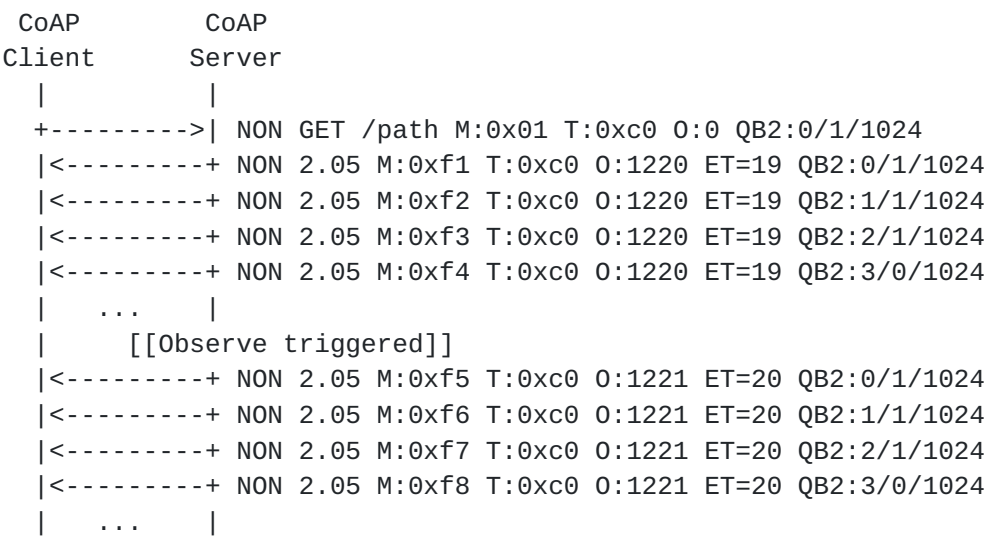

Figure 8: Example of NON Notifications with Q-Block2 Option (Without Loss)

10.2.2. Handling MAX_PAYLOADS Limits

Figure 9 illustrates the same as Figure 8 but this time has eleven (11) payloads which exceeds MAX_PAYLOADS. There is no loss experienced.

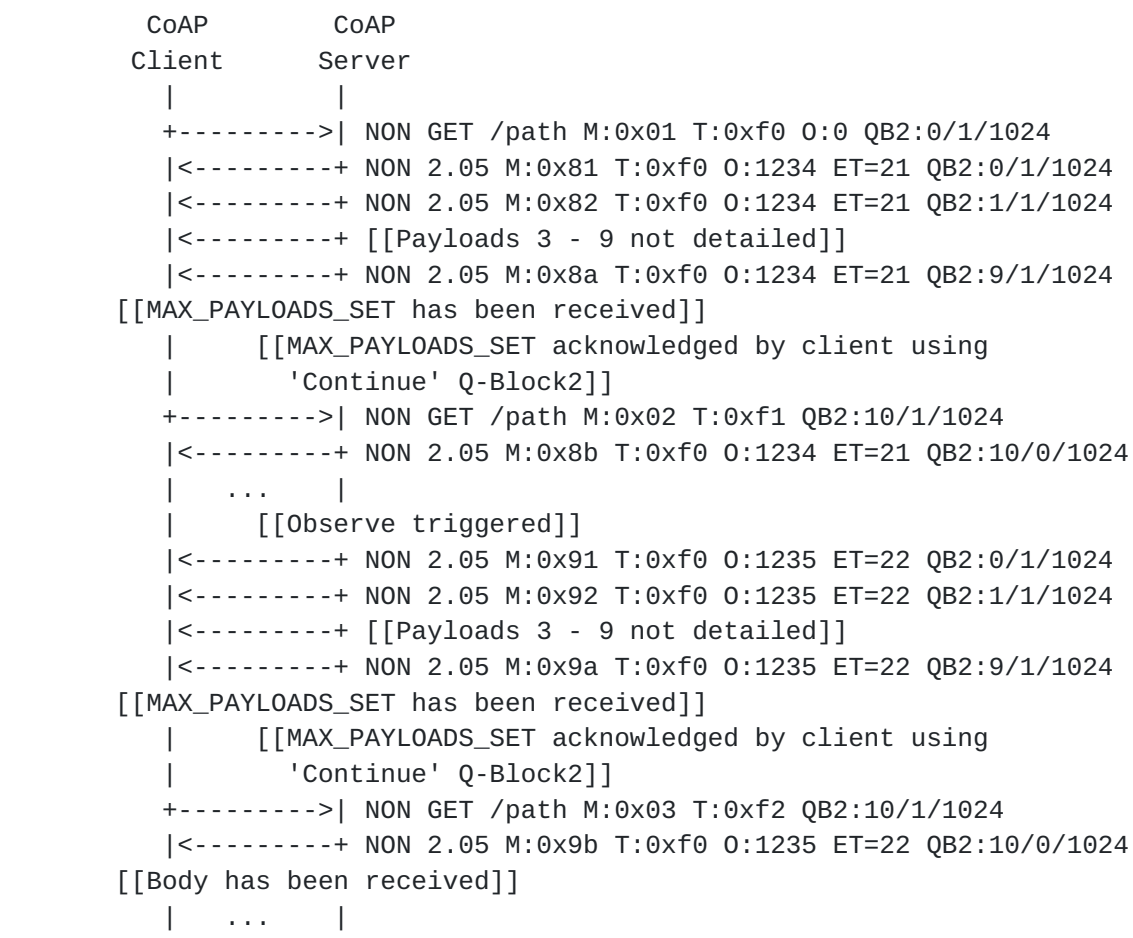

Figure 9: Example of NON Notifications with Q-Block2 Option (Without Loss)

10.2.3. Handling MAX_PAYLOADS with Recovery

Figure 10 shows the example of an Observe that is triggered but for which some notification blocks are lost. The client detects the missing blocks and requests their retransmission. It does so by indicating the blocks that are missing as one or more Q-Block2 Options.

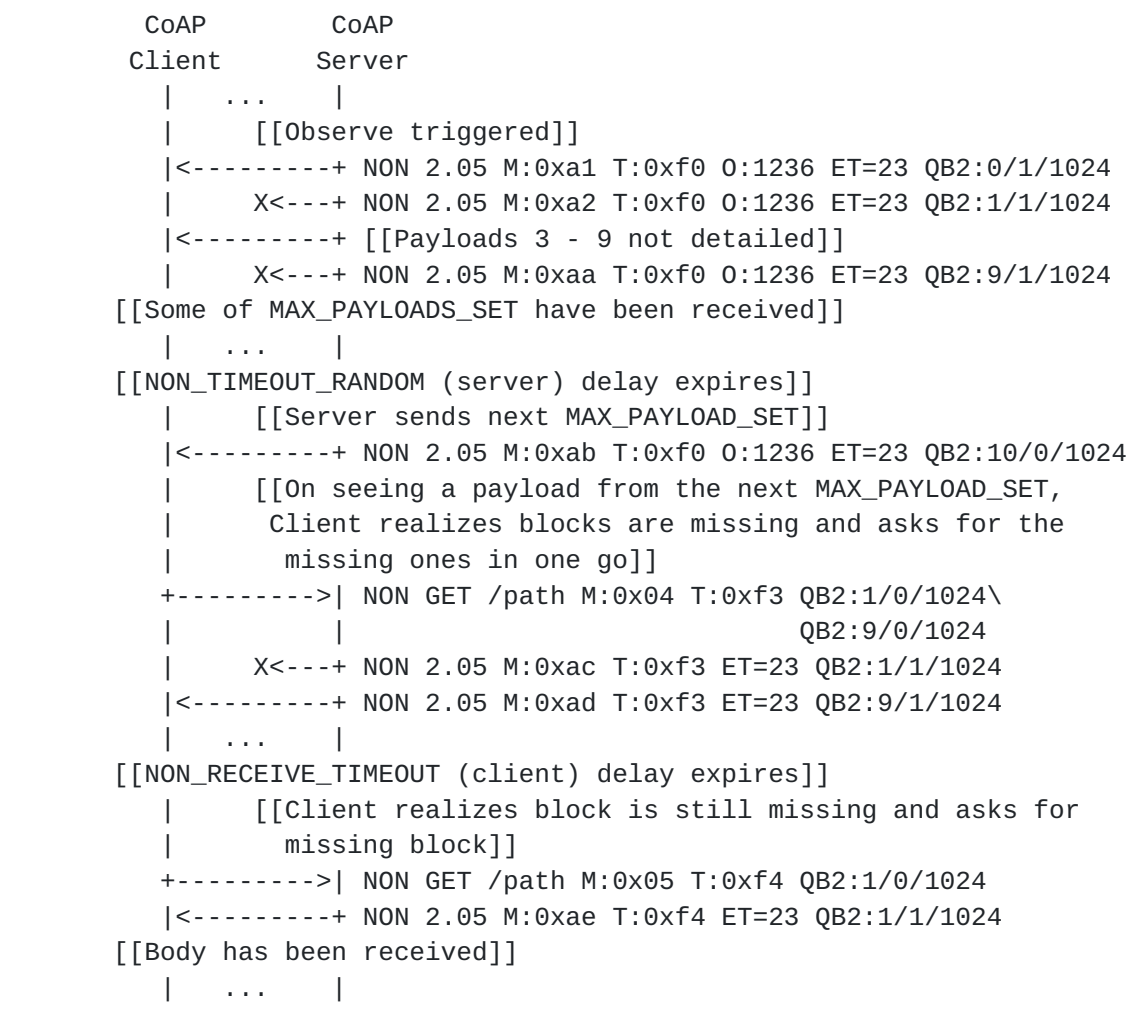

Figure 10: Example of NON Notifications with Q-Block2 Option (Blocks Recovery)

10.2.4. Handling Recovery using M-bit Set

Figure 11 shows the example of an Observe that is triggered but only the first two notification blocks reach the client. In order to retrieve the missing blocks, the client sends a request with a single Q-Block2 Option with the M bit set.

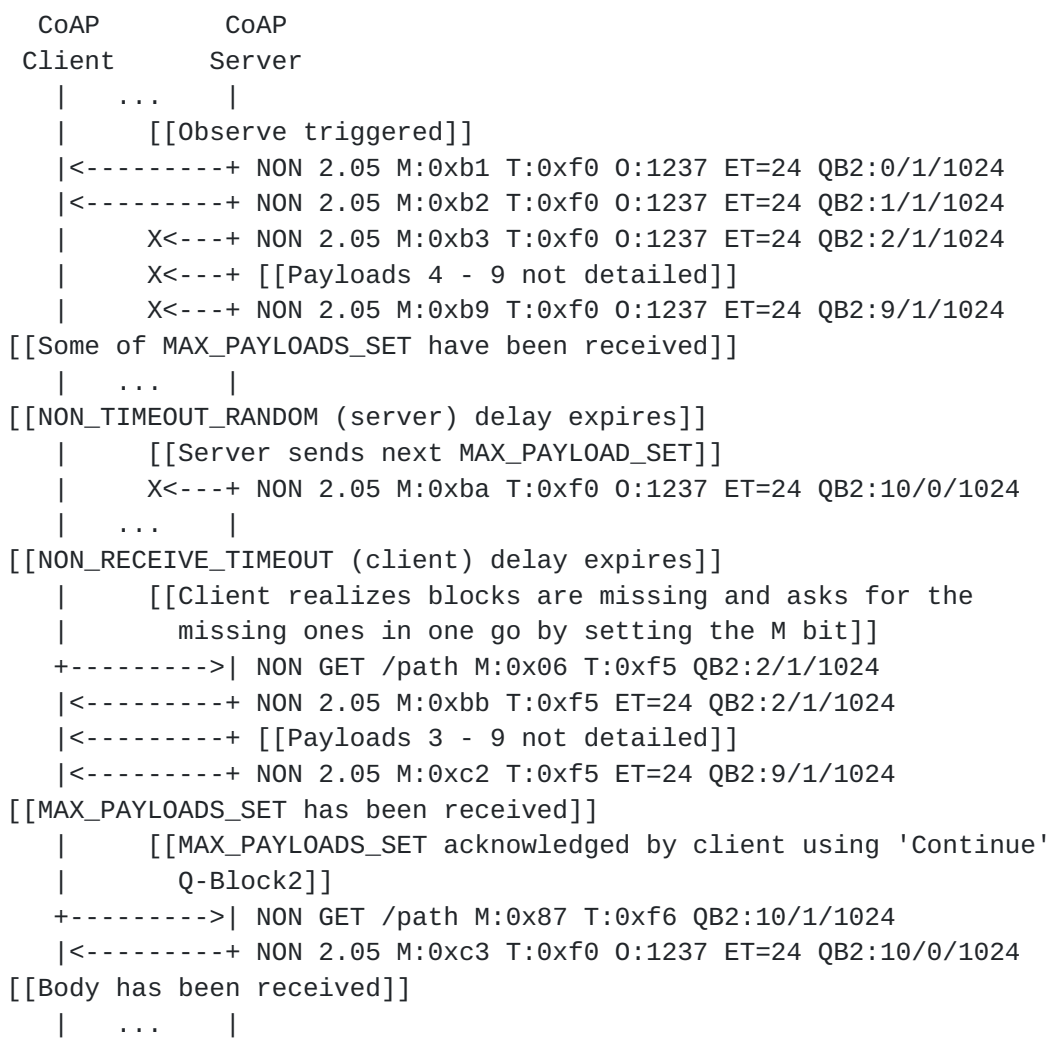

Figure 11: Example of NON Notifications with Q-Block2 Option (Blocks Recovery with M bit Set)

10.3. Q-Block1 and Q-Block2 Options

10.3.1. A Simple Example

Figure 12 illustrates the example of a FETCH using both Q-Block1 and Q-Block2 Options along with an Observe Option. No loss is experienced.

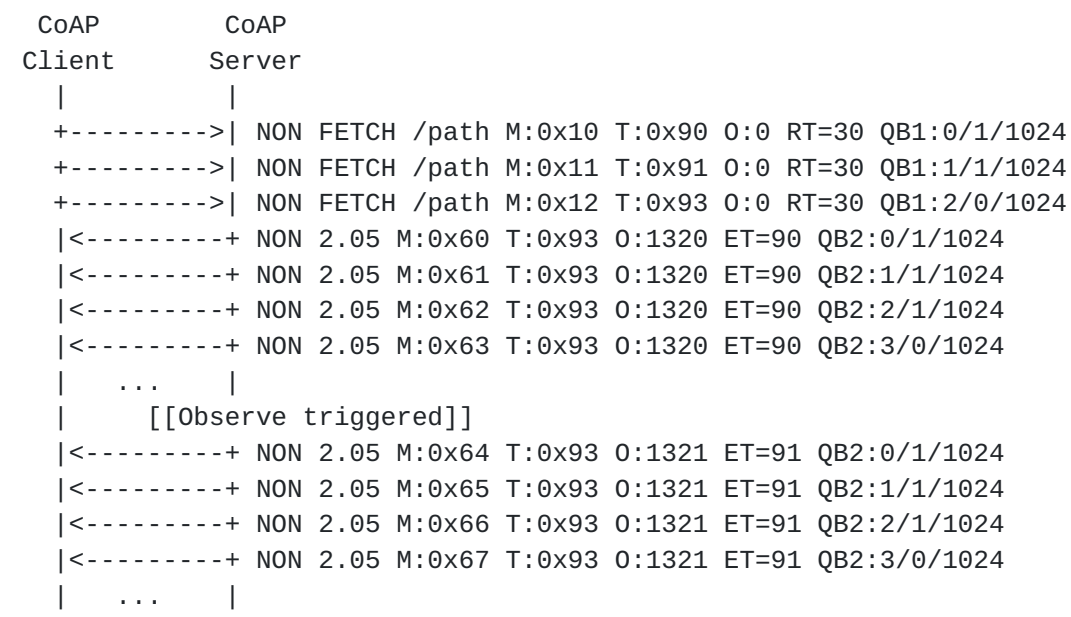

Figure 12: Example of NON FETCH with Q-Block1 and Q-Block2 Options (Without Loss)

10.3.2. Handling MAX_PAYLOADS Limits

Figure 13 illustrates the same as Figure 12 but this time has eleven (11) payloads in both directions which exceeds MAX_PAYLOADS. There is no loss experienced.

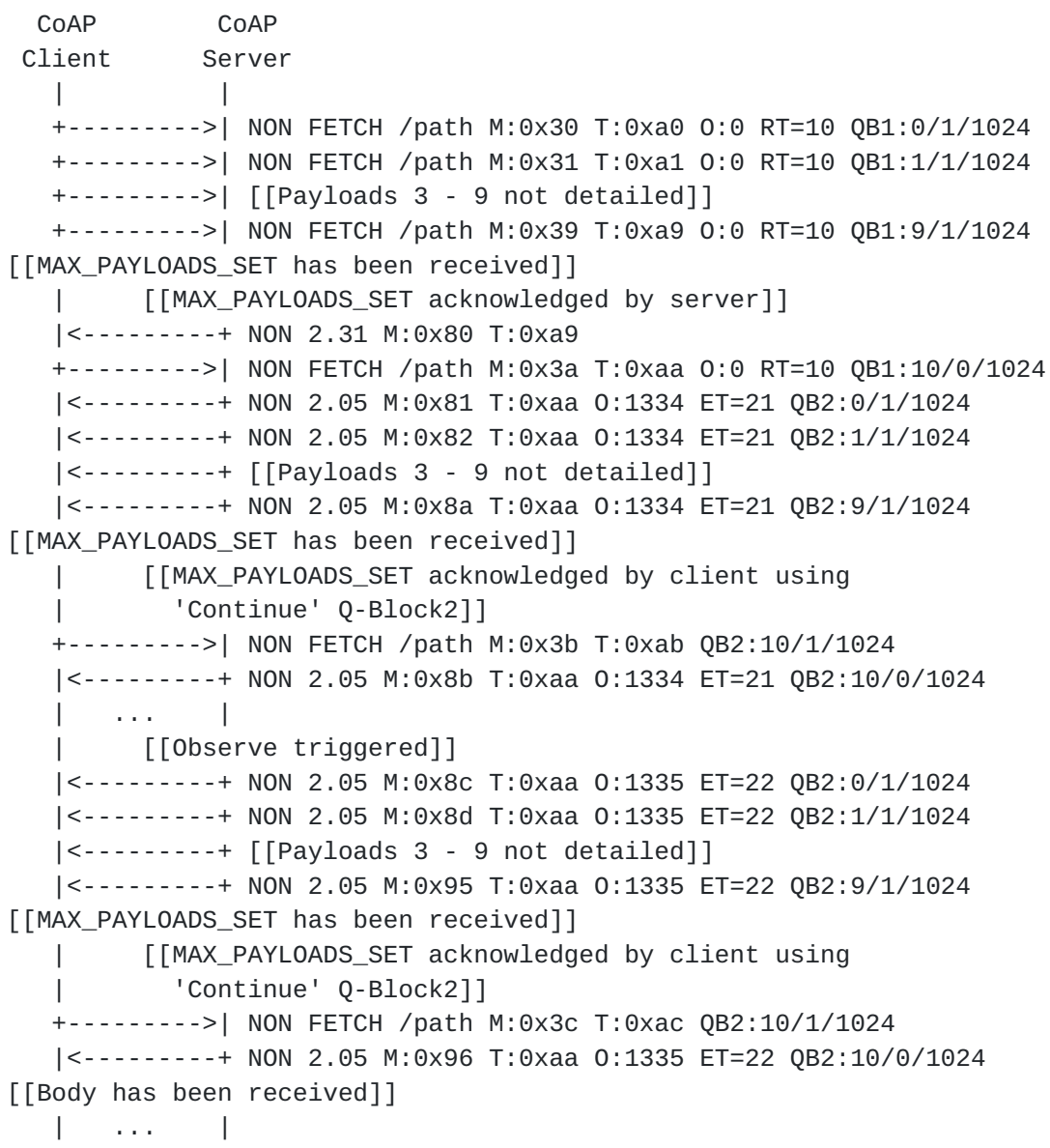

Figure 13: Example of NON FETCH with Q-Block1 and Q-Block2 Options (Without Loss)

Note that as 'Continue' was used, the server continues to use the same token (0xaa) since the 'Continue' is not being used as a request for a new set of packets, but rather is being used to instruct the server to continue its transmission (Section 7.2).

10.3.3. Handling Recovery

Consider now a scenario where some blocks are lost in transmission as illustrated in Figure 14.

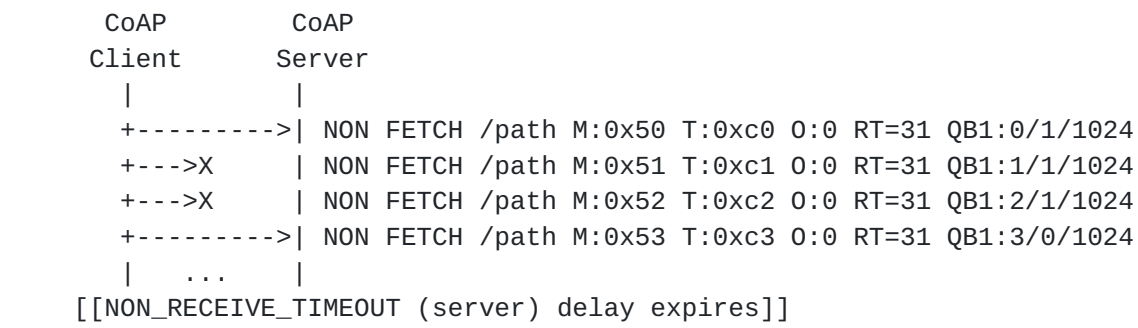

Figure 14: Example of NON FETCH with Q-Block1 and Q-Block2 Options (With Loss)

The server realizes that some blocks are missing and asks for the missing blocks in one go (Figure 15). It does so by indicating which blocks have not been received in the data portion of the response. The token used in the response is the token that was used in the last received payload. The client can then derive the Request-Tag by matching the token with the sent request.

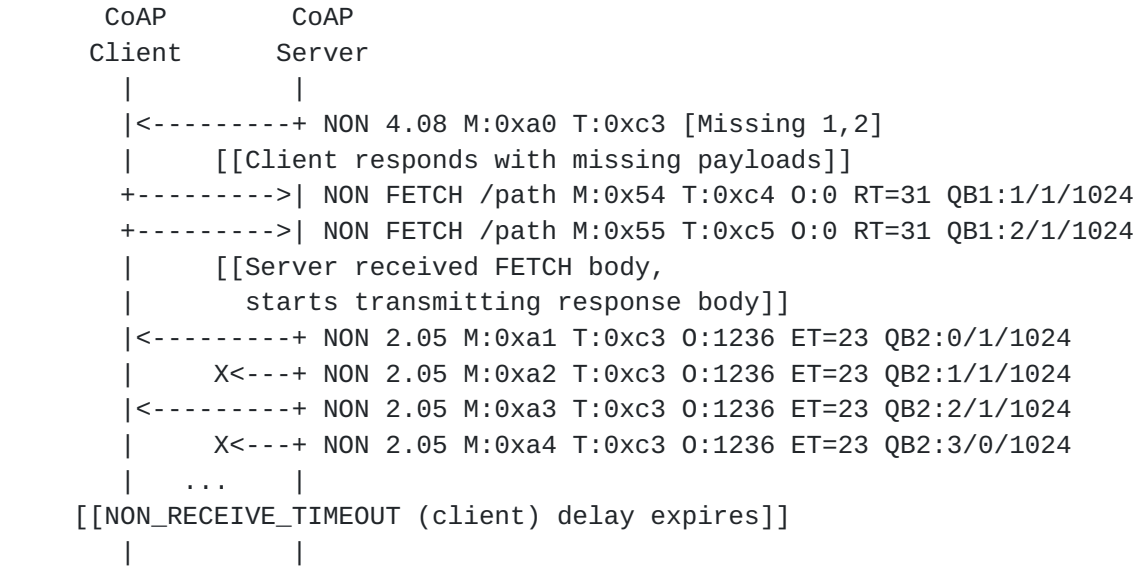

Figure 15: Example of NON Request with Q-Block1 Option (Server Recovery)

The client realizes that not all the payloads of the response have been returned. The client then asks for the missing blocks in one go (Figure 16). Note that, following Section 2.7 of [RFC7959], the FETCH request does not include the Q-Block1 or any payload.

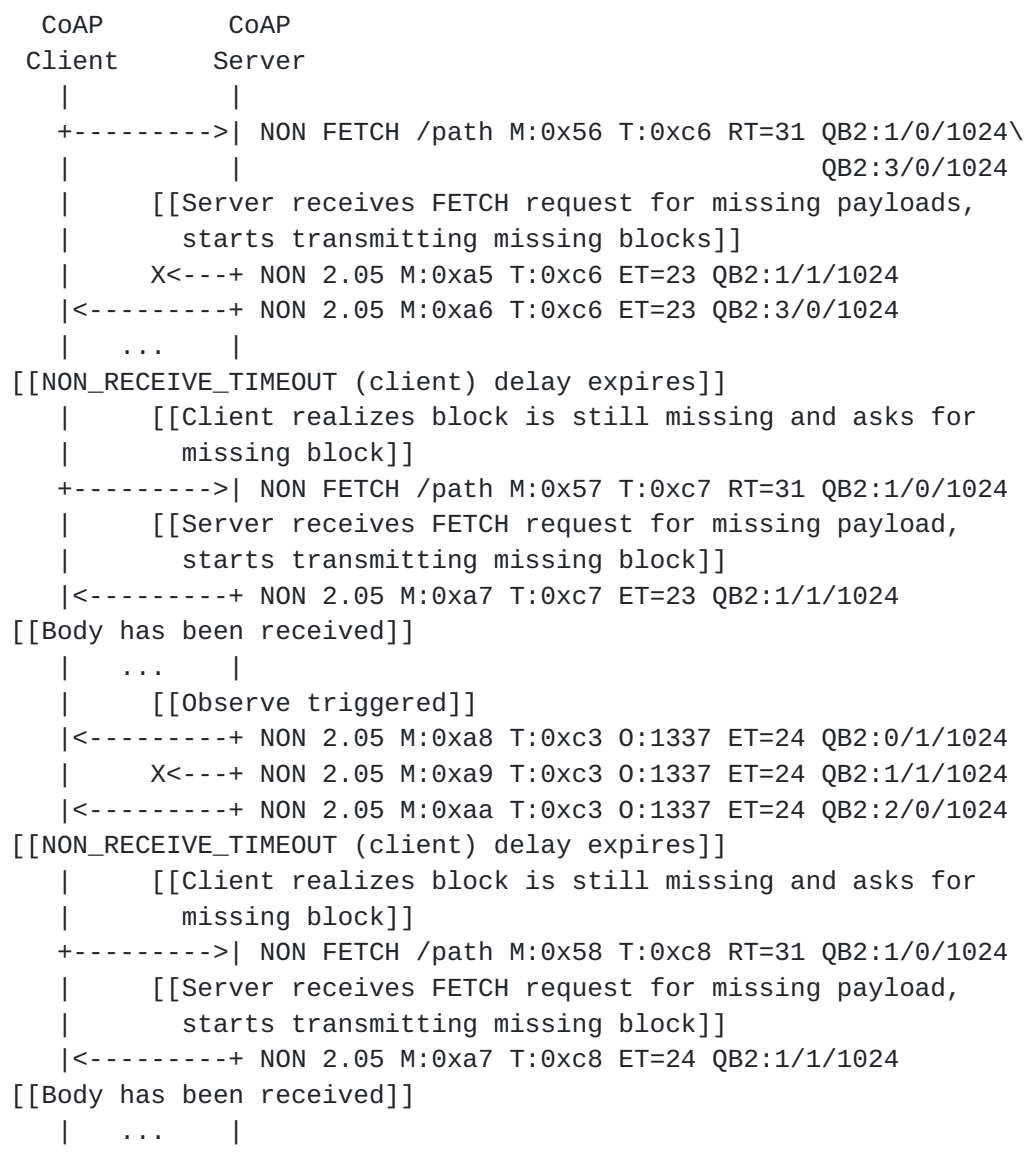

Figure 16: Example of NON Request with Q-Block1 Option (Client Recovery)

11. Security Considerations

Security considerations discussed in Section 7 of [RFC7959] should be taken into account.

Security considerations discussed in Sections 11.3 and 11.4 of [RFC7252] should be taken into account.

OSCORE provides end-to-end protection of all information that is not required for proxy operations and requires that a security context is set up (Section 3.1 of [RFC8613]). It can be trusted that the source endpoint is legitimate even if NoSec security mode is used. However,

an intermediary node can modify the unprotected outer Q-Block1 and/or Q-Block2 Options to cause a Q-Block transfer to fail or keep requesting all the blocks by setting the M bit and, thus, causing attack amplification. As discussed in Section 12.1 of [RFC8613], applications need to consider that certain message fields and messages types are not protected end-to-end and may be spoofed or manipulated. Therefore, it is NOT RECOMMENDED to use the NoSec security mode if either the Q-Block1 or Q-Block2 Options is used.

If OSCORE is not used, it is also NOT RECOMMENDED to use the NoSec security mode if either the Q-Block1 or Q-Block2 Options is used.

If NoSec is being used, Section D.5 of [RFC8613] discusses the security analysis and considerations for unprotected message fields even if OSCORE is not being used.

Security considerations related to the use of Request-Tag are discussed in Section 5 of [I-D.ietf-core-echo-request-tag].

12. IANA Considerations

RFC Editor Note: Please replace [RFCXXXX] with the RFC number to be assigned to this document.

12.1. CoAP Option Numbers Registry

IANA is requested to add the following entries to the "CoAP Option Numbers" sub-registry [Options] defined in [RFC7252] within the "Constrained RESTful Environments (CoRE) Parameters" registry:

Number	Name	Reference
TBA1	Q-Block1	[RFCXXXX]
TBA2	Q-Block2	[RFCXXXX]

Table 4: CoAP Q-Block1 and Q-Block2 Option Numbers

This document suggests 19 (TBA1) and 31 (TBA2) as values to be assigned for the new option numbers.

12.2. Media Type Registration

This document requests IANA to register the "application/missing-blocks+cbor-seq" media type in the "Media Types" registry [IANA-MediaTypes]. This registration follows the procedures specified in [RFC6838]:

Type name: application

Subtype name: missing-blocks+cbor-seq

Required parameters: N/A

Optional parameters: N/A

Encoding considerations: Must be encoded as a CBOR sequence [RFC8742], as defined in Section 4 of [RFCXXXX].

Security considerations: See Section 10 of [RFCXXXX].

Interoperability considerations: N/A

Published specification: [RFCXXXX]

Applications that use this media type: Data serialization and deserialization. In particular, the type is used by applications relying upon block-wise transfers, allowing a server to specify non-received blocks and request for their retransmission, as defined in Section 4 of [RFCXXXX].

Fragment identifier considerations: N/A

Additional information: N/A

Person & email address to contact for further information: IETF, iesg@ietf.org

Intended usage: COMMON

Restrictions on usage: none

Author: See Authors' Addresses section.

Change controller: IESG

Provisional registration? No

12.3. CoAP Content-Formats Registry

This document requests IANA to register the following CoAP Content-Format for the "application/missing-blocks+cbor-seq" media type in the "CoAP Content-Formats" registry [Format], defined in [RFC7252], within the "Constrained RESTful Environments (CoRE) Parameters" registry:

- Media Type: application/missing-blocks+cbor-seq
- Encoding: -
- Id: TBA3
- Reference: [RFCXXXX]

This document suggests 272 (TBA3) as a value to be assigned for the new content format number.

13. References

13.1. Normative References

[I-D.ietf-core-echo-request-tag]
Amsuess, C., Mattsson, J. P., and G. Selander, "CoAP: Echo, Request-Tag, and Token Processing", draft-ietf-core-echo-request-tag-12 (work in progress), February 2021.

[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>.

[RFC6838] Freed, N., Klensin, J., and T. Hansen, "Media Type Specifications and Registration Procedures", BCP 13, RFC 6838, DOI 10.17487/RFC6838, January 2013, <https://www.rfc-editor.org/info/rfc6838>.

[RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014, <https://www.rfc-editor.org/info/rfc7252>.

[RFC7641] Hartke, K., "Observing Resources in the Constrained Application Protocol (CoAP)", RFC 7641, DOI 10.17487/RFC7641, September 2015, <https://www.rfc-editor.org/info/rfc7641>.

[RFC7959] Bormann, C. and Z. Shelby, Ed., "Block-Wise Transfers in the Constrained Application Protocol (CoAP)", RFC 7959, DOI 10.17487/RFC7959, August 2016, <https://www.rfc-editor.org/info/rfc7959>.

[RFC8075] Castellani, A., Loreto, S., Rahman, A., Fossati, T., and E. Dijk, "Guidelines for Mapping Implementations: HTTP to the Constrained Application Protocol (CoAP)", RFC 8075, DOI 10.17487/RFC8075, February 2017, <https://www.rfc-editor.org/info/rfc8075>.

[RFC8132] van der Stok, P., Bormann, C., and A. Sehgal, "PATCH and FETCH Methods for the Constrained Application Protocol (CoAP)", RFC 8132, DOI 10.17487/RFC8132, April 2017, <https://www.rfc-editor.org/info/rfc8132>.

[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>.

[RFC8323] Bormann, C., Lemay, S., Tschofenig, H., Hartke, K., Silverajan, B., and B. Raymor, Ed., "CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets", RFC 8323, DOI 10.17487/RFC8323, February 2018, <https://www.rfc-editor.org/info/rfc8323>.

[RFC8610] Birkholz, H., Vigano, C., and C. Bormann, "Concise Data Definition Language (CDDL): A Notational Convention to Express Concise Binary Object Representation (CBOR) and JSON Data Structures", RFC 8610, DOI 10.17487/RFC8610, June 2019, <https://www.rfc-editor.org/info/rfc8610>.

[RFC8613] Selander, G., Mattsson, J., Palombini, F., and L. Seitz, "Object Security for Constrained RESTful Environments (OSCORE)", RFC 8613, DOI 10.17487/RFC8613, July 2019, <https://www.rfc-editor.org/info/rfc8613>.

[RFC8742] Bormann, C., "Concise Binary Object Representation (CBOR) Sequences", RFC 8742, DOI 10.17487/RFC8742, February 2020, <https://www.rfc-editor.org/info/rfc8742>.

[RFC8949] Bormann, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", STD 94, RFC 8949, DOI 10.17487/RFC8949, December 2020, <https://www.rfc-editor.org/info/rfc8949>.

13.2. Informative References

[Format] "CoAP Content-Formats", <https://www.iana.org/assignments/core-parameters/core-parameters.xhtml#content-formats>.

[I-D.bosh-dots-quick-blocks] Boucadair, M. and J. Shallow, "Distributed Denial-of-Service Open Threat Signaling (DOTS) Signal Channel Configuration Attributes for Faster Block Transmission", draft-bosh-dots-quick-blocks-01 (work in progress), January 2021.

[I-D.ietf-dots-telemetry]
Boucadair, M., Reddy, T., Doron, E., Chen, M., and J. Shallow, "Distributed Denial-of-Service Open Threat Signaling (DOTS) Telemetry", draft-ietf-dots-telemetry-15 (work in progress), December 2020.

[IANA-MediaTypes]
IANA, "Media Types", <https://www.iana.org/assignments/media-types>.

[Options] "CoAP Option Numbers", <https://www.iana.org/assignments/core-parameters/core-parameters.xhtml#option-numbers>.

[RFC6928] Chu, J., Dukkipati, N., Cheng, Y., and M. Mathis, "Increasing TCP's Initial Window", RFC 6928, DOI 10.17487/RFC6928, April 2013, <https://www.rfc-editor.org/info/rfc6928>.

[RFC7967] Bhattacharyya, A., Bandyopadhyay, S., Pal, A., and T. Bose, "Constrained Application Protocol (CoAP) Option for No Server Response", RFC 7967, DOI 10.17487/RFC7967, August 2016, <https://www.rfc-editor.org/info/rfc7967>.

[RFC8782] Reddy.K, T., Ed., Boucadair, M., Ed., Patil, P., Mortensen, A., and N. Teague, "Distributed Denial-of-Service Open Threat Signaling (DOTS) Signal Channel Specification", RFC 8782, DOI 10.17487/RFC8782, May 2020, <https://www.rfc-editor.org/info/rfc8782>.

[RFC8974] Hartke, K. and M. Richardson, "Extended Tokens and Stateless Clients in the Constrained Application Protocol (CoAP)", RFC 8974, DOI 10.17487/RFC8974, January 2021, <https://www.rfc-editor.org/info/rfc8974>.

Appendix A. Examples with Confirmable Messages

The following examples assume NSTART has been increased to 3.

The notations provided in Figure 2 are used in the following subsections.

A.1. Q-Block1 Option

Let's now consider the use of Q-Block1 Option with a CON request as shown in Figure 17. All the blocks are acknowledged (ACK).

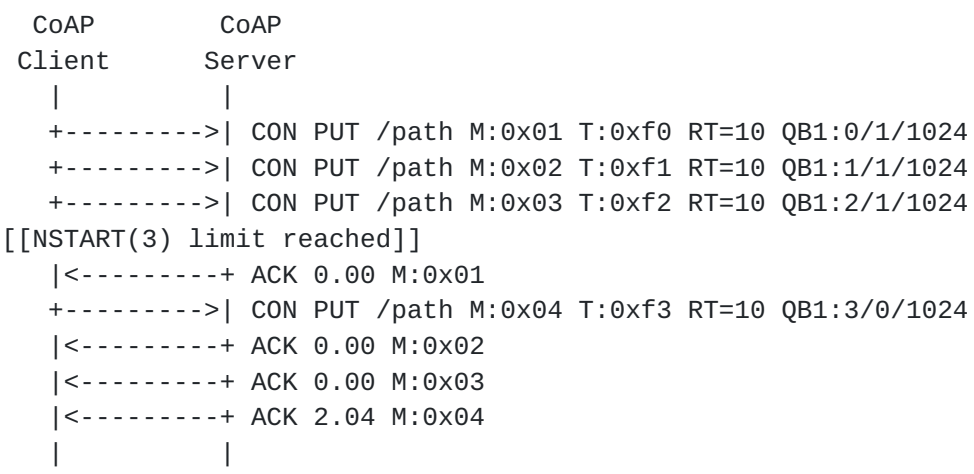

Figure 17: Example of CON Request with Q-Block1 Option (Without Loss)

Now, suppose that a new body of data is to be sent but with some blocks dropped in transmission as illustrated in Figure 18. The client will retry sending blocks for which no ACK was received.

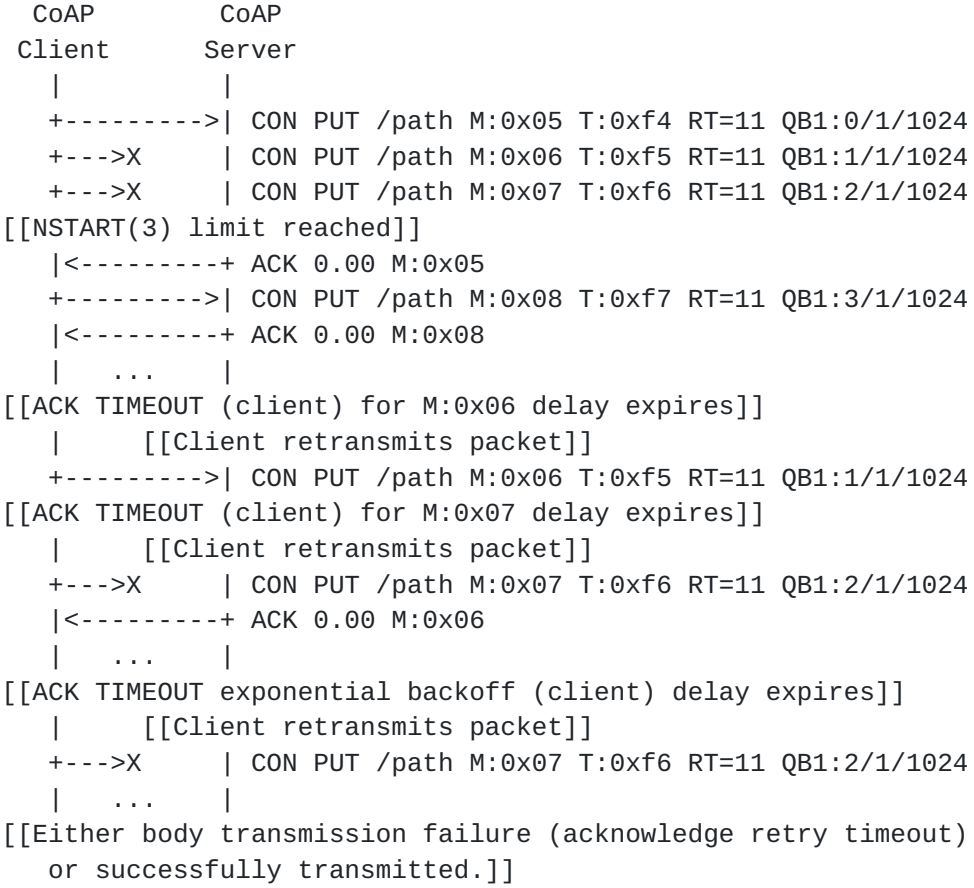

Figure 18: Example of CON Request with Q-Block1 Option (Blocks Recovery)

It is up to the implementation as to whether the application process stops trying to send this particular body of data on reaching MAX_RETRANSMIT for any payload, or separately tries to initiate the new transmission of the payloads that have not been acknowledged under these adverse traffic conditions.

If there is likely to be the possibility of transient network losses, then the use of NON should be considered.

A.2. Q-Block2 Option

An example of the use of Q-Block2 Option with Confirmable messages is shown in Figure 19.

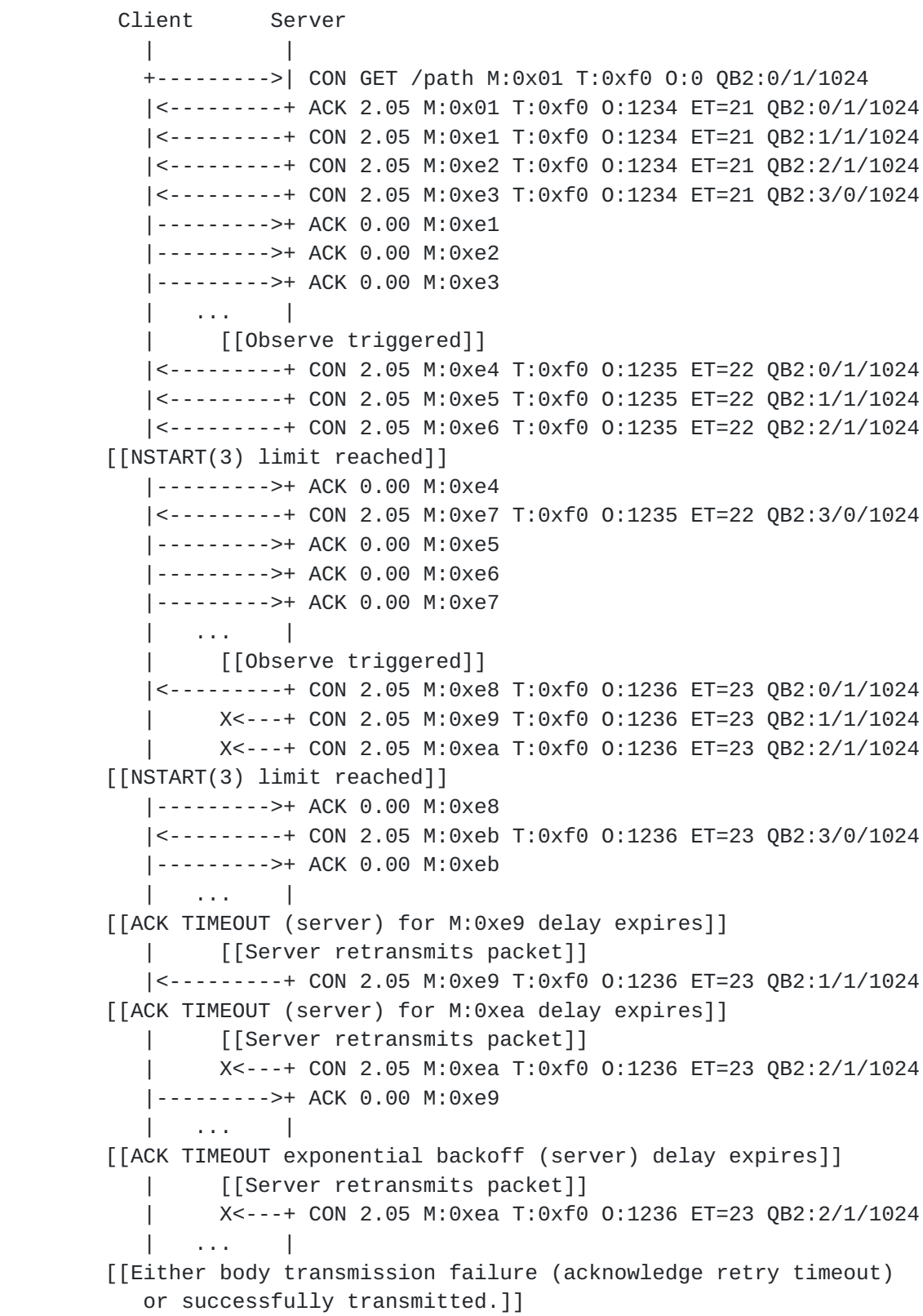

Figure 19: Example of CON Notifications with Q-Block2 Option

It is up to the implementation as to whether the application process stops trying to send this particular body of data on reaching MAX_RETRANSMIT for any payload, or separately tries to initiate the new transmission of the payloads that have not been acknowledged under these adverse traffic conditions.

If there is likely to be the possibility of transient network losses, then the use of NON should be considered.

Appendix B. Examples with Reliable Transports

The notations provided in Figure 2 are used in the following subsections.

B.1. Q-Block1 Option

Let's now consider the use of Q-Block1 Option with a reliable transport as shown in Figure 20. There is no acknowledgment of packets at the CoAP layer, just the final result.

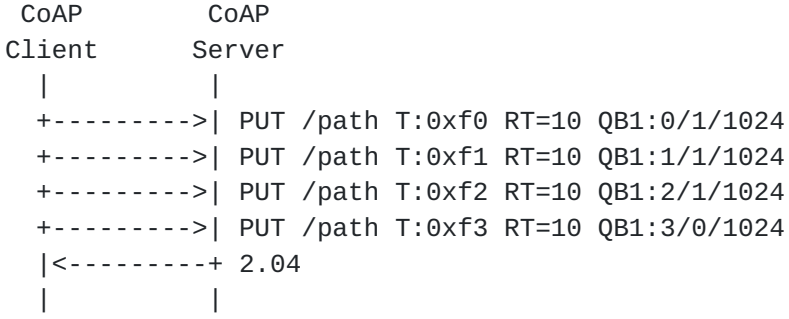

Figure 20: Example of Reliable Request with Q-Block1 Option

If there is likely to be the possibility of transient network losses, then the use of unreliable transport with NON should be considered.

B.2. Q-Block2 Option

An example of the use of Q-Block2 Option with a reliable transport is shown in Figure 21.

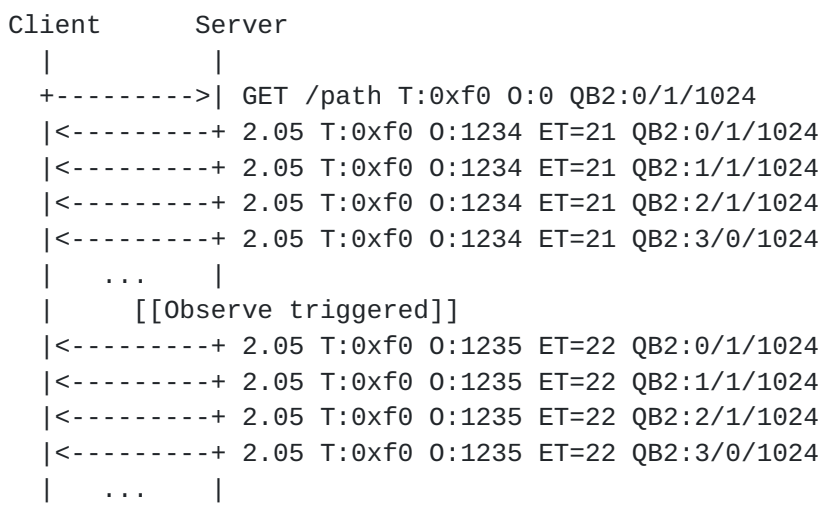

Figure 21: Example of Notifications with Q-Block2 Option

If there is likely to be the possibility of network transient losses, then the use of unreliable transport with NON should be considered.

Acknowledgements

Thanks to Achim Kraus, Jim Schaad, and Michael Richardson for their comments.

Special thanks to Christian Amsuess, Carsten Bormann, and Marco Tiloca for their suggestions and several reviews, which improved this specification significantly. Thanks to Francesca Palombini for the AD review.

Thanks to Pete Resnick for the Gen-ART review, Colin Perkins for the Tsvart review, and Emmanuel Baccelli for the Iotdir review. Thanks to Martin Duke, Eric Vyncke, Benjamin Kaduk, Roman Danyliw, John Scudder, and Lars Eggert for the IESG review.

Some text from [RFC7959] is reused for readers convenience.

Authors' Addresses

Mohamed Boucadair
Orange
Rennes 35000
France

Email: mohamed.boucadair@orange.com

Jon Shallow
United Kingdom

Email: supjps-ietf@jpshallow.com