6lo
Internet-Draft
Updates: 6775 (if approved)
Intended status: Standards Track
Expires: August 27, 2018

P. Thubert, Ed.
Cisco
B. Sarikaya
M. Sethi
Ericsson

February 23, 2018

Address Protected Neighbor Discovery for Low-power and Lossy Networks
draft-ietf-6lo-ap-nd-06

Abstract

This document defines an extension to 6LoWPAN Neighbor Discovery (ND) [RFC6775][I-D.ietf-6lo-rfc6775-update] called Address Protected ND (AP-ND); AP-ND protects the owner of an address against address theft and impersonation inside a low-power and lossy network (LLN). Nodes supporting this extension compute a cryptographic Owner Unique Interface ID and associate it with one or more of their Registered Addresses. The Cryptographic ID uniquely identifies the owner of the Registered Address and can be used for proof-of-ownership. It is used in 6LoWPAN ND in place of the EUI-64-based unique ID that is associated with the registration. Once an address is registered with a Cryptographic ID, only the owner of that ID can modify the anchor state information of the Registered Address, and Source Address Validation can be enforced.

Table of Contents

1. Introduction

"Neighbor Discovery Optimizations for 6LoWPAN networks" [RFC6775] (6LoWPAN ND) adapts the classical IPv6 ND protocol [RFC4861][RFC4862] (IPv6 ND) for operations over a constrained low-power and lossy network (LLN). In particular, 6LoWPAN ND introduces a unicast host address registration mechanism that contributes to reduce the use of multicast messages that are present in the classical IPv6 ND protocol. 6LoWPAN ND defines a new Address Registration Option (ARO) that is carried in the unicast Neighbor Solicitation (NS) and Neighbor Advertisement (NA) messages between the 6LoWPAN Node (6LN) and the 6LoWPAN Router (6LR). Additionally, it also defines the Duplicate Address Request (DAR) and Duplicate Address Confirmation (DAC) messages between the 6LR and the 6LoWPAN Border Router (6LBR). In LLN networks, the 6LBR is the central repository of all the registered addresses in its domain.

The registration mechanism in 6LoWPAN ND [RFC6775] prevents the use of an address if that address is already present in the subnet (first come first serve). In order to validate address ownership, the registration mechanism enables the 6LR and 6LBR to validate claims for a registered address with an associated Owner Unique Interface IDentifier (OUID). 6LoWPAN ND specifies that the OUID is derived from the MAC address of the device (using the 64-bit Extended Unique Identifier EUI-64 address format specified by IEEE), which can be spoofed. Therefore, any node connected to the subnet and aware of a registered-address-to-OUID mapping could effectively fake the OUID, steal the address and redirect traffic for that address towards a different 6LN. The "Update to 6LoWPAN ND" [I-D.ietf-6lo-rfc6775-update] defines an Extended ARO (EARO) option that allows to transport alternate forms of OUIDs, and is a prerequisite for this specification.

According to this specification, a 6LN generates a cryptographic ID (Crypto-ID) and places it in the OUID field in the registration of one (or more) of its addresses with the 6LR(s) that the 6LN uses as default router(s). Proof of ownership of the cryptographic ID (Crypto-ID) is passed with the first registration exchange to a new 6LR, and enforced at the 6LR. The 6LR validates ownership of the cryptographic ID before it can create a registration state, or a change the anchor information, that is the Link-Layer Address and associated parameters, in an existing registration state.

The protected address registration protocol proposed in this document enables the enforcement of Source Address Validation (SAVI) [RFC7039], which ensures that only the correct owner uses a registered address in the source address field in IPv6 packets. Consequently, a 6LN that sources a packet has to use a 6LR to which

the source address of the packet is registered to forward the packet. The 6LR maintains state information for the registered addressed, including the MAC address, and a link-layer cryptographic key associated with the 6LN. In SAVI-enforcement mode, the 6LR allows only packets from a connected Host if the connected Host owns the registration of the source address of the packet.

The 6lo adaptation layer framework ([RFC4944], [RFC6282]) expects that a device forms its IPv6 addresses based on Layer-2 address, so as to enable a better compression. This is incompatible with "Secure Neighbor Discovery (SeND)" [RFC3971] and "Cryptographically Generated Addresses (CGAs)" [RFC3972], which derive the Interface ID (IID) in the IPv6 addresses from cryptographic material. "Privacy Considerations for IPv6 Address Generation Mechanisms" [RFC7721] places additional recommendations on the way addresses should be formed and renewed.

This document specifies that a device may form and register addresses at will, without a constraint on the way the address is formed or the number of addresses that are registered in parallel. It enables to protect multiple addresses with a single cryptographic material and to send the proof only once to a given 6LR for multiple addresses and refresher registrations.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Readers are expected to be familiar with all the terms and concepts that are discussed in [RFC3971], [RFC3972], [RFC4861], [RFC4919], [RFC6775], and [I-D.ietf-6lo-backbone-router] which proposes an evolution of [RFC6775] for wider applicability.

This document defines Crypto-ID as an identifier of variable size which in most cases is 64 bits long. It is generated using cryptographic means explained later in this document Section 4.2. "Elliptic Curves for Security" [RFC7748] and "Edwards-Curve Digital Signature Algorithm (EdDSA)" [RFC8032] provides information on Elliptic Curve Cryptography (ECC) and a (twisted) Edwards curve, Ed25519, which can be used with this specification. "Alternative Elliptic Curve Representations" [I-D.struik-lwig-curve-representations] provides additional information on how to represent Montgomery curves and (twisted) Edwards curves as curves in short-Weierstrass form and illustrates how this can be used to implement elliptic curve computations using

existing implementations that already implement, e.g., ECDSA and ECDH using NIST [FIPS-186-4] prime curves.

The document also conforms to the terms and models described in [RFC5889] and uses the vocabulary and the concepts defined in [RFC4291] for the IPv6 Architecture. Finally, common terminology related to Low power And Lossy Networks (LLN) defined in [RFC7102] is also used.

3. Updating RFC 6775

This specification defines a cryptographic identifier (Crypto-ID) that can be used as a replacement to the MAC address in the OUID field of the EARO option; the computation of the Crypto-ID is detailed in Section 4.2. A node in possession of the necessary cryptographic material SHOULD use Crypto-ID by default as OUID in its registration. Whether a OUID is a Crypto-ID is indicated by a new "C" flag in the NS(EARO) message.

In order to prove its ownership of a Crypto-ID, the registering node needs to produce the parameters that where used to build it, as well as a nonce and a signature that will prove that it has the private key that corresponds to the public key that was used to build the Crypto-ID. This specification adds the capability to carry new options in the NS(EARO) and the NBA(EARO). These options are a variation of the CGA Option Section 4.4, a Nonce option and a variation of the RSA Signature option Section 4.6 in the NS(EARO) and a Nonce option in the NA(EARO).

4. New Fields and Options

In order to avoid an inflation of ND option types, this specification reuses / extends options defined in SEND [RFC3971] and 6LoWPAN ND [RFC6775][I-D.ietf-6lo-rfc6775-update]. This applies in particular to the CGA option and the RSA Signature Option. This specification provides aliases for the specific variations of those options as used in AP-ND. The presence of the EARO option in the NS/NA messages indicates that the options are to be understood as specified in this document. A router that would receive a NS(EARO) and try to process it as a SEND message will find that the signature does not match and drop the packet.

4.1. Encoding the Public Key

Public Key is the most important parameter in CGA Parameters (sent by 6LN in an NS message). ECC Public Key could be in uncompressed form or in compressed form where the first octet of the OCTET STRING is

0x04 and 0x02 or 0x03, respectively. Point compression can further reduce the key size by about 32 octets.

4.2. New Crypto-ID

Elliptic Curve Cryptography (ECC) is used to calculate the Crypto-ID. Each 6LN using a Crypto-ID for registration MUST have a public/ private key pair. The digital signature is constructed by using the 6LN's private key over its EUI-64 (MAC) address. The signature value is computed using the ECDSA signature algorithm and the hash function used is SHA-256 [RFC6234].

NIST P-256 [FIPS186-4] that MUST be supported by all implementations. To support cryptographic algorithm agility [RFC7696], Edwards-Curve Digital Signature Algorithm (EdDSA) curve Ed25519ph (pre-hashing) [RFC8032] MAY be supported as an alternate.

The Crypto-ID is computed as follows:

1. An 8-bits modifier is selected, for instance, but not necessarily, randomly; the modifier enables a device to form multiple Crypto-IDs with a single key pair. This may be useful for privacy reasons in order to avoid the correlation of addresses based on their Crypto-ID;

2. the modifier value and the DER-encoded public key (Section 4.1) are concatenated from left to right;

3. Digital signature (SHA-256 then either NIST P-256 or EdDSA) is executed on the concatenation

4. the leftmost bits of the resulting signature are used as the Crypto-ID;

With this specification, only 64 bits are retained, but it could be expanded to more bits in the future by increasing the size of the OUID field.

4.3. Updated EARO

This specification updates the EARO option as follows:

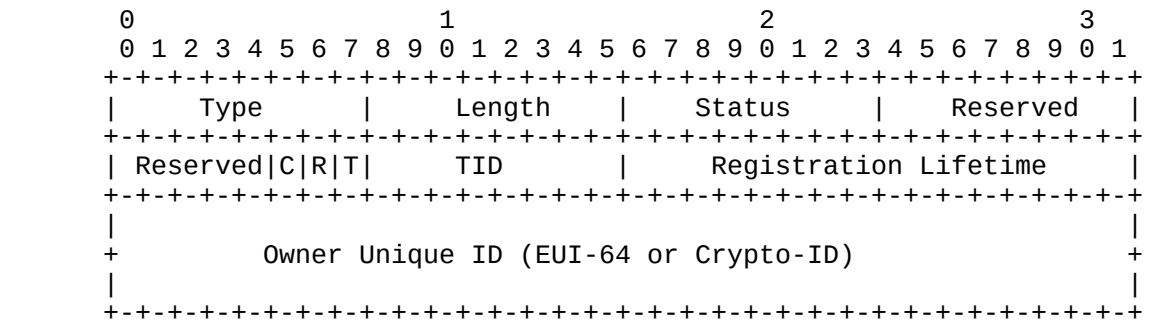


Figure 1: Enhanced Address Registration Option

Type: 33

Length: 8-bit unsigned integer. The length of the option (including the type and length fields) in units of 8 bytes.

Status: 8-bit unsigned integer. Indicates the status of a registration in the NA response. MUST be set to 0 in NS messages. This specification uses values introduced in the update to 6LoWPAN ND [I-D.ietf-6lo-rfc6775-update], such as "Validation Requested" and "Validation Failed". No additional value is defined.

Reserved: This field is unused. It MUST be initialized to zero by the sender and MUST be ignored by the receiver.

C: This "C" flag is set to indicate that the Owner Unique ID field contains a Crypto-ID and that the 6LN MAY be challenged for ownership as specified in this document.

R: Defined in [I-D.ietf-6lo-rfc6775-update].

T and TID: Defined in [I-D.ietf-6lo-rfc6775-update].

Owner Unique ID: When the "C" flag is set, this field contains a Crypto-ID.

4.4. Crypto-ID Parameters Option

This specification defines the Crypto-ID Parameters Option (CIPO), as a variation of the CGA Option that carries the parameters used to form a Crypto-ID. In order to provide cryptographic agility, AP-ND supports two possible signature algorithms, indicated by a Crypto-Type field. A value of 0 indicates that NIST P-256 is used for the signature operation and SHA-256 as the hash algorithm. NIST P-256 MUST be supported by all implement A value of 1 indicates that Ed25519ph is used for the signature operation and SHA-256 as the hash algorithm.

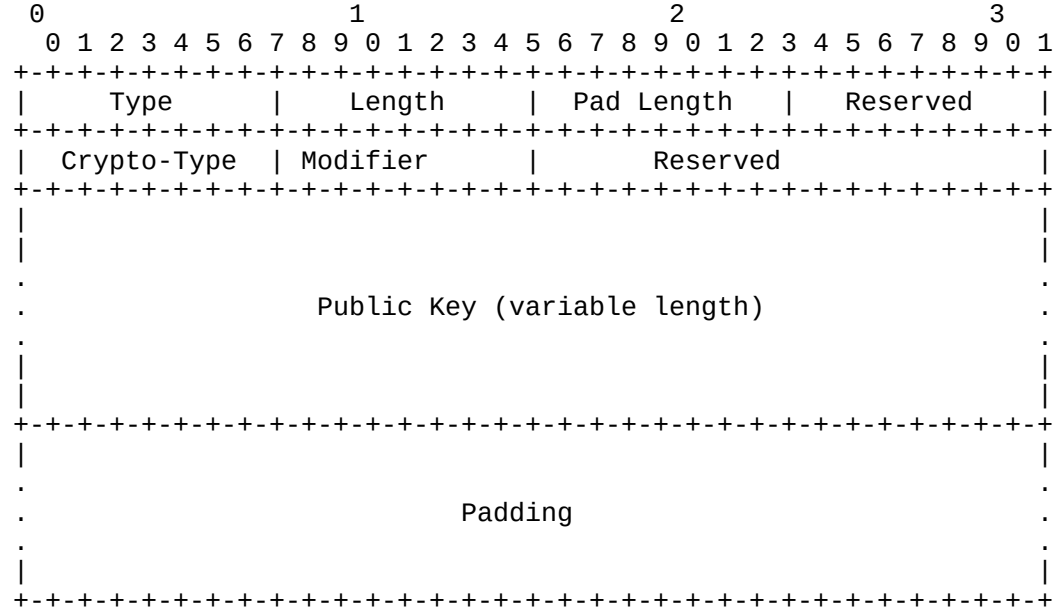


Figure 2: Crypto-ID Parameters Option

Type: 11. This is the same value as the CGA Option, CIPO is a particular case of the CGA option

Length: 8-bit unsigned integer. The length of the option in units of 8 octets.

Modifier: 8-bit unsigned integer.

Pad Length: 8-bit unsigned integer. The length of the Padding field.

Crypto-Type: The type of cryptographic algorithm used in calculation Crypto-ID. Default value of all zeros indicate NIST P-256. A value of 1 is assigned for Ed25519ph. New values may be defined later.

Public Key: Public Key of 6LN.

Padding: A variable-length field making the option length a multiple of 8, containing as many octets as specified in the Pad Length field.

4.5. Nonce Option

This document reuses the Nonce Option defined in section 5.3.2. of SEND [RFC3971] without a change.

4.6. NDP Signature Option

This document reuses the RSA Signature Option (RSAO) defined in section 5.2. of SEND [RFC3971]. Admittedly, the name is ill-chosen since the option is extended for non-RSA Signatures and this specification defines an alias to avoid the confusion.

The description of the operation on the option detailed in section 5.2. of SEND [RFC3971] apply, but for the following changes:

- The 128-bit CGA Message Type tag [RFC3972] for AP-ND is 0x8701 55c8 0cca dd32 6ab7 e415 f148 84d0. (The tag value has been generated by the editor of this specification on random.org).

- The signature is computed using the hash algorithm and the digital signature indicated in the Crypto-Type field of the CIPO option using the private key associated with the public key in the CIPO.

- The alias NDP Signature Option (NDPSO) can be used to refer to the RSAO when used as described in this specification.

5. Protocol Scope

The scope of the present work is a 6LoWPAN Low Power Lossy Network (LLN), typically a stub network connected to a larger IP network via a Border Router called a 6LBR per [RFC6775].

The 6LBR maintains a registration state for all devices in the attached LLN, and, in conjunction with the first-hop router (the 6LR), is in a position to validate uniqueness and grant ownership of an IPv6 address before it can be used in the LLN. This is a fundamental difference with a classical network that relies on IPv6

address auto-configuration [RFC4862], where there is no guarantee of ownership from the network, and any IPv6 Neighbor Discovery packet must be individually secured [RFC3971].

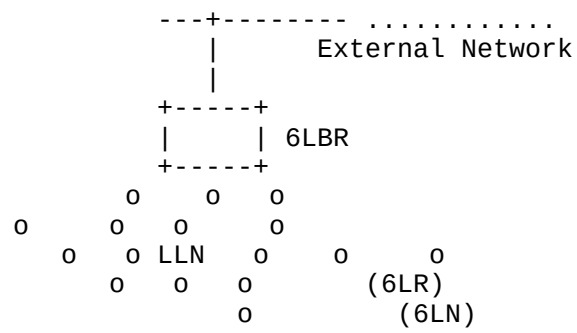


Figure 3: Basic Configuration

In a mesh network, the 6LR is directly connected to the host device. This specification expects that the peer-wise layer-2 security is deployed so that all the packets from a particular host are securely identifiable by the 6LR. The 6LR may be multiple hops away from the 6LBR. Packets are routed between the 6LR and the 6LBR via other 6LRs. This specification expects that a chain of trust is established so that a packet that was validated by the first 6LR can be safely routed by the next 6LRs to the 6LBR.

6. Protocol Flows

The 6LR/6LBR ensures first-come/first-serve by storing the EARO information including the Crypto-ID correlated to the node being registered. The node is free to claim any address it likes as long as it is the first to make such a claim. After a successful registration, the node becomes the owner of the registered address and the address is bound to the Crypto-ID in the 6LR/6LBR registry.

This specification enables to verify the ownership of the binding at any time assuming that the "C" flag is set. If it is not set, then the verification methods presented in this specification cannot be applied. The verification prevents other nodes from stealing the address and trying to attract traffic for that address or use it as their source address.

A node may use multiple IPv6 addresses at the same time. The node may use a same Crypto-ID, or multiple crypto-IDs derived from a same key pair, to protect multiple IPv6 addresses. The separation of the address and the cryptographic material avoids the constrained device

to compute multiple keys for multiple addresses. The registration process allows the node to bind all of its addresses to the same Crypto-ID.

6.1. First Exchange with a 6LR

A 6LN registers to a 6LR that is one hop away from it with the "C" flag set in the EARO, indicating that the Owner Unique ID field contains a Crypto-ID. The on-link (local) protocol interactions are shown in Figure 4 If the 6LR does not have a state with the 6LN that is consistent with the NS(EARO), then it replies with a challenge NA (EARO, status=Validation Requested) that contains a Nonce Option. The Nonce option MUST contain a Nonce value that was never used with this device.

The 6LN replies to the challenge with a proof-of-ownership NS(EARO) that includes the echoed Nonce option, the CIPO with all the parameters that where used to build EARO with a Crypto-ID, and as the last option the NDPSO with the signature. The information associated to a crypto-ID is passed to and stored by the 6LR on the first NS exchange where it appears. The 6LR SHOULD store the CIPO information associated with the crypto-ID so it can be used for more than one address.

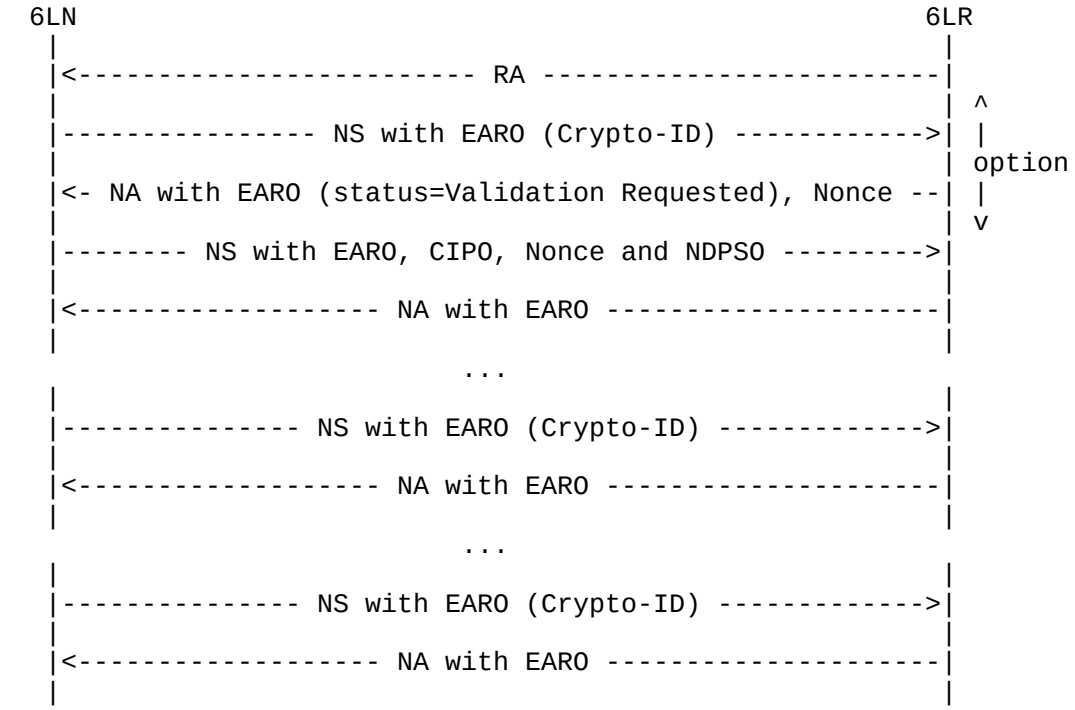


Figure 4: On-link Protocol Operation

The steps for the registration to the 6LR are as follows:

- Upon the first exchange with a 6LR, a 6LN may be challenged and have to produce the proof of ownership of the Crypto-ID. However, it is not expected that the proof is needed again in the periodic refresher registrations for that address, or when registering other addresses with the same OUID. When a 6LR receives a NS(EARO) registration with a new Crypto-ID as a OUID, it SHOULD challenge by responding with a NA(EARO) with a status of "Validation Requested". This process of validation MAY be skipped in networks where there is no mobility.

- The challenge MUST also be triggered in the case of a registration for which the Source Link-Layer Address is not consistent with a state that already exists either at the 6LR or the 6LBR. In the latter case, the 6LBR returns a status of "Validation Requested" in the DAR/DAC exchange, which is echoed by the 6LR in the NA (EARO) back to the registering node. This flow should not alter a preexisting state in the 6LR or the 6LBR.

- Upon receiving a NA(EARO) with a status of "Validation Requested", the registering node SHOULD retry its registration with a Crypto-ID Parameters Option (CIPO) Section 4.4 that contains all the necessary material for building the Crypto-ID, the Nonce and the NDP signature Section 4.6 options that prove its ownership of the Crypto-ID.

- In order to validate the ownership, the 6LR performs the same steps as the 6LN and rebuilds the Crypto-ID based on the parameters in the CIPO. If the result is different then the validation fails. Else, the 6LR checks the signature in the NDPSO using the public key in the CIPO. If it is correct then the validation passes, else it fails.

- If the 6LR fails to validate the signed NS(EARO), it responds with a status of "Validation Failed". After receiving a NA(EARO) with a status of "Validation Failed", the registering node SHOULD try an alternate Signature Algorithm and Crypto-ID. In any case, it MUST NOT use this Crypto-ID for registering with that 6LR again.

6.2. Multihop Operation

In a multihop 6LoWPAN, the registration with Crypto-ID is propagated to 6LBR as described in Section 6.2. If a chain of trust is present between the 6LR and the 6LBR, then there is no need to propagate the proof of ownership to the 6LBR. All the 6LBR needs to know is that this particular OUID is randomly generated, so as to enforce that any update via a different 6LR is also random.

A new device that joins the network auto-configures an address and performs an initial registration to an on-link 6LR with an NS message that carries an Address Registration Option (EARO) [RFC6775]. The 6LR validates the address with the central 6LBR using a DAR/DAC exchange, and the 6LR confirms (or denies) the address ownership with an NA message that also carries an Address Registration Option.

Figure 5 illustrates a registration flow all the way to a 6LowPAN Backbone Router (6BBR).

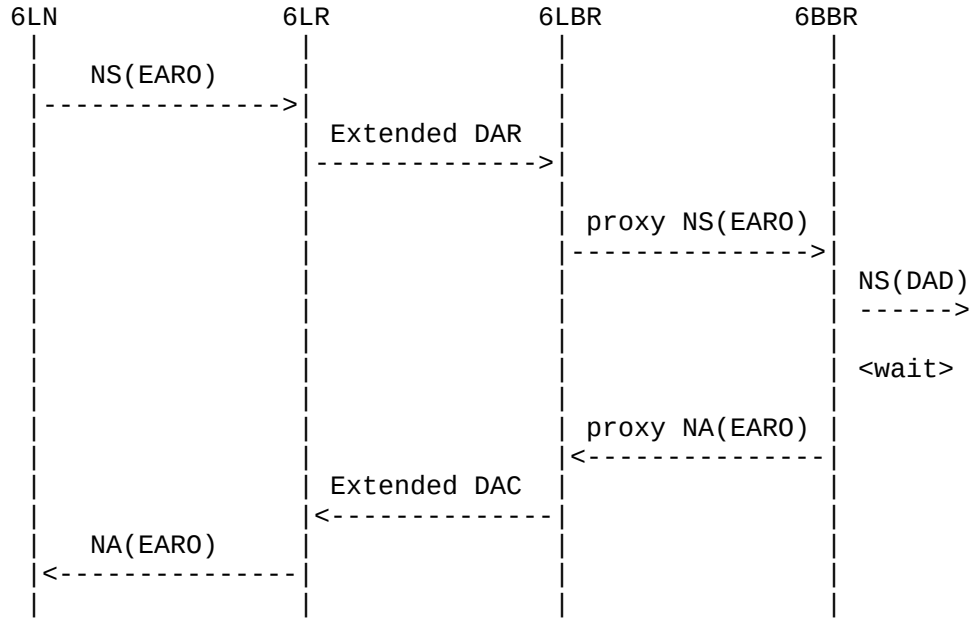


Figure 5: (Re-)Registration Flow

In a multihop 6LoWPAN, a 6LBR sends RAs with prefixes downstream and the 6LR receives and relays them to the nodes. 6LR and 6LBR communicate using ICMPv6 Duplicate Address Request (DAR) and Duplicate Address Confirmation (DAC) messages. The DAR and DAC use the same message format as NS and NA, but have different ICMPv6 type values.

In AP-ND we extend DAR/DAC messages to carry cryptographically generated OUID. In a multihop 6LoWPAN, the node exchanges the messages shown in Figure 5. The 6LBR must identify who owns an address (EUI-64) to defend it, if there is an attacker on another 6LR.

Occasionally, a 6LR might miss the node’s OUID (that it received in ARO). 6LR should be able to ask for it again. This is done by restarting the exchanges shown in Figure 4. The result enables 6LR to refresh the information that was lost. The 6LR MUST send DAR message with ARO to 6LBR. The 6LBR replies with a DAC message with the information copied from the DAR, and the Status field is set to zero. With this exchange, the 6LBR can (re)validate and store the information to make sure that the 6LR is not a fake.

In some cases, the 6LBR may use a DAC message to solicit a Crypto-ID from a 6LR and also requests 6LR to verify the EUI-64 6LR received from 6LN. This may happen when a 6LN node is compromised and a fake node is sending the Crypto-ID as if it is the node's EUI-64. Note that the detection in this case can only be done by 6LBR not by 6LR.

7. Security Considerations

7.1. Inheriting from RTC 3971

The observations regarding the threats to the local network in [RFC3971] also apply to this specification. Considering RFC3971 security section subsection by subsection:

Neighbor Solicitation/Advertisement Spoofing Threats in section 9.2.1 of RFC3971 apply. AP-ND counters the threats on NS(EARO) messages by requiring that the NDP Signature and CIPO options be present in these solicitations.

Neighbor Unreachability Detection Failure With RFC6775, a NUD can still be used by the endpoint to assess the liveliness of a device. The NUD request may be protected by SEND in which case the provision in section 92.2. of RFC 3972 applies. The response to the NUD may be proxied by a backbone router only if it has a fresh registration state for it. The registration being protected by this specification, the proxied NUD response provides a truthful information on the original owner of the address but it cannot be proven using SEND. If the NUD response is not proxied, the 6LR will pass the lookup to the end device which will respond with a traditional NA. If the 6LR does not have a cache entry associated for the device, it can issue a NA with EARO (status=Validation Requested) upon the NA from the device, which will trigger a NS that will recreate and revalidate the ND cache entry.

Duplicate Address Detection DoS Attack Inside the LLN, Duplicate Addresses are sorted out using the OUID, which differentiates it from a movement. DAD coming from the backbone are not forwarded over the LLN so the LLN is protected by the backbone routers. Over the backbone, the EARO option is present in NS/NA messages. This protects against misinterpreting a movement for a duplication, and enables to decide which backbone router has the freshest registration and thus most possibly the device attached to it. But this specification does not guarantee that the backbone router claiming an address over the backbone is not an attacker.

Router Solicitation and Advertisement Attacks This specification does not change the protection of RS and RA which can still be protected by SEND.

Replay Attacks A Nonce given by the 6LR in the NA with EARO (status=Validation Requested) and echoed in the signed NS guarantees against replay attacks of the NS(EARO). The NA(EARO) is not protected and can be forged by a rogue node that is not the 6LR in order to force the 6LN to rebuild a NS(EARO) with the proof of ownership, but that rogue node must have access to the L2 radio network next to the 6LN to perform the attack.

Neighbor Discovery DoS Attack A rogue node that managed to access the L2 network may form many addresses and register them using AP-ND. The perimeter of the attack os all the 6LRs in range of the attacker. The 6LR must protect itself against overflows and reject excessive registration with a status 2 "Neighbor Cache Full". This effectively blocks another (honest) 6LN from registering to the same 6LR, but the 6LN may register to other 6LRs that are in its range but not in that of the rogue.

7.2. Related to 6LoWPAN ND

The threats discussed in 6LoWPAN ND [RFC6775] and its update [I-D.ietf-6lo-rfc6775-update] also apply here. Compared with SeND, this specification saves about 1Kbyte in every NS/NA message. Also, this specification separates the cryptographic identifier from the registered IPv6 address so that a node can have more than one IPv6 address protected by the same cryptographic identifier. SeND forces the IPv6 address to be cryptographic since it integrates the CGA as the IID in the IPv6 address. This specification frees the device to form its addresses in any fashion, so as to enable the classical 6LoWPAN compression which derives IPv6 addresses from Layer-2 addresses, as well as privacy addresses. The threats discussed in Section 9.2 of [RFC3971] are countered by the protocol described in this document as well.

7.3. OUID Collisions

Collisions of Owner Unique Interface IDentifier (OUID) (which is the Crypto-ID in this specification) is a possibility that needs to be considered. The formula for calculating the probability of a collision is 1 - e^{-k^2/(2n)} where n is the maximum population size (2^64 here, 1.84E19) and K is the actual population (number of nodes). If the Crypto-ID is 64-bit long, then the chance of finding a collision is 0.01% when the network contains 66 million nodes. It is important to note that the collision is only relevant when this happens within one stub network (6LBR). A collision of Crypto-ID is

a rare event. In the case of a collision, an attacker may be able to claim the registered address of an another legitimate node. However for this to happen, the attacker would also need to know the address which was registered by the legitimate node. This registered address is however never broadcasted on the network and therefore it provides an additional entropy of 64-bits that an attacker must correctly guess. To prevent such a scenario, it is RECOMMENDED that nodes derive the address being registered independently of the OUID.

8. IANA considerations

8.1. CGA Message Type

This document defines a new 128-bit value under the CGA Message Type [RFC3972] namespace, 0x8701 55c8 0cca dd32 6ab7 e415 f148 84d0.

8.2. Crypto-Type Subregistry

IANA is requested to create a new subregistry "Crypto-Type Subregistry" in the "Internet Control Message Protocol version 6 (ICMPv6) Parameters". The registry is indexed by an integer 0..255 and contains a Signature Algorithm and a Hash Function as shown in Table 1. The following Crypto-Type values are defined in this document:

Crypto-Type value	Signature Algorithm	Hash Function	Defining Specification
0	NIST P-256 [FIPS186-4]	SHA-256 [RFC6234]	RFC THIS
1	Ed25519ph [RFC8032]	SHA-256 [RFC6234]	RFC THIS

Table 1: Crypto-Types

Assignment of new values for new Crypto-Type MUST be done through IANA with "Specification Required" and "IESG Approval" as defined in [RFC8126].

9. Acknowledgments

Many thanks to Charlie Perkins for his in-depth review and constructive suggestions. We are also especially grateful to Rene Struik and Robert Moskowitz for their comments that lead to many improvements to this document, in particular WRT ECC computation and references.

10. References

10.1. Normative References

[FIPS-186-4]
FIPS 186-4, "Digital Signature Standard (DSS), Federal Information Processing Standards Publication 186-4", US Department of Commerce/National Institute of Standards and Technology Gaithersburg, MD, July 2013.

[I-D.ietf-6lo-rfc6775-update]
Thubert, P., Nordmark, E., Chakrabarti, S., and C. Perkins, "An Update to 6LoWPAN ND", draft-ietf-6lo-rfc6775-update-13 (work in progress), February 2018.

[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>.

[RFC3279] Bassham, L., Polk, W., and R. Housley, "Algorithms and Identifiers for the Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 3279, DOI 10.17487/RFC3279, April 2002, <https://www.rfc-editor.org/info/rfc3279>.

[RFC3971] Arkko, J., Ed., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, DOI 10.17487/RFC3971, March 2005, <https://www.rfc-editor.org/info/rfc3971>.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, DOI 10.17487/RFC3972, March 2005, <https://www.rfc-editor.org/info/rfc3972>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC5758] Dang, Q., Santesson, S., Moriarty, K., Brown, D., and T. Polk, "Internet X.509 Public Key Infrastructure: Additional Algorithms and Identifiers for DSA and ECDSA", RFC 5758, DOI 10.17487/RFC5758, January 2010, <https://www.rfc-editor.org/info/rfc5758>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <https://www.rfc-editor.org/info/rfc6775>.

10.2. Informative references

[FIPS186-4] "FIPS Publication 186-4: Digital Signature Standard", July 2013, <http://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.186-4.pdf>.

[I-D.ietf-6lo-backbone-router] Thubert, P., "IPv6 Backbone Router", draft-ietf-6lo-backbone-router-05 (work in progress), January 2018.

[I-D.struik-lwig-curve-representations] Struik, R., "Alternative Elliptic Curve Representations", draft-struik-lwig-curve-representations-00 (work in progress), November 2017.

[RFC4919] Kushalnagar, N., Montenegro, G., and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals", RFC 4919, DOI 10.17487/RFC4919, August 2007, <https://www.rfc-editor.org/info/rfc4919>.

[RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <https://www.rfc-editor.org/info/rfc4944>.

[RFC5889] Baccelli, E., Ed. and M. Townsley, Ed., "IP Addressing Model in Ad Hoc Networks", RFC 5889, DOI 10.17487/RFC5889, September 2010, <https://www.rfc-editor.org/info/rfc5889>.

[RFC6234] Eastlake 3rd, D. and T. Hansen, "US Secure Hash Algorithms (SHA and SHA-based HMAC and HKDF)", RFC 6234, DOI 10.17487/RFC6234, May 2011, <https://www.rfc-editor.org/info/rfc6234>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC7039] Wu, J., Bi, J., Bagnulo, M., Baker, F., and C. Vogt, Ed., "Source Address Validation Improvement (SAVI) Framework", RFC 7039, DOI 10.17487/RFC7039, October 2013, <https://www.rfc-editor.org/info/rfc7039>.

[RFC7102] Vasseur, JP., "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102, DOI 10.17487/RFC7102, January 2014, <https://www.rfc-editor.org/info/rfc7102>.

[RFC7217] Gont, F., "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)", RFC 7217, DOI 10.17487/RFC7217, April 2014, <https://www.rfc-editor.org/info/rfc7217>.

[RFC7696] Housley, R., "Guidelines for Cryptographic Algorithm Agility and Selecting Mandatory-to-Implement Algorithms", BCP 201, RFC 7696, DOI 10.17487/RFC7696, November 2015, <https://www.rfc-editor.org/info/rfc7696>.

[RFC7721] Cooper, A., Gont, F., and D. Thaler, "Security and Privacy Considerations for IPv6 Address Generation Mechanisms", RFC 7721, DOI 10.17487/RFC7721, March 2016, <https://www.rfc-editor.org/info/rfc7721>.

[RFC7748] Langley, A., Hamburg, M., and S. Turner, "Elliptic Curves for Security", RFC 7748, DOI 10.17487/RFC7748, January 2016, <https://www.rfc-editor.org/info/rfc7748>.

[RFC8032] Josefsson, S. and I. Liusvaara, "Edwards-Curve Digital Signature Algorithm (EdDSA)", RFC 8032, DOI 10.17487/RFC8032, January 2017, <https://www.rfc-editor.org/info/rfc8032>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

Appendix A. Requirements Addressed in this Document

In this section we state requirements of a secure neighbor discovery protocol for low-power and lossy networks.

- The protocol MUST be based on the Neighbor Discovery Optimization for Low-power and Lossy Networks protocol defined in [RFC6775]. RFC6775 utilizes optimizations such as host-initiated interactions for sleeping resource-constrained hosts and elimination of multicast address resolution.

- New options to be added to Neighbor Solicitation messages MUST lead to small packet sizes, especially compared with existing protocols such as SEcure Neighbor Discovery (SEND). Smaller packet sizes facilitate low-power transmission by resource-constrained nodes on lossy links.

- The support for this registration mechanism SHOULD be extensible to more LLN links than IEEE 802.15.4 only. Support for at least the LLN links for which a 6lo "IPv6 over foo" specification exists, as well as Low-Power Wi-Fi SHOULD be possible.

- As part of this extension, a mechanism to compute a unique Identifier should be provided with the capability to form a Link Local Address that SHOULD be unique at least within the LLN connected to a 6LBR.

- The Address Registration Option used in the ND registration SHOULD be extended to carry the relevant forms of Unique Interface IDentifier.

- The Neighbour Discovery should specify the formation of a site-local address that follows the security recommendations from [RFC7217].

Authors' Addresses

Pascal Thubert (editor)
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
MOUGINS - Sophia Antipolis 06254
FRANCE

Phone: +33 497 23 26 34
Email: pthubert@cisco.com

Behcet Sarikaya
Plano, TX
USA

Email: sarikaya@ieee.org

Mohit Sethi
Ericsson
Hirsalantie
Jorvas 02420

Email: mohit@piuha.net

6lo
Internet-Draft
Intended status: Standards Track
Expires: August 27, 2018

P. Thubert, Ed.
cisco

February 23, 2018

IPv6 Backbone Router
draft-ietf-6lo-backbone-router-06

Abstract

This specification proposes proxy operations for IPv6 Neighbor Discovery on behalf of devices located on broadcast-inefficient wireless networks. A broadcast-efficient backbone running classical IPv6 Neighbor Discovery federates multiple wireless links to form a large MultiLink Subnet, but the broadcast domain does not need to extend to the wireless links for the purpose of ND operation. Backbone Routers placed at the wireless edge of the backbone proxy the ND operation and route packets from/to registered nodes, and wireless nodes register or are proxy-registered to the Backbone Router to setup proxy services in a fashion that is essentially similar to a classical Layer-2 association.

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 August 27, 2018.

Copyright Notice

Table of Contents

1. Introduction

One of the key services provided by IEEE std. 802.1 [IEEEstd8021] Ethernet Bridging is an efficient and reliable broadcast service, and multiple applications and protocols have been built that heavily depend on that feature for their core operation. But a wide range of wireless networks do not provide the solid and cheap broadcast capabilities of Ethernet Bridging, and protocols designed for bridged networks that rely on broadcast often exhibit disappointing behaviours when applied unmodified to a wireless medium.

IEEE std. 802.11 [IEEEstd80211] Access Points (APs) deployed in an Extended Service Set (ESS) effectively act as bridges, but, in order to ensure a solid connectivity to the devices and protect the medium against harmful broadcasts, they refrain from relying on broadcast-intensive protocols such as Transparent Bridging on the wireless side. Instead, an association process is used to register proactively the MAC addresses of the wireless device (STA) to the AP, and then the APs proxy the bridging operation and cancel the broadcasts.

Classical IPv6 [RFC8200] Neighbor Discovery [RFC4861] [RFC4862] Protocol (NDP) operations are reactive and rely heavily on multicast operations to locate an on-link correspondent and ensure address uniqueness, which is a pillar that sustains the whole IP architecture. When the Duplicate Address Detection [RFC4862] (DAD) mechanism was designed, it was a natural match with the efficient broadcast operation of Ethernet Bridging, but with the unreliable broadcast that is typical of wireless media, DAD is bound to fail to discover duplications [I-D.yourtchenko-6man-dad-issues]. In other words, because the broadcast service is unreliable, DAD appears to work on wireless media not because address duplication is detected and solved as designed, but because the duplication is a very rare event as a side effect of the sheer amount of entropy in 64-bits Interface IDs.

In the real world, IPv6 multicast messages are effectively broadcast, so they are processed by most if not all wireless nodes over the ESS fabric even when very few if any of the nodes is effectively listening to the multicast address. It results that a simple Neighbor Solicitation (NS) lookup message [RFC4861], that is supposedly targeted to a very small group of nodes, ends up polluting the whole wireless bandwidth across the fabric [I-D.vyncke-6man-mcast-not-efficient]. In other words, the reactive IPv6 ND operation leads to undesirable power consumption in battery-operated devices.

The inefficiencies of using radio broadcasts to support IPv6 NDP lead the community to consider (again) splitting the broadcast domain between the wired and the wireless access links. One classical way to achieve this is to split the subnet in multiple ones, and at the extreme provide a /64 per wireless device. Another is to proxy the Layer-3 protocols that rely on broadcast operation at the boundary of the wired and wireless domains, effectively emulating the Layer-2 association at layer-3. To that effect, the current IEEE std. 802.11 specifications require the capability to perform ARP and ND proxy [RFC4389] functions at the Access Points (APs).

But for the lack a comprehensive specification for the ND proxy and in particular the lack of an equivalent to an association process, implementations have to rely on snooping for acquiring the related state, which is unsatisfactory in a lossy and mobile conditions. With snooping, a state (e.g. a new IPv6 address) may not be discovered or a change of state (e.g. a movement) may be missed, leading to unreliable connectivity.

In the context of IEEE std. 802.15.4 [IEEEstd802154], the step of considering the radio as a medium that is different from Ethernet was already taken with the publication of Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs) [RFC6775]. RFC 6775 is updated as [I-D.ietf-6lo-rfc6775-update]; the update includes changes that are required by this document.

This specification applies that same thinking to other wireless links such as Low-Power IEEE std. 802.11 (Wi-Fi) and IEEE std. 802.15.1 (Bluetooth) [IEEEstd802151], and extends [RFC6775] to enable proxy operation by the 6BBR so as to decouple the broadcast domain in the backbone from the wireless links. The proxy operation can be maintained asynchronous so that low-power nodes or nodes that are deep in a mesh do not need to be bothered synchronously when a lookup is performed for their addresses, effectively implementing the ND contribution to the concept of a Sleep Proxy [I-D.nordmark-6man-dad-approaches].

2. Applicability and Requirements Served

Efficiency aware IPv6 Neighbor Discovery Optimizations [I-D.chakrabarti-nordmark-6man-efficient-nd] suggests that 6LoWPAN ND [RFC6775] can be extended to other types of links beyond IEEE std. 802.15.4 for which it was defined. The registration technique is beneficial when the Link-Layer technique used to carry IPv6 multicast packets is not sufficiently efficient in terms of delivery ratio or energy consumption in the end devices, in particular to enable energy-constrained sleeping nodes. The value of such extension is especially apparent in the case of mobile wireless nodes, to reduce the multicast operations that are related to classical ND ([RFC4861], [RFC4862]) and plague the wireless medium.

This specification updates and generalizes 6LoWPAN ND to a broader range of Low power and Lossy Networks (LLNs) with a solid support for Duplicate Address Detection (DAD) and address lookup that does not require broadcasts over the LLNs. The term LLN is used loosely in this specification to cover multiple types of WLANs and WPANs, including Low-Power Wi-Fi, BLUETOOTH(R) Low Energy, IEEE std. 802.11AH and IEEE std. 802.15.4 wireless meshes, so as to address the requirements listed in Appendix A.3

The scope of this draft is a Backbone Link that federates multiple LLNs as a single IPv6 MultiLink Subnet. Each LLN in the subnet is anchored at an IPv6 Backbone Router (6BBR). The Backbone Routers interconnect the LLNs over the Backbone Link and emulate that the LLN nodes are present on the Backbone using proxy-ND operations. This specification extends IPv6 ND over the backbone to discriminate address movement from duplication and eliminate stale state in the backbone routers and backbone nodes once a LLN node has roamed. This way, mobile nodes may roam rapidly from a 6BBR to the next and requirements in Appendix A.1 are met.

This specification can be used by any wireless node to associate at Layer-3 with a 6BBR and register its IPv6 addresses to obtain routing services including proxy-ND operations over the backbone, effectively providing a solution to the requirements expressed in Appendix A.4.

The Link Layer Address (LLA) that is returned as Target LLA (TLLA) in Neighbor Advertisements (NA) messages by the 6BBR on behalf of the Registered Node over the backbone may be that of the Registering Node, in which case the 6BBR needs to bridge the unicast packets (Bridging proxy), or that of the 6BBR on the backbone, in which case the 6BBRs needs to route the unicast packets (Routing proxy). In the latter case, the 6BBR may maintain the list of correspondents to which it has advertised its own MAC address on behalf of the LLN node and the IPv6 ND operation is minimized as the number of nodes scale up in the LLN. This enables to meet the requirements in Appendix A.6 as long has the 6BBRs are dimensioned for the number of registration that each needs to support.

In the context of the the TimeSlotted Channel Hopping (TSCH) mode of [IEEEstd802154], the 6TiSCH architecture [I-D.ietf-6tisch-architecture] introduces how a 6LoWPAN ND host could connect to the Internet via a RPL mesh Network, but this requires additions to the 6LOWPAN ND protocol to support mobility and reachability in a secured and manageable environment. This specification details the new operations that are required to implement the 6TiSCH architecture and serves the requirements listed in Appendix A.2.

In the case of Low-Power IEEE std. 802.11, a 6BBR may be collocated with a standalone AP or a CAPWAP [RFC5415] wireless controller, and the wireless client (STA) leverages this specification to register its IPv6 address(es) to the 6BBR over the wireless medium. In the case of a 6TiSCH LLN mesh, the RPL root is collocated with a 6LoWPAN Border Router (6LBR), and either collocated with or connected to the 6BBR over an IPv6 Link. The 6LBR leverages this specification to register the LLN nodes on their behalf to the 6BBR. In the case of

BTLE, the 6BBR is collocated with the router that implements the BTLE central role as discussed in section 2.2 of [RFC7668].

3. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Readers are expected to be familiar with all the terms and concepts that are discussed in "Neighbor Discovery for IP version 6" [RFC4861], "IPv6 Stateless Address Autoconfiguration" [RFC4862], "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals" [RFC4919], Neighbor Discovery Optimization for Low-power and Lossy Networks [RFC6775] and "Multi-link Subnet Support in IPv6" [I-D.ietf-ipv6-multilink-subnets].

Readers would benefit from reading "Multi-Link Subnet Issues" [RFC4903], ,"Mobility Support in IPv6" [RFC6275], "Neighbor Discovery Proxies (ND Proxy)" [RFC4389] and "Optimistic Duplicate Address Detection" [RFC4429] prior to this specification for a clear understanding of the art in ND-proxying and binding.

Additionally, this document uses terminology from [RFC7102], [I-D.ietf-6lo-rfc6775-update] and [I-D.ietf-6tisch-terminology], and introduces the following terminology:

Sleeping Proxy A 6BBR acts as a Sleeping Proxy if it answers ND Neighbor Solicitation over the backbone on behalf of the Registered Node whenever possible. This is the default mode for this specification but it may be overridden, for instance by configuration, into Unicasting Proxy.

Unicasting Proxy As a Unicasting Proxy, the 6BBR forwards NS messages to the Registering Node, transforming Layer-2 multicast into unicast whenever possible.

Routing proxy A 6BBR acts as a routing proxy if it advertises its own MAC address, as opposed to that of the node that performs the registration, as the TLLA in the proxied NAs over the backbone. In that case, the MAC address of the node is not visible at Layer-2 over the backbone and the bridging fabric is not aware of the addresses of the LLN devices and their mobility. The 6BBR installs a connected host route towards the registered node over the interface to the node, and acts as a Layer-3 router for unicast packets to the node. The 6BBR updates the ND Neighbor Cache Entries (NCE) in correspondent

nodes if the wireless node moves and registers to another 6BBR, either with a single broadcast, or with a series of unicast NA(O) messages, indicating the TLLA of the new router.

Bridging proxy A 6BBR acts as a bridging proxy if it advertises the MAC address of the node that performs the registration as the TLLA in the proxied NAs over the backbone. In that case, the MAC address and the mobility of the node is still visible across the bridged backbone fabric, as is traditionally the case with Layer-2 APs. The 6BBR acts as a Layer-2 bridge for unicast packets to the registered node. The MAC address exposed in the S/TLLA is that of the Registering Node, which is not necessarily the Registered Device. When a device moves within a LLN mesh, it may end up attached to a different 6LBR acting as Registering Node, and the LLA that is exposed over the backbone will change.

Primary BBR The BBR that will defend a Registered Address for the purpose of DAD over the backbone.

Secondary BBR A BBR to which the address is registered. A Secondary Router MAY advertise the address over the backbone and proxy for it.

4. Overview

An LLN node can move freely from an LLN anchored at a Backbone Router to an LLN anchored at another Backbone Router on the same backbone and conserve any of the IPv6 addresses that it has formed, transparently.

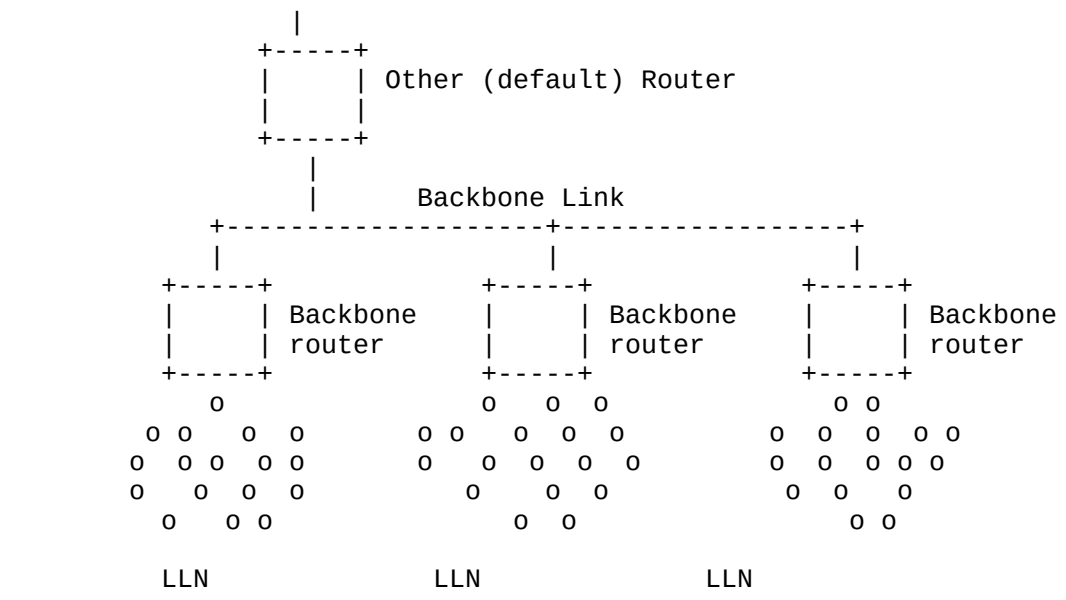


Figure 1: Backbone Link and Backbone Routers

The Backbone Routers maintain an abstract Binding Table of their Registered Nodes. The Binding Table operates as a distributed database of all the wireless Nodes whether they reside on the LLNs or on the backbone, and use an extension to the Neighbor Discovery Protocol to exchange that information across the Backbone in the classical ND reactive fashion.

The Extended Address Registration Option (EARO) defined in [I-D.ietf-6lo-rfc6775-update] is used to enable the registration for routing and proxy option is included in the ND exchanges over the backbone between the 6BBRs to sort out duplication from movement.

Address duplication is sorted out with the Owner Unique-ID field in the EARO, which is a generalization of the EUI-64 that allows different types of unique IDs beyond the name space derived from the MAC addresses. First-Come First-Serve rules apply, whether the duplication happens between LLN nodes as represented by their respective 6BBRs, or between an LLN node and a classical node that defends its address over the backbone with classical ND and does not include the EARO option.

In case of conflicting registrations to multiple 6BBRs from a same node, a sequence counter called Transaction ID (TID) in the EARO enables 6BBRs to sort out the latest anchor for that node.

Registrations with a same TID are compatible and maintained, but, in case of different TIDs, only the freshest registration is maintained and the stale state is eliminated. The EARO also transports a 'R' flag to be used by a 6LN when registering, to indicate that this 6LN is not a router and that it will not handle its own reachability.

With this specification, Backbone Routers perform a ND proxy operation over the Backbone Link on behalf of their Registered Nodes. The registration to the proxy service is done with a NS/NA(EARO) exchange. The EARO option with a 'R' flag is used in this specification to indicate to the 6BBR that it is expected to perform this proxy operation. The Backbone Router operation is essentially similar to that of a Mobile IPv6 (MIPv6) [RFC6275] Home Agent. This enables mobility support for LLN nodes that would move outside of the network delimited by the Backbone link attach to a Home Agent from that point on. This also enables collocation of Home Agent functionality within Backbone Router functionality on the same backbone interface of a router. Further specification may extend this be allowing the 6BBR to redistribute host routes in routing protocols that would operate over the backbone, or in MIPv6 or the Locator/ID Separation Protocol (LISP) [RFC6830] to support mobility on behalf of the nodes, etc...

The Optimistic Duplicate Address Detection [RFC4429] (ODAD) specification details how an address can be used before a Duplicate Address Detection (DAD) is complete, and insists that an address that is TENTATIVE should not be associated to a Source Link-Layer Address Option in a Neighbor Solicitation message. This specification leverages ODAD to create a temporary proxy state in the 6BBR till DAD is completed over the backbone. This way, the specification enables to distribute proxy states across multiple 6BBR and co-exist with classical ND over the backbone.

5. Backbone Router Routing Operations

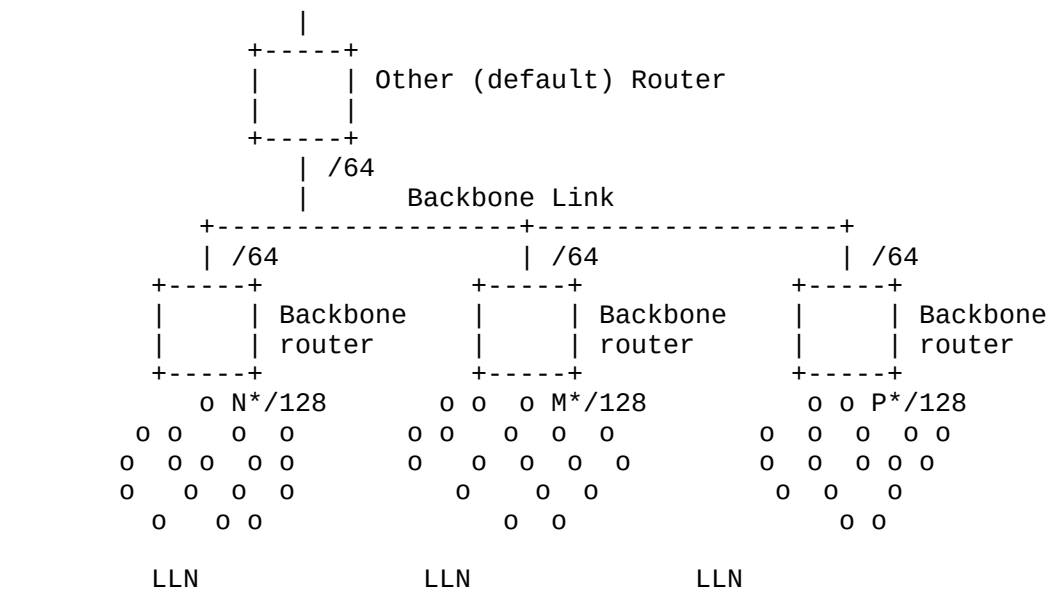


Figure 2: Routing Configuration in the ML Subnet

5.1. Over the Backbone Link

The Backbone Router is a specific kind of Border Router that performs proxy Neighbor Discovery on its backbone interface on behalf of the nodes that it has discovered on its LLN interfaces.

The backbone is expected to be a high speed, reliable Backbone link, with affordable and reliable multicast capabilities, such as a bridged Ethernet Network, and to allow a full support of classical ND as specified in [RFC4861] and subsequent RFCs. In other words, the backbone is not a LLN.

Still, some restrictions of the attached LLNs will apply to the backbone. In particular, it is expected that the MTU is set to the same value on the backbone and all attached LLNs, and the scalability of the whole subnet requires that broadcast operations are avoided as much as possible on the backbone as well. Unless configured otherwise, the Backbone Router MUST echo the MTU that it learns in RAs over the backbone in the RAs that it sends towards the LLN links.

As a router, the Backbone Router behaves like any other IPv6 router on the backbone side. It has a connected route installed towards the backbone for the prefixes that are present on that backbone and that it proxies for on the LLN interfaces.

As a proxy, the 6BBR uses an EARO option in the NS-DAD and the multicast NA messages that it generates over the Backbone Link on behalf of a Registered Node, and it places an EARO in its unicast NA messages, if and only if the NS/NA that stimulates it had an EARO in it and the 'R' bit set.

When possible, the 6BBR SHOULD use unicast or solicited-node multicast address (SNMA) [RFC4291] to defend its Registered Addresses over the backbone. In particular, the 6BBR MUST join the SNMA group that corresponds to a Registered Address as soon as it creates an entry for that address and as long as it maintains that entry, whatever the state of the entry. The expectation is that it is possible to get a message delivered to all the nodes on the backbone that listen to a particular address and support this specification - which includes all the 6BBRs in the MultiLink Subnet - by sending a multicast message to the associated SNMA over the backbone.

The support of Optimistic DAD (ODAD) [RFC4429] is recommended for all nodes in the backbone and followed by the 6BBRs in their proxy activity over the backbone. With ODAD, any optimistic node MUST join the SNMA of a Tentative address, which interacts better with this specification.

This specification allows the 6BBR in Routing Proxy mode to advertise the Registered IPv6 Address with the 6BBR Link Layer Address, and attempts to update Neighbor Cache Entries (NCE) in correspondent nodes over the backbone, using gratuitous NA(Override). This method may fail of the multicast message is not properly received, and correspondent nodes may maintain an incorrect neighbor state, which they will eventually discover through Neighbor Unreachability Detection (NUD). Because mobility may be slow, the NUD procedure defined in [RFC4861] may be too impatient, and the support of [RFC7048] is recommended in all nodes in the network.

Since the MultiLink Subnet may grow very large in terms of individual IPv6 addresses, multicasts should be avoided as much as possible even on the backbone. Though it is possible for plain hosts to participate with legacy IPv6 ND support, the support by all nodes connected to the backbone of [I-D.ietf-6man-rs-refresh] is recommended, and this implies the support of [RFC7559] as well.

5.2. Over the LLN Link

As a router, the Nodes and Backbone Router operation on the LLN follows [RFC6775]. Per that specification, LLN Hosts generally do not depend on multicast RAs to discover routers. It is still generally required for LLN nodes to accept multicast RAs [RFC7772], but those are rare on the LLN link. Nodes are expected to follow the

Simple Procedures for Detecting Network Attachment in IPv6 [RFC6059] (DNA procedures) to assert movements, and to support the Packet-Loss Resiliency for Router Solicitations [RFC7559] to make the unicast RS more reliable.

An LLN node signals that it requires IPv6 ND proxy services from a 6BBR by registering the corresponding IPv6 Address with an NS(EARO) message with the 'R' flag set. The LLN node that performs the registration (the Registering Node) may be the owner of the IPv6 Address (the Registered Node) or a 6LBR that performs the registration on its behalf.

When operating as a Routing Proxy, the router installs hosts routes (/128) to the Registered Addresses over the LLN links, via the Registering Node as identified by the Source Address and the SLLAO option in the NS(EARO) messages.

In that mode, the 6BBR handles the ND protocol over the backbone on behalf of the Registered Nodes, using its own MAC address in the TLLA and SLLA options in proxyed NS and NA messages. It results that for each Registered Address, a number of peer Nodes on the backbone have resolved the address with the 6BBR MAC address and keep that mapping stored in their Neighbor cache.

The 6BBR SHOULD maintain, per Registered Address, the list of the peers on the backbone to which it answered with it MAC address, and when a binding moves to a different 6BBR, it SHOULD send a unicast gratuitous NA(O) individually to each of them to inform them that the address has moved and pass the MAC address of the new 6BBR in the TLLAO option. If the 6BBR can not maintain that list, then it SHOULD remember whether that list is empty or not and if not, send a multicast NA(O) to all nodes to update the impacted Neighbor Caches with the information from the new 6BBR.

The Bridging Proxy is a variation where the BBR function is implemented in a Layer-3 switch or an wireless Access Point that acts as a Host from the IPv6 standpoint, and, in particular, does not operate the routing of IPv6 packets. In that case, the SLLAO in the proxied NA messages is that of the Registering Node and classical bridging operations take place on data frames.

If a registration moves from one 6BBR to the next, but the Registering Node does not change, as indicated by the S/TLLAO option in the ND exchanges, there is no need to update the Neighbor Caches in the peers Nodes on the backbone. On the other hand, if the LLAO changes, the 6BBR SHOULD inform all the relevant peers as described above, to update the impacted Neighbor Caches. In the same fashion,

if the Registering Node changes with a new registration, the 6BBR SHOULD also update the impacted Neighbor Caches over the backbone.

6. BackBone Router Proxy Operations

This specification enables a Backbone Router to proxy Neighbor Discovery operations over the backbone on behalf of the nodes that are registered to it, allowing any node on the backbone to reach a Registered Node as if it was on-link. The backbone and the LLNs are considered different Links in a MultiLink subnet but the prefix that is used may still be advertised as on-link on the backbone to support legacy nodes; multicast ND messages are link-scoped and not forwarded across the backbone routers.

ND Messages on the backbone side that do not match to a registration on the LLN side are not acted upon on the LLN side, which stands protected. On the LLN side, the prefixes associated to the MultiLink Subnet are presented as not on-link, so address resolution for other hosts do not occur.

The default operation in this specification is Sleeping proxy which means:

- creating a new entry in an abstract Binding Table for a new Registered Address and validating that the address is not a duplicate over the backbone
- defending a Registered Address over the backbone using NA messages with the Override bit set on behalf of the sleeping node whenever possible
- advertising a Registered Address over the backbone using NA messages, asynchronously or as a response to a Neighbor Solicitation messages.
- Looking up a destination over the backbone in order to deliver packets arriving from the LLN using Neighbor Solicitation messages.
- Forwarding packets from the LLN over the backbone, and the other way around.
- Eventually triggering a liveliness verification of a stale registration.

A 6BBR may act as a Sleeping Proxy only if the state of the binding entry is REACHABLE, or TENTATIVE in which case the answer is delayed.

In any other state, the Sleeping Proxy operates as a Unicasting Proxy.

As a Unicasting Proxy, the 6BBR forwards NS lookup messages to the Registering Node, transforming Layer-2 multicast into unicast whenever possible. This is not possible in UNREACHABLE state, so the NS messages are multicasted, and rate-limited to protect the medium with an exponential back-off. In other states, The messages are forwarded to the Registering Node as unicast Layer-2 messages. In TENTATIVE state, the NS message is either held till DAD completes, or dropped.

The draft introduces the optional concept of primary and secondary BBRs. The primary is the backbone router that has the highest EUI-64 address of all the 6BBRs that share a registration for a same Registered Address, with the same Owner Unique ID and same Transaction ID, the EUI-64 address being considered as an unsigned 64bit integer. The concept is defined with the granularity of an address, that is a given 6BBR can be primary for a given address and secondary or another one, regardless on whether the addresses belong to the same node or not. The primary Backbone Router is in charge of protecting the address for DAD over the Backbone. Any of the Primary and Secondary 6BBR may claim the address over the backbone, since they are all capable to route from the backbone to the LLN node, and the address appears on the backbone as an anycast address.

The Backbone Routers maintain a distributed binding table, using classical ND over the backbone to detect duplication. This specification requires that:

1. All addresses that can be reachable from the backbone, including IPv6 addresses based on burn-in EUI64 addresses MUST be registered to the 6BBR.

2. A Registered Node MUST include the EARO option in an NS message that used to register an addresses to a 6LR; the 6LR MUST propagate that option unchanged to the 6LBR in the DAR/DAC exchange, and the 6LBR MUST propagate that option unchanged in proxy registrations.

3. The 6LR MUST echo the same EARO option in the NA that it uses to respond, but for the status filed which is not used in NS messages, and significant in NA.

A false positive duplicate detection may arise over the backbone, for instance if the Registered Address is registered to more than one LBR, or if the node has moved. Both situations are handled gracefully unbeknownst to the node. In the former case, one LBR

becomes primary to defend the address over the backbone while the others become secondary and may still forward packets back and forth. In the latter case the LBR that receives the newest registration wins and becomes primary.

The expectation in this specification is that there is a single Registering Node at a time per Backbone Router for a given Registered Address, but that a Registered Address may be registered to Multiple 6BBRs for higher availability.

Over the LLN, and for any given Registered Address, it is REQUIRED that:

de-registrations (newer TID, same OUID, null Lifetime) are accepted and responded immediately with a status of 4; the entry is deleted;

newer registrations (newer TID, same OUID, non-null Lifetime) are accepted and responded with a status of 0 (success); the entry is updated with the new TID, the new Registration Lifetime and the new Registering Node, if any has changed; in TENTATIVE state the response is held and may be overwritten; in other states the Registration-Lifetime timer is restarted and the entry is placed in REACHABLE state.

identical registrations (same TID, same OUID) from a same Registering Node are not processed but responded with a status of 0 (success); they are expected to be identical and an error may be logged if not; in TENTATIVE state, the response is held and may be overwritten, but it MUST be eventually produced and it carries the result of the DAD process;

older registrations (not(newer or equal) TID, same OUID) from a same Registering Node are ignored;

identical and older registrations (not-newer TID, same OUID) from a different Registering Node are responded immediately with a status of 3 (moved); this may be rate limited to protect the medium;

and any registration for a different Registered Node (different OUID) are responded immediately with a status of 1 (duplicate).

6.1. Registration and Binding State Creation

Upon a registration for a new address with an NS(EARO) with the 'R' bit set, the 6BBR performs a DAD operation over the backbone placing the new address as target in the NS-DAD message. The EARO from the

registration MUST be placed unchanged in the NS-DAD message, and an entry is created in TENTATIVE state for a duration of TENTATIVE_DURATION. The NS-DAD message is sent multicast over the backbone to the SNMA address associated with the registered address. If that operation is known to be costly, and the 6BBR has an indication from another source (such as a NCE) that the Registered Address was present on the backbone, that information may be leveraged to send the NS-DAD message as a Layer-2 unicast to the MAC that was associated with the Registered Address.

In TENTATIVE state:

- the entry is removed if an NA is received over the backbone for the Registered Address with no EARO option, or with an EARO option with a status of 1 (duplicate) that indicates an existing registration for another LLN node. The OUID and TID fields in the EARO option received over the backbone are ignored. A status of 1 is returned in the EARO option of the NA back to the Registering Node;

- the entry is also removed if an NA with an ARO option with a status of 3 (moved), or a NS-DAD with an ARO option that indicates a newer registration for the same Registered Node, is received over the backbone for the Registered Address. A status of 3 is returned in the NA(EARO) back to the Registering Node;

- when a registration is updated but not deleted, e.g. from a newer registration, the DAD process on the backbone continues and the running timers are not restarted;

- Other NS (including DAD with no EARO option) and NA from the backbone are not responded in TENTATIVE state, but the list of their origins may be kept in memory and if so, the 6BBR may send them each a unicast NA with eventually an EARO option when the TENTATIVE_DURATION timer elapses, so as to cover legacy nodes that do not support ODAD.

- When the TENTATIVE_DURATION timer elapses, a status 0 (success) is returned in a NA(EARO) back to the Registering Node(s), and the entry goes to REACHABLE state for the Registration Lifetime; the DAD process is successful and the 6BBR MUST send a multicast NA(EARO) to the SNMA associated to the Registered Address over the backbone with the Override bit set so as to take over the binding from other 6BBRs.

6.2. Defending Addresses

If a 6BBR has an entry in REACHABLE state for a Registered Address:

- If the 6BBR is primary, or does not support the function of primary, it MUST defend that address over the backbone upon an incoming NS-DAD, either if the NS does not carry an EARO, or if an EARO is present that indicates a different Registering Node (different OUID). The 6BBR sends a NA message with the Override bit set and the NA carries an EARO option if and only if the NS-DAD did so. When present, the EARO in the NA(O) that is sent in response to the NS-DAD(EARO) carries a status of 1 (duplicate), and the OUID and TID fields in the EARO option are obfuscated with null or random values to avoid network scanning and impersonation attacks.

- If the 6BBR receives an NS-DAD(EARO) that reflect a newer registration, the 6BBR updates the entry and the routing state to forward packets to the new 6BBR, but keeps the entry REACHABLE. In that phase, it MAY use REDIRECT messages to reroute traffic for the Registered Address to the new 6BBR.

- If the 6BBR receives an NA(EARO) that reflect a newer registration, the 6BBR removes its entry and sends a NA(AERO) with a status of 3 (moved) to the Registering Node, if the Registering Node is different from the Registered Node. If necessary, the 6BBR cleans up ND cache in peers nodes as discussed in Section 5.1, by sending a series of unicast to the impacted nodes, or one broadcast NA(O) to all-nodes.

- If the 6BBR received a NS(LOOKUP) for a Registered Address, it answers immediately with an NA on behalf of the Registered Node, without polling it. There is no need of an EARO in that exchange.

- When the Registration-Lifetime timer elapses, the entry goes to STALE state for a duration of STABLE_STALE_DURATION in LLNs that keep stable addresses such as LWPANs, and UNSTABLE_STALE_DURATION in LLNs where addresses are renewed rapidly, e.g. for privacy reasons.

The STALE state is a chance to keep track of the backbone peers that may have an ND cache pointing on this 6BBR in case the Registered Address shows back up on this or a different 6BBR at a later time. In STALE state:

- If the Registered Address is claimed by another node on the backbone, with an NS-DAD or an NA, the 6BBR does not defend the address. Upon an NA(O), or the stale time elapses, the 6BBR

removes its entry and sends a NA(AERO) with a status of 4 (removed) to the Registering Node.

- o If the 6BBR received a NS(LOOKUP) for a Registered Address, the 6BBR MUST send an NS(NUD) following rules in [RFC7048] to the Registering Node targeting the Registered Address prior to answering. If the NUD succeeds, the operation in REACHABLE state applies. If the NUD fails, the 6BBR refrains from answering the lookup. The NUD expected to be mapped by the Registering Node into a liveliness validation of the Registered Node if they are in fact different nodes.

7. Security Considerations

This specification expects that the link layer is sufficiently protected, either by means of physical or IP security for the Backbone Link or MAC sublayer cryptography. In particular, it is expected that the LLN MAC provides secure unicast to/from the Backbone Router and secure Broadcast from the Backbone Router in a way that prevents tempering with or replaying the RA messages.

The use of EUI-64 for forming the Interface ID in the link local address prevents the usage of Secure ND ([RFC3971] and [RFC3972]) and address privacy techniques. This specification RECOMMENDS the use of additional protection against address theft such as provided by [I-D.ietf-6lo-ap-nd], which guarantees the ownership of the OUID.

When the ownership of the OUID cannot be assessed, this specification limits the cases where the OUID and the TID are multicasted, and obfuscates them in responses to attempts to take over an address.

8. Protocol Constants

This Specification uses the following constants:

TENTATIVE_DURATION: 800 milliseconds

STABLE_STALE_DURATION: 24 hours

UNSTABLE_STALE_DURATION: 5 minutes

DEFAULT_NS_POLLING: 3 times

9. IANA Considerations

This document has no request to IANA.

10. Acknowledgments

Kudos to Eric Levy-Abegnoli who designed the First Hop Security infrastructure at Cisco.

11. References

11.1. Normative References

[I-D.ietf-6lo-rfc6775-update]
Thubert, P., Nordmark, E., Chakrabarti, S., and C. Perkins, "An Update to 6LoWPAN ND", draft-ietf-6lo-rfc6775-update-13 (work in progress), February 2018.

[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>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4429] Moore, N., "Optimistic Duplicate Address Detection (DAD) for IPv6", RFC 4429, DOI 10.17487/RFC4429, April 2006, <https://www.rfc-editor.org/info/rfc4429>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC6059] Krishnan, S. and G. Daley, "Simple Procedures for Detecting Network Attachment in IPv6", RFC 6059, DOI 10.17487/RFC6059, November 2010, <https://www.rfc-editor.org/info/rfc6059>.

[RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <https://www.rfc-editor.org/info/rfc6550>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <https://www.rfc-editor.org/info/rfc6775>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

11.2. Informative References

[I-D.chakrabarti-nordmark-6man-efficient-nd] Chakrabarti, S., Nordmark, E., Thubert, P., and M. Wasserman, "IPv6 Neighbor Discovery Optimizations for Wired and Wireless Networks", draft-chakrabarti-nordmark-6man-efficient-nd-07 (work in progress), February 2015.

[I-D.delcarpio-6lo-wlanah] Vega, L., Robles, I., and R. Morabito, "IPv6 over 802.11ah", draft-delcarpio-6lo-wlanah-01 (work in progress), October 2015.

[I-D.ietf-6lo-ap-nd] Thubert, P., Sarikaya, B., and M. Sethi, "Address Protected Neighbor Discovery for Low-power and Lossy Networks", draft-ietf-6lo-ap-nd-06 (work in progress), February 2018.

[I-D.ietf-6lo-nfc] Choi, Y., Hong, Y., Youn, J., Kim, D., and J. Choi, "Transmission of IPv6 Packets over Near Field Communication", draft-ietf-6lo-nfc-09 (work in progress), January 2018.

[I-D.ietf-6man-rs-refresh] Nordmark, E., Yourtchenko, A., and S. Krishnan, "IPv6 Neighbor Discovery Optional RS/RA Refresh", draft-ietf-6man-rs-refresh-02 (work in progress), October 2016.

[I-D.ietf-6tisch-architecture] Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-13 (work in progress), November 2017.

[I-D.ietf-6tisch-terminology]
Palattella, M., Thubert, P., Watteyne, T., and Q. Wang, "Terminology in IPv6 over the TSCH mode of IEEE 802.15.4e", draft-ietf-6tisch-terminology-09 (work in progress), June 2017.

[I-D.ietf-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-08 (work in progress), September 2017.

[I-D.ietf-ipv6-multilink-subnets]
Thaler, D. and C. Huitema, "Multi-link Subnet Support in IPv6", draft-ietf-ipv6-multilink-subnets-00 (work in progress), July 2002.

[I-D.nordmark-6man-dad-approaches]
Nordmark, E., "Possible approaches to make DAD more robust and/or efficient", draft-nordmark-6man-dad-approaches-02 (work in progress), October 2015.

[I-D.popa-6lo-6loplc-ipv6-over-ieee19012-networks]
Popa, D. and J. Hui, "6LoPLC: Transmission of IPv6 Packets over IEEE 1901.2 Narrowband Powerline Communication Networks", draft-popa-6lo-6loplc-ipv6-over-ieee19012-networks-00 (work in progress), March 2014.

[I-D.vyncke-6man-mcast-not-efficient]
Vyncke, E., Thubert, P., Levy-Abegnoli, E., and A. Yourtchenko, "Why Network-Layer Multicast is Not Always Efficient At Datalink Layer", draft-vyncke-6man-mcast-not-efficient-01 (work in progress), February 2014.

[I-D.yourtchenko-6man-dad-issues]
Yourtchenko, A. and E. Nordmark, "A survey of issues related to IPv6 Duplicate Address Detection", draft-yourtchenko-6man-dad-issues-01 (work in progress), March 2015.

[RFC3810] Vida, R., Ed. and L. Costa, Ed., "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, DOI 10.17487/RFC3810, June 2004, <https://www.rfc-editor.org/info/rfc3810>.

[RFC3971] Arkko, J., Ed., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, DOI 10.17487/RFC3971, March 2005, <https://www.rfc-editor.org/info/rfc3971>.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, DOI 10.17487/RFC3972, March 2005, <https://www.rfc-editor.org/info/rfc3972>.

[RFC4389] Thaler, D., Talwar, M., and C. Patel, "Neighbor Discovery Proxies (ND Proxy)", RFC 4389, DOI 10.17487/RFC4389, April 2006, <https://www.rfc-editor.org/info/rfc4389>.

[RFC4903] Thaler, D., "Multi-Link Subnet Issues", RFC 4903, DOI 10.17487/RFC4903, June 2007, <https://www.rfc-editor.org/info/rfc4903>.

[RFC4919] Kushalnagar, N., Montenegro, G., and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals", RFC 4919, DOI 10.17487/RFC4919, August 2007, <https://www.rfc-editor.org/info/rfc4919>.

[RFC5415] Calhoun, P., Ed., Montemurro, M., Ed., and D. Stanley, Ed., "Control And Provisioning of Wireless Access Points (CAPWAP) Protocol Specification", RFC 5415, DOI 10.17487/RFC5415, March 2009, <https://www.rfc-editor.org/info/rfc5415>.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <https://www.rfc-editor.org/info/rfc6275>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, DOI 10.17487/RFC6830, January 2013, <https://www.rfc-editor.org/info/rfc6830>.

[RFC7048] Nordmark, E. and I. Gashinsky, "Neighbor Unreachability Detection Is Too Impatient", RFC 7048, DOI 10.17487/RFC7048, January 2014, <https://www.rfc-editor.org/info/rfc7048>.

[RFC7102] Vasseur, JP., "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102, DOI 10.17487/RFC7102, January 2014, <https://www.rfc-editor.org/info/rfc7102>.

[RFC7217] Gont, F., "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)", RFC 7217, DOI 10.17487/RFC7217, April 2014, <https://www.rfc-editor.org/info/rfc7217>.

[RFC7428] Brandt, A. and J. Buron, "Transmission of IPv6 Packets over ITU-T G.9959 Networks", RFC 7428, DOI 10.17487/RFC7428, February 2015, <https://www.rfc-editor.org/info/rfc7428>.

[RFC7559] Krishnan, S., Anipko, D., and D. Thaler, "Packet-Loss Resiliency for Router Solicitations", RFC 7559, DOI 10.17487/RFC7559, May 2015, <https://www.rfc-editor.org/info/rfc7559>.

[RFC7668] Nieminen, J., Savolainen, T., Isomaki, M., Patil, B., Shelby, Z., and C. Gomez, "IPv6 over BLUETOOTH(R) Low Energy", RFC 7668, DOI 10.17487/RFC7668, October 2015, <https://www.rfc-editor.org/info/rfc7668>.

[RFC7772] Yourtchenko, A. and L. Colitti, "Reducing Energy Consumption of Router Advertisements", BCP 202, RFC 7772, DOI 10.17487/RFC7772, February 2016, <https://www.rfc-editor.org/info/rfc7772>.

[RFC8105] Mariager, P., Petersen, J., Ed., Shelby, Z., Van de Logt, M., and D. Barthel, "Transmission of IPv6 Packets over Digital Enhanced Cordless Telecommunications (DECT) Ultra Low Energy (ULE)", RFC 8105, DOI 10.17487/RFC8105, May 2017, <https://www.rfc-editor.org/info/rfc8105>.

[RFC8163] Lynn, K., Ed., Martocci, J., Neilson, C., and S. Donaldson, "Transmission of IPv6 over Master-Slave/Token-Passing (MS/TP) Networks", RFC 8163, DOI 10.17487/RFC8163, May 2017, <https://www.rfc-editor.org/info/rfc8163>.

11.3. External Informative References

[IEEEstd8021]
IEEE standard for Information Technology, "IEEE Standard for Information technology-- Telecommunications and information exchange between systems Local and metropolitan area networks Part 1: Bridging and Architecture".

[IEEEstd80211]
IEEE standard for Information Technology, "IEEE Standard for Information technology-- Telecommunications and information exchange between systems Local and metropolitan area networks-- Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications".

[IEEEstd802151]
IEEE standard for Information Technology, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements. - Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs)".

[IEEEstd802154]
IEEE standard for Information Technology, "IEEE Standard for Local and metropolitan area networks-- Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)".

Appendix A. Requirements

This section lists requirements that were discussed at 6lo for an update to 6LoWPAN ND. This specification meets most of them, but those listed in Appendix A.5 which are deferred to a different specification such as [I-D.ietf-6lo-ap-nd].

A.1. Requirements Related to Mobility

Due to the unstable nature of LLN links, even in a LLN of immobile nodes a 6LoWPAN Node may change its point of attachment to a 6LR, say 6LR-a, and may not be able to notify 6LR-a. Consequently, 6LR-a may still attract traffic that it cannot deliver any more. When links to a 6LR change state, there is thus a need to identify stale states in a 6LR and restore reachability in a timely fashion.

Req1.1: Upon a change of point of attachment, connectivity via a new 6LR MUST be restored timely without the need to de-register from the previous 6LR.

Req1.2: For that purpose, the protocol MUST enable to differentiate between multiple registrations from one 6LoWPAN Node and registrations from different 6LoWPAN Nodes claiming the same address.

Req1.3: Stale states MUST be cleaned up in 6LRs.

Req1.4: A 6LoWPAN Node SHOULD also be capable to register its Address to multiple 6LRs, and this, concurrently.

A.2. Requirements Related to Routing Protocols

The point of attachment of a 6LoWPAN Node may be a 6LR in an LLN mesh. IPv6 routing in a LLN can be based on RPL, which is the routing protocol that was defined at the IETF for this particular purpose. Other routing protocols than RPL are also considered by Standard Defining Organizations (SDO) on the basis of the expected network characteristics. It is required that a 6LoWPAN Node attached via ND to a 6LR would need to participate in the selected routing protocol to obtain reachability via the 6LR.

Next to the 6LBR unicast address registered by ND, other addresses including multicast addresses are needed as well. For example a routing protocol often uses a multicast address to register changes to established paths. ND needs to register such a multicast address to enable routing concurrently with discovery.

Multicast is needed for groups. Groups MAY be formed by device type (e.g. routers, street lamps), location (Geography, RPL sub-tree), or both.

The Bit Index Explicit Replication (BIER) Architecture [I-D.ietf-bier-architecture] proposes an optimized technique to enable multicast in a LLN with a very limited requirement for routing state in the nodes.

Related requirements are:

Req2.1: The ND registration method SHOULD be extended in such a fashion that the 6LR MAY advertise the Address of a 6LoWPAN Node over the selected routing protocol and obtain reachability to that Address using the selected routing protocol.

Req2.2: Considering RPL, the Address Registration Option that is used in the ND registration SHOULD be extended to carry enough information to generate a DAO message as specified in [RFC6550] section 6.4, in particular the capability to compute a Path Sequence and, as an option, a RPLInstanceID.

Req2.3: Multicast operations SHOULD be supported and optimized, for instance using BIER or MPL. Whether ND is appropriate for the registration to the 6BBR is to be defined, considering the additional burden of supporting the Multicast Listener Discovery Version 2 [RFC3810] (MLDv2) for IPv6.

A.3. Requirements Related to the Variety of Low-Power Link types

6LoWPAN ND [RFC6775] was defined with a focus on IEEE std. 802.15.4 and in particular the capability to derive a unique Identifier from a globally unique MAC-64 address. At this point, the 6lo Working Group is extending the 6LoWPAN Header Compression (HC) [RFC6282] technique to other link types ITU-T G.9959 [RFC7428], Master-Slave/Token-Passing [RFC8163], DECT Ultra Low Energy [RFC8105], Near Field Communication [I-D.ietf-6lo-nfc], IEEE std. 802.11ah [I-D.delcarpio-6lo-wlanah], as well as IEEE1901.2 Narrowband Powerline Communication Networks [I-D.popa-6lo-6loplc-ipv6-over-ieee19012-networks] and BLUETOOTH(R) Low Energy [RFC7668].

Related requirements are:

Req3.1: The support of the registration mechanism SHOULD be extended to more LLN links than IEEE 802.15.4, matching at least the LLN links for which an "IPv6 over foo" specification exists, as well as Low-Power Wi-Fi.

Req3.2: As part of this extension, a mechanism to compute a unique Identifier should be provided, with the capability to form a Link-Local Address that SHOULD be unique at least within the LLN connected to a 6LBR discovered by ND in each node within the LLN.

Req3.3: The Address Registration Option used in the ND registration SHOULD be extended to carry the relevant forms of unique Identifier.

Req3.4: The Neighbour Discovery should specify the formation of a site-local address that follows the security recommendations from [RFC7217].

A.4. Requirements Related to Proxy Operations

Duty-cycled devices may not be able to answer themselves to a lookup from a node that uses classical ND on a backbone and may need a proxy. Additionally, the duty-cycled device may need to rely on the 6LBR to perform registration to the 6BBR.

The ND registration method SHOULD defend the addresses of duty-cycled devices that are sleeping most of the time and not capable to defend their own Addresses.

Related requirements are:

Req4.1: The registration mechanism SHOULD enable a third party to proxy register an Address on behalf of a 6LoWPAN node that may be sleeping or located deeper in an LLN mesh.

Req4.2: The registration mechanism SHOULD be applicable to a duty-cycled device regardless of the link type, and enable a 6BBR to operate as a proxy to defend the registered Addresses on its behalf.

Req4.3: The registration mechanism SHOULD enable long sleep durations, in the order of multiple days to a month.

A.5. Requirements Related to Security

In order to guarantee the operations of the 6LoWPAN ND flows, the spoofing of the 6LR, 6LBR and 6BBRs roles should be avoided. Once a node successfully registers an address, 6LoWPAN ND should provide energy-efficient means for the 6LBR to protect that ownership even when the node that registered the address is sleeping.

In particular, the 6LR and the 6LBR then should be able to verify whether a subsequent registration for a given Address comes from the original node.

In a LLN it makes sense to base security on layer-2 security. During bootstrap of the LLN, nodes join the network after authorization by a Joining Assistant (JA) or a Commissioning Tool (CT). After joining nodes communicate with each other via secured links. The keys for the layer-2 security are distributed by the JA/CT. The JA/CT can be part of the LLN or be outside the LLN. In both cases it is needed that packets are routed between JA/CT and the joining node.

Related requirements are:

Req5.1: 6LoWPAN ND security mechanisms SHOULD provide a mechanism for the 6LR, 6LBR and 6BBR to authenticate and authorize one another for their respective roles, as well as with the 6LoWPAN Node for the role of 6LR.

Req5.2: 6LoWPAN ND security mechanisms SHOULD provide a mechanism for the 6LR and the 6LBR to validate new registration of authorized nodes. Joining of unauthorized nodes MUST be impossible.

Req5.3: 6LoWPAN ND security mechanisms SHOULD lead to small packet sizes. In particular, the NS, NA, DAR and DAC messages for a re-registration flow SHOULD NOT exceed 80 octets so as to fit in a secured IEEE std. 802.15.4 frame.

Req5.4: Recurrent 6LoWPAN ND security operations MUST NOT be computationally intensive on the LoWPAN Node CPU. When a Key hash calculation is employed, a mechanism lighter than SHA-1 SHOULD be preferred.

Req5.5: The number of Keys that the 6LoWPAN Node needs to manipulate SHOULD be minimized.

Req5.6: The 6LoWPAN ND security mechanisms SHOULD enable CCM* for use at both Layer 2 and Layer 3, and SHOULD enable the reuse of security code that has to be present on the device for upper layer security such as TLS.

Req5.7: Public key and signature sizes SHOULD be minimized while maintaining adequate confidentiality and data origin authentication for multiple types of applications with various degrees of criticality.

Req5.8: Routing of packets should continue when links pass from the unsecured to the secured state.

Req5.9: 6LoWPAN ND security mechanisms SHOULD provide a mechanism for the 6LR and the 6LBR to validate whether a new registration for a given address corresponds to the same 6LoWPAN Node that registered it initially, and, if not, determine the rightful owner, and deny or clean-up the registration that is duplicate.

A.6. Requirements Related to Scalability

Use cases from Automatic Meter Reading (AMR, collection tree operations) and Advanced Metering Infrastructure (AMI, bi-directional communication to the meters) indicate the needs for a large number of LLN nodes pertaining to a single RPL DODAG (e.g. 5000) and connected to the 6LBR over a large number of LLN hops (e.g. 15).

Related requirements are:

Req6.1: The registration mechanism SHOULD enable a single 6LBR to register multiple thousands of devices.

Req6.2: The timing of the registration operation should allow for a large latency such as found in LLNs with ten and more hops.

Author's Address

Pascal Thubert (editor)
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
MOUGINS - Sophia Antipolis 06254
FRANCE

Phone: +33 497 23 26 34
Email: pthubert@cisco.com

6lo
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2020

Lijo Thomas
C-DAC
S. Anamalamudi
SRM University-AP
S.V.R.Anand
Malati Hegde
Indian Institute of Science
C. Perkins
Futurewei
July 8, 2019

Packet Delivery Deadline time in 6LoWPAN Routing Header
draft-ietf-6lo-deadline-time-05

Abstract

This document specifies a new type for the 6LoWPAN routing header containing the deadline time for data packets, designed for use over constrained networks. The deadline time enables forwarding and scheduling decisions for time critical IoT machine to machine (M2M) applications that operate within time-synchronized networks that agree on the meaning of the time representations used for the deadline time values.

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 January 9, 2020.

Copyright Notice

Table of Contents

1. Introduction

Low Power and Lossy Networks (LLNs) are likely to be deployed for real time industrial applications requiring end-to-end delay guarantees [I-D.ietf-detnet-use-cases]. A Deterministic Network ("detnet") typically requires some data packets to reach their receivers within strict time bounds. Intermediate nodes use the deadline information to make appropriate packet forwarding and scheduling decisions to meet the time bounds.

This document specifies a new type for the Elective 6LoWPAN Routing Header (6LoRHE), so that the deadline time (i.e., the time of latest acceptable delivery) of data packets can be included within the 6LoWPAN routing header. [RFC8138] specifies the 6LoWPAN Routing Header (6LoRH), compression schemes for RPL routing (source routing) operation [RFC6554], header compression of RPL Packet Information [RFC6553], and IP-in-IP encapsulation. This document also specifies handling of the deadline time when packets traverse between time-synchronized networks operating in different timezones or distinct reference clocks. Time synchronization techniques are outside the scope of this document. There are a number of standards available for this purpose, including IEEE 1588 [ieee-1588], IEEE 802.1AS [dot1AS-2011], IEEE 802.15.4-2015 TSCH [dot15-tsch], and more.

The Deadline-6LoRHE can be used in any time synchronized 6Lo network. A 6TiSCH network is used to describe the implementation of the Deadline-6LoRHE, but this does not preclude its use in scenarios other than 6TiSCH. For instance, there is a growing interest in using 6lo over a BLE mesh network [I-D.ietf-6lo-blemesh] in industrial IoT [dotBLEMesh]. BLE mesh time synchronization is being explored by the Bluetooth community. There are also cases under consideration in Wi-SUN [Wi-SUN_PHY], [dotWi-SUN].

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 [RFC2119] [RFC8174].

This document uses the terminology defined in [RFC6550] and [I-D.ietf-6tisch-terminology].

3. 6LoRHE Generic Format

Note: this section is not normative and is included for convenience. The generic header format of the 6LoRHE is specified in [I-D.ietf-roll-routing-dispatch]. Figure 1 illustrates the 6LoRHE generic format.

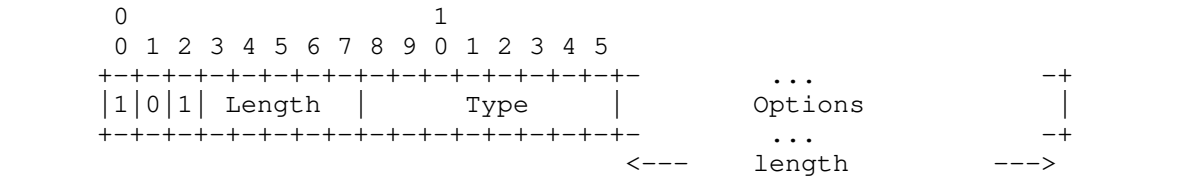


Figure 1: 6LoRHE format

- Length: Length of the 6LoRHE expressed in bytes, excluding the first 2 bytes. This enables a node to skip a 6LoRHE if the Type is not recognized/supported.

- Type (variable length): Type of the 6LoRHE (see Section 7)

4. Deadline-6LoRHE

The Deadline-6LoRHE (see Figure 3) is an elective 6LoRH (i.e., a 6LoRHE [RFC8138]) that provides the Deadline Time (DT) for an IPv6 datagram in a compressed form. Along with the deadline, the header can include the packet Origination Time Delta (OTD), the time at which the packet is enqueued for transmission (expressed as a value to be subtracted from DT); this enables a close estimate of the total delay incurred by a packet. The OTD field is initialized by the sender based on the current time at the outgoing network interface through which the packet is forwarded. Since the OTD is a delta, the length of the OTD field (i.e., OTL) will require fewer bits than the length of the DT field (i.e., DTL).

The deadline field contains the value of the deadline time for the packet -- in other words, the time by which the application expects the packet to be delivered to the Receiver.

```
packet_deadline_time = packet_origination_time + max_delay
```

In order to support delay-sensitive deterministic applications, all nodes within the network should process the Deadline-6LoRHE. The packet deadline time (DT) and origination time (OTD) are represented in time units determined by a scaling parameter in the routing header. The Network ASN (Absolute Slot Number) can be used as a time unit in a time slotted synchronized network (for instance a 6TiSCH network, where global time is maintained in the units of slot lengths of a certain resolution).

The delay experienced by packets in the network is a useful metric for network diagnostics and performance monitoring. Whenever a packet crosses into a network using a different reference clock, the Destination Time field is updated to represent the same Destination Time, but expressed using the reference clock of the interface into the new network. Then the origination time is the same as the current time when the packet is transmitted into the new network, minus the delay already experienced by the packet, say 'current_dly'. In this way, within the newly entered network, the packet will appear to have originated 'current_dly' time units earlier with respect to the reference clock of the new network.

```
new_network_origin_time = time_now_in_new_network - current_dly
```

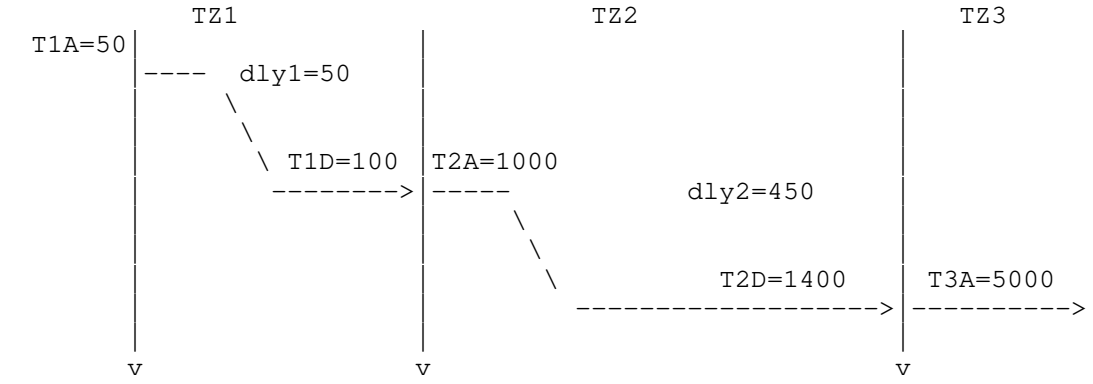
The following example illustrates these calculations when a packet travels between three networks, each in a different time zone. 'x' can be 1, 2 or 3. Suppose that the deadline time as measured in timezone 1 is 1050 and the origination time is 50. Suppose that the difference between TZ2 and TZ1 is 900, and the difference between TZ3 and TZ3 is 3600. In the figure, OT is the origination time as measured in the current timezone, and is equal to DT - OTD, that is, DT - 1000. Figure 2 uses the following abbreviations:

TxA : Time of arrival of packet in the network 'x'

TxD : Departure time of packet from the network 'x'

dlyx : Delay experienced by the packet in the previous network(s)

TZx : The time zone of network 'x'



dly0 = 0 dly1 = T1D-OT1 dly2 = T2D-OT2
= 100-50 = 1400 - 950
= 50 = 450

OT1 = T1A-dly0 OT2 = T2A-dly1 OT3 = T3A-dly2
= 50 = 1000-50 = 5000 - 450
= 950 = 4550

Figure 2: Destination Time Update example

There are multiple ways that a packet can be delayed, including queuing delay, MAC layer contention delay, serialization delay, and propagation delays. Sometimes there are processing delays as well. For the purpose of determining whether or not the deadline has already passed, these various delays are not distinguished.

5. Deadline-6LoRHE Format

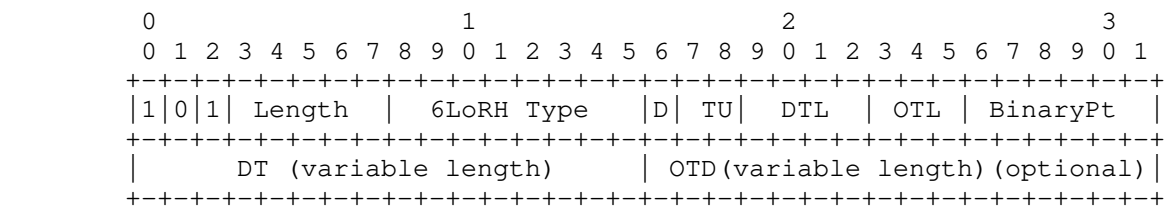


Figure 3: Deadline-6LoRHE format

- Length (5 bits): Length represents the total length of the Deadline-6LoRHE type measured in octets.
- 6LoRH Type: TBD (see Section 7)
- D flag (1 bit): The 'D' flag, set by the Sender, qualifies the action to be taken when a 6LR detects that the deadline time has elapsed. If 'D' bit is 1, then the 6LR MUST drop the packet if the deadline time is elapsed. If 'D' bit is 0, the packet MAY be forwarded on an exception basis, if the forwarding node is NOT in a situation of constrained resource, and if there are reasons to suspect that downstream nodes might find it useful (delay measurements, interpolations, etc.).
- TU (2 bits) : Indicates the time units for DT and OTD fields. The encodings for the DT and OTD fields use the same time units and precision.

  * 00 : Time represented in seconds and fractional seconds
  * 01 : Reserved
  * 10 : Network ASN
  * 11 : Reserved
- DTL (4 bits): Length of DT field as an unsigned 4-bit integer, encoding the length of the field in hex digits, minus one.
- OTL (3 bits) : Length of OTD field as an unsigned 3-bit integer, encoding the length of the field in hex digits. If OTL == 0, the OTD field is not present. The value of OTL MUST NOT exceed the value of DTL plus one.

  * For example, DTL = 0b0000 means the deadline time in the 6LoRHE is 1 hex digit (4 bits) long. OTL = 0b111 means the origination time is 7 hex digits (28 bits) long.
- Binary Pt (6 bits) : If zero, the number of bits of the integer part the DT is equal to the number of bits of the fractional part of the DT. if nonzero, the Binary Pt is a signed integer determining the position of the binary point within the value for the DT.

    * If BinaryPt value is positive, then the number of bits for the integer part of the DT is increased by the value of BinaryPt, and the number of bits for the fractional part of the DT is correspondingly reduced. This increases the range of DT.
    * If BinaryPt value is negative, then the number of bits for the integer part of the DT is decreased by the value of BinaryPt, and the number of bits for the fractional part of the DT is correspondingly increased. This increases the precision of the fractional seconds part of DT.
- o DT Value (8..64-bit) : An unsigned integer of DTL+1 hex digits giving the Deadline Time value
- o OTD Value (8..64-bit) : An unsigned integer of OTL hex digits giving the Origination Time as a negative offset from the DT value

Whenever a sender initiates the IP datagram, it includes the Deadline-6LoRHE along with other 6LoRH information. For information about the time synchronization requirements between sender and receiver see Section 8.

For the chosen time unit, a compressed time representation is available as follows. First, the application on the originating node has to determine how many time bits are needed to represent the difference between the time at which the packet is launched and the deadline time, including the representation of fractional time units. That number of bits (say, N_bits) determines DTL (the length of the Deadline Time (DT)) as follows:

```
DTL = (N_bits mod 4)
```

The number of bits determined by DTL allows counting any number of fractional time units in the range of interest determined by DT and the origination time OT. Denote this number of fractional time units to be Epoch_Range(DTL) (i.e., Epoch_Range is a function of DTL).

```
Epoch_Range(DTL) = (2^(4*(DTL+1))
```

Each point of time between OT and DT is represented by a time unit and a fractional time unit; in this section, this combined representation is called a rational time unit (RTU). 1 RTU measures the smallest fractional time that can be represented between two points of time in the epoch (i.e., within the range of interest).

DT - OT cannot exceed 2^(4*(DTL+1)) == 16^(DTL+1). A low value of DTL leads to a small Epoch_Range; if DTL = 0, there will only be 16 RTUs within the Epoch_Range (DTL) = 16^1 (for any time unit TU). The values that can be represented in the current epoch are in the range [0, (Epoch_Range(DTL) - 1)]. To minimize the required DTL,

wraparound is allowed but works naturally with the arithmetic modulo Epoch_Range.

By default, DTL determines t_0 in the chosen RTUs as follows:

t_0 = [current_time - (current_time mod Epoch_Range (DTL))].

Naturally, t_0 occurs at time 0 (or time 0.0000...) in the current epoch. The last possible origination time representable in the current epoch (counted in RTUs) is t_last = (t0 + (2^(4*(DTL+1))-1)). In the RTUs chosen, the current epoch resides at the underlying time interval [t_0, t_last]. If DT - OT is greater than t_last - OT, then wraparound within the Epoch_Range occurs naturally. In all cases, OT is represented by the value (OT mod Epoch_Range) and DT is represented by the value (DT mod Epoch_Range). All arithmetic is to be performed modulo (Epoch_Range(DTL)), yielding only positive values for DT - OT.

Example: Consider a 6TiSCH network with time-slot length of 10ms. Let the time units be ASNs (TU == (binary)0b10). Let the current ASN when the packet is originated be 54400, and the maximum allowable delay (max_delay) for the packet delivery be 1 second from the packet origination, then:

deadline_time = packet_origination_time + max_delay

= 0xD480 + 0x64 (Network ASNs)

= 0xD4E4 (Network ASNs)

Then, the Deadline-6LoRHE encoding with nonzero OTL is:

DTL = 3, OTL = 2, TU = 0b10, BinaryPt = 8, DT = 0xD4E4, OTD = 0x64

6. Deadline-6LoRHE in Three Network Scenarios

In this section, Deadline-6LoRHE operation is described for 3 network scenarios. Figure 4 depicts a constrained time-synchronized LLN that has two subnets N1 and N2, connected through LBRs [I-D.ietf-6lo-backbone-router] with different reference clock times T1 and T2.

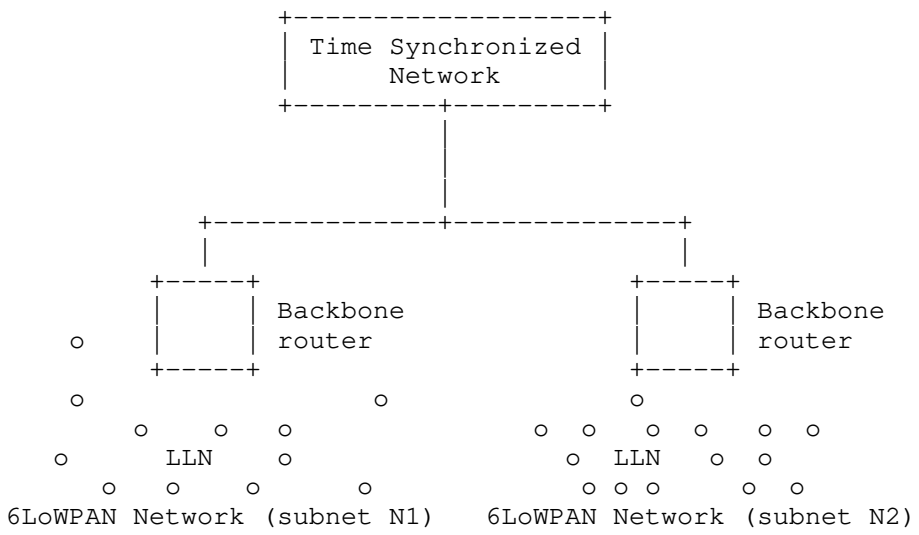


Figure 4: Intra-network Timezone Scenario

6.1. Scenario 1: Endpoints in the same DODAG (N1)

In scenario 1, shown in Figure 5, the Sender 'S' has an IP datagram to be routed to a Receiver 'R' within the same DODAG. For the route segment from Sender to 6LBR, the Sender includes a Deadline-6LoRHE by encoding the deadline time contained in the packet. Subsequently, each 6LR will perform hop-by-hop routing to forward the packet towards the 6LBR. Once 6LBR receives the IP datagram, it sends the packet downstream towards 'R'.

In case of a network running RPL non-storing mode, the 6LBR generates a IPv6-in-IPv6 encapsulated packet when sending the packet downwards to the Receiver [I-D.ietf-roll-useofrplinfo]. The 6LBR copies the Deadline-6LoRHE from the Sender originated IP header to the outer IP header. The Deadline-6LoRHE contained in the inner IP header is removed.

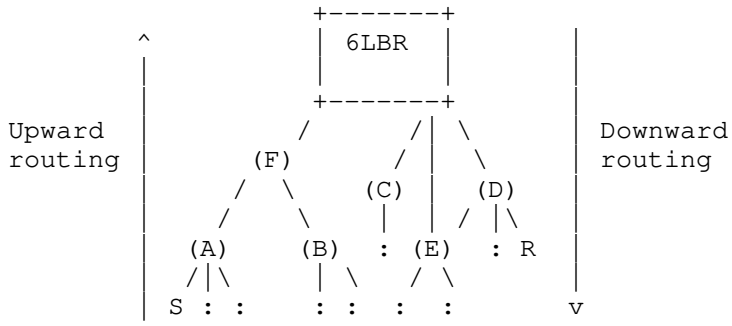


Figure 5: End points within same DODAG (subnet N1)

At the tunnel endpoint of the encapsulation, the Deadline-6LoRHE is copied back from the outer header to inner header, and the inner IP packet is delivered to 'R'.

6.2. Scenario 2: Endpoints in Networks with Dissimilar L2 Technologies.

In scenario 2, shown in Figure 6, the Sender 'S' (belonging to DODAG 1) has IP datagram to be routed to a Receiver 'R' over a time-synchronized IPv6 network. For the route segment from 'S' to 6LBR, 'S' includes a Deadline-6LoRHE. Subsequently, each 6LR will perform hop-by-hop routing to forward the packet towards the 6LBR. Once the Deadline Time information reaches the border router, the packet will be encoded according to the mechanism prescribed in the other time-synchronized network depicted as "Time Synchronized Network" in the figure 6. The specific data encapsulation mechanisms followed in the new network are beyond the scope of this document.

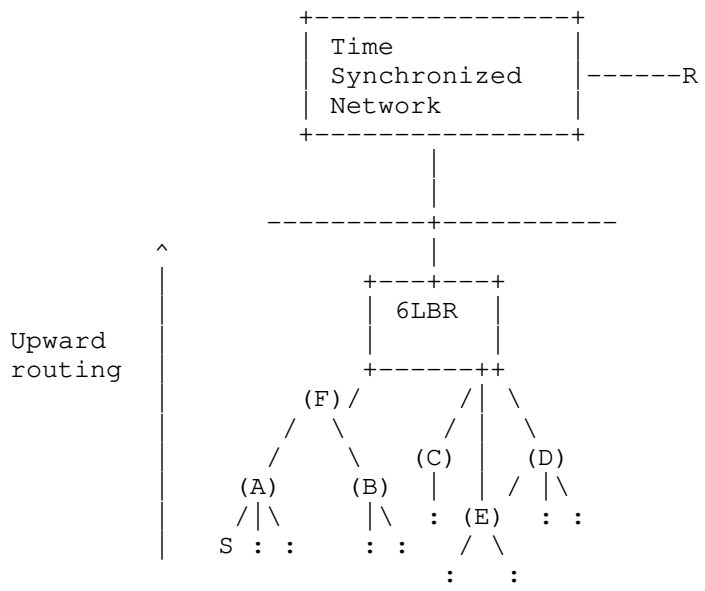


Figure 6: Packet transmission in Dissimilar L2 Technologies or Internet

For instance, the IP datagram could be routed to another time synchronized deterministic network using the mechanism specified in the In-band OAM [I-D.ietf-ippm-ioam-data], and then the deadline time would be updated according to the measurement of the current time in the new network.

6.3. Scenario 3: Packet transmission across different DODAGs (N1 to N2).

Consider the scenario depicted in Figure 7, in which the Sender 'S' (belonging to DODAG 1) has an IP datagram to be sent to Receiver 'R' belonging to another DODAG (DODAG 2). The operation of this scenario can be decomposed into combination of case 1 and case 2 scenarios. For the route segment from 'S' to 6LBR1, 'S' includes the Deadline-6LoRHE. Subsequently, each 6LR will perform hop-by-hop operation to forward the packet towards the 6LBR1. Once the IP datagram reaches 6LBR1 of DODAG1, it applies the same rule as described in Case 2 while routing the packet to 6LBR2 over a (likely) time synchronized wired backhaul. The wired side of 6LBR2 can be mapped to receiver of Case 2. Once the packet reaches 6LBR2, it updates the Deadline-6LoRHE by adding or subtracting the difference of time of DODAG2 and sends the packet downstream towards 'R'.

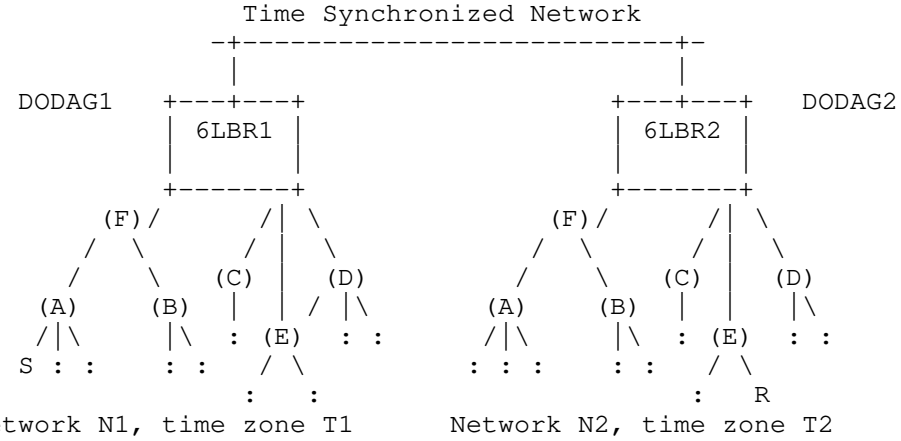


Figure 7: Packet transmission in different DODAGs(N1 to N2)

Consider an example of a 6TiSCH network in which S in DODAG1 generates the packet at ASN 20000 to R in DODAG2. Let the maximum allowable delay be 1 second. The time-slot length in DODAG1 and DODAG2 is assumed to be 10ms. Once the deadline time is encoded in Deadline-6LoRHE, the packet is forwarded to 6LBR of DODAG1. Suppose the packet reaches 6LBR of DODAG1 at ASN 20030.

```
current_time = ASN at LBR * slot_length_value

remaining_time = deadline_time - current_time
= ((packet_origination_time + max_delay) - current time)
= (20000 + 100) - 20030
= 30 (in Network ASNs)
= 30 * 10^3 milliseconds.
```

Once the Deadline Time information reaches the border router, the packet will be encoded according to the mechanism prescribed in the other time-synchronized network.

7. IANA Considerations

This document defines a new Elective 6LoWPAN Routing Header Type, and IANA is requested to assign a value (TBD) from the 6LoWPAN Dispatch Page1 number space for this purpose.

Elective 6LoRH Type	Value
Deadline-6LoRHE	TBD

Figure 8: Deadline-6LoRHE type

8. Synchronization Aspects

The document supports time representation of the deadline and origination times carried in the packets traversing through networks of different time zones having different time synchronization mechanisms. For instance, in a 6TiSCH network where the time is maintained as ASN time slots, the time synchronization is achieved through beaconing among the nodes as described in [RFC7554]. There could be 6lo networks that employ NTP where the nodes are synchronized with an external reference clock from an NTP server. The specification of the time synchronization method that need to be followed by a network is beyond the scope of the document.

The number of hex digits chosen to represent DT, and the portion of that field allocated to represent integer number of seconds, determines the meaning of t_0, i.e., the meaning of DT == 0 in the chosen representation. If DTL == 0, then there are only 4 bits that can be used to count the time units, so that DT == 0 can never be more than 16 time units (or fractional time units) in the past. This then requires that the time synchronization between sender and receiver has to be tighter than 16 units. If the binary point were moved so that all the bits were used for fractional time units (e.g., fractional seconds or fractional ASNs), the time synchronization requirement would be correspondingly tighter.

A 4-bit field for DT allows up to 16 hex digits, which is 64 bits. That is enough to represent the NTP [RFC5905] 64-bit timestamp format, which is more than enough for the purposes of establishing deadline times. Unless the binary point is moved, this is enough to represent time since year 1900.

For example, suppose that DTL = 0b0000 and the DT bits are split evenly; then we can count up to 3.75 seconds by quarter-seconds.

If DTL = 3 and the DT bits are again split evenly, then we can count up to 256 seconds (in steps of 1/256 of a second).

In all cases, t_0 is defined as specified in Section 5

```
t_0 = [current_time - (current_time mod (2^(4*(DTL+1))))]
```

regardless of the choice of TU.

For TU = 0b00, the time units are seconds. With DTL == 15, and Binary Pt == 0, the epoch is (by default) January 1, 1900 at 00:00 UTC. The resolution is then (2 ^ (- 32)) seconds, which is the maximum possible. This time format wraps around every 2^32 seconds, which is roughly 136 years.

For TU = 0b10, the time units are ASNs. The start time is relative, and updated by a mechanism out of scope for this document. With 10 ms slots, DTL = 15, and Binary Pt == 0, it would take over a year for the ASN to wrap around. Typically, the number of hex digits allocated for TU = 0b10 would be less than 15.

9. Security Considerations

The security considerations of [RFC4944], [RFC6282] and [RFC6553] apply. Using a compressed format as opposed to the full in-line format is logically equivalent and does not create an opening for a new threat when compared to [RFC6550], [RFC6553] and [RFC6554].

The protocol elements specified in this document are designed to work in controlled operational environments (e.g., industrial process control and automation). In order to avoid misuse of the deadline information that could potentially result in a Denial of Service (DoS) attack, proper functioning of this deadline time mechanism requires the provisioning and management of network resources for supporting traffic flows with deadlines, performance monitoring, and admission control policy enforcement. The network provisioning can be done either centrally or in a distributed fashion. For example, tracks in a 6tisch network could be established by a centralized PCE, as described in the 6tisch architecture [I-D.ietf-6tisch-architecture].

The Security Considerations of Detnet architecture [I-D.ietf-detnet-architecture] mostly apply to this document as well, as follows. To secure the request and control of resources allocated for tracks, authentication and authorization can be used for each device, and network controller devices. In the case of distributed control protocols, security is expected to be provided by the security properties of the protocols in use.

When deadline bearing flows are identified on a per-flow basis, which may provide attackers with additional information about the data flows, when compared to networks that do not include per-flow identification. The security implications of disclosing that additional information deserve consideration when implementing this deadline specification.

Because of the requirement of precise time synchronization, the accuracy, availability, and integrity of time synchronization is of critical importance. Extensive discussion of this topic can be found in [RFC7384].

10. Acknowledgements

The authors thank Pascal Thubert for suggesting the idea and encouraging the work. Thanks to Shwetha Bhandari's suggestions which were instrumental in extending the timing information to heterogeneous networks. The authors acknowledge the 6TiSCH WG members for their inputs on the mailing list. Special thanks to Jerry Daniel, Dan Frost (Routing Directorate) Charlie Kaufman (Security Directorate) Seema Kumar, Tal Mizrahi Avinash Mohan, Shalu Rajendran, Anita Varghese, and Dale Worley (Gen-ART review) for their support and valuable feedback.

11. References

11.1. Normative References

[I-D.ietf-6tisch-terminology]
Palattella, M., Thubert, P., Watteyne, T., and Q. Wang, "Terms Used in IPv6 over the TSCH mode of IEEE 802.15.4e", draft-ietf-6tisch-terminology-10 (work in progress), March 2018.

[I-D.ietf-detnet-architecture]
Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-13 (work in progress), May 2019.

[I-D.ietf-roll-routing-dispatch]
Thubert, P., Bormann, C., Toutain, L., and R. Cragie, "6LoWPAN Routing Header", draft-ietf-roll-routing-dispatch-05 (work in progress), October 2016.

[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>.

[RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <https://www.rfc-editor.org/info/rfc4944>.

[RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, DOI 10.17487/RFC5905, June 2010, <https://www.rfc-editor.org/info/rfc5905>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <https://www.rfc-editor.org/info/rfc6550>.

[RFC6553] Hui, J. and JP. Vasseur, "The Routing Protocol for Low-Power and Lossy Networks (RPL) Option for Carrying RPL Information in Data-Plane Datagrams", RFC 6553, DOI 10.17487/RFC6553, March 2012, <https://www.rfc-editor.org/info/rfc6553>.

[RFC6554] Hui, J., Vasseur, JP., Culler, D., and V. Manral, "An IPv6 Routing Header for Source Routes with the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6554, DOI 10.17487/RFC6554, March 2012, <https://www.rfc-editor.org/info/rfc6554>.

[RFC7384] Mizrahi, T., "Security Requirements of Time Protocols in Packet Switched Networks", RFC 7384, DOI 10.17487/RFC7384, October 2014, <https://www.rfc-editor.org/info/rfc7384>.

[RFC7554] Watteyne, T., Ed., Palattella, M., and L. Grieco, "Using IEEE 802.15.4e Time-Slotted Channel Hopping (TSCH) in the Internet of Things (IoT): Problem Statement", RFC 7554, DOI 10.17487/RFC7554, May 2015, <https://www.rfc-editor.org/info/rfc7554>.

[RFC8138] Thubert, P., Ed., Bormann, C., Toutain, L., and R. Cragie, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing Header", RFC 8138, DOI 10.17487/RFC8138, April 2017, <https://www.rfc-editor.org/info/rfc8138>.

[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>.

11.2. Informative References

[dot15-tsch]
"IEEE 802 Wireless", "IEEE Standard for Low-Rate Wireless Networks, Part 15.4, IEEE Std 802.15.4-2015", April 2016.

[dot1AS-2011]
"IEEE Standards", "IEEE Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks", March 2011.

[dotBLEMesh]
Leonardi, L., Pattim, G., and L. Lo Bello, "Multi-Hop Real-Time Communications Over Bluetooth Low Energy Industrial Wireless Mesh Networks", IEEE Access Vol 6, 26505-26519, May 2018.

[dotWi-SUN]
Harada, H., Mizutani, K., Fujiwara, J., Mochizuki, K., Obata, K., and R. Okumura, "IEEE 802.15.4g Based Wi-SUN Communication Systems", IEICE Transactions on Communications volume E100.B, Jan 2017.

[I-D.ietf-6lo-backbone-router]
Thubert, P., Perkins, C., and E. Levy-Abegnoli, "IPv6 Backbone Router", draft-ietf-6lo-backbone-router-11 (work in progress), February 2019.

[I-D.ietf-6lo-blemesh]
Gomez, C., Darroudi, S., Savolainen, T., and M. Spoerk, "IPv6 Mesh over BLUETOOTH(R) Low Energy using IPSP", draft-ietf-6lo-blemesh-05 (work in progress), March 2019.

[I-D.ietf-6tisch-architecture]
Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-24 (work in progress), July 2019.

[I-D.ietf-detnet-use-cases]
Grossman, E., "Deterministic Networking Use Cases", draft-ietf-detnet-use-cases-20 (work in progress), December 2018.

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-06 (work in progress), July 2019.

[I-D.ietf-roll-useofrplinfo]
Robles, I., Richardson, M., and P. Thubert, "Using RPL Option Type, Routing Header for Source Routes and IPv6-in-IPv6 encapsulation in the RPL Data Plane", draft-ietf-roll-useofrplinfo-31 (work in progress), July 2019.

[ieee-1588]
"IEEE Standards", "IEEE Std 1588-2008 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", July 2008.

[Wi-SUN_PHY]
Wi-SUN Alliance, "Wi-SUN PHY Specification V1.0", March 2016.

Appendix A. Changes from revision 04 to revision 05

This section lists the changes between draft-ietf-6lo-deadline-time revisions ...-04.txt and ...-05.txt.

- Included additional relevant material in Security Considerations regarding expected deployment scenarios and the effect of disclosing additional information during the travel of a packet.
- Reworked the specification for using time ranges shorter than the maximum allowed by the choice of TU, so that fewer bits are needed to represent DT and OT.
- Revised the figures and examples to use new parameters
- Reordered the field definitions for the Deadline-6LoRHE.
- Responded to numerous reviewer comments to improve terminology and editorial consistency.

Appendix B. Changes from revision 03 to revision 04

This section lists the changes between draft-ietf-6lo-deadline-time revisions ...-03.txt and ...-04.txt.

- Replaced OT (Origination Time) field by OTD (Origination Time Delta), allowing a more compressed representation that needs less processing during transitions between networks.

- Changed representation for DTL, OTL, DT, OTD. Eliminated EXP in favor of BinaryPt.

- Revised the figures and examples to use new parameters

- Added new section on Synchronization Aspects to supply pertinent information about how nodes agree on the meaning of t=0.

- Responded to numerous reviewer comments to improve editorial consistency and improve terminology.

Appendix C. Changes from revision 02 to revision 03

This section lists the changes between draft-ietf-6lo-deadline-time revisions ...-02.txt and ...-03.txt.

- Added non-normative 6LoRHE description, citing RFC 8138.

- Specified that the Origination Time (OT) is the time that packet is enqueued for transmission.

- Mentioned more sources of packet delay.

- Clarified reasons that packet MAY be forwarded if 'D' bit is 0.

- Clarified that DT, OT, DTL and OTL are unsigned integers.

- Updated bibliographic citations, including BLEmesh and Wi-SUN.

Appendix D. Changes from revision 01 to revision 02

This section lists the changes between draft-ietf-6lo-deadline-time revisions ...-01.txt and ...-02.txt.

- Replaced 6LoRHE description by reference to RFC 8138.

- Added figure to illustrate change to Origination Time when a packet crosses timezone boundaries.

- Clarified that use of 6tisch networks is descriptive, not normative.

- Clarified that In-Band OAM is used as an example and is not normative.

- Updated bibliographic citations.

- Alphabetized contributor names.

Appendix E. Changes between earlier versions

This section lists the changes between draft-ietf-6lo-deadline-time revisions ...-00.txt and ...-01.txt.

- Changed "SHOULD drop" to "MUST drop" a packet if the deadline is passed (see Section 5).

- Added explanatory text about how packet delays might arise. (see Section 4).

- Mentioned availability of time-synchronization protocols (see Section 1).

- Updated bibliographic citations.

- Alphabetized contributor names.

- Added this section.

Authors' Addresses

Lijo Thomas
C-DAC
Centre for Development of Advanced Computing (C-DAC), Vellayambalam
Trivandrum 695033
India

Email: lijo@cdac.in

Satish Anamalamudi
SRM University-AP
Amaravati Campus
Amaravati, Andhra Pradesh 522 502
India

Email: satishnaidu80@gmail.com

S.V.R Anand
Indian Institute of Science
Bangalore 560012
India

Email: anand@ece.iisc.ernet.in

Malati Hegde
Indian Institute of Science
Bangalore 560012
India

Email: malati@ece.iisc.ernet.in

Charles E. Perkins
Futurewei
2330 Central Expressway
Santa Clara 95050
Unites States

Email: charliep@computer.org

6Lo Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 12, 2018

Y. Choi, Ed.
Y-G. Hong
ETRI
J-S. Youn
Dongeui Univ
D-K. Kim
KNU
J-H. Choi
Samsung Electronics Co.,
January 8, 2018

Transmission of IPv6 Packets over Near Field Communication
draft-ietf-6lo-nfc-09

Abstract

Near field communication (NFC) is a set of standards for smartphones and portable devices to establish radio communication with each other by touching them together or bringing them into proximity, usually no more than 10 cm. NFC standards cover communications protocols and data exchange formats, and are based on existing radio-frequency identification (RFID) standards including ISO/IEC 14443 and FeliCa. The standards include ISO/IEC 18092 and those defined by the NFC Forum. The NFC technology has been widely implemented and available in mobile phones, laptop computers, and many other devices. This document describes how IPv6 is transmitted over NFC using 6LowPAN techniques.

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 July 12, 2018.

Table of Contents

1. Introduction

NFC is a set of short-range wireless technologies, typically requiring a distance of 10 cm or less. NFC operates at 13.56 MHz on ISO/IEC 18000-3 air interface and at rates ranging from 106 kbit/s to 424 kbit/s. NFC always involves an initiator and a target; the initiator actively generates an RF field that can power a passive target. This enables NFC targets to take very simple form factors such as tags, stickers, key fobs, or cards that do not require batteries. NFC peer-to-peer communication is possible, provided both devices are powered. NFC builds upon RFID systems by allowing two-way communication between endpoints, where earlier systems such as contactless smart cards were one-way only. It has been used in devices such as mobile phones, running Android operating system, named with a feature called "Android Beam". In addition, it is expected for the other mobile phones, running the other operating systems (e.g., iOS, etc.) to be equipped with NFC technology in the near future.

Considering the potential for exponential growth in the number of heterogeneous air interface technologies, NFC would be widely used as one of the other air interface technologies, such as Bluetooth Low Energy (BT-LE), Wi-Fi, and so on. Each of the heterogeneous air interface technologies has its own characteristics, which cannot be covered by the other technologies, so various kinds of air interface technologies would co-exist together. Therefore, it is required for them to communicate with each other. NFC also has the strongest ability (e.g., secure communication distance of 10 cm) to prevent a third party from attacking privacy.

When the number of devices and things having different air interface technologies communicate with each other, IPv6 is an ideal internet protocols owing to its large address space. Also, NFC would be one of the endpoints using IPv6. Therefore, this document describes how IPv6 is transmitted over NFC using 6LoWPAN techniques.

[RFC4944] specifies the transmission of IPv6 over IEEE 802.15.4. The NFC link also has similar characteristics to that of IEEE 802.15.4. Many of the mechanisms defined in [RFC4944] can be applied to the transmission of IPv6 on NFC links. This document specifies the details of IPv6 transmission over NFC links.

2. Conventions and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Overview of Near Field Communication Technology

NFC technology enables simple and safe two-way interactions between electronic devices, allowing consumers to perform contactless transactions, access digital content, and connect electronic devices with a single touch. NFC complements many popular consumer level wireless technologies, by utilizing the key elements in existing standards for contactless card technology (ISO/IEC 14443 A&B and JIS-X 6319-4). NFC can be compatible with existing contactless card infrastructure and it enables a consumer to utilize one device across different systems.

Extending the capability of contactless card technology, NFC also enables devices to share information at a distance that is less than 10 cm with a maximum communication speed of 424 kbps. Users can share business cards, make transactions, access information from a smart poster or provide credentials for access control systems with a simple touch.

NFC’s bidirectional communication ability is ideal for establishing connections with other technologies by the simplicity of touch. In addition to the easy connection and quick transactions, simple data sharing is also available.

3.1. Peer-to-peer Mode of NFC

NFC-enabled devices are unique in that they can support three modes of operation: card emulation, peer-to-peer, and reader/writer. Peer-to-peer mode enables two NFC-enabled devices to communicate with each other to exchange information and share files, so that users of NFC-enabled devices can quickly share contact information and other files with a touch. Therefore, an NFC-enabled device can securely send IPv6 packets to any corresponding node on the Internet when an NFC-enabled gateway is linked to the Internet.

3.2. Protocol Stacks of NFC

IP can use the services provided by the Logical Link Control Protocol (LLCP) in the NFC stack to provide reliable, two-way transport of information between the peer devices. Figure 1 depicts the NFC P2P protocol stack with IPv6 bindings to LLCP.

For data communication in IPv6 over NFC, an IPv6 packet SHALL be passed down to LLCP of NFC and transported to an Information Field in Protocol Data Unit (I PDU) of LLCP of the NFC-enabled peer device. LLCP does not support fragmentation and reassembly. For IPv6 addressing or address configuration, LLCP SHALL provide related information, such as link layer addresses, to its upper layer. The

LLCP to IPv6 protocol binding SHALL transfer the SSAP and DSAP value to the IPv6 over NFC protocol. SSAP stands for Source Service Access Point, which is a 6-bit value meaning a kind of Logical Link Control (LLC) address, while DSAP means an LLC address of the destination NFC-enabled device.

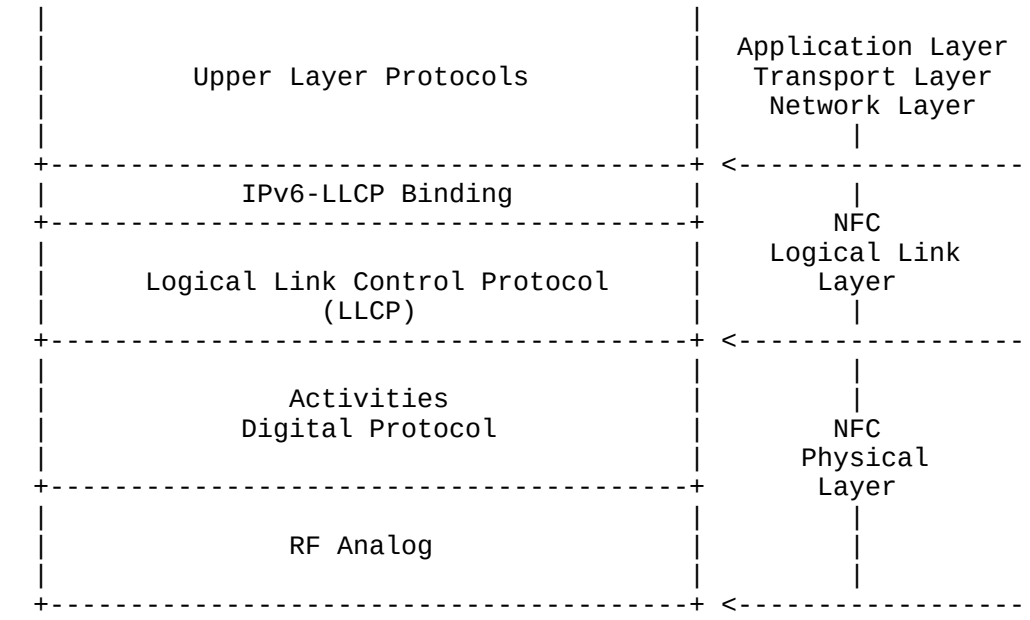


Figure 1: Protocol Stacks of NFC

The LLCP consists of Logical Link Control (LLC) and MAC Mapping. The MAC Mapping integrates an existing RF protocol into the LLCP architecture. The LLC contains three components, such as Link Management, Connection-oriented Transport, and Connection-less Transport. The Link Management component is responsible for serializing all connection-oriented and connection-less LLC PDU (Protocol Data Unit) exchanges and for aggregation and disaggregation of small PDUs. This component also guarantees asynchronous balanced mode communication and provides link status supervision by performing the symmetry procedure. The Connection-oriented Transport component is responsible for maintaining all connection-oriented data exchanges including connection set-up and termination. The Connectionless Transport component is responsible for handling unacknowledged data exchanges.

3.3. NFC-enabled Device Addressing

According to NFC Logical Link Control Protocol v1.3 [LLCP-1.3], NFC-enabled devices have two types of 6-bit addresses (i.e., SSAP and DSAP) to identify service access points. The several service access points can be installed on a NFC device. However, the SSAP and DSAP can be used as identifiers for NFC link connections with the IPv6 over NFC adaptation layer. Therefore, the SSAP can be used to generate an IPv6 interface identifier. Address values between 00h and 0Fh of SSAP and DSAP are reserved for identifying the well-known service access points, which are defined in the NFC Forum Assigned Numbers Register. Address values between 10h and 1Fh SHALL be assigned by the local LLC to services registered by local service environment. In addition, address values between 20h and 3Fh SHALL be assigned by the local LLC as a result of an upper layer service request. Therefore, the address values between 20h and 3Fh can be used for generating IPv6 interface identifiers.

3.4. MTU of NFC Link Layer

As mentioned in Section 3.2, an IPv6 packet SHALL passed down to LLCP of NFC and transported to an Unnumbered Information Protocol Data Unit (UI PDU) and an Information Field in Protocol Data Unit (I PDU) of LLCP of the NFC-enabled peer device.

The information field of an I PDU SHALL contain a single service data unit. The maximum number of octets in the information field is determined by the Maximum Information Unit (MIU) for the data link connection. The default value of the MIU for I PDUs SHALL be 128 octets. The local and remote LLCs each establish and maintain distinct MIU values for each data link connection endpoint. Also, an LLC MAY announce a larger MIU for a data link connection by transmitting an MIUX extension parameter within the information field. If no MIUX parameter is transmitted, the default MIU value of 128 SHALL be used. Otherwise, the MTU size in NFC LLCP SHALL calculate the MIU value as follows:

$$MIU = 128 + MIUX.$$

When the MIUX parameter is encoded as a TLV, the TLV Type field SHALL be 0x02 and the TLV Length field SHALL be 0x02. The MIUX parameter SHALL be encoded into the least significant 11 bits of the TLV Value field. The unused bits in the TLV Value field SHALL be set to zero by the sender and SHALL be ignored by the receiver. However, a maximum value of the TLV Value field can be 0x7FF, and a maximum size of the MTU in NFC LLCP is 2176 bytes.

4. Specification of IPv6 over NFC

NFC technology also has considerations and requirements owing to low power consumption and allowed protocol overhead. 6LoWPAN standards [RFC4944], [RFC6775], and [RFC6282] provide useful functionality for reducing overhead which can be applied to NFC. This functionality consists of link-local IPv6 addresses and stateless IPv6 address auto-configuration (see Section 4.3), Neighbor Discovery (see Section 4.5) and header compression (see Section 4.7).

4.1. Protocol Stacks

Figure 2 illustrates IPv6 over NFC. Upper layer protocols can be transport layer protocols (TCP and UDP), application layer protocols, and others capable running on top of IPv6.

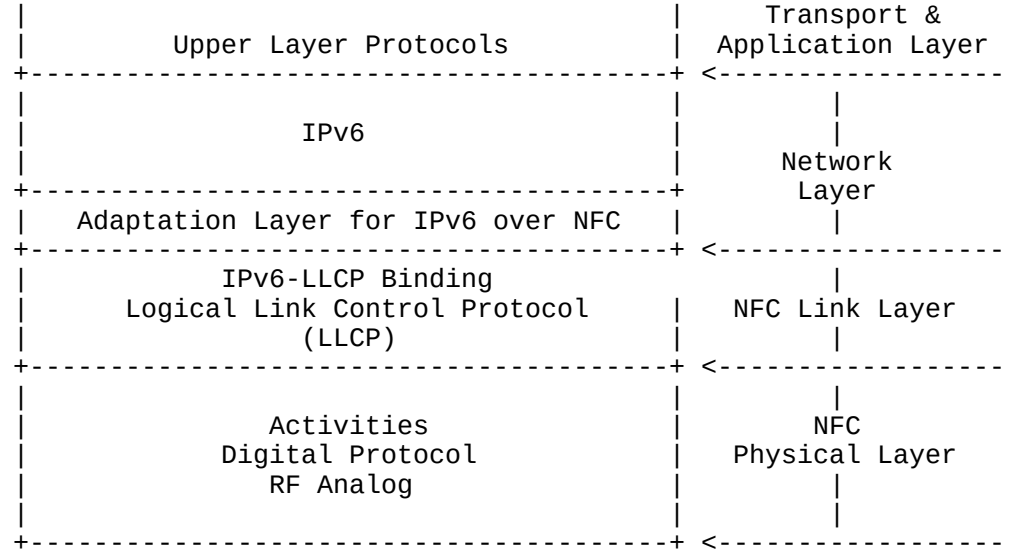


Figure 2: Protocol Stacks for IPv6 over NFC

The adaptation layer for IPv6 over NFC SHALL support neighbor discovery, stateless address auto-configuration, header compression, and fragmentation & reassembly.

4.2. Link Model

In the case of BT-LE, the Logical Link Control and Adaptation Protocol (L2CAP) supports fragmentation and reassembly (FAR) functionality; therefore, the adaptation layer for IPv6 over BT-LE does not have to conduct the FAR procedure. The NFC LLCP, in

contrast, does not support the FAR functionality, so IPv6 over NFC needs to consider the FAR functionality, defined in [RFC4944]. However, the MTU on an NFC link can be configured in a connection procedure and extended enough to fit the MTU of IPv6 packet (see Section 4.8).

The NFC link between two communicating devices is considered to be a point-to-point link only. Unlike in BT-LE, an NFC link does not support a star topology or mesh network topology but only direct connections between two devices. Furthermore, the NFC link layer does not support packet forwarding in link layer. Due to this characteristics, 6LoWPAN functionalities, such as addressing and auto-configuration, and header compression, need to be specialized into IPv6 over NFC.

4.3. Stateless Address Autoconfiguration

An NFC-enabled device (i.e., 6LN) performs stateless address autoconfiguration as per [RFC4862]. A 64-bit Interface identifier (IID) for an NFC interface is formed by utilizing the 6-bit NFC LLCP address (see Section 3.3). In the viewpoint of address configuration, such an IID SHOULD guarantee a stable IPv6 address because each data link connection is uniquely identified by the pair of DSAP and SSAP included in the header of each LLC PDU in NFC.

Following the guidance of [RFC7136], interface identifiers of all unicast addresses for NFC-enabled devices are 64 bits long and constructed by using the generation algorithm of random (but stable) identifier (RID) [RFC7217] (see Figure 3).

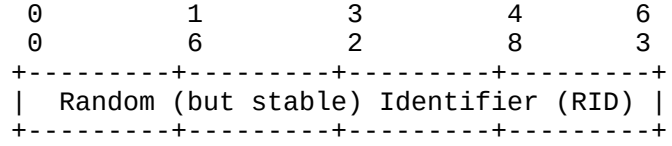


Figure 3: IID from NFC-enabled device

The RID is an output which MAY be created by the algorithm, F() with input parameters. One of the parameters is Net_IFace, and NFC Link Layer address (i.e., SSAP) MAY be a source of the NetIFace parameter. The 6-bit address of SSAP of NFC is easy and short to be targeted by attacks of third party (e.g., address scanning). The F() can provide secured and stable IIDs for NFC-enabled devices.

In addition, the "Universal/Local" bit (i.e., the ’u’ bit) of an NFC-enabled device address MUST be set to 0 [RFC4291].

4.4. IPv6 Link Local Address

Only if the NFC-enabled device address is known to be a public address, the "Universal/Local" bit be set to 1. The IPv6 link-local address for an NFC-enabled device is formed by appending the IID, to the prefix FE80::/64, as depicted in Figure 4.

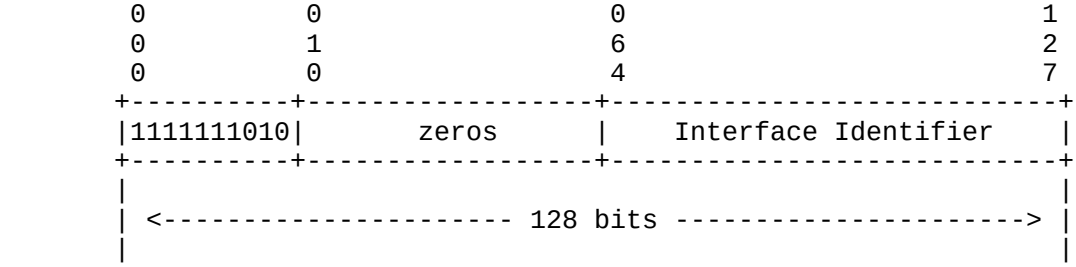


Figure 4: IPv6 link-local address in NFC

The tool for a 6LBR to obtain an IPv6 prefix for numbering the NFC network is can be accomplished via DHCPv6 Prefix Delegation ([RFC3633]).

4.5. Neighbor Discovery

Neighbor Discovery Optimization for 6LoWPANs ([RFC6775]) describes the neighbor discovery approach in several 6LoWPAN topologies, such as mesh topology. NFC does not support a complicated mesh topology but only a simple multi-hop network topology or directly connected peer-to-peer network. Therefore, the following aspects of RFC 6775 are applicable to NFC:

- When an NFC-enabled device (6LN) is directly connected to a 6LBR, an NFC 6LN MUST register its address with the 6LBR by sending a Neighbor Solicitation (NS) message with the Address Registration Option (ARO) and process the Neighbor Advertisement (NA) accordingly. In addition, if DHCPv6 is used to assign an address, Duplicate Address Detection (DAD) MAY not be required.

- When two or more NFC 6LNs meet, there MAY be two cases. One is that they meet with multi-hop connections, and the other is that they meet within a sigle hop range (e.g., isolated network). In a case of multi-hops, all of 6LNs, which have two or more connections with different neighbors, MAY be a router for 6LR/6LBR. In a case that they meet within a single hop and they have the same properties, any of them can be a router. Unless they are the same (e.g., different MTU, level of remaining energy, connectivity, etc.), a performance-outstanding device can become a

router. Also, they MAY deliver their own information (e.g., MTU and energy level, etc.) to neighbors with NFC LLCP protocols during connection initialization.

- For sending Router Solicitations and processing Router Advertisements, the NFC 6LNs MUST follow Sections 5.3 and 5.4 of RFC 6775.

4.6. Dispatch Header

All IPv6-over-NFC encapsulated datagrams are prefixed by an encapsulation header stack consisting of a Dispatch value followed by zero or more header fields. The only sequence currently defined for IPv6-over-NFC is the LOWPAN_IPHC header followed by payload, as depicted in Figure 5.

Figure 5: A IPv6-over-NFC Encapsulated 6LOWPAN_IPHC Compressed IPv6 Datagram

The dispatch value may be treated as an unstructured namespace. Only a single pattern is used to represent current IPv6-over-NFC functionality.

Pattern	Header Type	Reference
01 1xxxxx	6LOWPAN_IPHC	[RFC6282]

Figure 6: Dispatch Values

Other IANA-assigned 6LoWPAN Dispatch values do not apply to this specification.

4.7. Header Compression

Header compression as defined in [RFC6282], which specifies the compression format for IPv6 datagrams on top of IEEE 802.15.4, is REQUIRED in this document as the basis for IPv6 header compression on top of NFC. All headers MUST be compressed according to RFC 6282 encoding formats.

Therefore, IPv6 header compression in [RFC6282] MUST be implemented. Further, implementations MAY also support Generic Header Compression (GHC) of [RFC7400].

If a 16-bit address is required as a short address, it MUST be formed by padding the 6-bit NFC link-layer (node) address to the left with zeros as shown in Figure 7.

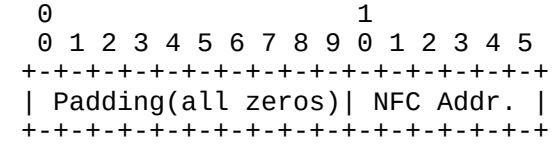


Figure 7: NFC short address format

4.8. Fragmentation and Reassembly

NFC provides fragmentation and reassembly (FAR) for payloads from 128 bytes up to 2176 bytes as mentioned in Section 3.4. The MTU of a general IPv6 packet can fit into a single NFC link frame. Therefore, the FAR functionality as defined in RFC 4944, which specifies the fragmentation methods for IPv6 datagrams on top of IEEE 802.15.4, MAY NOT be required as the basis for IPv6 datagram FAR on top of NFC. The NFC link connection for IPv6 over NFC MUST be configured with an equivalent MIU size to fit the MTU of IPv6 Packet. If NFC devices support extension of the MTU, the MIUX value is 0x480 in order to fit the MTU (1280 bytes) of a IPv6 packet.

4.9. Unicast Address Mapping

The address resolution procedure for mapping IPv6 non-multicast addresses into NFC link-layer addresses follows the general description in Section 7.2 of [RFC4861], unless otherwise specified.

The Source/Target link-layer Address option has the following form when the addresses are 6-bit NFC link-layer (node) addresses.

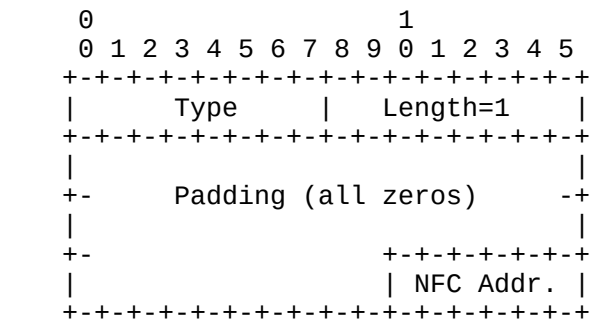


Figure 8: Unicast address mapping

Option fields:

Type:

1: for Source Link-layer address.

2: for Target Link-layer address.

Length:

This is the length of this option (including the type and length fields) in units of 8 octets. The value of this field is 1 for 6-bit NFC node addresses.

NFC address:

The 6-bit address in canonical bit order. This is the unicast address the interface currently responds to.

4.10. Multicast Address Mapping

All IPv6 multicast packets MUST be sent to NFC Destination Address, 0x3F (broadcast) and be filtered at the IPv6 layer. When represented as a 16-bit address in a compressed header, it MUST be formed by padding on the left with a zero. In addition, the NFC Destination Address, 0x3F, MUST NOT be used as a unicast NFC address of SSAP or DSAP.

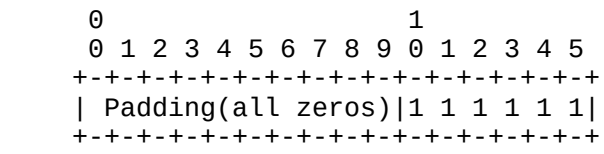


Figure 9: Multicast address mapping

5. Internet Connectivity Scenarios

As two typical scenarios, the NFC network can be isolated and connected to the Internet.

5.1. NFC-enabled Device Connected to the Internet

One of the key applications of using IPv6 over NFC is securely transmitting IPv6 packets because the RF distance between 6LN and 6LBR is typically within 10 cm. If any third party wants to hack into the RF between them, it must come to nearly touch them. Applications can choose which kinds of air interfaces (e.g., BT-LE, Wi-Fi, NFC, etc.) to send data depending on the characteristics of the data.

Figure 10 illustrates an example of an NFC-enabled device network connected to the Internet. The distance between 6LN and 6LBR is typically 10 cm or less. If there is any laptop computers close to a user, it will become the a 6LBR. Additionally, when the user mounts an NFC-enabled air interface adapter (e.g., portable NFC dongle) on the close laptop PC, the user's NFC-enabled device (6LN) can communicate with the laptop PC (6LBR) within 10 cm distance.

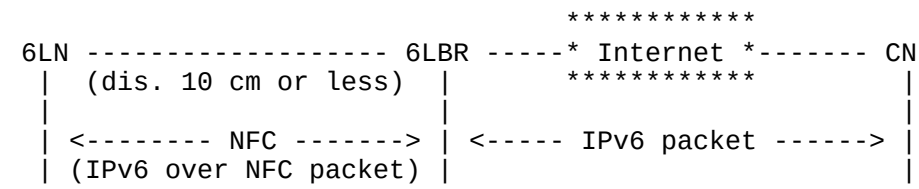


Figure 10: NFC-enabled device network connected to the Internet

5.2. Isolated NFC-enabled Device Network

In some scenarios, the NFC-enabled device network may transiently be a simple isolated network as shown in the Figure 11.

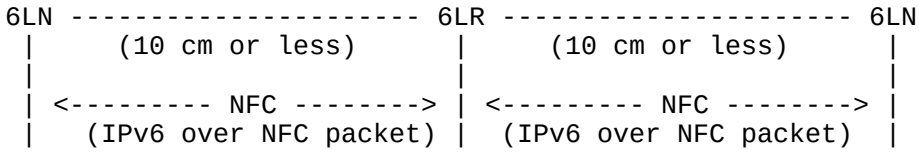


Figure 11: Isolated NFC-enabled device network

In mobile phone markets, applications are designed and made by user developers. They may image interesting applications, where three or more mobile phones touch or attach each other to accomplish outstanding performance.

6. IANA Considerations

There are no IANA considerations related to this document.

7. Security Considerations

When interface identifiers (IIDs) are generated, devices and users are required to consider mitigating various threats, such as correlation of activities over time, location tracking, device-specific vulnerability exploitation, and address scanning.

IPv6-over-NFC is, in practice, not used for long-lived links for big size data transfer or multimedia streaming, but used for extremely short-lived links (i.e., single touch-based approaches) for ID verification and mobile payment. This will mitigate the threat of correlation of activities over time.

IPv6-over-NFC uses an IPv6 interface identifier formed from a "Short Address" and a set of well-known constant bits (such as padding with ’0’s) for the modified EUI-64 format. However, the short address of NFC link layer (LLC) is not generated as a physically permanent value but logically generated for each connection. Thus, every single touch connection can use a different short address of NFC link with an extremely short-lived link. This can mitigate address scanning as well as location tracking and device-specific vulnerability exploitation.

8. Acknowledgements

We are grateful to the members of the IETF 6lo working group.

Michael Richardson, Suresh Krishnan, Pascal Thubert, Carsten Bormann, Alexandru Petrescu, James Woodyatt, Dave Thaler, Samita Chakrabarti, and Gabriel Montenegro have provided valuable feedback for this draft.

9. References

9.1. Normative References

[LLCP-1.3] "NFC Logical Link Control Protocol version 1.3", NFC Forum Technical Specification , March 2016.

[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>.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, DOI 10.17487/RFC3633, December 2003, <https://www.rfc-editor.org/info/rfc3633>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <https://www.rfc-editor.org/info/rfc4944>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <https://www.rfc-editor.org/info/rfc6775>.

[RFC7136] Carpenter, B. and S. Jiang, "Significance of IPv6 Interface Identifiers", RFC 7136, DOI 10.17487/RFC7136, February 2014, <https://www.rfc-editor.org/info/rfc7136>.

[RFC7217] Gont, F., "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)", RFC 7217, DOI 10.17487/RFC7217, April 2014, <https://www.rfc-editor.org/info/rfc7217>.

[RFC7400] Bormann, C., "6LoWPAN-GHC: Generic Header Compression for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 7400, DOI 10.17487/RFC7400, November 2014, <https://www.rfc-editor.org/info/rfc7400>.

9.2. Informative References

[ECMA-340] "Near Field Communication - Interface and Protocol (NFCIP-1) 3rd Ed.", ECMA-340 , June 2013.

Authors' Addresses

Younghwan Choi (editor)
Electronics and Telecommunications Research Institute
218 Gajeongno, Yuseung-gu
Daejeon 34129
Korea

Phone: +82 42 860 1429
Email: yhc@etri.re.kr

Yong-Geun Hong
Electronics and Telecommunications Research Institute
161 Gajeong-Dong Yuseung-gu
Daejeon 305-700
Korea

Phone: +82 42 860 6557
Email: yghong@etri.re.kr

Joo-Sang Youn
DONG-EUI University
176 Eomgwangno Busan_jin_gu
Busan 614-714
Korea

Phone: +82 51 890 1993
Email: joosang.youn@gmail.com

Dongkyun Kim
Kyungpook National University
80 Daehak-ro, Buk-gu
Daegu 702-701
Korea

Phone: +82 53 950 7571
Email: dongkyun@knu.ac.kr

JinHyouk Choi
Samsung Electronics Co.,
129 Samsung-ro, Youngdong-gu
Suwon 447-712
Korea

Phone: +82 2 2254 0114
Email: jinchoe@samsung.com

6Lo Working Group
Internet-Draft
Intended status: Informational
Expires: September 6, 2018

Y-G. Hong
ETRI
C. Gomez
UPC
Y-H. Choi
ETRI
D-Y. Ko
SKtelecom
AR. Sangi
Huaiyin Institute of Technology
T. Aanstoot
Modio AB
S. Chakrabarti
March 5, 2018

IPv6 over Constrained Node Networks (6lo) Applicability & Use cases
draft-ietf-6lo-use-cases-04

Abstract

This document describes the applicability of IPv6 over constrained node networks (6lo) and provides practical deployment examples. In addition to IEEE 802.15.4, various link layer technologies such as ITU-T G.9959 (Z-Wave), BLE, DECT-ULE, MS/TP, NFC, PLC (IEEE 1901.2), and IEEE 802.15.4e (6tisch) are used as examples. The document targets an audience who like to understand and evaluate running end-to-end IPv6 over the constrained node networks connecting devices to each other or to other devices on the Internet (e.g. cloud infrastructure).

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 September 6, 2018.

Table of Contents

1. Introduction

Running IPv6 on constrained node networks has different features from general node networks due to the characteristics of constrained node networks such as small packet size, short link-layer address, low bandwidth, network topology, low power, low cost, and large number of devices [RFC4919][RFC7228]. For example, some IEEE 802.15.4 link layers have a frame size of 127 octets and IPv6 requires the layer below to support an MTU of 1280 bytes, therefore an appropriate fragmentation and reassembly adaptation layer must be provided at the layer below IPv6. Also, the limited size of IEEE 802.15.4 frame and low energy consumption requirements make the need for header compression. The IETF 6LoPWAN (IPv6 over Low powerWPAN) working group published an adaptation layer for sending IPv6 packets over IEEE 802.15.4 [RFC4944], which includes a compression format for IPv6 datagrams over IEEE 802.15.4-based networks [RFC6282], and Neighbor Discovery Optimization for 6LoPWAN [RFC6775].

As IoT (Internet of Things) services become more popular, IPv6 over various link layer technologies such as Bluetooth Low Energy (Bluetooth LE), ITU-T G.9959 (Z-Wave), Digital Enhanced Cordless Telecommunications - Ultra Low Energy (DECT-ULE), Master-Slave/Token Passing (MS/TP), Near Field Communication (NFC), Power Line Communication (PLC), and IEEE 802.15.4e (TSCH), have been defined at [IETF_6lo] working group. IPv6 stacks for constrained node networks use a variation of the 6LoWPAN stack applied to each particular link layer technology.

In the 6LoPWAN working group, the [RFC6568], "Design and Application Spaces for 6LoWPANs" was published and it describes potential application scenarios and use cases for low-power wireless personal area networks. Hence, this 6lo applicability document aims to provide guidance to an audience who are new to IPv6-over-low-power networks concept and want to assess if variance of 6LoWPAN stack [6lo] can be applied to the constrained layer two (L2) network of their interest. This 6lo applicability document puts together various design space dimensions such as deployment, network size, power source, connectivity, multi-hop communication, traffic pattern, security level, mobility, and QoS requirements etc. In addition, it describes a few set of 6LoPWAN application scenarios and practical deployment as examples.

This document provides the applicability and use cases of 6lo, considering the following aspects:

- o 6lo applicability and use cases MAY be uniquely different from those of 6LoWPAN defined for IEEE 802.15.4.
- o It SHOULD cover various IoT related wire/wireless link layer technologies providing practical information of such technologies.
- o A general guideline on how the 6LoWPAN stack can be modified for a given L2 technology.
- o Example use cases and practical deployment examples.

2. Conventions and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. 6lo Link layer technologies and possible candidates

3.1. ITU-T G.9959 (specified)

The ITU-T G.9959 Recommendation [G.9959] targets low-power Personal Area Networks (PANs), and defines physical layer and link layer functionality. Physical layers of 9.6 kbit/s, 40 kbit/s and 100 kbit/s are supported. G.9959 defines how a unique 32-bit HomeID network identifier is assigned by a network controller and how an 8-bit NodeID host identifier is allocated to each node. NodeIDs are unique within the network identified by the HomeID. The G.9959 HomeID represents an IPv6 subnet that is identified by one or more IPv6 prefixes [RFC7428]. The ITU-T G.9959 can be used for smart home applications.

3.2. Bluetooth LE (specified)

Bluetooth LE was introduced in Bluetooth 4.0, enhanced in Bluetooth 4.1, and developed even further in successive versions. Bluetooth SIG has also published Internet Protocol Support Profile (IPSP). The IPSP enables discovery of IP-enabled devices and establishment of link-layer connection for transporting IPv6 packets. IPv6 over Bluetooth LE is dependent on both Bluetooth 4.1 and IPSP 1.0 or newer.

Many Devices such as mobile phones, notebooks, tablets and other handheld computing devices which support Bluetooth 4.0 or subsequent chipsets also support the low-energy variant of Bluetooth. Bluetooth

LE is also being included in many different types of accessories that collaborate with mobile devices such as phones, tablets and notebook computers. An example of a use case for a Bluetooth LE accessory is a heart rate monitor that sends data via the mobile phone to a server on the Internet [RFC7668]. A typical usage of Bluetooth LE is smartphone-based interaction with constrained devices. Bluetooth LE was originally designed to enable star topology networks. However, recent Bluetooth versions support the formation of extended topologies, and IPv6 support for mesh networks of Bluetooth LE devices is being developed [I-D.ietf-6lo-blemesh]

3.3. DECT-ULE (specified)

DECT ULE is a low power air interface technology that is designed to support both circuit switched services, such as voice communication, and packet mode data services at modest data rate.

The DECT ULE protocol stack consists of the PHY layer operating at frequencies in the 1880 - 1920 MHz frequency band depending on the region and uses a symbol rate of 1.152 Mbps. Radio bearers are allocated by use of FDMA/TDMA/TDD techniques.

In its generic network topology, DECT is defined as a cellular network technology. However, the most common configuration is a star network with a single Fixed Part (FP) defining the network with a number of Portable Parts (PP) attached. The MAC layer supports traditional DECT as this is used for services like discovery, pairing, security features etc. All these features have been reused from DECT.

The DECT ULE device can switch to the ULE mode of operation, utilizing the new ULE MAC layer features. The DECT ULE Data Link Control (DLC) provides multiplexing as well as segmentation and re-assembly for larger packets from layers above. The DECT ULE layer also implements per-message authentication and encryption. The DLC layer ensures packet integrity and preserves packet order, but delivery is based on best effort.

The current DECT ULE MAC layer standard supports low bandwidth data broadcast. However the usage of this broadcast service has not yet been standardized for higher layers [RFC8105]. DECT-ULE can be used for smart metering in a home.

3.4. MS/TP (specified)

Master-Slave/Token-Passing (MS/TP) is a Medium Access Control (MAC) protocol for the RS-485 [TIA-485-A] physical layer and is used primarily in building automation networks.

An MS/TP device is typically based on a low-cost microcontroller with limited processing power and memory. These constraints, together with low data rates and a small MAC address space, are similar to those faced in 6LoWPAN networks. MS/TP differs significantly from 6LoWPAN in at least three respects: a) MS/TP devices are typically mains powered, b) all MS/TP devices on a segment can communicate directly so there are no hidden node or mesh routing issues, and c) the latest MS/TP specification provides support for large payloads, eliminating the need for fragmentation and reassembly below IPv6.

MS/TP is designed to enable multidrop networks over shielded twisted pair wiring. It can support network segments up to 1000 meters in length at a data rate of 115.2 kbit/s or segments up to 1200 meters in length at lower bit rates. An MS/TP interface requires only a UART, an RS-485 [TIA-485-A] transceiver with a driver that can be disabled, and a 5 ms resolution timer. The MS/TP MAC is typically implemented in software.

Because of its superior "range" (˜1 km) compared to many low power wireless data links, MS/TP may be suitable to connect remote devices (such as district heating controllers) to the nearest building control infrastructure over a single link [RFC8163]. MS/TP can be used for building automation networks.

3.5. NFC (specified)

NFC technology enables simple and safe two-way interactions between electronic devices, allowing consumers to perform contactless transactions, access digital content, and connect electronic devices with a single touch. NFC complements many popular consumer level wireless technologies, by utilizing the key elements in existing standards for contactless card technology (ISO/IEC 14443 A&B and JIS-X 6319-4). NFC can be compatible with existing contactless card infrastructure and it enables a consumer to utilize one device across different systems.

Extending the capability of contactless card technology, NFC also enables devices to share information at a distance that is less than 10 cm with a maximum communication speed of 424 kbps. Users can share business cards, make transactions, access information from a smart poster or provide credentials for access control systems with a simple touch.

NFC’s bidirectional communication ability is ideal for establishing connections with other technologies by the simplicity of touch. In addition to the easy connection and quick transactions, simple data sharing is also available [I-D.ietf-6lo-nfc]. NFC can be used for secure transfer in healthcare services.

3.6. PLC (specified)

PLC is a data transmission technique that utilizes power conductors as medium. Unlike other dedicated communication infrastructure, power conductors are widely available indoors and outdoors. Moreover, wired technologies are more susceptible to cause interference but are more reliable than their wireless counterparts. PLC is a data transmission technique that utilizes power conductors as medium.

The below table shows some available open standards defining PLC.

PLC Systems	Frequency Range	Type	Data Rate	Distance
IEEE1901	<100MHz	Broadband	200Mbps	1000m
IEEE1901.1	<15MHz	PLC-IoT	10Mbps	2000m
IEEE1901.2	<500kHz	Narrowband	200Kbps	3000m

Table 1: Some Available Open Standards in PLC

[IEEE1901] defines a broadband variant of PLC but is effective within short range. This standard addresses the requirements of applications with high data rate such as: Internet, HDTV, Audio, Gaming etc. Broadband operates on OFDM (Orthogonal Frequency Division Multiplexing) modulation.

[IEEE1901.2] defines a narrowband variant of PLC with less data rate but significantly higher transmission range that could be used in an indoor or even an outdoor environment. It is applicable to typical IoT applications such as: Building Automation, Renewable Energy, Advanced Metering, Street Lighting, Electric Vehicle, Smart Grid etc. Moreover, IEEE 1901.2 standard is based on the 802.15.4 MAC sub-layer and fully endorses the security scheme defined in 802.15.4 [RFC8036]. A typical use case of PLC is smart grid.

3.7. IEEE 802.15.4e (specified)

The Time Slotted Channel Hopping (TSCH) mode was introduced in the IEEE 802.15.4-2015 standard. In a TSCH network, all nodes are synchronized. Time is sliced up into timeslots. The duration of a timeslot, typically 10ms, is large enough for a node to send a full-sized frame to its neighbor, and for that neighbor to send back an acknowledgment to indicate successful reception. Timeslots are grouped into one of more slotframes, which repeat over time.

All the communication in the network is orchestrated by a communication schedule which indicates to each node what to do in each of the timeslots of a slotframe: transmit, listen or sleep. The communication schedule can be built so that the right amount of link-layer resources (the cells in the schedule) are scheduled to satisfy the communication needs of the applications running on the network, while keeping the energy consumption of the nodes very low. Cells can be scheduled in a collision-free way, introducing a high level of determinism to the network.

A TSCH network exploits channel hopping: subsequent packet exchanges between neighbor nodes are done on a different frequency. This means that, if a frame isn't received, the transmitter node will re-transmitt the frame on a different frequency. The resulting "channel hopping" efficiently combats external interference and multi-path fading.

The main benefits of IEEE 802.15.4 TSCH are:

- ultra high reliability. Off-the-shelf commercial products offer over 99.999% end-to-end reliability.

- ultra low-power consumption. Off-the-shelf commercial products offer over a decade of battery lifetime.

- 6TiSCH at IETF defines communications of TSCH network and it uses 6LoWPAN stack [RFC7554].

IEEE 802.15.4e can be used for industrial automation.

3.8. LTE MTC (example of a potential candidate)

LTE category defines the overall performance and capabilities of the UE (User Equipment). For example, the maximum down rate of category 1 UE and category 2 UE are 10.3 Mbit/s and 51.0 Mbit/s respectively. There are many categories in LTE standards. 3GPP standards defined the category 0 to be used for low rate IoT service in release 12. Since category 1 and category 0 could be used for low rate IoT service, these categories are called LTE MTC (Machine Type Communication) [LTE_MTC]. And 3GPP standards defined the MTC Enhancements in release 13.

LTE MTC offer advantages in comparison to above category 2 and is appropriate to be used for low rate IoT services such as low power and low cost.

LTE MTC can be used for tracking services, such as asset tracker, bicycle/cat tracker and etc with national wide. LTE MTC can be also

used for monitoring & control service, such as car mobility record and weather observation that require much more traffic than other IoT services. Since the traffic collected by other IoT devices such as LoRa, Z-wave and BLE is small, LTE MTC can be used as a bachhaul of other IoT networks.

3.9. Comparison between 6lo Link layer technologies

In above clauses, various 6lo Link layer technologies and a possible candidate are described. The following table shows that dominant paramters of each use case corresponding to the 6lo link layer technology.

	Z-Wave	BLE	DECT-ULE	MS/TP	NFC	PLC	TSCH
Usage	Home Auto-mation	Interact w/ Smart Phone	Meter Reading	Building Auto-mation	Health-care Service	Smart Grid	Industr-ial Aut-mation
Topology & Subnet	L2-mesh or L3-mesh	Star & Mesh	Star No mesh	MS/TP No mesh	P2P L2-mesh	Star Tree Mesh	Mesh
Mobility Reqmt	No	Low	No	No	Moderate	No	No
Security Reqmt	High + Privacy required	Parti-ally	High + Privacy required	High + Authen. required	High	High + Encrypt. required	High + Privacy required
Buffering Reqmt	Low	Low	Low	Low	Low	Low	Low
Latency, QoS Reqmt	High	Low	Low	High	High	Low	High
Data Rate	Infrequ-ent	Infrequ-ent	Infrequ-ent	Frequent	Small	Infrequ-ent	Infrequ-ent
RFC # or Draft	RFC7428	RFC7668	RFC8105	RFC8163	draft-ietf-6lo-nfc	draft-hou-6lo-plc	RFC7554

Table 2: Comparison between 6lo Link layer technologies

4. 6lo Deployment Scenarios

4.1. jupitermesh in Smart Grid using 6lo in network layer

jupiterMesh is a multi-hop wireless mesh network specification designed mainly for deployment in large geographical areas. Each subnet in jupiterMesh is able to cover an entire neighborhood with thousands of nodes consisting of IPv6-enabled routers and end-points

(e.g. hosts). Automated network joining and load balancing allows a seamless deployment of a large number of subnets.

The main application domains targeted by jupiterMesh are smart grid and smart cities. This includes, but is not limited to the following applications:

- Automated meter reading
- Distribution Automation (DA)
- Demand-side management (DSM)
- Demand-side response (DSR)
- Power outage reporting
- Street light monitoring and control
- Transformer load management
- EV charging coordination
- Energy theft
- Parking space locator

jupiterMesh specification is based on the following technologies:

- The PHY layer is based on IEEE 802.15.4 SUN specification [IEEE 802.15.4-2015], supporting multiple operating modes for deployment in different regulatory domains and deployment scenarios in terms of density and bandwidth requirements. jupiterMesh supports bit rates from 50 kbps to 800 kbps, frame size up to 2048 bytes, up to 11 different RF bands and 3 modulation types (i.e., FSK, OQPSK and OFDM).
- The MAC layer is based on IEEE 802.15.4 TSCH specification [IEEE 802.15.4-2015]. With frequency hopping capability, TSCH MAC supports scheduling of dedicated timeslot enabling bandwidth management and QoS.
- The security layer consists of a certificate-based (i.e. X.509) network access authentication using EAP-TLS, with IEEE 802.15.9-based KMP (Key Management Protocol) transport, and PANA and link layer encryption using AES-128 CCM as specified in IEEE 802.15.4-2015 [IEEE 802.15.4-2015].

- Address assignment and network configuration are specified using DHCPv6 [RFC3315]. Neighbor Discovery (ND) [RFC6775] and stateless address auto-configuration (SLAAC) are not supported.

- The network layer consists of IPv6, ICMPv6 and 6lo/6LoPWAN header compression [RFC6282]. Multicast is supported using MPL. Two domains are supported, a delay sensitive MPL domain for low latency applications (e.g. DSM, DSR) and a delay insensitive one for less stringent applications (e.g. OTA file transfers).

- The routing layer uses RPL [RFC6550] in non-storing mode with the MRHOF objective function based on the ETX metric.

4.2. Wi-SUN usage of 6lo stacks

Wireless Smart Ubiquitous Network (Wi-SUN) is a technology based on the IEEE 802.15.4g standard. Wi-SUN networks support star and mesh topologies, as well as hybrid star/mesh deployments, but are typically laid out in a mesh topology where each node relays data for the network to provide network connectivity. Wi-SUN networks are deployed on both powered and battery-operated devices.

The main application domains targeted by Wi-SUN are smart utility and smart city networks. This includes, but is not limited to the following applications:

- Advanced Metering Infrastructure (AMI)

- Distribution Automation

- Home Energy Management

- Infrastructure Management

- Intelligent Transportation Systems

- Smart Street Lighting

- Agriculture

- Structural health (bridges, buildings etc)

- Monitoring and Asset Management

- Smart Thermostats, Air Conditioning and Heat Controls

- Energy Usage Information Displays

The Wi-SUN Alliance Field Area Network (FAN) covers primarily outdoor networks, and its specification is oriented towards meeting the more rigorous challenges of these environments. Examples include from meter to outdoor access point/router for AMI and DR, or between switches for DA. However, nothing in the profile restricts it to outdoor use. It has the following features;

- Open standards based on IEEE802, IETF, TIA, ETSI
- Architecture is an IPv6 frequency hopping wireless mesh network with enterprise level security
- Simple infrastructure which is low cost, low complexity
- Enhanced network robustness, reliability, and resilience to interference, due to high redundancy and frequency hopping
- Enhaced scalability, long range, and energy friendliness
- Supports multiple global license-exempt sub GHz bands
- Multi-vendor interoperability
- Very low power modes in development permitting long term battery operation of network nodes

In the Wi-SUN FAN specification, adaptation layer based on 6lo and IPv6 network layer are described. So, IPv6 protocol suite including TCP/UDP, 6lo Adaptation, Header Compression, DHCPv6 for IP address management, Routing using RPL, ICMPv6, and Unicast/Multicast forwarding is utilized.

4.3. G3-PLC usage of 6lo in network layer

G3-PLC [G3-PLC] is a narrow-band PLC technology that is based on ITU-T G.9903 Recommendation [G.9903]. G3-PLC supports multi-hop mesh network, and facilitates highly-reliable, long-range communication. With the abilities to support IPv6 and to cross transformers, G3-PLC is regarded as one of the next-generation NB-PLC technologies. G3-PLC has got massive deployments over several countries, e.g. Japan and France.

The main application domains targeted by G3-PLC are smart grid and smart cities. This includes, but is not limited to the following applications:

- Smart Metering

- Vehicle-to-Grid Communication
- Demand Response (DR)
- Distribution Automation
- Home/Building Energy Management Systems
- Smart Street Lighting
- Advanced Metering Infrastructure (AMI) backbone network
- Wind/Solar Farm Monitoring

In the G3-PLC specification, the 6lo adaptation layer utilizes the 6LoWPAN functions (e.g. header compression, fragmentation and reassembly) so as to enable IPv6 packet transmission. LOADng, which is a lightweight variant of AODV, is applied as the mesh-under routing protocol in G3-PLC networks. Address assignment and network configuration are based on the bootstrapping protocol specified in ITU-T G.9903. The network layer consists of IPv6 and ICMPv6 while the transport protocol UDP is used for data transmission.

4.4. Netricity usage of 6lo in network layer

The Netricity program in HomePlug Powerline Alliance [NETRICITY] promotes the adoption of products built on the IEEE 1901.2 Low-Frequency Narrow-Band PLC standard, which provides for urban and long distance communications and propagation through transformers of the distribution network using frequencies below 500 kHz. The technology also addresses requirements that assure communication privacy and secure networks.

The main application domains targeted by Netricity are smart grid and smart cities. This includes, but is not limited to the following applications:

- Utility grid modernization
- Distribution automation
- Meter-to-Grid connectivity
- Micro-grids
- Grid sensor communications
- Load control

- Demand response
- Net metering
- Street Lighting control
- Photovoltaic panel monitoring

Netricity system architecture is based on the PHY and MAC layers of IEEE 1901.2 PLC standard. Regarding the 6lo adaptation layer and IPv6 network layer, Netricity utilizes IPv6 protocol suite including 6lo/6LoWPAN header compression, DHCPv6 for IP address management, RPL routing protocol, ICMPv6, and unicast/multicast forwarding. Note that the layer 3 routing in Netricity uses RPL in non-storing mode with the MRHOF objective function based on the own defined Estimated Transmission Time (ETT) metric.

5. Design Space and Guidelines for 6lo Deployment

5.1. Design Space Dimensions for 6lo Deployment

The [RFC6568] lists the dimensions used to describe the design space of wireless sensor networks in the context of the 6LoWPAN working group. The design space is already limited by the unique characteristics of a LoWPAN (e.g. low power, short range, low bit rate). In [RFC6568], the following design space dimensions are described: Deployment, Network size, Power source, Connectivity, Multi-hop communication, Traffic pattern, Mobility, Quality of Service (QoS). However, in this document, the following design space dimensions are considered:

- Deployment/Bootstrapping: 6lo nodes can be connected randomly, or in an organized manner. The bootstrapping has different characteristics for each link layer technology.
- Topology: Topology of 6lo networks may inherently follow the characteristics of each link layer technology. Point-to-point, star, tree or mesh topologies can be configured, depending on the link layer technology considered.
- L2-Mesh or L3-Mesh: L2-mesh and L3-mesh may inherently follow the characteristics of each link layer technology. Some link layer technologies may support L2-mesh and some may not support.
- Multi-link subnet, single subnet: The selection of multi-link subnet and single subnet depends on connectivity and the number of 6lo nodes.

- Data rate: Typically, the link layer technologies of 6lo have low rate of data transmission. But, by adjusting the MTU, it can deliver higher upper layer data rate.

- Buffering requirements: Some 6lo use case may require more data rate than the link layer technology support. In this case, a buffering mechanism to manage the data is required.

- Security and Privacy Requirements: Some 6lo use case can involve transferring some important and personal data between 6lo nodes. In this case, high-level security support is required.

- Mobility across 6lo networks and subnets: The movement of 6lo nodes depends on the 6lo use case. If the 6lo nodes can move or moved around, a mobility management mechanism is required.

- Time synchronization requirements: The requirement of time synchronization of the upper layer service is dependent on the 6lo use case. For some 6lo use case related to health service, the measured data must be recorded with exact time and must be transferred with time synchronization.

- Reliability and QoS: Some 6lo use case requires high reliability, for example real-time service or health-related services.

- Traffic patterns: 6lo use cases may involve various traffic patterns. For example, some 6lo use case may require short data length and random transmission. Some 6lo use case may require continuous data and periodic data transmission.

- Security Bootstrapping: Without the external operations, 6lo nodes must have the security bootstrapping mechanism.

- Power use strategy: to enable certain use cases, there may be requirements on the class of energy availability and the strategy followed for using power for communication [RFC7228]. Each link layer technology defines a particular power use strategy which may be tuned [I-D.ietf-lwig-energy-efficient]. Readers are expected to be familiar with [RFC7228] terminology.

- Update firmware requirements: Most 6lo use cases will need a mechanism for updating firmware. In these cases support for over the air updates are required, probably in a broadcast mode when bandwith is low and the number of identical devices is high.

- Wired vs. Wireless: Plenty of 6lo link layer technologies are wireless, except MS/TP and PLC. The selection of wired or wireless link layer technology is mainly dependent on the

requirement of 6lo use cases and the characteristics of wired/ wireless technologies. For example, some 6lo use cases may require easy and quick deployment, whereas others may need a continuous source of power.

5.2. Guidelines for adopting IPv6 stack (6lo/6LoWPAN)

The following guideline targets new candidate constrained L2 technologies that may be considered for running modified 6LoWPAN stack on top. The modification of 6LoWPAN stack should be based on the following:

- Addressing Model: Addressing model determines whether the device is capable of forming IPv6 Link-local and global addresses and what is the best way to derive the IPv6 addresses for the constrained L2 devices. Whether the device is capable of forming IPv6 Link-local and global addresses, L2-address-derived IPv6 addresses are specified in [RFC4944], but there exist implications for privacy. For global usage, a unique IPv6 address must be derived using an assigned prefix and a unique interface ID. [RFC8065] provides such guidelines. For MAC derived IPv6 address, please refer to [RFC8163] for IPv6 address mapping examples. Broadcast and multicast support are dependent on the L2 networks. Most low-power L2 implementations map multicast to broadcast networks. So care must be taken in the design when to use broadcast and try to stick to unicast messaging whenever possible.

- MTU Considerations: The deployment SHOULD consider their need for maximum transmission unit (MTU) of a packet over the link layer and should consider if fragmentation and reassembly of packets are needed at the 6LoWPAN layer. For example, if the link layer supports fragmentation and reassembly of packets, then 6LoWPAN layer may skip supporting fragmentation/reassembly. In fact, for most efficiency, choosing a low-power link layer that can carry unfragmented application packets would be optimum for packet transmission if the deployment can afford it. Please refer to 6lo RFCs [RFC7668], [RFC8163], [RFC8105] for example guidance.

- Mesh or L3-Routing: 6LoWPAN specifications do provide mechanisms to support for mesh routing at L2. [RFC6550] defines layer three (L3) routing for low power lossy networks using directed graphs. 6LoWPAN is routing protocol agnostic and other L2 or L3 routing protocols can be run using a 6LoWPAN stack.

- Address Assignment: 6LoWPAN requires that IPv6 Neighbor Discovery for low power networks [RFC6775] be used for autoconfiguration of stateless IPv6 address assignment. Considering the energy sensitive networks [RFC6775] makes optimization from classical

IPv6 ND [RFC4861] protocol. It is the responsibility of the deployment to ensure unique global IPv6 addresses for the Internet connectivity. For local-only connectivity IPv6 ULA may be used. [RFC6775] specifies the 6LoWPAN border router(6LBR) which is responsible for prefix assignment to the 6lo/6LoWPAN network. 6LBR can be connected to the Internet or Enterprise network via its one of the interfaces. Please refer to [RFC7668] and [RFC8105] for examples of address assignment considerations. In addition, privacy considerations [RFC8065] must be consulted for applicability. In certain scenarios, the deployment may not support autoconfiguration of IPv6 addressing due to regulatory and business reasons and may choose to offer a separate address assignment service.

- o Header Compression: IPv6 header compression [RFC6282] is a vital part of IPv6 over low power communication. Examples of header compression for different link-layers specifications are found in [RFC7668], [RFC8163], [RFC8105]. A generic header compression technique is specified in [RFC7400].

- o Security and Encryption: Though 6LoWPAN basic specifications do not address security at the network layer, the assumption is that L2 security must be present. In addition, application level security is highly desirable. The working groups [ace] and [core] should be consulted for application and transport level security. 6lo working group is working on address authentication [6lo-ap-nd] and secure bootstrapping is also being discussed at IETF. However, there may be different levels of security available in a deployment through other standards such as hardware level security or certificates for initial booting process. Encryption is important if the implementation can afford it.

- o Additional processing: [RFC8066] defines guidelines for ESC dispatch octets use in the 6LoWPAN header. An implementation may take advantage of ESC header to offer a deployment specific processing of 6LoWPAN packets.

6. 6lo Use Case Examples

As IPv6 stacks for constrained node networks use a variation of the 6LoWPAN stack applied to each particular link layer technology, various 6lo use cases can be provided. In this clause, one 6lo use case example of Bluetooth LE (Smartphone-Based Interaction with Constrained Devices) is described. Other 6lo use case examples are described in Appendix.

The key feature behind the current high Bluetooth LE momentum is its support in a large majority of smartphones in the market. Bluetooth

LE can be used to allow the interaction between the smartphone and surrounding sensors or actuators. Furthermore, Bluetooth LE is also the main radio interface currently available in wearables. Since a smartphone typically has several radio interfaces that provide Internet access, such as Wi-Fi or 4G, the smartphone can act as a gateway for nearby devices such as sensors, actuators or wearables. Bluetooth LE may be used in several domains, including healthcare, sports/wellness and home automation.

Example: Use of Bluetooth LE-based Body Area Network for fitness

A person wears a smartwatch for fitness purposes. The smartwatch has several sensors (e.g. heart rate, accelerometer, gyrometer, GPS, temperature, etc.), a display, and a Bluetooth LE radio interface. The smartwatch can show fitness-related statistics on its display. However, when a paired smartphone is in the range of the smartwatch, the latter can report almost real-time measurements of its sensors to the smartphone, which can forward the data to a cloud service on the Internet. In addition, the smartwatch can receive notifications (e.g. alarm signals) from the cloud service via the smartphone. On the other hand, the smartphone may locally generate messages for the smartwatch, such as e-mail reception or calendar notifications.

The functionality supported by the smartwatch may be complemented by other devices such as other on-body sensors, wireless headsets or head-mounted displays. All such devices may connect to the smartphone creating a star topology network whereby the smartphone is the central component. Support for extended network topologies (e.g. mesh networks) is being developed as of the writing.

7. IANA Considerations

There are no IANA considerations related to this document.

8. Security Considerations

Security considerations are not directly applicable to this document. The use cases will use the security requirements described in the protocol specifications.

9. Acknowledgements

Carles Gomez has been funded in part by the Spanish Government through the Jose Castillejo CAS15/00336 grant, and through the TEC2016-79988-P grant. His contribution to this work has been carried out in part during his stay as a visiting scholar at the Computer Laboratory of the University of Cambridge.

Thomas Watteyne, Pascal Thubert, Xavier Vilajosana, Daniel Migault, and Jianqiang HOU have provided valuable feedback for this draft.

Das Subir and Michel Veillette have provided valuable information of jupiterMesh and Paul Duffy has provided valuable information of Wi-SUN for this draft. Also, Jianqiang Hou has provided valuable information of G3-PLC and Netricity for this draft. Kerry Lynn and Dave Robin have provided valuable information of MS/TP and practical use case of MS/TP for this draft.

10. References

10.1. Normative References

[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>.

[RFC4919] Kushalnagar, N., Montenegro, G., and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals", RFC 4919, DOI 10.17487/RFC4919, August 2007, <https://www.rfc-editor.org/info/rfc4919>.

[RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <https://www.rfc-editor.org/info/rfc4944>.

[RFC5826] Brandt, A., Buron, J., and G. Porcu, "Home Automation Routing Requirements in Low-Power and Lossy Networks", RFC 5826, DOI 10.17487/RFC5826, April 2010, <https://www.rfc-editor.org/info/rfc5826>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <https://www.rfc-editor.org/info/rfc6550>.

[RFC6568] Kim, E., Kaspar, D., and JP. Vasseur, "Design and Application Spaces for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6568, DOI 10.17487/RFC6568, April 2012, <https://www.rfc-editor.org/info/rfc6568>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <https://www.rfc-editor.org/info/rfc6775>.

[RFC7228] Bormann, C., Ersue, M., and A. Keranen, "Terminology for Constrained-Node Networks", RFC 7228, DOI 10.17487/RFC7228, May 2014, <https://www.rfc-editor.org/info/rfc7228>.

[RFC7400] Bormann, C., "6LoWPAN-GHC: Generic Header Compression for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 7400, DOI 10.17487/RFC7400, November 2014, <https://www.rfc-editor.org/info/rfc7400>.

[RFC7428] Brandt, A. and J. Buron, "Transmission of IPv6 Packets over ITU-T G.9959 Networks", RFC 7428, DOI 10.17487/RFC7428, February 2015, <https://www.rfc-editor.org/info/rfc7428>.

[RFC7554] Watteyne, T., Ed., Palattella, M., and L. Grieco, "Using IEEE 802.15.4e Time-Slotted Channel Hopping (TSCH) in the Internet of Things (IoT): Problem Statement", RFC 7554, DOI 10.17487/RFC7554, May 2015, <https://www.rfc-editor.org/info/rfc7554>.

[RFC7668] Nieminen, J., Savolainen, T., Isomaki, M., Patil, B., Shelby, Z., and C. Gomez, "IPv6 over BLUETOOTH(R) Low Energy", RFC 7668, DOI 10.17487/RFC7668, October 2015, <https://www.rfc-editor.org/info/rfc7668>.

[RFC8036] Cam-Winget, N., Ed., Hui, J., and D. Popa, "Applicability Statement for the Routing Protocol for Low-Power and Lossy Networks (RPL) in Advanced Metering Infrastructure (AMI) Networks", RFC 8036, DOI 10.17487/RFC8036, January 2017, <https://www.rfc-editor.org/info/rfc8036>.

[RFC8065] Thaler, D., "Privacy Considerations for IPv6 Adaptation-Layer Mechanisms", RFC 8065, DOI 10.17487/RFC8065, February 2017, <https://www.rfc-editor.org/info/rfc8065>.

[RFC8066] Chakrabarti, S., Montenegro, G., Droms, R., and J. Woodyatt, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) ESC Dispatch Code Points and Guidelines", RFC 8066, DOI 10.17487/RFC8066, February 2017, <https://www.rfc-editor.org/info/rfc8066>.

[RFC8105] Mariager, P., Petersen, J., Ed., Shelby, Z., Van de Logt, M., and D. Barthel, "Transmission of IPv6 Packets over Digital Enhanced Cordless Telecommunications (DECT) Ultra Low Energy (ULE)", RFC 8105, DOI 10.17487/RFC8105, May 2017, <https://www.rfc-editor.org/info/rfc8105>.

[RFC8163] Lynn, K., Ed., Martocci, J., Neilson, C., and S. Donaldson, "Transmission of IPv6 over Master-Slave/Token-Passing (MS/TP) Networks", RFC 8163, DOI 10.17487/RFC8163, May 2017, <https://www.rfc-editor.org/info/rfc8163>.

10.2. Informative References

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <https://www.rfc-editor.org/info/rfc3315>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[I-D.ietf-6lo-nfc] Choi, Y., Hong, Y., Youn, J., Kim, D., and J. Choi, "Transmission of IPv6 Packets over Near Field Communication", draft-ietf-6lo-nfc-09 (work in progress), January 2018.

[I-D.ietf-lwig-energy-efficient] Gomez, C., Kovatsch, M., Tian, H., and Z. Cao, "Energy-Efficient Features of Internet of Things Protocols", draft-ietf-lwig-energy-efficient-08 (work in progress), October 2017.

[I-D.ietf-roll-aodv-rpl] Anamalamudi, S., Zhang, M., Sangi, A., Perkins, C., and S. Anand, "Asymmetric AODV-P2P-RPL in Low-Power and Lossy Networks (LLNs)", draft-ietf-roll-aodv-rpl-02 (work in progress), September 2017.

[I-D.ietf-6tisch-6top-sfx]
Dujovne, D., Grieco, L., Palattella, M., and N. Accettura, "6TiSCH 6top Scheduling Function Zero / Experimental (SFX)", draft-ietf-6tisch-6top-sfx-00 (work in progress), September 2017.

[I-D.ietf-6lo-blemesh]
Gomez, C., Darroudi, S., and T. Savolainen, "IPv6 Mesh over BLUETOOTH(R) Low Energy using IPSP", draft-ietf-6lo-blemesh-02 (work in progress), September 2017.

[I-D.satish-6tisch-6top-sf1]
Anamalamudi, S., Liu, B., Zhang, M., Sangi, A., Perkins, C., and S. Anand, "Scheduling Function One (SF1): hop-by-hop Scheduling with RSVP-TE in 6tisch Networks", draft-satish-6tisch-6top-sf1-04 (work in progress), October 2017.

[I-D.hou-6lo-plc]
Hou, J., Hong, Y., and X. Tang, "Transmission of IPv6 Packets over PLC Networks", draft-hou-6lo-plc-03 (work in progress), December 2017.

[IETF_6lo]
"IETF IPv6 over Networks of Resource-constrained Nodes (6lo) working group", <https://datatracker.ietf.org/wg/6lo/charter/>.

[TIA-485-A]
"TIA, "Electrical Characteristics of Generators and Receivers for Use in Balanced Digital Multipoint Systems", TIA-485-A (Revision of TIA-485)", March 2003, <https://global.ihs.com/doc_detail.cfm?item_s_key=00032964>.

[G3-PLC] "G3-PLC Alliance", <http://www.g3-plc.com/home/>.

[NETRICITY]
"Netricity program in HomePlug Powerline Alliance", <http://groups.homeplug.org/tech/Netricity>.

[G.9959] "International Telecommunication Union, "Short range narrow-band digital radiocommunication transceivers - PHY and MAC layer specifications", ITU-T Recommendation", January 2015.

[G.9903] "International Telecommunication Union, "Narrowband orthogonal frequency division multiplexing power line communication transceivers for G3-PLC networks", ITU-T Recommendation", August 2017.

[LTE_MTC] "3GPP TS 36.306 V13.0.0, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio access capabilities (Release 13)", December 2015.

[IEEE1901] "IEEE Standard, IEEE Std. 1901-2010 - IEEE Standard for Broadband over Power Line Networks: Medium Access Control and Physical Layer Specifications", 2010, <https://standards.ieee.org/findstds/standard/1901-2010.html>.

[IEEE1901.1] "IEEE Standard (work-in-progress), IEEE-SA Standards Board", <http://sites.ieee.org/sagroups-1901-1/>.

[IEEE1901.2] "IEEE Standard, IEEE Std. 1901.2-2013 - IEEE Standard for Low-Frequency (less than 500 kHz) Narrowband Power Line Communications for Smart Grid Applications", 2013, <https://standards.ieee.org/findstds/standard/1901.2-2013.html>.

[BACnet] "ASHRAE, "BACnet-A Data Communication Protocol for Building Automation and Control Networks", ANSI/ASHRAE Standard 135-2016", January 2016, <http://www.techstreet.com/ashrae/standards/ashrae-135-2016?product_id=1918140#jumps>.

Appendix A. Other 6lo Use Case Examples

A.1. Use case of ITU-T G.9959: Smart Home

Z-Wave is one of the main technologies that may be used to enable smart home applications. Born as a proprietary technology, Z-Wave was specifically designed for this particular use case. Recently, the Z-Wave radio interface (physical and MAC layers) has been standardized as the ITU-T G.9959 specification.

Example: Use of ITU-T G.9959 for Home Automation

Variety of home devices (e.g. light dimmers/switches, plugs, thermostats, blinds/curtains and remote controls) are augmented with ITU-T G.9959 interfaces. A user may turn on/off or may control home appliances by pressing a wall switch or by pressing a button in a remote control. Scenes may be programmed, so that after a given event, the home devices adopt a specific configuration. Sensors may also periodically send measurements of several parameters (e.g. gas presence, light, temperature, humidity, etc.) which are collected at a sink device, or may generate commands for actuators (e.g. a smoke sensor may send an alarm message to a safety system).

The devices involved in the described scenario are nodes of a network that follows the mesh topology, which is suitable for path diversity to face indoor multipath propagation issues. The multihop paradigm allows end-to-end connectivity when direct range communication is not possible. Security support is required, specially for safety-related communication. When a user interaction (e.g. a button press) triggers a message that encapsulates a command, if the message is lost, the user may have to perform further interactions to achieve the desired effect (e.g. a light is turned off). A reaction to a user interaction will be perceived by the user as immediate as long as the reaction takes place within 0.5 seconds [RFC5826].

A.2. Use case of DECT-ULE: Smart Home

DECT is a technology widely used for wireless telephone communications in residential scenarios. Since DECT-ULE is a low-power variant of DECT, DECT-ULE can be used to connect constrained devices such as sensors and actuators to a Fixed Part, a device that typically acts as a base station for wireless telephones. Therefore, DECT-ULE is specially suitable for the connected home space in application areas such as home automation, smart metering, safety, healthcare, etc.

Example: Use of DECT-ULE for Smart Metering

The smart electricity meter of a home is equipped with a DECT-ULE transceiver. This device is in the coverage range of the Fixed Part of the home. The Fixed Part can act as a router connected to the Internet. This way, the smart meter can transmit electricity consumption readings through the DECT-ULE link with the Fixed Part, and the latter can forward such readings to the utility company using Wide Area Network (WAN) links. The meter can also receive queries from the utility company or from an advanced energy control system controlled by the user, which may also be connected to the Fixed Part via DECT-ULE.

A.3. Use case of MS/TP: Building Automation Networks

The primary use case for IPv6 over MS/TP (6LoBAC) is in building automation networks. [BACnet] is the open international standard protocol for building automation, and MS/TP is defined in [BACnet] Clause 9. MS/TP was designed to be a low cost multi-drop field bus to inter-connect the most numerous elements (sensors and actuators) of a building automation network to their controllers. A key aspect of 6LoBAC is that it is designed to co-exist with BACnet MS/TP on the same link, easing the ultimate transition of some BACnet networks to native end-to-end IPv6 transport protocols. New applications for 6LoBAC may be found in other domains where low cost, long distance, and low latency are required.

Example: Use of 6LoBAC in Building Automation Networks

The majority of installations for MS/TP are for "terminal" or "unitary" controllers, i.e. single zone or room controllers that may connect to HVAC or other controls such as lighting or blinds. The economics of daisy-chaining a single twisted-pair between multiple devices is often preferred over home-run Cat-5 style wiring.

A multi-zone controller might be implemented as an IP router between a traditional Ethernet link and several 6LoBAC links, fanning out to multiple terminal controllers.

The superior distance capabilities of MS/TP (˜1 km) compared to other 6lo media may suggest its use in applications to connect remote devices to the nearest building infrastructure. for example, remote pumping or measuring stations with moderate bandwidth requirements can benefit from the low cost and robust capabilities of MS/TP over other wired technologies such as DSL, and without the line-of-site restrictions or hop-by-hop latency of many low cost wireless solutions.

A.4. Use case of NFC: Alternative Secure Transfer

According to applications, various secured data can be handled and transferred. Depending on security level of the data, methods for transfer can be alternatively selected.

Example: Use of NFC for Secure Transfer in Healthcare Services with Tele-Assistance

A senior citizen who lives alone wears one to several wearable 6lo devices to measure heartbeat, pulse rate, etc. The 6lo devices are densely installed at home for movement detection. An LoWPAN Border Router (LBR) at home will send the sensed information to a connected

healthcare center. Portable base stations with LCDs may be used to check the data at home, as well. Data is gathered in both periodic and event-driven fashion. In this application, event-driven data can be very time-critical. In addition, privacy also becomes a serious issue in this case, as the sensed data is very personal.

While the senior citizen is provided audio and video healthcare services by a tele-assistance based on LTE connections, the senior citizen can alternatively use NFC connections to transfer the personal sensed data to the tele-assistance. At this moment, hidden hackers can overhear the data based on the LTE connection, but they cannot gather the personal data over the NFC connection.

A.5. Use case of PLC: Smart Grid

Smart grid concept is based on numerous operational and energy measuring sub-systems of an electric grid. It comprises of multiple administrative levels/segments to provide connectivity among these numerous components. Last mile connectivity is established over LV segment, whereas connectivity over electricity distribution takes place in HV segment.

Although other wired and wireless technologies are also used in Smart Grid (Advance Metering Infrastructure - AMI, Demand Response - DR, Home Energy Management System - HEMS, Wide Area Situational Awareness - WASA etc), PLC enjoys the advantage of existing (power conductor) medium and better reliable data communication. PLC is a promising wired communication technology in that the electrical power lines are already there and the deployment cost can be comparable to wireless technologies. The 6lo related scenarios lie in the low voltage PLC networks with most applications in the area of Advanced Metering Infrastructure, Vehicle-to-Grid communications, in-home energy management and smart street lighting.

Example: Use of PLC for Advanced Metering Infrastructure

Household electricity meters transmit time-based data of electric power consumption through PLC. Data concentrators receive all the meter data in their corresponding living districts and send them to the Meter Data Management System (MDMS) through WAN network (e.g. Medium-Voltage PLC, Ethernet or GPRS) for storage and analysis. Two-way communications are enabled which means smart meters can do actions like notification of electricity charges according to the commands from the utility company.

With the existing power line infrastructure as communication medium, cost on building up the PLC network is naturally saved, and more importantly, labor operational costs can be minimized from a long-

term perspective. Furthermore, this AMI application speeds up electricity charge, reduces losses by restraining power theft and helps to manage the health of the grid based on line loss analysis.

Example: Use of PLC (IEEE1901.1) for WASA in Smart Grid

Many sub-systems of Smart Grid require low data rate and narrowband variant (IEEE1901.2) of PLC fulfils such requirements. Recently, more complex scenarios are emerging that require higher data rates.

WASA sub-system is an appropriate example that collects large amount of information about the current state of the grid over wide area from electric substations as well as power transmission lines. The collected feedback is used for monitoring, controlling and protecting all the sub-systems.

A.6. Use case of IEEE 802.15.4e: Industrial Automation

Typical scenario of Industrial Automation where sensor and actuators are connected through the time-slotted radio access (IEEE 802.15.4e). For that, there will be a point-to-point control signal exchange in between sensors and actuators to trigger the critical control information. In such scenarios, point-to-point traffic flows are significant to exchange the controlled information in between sensors and actuators within the constrained networks.

Example: Use of IEEE 802.15.4e for P2P communication in closed-loop application

AODV-RPL [I-D.ietf-roll-aodv-rpl] is proposed as a standard P2P routing protocol to provide the hop-by-hop data transmission in closed-loop constrained networks. Scheduling Functions i.e. SF0 [I-D.ietf-6tisch-6top-sfx] and SF1 [I-D.satish-6tisch-6top-sf1] is proposed to provide distributed neighbor-to-neighbor and end-to-end resource reservations, respectively for traffic flows in deterministic networks (6TiSCH).

The potential scenarios that can make use of the end-to-end resource reservations can be in health-care and industrial applications. AODV-RPL and SF0/SF1 are the significant routing and resource reservation protocols for closed-loop applications in constrained networks.

Authors' Addresses

Yong-Geun Hong
ETRI
161 Gajeong-Dong Yuseung-Gu
Daejeon 305-700
Korea

Phone: +82 42 860 6557
Email: yghong@etri.re.kr

Carles Gomez
Universitat Politecnica de Catalunya/Fundacio i2cat
C/Esteve Terradas, 7
Castelldefels 08860
Spain

Email: carlesgo@entel.upc.edu

Younghwan Choi
ETRI
218 Gajeongno, Yuseong
Daejeon 305-700
Korea

Phone: +82 42 860 1429
Email: yhc@etri.re.kr

Deoknyong Ko
SKtelecom
9-1 Byundang-gu Sunae-dong, Seongnam-si
Gyeonggi-do 13595
Korea

Phone: +82 10 3356 8052
Email: engineer@sk.com

Abdur Rashid Sangi
Huaiyin Institute of Technology
No.89 North Beijing Road, Qinghe District
Huaian 223001
P.R. China

Email: sangi_bahrian@yahoo.com

Take Aanstoot
Modio AB
S:t Larsgatan 15, 582 24
Linkoping
Sweden

Email: take@modio.se

Samita Chakrabarti
San Jose, CA
USA

Email: samitac.ietf@gmail.com

6lo
Internet-Draft
Updates: 4944 (if approved)
Intended status: Standards Track
Expires: December 29, 2018

P. Thubert, Ed.
Cisco Systems
June 27, 2018

6LoWPAN Selective Fragment Recovery
draft-thubert-6lo-fragment-recovery-01

Abstract

This draft updates RFC 4944 with a simple protocol to recover individual fragments across a route-over mesh network, with a minimal flow control to protect the network against bloat.

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 December 29, 2018.

Copyright Notice

Table of Contents

1. Introduction

In most Low Power and Lossy Network (LLN) applications, the bulk of the traffic consists of small chunks of data (in the order few bytes to a few tens of bytes) at a time. Given that an IEEE Std. 802.15.4 [IEEE.802.15.4] frame can carry 74 bytes or more in all cases, fragmentation is usually not required. However, and though this happens only occasionally, a number of mission critical applications do require the capability to transfer larger chunks of data, for instance to support a firmware upgrades of the LLN nodes or an extraction of logs from LLN nodes. In the former case, the large chunk of data is transferred to the LLN node, whereas in the latter, the large chunk flows away from the LLN node. In both cases, the size can be on the order of 10Kbytes or more and an end-to-end reliable transport is required.

"Transmission of IPv6 Packets over IEEE 802.15.4 Networks" [RFC4944] defines the original 6LoWPAN datagram fragmentation mechanism for

LLNs. One critical issue with this original design is that routing an IPv6 [RFC8200] packet across a route-over mesh requires to reassemble the full packet at each hop, which may cause latency along a path and an overall buffer bloat in the network. The "6TiSCH Architecture" [I-D.ietf-6tisch-architecture] recommends to use a hop-by-hop fragment forwarding technique to alleviate those undesirable effects. "LLN Minimal Fragment Forwarding" [I-D.watteyne-6lo-minimal-fragment] proposes such a technique, in a fashion that is compatible with [RFC4944] without the need to define a new protocol. However, adding that capability alone to the local implementation of the original 6LoWPAN fragmentation would not address the bulk of the issues raised against it, and may create new issues like remnant state in the network.

Another issue against [RFC4944] is that it does not define a mechanism to first discover the loss of a fragment along a multi-hop path (e.g. having exhausted the link-layer retries at some hop on the way), and then to recover that loss. With RFC 4944, the forwarding of a whole datagram fails when one fragment is not delivered properly to the destination 6LoWPAN endpoint. End-to-end transport or application-level mechanisms may require a full retransmission of the datagram, wasting resources in an already constrained network.

In that situation, the source 6LoWPAN endpoint will not be aware that a loss occurred and will continue sending all fragments for a datagram that is already doomed. The original support is missing signaling to abort a multi-fragment transmission at any time and from either end, and, if the capability to forward fragments is implemented, clean up the related state in the network. It is also lacking flow control capabilities to avoid participating to a congestion that may in turn cause the loss of a fragment and trigger the retransmission of the full datagram.

This specification proposes a method to forward fragments across a multi-hop route-over mesh, and to recover individual fragments between LLN endpoints. The method is designed to limit congestion loss in the network and addresses the requirements that are detailed in Appendix B.

2. Updating RFC 4944

This specification updates the fragmentation mechanism that is specified in "Transmission of IPv6 Packets over IEEE 802.15.4 Networks" [RFC4944] for use in route-over LLNs by providing a model where fragments can be forwarded end-to-end across a 6LoWPAN LLN, and where fragments that are lost on the way can be recovered individually. A new format for fragment is introduces and new dispatch types are defined in Section 5.

3. Updating draft-watteyne-6lo-minimal-fragment

This specification updates the fragment forwarding mechanism specified in "LLN Minimal Fragment Forwarding" [I-D.watteyne-6lo-minimal-fragment] by providing additional events to iprove the management of the Virtual Reassembly Buffer (VRB).

At the time of this writing, [I-D.watteyne-6lo-minimal-fragment] allows for refragmenting in intermediate nodes, meaning that some bytes from a given fragment may be left in the VRB to be added to the next fragment. The reason for this to happen would be the need for space in the outgoing fragment that was not needed in the incoming fragment, for instance because the 6LoWPAN Header Compression is not as efficient on the outgoing link, e.g., if the IID of the source IPv6 address is elided on the first hop because it matches the MAC address, but cannot be on the next hops. This specification does not allow this since fragments are recovered end-to-end. This means that the fragments that contain 6LoWPAN-compressed data must have enough slack in them to enable a lesser efficient compression in the next hops to still fit in one frame.

4. Terminology

4.1. BCP 14

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.

4.2. References

In this document, readers will encounter terms and concepts that are discussed in the following documents:

- "Problem Statement and Requirements for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing" [RFC6606]

4.3. 6LoWPAN Acronyms

This document uses the following acronyms:

6BBR: 6LoWPAN Backbone Router

6LBR: 6LoWPAN Border Router

6LN: 6LoWPAN Node

6LR: 6LoWPAN Router

LLN: Low-Power and Lossy Network

4.4. Referenced Work

Past experience with fragmentation has shown that miss-associated or lost fragments can lead to poor network behavior and, occasionally, trouble at application layer. The reader is encouraged to read "IPv4 Reassembly Errors at High Data Rates" [RFC4963] and follow the references for more information.

That experience led to the definition of "Path MTU discovery" [RFC8201] (PMTUD) protocol that limits fragmentation over the Internet.

Specifically in the case of UDP, valuable additional information can be found in "UDP Usage Guidelines for Application Designers" [RFC8085].

Readers are expected to be familiar with all the terms and concepts that are discussed in "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals" [RFC4919] and "Transmission of IPv6 Packets over IEEE 802.15.4 Networks" [RFC4944].

"The Benefits of Using Explicit Congestion Notification (ECN)" [RFC8087] provides useful information on the potential benefits and pitfalls of using ECN.

Quoting the "Multiprotocol Label Switching (MPLS) Architecture" [RFC3031]: with MPLS, "packets are "labeled" before they are forwarded. At subsequent hops, there is no further analysis of the packet’s network layer header. Rather, the label is used as an index into a table which specifies the next hop, and a new label". The MPLS technique is leveraged in the present specification to forward fragments that actually do not have a network layer header, since the fragmentation occurs below IP.

"LLN Minimal Fragment Forwarding" [I-D.watteyne-6lo-minimal-fragment] introduces the concept of a Virtual Reassembly Buffer (VRB) and an associated technique to forward fragments as they come, using the datagram_tag as a label in a fashion similar to MLPS. This specification reuses that technique with slightly modified controls.

4.5. New Terms

This specification uses the following terms:

6LoWPAN endpoints

> The LLN nodes in charge of generating or expanding a 6LoWPAN header from/to a full IPv6 packet. The 6LoWPAN endpoints are the points where fragmentation and reassembly take place.

5. New Dispatch types and headers

This specification enables the 6LoWPAN fragmentation sublayer to provide an MTU up to 2048 bytes to the upper layer, which can be the 6LoWPAN Header Compression sublayer that is defined in the "Compression Format for IPv6 Datagrams" [RFC6282] specification. In order to achieve this, this specification enables the fragmentation and the reliable transmission of fragments over a multihop 6LoWPAN mesh network.

This specification provides a technique that is derived from MPLS in order to forward individual fragments across a 6LoWPAN route-over mesh. The datagram_tag is used as a label; it is locally unique to the node that is the source MAC address of the fragment, so together the MAC address and the label can identify the fragment globally. A node may build the datagram_tag in its own locally-significant way, as long as the selected tag stays unique to the particular datagram for the lifetime of that datagram. It results that the label does not need to be globally unique but also that it must be swapped at each hop as the source MAC address changes.

This specification extends RFC 4944 [RFC4944] with 4 new Dispatch types, for Recoverable Fragment (RFRAG) headers with or without Acknowledgment Request (RFRAG vs. RFRAG-ARQ), and for the RFRAG Acknowledgment back, with or without ECN Echo (RFRAG-ACK vs. RFRAG-ECHO).

(to be confirmed by IANA) The new 6LoWPAN Dispatch types use the Value Bit Pattern of 11 1010xx from page 0 [RFC8025], as follows:

Pattern	Header Type	
11	101000	RFRAG - Recoverable Fragment
11	101001	RFRAG-ARQ - RFRAG with Ack Request
11	101010	RFRAG-ACK - RFRAG Acknowledgment
11	101011	RFRAG-ECHO - RFRAG Ack with ECN Echo

Figure 1: Additional Dispatch Value Bit Patterns

In the following sections, the semantics of "datagram_tag" are unchanged from [RFC4944] Section 5.3. "Fragmentation Type and Header." and is compatible with the fragment forwarding operation described in [I-D.watteyne-6lo-minimal-fragment].

5.1. Recoverable Fragment Dispatch type and Header

In this specification, the size and offset of the fragments are expressed on the compressed packet form as opposed to the uncompressed - native - packet form.

The first fragment is recognized by a sequence of 0; it carries its fragment_size and the datagram_size of the compressed packet, whereas the other fragments carry their fragment_size and fragment_offset. The last fragment for a datagram is recognized when its fragment_offset and its fragment_size add up to the datagram_size.

Recoverable Fragments are sequenced and a bitmap is used in the RFRAG Acknowledgment to indicate the received fragments by setting the individual bits that correspond to their sequence.

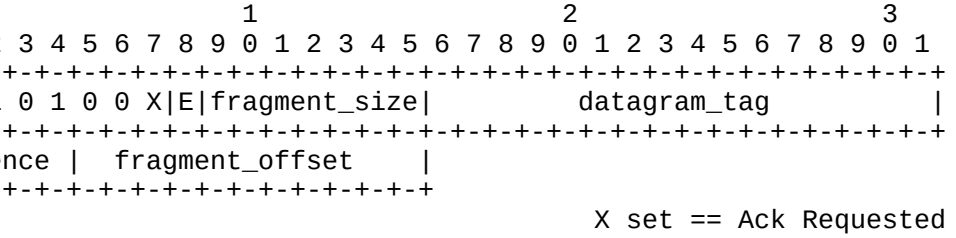


Figure 2: RFRAG Dispatch type and Header

X: 1 bit; Ack Requested: when set, the sender requires an RFRAG Acknowledgment from the receiver.

E: 1 bit; Explicit Congestion Notification; the "E" flag is reset by the source of the fragment and set by intermediate routers to signal that this fragment experienced congestion along its path.

Fragment_size: 7 bit unsigned integer; the size of this fragment in a unit that depends on the MAC layer technology. For IEEE Std. 802.15.4, the unit is octet, and the maximum fragment size, which is constrained by the maximum frame size of 128 octet minus the overheads of the MAC and Fragment Headers, is not limited by this encoding.

Sequence: 5 bit unsigned integer; the sequence number of the fragment. Fragments are sequence numbered [0..N] where N is in [0..31]. A sequence of 0 indicates the first fragment in a datagram. For IEEE Std. 802.15.4, as long as the overheads enable a fragment size of 64 octets or more, this enables to fragment a packet of 2047 octets.

Fragment_offset: 11 bit unsigned integer;

* When set to a non-0 value, the semantics of the Fragment_offset depends on the value of the Sequence.

  + When the Sequence is not 0, this field indicates the offset of the fragment in the compressed form. The fragment should be forwarded based on an existing VRB as described in Section 7.2, or silently dropped if none is found.

  + For a first fragment (i.e. with a sequence of 0), this field is overloaded to indicate the total_size of the compressed packet, to help the receiver allocate an adapted buffer for the reception and reassembly operations. This format limits the maximum MTU on a 6LoWPAN link to 2047 bytes, but 1280 bytes is the recommended value to avoid issues with IPV6 Path MTU Discovery [RFC8201]. The fragment should be routed based on the destination IPv6 address, and an VRB state should be installed as described in Section 7.1.

* When set to 0, this field indicates an abort condition and all state regarding the datagram should be cleaned up once the processing of the fragment is complete; the processing of the fragment depends on whether there is a VRB already established for this datagram, and the next hop is still reachable:

  + if a VRB already exists and is not broken, the fragment is to be forwarded along the associated Label Switched Path (LSP) as described in Section 7.2, but regardless of the value of the Sequence field;

  + else, if the Sequence is 0, then the fragment is to be routed as described in Section 7.1 but no state is conserved afterwards.

If the fragment cannot be forwarded or routed, then an abort RFRAG-ACK is sent back to the source.

5.2. RFRAG Acknowledgment Dispatch type and Header

This specification also defines a 4-octet RFRAG Acknowledgment bitmap that is used by the reassembling end point to confirm selectively the reception of individual fragments. A given offset in the bitmap maps one to one with a given sequence number.

The offset of the bit in the bitmap indicates which fragment is acknowledged as follows:

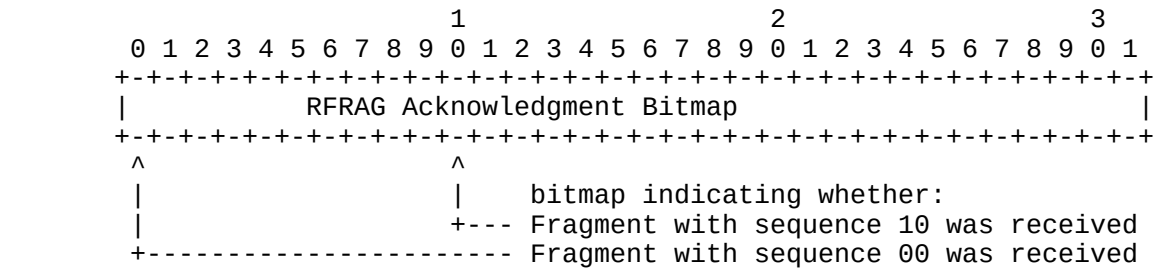


Figure 3: RFRAG Acknowledgment bitmap encoding

Figure 4 shows an example Acknowledgment bitmap which indicates that all fragments from sequence 0 to 20 were received, except for fragments 1, 2 and 16 that were either lost or are still in the network over a slower path.

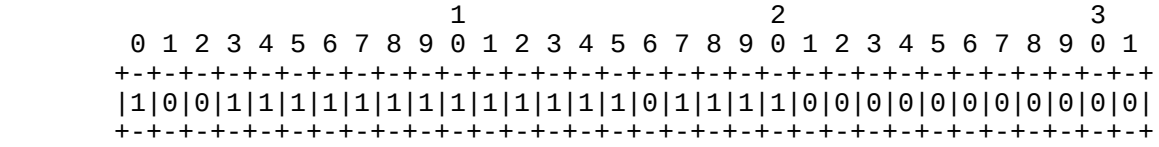


Figure 4: Expanding 3 octets encoding

The RFRAG Acknowledgment Bitmap is included in a RFRAG Acknowledgment header, as follows:

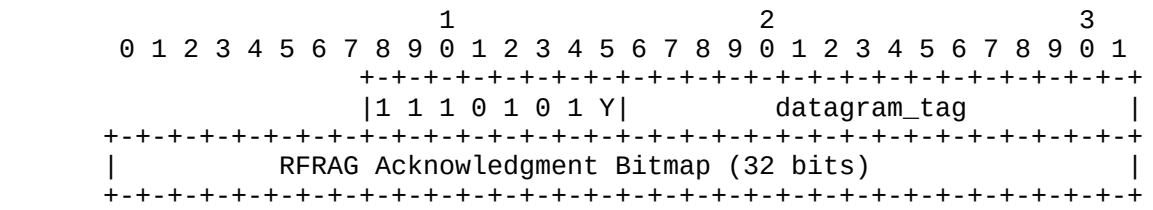


Figure 5: RFRAG Acknowledgment Dispatch type and Header

Y: 1 bit; Explicit Congestion Notification Echo

When set, the sender indicates that at least one of the acknowledged fragments was received with an Explicit Congestion Notification, indicating that the path followed by the fragments is subject to congestion.

RFRAG Acknowledgment Bitmap

An RFRAG Acknowledgment Bitmap, whereby setting the bit at offset x indicates that fragment x was received, as shown in Figure 3. All 0's is a NULL bitmap that indicates that the fragmentation process is aborted. All 1's is a FULL bitmap that indicates that the fragmentation process is complete, all fragments were received at the reassembly end point.

6. Fragments Recovery

The Recoverable Fragment headers RFRAG and RFRAG-ARQ are used to transport a fragment and optionally request an RFRAG Acknowledgment that will confirm the good reception of a one or more fragments. An RFRAG Acknowledgment can optionally carry an ECN indication; it is carried as a standalone header in a message that is sent back to the 6LoWPAN endpoint that was the source of the fragments, as known by its MAC address. The process ensures that at every hop, the source MAC address and the datagram_tag in the received fragment are enough information to send the RFRAG Acknowledgment back towards the source 6LoWPAN endpoint by reversing the MPLS operation.

The 6LoWPAN endpoint that fragments the packets at 6LoWPAN level (the sender) also controls when the reassembling end point sends the RFRAG Acknowledgments by setting the Ack Requested flag in the RFRAG packets. It may set the Ack Requested flag on any fragment to perform congestion control by limiting the number of outstanding fragments, which are the fragments that have been sent but for which reception or loss was not positively confirmed by the reassembling endpoint. When the sender of the fragment knows that an underlying link-layer mechanism protects the Fragments, it may refrain from

using the RFRAG Acknowledgment mechanism, and never set the Ack Requested bit. When it receives a fragment with the ACK Request flag set, the 6LoWPAN endpoint that reassembles the packets at 6LoWPAN level (the receiver) sends back an RFRAG Acknowledgment to confirm reception of all the fragments it has received so far.

The sender transfers a controlled number of fragments and MAY flag the last fragment of a series with an RFRAG Acknowledgment Request. The received MUST acknowledge a fragment with the acknowledgment request bit set. If any fragment immediately preceding an acknowledgment request is still missing, the receiver MAY intentionally delay its acknowledgment to allow in-transit fragments to arrive. Delaying the acknowledgment might defeat the round trip delay computation so it should be configurable and not enabled by default.

The receiver MAY issue unsolicited acknowledgments. An unsolicited acknowledgment signals to the sender endpoint that it can resume sending if it had reached its maximum number of outstanding fragments. Another use is to inform that the reassembling endpoint has canceled the process of an individual datagram. Note that acknowledgments might consume precious resources so the use of unsolicited acknowledgments should be configurable and not enabled by default.

An observation is that streamlining forwarding of fragments generally reduces the latency over the LLN mesh, providing room for retries within existing upper-layer reliability mechanisms. The sender protects the transmission over the LLN mesh with a retry timer that is computed according to the method detailed in [RFC6298]. It is expected that the upper layer retries obey the recommendations in "UDP Usage Guidelines" [RFC8085], in which case a single round of fragment recovery should fit within the upper layer recovery timers.

Fragments are sent in a round robin fashion: the sender sends all the fragments for a first time before it retries any lost fragment; lost fragments are retried in sequence, oldest first. This mechanism enables the receiver to acknowledge fragments that were delayed in the network before they are actually retried.

When a single frequency is used by contiguous hops, the sender should wait a reasonable amount of time between fragments so as to let a fragment progress a few hops and avoid hidden terminal issues. This precaution is not required on channel hopping technologies such as Time Slotted CHannel Hopping (TSCH) [RFC6554]

When the sender decides that a packet should be dropped and the fragmentation process canceled, it sends a pseudo fragment with the

fragment_offset, sequence and fragment_size all set to 0, and no data. Upon reception of this message, the receiver should clean up all resources for the packet associated to the datagram_tag. If an acknowledgment is requested, the receiver responds with a NULL bitmap.

The receiver might need to cancel the process of a fragmented packet for internal reasons, for instance if it is out of reassembly buffers, or considers that this packet is already fully reassembled and passed to the upper layer. In that case, the receiver SHOULD indicate so to the sender with a NULL bitmap in a RFRAG Acknowledgment. Upon an acknowledgment with a NULL bitmap, the sender endpoint MUST abort the transmission of the fragmented datagram.

7. Forwarding Fragments

It is assumed that the first Fragment is large enough to carry the IPv6 header and make routing decisions. If that is not so, then this specification MUST NOT be used.

This specification extends the Virtual Reassembly Buffer (VRB) technique to forward fragments with no intermediate reconstruction of the entire packet. The first fragment carries the IP header and it is routed all the way from the fragmenting end point to the reassembling end point. Upon the first fragment, the routers along the path install a label-switched path (LSP), and the following fragments are label-switched along that path. As a consequence, alternate routes not possible for individual fragments. The datagram_tag is used to carry the label, that is swapped at each hop. All fragments follow the same path and fragments are delivered in the order at which they are sent.

7.1. Upon the first fragment

In Route-Over mode, the source and destination MAC addressed in a frame change at each hop. The label that is formed and placed in the datagram_tag is associated to the source MAC and only valid (and unique) for that source MAC. Upon a first fragment (i.e. with a sequence of zero), a VRB and the associated LSP state are created for the tuple (source MAC address, datagram_tag) and the fragment is forwarded along the IPv6 route that matches the destination IPv6 address in the IPv6 header as prescribed by [I-D.watteyne-6lo-minimal-fragment]. The LSP state enables to match the (previous MAC address, datagram_tag) in an incoming fragment to the tuple (next MAC address, swapped datagram_tag) used in the forwarded fragment and points at the VRB. In addition, the router also forms a Reverse LSP state indexed by the MAC address of the next

hop and the swapped datagram_tag. This reverse LSP state also points at the VRB and enables to match the (next MAC address, swapped_datagram_tag) found in an RFRAG Acknowledgment to the tuple (previous MAC address, datagram_tag) used when forwarding a Fragment Acknowledgment (RFRAG-ACK) back to the sender endpoint.

7.2. Upon the next fragments

Upon a next fragment (i.e. with a non-zero sequence), the router looks up a LSP indexed by the tuple (MAC address, datagram_tag) found in the fragment. If it is found, the router forwards the fragment using the associated VRB as prescribed by [I-D.watteyne-6lo-minimal-fragment].

if the VRB for the tuple is not found, the router builds an RFRAG-ACK to abort the transmission of the packet. The resulting message has the following information:

- The source and destination MAC addresses are swapped from those found in the fragment
- The datagram_tag set to the datagram_tag found in the fragment
- A null bitmap is used to signal the abort condition

At this point the router is all set and can send the RFRAG-ACK back to the previous router. The RFRAG-ACK should normally be forwarded all the way to the source using the reverse LSP state in the VRBs in the intermediate routers as described in the next section.

7.3. Upon the RFRAG Acknowledgments

Upon an RFRAG-ACK, the router looks up a Reverse LSP indexed by the tuple (MAC address, datagram_tag), which are respectively the source MAC address of the received frame and the received datagram_tag. If it is found, the router forwards the fragment using the associated VRB as prescribed by [I-D.watteyne-6lo-minimal-fragment], but using the Reverse LSP so that the RFRAG-ACK flows back to the sender endpoint.

If the Reverse LSP is not found, the router MUST silently drop the RFRAG-ACK message.

Either way, if the RFRAG-ACK indicates either an error (NULL bitmap) or that the fragment was entirely received (FULL bitmap), arms a short timer, and upon timeout, the VRB and all associate state are destroyed. During that time, fragments of that datagram may still be received, e.g. if the RFRAG-ACK was lost on the way back and the

source retried the last fragment. In that case, the router sends an abort RFRAG-ACK along the Reverse LSP to complete the clean up.

8. Security Considerations

The process of recovering fragments does not appear to create any opening for new threat compared to "Transmission of IPv6 Packets over IEEE 802.15.4 Networks" [RFC4944].

9. IANA Considerations

Need extensions for formats defined in "Transmission of IPv6 Packets over IEEE 802.15.4 Networks" [RFC4944].

10. Acknowledgments

The author wishes to thank Thomas Watteyne and Michael Richardson for in-depth reviews and comments. Also many thanks to Jonathan Hui, Jay Werb, Christos Polyzois, Soumitri Kolavennu, Pat Kinney, Margaret Wasserman, Richard Kelsey, Carsten Bormann and Harry Courtice for their various contributions.

11. References

11.1. Normative References

[I-D.watteyne-6lo-minimal-fragment]
Watteyne, T., Bormann, C., and P. Thubert, "LLN Minimal Fragment Forwarding", draft-watteyne-6lo-minimal-fragment-01 (work in progress), March 2018.

[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>.

[RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <https://www.rfc-editor.org/info/rfc4944>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6554] Hui, J., Vasseur, JP., Culler, D., and V. Manral, "An IPv6 Routing Header for Source Routes with the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6554, DOI 10.17487/RFC6554, March 2012, <https://www.rfc-editor.org/info/rfc6554>.

[RFC8025] Thubert, P., Ed. and R. Cragie, "IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Paging Dispatch", RFC 8025, DOI 10.17487/RFC8025, November 2016, <https://www.rfc-editor.org/info/rfc8025>.

[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>.

11.2. Informative References

[I-D.ietf-6tisch-architecture] Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-14 (work in progress), April 2018.

[IEEE.802.15.4] IEEE, "IEEE Standard for Low-Rate Wireless Networks", IEEE Standard 802.15.4, DOI 10.1109/IEEE P802.15.4-REVd/D01, <http://ieeexplore.ieee.org/document/7460875/>.

[RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, DOI 10.17487/RFC2914, September 2000, <https://www.rfc-editor.org/info/rfc2914>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <https://www.rfc-editor.org/info/rfc3168>.

[RFC4919] Kushalnagar, N., Montenegro, G., and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals", RFC 4919, DOI 10.17487/RFC4919, August 2007, <https://www.rfc-editor.org/info/rfc4919>.

[RFC4963] Heffner, J., Mathis, M., and B. Chandler, "IPv4 Reassembly Errors at High Data Rates", RFC 4963, DOI 10.17487/RFC4963, July 2007, <https://www.rfc-editor.org/info/rfc4963>.

[RFC5681] Allman, M., Paxson, V., and E. Blanton, "TCP Congestion Control", RFC 5681, DOI 10.17487/RFC5681, September 2009, <https://www.rfc-editor.org/info/rfc5681>.

[RFC6298] Paxson, V., Allman, M., Chu, J., and M. Sargent, "Computing TCP's Retransmission Timer", RFC 6298, DOI 10.17487/RFC6298, June 2011, <https://www.rfc-editor.org/info/rfc6298>.

[RFC6606] Kim, E., Kaspar, D., Gomez, C., and C. Bormann, "Problem Statement and Requirements for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing", RFC 6606, DOI 10.17487/RFC6606, May 2012, <https://www.rfc-editor.org/info/rfc6606>.

[RFC7554] Watteyne, T., Ed., Palattella, M., and L. Grieco, "Using IEEE 802.15.4e Time-Slotted Channel Hopping (TSCH) in the Internet of Things (IoT): Problem Statement", RFC 7554, DOI 10.17487/RFC7554, May 2015, <https://www.rfc-editor.org/info/rfc7554>.

[RFC7567] Baker, F., Ed. and G. Fairhurst, Ed., "IETF Recommendations Regarding Active Queue Management", BCP 197, RFC 7567, DOI 10.17487/RFC7567, July 2015, <https://www.rfc-editor.org/info/rfc7567>.

[RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines", BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, <https://www.rfc-editor.org/info/rfc8085>.

[RFC8087] Fairhurst, G. and M. Welzl, "The Benefits of Using Explicit Congestion Notification (ECN)", RFC 8087, DOI 10.17487/RFC8087, March 2017, <https://www.rfc-editor.org/info/rfc8087>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6", STD 87, RFC 8201, DOI 10.17487/RFC8201, July 2017, <https://www.rfc-editor.org/info/rfc8201>.

Appendix A. Rationale

There are a number of uses for large packets in Wireless Sensor Networks. Such usages may not be the most typical or represent the largest amount of traffic over the LLN; however, the associated functionality can be critical enough to justify extra care for ensuring effective transport of large packets across the LLN.

The list of those usages includes:

Towards the LLN node:

- Firmware update: For example, a new version of the LLN node software is downloaded from a system manager over unicast or multicast services. Such a reflashing operation typically involves updating a large number of similar LLN nodes over a relatively short period of time.

- Packages of Commands: A number of commands or a full configuration can be packaged as a single message to ensure consistency and enable atomic execution or complete roll back. Until such commands are fully received and interpreted, the intended operation will not take effect.

From the LLN node:

- Waveform captures: A number of consecutive samples are measured at a high rate for a short time and then transferred from a sensor to a gateway or an edge server as a single large report.

- Data logs: LLN nodes may generate large logs of sampled data for later extraction. LLN nodes may also generate system logs to assist in diagnosing problems on the node or network.

- Large data packets: Rich data types might require more than one fragment.

Uncontrolled firmware download or waveform upload can easily result in a massive increase of the traffic and saturate the network.

When a fragment is lost in transmission, the lack of recovery in the original fragmentation system of RFC 4944 implies that all fragments

are resent, further contributing to the congestion that caused the initial loss, and potentially leading to congestion collapse.

This saturation may lead to excessive radio interference, or random early discard (leaky bucket) in relaying nodes. Additional queuing and memory congestion may result while waiting for a low power next hop to emerge from its sleeping state.

Considering that RFC 4944 defines an MTU is 1280 bytes and that in most incarnations (but 802.15.4g) a IEEE Std. 802.15.4 frame can limit the MAC payload to as few as 74 bytes, a packet might be fragmented into at least 18 fragments at the 6LoWPAN shim layer. Taking into account the worst-case header overhead for 6LoWPAN Fragmentation and Mesh Addressing headers will increase the number of required fragments to around 32. This level of fragmentation is much higher than that traditionally experienced over the Internet with IPv4 fragments. At the same time, the use of radios increases the probability of transmission loss and Mesh-Under techniques compound that risk over multiple hops.

Mechanisms such as TCP or application-layer segmentation could be used to support end-to-end reliable transport. One option to support bulk data transfer over a frame-size-constrained LLN is to set the Maximum Segment Size to fit within the link maximum frame size. Doing so, however, can add significant header overhead to each 802.15.4 frame. In addition, deploying such a mechanism requires that the end-to-end transport is aware of the delivery properties of the underlying LLN, which is a layer violation, and difficult to achieve from the far end of the IPv6 network.

Appendix B. Requirements

For one-hop communications, a number of Low Power and Lossy Network (LLN) link-layers propose a local acknowledgment mechanism that is enough to detect and recover the loss of fragments. In a multihop environment, an end-to-end fragment recovery mechanism might be a good complement to a hop-by-hop MAC level recovery. This draft introduces a simple protocol to recover individual fragments between 6LoWPAN endpoints that may be multiple hops away. The method addresses the following requirements of a LLN:

Number of fragments

> The recovery mechanism must support highly fragmented packets, with a maximum of 32 fragments per packet.

Minimum acknowledgment overhead

Because the radio is half duplex, and because of silent time spent in the various medium access mechanisms, an acknowledgment consumes roughly as many resources as data fragment.

The new end-to-end fragment recovery mechanism should be able to acknowledge multiple fragments in a single message and not require an acknowledgment at all if fragments are already protected at a lower layer.

Controlled latency

The recovery mechanism must succeed or give up within the time boundary imposed by the recovery process of the Upper Layer Protocols.

Optional congestion control

The aggregation of multiple concurrent flows may lead to the saturation of the radio network and congestion collapse.

The recovery mechanism should provide means for controlling the number of fragments in transit over the LLN.

Appendix C. Considerations On Flow Control

Considering that a multi-hop LLN can be a very sensitive environment due to the limited queuing capabilities of a large population of its nodes, this draft recommends a simple and conservative approach to congestion control, based on TCP congestion avoidance.

Congestion on the forward path is assumed in case of packet loss, and packet loss is assumed upon time out. The draft allows to control the number of outstanding fragments, that have been transmitted but for which an acknowledgment was not received yet. It must be noted that the number of outstanding fragments should not exceed the number of hops in the network, but the way to figure the number of hops is out of scope for this document.

Congestion on the forward path can also be indicated by an Explicit Congestion Notification (ECN) mechanism. Though whether and how ECN [RFC3168] is carried out over the LoWPAN is out of scope, this draft provides a way for the destination endpoint to echo an ECN indication back to the source endpoint in an acknowledgment message as represented in Figure 5 in Section 5.2.

It must be noted that congestion and collision are different topics. In particular, when a mesh operates on a same channel over multiple hops, then the forwarding of a fragment over a certain hop may

collide with the forwarding of a next fragment that is following over a previous hop but in a same interference domain. This draft enables an end-to-end flow control, but leaves it to the sender stack to pace individual fragments within a transmit window, so that a given fragment is sent only when the previous fragment has had a chance to progress beyond the interference domain of this hop. In the case of 6TiSCH [I-D.ietf-6tisch-architecture], which operates over the TimeSlotted Channel Hopping [RFC7554] (TSCH) mode of operation of IEEE802.14.5, a fragment is forwarded over a different channel at a different time and it makes full sense to transmit the next fragment as soon as the previous fragment has had its chance to be forwarded at the next hop.

From the standpoint of a source 6LoWPAN endpoint, an outstanding fragment is a fragment that was sent but for which no explicit acknowledgment was received yet. This means that the fragment might be on the way, received but not yet acknowledged, or the acknowledgment might be on the way back. It is also possible that either the fragment or the acknowledgment was lost on the way.

From the sender standpoint, all outstanding fragments might still be in the network and contribute to its congestion. There is an assumption, though, that after a certain amount of time, a frame is either received or lost, so it is not causing congestion anymore. This amount of time can be estimated based on the round trip delay between the 6LoWPAN endpoints. The method detailed in [RFC6298] is recommended for that computation.

The reader is encouraged to read through "Congestion Control Principles" [RFC2914]. Additionally [RFC7567] and [RFC5681] provide deeper information on why this mechanism is needed and how TCP handles Congestion Control. Basically, the goal here is to manage the amount of fragments present in the network; this is achieved by to reducing the number of outstanding fragments over a congested path by throttling the sources.

Section 6 describes how the sender decides how many fragments are (re)sent before an acknowledgment is required, and how the sender adapts that number to the network conditions.

Author’s Address

Pascal Thubert (editor)
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
MOUGINS - Sophia Antipolis 06254
FRANCE

Phone: +33 497 23 26 34
Email: pthubert@cisco.com

ROLL
Internet-Draft
Updates: 6550, 6775 (if approved)
Intended status: Standards Track
Expires: August 27, 2018

P. Thubert, Ed.
Cisco

February 23, 2018

Routing for RPL Leaves
draft-thubert-roll-unaware-leaves-03

Abstract

This specification updates RFC 6550 and RFC 6775 unicast routing service in a RPL domain to 6LoWPAN ND nodes that do not participate to the routing protocol.

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 August 27, 2018.

Copyright Notice

Table of Contents

1. Introduction

The design of Low Power and Lossy Networks (LLNs) is generally focused on saving energy, which is the most constrained resource of all. Other design constraints, such as a limited memory capacity, duty cycling of the LLN devices and low-power lossy transmissions, derive from that primary concern.

The IETF produced the "Routing Protocol for Low Power and Lossy Networks" [RFC6550] (RPL) to provide routing services within such constraints. RPL is a Distance-Vector protocol, which, compared to link-state protocols, limits the amount of topological knowledge that needs to be installed and maintained in each node. In order to operate in constrained networks, RPL allows a Routing Stretch (see [RFC6687]), whereby routing is only performed along a DODAG as opposed to straight along a shortest path between 2 peers, whatever that would mean in a given LLN. This trades the quality of peer-to-peer (P2P) paths for a vastly reduced amount of control traffic and routing state that would be required to operate a any-to-any shortest path protocol. Finally, broken routes may be fixed lazily and on-demand, based on dataplane inconsistency discovery, which avoids wasting energy in the proactive repair of unused paths.

In order to cope with lossy transmissions, RPL forms Direction-Oriented Directed Acyclic Graphs (DODAGs) using DODAG Information Solicitation (DIS) and DODAG Information Object (DIO) messages. For

most of the nodes, though not all, a DODAG provides multiple forwarding solutions towards the Root of the topology via so-called parents. RPL is designed to adapt to fuzzy connectivity, whereby the physical topology cannot be expected to reach a stable state, with a lazy control that creates routes proactively but only fixes them when they are used by actual traffic. It results that RPL provides reachability for most of the LLN nodes, most of the time, but does not really converge in the classical sense. RPL provides unicast and multicast routing services back to RPL-Aware nodes. A RPL-Aware Node will inject routes to self using Destination Advertisement Object (DAO) messages sent to either their parents in Storing Mode or to the Root indicating their parent in Non-Storing mode. This process effectively forms a DODAG back to the device that is a subset of the DODAG to the Root with all links reversed.

The IPv6 [RFC8200] Neighbor Discovery (IPv6 ND) Protocol (NDP) suite [RFC4861] [RFC4862] defined for fast media such a Ethernet, relies heavily on multicast operations for address discovery and duplicate address detection (DAD).

"Neighbor Discovery Optimizations for 6LoWPAN networks" [RFC6775] (6LoWPAN ND) adapts IPv6 ND for operations over energy-constrained LLNs. In particular, 6LoWPAN ND introduces a unicast host address registration mechanism that contributes to reduce the use of multicast messages that are present in the classical IPv6 ND protocol. 6LoWPAN ND defines a new Address Registration Option (ARO) that is carried in the unicast Neighbor Solicitation (NS) and Neighbor Advertisement (NA) messages between the 6LoWPAN Node (6LN) and the 6LoWPAN Router (6LR). 6LoWPAN ND also defines the Duplicate Address Request (DAR) and Duplicate Address Confirmation (DAC) messages between the 6LR and the 6LoWPAN Border Router (6LBR). In an LLN, the 6LBR is the central repository of all the registered addresses in its domain.

When a routing protocol such as RPL is used to maintain reachability within a Non-Broadcast Multi-Access (NBMA) subnet, some nodes may act as routers and participate to the routing operations whereas others may be plain hosts. In RPL terms, a plain host that does not participate to the routing protocol is called a Leaf. It must be noted that a 6LN could participate to RPL and inject DAO routes to self, but refrain from advertising DIO and get children. In that case, the 6LN is still a host but not a Leaf.

An Update to 6LoWPAN ND [I-D.ietf-6lo-rfc6775-update] defines an Extended ARO (EARO) with a 'R' flag to be used by a 6LN when registering, to indicate that this 6LN is not a router and that it will not handle its own reachability. The EARO also includes a sequence counter called Transaction ID (TID), which maps to the Path

Sequence Field found in Transit Options in RPL DAO messages. It is a prerequisite for this specification. The DAR and DAC messages are also extended as EDAR and EDAC messages respectively.

With this specification, a 6LN that operates as a Leaf uses the 'R' flag to declare itself as such and the 6LR that accepts the registration will inject routing information on behalf of the 6LN in the RPL domain. The packet forwarding operation by the 6LR serving a Leaf 6LN is described in "When to use RFC 6553, 6554 and IPv6-in-IPv6" [I-D.ietf-roll-useofrplinfo]. This document adds the capability by a 6LR to advertise the IPv6 address(es) of the 6LN in the RPL protocol. Examples of routing-agnostic 6LN may include lightly-powered sensors such as window smash sensor (alarm system), or the kinetically powered light switch.

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 [RFC2119].

The Terminology used in this document is consistent with and incorporates that described in Terms Used in Routing for Low-Power and Lossy Networks (LLNs). [RFC7102].

Other terms in use in LLNs are found in Terminology for Constrained-Node Networks [RFC7228].

A glossary of classical 6LoWPAN acronyms is given in Appendix A.

The term "byte" is used in its now customary sense as a synonym for "octet".

"RPL", "RPL Packet Information" (RPI) and "RPL Instance", DIO, DAO and DIS messages are defined in the "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks" [RFC6550] specification.

3. Updating RFC 6550

This document specifies a new behavior whereby a 6LR injects DAO messages for unicast addresses registered through the updated 6LoWPAN ND [I-D.ietf-6lo-rfc6775-update] on behalf of 6LN nodes that are not RPL-aware.

Upon the renewal of a 6lowPAN ND registration, this specification changes the behavior of the 6LR as follows. If the 'R' flag is set, the 6LR injects a DAO targeting the Registered Address, and refrains

from sending a DAR message. the DAR/DAC exchange that refreshes the state in the 6LBR happens instead between the RPL Root and the 6LBR. In that flow, the RPL Root acts as a proxy on behalf of the 6LR upon the reception of the DAO propagation initiated at the 6LR.

4. Updating RFC 6775 Update

This document makes use of the 'R' flag in the EARO option, used by a 6LN, when registering, to indicate that this 6LN is a Leaf, not aware of the RPL operation in the network, and thus does not participate to it. The behavior defined in this specification whereby the 6LR that processes the registration advertises the Registered Address in DAO messages and bypasses the DAR/DAC process for the renewal of a registration, is only triggered by an NS(EARO) that has the 'R' flag set. A RPL Leaf SHOULD set the 'R' flag.

If the 'R' flag is not set, then the Registering Node is expected to be a RPL router that handles the reachability of the Registered Address by itself. This document also specifies a keep-alive EDAR message that the RPL Root may use to maintain an existing state in the 6LBR upon receiving DAO messages. The EDAR message may only act as a refresher and can only update the Lifetime and the TID of the state in the 6LBR. A RPL router SHOULD NOT set the 'R' flag.

5. Protocol Operations

5.1. General Flow

This specification enables to save the exchange of Extended Duplicate Address messages, EDAR and EDAC, from a 6LN all the way to the 6LBR across a RPL mesh, for the sole purpose of refreshing an existing state in the 6LBR. Instead, the EDAR/EDAC exchange is proxied by the RPL Root upon a DAO message that refreshes the RPL routing state. To achieve this, the lifetimes and sequence counters in 6LoWPAN ND and RPL are aligned. In other words, the Path Sequence and the Path Lifetime in the DAO message are derived from the Transaction ID and the registration lifetime in the NS(EARO) message from the 6LN.

From the perspective of the 6LN, the registration flow happens transparently; it is not delayed by the proxy RPL operation, so the device does not need to wait more whether RPL proxy operation happens or not. The flows below are RPL Non-Storing Mode examples. In Storing Mode, the DAO ACK may not be present, and the DAO messages cascade from child to parent all the way to the DODAG Root.

On the first registration, illustrated in Figure 1, from the perspective of the 6LR, the Extended Duplicate Address message takes place as prescribed by [I-D.ietf-6lo-rfc6775-update]. When

successful, the flow creates a Neighbor Cache Entry (NCE) in the 6LR, and the 6LR injects the registered address in RPL using DAO/DAO-ACK exchanges all the way to the RPL DODAG Root. The protocol does not carry a specific information that the Extended Duplicate Address messages were already exchanged, so the Root proxies them anyway.

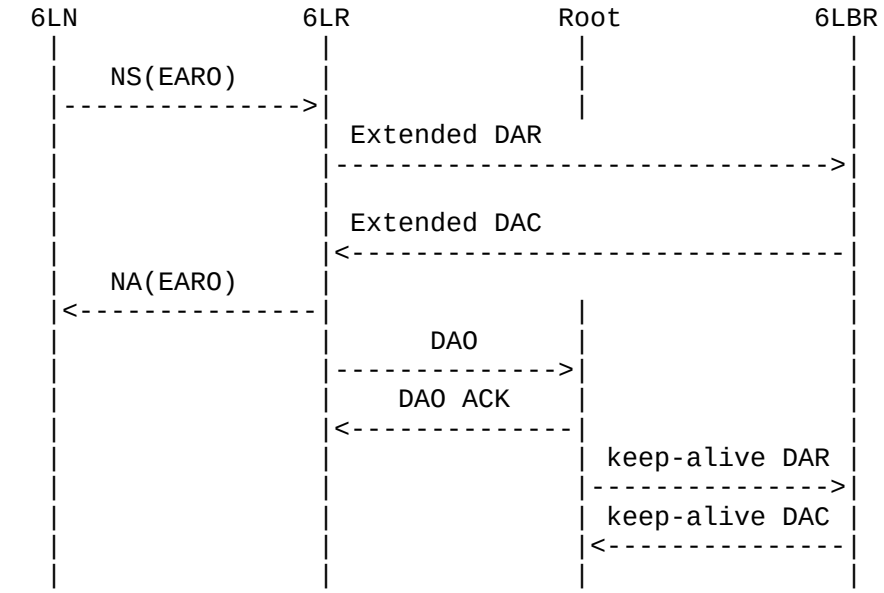


Figure 1: First Registration Flow

A re-registration is performed by the 6LN to maintain the NCE in the 6LR alive before lifetime expires. Upon a re-registration, as illustrated in Figure 1, the 6LR redistributes the NS(EARO) in RPL. This causes the RPL DODAG Root to refresh the state in the 6LBR with a keep-alive EDAR message.

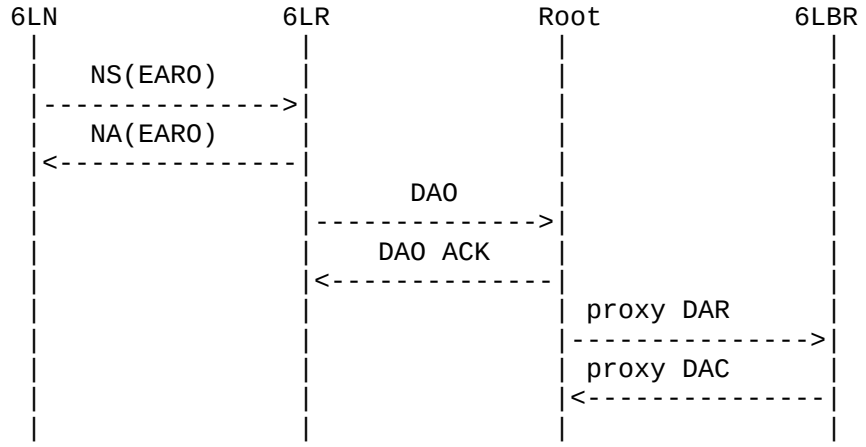


Figure 2: Next Registration Flow

Note that any of the functions 6LR, Root and 6LBR might be collapsed in a single node, in which case the flow above happens internally, and possibly through internal API calls as opposed to messaging.

5.2. 6LN Operation

This specification does not alter the operation of a 6LowpAN ND-compliant 6LN, which is expected to operate as follows:

- The 6LN obtains an IPv6 global address, for instance using autoconfiguration [RFC4862] based on a Prefix Information Option (PIO) [RFC4861] found in a Router Advertisement message or by some other means such as DHCPv6 [RFC3315].
- Once it has formed an address, the 6LN (re)registers its address periodically, within the Lifetime of the previous registration, as prescribed by [I-D.ietf-6lo-rfc6775-update].
- Upon each consecutive registration, the 6LN increases the TID field.
- The 6LN MAY register to more than one 6LR at the same time. In that case, a same value of TID is used for each registration.
- The 6LN MAY use any of the 6LRs to which it register to forward its packets.

5.3. 6LR Operation

Also as prescribed by [I-D.ietf-6lo-rfc6775-update], the 6LR generates a DAR message upon reception of a valid NS(EARO) message for the registration of a new IPv6 Address by a 6LN. If the Duplicate Address exchange succeeds, then the 6LR installs a Neighbor Cache Entry (NCE). If the ’R’ flag was set in the EARO of the NS

message, and this 6LR can manage the reachability of Registered Address, then the 6LR sets the 'R' flag in the ARO of the response NA message.

From then on, the 6LN periodically sends a new NS(EARO) to refresh the NCE state before the lifetime indicated in the EARO expires, with TID that is incremented each time till it wraps in a lollipop fashion. As long as the 'R' flag is set and this router can still manage the reachability of Registered Address, the 6LR keeps setting the 'R' flag in the ARO of the response NA message, but the exchange of Duplicate Address messages is skipped.

Upon a successful NS/NA(EARO) exchange: if the 'R' flag was set in the EARO of the NS message, then the 6LR SHOULD inject the Registered Address in RPL by sending a DAO message on behalf of the 6LN; else the 6LR SHOULD refrain from injecting the registered address into RPL.

The DAO message advertising the Registered Address MUST be constructed as follows:

- The registered address is placed in a RPL Target Option in the DAO message as the Target Prefix, and the Prefix Length is set to 128
- the External 'E' flag in the Transit Information Option (TIO) associated to the Target Option is set to indicate that the 6LR redistributes an external target into the RPL network
- the Path Lifetime in the TIO is computed from the Lifetime in the EARO Option to adapt it to the Lifetime Units used in the RPL operation. Note that if the lifetime is 0, then the 6LR generates a No-Path DAO message that cleans up the routes down to the Address of the 6LN.
- the Path Sequence in the TIO is set to the TID value found in the EARO option.
- Additionally, in Non-Storing Mode the 6LR indicates one of its global IPv6 unicast addresses as the Parent Address in the TIO.

If a 6LR receives a valid NS(EARO) message with the 'R' flag reset and the 6LR was redistributing the registered address due to previous NS(EARO) messages with the flag set, then it MUST stop injecting the address. It is up to the Registering Node to maintain the corresponding route from then on, either keeping it active by sending further DAO messages, or destroying it using a No-Path DAO.

5.4. RPL Root Operation

In RPL Storing Mode of Operation (MOP), the DAO message is propagated from child to parent all the way to the Root along the DODAG, populating routing state as it goes. In Non-Storing Mode, The DAO

message is sent directly to the route. Upon reception of a DAO message that creates or updates an existing RPL state, the Root notifies the 6LR using an internal API if they are collocated, or a proxied DAR/DAC exchange on behalf of the registering node if they are separated.

In the latter case, the DAR message MUST be constructed as follows:

- The registered address from in the Target Option is placed in the Registered Address field
- the Owner Unique ID field is set to all ones to indicate that it is not provided
- the Registration Lifetime in the DAR message is adapted from the Path Lifetime in the TIO.
- the TID value is set to the Path Sequence in the TIO.

Upon a status in a DAC message that is not "Success", the Root MAY destroy the formed paths using a No-Path DAO downwards as specified in [I-D.ietf-roll-efficient-npdao].

In Non-Storing Mode, the outer IPv6 header that is used by the Root to transport the source routing information in data packets down the DODAG has the 6LR that serves the 6LN as final destination. This way, when the final 6LR decapsulates the outer header, it also removes all the RPL artifacts from the packet.

5.5. 6LBR Operation

Upon reception of a DAR message with the Owner Unique ID field is set to all ones, the 6LBR checks whether an entry exists for the and computes whether the TID in the DAR message is fresher than that in the entry as prescribed in section 4.2.1. of [I-D.ietf-6lo-rfc6775-update].

If the entry does not exist, the 6LBR does not create the entry, and answers with a Status "Removed" in the DAC message.

If the entry exists but is not fresher, the 6LBR does not update the entry, and answers with a Status "Success" in the DAC message.

If the entry exists and the TID in the DAR message is fresher, the 6LBR updates the TID in the entry, and if the lifetime of the entry is extended by the Registration Lifetime in the DAR message, it also updates the lifetime of the entry. In that case, the 6LBR replies with a Status "Success" in the DAC message.

6. Implementation Status

7. Security Considerations

The LLN nodes depend on the 6LBR and the RPL participants for their operation. A trust model must be put in place to ensure that the right devices are acting in these roles, so as to avoid threats such as black-holing, or bombing attack whereby an impersonated 6LBR would destroy state in the network by using the "Removed" Status code. This trust model could be at a minimum based on a Layer-2 access control, or could provide role validation as well. This is a generic 6LoWPAN requirement, see Req5.1 in Appendix of [I-D.ietf-6lo-rfc6775-update].

The keep-alive EDAR message does not carry a valid Registration Unique ID [I-D.ietf-6lo-rfc6775-update] and it cannot be used to create a binding state in the 6LBR. The 6LBR MUST NOT create an entry based on a keep-alive EDAR that does not match an existing entry. All it can do is refresh the lifetime and the TID of an existing entry.

8. IANA Considerations

This specification has no requirement on IANA.

9. Acknowledgments

The author wishes to thank Michael Richardson and Georgios Papadopoulos for their early reviews of and contributions to this document

10. References

10.1. Normative References

[I-D.ietf-6lo-rfc6775-update]
Thubert, P., Nordmark, E., Chakrabarti, S., and C. Perkins, "An Update to 6LoWPAN ND", draft-ietf-6lo-rfc6775-update-13 (work in progress), February 2018.

[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>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <https://www.rfc-editor.org/info/rfc6550>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <https://www.rfc-editor.org/info/rfc6775>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

10.2. Informative References

[I-D.ietf-6lo-ap-nd] Thubert, P., Sarikaya, B., and M. Sethi, "Address Protected Neighbor Discovery for Low-power and Lossy Networks", draft-ietf-6lo-ap-nd-05 (work in progress), January 2018.

[I-D.ietf-roll-efficient-npdao] Jadhav, R., Sahoo, R., and Z. Cao, "No-Path DAO modifications", draft-ietf-roll-efficient-npdao-01 (work in progress), October 2017.

[I-D.ietf-roll-useofrplinfo] Robles, I., Richardson, M., and P. Thubert, "When to use RFC 6553, 6554 and IPv6-in-IPv6", draft-ietf-roll-useofrplinfo-21 (work in progress), February 2018.

[IEEEstd802154]
IEEE standard for Information Technology, "IEEE Standard for Local and metropolitan area networks-- Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)".

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <https://www.rfc-editor.org/info/rfc3315>.

[RFC6687] Tripathi, J., Ed., de Oliveira, J., Ed., and JP. Vasseur, Ed., "Performance Evaluation of the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6687, DOI 10.17487/RFC6687, October 2012, <https://www.rfc-editor.org/info/rfc6687>.

[RFC7102] Vasseur, JP., "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102, DOI 10.17487/RFC7102, January 2014, <https://www.rfc-editor.org/info/rfc7102>.

[RFC7228] Bormann, C., Ersue, M., and A. Keranen, "Terminology for Constrained-Node Networks", RFC 7228, DOI 10.17487/RFC7228, May 2014, <https://www.rfc-editor.org/info/rfc7228>.

Appendix A. Subset of a 6LoWPAN Glossary

This document often uses the followng acronyms:

6BBR: 6LoWPAN Backbone Router (proxy for the registration)
6LBR: 6LoWPAN Border Router (authoritative on DAD)
6LN: 6LoWPAN Node
6LR: 6LoWPAN Router (relay to the registration process)
6CIO: Capability Indication Option
(E)ARO: (Extended) Address Registration Option
DAD: Duplicate Address Detection
LLN: Low Power Lossy Network (a typical IoT network)
NCE: Neighbor Cache Entry
RUID: Registration Unique ID
TID: Transaction ID (a sequence counter in the EARO)

Author’s Address

Pascal Thubert (editor)
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
MOUGINS - Sophia Antipolis 06254
FRANCE

Phone: +33 497 23 26 34
Email: pthubert@cisco.com

6lo
Internet-Draft
Intended status: Informational
Expires: January 17, 2019

T. Watteyne, Ed.
Analog Devices
C. Bormann
Universitaet Bremen TZI
P. Thubert
Cisco

July 16, 2018

LLN Minimal Fragment Forwarding
draft-watteyne-6lo-minimal-fragment-02

Abstract

This document gives an overview of LLN Minimal Fragment Forwarding. When employing adaptation layer fragmentation in 6LoWPAN, it may be beneficial for a forwarder not to have to reassemble each packet in its entirety before forwarding it. This has been always possible with the original fragmentation design of RFC4944. This document is a companion document to [I-D.ietf-lwig-6lowpan-virtual-reassembly], which details the virtual Reassembly Buffer (VRB) implementation technique which reduces the latency and increases end-to-end reliability in route-over forwarding.

Table of Contents

1. Overview of 6LoWPAN Fragmentation

6LoWPAN fragmentation is defined in [RFC4944]. Although [RFC6282] updates [RFC4944], it does not redefine 6LoWPAN fragmentation.

We use Figure 1 to illustrate 6LoWPAN fragmentation. We assume node A forwards a packet to node B, possibly as part of a multi-hop route between IPv6 source and destination nodes which are neither A nor B.

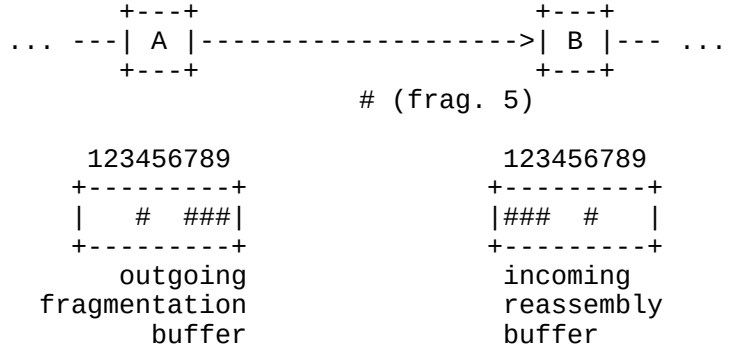


Figure 1: Fragmentation at node A, reassembly at node B.

Node A starts by compacting the IPv6 packet using header compression defined in [RFC6282]. If the resulting 6LoWPAN packet does not fit into a single link-layer frame, node A's 6LoWPAN sublayer cuts it

into multiple 6LoWPAN fragments, which it transmits as separate link-layer frames to node B. Node B's 6LoWPAN sublayer reassembles these fragments, inflates the compressed header fields back to the original IPv6 header, and hands over the full IPv6 packet to its IPv6 layer.

In Figure 1, a packet forwarded by node A to node B is cut into nine fragments, numbered 1 to 9. Each fragment is represented by the '#' symbol. Node A has sent fragments 1, 2, 3, 5, 6 to node B. Node B has received fragments 1, 2, 3, 6 from node A. Fragment 5 is still being transmitted at the link layer from node A to node B.

A reassembly buffer for 6LoWPAN contains:

- o datagram_size,
- o datagram_tag and link-layer sender and receiver addresses (to which the datagram_tag is local),
- o actual packet data from the fragments received so far, in a form that makes it possible to detect when the whole packet has been received and can be processed or forwarded,
- o a timer that allows discarding the partial packet after a timeout.

A fragmentation header is added to each fragment; it indicates what portion of the packet that fragment corresponds to. Section 5.3 of [RFC4944] defines the format of the header for the first and subsequent fragments. All fragments are tagged with a 16-bit "datagram_tag", used to identify which packet each fragment belongs to. Each fragment can be uniquely identified by the source and destination link-layer addresses of the frame that carries it, and the datagram_tag. The value of the datagram_tag only needs to be locally unique to nodes A and B.

Node B's typical behavior, per [RFC4944], is as follows. Upon receiving a fragment from node A with a datagram_tag previously unseen from node A, node B allocates a buffer large enough to hold the entire packet. The length of the packet is indicated in each fragment (the datagram_size field), so node B can allocate the buffer even if the first fragment it receives is not fragment 1. As fragments come in, node B fills the buffer. When all fragments have been received, node B inflates the compressed header fields into an IPv6 header, and hands the resulting IPv6 packet to the IPv6 layer.

This behavior typically results in per-hop fragmentation and reassembly. That is, the packet is fully reassembled, then (re)fragmented, at every hop.

2. Limits of Per-Hop Fragmentation and Reassembly

There are at least 2 limits to doing per-hop fragmentation and reassembly:

2.1. Latency

When reassembling, a node needs to wait for all the fragments to be received before being able to generate the IPv6 packet, and possibly forward it to the next hop. This repeats at every hop.

This may result in increased end-to-end latency compared to the case where each fragment would be forwarded without per-hop reassembly.

2.2. Memory Management and Reliability

Constrained nodes have limited memory. Assuming 1 kB reassembly buffers, typical nodes only have enough memory for 1-3 reassembly buffers.

Assuming the topology from Figure 2, where nodes A, B, C and D all send packets through node E. We further assume that node E's memory can only hold 3 reassembly buffers.

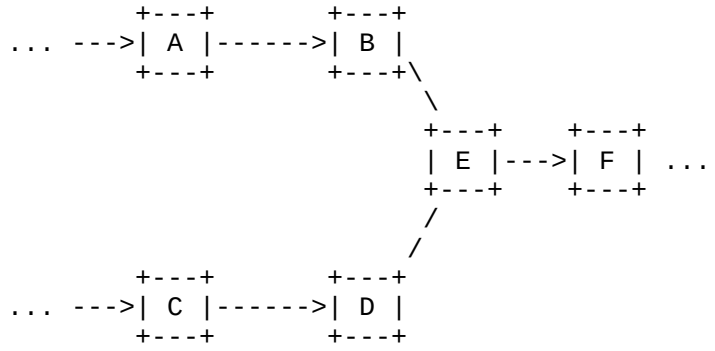


Figure 2: Illustrating the Memory Management Issue.

When nodes A, B and C concurrently send fragmented packets, all 3 reassembly buffers in node E are occupied. If, at that moment, node D also sends a fragmented packet, node E has no option but to drop one of the packets, lowering end-to-end reliability.

3. Virtual Reassembly Buffer (VRB) Implementation

Virtual Reassembly Buffer (VRB) is the implementation technique described in [I-D.ietf-lwig-6lowpan-virtual-reassembly] in which a forwarder does not reassemble each packet in its entirety before forwarding it.

VRB overcomes the limits listed in Section 2. Nodes don't wait for the last fragment before forwarding, reducing end-to-end latency. Similarly, the memory footprint of VRB is just the VRB table, reducing the packet drop probability significantly.

There are, however, limits:

Non-zero Packet Drop Probability: Each VRB table entry can be 12 B (assuming 16-bit link-layer addresses). This is a footprint 2 orders of magnitude smaller compared to needing a 1280-byte reassembly buffer for each packet. Yet, the size of the VRB table necessarily remains finite. In the extreme case where a node is required to concurrently forward more packets that it has entries in its VRB table, packets are dropped.

No Fragment Recovery: There is no mechanism in VRB for the node that reassembles a packet to request a single missing fragment. Dropping a fragment requires the whole packet to be resent. This causes unnecessary traffic, as fragments are forwarded even when the destination node can never construct the original IPv6 packet.

No Per-Fragment Routing: All subsequent fragments follow the same sequence of hops from the source to the destination node as fragment 1.

The severity and occurrence of these limits depends on the link-layer used. Whether these limits are acceptable depends entirely on the requirements the application places on the network.

If the limits are both present and not accepted by the application, future specifications may define new protocols to overcome these limits. One example is [I-D.thubert-6lo-fragment-recovery] which defines a protocol which allows fragment recovery.

4. Security Considerations

An attacker can perform a DoS attack on a node implementing VRB by generating a large number of bogus "fragment 1" fragments without sending subsequent fragments. This causes the VRB table to fill up.

Secure joining and the link-layer security that it sets up protects against those attacks from network outsiders.

5. IANA Considerations

No requests to IANA are made by this document.

6. Acknowledgments

The authors would like to thank Yasuyuki Tanaka for his in-depth review of this document.

7. Informative References

[BOOK] Shelby, Z. and C. Bormann, "6LoWPAN", John Wiley & Sons, Ltd monograph, DOI 10.1002/9780470686218, November 2009.

[I-D.ietf-lwig-6lowpan-virtual-reassembly] Bormann, C. and T. Watteyne, "Virtual reassembly buffers in 6LoWPAN", draft-ietf-lwig-6lowpan-virtual-reassembly-00 (work in progress), July 2018.

[I-D.thubert-6lo-fragment-recovery] Thubert, P., "6LoWPAN Selective Fragment Recovery", draft-thubert-6lo-fragment-recovery-01 (work in progress), June 2018.

[RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <https://www.rfc-editor.org/info/rfc4944>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

Authors' Addresses

Thomas Watteyne (editor)
Analog Devices
32990 Alvarado-Niles Road, Suite 910
Union City, CA 94587
USA

Email: thomas.watteyne@analog.com

Carsten Bormann
Universitaet Bremen TZI
Postfach 330440
Bremen D-28359
Germany

Email: cabo@tzi.org

Pascal Thubert
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
MOUGINS - Sophia Antipolis 06254
France

Email: pthubert@cisco.com