Network Working Group
Internet-Draft
Intended status: Informational
Expires: February 13, 2015

T. Clausen
U. Herberg
Fujitsu Laboratories of America
J. Yi
LIX, Ecole Polytechnique

August 12, 2014

Security Threats for the Optimized Link State Routing Protocol version 2 (OLSRv2)
draft-clausen-manet-olsrv2-sec-threats-01

Abstract

This document analyzes common security threats of the Optimized Link State Routing Protocol version 2 (OLSRv2) and describes their potential impacts on Mobile Ad Hoc Network (MANET) operations. It then analyzes which of these security vulnerabilities can be mitigated when using the mandatory-to-implement security mechanisms for OLSRv2, and how the vulnerabilities are mitigated.

Table of Contents

1. Introduction

The Optimized Link State Routing Protocol version 2 (OLSRv2) [RFC5148], [RFC5444], [RFC5497], [RFC6130], [RFC7181], [RFC7182], [RFC7183], [RFC7187], [RFC7188] is a successor to OLSR [RFC3626] as a routing protocol for MANETs (Mobile Ad hoc NETworks). OLSRv2 retains the same basic algorithms as its predecessor, however offers various improvements, e.g., a modular and flexible architecture allowing extensions, such as for security, to be developed as add-ons to the basic protocol.

The developments reflected in OLSRv2 have been motivated by increased real-world deployment experiences, e.g., from networks such as FunkFeuer [FUNKFEUER], and the requirements presented for continued successful operation of these networks. With participation in such networks increasing (the FunkFeuer community network has, e.g., roughly 400 individual participants), operating with the assumption, that participants can be "trusted" to behave in a non-destructive way, is utopia. Taking the Internet as an example, as participation in the network increases and becomes more diverse, more efforts are required to preserve the integrity and operation of the network. Most SMTP-servers were, e.g., initially available for use by everyone on the Internet - with an increased populace on the Internet, attacks and abuses caused the recommended practice is today to require authentication and accounting for users of such SMTP servers [RFC5068].

As OLSRv2 often is used in wireless environments, it is potentially exposed to different kinds of security threats, some of which are of particular significance as compared to wired networks. As radio signals can be received as well as transmitted by any compatible wireless device within radio range, there is commonly no physical protection as otherwise known for wired networks.

A first step towards hardening against attacks disrupting the connectivity of a network, is to understand the vulnerabilities of routing protocol, managing the connectivity. This document therefore analyzes OLSRv2, to understand its inherent vulnerabilities and resiliences. The authors do not claim completeness of the analysis, but hope that the identified attacks, as presented, form a meaningful starting-point for developing secured OLSRv2 networks.

This document first describes security vulnerabilities to OLSRv2 when it is used without the mandatory-to-implement security mechanisms specified in Section 23.5 of [RFC7181]. It then analyzes which of these security vulnerabilities can be mitigated when using the mandatory-to-implement security mechanisms for OLSRv2, and how the vulnerabilities are mitigated. This separation is important since

other security mechanisms than the mandatory-to-implement ones may be used in a deployment, as stated in [RFC7181]:

> "Any deployment of OLSRv2 SHOULD use the security mechanism specified in [RFC7183] but MAY use another mechanism if more appropriate in an OLSRv2 deployment. For example, for longer-term OLSRv2 deployments, alternative security mechanisms (e.g., rekeying) SHOULD be considered."

Moreover, this document is also based on the assumption that no additional security mechanism such as IPsec is used in the IP layer or other mechanisms on lower layers, as not all MANET deployments may be able to accommodate such common protection mechanisms (e.g., because of limited resources of MANET routers).

The threats related to NHDP (Neighborhood Discovery Protocol) have been discussed in [RFC7186]. As NHDP is a fundamental block of OLSRv2, the vulnerabilities of NHDP apply also to OLSRv2.

It should be noted that many OLSRv2 implementations are configurable, and so an attack on the configuration system (such as [RFC6779] and [RFC7184]) can be used to adversely affect the operation of an NHDP implementation.

The NHDP and OLSRv2 MIB modules [RFC6779] and [RFC7184] might help monitoring some of the security attacks mentioned in this document. [MGMT-SNAP] provides a snapshot of OLSRv2-routed MANET management as currently deployed.

1.1. OLSRv2 Overview

OLSRv2 contains three basic processes: Neighborhood Discovery, MPR Flooding and Link State Advertisements, described in the below with sufficient details for elaborating the analyses in this document.

1.1.1. Neighborhood Discovery

Neighborhood Discovery is the process, whereby each router discovers the routers which are in direct communication range of itself (1-hop neighbors), and detects with which of these it can establish bi-directional communication. Each router sends HELLO messages periodically, listing the identifiers of all the routers from which it has recently received a HELLO message, as well as the "status" of the link (heard or verified bi-directional). A router a receiving a HELLO message from a neighbor b, in which b indicates to have recently received a HELLO message from a, considers the link a-b to be bi-directional. As b lists identifiers of all its neighbors in its HELLO message, a learns the "neighbors of its neighbors" (2-hop

neighbors) through this process. HELLO messages are sent periodically, however certain events may trigger non-periodic HELLOs. OLSRv2 [RFC7181] uses NHDP [RFC6130] as its neighborhood discovery mechanism. The vulnerabilities of NHDP are analyzed in [RFC7186].

1.1.2. MPR Flooding

Multi Point Relay (MPR) Flooding is the process whereby each router is able to, efficiently, conduct network-wide broadcasts. Each router designates, from among its bi-directional neighbors, a subset (MPR set) such that a multicast message transmitted by the router and relayed by the MPR set can received by all its 2-hop neighbors. MPR selection is encoded in outgoing HELLO messages.

Routers may express, in their HELLO messages, their "willingness" (integer between 1 "will never" and 7 "will always") to be selected as MPR, which is taken into consideration for the MPR calculation, and which is useful for example when an OLSRv2 network is "planned". The set of routers having selected a given router as MPR is the MPR-selector-set of that router. A study of the MPR flooding algorithm can be found in [MPR-FLOODING].

1.1.3. Link State Advertisement

Link State Advertisement is the process whereby routers are determining which link state information to advertise through the network. Each router must advertise, at least, all links between itself and its MPR selectors, in order to allow all routers to calculate shortest paths. Such link state advertisements are carried in Topology Control (TC) messages, broadcast through the network using the MPR flooding process described above. As a router selects MPRs only from among bi-directional neighbors, links advertised in TC are also bi-directional and routing paths calculated by OLSRv2 contain only bi-directional links. TCs are sent periodically, however certain events may trigger non-periodic TCs.

1.2. Link State Vulnerability Taxonomy

Proper functioning of OLSRv2 assumes that (i) each router signals its presence in the network and the topology information that it obtained honestly, (ii) each router can acquire and maintain a topology map, accurately reflecting the effective network topology; and (iii) that the network converges, i.e., that all routers in the network will have sufficiently identical topology maps.

An OLSRv2 network can be disrupted by breaking either of these assumptions, specifically (a) routers may be prevented from acquiring a topology map of the network; (b) routers may acquire a topology map

that does not reflect the effective network topology; and (c) two or more routers may acquire inconsistent topology maps.

1.3. OLSRv2 Attack Vectors

Besides "radio jamming", attacks on OLSRv2 consist of a compromised OLSRv2 router injecting "correctly looking, but invalid, control traffic" (TCs, HELLOs) into the network. A compromised OLSRv2 router can either (a) lie about itself (its identification, its willingness to serve as MPR), henceforth Identity Spoofing, or (b) lie about its relationship to other routers (pretend existence of links to other routers), henceforth Link Spoofing. Such attacks may disrupt the the Link State Advertisement process, through targeting the MPR Flooding mechanism, or by causing incorrect link state information to be included in TCs, causing routers to have incomplete, inaccurate or inconsistent topology maps. In a different class of attacks, a compromised OLSRv2 router injects control traffic, designed so as to cause an in-router resource exhaustion, e.g., by causing the algorithms calculating routing tables or MPR sets to be invoked continuously, preventing the internal state of a router from converging.

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].

This document uses the terminology and notation defined in [RFC5444], [RFC6130] and [RFC7181]. Additionally, it defines the following terminology:

Compromised OLSRv2 router: - An attacker that is present in the network and generates syntactically correct OLSRv2 control messages. Control messages emitted by a compromised OLSRv2 router may contain additional information, or omit information, as compared to a control message generated by a non-compromised OLSRv2 router located in the same topological position in the network.

Legitimate OLSRv2 router: - An OLSRv2 router that is not a compromised OLSRv2 router.

3. Topology Map Acquisition

Topology Map Acquisition relates to the ability for any given router in the network to acquire a representation of the network connectivity. A router, unable to acquire a topology map, is incapable of calculating routing paths and participating in forwarding data. Topology map acquisition can be hindered by (i) TCs to not being delivered to (all) routers in the network, such as what happens in case of Flooding Disruption, or (ii) in case of "jamming" of the communication channel.

The jamming and flooding disruption due to identity spoofing and link spoofing have been discussed in [RFC7186].

3.1. Attack on Jittering

OLSRv2 incorporates a jittering: a random, but bounded, delay on outgoing control traffic [RFC5148]. This may be necessary when link layers (such as 802.11 [IEEE802.11]) are used that do not guarantee collision-free delivery of frames, and where jitter can reduce the probability of collisions of frames on lower layers.

In OLSRv2, TC forwarding is jittered by a value between 0 and MAX_JITTER. In order to reduce the number of transmissions, when a control message is due for transmission, OLSRv2 piggybags all queued messages into a single transmission. Thus, if a compromised OLSRv2 router sends many TCs within a very short time interval, the jitter time of the attacked router tends to 0. This renders jittering ineffective and can lead to collisions on the link layer.

In addition to causing more collisions, forwarding a TC with little or no jittering can make sure that the TC message forwarded by a compromised router arrives before the message forwarded by legitimate routers. The compromised router can thus inject malicious content in the TC, and the legitimate message will be discarded as duplicate message. This preemptive action is important for some of the attacks introduced in the following sections.

3.2. Hop-count and Hop-limit Attacks

The hop-count and hop-limit fields are the only parts of a TC that are modified when forwarding. A compromised OLSRv2 router can modify either of these when forwarding TCs.

3.2.1. Modifying the Hop Limit

A compromised OLSRv2 router can decrease the hop limit when forwarding a TC. This will reduce the scope of forwarding the

message, and may lead to some routers in the network not receiving that TC. Note that this is not necessarily the same as not relaying the message (i.e., setting the hop limit to 0), as illustrated in Figure 1.

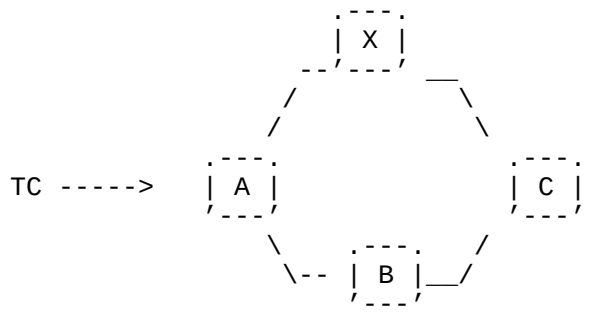


Figure 1: Hop Limit Attack.

A TC arrives at and is forwarded by router A, such that it is received by both B and the malicious X. X can forward the TC without any delay (including without jitter) such that its transmissions arrives before that of B at C. Before forwarding, it significantly reduces the hop limit of the message. Router C receives the TC, processes (and forwards) it, and marks it as already received - causing it to discard further copies received from B. Thus, if the TC is forwarded by C, it has a very low hop limit and will not reach the whole network.

3.2.2. Modifying the Hop Count

A compromised OLSRv2 router can modify the hop count when forwarding a TC. This may have two consequences: (i) if the hop count is set to the maximum value, then the TC will be forwarded no further by, or (ii) artificially manipulating the hop count may affect the validity time as calculated by recipients, when using distance-dependent validity times as defined in [RFC5497] (e.g., as part of a fish-eye extension to OLSR2 [OLSR-FSR]).

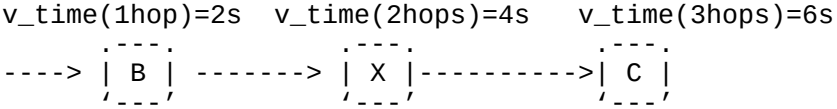


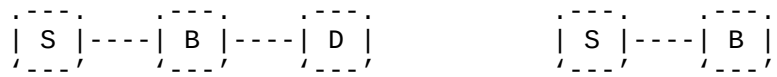
Figure 2: Different validity times based on the distance in hops.

In Figure 2, router A sends a TC with a validity time of two seconds for neighbors that are one hop away, four seconds for routers in a two-hop distance and six seconds in a three-hop distance. If X is a compromised OLSRv2 router and modifies the hop count (say, by decreasing it to 0), then C will calculate the validity time of received information to two seconds - after which it expires unless refreshed. If TCs from A are sent less frequently than that up to 3 hops, this causes links advertised in such TCs to be only intermittently available to C.

4. Effective Topology

Link-state protocols assume that each router can acquire an accurate topology map, reflecting the effective network topology. This implies that the routing protocol, through its message exchange, identifies a path from a source to a destination, and this path is valid for forwarding data traffic. If an attacker disturbs the correct protocol behavior, the perceived topology map of a router can permanently differ from the effective topology.

Considering the example in Figure 3(a), which illustrates the topology map as acquired by router S. This topology map indicates that the routing protocol has identified that for S, a path exists to D via B, which it therefore assumes can be used for transmitting data. If, effectively, B does not forward data traffic from S, then the topology map in S does not accurately reflect the effective network topology. Rather, the effective network topology from the point of view of S would be as indicated in Figure 3(b): D is not part of the network reachable from router S.



(a) (b)

Figure 3: Incorrect Data Traffic Forwarding.

Some of the attacks related to NHDP, such as message timing attack, indirect channel overloading have been discussed in [RFC7186]. Other threats specific to OLSRv2 are further detailed in this section.

4.1. Incorrect Forwarding

OLSRv2 routers exchange information using link-local transmissions (link-local multicast or limited broadcast) for their control messages, with the routing process in each router retransmitting received messages destined for network-wide diffusion. Thus, if the operating system in a router is not configured to enable forwarding, this will not affect the operating of the routing protocol, or the topology map acquired by the routing protocol. It will, however, cause a discrepancy between the effective topology and the topology map, as indicated in Figure 3(a) and Figure 3(b).

This situation is not hypothetical. A common error seen when deploying OLSRv2-based networks using Linux-based computers as router is to neglect enabling IP forwarding, which effectively becomes an accidental attack of this type.

4.2. Wormholes

A wormhole, depicted in the example in Figure 4, may be established between two collaborating devices, connected by an out-of-band channel. These devices send traffic through the "tunnel" to their alter-ego, which "replays" the traffic. Thus, routers D and S appear as if direct neighbors and reachable from each other in 1 hop through the tunnel, with the path through the MANET being 100 hops long.

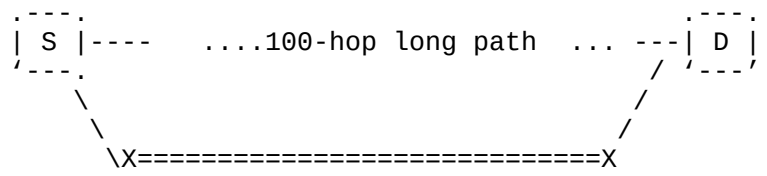


Figure 4: Wormholing between two collaborating devices not participating in the routing protocol.

The consequences of such a wormhole in the network depends on the detailed behavior of the wormhole. If the wormhole relays only control traffic, but not data traffic, the same considerations as in Section 4.1 applies. If, however, the wormhole relays all traffic, control and data alike, it is connectivity-wise identical to a usable link - and the routing protocol will correctly generate a topology

map reflecting the effective network topology. The efficiency of the topology so obtained depends on (i) the wormhole characteristics, (ii) how the wormhole presents itself, and (iii) how paths are calculated.

Assuming that paths are calculated with unit-cost for all links, including the "link" presented by the wormhole: if the real characteristics of the wormhole are as-if it was a path of more than 100 hops (e.g., with respect to delay, bandwidth, etc.), then the presence of the wormhole results in a degradation in performance as compared to using the non-wormhole path. Conversely, if the "link" presented by the wormhole has better characteristics, the wormhole results in improved performance.

If paths are calculated using non-unit-costs for all links, and if the cost of the "link" presented by the wormhole correctly represents the actual cost (e.g., if the cost is established through measurements across the wormhole), then the wormhole may in the worst case cause no degradation in performance, in the best case improve performance by offering a better path. If the cost of the "link" presented by the wormhole is misrepresented, then the same considerations as for unit-cost links apply.

An additional consideration with regards to wormholes is, that such may present topologically attractive paths for the network - however it may be undesirable to have data traffic transit such a path: an attacker could, by virtue of introducing a wormhole, acquire the ability to record and inspect transiting data traffic.

4.3. Sequence Number Attacks

OLSRv2 uses two different sequence numbers in TCs, to (i) avoid processing and forwarding the same message more than once (Message Sequence Number), and (ii) to ensure that old information, arriving late due to, e.g., long paths or other delays, is not allowed to overwrite fresher information (Advertised Neighbor Sequence Number - ANSN).

4.3.1. Message Sequence Number

An attack may consist of a compromised OLSRv2 router spoofing the identity of another router in the network, and transmitting a large number of TCs, each with different Message Sequence Numbers. Subsequent TCs with the same sequence numbers, originating from the router whose identity was spoofed, would hence be ignored, until eventually information concerning these "spoofed" TCs expires.

4.3.2. Advertised Neighbor Sequence Number (ANSN)

An attack may consist of a compromised OLSRv2 router spoofing the identity of another router in the network, and transmitting a single TC, with an ANSN significantly larger than that which was last used by the legitimate router. Routers will retain this larger ANSN as "the most fresh information" and discard subsequent TCs with lower sequence numbers as being "old".

4.4. Indirect Jamming

Indirect Jamming is an attack in which a compromised OLSRv2 router is, by its actions, causing legitimate routers to generate inordinate amounts of control traffic, thereby increasing both channel occupation and the overhead incurred in each router for processing this control traffic. This control traffic will be originated from legitimate routers, thus to the wider network, the malicious device may remain undetected.

The general mechanism whereby a malicious router can cause indirect jamming is for it to participate in the protocol by generating plausible control traffic, and to tune this control traffic to in turn trigger receiving routers to generate additional traffic. For OLSRv2, such an indirect attack can be directed at, respectively, the Neighborhood Discovery mechanism and the Link State Advertisement mechanism.

The most efficient Indirect Jamming attack in OLSRv2 is to target control traffic, destined for network-wide diffusion. This is illustrated in Figure 5.

The malicious router X selects router A as MPR at time t0 in a HELLO. This causes X to appear as MPR selector for A and, consequently, A sets X to be advertised in its "Neighbor Set" and increments the associated "Advertised Neighbor Sequence Number" (ANSN). A must, then, advertise the link between itself and X in subsequent outgoing TCs (t1), also including the ANSN in such TCs. Upon X having received this TC, it declares the link between itself and A as no longer valid (t2) in a HELLO (indicating the link to a as LOST). Since only symmetric links are advertised by OLSRv2 routers, A will upon receipt hereof remove X from the set of advertised neighbors and increment the ANSN. A will then in subsequent TCs advertise the remaining set of advertised neighbors (i.e., with X removed) and the corresponding ANSN (t3). Upon X having received this information in another TC from A, it may repeat this cycle, alternating advertising the link A-X as "LOST" and as "MPR".

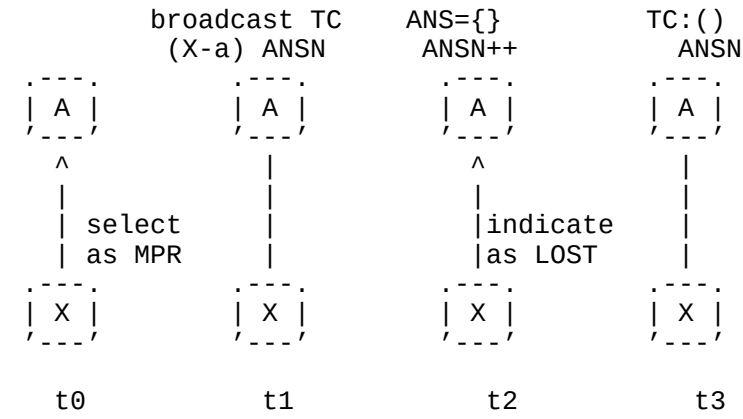


Figure 5: Indirect Jamming in Link State Advertisement: the malicious X flips between link status MPR and LOST.

Routers receiving a TC will parse and process this message, specifically updating their topology map as a consequence of successful receipt. If the ANSN between two successive TCs from the same router has incremented, then the topology has changed and routing tables are to be recalculated. This is a potentially computationally costly operation [DSP-OLSRv2].

A compromised OLSRv2 router may chose to conduct this attack against all its neighbors, thus attaining maximum disruptive impact on the network with relatively little overhead of its own: other than participating in the Neighborhood Discovery procedure, the compromised OLSRv2 router will monitor TCs generated by its neighbors and alternate the advertised status for each such neighbor, between "MPR" and "LOST". The compromised OLSRv2 router will indicate its willingness to be zero (thus, avoid being selected as MPR) and may ignore all other protocol operations, while still remaining effective as an attacker.

The basic operation of OLSRv2 employs periodic message emissions, and by this attack it can be ensured that each such periodic message will entail routing table recalculation in all routers in the network.

If the routers in the network have "triggered TCs" enabled, this attack may also cause an increased TC frequency. Triggered TCs are intended to allow a (stable) network to have relatively low TC emission frequencies, yet still allow link breakage or link emergence to be advertised through the network rapidly. A minimum message interval (typically much smaller than the regular periodic message interval) is imposed, to rate-limit worst-case message emissions. This attack can cause the TC interval to, permanently, become equal

to the minimum message interval. [RFC7181] proposes as default that the minimum TC interval be 0.25 x TC interval.

Indirect Jamming by a compromised OLSRv2 router can thus have two effects: it may cause increased frequency of TC generation and transmission, and it will cause additional routing table recalculation in all routers in the network.

5. Inconsistent Topology

Inconsistent topology maps can occur by a compromised OLSRv2 router employing either of identity spoofing or link spoofing for conducting an attack against an OLSRv2 network. The threats related to NHDP, such as identity spoofing in NHDP, link spoofing in NHDP and creating loops have been illustrated in [RFC7186]. This section mainly addresses the vulnerabilities in [RFC7181].

5.1. Identity Spoofing

Identity spoofing can be employed by a compromised OLSRv2 router via the Neighborhood Discovery process and via the Link State Advertisement process. Either of them causes inconsistent topology maps in routers in the network. The inconsistent topology maps due to neighborhood discovery has been discussed in [RFC7186]. For OLSRv2, the attack on link state advertisements can also cause inconsistent topology maps.

An inconsistent topology map may occur when the compromised OLSRv2 router takes part in the Link State Advertisement (LSA) procedure, by selecting a neighbor as MPR, which in turn advertises the spoofed identities of the compromised OLSRv2 router. This attack will alter the topology maps all routers of the network.

Figure 6: Identity Spoofing: compromised OLSRv2 router X spoofs the identity of A, leading to a wrongly perceived topology.

In Figure 6, router X spoofs the address of router A. If X selects F as MPR, all routers in the network will be informed about the link F-A by the TCs originating from F. Assuming that (the real) A selects B as MPR, the link B-A will also be advertised in the network.

When calculating paths, B and C will calculate paths to A via B, as illustrated in Figure 7(a); for these routers, the shortest path to A is via B. E and F will calculate paths to A via F, as illustrated in Figure 7(b); for these routers, the shortest path to A is via the compromised OLSRv2 router X, and these are thus disconnected from the real A. D will have a choice: the path calculated to A via B is of the same length as the path via the compromised OLSRv2 router X, as illustrated in Figure 7(b).

In general, the following observations can be made:

- The network will be split in two, with those routers closer to B than to X reaching A, whereas those routers closer to X than to B will be unable to reach A.

- Routers beyond B, i.e., routers beyond one hop away from A will be unable to detect this identity spoofing.

The identity spoofing attack via the Link State Advertisement procedure has a higher impact than the attack on the neighborhood discovery procedure, since it alters the topology maps of all routers in the network, and not only in the 2-hop neighborhood. However, the attack is easier to detect by other routers in the network. Since the compromised OLSRv2 router is advertised in the whole network, routers whose identities are spoofed by the compromised OLSRv2 router can detect the attack. For example, when a receives a TC from F advertising the link F-A, it can deduce that some entity is injecting incorrect Link State information as it does not have F as one of its direct neighbors.

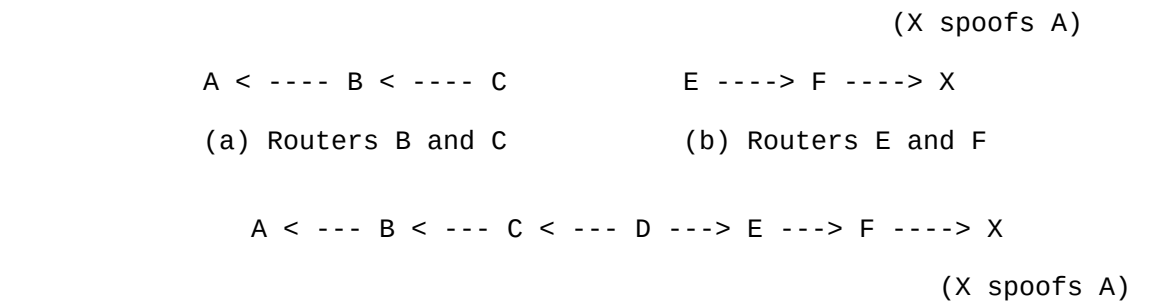


Figure 7: Routing paths towards A, as calculated by the different routers in the network in presence of a compromised OLSRv2 router X, spoofing the address of A.

As the compromised OLSRv2 router X does not itself send the TCs, but rather, by virtue of MPR selection, ensures that the addresses it

spoofs are advertised in TCs from its MPR selector F, the attack may be difficult to counter: simply ignoring TCs that originate from F may also suppress the link state information for other, legitimate, MPR selectors of F.

Identity spoofing by a compromised OLSRv2 router, participating in the Link State Advertisement process by selecting MPRs only, thus, creates a situation wherein two or more routers have substantially inconsistent topology maps: traffic for an identified destination is, depending on where in the network it appears, delivered to different routers.

5.2. Link Spoofing

Link Spoofing is a situation in which a router advertises non-existing links to another router (possibly not present in the network). Essentially, TCs and HELLOs both advertise links to direct neighbor routers, with the difference being the scope of the advertisement. Thus, link spoofing consists of a compromised OLSRv2 router, reporting that it has as as neighbors routers which are, either, not present in the network, or which are effectively not neighbors of the compromised OLSRv2 router.

It can be noted that a situation similar to Link Spoofing may occur temporarily in an OLSR or OLSRv2 network without compromised OLSRv2 routers: if a was, but is no more, a neighbor of B, then A may still be advertising a link to B for the duration of the time it takes for the the Neighborhood Discovery process to determine this changed neighborhood.

In the context of this document, Link Spoofing refers to a persistent situation where a compromised OLSRv2 router intentionally advertises links to other routers, for which it is not a direct neighbor.

5.2.1. Inconsistent Topology Maps due to Link State Advertisements

Figure 8 illustrates a network, in which the compromised OLSRv2 router X spoofs links to the existing router A by participating in the Link State Advertisement process and including this non-existing link in its advertisements.

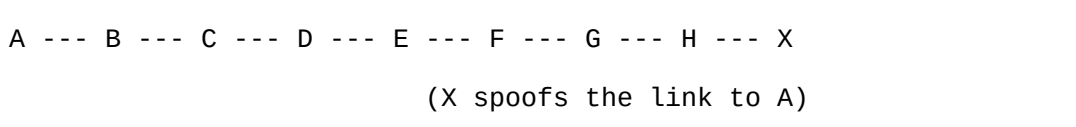


Figure 8: Link Spoofing: The compromised OLSRv2 router X advertises a

spoofed link to A in its TCs, thus all routers will record both of the links X-A and B-A.

As TCs are flooded through the network, all routers will receive and record information describing a link X-A in this link state information. If A has selected router B as MPR, A will likewise flood this link state information through the network, thus all routers will receive and record information describing a link B-A.

When calculating routing paths, B, C and D will calculate paths to A via B, as illustrated in Figure 9(a); for these routers, the shortest path to A is via B. F and G will calculate paths to A via X, as illustrated in Figure 9(b); for these routers, the shortest path to A is via X, and these are thus disconnected from the real router A. E will have a choice: the path calculated to A via B is of the same length as the path via X, as illustrated in Figure 9(b).

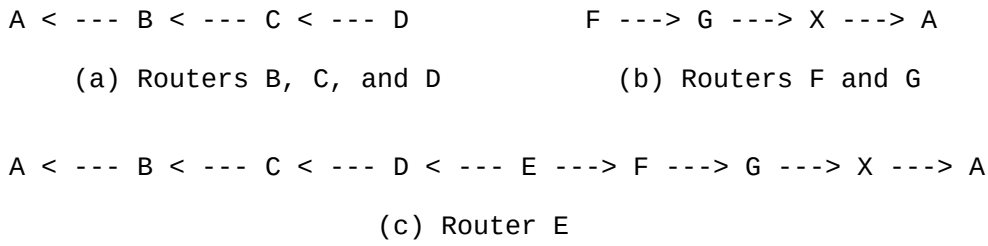


Figure 9: Routing paths towards router A, as calculated by the different routers in the network in presence of a compromised OLSRv2 router X, spoofing a link to router A.

In general, the following observations can be made:

- The network will be separated in two, with those routers closer to B than to X reaching A, whereas those routers closer to X than to B unable to reach A.
- Routers beyond B, i.e., routers beyond one hop away from A will be unable to detect this link spoofing.

The impact of this attack is similar to that presented in Section 5.2.1, however, is easier to detect as the compromised OLSRv2 router is generating control traffic reaching the entire network.

6. Mitigation of Security Vulnerabilities for OLSRv2

As described in Section 1, [RFC7183] specifies a security mechanism for OLSRv2 that is mandatory to implement. However, deployments may choose to use different security mechanisms if more appropriate. Therefore, it is important to understand both the inherent resilience of OLSRv2 against security vulnerabilities when not using the mechanisms specified in [RFC7183], as well as the protection that [RFC7183] provides when used in a deployment.

6.1. Inherent OLSRv2 Resilience

OLSRv2 (without the mandatory-to-implement security mechanisms in [RFC7183]) provides some inherent resilience against part of the attacks described in this document. In particular, it provides the following resilience:

- Sequence numbers: OLSRv2 employs message sequence numbers, specific per router identity and message type. Routers keep an "information freshness" number (ANSN), incremented each time the content of a Link State Advertisement from a router changes. This allows rejecting "old" information and duplicate messages, and provides some protection against "message replay". This, however, also presents an attack vector (Section 4.3).

- Ignoring uni-directional links: The Neighborhood Discovery process detects and admits only bi-directional links for use in MPR selection and Link State Advertisement. Jamming attacks may affect only reception of control traffic, however OLSRv2 will correctly recognize, and ignore, such a link as not bi-directional.

- Message interval bounds: The frequency of control messages, with minimum intervals imposed for HELLO and TCs. This may limit the impact from an indirect jamming attack (Section 4.4).

- Additional reasons for rejecting control messages: The OLSRv2 specification includes a list of reasons, for which an incoming control message should be rejected as malformed - and allows that a protocol extension may recognize additional reasons for OLSRv2 to consider a message malformed. This allows - together with the flexible message format [RFC5444] - addition of security mechanisms, such as digital signatures, while remaining compliant with the OLSRv2 standard specification.

6.2. Resilience by using RFC7183 with OLSRv2

[RFC7183] specifies mechanisms for integrity and replay protection for NHDP and OLSRv2, using the generalized packet/message format described in [RFC5444] and the TLV definitions in [RFC7182]. The specification describes how to add an Integrity Check Value (ICV) in a TLV to each control message, providing a digital signature of the content of the message using HMAC/SHA-256. In addition, a timestamp TLV is added to the message prior to creating the ICV, enabling replay protection of messages. The document specifies how to sign outgoing messages and how to verify incoming messages, as well as under which circumstances a non-valid message is rejected. Because of the HMAC/SHA-256 ICV, a shared key between all routers in the MANET is assumed. A router without valid credentials is not able to create an ICV that can be correctly verified by other routers in the MANET; therefore, such an incorrectly signed message will be rejected by other MANET routers, and the router cannot participate in the OLSRv2 routing process (i.e., the malicious router will be ignored by other, legitimate routers). [RFC7183] does not address the case where a router with valid credentials has been compromised. Such a compromised router will not be excluded from the routing process, and other means of detecting such a router are necessary if required in a deployment (in addition to using asymmetric keys, allowing to revoke access to one particular router instead of revoking the shared key used by all routers in the MANET).

In the following sections, each of the vulnerabilities described earlier in this document will be evaluated in terms of whether OLSRv2 with the mechanisms in [RFC7183] provides sufficient protection against the attack. It is implicitly assumed in each of the following sections that [RFC7183] is used with OLSRv2.

6.2.1. Topology Map Acquisition

Attack on Jittering - As only OLSRv2 routers with valid credentials can participate in the routing process, a malicious router cannot reduce the jitter time of an attacked router to 0 by sending many TC messages in a short time. The attacked router would reject all the incoming messages as "invalid" and not forward them. The same applies for the case where a malicious routers wants to assure that by forcing a zero jitter interval, the message arrives before the same message forwarded by legitimate routers.

Modifying the Hop Limit - As the hop limit is not protected by [RFC7183] (since it is a mutual field, changing at every hop), this attack is still feasible.

Modifying the Hop Count - Similarly to the hop limit, as the hop count is not protected by [RFC7183] (since it is a mutual field, changing at every hop), this attack is still feasible.

6.2.2. Effective Topology

Incorrect Forwarding - As only OLSRv2 routers with valid credentials can participate in the routing process, a malicious router will not be part of the topology of other legitimate OLSRv2 routers. Therefore, no data traffic will be sent to the malicious router for forwarding.

Wormholes - Since a wormhole consists of at least two devices forwarding (unmodified) traffic, this attack is still feasible and undetectable by the OLSRv2 routing process since the attack does not involve the OLSRv2 protocol itself (but rather lower layers). By using [RFC7183], it can at least be assured that the content of the control messages is not modified while being forwarded via the wormhole. Moreover, the timestamp TLV assures that the forwarding can only be done in a short time window after the actual TC message has been sent.

Message Sequence Number - As the message sequence number is included in the ICV calculation, OLSRv2 is protected against this attack.

Advertised Neighbor Sequence Number (ANSN) - As the ANSN is included in the ICV calculation, OLSRv2 is protected against this attack.

Indirect Jamming - Since the control messages of a malicious router will be rejected by other legitimate OLSRv2 routers in the MANET, this attack is mitigated.

6.2.3. Inconsistent Topology

Identity Spoofing - Since the control messages of a malicious router will be rejected by other legitimate OLSRv2 routers in the MANET, a router without valid credentials may spoof its identity (e.g., IP source address or message originator address), but the messages will be ignored by other routers. As [RFC7183] uses shared keys amongst all MANET routers, a single compromised router may spoof its identify and cause harm to the network stability. Removing this one malicious router once detected implies rekeying all other routers in the MANET. Asymmetric keys, in particular when using identity based signatures, such as specified in [IBS] may further allow to revoke single routers and to verify their identity based on the ICV itself.

Link Spoofing - Similar to identity spoofing, a malicious router without valid credential may spoof links, but its control messages will be rejected by other routers, thereby mitigating the attack.

Inconsistent Topology Maps due to Link State Advertisements - The same considerations as for link spoofing apply.

7. Security Considerations

This document does not specify a protocol or a procedure. The document, however, reflects on security considerations for OLSRv2, and its constituent parts, including NHDP.

8. IANA Considerations

This document has no actions for IANA.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, BCP 14, March 1997.

[RFC7181] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol Version 2", RFC 7181, April 2014.

[RFC7186] Yi, J., Herberg, U., and T. Clausen, "Security Threats for the Neighborhood Discovery Protocol (NHDP)", RFC 7186, April 2014.

9.2. Informative References

[DSP-OLSRv2] Herberg, U., "Performance Evaluation of using a Dynamic Shortest Path Algorithm in OLSRv2, Proceedings of the 8th Eighth Annual Conference on Communication Networks and Services Research", 2010.

[FUNKFEUER] "http://www.funkfeuer.at/".

[IBS] Dearlove, C., "Identity-Based Signatures for MANET Routing Protocols", work in progress draft-ietf-manet-ibs-02,

July 2014.

[IEEE802.11]
"IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Spec.", 2007.

[MGMT-SNAP]
Herberg, U. and T. Clausen, "Snapshot of OLSRv2-Routed MANET Management", work in progress draft-ietf-manet-olsrv2-management-snapshot, July 2014.

[MPR-FLOODING]
Qayyum, A., Viennot, L., and A. Laouiti, "Multipoint relaying: An efficient technique for flooding in mobile wireless networks.", 2001.

[OLSR-FSR]
Adjih, C., Baccelli, E., Clausen, T., Jacquet, P., and G. Rodolakis, "Fish eye OLSR scaling properties, IEEE Journal of Communication and Networks (JCN), Special Issue on Mobile Ad Hoc Wireless Networks", 2004.

[RFC3626] Clausen, T. and P. Jacquet, "The Optimized Link State Routing Protocol", RFC 3626, October 2003.

[RFC5068] Hutzler, C., Crocker, D., Resnick, P., Allman, E., and T. Finch, "Email Submission Operations: Access and Accountability Requirements", RFC 5068, BCP 134, October 2007.

[RFC5148] Clausen, T., Dearlove, C., and B. Adamson, "Jitter Considerations in Mobile Ad Hoc Networks (MANETs)", RFC 5148, February 2008.

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized MANET Packet/Message Format", RFC 5444, February 2009.

[RFC5497] Clausen, T. and C. Dearlove, "Representing Multi-Value Time in Mobile Ad Hoc Networks (MANETs)", RFC 5497, March 2009.

[RFC6130] Clausen, T., Dearlove, C., and J. Dean, "Mobile Ad Hoc Network (MANET) Neighborhood Discovery Protocol (NHDP)", RFC 6130, April 2011.

[RFC6779] Herberg, U., Cole, R., and I. Chakeres, "Definition of

Managed Objects for the Neighborhood Discovery Protocol", RFC 6779, May 2012.

[RFC7182] Herberg, U., Clausen, T., and C. Dearlove, "Integrity Check Value and Timestamp TLV Definitions for Mobile Ad Hoc Networks (MANETs)", RFC 7182, April 2014.

[RFC7183] Herberg, U., Dearlove, C., and T. Clausen, "Integrity Protection for the Neighborhood Discovery Protocol (NHDP) and Optimized Link State Routing Protocol Version 2 (OLSRv2)", RFC 7183, April 2014.

[RFC7184] Herberg, U., Cole, R., and T. Clausen, "Definition of Managed Objects for the Optimized Link State Routing Protocol Version 2", RFC 7184, April 2014.

[RFC7187] Dearlove, C. and T. Clausen, "Routing Multipoint Relay Optimization for the Optimized Link State Routing Protocol Version 2 (OLSRv2)", RFC 7187, April 2014.

[RFC7188] Dearlove, C. and T. Clausen, "Optimized Link State Routing Protocol Version 2 (OLSRv2) and MANET Neighborhood Discovery Protocol (NHDP) Extension TLVs", RFC 7188, April 2014.

Authors' Addresses

Thomas Clausen

Phone: +33-6-6058-9349
Email: T.Clausen@computer.org
URI: http://www.thomasclausen.org

Ulrich Herberg
Fujitsu Laboratories of America
1240 E Arques Ave
Sunnyvale CA 94086,
US

Phone:
Email: ulrich@herberg.name
URI: http://www.herberg.name

Jiazi Yi
LIX, Ecole Polytechnique
91128 Palaiseau Cedex,
France

Phone: +33 1 77 57 80 85
Email: jiazi@jiaziyi.com
URI: http://www.jiaziyi.com/

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: August 29, 2015

Y. Yi
S. Lee
University of California, Los Angeles
W. Su
The Boeing Company
M. Gerla
A. Colin de Verdiere
University of California, Los Angeles
February 25, 2015

On-Demand Multicast Routing Protocol (ODMRP) for Ad Hoc Networks
draft-gerla-manet-odmrp-05

Abstract

The On-Demand Multicast Routing Protocol (ODMRP) is a multicast routing protocol designed for ad hoc networks with mobile hosts. ODMRP is a mesh-based, rather than a conventional tree-based, multicast scheme and uses a forwarding group concept (only a subset of nodes forwards the multicast packets via scoped flooding). It applies on-demand procedures to dynamically build routes and maintain multicast group membership, without relying on pre-existing unicast routing protocols. ODMRP is well suited for ad hoc wireless networks with mobile hosts where bandwidth is limited, topology changes frequently and rapidly, and power is constrained.

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 http://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 29, 2015.

Copyright Notice

Table of Contents

1. Introduction

This document describes the On-Demand Multicast Routing Protocol (ODMRP) [ODMRP-Journal]. ODMRP applies "on-demand" routing techniques to avoid channel overhead and improve scalability. It uses the concept of "forwarding group" [FGMP], a set of nodes responsible for forwarding multicast data, to build a forwarding mesh for each multicast group. By maintaining and using a mesh instead of a tree, the drawbacks of multicast trees in mobile wireless networks (e.g., intermittent connectivity, traffic concentration, frequent tree reconfiguration, non-shortest path in a shared tree, etc.) are avoided. A soft-state approach is taken to maintain multicast group members, meaning that no explicit control message is required to leave the group. ODMRP does not rely on any unicast routing protocol: in particular, it can operate in conjunction with both reactive and proactive unicast routing protocols.

1.1. Motivation and Experiments

The main rationale for ODMRP is its potential to reduce control and traffic overhead in certain MANET deployments, typically where multicast traffic is relatively sparse. While this protocol has been extensively studied in simulations, it does not yet benefit from sufficient operational experience in order to be considered for Standards Track. In addition to general operational experience such as interoperability testing, this specification is intended to collect data on the following points:

- As a multicast routing protocol for MANET, ODMRP can be compared with [RFC6621], but can also be used in conjunction, taking advantage of its Duplicate Packet Detection and optimized flooding mechanisms. The rationale behind ODMRP is that, with sparser traffic, and in particular less sources, ODMRP should reduce the control overhead and number of data packets transmitted by making use of Forwarding Groups. This hypothesis should be validated, and experiments and operational deployments demonstrating the scenarios in which ODMRP performs better, or worse, than [RFC6621] should be performed.

- The potential scope of deployment of ODMRP should be assessed, particularly in comparison to other MANET protocols.

- Default values and guidelines for the parameters described in Section 5 should be provided, based on operational experience gathered from implementing and deploying this specification.

- The feasability of implementing ODMRP in common MANET situations should be examined. In particular, it should be determined if a

linux user space implementation is possible.

2. Terminology and Notation

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].

This document also makes use of the terminology defined in [RFC5444]. Additionally, it uses the notation defined in Section 2.1, and the terminology defined in Section 2.2.

2.1. Notation

ODMRP Routers generate and process messages, each of which has a number of distinct fields. For describing the protocol operations, specifically the generation and processing of such messages, the following notation is employed:

MsgType.field

where:

MsgType - is the type of message (e.g., JQ or JR);

field - is the field in the message (e.g., SourceAddress).

Furthermore, the following notational conventions are used:

a := b an assignment operator, whereby the left side (a) is assigned the value of the right side (b)

c = d a comparison operator, returning TRUE if and only if the value of the left side (c) is equal to the value of the right side (d)

[x] a list containing x as its only element

The different messages, their fields and their meaning are described in Section 7.

2.2. Terminology

ODMRP Router - A router that implements this protocol. An ODMRP Router is equipped with at least one, and possibly more, ODMRP Interfaces.

ODMRP Interface - An ODMRP Router's attachment to a communication medium, over which it receives and generates control messages, according to this specification. An ODMRP Interface is assigned one or more addresses.

Neighbor (ODMRP) Router - An ODMRP Router A is a neighbor of another ODMRP Router B if B can receive control messages from A according to this specification. This relationship is not necessarily symmetrical.

Neighbor (ODMRP) Interface - An interface X of an ODMRP Router A is a neighbor relative to interface Y of ODMRP Router B if B can receive control messages sent by A over the link X - Y. The link, and this relationship, are not necessarily symmetrical.

Multicast session - The entity defined by a (multicast group, source) pair, representing the group to which a source sends multicast packets.

Forwarding group - A group of ODMRP Routers participating in multicast packet forwarding for a given Multicast Session. In particular, the Forwarding Group is constituted of the Multicast Source, the Multicast Receivers and the Intermediate Routers.

Multicast Receiver - An ODMRP Router is a Multicast Receiver, relative to a given Multicast Session, if it subscribes to the Multicast Session in order to receive data packets sent by the Multicast Source.

Intermediate Router - An Intermediate Router is an ODMRP Router that is a member of a Forwarding Group without being a Multicast Receiver. In other words, it joined the Forwarding Group to transmit control and data traffic between the Multicast Source and the Multicast Receivers.

Join Query - The control message sent by Multicast Sources to establish and update group memberships and routes.

Join Reply - The control message sent by Multicast Receivers and forwarded by Intermediate Routers to build the Forwarding Group according to group membership information.

Upstream - An ODMRP Router (A) is said to be "upstream" compared to another ODMRP Router (B), relatively to a given Multicast Session, if A is on the path, which is discovered by a Join Query-Join Reply exchange and used by data packets, between B and the Multicast Source. In other words, any data packet sent within the Multicast Session has to transit through A before reaching B.

Downstream - An ODMRP Router (A) is said to be "downstream" compared to another ODMRP Router (B), relatively to a given Multicast Session, if B is on the path which is discovered by a Join Query-Join Reply exchange and used by data packets, between A and the Multicast Source. In other words, A is "downstream" from B if B is "upstream" from A.

3. Applicability Statement

This protocol is a multicast routing protocol, intended for use in Mobile Ad Hoc Networks (MANETs). MANETs generally have constrained resources (processing power, battery, etc.) and very dynamic topologies. With ODMRP, routing state is installed and maintained in an on-demand fashion, which avoids the issue of frequent tree reconfiguration seen with more classic multicast routing protocols.

ODMRP does not rely on the use of any unicast routing protocol, whether reactive or proactive, but MAY be used coinjointly with such protocols, such as [RFC7181] or [RFC3561]. Additionally, ODMRP can run in conjunction with [RFC6621], and take advantage of any optimized flooding mechanism used in the network, such as those offered by SMF, to disseminate Join Query messages as described in Section 12.

4. Protocol Overview and Functioning

The objective of this protocol is to allow each ODMRP Router to:

- Build a Forwarding Group only when it has data traffic to send to a Multicast group.
- Maintain the Forwarding Group for as long as necessary, until there is no more data to be sent to the Multicast group.
- Join any Forwarding Group, in order to receive multicast data packets from the corresponding multicast source. The decision to join a given Forwarding Group is triggered by Multicast membership information relative to the corresponding Multicast session. Such information can be received from other protocols, such as IGMP

[RFC3376] and MLD [RFC3810].

4.1. Routers and Interfaces

Each ODMRP Router MUST be provisioned with at least one ODMRP Interface, and keep a list of all these interfaces, as described in Section 9. The management of these interfaces (addition, deletion, re-addressing of any interface) is out of scope for this document.

4.2. Information Base Overview

Protocol state is recorded in eight distinct information sets: the Local Interface Set, the Neighbor Interface Set, the Multicast Routing Set, the Forwarding Table, the Pending Acknowledgement Set, the Pre-acknowledgement Set, the Blacklist and the Sent JQ Set. With the exception of the Local Interface Set, all these information sets are both used and updated by this protocol.

The Local Interface Set records a list matching each ODMRP interface of this router to the addresses in use for this interface. This set is used, but not updated by this protocol.

The Neighbor Interface Set records all the known addresses of neighbor ODMRP interfaces, by way of recording data from received JQ messages. This set can also be updated by other protocols with knowledge of neighbor interfaces, such as [RFC6130].

The Multicast Routing Set contains tuples, each representing the address of a multicast group, the address of a source sending data to this multicast group, and the next hop towards the multicast source.

The Forwarding Table contains tuples, each representing a given Multicast session for which the ODMRP Router forwards packets.

The Pending Acknowlegement Set contains tuples, each corresponding to a Join Reply message which has been sent by this Router and is waiting for an acknowledgement from a chosen upstream Router.

The Pre-acknowledgement Set contains tuples, representing overheard Join Reply messages, that are not destined to this Router but may pre-acknowledge a future Join Reply from this Router.

The Blacklist contains tuples, corresponding to neighbor ODMRP Routers, with which connectivity has been detected to be unidirectional.

The Sent JQ Set matches interfaces of this Router with the address carried by the last JQ message to have transited through that

interface.

4.3. Signaling Overview

This protocol generates and processes the following routing messages:

Join Query - Generated by an ODMRP Router while it has data packets to send to a multicast group, and flooded periodically to maintain the Forwarding Group necessary to deliver these data packets. A Join Query message hence advertises a Multicast Session, and contains:

* The multicast group address
* The source address
* A sequence number
* The last address used by this interface to send a JQ message

Join Reply - Generated by an ODMRP Router belonging to a multicast session, in reply to a Join Query message advertising this multicast session (corresponding Join Query), then forwarded by ODMRP Routers belonging to the same multicast session along the reverse path to the multicast source. A Join Reply message contains:

* The multicast group address and source address, identifying the multicast session
* The sequence number carried by the corresponding Join Query
* The address of the next hop on the path towards the multicast session source

4.4. Overview

The objectives of this protocol are achieved, for each ODMRP Router, by the following operations:

o When having data to send to a multicast group, for which no Forwarding Group is already established, an ODMRP Router generates a Join Query and transmits it over all of its ODMRP Interfaces. It then periodically repeat this process, until it has no more data to send to the multicast group.

o Upon receiving a Join Query, an ODMRP Router installs or refreshes a tuple in the Multicast Routing Set indicating the reverse path

towards the source of the Join Query, then considers it for forwarding, according to the forwarding mechanism specified for the network.

- If this Router belongs to, i.e., has attached hosts which have subscribed to, the multicast session that the Join Query advertises, it originates a Join Reply and transmits it over all of its ODMRP Interfaces.

- Upon receiving a Join Reply, an ODMRP Router inspects the next hop address carried by this packet:

  * If it corresponds to the address of an interface of this Router, and if this Router has a tuple in its Multicast Routing Set, corresponding to the advertised source, the ODMRP Router belongs to the Forwarding Group for the Multicast Session. Consequently, it installs or refreshes the corresponding entry in its Forwarding Table. It then considers the Join Reply for forwarding, according to the forwarding mechanism specified for the network.

  * Otherwise, it verifies if any Pending Acknowledgement tuple corresponds to this Join Reply and marks each such tuple as acknowledged. It silently discards the Join Reply.

- After sending a Join Reply, addressed to an upstream router A, an ODMRP Router looks in its Pre-acknowledgement Set for a corresponding Overheard tuple.

  * If such a tuple exists, the Overheard tuple is discarded and no further action is taken.

  * Otherwise, i.e., if the Pre-acknowledgement Set does not contain any corresponding Overheard tuple, it creates a Pending Acknowledgement tuple in the Pending Acknowledgement Set. If this tuple expires without being acknowledged, the link with router A is considered unidirectional: it is blacklisted, and the current router MAY try other means of joining the Forwarding Group.

- While it has data to send to the multicast group, an ODMRP Router periodically originates a Join Query and transmits it to all of its neighbors, in order to maintain the Forwarding Group.

5. Parameters and Constants

This specification uses both Router parameters, described in Section 5.1, and per-interface parameters, described in Section 5.2.

5.1. Router Parameters

This specification uses the following Router parameters:

ROUTE_REFRESH_INTERVAL - is the interval between two perodic Join Queries sent by a Multicast Source

FG_TIMEOUT - is the minimum time a Forwarding Tuple SHOULD be kept in the Forwarding Table after it was last refreshed

5.2. Interface Parameters

This specification uses the following interface parameters:

ROUTE_TIMEOUT - is the minimum time a Routing Tuple SHOULD be kept in the Routing Set after it was last refreshed

JR_RETRIES - is the number of times an ODMRP Router SHOULD attempt to retransmit a given Join Reply before declaring the link with the upstream neighbor interface unidirectional

ACK_TIMEOUT - is the time after which a Pending Tuple expires and MUST be considered invalid, as well as trigger the appropriate action according to Section 11

PRE_ACK_TIMEOUT - is the time after which an Overheard Tuple expires and MUST be considered invalid

LOCAL_ADDRESS_TIMEOUT - is the time after which a sent JQ tuple expires and MUST be considered invalid. This parameter SHOULD be less than the time an interface address is expected to be in use for the corresponding communication medium

NEIGHBOR_ADDRESS_TIMEOUT - is the time after which an address tuple of a neighbor interface tuple expires and MUST be considered invalid. This parameter SHOULD be less than the time an interface address is expected to be in use for the corresponding communication medium

6. Sequence Numbers

Each ODMRP Router maintains a single sequence number, which must be included in each Join Query message it generates. Each ODMRP Router MUST make sure that no two Join Query messages are generated with the same sequence number, and MUST generate sequence numbers such that these are monotonically increasing. This sequence number is used as freshness information for when comparing routes to the ODMRP Router having generated the message.

However, with a limited number of bits for representing sequence numbers, wrap-around (that the sequence number is incremented from the maximum possible value to zero) will occur. To prevent this from interfering with the operation of the protocol, the following MUST be observed. The term MAX_SEQ_NUM designates in the following the largest possible value for a sequence number. The sequence number S1 is said to be "greater than" (denoted ’>’) the sequence number S2 if:

S2 < S1 AND S1 - S2 <= MAX_SEQ_NUM/2 OR

S1 < S2 AND S2 - S1 > MAX_SEQ_NUM/2

7. Packets and Messages

This section describes the protocol messages generated and processed by ODMRP, according to the notations defined in Section 2. The objective of this section is to specify the content and meaning of each message. The specifics of the encoding of these messages, including the exact type and length of each field, in accordance with [RFC5444], are described in Section 8.

7.1. Join Query Format

A Join Query (JQ) message has the following fields:

JQ.AddressLength encodes the length of the addresses carried by this message as follows:

JQ.AddressLength := the length of an address in octets - 1

JQ.MulticastGroupAddress encodes the address of the Multicast Group, to which this Join Query is addressed

JQ.SourceAddress encodes an address of the source of this Join Query

JQ.SequenceNumber encodes the sequence number (see Section 6) of the ODMRP Router, generating the Join Query message

JQ.LastAddress encodes the address set as the source address of the last IP datagram sent through the same interface and containing a JQ message, or of the IP datagram carrying this JQ message if no such datagram is known

7.2. Join Reply Format

A Join Reply (JR) message has the following fields:

JR.AddressLength encodes the length of the addresses carried by this message as follows:

JR.AddressLength := the length of an address in octets - 1

JR.MulticastGroupAddress encodes the address of the Multicast Group, to which this Join Reply is addressed

JR.AckRequired is a boolean flag. When set (’1’), it specifies that the recipient of the Join Reply MUST acknowledge its reception by a sending Join Reply message. If cleared (’0’), the recipient of this message MAY suppress it’s Join Reply transmission, according to Section 10

JR.SourceAddress encodes the address of the Source of the Multicast Session

JR.SequenceNumber encodes the sequence number (see Section 6) of the corresponding Join Query message

JR.NextHopAddress encodes the the address of the next hop on the path towards the source of the multicast session

8. RFC5444 Encoding

This section describes the encoding of ODMRP messages using [RFC5444].

8.1. Join Query Encoding

This protocol defines the Join Query message type. Hence, according to [RFC5444], all Join Query messages are generated, processed and transmitted following this specification. Table 1 shows the mapping between the Join Query elements described in Section 7.1 and their encoding. All elements described in Table 1 MUST be included in every Join Query message, with the exception of the JQ.LastAddress element. JQ.LastAddress MAY be omitted in a given JQ message if it corresponds to the source address of the IP datagram containing this message.

JQ Element	RFC5444 Element
JQ.AddressLength	<msg-addr-length>
JQ.SourceAddress	<msg-orig-addr>
JQ.MulticastGroupAddress	Address in address block + TLV
JQ.SequenceNumber	<msg-seq-num>
JQ.LastAddress	Address in address block + TLV

Table 1: Join Query Message Elements

8.2. Join Reply Encoding

This protocol defines the Join Reply message type. Hence, according to [RFC5444], all Join Reply messages are generated, processed and transmitted following this specification. Table 2 shows the mapping between the Join Reply elements described in Section 7.2 and their encoding. With the exception of the ACKREQUIRED TLV, all elements described in Table 2 MUST be included in every Join Reply message.

JR Element	RFC5444 Element
JR.AddressLength	<msg-addr-length>
JR.SourceAddress	<msg-orig-addr>
JR.MulticastGroupAddress	Address in address block + TLV
JR.SequenceNumber	<msg-seq-num>
JR.NextHopAddress	Address in address block + TLV
JR.AckRequired	ACKREQUIRED TLV

Table 2: Join Reply Message Elements

9. Information Bases

Each router maintains an Information Base, containing a Local Interface Set, a Neighbor Interface Set, a Multicast Routing Set, a Forwarding Table, a Pending Acknowledgement Set, a Pre-Acknowledgement Set, a Blacklist and a Sent JQ Set, as described in the following sections. These information sets are given so as to facilitate description of message generation, forwarding and processing rules. In particular, an implementation may chose any representation or structure for when maintaining this information.

9.1. Local Interface Set

The Local Interface Set records a list of all the interfaces of the ODMRP Router, which participate in the operations of this protocol; that is, over which ODMRP control messages are exchanged, according to this specification. Each tuple of the Interface Set, or Interface Tuple, is as follows:

(I_interface, I_interface_address_list)

Where:

I_interface - The local ODMRP Interface

I_interface_address_list - The list of addresses used by the local interface in the operations of this protocol.

9.2. Neighbor Interface Set

The Neighbor Interface Set records the known addresses of Neighbor ODMRP Interfaces. It is used by this protocol, and can be updated by this protocol or by any other suitable protocol in operation that provides the necessary information, such as NHDP [RFC6130]. Each neighbor interface tuple is as follows:

(N_interface_address_list)

Where:

N_interface_address_list - Is an unordered list of at least one address tuple (i_addr, i_addr_exp_time), where:

i_addr - is a known address of the Neighbor Interface, i.e., an address that was set as the sender of an IP datagram sent through this interface

i_addr_exp_time - is the time at which the address tuple MUST be considered expired and thus MUST NOT be taken in considerations for the operations of this protocol

A neighbor interface tuple that contains no valid (i.e., non-expired) address tuple MUST considered expired and MUST NOT be taken in considerations for the operations of this protocol

9.3. Multicast Routing Set

The Multicast Routing Set contains Routing Tuples, indicating the path towards Multicast Sources, and containing the following fields:

(R_source, R_next_hop, R_local_interface, R_seq_num, R_exp_time)

Where:

R_source - is the address of the Multicast Source.

R_next_hop - is an address of the next hop along the path to the Multicast Source, i.e., an address of one of the interfaces of the neighbor ODMRP Router, from which the last valid Join Query message from this source was received, as recorded by the packet containing this Join Query.

R_local_interface - is the local interface, through which the next hop can be reached.

R_seq_num - corresponds to the JQ.SequenceNumber of the last valid Join Query originated by the Multicast Source and received by this ODMRP Router.

R_exp_time - is the time at which the tuple MUST be considered expired and thus MUST NOT be taken into consideration by the operations of this protocol.

9.4. Forwarding Table

The Forwarding Table contains Forwarding Tuples, representing Multicast Sessions for which the ODMRP Router forwards messages, i.e., the ODMRP Router is part of these Multicast Sessions' Forwarding Groups. These tuples are as follows:

(F_multicast_group, F_multicast_source, F_seq_num, F_exp_time)

Where:

F_multicast_group - is the address of the Multicast Group of the Multicast Session, for which the ODMRP Router forwards messages.

F_source - is the address of the Multicast Source of the Multicast Session, for which the ODMRP Router forwards messages.

F_seq_num - is the sequence number, corresponding to the last Join Query sent by the multicast source for the multicast session.

F_exp_time - is the time at which the tuple MUST be considered expired and thus MUST NOT be taken into consideration by the operations of this protocol.

9.5. Pending Acknowledgements

The Pending Acknowledgements Set contains Pending Acknowledgement tuples, representing Join Reply messages that are waiting to be acknowledged by the selected upstream Forwarding Group member. These tuples are as follows:

(P_multicast_group, P_multicast_source, P_seq_num, P_local_interface, P_next_hop, P_nth_time, P_exp_time)

Where:

P_multicast_group - is the JR.MulticastGroupAddress carried in the Join Reply awaiting acknowledgement (henceforth corresponding Join Reply).

P_multicast_source - is the JR.SourceAddress field carried in the corresponding Join Reply.

P_seq_num - is the JR.SequenceNumber field of the corresponding Join Reply.

P_next_hop - is the JR.NextHopAddress field of the corresponding Join Reply.

P_local_interface - is the local interface, through which the Join Reply was sent.

P_nth_time - corresponds to the number of times this Join Reply has been previously sent without being acknowledged.

P_exp_time - is the time at which this tuple MUST be considered expired.

P_acknowledged - is a boolean indicating whether the corresponding Join Reply has been acknowledged.

9.6. Pre-acknowledgements

The Pre-acknowledgements Set contains Overheard Tuples, corresponding to Join Reply messages, which have been sent by neighbors of this ODMRP Router but do not contain an address of this Router and do not acknowledge any tuple in the Pending Acknowledgement Set. The Overheared Tuples are as follows:

(O_multicast_group, O_multicast_source, O_seq_num, O_originator, O_exp_time)

Where:

O_multicast_group - is the JR.MulticastGroupAddress carried in the overheard Join Reply.

O_multicast_source - is the JR.SourceAddress field carried in the corresponding Join Reply.

O_seq_num - is the JR.SequenceNumber field of the corresponding Join Reply.

O_originator - is the address of the ODMRP Router’s interface which has sent the Join Reply.

O_exp_time - is the time at which this tuple expires MUST be considered invalid.

9.7. Blacklist

The Blacklist contains Blacklisted Tuples, corresponding to neighbor ODMRP Router interfaces, with which connectivity has been detected to be unidirectional, e.g., which have not acknowledged Join Replies from this Router, as specified in Section 10. In other words, a Blacklisted Tuple corresponds to a link between one local interface and one neighbor interface which has been detected to be unidirectional or broken. The Blacklist Tuples are as follows:

(B_neighbor_interface, B_local_interface, B_exp_time)

Where:

B_neighbor_interface_address_list - is a list of addresses belonging to the blacklisted interface.

B_local_interface - is the interface of this ODMRP router over which packets from the blacklisted interface were received.

B_exp_time - is the time at which this tuple expires and MUST be considered invalid.

9.8. Sent JQ set

The Sent JQ Set contains tuples matching transmitted (generated or relayed) Join Queries with interfaces addresses. Each of its tuples contains the source address, multicast group address and sequence number uniquely identifying a JQ message, as well as an interface and the address of that interface that was advertised when transmitting the packet containing the Join Query. More precisely, given a transmitted Join Query and an interface over which it was transmitted, a tuple of this set, or Sent JQ Tuple, is as follows:

Where:

S_interface - is the local interface, through which the JQ message was sent

S_interface_address - is the address of the ODMRP interface that was set as the packet source

S_exp_time - is the time at which this tuple expires and MUST be considered invalid.

10. Protocol Details

This protocol generates and processes Join Query and Join Reply messages, according to the operations described in the following sections. This section uses the additional notation and variables:

previous-hop-address - refers to the address of the neighbor ODMRP interface recorded by the source address field of the IP datagram carrying the message currently being processed (Join Query or Join Reply)

this Router - refers to the ODMRP Router generating, processing or forwarding the message (Join Query or Join Reply)

Receiving Interface (receiving-interface) - refers to the local ODMRP Interface, over which the message currently being processed was received

10.1. Join Query

A Join Query is generated by an ODMRP Router, which has data to send to a multicast group, for which no multicast session has been initialized. Join Queries are then periodically originated by the ODMRP Router while it has data to send to the multicast group.

10.1.1. Invalid Join Queries

A Join Query, received by an ODMRP Router, is invalid and MUST be discarded without processing (and in particular, MUST NOT be considered for forwarding) if any of these conditions applies:

- The address length carried by the Join Query (see Section 7) differs from the length of the addresses of this Router
- The Multicast Routing Set of this Router contains a Multicast Routing tuple, for which:
  - R_multicast_source = JQ.SourceAddress, and
  - R_seqnum > JQ.SequenceNumber or R_seqnum = JQ.SequenceNumber
- JQ.SourceAddress is an address of an interface of this Router
- The Blacklist contains a Blacklisted Tuple, for which
  - previous-hop-address is contained in B_neighbor_interface_address_list
  - B_local_interface = receiving-interface

10.1.2. Join Query Generation

A Join Query is generated according to Section 7 with the following content:

- JQ.AddressLength set to the length of the addresses of this Router minus 1, as specified in Section 7

- JQ.MulticastGroupAddress set to the address of the multicast group, to which this Router is sending data
- JQ.SourceAddress set to an address of this ODMRP Router
- JQ.SequenceNumber set to the current sequence number of this Router, as specified in Section 6

10.1.3. Join Query Processing

Upon receiving a valid Join Query message, an ODMRP Router proceeds as follows:

1. Find the neighbor interface tuple such as N_interface_address_list contains an address tuple with i_addr = previous-hop-address, and update the address tuple such as i_addr_exp_time := current-time + NEIGHBOR_ADDRESS_TIMEOUT
2. If no such tuple exists, create one with:
   * N_interface_address_list := [(previous-hop-address, current-time + NEIGHBOR_ADDRESS_TIMEOUT)]
3. Find the Routing Tuple which satisfies: R_source = JQ.SourceAddress
4. If no such tuple exists, create a Routing Tuple with the following fields:
   * R_source := JQ.SourceAddress
   * R_next_hop := previous-hop
   * R_local_interface := receiving-interface
   * R_seq_num := JQ.SequenceNumber
   * R_exp_time := current-time + ROUTE_TIMEOUT

   and insert this tuple in the Routing Set
5. Else, i.e., if such a tuple exist, update it as follows:
   * R_next_hop := previous-hop
   * R_seq_num := JQ.SequenceNumber

* R_local_interface := receiving-interface

* R_exp_time := current-time + ROUTE_TIMEOUT

6. Consider the Join Query for forwarding, according to Section 10.1.4

7. If this Router is a member of the Multicast Group, addressed by JQ.MulticastGroupAddress, create a new Join Reply according to Section 10.2 and transmit it to all of this Router's neighbors

10.1.4. Join Query Forwarding

This section defines the following additional variables:

this-interface - is the ODMRP interface being considered

packet-source-address - is the source address of the outbound IP datagram carrying the JQ message being transmitted

For each ODMRP interface over which a JQ message is to be transmitted, a router MUST proceed as follows:

- Find the corresponding sent JQ tuple in the sent JQ set, such as S_interface = this-interface
- If no such tuple exists, create one with:
  * S_interface := this-interface
  * S_interface_address := packet-source-address
  * S_exp_time := current-time + LOCAL_ADDRESS_TIMEOUT
- Set JQ.LastAddress to S_interface_address
- Update the corresponding sent JQ tuple such as:
  * S_interface_address = packet-source-address
  * S_exp_time = current-time + LOCAL_ADDRESS_TIMEOUT

10.2. Join Reply

A Join Reply is generated by an ODMRP Router when it receives a Join Query such that at least one host attached to the ODMRP Router is a member of the Multicast Session advertised by the Join Query. This section makes use of the variable "new-jr", which is a boolean flag

set to TRUE if the Join Reply being processed contains more recent data than in the current information base. It has an initial value of FALSE.

10.2.1. Invalid Join Replies

A Join Reply, received by an ODMRP Router, is invalid and MUST be discarded without processing (and in particular, MUST NOT be considered for forwarding) if:

- The address length carried by the Join Reply (see Section 7) differs from the length of the address of the ODMRP Router
- There exists a Forwarding Tuple in this Router’s Forwarding Group table, such as:
  - F_source = JR.MulticastSourceAddress
  - F_seq_num > JR.SequenceNumber

10.2.2. Join Reply Generation

An ODMRP Router MUST generate a Join Reply in response to a received Join Query (henceforth "corresponding Join Query"), if at least one host attached to this Router is a member of the Multicast Session, advertised by the Join Query. A Join Reply is generated according to Section 7 with the following content:

- JR.AddressLength is set to the length of the address of this router minus 1, as specified in Section 7
- JR.MulticastGroupAddress is set to JQ.MulticastGroupAddress for the corresponding Join Query
- JR.SourceAddress is set to JQ.SourceAddress for the corresponding Join Query
- JR.SequenceNumber is set to JQ.SequenceNumber for the corresponding Join Query
- JR.NextHopAddress is set to the source address of the IP datagram containing the Join Query message

10.2.3. Join Reply Processing

Upon receiving a valid Join Reply, an ODMRP Router proceeds as follows:

1. If JR.NextHopAddress corresponds to an address recorded in the Local Interface Set of this ODMRP Router:

    1. Find the Forwarding Tuple (henceforth Matching Forwarding Tuple) such that:

        + F_multicast_group = JR.MulticastGroupAddress

        + F_multicast_source = JR.MulticastSourceAddress

    2. If no such tuple exists, insert in the Forwarding Table a new Forwarding Tuple such that:

        + F_multicast_group = JR.MulticastGroupAddress

        + F_multicast_source = JR.MulticastSourceAddress

        + F_seq_num = JR.SequenceNumber

        + F_exp_time = current-time + FG_TIMEOUT

        And set new-jr to TRUE

    3. Otherwise, the variable "new-jr" is set to TRUE if JR.SequenceNumber > F_seq_num, and to FALSE otherwise. Then, the pre-existing Matching Forwarding Tuple is updated as follows:

        + F_seq_num := JR.SequenceNumber

        + F_exp_time := current-time + FG_TIMEOUT

    4. If new-jr = TRUE or if JR.AckRequired is set the Join Reply is considered for forwarding. Otherwise, it is not processed further; in particular, it MUST NOT be considered for forwarding.

2. Otherwise, find the Multicast Routing Tuple in the Routing Set (henceforth "Matching Multicast Routing Tuple"), such as:

    * R_source = JR.SourceAddress

    * R_seq_num <= JR.SequenceNumber

    If previous-hop-address = R_next_hop, then:

    3. If the Pending Acknowledgement Set contains a Pending Tuple (henceforth "Matching Pending Tuple") such as:

+ P_multicast_group = JR.MulticastAddress

+ P_multicast_source = JR.SourceAddress

+ P_seq_num = JR.SequenceNumber

+ P_next_hop = previous-hop-address

The Matching Pending Tuple MUST be updated as follows:

+ P_acknowledged = TRUE

+ P_exp_time = EXPIRED

The Join Reply is not processed further, and in particular MUST NOT be considered for forwarding

4. Otherwise, if the Pre-Acknowledgement Set does not contain any Overheard Tuple such as:

+ O_multicast_group = JR.MulticastGroupAddress

+ O_multicast_source = JR.SourceAddress

+ O_seq_num = JR.SequenceNumber

+ O_originator = previous-hop-address

Insert a tuple with these fields, and O_exp_time = current-time + PRE_ACK_TIMEOUT in the Pre-Acknowledgement Set. The Join Reply is not processed further, and in particular MUST NOT be considered for forwarding

3. Otherwise, the Join Reply is silently discarded without further processing

10.2.4. Join Reply Forwarding

A Join Reply, considered for forwarding, MUST be updated as follows:

o Find the Matching Routing Tuple, such that:

  * R_source = JR.MulticastSourceAddress

  * R_seq_num <= JR.SequenceNumber

- If no such tuple exists, then the Join Reply is not processed further, and in particular MUST NOT be forwarded
- Otherwise, set JR.NextHop to R_next_hop

The Join Reply is then transmitted according to Section 10.2.5

10.2.5. Join Reply Transmission

A Join Reply is transmitted to all of an ODMRP Router's neighbors, in order to achieve two objectives:

- Set up or refresh the corresponding Forwarding Tuple for the upstream ODMRP neighbor
- If the Join Reply was not originated by this router, acknowledge its reception to the previous hop

Before transmitting the Join Reply, the Information Base is updated as follows:

1. If the Pre-acknowledgement Set contains a tuple, such that:
    * O_multicast_group = JR.MulticastGroupAddress
    * O_multicast_source = JR.SourceAddress
    * O_seq_num = JR.SequenceNumber
    * O_originator = JR.NextHopAddress

    Then clear the JR.AckRequired flag, and set O_exp_time to EXPIRED
2. Otherwise, if the Pending Acknowledgement Set contains a Pending Tuple such as:
    * P_multicast_group = JR.MulticastGroupAddress
    * P_multicast_source = JR.SourceAddress
    * P_seq_num = JR.SequenceNumber
    * P_next_hop = JR.NextHopAddress

    Then set JR.AckRequired, and increase P_nth_time by 1
3. Finally, if neither the Pre-acknowledgement Set nor the Pending Acknowledgement Set contain a corresponding tuple:

1. Insert a Pending Tuple in the Pending Acknowledgement Set, such as:

   + P_multicast_group = JR.MulticastGroupAddress

   + P_multicast_source = JR.SourceAddress

   + P_seq_num = JR.SequenceNumber

   + P_next_hop = JR.NextHopAddress

   + P_nth_time = 1

   + P_acknowledged = FALSE

   + P_expiration_time = current-time + ACK_TIMEOUT

2. Clear the JR.AckRequired flag

10.3. Forwarding Group Maintenance

While an ODMRP Router has data to send to a Multicast Group (on behalf of the Multicast Source), it MUST maintain the Forwarding Group generated by the initial Join Query. To this end, it MUST periodically generate JQ messages, according to Section 10.1.2. The interval between two Join Queries SHOULD be no less than ROUTE_REFRESH_INTERVAL. Note should be taken that, if the Multicast Session has no member other than the source, the Forwarding Group may contain only the designated ODMRP Router for the Multicast Source. That Router still needs to periodically flood Join Queries in order to rebuild a Forwarding Group if necessary.

10.4. Message Transmission

When using physical media subject to collisions and packet loss, both Join Query and Join Reply messages SHOULD be jittered to minimize the effect of collisions, as described in [RFC5148]

11. Unidirectional Links Handling

After sending a Join Reply, an ODMRP Router MUST verify that the upstream neighbor has joined the Forwarding Group. To this end, the following three mechanisms are used after transmitting a given Join Reply:

- If the ODMRP Router overhears a corresponding Join Reply from the upstream neighbor (see Section 10.2.3), this verifies that the

link is bidirectional and that the upstream neighbor has joined the Forwarding Group (passive acknowledgement)

- If the ODMRP Router has already overheard a corresponding Join Reply from the upstream neighbor prior to transmitting its own Join Reply, this means that the upstream neighbor has already joined the Forwarding Group (see Section 10.2.3) (pre-acknowledgement)

- Otherwise, i.e., if neither the pre-acknowledgement nor the passive acknowledgement have verified that the upstream neighbor joined the Forwarding Group (i.e., if the corresponding Pending Tuple expires with P_acknowledged set to False), then the ODMRP Router MUST proceed as follows:

  1. If the corresponding Pending tuple verifies P_nth_time < JR_RETRIES, then the ODMRP Router MUST retransmit the Join Reply with the JR.AckRequired flag set

  2. Otherwise, the link between the local interface and the interface of the upstream ODMRP Router identified by JR.NextHopAddress is considered unidirectional. In that case, the ODMRP Router SHOULD proceed as follows:

     + Find the neighbor interface tuple such that N_address_list contains an address tuple with i_addr = JR.NextHopAddress, and set the variable blacklisted-addresses to the list of addresses contained in N_address_list

     + Otherwise, if no such tuple exists, set the variable blacklisted-addresses to [JR.NextHopAddress]

     + Add a tuple in the Blacklist such as:

       - B_neighbor_interface_address_list := blacklisted-addresses

       - B_local_interface := P_local_interface

       - B_exp_time = current-time + BLACKLIST_TIMEOUT

     An ODMRP Router MAY attempt to use other mechanisms, such as [I-D.gerla-manet-odmrp-asym], to resume the Forwarding Group building process, instead of or in addition to using the Blacklist

12. SMF considerations

This protocol MAY be run in conjunction with SMF [RFC6621], and benefit from some of its features. In particular, if SMF is in use, it is RECOMMENDED that its duplicate packet detection feature described in Section 6 be used for multicast packet forwarding. Additionally, optimized flooding mechanisms, such as E-CDS or S-MPR, as specified in Appendices A through C of [RFC6621], MAY be used to flood Join Query messages throughout the network.

13. IGMP and MLD considerations

In order to determine whether or not it needs to reply to a Join Query message with a Join Reply message (as specified in Section 10.1.3), an ODMRP Router needs Multicast Group membership information. Such information can be provided by protocols such as IGMP [RFC3376] and/or MLD [RFC3810]. In particular, an ODMRP Router MUST reply with a Join Reply message to a valid Join Query messages advertising a Multicast Session if any of those conditions apply:

- This Router is subscribed to the corresponding Multicast Group.
- A host attached to this Router has signaled, for example using IGMP, that it has subscribed to the corresponding Multicast Group.

14. Multicast Packet Forwarding

ODMRP Routers originating and forwarding multicast packets MUST implement a duplicate packet detection (DPD) mechanism. If using IPv4 or IPV6 addresses, the use of SMF [RFC6621] is RECOMMENDED, as described in Section 12.

An ODMRP Router, receiving a non-duplicate multicast data packet, transmits it over all of its interfaces if it is a member of the forwarding group for this data packet, i.e., there exists a tuple in the Forwarding Group Table such as:

F_multicast_group correspond to the multicast address of this packet

F_multicast_source corresponds to the source of this packet

15. Security Considerations

As a multicast routing protocol, this protocol is potentially vulnerable to a number of attacks. This section attempts to describe the envisioned threats to the protocol, as well as some guidelines as to how to ensure confidentiality and integrity of the operations of ODMRP, and to mitigate threats of network overload.

This protocol relies on the use of a Duplicate Packet Detection (DPD) mechanism, such as one described in [RFC6621] (SMF), and suggests the use of optimized flooding to disseminate JQ messages. Some deployments of ODMRP are thus expected to function on top of [RFC6621] by taking advantage of the DPD and optimized flooding mechanisms provided by SMF. Such deployments are thus subject to the same security threats as SMF, such as those described in [I-D.ietf-manet-smf-threats].

15.1. Confidentiality

ODMRP routers which forward packets for multicast data source have to periodically transmit JQ messages throughout the network. In an unsecured network, an attacker could then eavesdrop on those messages and learn part or all of the network topology, depending on the traffic pattern.

15.2. Integrity

ODMRP relies on routers, in particular intermediate routers, to correctly transmit JQ and JR messages. An ODMRP router could, by malice or malfunction, originate JQ messages on behalf of a target multicast source with high enough sequence numbers to replace routing information in other routers. Such behavior would prevent the interested multicast receivers from receiving data packets sent by the target multicast source. An ODMRP router could also forward JQ messages with altered sequence numbers, thus preventing future routing updates. Both behaviors can be mitigated by end-to-end authentication of routing messages.

If NHDP [RFC6130] is not in use to update the Neighbor Interface Set, ODMRP relies on routers correctly informing their neighbors of the addresses they use via the JQ.LastAddress field. Upon transmission of a JQ message, an ODMRP router could, by malice or malfunction, set JQ.LastAddress to a network address that does not belong to this router (address spoofing). This could force neighbor ODMRP routers to blacklist this address in case the malicious router simulate unidirectional links by withholding JR messages. This behavior would break or slow down protocol convergence, potentially triggering data packet loss for multicast receivers. If NHDP is in use, the

deployment is subject to its own security vulnerabilities, such as those described in [RFC7186].

15.3. Channel Overload

ODMRP's main construct, the forwarding group, is built and maintained by having the source ODMRP router periodically flood JQ messages, which can be a costly operation in terms of bandwidth, processing and battery life, if applicable. A malicious router could flood JQ messages at a very high rate to overload the network. It is thus RECOMMENDED that ODMRP routers in a given deployment implement a rate-limit mechanism to prevent such behavior.

The efficiency (in terms of number of multicast data packets transmitted) of forwarding groups depends on routers actually sending JR messages only when necessary, in order to build a graph as sparse as possible and avoid redundant transmissions. Thus an ODMRP router which replies to JQ messages by transmitting one JR message for each of its known neighbors and with JR.NextHopAddress set to an address of this neighbor, would severely harm the efficiency of ODMRP by forcing the routers to build a forwarding group with unnecessary redundancy. Such behavior could also result in routing loops.

16. IANA Considerations

This specification defines two new Message Types, Join Query and Join Reply, which must be allocated from the "Message Type" repository of [RFC5444].

16.1. Join Query Registries

IANA is requested to create a registry for Message-Type-specific Message TLV Types for Join Query messages, with initial assignments according to Table 3.

Type	Description	Allocation Policy
128-223	Unassigned	Expert Review

Table 3: Join Query Message-Type-specific Message TLV Types

IANA is requested to create a registry for Message-Type-specific Address Block TLV Types for Join Query messages, with initial assignments according to Table 4.

Type	Description	Allocation Policy
128	ADDR-TYPE	
129-223	Unassigned	Expert Review

Table 4: Join Query Message-Type-specific Address Block TLV Types

Allocation of the ADDR-TYPE TLV from the Join Query specific Address Block TLV Types will create a new Type Extension Registry with initial assignments as specified in Table 5.

Name	Type	Type Ext.	Description	Al. Policy
ADDR-TYPE	128	0	MULTICAST-GROUP-ADDRESS	
ADDR-TYPE	128	1	LAST-ADDRESS	
ADDR-TYPE	128	2-255	Unassigned	Expert Review

Table 5: Address Block TLV Type assignment for ADDR-TYPE

16.2. Join Reply Registries

IANA is requested to create a registry for Message-Type-Specific Message TLV Types for Join Reply messages, with initial assignments according to Table 6.

Type	Description	Allocation Policy
128	ACKREQUIRED	
129-223	Unassigned	Expert Review

Table 6: Join Reply Message-Type-specific Message TLV Types

IANA is requested to create a registry for Message-Type-specific Address Block TLV Types for Join Reply messages, with initial assignments according to Table 7.

Type	Description	Allocation Policy
128	ADDR-TYPE	
129-223	Unassigned	Expert Review

Table 7: Join Reply Message-Type-specific Addres Block TLV Types

Allocation of the ADDR-TYPE TLV from the Join Reply specific Address Block TLV Types will create a new Type Extension Registry with initial assignments as specified in Table 8.

Name	Type	Type Ext.	Description	Al. Policy
ADDR-TYPE	128	0	MULTICAST-GROUP-ADDRESS	
ADDR-TYPE	128	1	NEXT-HOP-ADDRESS	Expert Review
ADDR-TYPE	128	2-255	Unassigned	Expert Review

Table 8: Address Block TLV Type assignment for ADDR-TYPE

17. Acknowledgements

The authors would like to thank Thomas Clausen and Justin Dean for their insigthful reviews and comments.

18. References

18.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, BCP 14, March 1997.

[RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

[RFC3810] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

[RFC5148] Clausen, T., Dearlove, C., and B. Adamson, "Jitter

Considerations in Mobile Ad Hoc Networks (MANETs)", RFC 5148, February 2008.

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized MANET Packet/Message Format", RFC 5444, February 2009.

[RFC6621] Macker, J., "Simplified Multicast Forwarding", RFC 6621, May 2012.

18.2. Informative References

[FGMP] Chiang, C., Gerla, M., and L. Zhang, "Forwarding Group Multicast Protocol (FGMP) for Multihop, Mobile Wireless Networks", Avril 1998.

[I-D.gerla-manet-odmrp-asym] Gerla, M., Oh, S., and A. Colin de Verdiere, "ODMRP_ASYM", draft-gerla-manet-odmrp-asym-00 (work in progress).

[I-D.ietf-manet-smf-threats] Yi, J., Clausen, T., and U. Herberg, "Security Threats for Simplified Multicast Forwarding (SMF)", draft-ietf-manet-smf-threats-00 (work in progress), August 2014.

[ODMRP-Journal] Lee, S., Su, W., and M. Gerla, "On-Demand Multicast Routing Protocol in Multihop Wireless Networks", Journal of Mobile Networks and Applications, Volume 7 Issue 6, Pages 441 - 453, December 2002.

[RFC3561] Perkins, C., Belding-Royer, E., and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", RFC 3561, July 2003.

[RFC6130] Clausen, T., Dearlove, C., and J. Dean, "Mobile Ad Hoc Network (MANET) Neighborhood Discovery Protocol (NHDP)", RFC 6130, April 2011.

[RFC7181] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol Version 2", RFC 7181, April 2014.

[RFC7186] Yi, J., Herberg, U., and T. Clausen, "Security Threats for the Neighborhood Discovery Protocol (NHDP)", RFC 7186, April 2014.

Appendix A. Illustrations

This section shows examples of ODMRP control messages encoded using [RFC5444]. [RFC5444] specifies that a packet is formed by a packet header, an optional TLV block and zero or more messages. This specification does not use or require any packet TLV. Additionally, the minimal packet header required by ODMRP is shown in Figure 1.

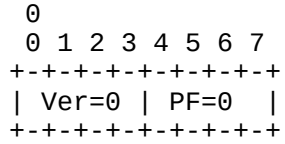


Figure 1: Packet Header

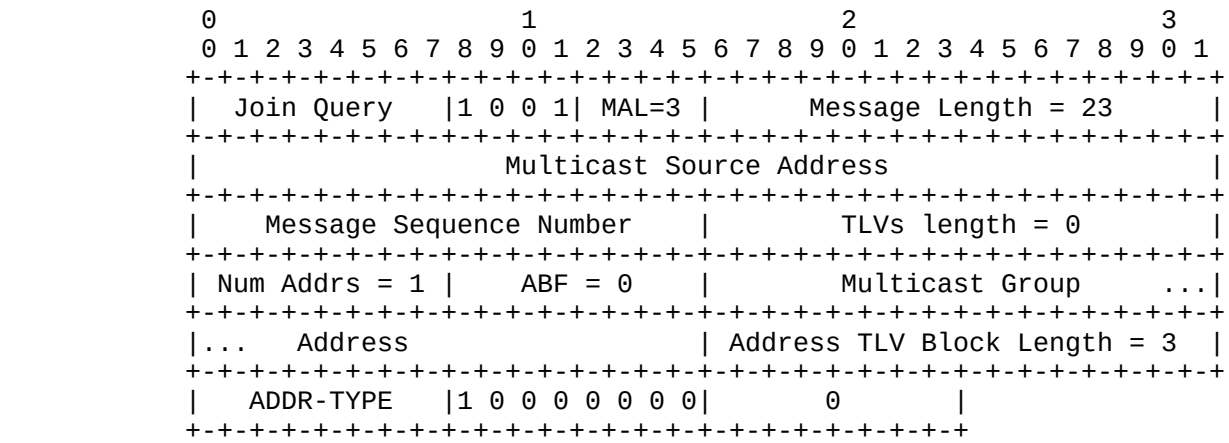
A.1. Join Query Message

JQ messages are instances of [RFC5444] messages. This section illustrates an example of one such message.

The JQ message’s header’s flag octet has a value of 9, meaning that the sequence number and source address fields are present, encoding respectively the sequence number and the address of the multicast source that originated the message. Additionally, the address length field (MAL) is set to 3, corresponding to an address length of 4 octets (i.e., the length of an IPv4 address). The overall message size is 23 octets.

An additional Message-Type specific address block is present, with one address and a flag octet (ABF) having value 0, meaning that the address block has no Head or Tail element. The Mid element encodes the Multicast group address. The associated TLV is of type ADDR-TYPE and value 0, i.e. MULTICAST-GROUP-ADDRESS.

The LastAddress element is omitted, meaning that the last JQ message from this interface was transmitted using the same address as this one.

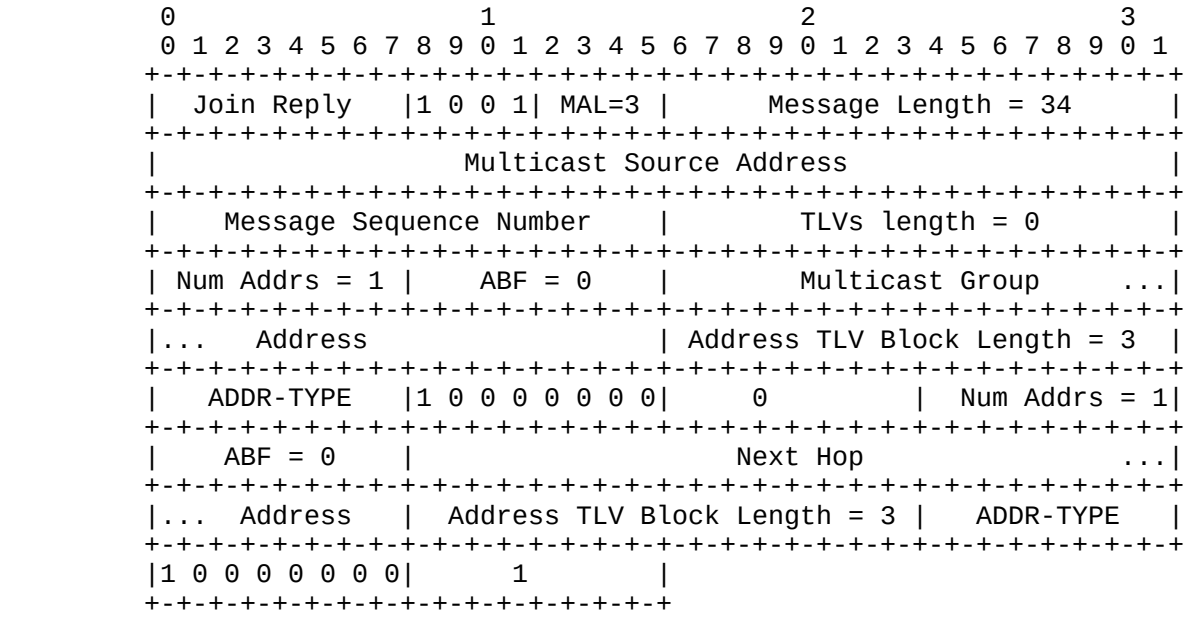


A.2. Join Reply Message

JR messages are instances of [RFC5444] messages. This section illustrates an example of one such message.

The JR message's header's flag octet has a value of 9, meaning that the sequence number and source address fields are present, encoding respectively the sequence number and the address of the multicast source that originated the message. Additionally, the address length field (MAL) is set to 3, corresponding to an address length of 4 octets (i.e., the length of an IPv4 address). The overall message size is 34 octets.

Two additional Message-Type specific address blocks are present, both with one address and a flag octet (ABF) having value 0, meaning that the address block has no Head or Tail element. For the first address block, the Mid element encodes the Multicast group address; the associated Message-Type-specific TLV is of type ADDR-TYPE and value 0, i.e. MULTICAST-GROUP-ADDRESS. The second address block's Mid element encodes the Next Hop address; its associated Message-Type-specific TLV is of type ADDR-TYPE and value 1, i.e., NEXT-HOP-ADDRESS.



Authors' Addresses

Yunjung Yi
University of California, Los Angeles

Sung-Ju Lee
University of California, Los Angeles

William Su
The Boeing Company

Email: william.su@boeing.com

Mario Gerla
University of California, Los Angeles
3732F Boelter Hall
Computer Science Department
University of California
Los Angeles, CA 90095-1596,
USA

Phone: +1 310 825-4367
Email: gerla@cs.ucla.edu

Axel Colin de Verdiere
University of California, Los Angeles

Email: axel@axelcdv.com

Mobile Ad hoc Networks Working Group
Internet-Draft
Intended status: Experimental
Expires: November 4, 2016

C. Perkins
Futurewei
S. Ratliff
Idirect
J. Dowdell
Airbus Defence and Space
L. Steenbrink
HAW Hamburg, Dept. Informatik
V. Mercieca
Airbus Defence and Space
May 3, 2016

Ad Hoc On-demand Distance Vector Version 2 (AODVv2) Routing
draft-ietf-manet-aodvv2-16

Abstract

The Ad Hoc On-demand Distance Vector Version 2 (AODVv2) routing protocol is intended for use by mobile routers in wireless, multihop networks. AODVv2 determines unicast routes among AODVv2 routers within the network in an on-demand fashion.

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 http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 4, 2016.

Table of Contents

1. Overview

The Ad hoc On-Demand Distance Vector Version 2 (AODVv2) protocol enables dynamic, multihop routing between participating mobile nodes wishing to establish and maintain an ad hoc network. The basic operations of the AODVv2 protocol are route discovery and route maintenance. AODVv2 does not require nodes to maintain routes to destinations that are not in active communication. AODVv2 allows mobile nodes to respond to link breakages and changes in network topology in a timely manner. The operation of AODVv2 is loop-free, and by avoiding the Bellman-Ford "counting to infinity" problem offers quick convergence when the ad hoc network topology changes (typically, when a node moves in the network). When links break, AODVv2 causes the affected set of nodes to be notified so that they are able to invalidate the routes using the lost link.

One distinguishing feature of AODVv2 is its use of a destination sequence number for each route entry. The destination sequence number is created by the destination to be included along with any route information it sends to requesting nodes. Using destination sequence numbers ensures loop freedom and is simple to program. Given the choice between two routes to a destination, a requesting node is required to select the one with the greatest sequence number.

Compared to AODV [RFC3561], AODVv2 has moved some features out of the scope of the document, notably intermediate route replies, expanding ring search, route lifetimes and precursor lists. However, the document has been designed to allow their specification in a separate document. Hello messages and local repair have been removed. AODVv2 provides a mechanism for the use of multiple metric types. Message formats have been updated and made compliant with [RFC5444]. AODVv2

control messages are defined as sets of data, which are mapped to message elements using the Generalized MANET Packet/Message Format defined in [RFC5444] and sent using the parameters in [RFC5498]. Verification of link bidirectionality has been substantially improved, and additional refinements made for route timeouts and state management.

The basic protocol mechanisms are as follows. Since AODVv2 is a reactive protocol, route discovery is initiated only when a route to the target is needed (i.e. when a router' client wants to send data). AODVv2 does this with the help of a Route Request (RREQ) and Route Reply (RREP) cycle: an RREQ is distributed across the network until it arrives at the target. When forwarding an RREQ, all routers across the network store the neighbor they've received the RREQ from, memorizing a possible route back to the originator of the RREQ. When the target receives the RREQ, it answers with an RREP, which then travels back to the originator along the path memorized by the intermediate routers. A metric value is included within the messages to record the cost of the route. AODVv2 uses sequence numbers to identify stale routing information, and compares route metric values to determine if advertised routes could form loops.

Route maintenance includes confirming bidirectionality of links to next hop AODVv2 routers and issuing Route Error (RERR) messages informing other routers of broken links. It also includes reacting to received Route Error messages, and extending and enforcing route timeouts.

The on-demand nature of AODVv2 requires signals to be exchanged between AODVv2 and the forwarding plane. These signals indicate when: * a packet is to be forwarded, in order to initiate route discovery * packet forwarding fails, in order to initiate route error reporting * a packet is successfully forwarded, for route maintenance.

Security for authentication of AODVv2 routers and encryption of control messages is accomplished using the TIMESTAMP and ICV TLVs defined in [RFC7182].

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]. In addition, this document uses terminology from [RFC5444], and defines the following terms:

AddressList

A list of IP addresses as used in AODVv2 messages.

AckReq
Used in a Route Reply Acknowledgement message to indicate that a Route Reply Acknowledgement is expected in return.

AdvRte
A route advertised in an incoming route message.

AODVv2 Router
An IP addressable device in the ad hoc network that performs the AODVv2 protocol operations specified in this document.

CurrentTime
The current time as maintained by the AODVv2 router.

ENAR (External Network Access Router)
An AODVv2 router with an interface to an external, non-AODVv2 network.

InterfaceSet
The set of all network interfaces supporting AODVv2.

Invalid route
A route that cannot be used for forwarding but still contains useful sequence number information.

LocalRoute
An entry in the Local Route Set as defined in Section 5.5.

MANET
A Mobile Ad Hoc Network as defined in [RFC2501].

MetricType
The metric type for a metric value included in a message.

MetricTypeList
A list of metric types associated with the addresses in the AddressList of a Route Error message.

Neighbor
An AODVv2 router from which an RREQ or RREP message has been received. Neighbors exchange routing information and verify bidirectionality of the link to a neighbor before installing a route via that neighbor into the Local Route Set.

OrigAddr
The source IP address of the IP packet triggering route discovery.

OrigMetric
The metric value associated with the route to OrigPrefix.

OrigPrefix
The prefix configured in the Router Client entry which includes OrigAddr.

OrigPrefixLen
The prefix length, in bits, configured in the Router Client entry which includes OrigAddr.

OrigSeqNum
The sequence number of the AODVv2 router which originated the Route Request on behalf of OrigAddr.

PktSource
The source address of the IP packet that triggered a Route Error message.

PrefixLengthList
A list of routing prefix lengths associated with the addresses in the AddressList of a message.

Reactive
Performed only in reaction to specific events. In AODVv2, routes are requested only when data packets need to be forwarded. In this document, "reactive" is synonymous with "on-demand".

RERR (Route Error)
The AODVv2 message type used to indicate that an AODVv2 router does not have a valid LocalRoute toward one or more particular destinations.

RERR_Gen (RERR Generating Router)
The AODVv2 router generating a Route Error message.

RerrMsg (RERR Message)
A Route Error (RERR) message.

Routable Unicast IP Address
A routable unicast IP address is a unicast IP address that is scoped sufficiently to be forwarded by a router. Globally-scoped unicast IP addresses and Unique Local Addresses (ULAs) [RFC4193] are examples of routable unicast IP addresses.

Router Client
An address or address range configured on an AODVv2 router, on behalf of which that router will initiate and respond to route

discoveries. These addresses may be used by the AODVv2 router itself or by its Router Clients that are reachable without traversing another AODVv2 router.

RREP (Route Reply)
The AODVv2 message type used to reply to a Route Request message.

RREP_Gen (RREP Generating Router)
The AODVv2 router that generates the Route Reply message, i.e., the router configured with TargAddr as a Router Client.

RREQ (Route Request)
The AODVv2 message type used to discover a route to TargAddr and distribute information about a route to OrigPrefix.

RREQ_Gen (RREQ Generating Router)
The AODVv2 router that generates the Route Request message, i.e., the router configured with OrigAddr as a Router Client.

RteMsg (Route Message)
A Route Request (RREQ) or Route Reply (RREP) message.

SeqNum
The sequence number maintained by an AODVv2 router to indicate freshness of route information.

SeqNumList
A list of sequence numbers associated with the addresses in the AddressList of a message.

TargAddr
The target address of a route request, i.e., the destination address of the IP packet triggering route discovery.

TargMetric
The metric value associated with the route to TargPrefix.

TargPrefix
The prefix configured in the Router Client entry which includes TargAddr.

TargPrefixLen
The prefix length, in bits, configured in the Router Client entry which includes TargAddr.

TargSeqNum
The sequence number of the AODVv2 router which originated the Route Reply on behalf of TargAddr.

Unreachable Address
An address reported in a Route Error message, as described in Section 8.4.1.

Upstream
In the direction from destination to source (from TargAddr to OrigAddr).

Valid route
A route that can be used for forwarding, as described in Section 8.4.1.

This document uses the notational conventions in Table 1 to simplify the text.

Notation	Meaning
Route[Address]	A route toward Address
Route[Address].Field	A field in a route toward Address
RteMsg.Field	A field in either RREQ or RREP

Table 1: Notational Conventions

3. Applicability Statement

The AODVv2 routing protocol is a reactive routing protocol intended for use in mobile ad hoc wireless networks. A reactive protocol only sends messages to discover a route when there is data to send on that route. Therefore, a reactive routing protocol requires certain interactions with the forwarding plane (for example, to indicate when a packet is to be forwarded, in order to initiate route discovery). The set of signals exchanged between AODVv2 and the forwarding plane are discussed in Section 7.4.

AODVv2 is designed for stub or disconnected mobile ad hoc networks, i.e., non-transit networks or those not connected to the internet. AODVv2 can, however, be configured to perform gateway functions when attached to external networks, as discussed in Section 10.

AODVv2 handles a wide variety of mobility and traffic patterns by determining routes on-demand. In networks with a large number of routers, AODVv2 is best suited for relatively sparse traffic scenarios where each router forwards IP packets to a small percentage of other AODVv2 routers in the network. In this case fewer routes are needed, and therefore less control traffic is produced. In large networks with very frequent or bursty traffic, AODVv2 control

messages may cause a broadcast storm, overwhelming the network with control messages and preventing routes from being established. This especially applies to networks with point-to-point or point-to-multipoint traffic. In this case, the transmission priorities described in Section 7.5 prioritize route maintenance traffic over route discovery traffic.

Data packets may be buffered until a route to their destination is available, as described in Section 7.6.

AODVv2 provides for message integrity and security against replay attacks by using integrity check values, timestamps and sequence numbers, as described in Section 13. If security associations can be established, encryption can be used for AODVv2 messages to ensure that only trusted routers participate in routing operations.

Since the route discovery process aims for a route to be established in both directions along the same path, uni-directional links are not suitable. AODVv2 will detect and exclude those links from route discovery. The route discovered is optimised for the requesting router, and the return path may not be the optimal route.

AODVv2 is applicable to memory constrained devices, since only a little routing state is maintained in each AODVv2 router. AODVv2 routes that are not needed for forwarding data do not need to be maintained. On routers unable to store persistent AODVv2 state, recovery can impose a performance penalty (e.g., in case of AODVv2 router reboot), since if a router loses its sequence number, there is a delay before the router can resume full operations. This is described in Section 7.1.

AODVv2 supports routers with multiple interfaces and multiple IP addresses per interface. A router may also use the same IP address on multiple interfaces. AODVv2 requires only that each interface configured for AODVv2 has at least one unicast IP address. Address assignment procedures are out of scope for AODVv2.

AODVv2 supports Router Clients with multiple interfaces, as long as each interface is configured with its own unicast IP address. Multi-homing of a Router Client IP address is not supported by AODVv2, and therefore an IP address SHOULD NOT be configured as a Router Client on more than one AODVv2 router at any one time.

The routing algorithm in AODVv2 MAY be operated at layers other than the network layer, using layer-appropriate addresses.

4. Purpose of the Experiment

AODVv2 is an Experimental protocol. While it is based on AODV [RFC3561], important protocol mechanisms have changed: * Bidirectionality is ensured using a new mechanism * Alternate metrics may be used to determine route quality * Support for multiple interfaces has been improved * Support for multi-interface IP addresses has been added * A new security model allowing end to end integrity checks has been added * A new message format ([RFC5444]) is used.

Many of these changes have been made quite recently, after a protocol development hiatus of several years.

Thus, the purpose of the experiment is to gain information on the behavior of these significant changes in real-world deployments, not only to learn about AODVv2 in particular, but also to further the knowledge base of reactive protocols in general.

Suitable future experiments could be:

- Evaluation of the new features mentioned above with regard to performance and functionality
- determining default values for configuration parameters such as timeouts, numbers of retries, buffer sizes, control message limits (ensuring the level of multicast traffic does not interfere with data traffic throughput)
- specification of optimisations / verification of minimum requirements for low-power or low-memory routers
- developing security strategies for different environments
- Quantification of effectiveness and performance of precursors
- Evaluation of different metric types and their suitability for reactive distance vector protocols
- Evaluation of use of an AODVv2 router as an External Network Attached Router or gateway router, including network topologies including multiple gateways.
- Achieving implementations
- multiple and interoperable
- deployments in different network types

- Analysis of the effects of buffering traffic while route discovery is in progress

- Specification of extensions to deal with timed routes, expanding ring multicast, unicast RERR to specific route precursors, accurate bidirectional metric discovery, dealing with and allowing uni-directional links and routes

The final goal of the experiment is to determine if sufficient demand exists for the AODVv2 protocol to prompt an effort to bring the protocol to Standards Track.

5. Data Structures

5.1. InterfaceSet

The InterfaceSet is a conceptual data structure which contains information about all interfaces configured for use by AODVv2. Any interface with an IP address can be used. Multiple interfaces on a single router can be used. Multiple interfaces on the same router may be configured with the same IP address.

Each element in the InterfaceSet MUST contain the following:

Interface.Id
: An identifier that is unique in node-local scope and that allows the AODVv2 implementation to identify exactly one local network interface.

If multiple interfaces of the AODVv2 router are configured for use by AODVv2, they MUST be configured in the InterfaceSet.

Implementations for constrained devices using only one interface MAY choose not to use the InterfaceSet.

5.2. Router Client Set

An AODVv2 router provides route discovery services for its own local applications and for its Router Clients that are reachable without traversing another AODVv2 router. The addresses used by these devices, and the AODVv2 router itself, are configured in the Router Client Set. An AODVv2 router will only originate Route Request and Route Reply messages on behalf of configured Router Client addresses.

Router Client Set entries MUST contain:

RouterClient.IPAddress

An IP address or the start of an address range that requires route discovery services from the AODVv2 router.

RouterClient.PrefixLength
The length, in bits, of the routing prefix associated with the RouterClient.IPAddress. If the prefix length is not equal to the address length of RouterClient.IPAddress, the AODVv2 router MUST participate in route discovery on behalf of all addresses within that prefix.

RouterClient.Cost
The cost associated with reaching this address or address range.

A Router Client address MUST NOT be served by more than one AODVv2 router at any one time. To shift responsibility for a Router Client to a different AODVv2 router, correct AODVv2 routing behavior MUST be observed; The AODVv2 router adding the Router Client MUST wait for any existing routing information about this Router Client to be purged from the network, i.e., at least MAX_SEQNUM_LIFETIME since the last SeqNum update on the router that is removing this Router Client.

5.3. Neighbor Set

A Neighbor Set MUST be maintained with information about neighboring AODVv2 routers. Neighbor Set entries are stored when AODVv2 messages are received. If the Neighbor is chosen as a next hop on an installed route, the link to the Neighbor MUST be tested for bidirectionality and the result stored in this set. A route will only be considered valid when the link is confirmed to be bidirectional.

Neighbor Set entries MUST contain:

Neighbor.IPAddress
An IP address of the neighboring router, learned from the source IP address of a received route message.

Neighbor.State
Indicates whether the link to the neighbor is bidirectional. There are three possible states: Confirmed, Heard, and Blacklisted. Heard is the initial state. Confirmed indicates that the link to the neighbor has been confirmed as bidirectional. Blacklisted indicates that the link to the neighbor is uni-directional. Section 7.2 discusses how to monitor link bidirectionality.

Neighbor.Timeout
Indicates at which time the Neighbor.State should be updated:

- If the value of Neighbor.State is Blacklisted, this indicates the time at which Neighbor.State will revert to Heard. By default this value is calculated at the time the router is blacklisted and is equal to CurrentTime + MAX_BLACKLIST_TIME.

- If Neighbor.State is Heard, and an RREP_Ack has been requested from the neighbor, it indicates the time at which Neighbor.State will be set to Blacklisted, if an RREP_Ack has not been received.

- If the value of Neighbor.State is Heard and no RREP_Ack has been requested, or if Neighbor.State is Confirmed, this time is set to INFINITY_TIME.

Neighbor.Interface
: The interface on which the link to the neighbor was established.

Neighbor.AckSeqNum
: The next sequence number to use for the TIMESTAMP value in an RREP_Ack request, in order to detect replay of an RREP_Ack response. Initially set to a random value.

Neighbor.HeardRERRSeqNum
: The last heard sequence number used as the TIMESTAMP value in a RERR received from this neighbor, saved in order to detect replay of a RERR message. Initially set to zero.

See Section 13.4.4.3 and Section 13.4.4.4 for more information on how Neighbor.AckSeqNum and Neighbor.HeardRERRSeqNum are used.

5.4. Sequence Numbers

Sequence Numbers enable AODVv2 routers to determine the temporal order of route discovery messages, identifying stale routing information so that it can be discarded. The sequence number fulfills the same roles as the "Destination Sequence Number" of DSDV [Perkins94], and the AODV Sequence Number in [RFC3561].

Each AODVv2 router in the network MUST maintain its own sequence number. All RREQ and RREP messages created by an AODVv2 router include the router’s sequence number, reported as a 16-bit unsigned integer. Each AODVv2 router MUST ensure that its sequence number is strictly increasing, and that it is incremented by one (1) whenever an RREQ or RREP is created, except when the sequence number is 65,535 (the maximum value of a 16-bit unsigned integer), in which case it MUST be reset to one (1) to achieve wrap around. The value zero (0) is reserved to indicate that the sequence number is unknown.

An AODVv2 router MUST only attach its own sequence number to information about a route to one of its configured Router Clients, all route messages forwarded by other routers retain the originator's sequence number.

To determine if newly received information is stale and therefore redundant, the sequence number attached to the information is compared to the sequence number of existing information about the same route. The comparison is carried out by subtracting the existing sequence number from the newly received sequence number, using unsigned arithmetic. The result of the subtraction is to be interpreted as a signed 16-bit integer.

- If the result is negative, the newly received information is considered older than the existing information and is considered stale and redundant and MUST therefore be discarded.

- If the result is positive, the newly received information is considered newer than the existing information and is not considered stale or redundant and MUST therefore be processed.

- If the result is zero, the newly received information is not considered stale, and therefore MUST be processed further to determine if it is redundant. For example, it is considered redundant if the metric attached to the newly received information is higher than the metric of existing information about the same route (see Section 7.7.1 and Section 7.8).

This, along with the processes in Section 7.7.1, ensures loop freedom.

An AODVv2 router SHOULD maintain its sequence number in persistent storage. If the sequence number is lost, the router MUST follow the procedure in Section 7.1 to safely resume routing operations with a new sequence number.

5.5. Local Route Set

All AODVv2 routers MUST maintain a Local Route Set, containing information about routes learned from AODVv2 route messages. The Local Route Set is stored separately from the forwarding plane's routing table (referred to as Routing Information Base (RIB)), which may be updated by other routing protocols operating on the AODVv2 router as well. The Routing Information Base is updated using information from the Local Route Set. Alternatively, implementations MAY choose to modify the Routing Information Base directly.

Routes learned from AODVv2 route messages are referred to in this document as LocalRoutes, and MUST contain the following information:

LocalRoute.Address
: An address, which, when combined with LocalRoute.PrefixLength, describes the set of destination addresses this route includes.

LocalRoute.PrefixLength
: The prefix length, in bits, associated with LocalRoute.Address.

LocalRoute.SeqNum
: The sequence number associated with LocalRoute.Address, obtained from the last route message that successfully updated this entry.

LocalRoute.NextHop
: The source IP address of the IP packet containing the AODVv2 message advertising the route to LocalRoute.Address, i.e. an IP address of the AODVv2 router used for the next hop on the path toward LocalRoute.Address.

LocalRoute.NextHopInterface
: The interface used to send IP packets toward LocalRoute.Address.

LocalRoute.LastUsed
: If this route is installed in the Routing Information Base, the time it was last used to forward an IP packet.

LocalRoute.LastSeqNumUpdate
: The time LocalRoute.SeqNum was last updated.

LocalRoute.MetricType
: The type of metric associated with this route.

LocalRoute.Metric
: The cost of the route toward LocalRoute.Address expressed in units consistent with LocalRoute.MetricType.

LocalRoute.State
: The last known state (Unconfirmed, Idle, Active, or Invalid) of the route.

There are four possible states for a LocalRoute:

Unconfirmed
: A route learned from a Route Request message, which has not yet been confirmed as bidirectional. It MUST NOT be used for forwarding IP packets, and therefore it is not referred to as a

valid route. This state only applies to routes learned through RREQ messages.

Idle
A route which has been learned from a route message, and has also been confirmed, but has not been used in the last ACTIVE_INTERVAL. It is able to be used for forwarding IP packets, and therefore it is referred to as a valid route.

Active
A route which has been learned from a route message, and has also been confirmed, and has been used in the last ACTIVE_INTERVAL. It is able to be used for forwarding IP packets, and therefore it is referred to as a valid route.

Invalid
A route which has expired or been lost. It MUST NOT be used for forwarding IP packets, and therefore it is not referred to as a valid route. Invalid routes contain sequence number information which allows incoming information to be assessed for freshness.

When the Local Route Set is stored separately from the Routing Information Base, routes are added to the Routing Information Base when LocalRoute.State is valid (set to Active or Idle), and removed from the Routing Information Base when LocalRoute.State becomes Invalid.

Changes to LocalRoute state are detailed in Section 7.10.1.

Multiple valid routes for the same address and prefix length but for different metric types may exist in the Local Route Set, but the decision of which of these routes to install in the Routing Information Base to use for forwarding is outside the scope of AODVv2.

5.6. Multicast Route Message Set

Route Request (RREQ) messages are multicast by default and forwarded multiple times. This set stores recently received RREQs in order that received RREQs can be tested for redundancy to avoid unnecessary processing and forwarding.

The Multicast Route Message Set is a conceptual set which contains information about previously received multicast route messages, so that incoming route messages can be compared with previously received messages to determine if the incoming information is redundant or stale, and the router can avoid sending redundant control traffic.

Multicast Route Message Set entries MUST contain the following information:

RteMsg.OrigPrefix
: The prefix associated with OrigAddr, the source address of the IP packet triggering the route request.

RteMsg.OrigPrefixLen
: The prefix length associated with RteMsg.OrigPrefix, originally from the Router Client entry on RREQ_Gen which includes OrigAddr.

RteMsg.TargPrefix
: The prefix associated with TargAddr, the destination address of the IP packet triggering the route request. In an RREQ this MUST be set to TargAddr.

RteMsg.OrigSeqNum
: The sequence number associated with the route to OrigPrefix, if RteMsg is an RREQ.

RteMsg.TargSeqNum
: The sequence number associated with the route to TargPrefix.

RteMsg.MetricType
: The metric type of the route requested.

RteMsg.Metric
: The metric value received in the RteMsg.

RteMsg.Timestamp
: The last time this Multicast Route Message Set entry was updated.

RteMsg.RemoveTime
: The time at which this entry MUST be removed from the Multicast Route Message Set. This is set to CurrentTime + MAX_SEQNUM_LIFETIME, whenever the RteMsg.OrigSeqNum of this entry is updated.

RteMsg.Interface
: The interface on which the message was received.

The Multicast Route Message Set is maintained so that no two entries have the same OrigPrefix, OrigPrefixLen, TargPrefix, and MetricType. See Section 7.8 for details about updating this set.

5.7. Route Error (RERR) Set

Each RERR message sent because no route exists for packet forwarding SHOULD be recorded in a conceptual set called the Route Error (RERR) Set. Each entry contains the following information:

RerrMsg.Timeout
: The time after which the entry SHOULD be deleted.

RerrMsg.UnreachableAddress
: The UnreachableAddress reported in the AddressList of the RERR.

RerrMsg.PktSource:
: The PktSource of the RERR.

See section Section 7.9 for instructions on how to update the set.

6. Metrics

Metrics measure a cost or quality associated with a route or a link, e.g., latency, delay, financial cost, energy, etc. Metric values are reported in Route Request and Route Reply messages.

In Route Request messages, the metric describes the cost of the route from OrigPrefix to the router sending the Route Request. For RREQ_Gen, this is the cost associated with the Router Client entry which includes OrigAddr. For routers which forward the RREQ, this is the cost from OrigPrefix to the forwarding router, combining the metric value from the received RREQ message with knowledge of the link cost from the sender to the receiver, i.e., the incoming link cost. This updated route cost is included when forwarding the Route Request message, and used to install a route to OrigPrefix.

Similarly, in Route Reply messages, the metric reflects the cost of the route from TargPrefix to the router sending the Route Reply. For RREP_Gen, this is the cost associated with the Router Client entry which includes TargAddr. For routers which forward the RREP, this is the cost from TargPrefix to the forwarding router, combining the metric value from the received RREP message with knowledge of the link cost from the sender to the receiver, i.e., the incoming link cost. This updated route cost is included when forwarding the Route Reply message, and used to install a route to TargPrefix.

Assuming link metrics are symmetric, the cost of the routes installed in the Local Route Set at each router will be correct. While this assumption is not always correct, calculating incoming/outgoing metric data is outside of scope of this document. The route

discovered is optimised for the requesting router, and the return path may not be the optimal route.

AODVv2 enables the use of multiple metric types. Each route discovery attempt indicates the metric type which is requested for the route. Only one metric type MUST be used in each route discovery attempt.

For each MetricType, AODVv2 requires:

- A MetricType number, to indicate the metric type of a route. MetricType numbers allocated are detailed in Section 11.5.

- A maximum value, denoted MAX_METRIC[MetricType]. This MUST always be the maximum expressible metric value of type MetricType. Field lengths associated with metric values are found in Section 11.5. If the cost of a route exceeds MAX_METRIC[MetricType], the route is ignored.

- A function for incoming link cost, denoted Cost(L). Using incoming link costs means that the route learned has a path optimized for the direction from OrigAddr to TargAddr.

- A function for route cost, denoted Cost(R).

- A function to analyze routes for potential loops based on metric information, denoted LoopFree(R1, R2). LoopFree verifies that a route R2 is not a sub-section of another route R1. An AODVv2 router invokes LoopFree() as part of the process in Section 7.7.1, when an advertised route (R1) and an existing LocalRoute (R2) have the same destination address, metric type, and sequence number. LoopFree returns FALSE to indicate that an advertised route is not to be used to update a stored LocalRoute, as it may cause a routing loop. In the case where the existing LocalRoute is Invalid, it is possible that the advertised route includes the existing LocalRoute and came from a router which did not yet receive notification of the route becoming Invalid, so the advertised route should not be used to update the Local Route Set, in case it forms a loop to a broken route.

AODVv2 currently supports cost metrics where Cost(R) is strictly increasing, by defining:

- Cost(R) := Sum of Cost(L) of each link in the route

- LoopFree(R1, R2) := (Cost(R1) <= Cost(R2))

Implementers MAY consider other metric types, but the definitions of Cost and LoopFree functions for such types are undefined, and interoperability issues need to be considered.

7. AODVv2 Protocol Operations

The AODVv2 protocol's operations include managing sequence numbers, monitoring next hop AODVv2 routers on discovered routes and updating the Neighbor Set, performing route discovery and dealing with requests from other routers, processing incoming route information and updating the Local Route Set, updating the Multicast Route Message Set and suppressing redundant messages, and reporting broken routes. These processes are discussed in detail in the following sections.

7.1. Initialization

During initialization where an AODVv2 router does not have information about its previous sequence number, or if its sequence number is lost at any point, the router resets its sequence number to one (1). However, other AODVv2 routers may still hold sequence number information that this router previously issued. Since sequence number information is removed if there has been no update to the sequence number in MAX_SEQNUM_LIFETIME, the initializing router MUST wait for MAX_SEQNUM_LIFETIME before it creates any messages containing its new sequence number. It can then be sure that the information it sends will not be considered stale.

During this wait period, the router is permitted to do the following:

- Process information in a received RREQ or RREP message to learn a route to the originator or target of that route discovery
- Forward a received RREQ or RREP
- Send an RREP_Ack
- Maintain valid routes in the Local Route Set
- Create, process and forward RERR messages

7.2. Next Hop Monitoring

To ensure AODVv2 routers Routers do not establish routes over uni-directional links, AODVv2 routers MUST verify that the link to the next hop router is bidirectional before marking a route as valid in the Local Route Set.

AODVv2 provides a mechanism for testing bidirectional connectivity during route discovery, and blacklisting routers where bidirectional connectivity is not available. If a route discovery is retried by RREQ_Gen, the blacklisted routers can be excluded from the process, and a different route can be discovered. Further, a route is not to be used for forwarding until the bidirectionality of the link to the next hop is confirmed. AODVv2 routers do not need to monitor bidirectionality for links to neighboring routers which are not used as next hops on routes in the Local Route Set.

- Bidirectional connectivity to upstream routers is tested by requesting acknowledgement of RREP messages by also sending an RREP_Ack, including an AckReq element to indicate that an acknowledgement is requested. This MUST be answered by sending an RREP_Ack in response. Receipt of an RREP_Ack within RREP_Ack_SENT_TIMEOUT proves that bidirectional connectivity exists. Otherwise, a link is determined to be unidirectional. All AODVv2 routers MUST support this process, which is explained in Section 8.2 and Section 8.3.

- For the downstream router, receipt of an RREP message containing the route to TargAddr is confirmation of bidirectionality , since an RREP message is a reply to a RREQ message which previously crossed the link in the opposite direction.

To assist with next hop monitoring, a Neighbor Set (Section 5.3) is maintained. When an RREQ or RREP is received, search for an entry in the Neighbor Set where all of the following conditions are met:

- Neighbor.IPAddress == IP address from which the RREQ or RREP was received

- Neighbor.Interface == Interface on which the RREQ or RREP was received.

If such an entry does not exist, a new entry is created as described in Section 7.3. While the value of Neighbor.State is Heard, acknowledgement of RREP messages sent to that neighbor MUST be requested. If an acknowledgement is not received within the timeout period, the neighbor MUST have Neighbor.State set to Blacklisted. If an acknowledgement is received within the timeout period, Neighbor.State is set to Confirmed. While the value of Neighbor.State is Confirmed, the request for an acknowledgement of any other RREP message is unnecessary.

When routers perform other operations such as those from the list below, these MAY be used as additional indications of connectivity:

- NHDP HELLO Messages [RFC6130]
- Route timeout
- Lower layer triggers, e.g. message reception or link status notifications
- TCP timeouts
- Promiscuous listening
- Other monitoring mechanisms or heuristics

If such an external process signals that the link to a neighbor is bidirectional, the AODVv2 router MAY update the matching Neighbor Set entry by changing the value of Neighbor.State to Confirmed, e.g. receipt of a Neighborhood Discovery Protocol HELLO message with the receiving router listed as a neighbor. If an external process signals that a link is not bidirectional, the the value of Neighbor.State MAY be changed to Blacklisted, e.g. notification of a TCP timeout.

7.3. Neighbor Set Update

On receipt of an RREQ or RREP message, the Neighbor Set MUST be checked for an entry with Neighbor.IPAddress which matches the source IP address of a packet containing the AODVv2 message. If no matching entry is found, a new entry is created.

A new Neighbor Set entry is created as follows:

- Neighbor.IPAddress := Source IP address of the received route message
- Neighbor.State := Heard
- Neighbor.Timeout := INFINITY_TIME
- Neighbor.Interface := Interface on which the RREQ or RREP was received. MUST equal Interface.Id of one of the entries in the InterfaceSet (see Section 5.1).

When an RREP_Ack is sent to a neighbor, the Neighbor Set entry is updated as follows:

- Neighbor.Timeout := CurrentTime + RREP_Ack_SENT_TIMEOUT

When a received message is one of the following:

- an RREP which answers an RREQ sent within RREQ_WAIT_TIME over the same interface as Neighbor.Interface

- an RREP_Ack response received from a Neighbor with Neighbor.State set to Heard, where Neighbor.Timeout > CurrentTime

the link to the neighbor is bidirectional and the Neighbor Set entry is updated as follows:

- Neighbor.State := Confirmed

- Neighbor.Timeout := INFINITY_TIME

When the Neighbor.Timeout is reached and Neighbor.State is Heard, then an RREP_Ack response has not been received from the neighbor within RREP_Ack_SENT_TIMEOUT of sending the RREP_Ack request. The link is considered to be uni-directional and the Neighbor Set entry is updated as follows:

- Neighbor.State := Blacklisted

- Neighbor.Timeout := CurrentTime + MAX_BLACKLIST_TIME

When the Neighbor.Timeout is reached and Neighbor.State is Blacklisted, the Neighbor Set entry is updated as follows:

- Neighbor.State := Heard

If an external mechanism reports a link as broken, the Neighbor Set entry SHOULD be removed.

Route requests from neighbors with Neighbor.State set to Blacklisted are ignored to avoid persistent IP packet loss or protocol failures. Neighbor.Timeout allows the neighbor to again be allowed to participate in route discoveries after MAX_BLACKLIST_TIME, in case the link between the routers has become bidirectional.

7.4. Interaction with the Forwarding Plane

The signals descried in the following are conceptual signals, and can be implemented in various ways. Conformant implementations of AODVv2 are not mandated to implement the forwarding plane separately from the control plane or data plane; these signals and interactions are identified simply as assistance for implementers who may find them useful.

AODVv2 requires signals from the forwarding plane:

- A packet cannot be forwarded because a route is unavailable: AODVv2 needs to know the source and destination IP addresses of the packet. If the source of the packet is configured as a Router Client, the router should initiate route discovery to the destination. If it is not a Router Client, the router should create a Route Error message.

- A packet is to be forwarded: AODVv2 needs to check the state of the route to ensure it is still valid.

- Packet forwarding succeeds: AODVv2 needs to update the record of when a route was last used to forward a packet.

- Packet forwarding failure occurs: AODVv2 needs to create a Route Error message.

AODVv2 needs to send signals to the forwarding plane:

- A route discovery is in progress: buffering might be configured for packets requiring a route, while route discovery is attempted.

- A route discovery failed: any buffered packets requiring that route should be discarded, and the source of the packet should be notified that the destination is unreachable (using an ICMP Destination Unreachable message). Route discovery fails if an RREQ cannot be generated because the control message generation limit has been reached, or if an RREP is not received within RREQ_WAIT_TIME (see Section 7.6).

- A route discovery is not permitted: any buffered packets requiring that route should be discarded. A route discovery will not be attempted if the source address of the packet needing a route is not configured as a Router Client.

- A route discovery succeeded: install a corresponding route into the Routing Information Base and begin transmitting any buffered packets.

- A route has been made invalid: remove the corresponding route from the Routing Information Base.

- A route has been updated: update the corresponding route in the Routing Information Base.

7.5. Message Transmission

AODVv2 sends [RFC5444] formatted messages using the parameters for port number and IP protocol specified in [RFC5498]. Mapping of AODVv2 data to [RFC5444] messages is detailed in Section 9. AODVv2 multicast messages are sent to the link-local multicast address LL-MANET-Routers [RFC5498]. All AODVv2 routers MUST subscribe to LL-MANET-Routers on all AODVv2 interfaces [RFC5498] to receive AODVv2 messages. Note that multicast messages MAY be sent via unicast. For example, this may occur for certain link-types (non-broadcast media), for manually configured router adjacencies, or in order to improve robustness.

When multiple interfaces are available, an AODVv2 router transmitting a multicast message to LL-MANET-Routers MUST send the message on all interfaces that have been configured for AODVv2 operation, as given in the InterfaceSet (Section 5.1).

To avoid congestion, each AODVv2 router's rate of message generation SHOULD be limited (CONTROL_TRAFFIC_LIMIT) and administratively configurable. Messages SHOULD NOT be sent more frequently than one message per (1 / CONTROL_TRAFFIC_LIMIT)th of a second. If this threshold is reached, messages MUST be sent based on their priority:

- Highest priority SHOULD be given to RREP_Ack messages. This allows links between routers to be confirmed as bidirectional and avoids undesired blacklisting of next hop routers.

- Second priority SHOULD be given to RERR messages for undeliverable IP packets. This avoids repeated forwarding of packets over broken routes that are still in use by other routers.

- Third priority SHOULD be given to RREP messages in order that RREQs do not time out.

- Fourth priority SHOULD be given to RREQ messages.

- Fifth priority SHOULD be given to RERR messages for newly invalidated routes.

- Lowest priority SHOULD be given to RERR messages generated in response to RREP messages which cannot be forwarded. In this case the route request will be retried at a later point.

To implement the congestion control, a queue length is set. If the queue is full, in order to queue a new message, a message of lower priority must be removed from the queue. If this is not possible,

the new message MUST be discarded. The queue should be sorted in order of message priority

7.6. Route Discovery, Retries and Buffering

AODVv2's RREQ and RREP messages are used for route discovery. RREQ messages are multicast to solicit an RREP, whereas RREP are unicast. The constants used in this section are defined in Section 11.

When an AODVv2 router needs to forward an IP packet (with source address OrigAddr and destination address TargAddr) from one of its Router Clients, it needs a route to TargAddr in its Routing Information Base. If no route exists, the AODVv2 router generates (RREQ_Gen) and multicasts a Route Request message (RREQ), on all configured interfaces, containing information about the source and destination. The procedure for this is described in Section 8.1.1. Each generated RREQ results in an increment to the router's sequence number. The AODVv2 router generating an RREQ is referred to as RREQ_Gen.

Buffering might be configured for IP packets awaiting a route for forwarding by RREQ_Gen, if sufficient memory is available. Buffering of IP packets might have both positive and negative effects. Real-time traffic, voice, and scheduled delivery may suffer if packets are buffered and subjected to delays, but TCP connection establishment will benefit if packets are queued while route discovery is performed [Koodli01]. Recommendations for appropriate buffer methods are out of scope for this specification. Determining which packets to discard first when the buffer is full is a matter of policy at each AODVv2 router. Note that using different or no buffer methods does not affect interoperability.

RREQ_Gen awaits reception of a Route Reply message (RREP) containing a route toward TargAddr. This can be achieved by monitoring the entry in the Multicast Route Message Table that corresponds to the generated RREQ. When CurrentTime exceeds RteMsg.Timestamp + RREQ_WAIT_TIME and no RREP has been received, RREQ_Gen will retry the route discovery.

To reduce congestion in a network, repeated attempts at route discovery for a particular target address utilize a binary exponential backoff: for each additional attempt, the time to wait for receipt of the RREP is multiplied by 2. If the requested route is not learned within the wait period, another RREQ is sent, up to a total of DISCOVERY_ATTEMPTS_MAX. This is the same technique used in AODV [RFC3561].

Through the use of bidirectional link monitoring and blacklists (see Section 7.2) uni-directional links on initial selected route will be ignored on subsequent route discovery attempts.

Route discovery is considered to have failed after DISCOVERY_ATTEMPTS_MAX and the corresponding wait time for an RREP response to the final RREQ. After the attempted route discovery has failed, RREQ_Gen waits at least RREQ_HOLDDOWN_TIME before attempting another route discovery to the same destination, in order to avoid repeatedly generating control traffic that is unlikely to discover a route. Any IP packets buffered for TargAddr are also dropped and a Destination Unreachable ICMP message (Type 3) with a code of 1 (Host Unreachable Error) is delivered to the source of the packet, so that the application knows about the failure.

If RREQ_Gen does receive a route message containing a route to TargAddr within the timeout, it processes the message according to Section 8. When a valid LocalRoute entry is created in the Local Route Set, the route is also installed in the Routing Information Base, and the router will begin sending the buffered IP packets. Any retry timers for the corresponding RREQ are then cancelled.

During route discovery, all routers on the path learn a route to both OrigPrefix and TargPrefix, so that routes are constructed in both directions. The route is optimized for the forward route.

7.7. Processing Received Route Information

All AODVv2 route messages contain a route. A Route Request (RREQ) contains a route to OrigPrefix, and a Route Reply (RREP) contains a route to TargPrefix. All AODVv2 routers that receive a route message are able to store the route contained within it in their Local Route Set. Incoming information is first checked to verify that it is both safe to use and offers an improvement to existing information, as explained in Section 7.7.1. If these checks pass, the Local Route Set MUST be updated according to Section 7.7.2.

In the processes below, RteMsg is used to denote the route message, AdvRte is used to denote the route contained within it, and LocalRoute denotes an existing entry in the Local Route Set which matches AdvRte on address, prefix length, and metric type.

AdvRte has the following properties:

- AdvRte.Address := RteMsg.OrigPrefix (in RREQ) or RteMsg.TargPrefix (in RREP)

- AdvRte.PrefixLength := RteMsg.OrigPrefixLen (in RREQ) or RteMsg.TargPrefixLen (in RREP). If no prefix length was included in RteMsg, prefix length is the address length, in bits, of RteMsg.OrigPrefix (in RREQ) or RteMsg.TargPrefix (in RREP)

- AdvRte.SeqNum := RteMsg.OrigSeqNum (in RREQ) or RteMsg.TargSeqNum (in RREP)

- AdvRte.NextHop := RteMsg.IPSourceAddress (an address of the sending interface of the router from which the RteMsg was received)

- AdvRte.MetricType := RteMsg.MetricType

- AdvRte.Metric := RteMsg.Metric

- AdvRte.Cost := Cost(R) using the cost function associated with the route's metric type, i.e. Cost(R) = AdvRte.Metric + Cost(L), as described in Section 6, where L is the link from the advertising router

7.7.1. Evaluating Route Information

An incoming advertised route (AdvRte) is compared to existing LocalRoutes to determine whether the advertised route is to be used to update the AODVv2 Local Route Set. The incoming route information MUST be processed as follows:

1. Search for LocalRoutes in the Local Route Set matching AdvRte's address, prefix length and metric type

    * If no matching LocalRoute exists, AdvRte MUST be used to update the Local Route Set and no further checks are required.

    * If matching LocalRoutes are found, continue to Step 2.

2. Compare sequence numbers using the technique described in Section 5.4

    * If AdvRte is more recent than all matching LocalRoutes, AdvRte MUST be used to update the Local Route Set and no further checks are required.

    * If AdvRte is stale, AdvRte MUST NOT be used to update the Local Route Set. Ignore AdvRte for further processing.

    * If the sequence numbers are equal, continue to Step 3.

3. Check that AdvRte is safe against routing loops compared to all matching LocalRoutes (see Section 6)

   * If LoopFree(AdvRte, LocalRoute) returns FALSE, ignore AdvRte for further processing. AdvRte MUST NOT be used to update the Local Route Set because using the incoming information might cause a routing loop.

   * If LoopFree(AdvRte, LocalRoute) returns TRUE, continue to Step 4.

4. Compare route costs

   * If AdvRte is better than all matching LocalRoutes, it MUST be used to update the Local Route Set because it offers improvement.

   * If AdvRte is equal in cost and LocalRoute is valid, AdvRte SHOULD NOT be used to update the Local Route Set because it will offer no improvement.

   * If AdvRte is worse and LocalRoute is valid, ignore AdvRte for further processing. AdvRte MUST NOT be used to update the Local Route Set because it does not offer any improvement.

   * If AdvRte is not better (i.e., it is worse or equal) but LocalRoute is Invalid, AdvRte SHOULD be used to update the Local Route Set because it can safely repair the existing Invalid LocalRoute.

If the advertised route is to be used to update the Local Route Set, the procedure in Section 7.7.2 MUST be followed. If not, non-optimal routes will remain in the Local Route Set.

For information on how to apply these changes to the Routing Information Base, see Section 5.5.

7.7.2. Applying Route Updates

After determining that AdvRte is to be used to update the Local Route Set (as described in Section 7.7.1), the following procedure applies.

If AdvRte is learned from an RREQ message, the link to the next hop neighbor may not be confirmed as bidirectional (see Section 5.3). If there is no existing matching route in the Local Route Set, AdvRte MUST be installed to allow a corresponding RREP to be sent. If a matching entry already exists, AdvRte offers potential improvement, if the link to the neighbor can be confirmed as bidirectional.

The route update is applied as follows:

1. If no existing entry in the Local Route Set matches AdvRte's address, prefix length and metric type, continue to Step 4 and create a new entry in the Local Route Set.

2. If two matching LocalRoutes exist in the Local Route Set, one is a valid route, and one is an Unconfirmed route, AdvRte may offer further improvement to the Unconfirmed route, or may offer an update to the valid route.

   * If AdvRte.NextHop's Neighbor.State is Heard, the advertised route may offer improvement to the existing valid route, if the link to the next hop can be confirmed as bidirectional. Continue processing from Step 5 to update the existing Unconfirmed LocalRoute.

   * If AdvRte.NextHop's Neighbor.State is Confirmed, the advertised route offers an update or improvement to the existing valid route. Continue processing from Step 5 to update the existing valid LocalRoute.

3. If only one matching LocalRoute exists in the Local Route Set:

   * If AdvRte.NextHop's Neighbor.State is Confirmed, continue processing from Step 5 to update the existing LocalRoute.

   * If AdvRte.NextHop's Neighbor.State is Heard, AdvRte may offer improvement the existing LocalRoute, if the link to AdvRte.NextHop can be confirmed as bidirectional.

   * If LocalRoute.State is Unconfirmed, AdvRte is an improvement to an existing Unconfirmed route. Continue processing from Step 5 to update the existing LocalRoute.

   * If LocalRoute.State is Invalid, AdvRte can replace the existing LocalRoute. Continue processing from Step 5 to update the existing LocalRoute.

   * If LocalRoute.State is Active or Idle, AdvRte SHOULD be stored as an additional entry in the Local Route Set, with LocalRoute.State set to Unconfirmed. Continue processing from Step 4 to create a new LocalRoute.

4. Create an entry in the Local Route Set and initialize as follows:

   * LocalRoute.Address := AdvRte.Address

* LocalRoute.PrefixLength := AdvRte.PrefixLength

* LocalRoute.MetricType := AdvRte.MetricType

5. Update the LocalRoute as follows:

   * LocalRoute.SeqNum := AdvRte.SeqNum

   * LocalRoute.NextHop := AdvRte.NextHop

   * LocalRoute.NextHopInterface := interface on which RteMsg was received

   * LocalRoute.Metric := AdvRte.Cost

   * LocalRoute.LastUsed := CurrentTime

   * LocalRoute.LastSeqNumUpdate := CurrentTime

6. If a new LocalRoute was created, or if the existing LocalRoute.State is Invalid or Unconfirmed, update LocalRoute as follows:

   * LocalRoute.State := Unconfirmed (if the next hop’s Neighbor.State is Heard)

   * LocalRoute.State := Idle (if the next hop’s Neighbor.State is Confirmed)

7. If an existing LocalRoute.State changed from Invalid or Unconfirmed to become Idle, any matching Unconfirmed LocalRoute with worse metric value SHOULD be expunged.

8. If an existing LocalRoute was updated with a better metric value, any matching Unconfirmed LocalRoute with worse metric value SHOULD be expunged.

9. If this update results in LocalRoute.State of Active or Idle, which matches a route request which is still in progress, the associated route request retry timers MUST be cancelled.

If this update to the Local Route Set results in two LocalRoutes to the same address, the best LocalRoute will be Unconfirmed. In order to improve the route used for forwarding, the router SHOULD try to determine if the link to the next hop of that LocalRoute is bidirectional, by using that LocalRoute to forward future RREPs and request acknowledgements (see Section 8.2.1 and Section 8.3.

7.8. Suppressing Redundant Messages Using the Multicast Route Message Set

When route messages are flooded in a MANET, an AODVv2 router may receive several instances of the same message. Forwarding every one of these gives little additional benefit, and generates unnecessary signaling traffic and might generate unnecessary interference.

Each AODVv2 router stores information about recently received route messages in the AODVv2 Multicast Route Message Set (Section 5.6).

Entries in the Multicast Route Message Set SHOULD be maintained for at least RteMsg_ENTRY_TIME after the last Timestamp update in order to account for long-lived RREQs traversing the network. An entry MUST be deleted when the sequence number is no longer valid, i.e., after MAX_SEQNUM_LIFETIME. Memory-constrained devices MAY remove the entry before this time.

Received route messages are tested against previously received route messages, and if determined to be redundant, forwarding or response can be avoided.

To determine if a received message is redundant:

1. Search for an entry in the Multicast Route Message Set with the same OrigPrefix, OrigPrefixLen, TargPrefix, Interface and MetricType

    * If there is no entry, the message is not redundant.

    * If there is an entry, continue to Step 2.

2. Compare sequence numbers using the technique described in Section 5.4

    * Use OrigSeqNum of the entry for comparison.

    * If the entry has an older sequence number than the received message, the message is not redundant.

    * If the entry has a newer sequence number than the received message, the message is redundant.

    * If the entry has the same sequence number, continue to Step 3.

3. Compare the metric values

  * If the entry has a Metric value that is worse than or equal to the metric in the received message, the message is redundant.

  * If the entry has a Metric value that is better than the metric in the received message, the message is not redundant.

If the message is redundant, update the entry as follows:

- RteMsg.Timestamp := CurrentTime

- RteMasg.RemoveTime := CurrentTime + MAX_SEQNUM_LIFETIME

since matching route messages are still traversing the network and this entry should be maintained. This message MUST NOT be forwarded or responded to.

If the message is not redundant, create an entry or update the existing entry.

To update a Multicast Route Message Set entry, set:

- RteMsg.OrigPrefix := OrigPrefix from the message

- RteMsg.OrigPrefixLen := the prefix length associated with OrigPrefix

- RteMsg.TargPrefix := TargPrefix from the message

- RteMsg.OrigSeqNum := the sequence number associated with OrigPrefix, if RteMsg is an RREQ

- RteMsg.TargSeqNum := the sequence number associated with TargPrefix, if RteMsg is an RREP

- RteMsg.Metric := the metric value associated with OrigPrefix in a received RREQ

- RteMsg.MetricType := the metric type associated with RteMsg.Metric

- RteMsg.Timestamp := CurrentTime

- RteMsg.RemoveTime := CurrentTime + MAX_SEQNUM_LIFETIME

Where the message is determined not redundant before Step 3, it MUST be forwarded or responded to. When a message is determined to be not redundant in Step 3, it MAY be suppressed to avoid extra control traffic. However, since the processing of the message will result in an update to the Local Route Set, the message SHOULD be forwarded or

responded to, to ensure other routers have up-to-date information and the best metrics. If the message is not forwarded, the best route may not be found. Forwarding or response is to be performed using the processes outlined in Section 8.

7.9. Suppressing Redundant Route Error Messages using the Route Error Set

In order to avoid flooding the network with RERR messages when a stream of IP packets to an unreachable address arrives, an AODVv2 router SHOULD avoid creating duplicate messages by determining whether an equivalent RERR has recently been sent. This is achieved with the help of the Route Error Set (see Section 5.7).

To determine if a RERR should be created:

1. Search for an entry in the Route Error Set where:

   * RerrMsg.UnreachableAddress == UnreachableAddress to be reported

   * RerrMsg.PktSource == PktSource to be included in the RERR

   If a matching entry is found, no further processing is required and the RERR SHOULD NOT be sent.

2. If no matching entry is found, a new entry with the following properties is created, and the RERR is created and sent as described in Section 8.4.1:

   * RerrMsg.Timeout := CurrentTime + RERR_TIMEOUT

   * RerrMsg.UnreachableAddress == UnreachableAddress to be reported

   * RerrMsg.PktSource == PktSource to be included in the RERR

7.10. Local Route Set Maintenance

Route maintenance involves monitoring LocalRoutes in the Local Route Set, updating LocalRoute.State to handle route timeouts and reporting routes that become Invalid.

7.10.1. LocalRoute State Changes

During normal operation, AODVv2 does not require any explicit timeouts to manage the lifetime of a route. At any time, any LocalRoute MAY be examined and updated according to the rules below.

If timers are not used to prompt updates of LocalRoute.State, the LocalRoute.State MUST be checked before IP packet forwarding and before any operation based on LocalRoute.State.

Route timeout behaviour is as follows:

- An Unconfirmed route MUST be expunged at MAX_SEQNUM_LIFETIME after LocalRoute.LastSeqNumUpdate.

- An Idle route MUST become Active when used to forward an IP packet. If the route is not used to forward an IP packet within MAX_IDLETIME, LocalRoute.State MUST become Invalid.

- An Invalid route SHOULD remain in the Local Route Set, since LocalRoute.SeqNum is used to classify future information about LocalRoute.Address as stale or fresh.

- In all cases, if the time since LocalRoute.LastSeqNumUpdate exceeds MAX_SEQNUM_LIFETIME, LocalRoute.SeqNum must be set to

1. This is required to ensure that any AODVv2 routers following the initialization procedure can safely begin routing functions using a new sequence number. A LocalRoute with LocalRoute.State set to Active or Idle can remain in the Local Route Set after the sequence number has been set to 0, for example if the route is reliably carrying traffic. If LocalRoute.State is Invalid, or later becomes Invalid, the LocalRoute MUST be expunged from the Local Route Set.

LocalRoutes can become Invalid before a timeout occurs:

- If an external mechanism reports a link as broken, all LocalRoutes using that link for LocalRoute.NextHop MUST immediately have LocalRoute.State set to Invalid.

- LocalRoute.State MUST immediately be set to Invalid if a Route Error (RERR) message is received where:

  * The sender is LocalRoute.NextHop or PktSource is a Router Client address

  * There is an Address in AddressList which matches LocalRoute.Address, and:

    + The prefix length associated with this Address, if any, matches LocalRoute.PrefixLength

+ The sequence number associated with this Address, if any, is newer or equal to LocalRoute.SeqNum (see Section 5.4)

+ The metric type associated with this Address matches LocalRoute.MetricType

LocalRoutes are also updated when Neighbor.State is updated:

- While the value of Neighbor.State is set to Heard, any routes in the Local Route Set using that neighbor as a next hop MUST have LocalRoute.State set to Unconfirmed.

- When the value of Neighbor.State is set to Confirmed, the Unconfirmed routes in the Local Route Set using that neighbor as a next hop MUST have LocalRoute.State set to Idle. Any other matching LocalRoutes with metric values worse than LocalRoute.Metric MUST be expunged from the Local Route Set.

- When the value of Neighbor.State is set to Blacklisted, any valid routes in the Local Route Set using that neighbor for their next hop MUST have LocalRoute.State set to Invalid.

- When a Neighbor Set entry is removed, all routes in the Local Route Set using that neighbor as next hop MUST have LocalRoute.State set to Invalid.

Memory constrained devices MAY choose to expunge routes from the AODVv2 Local Route Set at other times, but MUST adhere to the following rules:

- An Active route MUST NOT be expunged, as it is in use. If deleted, IP traffic forwarded to this router will prompt generation of a Route Error message, and it will be necessary for a Route Request to be generated by the originator’s router to re-establish the route.

- An Idle route SHOULD NOT be expunged, as it is still valid for forwarding IP traffic. If deleted, this could result in dropped IP packets and a Route Request could be generated to re-establish the route.

- Any Invalid route MAY be expunged. Least recently used Invalid routes SHOULD be expunged first, since the sequence number information is less likely to be useful.

- An Unconfirmed route MUST NOT be expunged if it was installed within the last RREQ_WAIT_TIME, because it may correspond to a route discovery in progress. A Route Reply message might be

received which needs to use the LocalRoute.NextHop information. Otherwise, it MAY be expunged.

7.10.2. Reporting Invalid Routes

When LocalRoute.State changes from Active to Invalid as a result of a broken link or a received Route Error (RERR) message, other AODVv2 routers MUST be informed by sending an RERR message containing details of the invalidated route.

An RERR message MUST also be sent when an AODVv2 router receives an RREP message to forward, but the LocalRoute to the OrigPrefix in the RREP has been lost or is marked as Invalid.

An RERR message MUST also be sent when an AODVv2 router receives an RREP message to forward, but the LocalRoute to the OrigAddr in the RREP has been lost or is marked as Invalid.

The packet or message triggering the RERR MUST be discarded.

Generation of an RERR message is described in Section 8.4.1.

8. AODVv2 Protocol Messages

AODVv2 defines four message types: Route Request (RREQ), Route Reply (RREP), Route Reply Acknowledgement (RREP_Ack), and Route Error (RERR).

Each AODVv2 message is defined as a set of data. Rules for the generation, reception and forwarding of each message type are described in the following sections. Section 9 discusses how the data is mapped to [RFC5444] Message TLVs, Address Blocks, and Address TLVs.

8.1. Route Request (RREQ) Message

Route Request messages are used in route discovery operations to request a route to a specified target address. RREQ messages have the following contents:

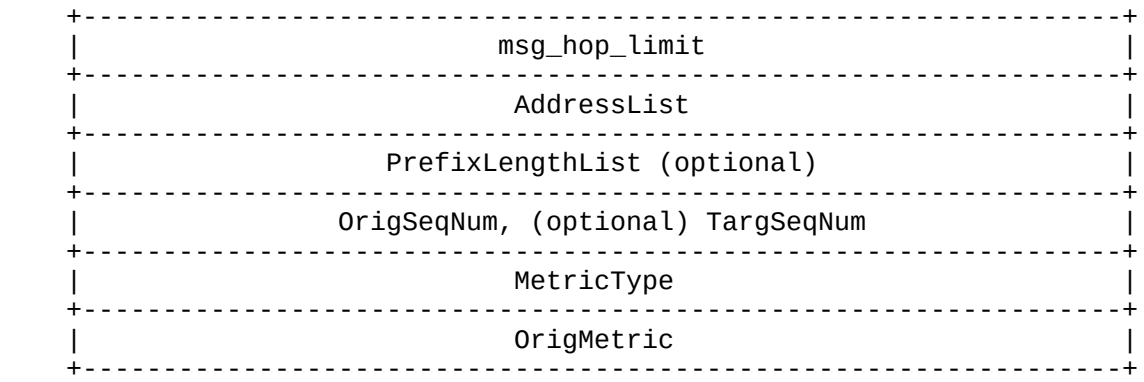


Figure 1: RREQ message contents

msg_hop_limit
The remaining number of hops allowed for dissemination of the RREQ message.

AddressList
Contains OrigPrefix, from the Router Client entry which includes OrigAddr, the source address of the IP packet for which a route is requested, and TargPrefix, set to TargAddr, the destination address of the IP packet for which a route is requested.

PrefixLengthList
Contains OrigPrefixLen, i.e., the length, in bits, of the prefix associated with the Router Client entry which includes OrigAddr. If omitted, the prefix length is equal to OrigAddr's address length in bits.

OrigSeqNum
The sequence number associated with OrigPrefix.

TargSeqNum
A sequence number associated with an existing Invalid route to TargAddr. This MAY be included if available.

MetricType
The metric type associated with OrigMetric.

OrigMetric
The metric value associated with the route to OrigPrefix, as seen from the sender of the message.

8.1.1. RREQ Generation

An RREQ is generated when an IP packet needs to be forwarded for a Router Client, and no valid route currently exists for the packet's destination in the Routing Information Base.

Before creating an RREQ, the router SHOULD check the Multicast Route Message Set to see if an RREQ has recently been sent for the requested destination. If so, and the wait time for a reply has not yet been reached, the router SHOULD continue to await a response without generating a new RREQ. If the timeout has been reached, a new RREQ MAY be generated. If buffering is configured, incoming IP packets awaiting this route SHOULD be buffered until the route discovery is completed.

If the limit for the rate of AODVv2 control message generation has been reached, no message SHOULD be generated.

To generate the RREQ, the router (referred to as RREQ_Gen) follows this procedure:

1. Set msg_hop_limit := MAX_HOPCOUNT

2. Set AddressList := {OrigPrefix, TargPrefix}

3. For the PrefixLengthList:

    * If OrigAddr is part of an address range configured as a Router Client, set PrefixLengthList := {RouterClient.PrefixLength, null}.

    * Otherwise, omit PrefixLengthList.

4. For OrigSeqNum:

    * Increment the router Sequence Number as specified in Section 5.4.

    * Set OrigSeqNum := router Sequence Number.

5. For TargSeqNum:

    * If an Invalid route exists in the Local Route Set matching TargAddr using longest prefix matching and has a valid sequence number, set TargSeqNum := LocalRoute.SeqNum.

* If no Invalid route exists in the Local Route Set matching TargAddr, or the route doesn't have a sequence number, omit TargSeqNum.

6. Include MetricType and set the type accordingly

7. Find the Router Client Set Entry where RouterClient.IPAddress == OrigPrefix:

   * Set OrigMetric := RouterClient.Cost

This AODVv2 message is used to create a corresponding [RFC5444] message (see Section 9) which is handed to the RFC5444 multiplexer for further processing. By default, the multiplexer is instructed to multicast the message to LL-MANET- Routers on all interfaces configured for AODVv2 operation. The RREP MUST be sent over LocalRoute[OrigPrefix].NextHopInterface.

8.1.2. RREQ Reception

Upon receiving a Route Request, an AODVv2 router performs the following steps:

1. Check and update the Neighbor Set according to Section 7.3

   * If the sender has Neighbor.State set to Blacklisted, ignore this RREQ for further processing.

2. Verify that the message contains the required data: msg_hop_limit, OrigPrefix, TargPrefix, OrigSeqNum, and OrigMetric, and that OrigPrefix and TargPrefix are valid addresses

   * If not, ignore this RREQ for further processing.

3. Check that the MetricType is supported and configured for use

   * If not, ignore this RREQ for further processing.

4. Verify that the cost of the advertised route will not exceed the maximum allowed metric value for the metric type (Metric <= MAX_METRIC[MetricType] - Cost(L))

   * If it will, ignore this RREQ for further processing.

5. Process the route to OrigPrefix as specified in Section 7.7

6. Check if the information in the message is redundant by comparing to entries in the Multicast Route Message Set, following the procedure in Section 7.8

   * If redundant, ignore this RREQ for further processing.

   * If not redundant, create a new entry in the Multicast Route Message Set and continue processing.

7. Check if the TargPrefix matches an entry in the Router Client Set

   * If so, generate an RREP as specified in Section 8.2.1.

   * If not, continue to RREQ forwarding.

8.1.3. RREQ Forwarding

By forwarding an RREQ, a router advertises that it will forward IP packets to the OrigPrefix contained in the RREQ according to the information enclosed.The router MAY choose not to forward the RREQ, for example if the router is heavily loaded or low on energy and therefore unwilling to advertise routing capability for more traffic. This could, however, decrease connectivity in the network or result in non-optimal paths.

The RREQ SHOULD NOT be forwarded if the limit for the rate of AODVv2 control message generation has been reached.

The procedure for RREQ forwarding is as follows:

1. Set msg_hop_limit := received msg_hop_limit - 1

2. If msg_hop_limit is now zero, do not continue the forwarding process

3. Set OrigMetric := LocalRoute[OrigPrefix].Metric

This modified message is handed to the [RFC5444] multiplexer for further processing. By default, the multiplexer is instructed to multicast the message to LL-MANET-Routers on all interfaces configured for AODVv2 operation.

8.2. Route Reply (RREP) Message

When a Route Request message is received, requesting a route to a target address (TargAddr) which is configured as part of a Router Client entry, a Route Reply message is sent in response. The RREP offers a route to TargPrefix.

RREP messages have the following contents:

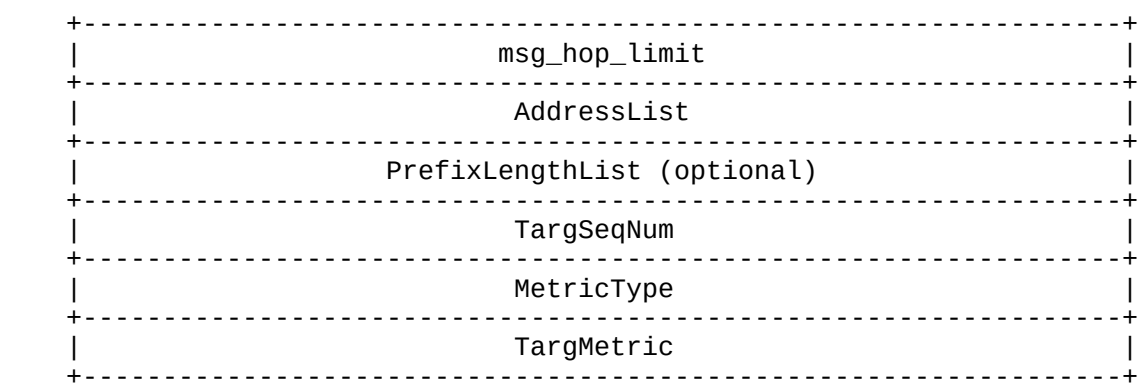


Figure 2: RREP message contents

msg_hop_limit
The remaining number of hops allowed for dissemination of the RREP message.

AddressList
Contains OrigPrefix and TargPrefix, the prefixes of the source and destination addresses of the IP packet for which a route is requested.

PrefixLengthList
Contains TargPrefixLen, i.e., the length, in bits, of the prefix associated with the Router Client entry which includes TargAddr. If omitted, the prefix length is equal to TargAddr's address length, in bits.

TargSeqNum
The sequence number associated with TargPrefix.

MetricType
The metric type associated with TargMetric.

TargMetric
The metric value associated with the route to TargPrefix, as seen from the sender of the message.

8.2.1. RREP Generation

A Route Reply message is generated when a Route Request for a Router Client of the AODVv2 router arrives. This is the case when

RteMsg.TargPrefix matches an entry in the Router Client Set of the AODVv2 router.

Before creating an RREP, the router SHOULD check if CONTROL_TRAFFIC_LIMIT has been reached. If so, the RREP SHOULD NOT be created.

The RREP will follow the path of the route to OrigPrefix. If the best route to OrigPrefix in the Local Route Set is Unconfirmed, the link to the next hop neighbor is not yet confirmed as bidirectional (as described in Section 7.2). In this case an RREP_Ack MUST also be sent as described in Section 8.3, in order to request an acknowledgement message from the next hop router to prove that the link is bidirectional. If the best route to OrigPrefix in the Local Route Set is valid, the link to the next hop neighbor is already confirmed as bidirectional, and no acknowledgement is required.

Implementations MAY allow a number of retries of the RREP if a requested acknowledgement is not received within RREP_Ack_SENT_TIMEOUT, doubling the timeout with each retry, up to a maximum of RREP_RETRIES, using the same exponential backoff described in Section 7.6 for RREQ retries. The acknowledgement MUST be considered to have failed after the wait time for an RREP_Ack response to the final RREP.

To generate the RREP, the router (also referred to as RREP_Gen) follows this procedure:

1. Set msg_hop_limit := MAX_HOPCOUNT - msg_hop_limit from the received RREQ message

2. Set Address List := {OrigPrefix, TargPrefix}

3. For the PrefixLengthList:

    * If TargAddr is part of an address range configured as a Router Client, set PrefixLengthList := {null, RouterClient.PrefixLength}.

    * Otherwise, omit PrefixLengthList.

4. For the TargSeqNum:

    * Increment the router Sequence Number as specified in Section 5.4.

    * Set TargSeqNum := router Sequence Number.

5. Include MetricType and set the type to match the MetricType in the received RREQ message

6. Set TargMetric := RouterClient.Cost for the Router Client entry which includes TargAddr

This AODVv2 message is used to create a corresponding [RFC5444] message (see Section 9) which is handed to the RFC5444 multiplexer for further processing. The multiplexer is instructed to unicast the RREP to LocalRoute[OrigPrefix].NextHop. The RREP MUST be sent over LocalRoute[OrigPrefix].NextHopInterface.

8.2.2. RREP Reception

Upon receiving a Route Reply, an AODVv2 router performs the following steps:

1. Verify that the message contains the required data: msg_hop_limit, OrigPrefix, TargPrefix, TargSeqNum, and TargMetric, and that OrigPrefix and TargPrefix are valid addresses

    * If not, ignore this RREP for further processing.

2. Check that the MetricType is supported and configured for use

    * If not, ignore this RREP for further processing. <!--

3. If this RREP does not correspond to an RREQ generated or forwarded in the last RREQ_WAIT_TIME, ignore for further processing. -->

4. If the Multicast Route Message Set does not contain an entry where:

- RteMsg.OrigPrefix == RREP.OrigPrefix
- RteMsg.OrigPrefixLen == RREP.OrigPrefixLen
- RteMsg.TargAddr exists within RREP.TargPrefix
- RteMsg.OrigSeqNum <= RREP.OrigSeqNum
- RteMsg.MetricType == RREP.MetricType
- RteMsg.Timestamp > CurrentTime - RREQ_WAIT_TIME
- RteMsg.Interface == The interface on which the RREP was received

ignore this RREP for further processing, since it does not correspond to a previously sent RREQ.

1. Update the Neighbor Set according to Section 7.3

2. Verify that the cost of the advertised route does not exceed the maximum allowed metric value for the metric type (Metric <= MAX_METRIC[MetricType] - Cost(L))

   * If it does, ignore this RREP for further processing.

3. Process the route to TargPrefix as specified in Section 7.7

4. Check if the message is redundant by comparing to entries in the Multicast Route Message Set (Section 7.8)

   * If redundant, ignore this RREP for further processing.

   * If not redundant, save the information in the Multicast Route Message Set to identify future redundant RREP messages and continue processing.

5. Check if the OrigPrefix matches an entry in the Router Client Set

   * If so, no further processing is necessary.

   * If not, continue to Step 10.

6. Check if a valid (Active or Idle) or Unconfirmed LocalRoute exists to OrigPrefix

   * If so, continue to RREP forwarding.

   * If not, a Route Error message SHOULD be transmitted toward TargPrefix according to Section 8.4.1 and the RREP SHOULD be discarded and not forwarded.

8.2.3. RREP Forwarding

A received Route Reply message is forwarded toward OrigPrefix. By forwarding an RREP, a router advertises that it will forward IP packets to TargPrefix.

The RREP SHOULD NOT be forwarded if CONTROL_TRAFFIC_LIMIT has been reached. Otherwise, the router MUST forward the RREP.

The procedure for RREP forwarding is as follows:

1. Set msg_hop_limit := received msg_hop_limit - 1

2. If msg_hop_limit is now zero, do not continue the forwarding process

3. Set TargMetric := LocalRoute[TargPrefix].Metric

This modified message is handed to the [RFC5444] multiplexer for further processing. The multiplexer is instructed to unicast the RREP to LocalRoute[OrigPrefix].NextHop. The RREP MUST be sent over LocalRoute[OrigPrefix].NextHopInterface.

8.3. Route Reply Acknowledgement (RREP_Ack) Message

The Route Reply Acknowledgement is used as both a request and a response message to test bidirectionality of a link over which a Route Reply has also been sent. The router which forwards the RREP MUST send a Route Reply Acknowledgement message to the intended next hop, if the link to the next hop neighbor is not yet confirmed as bidirectional.

The receiving router MUST then reply with a Route Reply Acknowledgement response message.

When the Route Reply Acknowledgement response message is received by the sender of the RREP, it confirms that the link between the two routers is bidirectional (see Section 7.2).

If the Route Reply Acknowledgement is not received within RREP_Ack_SENT_TIMEOUT, the link is determined to be unidirectional.

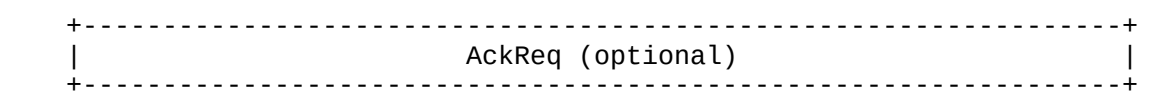


Figure 3: RREP_Ack message contents

8.3.1. RREP_Ack Request Generation

An RREP_Ack MUST be generated if a Route Reply is sent over a link which is not known to be bidirectional. It includes an AckReq element to indicate that it is a request for acknowledgement.

The RREP_Ack SHOULD NOT be generated if the limit for the rate of AODVv2 control message generation has been reached.

The [RFC5444] representation of the RREP_Ack is discussed in Section 9.

The RREP_Ack request MUST be sent unicast to the LocalRoute[OrigPrefix].NextHop via LocalRoute[OrigPrefix].NextHopInterface. The multiplexer MAY be instructed to send the RREP_Ack in the same [RFC5444] packet as the RREP.

The Neighbor Set entry for LocalRoute[OrigPrefix].NextHop MUST also be updated to indicate that an RREP_Ack is required (see Section 7.3).

8.3.2. RREP_Ack Reception

Upon receiving an RREP_Ack, an AODVv2 router performs the following steps:

1. Check if an AckReq element is included:
    * If so, create an RREP_Ack Response as described in Section 8.3.3. No further processing is required.
    * If not, continue to step 2.
2. Check if the RREP_Ack was expected:
    * Check if the Neighbor Set contains an entry where:
        + Neighbor.IPAddress == IP.SourceAddress of the RREP_Ack message
        + Neighbor.State == Heard
        + Neighbor.Timeout < CurrentTime
        + Neighbor.Interface matches the interface on which the RREP_Ack was received
    * If it does, the router sets Neighbor.Timeout to INFINITY_TIME, and processing continues to Step 3.
    * Otherwise no actions are required and processing ends.
3. Update the Neighbor Set according to Section 7.3, including updating routes using this Neighbor as LocalRoute.NextHop.

8.3.3. RREP_Ack Response Generation

An RREP_Ack response MUST be generated if a received RREP_Ack includes an AckReq.

The RREP_Ack response SHOULD NOT be generated if the limit for the rate of AODVv2 control message generation has been reached.

There is no further data in an RREP_Ack response. The [RFC5444] representation is discussed in Section 9. In this case, the multiplexer is instructed to unicast the RREP_Ack to the source IP address of the RREP_Ack message that requested it, over the same interface on which the RREP_Ack was received.

8.4. Route Error (RERR) Message

A Route Error message is generated by an AODVv2 router to notify other AODVv2 routers of routes that are no longer available. An RERR message has the following contents:

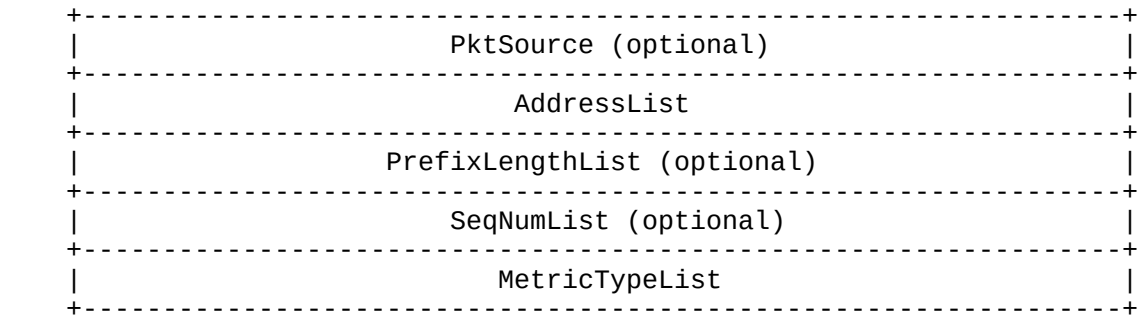


Figure 4: RERR message contents

PktSource
: The source address of the IP packet triggering the RERR. If the RERR is triggered by a broken link, PktSource is not required.

AddressList
: The addresses of the routes not available through RERR_Gen.

PrefixLengthList
: The prefix lengths, in bits, associated with the routes not available through RERR_Gen. These values indicate whether routes represent a single device or an address range.

SeqNumList

The sequence numbers of the routes not available through RERR_Gen (where known).

MetricTypeList
The metric types associated with the routes not available through RERR_Gen.

8.4.1. RERR Generation

A Route Error message is generated when an AODVv2 router (also referred to as RERR_Gen) needs to report that a destination is not reachable. There are three events that cause this response:

- When an IP packet that has been forwarded from another router, but cannot be forwarded further because there is no valid route in the Routing Information Base for its destination, the source of the packet needs to be informed that the route to the destination of the packet does not exist. The RERR generated MUST include PktSource set to the source address of the IP packet, and MUST contain only one unreachable address in the AddressList, i.e., the destination address of the IP packet. RERR_Gen MUST discard the IP packet that triggered generation of the RERR. The prefix length, sequence number and metric type SHOULD be included if known from an existing Invalid LocalRoute to the unreachable address.

- When an RREP message cannot be forwardeded because the LocalRoute to OrigPrefix has been lost or is Invalid, RREP_Gen needs to be informed that the route to OrigPrefix does not exist. The RERR generated MUST include PktSource set to the TargPrefix of the RREP, and MUST contain only one unreachable address in the AddressList, the OrigPrefix from the RREP. RERR_Gen MUST discard the RREP message that triggered generation of the RERR. The prefix length, sequence number and metric type SHOULD be included if known from an Invalid LocalRoute to the unreachable address.

- When a link breaks, multiple LocalRoutes may become Invalid, and the RERR generated MAY contain multiple unreachable addresses. The RERR MUST include MetricTypeList. PktSource is omitted. All previously Active LocalRoutes that used the broken link MUST be reported. The AddressList, PrefixLengthList, SeqNumList, and MetricTypeList will contain entries for each LocalRoute which has become Invalid. An RERR message is only sent if an Active LocalRoute becomes Invalid, though an AODVv2 router can also include Idle LocalRoutes that become Invalid if the configuration parameter ENABLE_IDLE_IN_RERR is set (see Section 11.3).

The RERR SHOULD NOT be generated if CONTROL_TRAFFIC_LIMIT has been reached. The RERR also SHOULD NOT be generated if it is a duplicate, as determined by Section 7.9.

Incidentally, if an AODVv2 router receives an ICMP error packet to or from the address of one of its Router Clients, it forwards the ICMP packet in the same way as any other IP packet, and will not generate any RERR message based on the contents of the ICMP packet.

To generate the RERR, the router follows this procedure:

1. If necessary, include PktSource and set the value as given above

2. For each LocalRoute that needs to be reported:

    * Insert LocalRoute.Address into the AddressList.

    * Insert LocalRoute.PrefixLength into PrefixLengthList, if known and not equal to the address length.

    * Insert LocalRoute.SeqNum into SeqNumList, if known.

    * Insert LocalRoute.MetricType into MetricTypeList.

The AODVv2 message is used to create a corresponding [RFC5444] message (see Section 9).

If the RERR is sent in response to an undeliverable IP packet or RREP message, i.e., if PktSource is included, the RERR SHOULD be sent unicast to the next hop on the route to PktSource. It MUST be sent over the same interface on which the undeliverable IP packet was received. If there is no route to PktSource, the RERR SHOULD be multicast to LL-MANET-Routers. If the RERR is sent in response to a broken link, i.e., PktSource is not included, the RERR is, by default, multicast to LL-MANET-Routers.

8.4.2. RERR Reception

Upon receiving a Route Error, an AODVv2 router performs the following steps:

1. Verify that the message contains the required data: at least one unreachable address

    * If not, ignore this RERR for further processing.

2. For each address in the AddressList, check that:

* The address is valid (routable and unicast)

* The MetricType is supported and configured for use

* There is a LocalRoute with the same MetricType matching the address using longest prefix matching

* Either the LocalRoute’s next hop is the sender of the RERR and the next hop interface is the interface on which the RERR was received, or PktSource is present in the RERR and is a Router Client address

* The unreachable address’ sequence number is either unknown, or is greater than the LocalRoute’s sequence number

If any of the above are false the address does not match a LocalRoute and MUST NOT be processed or regenerated in a RERR.

If all of the above are true, the LocalRoute which matches the address is no longer valid. If the LocalRoute was previously Active, it MUST be reported in a regenerated RERR. If the LocalRoute was previously Idle, it MAY be reported in a regenerated RERR, if ENABLE_IDLE_IN_RERR is configured. The Local Route Set MUST be updated according to these rules:

* If the LocalRoute’s prefix length is the same as the unreachable address’ prefix length, set LocalRoute.State to Invalid.

* If the LocalRoute’s prefix length is longer than the unreachable address’ prefix length, the LocalRoute MUST be expunged from the Local Route Set, since it is a sub-route of the route which is reported to be Invalid.

* If the prefix length is different, create a new LocalRoute with the unreachable address, and its prefix length and sequence number, and set LocalRoute.State to Invalid. These Invalid routes are retained to avoid processing stale messages.

* Update the sequence number on the existing LocalRoute, if the reported sequence number is determined to be newer using the comparison technique described in Section 5.4.

3. If there are previously Active LocalRoutes that MUST be reported, as identified in step 2.:

   * Regenerate the RERR as detailed in Section 8.4.3.

8.4.3. RERR Regeneration

The Route Error message SHOULD NOT be regenerated if CONTROL_TRAFFIC_LIMIT has been reached.

The procedure for RERR regeneration is as follows:

1. If PktSource was included in the original RERR, and PktSource is not a Router Client, copy it into the regenerated RERR

2. For each LocalRoute that needs to be reported as identified in Section 8.4.1:

   * Insert LocalRoute.Address into the AddressList.

   * Insert LocalRoute.PrefixLength into PrefixLengthList, if known and not equal to the address length.

   * Insert LocalRoute.SeqNum into SeqNumList, if known.

   * Insert LocalRoute.MetricType into MetricTypeList.

The AODVv2 message is used to create a corresponding [RFC5444] message (see Section 9). If the RERR contains PktSource, the regenerated RERR SHOULD be sent unicast to the next hop on the LocalRoute to PktSource. It MUST be sent over the same interface on which the undeliverable IP packet was received. If there is no route to PktSource, or PktSource is a Router Client, it SHOULD be multicast to LL-MANET-Routers. If the RERR is sent in response to a broken link, the RERR is, by default, multicast to LL-MANET-Routers.

9. RFC 5444 Representation

AODVv2 specifies that all control messages between routers MUST use the Generalized Mobile Ad Hoc Network Packet/Message Format [RFC5444], and therefore AODVv2’s route messages comprise data which is mapped to message elements in [RFC5444].

[RFC5444] provides a multiplexed transport for multiple protocols. An [RFC5444] implementation MAY choose to optimize the content of certain elements during message creation to reduce control message overhead.

A brief summary of the [RFC5444] format:

1. A packet contains zero or more messages

2. A message contains a Message Header, one Message TLV Block, zero or more Address Blocks, and one Address Block TLV Block per Address Block

3. The Message TLV Block MAY contain zero or more Message TLVs

4. An Address Block TLV Block MAY include zero or more Address Block TLVs

5. Each TLV value in an Address Block TLV Block can be associated with all of the addresses, or with a contiguous set of addresses, or with a single address in the Address Block

AODVv2 does not require access to the [RFC5444] packet header.

In the message header, AODVv2 uses <msg-type>, <msg-hop-limit> and <msg-addr-length>. The <msg-addr-length> field indicates the length of any addresses in the message, using <msg-addr-length> := (address length in octets - 1), i.e. 3 for IPv4 and 15 for IPv6.

The addresses in an Address Block MAY appear in any order, and values in a TLV in the Address Block TLV Block must be associated with the correct address in the Address Block by the [RFC5444] implementation. To indicate which value is associated with each address, the AODVv2 message representation uses lists where the order of the addresses in the AODVv2 AddressList matches the order of values in other data lists, e.g., the order of SeqNums in the SeqNumList in an RERR. [RFC5444] maps this information to Address Block TLVs associated with the relevant addresses in the Address Block.

Each address included in the Address Block is identified as OrigPrefix, TargPrefix, PktSource, or Unreachable Address by including an ADDRESS_TYPE TLV in the Address Block TLV Block.

The following sections show how AODVv2 data is represented in [RFC5444] messages. AODVv2 defines (in Section 11.8) a number of new TLVs.

Where the extension type of a TLV is set to zero, this is the default [RFC5444] value and the extension type will not be included in the message.

9.1. Route Request Message Representation

9.1.1. Message Header

Data	Header Field	Value
None	<msg-type>	RREQ
msg_hop_limit	<msg-hop-limit>	MAX_HOPCOUNT, reduced by number of hops traversed so far by the message.

9.1.2. Message TLV Block

AODVv2 does not define any Message TLVs for an RREQ message.

9.1.3. Address Block

An RREQ contains OrigPrefix and TargPrefix, and each of these addresses has an associated prefix length. If the prefix length has not been included in the AODVv2 message, it is equal to the address length in bits.

Data	Address Block
OrigPrefix/OrigPrefixLen	<address> + <prefix-length>
TargPrefix/TargPrefixLen	<address> + <prefix-length>

9.1.4. Address Block TLV Block

Address Block TLVs are always associated with one or more addresses in the Address Block. The following sections show the TLVs that apply to each address.

9.1.4.1. Address Block TLVs for OrigPrefix

Data	TLV Type	Extension Type	Value
None	ADDRESS_TYPE	0	ORIGPREFIX
OrigSeqNum	SEQ_NUM	0	Sequence number of RREQ_Gen, the router which initiated route discovery.
OrigMetric /MetricType	PATH_METRIC	MetricType	Metric value for the route to OrigPrefix, using MetricType.

9.1.4.2. Address Block TLVs for TargPrefix

Data	TLV Type	Extension Type	Value
None	ADDRESS_TYPE	0	TARGPREFIX
TargSeqNum	SEQ_NUM	0	The last known TargSeqNum for TargPrefix.

9.2. Route Reply Message Representation

9.2.1. Message Header

Data	Header Field	Value
None	<msg-type>	RREP
msg_hop_limit	<msg-hop-limit>	MAX_HOPCOUNT - msg_hop_limit from the corresponding RREQ, reduced by number of hops traversed so far by the message.

9.2.2. Message TLV Block

AODVv2 does not define any Message TLVs for an RREP message.

9.2.3. Address Block

An RREP contains OrigPrefix and TargPrefix, and each of these addresses has an associated prefix length. If the prefix length has not been included in the AODVv2 message, it is equal to the address length in bits.

Data	Address Block
OrigPrefix/OrigPrefixLen	<address> + <prefix-length>
TargPrefix/TargPrefixLen	<address> + <prefix-length>

9.2.4. Address Block TLV Block

Address Block TLVs are always associated with one or more addresses in the Address Block. The following sections show the TLVs that apply to each address.

9.2.4.1. Address Block TLVs for OrigPrefix

Data	TLV Type	Extension Type	Value
None	ADDRESS_TYPE	0	ORIGPREFIX

9.2.4.2. Address Block TLVs for TargPrefix

Data	TLV Type	Extension Type	Value
None	ADDRESS_TYPE	0	TARGPREFIX
TargSeqNum	SEQ_NUM	0	Sequence number of RREP_Gen, the router which created the RREP.
TargMetric /MetricType	PATH_METRIC	MetricType	Metric value for the route to TargPrefix, using MetricType.

9.3. Route Reply Acknowledgement Message Representation

9.3.1. Message Header

Data	Header Field	Value
None	<msg-type>	RREP_Ack

9.3.2. Message TLV Block

AODVv2 defines an AckReq Message TLV, included when an acknowledgement of this message is required, in order to monitor adjacency, as described in Section 7.2.

Data	TLV Type	Extension Type	Value
AckReq	ACK_REQ	0	None

9.3.3. Address Block

AODVv2 does not define an Address Block for an RREP_Ack message.

9.3.4. Address Block TLV Block

AODVv2 does not define any Address Block TLVs for an RREP_Ack message.

9.4. Route Error Message Representation

Route Error Messages MAY be split into multiple [RFC5444] messages when the desired contents would exceed the MTU. However, all of the resulting messages MUST have the same message header as described below. If PktSource is included in the AODVv2 message, it MUST be included in all of the resulting [RFC5444] messages.

9.4.1. Message Header

Data	Header Field	Value
None	<msg-type>	RERR

9.4.2. Message TLV Block

AODVv2 does not define any Message TLVs for an RERR message.

9.4.3. Address Block

The Address Block in an RERR MAY contain PktSource, the source address of the IP packet triggering RERR generation, as detailed in Section 8.4. The prefix length associated with PktSource is equal to the address length in bits.

Address Block always contains one address per route that is no longer valid, and each address has an associated prefix length. If a prefix length has not been included for this address, it is equal to the address length in bits.

Data	Address Block
PktSource	<address> + <prefix-length> for PktSource
AddressList/PrefixLengthList	<address> + <prefix-length> for each unreachable address in AddressList

9.4.4. Address Block TLV Block

Address Block TLVs are always associated with one or more addresses in the Address Block. The following sections show the TLVs that apply to each type of address in the RERR.

9.4.4.1. Address Block TLVs for PktSource

Data	TLV Type	Extension Type	Value
PktSource	ADDRESS_TYPE	0	PKTSOURCE

9.4.4.2. Address Block TLVs for Unreachable Addresses

Data	TLV Type	Extension Type	Value
None	ADDRESS_TYPE	0	UNREACHABLE
SeqNumList	SEQ_NUM	0	Sequence number associated with invalid route to the unreachable address.
MetricTypeList	PATH_METRIC	MetricType	None. Extension Type set to MetricType of the route to the unreachable address.

10. Simple External Network Attachment

Figure 5 shows a stub (i.e., non-transit) network of AODVv2 routers which is attached to an external network via a single External Network Access Router (ENAR). The interface to the external network MUST NOT be configured in the InterfaceSet.

As in any externally-attached network, AODVv2 routers and Router Clients that wish to be reachable from the external network MUST have IP addresses within the ENAR’s routable and topologically correct prefix (e.g., 191.0.2.0/24 in Figure 5). This AODVv2 network and networks attached to routers within it will be advertised to the external network using procedures which are out of scope for this specification.

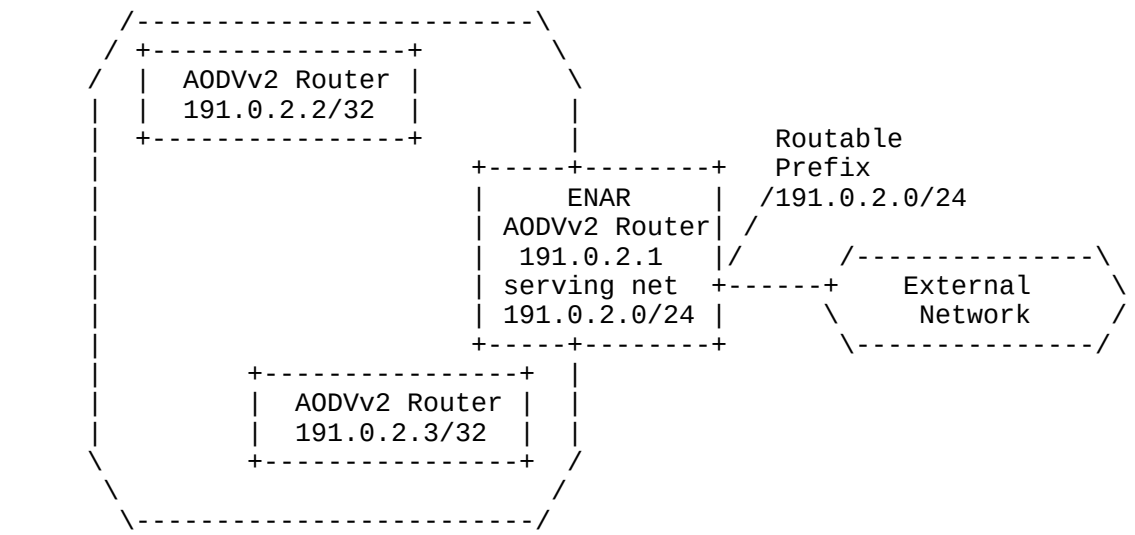


Figure 5: Simple External Network Attachment Example

When an AODVv2 router within the AODVv2 MANET wants to discover a route toward an address on the external network, it uses the normal AODVv2 route discovery for that IP Destination Address. The ENAR MUST respond to RREQ on behalf of all external network destinations, e.g., destinations not on the configured 191.0.2.0/24 network. The ENAR MAY respond with a TargPrefix and TargPrefixLen that represent a prefix including more addresses than just TargAddr, but MUST NOT respond with a TargPrefix and TargPrefixLen which includes any of the networks configured as part of the AODVv2 network. This does result in some inefficiencies in the way external routes are discovered. Sending a Route Request for a gateway is not currently supported.

RREQs for addresses inside the AODVv2 network, e.g. destinations on the configured 191.0.2.0/24 network, are handled using the standard processes described in Section 8. Note that AODvv2 does not support RREQs for prefixes that do not equal address length, but RREPs do advertise the prefix on which TargAddr resides.

When an IP packet from an address on the external network destined for an address in the AODVv2 MANET reaches the ENAR, if the ENAR does not have a route toward that destination in its Routing Information Base, it will perform normal AODVv2 route discovery for that destination.

Configuring the ENAR as a default router is outside the scope of this specification.

11. Configuration

AODVv2 uses various parameters which can be grouped into the following categories:

- Timers
- Protocol constants
- Administrative parameters and controls

This section show the parameters along with their definitions and default values (if any).

Note that several fields have limited size (bits or bytes). These sizes and their encoding may place specific limitations on the values that can be set.

11.1. Timers

AODVv2 requires certain timing information to be associated with Local Route Set entries and message replies. The default values are as follows:

Name	Default Value
ACTIVE_INTERVAL	5 second
MAX_IDLETIME	200 seconds
MAX_BLACKLIST_TIME	200 seconds
MAX_SEQNUM_LIFETIME	300 seconds
RERR_TIMEOUT	3 seconds
RteMsg_ENTRY_TIME	12 seconds
RREQ_WAIT_TIME	2 seconds
RREP_Ack_SENT_TIMEOUT	1 second
RREQ_HOLDDOWN_TIME	10 seconds

Table 2: Timing Parameter Values

The above timing parameter values have worked well for small and medium well-connected networks with moderate topology changes. The timing parameters SHOULD be administratively configurable. Ideally, for networks with frequent topology changes the AODVv2 parameters SHOULD be adjusted using experimentally determined values or dynamic adaptation. For example, in networks with infrequent topology changes MAX_IDLETIME MAY be set to a much larger value. If the

values were configured differently, the following consequences may be observed:

- If MAX_SEQNUM_LIFETIME was configured differently across the network, and any of the routers lost their sequence number or rebooted, this could result in their next route messages being classified as stale at any AODVv2 router using a greater value for MAX_SEQNUM_LIFETIME. This would delay route discovery from and to the re-initializing router.

- Routers with lower values for ACTIVE_INTERVAL + MAX_IDLETIME will invalidate routes more quickly and free resources used to maintain them. This can affect bursty traffic flows which have quiet periods longer than ACTIVE_INTERVAL + MAX_IDLETIME. A route which has timed out due to perceived inactivity is not reported. When the bursty traffic resumes, it would cause a RERR to be generated, and the traffic itself would be dropped. This route would be removed from all upstream routers, even if those upstream routers had larger ACTIVE_INTERVAL or MAX_IDLETIME values. A new route discovery would be required to re-establish the route, causing extra routing protocol traffic and disturbance to the bursty traffic.

- Routers with lower values for MAX_BLACKLIST_TIME would allow neighboring routers to participate in route discovery sooner than routers with higher values. This could result in failed route discoveries if un-blacklisted links are still uni-directional. Since RREQs are retried, this would not affect success of route discovery unless this value was so small as to un-blacklist the router before the RREQ is retried. This value need not be consistent across the network since it is used for maintaining a 1-hop blacklist. However it MUST be greater than RREQ_WAIT_TIME.

- Routers with lower values for RERR_TIMEOUT may create more RERR messages than routers with higher values. This value should be large enough that a RERR will reach all routers using the route reported within it before the timer expires, so that no further data traffic will arrive, and no duplicated RERR messages will be generated.

- Routers with lower values for RteMsg_ENTRY_TIME may not consider received redundant multicast route messages as redundant, and may forward these messages unnecessarily.

- Routers with lower values for RREQ_WAIT_TIME may send more frequent RREQ messages and wrongly determine that a route does not exist, if the delay in receiving an RREP is greater than this interval.

- Routers with lower values for RREP_Ack_SENT_TIMEOUT may wrongly determine links to neighbors to be unidirectional if an RREP_Ack is delayed longer than this timeout.

- Routers with lower values for RREQ_HOLDDOWN_TIME will retry failed route discoveries sooner than routers with higher values. This may be an advantage if the network topology is frequently changing, or may unnecessarily cause more routing protocol traffic.

MAX_SEQNUM_LIFETIME MUST be configured to have the same values for all AODVv2 routers in the network.

11.2. Protocol Constants

AODVv2 protocol constants typically do not require changes. The following table lists these constants, along with their values and a reference to the section describing their use.

Name	Default	Description
DISCOVERY_ATTEMPTS_MAX	3	Section 7.6
RREP_RETRIES	2	Section 8.2.1
MAX_METRIC[MetricType]	[TBD]	Section 6
MAX_METRIC[HopCount]	255	Section 6 and Section 8
MAX_HOPCOUNT	20	Limit to number of hops an RREQ or RREP message can traverse
INFINITY_TIME	[TBD]	Maximum expressible clock time (Section 7.7.2)

Table 3: AODVv2 Constants

MAX_HOPCOUNT cannot be larger than 255.

MAX_METRIC[MetricType] MUST always be the maximum expressible metric value of type MetricType. Field lengths associated with metric values are found in Section 11.5.

These protocol constants MUST have the same values for all AODVv2 routers in the ad hoc network. If the values were configured differently, the following consequences may be observed:

- DISCOVERY_ATTEMPTS_MAX: Routers with higher values are likely to be more successful at finding routes, at the cost of additional control traffic.

- RREP_RETRIES: Routers with lower values are more likely to blacklist neighbors when there is a temporary fluctuation in link quality.

- MAX_METRIC[MetricType]: No interoperability problems due to variations on different routers, but routers with lower values may exhibit overly restrictive behavior during route comparisons.

- MAX_HOPCOUNT: Routers with a value too small would not be able to discover routes to distant addresses.

- INFINITY_TIME: No interoperability problems due to variations on different routers, but if a lower value is used, route state management may exhibit overly restrictive behavior.

11.3. Local Settings

The following table lists AODVv2 parameters which SHOULD be administratively configured for each router:

Name	Default Value	Description
InterfaceSet		Section 5.1
Router Client Set		Section 5.2
BUFFER_SIZE_PACKETS	2	Section 7.6
BUFFER_SIZE_BYTES	MAX_PACKET_SIZE [TBD]	Section 7.6
CONTROL_TRAFFIC_LIMIT	[TBD - 50 pkts/sec?]	Section 8

Table 4: Configuration for Local Settings

11.4. Network-Wide Settings

The following administrative controls MAY be used to change the operation of the network. The same settings SHOULD be used across the network. Inconsistent settings at different routers in the network will not result in protocol errors, but poor performance may result.

Name	Default	Description
ENABLE_IDLE_IN_RERR	Disabled	Section 8.4.1

Table 5: Configuration for Network-Wide Settings

11.5. MetricType Allocation

The metric types used by AODVv2 are identified according to Table 6. All implementations MUST use these values.

Name of MetricType	Type	Metric Value Size
Unassigned	0	Undefined
Hop Count	1	1 octet
Unallocated	2 - 254	TBD
Reserved	255	Undefined

Table 6: AODVv2 Metric Types

11.6. RFC 5444 Message Type Allocation

This specification defines four Message Types, to be allocated from the Experimental range of the "Message Types" namespace defined in [RFC5444], as specified in Table 7.

Name of Message	Type
Route Request (RREQ)	224
Route Reply (RREP)	225
Route Error (RERR)	226
Route Reply Acknowledgement (RREP_Ack)	227

Table 7: AODVv2 Message Types

If the AODVv2 experiment proves to be successful, types from the 0-223 range can be allocated in the future.

11.7. RFC 5444 Message TLV Types

This specification defines one Message TLV Type, to be allocated from the Message-Type-specific "Message TLV Types" namespace defined in [RFC5444], as specified in Table 8.

Name of TLV	Type	Length (octets)	Reference
ACK_REQ	128 (TBD)	0	Section 7.2

Table 8: AODVv2 Message TLV Types

11.8. RFC 5444 Address Block TLV Type Allocation

This specification defines three Address Block TLV Types, to be allocated from the Message-Type-specific "Address Block TLV Types" namespace defined in [RFC5444], as specified in Table 9.

Name of TLV	Type	Length (octets)	Reference
PATH_METRIC	129 (TBD)	depends on MetricType	Section 8
SEQ_NUM	130 (TBD)	2	Section 8
ADDRESS_TYPE	131 (TBD)	1	Section 9

Table 9: AODVv2 Address Block TLV Types

11.9. ADDRESS_TYPE TLV Values

These values are used in the [RFC5444] Address Type TLV discussed in Section 9. All implementations MUST use these values.

Address Type	Value
ORIGPREFIX	0
TARGPREFIX	1
UNREACHABLE	2
PKTSOURCE	3
UNSPECIFIED	255

Table 10: AODVv2 Address Types

12. IANA Considerations

This document has no IANA actions.

13. Security Considerations

This section describes various security considerations and potential avenues to secure AODVv2 routing. The main objective of the AODVv2 protocol is for each router to communicate reachability information about addresses for which it is responsible, and for routes it has learned from other AODVv2 routers.

Networks using AODVv2 to maintain connectivity and establish routes on demand may be vulnerable to certain well-known types of threats, which will be detailed in the following. Some of the threats described can be mitigated or eliminated. Tools to do so will be described also.

With the exception of metric values, AODVv2 assures the integrity of all RteMsg data end-to-end though the use of ICVs (see Section 13.4.2.

The on-demand nature of AODVv2 route discovery automatically reduces the vulnerability to route disruption. Since control traffic for updating route tables is diminished, there is less opportunity for attack and failure.

13.1. Availability

Threats to AODVv2 which reduce availability are considered below.

13.1.1. Denial of Service

Flooding attacks using RREQ amount to a (BLIND) denial of service for route discovery: By issuing RREQ messages for targets that don’t exist, an attacker can flood the network, blocking resources and drowning out legitimate traffic. By triggering the generation of CONTROL_TRAFFIC_LIMIT amount of messages (for example by sending RREQs for many non-existent destinations), an attacker can prevent legitimate messages from being generated. The effect of this attack is dampened by the fact that duplicate RREQ messages are dropped (preventing the network from DDoSing itself). Processing requirements for AODVv2 messages are typically quite small, however AODVv2 routers receiving RREQs do allocate resources in the form of Neighbor Set, Local Route Set and Multicast Route Message Set entries. The attacker can maximize their impact on set growth by changing OrigPrefix or OrigPrefixLen for each RREQ. If a specific node is to be targeted, this attack may be carried out in a

DISTRIBUTED fashion, either by compromising its direct neighbors or by specifying the target's address with TargPrefix and TargPrefixLen. Note that it might be more economical for the attacker to simply jam the medium; an attack which AODVv2 cannot defend itself against.

Mitigation:

- If AODVv2 routers always verify that the sender of the RERR message is trusted, this threat is reduced. Processing requirements would typically be dominated by calculations to verify integrity. This has the effect of reducing (but by no means eliminating) AODVv2's vulnerability to denial of service attacks.

- Authentication of senders can prevent unauthenticated routers from launching a Denial of Service attack on another AODVv2 router. However, this does not protect the network if an attacker has access to an already authenticated router.

13.1.2. Malicious RERR messages

RERR messages are designed to cause removal of installed routes. A malicious node could send an RERR message with false information to attempt to get other routers to remove a route to one or more specific destinations, therefore disrupting traffic to the advertised destinations.

Routes will be deleted if an RERR is received, withdrawing a route for which the sender is the receiver's next hop, and when the RERR includes the MetricType of the installed route, and includes either no sequence number for the route, or includes a greater sequence number than the sequence number stored with that route in the receiver's Local Route Set. Routes will also be deleted if a received RERR contains a PktSource address corresponding to a Router Client.

The information necessary to construct a malicious RERR could be learned by eavesdropping, either by listening to AODVv2 messages or by watching data packet flows.

When the RERR is multicast, it can be received by many routers in the ad hoc network, and will be regenerated when processing results in an active route being removed. This threat could have serious impact on applications communicating by way of the sender of the RERR message.

- The set of routers which use the malicious router as a next hop may be targeted with a malicious RERR with no PktSource address included, if the RERR contains routes for which the malicious router is a next hop from the receiving router. However, since

the sender of the RERR message is either malicious or broken, it is better that it is not used as a next hop for these routes anyway.

- A single router which does not use the malicious router as part of its route may be targeted with a malicious RERR with a PktSource address included.

- Replayed RERR messages could be used to disrupt active routes.

Mitigation:

- Protection against eavesdropping of AODVv2 messages would mitigate this attack to some extent, but eavesdropping of data packets can also be used to deduce the information about which routes could be targeted.

- Protection against a malicious router becoming part of a route will mitigate the attack where a set of routers are targeted. This will not protect against the attack if a PktSource address is included.

- By only regenerating RERR messages where active routes are removed, the spread of the malicious RERR is limited.

- Including sequence numbers in RERR messages offers protection against attacks using replays of these RERR messages.

- If AODVv2 routers always verify that the sender of the RERR message is trusted, this threat is reduced.

13.1.3. False Confirmation of Link Bidirectionality

Links could be erroneously treated as bidirectional if malicious unsolicited or spoofed RREP messages were to be accepted. This would result in a route being installed which could not in fact be used to forward data to the destination, and may divert data packets away from the intended destination.

There is a window of RREQ_WAIT_TIME after an RREQ is sent, in which any malicious router could send an RREP in response, in order for the link to the malicious router to be deemed as bidirectional.

Mitigation:

- Ignoring unsolicited RREP and RREP_Ack messages partially mitigates against this threat.

- o If AODVv2 routers always verify that the sender of the RERR message is trusted, this threat is reduced.

13.1.4. Message Deletion

A malicious router could decide not to forward an RREQ or RREP or RERR message. Not forwarding a RERR or RREP message would disrupt route discovery. Not regenerating a RERR message would result in the source of data packets continuing to maintain and use the route, and further RERR messages being generated by the sender of the non-regenerated RERR. A malicious router could intentionally disrupt traffic flows by not allowing the source of data traffic to re-discover a new route when one breaks.

Failing to send an RREP_Ack would also disrupt route establishment, by not allowing the reverse route to be validated. Return traffic which needs that route will prompt a new route discovery, wasting resources and incurring a slight delay but not disrupting the ability for applications to communicate.

Mitigation:

- o None. also note that malicious router would have to wait for a route to break before it could perform this attack.

13.2. Confidentiality

Passive inspection (eavesdropping) of AODVv2 control messages could enable unauthorized devices to gain information about the network topology, since exchanging such information is the main purpose of AODVv2.

Eavesdropping of data traffic could allow a malicious device to obtain information about how data traffic is being routed. With knowledge of source and destination addresses, malicious messages could be constructed to disrupt normal operation.

13.3. Integrity

Integrity of route information can be compromised in the following types of attack:

13.3.1. Message Insertion

Valid route set entries can be replaced or modified by maliciously constructed AODVv2 messages, destroying existing routes and the network’s integrity. Any router may pose as another router by sending RREQ, RREP, RREP_Ack and RERR messages in its name.

- Sending an RREQ message with false information can disrupt traffic to OrigPrefix, if the sequence number attached is not stale compared to any existing information about OrigPrefix. Since RREQ is multicast and likely to be received by all routers in the ad hoc network, this threat could have serious impact on applications communicating with OrigPrefix. The actual threat to disrupt routes to OrigPrefix is reduced by the AODVv2 mechanism of marking RREQ-derived routes as "Unconfirmed" until the link to the next hop is confirmed.

- Sending an RREP message with false information can disrupt traffic to TargPrefix. Since RREP is unicast, and ignored if a corresponding RREQ was not recently sent, this threat is minimized, and is restricted to receivers along the path from OrigAddr to TargAddr.

- Sending an RREP_Ack response message with false information can cause the route to an originator address to be erroneously accepted even though the route would contain a unidirectional link and thus not be suitable for most traffic. Since the RREP_Ack response is unicast, and ignored if a RREP_Ack was not sent recently to the sender of this RREP_Ack response, this threat is minimized and is strictly local to the RREP transmitter expecting the acknowledgement. Unsolicited RREP_Acks are ignored.

- Sending an RERR message with false information is discussed in Section 13.1.2.

Mitigation:

- If AODVv2 routers always verify that the sender of a message is trusted, this threat is reduced.

13.3.2. Message Modification - Man in the Middle

Any AODVv2 router can forward messages with modified data.

Mitigation:

- If AODVv2 routers verify the integrity of AODVv2 messages, then the threat of disruption is minimized. A man in the middle with no knowledge of the key used to calculate an integrity check value may modify a message but the message will be rejected when it fails an integrity check.

13.3.3. Replay Attacks

Replaying of RREQ or RREP messages would be of less use to an attacker, since they would be dropped immediately due to their stale sequence number. RERR messages may or may not include sequence numbers and are therefore susceptible to replay attacks. RREP_Ack messages do not include sequence numbers and are therefore susceptible to replay attacks.

Mitigation:

- Use of timestamps or sequence numbers prevents replay attacks.

13.4. Protection Mechanisms

13.4.1. Confidentiality and Authentication

Encryption MAY be used for AODVv2 messages. If the routers share a packet-level security association, the message data can be encrypted prior to message transmission. The establishment of such security associations is outside the scope of this specification. Encryption will not only protect against unauthorized devices obtaining information about network topology (eavesdropping) but will ensure that only trusted routers participate in routing operations.

13.4.2. Integrity and Trust using ICVs

Cryptographic Integrity Check Values (ICVs) can be used to ensure integrity of received messages, protecting against man in the middle attacks. Further, by using ICVs, only those routers with knowledge of a shared secret key are allowed to participate in routing information exchanges. [RFC7182] defines ICV TLVs for use with [RFC5444].

The data contained in AODVv2 routing protocol messages MUST be verified using Integrity Check Values, to avoid the use of message data if the message has been tampered with.

13.4.3. Replay Protection using Timestamps

Replay attacks MUST be prevented by using timestamps or sequence numbers in messages. [RFC7182] defines a TIMESTAMP TLV for use with [RFC5444].

The data contained in AODVv2 routing protocol messages MUST be protected with a TIMESTAMP value to ensure the protection against replaying of the message. Sequence numbers can be used as timestamps, since they are known to be strictly increasing.

13.4.4. Application to AODVv2

AODVv2 implementations MUST support ICV and TIMESTAMP TLVs, unless the implementation is intended solely for an environment in which security is unnecessary. AODVv2 deployments SHOULD be configured to use these TLVs to secure messages.

Implementations of AODVv2 MUST support ICV TLVs using type-extensions 1 and 2, hash-function HASH_FUNCTION, and cryptographic function CRYPTOGRAPHIC_FUNCTION. An ICV MUST be included with every message. The ICV value MAY be truncated as specified in [RFC7182].

Since the msg-hop-limit and PATH_METRIC values are mutable when included in AODVv2 messages, these values MUST be set to zero before calculating an ICV. This means that these values are not protected end-to-end and are therefore susceptible to manipulation. This form of attack is described in Section 13.3.2.

Implementations of AODVv2 MUST support a TIMESTAMP TLV using type-extension 0. The timestamp used is a sequence number, and therefore the length of the <TIMESTAMP-value> field matches the AODVv2 sequence number defined in Section 5.4. The TIMESTAMP TLV MUST be included in RREP_Ack and RERR messages.

When more than one message is included in an RFC5444 packet, using a single ICV Packet TLV or single TIMESTAMP Packet TLV is more efficient than including ICV and TIMESTAMP Message TLVs in each message created. If the RFC5444 multiplexer is capable of adding the Packet TLVs, it SHOULD be instructed to include the Packet TLVs in packets containing AODVv2 messages. However, if the multiplexer is not capable of adding the Packet TLVs, the TLVs MUST be included as Message TLVs in each AODVv2 message in the packet.

After message generation but before transmission, the ICV and TIMESTAMP TLVs MUST be added according to each message type as detailed in the following sections. The following steps list the procedure to be performed:

1. If the TIMESTAMP is to be included, depending on AODVv2 message type as specified below, add the TIMESTAMP TLV.

- When a TIMESTAMP Packet TLV is being added, the Packet TLV Block size field MUST be updated.

- When a TIMESTAMP Message TLV is being added, the Message TLV Block size field MUST be updated.

1. The considerations in Section 8 and section 9 of [RFC7182] are followed, removing existing ICV TLVs and adjusting the size and flags fields as appropriate:

- When an ICV Packet TLV is being added, existing ICV Packet TLVs MUST be removed and the Packet TLV Block size MUST be updated. If the Packet TLV Block now contains no TLVs, the phastlv bit in the <pkt-flags> field in the Packet Header MUST be cleared.

- When an ICV Message TLV is being added, existing ICV Message TLVs are removed and the Message TLV Block Size MUST be updated.

1. Mutable fields in the message MUST have their mutable values set to zero before calculating the ICV.

- If the msg-hop-limit field is included in the [RFC5444] message header, msg-hop-limit MUST be set to zero before calculating the ICV.

- If a PATH_METRIC TLV is included, any values present in the TLV MUST be set to zero before calculating the ICV value.

1. Depending on the message type, the ICV is calculated over the appropriate fields (as specified in sections Section 13.4.4.1, Section 13.4.4.2, Section 13.4.4.3 and Section 13.4.4.4) to include the fields <hash-function>, <cryptographic-function>, <key-id-length>, and, if present, <key-id> (in that order), followed by the entire packet or message. This value MAY be truncated (as specified in [RFC7182]).

2. Add the ICV TLV, updating size fields as necessary.

3. The changes made in Step 2 and Step 3 are reversed to re-add any existing ICV TLVs, re-adjust the relevant size and flags fields, and set the msg-hop-limit and PATH_METRIC TLV values.

On message reception, and before message processing, verification of the received message MUST take place:

1. The considerations in Section 8 and Section 9 of [RFC7182] are followed, removing existing ICV TLVs and adjusting the size and flags fields as appropriate.

- When verifying the ICV value in an ICV Packet TLV, all ICV Packet TLVs present in the Packet TLV Block MUST be removed before calculating the ICV, and the Packet TLV Block size MUST be updated. If there are no remaining Packet TLVs, the Packet TLV

Block MUST be removed and the phastlv bit in the <pkt-flags> field MUST be cleared.

- When verifying the ICV value in an ICV Message TLV, all ICV Message TLVs present in the Message TLV Block MUST be removed before calculating the ICV, and the Message TLV Block size MUST be updated.

1. Mutable fields in the message MUST have their mutable values set to zero before calculating the ICV.

- If the msg-hop-limit field is included in the [RFC5444] message header, msg-hop-limit MUST be set to zero before calculating the ICV.

- If a PATH_METRIC TLV is included, any values present in the TLV MUST be set to zero before calculating the ICV value.

1. The ICV is calculated following the considerations in Section 12.2 of [RFC7182], to include the fields <hash-function>, <cryptographic-function>, <key-id-length>, and, if present, <key-id> (in that order), followed by the entire packet or message.

- If the received ICV value is truncated, the calculated ICV value MUST also be truncated (as specified in [RFC7182]), before comparing.

- If the ICV value calculated from the received message or packet does not match the value of <ICV-data> in the received message or packet, the validation fails and the AODVv2 message MUST be discarded and NOT processed or forwarded.

- If the ICV values do match, the values set to zero before calculating the ICV are reset to the received values, and processing continues to Step 4.

1. Verification of a received TIMESTAMP value MUST be performed. The procedure depends on message type as specified in the following sub sections.

- If the TIMESTAMP value in the received message is not valid, the AODVv2 message MUST be discarded and NOT processed or forwarded.

- If the TIMESTAMP value is valid, processing continues as defined in Section 7.

13.4.4.1. RREQ Generation and Reception

Since OrigPrefix is included in the RREQ, the ICV can be calculated and verified using the [RFC5444] contents.The ICV TLV has type extension := 1. Inclusion of an ICV TLV message integrity and endpoint authentication, because trusted routers MUST hold the shared key in order to calculate the ICV value, both to include when creating a message, and to validate the message by checking that the ICV is correct.

Since RREQ_Gen's sequence number is incremented for each new RREQ, replay protection is already afforded and no extra TIMESTAMP TLV is required.

After message generation and before message transmission:

1. Add the ICV TLV as described above.

On message reception and before message processing:

1. Verify the received ICV value as described above.

2. Verification of the sequence number is handled according to Section 7.

13.4.4.2. RREP Generation and Reception

Since TargPrefix is included in the RREP, the ICV can be calculated and verified using the [RFC5444] contents. The ICV TLV has type extension := 1. Inclusion of an ICV provides message integrity and endpoint authentication, because trusted routers MUST hold a valid key in order to calculate the ICV value, both to include when creating a message, and to validate the message by checking that the ICV is correct.

Since RREP_Gen's sequence number is incremented for each new RREP, replay protection is already afforded and no extra TIMESTAMP TLV is required.

After message generation and before message transmission:

1. Add the ICV TLV as described above.

On message reception and before message processing:

1. Verify the received ICV value as described above.

2. Verification of the sequence number is handled according to Section 7.

13.4.4.3. RREP_Ack Generation and Reception

Since no sequence number is included in the RREP_Ack, a TIMESTAMP TLV MUST be included to protect against replay attacks. The value in the TIMESTAMP TLV is set as follows:

- For RREP_Ack request, use Neighbor.AckSeqNum.
- For RREP_Ack response, use the sequence number from the TIMESTAMP TLV in the received RREP_Ack request.

Since no addresses are included in the RREP_Ack, and the receiver of the RREP_Ack uses the source IP address of a received RREP_Ack to identify the sender, the ICV MUST be calculated using the message contents and the IP source address. The ICV TLV has type extension := 2 in order to accomplish this. This provides message integrity and endpoint authentication, because trusted routers MUST hold the correct key in order to calculate the ICV value.

After message generation and before message transmission:

1. Add the TIMESTAMP TLV and ICV TLV as described above.

On message reception and before message processing:

1. Verify the received ICV value as described above.
2. Verify the received TIMESTAMP value by comparing the sequence number in the value field of the TIMESTAMP TLV as follows:

- For a received RREP_Ack request, there is no need to verify the timestamp value. Proceed to message processing as defined in Section 7.
- For a received RREP_Ack response, compare with the Neighbor.AckSeqNum of the Neighbor Set entry for sender of the RREP_Ack request.
- If the sequence number does not match, the AODVv2 message MUST be discarded. Otherwise, Neighbor.AckSeqNum is incremented by 1 and processing continues according to Section 7.

13.4.4.4. RERR Generation and Reception

Since the sender's sequence number is not contained in the RERR, a TIMESTAMP TLV MUST be included to protect against replay attacks. The value in the TIMESTAMP TLV is set by incrementing and using RERR_Gen's sequence number.

Since the receiver of the RERR MUST use the source IP address of the RERR to identify the sender, the ICV MUST be calculated using the message contents and the IP source address. The ICV TLV has type extension := 2 in order to accomplish this. This provides message integrity and endpoint authentication, because trusted routers MUST hold the shared key in order to calculate the ICV value.

After message generation and before message transmission:

1. Add the TIMESTAMP TLV and ICV TLV as described above.

On message reception and before message processing:

1. Verify the received ICV value as described above.

2. Verify the received TIMESTAMP value by comparing the sequence number in the value field of the TIMESTAMP TLV with the Neighbor.HeardRERRSeqNum. If the sequence number in the message is lower than the stored value, the AODVv2 message MUST be discarded. Otherwise, the Neighbor.HeardRERRSeqNum MUST be set to the received value and processing continues according to Section 7.

13.5. Key Management

The method of distribution of shared secret keys is out of the scope of this protocol. Key management is not specified for the following reasons:

Against [RFC4107], an analysis as to whether automated or manual key management should be used shows a compelling case for automated management. In particular:

- a potentially large number of routers may have to be managed, belonging to several organisations, for example in vehicular applications.

- a stream cipher is likely to be used, such as an AES variant.

- o long term session keys might be used by more than two parties, including multicast operations. AODVv2 makes extensive use of multicast.

- o there may be frequent turnover of devices.

On reviewing the case for manual key management against the same document, it can be seen that manual management might be advantageous in environments with limited bandwidth or high round trip times. AODVv2 lends itself to sparse ad hoc networks where transmission conditions may indeed be limited, depending on the bearers selected for use.

However, [RFC4107] assumes that the connectivity between endpoints is already available. In AODVv2, no route is available to a given destination until a router client requests that user traffic be transmitted. It is required to secure the signalling path of the routing protocol that will establish the path across which key exchange functions might subsequently be applied, which is clearly the reverse of the expected functionality. A different strategy is therefore required.

There are two possible solutions. In each case, it is assumed that a defence in depth security posture is being adopted by the system integrator, such that each function in the network as a whole is appropriately secured or defended as necessary, and that there is not complete reliance on security mechanisms built in to AODVv2. Such additional mechanisms could include a suitable wireless device security technology, so that wireless devices are authenticated and secured by their peers prior to exchanging user data, which in this case would include AODVV2 signalling traffic as a payload, and mechanisms which verify the authenticity and/or integrity of application-layer user data transported once a route has been established.

1. In the case that no AODVv2 routers have any detailed prior knowledge of any other AODVv2 router, but does have knowledge of the credentials of other organisations in which the router has been previously configured to trust, it is possible for an AODVv2 router to send an initialisation vector as part of an exchange, which could be verified against such credentials. Such an exchange could make use of Identity-Based Signatures ([I-D.ietf-manet-ibs]), based on Elliptic Curve-Based Certificateless Signatures for Identity-Based Encryption [RFC6507], which eliminate the need for a handshake process to establish trust.

2. If it is impossible to use Identity-Based Signatures, and the risk to the AODVv2 signalling traffic is considered to be low due to the use of security countermeasures elsewhere in the system, a simple pre-placed shared secret could be used between routers, which is used as-is or is used to generate some ephemeral secret based on another known variable, such as time of day if that is universally available at a level of accuracy sufficient to make such a system viable.

14. Acknowledgments

AODVv2 is a descendant of the design of previous MANET on-demand protocols, especially AODV [RFC3561] and DSR [RFC4728]. Changes to previous MANET on-demand protocols stem from research and implementation experiences. Thanks to Elizabeth Belding and Ian Chakeres for their long time authorship of AODV. Additional thanks to Derek Atkins, Emmanuel Baccelli, Abdussalam Baryun, Ramon Caceres, Justin Dean, Christopher Dearlove, Fatemeh Ghassemi, Ulrich Herberg, Henner Jakob, Ramtin Khosravi, Luke Klein-Berndt, Lars Kristensen, Tronje Krop, Koojana Kuladinithi, Kedar Namjoshi, Keyur Patel, Alexandru Petrescu, Henning Rogge, Fransisco Ros, Pedro Ruiz, Christoph Sommer, Romain Thouvenin, Richard Trefler, Jiazi Yi, Seung Yi, Behnaz Yousefi, and Cong Yuan, for their reviews of AODVv2 and DYMO, as well as numerous specification suggestions.

15. References

15.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3561] Perkins, C., Belding-Royer, E., and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", RFC 3561, DOI 10.17487/RFC3561, July 2003, <http://www.rfc-editor.org/info/rfc3561>.

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized Mobile Ad Hoc Network (MANET) Packet/Message Format", RFC 5444, DOI 10.17487/RFC5444, February 2009, <http://www.rfc-editor.org/info/rfc5444>.

[RFC5498] Chakeres, I., "IANA Allocations for Mobile Ad Hoc Network (MANET) Protocols", RFC 5498, DOI 10.17487/RFC5498, March 2009, <http://www.rfc-editor.org/info/rfc5498>.

[RFC7182] Herberg, U., Clausen, T., and C. Dearlove, "Integrity Check Value and Timestamp TLV Definitions for Mobile Ad Hoc Networks (MANETs)", RFC 7182, DOI 10.17487/RFC7182, April 2014, <http://www.rfc-editor.org/info/rfc7182>.

15.2. Informative References

[I-D.ietf-manet-ibs]
Dearlove, C., "Identity-Based Signatures for MANET Routing Protocols", draft-ietf-manet-ibs-05 (work in progress), March 2016.

[Koodli01]
Koodli, R. and C. Perkins, "Fast handovers and context transfers in mobile networks", Proceedings of the ACM SIGCOMM Computer Communication Review 2001, Volume 31 Issue 5, 37-47, October 2001.

[Perkins94]
Perkins, C. and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers", Proceedings of the ACM SIGCOMM ’94 Conference on Communications Architectures, Protocols and Applications, London, UK, pp. 234-244, August 1994.

[RFC2501] Corson, S. and J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, DOI 10.17487/RFC2501, January 1999, <http://www.rfc-editor.org/info/rfc2501>.

[RFC4107] Bellovin, S. and R. Housley, "Guidelines for Cryptographic Key Management", BCP 107, RFC 4107, DOI 10.17487/RFC4107, June 2005, <http://www.rfc-editor.org/info/rfc4107>.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, DOI 10.17487/RFC4193, October 2005, <http://www.rfc-editor.org/info/rfc4193>.

[RFC4728] Johnson, D., Hu, Y., and D. Maltz, "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4", RFC 4728, DOI 10.17487/RFC4728, February 2007, <http://www.rfc-editor.org/info/rfc4728>.

[RFC6130] Clausen, T., Dearlove, C., and J. Dean, "Mobile Ad Hoc Network (MANET) Neighborhood Discovery Protocol (NHDP)", RFC 6130, DOI 10.17487/RFC6130, April 2011, <http://www.rfc-editor.org/info/rfc6130>.

[RFC6507] Groves, M., "Elliptic Curve-Based Certificateless Signatures for Identity-Based Encryption (ECCSI)", RFC 6507, DOI 10.17487/RFC6507, February 2012, <http://www.rfc-editor.org/info/rfc6507>.

Appendix A. AODVv2 Draft Updates

This section lists the changes between AODVv2 revisions ...-15.txt and ...-16.txt.

- Changed 'regeneration' language in favor of 'forwarding'.
- Reintroduced use of msg-hop-limit in 5444 message header.
- Use OrigPrefix rather than OrigAddr and TargPrefix rather than TargAddr where appropriate
- Removed validity time
- Removed AckReq from RREP messages, use two-way RREP_ack to check for bidirectionality
- Unicast RREP messages
- Removed orphaned references
- Clarified language
- Improved Sequence Number instructions
- Changed 'Unknown' terminology to 'Heard'
- Extended experiment description
- Added detailed description of which steps to take when calculating and evaluating ICVs, particularly how to zero out the metric value

Authors' Addresses

Charles E. Perkins
Futurewei Inc.
2330 Central Expressway
Santa Clara, CA 95050
USA

Phone: +1-408-330-4586
Email: charliep@computer.org

Stan Ratliff
Idirect
13861 Sunrise Valley Drive, Suite 300
Herndon, VA 20171
USA

Email: ratliffstan@gmail.com

John Dowdell
Airbus Defence and Space
Celtic Springs
Newport, Wales NP10 8FZ
United Kingdom

Email: john.dowdell@airbus.com

Lotte Steenbrink
HAW Hamburg, Dept. Informatik
Berliner Tor 7
D-20099 Hamburg
Germany

Email: lotte.steenbrink@haw-hamburg.de

Victoria Mercieca
Airbus Defence and Space
Celtic Springs
Newport, Wales NP10 8FZ
United Kingdom

Email: victoria.mercieca@airbus.com

Mobile Ad hoc Networks Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 29, 2017

S. Ratliff
VT iDirect
S. Jury
Cisco Systems
D. Satterwhite
Broadcom
R. Taylor
Airbus Defence & Space
B. Berry
March 28, 2017

Dynamic Link Exchange Protocol (DLEP)
draft-ietf-manet-dlep-29

Abstract

When routing devices rely on modems to effect communications over wireless links, they need timely and accurate knowledge of the characteristics of the link (speed, state, etc.) in order to make routing decisions. In mobile or other environments where these characteristics change frequently, manual configurations or the inference of state through routing or transport protocols does not allow the router to make the best decisions. DLEP describes a new protocol for a bidirectional, event-driven communication channel between the router and the modem to facilitate communication of changing link characteristics.

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 http://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 29, 2017.

Table of Contents

1. Introduction

There exist today a collection of modem devices that control links of variable datarate and quality. Examples of these types of links include line-of-sight (LOS) terrestrial radios, satellite terminals, and broadband modems. Fluctuations in speed and quality of these links can occur due to configuration, or on a moment-to-moment basis, due to physical phenomena like multipath interference, obstructions, rain fade, etc. It is also quite possible that link quality and datarate vary with respect to individual destinations on a link, and with the type of traffic being sent. As an example, consider the case of an IEEE 802.11 access point, serving two associated laptop computers. In this environment, the answer to the question "What is the datarate on the 802.11 link?" is "It depends on which associated laptop we're talking about, and on what kind of traffic is being sent." While the first laptop, being physically close to the access point, may have a datarate of 54Mbps for unicast traffic, the other laptop, being relatively far away, or obstructed by some object, can

simultaneously have a datarate of only 32Mbps for unicast. However, for multicast traffic sent from the access point, all traffic is sent at the base transmission rate (which is configurable, but depending on the model of the access point, is usually 24Mbps or less).

In addition to utilizing variable datarate links, mobile networks are challenged by the notion that link connectivity will come and go over time, without an effect on a router's interface state (Up or Down). Effectively utilizing a relatively short-lived connection is problematic in IP routed networks, as IP routing protocols tend to rely on interface state and independent timers to maintain network convergence (e.g., HELLO messages and/or recognition of DEAD routing adjacencies). These dynamic connections can be better utilized with an event-driven paradigm, where acquisition of a new neighbor (or loss of an existing one) is signaled, as opposed to a paradigm driven by timers and/or interface state. DLEP not only implements such an event-driven paradigm, but does so over a local (1 hop) TCP session, which guarantees delivery of the event messages.

Another complicating factor for mobile networks are the different methods of physically connecting the modem devices to the router. Modems can be deployed as an interface card in a router's chassis, or as a standalone device connected to the router via Ethernet or serial link. In the case of Ethernet attachment, with existing protocols and techniques, routing software cannot be aware of convergence events occurring on the radio link (e.g., acquisition or loss of a potential routing neighbor), nor can the router be aware of the actual capacity of the link. This lack of awareness, along with the variability in datarate, leads to a situation where finding the (current) best route through the network to a given node is difficult to establish and properly maintain. This is especially true of demand-based access schemes such as Demand Assigned Multiple Access (DAMA) implementations used on some satellite systems. With a DAMA-based system, additional datarate may be available, but will not be used unless the network devices emit traffic at a rate higher than the currently established rate. Increasing the traffic rate does not guarantee additional datarate will be allocated; rather, it may result in data loss and additional retransmissions on the link.

Addressing the challenges listed above, the Dynamic Link Exchange Protocol, or DLEP, has been developed. The DLEP protocol runs between a router and its attached modem devices, allowing the modem to communicate link characteristics as they change, and convergence events (acquisition and loss of potential routing next-hops). The following diagrams are used to illustrate the scope of DLEP packets.

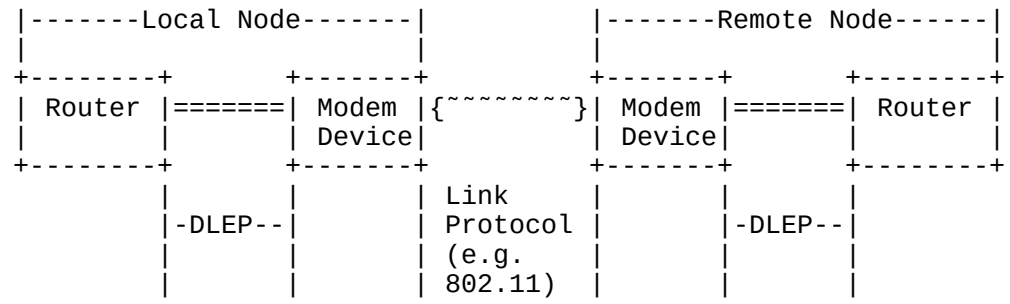


Figure 1: DLEP Network

In Figure 1, when the local modem detects the presence of a remote node, it (the local modem) sends a message to its router via the DLEP protocol. The message consists of an indication of what change has occurred on the link (e.g., presence of a remote node detected), along with a collection of DLEP-defined data items that further describe the change. Upon receipt of the message, the local router may take whatever action it deems appropriate, such as initiating discovery protocols, and/or issuing HELLO messages to converge the network. On a continuing, as-needed basis, the modem devices use DLEP to report any characteristics of the link (datarate, latency, etc.) that have changed. DLEP is independent of the link type and topology supported by the modem. Note that the DLEP protocol is specified to run only on the local link between router and modem. Some over the air signaling may be necessary between the local and remote modem in order to provide some parameters in DLEP messages between the local modem and local router, but DLEP does not specify how such over the air signaling is carried out. Over the air signaling is purely a matter for the modem implementer.

Figure 2 shows how DLEP can support a configuration where routers are connected with different link types. In this example, Modem A implements a point-to-point link, and Modem B is connected via a shared medium. In both cases, the DLEP protocol is used to report the characteristics of the link (datarate, latency, etc.) to routers. The modem is also able to use the DLEP session to notify the router when the remote node is lost, shortening the time required to re-converge the network.

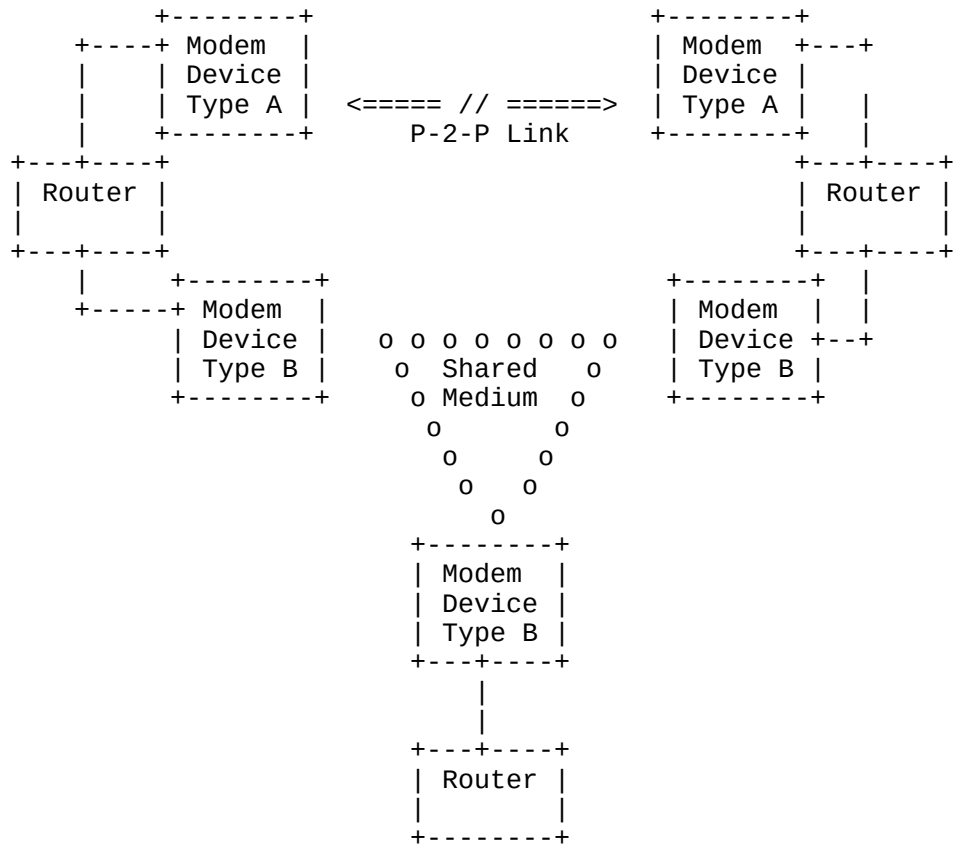


Figure 2: DLEP Network with Multiple Modem Devices

2. Protocol Overview

DLEP defines a set of Messages used by modems and their attached routers to communicate events that occur on the physical link(s) managed by the modem: for example, a remote node entering or leaving the network, or that the link has changed. Associated with these Messages are a set of Data Items - information that describes the remote node (e.g., address information), and/or the characteristics of the link to the remote node. Throughout this document, we refer to a modems/routers participating in a DLEP session as 'DLEP Participants', unless a specific distinction (e.g. modem or router) is required.

DLEP uses a session-oriented paradigm between the modem device and its associated router. If multiple modem devices are attached to a router (as in Figure 2), or the modem supports multiple connections

(via multiple logical or physical interfaces), then separate DLEP sessions exist for each modem or connection. A router and modem form a session by completing the discovery and initialization process. This router-modem session persists unless or until it either (1) times out, based on the absence of DLEP traffic (including heartbeats), or (2) is explicitly torn down by one of the DLEP participants.

While this document represents the best efforts of the working group to be functionally complete, it is recognized that extensions to DLEP will in all likelihood be necessary as more link types are used. Such extensions are defined as additional Messages, Data Items and/or status codes, and associated rules of behavior, that are not defined in this document. DLEP contains a standard mechanism for router and modem implementations to negotiate the available extensions to use on a per-session basis.

2.1. Destinations

The router/modem session provides a carrier for information exchange concerning 'destinations' that are available via the modem device. Destinations can be identified by either the router or the modem, and represent a specific, addressable location that can be reached via the link(s) managed by the modem.

The DLEP Messages concerning destinations thus become the way for routers and modems to maintain, and notify each other about, an information base representing the physical and logical destinations accessible via the modem device, as well as the link characteristics to those destinations.

A destination can be either physical or logical. The example of a physical destination would be that of a remote, far-end router attached via the variable-quality network. It should be noted that for physical destinations the MAC address is the address of the far-end router, not the modem.

The example of a logical destination is Multicast. Multicast traffic destined for the variable-quality network (the network accessed via the modem) is handled in IP networks by deriving a Layer 2 MAC address based on the Layer 3 address. Leveraging on this scheme, multicast traffic is supported in DLEP simply by treating the derived MAC address as any other destination in the network.

To support these logical destinations, one of the DLEP participants (typically, the router) informs the other as to the existence of the logical destination. The modem, once it is aware of the existence of this logical destination, reports link characteristics just as it

would for any other destination in the network. The specific algorithms a modem would use to derive metrics on logical destinations are outside the scope of this specification, and is left to specific implementations to decide.

In all cases, when this specification uses the term destination, it refers to the addressable locations, either logical or physical, that are accessible by the radio link(s).

2.2. Conventions and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [RFC2119].

3. Requirements

DLEP MUST be implemented on a single Layer 2 domain. The protocol identifies next-hop destinations by using the MAC address for delivering data traffic. No manipulation or substitution is performed; the MAC address supplied in all DLEP Messages is used as the Destination MAC address for frames emitted by the participating router. MAC addresses MUST be unique within the context of router-modem session.

To enforce the single Layer 2 domain, implementations MUST support The Generalized TTL Security Mechanism [RFC5082], and implementations MUST adhere to this specification for all DLEP Messages.

DLEP specifies UDP multicast for single-hop discovery signaling, and TCP for transport of the Messages. Modems and routers participating in DLEP sessions MUST have topologically consistent IP addresses assigned. It is RECOMMENDED that DLEP implementations utilize IPv6 link-local addresses to reduce the administrative burden of address assignment.

DLEP relies on the guaranteed delivery of its Messages between router and modem, once the 1 hop discovery process is complete, hence, the specification of TCP to carry the Messages. Other reliable transports for the protocol are possible, but are outside the scope of this document.

4. Implementation Scenarios

During development of this specification, two types of deployments were discussed.

The first can be viewed as a "dedicated deployment". In this mode, DLEP routers and modems are either directly connected (e.g., using cross-over cables to connect interfaces), or are connected to a dedicated switch. An example of this type of deployment would be a router with a line-of- sight radio connected into one interface, with a satellite modem connected into another interface. In mobile environments, the router and the connected modem(s) are placed into a mobile platform (e.g., a vehicle, boat, or airplane). In this mode, when a switch is used, it is possible that a small number of ancillary devices (e.g., a laptop) are also plugged into the switch. But in either event, the resulting network segment is constrained to a small number of devices, and is not generally accessible from anywhere else in the network.

The other type of deployment envisioned can be viewed as a "networked deployment". In this type of scenario, the DLEP router and modem(s) are placed on a segment that is accessible from other points in the network. In this scenario, not only are the DLEP router and modem(s) accessible from other points in the network; the router and a given modem could be multiple physical hops away from each other. This scenario necessitates the use of Layer 2 tunneling technology to enforce the single-hop requirement of DLEP.

5. Assumptions

DLEP assumes that a signaling protocol exists between modems participating in a network. The specification does not define the character or behavior of this over-the-air signaling, but does expect some information to be carried (or derived) by the signaling, such as the arrival and departure of modems from this network, and the variation of the link characteristics between modems. This information is then assumed to be used by the modem to implement the DLEP protocol.

The specification assumes that the link between router and modem is static with respect to datarate and latency, and that this link is not likely to be the cause of a performance bottleneck. In deployments where the router and modem are physically separated by multiple network hops, served by Layer 2 tunneling technology, DLEP statistics on the RF links could be insufficient for routing protocols to make appropriate routing decisions. This would especially become an issue in cases where the Layer 2 tunnel between router and modem is itself served in part (or in total) with a wireless back-haul link.

6. Metrics

DLEP includes the ability for the router and modem to communicate metrics that reflect the characteristics (e.g., datarate, latency) of the variable-quality link in use. DLEP does not specify how a given metric value is to be calculated, rather, the protocol assumes that metrics have been calculated by a 'best effort', incorporating all pertinent data that is available to the modem device. Metrics based on large enough sample sizes will preclude short traffic bursts from adversely skewing reported values.

DLEP allows for metrics to be sent within two contexts - metrics for a specific destination within the network (e.g., a specific router), and per-session (those that apply to all destinations accessed via the modem). Most metrics can be further subdivided into transmit and receive metrics. In cases where metrics are provided at session level, the router propagates the metrics to all entries in its information base for destinations that are accessed via the modem.

DLEP modems announce all metric Data Items that will be reported during the session, and provide default values for those metrics, in the Session Initialization Response Message (Section 12.6). In order to use a metric type that was not included in the Session Initialization Response Message, modem implementations terminate the session with the router (via the Session Terminate Message (Section 12.9)), and establish a new session.

A DLEP modem can send metrics both in a session context, via the Session Update Message (Section 12.7), and a specific destination context, via the Destination Update Message (Section 12.17), at any time. The most recently received metric value takes precedence over any earlier value, regardless of context - that is:

1. If the router receives metrics in a specific destination context (via the Destination Update Message), then the specific destination is updated with the new metric.

2. If the router receives metrics in a session-wide context (via the Session Update Message), then the metrics for all destinations accessed via the modem are updated with the new metric.

It is left to implementations to choose sensible default values based on their specific characteristics. Modems having static (non-changing) link metric characteristics can report metrics only once for a given destination (or once on a session-wide basis, if all connections via the modem are of this static nature).

In addition to communicating existing metrics about the link, DLEP provides a Message allowing a router to request a different datarate or latency from the modem. This Message is the Link Characteristics Request Message (Section 12.18), and gives the router the ability to deal with requisite increases (or decreases) of allocated datarate/ latency in demand-based schemes in a more deterministic manner.

7. DLEP Session Flow

All DLEP participants of a session transition through a number of distinct states during the lifetime of a DLEP session:

- Peer Discovery
- Session Initialization
- In-Session
- Session Termination
- Session Reset

Modems, and routers supporting DLEP discovery, transition through all five (5) of the above states. Routers that rely on preconfigured TCP address/port information start in the Session Initialization state.

Modems MUST support the Peer Discovery state.

7.1. Peer Discovery State

Modems MUST support DLEP Peer Discovery; routers MAY support the discovery signals, or rely on a priori configuration to locate modems. If a router chooses to support DLEP discovery, all signals MUST be supported.

In the Peer Discovery state, routers that support DLEP discovery MUST send Peer Discovery Signals (Section 12.3) to initiate modem discovery.

The router implementation then waits for a Peer Offer Signal (Section 12.4) response from a potential DLEP modem. While in the Peer Discovery state, Peer Discovery Signals MUST be sent repeatedly by a DLEP router, at regular intervals. It is RECOMMENDED that this interval be set to 60 seconds. The interval MUST be a minimum of one second; it SHOULD be a configurable parameter. Note that this operation (sending Peer Discovery and waiting for Peer Offer) is outside the DLEP Transaction Model (Section 8), as the Transaction Model only describes Messages on a TCP session.

Routers receiving a Peer Offer Signal MUST use one of the modem address/port combinations from the Peer Offer Signal to establish a TCP connection to the modem, even if a priori configuration exists. If multiple connection point Data Items exist in the received Peer Offer Signal, routers SHOULD prioritize IPv6 connection points over IPv4 connection points. If multiple connection points exist with the same transport (e.g. IPv6 or IPv4), implementations MAY use their own heuristics to determine the order in which they are tried. If a TCP connection cannot be achieved using any of the address/port combinations and the Discovery mechanism is in use, then the router SHOULD resume issuing Peer Discovery Signals. If no Connection Point Data Items are included in the Peer Offer Signal, the router MUST use the source address of the UDP packet containing the Peer Offer Signal as the IP address, and the DLEP well-known port number.

In the Peer Discovery state, the modem implementation MUST listen for incoming Peer Discovery Signals on the DLEP well-known IPv6 and/or IPv4 link-local multicast address and port. On receipt of a valid Peer Discovery Signal, it MUST reply with a Peer Offer Signal.

Modems MUST be prepared to accept a TCP connection from a router that is not using the Discovery mechanism, i.e. a connection attempt that occurs without a preceding Peer Discovery Signal.

Implementations of DLEP SHOULD implement, and use, TLS [RFC5246] to protect the TCP session. The "dedicated deployments" discussed in Implementation Scenarios (Section 4) MAY consider use of DLEP without TLS. For all "networked deployments" (again, discussed in Implementation Scenarios), implementation and use of TLS is STRONGLY RECOMMENDED. If TLS is to be used then the TLS session MUST be established before any Messages are passed between peers. Routers supporting TLS MUST prioritize connection points using TLS over those that do not.

Upon establishment of a TCP connection, and TLS session if TLS is in use, both modem and router enter the Session Initialization state. It is up to the router implementation if Peer Discovery Signals continue to be sent after the device has transitioned to the Session Initialization state. Modem implementations MUST silently ignore Peer Discovery Signals from a router with which it already has a TCP connection.

7.2. Session Initialization State

On entering the Session Initialization state, the router MUST send a Session Initialization Message (Section 12.5) to the modem. The router MUST then wait for receipt of a Session Initialization Response Message (Section 12.6) from the modem. Receipt of the

Session Initialization Response Message containing a Status Data Item (Section 13.1) with status code set to 0 'Success', see Table 2, indicates that the modem has received and processed the Session Initialization Message, and the router MUST transition to the In-Session state.

On entering the Session Initialization state, the modem MUST wait for receipt of a Session Initialization Message from the router. Upon receipt of a Session Initialization Message, the modem MUST send a Session Initialization Response Message, and the session MUST transition to the In-Session state. If the modem receives any Message other than Session Initialization, or it fails to parse the received Message, it MUST NOT send any Message, and MUST terminate the TCP connection and transition to the Session Reset state.

DLEP provides an extension negotiation capability to be used in the Session Initialization state, see Section 9. Extensions supported by an implementation MUST be declared to potential DLEP participants using the Extensions Supported Data Item (Section 13.6). Once both DLEP participants have exchanged initialization Messages, an implementation MUST NOT emit any Message, Signal, Data Item or status code associated with an extension that was not specified in the received initialization Message from its peer.

7.3. In-Session State

In the In-Session state, Messages can flow in both directions between DLEP participants, indicating changes to the session state, the arrival or departure of reachable destinations, or changes of the state of the links to the destinations.

The In-Session state is maintained until one of the following conditions occur:

- The implementation terminates the session by sending a Session Termination Message (Section 12.9), or,
- Its peer terminates the session, indicated by receiving a Session Termination Message.

The implementation MUST then transition to the Session Termination state.

7.3.1. Heartbeats

In order to maintain the In-Session state, periodic Heartbeat Messages (Section 12.20) MUST be exchanged between router and modem. These Messages are intended to keep the session alive, and to verify

bidirectional connectivity between the two DLEP participants. It is RECOMMENDED that the interval timer between heartbeat messages be set to 60 seconds. The interval MUST be a minimum of one second; it SHOULD be a configurable parameter.

Each DLEP participant is responsible for the creation of Heartbeat Messages.

Receipt of any valid DLEP Message MUST reset the heartbeat interval timer (i.e., valid DLEP Messages take the place of, and obviate the need for, additional Heartbeat Messages).

Implementations MUST allow a minimum of two (2) heartbeat intervals to expire with no Messages from its peer before terminating the session. When terminating the session, a Session Termination Message containing a Status Data Item (Section 13.1) with status code set to 132 'Timed Out', see Table 2, MUST be sent, and then the implementation MUST transition to the Session Termination state.

7.4. Session Termination State

When an implementation enters the Session Termination state after sending a Session Termination Message (Section 12.9) as the result of an invalid Message or error, it MUST wait for a Session Termination Response Message (Section 12.10) from its peer. Senders SHOULD allow four (4) heartbeat intervals to expire before assuming that its peer is unresponsive, and continuing with session termination. Any other Message received while waiting MUST be silently ignored.

When the sender of the Session Termination Message receives a Session Termination Response Message from its peer, or times out, it MUST transition to the Session Reset state.

When an implementation receives a Session Termination Message from its peer, it enters the Session Termination state and then it MUST immediately send a Session Termination Response and transition to the Session Reset state.

7.5. Session Reset state

In the Session Reset state the implementation MUST perform the following actions:

- Release all resources allocated for the session.
- Eliminate all destinations in the information base represented by the session. Destination Down Messages (Section 12.15) MUST NOT be sent.

- Terminate the TCP connection.

Having completed these actions the implementation SHOULD return to the relevant initial state: Peer Discovery for modems; either Peer Discovery or Session Initialization for routers, depending on configuration.

7.5.1. Unexpected TCP connection termination

If the TCP connection between DLEP participants is terminated when an implementation is not in the Session Reset state, the implementation MUST immediately transition to the Session Reset state.

8. Transaction Model

DLEP defines a simple Message transaction model: Only one request per destination may be in progress at a time per session. A Message transaction is considered complete when a response matching a previously issued request is received. If a DLEP participant receives a request for a destination for which there is already an outstanding request, the implementation MUST terminate the session by issuing a Session Termination Message (Section 12.9) containing a Status Data Item (Section 13.1) with status code set to 129 'Unexpected Message', see Table 2, and transition to the Session Termination state. There is no restriction to the total number of Message transactions in progress at a time, as long as each transaction refers to a different destination.

It should be noted that some requests may take a considerable amount of time for some DLEP participants to complete, for example, a modem handling a multicast destination up request may have to perform a complex network reconfiguration. A sending implementation MUST be able to handle such long running transactions gracefully.

Additionally, only one session request, e.g. a Session Initialization Message (Section 12.5), may be in progress at a time per session. As above, a session transaction is considered complete when a response matching a previously issued request is received. If a DLEP participant receives a session request while there is already a session request in progress, it MUST terminate the session by issuing a Session Termination Message containing a Status Data Item with status code set to 129 'Unexpected Message', and transition to the Session Termination state. Only the Session Termination Message may be issued when a session transaction is in progress. Heartbeat Messages (Section 12.20) MUST NOT be considered part of a session transaction.

DLEP transactions do not time out and are not cancellable, except for transactions in-flight when the DLEP session is reset. If the session is terminated, canceling transactions in progress MUST be performed as part of resetting the state machine. An implementation can detect if its peer has failed in some way by use of the session heartbeat mechanism during the In-Session state, see Section 7.3.

9. Extensions

Extensions MUST be negotiated on a per-session basis during session initialization via the Extensions Supported mechanism. Implementations are not required to support any extension in order to be considered DLEP compliant.

If interoperable protocol extensions are required, they will need to be standardized either as an update to this document, or as an additional stand-alone specification. The requests for IANA-controlled registries in this document contain sufficient Reserved space for DLEP Signals, Messages, Data Items and status codes to accommodate future extensions to the protocol.

As multiple protocol extensions MAY be announced during session initialization, authors of protocol extensions need to consider the interaction of their extension with other published extensions, and specify any incompatibilities.

9.1. Experiments

This document requests Private Use numbering space in the DLEP Signal, Message, Data Item and status code registries for experimental extensions. The intent is to allow for experimentation with new Signals, Messages, Data Items, and/or status codes, while still retaining the documented DLEP behavior.

Use of the Private Use Signals, Messages, Data Items, status codes, or behaviors MUST be announced as DLEP Extensions, during session initialization, using extension identifiers from the Private Use space in the Extensions Supported registry (Table 3), with a value agreed upon (a priori) between the participants. DLEP extensions using the Private Use numbering space are commonly referred to as Experiments.

Multiple experiments MAY be announced in the Session Initialization Messages. However, use of multiple experiments in a single session could lead to interoperability issues or unexpected results (e.g., clashes of experimental Signals, Messages, Data Items and/or status code types), and is therefore discouraged. It is left to implementations to determine the correct processing path (e.g., a

decision on whether to terminate the session, or to establish a precedence of the conflicting definitions) if such conflicts arise.

10. Scalability

The protocol is intended to support thousands of destinations on a given modem/router pair. At large scale, implementations should consider employing techniques to prevent flooding its peer with a large number of Messages in a short time. For example, a dampening algorithm could be employed to prevent a flapping device from generating a large number of Destination Up/Destination Down Messages.

Also, use of techniques such as a hysteresis can lessen the impact of rapid, minor fluctuations in link quality. The specific algorithms for handling flapping destinations and minor changes in link quality are outside the scope of this specification.

11. DLEP Signal and Message Structure

DLEP defines two protocol units used in two different ways: Signals and Messages. Signals are only used in the Discovery mechanism and are carried in UDP datagrams. Messages are used bidirectionally over a TCP connection between the participants, in the Session Initialization, In-Session and Session Termination states.

Both Signals and Messages consist of a Header followed by an unordered list of Data Items. Headers consist of Type and Length information, while Data Items are encoded as TLV (Type-Length-Value) structures. In this document, the Data Items following a Signal or Message Header are described as being 'contained in' the Signal or Message.

There is no restriction on the order of Data Items following a Header, and the acceptability of duplicate Data Items is defined by the definition of the Signal or Message declared by the type in the Header.

All integers in Header fields and values MUST be in network byte-order.

11.1. DLEP Signal Header

The DLEP Signal Header contains the following fields:

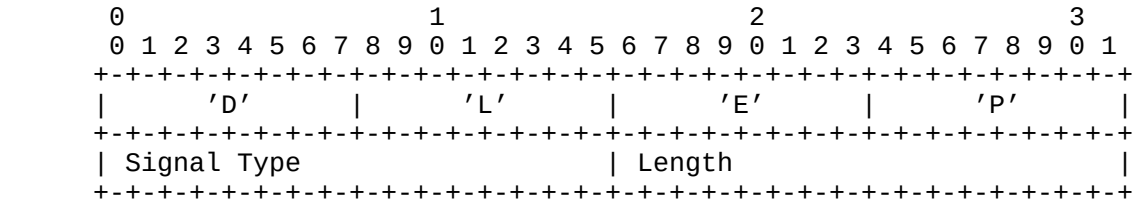


Figure 3: DLEP Signal Header

"DLEP": Every Signal MUST start with the characters: U+0044, U+004C, U+0045, U+0050.

Signal Type: A 16-bit unsigned integer containing one of the DLEP Signal Type values defined in this document.

Length: The length in octets, expressed as a 16-bit unsigned integer, of all of the DLEP Data Items contained in this Signal. This length MUST NOT include the length of the Signal Header itself.

The DLEP Signal Header is immediately followed by zero or more DLEP Data Items, encoded in TLVs, as defined in this document.

11.2. DLEP Message Header

The DLEP Message Header contains the following fields:

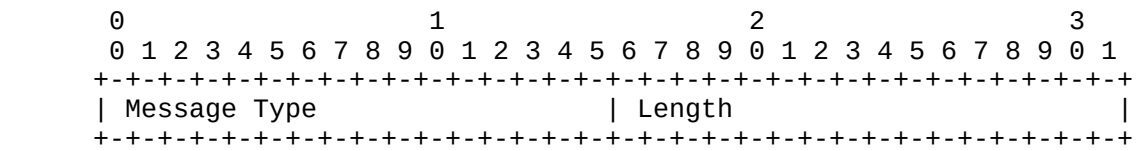


Figure 4: DLEP Message Header

Message Type: A 16-bit unsigned integer containing one of the DLEP Message Type values defined in this document.

Length: The length in octets, expressed as a 16-bit unsigned integer, of all of the DLEP Data Items contained in this Message. This length MUST NOT include the length of the Message Header itself.

The DLEP Message Header is immediately followed by zero or more DLEP Data Items, encoded in TLVs, as defined in this document.

11.3. DLEP Generic Data Item

All DLEP Data Items contain the following fields:

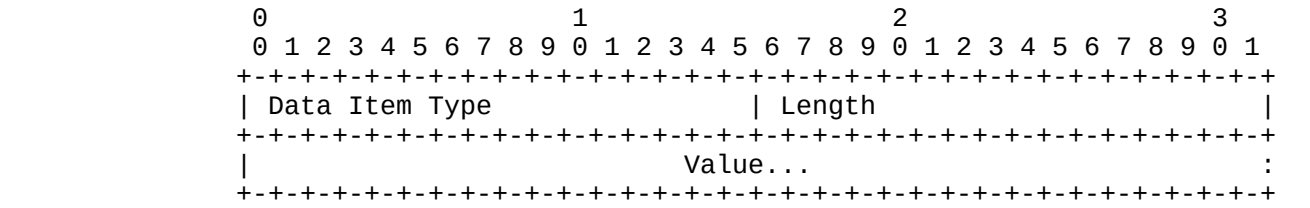


Figure 5: DLEP Generic Data Item

Data Item Type: A 16-bit unsigned integer field specifying the type of Data Item being sent.

Length: The length in octets, expressed as a 16-bit unsigned integer, of the Value field of the Data Item. This length MUST NOT include the length of the Data Item Type and Length fields.

Value: A field of <Length> octets, which contains data specific to a particular Data Item.

12. DLEP Signals and Messages

12.1. General Processing Rules

If an unrecognized, or unexpected Signal is received, or a received Signal contains unrecognized, invalid, or disallowed duplicate Data Items, the receiving implementation MUST ignore the Signal.

If a Signal is received with a TTL value that is NOT equal to 255, the receiving implementation MUST ignore the Signal.

If an unrecognized Message is received, the receiving implementation MUST issue a Session Termination Message (Section 12.9) containing a Status Data Item (Section 13.1) with status code set to 128 'Unknown Message', see Table 2, and transition to the Session Termination state.

If an unexpected Message is received, the receiving implementation MUST issue a Session Termination Message containing a Status Data Item with status code set to 129 'Unexpected Message', and transition to the Session Termination state.

If a received Message contains unrecognized, invalid, or disallowed duplicate Data Items, the receiving implementation MUST issue a

Session Termination Message containing a Status Data Item with status code set to 130 'Invalid Data', and transition to the Session Termination state.

If a packet in the TCP stream is received with a TTL value other than 255, the receiving implementation MUST immediately transition to the Session Reset state.

Prior to the exchange of Destination Up (Section 12.11) and Destination Up Response (Section 12.12) Messages, or Destination Announce (Section 12.13) and Destination Announce Response (Section 12.14) Messages, no Messages concerning a destination may be sent. An implementation receiving any Message with such an unannounced destination MUST terminate the session by issuing a Session Termination Message containing a Status Data Item with status code set to 131 'Invalid Destination', and transition to the Session Termination state.

After exchanging Destination Down (Section 12.15) and Destination Down Response (Section 12.16) Messages, no Messages concerning a destination may be a sent until a new Destination Up or Destination Announce Message is sent. An implementation receiving a Message about a destination previously announced as 'down' MUST terminate the session by issuing a Session Termination Message containing a Status Data Item with status code set to 131 'Invalid Destination', and transition to the Session Termination state.

12.2. Status code processing

The behavior of a DLEP participant receiving a Message containing a Status Data Item (Section 13.1) is defined by the failure mode associated with the value of the status code field, see Table 2. All status code values less than 100 have a failure mode of 'Continue', all other status codes have a failure mode of 'Terminate'.

A DLEP participant receiving any Message apart from Session Termination Message (Section 12.9) containing a Status Data Item with a status code value with failure mode 'Terminate' MUST immediately issue a Session Termination Message echoing the received Status Data Item, and then transition to the Session Termination state.

A DLEP participant receiving a Message containing a Status Data Item with a status code value with failure mode 'Continue' can continue normal operation of the session.

12.3. Peer Discovery Signal

A Peer Discovery Signal SHOULD be sent by a DLEP router to discover DLEP modems in the network, see Section 7.1.

A Peer Discovery Signal MUST be encoded within a UDP packet. The destination MUST be set to the DLEP well-known address and port number. For routers supporting both IPv4 and IPv6 DLEP operation, it is RECOMMENDED that IPv6 be selected as the transport. The source IP address MUST be set to the router IP address associated with the DLEP interface. There is no DLEP-specific restriction on source port.

To construct a Peer Discovery Signal, the Signal Type value in the Signal Header is set to 1 (see Signal Type Registration (Section 15.2)).

The Peer Discovery Signal MAY contain a Peer Type Data Item (Section 13.4).

12.4. Peer Offer Signal

A Peer Offer Signal MUST be sent by a DLEP modem in response to a properly formatted and addressed Peer Discovery Signal (Section 12.3).

A Peer Offer Signal MUST be encoded within a UDP packet. The IP source and destination fields in the packet MUST be set by swapping the values received in the Peer Discovery Signal. The Peer Offer Signal completes the discovery process, see Section 7.1.

To construct a Peer Offer Signal, the Signal Type value in the Signal Header is set to 2 (see Signal Type Registration (Section 15.2)).

The Peer Offer Signal MAY contain a Peer Type Data Item (Section 13.4).

The Peer Offer Signal MAY contain one or more of any of the following Data Items, with different values:

- IPv4 Connection Point (Section 13.2)
- IPv6 Connection Point (Section 13.3)

The IP Connection Point Data Items indicate the unicast address the router MUST use when connecting the DLEP TCP session.

12.5. Session Initialization Message

A Session Initialization Message MUST be sent by a DLEP router as the first Message of the DLEP TCP session. It is sent by the router after a TCP connect to an address/port combination that was obtained either via receipt of a Peer Offer, or from a priori configuration.

To construct a Session Initialization Message, the Message Type value in the Message Header is set to 1 (see Message Type Registration (Section 15.3)).

The Session Initialization Message MUST contain one of each of the following Data Items:

- Heartbeat Interval Data Item (Section 13.5)
- Peer Type (Section 13.4)

The Session Initialization Message MUST contain an Extensions Supported Data Item (Section 13.6), if DLEP extensions are supported.

The Session Initialization Message MAY contain one or more of each of the following Data Items, with different values, and the data item Add flag set to 1:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)
- IPv4 Attached Subnet (Section 13.10)
- IPv6 Attached Subnet (Section 13.11)

If any optional extensions are supported by the implementation, they MUST be enumerated in the Extensions Supported Data Item. If an Extensions Supported Data Item does not exist in a Session Initialization Message, the modem MUST conclude that there is no support for extensions in the router.

DLEP Heartbeats are not started until receipt of the Session Initialization Response Message (Section 12.6), and therefore implementations MUST use their own timeout heuristics for this Message.

As an exception to the general rule governing an implementation receiving an unrecognized Data Item in a Message, see Section 12.1, if a Session Initialization Message contains one or more Extension Supported Data Items announcing support for extensions that the

implementation does not recognize, then the implementation MAY ignore Data Items it does not recognize.

12.6. Session Initialization Response Message

A Session Initialization Response Message MUST be sent by a DLEP modem in response to a received Session Initialization Message (Section 12.5).

To construct a Session Initialization Response Message, the Message Type value in the Message Header is set to 2 (see Message Type Registration (Section 15.3)).

The Session Initialization Response Message MUST contain one of each of the following Data Items:

- Status (Section 13.1)
- Peer Type (Section 13.4)
- Heartbeat Interval (Section 13.5)
- Maximum Data Rate (Receive) (Section 13.12)
- Maximum Data Rate (Transmit) (Section 13.13)
- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

The Session Initialization Response Message MUST contain one of each of the following Data Items, if the Data Item will be used during the lifetime of the session:

- Resources (Section 13.17)
- Relative Link Quality (Receive) (Section 13.18)
- Relative Link Quality (Transmit) (Section 13.19)
- Maximum Transmission Unit (MTU) (Section 13.20)

The Session Initialization Response Message MUST contain an Extensions Supported Data Item (Section 13.6), if DLEP extensions are supported.

The Session Initialization Response Message MAY contain one or more of each of the following Data Items, with different values, and the data item Add flag set to 1:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)
- IPv4 Attached Subnet (Section 13.10)
- IPv6 Attached Subnet (Section 13.11)

The Session Initialization Response Message completes the DLEP session establishment; the modem should transition to the In-Session state when the Message is sent, and the router should transition to the In-Session state upon receipt of an acceptable Session Initialization Response Message.

All supported metric Data Items MUST be included in the Session Initialization Response Message, with default values to be used on a session-wide basis. This can be viewed as the modem 'declaring' all supported metrics at DLEP session initialization. Receipt of any further DLEP Message containing a metric Data Item not included in the Session Initialization Response Message MUST be treated as an error, resulting in the termination of the DLEP session between router and modem.

If any optional extensions are supported by the modem, they MUST be enumerated in the Extensions Supported Data Item. If an Extensions Supported Data Item does not exist in a Session Initialization Response Message, the router MUST conclude that there is no support for extensions in the modem.

After the Session Initialization/Session Initialization Response Messages have been successfully exchanged, implementations MUST only use extensions that are supported by both DLEP participants, see Section 7.2.

12.7. Session Update Message

A Session Update Message MAY be sent by a DLEP participant to indicate local Layer 3 address changes, or metric changes on a session-wide basis.

To construct a Session Update Message, the Message Type value in the Message Header is set to 3 (see Message Type Registration (Section 15.3)).

The Session Update Message MAY contain one or more of each of the following Data Items, with different values:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)
- IPv4 Attached Subnet (Section 13.10)
- IPv6 Attached Subnet (Section 13.11)

When sent by a modem, the Session Update Message MAY contain one of each of the following Data Items:

- Maximum Data Rate (Receive) (Section 13.12)
- Maximum Data Rate (Transmit) (Section 13.13)
- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

When sent by a modem, the Session Update Message MAY contain one of each of the following Data Items, if the Data Item is in use by the session:

- Resources (Section 13.17)
- Relative Link Quality (Receive) (Section 13.18)
- Relative Link Quality (Transmit) (Section 13.19)
- Maximum Transmission Unit (MTU) (Section 13.20)

If metrics are supplied with the Session Update Message (e.g., Maximum Data Rate), these metrics are considered to be session-wide, and therefore MUST be applied to all destinations in the information base associated with the DLEP session. This includes destinations for which metrics may have been stored based on received Destination Update messages.

It should be noted that Session Update Messages can be sent by both routers and modems. For example, addition of an IPv4 address on the router MAY prompt a Session Update Message to its attached modems. Also, for example, a modem that changes its Maximum Data Rate

(Receive) for all destinations MAY reflect that change via a Session Update Message to its attached router(s).

Concerning Layer 3 addresses and subnets: If the modem is capable of understanding and forwarding this information (via mechanisms not defined by DLEP), the update would prompt any remote DLEP-enabled modems to issue a Destination Update Message (Section 12.17) to their local routers with the new (or deleted) addresses and subnets.

12.8. Session Update Response Message

A Session Update Response Message MUST be sent by a DLEP participant when a Session Update Message (Section 12.7) is received.

To construct a Session Update Response Message, the Message Type value in the Message Header is set to 4 (see Message Type Registration (Section 15.3)).

The Session Update Response Message MUST contain a Status Data Item (Section 13.1).

12.9. Session Termination Message

When a DLEP participant determines the DLEP session needs to be terminated, the participant MUST send (or attempt to send) a Session Termination Message.

To construct a Session Termination Message, the Message Type value in the Message Header is set to 5 (see Message Type Registration (Section 15.3)).

The Session Termination Message MUST contain Status Data Item (Section 13.1).

It should be noted that Session Termination Messages can be sent by both routers and modems.

12.10. Session Termination Response Message

A Session Termination Response Message MUST be sent by a DLEP participant when a Session Termination Message (Section 12.9) is received.

To construct a Session Termination Response Message, the Message Type value in the Message Header is set to 6 (see Message Type Registration (Section 15.3)).

There are no valid Data Items for the Session Termination Response Message.

Receipt of a Session Termination Response Message completes the tear-down of the DLEP session, see Section 7.4.

12.11. Destination Up Message

Destination Up Messages MAY be sent by a modem to inform its attached router of the presence of a new reachable destination.

To construct a Destination Up Message, the Message Type value in the Message Header is set to 7 (see Message Type Registration (Section 15.3)).

The Destination Up Message MUST contain a MAC Address Data Item (Section 13.7).

The Destination Up Message SHOULD contain one or more of each of the following Data Items, with different values:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)

The Destination Up Message MAY contain one of each of the following Data Items:

- Maximum Data Rate (Receive) (Section 13.12)
- Maximum Data Rate (Transmit) (Section 13.13)
- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

The Destination Up Message MAY contain one of each of the following Data Items, if the Data Item is in use by the session:

- Resources (Section 13.17)
- Relative Link Quality (Receive) (Section 13.18)
- Relative Link Quality (Transmit) (Section 13.19)
- Maximum Transmission Unit (MTU) (Section 13.20)

The Destination Up Message MAY contain one or more of each of the following Data Items, with different values:

- IPv4 Attached Subnet (Section 13.10)
- IPv6 Attached Subnet (Section 13.11)

A router receiving a Destination Up Message allocates the necessary resources, creating an entry in the information base with the specifics (MAC Address, Latency, Data Rate, etc.) of the destination. The information about this destination will persist in the router's information base until a Destination Down Message (Section 12.15) is received, indicating that the modem has lost contact with the remote node, or the implementation transitions to the Session Termination state.

12.12. Destination Up Response Message

A router MUST send a Destination Up Response Message when a Destination Up Message (Section 12.11) is received.

To construct a Destination Up Response Message, the Message Type value in the Message Header is set to 8 (see Message Type Registration (Section 15.3)).

The Destination Up Response Message MUST contain one of each of the following Data Items:

- MAC Address (Section 13.7)
- Status (Section 13.1)

A router that wishes to receive further information concerning the destination identified in the corresponding Destination Up Message MUST set the status code of the included Status Data Item to 0 'Success', see Table 2.

If the router has no interest in the destination identified in the corresponding Destination Up Message, then it MAY set the status code of the included Status Data Item to 1 'Not Interested'.

A modem receiving a Destination Up Response Message containing a Status Data Item with status code of any value other than 0 'Success' MUST NOT send further Destination messages about the destination, e.g. Destination Down (Section 12.15) or Destination Update (Section 12.17) with the same MAC address.

12.13. Destination Announce Message

Usually a modem will discover the presence of one or more remote router/modem pairs and announce each destination's arrival by sending a corresponding Destination Up Message (Section 12.11) to the router. However, there may be times when a router wishes to express an interest in a destination that has yet to be announced, typically a multicast destination. Destination Announce Messages MAY be sent by a router to announce such an interest.

A Destination Announce Message MAY also be sent by a router to request information concerning a destination in which it has previously declined interest, via the 1 'Not Interested' status code in a Destination Up Response Message (Section 12.12), see Table 2, or declared as 'down', via the Destination Down Message (Section 12.15).

To construct a Destination Announce Message, the Message Type value in the Message Header is set to 9 (see Message Type Registration (Section 15.3)).

The Destination Announce Message MUST contain a MAC Address Data Item (Section 13.7).

The Destination Announce Message MAY contain zero or more of the following Data Items, with different values:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)

One of the advantages of implementing DLEP is to leverage the modem's knowledge of the links between remote destinations allowing routers to avoid using probed neighbor discovery techniques, therefore modem implementations SHOULD announce available destinations via the Destination Up Message, rather than relying on Destination Announce Messages.

12.14. Destination Announce Response Message

A modem MUST send a Destination Announce Response Message when a Destination Announce Message (Section 12.13) is received.

To construct a Destination Announce Response Message, the Message Type value in the Message Header is set to 10 (see Message Type Registration (Section 15.3)).

The Destination Announce Response Message MUST contain one of each of the following Data Items:

- MAC Address (Section 13.7)
- Status (Section 13.1)

The Destination Announce Response Message MAY contain one or more of each of the following Data Items, with different values:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)
- IPv4 Attached Subnet (Section 13.10)
- IPv6 Attached Subnet (Section 13.11)

The Destination Announce Response Message MAY contain one of each of the following Data Items:

- Maximum Data Rate (Receive) (Section 13.12)
- Maximum Data Rate (Transmit) (Section 13.13)
- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

The Destination Announce Response Message MAY contain one of each of the following Data Items, if the Data Item is in use by the session:

- Resources (Section 13.17)
- Relative Link Quality (Receive) (Section 13.18)
- Relative Link Quality (Transmit) (Section 13.19)
- Maximum Transmission Unit (MTU) (Section 13.20)

If a modem is unable to report information immediately about the requested information, if the destination is not currently reachable, for example, the status code in the Status Data Item MUST be set to 2 'Request Denied', see Table 2.

After sending a Destination Announce Response Message containing a Status Data Item with status code of 0 'Success', a modem then announces changes to the link to the destination via Destination Update Messages.

When a successful Destination Announce Response Message is received, the router should add knowledge of the available destination to its information base.

12.15. Destination Down Message

A modem MUST send a Destination Down Message to report when a destination (a remote node or a multicast group) is no longer reachable.

A router MAY send a Destination Down Message to report when it no longer requires information concerning a destination.

To construct a Destination Down Message, the Message Type value in the Message Header is set to 11 (see Message Type Registration (Section 15.3)).

The Destination Down Message MUST contain a MAC Address Data Item (Section 13.7).

It should be noted that both modem and router may send a Destination Down Message to their peer, regardless of which participant initially indicated the destination to be 'up'.

12.16. Destination Down Response Message

A Destination Down Response MUST be sent by the recipient of a Destination Down Message (Section 12.15) to confirm that the relevant data concerning the destination has been removed from the information base.

To construct a Destination Down Response Message, the Message Type value in the Message Header is set to 12 (see Message Type Registration (Section 15.3)).

The Destination Down Response Message MUST contain one of each of the following Data Items:

- MAC Address (Section 13.7)
- Status (Section 13.1)

12.17. Destination Update Message

A modem SHOULD send the Destination Update Message when it detects some change in the information base for a given destination (remote node or multicast group). Some examples of changes that would prompt a Destination Update Message are:

- Change in link metrics (e.g., Data Rates)
- Layer 3 addressing change

To construct a Destination Update Message, the Message Type value in the Message Header is set to 13 (see Message Type Registration (Section 15.3)).

The Destination Update Message MUST contain a MAC Address Data Item (Section 13.7).

The Destination Update Message MAY contain one of each of the following Data Items:

- Maximum Data Rate (Receive) (Section 13.12)
- Maximum Data Rate (Transmit) (Section 13.13)
- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

The Destination Update Message MAY contain one of each of the following Data Items, if the Data Item is in use by the session:

- Resources (Section 13.17)
- Relative Link Quality (Receive) (Section 13.18)
- Relative Link Quality (Transmit) (Section 13.19)
- Maximum Transmission Unit (MTU) (Section 13.20)

The Destination Update Message MAY contain one or more of each of the following Data Items, with different values:

- IPv4 Address (Section 13.8)
- IPv6 Address (Section 13.9)
- IPv4 Attached Subnet (Section 13.10)
- IPv6 Attached Subnet (Section 13.11)

Metrics supplied in this message overwrite metrics provided in a previously received Session or Destination Up Messages.

It should be noted that this Message has no corresponding response.

12.18. Link Characteristics Request Message

The Link Characteristics Request Message MAY be sent by a router to request that the modem initiate changes for specific characteristics of the link. The request can reference either a real destination (e.g., a remote node), or a logical destination (e.g., a multicast group) within the network.

To construct a Link Characteristics Request Message, the Message Type value in the Message Header is set to 14 (see Message Type Registration (Section 15.3)).

The Link Characteristics Request Message MUST contain one of the following Data Items:

- MAC Address (Section 13.7)

The Link Characteristics Request Message MUST contain at least one of each of the following Data Items:

- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

The Link Characteristics Request Message MAY contain either a Current Data Rate (CDRR or CDRT) Data Item to request a different datarate than is currently allocated, a Latency Data Item to request that traffic delay on the link not exceed the specified value, or both.

The router sending a Link Characteristics Request Message should be aware that a request may take an extended period of time to complete.

12.19. Link Characteristics Response Message

A modem MUST send a Link Characteristics Response Message when a Link Characteristics Request Message (Section 12.18) is received.

To construct a Link Characteristics Response Message, the Message Type value in the Message Header is set to 15 (see Message Type Registration (Section 15.3)).

The Link Characteristics Response Message MUST contain one of each of the following Data Items:

- MAC Address (Section 13.7)
- Status (Section 13.1)

The Link Characteristics Response Message SHOULD contain one of each of the following Data Items:

- Maximum Data Rate (Receive) (Section 13.12)
- Maximum Data Rate (Transmit) (Section 13.13)
- Current Data Rate (Receive) (Section 13.14)
- Current Data Rate (Transmit) (Section 13.15)
- Latency (Section 13.16)

The Link Characteristics Response Message MAY contain one of each of the following Data Items, if the Data Item is in use by the session:

- Resources (Section 13.17)
- Relative Link Quality (Receive) (Section 13.18)
- Relative Link Quality (Transmit) (Section 13.19)
- Maximum Transmission Unit (MTU) (Section 13.20)

The Link Characteristics Response Message MUST contain a complete set of metric Data Items, referencing all metrics declared in the Session Initialization Response Message (Section 12.6). The values in the metric Data Items in the Link Characteristics Response Message MUST reflect the link characteristics after the request has been processed.

If an implementation is not able to alter the characteristics of the link in the manner requested, then the status code of the Status Data Item MUST be set to 2 ’Request Denied’, see Table 2.

12.20. Heartbeat Message

A Heartbeat Message MUST be sent by a DLEP participant every N milliseconds, where N is defined in the Heartbeat Interval Data Item (Section 13.5) of the Session Initialization Message (Section 12.5) or Session Initialization Response Message (Section 12.6).

To construct a Heartbeat Message, the Message Type value in the Message Header is set to 16 (see Message Type Registration (Section 15.3)).

There are no valid Data Items for the Heartbeat Message.

The Message is used by DLEP participants to detect when a DLEP session peer (either the modem or the router) is no longer communicating, see Section 7.3.1.

13. DLEP Data Items

The core DLEP Data Items are:

Type Code	Description
0	Reserved
1	Status (Section 13.1)
2	IPv4 Connection Point (Section 13.2)
3	IPv6 Connection Point (Section 13.3)
4	Peer Type (Section 13.4)
5	Heartbeat Interval (Section 13.5)
6	Extensions Supported (Section 13.6)
7	MAC Address (Section 13.7)
8	IPv4 Address (Section 13.8)
9	IPv6 Address (Section 13.9)
10	IPv4 Attached Subnet (Section 13.10)
11	IPv6 Attached Subnet (Section 13.11)
12	Maximum Data Rate (Receive) (MDRR) (Section 13.12)
13	Maximum Data Rate (Transmit) (MDRT) (Section 13.13)
14	Current Data Rate (Receive) (CDRR) (Section 13.14)
15	Current Data Rate (Transmit) (CDRT) (Section 13.15)
16	Latency (Section 13.16)
17	Resources (RES) (Section 13.17)
18	Relative Link Quality (Receive) (RLQR) (Section 13.18)
19	Relative Link Quality (Transmit) (RLQT) (Section 13.19)
20	Maximum Transmission Unit (MTU) (Section 13.20)
21-65407	Reserved for future extensions
65408-65534	Private Use. Available for experiments
65535	Reserved

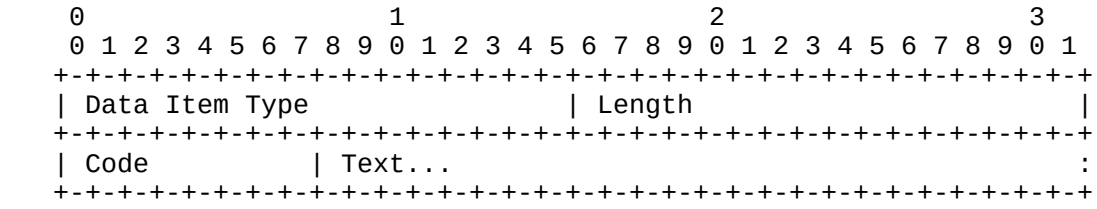
Table 1: DLEP Data Item types

13.1. Status

For the Session Termination Message (Section 12.9), the Status Data Item indicates a reason for the termination. For all response Messages, the Status Data Item is used to indicate the success or failure of the previously received Message.

The Status Data Item includes an optional Text field that can be used to provide a textual description of the status. The use of the Text field is entirely up to the receiving implementation, e.g., it could be output to a log file or discarded. If no Text field is supplied with the Status Data Item, the Length field MUST be set to 1.

The Status Data Item contains the following fields:



Data Item Type: 1

Length: 1 + Length of text, in octets

Status Code: One of the codes defined in Table 2 below.

Text: UTF-8 encoded string of UNICODE [RFC3629] characters, describing the cause, used for implementation defined purposes. Since this field is used for description, implementations SHOULD limit characters in this field to printable characters.

An implementation MUST NOT assume the Text field is a NUL-terminated string of printable characters.

Status Code	Failure Mode	Description	Reason
0	Continue	Success	The Message was processed successfully.
1	Continue	Not Interested	The receiver is not interested in this Message subject, e.g. in a Destination Up

			Response Message (Section 12.12) to indicate no further Messages about the destination.
2	Continue	Request Denied	The receiver refuses to complete the request.
3	Continue	Inconsistent Data	One or more Data Items in the Message describe a logically inconsistent state in the network. For example, in the Destination Up Message (Section 12.11) when an announced subnet clashes with an existing destination subnet.
4-111	Continue	<Reserved>	Reserved for future extensions.
112-127	Continue	<Private Use>	Available for experiments.
128	Terminate	Unknown Message	The Message was not recognized by the implementation.
129	Terminate	Unexpected Message	The Message was not expected while the device was in the current state, e.g., a Session Initialization Message (Section 12.5) in the In-Session state.
130	Terminate	Invalid Data	One or more Data Items in the Message are invalid, unexpected or incorrectly duplicated.
131	Terminate	Invalid Destination	The destination included in the Message does not match a previously announced

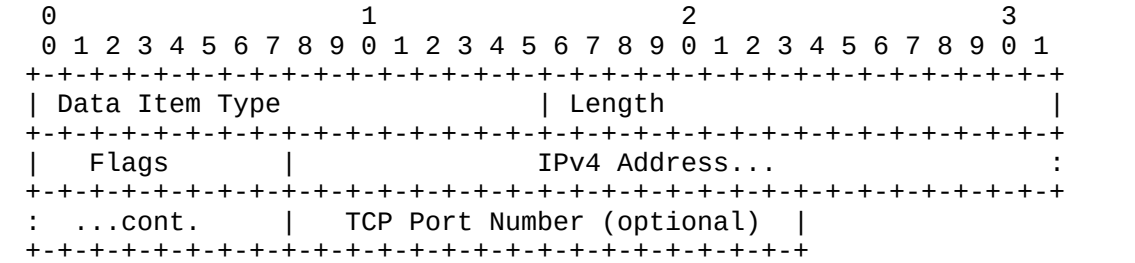
			destination. For example, in the Link Characteristic Response Message (Section 12.19).
132	Terminate	Timed Out	The session has timed out.
133-239	Terminate	<Reserved>	Reserved for future extensions.
240-254	Terminate	<Private Use>	Available for experiments.
255	Terminate	<Reserved>	Reserved.

Table 2: DLEP Status Codes

13.2. IPv4 Connection Point

The IPv4 Connection Point Data Item indicates the IPv4 address and, optionally, the TCP port number on the modem available for connections. If provided, the router MUST use this information to initiate the TCP connection to the modem.

The IPv4 Connection Point Data Item contains the following fields:



Data Item Type: 2

Length: 5 (or 7 if TCP Port included)

Flags: Flags field, defined below.

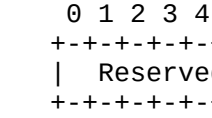
IPv4 Address: The IPv4 address listening on the modem.

TCP Port Number: TCP Port number on the modem.

If the Length field is 7, the port number specified MUST be used to establish the TCP session. If the TCP Port Number is omitted, i.e.

the Length field is 5, the router MUST use the DLEP well-known port number (Section 15.14) to establish the TCP connection.

The Flags field is defined as:



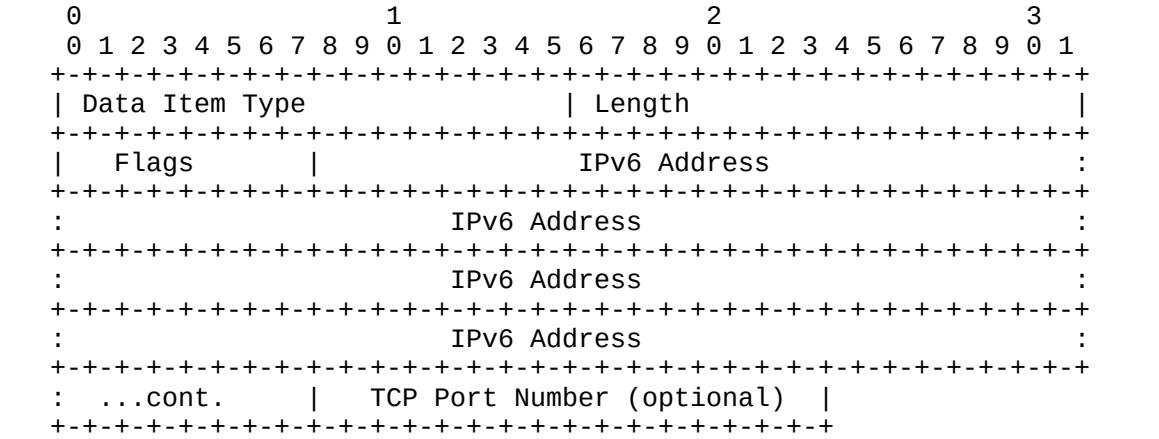
T: Use TLS flag, indicating whether the TCP connection to the given address and port requires the use of TLS [RFC5246] (1), or not (0).

Reserved: MUST be zero. Left for future assignment.

13.3. IPv6 Connection Point

The IPv6 Connection Point Data Item indicates the IPv6 address and, optionally, the TCP port number on the modem available for connections. If provided, the router MUST use this information to initiate the TCP connection to the modem.

The IPv6 Connection Point Data Item contains the following fields:



Data Item Type: 3

Length: 17 (or 19 if TCP Port included)

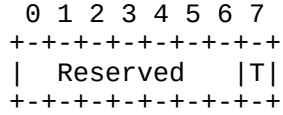
Flags: Flags field, defined below.

IPv6 Address: The IPv6 address listening on the modem.

TCP Port Number: TCP Port number on the modem.

If the Length field is 19, the port number specified MUST be used to establish the TCP session. If the TCP Port Number is omitted, i.e. the Length field is 17, the router MUST use the DLEP well-known port number (Section 15.14) to establish the TCP connection.

The Flags field is defined as:



T: Use TLS flag, indicating whether the TCP connection to the given address and port requires the use of TLS [RFC5246] (1), or not (0).

Reserved: MUST be zero. Left for future assignment.

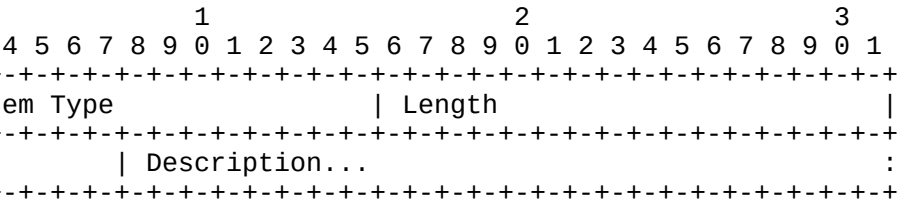
13.4. Peer Type

The Peer Type Data Item is used by the router and modem to give additional information as to its type and the properties of the over-the-air control-plane.

With some devices, access to the shared RF medium is strongly controlled. One example of this would be satellite modems - where protocols, proprietary in nature, have been developed to insure a given modem has authorization to connect to the shared medium. Another example of this class of modems is governmental/military devices, where elaborate mechanisms have been developed to ensure that only authorized devices can connect to the shared medium. Contrasting with the above, there are modems where no such access control is used. An example of this class of modem would be one that supports the 802.11 ad-hoc mode of operation. The Secured Medium flag is used to indicate if access control is in place.

The Peer Type Data Item includes a textual description of the peer that is envisioned to be used for informational purposes (e.g., as output in a display command).

The Peer Type Data Item contains the following fields:



Data Item Type: 4

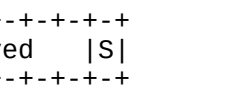
Length: 1 + Length of Peer Type string, in octets.

Flags: Flags field, defined below.

Description: UTF-8 encoded string of UNICODE [RFC3629] characters. For example, a satellite modem might set this variable to "Satellite terminal". Since this Data Item is intended to provide additional information for display commands, sending implementations SHOULD limit the data to printable characters.

An implementation MUST NOT assume the Description field is a NUL-terminated string of printable characters.

The Flags field is defined as:



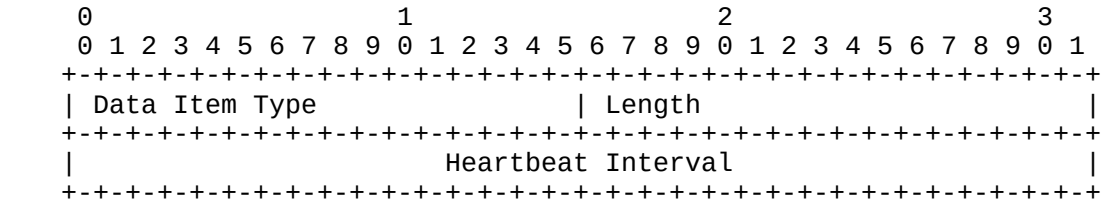
S: Secured Medium flag, used by a modem to indicate if the shared RF medium implements access control (1), or not (0). The Secured Medium flag only has meaning in Signals and Messages sent by a modem.

Reserved: MUST be zero. Left for future assignment.

13.5. Heartbeat Interval

The Heartbeat Interval Data Item is used to specify a period in milliseconds for Heartbeat Messages (Section 12.20).

The Heartbeat Interval Data Item contains the following fields:



Data Item Type: 5

Length: 4

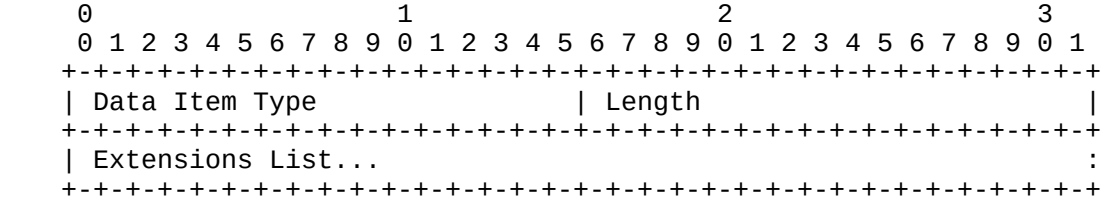
Heartbeat Interval: The interval in milliseconds, expressed as a 32-bit unsigned integer, for Heartbeat Messages. This value MUST NOT be 0.

As mentioned before, receipt of any valid DLEP Message MUST reset the heartbeat interval timer (e.g., valid DLEP Messages take the place of, and obviate the need for, additional Heartbeat Messages).

13.6. Extensions Supported

The Extensions Supported Data Item is used by the router and modem to negotiate additional optional functionality they are willing to support. The Extensions List is a concatenation of the types of each supported extension, found in the IANA DLEP Extensions repository. Each Extension Type definition includes which additional Signals and Data Items are supported.

The Extensions Supported Data Item contains the following fields:



Data Item Type: 6

Length: Length of the extensions list in octets. This is twice (2x) the number of extensions.

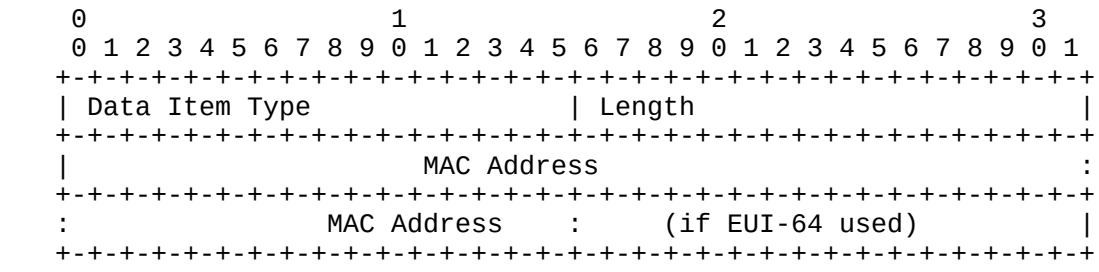
Extension List: A list of extensions supported, identified by their 2-octet value as listed in the extensions registry.

13.7. MAC Address

The MAC Address Data Item contains the address of the destination on the remote node.

DLEP can support MAC addresses in either EUI-48 or EUI-64 format, with the restriction that all MAC addresses for a given DLEP session MUST be in the same format, and MUST be consistent with the MAC address format of the connected modem (e.g., if the modem is connected to the router with an EUI-48 MAC, all destination addresses via that modem MUST be expressed in EUI-48 format).

Examples of a virtual destination would be a multicast MAC address, or the broadcast MAC (FF:FF:FF:FF:FF:FF).



Data Item Type: 7

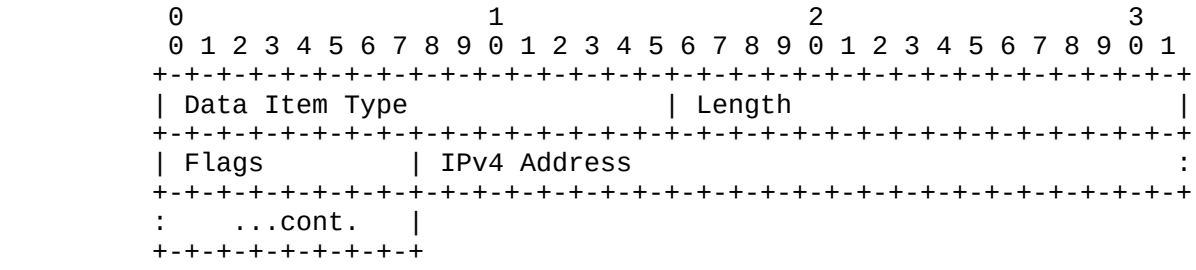
Length: 6 for EUI-48 format, or 8 for EUI-64 format

MAC Address: MAC Address of the destination.

13.8. IPv4 Address

When included in the Session Update Message, this Data Item contains the IPv4 address of the peer. When included in Destination Messages, this Data Item contains the IPv4 address of the destination. In either case, the Data Item also contains an indication of whether this is a new or existing address, or is a deletion of a previously known address.

The IPv4 Address Data Item contains the following fields:



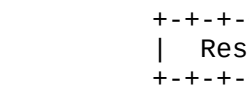
Data Item Type: 8

Length: 5

Flags: Flags field, defined below.

IPv4 Address: The IPv4 address of the destination or peer.

The Flags field is defined as:



A: Add/Drop flag, indicating whether this is a new or existing address (1), or a withdrawal of an address (0).

Reserved: MUST be zero. Reserved for future use.

13.8.1. IPv4 Address Processing

Processing of the IPv4 Address Data Item MUST be done within the context of the DLEP Peer session on which it is presented.

The handling of erroneous or logically inconsistent conditions depends upon the type of the message that contains the data item:

If the containing message is a Session Message, e.g., Session Initialization Message (Section 12.5), or Session Update Message (Section 12.7), the receiver of inconsistent information MUST issue a Session Termination Message (Section 12.9) containing a Status Data Item (Section 13.1) with status code set to 130 'Invalid Data', and transition to the Session Termination state. Examples of such conditions are:

- An address Drop operation referencing an address that is not associated with the peer in the current session.

- An address Add operation referencing an address that has already been added to the peer in the current session.

If the containing message is a Destination Message, e.g., Destination Up Message (Section 12.11), or Destination Update Message (Section 12.17), the receiver of inconsistent information MAY issue the appropriate response message containing a Status Data Item, with status code set to 3 'Inconsistent Data', but MUST continue with session processing. Examples of such conditions are:

- An address Add operation referencing an address that has already been added to the destination in the current session.
- An address Add operation referencing an address that is associated with a different destination or the peer in the current session.
- An address Add operation referencing an address that makes no sense, for example defined as not forwardable in [RFC6890].
- An address Drop operation referencing an address that is not associated with the destination in the current session.

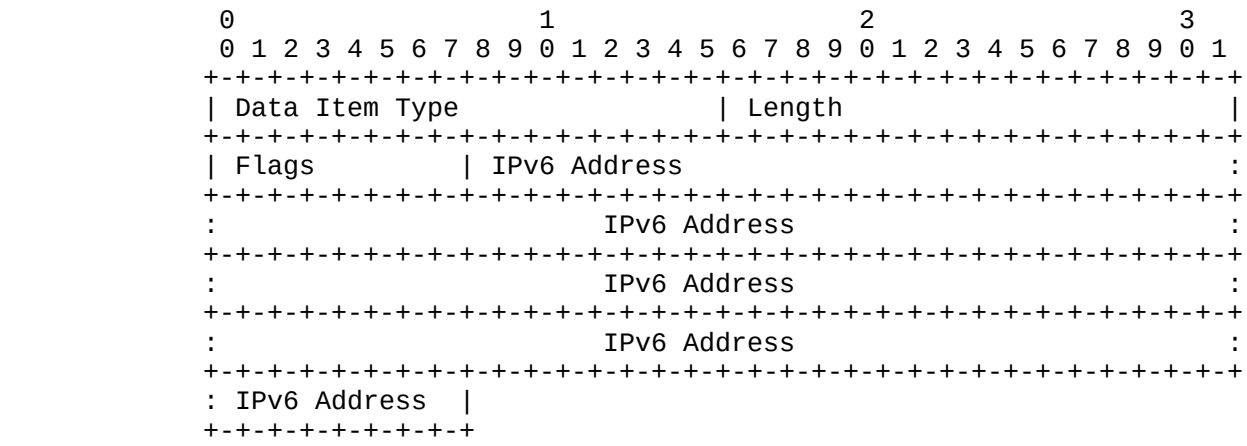
If no response message is appropriate, for example, the Destination Update Message, then the implementation MUST continue with session processing.

Modems that do not track IPv4 addresses MUST silently ignore IPv4 Address Data Items.

13.9. IPv6 Address

When included in the Session Update Message, this Data Item contains the IPv6 address of the peer. When included in Destination Messages, this Data Item contains the IPv6 address of the destination. In either case, the Data Item also contains an indication of whether this is a new or existing address, or is a deletion of a previously known address.

The IPv6 Address Data Item contains the following fields:



Data Item Type: 9

Length: 17

Flags: Flags field, defined below.

IPv6 Address: IPv6 Address of the destination or peer.

The Flags field is defined as:

A: Add/Drop flag, indicating whether this is a new or existing address (1), or a withdrawal of an address (0).

Reserved: MUST be zero. Reserved for future use.

13.9.1. IPv6 Address Processing

Processing of the IPv6 Address Data Item MUST be done within the context of the DLEP Peer session on which it is presented.

The handling of erroneous or logically inconsistent conditions depends upon the type of the message that contains the data item:

If the containing message is a Session Message, e.g., Session Initialization Message (Section 12.5), or Session Update Message (Section 12.7), the receiver of inconsistent information MUST issue a Session Termination Message (Section 12.9) containing a Status Data

Item (Section 13.1) with status code set to 130 'Invalid Data', and transition to the Session Termination state. Examples of such conditions are:

- An address Drop operation referencing an address that is not associated with the peer in the current session.
- An address Add operation referencing an address that has already been added to the peer in the current session.

If the containing message is a Destination Message, e.g., Destination Up Message (Section 12.11), or Destination Update Message (Section 12.17), the receiver of inconsistent information MAY issue the appropriate response message containing a Status Data Item, with status code set to 3 'Inconsistent Data', but MUST continue with session processing. Examples of such conditions are:

- An address Add operation referencing an address that has already been added to the destination in the current session.
- An address Add operation referencing an address that is associated with a different destination or the peer in the current session.
- An address Add operation referencing an address that makes no sense, for example defined as not forwardable in [RFC6890].
- An address Drop operation referencing an address that is not associated with the destination in the current session.

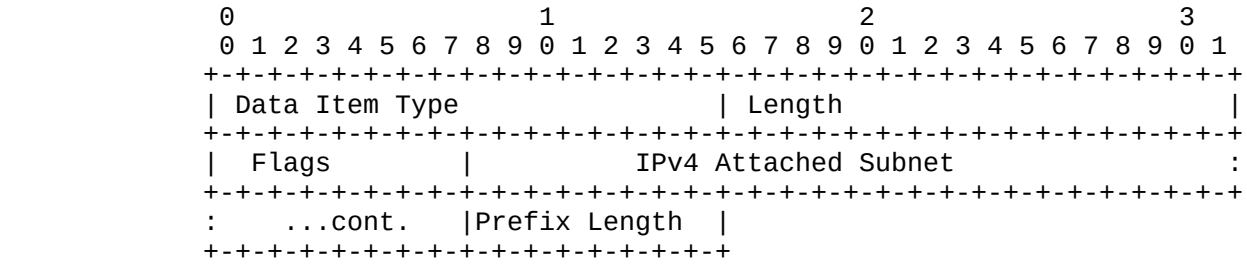
If no response message is appropriate, for example, the Destination Update Message, then the implementation MUST continue with session processing.

Modems that do not track IPv6 addresses MUST silently ignore IPv6 Address Data Items.

13.10. IPv4 Attached Subnet

The DLEP IPv4 Attached Subnet allows a device to declare that it has an IPv4 subnet (e.g., a stub network) attached, that it has become aware of an IPv4 subnet being present at a remote destination, or that it has become aware of the loss of a subnet at the remote destination.

The DLEP IPv4 Attached Subnet Data Item contains the following fields:



Data Item Type: 10

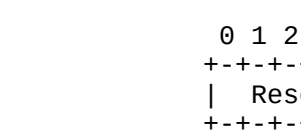
Length: 6

Flags: Flags field, defined below.

IPv4 Subnet: The IPv4 subnet reachable at the destination.

Prefix Length: Length of the prefix (0-32) for the IPv4 subnet. A prefix length outside the specified range MUST be considered as invalid.

The Flags field is defined as:



A: Add/Drop flag, indicating whether this is a new or existing subnet address (1), or a withdrawal of a subnet address (0).

Reserved: MUST be zero. Reserved for future use.

13.10.1. IPv4 Attached Subnet Processing

Processing of the IPv4 Attached Subnet Data Item MUST be done within the context of the DLEP Peer session on which it is presented.

If the containing message is a Session Message, e.g., Session Initialization Message (Section 12.5), or Session Update Message (Section 12.7), the receiver of inconsistent information MUST issue a Session Termination Message (Section 12.9) containing a Status Data Item (Section 13.1) with status code set to 130 'Invalid Data', and transition to the Session Termination state. Examples of such conditions are:

- A subnet Drop operation referencing a subnet that is not associated with the peer in the current session.
- A subnet Add operation referencing a subnet that has already been added to the peer in the current session.

If the containing message is a Destination Message, e.g., Destination Up Message (Section 12.11), or Destination Update Message (Section 12.17), the receiver of inconsistent information MAY issue the appropriate response message containing a Status Data Item, with status code set to 3 'Inconsistent Data', but MUST continue with session processing. Examples of such conditions are:

- A subnet Add operation referencing a subnet that has already been added to the destination in the current session.
- A subnet Add operation referencing a subnet that is associated with a different destination in the current session.
- An subnet Add operation referencing an subnet that makes no sense, for example defined as not forwardable in [RFC6890].
- A subnet Drop operation referencing a subnet that is not associated with the destination in the current session.

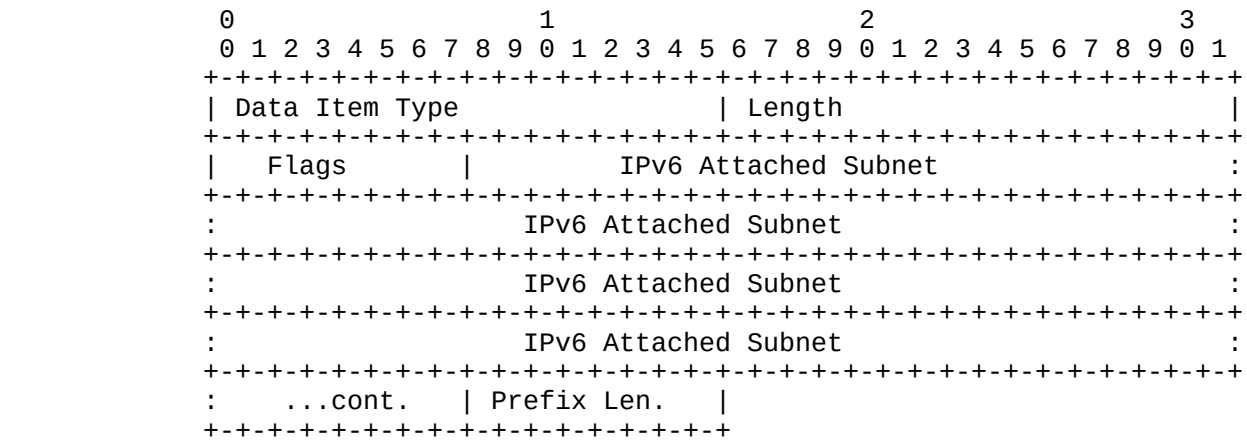
If no response message is appropriate, for example, the Destination Update Message, then the implementation MUST continue with session processing.

Modems that do not track IPv4 subnets MUST silently ignore IPv4 Attached Subnet Data Items.

13.11. IPv6 Attached Subnet

The DLEP IPv6 Attached Subnet allows a device to declare that it has an IPv6 subnet (e.g., a stub network) attached, that it has become aware of an IPv6 subnet being present at a remote destination, or that it has become aware of the loss of a subnet at the remote destination.

The DLEP IPv6 Attached Subnet Data Item contains the following fields:



Data Item Type: 11

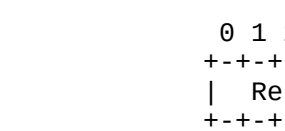
Length: 18

Flags: Flags field, defined below.

IPv6 Attached Subnet: The IPv6 subnet reachable at the destination.

Prefix Length: Length of the prefix (0-128) for the IPv6 subnet. A prefix length outside the specified range MUST be considered as invalid.

The Flags field is defined as:



A: Add/Drop flag, indicating whether this is a new or existing subnet address (1), or a withdrawal of a subnet address (0).

Reserved: MUST be zero. Reserved for future use.

13.11.1. IPv6 Attached Subnet Processing

Processing of the IPv6 Attached Subnet Data Item MUST be done within the context of the DLEP Peer session on which it is presented.

If the containing message is a Session Message, e.g., Session Initialization Message (Section 12.5), or Session Update Message (Section 12.7), the receiver of inconsistent information MUST issue a

Session Termination Message (Section 12.9) containing a Status Data Item (Section 13.1) with status code set to 130 'Invalid Data', and transition to the Session Termination state. Examples of such conditions are:

- A subnet Drop operation referencing a subnet that is not associated with the peer in the current session.
- A subnet Add operation referencing a subnet that has already been added to the peer in the current session.

If the containing message is a Destination Message, e.g., Destination Up Message (Section 12.11), or Destination Update Message (Section 12.17), the receiver of inconsistent information MAY issue the appropriate response message containing a Status Data Item, with status code set to 3 'Inconsistent Data', but MUST continue with session processing. Examples of such conditions are:

- A subnet Add operation referencing a subnet that has already been added to the destination in the current session.
- A subnet Add operation referencing a subnet that is associated with a different destination in the current session.
- An subnet Add operation referencing an subnet that makes no sense, for example defined as not forwardable in [RFC6890].
- A subnet Drop operation referencing a subnet that is not associated with the destination in the current session.

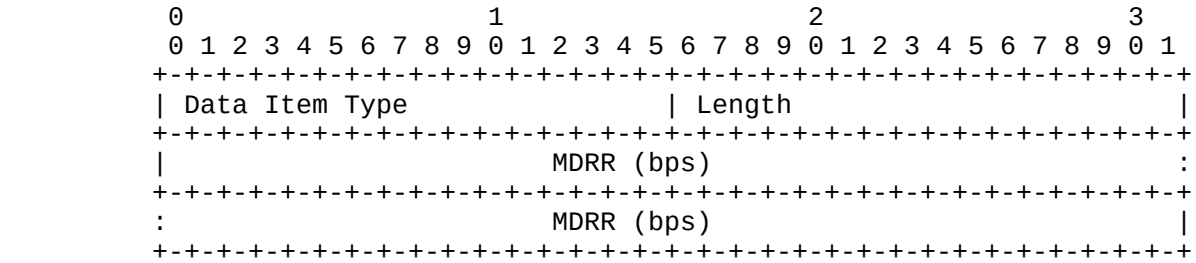
If no response message is appropriate, for example, the Destination Update Message, then the implementation MUST continue with session processing.

Modems that do not track IPv6 subnets MUST silently ignore IPv6 Attached Subnet Data Items.

13.12. Maximum Data Rate (Receive)

The Maximum Data Rate (Receive) (MDRR) Data Item is used to indicate the maximum theoretical data rate, in bits per second, that can be achieved while receiving data on the link.

The Maximum Data Rate (Receive) Data Item contains the following fields:



Data Item Type: 12

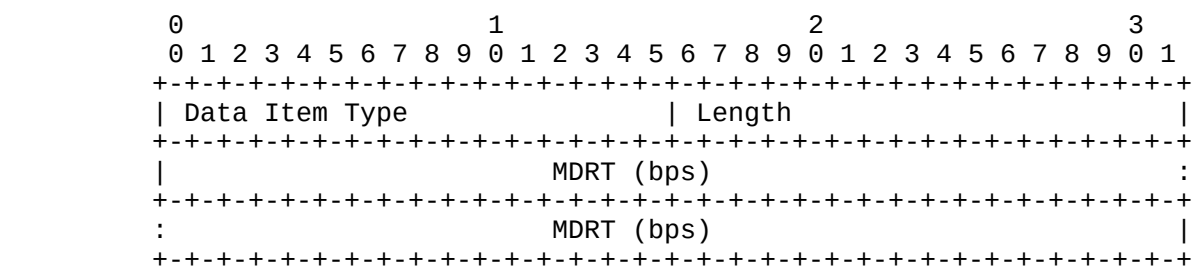
Length: 8

Maximum Data Rate (Receive): A 64-bit unsigned integer, representing the maximum theoretical data rate, in bits per second (bps), that can be achieved while receiving on the link.

13.13. Maximum Data Rate (Transmit)

The Maximum Data Rate (Transmit) (MDRT) Data Item is used to indicate the maximum theoretical data rate, in bits per second, that can be achieved while transmitting data on the link.

The Maximum Data Rate (Transmit) Data Item contains the following fields:



Data Item Type: 13

Length: 8

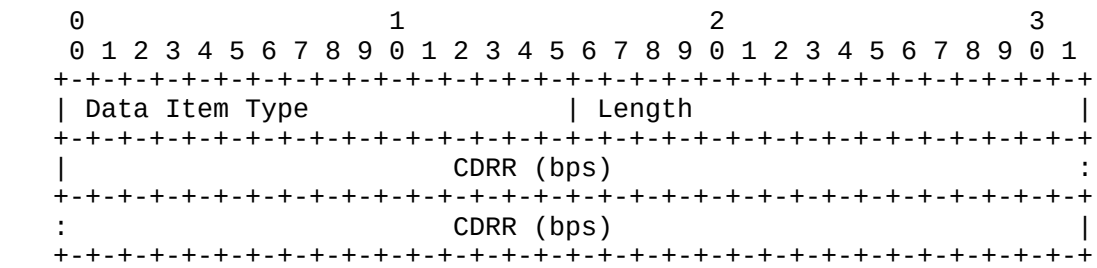
Maximum Data Rate (Transmit): A 64-bit unsigned integer, representing the maximum theoretical data rate, in bits per second (bps), that can be achieved while transmitting on the link.

13.14. Current Data Rate (Receive)

The Current Data Rate (Receive) (CDRR) Data Item is used to indicate the rate at which the link is currently operating for receiving traffic.

When used in the Link Characteristics Request Message (Section 12.18), Current Data Rate (Receive) represents the desired receive rate, in bits per second, on the link.

The Current Data Rate (Receive) Data Item contains the following fields:



Data Item Type: 14

Length: 8

Current Data Rate (Receive): A 64-bit unsigned integer, representing the current data rate, in bits per second, that can currently be achieved while receiving traffic on the link.

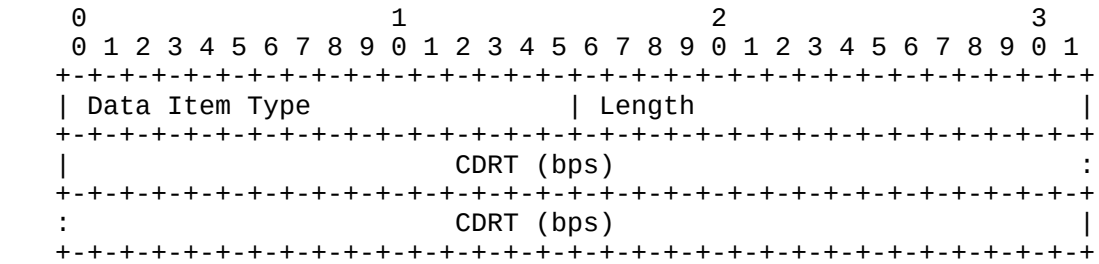
If there is no distinction between Current Data Rate (Receive) and Maximum Data Rate (Receive) (Section 13.12), Current Data Rate (Receive) MUST be set equal to the Maximum Data Rate (Receive). The Current Data Rate (Receive) MUST NOT exceed the Maximum Data Rate (Receive).

13.15. Current Data Rate (Transmit)

The Current Data Rate (Transmit) (CDRT) Data Item is used to indicate the rate at which the link is currently operating for transmitting traffic.

When used in the Link Characteristics Request Message (Section 12.18), Current Data Rate (Transmit) represents the desired transmit rate, in bits per second, on the link.

The Current Data Rate (Transmit) Data Item contains the following fields:



Data Item Type: 15

Length: 8

Current Data Rate (Transmit): A 64-bit unsigned integer, representing the current data rate, in bits per second, that can currently be achieved while transmitting traffic on the link.

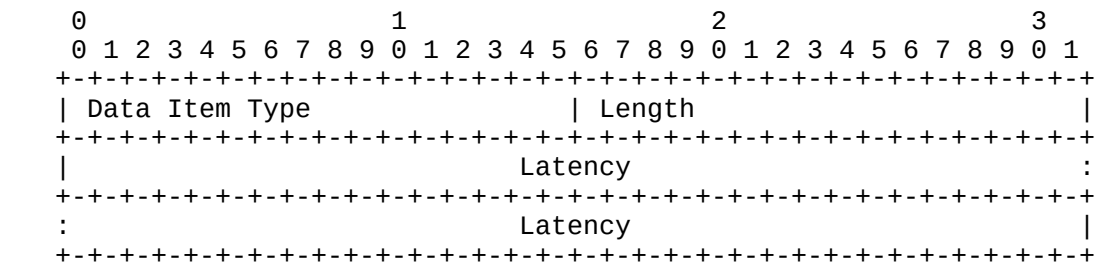
If there is no distinction between Current Data Rate (Transmit) and Maximum Data Rate (Transmit) (Section 13.13), Current Data Rate (Transmit) MUST be set equal to the Maximum Data Rate (Transmit). The Current Data Rate (Transmit) MUST NOT exceed the Maximum Data Rate (Transmit).

13.16. Latency

The Latency Data Item is used to indicate the amount of latency, in microseconds, on the link.

The Latency value is reported as transmission delay. The calculation of latency is implementation dependent. For example, the latency may be a running average calculated from the internal queuing.

The Latency Data Item contains the following fields:



Data Item Type: 16

Length: 8

Latency: A 64-bit unsigned integer, representing the transmission delay, in microseconds, that a packet encounters as it is transmitted over the link.

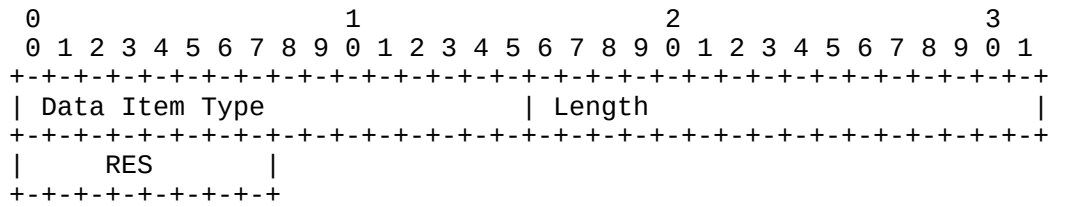
13.17. Resources

The Resources (RES) Data Item is used to indicate the amount of finite resources available for data transmission and reception at the destination as a percentage, with 0 meaning 'no resources remaining', and 100 meaning 'a full supply', assuming that when Resources reaches 0 data transmission and/or reception will cease.

An example of such resources might be battery life, but could equally be magic beans. The list of resources that might be considered is beyond the scope of this document, and is left to implementations to decide.

This Data Item is designed to be used as an indication of some capability of the modem and/or router at the destination.

The Resources Data Item contains the following fields:



Data Item Type: 17

Length: 1

Resources: An 8-bit unsigned integer percentage, 0-100, representing the amount of resources available. Any value greater than 100 MUST be considered as invalid.

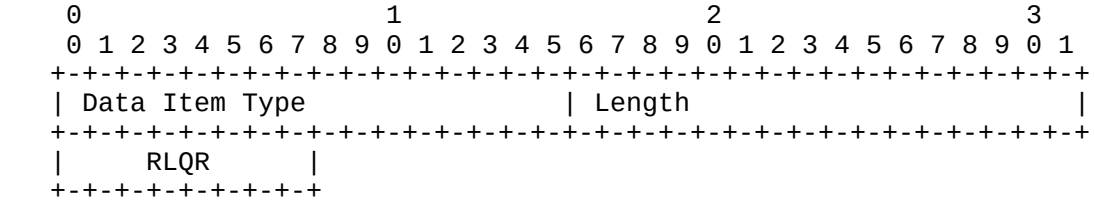
If a device cannot calculate Resources, this Data Item MUST NOT be issued.

13.18. Relative Link Quality (Receive)

The Relative Link Quality (Receive) (RLQR) Data Item is used to indicate the quality of the link to a destination for receiving traffic, with 0 meaning 'worst quality', and 100 meaning 'best quality'.

Quality in this context is defined as an indication of the stability of a link for reception; a destination with high Relative Link Quality (Receive) is expected to have generally stable DLEP metrics, and the metrics of a destination with low Relative Link Quality (Receive) can be expected to rapidly fluctuate over a wide range.

The Relative Link Quality (Receive) Data Item contains the following fields:



Data Item Type: 18

Length: 1

Relative Link Quality (Receive): A non-dimensional unsigned 8-bit integer, 0-100, representing relative quality of the link for receiving traffic. Any value greater than 100 MUST be considered as invalid.

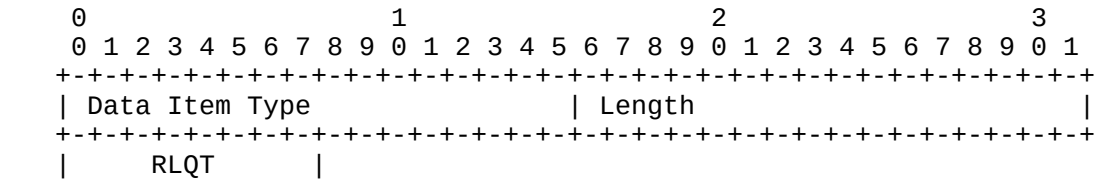
If a device cannot calculate the Relative Link Quality (Receive), this Data Item MUST NOT be issued.

13.19. Relative Link Quality (Transmit)

The Relative Link Quality (Transmit) (RLQT) Data Item is used to indicate the quality of the link to a destination for transmitting traffic, with 0 meaning 'worst quality', and 100 meaning 'best quality'.

Quality in this context is defined as an indication of the stability of a link for transmission; a destination with high Relative Link Quality (Transmit) is expected to have generally stable DLEP metrics, and the metrics of a destination with low Relative Link Quality (Transmit) can be expected to rapidly fluctuate over a wide range.

The Relative Link Quality (Transmit) Data Item contains the following fields:



Data Item Type: 19

Length: 1

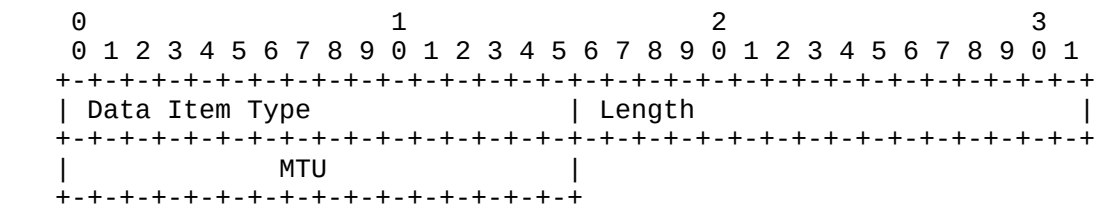
Relative Link Quality (Transmit): A non-dimensional unsigned 8-bit integer, 0-100, representing relative quality of the link for transmitting traffic. Any value greater than 100 MUST be considered as invalid.

If a device cannot calculate the Relative Link Quality (Transmit), this Data Item MUST NOT be issued.

13.20. Maximum Transmission Unit (MTU)

The Maximum Transmission Unit (MTU) Data Item is used to indicate the maximum size, in octets, of an IP packet that can be transmitted without fragmentation, including headers, but excluding any lower layer headers.

The Maximum Transmission Unit Data Item contains the following fields:



Data Item Type: 20

Length: 2

Maximum Transmission Unit: The maximum size, in octets, of an IP packet that can be transmitted without fragmentation, including headers, but excluding any lower layer headers.

If a device cannot calculate the Maximum Transmission Unit, this Data Item MUST NOT be issued.

14. Security Considerations

The potential security concerns when using DLEP are:

1. An attacker might pretend to be a DLEP participant, either at DLEP session initialization, or by injection of DLEP Messages once a session has been established.

2. DLEP Data Items could be altered by an attacker, causing the receiving implementation to inappropriately alter its information base concerning network status.

3. An attacker could join an unsecured radio network and inject over-the-air signals that maliciously influence the information reported by a DLEP modem, causing a router to forward traffic to an inappropriate destination.

The implications of attacks on DLEP peers are directly proportional to the extent to which DLEP data is used within the control plane. While the use of DLEP data in other control plane components is out of scope for this document, as an example, if DLEP statistics are incorporated into route cost calculations, adversaries masquerading as a DLEP peer, and injecting malicious data via DLEP, could cause suboptimal route selection, adversely impacting network performance. Similar issues can arise if DLEP data is used as an input to policing algorithms - injection of malicious data via DLEP can cause those policing algorithms to make incorrect decisions, degrading network throughput.

For these reasons, security of the DLEP transport must be considered at both the transport layer, and at Layer 2.

At the transport layer, when TLS is in use, each peer SHOULD check the validity of credentials presented by the other peer during TLS session establishment. Implementations following the "dedicated deployments" model attempting to use TLS MAY need to consider use of pre-shared keys for credentials, and provide specialized techniques for peer identity validation, and MAY refer to [RFC5487] for additional details. Implementations following the "networked deployment" model described in Implementation Scenarios SHOULD refer to [RFC7525] for additional details.

At layer 2 - since DLEP is restricted to operation over a single (possibly logical) hop, implementations SHOULD also secure the Layer

2 link. Examples of technologies that can be deployed to secure the Layer 2 link include [IEEE-802.1AE] and [IEEE-802.1X].

By examining the Secured Medium flag in the Peer Type Data Item (Section 13.4), a router can decide if it is able to trust the information supplied via a DLEP modem. If this is not the case, then the router SHOULD consider restricting the size of attached subnets, announced in IPv4 Attached Subnet Data Items (Section 13.10) and/or IPv6 Attached Subnet Data Items (Section 13.11), that are considered for route selection.

To avoid potential denial of service attack, it is RECOMMENDED that implementations using the Peer Discovery mechanism maintain an information base of hosts that persistently fail Session Initialization having provided an acceptable Peer Discovery Signal, and ignore subsequent Peer Discovery Signals from such hosts.

This specification does not address security of the data plane, as it (the data plane) is not affected, and standard security procedures can be employed.

15. IANA Considerations

15.1. Registrations

Upon approval of this document, IANA is requested to create a new protocol registry for Dynamic Link Exchange Protocol (DLEP). The remainder of this section requests the creation of new DLEP specific registries.

15.2. Signal Type Registration

Upon approval of this document, IANA is requested to create a new DLEP registry, named "Signal Type Values".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Type Code	Description/Policy
0	Reserved
1	Peer Discovery Signal
2	Peer Offer Signal
3-65519	Specification Required
65520-65534	Private Use
65535	Reserved

15.3. Message Type Registration

Upon approval of this document, IANA is requested to create a new DLEP registry, named "Message Type Values".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Type Code	Description/Policy
0	Reserved
1	Session Initialization Message
2	Session Initialization Response Message
3	Session Update Message
4	Session Update Response Message
5	Session Termination Message
6	Session Termination Response Message
7	Destination Up Message
8	Destination Up Response Message
9	Destination Announce Message
10	Destination Announce Response Message
11	Destination Down Message
12	Destination Down Response Message
13	Destination Update Message
14	Link Characteristics Request Message
15	Link Characteristics Response Message
16	Heartbeat Message
17-65519	Specification Required
65520-65534	Private Use
65535	Reserved

15.4. DLEP Data Item Registrations

Upon approval of this document, IANA is requested to create a new DLEP registry, named "Data Item Type Values".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Type Code	Description/Policy
0	Reserved
1	Status
2	IPv4 Connection Point
3	IPv6 Connection Point
4	Peer Type
5	Heartbeat Interval
6	Extensions Supported
7	MAC Address
8	IPv4 Address
9	IPv6 Address
10	IPv4 Attached Subnet
11	IPv6 Attached Subnet
12	Maximum Data Rate (Receive) (MDRR)
13	Maximum Data Rate (Transmit) (MDRT)
14	Current Data Rate (Receive) (CDRR)
15	Current Data Rate (Transmit) (CDRT)
16	Latency
17	Resources (RES)
18	Relative Link Quality (Receive) (RLQR)
19	Relative Link Quality (Transmit) (RLQT)
20	Maximum Transmission Unit (MTU)
21-65407	Specification Required
65408-65534	Private Use
65535	Reserved

15.5. DLEP Status Code Registrations

Upon approval of this document, IANA is requested to create a new DLEP registry, named "Status Code Values".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Status Code	Failure Mode	Description/Policy
0	Continue	Success
1	Continue	Not Interested
2	Continue	Request Denied
3	Continue	Inconsistent Data
4-111	Continue	Specification Required
112-127	Continue	Private Use
128	Terminate	Unknown Message
129	Terminate	Unexpected Message
130	Terminate	Invalid Data
131	Terminate	Invalid Destination
132	Terminate	Timed Out
133-239	Terminate	Specification Required
240-254	Terminate	Private Use
255	Terminate	Reserved

15.6. DLEP Extensions Registrations

Upon approval of this document, IANA is requested to create a new DLEP registry, named "Extension Type Values".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Code	Description/Policy
0	Reserved
1-65519	Specification Required
65520-65534	Private Use
65535	Reserved

Table 3: DLEP Extension types

15.7. DLEP IPv4 Connection Point Flags

Upon approval of this document, IANA is requested to create a new DLEP registry, named "IPv4 Connection Point Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Use TLS [RFC5246] indicator

15.8. DLEP IPv6 Connection Point Flags

Upon approval of this document, IANA is requested to create a new DLEP registry, named "IPv6 Connection Point Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Use TLS [RFC5246] indicator

15.9. DLEP Peer Type Flag

Upon approval of this document, IANA is requested to create a new DLEP registry, named "Peer Type Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Secured Medium indicator

15.10. DLEP IPv4 Address Flag

Upon approval of this document, IANA is requested to create a new DLEP registry, named "IPv4 Address Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Add/Drop indicator

15.11. DLEP IPv6 Address Flag

Upon approval of this document, IANA is requested to create a new DLEP registry, named "IPv6 Address Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Add/Drop indicator

15.12. DLEP IPv4 Attached Subnet Flag

Upon approval of this document, IANA is requested to create a new DLEP registry, named "IPv4 Attached Subnet Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Add/Drop indicator

15.13. DLEP IPv6 Attached Subnet Flag

Upon approval of this document, IANA is requested to create a new DLEP registry, named "IPv6 Attached Subnet Flags".

The following table provides initial registry values and the [RFC5226] defined policies that should apply to the registry:

Bit	Description/Policy
0-6	Unassigned/Specification Required
7	Add/Drop indicator

15.14. DLEP Well-known Port

Upon approval of this document, IANA is requested to assign a single value in the "Service Name and Transport Protocol Port Number Registry" found at https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml for use by "DLEP", as defined in this document. This assignment should be valid for TCP and UDP.

15.15. DLEP IPv4 Link-local Multicast Address

Upon approval of this document, IANA is requested to assign an IPv4 multicast address registry found at http://www.iana.org/assignments/multicast-addresses for use as the "IPv4 DLEP Discovery Address".

15.16. DLEP IPv6 Link-local Multicast Address

Upon approval of this document, IANA is requested to assign an IPv6 multicast address registry found at http://www.iana.org/assignments/multicast-addresses for use as the "IPv6 DLEP Discovery Address".

16. Acknowledgments

We would like to acknowledge and thank the members of the DLEP design team, who have provided invaluable insight. The members of the design team are: Teco Boot, Bow-Nan Cheng, John Dowdell, and Henning Rogge.

We would also like to acknowledge the influence and contributions of Greg Harrison, Chris Olsen, Martin Duke, Subir Das, Jaewon Kang, Vikram Kaul, Nelson Powell, Lou Berger, and Victoria Pritchard.

17. References

17.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, DOI 10.17487/RFC3629, November 2003, <http://www.rfc-editor.org/info/rfc3629>.

[RFC5082] Gill, V., Heasley, J., Meyer, D., Savola, P., Ed., and C. Pignataro, "The Generalized TTL Security Mechanism (GTSM)", RFC 5082, DOI 10.17487/RFC5082, October 2007, <http://www.rfc-editor.org/info/rfc5082>.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <http://www.rfc-editor.org/info/rfc5246>.

17.2. Informative References

[IEEE-802.1AE] "IEEE Standards for Local and Metropolitan Area Networks: Media Access Control (MAC) Security", DOI 10.1109/IEEESTD.2006.245590, August 2006.

[IEEE-802.1X] "IEEE Standards for Local and Metropolitan Area Networks: Port based Network Access Control", DOI 10.1109/IEEESTD.2010.5409813, February 2010.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

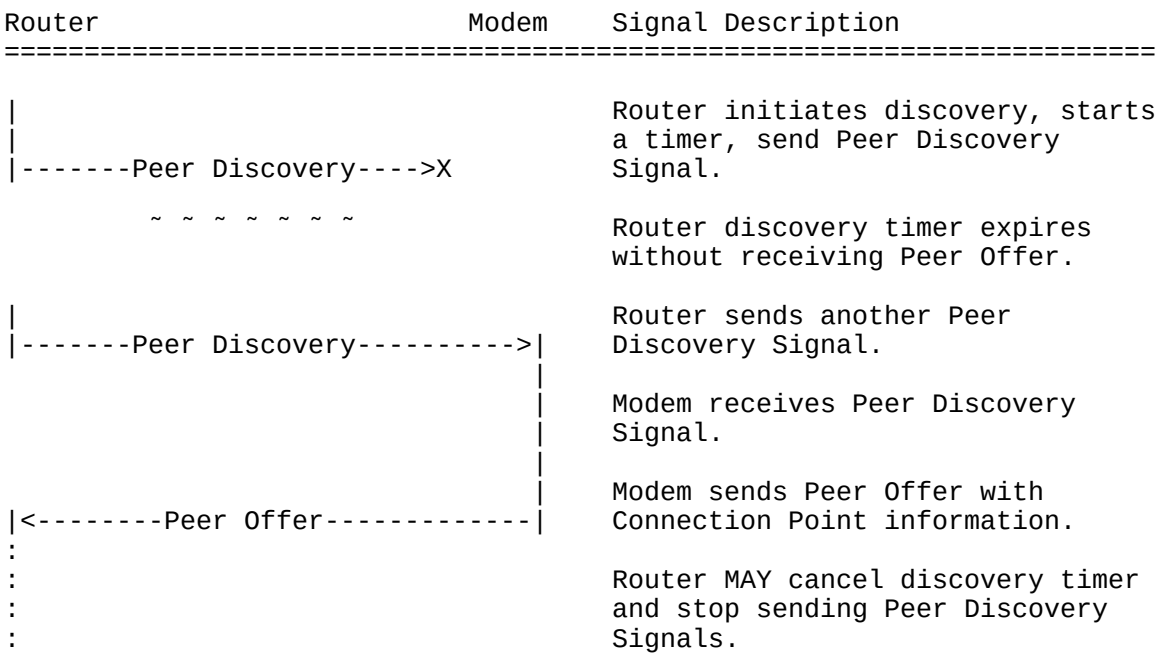
[RFC5487] Badra, M., "Pre-Shared Key Cipher Suites for TLS with SHA-256/384 and AES Galois Counter Mode", RFC 5487, DOI 10.17487/RFC5487, March 2009, <http://www.rfc-editor.org/info/rfc5487>.

[RFC5578] Berry, B., Ed., Ratliff, S., Paradise, E., Kaiser, T., and M. Adams, "PPP over Ethernet (PPPoE) Extensions for Credit Flow and Link Metrics", RFC 5578, DOI 10.17487/RFC5578, February 2010, <http://www.rfc-editor.org/info/rfc5578>.

[RFC6890] Cotton, M., Vegoda, L., Bonica, R., Ed., and B. Haberman, "Special-Purpose IP Address Registries", BCP 153, RFC 6890, DOI 10.17487/RFC6890, April 2013, <http://www.rfc-editor.org/info/rfc6890>.

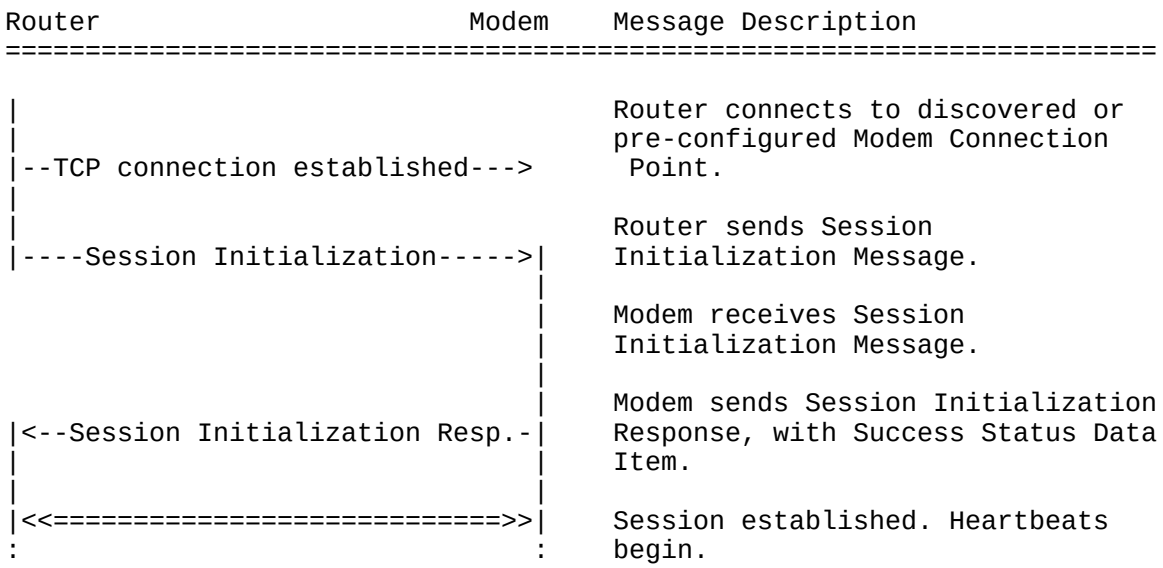
[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <http://www.rfc-editor.org/info/rfc7525>.

Appendix A. Discovery Signal Flows

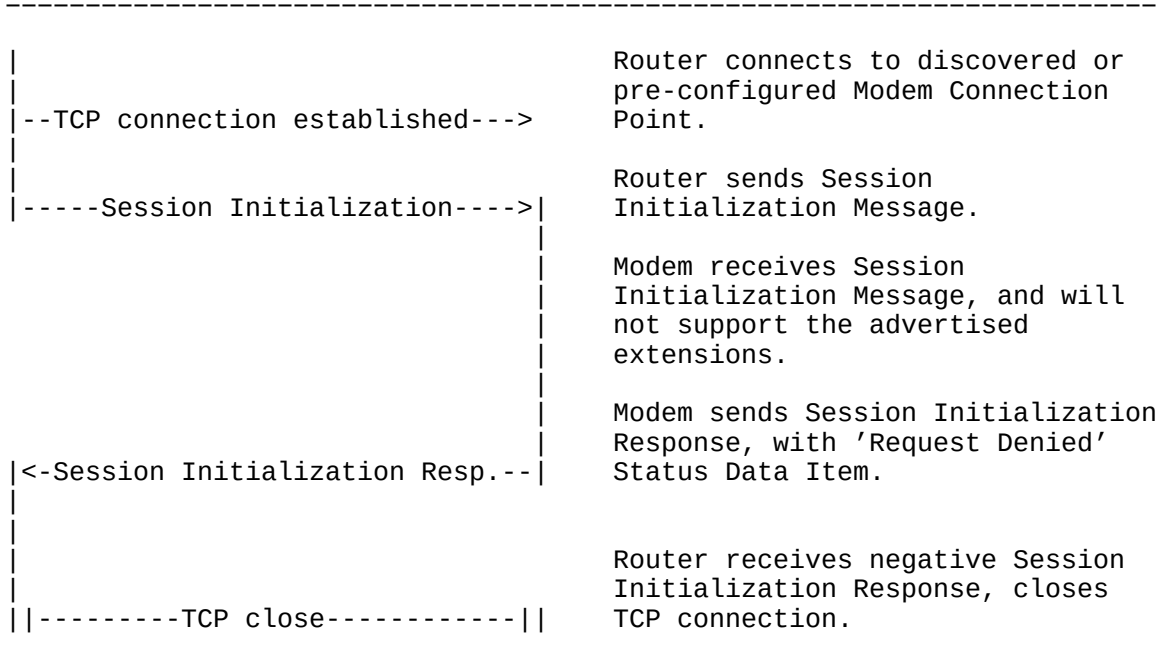


Appendix B. Peer Level Message Flows

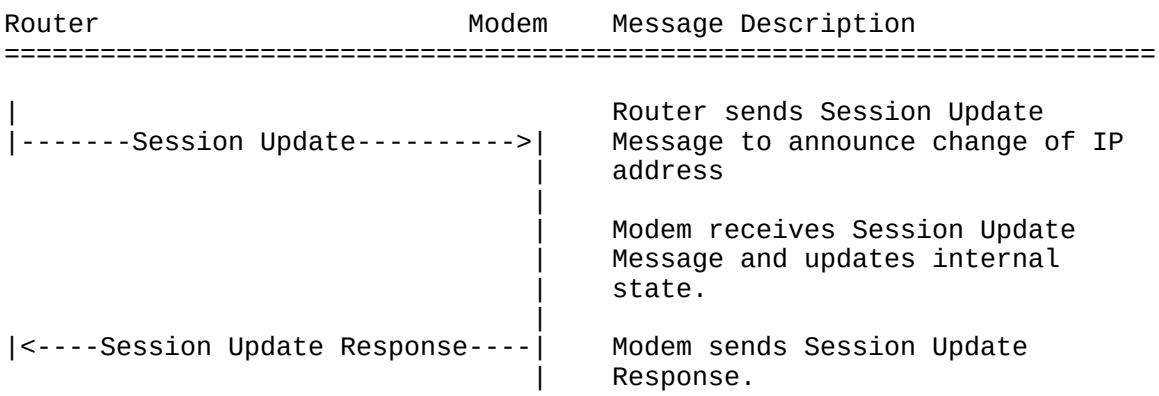
B.1. Session Initialization



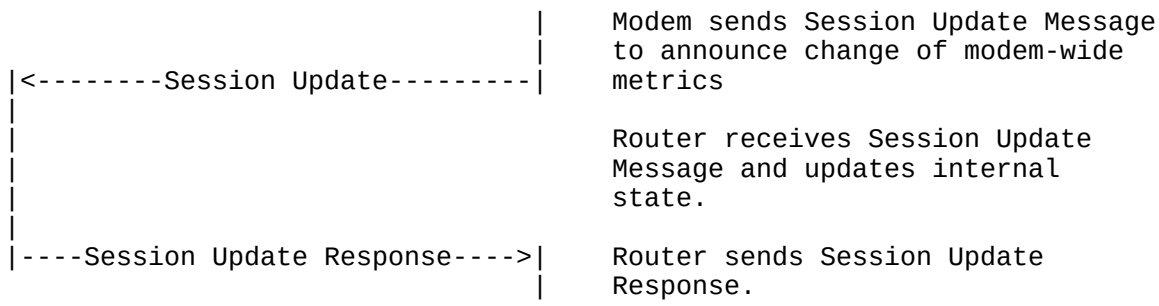
B.2. Session Initialization - Refused



B.3. Router Changes IP Addresses



B.4. Modem Changes Session-wide Metrics



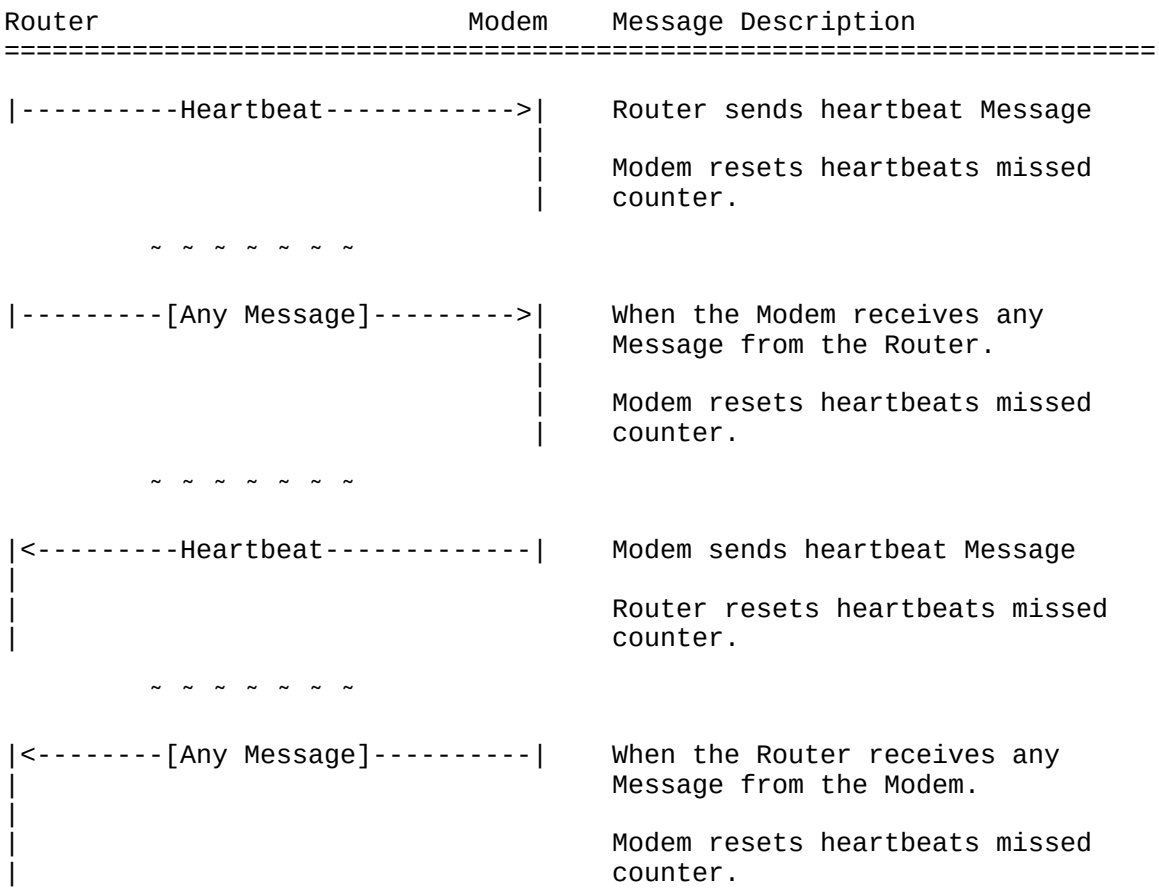
B.5. Router Terminates Session

```
Router                          Modem    Message Description
=============================================================================

|                                        Router sends Session Termination
|------Session Termination------>|       Message with Status Data Item.
|                                |
|-------TCP shutdown (send)--->  |       Router stops sending Messages.
                                 |
                                 |       Modem receives Session
                                 |       Termination, stops counting
                                 |       received heartbeats and stops
                                 |       sending heartbeats.
                                 |
                                 |       Modem sends Session Termination
|<---Session Termination Resp.---|       Response with Status 'Success'.
|
|                                        Modem stops sending Messages.
|
||---------TCP close------------||       Session terminated.
```

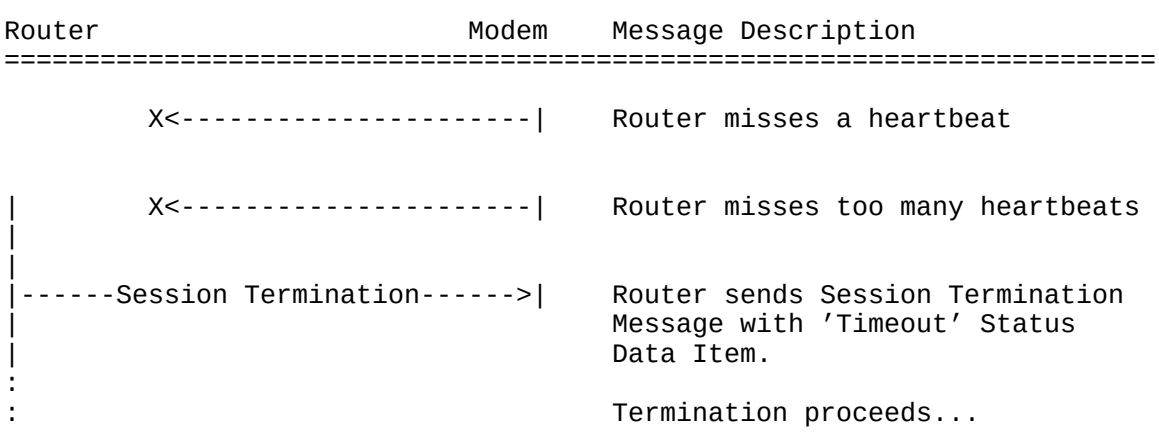
B.6. Modem Terminates Session

```
Router                          Modem    Message Description
=============================================================================

                                 |       Modem sends Session Termination
|<----Session Termination--------|       Message with Status Data Item.
|
|                                        Modem stops sending Messages.
|
|                                        Router receives Session
|                                        Termination, stops counting
|                                        received heartbeats and stops
|                                        sending heartbeats.
|
|                                        Router sends Session Termination
|---Session Termination Resp.--->|       Response with Status 'Success'.
                                 |
                                 |       Router stops sending Messages.
                                 |
||---------TCP close------------||       Session terminated.
```

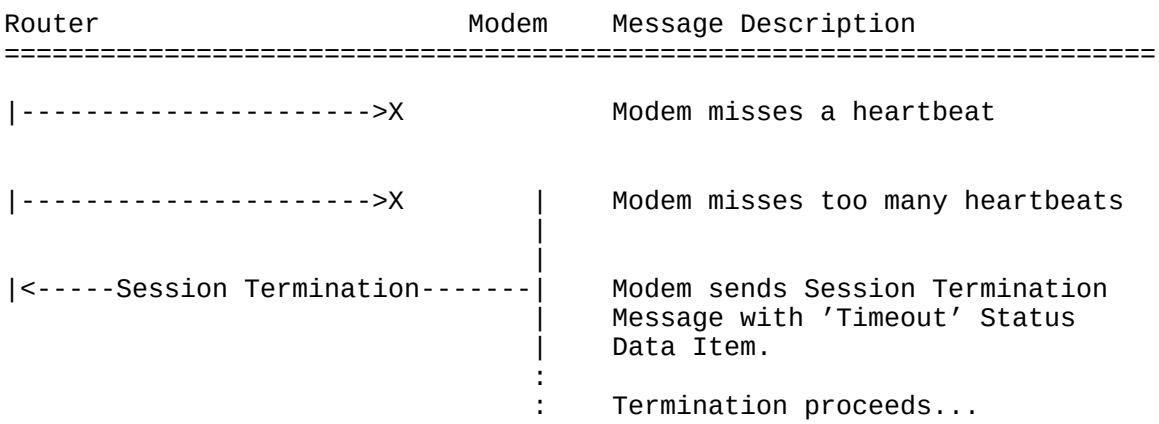
B.7. Session Heartbeats



B.8. Router Detects a Heartbeat timeout

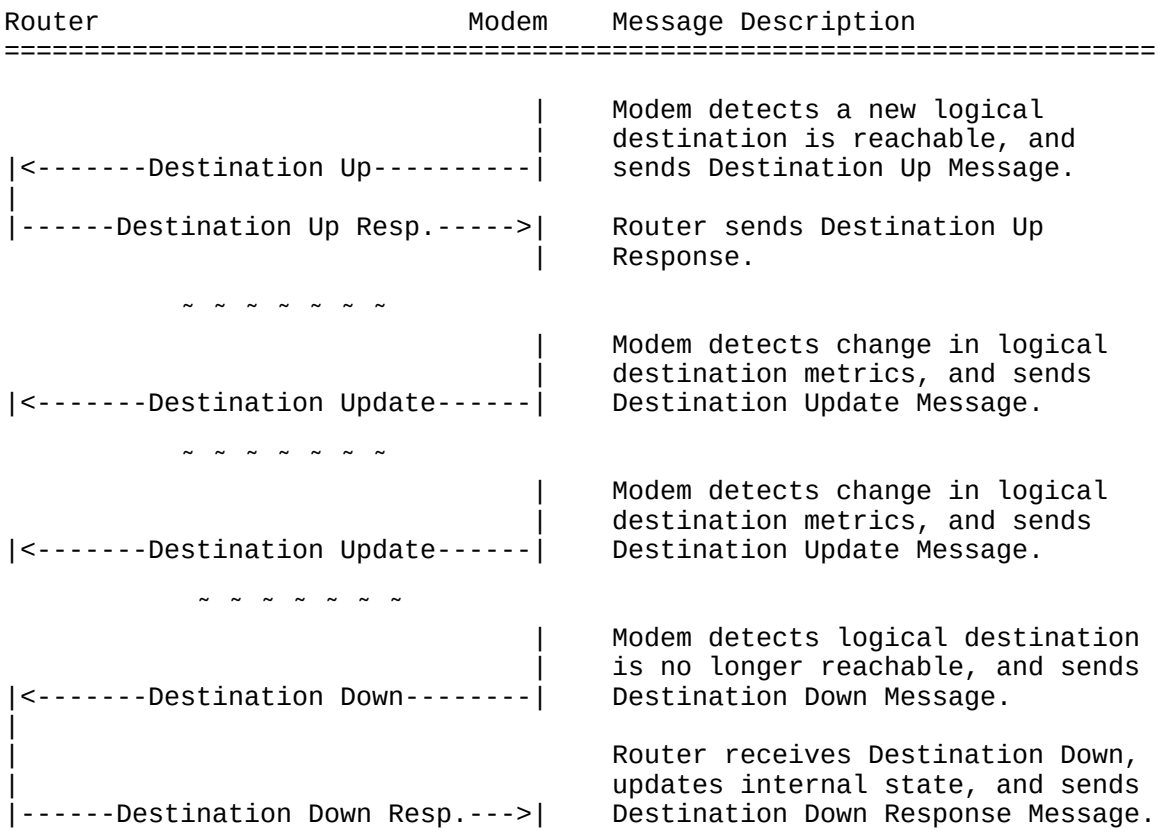


B.9. Modem Detects a Heartbeat timeout

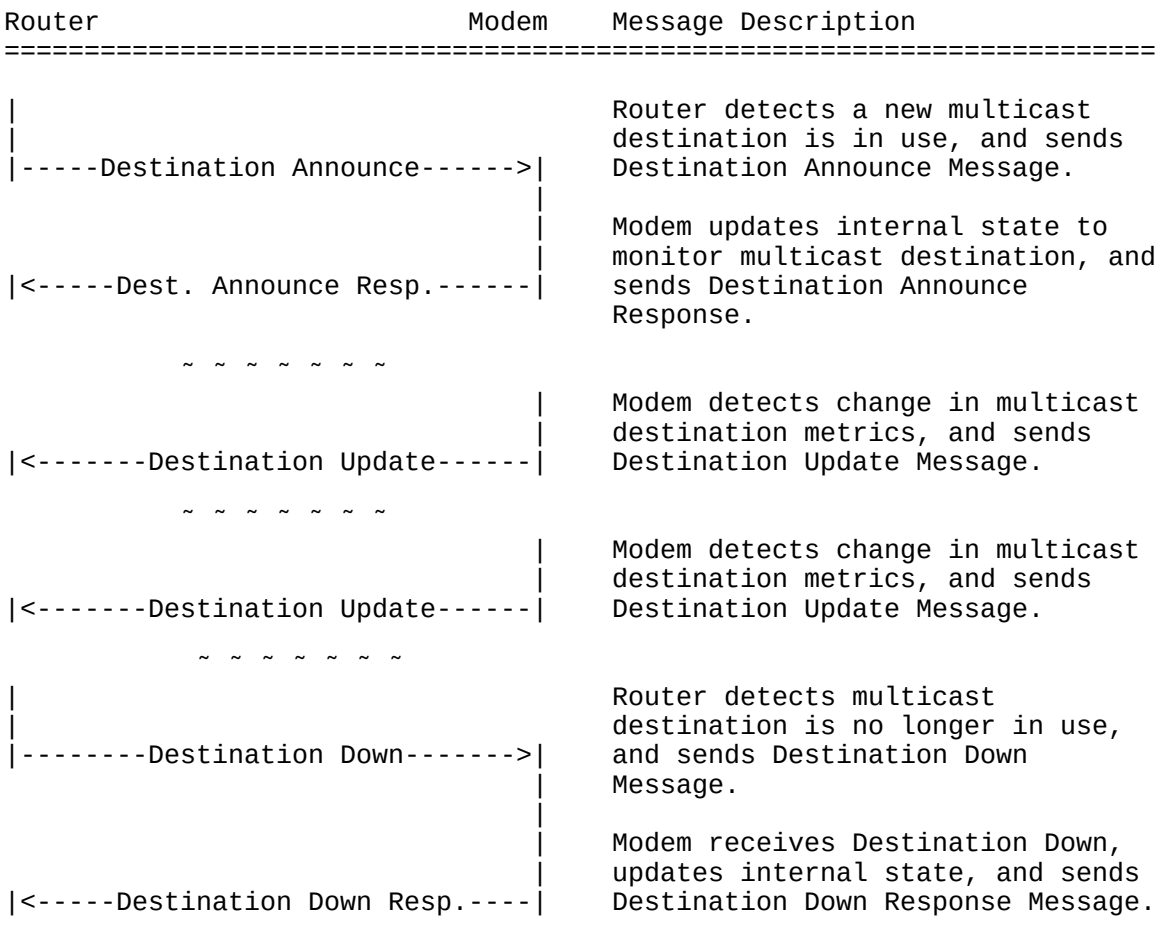


Appendix C. Destination Specific Message Flows

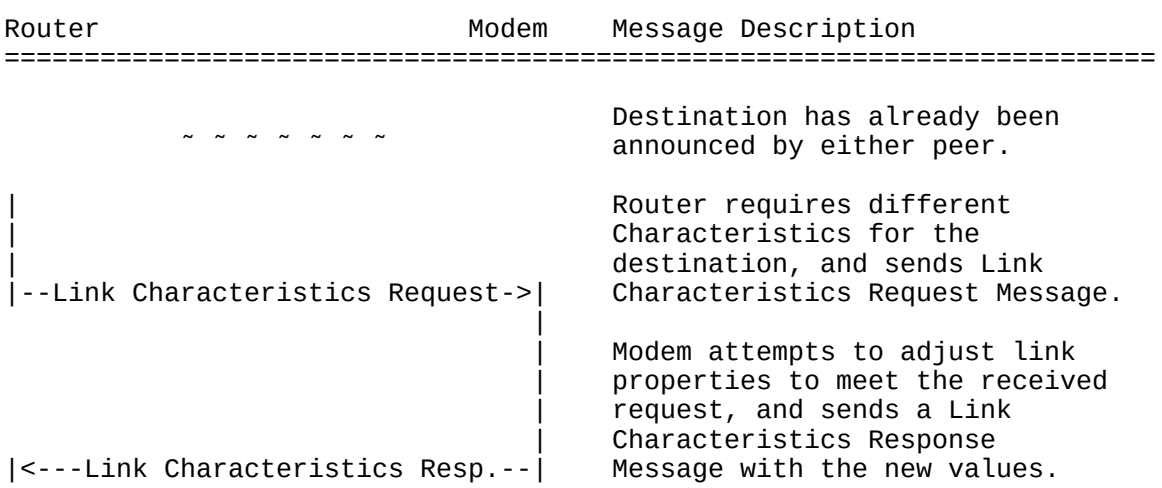
C.1. Common Destination Notification



C.2. Multicast Destination Notification



C.3. Link Characteristics Request



Authors' Addresses

Stan Ratliff
VT iDirect
13861 Sunrise Valley Drive, Suite 300
Herndon, VA 20171
USA

Email: sratliff@idirect.net

Shawn Jury
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
USA

Email: sjury@cisco.com

Darryl Satterwhite
Broadcom

Email: dsatterw@broadcom.com

Rick Taylor
Airbus Defence & Space
Quadrant House
Celtic Springs
Coedkernew
Newport NP10 8FZ
UK

Email: rick.taylor@airbus.com

Bo Berry

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: November 25, 2017

J. Yi
Ecole Polytechnique
B. Parrein
University of Nantes
May 24, 2017

Multipath Extension for the Optimized Link State Routing Protocol version 2 (OLSRv2) draft-ietf-manet-olsrv2-multipath-15

Abstract

This document specifies a multipath extension for the Optimized Link State Routing Protocol version 2 (OLSRv2) to discover multiple disjoint paths for Mobile Ad Hoc Networks (MANETs). Considering the characteristics of MANETs, especially the dynamic network topology, using multiple paths can increase aggregated throughput and improve the reliability by avoiding single route failures. The interoperability with OLSRv2 is retained.

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 http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 25, 2017.

Copyright Notice

Table of Contents

1. Introduction

The Optimized Link State Routing Protocol version 2 (OLSRv2) [RFC7181] is a proactive link state protocol designed for use in mobile ad hoc networks (MANETs). It generates routing messages periodically to create and maintain a Routing Set, which contains routing information to all the possible destinations in the routing domain. For each destination, there exists a unique Routing Tuple, which indicates the next hop to reach the destination.

This document specifies an extension of the OLSRv2 protocol [RFC7181], to provide multiple disjoint paths when appropriate for a source-destination pair. Because of the characteristics of MANETs [RFC2501], especially the dynamic topology, having multiple paths is helpful for increasing network throughput, improving forwarding reliability, and load balancing.

Multipath OLSRv2 (MP-OLSRv2) specified in this document uses the Multipath Dijkstra algorithm by default to explore multiple disjoint paths from a source router to a destination router based on the topology information obtained through OLSRv2, and to forward the datagrams in a load-balancing manner using source routing. MP-OLSRv2 is designed to be interoperable with OLSRv2.

1.1. Motivation and Experiments to Be Conducted

This document is an experimental extension of OLSRv2 that can increase the data forwarding reliability in dynamic and high-load MANET scenarios by transmitting datagrams over multiple disjoint paths using source routing. This mechanism is used because:

- Disjoint paths can avoid single route failures.
- Transmitting datagrams through parallel paths can increase aggregated throughput.
- Some scenarios may require some routers must (or must not) be used.
- Having control of the paths at the source benefits the load balancing and traffic engineering.
- An application of this extension is in combination with Forward Error Correction (FEC) coding applied across packets (erasure coding) [WPMC11]. Because the packet drops are normally bursty in a path (for example, due to route failure), erasure coding is less effective in single path routing protocols. By providing multiple disjoint paths, the application of erasure coding with multipath

protocol is more resilient to routing failures.

While in existing deployments, running code and simulations have proven the interest of multipath extension for OLSRv2 in certain networks, more experiments and experiences are still needed to understand the effects of the protocol specified in this experimental document. The multipath extension for OLSRv2 is expected to be revised documented as a Standard Track document once sufficient operational experience is obtained. Other than general experiences, including the protocol specification and interoperability with base OLSRv2 implementations, experiences in the following aspects are highly appreciated:

- Optimal values for the number of multiple paths (NUMBER_OF_PATHS, Section 5) to be used. This depends on the network topology and router density.

- Optimal values used in the metric functions. Metric functions are applied to increase the metric of used links and nodes so as to obtain disjoint paths. What kind of disjointness is desired (node-disjoint or link-disjoint) may depend on the layer 2 protocol used, and can be achieved by applying different sets of metric functions.

- Use of different metric types. This multipath extension can be used with metric types that meet the requirement of OLSRv2, such as [RFC7779]. The metric type used has also impact to the choice of metric functions as indicated in the previous bullet point.

- The impact of partial topology information to multipath calculation. OLSRv2 maintains a partial topology information base to reduce protocol overhead. Experience has shown that multiple paths can be obtained even with such partial information, however, depending on the Multi-Point Relay (MPR) selection algorithm used, the disjointness of the multiple paths might be impacted depending on the Multi-Point Relay (MPR) selection algorithm used.

- Use of IPv6 loose source routing. In the current specification, only strict source routing is used for IPv6 based on [RFC6554]. In [I-D.ietf-6man-segment-routing-header], the use of the loose source routing is also proposed in IPv6. In scenarios where the length of the source routing header is critical, the loose source routing can be considered.

- Optimal choice of "key" routers for loose source routing. In some cases, loose source routing is used to reduce overhead or for interoperability with OLSRv2 routers. Other than the basic rules defined in the following parts of this document, optimal choices

of routers to put in the loose source routing header can be further studied.

- Different path-selection schedulers. Depending on the application type and transport layer type, either per-flow scheduler or per-datagram scheduler is applied. By default, the traffic load should be equally distributed in multiple paths. In some scenarios, weighted scheduling can be considered: for example, the paths with lower metrics (i.e., higher quality) can transfer more datagrams or flows compared to paths with higher metrics.

- The impacts of the delay variation due to multipath routing. [RFC2991] brings out some concerns of multipath routing, especially variable latencies when per-datagram scheduling is applied. Although current experiment results show that multipath routing can reduce the jitter in dynamic scenarios, some transport protocols or applications may be sensitive to the datagram re-ordering.

- The disjoint multipath protocol has interesting application with erasure coding, especially for services like video/audio streaming [WPMC11]. The combination of erasure coding mechanisms and this extension is thus encouraged.

- Different algorithms to obtain multiple paths, other than the default Multipath Dijkstra algorithm introduced in Section 8.5.2 of this specification.

- The use of multi-topology information. By using [RFC7722], multiple topologies using different metric types can be obtained. Although there is no work defining how this extension can make use of the multi-topology information base yet, it is encouraged to experiment with the use of multiple metrics for building multiple paths.

Comments are solicited and should be addressed to the MANET working group’s mailing list at manet@ietf.org and/or the authors."

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].

This document uses the terminology and notation defined in [RFC5444], [RFC6130], [RFC7181]. Additionally, it defines the following

terminology:

OLSRv2 Routing Process - A routing process based on [RFC7181], without multipath extension specified in this document.

MP-OLSRv2 Routing Process - A multipath routing process based on this specification as an extension to [RFC7181].

SR-OLSRv2 Routing Process - A OLSRv2 Routing Process that supports source routing (SR), or an MP-OLSRv2 Routing Process.

3. Applicability Statement

As an extension of OLSRv2, this specification is applicable to MANETs for which OLSRv2 is applicable (see [RFC7181]). It can operate on single or multiple interfaces to discover multiple disjoint paths from a source router to a destination router. MP-OLSRv2 is designed for networks with dynamic topology to avoid single route failure. It can also provide higher aggregated throughput and load balancing.

In a router supporting MP-OLSRv2, MP-OLSRv2 does not necessarily replace OLSRv2 completely. The extension can be applied for certain applications that are suitable for multipath routing (mainly video or audio streams), based on information such as a DiffServ codepoint [RFC2474].

Compared to OLSRv2, this extension does not introduce any new message type. A new Message TLV Type is introduced to identify the routers that support forwarding based on source routing header. It is interoperable with OLSRv2 implementations that do not have this extension: as the MP-OLSRv2 uses source routing, in IPv4 networks the interoperability is achieved using loose source routing headers; in IPv6 networks, it is achieved by eliminating routers that do not support IPv6 strict source routing.

MP-OLSRv2 supports two different, but interoperable multipath calculation approaches: proactive and reactive. In the proactive calculation, the paths to all the destinations are calculated before needed. In the reactive calculation, only the paths to desired destination(s) are calculated on demand. The proactive approach requires more computational resources than the reactive one. The reactive approach requires the IP forwarding plane to trigger the multipath calculation.

MP-OLSRv2 forwards datagrams using the source routing header. As there are multiple paths to each destination, MP-OLSRv2 requires the IP forwarding plane to be able to choose which source route to be put

in the source routing header based on the path scheduler defined by MP-OLSRv2. For IPv4 networks, implementation of loose source routing is required following [RFC0791]. For IPv6 networks, implementation of strict source routing is required following the source routing header generation and processing defined in [RFC6554].

4. Protocol Overview and Functioning

This specification uses OLSRv2 [RFC7181] to:

- Identify all the reachable routers in the network.
- Identify a sufficient subset of links in the networks, so that routes can be calculated to all reachable destinations.
- Provide a Routing Set containing the shortest routes from this router to all destinations.

In addition, the MP-OLSRv2 Routing Process identifies the routers that support source routing by adding a new Message TLV in HELLO and Topology Control (TC) messages. Based on the above information acquired, every MP-OLSRv2 Routing Process is aware of a reduced topology map of the network and the routers supporting source routing.

A Multipath Routing Set containing the multipath information is maintained. It may either be proactively calculated or reactively calculated:

- In the proactive approach, multiple paths to all possible destinations are calculated and updated based on control message exchange. The routes are thus available before they are actually needed.
- In the reactive approach, a multipath algorithm is invoked on demand, i.e., only when there is a datagram to be sent from the source to the destination, and there is no available Routing Tuple in the Multipath Routing Set. This requires the IP forwarding information base to trigger the multipath calculation specified in Section 8.5 when no Multipath Routing Tuple is available. The reactive operation is local to the router and no additional routing control messages exchange is required. When the paths are being calculated, the datagrams SHOULD be buffered unless the router does not have enough memory.

Routers in the same network may choose either proactive or reactive multipath calculation independently according to their computation

resources. The Multipath Dijkstra algorithm (defined in Section 8.5) is introduced as the default algorithm to generate multiple disjoint paths from a source to a destination, and such information is kept in the Multipath Routing Set.

The datagram is forwarded based on source routing. When there is a datagram to be sent to a destination, the source router acquires a path from the Multipath Routing Set. The path information is stored in the datagram header using the source routing header.

5. Parameters and Constants

In addition to the parameters and constants defined in [RFC7181], this specification uses the parameters and constants described in this section.

5.1. Router Parameters

NUMBER_OF_PATHS The number of paths desired by the router.

MAX_SRC_HOPS The maximum number of hops allowed to be put in the source routing header. A value set zero means there is no limitation on the maximum number of hops. In an IPv6 network, it MUST be set to 0 because [RFC6554] supports only strict source routing. All the intermediate routers MUST be included in the source routing header, which a various number of hops. In an IPv4 network, it MUST be strictly less than 11 and greater than 0 due to the length limit of the IPv4 header.

CUTOFF_RATIO The ratio that defines the maximum metric of a path compared to the shortest path kept in the OLSRv2 Routing Set. For example, the metric to a destination is R_metric based on the Routing Set. Then the maximum metric allowed for a path is CUTOFF_RATIO * R_metric. CUTOFF_RATIO MUST be greater than or equal to 1. Setting the number low makes it less likely that additional paths will be found -- for example, setting it to 1 will only consider equal length paths.

SR_TC_INTERVAL The maximum time between the transmission of two successive TC messages by an MP-OLSRv2 Routing Process.

SR_HOLD_TIME The minimum value in the TLV with Type = VALIDITY_TIME included in TC messages generated based on SR_TC_INTERVAL.

6. Packets and Messages

This extension employs the routing control messages HELLO and TC (Topology Control) as defined in OLSRv2 [RFC7181] to obtain network topology information. For the datagram to support source routing, a source routing header is added to each datagram routed by this extension. Depending on the IP version used, the source routing header is defined in this section.

6.1. HELLO and TC messages

HELLO and TC messages used by the MP-OLSRv2 Routing Process use the same format as defined in [RFC7181]. In addition, a new Message TLV type is defined, to identify the originator of the HELLO or TC message that supports source route forwarding. The new Message TLV type is introduced for enabling MP-OLSRv2 as an extension of OLSRv2: only the routers supporting source-route forwarding can be used in the source routing header of a datagram, because adding a router that does not understand the source routing header will cause routing failure.

6.1.1. SOURCE_ROUTE TLV

SOURCE_ROUTE TLV is a Message TLV signaling that the message is generated by a router that supports source-route forwarding. It can be an MP-OLSRv2 Routing Process, or an OLSRv2 Routing Process that supports source-route forwarding.

Every HELLO or TC message generated by a MP-OLSRv2 Routing Process MUST have exactly one SOURCE_ROUTE TLV without value.

Every HELLO or TC message generated by an OLSRv2 Routing Process MUST have exactly one SOURCE_ROUTE TLV, if the OLSRv2 Routing Process supports source-route forwarding, and is willing to join the source route generated by other MP-OLSRv2 Routing Processes. The existence of SOURCE_ROUTE TLV MUST be consistent for a specific OLSRv2 Routing Process, i.e., either it adds SOURCE_ROUTE TLV to all its HELLO/TC messages, or it does not add SOURCE_ROUTE TLV to any HELLO/TC messages.

6.2. Datagram

6.2.1. Source Routing Header in IPv4

In IPv4 [RFC0791] networks, the MP-OLSRv2 Routing Process employs the loose source routing header, as defined in [RFC0791]. It exists as an option header, with option class 0, and option number 3.

The source route information is kept in the "route data" field of the loose source route header.

6.2.2. Source Routing Header in IPv6

In IPv6 [RFC2460] networks, the MP-OLSRv2 Routing Process employs the source routing header as defined in section 3 of [RFC6554], with IPv6 Routing Type 3.

The source route information is kept in the "Addresses" field of the routing header.

7. Information Bases

Each MP-OLSRv2 Routing Process maintains the information bases as defined in [RFC7181]. Additionally, a Multipath Information Base is used for this specification. It includes the protocol sets as defined below.

7.1. SR-OLSRv2 Router Set

The SR-OLSRv2 Router Set records the routers that support source-route forwarding. This includes routers that run the MP-OLSRv2 Routing Process or the OLSRv2 Routing Process with source-route forwarding support. The set consists of SR-OLSRv2 Router Tuples:

(SR_addr, SR_time)

where:

SR_addr - is the originator address of the router that supports source-route forwarding;

SR_time - is the time until which the SR-OLSRv2 Router Tuple is considered valid.

7.2. Multipath Routing Set

The Multipath Routing Set records the full path information of different paths to the destination. It consists of Multipath Routing Tuples:

(MR_dest_addr, MR_path_set)

where:

MR_dest_addr - is the network address of the destination, either the network address of an interface of a destination router or the network address of an attached network;

MP_path_set - contains the multiple paths to the destination. It consists of a set of Path Tuples.

Each Path Tuple is defined as:

(PT_metric, PT_address[1], PT_address[2], ..., PT_address[n])

where:

PT_metric - is the metric of the path to the destination, measured in LINK_METRIC_TYPE defined in [RFC7181];

PT_address[1, ..., n-1] - are the addresses of intermediate routers to be visited numbered from 1 to n-1, where n is the number of routers in the path, i.e., the hop count.

8. Protocol Details

This protocol is based on OLSRv2, and extended to discover multiple disjoint paths from a source router to a destination router. It retains the basic routing control packets formats and processing of OLSRv2 to obtain the topology information of the network. The main differences from the OLSRv2 Routing Process are the datagram processing at the source router and datagram forwarding.

8.1. HELLO and TC Message Generation

HELLO messages are generated according to Section 15.1 of [RFC7181], plus a single message TLV with Type := SOURCE_ROUTE included.

TC messages are generated according to Section 16.1 of [RFC7181] plus a single message TLV with Type := SOURCE_ROUTE included.

For the routers that do not generate TC messages according to [RFC7181], at least one TC message MUST be generated by an MP-OLSRv2 Routing Process during the SR_TC_INTERVAL (Section 5), which MUST be greater than or equal to TC_INTERVAL. Those TC messages MUST NOT carry any advertised neighbor addresses. This serves for those routers to advertise the SOURCE_ROUTE TLV so that the other routers can be aware of the source-route enabled routers so as to be used as destinations of multipath routing. The validity time associated with the VALIDITY_TIME TLV in such TC messages equals SR_HOLD_TIME, which MUST be greater than the SR_TC_INTERVAL. If the TC message carries

an optional INTERVAL_TIME TLV, it MUST have a value encoding the SR_TC_INTERVAL.

8.2. HELLO and TC Message Processing

HELLO and TC messages are processed according to section 15.3 and 16.3 of [RFC7181].

In addition to the reasons specified in [RFC7181] for discarding a HELLO message or a TC message on reception, a HELLO or TC message received MUST be discarded if it has more than one Message TLV with Type = SOURCE_ROUTE.

For every HELLO or TC message received, if there is a Message TLV with Type := SOURCE_ROUTE, create or update (if the Tuple exists already) the SR-OLSR Router Tuple with

- SR_addr := originator address of the HELLO or TC message
- SR_time := current_time + validity time of the TC or HELLO message defined in [RFC7181].

8.3. MPR Selection

Each MP-OLSRv2 Routing Process selects routing MPRs and flooding MPRs following Section 18 of [RFC7181]. In a mixed network with OLSRv2-only routers, the following considerations apply when calculating MPRs:

- MP-OLSRv2 routers SHOULD be preferred as routing MPRs to increase the possibility of finding disjoint paths using MP-OLSRv2 routers.
- The number of routing MPRs that run MP-OLSRv2 Routing Process MUST be equal or greater than NUMBER_OF_PATHS if there are enough MP-OLSRv2 symmetric neighbors. Otherwise all the MP-OLSRv2 routers are selected as routing MPRs, expect the routers with willingness WILL_NEVER.

8.4. Datagram Processing at the MP-OLSRv2 Originator

If datagrams without source routing header need to be forwarded using multiple paths (for example, based on the information of a DiffServ codepoint [RFC2474]), the MP-OLSRv2 Routing Process will try to find the Multipath Routing Tuple where:

- MR_dest_addr = destination of the datagram

If no matching Multipath Routing Tuple is found and the Multipath

Routing Set is maintained proactively, it indicates that there is no multipath route available to the desired destination. The datagram is forwarded following the OLSRv2 Routing Process.

If no matching Multipath Routing Tuple is found and the Multipath Routing Set is maintained reactively, the multipath algorithm defined in Section 8.5 is invoked, to calculate the Multipath Routing Tuple to the destination. If the calculation does not return any Multipath Routing Tuple, the following steps are aborted and the datagram is forwarded following the OLSRv2 Routing Process.

If a matching Multipath Routing Tuple is obtained, the Path Tuples of the Multipath Routing Tuple are applied to the datagrams using either per-flow scheduling or per-datagram scheduling, depending on the transport layer protocol and the application used. By default, per-flow scheduling is used, especially for the transport protocols that are sensitive to reordering, such as TCP. The path selection decision is made on the first datagram and all subsequent datagrams of the same flow use the same path. If the path is detected broken before the flow is closed, another path with the most similar metric is used. Per-datagram scheduling is recommended if the traffic is insensitive to reordering such as non-reliable transmission of media traffic, or when erasure coding is applied. In such case, each datagram selects its paths independently.

By default, the traffic load should be equally distributed in multiple paths. Other path scheduling mechanisms (e.g., assigning more traffic over better paths) are also possible and will not impact the interoperability of different implementations.

The addresses in PT_address[1, ..., n-1] of the chosen Path Tuple are thus added to the datagram header as the source routing header. For IPv6 networks, strict source routing is used, thus all the intermediate routers in the path are stored in the source routing header following the format defined in section 3 of [RFC6554] with Routing Type set to 3.

For IPv4 networks, loose source routing is used, with the following rules:

- Only the addresses that exist in SR-OLSR Router Set can be added to the source routing header.

- If the length of the path (n) is greater than MAX_SRC_HOPS (Section 5) or adding the whole path information exceeds the MTU, only the "key" routers in the path are kept. By default, the key routers are uniformly chosen in the path. If further information such as capacity of the routers (e.g., battery life) or the

routers' willingness in forwarding data is available, the routers with higher capacity and willingness are preferred.

- The routers that are considered not appropriate for forwarding indicated by external policies should be avoided.

It is not recommended to fragment the IP packet if the packet with the source routing header would exceed the minimum MTU along the path. Depending on the size of the routing domain, the MTU should be at least 1280 + 40 (for the outer IP header) + 16 * diameter of the network in number of hops (for the source routing header). If the links in the network have different MTU sizes, by using technologies like Path MTU Discovery, the routers are able to be aware of the MTU along the path. The size of the datagram plus the size of IP headers (including the source routing header) should not exceed the minimum MTU along the path, otherwise, the source routing should not be used.

If the destination of the datagrams is out the MP-OLSRv2 routing domain, the datagram must be source routed to the gateway between the MP-OLSRv2 routing domain and the rest of the Internet. The gateway MUST remove the source routing header before forwarding the datagram to the rest of the Internet.

8.5. Multipath Calculation

8.5.1. Requirements of Multipath Calculation

The Multipath Routing Set maintains the information of multiple paths to the destination. The Path Tuples of the Multipath Routing Set (Section 7.2) are generated based on a multipath algorithm.

For each path to a destination, the algorithm must provide:

- The metric of the path to the destination,

- The list of intermediate routers on the path.

For IPv6 networks, as strict source routing is used, only the routers that exist in the SR-OLSRv2 Router Set are considered in the path calculation, i.e., only the source-routing supported routers can exist in the path.

After the calculation of multiple paths, the metric of paths (denoted c_i for path i) to the destination is compared to the R_metric of the the OLSRv2 Routing Tuple ([RFC7181]) to the same destination. If the metric c_i is greater than R_metric * CUTOFF_RATIO (Section 5), the corresponding path i SHOULD NOT be used. If less than 2 paths are found with metrics less than R_metric * CUTOFF_RATIO, the router

SHOULD fall back to OLSRv2 Routing Process without using multipath routing. This can happen if there are too many OLSRv2-only routers in the network, and requiring multipath routing may result in inferior paths.

By invoking the multipath algorithm, up to NUMBER_OF_PATHS paths are obtained and added to the Multipath Routing Set by creating a Multipath Routing Tuple with:

- MR_dest_addr := destination of the datagram
- An MP_path_set with calculated Path Tuples. Each Path Tuple corresponds to a path obtained in the Multipath Dijkstra algorithm, with PT_metric := metric of the calculated path and PT_address[1, ..., n-1] := list of intermediate routers.

8.5.2. Multipath Dijkstra Algorithm

This section introduces the Multipath Dijkstra Algorithm as a default algorithm. It tries to obtain disjoint paths when appropriate, but does not guarantee strict disjoint paths. The use of other algorithms is not prohibited, as long as the requirements described in Section 8.5.1 are met. Using different multipath algorithms will not impact the interoperability.

The general principle of the Multipath Dijkstra Algorithm [ADHOC11] is using Dijkstra algorithm for multiple iterations, and at iteration i to look for the shortest path P[i] to the destination d. After each iteration, the metric of used links is increased. Compared to the original Dijkstra’s algorithm, the main modification consists in adding two incremental functions named metric functions fp and fe in order to prevent the next steps resulting in similar paths:

- fp(c) is used to increase metrics of arcs belonging to the previous path P[i-1] (with i>1), where c is the value of the previous metric. This encourages future paths to use different arcs but not different vertices.
- fe(c) is used to increase metrics of the arcs that lead to intermediate vertices of the previous path P[i-1] (with i>1), where c is the value of the previous metric. The "lead to" means that only one vertex of the arc belongs to the previous path P[i-1], while the other vertex does not. The "intermediate" means that the source and destination vertices are not considered.

Considering the simple example in Figure 1: a path P[i] S--A--D is obtained at step i. For the next step, the metric of link S--A and A--D are to be increased using fp(c), because they belong to the path

P[i]. A--B is to be increased using fe(c), because A is an intermediate vertex of path P[i], and B is not part of P[i]. B--D is unchanged.

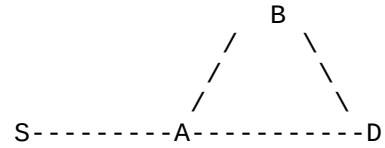


Figure 1

It is possible to choose different fp and fe to get link-disjoint paths or node-disjoint paths as desired. A recommendation for configuration of fp and fe is given in Section 9.

To get NUMBER_OF_PATHS different paths, for each path P[i] (i = 1, ..., NUMBER_OF_PATHS) do:

1. Run Dijkstra's algorithm to get the shortest path P[i] for the destination d.

2. Apply metric function fp to the metric of links (in both directions) in P[i].

3. Apply metric function fe to the metric of links (in both directions) that lead to routers used in P[i].

A simple example of the Multipath Dijkstra Algorithm is illustrated in Appendix A.

8.6. Multipath Routing Set Updates

The Multipath Routing Set MUST be updated when the Local Information Base, the Neighborhood Information Base, or the Topology Information Base indicate a change (including of any potentially used outgoing neighbor metric values) of the known symmetric links and/or attached networks in the MANET, hence changing the Topology Graph, as described in section 17.7 of [RFC7181]. How the Multipath Routing Set is updated depends on whether the set is maintained reactively or proactively:

- In reactive mode, all the Tuples in the Multipath Routing Set are removed. The new arriving datagrams will be processed as specified in Section 8.4;

- In proactive mode, the route to all the destinations are updated according to Section 8.5.

8.7. Datagram Forwarding

In IPv4 networks, datagrams are forwarded using loose source routing as specified in Section 3.1 of [RFC0791].

In IPv6 networks, datagrams are forwarded using strict source routing as specified in Section 4.2 of [RFC6554], except the applied routers are MP-OLSRv2 routers rather than RPL routers. The last hop of the source route MUST remove the source routing header.

9. Configuration Parameters

This section gives default values and guidelines for setting parameters defined in Section 5. Network administrators may wish to change certain or all the parameters for different network scenarios. As an experimental protocol, the users of this protocol are also encouraged to explore different parameter setting in various network environments, and provide feedback.

- NUMBER_OF_PATHS := 3. This parameter defines the number of parallel paths used in datagram forwarding. Setting it to one makes the specification identical to OLSRv2. Setting it to too large values may lead to unnecessary computational overhead and inferior paths.

- MAX_SRC_HOPS := 10, for IPv4 networks. For IPv6 networks, it MUST be set to 0, i.e., no constraint on maximum number of hops.

- CUTOFF_RATIO := 1.5. It MUST be greater or equal than 1.

- SR_TC_INTERVAL := 10 x TC_INTERVAL. It MUST be greater than or equal to TC_INTERVAL. It SHOULD be significantly greater than TC_INTERVAL to reduce unnecessary TC message generations.

- SR_HOLD_TIME := 3 x SR_TC_INTERVAL. It MUST be greater than SR_TC_INTERVAL and SHOULD allow for a small number of lost messages.

If Multipath Dijkstra Algorithm is applied:

- fp(c) := 4*c, where c is the original metric of the link.

- fe(c) := 2*c, where c is the original metric of the link.

The setting of metric functions fp and fc defines the preference of obtained multiple disjoint paths. If id is the identity function, i.e., fp(c)=c, 3 cases are possible:

- if id=fe<fp: only increase the metric of related links;
- if id<fe=fp: apply equal increase to the metric of related nodes and links;
- if id<fe<fp: apply greater increase to the metric of related links.

Increasing the metric of related links or nodes means avoiding the use of such links or nodes in the next path to be calculated.

10. Implementation Status

The RFC Editor is advised to remove the entire section before publication, as well as the reference to RFC 7942.

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

Until April 2015, there are 3 open source implementations of the protocol specified in this document, for both testbed and simulation use.

10.1. Multipath extension based on nOLSRv2

The implementation is conducted by University of Nantes, France, and is based on Niigata University's nOLSRv2 implementation. It is an open source implementation. The code is available at https://github.com/yijiazi/mpolsr_qualnet and http://jiaziyi.com/index.php/research-projects/mp-olsr .

It can be used for Qualnet simulations, and be exported to run in a testbed. All the specification is implemented in this implementation.

Implementation experience and test data can be found at [ADHOC11].

10.2. Multipath extension based on olsrd

The implementation is conducted under SEREADMO (Securite des Reseaux Ad Hoc & Mojette) project, and supported by French research agency (RNRT2803). It is based on olsrd (http://www.olsr.org/) implementation, and is open sourced. The code is available at https://github.com/yijiazi/mpolsr_testbed and http://jiaziyi.com/index.php/research-projects/sereadmo .

The implementation is for testing the specification in the field. All the specification is implemented in this implementation.

Implementation experience and test data can be found at [ADHOC11] and [GIIS14].

10.3. Multipath extension based on umOLSR

The implementation is conducted by University of Nantes, France, and is based on um-olsr implementation (http://masimum.inf.um.es/fjrm/development/um-olsr/). The code is available at https://github.com/yijiazi/mpolsr_ns2 and http://jiaziyi.com/index.php/research-projects/mp-olsr under GNU GPL license.

The implementation is for network simulation for NS2 network simulator. All the specification is implemented in this implementation.

Implementation experience and test data can be found at [WCNC08].

11. Security Considerations

As an extension of [RFC7181], the security considerations and

security architecture illustrated in [RFC7181] are applicable to this MP-OLSRv2 specification. The implementations without security mechanisms are vulnerable to threats discussed in [I-D.ietf-manet-olsrv2-sec-threats].

In a mixed network with OLSRv2-only routers, a compromised router can add SOURCE_ROUTE TLVs in its TC and HELLO messages, which will make other MP-OLSRv2 Routing Processes believe that it supports source routing. This will increase the possibility of being chosen as MPRs and put into the source routing header. The former will make it possible to manipulate the flooding of TC messages and the latter will make the datagram pass through the compromised router.

As with [RFC7181], a conformant implementation of MP-OLSRv2 MUST, at minimum, implement the security mechanisms specified in [RFC7183] to provide integrity and replay protection of routing control messages.

The MP-OLSRv2 Routing Process MUST drop datagrams entering or exiting a OLSRv2/MP-OLSRv2 routing domain that contain a source routing header. Compared to OLSRv2, the use of the source routing header in this specification introduces vulnerabilities related to source routing attacks, which include bypassing filtering devices, bandwidth exhaustion of certain routers, etc. Those attacks are discussed in Section 5 of [RFC6554] and [RFC5095]. The influence is limited to the OLSRv2/MP-OLSRv2 routing domain, because the source routing header is used only in the current routing domain.

If the multiple paths are calculated reactively, the datagrams SHOULD be buffered while the paths are being calculated. Because the path calculation is local and no control message is exchanged, the buffering time should be trivial. However, depending on the CPU power and memory of the router, a maximum buffer size SHOULD be set to avoid occupying too much memory of the router. When the buffer is full, the oldest datagrams are dropped. A possible attack that a malicious application could launch is that it initiates a large amount of datagrams to all the other routers in the network, thus triggering path calculation to all the other routers and during which the datagrams are buffered. This might flush other legitimate datagrams. But the impact of the attack is transient: once the path calculation is finished, the datagrams are forwarded and the buffer goes back to empty.

12. IANA Considerations

This section adds one new Message TLV, allocated as a new Type Extension to an existing Message TLV.

12.1. Message TLV Types

This specification updates the IANA registry "Message TLV Types" -- Message Type 7 by adding the new Type Extension SOURCE_ROUTE, as illustrated in Table 1.

Type Extension	Name	Description	Reference
TBD	SOURCE_ROUTE	Indicates that the originator of the message supports source route forwarding. No value.	This specification

Table 1: SOURCE_ROUTE type for RFC 5444 Type 7 Message TLV Type Extensions

13. Acknowledgments

The authors would like to thank Sylvain David, Asmaa Adnane, Eddy Cizeron, Salima Hamma, Pascal Lesage and Xavier Lecourtier for their efforts in developing, implementing and testing the specification. The authors also appreciate valuable discussions with Thomas Clausen, Ulrich Herberg, Justin Dean, Geoff Ladwig, Henning Rogge , Marcus Barkowsky and especially Christopher Dearlove for his multiple rounds of reviews during the working group last calls.

14. References

14.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <http://www.rfc-editor.org/info/rfc791>.

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

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized Mobile Ad Hoc Network (MANET) Packet/Message Format", RFC 5444, DOI 10.17487/RFC5444, February 2009,

<http://www.rfc-editor.org/info/rfc5444>.

[RFC6130] Clausen, T., Dearlove, C., and J. Dean, "Mobile Ad Hoc Network (MANET) Neighborhood Discovery Protocol (NHDP)", RFC 6130, DOI 10.17487/RFC6130, April 2011, <http://www.rfc-editor.org/info/rfc6130>.

[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, <http://www.rfc-editor.org/info/rfc6554>.

[RFC7181] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol Version 2", RFC 7181, DOI 10.17487/RFC7181, April 2014, <http://www.rfc-editor.org/info/rfc7181>.

[RFC7183] Herberg, U., Dearlove, C., and T. Clausen, "Integrity Protection for the Neighborhood Discovery Protocol (NHDP) and Optimized Link State Routing Protocol Version 2 (OLSRv2)", RFC 7183, DOI 10.17487/RFC7183, April 2014, <http://www.rfc-editor.org/info/rfc7183>.

14.2. Informative References

[ADHOC11] Yi, J., Adnane, A-H., David, S., and B. Parrein, "Multipath optimized link state routing for mobile ad hoc networks", In Elsevier Ad Hoc Journal, vol.9, n. 1, 28-47, January, 2011.

[GIIS14] Macedo, R., Melo, R., Santos, A., and M. Nogueria, "Experimental performance comparison of single-path and multipath routing in VANETs", In Global Information Infrastructure and Networking Symposium (GIIS), 2014 , vol. 1, no. 6, pp. 15-19, 2014.

[I-D.ietf-6man-segment-routing-header] Previdi, S., Filsfils, C., Raza, K., Leddy, J., Field, B., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Matsushima, S., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., Lebrun, D., Steinberg, D., and R. Raszuk, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-06 (work in progress), March 2017.

[I-D.ietf-manet-olsrv2-sec-threats] Clausen, T., Herberg, U., and J. Yi, "Security Threats to

the Optimized Link State Routing Protocol version 2 (OLSRv2)", draft-ietf-manet-olsrv2-sec-threats-04 (work in progress), January 2017.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <http://www.rfc-editor.org/info/rfc2460>.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <http://www.rfc-editor.org/info/rfc2474>.

[RFC2501] Corson, S. and J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, DOI 10.17487/RFC2501, January 1999, <http://www.rfc-editor.org/info/rfc2501>.

[RFC2991] Thaler, D. and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection", RFC 2991, DOI 10.17487/RFC2991, November 2000, <http://www.rfc-editor.org/info/rfc2991>.

[RFC5095] Abley, J., Savola, P., and G. Neville-Neil, "Deprecation of Type 0 Routing Headers in IPv6", RFC 5095, DOI 10.17487/RFC5095, December 2007, <http://www.rfc-editor.org/info/rfc5095>.

[RFC7722] Dearlove, C. and T. Clausen, "Multi-Topology Extension for the Optimized Link State Routing Protocol Version 2 (OLSRv2)", RFC 7722, DOI 10.17487/RFC7722, December 2015, <http://www.rfc-editor.org/info/rfc7722>.

[RFC7779] Rogge, H. and E. Baccelli, "Directional Airtime Metric Based on Packet Sequence Numbers for Optimized Link State Routing Version 2 (OLSRv2)", RFC 7779, DOI 10.17487/RFC7779, April 2016, <http://www.rfc-editor.org/info/rfc7779>.

[RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <http://www.rfc-editor.org/info/rfc7942>.

[WCNC08] Yi, J., Cizeron, E., Hamma, S., and B. Parrein, "Simulation and performance analysis of MP-OLSR for mobile

ad hoc networks", In Proceeding of IEEE Wireless Communications and Networking Conference, 2008.

[WPMC11] Yi, J., Parrein, B., and D. Radu, "Multipath routing protocol for manet: Application to H.264/SVC video content delivery", In Proceeding of 14th International Symposium on Wireless Personal Multimedia Communications.

Appendix A. Examples of Multipath Dijkstra Algorithm

This appendix gives two examples of Multipath Dijkstra algorithm.

A network topology is depicted in Figure 2.

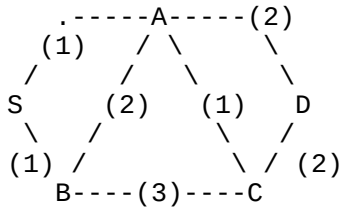


Figure 2

The capital letters are the names of routers. An arbitrary metric with value between 1 and 3 is used. The initial metrics of all the links are indicated in the parentheses. The incremental functions fp(c)=4c and fe(c)=2c are used in this example. Two paths from router S to router D are demanded.

On the first run of the Dijkstra algorithm, the shortest path S->A->D with metric 3 is obtained.

The incremental function fp is applied to increase the metric of the link S-A and A-D. fe is applied to increase the metric of the link A-B and A-C. Figure 3 shows the link metrics after the increment.

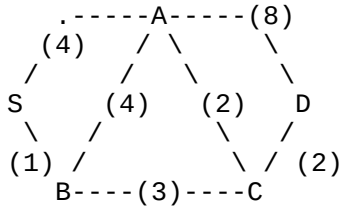


Figure 3

On the second run of the Dijkstra algorithm, the second path S->B->C->D with metric 6 is obtained.

As mentioned in Section 8.5, the Multipath Dijkstra Algorithm does not guarantee strict disjoint paths in order to avoid choosing inferior paths. For example, given the topology in Figure 4, two paths from node S to D are desired. On the top of the figure, there is a high cost path between S and D.

If a algorithm tries to obtain strict disjoint paths, the two paths obtained will be S--B--D and S--(high cost path)--D, which are extremely unbalanced. It is undesirable because it will cause huge delay variance between the paths. By using the Multipath Dijkstra algorithm, which is based on the punishing scheme, S--B--D and S--B--C--D will be obtained.

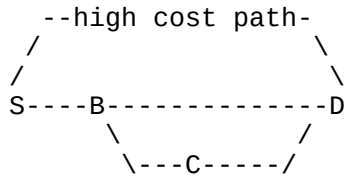


Figure 4

Authors' Addresses

Jiazi Yi
Ecole Polytechnique
91128 Palaiseau Cedex,
France

Phone: +33 (0) 1 77 57 80 85
Email: jiazi@jiaziyi.com
URI: http://www.jiaziyi.com/

Benoit Parrein
University of Nantes
IRCCyN lab - IVC team
Polytech Nantes, rue Christian Pauc, BP50609
44306 Nantes cedex 3
France

Phone: +33 (0) 2 40 68 30 50
Email: Benoit.Parrein@polytech.univ-nantes.fr
URI: http://www.irccyn.ec-nantes.fr/~parrein

Mobile Ad hoc Networking (MANET)
Internet-Draft
Updates: 7186 (if approved)
Intended status: Informational
Expires: February 28, 2017

J. Yi
T. Clausen
Ecole Polytechnique
U. Herberg

August 27, 2016

Security Threats for Simplified Multicast Forwarding (SMF) draft-ietf-manet-smf-sec-threats-06

Abstract

This document analyzes security threats of the Simplified Multicast Forwarding (SMF) mechanism, including the vulnerabilities of duplicate packet detection and relay set selection mechanisms. This document is not intended to propose solutions to the threats described.

This document also updates RFC7186 regarding the threats to relay set selection mechanisms using RFC6130.

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 http://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 February 28, 2017.

Copyright Notice

Table of Contents

1. Introduction

This document analyzes security threats to the Simplified Multicast Forwarding (SMF) mechanism [RFC6621]. SMF aims at providing basic Internet Protocol (IP) multicast forwarding, in a way that is suitable for limited wireless mesh and Mobile Ad hoc NETworks (MANET). SMF is constituted of two major functional components: Duplicate Packet Detection and Relay Set Selection.

SMF is typically used in decentralized wireless environments, and is potentially exposed to various attacks and misconfigurations. Some of these attacks and misconfigurtions, in a wireless enviroment, represent threats of particular significance as compared to what they would do in wired networks. [RFC6621] briefly discusses several of these, but does not define any explicit security measures for protecting the integrity of the protocol.

This document is based on the assumption that no additional security mechanism such as IPsec is used in the IP layer, as not all MANET deployments may be suitable to deploy common IP protection mechanisms (e.g., because of limited resources of MANET routers to support the IPsec stack). It assumes that there is no lower-layer protection either. The document analyzes possible attacks on and mis-configurations of SMF and outlines the consequences of such attacks/ mis-configurations to the state maintained by SMF in each router.

In the Security Considerations section of [RFC6621], denial-of-service attack scenarios are briefly discussed. This document further analyzes and describes the potential vulnerabilities of and attack vectors for SMF. While completeness in such analysis is always a goal, no claims of being complete are made. The goal of this document is to be helpful for when deploying SMF in a network and needing to understand the risks thereby incurred - as well as for providing a reference and documented experience with SMF as input for possibly future developments of SMF.

This document is not intended to propose solutions to the threats described. [RFC7182] provides a framework that can be used with SMF, and depending on how it is used - may offer some degree of protection against the threats described in this document related to identity spoofing.

This document also updates [RFC7186], specifically with respect to threats to relay set selection mechanisms which are using [RFC6130].

2. Terminology

This document uses the terminology and notation defined in [RFC5444], [RFC6130], [RFC6621] and [RFC4949].

Additionally, this document introduces the following terminology:

SMF router: A MANET router, running SMF as specified in [RFC6621].

Attacker: A device that is present in the network and intentionally seeks to compromise the information bases in SMF routers. It may generate syntactically correct SMF control messages.

Legitimate SMF router: An SMF router that is correctly configured and not compromised by an attacker.

3. SMF Threats Overview

SMF requires an external dynamic neighborhood discovery mechanism in order to maintain suitable topological information describing its immediate neighborhood, and thereby allowing it to select reduced relay sets for forwarding multicast data traffic. Such an external dynamic neighborhood discovery mechanism may be provided by lower-layer interface information, by a concurrently operating MANET routing protocol that already maintains such information such as [RFC7181], or by explicitly using MANET Neighborhood Discovery Protocol (NHDP) [RFC6130]. If NHDP is used for both 1-hop and 2-hop neighborhood discovery by SMF, SMF implicitly inherits the vulnerabilities of NHDP discussed in [RFC7186]. As SMF relies on NHDP to assist in network layer 2-hop neighborhood discovery (no matter if other lower-layer mechanisms are used for 1-hop neighborhood discovery), this document assumes that NHDP is used in SMF. The threats that are NHDP-specific are indicated explicitly.

Based on neighborhood discovery mechanisms, [RFC6621] specifies two principal functional components: Duplicate Packet Detection (DPD) and Relay Set Selection (RSS).

DPD is required by SMF in order to be able to detect duplicate packets and eliminate their redundant forwarding. An Attacker has two ways in which to harm the DPD mechanisms, specifically it can:

- "deactivate" DPD, so as to make it such that duplicate packets are not correctly detected, and that as a consequence they are (redundantly) transmitted, increasing the load on the network, draining the batteries of the routers involved, etc.

- o "pre-activate" DPD, so as to make DPD detect a later arriving (valid) packet as being a duplicate, which therefore won't be forwarded.

Attacks on DPD can be achieved by replaying existing packets, by wrangling sequence numbers, by manipulating hash values, etc., and are detailed in Section 4.

RSS produces a reduced relay set for forwarding multicast data packets across the MANET. [RFC6621] specifies several relay set algorithms, including E-CDS (Essential Connected Dominating Set) [RFC5614], S-MPR (Source-based Multi-point Relay, as known from [RFC3626] and [RFC7181]), or MPR-CDS [MPR-CDS], for use in SMF. An Attacker can disrupt the RSS algorithm, and thereby SMF operation, by degrading it to classical flooding, or by "masking" certain parts of the network from the multicasting domain. Attacks on RSS algorithms are detailed in Section 5.

Other than the attacks on DPD and RSS, a common vulnerability of MANETs is "jamming", i.e., a device generates massive amounts of interfering radio transmissions, which will prevent legitimate traffic (e.g., control traffic as well as data traffic) on part of a network. The attacks on DPD and RSS can be further enhanced by jamming.

4. Threats to Duplicate Packet Detection

Duplicate Packet Detection (DPD) is required for packet dissemination in MANETs because: (1) packets may be transmitted via the same physical interface as the one over which they were received, and (2) a router may receive multiple copies of the same packet (on the same, or on different interfaces) from different neighbors. DPD is thus used to check if an incoming packet has been previously received or not.

DPD is achieved by maintaining a record of recently processed multicast packets, and comparing later received multicast packets herewith. A duplicate packet detected is silently dropped and is not inserted into the forwarding path of that router, nor is it delivered to an application. DPD, as proposed by SMF, supports both IPv4 and IPv6 and for each suggests two duplicate packet detection mechanisms: 1) header content identification-based DPD (I-DPD), using packet headers, in combination with flow state, to estimate temporal uniqueness of a packet, and 2) hash-based DPD (H-DPD), employing hashing of selected header fields and payload for the same effect.

In the Security Considerations section of [RFC6621], a selection of

threats to DPD are briefly introduced. This section expands on that discussion, and describes how to effectively launch the attacks on DPD - for example, by way of manipulating jitter and/or the Hash-Assistant Value. In the remainder of this section, common threats to packet detection mechanisms are first discussed. Then the threats to I-DPD and H-DPD are introduced separately. The threats described in this section are applicable to general SMF implementations, no matter if NHDP is used or not.

4.1. Attack to The Hop Limit Field

One immediate DoS attack is based on manipulating the Time-to-Live (TTL, for IPv4) or hop limit (for IPv6) field. As routers only forward packets with TTL > 1, an attacker can forward an otherwise valid packet, while drastically reducing the TTL hereof. This will inhibit recipient routers from later forwarding the same multicast packet, even if received with a different TTL - essentially an attacker thus can instruct its neighbors to block forwarding of valid multicast packets.

For example, in Figure 1, router A forwards a multicast packet with a TTL of 64 to the network. A, B, and C are legitimate SMF routers, and X is an attacker. In a wireless environment, jitter is commonly used to avoid systematic collisions in MAC protocols [RFC5148]. An attacker can thus increase the probability that its invalid packets arrive first by retransmitting them without applying jitter. In this example, router X forwards the packet without applying jitter and reduces the TTL to 1. Router C thus records the duplicate detection value (hash value for H-DPD, or the header content of the packets for I-DPD) but does (due to TTL == 1) not forward. When a second copy the same packet, with a non-maliciously manipulated TTL value (63 in this case), arrives from router B, it will be discarded as duplicate packet.

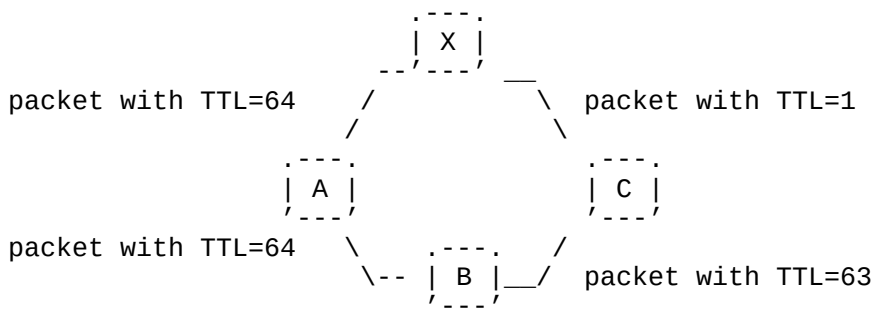


Figure 1

As the TTL of a packet is intended to be manipulated by intermediaries forwarding it, classic methods such as integrity check values (e.g., digital signatures) are typically calculated with setting TTL fields to some pre-determined value (e.g., 0) - such is for example the case for IPsec Authentication Headers - rendering such an attack more difficult to both detect and counter.

If the attacker has access to a "wormhole" through the network (a directional antenna, a tunnel to a collaborator or a wired connection, allowing it to bridge parts of a network otherwise distant), it can make sure that the packets with such an artificially reduced TTL arrive before their unmodified counterparts.

4.2. Threats to Identification-based Duplicate Packet Detection

I-DPD uses a specific DPD identifier in the packet header to identify a packet. By default, such packet identification is not provided by the IP packet header (for both IPv4 and IPv6). Therefore, additional identification headers, such as the fragment header, a hop-by-hop header option, or IPSec sequencing, must be employed in order to support I-DPD. The uniqueness of a packet can then be identified by the source IP address of the packet originator and the sequence number (from the fragment header, hop-by-hop header option, or IPsec). By doing so, each intermediate router can keep a record of recently received packets and determine whether the incoming packet has been received or not.

4.2.1. Pre-activation Attacks (Pre-Play)

In a wireless environment, or across any other shared channel, an attacker can perceive the identification tuple (source IP address, sequence number) of a packet. It is possible to generate a packet with the same (source IP address, sequence number) pair with invalid

content. If sequence number progression is predictable, then it is trivial to generate and inject invalid packets with "future" identification information into the network. If these invalid packets arrive before the legitimate packets that they are spoofing, the latter will be treated as a duplicate and discarded. This can prevent multicast packets from reaching parts of the network.

Figure 2 gives an example of pre-activation attack. A, B and C are legitimate SMF routers, and X is the attacker. The line between the routers presents the packet forwarding. Router A is the source and originates a multicast packet with sequence number n. When router X receives the packet, it generates an invalid packet with the source address of A and sequence number n. If the invalid packet arrives at router C before the forwarding of router B, the valid packet will be dropped by C as a duplicate packet. An attacker can manipulate jitter to make sure that the invalid packets arrive first. Router X can even generate packets with future sequence numbers (if they are predictable), so that the future legitimate packets with the same sequence numbers will be dropped as duplicate ones.

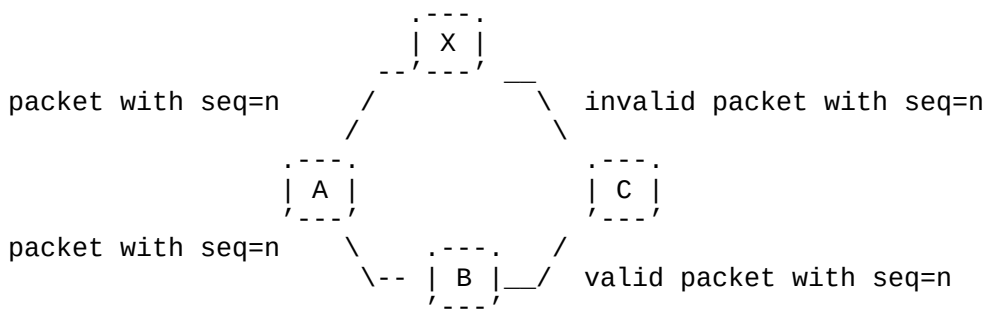


Figure 2

As SMF currently does not have any timestamp mechanisms to protect data packets, there is no viable way to detect such pre-play attacks by way of timestamps. Especially, if the attack is based on manipulation of jitter, the validation of timestamp would not be helpful because the timing is still valid (but with much less value).

4.2.2. De-activation Attacks (Sequence Number wrangling)

An attacker can also seek to de-activate DPD, by modifying the sequence number in packets that it forwards. Thus, routers will not be able to detect an actual duplicate packet as a duplicate - rather, they will treat them as new packets, i.e., process and forward them. This is similar to DoS attacks, as each packet that is considered

unique will be multicasted: for a network with n routers, there will be n-1 retransmissions. This can easily cause the "broadcast storm" problem discussed in [MOBICOM99]. The consequence of this attack is an increased channel load, the origin of which appears to be a router other than the attacker.

Given the topology shown in Figure 2, on receiving a packet with seq=n, the attacker X can forward the packet with modified sequence number n+i. This has two consequences: firstly, router C will not be able to detect the packet forwarded by X is a duplicate packet; secondly, the consequent packet with seq=n+i generated by router A probably will be treated as duplicate packet, and dropped by router C.

4.3. Threats to Hash-based Duplicate Packet Detection

When explicit sequence numbers in packet headers is undesired, hash-based DPD can be used. A hash of the non-mutable fields in the header of and the data payload can be generated, and recorded at the intermediate routers. A packet can thus be uniquely identified by the source IP address of the packet and its hash-value.

The hash algorithm used by SMF is being applied only to provide a reduced probability of collision and is not being used for cryptographic or authentication purposes. Consequently, a digest collision is still possible. In case the source router or gateway identifies that it recently has generated or injected a packet with the same hash-value, it inserts a "Hash-Assist Value (HAV)" IPv6 header option into the packet, such that calculating the hash also over this HAV will render the resulting value unique.

4.3.1. Attack on Hash-Assistant Value

The HAV header is helpful when a digest collision happens. However, it also introduces a potential vulnerability. As the HAV option is only added when the source or the ingress SMF router detects that the coming packet has digest collision with previously generated packets, it actually can be regarded as a "flag" of potential digest collision. An attacker can discover the HAV header, and be able to conclude that a hash collision is possible if the HAV header is removed. By doing so, the modified packet received by other SMF routers will be treated as duplicate packets, and be dropped because they have the same hash value with the precedent packet.

In the example of Figure 3, Router A and B are legitimate SMF routers; X is an attacker. A generates two packets P1 and P2, with the same hash value h(P1)=h(P2)=x. Based on the SMF specification, a hash-assistant value (HAV) is added to the latter packet P2, so that

h(P2+HAV)=x', to avoid digest collision. When the attacker X detects the HAV of P2, it is able to conclude that a collision is possible by removing the HAV header. By doing so, packet P2 will be treated as duplicate packet by router B, and be dropped.

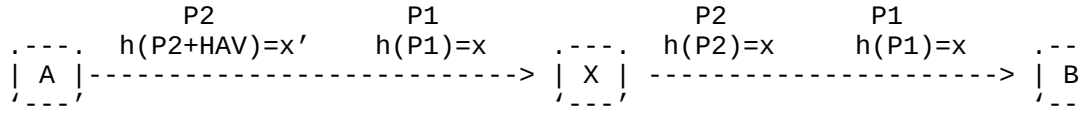


Figure 3

5. Threats to Relay Set Selection

A framework for RSS mechanism, rather than a specific RSS algorithm is provided by SMF. It is normally achieved by distributed algorithms that can dynamically generate a topological Connected Dominating Set based on 1-hop and 2-hop neighborhood information. In this section, the common threats to the RSS framework are first discussed. Then the three commonly used algorithms: Essential Connection Dominating Set (E-CDS) algorithm, Source-based Multipoint Relay (S-MPR) and Multipoint Relay Connected Dominating Set (MPR-CDS) are analyzed. As the relay set selection is based on 1-hop and 2-hop neighborhood information, which rely on NHDP, the threats described in this section are NHDP-specific.

5.1. Relay Set Selection Common Threats

Common (i.e., non algorithm specific) threats to RSS algorithms, including Denial of Service attack, eavesdropping, message timing attack and broadcast storm have been discussed in [RFC7186].

5.2. Threats to E-CDS Algorithm

The "Essential Connected Dominating Set" (E-CDS) algorithm [RFC5614] forms a single CDS mesh for the SMF operating region. It requires 2-hop neighborhood information (the identify of the neighbors, the link to the neighbors and neighbors' priority information) collected through NHDP or another process.

An SMF Router will select itself as a relay, if:

- The SMF Router has a higher priority than all of its symmetric neighbors, or

- o There does not exist a path from the neighbor with largest priority to any other neighbor, via neighbors with greater priority.

An attacker can disrupt the E-CDS algorithm by link spoofing or identity spoofing.

5.2.1. Link Spoofing

Link spoofing implies that an attacker advertises non-existing links to another router (present in the network or not).

An attacker can declare itself with high route priority, and spoofs the links to as many legitimate SMF Routers as possible to declare high connectivity. By doing so, it can prevent legitimate SMF Routers from self-selecting as relays. As the "super" relay in the network, the attacker can manipulate the traffic relayed by it.

5.2.2. Identity Spoofing

Identity spoofing implies that an attacker determines and makes use of the identity of other legitimate routers, without being authorized to do so. The identity of other routers can be obtained by overhearing the control messages or the source/destination address from datagrams. The attacker can then generate control or datagram traffic, pretending to be a legitimate router.

Because E-CDS self-selection is based on the router priority value, an attacker can spoof the identity of other legitimate routers, and declares a different router priority value. If it declares a higher priority of a spoofed router, it can prevent other routers from selecting themselves as relays. On the other hand, if the attacker declares lower priority of a spoofed router, it can force other routers to selecting themselves as relays, to degrade the multicast forwarding to classical flooding.

5.3. Threats to S-MPR Algorithm

The source-based multipoint relay (S-MPR) set selection algorithm enables individual routers, using 2-hop topology information, to select relays from their set of neighboring routers. MPRs are selected so that forwarding to the router’s complete 2-hop neighbor set is covered.

An SMF router forwards a multicast packet if and only if:

- o the packet has not been received before, and

- o the neighbor from which the packet was received has selected the router as MPR.

Because MPR calculation is based on the willingness declared by the SMF routers, and the connectivity of the routers, it can be disrupted by both link spoofing and identity spoofing. The threats and its impacts have been illustrated in section 5.1 of [RFC7186].

5.4. Threats to MPR-CDS Algorithm

MPR-CDS is a derivative from S-MPR. The main difference between S-MPR and MPR-CDS is that while S-MPR forms a different broadcast tree for each source in the network, MPR-CDS forms a unique broadcast tree for all sources in the network.

As MPR-CDS combines E-CDS and S-MPR and the simple combination of the two algorithms does not address the weakness, the vulnerabilities of E-CDS and S-MPR that discussed in Section 5.2 and Section 5.3 apply to MPR-CDS also.

6. Security Considerations

This document does not specify a protocol or a procedure. The whole document, however, reflects on security considerations for SMF for packet dissemination in MANETs. Possible attacks to the two main functional components of SMF, duplicate packet detection and relay set selection, are analyzed and documented.

Although [RFC6621] nor this document propose mechanisms to secure the SMF protocol, there are several possibilities to secure the protocol in the future and driving new work by suggesting which threats discussed in the previous sections could be addressed.

For the I-DPD mechanism, employing randomized packet sequence numbers can avoid some pre-activation attacks based on sequence number prediction. If predicable sequence numbers have to be used, applying timestamps can mitigate pre-activation attacks.

For the H-DPD mechanism, applying cryptographically strong hashes can make the digest collisions effectively impossible, and avoid the use of hash-assistant value.

[RFC7182] specifies a framework for representing cryptographic Integrity Check Values (ICVs) and timestamps in MANETs. Based on [RFC7182], [RFC7183] specifies integrity and replay protection for NHDP using shared keys, as a mandatory-to-implement security mechanism. If SMF is using NHDP as neighborhood discovery protocol,

implementing [RFC7183] remains advisable so as to enable integrity protection for NHDP control messages. This can help mitigating threats related to identity spoofing through the exchange of HELLO messages, and provides some general protection against identity spoofing by admitting only trusted routers to the network using ICVs in HELLO messages.

Using ICVs does, of course, not address the problem of attackers, able to also generate valid ICVs. Detection and exclusion of such attackers is, in general, a challenge, which is not unrelated to how [RFC7182] is used. If, for example, it is used with a shared key (as per [RFC7183]), excluding single attackers generally is not aided by the use of ICVs. However if routers have sufficient capabilities to support the use of asymmetric keys (as per [RFC7859]), part of addressing this challenge becomes one of providing key revocation, in a way that does not in itself introduce additional vulnerabilities.

As [RFC7183] does not protect the integrity of the multicast user datagram, and as no mechanism is specified by SMF for doing so, duplicate packet detection remains vulnerable to the threats introduced in Section 4.

If pre-activation/de-activation attacks and attack on hash-assistant value of the multicast datagrams are to be mitigated, a datagram-level integrity protection mechanism is desired, by taking consideration of the identity field or hash-assistant value. However, this would not be helpful for the attacks on the TTL (or hop limit for IPv6) field, because the mutable fields are generally not considered when ICV is calculated.

7. IANA Considerations

This document contains no actions for IANA.

[RFC Editor: please remove this section prior to publication.]

8. Acknowledgments

The authors would like to thank Christopher Dearlove (BAE Systems ATC) who provided detailed review and valuable comments.

9. References

9.1. Normative References

[RFC6130] Clausen, T., Dean, J., and C. Dearlove, "Mobile Ad Hoc Network (MANET) Neighborhood Discovery Protocol (NHDP)", RFC 6130, April 2011.

[RFC6621] Macker, J., "Simplified Multicast Forwarding", RFC 6621, May 2012.

[RFC7186] Yi, J., Herberg, U., and T. Clausen, "Security Threats for the Neighborhood Discovery Protocol (NHDP)", RFC 7186, April 2014.

9.2. Informative References

[MOBICOM99] Ni, S., Tseng, Y., Chen, Y., and J. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network", Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking, 1999.

[MPR-CDS] Adjih, C., Jacquet, P., and L. Viennot, "Computing Connected Dominating Sets with Multipoint Relays", Journal of Ad Hoc and Sensor Wireless Networks 2002, January 2002.

[RFC3626] Clausen, T. and P. Jacquet, "The Optimized Link State Routing Protocol", RFC 3626, October 2003.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", RFC 4949, August 2007.

[RFC5148] Clausen, T., Dearlove, C., and B. Adamson, "Jitter Considerations in Mobile Ad Hoc Networks (MANETs)", RFC 5148, February 2008.

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized MANET Packet/Message Format", RFC 5444, February 2009.

[RFC5614] Ogier, R. and P. Spagnolo, "Mobile Ad Hoc Network (MANET) Extension of OSPF Using Connected Dominating Set (CDS) Flooding", RFC 5614, August 2009.

[RFC7181] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol version 2", RFC 7181, April 2014.

[RFC7182] Herberg, U., Clausen, T., and C. Dearlove, "Integrity

Check Value and Timestamp TLV Definitions for Mobile Ad Hoc Networks (MANETs)", RFC 7182, April 2014.

[RFC7183] Herberg, U., Dearlove, C., and T. Clausen, "Integrity Protection for the Neighborhood Discovery Protocol (NHDP) and Optimized Link State Routing Protocol Version 2 (OLSRv2)", RFC 7183, April 2014.

[RFC7859] Dearlove, C., "Identity-Based Signatures for Mobile Ad Hoc Network (MANET) Routing Protocols", RFC 7859, DOI 10.17487/RFC7859, May 2016, <http://www.rfc-editor.org/info/rfc7859>.

Authors' Addresses

Jiazi Yi
Ecole Polytechnique
91128 Palaiseau Cedex,
France

Phone: +33 1 77 57 80 85
Email: jiazi@jiaziyi.com
URI: http://www.jiaziyi.com/

Thomas Heide Clausen
Ecole Polytechnique
91128 Palaiseau Cedex,
France

Phone: +33 6 6058 9349
Email: T.Clausen@computer.org
URI: http://www.thomasclausen.org/

Ulrich Herberg

Email: ulrich@herberg.name
URI: http://www.herberg.name/