Network Working Group
Internet-Draft
Intended status: Experimental
Expires: May 09, 2014

B. Adamson
Naval Research Laboratory
C. Danilov
Boeing Corporation
J. Macker
Naval Research Laboratory

November 05, 2013

Elastic Multicast Routing Protocol
draft-adamson-elasticmcast-00

Abstract

This document describes an Internet Protocol (IP) multicast routing protocol suitable for dynamic network environments including mobile wireless. To handle high dynamics, this routing mechanism uses redundant forwarding, based upon the Simplified Multicast Forwarding (SMF) approach of [RFC6621], while converging to regular multicast distribution trees where or when the network becomes relatively stable. The rationale is that intermittent connectivity directly affects the ability of routers to synchronize on their view of the network, thus making it difficult to converge on efficient distribution trees, while network wide broadcast may be prohibitively expensive for relatively sparse groups. A hybrid approach, called Elastic Multicast, is specified which dynamically switches between limited scope broadcast and tree path forwarding independently at each node. The trees created during stable periods and portions of the network are pruned from the SMF efficient flooding mesh.

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 May 09, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 2
2. Terminology 4
3. Applicability Statement 5
4. Overview 5
5. Control Message Formats 6
5.1. Elastic Multicast Acknowledgment (EM-ACK) 7
5.2. Elastic Multicast Advertisement (EM-ADV) 7
6. Detailed Protocol Operation 8
6.1. Elastic Multicast Timers 11
7. SMF Relay Set Algorithm Considerations 11
8. IGMP and MLD Considerations 12
9. Border Gateway Considerations 13
10. Security Considerations 13
11. IANA Considerations 13
12. Acknowledgments 14
13. References 14
13.1. Normative References 14
13.2. Informative References 15
Authors' Addresses 16

1. Introduction

IP Multicast is an efficient way to distribute data to multiple receivers in networks. And, in wireless networks, the often broadcast nature of the communication media makes this even more advantageous and often important, given capacity limits. In relatively stable networks, multicast forwarding trees can be formed by protocols such as Protocol Independent Multicast (PIM) Dense Mode (DM)[RFC3973] or PIM Sparse Mode (SM)[RFC4601] to route and replicate packets from a source to multiple destinations, thus reducing the overhead of multiple point-to-point forwarding paths run in parallel,

one for each receiver of the multicast traffic. Building and maintaining multicast distribution trees requires some coordination between the network routers in order to avoid loops and to make sure that the dissemination paths are valid. In highly dynamic environments, the cost of maintaining the distribution trees can become prohibitive and the routers may be unable to synchronize their view of the network, leading to dropped traffic or transmission over invalid paths _(TBD-reference?)._ Additionally, these existing IP multicast routing protocols do not provide for the case in wireless networks where a received packet must be retransmitted via the same interface as which it was received to reach adjacent hosts or routers that were not in reception range of the previous-hop transmitter. This is a fundamental difference from wired network connectivity.

A practical solution that proves very effective in relatively small networks with dense groups is Simplified Multicast Forwarding (SMF) [RFC6621] which can efficiently flood IP multicast traffic to the entire network, regardless of group membership. The SMF approach can provide efficient flooding when a subset of well-connected nodes that form a Connected Dominating Set (CDS) is selected to broadcast the data and relay it to the entire network. The SMF specification describes distributed relay set selection algorithms that can from CDS with local (2-hop neighborhood) information that can be collected via the Neighborhood Discovery Protocol (NHDP) of [RFC6130]. SMF utilizes Duplicate Packet Detection (DPD) forwarding that can be applied to the wireless interface case as needed instead of the typical reverse path forwarding checks used by other IP multicast routing protocols. The approach specified herein, called Elastic Multicast[CH2012], builds off the group forwarding mesh established by SMF with distributed relay set selection and dynamically prunes the mesh as any network stability permits. This hybrid approach allows for highly efficient, group-based multicast datagram distribution when and where the network is relatively stable and provides for SMF-based flooding in times and network portions where unstable connectivity and high dynamics occur. The pruning is implemented more as a "grafting" process in that acknowledgments from neighboring forwarders keep relevant portions of the SMF forwarding mesh active. In dynamic environments, a richer portion of the forwarding mesh is kept active as the acknowledgments are

Starting with the SMF mesh and driven by IP multicast datagram transmissions, the Elastic Multicast protocol forms multicast distribution trees whenever possible to minimize overhead of flooding to unnecessary parts of the topology and adaptively re-expands the forwarding base to additional redundant nodes when needed, when and where these trees become unstable. The data-driven aspect is based on active presence of IP multicast traffic flows among the routers where the flows are timed out due to lack of transmission activity.

This mechanism of automatically expanding the forwarding base and then reducing it when not needed is based on acknowledgements in response to detected IP multicast traffic activity. The acknowledgements are aggregated and repeated at the intermediate forwarding nodes for active multicast traffic flows. The pruned SMF forwarding state is periodically timed out for active flows and flooding is reinstated. Thus, at a low duty cycle, active traffic flows are flooded (efficiently per the configured SMF algorithm) throughout the SMF domain to excite routers for Elastic Multicast acknowledgments. An OPTIONAL mechanism is described where periodic data traffic flooding may be economized by use of a control plane message that is periodically hop-by-hop disseminated throughout the network to advertise the aggregate set of active, but unacknowledged, IP multicast flows instead of the actual user traffic. The default flooding behavior of Elastic Multicast traffic is configurable on a per-flow (protocol, source, destination, traffic class) basis.

Elastic Multicast routing trades off pure efficiency in favor of robust datagram delivery in these types of network environments. However, it should be noted that in wireless network systems, the impact of Layer 2 Media Access Control (MAC) mechanisms and radio broadcast contention can extend multiple hops and, in limited scope wireless networks, SMF flooding with CDS relay sets is not necessarily much less efficient than group-specific forwarding in many cases, particularly in dynamic topologies. Since Elastic Multicast builds upon this SMF basis, it is expected to perform similarly and provide additional utility for cases of sparse group membership and stable connectivity.

2. Terminology

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

The terms introduced in [RFC5444], including "packet", "message", "TLV Block", "TLV", and "address" are to be interpreted as described therein.

The following abbreviations are used throughout this document:

Abbreviation	Definition
MANET	Mobile Ad hoc Network
SMF	Simplified Multicast Forwarding
CDS	Connected Dominating Set

NHDP	Neighborhood Discovery Protocol
DPD	Duplicate Packet Detection
FIB	Forwarding Information Base
TLV	type-length-value encoding

3. Applicability Statement

Within dynamic wireless routing topologies, maintaining traditional forwarding trees to support a multicast routing protocol is often not as effective as in wired networks due to the reduced reliability and increased dynamics of mesh topologies [MGL04] [GM99].

The SMF [RFC6621] specification provides for efficient forwarding of non-group-specific, IP multicast datagrams in wireless networks. This specification extends upon SMF to dynamically establish group-specific forwarding trees that are pared down from the base flooding mesh when and where network conditions of stability permit.

Elastic Multicast is compatible with different relay set selection algorithms and forwarding heuristics as described in the SMF specification[RFC6621] appendices. Elastic Multicast is also compatible with both the Any Source Multicast (ASM) [RFC1112]

and Single Source Multicast (SSM) [RFC4607] models of multicast group membership.

Elastic Multicast deployments are able to connect and interoperate with existing standard multicast protocols operating within more conventional Internet infrastructures. To this end, a multicast border router or proxy mechanism MUST be used when deployed alongside more fixed-infrastructure IP multicast routing such as the PIM variants [RFC3973] and [RFC4601].

This document does not presently support forwarding of packets with directed broadcast addresses as a destination [RFC2644].

4. Overview

The Elastic Multicast protocol extends upon the SMF specification by providing for controlled forwarding of group-specific multicast traffic flows. Multicast flows are identified by the source address, group destination address, and optionally, protocol, port, and/or traffic class information. By default, newly detected flows are flooded per SMF to the entire network, but with a preset token bucket limited traffic shaping. Elastic multicast acknowledgment messages (EM-ACK) for flows of interest sent in response to the initially

flooded traffic keep relevant portions of the SMF mesh "active" where token bucket limitations are lifted and other portions of the SMF mesh are "deactivated" providing only low rate forwarding until EM-ACK activity potentially reactivates it. As a more efficient alternative to gratuitous forwarding of the user data traffic, an Elastic Multicast advertisement message (EM-ADV) is specified that can allow for a "bundled" listing of active flow descriptions that is flooded within the routing area. The use of EM-ADV messages or low duty cycle flooding of the user traffic SHOULD be configurable on a per-flow basis. Initial forwarding/flooding of user traffic may be preferable depending upon the application and network use case.

Elastic Multicast monitors group membership of local hosts (collocated applications, directly attached hosts/devices, or neighboring non-routing hosts) to determine its transmission of EM-ACK messages and forwarding behaviors. When a local host is joined to a group, the Elastic Multicast router relieves any token bucket restrictions for the given group and provides periodic EM-ACK to neighboring routers upon receipt of applicable multicast traffic (or equivalent EM-ADV messages).

An Elastic Multicast router responds to IGMP joins and leaves issued by directly connected hosts. Based on such joins and leaves from connected hosts, an Elastic Multicast router may subscribe to different multicast groups. A router will be a subscriber for a multicast group as long as it has at least one directly connected host that joins that group.

The Elastic Multicast protocol may be deployed in a stand-alone MANET, or can be interfaced with PIM-DM enabled networks through Elastic Multicast gateways.

5. Control Message Formats

There is one principal and one optional control message used by the Elastic Multicast protocol. The first message is the Elastic Multicast Acknowledgment (EM-ACK) that is sent in response to received, non-duplicative, multicast user traffic from neighboring "upstream" routers. This message is directed to the given "upstream" router, informing it to make the flow "active" with respect to forwarding and similarly signal (via EM-ACK messages) its neighboring "upstream" routers that the flow is actively subscribed. The second, optional message is the Elastic Multicast Advertisement (EM-ADV) message that can be used as a surrogate instead of forwarding user traffic for unacknowledged flows. Multiple flow descriptions can be "bundled" into individual EM-ADV messages that can reduce the flooding overhead of advertising active multicast flows.

For both of these message types, standard MANET packet building block [RFC5444] formats will be defined. Note, as such, these messages may be sent independently or possibly opportunistically piggy-backed with other MANET link-local protocol messages such as NHDP [RFC6130].

5.1. Elastic Multicast Acknowledgment (EM-ACK)

The principal control message used in the Elastic Multicast protocol is acknowledgment message (EM-ACK) that is used to convey subscription interest in active multicast groups among routers in response to detected (transmitted) multicast traffic or the equivalent advertisement (EM-ADV) as described below. The EM-ACK message is directed to the specific neighboring router from which the local router received non-duplicative packets for groups of interest (i.e. joined by locally associate applications or hosts, or which are active due to directed EM-ACK messages received from other neighboring routers). Each EM-ACK message contains the following fields:

- o router-id: A unique, addressable identifier for the intended "upstream" router. For example, in a broadcast wireless medium it can be the source Ethernet address of the original data multicast packet that is acknowledged.

- o source-addr: The source IP address of the multicast packet being acknowledged

- o group-addr: The destination IP multicast address of the packet being acknowledged

A standard MANET packet building block [RFC5444] message format for the EM-ACK message will be specified in a future revision of this document.

5.2. Elastic Multicast Advertisement (EM-ADV)

The Elastic Multicast Advertisement (EM-ADV) message provides an alternative means to convey the set of active, but unacknowledged, traffic flows of which the router is aware. Information regarding multiple flows can be bundled into a single EM-ADV message, thus reducing the overhead of advertising active traffic flows for which EM-ACK responses might be generated. Note that the use of EM-ADV message should be optional and configurable on a per-flow basis as, for some use cases, it may be preferable to flood the user traffic directly instead. The concept of using the EM-ADV message to proactively establish Elastic Multicast routing state prior to traffic transmission is being explored and may be described in a future revision of this draft.

Each EM-ADV message consists of a list of active multicast flow descriptions, each with the following fields:

- o group-addr: IP multicast group destination address of active flow being advertised
- o Source address: IP source address of flow
- o Protocol: IP protocol id of flow
- o Traffic class: IPv4 TOS / IPv6 traffic class of flow
- o Token bucket parameters: bucket depth and rate
- o Tagger identifier (Tagger-ID): address of router originating this flow description
- o Duplicate packet detection identifier (DPD-ID): DPD-ID of packet triggering this flow description

The Tagger-ID and DPD-ID fields are set by the originator of the given flow description. The Tagger-ID is a unique router identifier of the Elastic Multicast router originating the EM-ADV flow description and the DPD-ID is the duplicate packet detection identifier for the packet that triggered the EM-ADV flow description generation. A standard MANET packet building block [RFC5444] message format for the EM-ADV message will be specified in a future revision of this document. Additionally, provisions for "gateway address" and /or "wildcard" address descriptors may be provided to potentially facilitate border gateway and routing area group membership collection in future revisions.

6. Detailed Protocol Operation

Elastic Multicast works as an extension of SMF. Building on the distributed relay set selection and robust, efficient flooding, the goal of Elastic Multicast is to keep multicast forwarding active only for the subset of SMF relays needed to reach multicast group subscribers. The mechanism described here provides for this relay set pruning in stable parts of the network while still relying on the SMF flooding behavior in the highly dynamic parts of the network where it is hard or impossible to determine which nodes are on the critical path to the destination(s). By eliminating from the set of forwarding nodes those that do not move multicast traffic closer to the subscriber destinations, the number of retransmissions, and consequently the total overhead of the protocol will decrease.

By default, the SMF protocol does not maintain or use group membership information. In its simplest form, called _classic flooding_ (CF), all SMF-enabled nodes rebroadcast all multicast packets, with the end result that all connected nodes get all multicast data. Duplicate Packet Detection (DPD), based on a unique packet identifiers and/or a hash of the packet content, is used to make sure that the same packet is not broadcasted more than once by a node, thus eliminating forwarding loops. In order to reduce overhead, SMF can use inputs from an external protocol, such as OSPF-MDR [RFC5614], OLSR [RFC3626], or NHDP [RFC6130] to enable forwarding by a selected Connected Dominating Set (CDS) of relays through which the entire MANET can be reached. This mechanism works well in relatively dense networks, significantly reducing the number of retransmissions. However, it still propagates multicast data throughout the entire MANET to all nodes, regardless of whether they subscribe to the specific multicast groups or not. Elastic Multicast works to leverage the distributed nature and robustness of SMF, but allow for improved efficiency by attempting to minimize the set of nodes forwarding traffic where it can for specific multicast groups.

To achieve this, Elastic Multicast treats low and high data rate flows of multicast traffic differently. Low data rate flows are flooded throughout the MANET area per SMF as usual. High data rate flows are limited to low data rate forwarding unless Elastic Multicast acknowledgment (EM-ACK) messages are received to "activate" high data rate forwarding of the given flow. Consistent acknowledgment under stable topology conditions to a consistent subset of the SMF relays serves to prune the SMF mesh towards a more efficient distribution tree for the flow while more dynamic acknowledgements to a varied (with topology changes) set of forwarders under less stable topologies serves to activate a larger, more redundant portion of the SMF mesh. The threshold that differentiates low and high data rate flows can be configurable. But, for an example, let’s consider that a low data rate threshold is set at 1 packet per second. Thus, flows of less than 1 packet per second are disseminated throughout the entire MANET, while the protocol attempts to prune the distribution only to subscriber nodes for flows of more than 1 packet per second.

Elastic Multicast routers only maintain a local, temporary subscription for the multicast groups for which they are supposed to rebroadcast data packets, and do not maintain any information about the global network topology or global node membership, thus sharing the scalability benefits and most of the simplicity of SMF. However, routers that use Elastic Multicast need to be aware of the multicast groups to which locally attached devices subscribe. The standard protocols used in practice today for multicast join and leave operation are IGMP [RFC3376] and MLD [RFC3810]. The Elastic

Multicast protocol does not change the functionality of the IGMP or MLD joins and leaves; it simply uses them to maintain the local membership of the attached devices or collocated processes. The considerations for IGMP and MLD operation are described further in IGMP (Section 8).

Elastic Multicast routers maintain a traffic shaping token bucket for each multicast data flow they observe. A multicast data flow is defined as a <sourceAddress:groupAddress> pair, although other parameters such as protocol, source/destination ports, or traffic class value may also be considered. The token bucket limits the rate at which multicast data for a certain flow is broadcasted throughout the network, in effect acting as a low rate enforcer per flow for SMF network broadcast. The bucket depth allows for some initial limited, gratuitous flooding of traffic for new flows and for bursty, but low average rate, traffic sources. When a flow is active, because of received local group membership or forwarding acknowledgments from adjacent Elastic Multicast routers, the token bucket enforcement is relaxed until the group membership is dropped and/or the acknowledgment activity for the flow has timed out.

The Elastic Multicast group-specific behavior is initiated by IP multicast group membership (e.g. IGMP join messages) by locally attached subscribers. For example, suppose an Elastic Multicast router becomes subscribed to a certain group as a result of one of its attached devices joining that group. Upon receiving a non-duplicative multicast data packet addressed to the subscribed group, the router sends an Elastic Multicast Acknowledgement (EM-ACK) to the upstream router from which it received the multicast packet. The EM-ACK contains the source IP and the destination group of the original multicast packet. We call an upstream router that receives such a EM-ACK for a flow a forwarding router for that flow. A forwarding router disables the token bucket for a preconfigured period of time, and for that time sends unicast EM-ACKs upstream when it receives multicast packets for that flow. The difference between a forwarding router for a flow and a subscribing router for a multicast group due to some of its locally attached devices joining that group is that the forwarding router only sends EM-ACKs for the flows for which it is receiving EM-ACKs, while the end subscribing router sends EM-ACKs for all flows addressed to that group, regardless of source of the flows. Note, in the case of Single Source Multicast (SSM), the group identifier includes the source address and thus MAY coincide with the flow identifier.

In a static network, trees of temporary multicast forwarders for higher rate traffic is formed from the sources to the subscribers of the groups through the mechanism described above. Forwarding routers send EM-ACKs either periodically or after a certain number of packets

for that flow was received, whichever occurs first, in order to refresh the forwarding state at upstream routers. If a forwarding router does not receive EM-ACKs for a multicast flow during a certain timeout, it simply re-enables the rate-limiting token bucket and no longer participates in the high rate forwarding or sending EM-ACKs upstream for that flow. It continues, however, to forward data at the lower rate, as limited by the token bucket. Other than EM-ACKs, no other control messages need to be used in the Elastic Multicast mechanism. However, the optional EM-ADV message MAY be used as a surrogate for low rate forwarding and a single EM-ADV can represent multiple traffic flows with a single message. Elastic Multicast routers use EM-ADV messaging should have its use configurable on a per-flow basis.

When the network topology becomes dynamic, or in parts of the network that become dynamic, more routers will become high load forwarders, as they receive EM-ACKs from downstream routers. Note that downstream routers send unicast EM-ACKs to the upstream router form which they received multicast data. If a forwarder router moves out of range, the next closest router will still forward data at a lower rate and will become activated when an EM-ACK is received. This mechanism expands the forwarding base and add redundancy in the dynamic parts of the network where more routers will be forwarding at a high rate, duplicating traffic, while in the stable parts of the network the protocol will form single paths or trees.

The Elastic Multicast protocol is similar to PIM Dense Mode (PIM-DM), in that it does not maintain topology information, it functions completely decentralized, and defaults in sending multicast data to all nodes in the network at a low rate. However, we believe that Elastic Multicast is more suitable to dynamic networks than PIM-DM mainly because Elastic Multicast routers do not depend on a unicast routing protocol for checking reverse paths in order to avoid loops, nor maintain point-to-point relationships with their neighbors. This allows them to broadcast multicast data, instead of streaming it point-to-point to downstream routers, and because of this the end to end packet delivery is less likely to be affected when network connectivity changes.

6.1. Elastic Multicast Timers

TBD - describe the timers that control EM-ACK transmission and per-flow high rate forwarding activation / deactivation.

7. SMF Relay Set Algorithm Considerations

SMF has the ability to select only some of the nodes to participate in the forwarding mechanism, such that in relatively dense areas, if

only a subset of the nodes act as relays, the data forwarded still reaches all nodes in the MANET. Such a subset of forwarder nodes is called a Connected Dominant Set (CDS). Several CDS selection algorithms are currently implemented, and in particular SMF has been proven to perform well with the source-based multipoint relay (S-MPR) and essential connected dominating set (ECDS) algorithms described in [RFC6621].

For CF and ECDS operation where the SMF forwarding rules are simple, the Elastic Multicast operation is straightforward where EM-ACK messages are directed to the previous hop forwarder whenever a non-duplicative multicast packet is received for a flow and its activation is due to be refreshed due to timeout. Other SMF relay set forwarding rules may require special consideration. For example, the forwarding rules of the S-MPR algorithm necessitate slightly different handling. In the case of S-MPR, the EM-ACK message should be sent to the upstream router only when the received packet is non-duplicative _and_ received from an MPR selector for the local router. It is expected that different relay set selection algorithms and selection criteria (e.g. metrics) will have impact on the utility of Elastic Multicast traffic for different application purposes. It is also expected that augmented relay set selection algorithms for efficient multi-interface operation will also impact the requirements for EM-ACK generation for correct protocol operation, but it is anticipated that Elastic Multicast can be adapted to these changes. Further experimentation and analysis should be conducted to determine the tradeoffs for network deployment and operations.

8. IGMP and MLD Considerations

The Elastic Multicast nodes act as regular multicast routers for local hosts subscribing to multicast groups. A local host can be:

1. an application residing on the same network node as the router,

2. a device attached directly to one of the non-MANET router interfaces, or

3. a separate neighboring host in the wireless network that does not have multicast routing capability but rather relies on the Elastic Multicast routers to forward their data appropriately.

In the case #1, the Elastic Multicast router may directly receive group membership join and leave notifications via operating system or implementation-specific mechanisms. In the case #2, the Elastic Multicast router should perform IGMP or MLD query functions per are IGMP [RFC3376] and MLD [RFC3810], respectively. For the case #3, standard IGMP or MLD operation is not sufficient for operation on

MANET interfaces since adjacencies on wireless interfaces are incongruent for different routers due to radio propagation. An alternative form of IGMP/MLD query and response might be implemented where all routers send queries without observing the Querier Election provisions and MANET hosts would select (e.g., based on querier address or link quality) the router to which to direct their responses (i.e., address the reply to a specific router). This issue needs further study and may be addressed in a later version of this draft or other documents.

In any case, once an Elastic Multicast router has determined local host subscription(s), it will provide EM-ACK messages for subscribed groups and perform multicast forwarding per this specification.

9. Border Gateway Considerations

Elastic Multicast can be compatible with other multicast routing protocols (e.g. PIM) if border gateway provisions are met. Given the similarities to PIM-DM, this allows a relatively straight forward interconnection with PIM-DM domains. Elastic Multicast gateways can forward multicast traffic received from their PIM-DM neighbors towards the MANET and send PIM join and leave messages to the upstream PIM routers based on the subscription of other routers in the MANET. Similarly, on the egress side Elastic Multicast gateways can act as subscribers to groups for which they receive PIM join messages from their downstream PIM routers. In the case where the Elastic Multicast routing area is a stub network, the adjacent gateway routing protocol membership information can be used to manage Elastic Multicast behavior in the same manner as host group membership for traffic that egresses the Elastic Multicast area. Group membership subscription needs to be aggregated from within the Elastic Multicast area and provided to the gateway router so that appropriate multicast traffic from external networks will be routed to the Elastic Multicast area. An approach to do this is described in [CDHM07]. Additionally, an extended approach that allows for multiple gateways is describe in [DHS08]. Further consider of border gateway operation, including cases where the Elastic Multicast area is a transit domain, may be addressed in future revisions of this draft.

10. Security Considerations

(TBD - We can refer to the SMF Security Considerations as applicable, the PacketBB-Sec document, etc. Some security for the EM-ACK messages should be provided via PacketBB-Sec)

11. IANA Considerations

(TBD - Call out any IANA assignments that need to be made. Most notably this will be the Elastic Multicast message types in the PacketBB message type namespace, a registry for any Elastic Multicast message TLVs, and those specific TLVs.)

12. Acknowledgments

(TBD)

Some core contributors/authors in alphabetical order:

Tom Henderson, et al

The RFC text was produced using Marshall Rose’s xml2rfc tool and Bill Fenner’s XMLmind add-ons.

13. References

13.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, RFC 1112, August 1989.

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

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC2644] Senie, D., "Changing the Default for Directed Broadcasts in Routers", BCP 34, RFC 2644, August 1999.

[RFC2780] Bradner, S. and V. Paxson, "IANA Allocation Guidelines For Values In the Internet Protocol and Related Headers", BCP 37, RFC 2780, March 2000.

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

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

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

[RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, August 2006.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized Mobile Ad Hoc Network (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.

[RFC5771] Cotton, M., Vegoda, L., and D. Meyer, "IANA Guidelines for IPv4 Multicast Address Assignments", BCP 51, RFC 5771, March 2010.

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

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

13.2. Informative References

[CDHM07] Chakeres, I., Danilov, C., and T. Henderson, "Connecting MANET Multicast", IEEE MILCOM 2007 Proceedings , 2007.

[CH2012] Danilov, C., Henderson, T., Brewer, O., Kim, J., Macker, J., and B. Adamson, "Elastic Multicast for Tactical Networks", IEEE MILCOM 2012 , October 2012.

[DHS08] Danilov, C., Henderson, T., and T. Spagnolo, "MANET Multicast with Multiple Gateways", IEEE MILCOM 2008 Proceedings , 2008.

[GM99] Garcia-Luna-Aceves, JJ. and E. Madruga, "The core-assisted mesh protocol", Selected Areas in Communications, IEEE Journal on Volume 17, Issue 8, August 1999.

[JLMV02] Jacquet, P., Laouiti, V., Minet, P., and L. Viennot, "Performance of multipoint relaying in ad hoc mobile routing protocols", Networking , 2002.

[MDC04] Macker, J., Dean, J., and W. Chao, "Simplified Multicast Forwarding in Mobile Ad hoc Networks", IEEE MILCOM 2004 Proceedings , 2004.

[MDDA07] Macker, J., Downard, I., Dean, J., and R. Adamson, "Evaluation of distributed cover set algorithms in mobile ad hoc network for simplified multicast forwarding", ACM SIGMOBILE Mobile Computing and Communications Review Volume 11 , Issue 3, July 2007.

[MGL04] Mohapatra, P., Gui, C., and J. Li, "Group Communications in Mobile Ad hoc Networks", IEEE Computer Vol. 37, No. 2, February 2004.

[RFC2501] Corson, M. and J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, January 1999.

[RFC3973] Adams, A., Nicholas, J., and W. Siadak, "Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)", RFC 3973, January 2005.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.

Authors' Addresses

Brian Adamson
Naval Research Laboratory
Washington, DC 20375
USA

Email: brian.adamson@nrl.navy.mil

Claudiu Danilov
Boeing Corporation
Seattle, WA
USA

Email: Claudiu.B.Danilov@boeing.com

Joseph Macker
Naval Research Laboratory
Washington, DC 20375
USA

Email: joseph.macker@nrl.navy.mil

Mobile Ad hoc Networking (MANET)
Internet-Draft
Intended status: Experimental
Expires: June 23, 2014

C. Dearlove
BAE Systems ATC
T. Clausen
LIX, Ecole Polytechnique

December 20, 2013

Multi-Topology Extension for the Optimized Link State Routing Protocol version 2 (OLSRv2)
draft-dearlove-manet-olsrv2-multitopology-02

Abstract

This specification describes an extension to the Optimized Link State Routing Protocol version 2 (OLSRv2) to support multiple routing topologies, while retaining interoperability with OLSRv2 routers that do not implement this extension.

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 June 23, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Terminology and Notation 3
3. Applicability Statement 4
4. Protocol Overview and Functioning 4
5. Parameters 5
6. Information Bases 5
6.1. Local Attached Network Set 6
6.2. Link Sets 6
6.3. 2-Hop Sets 6
6.4. Neighbor Set 6
6.5. Router Topology Set 7
6.6. Routable Address Topology Set 7
6.7. Attached Network Set 7
6.8. Routing Sets 7
7. TLVs 8
7.1. Message TLVs 8
7.1.1. MPR_TYPES TLV 8
7.1.2. MPR_WILLING TLV 8
7.2. Address Block TLVs 9
7.2.1. LINK_METRIC TLV 9
7.2.2. MPR TLV 9
7.2.3. GATEWAY TLV 10
8. HELLO Messages 11
8.1. HELLO Message Generation 11
8.2. HELLO Message Processing 11
9. TC Messages 12
9.1. TC Message Generation 12
9.2. TC Message Processing 12
10. MPR Calculation 12
11. Routing Set Calculation 13
12. Management Considerations 13
13. IANA Considerations 14
13.1. Expert Review: Evaluation Guidelines 14
13.2. Message TLV Types 14
13.3. Address Block TLV Types 16
14. Security Considerations 16
15. Acknowledgments 16
16. References 17
16.1. Normative References 17
16.2. Informative References 17
Authors’ Addresses 17

1. Introduction

The Optimized Link State Routing Protocol, version 2 [OLSRv2] is a proactive link state routing protocol designed for use in mobile ad hoc networks (MANETs) [RFC2501]. One of the significant improvements of OLSRv2 over its Experimental precursor [RFC3626] is the ability of OLSRv2 to route over other than minimum hop routes, using a link metric.

A limitation that remains in OLSRv2 is that it uses a single link metric type for all routes. However in some MANETs it would be desirable to be able to use alternative metrics for different packet routing. This specification describes an extension to OLSRv2, that is designed to permit this, while maintaining maximal interoperability with OLSRv2 routers not implementing this extension.

The purpose of OLSRv2 can be described as to create and maintain a Routing Set, which contains all the necessary information to populate an IP routing table. In a similar way, the role of this extension can be described as to create and maintain multiple Routing Sets, one for each link metric type supported by the router maintaining the sets.

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 specification uses the terminology of [RFC5444], [RFC6130] and [OLSRv2], which is to be interpreted as described in those specifications.

Additionally, this specification uses the following terminology:

Router - A MANET router that implements [OLSRv2].

MT-OLSRv2 - The protocol defined in this specification as an extension to [OLSRv2].

This specification introduces the notation map[range -> type] to represent an associative map from elements of the range, which in this specification is always a set of link metric types that the router supports (either IFACE_METRIC_TYPES or ROUTER_METRIC_TYPES, as defined in Section 5), to a type, which may be a boolean, a willingness (a 4 bit unsigned integer from 0 to 15), a number of hops

(an 8 bit unsigned integer value from 0 to 255), or a metric-value (either a representable link metric value, as described in [OLSRv2], or UNKNOWN_METRIC).

3. Applicability Statement

The protocol described in this specification is applicable to a MANET for which OLSRv2 is otherwise applicable (see [OLSRv2]), but in which multiple topologies are maintained, each characterized by a different choice of link metric type. It is assumed, but outside the scope of this specification, that the network layer is able to choose which topology to use for each packet, for example using the DiffServ Code Point (DSCP) defined in [RFC2474].

4. Protocol Overview and Functioning

The purpose of this specification is to extend [OLSRv2] so as to enable a router to establish and maintain multiple routing topologies in a MANET, each topology associated with a link metric type. Routers in the MANET may each form part of some or all of these topologies, and each router will maintain a Routing Set for each topology that it forms part of, allowing separate routing of packets for each topology.

Each router implementing this specification selects a set of link metric types for each of its OLSRv2 interfaces. If all routers in the MANET implement MT-OLSRv2, then there are no restrictions on how these sets of link metrics are selected. However there may be deployments where routers, that do not implement MT-OLSRv2 (non-MT-OLSRv2 routers), are to participate in a MANET with MT-OLSRv2 routers. In this case, the single link metric used by these non-MT-OLSRv2 routers must be included in the set of link metrics for each OLSRv2 interface of an MT-OLSRv2 router thatEmay be heard on an OLSRv2 interface of a non-MT-OLSRv2 router in the MANET.

Each router then determines an incoming link metric for each link metric type selected for each of its OLSRv2 interfaces. These link metrics are distributed using link metric TLVs contained in all HELLO messages sent on OLSRv2 interfaces, and in all TC messages.

In addition to link and neighbor metric values for each link metric type, router MPR (multipoint relay) and MPR selector status, and advertised neighbor status, is maintained per supported neighbor metric type for each symmetric 1-hop neighbor. Each router may choose a different willingness to be a routing MPR for each link metric type that it supports.

More so than OLSRv2, the use of multiple metric types across the MANET must be managed, by administrative configuration or otherwise. Similarly to other decisions that may be made using OLSRv2, a bad collective choice will make the MANET anywhere from inefficient to non-functional, so care will be needed in selecting supported link metric types across the MANET.

5. Parameters

The parameters used in [OLSRv2], including from its normative references, are used in this specification with the following changes.

Each OLSRv2 interface will support a number of link metric types, corresponding to Type Extensions of the LINK_METRIC TLV defined in [OLSRv2]. The router parameter LINK_METRIC_TYPE, used by routers that do not implement MT-OLSRv2, and used with that definition in this specification, is replaced in routers implementing MT-OLSRv2 by an interface parameter array IFACE_METRIC_TYPES and a router parameter array ROUTER_METRIC_TYPES. Each element in these arrays is a link metric type (i.e., a type extension used by the LINK_METRIC TLV [OLSRv2]).

The interface parameter array IFACE_METRIC_TYPES contains the link metric types supported on that OLSRv2 interface. The router parameter array ROUTER_METRIC_TYPES is the union of all of the IFACE_METRIC_TYPES. Both arrays MUST be without repetitions.

If in a given deployment there may be any routers that do not implement MT-OLSRv2, then IFACE_METRIC_TYPES MUST include LINK_METRIC_TYPE if that OLSRv2 interface may be able to communicate with any routers that do not implement MT-OLSRv2. In that case, ROUTER_METRIC_TYPES MUST also include LINK_METRIC_TYPE.

In addition, the router parameter WILL_ROUTING is extended to an array of values, one each for each link metric type in the router parameter list ROUTER_METRIC_TYPES.

6. Information Bases

The Information Bases specified in [OLSRv2], which extend those specified in in [RFC6130], are further extended in this specification. With the exception of the Routing Set, the extensions in this specification are the replacement of single values (boolean, willingness, number of hops, or link-metric) from [OLSRv2] with elements representing multiple values (associative maps from a set of

metric types to their corresponding values). The following subsections detail these extensions.

Note that, as in [OLSRv2], an implementation is free to organize its internal data in any manner it chooses, it needs only to behave as if it were organized as described in [OLSRv2] and this specification.

6.1. Local Attached Network Set

Each element AL_dist becomes a map[ROUTER_METRIC_TYPES -> number of hops].

Each element AL_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

6.2. Link Sets

Each element L_in_metric becomes a map[IFACE_METRIC_TYPES -> link-metric].

Each element L_out_metric becomes a map[IFACE_METRIC_TYPES -> link-metric].

The elements of L_in_metric MUST be set following the same rules that apply to the setting of the single element L_in_metric in [OLSRv2].

6.3. 2-Hop Sets

Each element N2_in_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

Each element N2_out_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

6.4. Neighbor Set

Each element N_in_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

Each element N_out_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

Each element N_will_routing becomes a map[ROUTER_METRIC_TYPES -> willingness].

Each element N_routing_mpr becomes a map[ROUTER_METRIC_TYPES -> boolean].

Each element N_mpr_selector becomes a map[ROUTER_METRIC_TYPES -> boolean].

Each element N_advertised becomes a map[ROUTER_METRIC_TYPES -> boolean].

6.5. Router Topology Set

Each element TR_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

Note that some values of TR_metric may now take the value UNKNOWN_METRIC. When used to construct a Routing Set, where just the corresponding value from this map is used, Router Topology Tuples whose corresponding value of TR_metric is UNKNOWN_METRIC are ignored.

6.6. Routable Address Topology Set

Each element TA_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

Note that some values of TA_metric may now take the value UNKNOWN_METRIC. When used to construct a Routing Set, where just the corresponding value from this map is used, Routable Address Topology Tuples whose corresponding value of TA_metric is UNKNOWN_METRIC are ignored.

6.7. Attached Network Set

Each element AN_dist becomes a map[ROUTER_METRIC_TYPES -> number of hops].

Each element AN_metric becomes a map[ROUTER_METRIC_TYPES -> link-metric].

Note that some values of AN_metric may now take the value UNKNOWN_METRIC. When used to construct a Routing Set, where just the corresponding value from this map is used, Attached Network Tuples whose corresponding value of AN_metric is UNKNOWN_METRIC are ignored.

6.8. Routing Sets

There is a separate Routing Set for each link metric type in ROUTER_METRIC_TYPES.

7. TLVs

This specification makes the following additions and extensions to the TLVs defined in [OLSRv2].

7.1. Message TLVs

One new Message TLV is defined in this specification, and one existing Message TLV is extended by this specification.

7.1.1. MPR_TYPES TLV

The MPR_TYPES TLV is used in HELLO messages, and may be used in TC messages. A message MUST NOT contain more than one MPR_TYPES TLV.

The presence of this TLV in a HELLO message is used to indicate that the router supports MT-OLSRv2, in the same way that the presence of the MPR_WILLING TLV is used to indicate that the router supports OLSRv2, as specified in [OLSRv2]. For this reason, the MPR_TYPES TLV has been defined with the same Type as the MPR_WILLING TLV, but with Type Extension == 1. (The different symbolic name is used for convenience, any reference to a MPR_TYPES TLV means to this TLV, with this Type and Type Extension.)

This TLV may take a Value field of any size. Each octet in its Value field will contain a link metric type that is supported for the OLSRv2 interface over which the HELLO message containing this TLV is sent. These octets MAY be in any order, except that if there may be any routers in the MANET not implementing MT-OLSRv2, then the first octet MUST be LINK_METRIC_TYPE.

7.1.2. MPR_WILLING TLV

The MPR_WILLING TLV, which is used in HELLO messages, is specified in [OLSRv2], and extended in this specification as enabled by [TLV-Extensions].

The interpretation of this TLV, specified by [OLSRv2], and which uses all of its single octet Value field, is unchanged. That interpretation uses bits 0-3 of its Value field to specify its willingness to be a flooding TLV, and bits 4-7 of its Value field to be a routing TLV. Those latter bits are, when using this specification, interpreted as its willingness to be a routing TLV using the link metric type LINK_METRIC_TYPE.

The extended use of this message TLV, as defined by this specification, defines additional 4 bit sub-fields of the Value field, starting with bits 4-7 of the first octet and continuing with

bits 0-3 of the second octet, to represent willingness to be a routing MPR using the link metric types specified in this OLSRv2 interface's IFACE_METRIC_TYPES parameter, ordered as reported in the included MPR_TYPES Message TLV. (If there is no such TLV included, then the router does not support MT-OLSRv2, and only the first octet of the Value field will be used.)

If the number of link metric types in this OLSRv2 interface's IFACE_METRIC_TYPES parameter is even, then there will be an unused 4 bit sub-field in bits 4-7 of the last octet of a full sized Value field. These bits will not be used, they SHOULD all be cleared ('0').

If the Value field in an MPR_WILLING TLV is shorter than its full length, then, as specified in [TLV-Extensions], missing Value octets, i.e., missing willingness values, are considered as zero, i.e., as WILL_NEVER. This is the correct behaviour. (In particular it means that an OLSRv2 router that is not implementing MT-OLSRv2 will not act as a routing MPR for any link metric that it does not recognise.)

7.2. Address Block TLVs

New Type Extensions are defined for the LINK_METRIC TLV defined in [OLSRv2], and the Value fields of the MPR TLV and the GATEWAY TLV, both defined in [OLSRv2], are extended, as enabled by [TLV-Extensions].

7.2.1. LINK_METRIC TLV

The LINK_METRIC TLV is used in HELLO messages and TC messages. This TLV is unchanged from the definition in [OLSRv2].

Only a single Type Extension was specified by [OLSRv2] (link metric type) 0 as defined by administrative action. This specification extends this range, it is suggested either to 0-7 or to 0-15. This specification will work with any combination of Type Extensions both within and without that range (assuming that the latter are defined as specified in [OLSRv2]).

7.2.2. MPR TLV

The MPR TLV is used in HELLO messages, and indicates that an address with which it is associated is of a symmetric 1-hop neighbor that has been selected as an MPR.

The Value field of this address block TLV is, in [OLSRv2], defined to be one octet long, with the values 1, 2 and 3 defined. [TLV-Extensions] redefines this Value field to be a bitfield where

bit 7 (the lsb) denotes flooding status, bit 6 denotes routing MPR status, and bits 5-0 are unallocated (respecting the semantics of the bits/values 1, 2 and 3 from [OLSRv2]).

This specification, as enabled by [TLV-Extensions], extends the MPR TLV to have a variable-length Value field. For interoperability with a router not implementing MT-OLSRv2, the two least significant bits of the first octet in the Value field of this TLV MUST be the TLV Value of the MPR TLV, generated according to [OLSRv2].

Subsequent bits (in increasing significance within an octet, then continuing with the least significant bit in the next octet, if required) in the TLV Value field indicate which link metric types, for which the corresponding address is selected as a routing MPR, link metric types (including the first) being indicated in the Value field of an MPR_TYPES Message TLV.

7.2.3. GATEWAY TLV

The GATEWAY TLV is used in TC messages to indicate that a network address is of an attached network.

The Value field of this address block TLV is, in [OLSRv2] defined to be one octet long, containing the number of hops to that attached network.

This specification, as enabled by [TLV-Extensions], allows the extension the GATEWAY TLV to have a variable-length Value field when the number of hops to each attached network is different for different link metric types. For interoperability with a router not implementing MT-OLSRv2, the first octet in the Value field of this TLV MUST be the TLV Value of the GATEWAY TLV generated according to [OLSRv2].

Any subsequent octets in the TLV Value field indicate the number of hops to the attached network for each other link metric type, link metric types (including the first) being indicated in the Value field of an MPR_TYPES Message TLV.

Type	Value
GATEWAY	Number of hops to attached network for each link metric type.

Table 1: GATEWAY TLV definition

8. HELLO Messages

The following changes are made to the generation and processing of HELLO messages compared to that described in [OLSRv2] by routers that implement MT-OLSRv2.

8.1. HELLO Message Generation

A generated HELLO message to be sent on an OLSRv2 interface is extended by:

- o Adding an MPR_TYPES TLV. The value octets will be the link metric types in IFACE_METRIC_TYPES.

- o Extending the MPR_WILLING TLV Value field to report the willingness values from the WILL_ROUTING parameter list that correspond to the link metric types in IFACE_METRIC_LIST, in the same order as reported in the MPR_TYPES TLV, each value (also including one representing WILL_FLOODING) occupying 4 bits.

- o Including LINK_METRIC TLVs that report all values of L_in_metric, L_out_metric, N_in_metric and N_out_metric that are not equal to UNKNOWN_METRIC, with the TLV Type Extension being the link metric type, and otherwise following the rules for such inclusions specified in [OLSRv2].

- o Including MPR TLVs such that for each link metric type in IFACE_METRIC_TYPES, and for the choice of flooding MPRs, these MUST be an MPR set as specified for a single link metric type in [OLSRv2].

8.2. HELLO Message Processing

On receipt of a HELLO message, a router implementing MT-OLSRv2 MUST, in addition to the processing described in [OLSRv2]:

1. Determine the list of link metric types supported by the sending router on the relevant OLSRv2 interface, either from an MPR_TYPES TLV or, if not present, the type LINK_METRIC_TYPE supported by a router not implementing the extension described in this specification.

2. For those link metric types supported by both routers, set the appropriate L_out_metric, N_in_metric, N_out_metric, N_will_routing, N_mpr_selector, N_advertised, N2_in_metric and N2_out_metric values as described for the single such elements in [OLSRv2].

3. For any other metric types supported by the receiving router only, set those elements to their default value: UNKNOWN_METRIC, WILL_NEVER (not WILL_DEFAULT), or false.

9. TC Messages

The following changes are made to the generation and processing of TC messages compared to that described in [OLSRv2] by routers that implement MT-OLSRv2.

9.1. TC Message Generation

A generated TC message is extended by:

- If any GATEWAY TLVs are included requiring more than one number of hops value, then adding an MPR_TYPES TLV with Value octets being the link metric types in ROUTER_METRIC_TYPES.

- Including LINK_METRIC TLVs that report all values of N_out_metric that are not equal to UNKNOWN_METRIC, with the TLV Type Extension being the link metric type, and otherwise following the rules for such inclusions specified in [OLSRv2].

- When not all the same, including a number of hops per reported (in an MPR_TYPES Messsage TLV) link metric type in the Value field of each GATEWAY TLV included.

9.2. TC Message Processing

On receipt of a TC message, a router implementing this extension MUST, in addition to the processing specified in [OLSRv2]:

- Set the appropriate TR_metric, TA_metric, AN_dist and AN_metric elements using the rules for setting the single elements of those types specified in [OLSRv2].

- For any other metric types supported by the receiving router that do not have an advertised outgoing neighbor metric of that type, set the corresponding elements of TR_metric, TA_metric and AN_metric to UNKNOWN_METRIC. (The corresponding element of AN_dist may be set to any value.)

10. MPR Calculation

Routing MPRs are calculated for each link metric type in ROUTER_METRIC_TYPES. Links to symmetric 1-hop neighbors via OLSRv2

interfaces that do not support that link metric type are not considered. The determined status (routing MPR or not routing MPR) for each link metric type is recorded in the relevant element of N_routing_mpr.

Each router may make its own decision as to whether or not to use a link metric, or link metrics, for flooding MPR calculation, and if so which and how. This decision MUST be made in a manner that ensures that flooded messages will reach the same symmetric 2-hop neighbors as would be the case for a router not supporting MT-OLSRv2.

Note that it is possible that a 2-Hop Tuple in the Information Base for a given OLSRv2 interface does not support any of the link metric types that are in the router’s corresponding IFACE_METRIC_TYPES, but nevertheless that 2-Hop Tuple MUST be considered when determining flooding MPRs.

11. Routing Set Calculation

A Routing Set is calculated for each link metric type in ROUTER_METRIC_TYPES. The calculation may be as for [OLSRv2], except that where an element is now represented by a map, the value from the map for the selected link metric type is used. Where this is a link metric of value UNKNOWN_METRIC, that protocol Tuple is ignored for the calculation.

12. Management Considerations

MT-OLSRv2 may require greater management than unextended OLSRv2. In particular MT-OLSRv2 requires the following management considerations:

- o Selecting which link metrics to support on each OLSRv2 interface and implementing that decision. (Different interfaces may have different physical and data link layer properties, and this may inform the selection of link metrics to support, and their values.)

- o Ensuring that the MANET is sufficiently connected. Note that if there is any possibility that there are any routers not implementing MT-OLSRv2, then the MANET will be connected, to the maximum extent possible, using the link metric type LINK_METRIC_TYPE.

- o Deciding which link metric, and hence which Routing Set to use, for received packets, hence how to use the Routing Sets to

configure the network layer (IP). An obvious approach is to map each DiffServ Code Point (DSCP) [RFC2474] to a single link metric. (This may be a many to one mapping.)

- o Note that there could be cases where a router that is not implementing MT-OLSRv2 is the source or destination of an IP packet that is mapped to a link metric that is not the link metric LINK_METRIC_TYPE used by that router.

- o If such a router is the source, then routing may work if the first router implementing MT-OLSRv2 to receive the packet supports the appropriate link metric type. At worst the packet will be dropped, it will not loop.

- o If such a router is the destination, then the packet will never reach its destination, as the source will not have a suitable routing table entry for the destination. Network management may be required to ensure that the MANET still functions in these cases.

13. IANA Considerations

This specification adds one new Message TLV, allocated as a new Type Extension to an existing Message TLV, using a new name. It also modifies the Value field of an existing Message TLV, and of an existing Address Block TLV.

13.1. Expert Review: Evaluation Guidelines

For the registry where an Expert Review is required, the designated expert SHOULD take the same general recommendations into consideration as are specified by [RFC5444].

13.2. Message TLV Types

This specification replaces Table 11 of [OLSRv2]. That specified a Message MPR Type described as MPR_WILLING, for which only Type Extension 0 was defined. This specification reserves that name MPR_WILLING for Type Extension 0, defines a new Type Extension 1, with a new name MPR_TYPES, and leaves the remaining Type Extensions of this TLV Type unnamed. It also changes the Value field specification of the MPR_WILLING TLV.

Note: The Type number TBD2 will be replaced by the value assigned by IANA when [OLSRv2] is published as an RFC, and this note will be removed.

Specifications of these TLVs are in Table 2. Each of these TLVs MUST NOT be included more than once in a Message TLV Block.

Name	Type	Type Extension	Description	Allocation Policy
MPR_WILLING	TBD2	0	Bits 0-3 specify the originating router's willingness to act as a flooding MPR. Each following 4 bit subfield (using bits 0-3 of an octet before bits 4-7) specifies the originating router's willingness to act as a routing MPR for a link metric, either a single such field (bits 4-7) when no MPR_TYPES Message TLV is present, or one subfield per type reported in an MPR_TYPES Message TLV Value field (in the same order).	
MPR_TYPES	TBD2	1	The link metric types supported on this OLSRv2 interface of this router (one octet each).	
Unnamed	TBD2	2-255	Unassigned.	Expert Review

Table 2: Message TLV Type assignment: MPR_WILLING and MPR_TYPES

13.3. Address Block TLV Types

Table 16 of [OLSRv2] is replaced by Table 3. Note that the only change is to the description of the Value field.

Note: The Type number TBD7 will be replaced by the value assigned by IANA when [OLSRv2] is published as an RFC, and this note will be removed.

Name	Type	Type extension	Description	Allocation Policy
GATEWAY	TBD7	0	Specifies that a given network address is reached via a gateway on the originating router. The number of hops is indicated by the Value field, either using a single octet (if no MPR_TYPES Message TLV is present) or one octet per type reported in an MPR_TYPES Message TLV (in the same order).	
GATEWAY	TBD7	1-255		Expert Review

Table 3: Address Block TLV Type assignment: GATEWAY

14. Security Considerations

TBD.

15. Acknowledgments

TBD.

16. References

16.1. Normative References

[OLSRv2] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol version 2", work in progress draft-ietf-manet-olsrv2-19, March 2013.

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

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

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

[TLV-Extensions] Dearlove, C. and T. Clausen, "Optimized Link State Routing Protocol version 2 (OLSRv2) and MANET Neighborhood Disovery Protocol (NHDP) Extension TLVs", work in progress draft-ietf-manet-nhdp-olsrv2-tlv-extensions-00, September 2013.

16.2. Informative References

[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, December 1998.

[RFC2501] Macker, J. and S. Corson, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, January 1999.

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

Authors' Addresses

Christopher Dearlove
BAE Systems Advanced Technology Centre
West Hanningfield Road
Great Baddow, Chelmsford
United Kingdom

Phone: +44 1245 242194
Email: chris.dearlove@baesystems.com
URI: http://www.baesystems.com/

Thomas Heide Clausen
LIX, Ecole Polytechnique

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

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.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Overview 4
2. Terminology 5
3. Applicability Statement 9
4. Purpose of the Experiment 11
5. Data Structures 12
5.1. InterfaceSet 12
5.2. Router Client Set 12
5.3. Neighbor Set 13
5.4. Sequence Numbers 14
5.5. Local Route Set 15
5.6. Multicast Route Message Set 17
5.7. Route Error (RERR) Set 19
6. Metrics 19
7. AODVv2 Protocol Operations 21
7.1. Initialization 21
7.2. Next Hop Monitoring 21
7.3. Neighbor Set Update 23
7.4. Interaction with the Forwarding Plane 24
7.5. Message Transmission 26
7.6. Route Discovery, Retries and Buffering 27
7.7. Processing Received Route Information 28
7.7.1. Evaluating Route Information 29
7.7.2. Applying Route Updates 30
7.8. Suppressing Redundant Messages Using the Multicast Route Message Set 33
7.9. Suppressing Redundant Route Error Messages using the Route Error Set 35
7.10. Local Route Set Maintenance 35
7.10.1. LocalRoute State Changes 35
7.10.2. Reporting Invalid Routes 38
8. AODVv2 Protocol Messages 38
8.1. Route Request (RREQ) Message 38
8.1.1. RREQ Generation 40
8.1.2. RREQ Reception 41
8.1.3. RREQ Forwarding 42
8.2. Route Reply (RREP) Message 42
8.2.1. RREP Generation 43
8.2.2. RREP Reception 45
8.2.3. RREP Forwarding 46

8.3. Route Reply Acknowledgement (RREP_Ack) Message 47
8.3.1. RREP_Ack Request Generation 47
8.3.2. RREP_Ack Reception 48
8.3.3. RREP_Ack Response Generation 49
8.4. Route Error (RERR) Message 49
8.4.1. RERR Generation 50
8.4.2. RERR Reception 51
8.4.3. RERR Regeneration 53
9. RFC 5444 Representation 53
9.1. Route Request Message Representation 54
9.1.1. Message Header 55
9.1.2. Message TLV Block 55
9.1.3. Address Block 55
9.1.4. Address Block TLV Block 55
9.2. Route Reply Message Representation 56
9.2.1. Message Header 56
9.2.2. Message TLV Block 56
9.2.3. Address Block 57
9.2.4. Address Block TLV Block 57
9.3. Route Reply Acknowledgement Message Representation . . . 58
9.3.1. Message Header 58
9.3.2. Message TLV Block 58
9.3.3. Address Block 58
9.3.4. Address Block TLV Block 58
9.4. Route Error Message Representation 58
9.4.1. Message Header 58
9.4.2. Message TLV Block 59
9.4.3. Address Block 59
9.4.4. Address Block TLV Block 59
10. Simple External Network Attachment 60
11. Configuration 62
11.1. Timers 62
11.2. Protocol Constants 64
11.3. Local Settings 65
11.4. Network-Wide Settings 65
11.5. MetricType Allocation 66
11.6. RFC 5444 Message Type Allocation 66
11.7. RFC 5444 Message TLV Types 66
11.8. RFC 5444 Address Block TLV Type Allocation 67
11.9. ADDRESS_TYPE TLV Values 67
12. IANA Considerations 68
13. Security Considerations 68
13.1. Availability 68
13.1.1. Denial of Service 68
13.1.2. Malicious RERR messages 69
13.1.3. False Confirmation of Link Bidirectionality 70
13.1.4. Message Deletion 71
13.2. Confidentiality 71

13.3. Integrity 71
13.3.1. Message Insertion 71
13.3.2. Message Modification - Man in the Middle 72
13.3.3. Replay Attacks 73
13.4. Protection Mechanisms 73
13.4.1. Confidentiality and Authentication 73
13.4.2. Integrity and Trust using ICVs 73
13.4.3. Replay Protection using Timestamps 73
13.4.4. Application to AODVv2 74
13.5. Key Management 79
14. Acknowledgments 81
15. References 81
15.1. Normative References 81
15.2. Informative References 82
Appendix A. AODVv2 Draft Updates 83
Authors' Addresses 83

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

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

- o 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:

- o 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.

- o 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.

- o 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.

- o 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.

- o 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.

- o 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:

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

- o 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.

- o 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.

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

- o 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:

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

- o 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:

- o Process information in a received RREQ or RREP message to learn a route to the originator or target of that route discovery

- o Forward a received RREQ or RREP

- o Send an RREP_Ack

- o Maintain valid routes in the Local Route Set

- o 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.

- o 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.

- o 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:

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

- o 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:

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

- o Neighbor.IPAddress := Source IP address of the received route message
- o Neighbor.State := Heard
- o Neighbor.Timeout := INFINITY_TIME
- o 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:

- o Neighbor.Timeout := CurrentTime + RREP_Ack_SENT_TIMEOUT

When a received message is one of the following:

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

- o 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:

- o Neighbor.State := Confirmed

- o 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:

- o Neighbor.State := Blacklisted

- o 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:

- o 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:

- o 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.

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

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

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

AODVv2 needs to send signals to the forwarding plane:

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

- o 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).

- o 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.

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

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

- o 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)

- o 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)

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

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

- o AdvRte.MetricType := RteMsg.MetricType

- o AdvRte.Metric := RteMsg.Metric

- o 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:

o RteMsg.Timestamp := CurrentTime

o 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:

o RteMsg.OrigPrefix := OrigPrefix from the message

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

o RteMsg.TargPrefix := TargPrefix from the message

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

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

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

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

o RteMsg.Timestamp := CurrentTime

o 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:

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

- o 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.

- o 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.

- o 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:

- o 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.

- o 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:

o 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.

o 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.

o 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.

o 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:

o 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.

o 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.

o 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.

o 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:

```
+-----------------------------------------------------------------+
|                          msg_hop_limit                          |
+-----------------------------------------------------------------+
|                           AddressList                           |
+-----------------------------------------------------------------+
|                   PrefixLengthList (optional)                   |
+-----------------------------------------------------------------+
|                OrigSeqNum, (optional) TargSeqNum                |
+-----------------------------------------------------------------+
|                           MetricType                            |
+-----------------------------------------------------------------+
|                           OrigMetric                            |
+-----------------------------------------------------------------+
```

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:

```
+-----------------------------------------------------------------+
|                          msg_hop_limit                          |
+-----------------------------------------------------------------+
|                           AddressList                           |
+-----------------------------------------------------------------+
|                   PrefixLengthList (optional)                   |
+-----------------------------------------------------------------+
|                           TargSeqNum                            |
+-----------------------------------------------------------------+
|                           MetricType                            |
+-----------------------------------------------------------------+
|                           TargMetric                            |
+-----------------------------------------------------------------+
```

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:

o RteMsg.OrigPrefix == RREP.OrigPrefix

o RteMsg.OrigPrefixLen == RREP.OrigPrefixLen

o RteMsg.TargAddr exists within RREP.TargPrefix

o RteMsg.OrigSeqNum <= RREP.OrigSeqNum

o RteMsg.MetricType == RREP.MetricType

o RteMsg.Timestamp > CurrentTime - RREQ_WAIT_TIME

o 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.

```
+-------------------------------------------------------------------+
|                        AckReq (optional)                          |
+-------------------------------------------------------------------+
```

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:

```
+-----------------------------------------------------------------+
|                       PktSource (optional)                      |
+-----------------------------------------------------------------+
|                           AddressList                           |
+-----------------------------------------------------------------+
|                   PrefixLengthList (optional)                   |
+-----------------------------------------------------------------+
|                       SeqNumList (optional)                     |
+-----------------------------------------------------------------+
|                          MetricTypeList                         |
+-----------------------------------------------------------------+
```

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:

- o 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.

- o 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.

- o 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.

```
       /-------------------------\
      / +----------------+        \
     /  |  AODVv2 Router |         \
     |  |  191.0.2.2/32  |         |
     |  +----------------+         |            Routable
     |                       +-----+--------+   Prefix
     |                       |     ENAR     |  /191.0.2.0/24
     |                       | AODVv2 Router| /
     |                       |  191.0.2.1   |/      /---------------\
     |                       | serving net  +------+    External     \
     |                       | 191.0.2.0/24 |      \     Network     /
     |                       +-----+--------+       \---------------/
     |         +----------------+  |
     |         |  AODVv2 Router |  |
     |         |  191.0.2.3/32  |  |
     \         +----------------+  /
      \                           /
       \-------------------------/
```

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:

- o Timers
- o Protocol constants
- o 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:

- o 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.

- o 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.

- o 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.

- o 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.

- o Routers with lower values for RteMsg_ENTRY_TIME may not consider received redundant multicast route messages as redundant, and may forward these messages unnecessarily.

- o 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.

- o 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.

- o 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:

- o DISCOVERY_ATTEMPTS_MAX: Routers with higher values are likely to be more successful at finding routes, at the cost of additional control traffic.

- o RREP_RETRIES: Routers with lower values are more likely to blacklist neighbors when there is a temporary fluctuation in link quality.

- o 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.

- o MAX_HOPCOUNT: Routers with a value too small would not be able to discover routes to distant addresses.

- o 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:

- o 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.

- o 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.

- o 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.

- o 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.

- o Replayed RERR messages could be used to disrupt active routes.

Mitigation:

- o 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.

- o 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.

- o By only regenerating RERR messages where active routes are removed, the spread of the malicious RERR is limited.

- o Including sequence numbers in RERR messages offers protection against attacks using replays of these RERR messages.

- o 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:

- o 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.

- o 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.

- o 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.

- o 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.

- o Sending an RERR message with false information is discussed in Section 13.1.2.

Mitigation:

- o 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:

- o 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:

- o 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.

- o When a TIMESTAMP Packet TLV is being added, the Packet TLV Block size field MUST be updated.

- o 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:

o 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.

o 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.

o 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.

o 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.

o 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.

- o 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.

- o 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.

- o 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.

- o If the received ICV value is truncated, the calculated ICV value MUST also be truncated (as specified in [RFC7182]), before comparing.

- o 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.

- o 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.

- o If the TIMESTAMP value in the received message is not valid, the AODVv2 message MUST be discarded and NOT processed or forwarded.

- o 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.

- o Changed ’regeneration’ language in favor of ’forwarding’.
- o Reintroduced use of msg-hop-limit in 5444 message header.
- o Use OrigPrefix rather than OrigAddr and TargPrefix rather than TargAddr where appropriate
- o Removed validity time
- o Removed AckReq from RREP messages, use two-way RREP_ack to check for bidirectionality
- o Unicast RREP messages
- o Removed orphaned references
- o Clarified language
- o Improved Sequence Number instructions
- o Changed ’Unknown’ terminology to ’Heard’
- o Extended experiment description
- o 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.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 4
2. Protocol Overview 7
  2.1. Destinations 8
  2.2. Conventions and Terminology 9
3. Requirements 9
4. Implementation Scenarios 9
5. Assumptions 10
6. Metrics 11
7. DLEP Session Flow 12
  7.1. Peer Discovery State 12
  7.2. Session Initialization State 13
  7.3. In-Session State 14
    7.3.1. Heartbeats 14
  7.4. Session Termination State 15
  7.5. Session Reset state 15
    7.5.1. Unexpected TCP connection termination 16
8. Transaction Model 16
9. Extensions 17
  9.1. Experiments 17
10. Scalability 18
11. DLEP Signal and Message Structure 18
  11.1. DLEP Signal Header 18
  11.2. DLEP Message Header 19
  11.3. DLEP Generic Data Item 20
12. DLEP Signals and Messages 20
  12.1. General Processing Rules 20
  12.2. Status code processing 21
  12.3. Peer Discovery Signal 22
  12.4. Peer Offer Signal 22
  12.5. Session Initialization Message 23
  12.6. Session Initialization Response Message 24

12.7. Session Update Message 25
12.8. Session Update Response Message 27
12.9. Session Termination Message 27
12.10. Session Termination Response Message 27
12.11. Destination Up Message 28
12.12. Destination Up Response Message 29
12.13. Destination Announce Message 30
12.14. Destination Announce Response Message 30
12.15. Destination Down Message 32
12.16. Destination Down Response Message 32
12.17. Destination Update Message 32
12.18. Link Characteristics Request Message 34
12.19. Link Characteristics Response Message 34
12.20. Heartbeat Message 35
13. DLEP Data Items 36
13.1. Status 37
13.2. IPv4 Connection Point 39
13.3. IPv6 Connection Point 40
13.4. Peer Type 41
13.5. Heartbeat Interval 42
13.6. Extensions Supported 43
13.7. MAC Address 44
13.8. IPv4 Address 44
13.8.1. IPv4 Address Processing 45
13.9. IPv6 Address 46
13.9.1. IPv6 Address Processing 47
13.10. IPv4 Attached Subnet 48
13.10.1. IPv4 Attached Subnet Processing 49
13.11. IPv6 Attached Subnet 50
13.11.1. IPv6 Attached Subnet Processing 51
13.12. Maximum Data Rate (Receive) 52
13.13. Maximum Data Rate (Transmit) 53
13.14. Current Data Rate (Receive) 54
13.15. Current Data Rate (Transmit) 54
13.16. Latency 55
13.17. Resources 56
13.18. Relative Link Quality (Receive) 57
13.19. Relative Link Quality (Transmit) 57
13.20. Maximum Transmission Unit (MTU) 58
14. Security Considerations 59
15. IANA Considerations 60
15.1. Registrations 60
15.2. Signal Type Registration 60
15.3. Message Type Registration 61
15.4. DLEP Data Item Registrations 61
15.5. DLEP Status Code Registrations 62
15.6. DLEP Extensions Registrations 63
15.7. DLEP IPv4 Connection Point Flags 63

15.8. DLEP IPv6 Connection Point Flags 64
15.9. DLEP Peer Type Flag 64
15.10. DLEP IPv4 Address Flag 64
15.11. DLEP IPv6 Address Flag 65
15.12. DLEP IPv4 Attached Subnet Flag 65
15.13. DLEP IPv6 Attached Subnet Flag 65
15.14. DLEP Well-known Port 66
15.15. DLEP IPv4 Link-local Multicast Address 66
15.16. DLEP IPv6 Link-local Multicast Address 66
16. Acknowledgments 66
17. References 66
17.1. Normative References 66
17.2. Informative References 67
Appendix A. Discovery Signal Flows 68
Appendix B. Peer Level Message Flows 68
B.1. Session Initialization 68
B.2. Session Initialization - Refused 69
B.3. Router Changes IP Addresses 70
B.4. Modem Changes Session-wide Metrics 70
B.5. Router Terminates Session 70
B.6. Modem Terminates Session 71
B.7. Session Heartbeats 71
B.8. Router Detects a Heartbeat timeout 72
B.9. Modem Detects a Heartbeat timeout 73
Appendix C. Destination Specific Message Flows 73
C.1. Common Destination Notification 73
C.2. Multicast Destination Notification 74
C.3. Link Characteristics Request 75
Authors’ Addresses 76

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.

```
      |-------Local Node-------|          |-------Remote Node------|
      |                        |          |                        |
      +--------+       +-------+          +-------+       +--------+
      | Router |=======| Modem |{~~~~~~~~}| Modem |=======| Router |
      |        |       | Device|          | Device|       |        |
      +--------+       +-------+          +-------+       +--------+
               |       |       | Link     |       |       |
               |-DLEP--|       | Protocol |       |-DLEP--|
               |       |       | (e.g.    |       |       |
               |       |       | 802.11)  |       |       |
```

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.

```
                 +--------+                      +--------+
            +----+ Modem  |                      | Modem  +---+
            |    | Device |                      | Device |
            |    | Type A |  <===== // ======>   | Type A |   |
            |    +--------+      P-2-P Link      +--------+   |
        +---+----+                                        +---+----+
        | Router |                                        | Router |
        |        |                                        |        |
        +---+----+                                        +---+----+
            |     +--------+                      +--------+  |
            +-----+ Modem  |                      | Modem  |  |
                  | Device |   o o o o o o o o   | Device +--+
                  | Type B |    o  Shared   o    | Type B |
                  +--------+     o Medium  o     +--------+
                                  o       o
                                   o     o
                                    o   o
                                      o
                                 +--------+
                                 | Modem  |
                                 | Device |
                                 | Type B |
                                 +---+----+
                                     |
                                     |
                                 +---+----+
                                 | Router |
                                 |        |
                                 +--------+
```

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:

- o Peer Discovery
- o Session Initialization
- o In-Session
- o Session Termination
- o 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:

- o The implementation terminates the session by sending a Session Termination Message (Section 12.9), or,

- o 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:

- o Release all resources allocated for the session.

- o Eliminate all destinations in the information base represented by the session. Destination Down Messages (Section 12.15) MUST NOT be sent.

o 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      'D'      |      'L'      |      'E'      |      'P'      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Signal Type                   | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Message Type                  | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Value...                            :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

- o Heartbeat Interval Data Item (Section 13.5)
- o 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:

- o IPv4 Address (Section 13.8)
- o IPv6 Address (Section 13.9)
- o IPv4 Attached Subnet (Section 13.10)
- o 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:

- o IPv4 Address (Section 13.8)
- o IPv6 Address (Section 13.9)
- o IPv4 Attached Subnet (Section 13.10)
- o 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:

- o IPv4 Address (Section 13.8)
- o IPv6 Address (Section 13.9)
- o IPv4 Attached Subnet (Section 13.10)
- o 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:

- o Maximum Data Rate (Receive) (Section 13.12)
- o Maximum Data Rate (Transmit) (Section 13.13)
- o Current Data Rate (Receive) (Section 13.14)
- o Current Data Rate (Transmit) (Section 13.15)
- o 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:

- o Resources (Section 13.17)
- o Relative Link Quality (Receive) (Section 13.18)
- o Relative Link Quality (Transmit) (Section 13.19)
- o 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:

- o IPv4 Address (Section 13.8)
- o IPv6 Address (Section 13.9)

The Destination Up Message MAY contain one of each of the following Data Items:

- o Maximum Data Rate (Receive) (Section 13.12)
- o Maximum Data Rate (Transmit) (Section 13.13)
- o Current Data Rate (Receive) (Section 13.14)
- o Current Data Rate (Transmit) (Section 13.15)
- o 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:

- o Resources (Section 13.17)
- o Relative Link Quality (Receive) (Section 13.18)
- o Relative Link Quality (Transmit) (Section 13.19)
- o 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:

- o IPv4 Attached Subnet (Section 13.10)

- o 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:

- o MAC Address (Section 13.7)

- o 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:

- o IPv4 Address (Section 13.8)

- o 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:

- o MAC Address (Section 13.7)
- o Status (Section 13.1)

The Destination Announce Response Message MAY contain one or more of each of the following Data Items, with different values:

- o IPv4 Address (Section 13.8)
- o IPv6 Address (Section 13.9)
- o IPv4 Attached Subnet (Section 13.10)
- o IPv6 Attached Subnet (Section 13.11)

The Destination Announce Response Message MAY contain one of each of the following Data Items:

- o Maximum Data Rate (Receive) (Section 13.12)
- o Maximum Data Rate (Transmit) (Section 13.13)
- o Current Data Rate (Receive) (Section 13.14)
- o Current Data Rate (Transmit) (Section 13.15)
- o 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:

- o Resources (Section 13.17)
- o Relative Link Quality (Receive) (Section 13.18)
- o Relative Link Quality (Transmit) (Section 13.19)
- o 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:

- o MAC Address (Section 13.7)
- o 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:

- o Change in link metrics (e.g., Data Rates)

- o 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:

- o Maximum Data Rate (Receive) (Section 13.12)

- o Maximum Data Rate (Transmit) (Section 13.13)

- o Current Data Rate (Receive) (Section 13.14)

- o Current Data Rate (Transmit) (Section 13.15)

- o 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:

- o Resources (Section 13.17)

- o Relative Link Quality (Receive) (Section 13.18)

- o Relative Link Quality (Transmit) (Section 13.19)

- o 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:

- o IPv4 Address (Section 13.8)

- o IPv6 Address (Section 13.9)

- o IPv4 Attached Subnet (Section 13.10)

- o 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:

- o MAC Address (Section 13.7)

The Link Characteristics Request Message MUST contain at least one of each of the following Data Items:

- o Current Data Rate (Receive) (Section 13.14)
- o Current Data Rate (Transmit) (Section 13.15)
- o 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:

- o MAC Address (Section 13.7)
- o Status (Section 13.1)

The Link Characteristics Response Message SHOULD contain one of each of the following Data Items:

- o Maximum Data Rate (Receive) (Section 13.12)
- o Maximum Data Rate (Transmit) (Section 13.13)
- o Current Data Rate (Receive) (Section 13.14)
- o Current Data Rate (Transmit) (Section 13.15)
- o 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:

- o Resources (Section 13.17)
- o Relative Link Quality (Receive) (Section 13.18)
- o Relative Link Quality (Transmit) (Section 13.19)
- o 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Code          | Text...                                       :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Flags       |               IPv4 Address...                 :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:  ...cont.     |   TCP Port Number (optional)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |T|
+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Flags       |                IPv6 Address                   :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        IPv6 Address                           :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        IPv6 Address                           :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        IPv6 Address                           :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:  ...cont.     |   TCP Port Number (optional)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |T|
+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Flags         | Description...                                :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |S|
+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Heartbeat Interval                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Extensions List...                                            :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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).

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       MAC Address                             :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                MAC Address    :     (if EUI-64 used)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Flags         | IPv4 Address                                  :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:    ...cont.   |
+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |A|
+-+-+-+-+-+-+-+-+
```

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:

- o An address Drop operation referencing an address that is not associated with the peer in the current session.

- o 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:

- o An address Add operation referencing an address that has already been added to the destination in the current session.
- o An address Add operation referencing an address that is associated with a different destination or the peer in the current session.
- o An address Add operation referencing an address that makes no sense, for example defined as not forwardable in [RFC6890].
- o 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Flags         | IPv6 Address                                  :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                         IPv6 Address                          :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                         IPv6 Address                          :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                         IPv6 Address                          :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
: IPv6 Address  |
+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |A|
+-+-+-+-+-+-+-+-+
```

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:

- o An address Drop operation referencing an address that is not associated with the peer in the current session.

- o 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:

- o An address Add operation referencing an address that has already been added to the destination in the current session.

- o An address Add operation referencing an address that is associated with a different destination or the peer in the current session.

- o An address Add operation referencing an address that makes no sense, for example defined as not forwardable in [RFC6890].

- o 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Flags        |          IPv4 Attached Subnet                 :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:    ...cont.   |Prefix Length  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |A|
+-+-+-+-+-+-+-+-+
```

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:

- o A subnet Drop operation referencing a subnet that is not associated with the peer in the current session.

- o 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:

- o A subnet Add operation referencing a subnet that has already been added to the destination in the current session.

- o A subnet Add operation referencing a subnet that is associated with a different destination in the current session.

- o An subnet Add operation referencing an subnet that makes no sense, for example defined as not forwardable in [RFC6890].

- o 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Flags       |         IPv6 Attached Subnet                  :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                   IPv6 Attached Subnet                        :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                   IPv6 Attached Subnet                        :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                   IPv6 Attached Subnet                        :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:    ...cont.   | Prefix Len.   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|  Reserved   |A|
+-+-+-+-+-+-+-+-+
```

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:

- o A subnet Drop operation referencing a subnet that is not associated with the peer in the current session.

- o 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:

- o A subnet Add operation referencing a subnet that has already been added to the destination in the current session.

- o A subnet Add operation referencing a subnet that is associated with a different destination in the current session.

- o An subnet Add operation referencing an subnet that makes no sense, for example defined as not forwardable in [RFC6890].

- o 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        MDRR (bps)                             :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        MDRR (bps)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        MDRT (bps)                             :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        MDRT (bps)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        CDRR (bps)                             :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        CDRR (bps)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        CDRT (bps)                             :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                        CDRT (bps)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Latency                            :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                            Latency                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     RES       |
+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     RLQR      |
+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     RLQT      |
+-+-+-+-+-+-+-+-+
```

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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Item Type                | Length                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             MTU               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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

```
Router                        Modem    Signal Description
==========================================================================

|                                      Router initiates discovery, starts
|                                      a timer, send Peer Discovery
|-------Peer Discovery---->X           Signal.

         ~ ~ ~ ~ ~ ~ ~                 Router discovery timer expires
                                       without receiving Peer Offer.

|                                      Router sends another Peer
|-------Peer Discovery---------->|     Discovery Signal.
                                 |
                                 |     Modem receives Peer Discovery
                                 |     Signal.
                                 |
                                 |     Modem sends Peer Offer with
|<--------Peer Offer-------------|     Connection Point information.
:
:                                      Router MAY cancel discovery timer
:                                      and stop sending Peer Discovery
:                                      Signals.
```

Appendix B. Peer Level Message Flows

B.1. Session Initialization

```
Router                        Modem    Message Description
==========================================================================

|                                      Router connects to discovered or
|                                      pre-configured Modem Connection
|--TCP connection established--->       Point.
|
|                                      Router sends Session
|----Session Initialization----->|     Initialization Message.
                                 |
                                 |     Modem receives Session
                                 |     Initialization Message.
                                 |
                                 |     Modem sends Session Initialization
|<--Session Initialization Resp.-|     Response, with Success Status Data
|                                 |     Item.
|                                 |
|<<============================>>|     Session established. Heartbeats
:                                 :     begin.
```

B.2. Session Initialization - Refused

```
Router                        Modem    Message Description
==========================================================================

|                                      Router connects to discovered or
|                                      pre-configured Modem Connection
|--TCP connection established--->      Point.
|
|                                      Router sends Session
|-----Session Initialization---->|     Initialization Message.
                                 |
                                 |     Modem receives Session
                                 |     Initialization Message, and will
                                 |     not support the advertised
                                 |     extensions.
                                 |
                                 |     Modem sends Session Initialization
                                 |     Response, with 'Request Denied'
|<-Session Initialization Resp.--|     Status Data Item.
|
|
|                                      Router receives negative Session
|                                      Initialization Response, closes
||---------TCP close------------||     TCP connection.
```

B.3. Router Changes IP Addresses

```
Router                        Modem    Message Description
==========================================================================

|                                      Router sends Session Update
|-------Session Update---------->|     Message to announce change of IP
                                 |     address
                                 |
                                 |     Modem receives Session Update
                                 |     Message and updates internal
                                 |     state.
                                 |
|<----Session Update Response----|     Modem sends Session Update
                                 |     Response.
```

B.4. Modem Changes Session-wide Metrics

```
Router                        Modem    Message Description
==========================================================================

                                 |     Modem sends Session Update Message
                                 |     to announce change of modem-wide
|<--------Session Update---------|     metrics
|
|                                      Router receives Session Update
|                                      Message and updates internal
|                                      state.
|
|----Session Update Response---->|     Router sends Session Update
                                 |     Response.
```

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

```
Router                        Modem    Message Description
==========================================================================

|----------Heartbeat------------>|    Router sends heartbeat Message
                                 |
                                 |    Modem resets heartbeats missed
                                 |    counter.

         ~ ~ ~ ~ ~ ~ ~

|---------[Any Message]--------->|    When the Modem receives any
                                 |    Message from the Router.
                                 |
                                 |    Modem resets heartbeats missed
                                 |    counter.

         ~ ~ ~ ~ ~ ~ ~

|<---------Heartbeat-------------|    Modem sends heartbeat Message
|
|                                     Router resets heartbeats missed
|                                     counter.

         ~ ~ ~ ~ ~ ~ ~

|<--------[Any Message]----------|    When the Router receives any
|                                     Message from the Modem.
|
|                                     Modem resets heartbeats missed
|                                     counter.
```

B.8. Router Detects a Heartbeat timeout

```
Router                        Modem    Message Description
==========================================================================

         X<----------------------|    Router misses a heartbeat


|        X<----------------------|    Router misses too many heartbeats
|
|
|------Session Termination------>|    Router sends Session Termination
|                                     Message with 'Timeout' Status
|                                     Data Item.
:
:                                     Termination proceeds...
```

B.9. Modem Detects a Heartbeat timeout

```
Router                        Modem    Message Description
==========================================================================

|---------------------->X             Modem misses a heartbeat


|---------------------->X        |    Modem misses too many heartbeats
                                 |
                                 |
|<-----Session Termination-------|    Modem sends Session Termination
                                 |    Message with 'Timeout' Status
                                 |    Data Item.
                                 :
                                 :    Termination proceeds...
```

Appendix C. Destination Specific Message Flows

C.1. Common Destination Notification

```
Router                        Modem    Message Description
==========================================================================

                                  |    Modem detects a new logical
                                  |    destination is reachable, and
|<-------Destination Up----------|    sends Destination Up Message.
|
|------Destination Up Resp.----->|    Router sends Destination Up
                                  |    Response.

           ~ ~ ~ ~ ~ ~ ~

                                  |    Modem detects change in logical
                                  |    destination metrics, and sends
|<-------Destination Update------|    Destination Update Message.

           ~ ~ ~ ~ ~ ~ ~

                                  |    Modem detects change in logical
                                  |    destination metrics, and sends
|<-------Destination Update------|    Destination Update Message.

            ~ ~ ~ ~ ~ ~ ~

                                  |    Modem detects logical destination
                                  |    is no longer reachable, and sends
|<-------Destination Down--------|    Destination Down Message.
|
|                                     Router receives Destination Down,
|                                     updates internal state, and sends
|------Destination Down Resp.--->|    Destination Down Response Message.
```

C.2. Multicast Destination Notification

```
Router                        Modem    Message Description
==========================================================================

|                                      Router detects a new multicast
|                                      destination is in use, and sends
|-----Destination Announce------>|     Destination Announce Message.
                                 |
                                 |     Modem updates internal state to
                                 |     monitor multicast destination, and
|<-----Dest. Announce Resp.------|     sends Destination Announce
                                       Response.

           ~ ~ ~ ~ ~ ~ ~

                                 |     Modem detects change in multicast
                                 |     destination metrics, and sends
|<-------Destination Update------|     Destination Update Message.

           ~ ~ ~ ~ ~ ~ ~

                                 |     Modem detects change in multicast
                                 |     destination metrics, and sends
|<-------Destination Update------|     Destination Update Message.

            ~ ~ ~ ~ ~ ~ ~

|                                      Router detects multicast
|                                      destination is no longer in use,
|--------Destination Down------->|     and sends Destination Down
                                 |     Message.
                                 |
                                 |     Modem receives Destination Down,
                                 |     updates internal state, and sends
|<-----Destination Down Resp.----|     Destination Down Response Message.
```

C.3. Link Characteristics Request

```
Router                        Modem    Message Description
==========================================================================

                                       Destination has already been
           ~ ~ ~ ~ ~ ~ ~               announced by either peer.

|                                      Router requires different
|                                      Characteristics for the
|                                      destination, and sends Link
|--Link Characteristics Request->|     Characteristics Request Message.
                                 |
                                 |     Modem attempts to adjust link
                                 |     properties to meet the received
                                 |     request, and sends a Link
                                 |     Characteristics Response
|<---Link Characteristics Resp.--|     Message with the new values.
```

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

Mobile Ad hoc Networking (MANET)
Internet-Draft
Updates: RFC6130, OLSRv2
(if approved)
Intended status: Standards Track
Expires: September 6, 2014

C. Dearlove
BAE Systems ATC
T. Clausen
LIX, Ecole Polytechnique

March 5, 2014

Optimized Link State Routing Protocol version 2 (OLSRv2) and MANET Neighborhood Discovery Protocol (NHDP) Extension TLVs
draft-ietf-manet-nhdp-olsrv2-tlv-extension-05

Abstract

This specification describes extensions to definitions of TLVs used by the Optimized Link State Routing Protocol version 2 (OLSRv2) and the MANET Neighborhood Discovery Protocol (NHDP), to increase their abilities to accommodate protocol extensions. This document updates OLSRv2 and 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 September 6, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Applicability Statement 3
4. TLV Values 4
4.1. Unrecognized TLV Values 4
4.2. TLV Value Lengths 5
4.3. Undefined TLV Values 6
4.3.1. NHDP TLVs: LOCAL_IF, LINK_STATUS and OTHER_NEIGHB . . 6
4.3.2. OLSRv2 TLVs: MPR and NBR_ADDR_TYPE 6
4.3.3. Unspecified TLV Values 7
5. IANA Considerations 8
5.1. LOCAL_IF Address Block TLVs 8
5.1.1. Create New Registry 8
5.1.2. Modification to Existing Registry 9
5.2. LINK_STATUS Address Block TLVs 10
5.2.1. Create New Registry 10
5.2.2. Modification to Existing Registry 11
5.3. OTHER_NEIGHB Address Block TLVs 12
5.3.1. Create New Registry 12
5.3.2. Modification to Existing Registry 13
5.4. MPR Address Block TLVs 13
5.4.1. Create New Registry 13
5.4.2. Modification to Existing Registry 14
5.5. NBR_ADDR_TYPE Address Block TLVs 15
5.5.1. Create New Registry 15
5.5.2. Modification to Existing Registry 16
6. Security Considerations 16
7. Acknowledgments 17
8. References 17
8.1. Normative References 17
8.2. Informative References 17
Authors' Addresses 18

1. Introduction

The MANET Neighborhood Discovery Protocol (NHDP) [RFC6130] and the Optimized Link State Routing Protocol, version 2 (OLSRv2) [OLSRv2] are protocols for use in mobile ad hoc networks (MANETs) [RFC2501], based on the Generalized Mobile Ad Hoc Network (MANET) Packet/Message Format [RFC5444].

This document updates [RFC6130] and [OLSRv2], specifically their use of TLV (Type-Length-Value) elements, to increase the extensibility of these protocols, and to enable some improvements in their implementation.

This specification reduces the latitude of implementations of [OLSRv2] and [RFC6130] to consider some messages, which will not be created by implementations simply following those specifications, as a reason to consider the message as "badly formed", and thus as a reason to reject the message. This gives greater latitude to the creation of extensions of these protocols, in particular extensions that will interoperate with unextended implementations of those protocols. As part of that, it indicates how TLVs (Type-Length-Value elements) [RFC5444] with unexpected value fields must be handled, and adds some additional options to those TLVs.

Note that TLVs with unknown type or type extension are already specified as to be ignored by [RFC6130] and [OLSRv2], and also are not a reason to reject a message.

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].

Additionally, this document uses the terminology of [RFC5444], [RFC6130], and [OLSRv2].

3. Applicability Statement

This document updates the specification of the protocols [OLSRv2] and [RFC6130].

Specifically, this specification updates [RFC6130] and [OLSRv2] in the following way:

- o Removes the latitude of rejecting a message with a TLV with a known type, but with an unexpected TLV Value field, for the TLV Types defined in [RFC6130] and [OLSRv2].

- o Specifies the handling of a TLV Value field with unexpected length.

- o Sets up IANA registries for TLV Values for the Address Block TLVs:

  - * LOCAL_IF, defined in [RFC6130].

  - * LINK_STATUS, defined in [RFC6130].

  - * OTHER_NEIGHB, defined in [RFC6130].

  - * MPR, defined in [OLSRv2], now considered as a bit field.

  - * NBR_ADDR_TYPE, defined in [OLSRv2], now considered as a bit field.

- o Defines a well-known TLV Value for "UNSPECIFIED" for the Address Block TLV Types LOCAL_IF, LINK_STATUS, and OTHER_NEIGHB, all defined in [RFC6130].

4. TLV Values

NHDP [RFC6130] and OLSRv2 [OLSRv2] define a number of TLVs within the framework of [RFC5444]. These TLVs define the meaning of only some of the contents that can be found in a TLV Value field. This limitation may be either only defining certain TLV Values, or considering only some lengths of the TLV Value fields (or single value field in a multi value Address-Block TLV). This specification describes how NHDP [RFC6130] and OLSRv2 [OLSRv2] are to handle TLVs with other TLV Value fields.

4.1. Unrecognized TLV Values

NHDP and OLSRv2 specify that, in addition to well-defined reasons (in the respective protocol specifications), an implementation of these protocols MAY recognize a message as "badly formed" and therefore "invalid for processing" for other reasons (Section 12.1 of [RFC6130] and Section 16.3.1 of [OLSRv2]). These sections could be interpreted as allowing rejection of a message because a TLV Value field is unrecognized. This specification removes that latitude:

- o An implementation MUST NOT reject a message because it contains an unrecognized TLV value. Instead, any unrecognised TLV Value field

MUST be processed or ignored by an unextended implementation of NHDP or OLSRv2, as described in the following sections.

- o Hence, this specification removes the 7th, 10th, and 11th bullets in Section 12.1 of [RFC6130].

It should be stressed that this is not a change to [RFC6130] or [OLSRv2], except with regard to not allowing this to be a reason for rejection of a message. [RFC6130] or [OLSRv2] are specified in terms such as "if an address is associated with a value of LOST by a LINK_STATUS TLV". Association with an unrecognized value has no effect on any implementation strictly following such a specification.

4.2. TLV Value Lengths

The TLVs specified in [RFC6130] and [OLSRv2] may be either single-value or multi-value TLVs. In either case, the length of each item of information encoded in the TLV Value field is the "single-length", defined and calculated as in section 5.4.1 in [RFC5444]. All TLVs specified in [RFC6130] and [OLSRv2] have a one or two octet single-length. These are considered the expected single-lengths of such a received TLV.

Other single-length TLV Value fields may be introduced by extensions to [RFC6130] and [OLSRv2]. This document specifies how implementations of [RFC6130] and [OLSRv2], or extensions thereof, MUST behave on receiving TLVs of the TLV types defined in [RFC6130] and [OLSRv2], but with TLV Value fields with other single-length values.

The following principles apply:

- o If the received single-length is greater than the expected single-length, then the excess octets MUST be ignored.
- o If the received single-length is less than the expected single-length, then the absent octets MUST considered to have all bits cleared (0).

Exceptions:

- o A received CONT_SEQ_NUM with a single-length < 2 SHOULD be considered an error.

4.3. Undefined TLV Values

[RFC6130] and [OLSRv2] define a number of TLVs, but for some of these TLVs specify meanings for only some TLV Values. This document establishes IANA registries for these TLV Values, with initial registrations reflecting those used by [RFC6130] and [OLSRv2], and as specified in Section 4.3.3.

There are different cases of TLV Values with different characteristics. These cases are considered in this section.

4.3.1. NHDP TLVs: LOCAL_IF, LINK_STATUS and OTHER_NEIGHB

For the Address-Block TLVs LOCAL_IF, LINK_STATUS and OTHER_NEIGHB TLVs, defined in [RFC6130], only a limited number of values are specified for each. These are converted, by this specification, into extensible registries with initial registrations for values defined and used by [RFC6130] - see Section 5.

An implementation of [RFC6130], receiving a LOCAL_IF, LINK_STATUS, or OTHER_NEIGHB TLV with any TLV Value other than the values which are defined in [RFC6130] MUST ignore that TLV Value, as well as any corresponding attribute association to the address.

4.3.2. OLSRv2 TLVs: MPR and NBR_ADDR_TYPE

The Address-Block TLVs MPR and NBR_ADDR_TYPE, defined in [OLSRv2], are similar to those defined in [RFC6130] in having only limited values specified (1, 2 and 3): 1 and 2, represent presence of two different attributes associated to an address, and 3 represents "both 1 and 2".

These TLV Value fields, are by this specification, converted to bit fields, and MUST be interpreted as such. As the existing definitions of values 1, 2, and 3 behave in that manner, it is likely that this will involve no change to an implementation, but any test of (for example) Value = 1 or Value = 3 MUST be converted to a test of (for example) Value bitand 1 = 1, where "bitand" denotes a bitwise and operation.

This specification creates registries for recording reservations of the individual bits in these bitfields, with initial registrations for values defined and used by [OLSRv2] - see Section 5.

Other TLVs defined by [OLSRv2] are not affected by this specification.

4.3.3. Unspecified TLV Values

The registries defined in Section 5 for the LOCAL_IF, LINK_STATUS and OTHER_NEIGHB TLVs each include an additional TLV Value UNSPECIFIED. This TLV Value represents a defined value that, like currently undefined TLV Values, indicates that no information is associated with this address, but will always have this meaning. Such a TLV Value may be used to enable the creation of more efficient multivalue Address Block TLVs, or to simplify an implementation.

The similar requirement for the MPR and NBR_ADDR_TYPES TLVs is already satisfied by the TLV Value zero, provided that each bit in the TLV Value is defined as set ('1') when indicating the presence of an attribute, or clear ('0') when indicating the absence of an attribute; this is therefore required for registrations from the relevant registries, see Section 5.

For the LINK_METRIC TLV, this is already possible by clearing the most significant bits (0 to 3) of the first octet of the TLV Value. It is RECOMMENDED that in this case the remaining bits of the TLV Value are either all clear ('0') or all set ('1').

5. IANA Considerations

IANA is requested to take a total of ten actions as set out in the following sections.

5.1. LOCAL_IF Address Block TLVs

5.1.1. Create New Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters". IANA is requested to create a new sub-registry called "LOCAL_IF TLV Values".

IANA is requested to populate this registry as specified in Table 1.

Value	Name	Description	Reference
0	THIS_IF	The network address is associated with this local interface of the sending router	[This.I-D]
1	OTHER_IF	The network address is associated with another local interface of the sending router	[This.I-D]
2-223		Unallocated: Expert Review	
224-254		Experimental Use	[This.I-D]
255	UNSPECIFIED	No information about this network address is provided	[This.I-D]

Table 1: LOCAL_IF TLV Values

New assignments are to be made by Expert Review [RFC5226].

The Designated Experts are required to use the guidelines specified in [RFC6130] and [OLSRv2]. IANA is not expected to record this fact in the registry.

5.1.2. Modification to Existing Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters" with a sub-registry called "LOCAL_IF Address Block TLV Type Extensions". This sub-registry currently has an entry for value 0. IANA is requested to replace the entry in the Description column for this value with the text "The value is to be interpreted according to the registry LOCAL_IF TLV Values". The resulting table should look as specified in Table 2.

Type Extension	Description	Reference
0	The value is to be interpreted according to the registry LOCAL_IF TLV Values	[RFC6130] [This.I-D]
1-255	Unassigned	[This.I-D]

Table 2: LOCAL_IF Address Block TLV Type Extensions Modifications

5.2. LINK_STATUS Address Block TLVs

5.2.1. Create New Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters". IANA is requested to create a new sub-registry called "LINK_STATUS TLV Values".

IANA is requested to populate this registry as specified in Table 3.

Value	Name	Description	Reference
0	LOST	The link on this interface from the router with that network address has been lost	[This.I-D]
1	SYMMETRIC	The link on this interface from the router with that network address has the status of symmetric	[This.I-D]
2	HEARD	The link on this interface from the router with that network address has the status of heard	[This.I-D]
3-223		Unallocated: Expert Review	
224-254		Experimental Use	[This.I-D]
255	UNSPECIFIED	No information about this network address is provided	[This.I-D]

Table 3: LINK_STATUS TLV Values

New assignments are to be made by Expert Review [RFC5226].

The Designated Experts are required to use the guidelines specified in [RFC6130] and [OLSRv2]. IANA is not expected to record this fact in the registry.

5.2.2. Modification to Existing Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters" with a sub-registry called "LINK_STATUS Address Block TLV Type Extensions". This sub-registry currently has an entry for value 0. IANA is requested to replace the entry in the Description column for this value with the text "The value is to be interpreted according to the registry LINK_STATUS TLV Values". The resulting table should look as specified in Table 4.

Type Extension	Description	Reference
0	The value is to be interpreted according to the registry LINK_STATUS TLV Values	[RFC6130] [This.I-D]
1-255	Unassigned	[This.I-D]

Table 4: LINK_STATUS Address Block TLV Type Extensions Modifications

5.3. OTHER_NEIGHB Address Block TLVs

5.3.1. Create New Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters". IANA is requested to create a new sub-registry called "OTHER_NEIGHB TLV Values".

IANA is requested to populate this registry as specified in Table 5.

Value	Name	Description	Reference
0	LOST	The neighbor relationship with the router with that network address has been lost	[This.I-D]
1	SYMMETRIC	The neighbor relationship with the router with that network address is symmetric	[This.I-D]
2-223		Unallocated: Expert Review	
224-254		Experimental Use	[This.I-D]
255	UNSPECIFIED	No information about this network address is provided	[This.I-D]

Table 5: OTHER_NEIGHB Address Block TLV Values

New assignments are to be made by Expert Review [RFC5226].

The Designated Experts are required to use the guidelines specified in [RFC6130] and [OLSRv2]. IANA is not expected to record this fact in the registry.

5.3.2. Modification to Existing Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters" with a sub-registry called "OTHER_NEIGHB Address Block TLV Type Extensions". This sub-registry currently has an entry for value 0. IANA is requested to replace the entry in the Description column for this value with the text "The value is to be interpreted according to the registry OTHER_NEIGHB TLV Values". The resulting table should look as specified in Table 6.

Type Extension	Description	Reference
0	The value is to be interpreted according to the registry OTHER_NEIGHB TLV Values	[RFC6130] [This.I-D]
1-255	Unassigned	[This.I-D]

Table 6: OTHER_NEIGHB Address Block TLV Type Extensions Modifications

5.4. MPR Address Block TLVs

5.4.1. Create New Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters". IANA is requested to create a new sub-registry called "MPR TLV Bit Values".

IANA is requested to populate this registry as specified in Table 7.

Bit	Value	Name	Description	Reference
7	0x01	Flooding	The neighbor with that network address has been selected as flooding MPR	[This.I-D]
6	0x02	Routing	The neighbor with that network address has been selected as routing MPR	[This.I-D]
0-5			Unallocated: Expert Review	

Table 7: MPR Address Block TLV Bit Values

New assignments are to be made by Expert Review [RFC5226].

The Designated Experts are required to use the guidelines specified in [RFC6130] and [OLSRv2]. Additionally, the Designated Experts are required to ensure that the following sense is preserved:

- o For each bit in the field, a set bit (1) means that the address has the designated property, while an unset bit (0) means that no information about the designated property is provided. In particular, an unset bit must not be used to convey any specific information about the designated property. IANA is not expected to record these facts in the registry.

5.4.2. Modification to Existing Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters" with a sub-registry called "MPR Address Block TLV Type Extensions". This sub-registry currently has an entry for value 0. IANA is requested to replace the entry in the Description column for this value with the text "The value is to be interpreted according to the registry MPR TLV Bit Values". The resulting table should look as specified in Table 8.

Type Extension	Description	Reference
0	The value is to be interpreted according to the registry MPR TLV Bit Values	[OLSRv2] [This.I-D]
1-255	Unassigned	[This.I-D]

Table 8: MPR Address Block TLV Type Extensions Modifications

5.5. NBR_ADDR_TYPE Address Block TLVs

5.5.1. Create New Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters". IANA is requested to create a new sub-registry called "NBR_ADDR_TYPE Address Block TLV Bit Values".

IANA is requested to populate this registry as specified in Table 9.

Bit	Value	Name	Description	Reference
7	0x01	ORIGINATOR	The network address is an originator address reachable via the originating router	[This.I-D]
6	0x02	ROUTABLE	The network address is a routable address reachable via the originating router	[This.I-D]
0-5			Unallocated: Expert Review	

Table 9: NBR_ADDR_TYPE Address Block TLV Bit Values

New assignments are to be made by Expert Review [RFC5226].

The Designated Experts are required to use the guidelines specified in [RFC6130] and [OLSRv2]. Additionally, the Designated Experts are required to ensure that the following sense is preserved:

- For each bit in the field, a set bit (1) means that the address has the designated property, while an unset bit (0) means that no information about the designated property is provided. In particular, an unset bit must not be used to convey any specific information about the designated property. IANA is not expected to record these facts in the registry.

5.5.2. Modification to Existing Registry

IANA maintains a registry called "Mobile Ad hoc NETwork (MANET) Parameters" with a sub-registry called "NBR_ADDR_TYPE Address Block TLV Type Extensions". This sub-registry currently has an entry for value 0. IANA is requested to replace the entry in the Description column for this value with the text "The value is to be interpreted according to the registry NBR_ADDR_TYPE TLV Bit Values". The resulting table should look as specified in Table 10.

Type Extension	Description	Reference
0	The value is to be interpreted according to the registry NBR_ADDR_TYPE Address Block TLV Bit Values	[OLSRv2] [This.I-D]
1-255	Unassigned	[This.I-D]

Table 10: NBR_ADDR_TYPE Address Block TLV Type Extensions Modifications

6. Security Considerations

The presented updates to [RFC6130] and [OLSRv2]:

- o Create IANA registries for retaining TLV values for TLVs, already defined in the already published specifications of the two protocols, and with initial registrations for the TLV values defined by these specifications. This does not give rise to any additional security considerations.

- o Enable protocol extensions to be able to register TLV values in the created IANA registries. Such extensions MUST specify appropriate security considerations.

- o Create, in some registries, a registration for "UNSPECIFIED" values, for more efficient use of multi-value Address Block TLVs. The interpretation of an address being associated with a TLV of a given type and with the value "UNSPECIFIED" is identical to that address not being associated with a TLV of that type. Thus, this update does not give rise to any additional security considerations.

- o Reduces the latitude of implementations of the two protocols to reject a message as "badly formed", due to the value field of a

TLV being unexpected. These protocols are specified in terms such as "if an address is associated with a value of LOST by a LINK_STATUS TLV". Association with an unknown value (or a value newly defined to mean no link status information) has no effect on such a specification. Thus, this update does not give rise to any additional security considerations.

- o Do not introduce any opportunities for attacks on the protocols through signal modification that are not already present in the two protocols.

7. Acknowledgments

The authors would like to gratefully acknowledge the following people for intense technical discussions, early reviews, and comments on the specification (listed alphabetically): Ulrich Herberg (Fujitsu Laboratories of America) and Henning Rogge (Frauenhofer FKIE).

The authors would also like to express their gratitude to Adrian Farrel, for his assistance and contributions to successful and timely completion of this specification.

8. References

8.1. Normative References

[OLSRv2] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol version 2", work in progress draft-ietf-manet-olsrv2-19, March 2013.

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

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

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

8.2. Informative References

[RFC2501] Macker, J. and S. Corson, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, January 1999.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

Authors' Addresses

Christopher Dearlove
BAE Systems Advanced Technology Centre
West Hanningfield Road
Great Baddow, Chelmsford
United Kingdom

Phone: +44 1245 242194
Email: chris.dearlove@baesystems.com
URI: http://www.baesystems.com/

Thomas Heide Clausen
LIX, Ecole Polytechnique

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

Mobile Ad hoc Networking (MANET)
Internet-Draft
Updates: OLSRv2 (if approved)
Intended status: Standards Track
Expires: August 25, 2014

C. Dearlove
BAE Systems ATC
T. Clausen
LIX, Ecole Polytechnique

February 21, 2014

Routing MPR Optimization for the Optimized Link State Routing Protocol version 2 (OLSRv2)

draft-ietf-manet-olsrv2-rmpr-optimization-01

Abstract

This specification updates the Optimized Link State Routing Protocol version 2 (OLSRv2) with an optimization to improve the selection of routing Multi-Point Relays. The optimization retains full interoperability between implementations of OLSRv2 with and without this optimization.

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 25, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Applicability Statement 3
4. Routing MPR Selection 4
5. IANA Considerations 4
6. Security Considerations 4
7. Acknowledgments 5
8. References 5
8.1. Normative References 5
8.2. Informative References 5
Authors' Addresses 5

1. Introduction

The Optimized Link State Routing Protocol, version 2 [OLSRv2] is a proactive link state routing protocol designed for use in mobile ad hoc networks (MANETs) [RFC2501]. This document improves one area of that specification.

One improvement included in OLSRv2, compared to its predecessor described in [RFC3626], is the use of link metrics, rather than minimum hop routing. A rationale for how link metrics were included in OLSRv2 is documented in [RFC6966-to-be]. However, one aspect of the use of link metrics described in [RFC6966-to-be], the removal of some unnecessarily selected routing Multi-Point Relays (MPRs), was not included in [OLSRv2]. This specification updates OLSRv2 to include this optimization.

Note that this optimization does not impact interoperability: implementations which do, and which do not, implement this optimization will interoperate seamlessly.

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].

Additionally, this document uses the terminology of [OLSRv2].

3. Applicability Statement

This specification updates [OLSRv2]. As such it is applicable to all implementations of this protocol. The optimization presented in this specification is simply permissive, it allows an additional optimization, and there is no requirement for any implementation to include it. However inclusion of this optimization is advised; it can, in some cases, create smaller and fewer messages, without ever having the opposite effect.

[OLSRv2] defines the properties for the selection of routing MPRs (multipoint relays) from among a router’s symmetric 1-hop neighbors as properties that correspond first to that the selected MPRs consist of a set of symmetric 1-hop neighbors that cover all the symmetric 2-hop neighbors, and second that they do so retaining a minimum distance route (1-hop, if present, or 2-hop) to each symmetric 2-hop neighbor. The discussion in the latter part of Section 6.2 of

[RFC6966-to-be] indicates that this requirement is over-prescriptive for routing MPR selection. The update to [OLSRv2] described in this specification permits a router to use the described optimization, while still being considered compliant.

Note that, whether considered compliant or not, a router that implements the optimization, described in this specification, will interoperate successfully with routers not implementing this optimization.

4. Routing MPR Selection

A set of routing MPRs created as specified in [OLSRv2] MAY be optimized in the following manner. Note that this uses the notation of Section 18.2 of [OLSRv2]:

1. If there is a sequence $x_0, \ldots, x_n$ of elements of N1 such that:

   * x_0 is a routing MPR,

   * $x_1, \ldots, x_n$ have corresponding elements $y_1, \ldots, y_n$ of N2, and

   * $d1(x_0) + d2(x_0,y_1) + \ldots + d2(x_{m-1},y_m) < d1(x_m)$ for $m = 1, \ldots, n$,

   then x_1 to x_n may be removed from the set of routing MPRs, if selected.

Note that "corresponding elements" in N1 and N2 means that these elements represent the same router. All of this information is available from information gathered by NHDP [RFC6130].

5. IANA Considerations

This document has no actions for IANA.

[This section may be removed by the RFC Editor.]

6. Security Considerations

The update to [OLSRv2] does not introduce any new protocol signals, compared to the already published specifications of the protocol, nor does it change the processing of any received protocol signals.

This update to [OLSRv2] permits a compliant implementation of OLSRv2 to (potentially) eliminate some redundant information from the routing MPR sets otherwise generated by the algorithms described in [OLSRv2], and therefore also eliminate the need for including that information in generated TC messages. Because this information is not used when included, this update to [OLSRv2] does not present any additional security considerations, beyond those described in [OLSRv2].

7. Acknowledgments

The authors would like to gratefully acknowledge Philippe Jacquet (Alcatel-Lucent) for intense technical discussions and comments.

8. References

8.1. Normative References

[OLSRv2] Clausen, T., Dearlove, C., Jacquet, P., and U. Herberg, "The Optimized Link State Routing Protocol version 2", work in progress draft-ietf-manet-olsrv2-19, March 2013.

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

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

8.2. Informative References

[RFC2501] Macker, J. and S. Corson, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, January 1999.

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

[RFC6966-to-be] Clausen, T., Dearlove, C., and P. Jacquet, "Rationale for the Use of Link Metrics in the Optimized Link State Routing Protocol Version 2 (OLSRv2)", RFC 6966, TBD 2013.

Authors' Addresses

Christopher Dearlove
BAE Systems Advanced Technology Centre
West Hanningfield Road
Great Baddow, Chelmsford
United Kingdom

Phone: +44 1245 242194
Email: chris.dearlove@baesystems.com
URI: http://www.baesystems.com/

Thomas Heide Clausen
LIX, Ecole Polytechnique

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

MANET
Internet-Draft
Intended status: Informational
Expires: August 18, 2014

H. Rogge
Fraunhofer FKIE
E. Baccelli
INRIA

February 14, 2014

Packet Sequence Number based directional airtime metric for OLSRv2
draft-rogge-baccelli-olsrv2-ett-metric-04

Abstract

This document specifies an directional airtime link metric for usage in OLSRv2.

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 18, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Applicability Statement 4
4. Directional Airtime Metric Rational 5
5. Metric Functioning & Overview 6
6. Protocol Parameters 7
6.1. Recommended Values 7
7. Protocol Constants 7
8. Data Structures 8
8.1. Initial Values 9
9. Packets and Messages 9
9.1. Definitions 9
9.2. Requirements 9
9.3. Link Loss Data Gathering 10
9.3.1. Packet Sequence based link loss 10
9.3.2. HELLO based Link Loss 11
9.3.3. Other Measurement of Link Loss 11
9.4. HELLO Message Processing 11
10. HELLO Timeout Processing 12
11. Metric Update 12
12. IANA Considerations 13
13. Security Considerations 13
14. Acknowledgements 13
15. References 14
15.1. Normative References 14
15.2. Informative References 14
Appendix A. OLSR.org metric history 15
Authors' Addresses 16

1. Introduction

One of the major shortcomings of OLSR [RFC3626] is the missing of a link cost metric between mesh nodes. Operational experience with mesh networks gathered since the standardization of OLSR has revealed that wireless networks links can have highly variable and heterogeneous properties. This makes a hopcount metric insufficient for effective mesh routing.

Based on this experience, OLSRv2 [OLSRV2] integrates the concept of link metrics directly into the core specification of the routing protocol. The OLSRv2 routing metric is an external process, it can be any kind of dimensionless additive cost function which reports to the OLSRv2 protocol.

Since 2004 the OLSR.org [OLSR.org] implementation of OLSR included an Estimated Transmission Count (ETX) metric [MOBICOM04] as a proprietary extension. While this metric is not perfect, it proved to be sufficient for a long time for Community Mesh Networks (Appendix A). But the increasing maximum data rate of IEEE 802.11 made the ETX metric less efficient than in the past, which is one reason to move to a different metric.

This document describes a Directional Airtime routing metric for OLSRv2, a successor of the OLSR.org routing metric for [RFC3626]. It takes both the loss rate and the link speed into account to provide a more accurate picture of the mesh network links.

2. Terminology

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

The terminology introduced in [RFC5444], [OLSRV2] and [RFC6130], including the terms "packet", "message" and "TLV" are to be interpreted as described therein.

Additionally, this document uses the following terminology and notational conventions:

QUEUE - a first in, first out queue of integers.

QUEUE[TAIL] - the most recent element in the queue.

add(QUEUE, value) - adds a new element to the TAIL of the queue.

remove(QUEUE) - removes the HEAD element of the queue

sum(QUEUE) - an operation which returns the sum of all elements in a QUEUE.

diff_seqno(new, old) - an operation which returns the positive distance between two elements of the circular sequence number space defined in section 5.1 of [RFC5444]. Its value is either (new - old) if this result is positive, or else its value is (new - old + 65536).

MAX(a,b) - the maximum of a and b.

UNDEFINED - a value not in the normal value range of a variable. Might be -1 for this protocol.

airtime - the time a transmitted packet blocks the link layer, e.g., a wireless link.

ETX - Expected Transmission Count, a link metric proportional to the number of transmissions to successfully send an IP packet over a link.

ETT - Estimated Travel Time, a link metric proportional to the amount of airtime needed to transmit an IP packet over a link, not considering layer-2 overhead created by preamble, backoff time and queuing.

DAT - Directional Airtime Metric, the link metric described in this document, which is a directional variant of ETT. It does not take reverse path loss into account.

3. Applicability Statement

The Directional Airtime Metric was designed and tested in wireless IEEE 802.11 mesh networks. These networks employ link layer retransmission to increase the delivery probability and multiple unicast data rates.

The metric must learn about the unicast data rate towards each one-hop neighbor from an external process, either by configuration or by an external measurement process. This measurement could be done by gathering cross-layer data from the operating system or an external

daemon like DLEP [DLEP], but also by indirect layer-3 measurements like packet-pair.

If [RFC5444] control traffic is used to determine the link packet loss, the administrator should take care that link layer multicast transmission do not not have a higher reception probability than the slowest unicast transmission. It might be necessary to increase the data-rate of the multicast transmissions, e.g. set the multicast data-rate to 6 MBit/s if you use IEEE 802.11g only.

The metric can only handle a certain range of packet loss and unicast data-rate. Maximum packet loss is "ETX 4" (1 of 4 packets is successfully sent to the receiver, without link layer retransmissions), the unicast data-rate can be between 1024 Bit/s and 4 GBit/s. The metric has been designed for data-rates of 1 MBit/s and hundreds of MBit/s.

4. Directional Airtime Metric Rational

The Directional Airtime Metric has been inspired by the publications on the ETX [MOBICOM03] and ETT [MOBICOM04] metric, but has several key differences.

Instead of measuring the combined loss probability of a bidirectional transmission of a packet over a link in both directions, the Directional Airtime Metric measures the incoming loss rate and integrates the incoming linkspeed into the metric cost. There are multiple reasons for this decision:

- o OLSRv2 [OLSRV2] defines the link metric as directional costs between nodes.

- o Not all link layer implementations use acknowledgement mechanisms. Most link layer implementations who do use them use less airtime and a more robust modulation for the acknowledgement than the data transmission, which makes it more likely for the data transmission to be disrupted compared to the acknowledgement.

- o Incoming packet loss and linkspeed can be measured locally, symmetric link loss would need an additional signaling TLV in the [RFC6130] HELLO and would delay metric calculation by up to one HELLO interval.

The Directional Airtime Metric does not integrate the packet size into the link cost. Doing so is not feasible in most link-state routing protocol implementations. The routing decision of most operation systems don’t take packet size into account. Multiplying

all link costs of a topology with the size of a data-plane packet would never change the dijkstra result anyways.

The queue based packet loss estimator has been tested extensively in the OLSR.org ETX implementation, see Appendix A. The output is the average of the packet loss over a configured time period.

5. Metric Functioning & Overview

The Directional Airtime Metric is calculated for each link set entry, as defined in [RFC6130] section 7.1.

The metric processes two kinds of data into the metric value, namely packet loss rate and link-speed. While the link-speed is taken from an external process, the current packet loss rate is calculated by keeping track of packet reception and packet loss events.

Multiple incoming packet loss/reception events must be combined into a loss rate to get a smooth metric. Experiments with exponential weighted moving average (EWMA) lead to a highly fluctuating or a slow converging metric (or both). To get a smoother and more controllable metric result, this metric uses two fixed length queues to measure and average the incoming packet events, one queue for received packets and one for the estimated number of packets sent by the other side of the link.

Because the rate of incoming packets is not uniform over time, the queue contains a number of counters, each representing a fixed time interval. Incoming packet loss and packet reception event are accumulated in the current queue element until a timer adds a new empty counter to both queues and remove the oldest counter from both.

In addition to the packet loss stored in the queue, this metric uses a timer to detect a total link-loss. For every NHDP HELLO interval in which the metric received no packet from a neighbor, it scales the number of received packets in the queue based on the total time interval the queue represents compared to the total time of the lost HELLO intervals.

The average packet loss ratio is calculated as the sum of the ’total packets’ counters divided by the sum of the ’packets received’ counters. This value is then divided through the current link-speed and then scaled into the range of metrics allowed for OLSRv2.

The metric value is then used as L_in_metric of the Link Set (as defined in section 8.1. of [OLSRV2]).

6. Protocol Parameters

This specification defines the following parameters, which can be changed without making the metric outputs incomparable with each other:

DAT_MEMORY_LENGTH - Queue length for averaging packet loss. All received and lost packets within the queue are used to calculate the cost of the link.

DAT_REFRESH_INTERVAL - interval in seconds between two metric recalculations as described in Section 11. This value SHOULD be smaller than a typical HELLO interval.

DAT_HELLO_TIMEOUT_FACTOR - timeout factor for HELLO interval at which point a HELLO is definitely considered lost. The value must be a floating point number between 1.0 and 2.0, large enough to take the delay and jitter for message aggregation into account.

DAT_SEQNO_RESTART_DETECTION - threshold in number of missing packets (based on received packet sequence numbers) at which point the router considers the neighbor has restarted. This parameter is only used for packet sequence number based loss estimation. This number MUST be larger than DAT_MAXIMUM_LOSS.

6.1. Recommended Values

The proposed values of the protocol parameters are for Community Mesh Networks, which mostly use immobile mesh nodes. Using this metric for mobile networks might require shorter DAT_REFRESH_INTERVAL and/or DAT_MEMORY_LENGTH.

DAT_MEMORY_LENGTH := 64

DAT_REFRESH_INTERVAL := 1

DAT_HELLO_TIMEOUT_FACTOR := 1.2

DAT_SEQNO_RESTART_DETECTION := 256

7. Protocol Constants

This specification defines the following constants, which cannot be changed without making the metric outputs incomparable:

DAT_MAXIMUM_LOSS - Fraction of the loss rate used in this routing metric. Loss rate will be between 0/DAT_MAXIMUM_LOSS and (DAT_MAXIMUM_LOSS-1)/DAT_MAXIMUM_LOSS: 4.

DAT_MINIMUM_BITRATE - Minimal bit-rate in Bit/s used by this routing metric: 1024.

8. Data Structures

This specification extends the Link Set Tuples of the Interface Information Base, as defined in [RFC6130] section 7.1, by the following additional elements for each link tuple when being used with this metric:

L_DAT_received is a QUEUE with DAT_MEMORY_LENGTH integer elements. Each entry contains the number of successfully received packets within an interval of DAT_REFRESH_INTERVAL.

L_DAT_total is a QUEUE with DAT_MEMORY_LENGTH integer elements. Each entry contains the estimated number of packets transmitted by the neighbor, based on the received packet sequence numbers within an interval of DAT_REFRESH_INTERVAL.

L_DAT_hello_time is the time when the next hello will be expected.

L_DAT_hello_interval is the interval between two hello messages of the links neighbor as signaled by the INTERVAL_TIME TLV [RFC5497] of NHDP messages [RFC6130].

L_DAT_lost_hello_messages is the estimated number of lost hello messages from this neighbor, based on the value of the hello interval.

L_DAT_rx_bitrate is the current bitrate of incoming unicast traffic for this neighbor.

Methods to obtain the value of L_DAT_rx_bitrate are out of the scope of this specification. Such methods may include static configuration via a configuration file or dynamic measurement through mechanisms described in a separate specification (e.g. [DLEP]). Any Link tuple with L_status = HEARD or L_status = SYMMETRIC MUST have a specified value of L_DAT_rx_bitrate if it is to be used by this routing metric.

When using packet sequence numbers to estimate the loss rate, the Link Set Tuples get another field:

L_DAT_last_pkt_seqno is the last received packet sequence number received from this link.

8.1. Initial Values

When generating a new tuple in the Link Set, as defined in [RFC6130] section 12.5 bullet 3, the values of the elements specified in Section 8 are set as follows:

- o L_DAT_received := 0, ..., 0. The queue always has DAT_MEMORY_LENGTH elements.
- o L_DAT_total := 0, ..., 0. The queue always has DAT_MEMORY_LENGTH elements.
- o L_DAT_last_pkt_seqno := UNDEFINED (no earlier packet received).
- o L_DAT_hello_time := EXPIRED (no earlier NHDP HELLO received).
- o L_DAT_hello_interval := UNDEFINED (no earlier NHDP HELLO received).
- o L_DAT_lost_hello_messages := 0 (no HELLO interval without packets).

9. Packets and Messages

9.1. Definitions

For the purpose of this section, note the following definitions:

- o "pkt_seqno" is defined as the [RFC5444] packet sequence number of the received packet.
- o "interval_time" is the time encoded in the INTERVAL_TIME message TLV of a received [RFC6130] HELLO message.

9.2. Requirements

An implementation of OLSRv2 using the metric specified by this document MUST include the following parts into its [RFC5444] output:

- o an INTERVAL_TIME message TLV in each HELLO message, as defined in [RFC6130] section 4.3.2.

9.3. Link Loss Data Gathering

While this metric was designed for measuring the packet loss based on the [RFC5444] packet sequence number, some implementations might not be able to add the packet sequence number to their output.

9.3.1. Packet Sequence based link loss

An implementation of OLSRv2, using the metric specified by this document with packet sequence based link loss, MUST include the following element into its [RFC5444] output:

- o an interface specific packet sequence number as defined in [RFC5444] section 5.1 which is incremented by 1 for each outgoing [RFC5444] packet on the interface.

For each incoming [RFC5444] packet, additional processing MUST be carried out after the packet messages have been processed as specified in [RFC6130] and [OLSRV2].

[RFC5444] packets without packet sequence number MUST NOT be processed in this way by this metric.

The router MUST update the Link Set Tuple corresponding to the originator of the packet:

1. If L_DAT_last_pkt_seqno = UNDEFINED, then:
    1. L_DAT_received[TAIL] := 1.
    2. L_DAT_total[TAIL] := 1.
2. Otherwise:
    1. L_DAT_received[TAIL] := L_DAT_received[TAIL] + 1.
    2. diff := seq_diff(pkt_seqno, L_DAT_last_pkt_seqno).
    3. If diff > DAT_SEQNO_RESTART_DETECTION, then:
        1. diff := 1.
    4. L_DAT_total[TAIL] := L_DAT_total[TAIL] + diff.
3. L_DAT_last_pkt_seqno := pkt_seqno.
4. If L_DAT_hello_interval != UNDEFINED, then:

1. L_DAT_hello_time := current time + (L_DAT_hello_interval * DAT_HELLO_TIMEOUT_FACTOR).

5. L_DAT_lost_hello_messages := 0.

9.3.2. HELLO based Link Loss

A metric might just use the incoming NHDP HELLO messages of a neighbor to calculate the link loss. Because this method uses fewer events to calculate the metric, the variance of the output will increase. It might be necessary to increase the value of DAT_MEMORY_LENGTH to compensate for this.

For each incoming HELLO message, after it has been processed as defined in [RFC6130] section 12, the Link Set Tuple as defined in section 7.1 corresponding to the incoming HELLO message must be updated.

1. L_DAT_received[TAIL] := L_DAT_received[TAIL] + 1.

2. L_DAT_total[TAIL] := L_DAT_total[TAIL] + 1.

3. L_DAT_lost_hello_messages := 0.

9.3.3. Other Measurement of Link Loss

Instead of using incoming [RFC5444] packets or [RFC6130] messages, the routing daemon can also use other sources to measure the link layer lossrate (e.g. [DLEP]).

To use a source like this with the DAT metric, the routing daemon has to add incoming total traffic (or the sum of received and lost traffic) and lost traffic to the queued elements in the extension of the Link Set Tuple defined in Section 8 corresponding to originator of the traffic.

The routing daemon should also set L_DAT_lost_hello_messages to zero every times new packages arrive.

9.4. HELLO Message Processing

For each incoming HELLO Message, after it has been processed as defined in [RFC6130] section 12, the Link Set Tuple corresponding to the incoming HELLO message must be updated.

Only HELLO messages with an INTERVAL_TIME message TLVs must be processed.

1. L_DAT_hello_interval := interval_time.

10. HELLO Timeout Processing

When L_DAT_hello_time has timed out, the following step MUST be done:

1. L_DAT_lost_hello_messages := L_DAT_lost_hello_messages + 1.

2. L_DAT_hello_time := L_DAT_hello_time + L_DAT_hello_interval.

11. Metric Update

Once every DAT_REFRESH_INTERVAL, all L_in_metric values in all Link Set entries MUST be recalculated:

1. sum_received := sum(L_DAT_total).

2. sum_total := sum(L_DAT_received).

3. If L_DAT_hello_interval != UNDEFINED and L_DAT_lost_hello_messages > 0, then:

    1. lost_time_proportion := L_DAT_hello_interval * L_DAT_lost_hello_messages / DAT_MEMORY_LENGTH.

    2. sum_received := sum_received * MAX (0, 1 - lost_time_proportion);

4. If sum_received < 1, then:

    1. L_in_metric := MAXIMUM_METRIC, as defined in [OLSRV2] section 5.6.1.

5. Otherwise:

    1. loss := sum_total / sum_received.

    2. If loss > DAT_MAXIMUM_LOSS, then:

        1. loss := DAT_MAXIMUM_LOSS.

    3. bitrate := L_DAT_rx_bitrate.

    4. If bitrate < DAT_MINIMUM_BITRATE, then:

1. bitrate := DAT_MINIMUM_BITRATE.

5. L_in_metric := (2^24 / DAT_MAXIMUM_LOSS) * loss / (bitrate / DAT_MINIMUM_BITRATE).

6. remove(L_DAT_total)

7. add(L_DAT_total, 0)

8. remove(L_DAT_received)

9. add(L_DAT_received, 0)

12. IANA Considerations

This document contains no actions for IANA.

13. Security Considerations

Artificial manipulation of metrics values can drastically alter network performance. In particular, advertising a higher L_in_metric value may decrease the amount of incoming traffic, while advertising lower L_in_metric may increase the amount of incoming traffic. By artificially increasing or decreasing the L_in_metric values it advertises, a rogue router may thus attract or repulse data traffic. A rogue router may then potentially degrade data throughput by not forwarding data as it should or redirecting traffic into routing loops or bad links.

An attacker might also inject packets with incorrect packet level sequence numbers, pretending to be somebody else. This attack could be prevented by the true originator of the RFC5444 packets by adding a [RFC6622] ICV Packet TLV and TIMESTAMP Packet TLV to each packet. This allows the receiver to drop all incoming packets which have a forged packet source, both packets generated by the attacker or replayed packets.

14. Acknowledgements

The authors would like to acknowledge the network administrators from Freifunk Berlin [FREIFUNK] and Funkfeuer Vienna [FUNKFEUER] for endless hours of testing and suggestions to improve the quality of the original ETX metric for the OLSR.org routing daemon.

This effort/activity is supported by the European Community Framework

Program 7 within the Future Internet Research and Experimentation Initiative (FIRE), Community Networks Testbed for the Future Internet ([CONFINE]), contract FP7-288535.

The authors would like to gratefully acknowledge the following people for intense technical discussions, early reviews and comments on the specification and its components (listed alphabetically): Teco Boot (Infinity Networks), Juliusz Chroboczek (PPS, University of Paris 7), Thomas Clausen, Christopher Dearlove (BAE Systems Advanced Technology Centre), Ulrich Herberg (Fujitsu Laboratories of America), Markus Kittenberger (Funkfeuer Vienna), Joseph Macker (Naval Research Laboratory) and Stan Ratliff (Cisco Systems).

15. References

15.1. Normative References

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

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

[RFC5444] Clausen, T., Dearlove, C., Dean, J., and C. Adjih, "Generalized Mobile Ad Hoc Network (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.

[RFC6622] Ulrich, U. and T. Clausen, "Integrity Check Value and Timestamp TLV Definitions for Mobile Ad Hoc Networks (MANETs)", RFC 6622, May 2012.

[OLSRV2] Clausen, T., Jacquet, P., and C. Dearlove, "The Optimized Link State Routing Protocol version 2", draft-ietf-manet-olsrv2-19 , March 2013.

15.2. Informative References

[CONFINE] "Community Networks Testbed for the Future Internet (CONFINE)", 2013, <http://www.confine-project.eu>.

[DLEP] Ratliff, S., Berry, B., Harrison, G., Jury, S., and D. Satterwhite, "Dynamic Link Exchange Protocol (DLEP)", draft-ietf-manet-dlep-04 , March 2013.

[MOBICOM03] De Couto, D., Aguayo, D., Bicket, J., and R. Morris, "A High-Throughput Path Metric for Multi-Hop Wireless Routing", Proceedings of the MOBICOM Conference , 2003.

[MOBICOM04] Richard, D., Jitendra, P., and Z. Brian, "Routing in Multi-Radio, Multi-Hop Wireless Mesh Networks", Proceedings of the MOBICOM Conference , 2004.

[OLSR.org] "The OLSR.org OLSR routing daemon", 2013, <http://www.olsr.org/>.

[FREIFUNK] "Freifunk Wireless Community Networks", 2013, <http://www.freifunk.net>.

[FUNKFEUER] "Austria Wireless Community Network", 2013, <http://www.funkfeuer.at>.

Appendix A. OLSR.org metric history

The Funkfeuer [FUNKFEUER] and Freifunk networks [FREIFUNK] are OLSR-based [RFC3626] or B.A.T.M.A.N. based wireless community networks with hundreds of routers in permanent operation. The Vienna Funkfeuer network in Austria, for instance, consists of 400 routers (around 600 routes) covering the whole city of Vienna and beyond, spanning roughly 40km in diameter. It has been in operation since 2003 and supplies its users with Internet access. A particularity of the Vienna Funkfeuer network is that it manages to provide Internet access through a city wide, large scale Wi-Fi mesh network, with just a single Internet uplink.

Operational experience of the OLSR project [OLSR.org] with these networks have revealed that the use of hop-count as routing metric leads to unsatisfactory network performance. Experiments with the ETX metric [MOBICOM03] were therefore undertaken in parallel in the Berlin Freifunk network as well as in the Vienna Funkfeuer network in 2004, and found satisfactory, i.e., sufficiently easy to implement and providing sufficiently good performance. This metric has now been in operational use in these networks for several years.

The ETX metric of a link is the estimated number of transmissions required to successfully send a packet (each packet equal to or smaller than MTU) over that link, until a link layer acknowledgement is received. The ETX metric is additive, i.e., the ETX metric of a path is the sum of the ETX metrics for each link on this path.

While the ETX metric delivers a reasonable performance, it doesn't handle well networks with heterogeneous links that have different bitrates. Since every wireless link, when using ETX metric, is characterized only by its packet loss ratio, the ETX metric prefers long-ranged links with low bitrate (with low loss ratios) over short-ranged links with high bitrate (with higher but reasonable loss ratios). Such conditions, when they occur, can degrade the performance of a network considerably by not taking advantage of higher capacity links.

Because of this the OLSR.org project has implemented the Directional Airtime Metric for OLSRv2, which has been inspired by the Estimated Travel Time (ETT) metric [MOBICOM04]. This metric uses an unidirectional packet loss, but also takes the bitrate into account to create a more accurate description of the relative costs or capabilities of mesh links.

Authors' Addresses

Henning Rogge
Fraunhofer FKIE

Email: henning.rogge@fkie.fraunhofer.de
URI: http://www.fkie.fraunhofer.de

Emmanuel Baccelli
INRIA

Email: Emmanuel.Baccelli@inria.fr
URI: http://www.emmanuelbaccelli.org/

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

J. Yi
LIX, Ecole Polytechnique
JA. Cordero
ICTEAM, Universite catholique de Louvain
T. Clausen
LIX, Ecole Polytechnique

January 9, 2015

Jitter Consideration for Reactive Protocols in Mobile Ad Hoc Networks (MANETs)
draft-yi-manet-reactive-jitter-04

Abstract

This document provides recommendations for jittering (randomly timing) of routing control message transmission, especially route request dissemination, in reactive protocols of Mobile Ad Hoc Networks, to reduce the probability of collisions, decrease routing overhead, and help finding the optimum paths in the network.

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 July 13, 2015.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Applicability Statement 4
4. Problem Statement 4
5. Reactive Jitter 6
5.1. Window Jitter for Hop-count Metric 7
5.2. Window Jitter for Generic Metric 7
6. Implementation Status 7
6.1. Implementation by Ecole Polytechnique 8
7. IANA Considerations 8
8. References 8
8.1. Normative References 8
8.2. Informative References 8
Authors' Addresses 10

1. Introduction

Jitter - randomly modifying timing of packet transmissions - is RECOMMENDED to be used in MANETs [RFC5148], to avoid simultaneous packet transmission by neighboring routers - something which might result packet losses due to link-layer collisions.

In [RFC5148], it is RECOMMENDED that in a protocol with regularly scheduled messages, event-triggered message, schedule reset, forwarding, etc, a deliberated random variation in time (jitter) SHOULD be employed. If a message transmission is scheduled, or triggered at time t, a random value between zero and maximum timing variation (denoted MAXJITTER) is chosen to reduce or increase the time of that transmission.

Jitter has been used in NHDP [RFC6130], for periodic HELLO message transmission, and OLSRv2 [RFC7181], for triggered and periodic Topology Control (TC) message transmissions.

In reactive protocols such as AODV [RFC3561], DSR [RFC4728] and LOADng [I-D.clausen-lln-loadng], packet loss due to concurrent transmissions by neighboring routers are also a concern, in particular for Route Request message (RREQ) dissemination. This, because RREQ transmissions in neighbor routers are triggered by a single event: reciept of RREQ message to be flooded through the network as part of the route disovery process. However, unlike TC message dissemination in OLSRv2, forwarding of RREQ message has another objective: to discover the best path from the source to the destination. It has been observed, however, that the jitter mechanism as, defined in [RFC5148] and if applied directly, in some cases may result in inferier paths, or unnecessary RREQ retransmissions.

This document analyzes the limitation of [RFC5148] when it is applied to reactive protocols, and then introduces a "window" jitter mechanism, which can help reducing RREQ message retransmission and finding the optimum paths.

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 [RFC5148].

3. Applicability Statement

This document describes a jitter mechanism that is applicable to the Route Discovery process in reactive MANET protocols, such as Route Request (RREQ) flooding in AODV [RFC3561], DSR [RFC4728], LOADng [I-D.clausen-lln-loadng], or AODVv2 [I-D.ietf-manet-aodvv2].

The jitter mechanism, as described in [RFC5148], is originally intended for application where the underlying medium access control and lower layers do not provide effective mechanisms to avoid packet collisions when faced with concurrent transmission by neighboring routers. This document handles the situation of "Message Forwarding", described in section 5.3 of [RFC5148]. In addition to collision avoidance by way of a random delay in transmission of RREQ messages, the document also considers:

Route Discovery of Optimal Paths When RREQ messages are flooded through the network, the path cost (e.g., hop count or any other link metrics) is also accumulated and recorded. The destination of the RREQ will reply it with a Route Reply (RREP) message. However, the RREQ copy that arrives first may not always be the one which has traversed the optimal path, with respect to the metric used. It has been observed that, in some cases, this is exacerbated by the use of [RFC5148].

Route Discovery Overhead In classic flooding, duplicate message are dropped by intermediate routers, and not retransmitted. However, for RREQ flooding, in which the cumulated path cost is carried in the RREQ, intermediate routers may need to transmit the same RREQ message multiple times, when the shortest (according to the metric in use) path is desired. For example, when an RREQ arrives from the same source to the same destination, and with same sequence number as previously forwarded RREQ, but with a lower path cost. Again, this is exacerbated by the used of [RFC5148].

This document suggest a "window jitter" mechanism, which can help discovering the optimal paths in reactive protocols, and, simultaneously, can reduce the route discovery overhead, with the cost of slightly increasing the route discovery delay.

4. Problem Statement

[RFC5148] recommends applying jitter to a forwarded message by reducing the time of its emission by a small, random duration in the mediums where transmission collisions are possible. This delay is recommended to be generated uniformly in an interval between zero and MAXJITTER. This has been show to work well in message flooding,

where the goal simply is that all routers get a copy of the unmodified message, such as is the case for TC messages in OLSRv2 [RFC7181].

In reactive protocols, RREQ message from a source are flooded through the network, carrying a "path cost" field, modified by intermediate routers when the message is forwarded. This allows the destination sought through the Route Discovery process to identify which copy of the RREQ has traveled through the "least cost path" (according to the metric in use in the network), and select that path for generating a RREP and installing a routing path. It is, therefore, unfortunate if the copy of the RREQ arriving via the "least cost path" is received later than a RREQ over a path with a higher cost due to inappropriate application of a jitter mechanism.

Consider the topology shown in Figure 1, and assume that router A floods an RREQ to identify a path towards router D. Hop count metric is used in this example. If no jitter is used, the RREQ would reach router D through path {A-E-D} faster than the path {A-B-C-D}, assuming that processing time and transmission time at each intermediate router (Ti) are similar.

If [RFC5148] is applied, a uniform random distribution [0, MAXJITTER] is used at each hop to determine an additional delay before retransmission, see [RFC5148] section 5.3, the RREQ copy sent through the longer path (in number of hops), may reach the destination faster than the RREQ over shorter path. For example, in Figure 1, the MAXJITTER is 500ms (MAXJITTER is normally chosen to be much greater than transmission time Ti, to avoid collision. Therefore, Ti is neglectable if jitter is used). If Jitter at E (JitterE) is chosen to be 300ms, JitterB is 100ms, and JitterC is 150ms, the RREQ though the longer path (A-B-C-D) would reach D faster than the shorter path (A-E-D). This phenomenon is called "delay inversion".

```
                    JitterE=300ms
             /----- E------\
            /               \
          A                  D
            \               /
              B--------C---/
          JitterB=100ms  JitterC=150ms
```

Figure 1: An example of delay inversion. The RREQ through longer paths may traverl faster, if jitter in RFC5148 is used.

In this case, router D would reply to the RREQ with an RREP that

advertises path (A-B-C-D), which is suboptimal. When the RREQ traversing (A-E-D) reaches router D, router D would reply again to the cope of that same RREQ again with the shorter path (assuming shorter path is preferred).

If router D is not the final destination of the RREQ, but only an intermediate router that forwards the RREQ message, there are two possible approches used in different protocols:

- o For AODV [RFC3561] and DSR [RFC4728], the intermediate routers only forward the first copy of the RREQ received - all the later ones are discarded, even if the later one carries paths with lower costs. This would lead to sub-optimal paths discovered.

- o For LOADng [I-D.clausen-lln-loadng], the intermediate routers would always retransmit the RREQ carrying better paths that comes later. In the example of Figure 1, router D would retransmit the RREQ received from JitterE also, which result in one more flooding (not only at router D, but to the whole network).

The example above illustrates that a "naive" application of [RFC5148] for reactive protocols may present some drawbacks, in terms of path sub-optimality and/or control traffic inefficiency.

The "delay inversion" phenomenon stems from the fact that the delay imposed by intermediate routers can not really reflect the link metrics of related routers. Even jitter is not used at all, it will happen -- and the application of jitter amplified such phenomenon. Using other link metrics, or reduced topology mechanisms (like Connected Dominating Set) will not mitigate the problem. On the other hand, the delay inversion has direct relationship with the network size -- it is more often as the network grows [NBis2013].

5. Reactive Jitter

In order to reduce the impact of the "delay inversion" phenomenon, described in Section 4, the notion of window jitter introduced in this section. The purpose of "window jitter" is to attempt at "penalizing long paths" more than short paths (in the aspect of hop count, or other metrics), and it is RECOMMENDED that this be employed for Route Discovery (RREQ flooding). In addition to the MAXJITTER, a lower bound of jitter is applied, with this lower bound depending on the metrics used.

5.1. Window Jitter for Hop-count Metric

For protocols like AODV [RFC3561], DSR [RFC4728], LOADng [I-D.clausen-lln-loadng] or AODVv2 [I-D.ietf-manet-aodvv2], the hop-count metric is supported by default, i.e., a path with a lower hop count is better than a path with more hop counts.

When a router forwards an RREQ message, it SHOULD be jittered by delaying it by a random duration. This delay SHOULD be generated uniformly in an interval between MAXJITTER/2 and MAXJITTER.

5.2. Window Jitter for Generic Metric

While the hop count metric is straightforward and easy to implement, operational experience has revealed that the used of hop-count as routing metric often leads to unsatisfactory network performance. Reactive protocols like LOADng [I-D.clausen-lln-loadng] thus support using metrics other than hop count, such as ETX [I-D.funkfeuer-manet-olsrv2-etx] or DAT [I-D.ietf-manet-olsrv2-dat-metric].

For those generic metrics, given a link quality indicator LQ between (0,1) (1 indicates highest quality links), jitter values SHOULD be assigned under a generalized window jitter distribution uniformly within the interval between (1-LQ)MAXJITTER and MAXJITTER.

6. Implementation Status

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 [RFC6982]. 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 [RFC6982], "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".

There is currently one publicly-known implementation of window jitter specified in this document.

6.1. Implementation by Ecole Polytechnique

This implementation is developed by the Networking Group at Ecole Polytechnique and applied to LOADng [I-D.clausen-lln-loadng] for RREQ message flooding. It can run over real network interfaces, and can also be integrated with the network simulator NS2. It is a Java implementation, and can be used on any platform that includes a Java virtual machine.

The implementation is based on the -04 revision of this document, and makes up only a handful of lines of code - in addition to the core LOADng protocol implementation. Both analytical and simulation results have been published in [IEEE_WiOpt2013] and [NBis2013]. The results show that if the shortest paths are desired, the window jitter can reduce the RREQ flooding overhead by 50%, as compared to a naive application of [RFC5148].

7. IANA Considerations

This document contains no actions for IANA.

8. References

8.1. Normative References

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

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

8.2. Informative References

[I-D.clausen-lln-loadng]
Clausen, T., Verdiere, A., Yi, J., Niktash, A., Igarashi, Y., Satoh, H., Herberg, U., Lavenu, C., Lys, T., and J. Dean, "The Lightweight On-demand Ad hoc Distance-vector Routing Protocol - Next Generation (LOADng)", draft-clausen-lln-loadng-12 (work in progress), October 2014.

[I-D.funkfeuer-manet-olsrv2-etx]

Rogge, H., Baccelli, E., and A. Kaplan, "Packet Sequence Number based ETX Metric for Mobile Ad Hoc Networks", draft-funkfeuer-manet-olsrv2-etx-01 (work in progress), March 2010.

[I-D.ietf-manet-aodvv2]
Perkins, C., Ratliff, S., Dowdell, J., and L. Steenbrink, "Dynamic MANET On-demand (AODVv2) Routing", draft-ietf-manet-aodvv2-06 (work in progress), December 2014.

[I-D.ietf-manet-olsrv2-dat-metric]
Rogge, H. and E. Baccelli, "Packet Sequence Number based directional airtime metric for OLSRv2", draft-ietf-manet-olsrv2-dat-metric-04 (work in progress), December 2014.

[IEEE_WiOpt2013]
Yi, J., Cordero, J., and T. Clausen, "Optimization of Jitter Configuration for Reactive Route Discovery in Wireless Mesh Networks", Proceedings of IEEE WiOpt 2013, IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2013.

[NBis2013]
Cordero, J., Yi, J., and T. Clausen, "Jitter Considerations in On-demand Route Discovery for Mobile Ad Hoc Networks", Proceedings of the 16th International Conference on Netwokr-Based Information Systems, 2013.

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

[RFC4728] Johnson, D., Hu, Y., and D. Maltz, "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4", RFC 4728, February 2007.

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

[RFC6982] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", RFC 6982, July 2013.

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

RFC 7181, April 2014.

Authors' Addresses

Jiazi Yi
LIX, Ecole Polytechnique

Phone: +33 1 6933 4031
Email: jiazi@jiaziyi.com
URI: http://www.jiaziyi.com/

Juan Antonio Cordero Fuertes
ICTEAM, Universite catholique de Louvain

Email: j.a.cordero@gmail.com

Thomas Clausen
LIX, Ecole Polytechnique
91128 Palaiseau Cedex,
France

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