roll
Internet-Draft
Intended status: Standards Track
Expires: June 10, 2017

P. van der Stok
consultant
AR. Sangi
Huawei Technologies

December 7, 2016

MPL Forwarder Select (MPLFS)
draft-ietf-roll-mpl-forw-select-00

Abstract

This document describes a Forwarder Selection (MPLFS) protocol for the Multicast Protocol for Low-Power and lossy Networks (MPL) to reduce the density of forwarders such that the number of forwarded messages is reduced.

The protocol uses Trickle to distribute link-local information about the identity of the neighbours of the nodes that have MPL-enabled interfaces. In the end-state all nodes are connected to a minimum number, N_DUPLICATE, of forwarders, where N_DUPLICATE is application dependent, and there is a path between any two forwarders.

Note

Discussion and suggestions for improvement are requested, and should be sent to roll@ietf.org.

Copyright Notice

Table of Contents

1. Introduction

The Multicast Protocol for Low-Power and Lossy Networks (MPL) [RFC7731] is designed for small devices interconnected by a lossy wireless network such as IEEE 802.15.4. A seed sends a multicast message with a realm-local scope, admin-local scope or higher as specified in [RFC4291].

Forwarders forward these messages with an increasing interval size. When the density of the forwarders is high, the DATA_MESSAGE_K (k) parameter stops the retransmission of messages in a given Trickle interval when more than k copies of a message have been received. This mechanism prevents network saturation but leads to difficult to predict end-to-end data transmission delays. For IoT short predictable delays are wanted and the k parameter is set to a large

value. When the density of forwarders is high and k is large, the message may be forwarded by a high number of forwarders that conflict on the link. With extreme forwarder densities, small Trickle intervals, and k is infinite, just sending one multicast message may lead to an overload of the communication medium.

The number of forwarded messages can be reduced by selecting a minimal set of forwarders. However, for large networks, manually selecting the forwarders is much work, and changing network conditions and configurations make the manual selection an unwanted burden to the network management.

This document specifies a protocol that selects the forwarders such that each MPL-enabled device is connected to N_DUPLICATE forwarders, where N_DUPLICATE > 0 can be set. The parameter N_DUPLICATE determines how much path redundancy there is for each MPL message. The value of N_DUPLICATE should be at least 1, because a value of 0 has as result that no forwarder exists in the network during the protocol execution. Moreover, the protocol is distributed and dynamic in nature to face a continuously changing topology.

The protocol is inspired by the work described for NeighbourHood Discovery (NHDP) [RFC6130] and Simple Multicast Forwarding (SMF) [RFC6621]. In contrast to the "HELLO" messages described in [RFC6130], this protocol uses the Trickle protocol [RFC6206] to multicast link-local messages, containing a CBOR payload [RFC7049].

A simulation of the protocol has been done for some example topologies. The number of forwarders allocated by the protocol are compared in Appendix A with the optimum (lowest) number of forwarders for these topologies.

1.1. Terminology

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

Readers of this specification should be familiar with all the terms and concepts discussed in [RFC7731]. The following terms are defined in this document:

synchronization time The moment that a node can change its state at messages reception.

The following list contains the abbreviations used in this document.

XXXX: TODO, and others to follow.

2. Protocol overview

Nodes participating in MPLFS exchange messages with a format that is described in Section 6. A participating node communicates to all its neighbours with link-local multicast messages as described in Section 4.

Failing links provide a lot of instability. Only messages sent over stable links are accepted. Section 4 describes a mechanism to refuse messages from unstable links.

Each node maintains a set of 1-hop neighbours where each neighbour contains information about its own 1-hop neighbours. On the basis of the contents of the set, the node can decide to become a Forwarder or not, as explained in Section 5.

The protocol never ends, with a minimum frequency of exchanging maintenance messages specified by an interval size of I_MAX_SELECT. When the set of links is stable, the protocol stabilizes such that there is a path between any two forwarders, and every MPL-enabled node is connected to at least N_DUPLICATE MPL forwarders (when existing), where N_DUPLICATE > 0. N_DUPLICATE can be set dependent on the application requirements. With N_DUPLICATE = 2, it is expected that a multicast message arrives at an intended recipient with very high probability.

Nodes have a state that determines whether they are forwarder or not. The state of a node can only be changed by the node itself. To avoid race conditions, (e.g. two nodes simultaneously decide to be forwarder, while only one is intended) the node with the highest address of all 1-hop neighbours is the only one allowed to change state. Unlike [RFC5614], that considers 3-tuple (Router Priority, MDR Level and Router ID) to allow self state change, this approach only takes into account the node address. Consequently, only k-hop neighbours, with k > 2, can change state simultaneously, and the 1-hop and 2-hop neighbours of a given node can change state one by one.

3. Data sets

Each node, n_0, maintains a state with two values: Fixed Forwarder (FF) and No Forwarder (NF). Each node also maintains a set, S1_0, containing information about n_0's 1-hop neighbours and n_0 itself. Each entry, n_i, in S1_0 has the following attributes:

address of n_i: the address can be the 64 bit IPv6 address or the short 16 bit address.

average-rssi-in: the average rssi of the messages received by n_0 from n_i.

average-rssi-out: the average rssi of the messages received by n_i from n_0.

nr_FF: the number of neighbours, n_ij, of n_i (including n_i) with state = FF.

nr_Under: the number of neighbours, n_ij, of n_i with nr_FF < N_DUPLICATE.

nr_Above: the number of neighbours, n_ij, of n_i with nr_FF > N_DUPLICATE.

size: the size of S1_i, the set of 1-hop neighbours of n_i.

state: the state of n_i.

4. Neighbor distribution

A participating node multicasts link-local so-called "neighbour messages" with the Trickle protocol. It uses the multicast address LINK_LOCAL_ALL_NODES as destination. The message sent by n_0 contains the contents of S1_0. The contents of a "neighbour message" from n_i received by n_0 is called M_i. The rssi value associated with the reception of the "neighbour message" is called new_rssi. The message M_i contains information about the set S1_i with the following attributes for all nodes in S1_i:

- address
- average-rssi-in
- nr_FF
- nr_Under
- nr_Above
- size
- state

On reception of M_i from n_i for the first time, the receiving node adds n_i to S1_0, and sets average-rssi-in of n_i in S1_0 to new_rssi. For all following messages from n_i, the average-rssi-in

for n_i is calculated in the following way: average-rssi-in := (average-rssi-in*WEIGHT_AVERAGE + new_rssi)/(WEIGHT_AVERAGE+1).

The neighbour nodes of M_i are called n_ij. For the n_ij with an address that is equal to the address of n_0: the value of average-rssi-out of n_0 is set equal to the value of average-rssi-in of n_ij.

The contents of n_0 is updated with the contents of M_i. Updating includes the following actions:

- Add n_i to S1_0, if n_i not present in S1_0.

- Set size of n_i equal to the number of entries in M_i.

- When n_ij.address = n_j.address, copy the values of nr_Under, nr_Above, nr_FF, and state of n_ij to n_j.

When the average-rssi-in and average-rssi-out values of n_i have been averaged over more than WEIGHT_AVERAGE messages, and the averaged RSSI values are smaller than MAXIMUM_RSSI, n_i is called "valid".

5. Selection Algorithm

The protocol aims at allocating forwarders in the densest part of the network. A dense network is characterized by a high number of neighbours. Therefore, the protocol attempts to assign status FF to the nodes with the highest number of neighbours that have less than N_DUPLICATE neighbours with state = FF (nr_FF < N_DUPLICATE).

It is required that a path exists between every two forwarders to prevent network partitioning. Therefore, a node can become forwarder iff one of its neighbours is a forwarder. The consequence of this rule is that one so-called "source-forwarder" must be selected by the network administrator. A likely choice for the "source-forwarder" is the border router.

At the start of the selection protocol the node, n_0, sets its state to No Forwarder (NF). It sets the Trickle timer to its minimum interval, I_MIN_SELECT, and starts multicasting M_0 to its neighbours. Every time entries are added to, or removed from, S1_0, the Trickle interval timer is set to I_MIN_SELECT.

The executing node, n_0, calculates the following parameter values:

- max-under is the maximum of the nr_Under attribute of all valid n_i in S1_0.

o max_address_u is the maximum of the addresses of valid n_i with nr_Under = max-under.

o max_address_a is the maximum of the addresses of all valid n_i.

o connected is true iff nr_FF of all neighbouring forwarders is equal to nr_FF of n_0.

The information about the state and the nr_Under value of the neighbours comes in asynchronously. Time is needed before the state in a node correctly reflects the state changes of the network. A node can change its state when during the reception of messages of all neighbours, the value of nr_Under has not changed.

To calculate its new state, n_0 does the following:

When the state is NF, a neighbour with state = FF exists, and address = max-address_u:
set state to FF.

When the state is FF, nr_Above = size S1_0, connected is true, and address = max_address_a:
set state to NF.

6. CBOR payload

The payload format is /application/cbor [RFC7049]. The contents of the message is a CBOR array (Major type 4) of CBOR arrays composed of neighbour address, rssi value, size of S1_i, forwarder state, nr_FF, nr_Under, and nr_Above. Assuming two neighbours, in diagnostic JSON the payload looks like:

```
[
[address_1, average-rssi-in_1, size_1, state_1,
    nr_FF_1, nr_Under_1, nr_Above_1],
[address_2, average-rssi-in_2, size_2, state_2,
    nr_FF_2, nr_Under_2, nr_Above_2]
]
```

Figure 1: CBOR payload

7. Default parameter values

The following text recommends default values for the MPLFS protocols.

I_MIN_SELECT = 0,2; minimum Trickle timer interval.

I_MAX_SELECT = 10; maximum Trickle timer interval.

DATA_MESSAGE_K > 10; the redundancy constant.

WEIGHT_AVERAGE = 10; number of values to average rssi.

MAXIMUM_RSSI = 3; maximum acceptable average rssi value.

N_DUPLICATE = 2; requested number of MPL forwarder neighbours for every MPL enabled node.

8. Acknowledgements

We are very grateful to Yasuyuki Tanaka for pointing out the missing discussion on k-values.

9. Changelog

Changes from vanderstok version to ietf version 00

- o copied from vanderstok-roll-forw-select-01
- o added discussion on DATA_MESSAGE_K
- o compared generated forwarder set with optmal set

10. References

10.1. Normative References

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

[RFC6206] Levis, P., Clausen, T., Hui, J., Gnawali, O., and J. Ko, "The Trickle Algorithm", RFC 6206, DOI 10.17487/RFC6206, March 2011, <http://www.rfc-editor.org/info/rfc6206>.

[RFC7049] Bormann, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", RFC 7049, DOI 10.17487/RFC7049, October 2013, <http://www.rfc-editor.org/info/rfc7049>.

[RFC7731] Hui, J. and R. Kelsey, "Multicast Protocol for Low-Power and Lossy Networks (MPL)", RFC 7731, DOI 10.17487/RFC7731, February 2016, <http://www.rfc-editor.org/info/rfc7731>.

10.2. Informative References

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

[RFC5614] Ogier, R. and P. Spagnolo, "Mobile Ad Hoc Network (MANET) Extension of OSPF Using Connected Dominating Set (CDS) Flooding", RFC 5614, DOI 10.17487/RFC5614, August 2009, <http://www.rfc-editor.org/info/rfc5614>.

[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>.

[RFC6621] Macker, J., Ed., "Simplified Multicast Forwarding", RFC 6621, DOI 10.17487/RFC6621, May 2012, <http://www.rfc-editor.org/info/rfc6621>.

Appendix A. example forwarder allocations

The protocol has been simulated with omnet++ on four different network topologies. The value of N_DUPLICATE is set to two. Each topology is a rectangular grid, where each grid point is occupied by a node. The distance, D, between two horizontal and vertical direct neighbours is the same. The topologies differ in number of grid points and D. A topology of 9x9 grid points shows how the selection criteria of the protocol affect the distribution of the forwarders when the number of edge nodes is small compared to the number of interior nodes. A topology of 3x20 nodes shows the distribution with a majority of edge nodes. Two radio ranges of the link have been selected: (1) a range of 3,5 times D, and (2) a range of 7 times D. The node density of case 2 is four times higher than for case 1. The two example ranges nicely show how the average distance between the forwarders increases with increasing range.

Table 1 shows the comparison results. Column one indicates the grid topology, Columns 2 indicates the radio range, column 3 indicates the number of protocol forwarders as delivered by the simulation, column four indicates the minimum number of forwarders needed to arrive at at least N-DUPLICATE forwarders for all nodes, column five cites the number of nodes.

topology	range	forwarders	optimum	nodes
9x9	3,5 D	10	9	81
9x9	7 D	3	3	81
3x20	3,5 D	8	8	60
3x20	7 D	5	5	60

Table 1: comparison of protocol set size and optimal set

The table gives confidence that for many cases the size of the protocol forwarder set will be close to the size of an optimal set.

Authors' Addresses

Peter van der Stok
consultant

Phone: +31-492474673 (Netherlands), +33-966015248 (France)
Email: consultancy@vanderstok.org
URI: www.vanderstok.org

Abdur Rashid Sangi
Huawei Technologies
No.156 Beiqing Rd. Haidian District
Beijing 100095
P.R. China

Email: rashid.sangi@huawei.com