

Network Working Group
Internet Draft
Expiration Date: September 2003
File Name: [draft-retana-marp-02.txt](#)

Alvaro Retana
Russ White
Cisco Systems, Inc.
March 2003

MultiAccess Reachability Protocol (MARP)

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

Internet Drafts are working documents of the Internet Engineering Task Force (IETF), its Areas, and its Working Groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet Drafts are draft documents valid for a maximum of six months. Internet Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "work in progress".

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Abstract

This document defines a protocol to quickly determine the existence or aliveness of devices attached to a shared media (broadcast) subnet. While the examples used are narrowly defined for simplicity, the protocol could be applied to other situations as well.

[1](#). Specification of Requirements

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](#)].

2. Motivation

There is a great deal of interest in discovering when a device drops off of a broadcast (shared) media link for various purposes, not limited to:

- o Loss of routing protocol neighbors. Routing protocols would like to discover the loss of a neighbor as quickly as possible so they can reconverge around the topology change, dropping as little traffic as possible.
- o Loss of a server. If multiple servers, offering the same service, exist on a segment, a device which is load balancing traffic between those servers would like to know as soon as one of them fails.

Towards this end, several solutions ([\[ISIS SHORT\]](#), [\[LSP PING\]](#), [\[FLIP\]](#) and [\[PLP\]](#), for example) have been designed, most (or all) of which rely on some sort of "fast aliveness" or "fast hello" protocol to quickly determine the failure of a node on a shared media segment. There is some question about the scalability of such protocols, since there could be hundreds of devices on a single high speed broadcast network, and a single device could be connected to hundreds of broadcast networks.

Most devices in today's networks are not connected to a true broadcast segment (such as a 10base5 coax cable), but are instead connected to a layer 2 switch (using point-to-point connections) that can determine if a device is still alive based on the carrier detect circuitry at the physical or data link layers. It should be possible to somehow harness this immediate and constant status information to inform other network devices about state changes for a particular device.

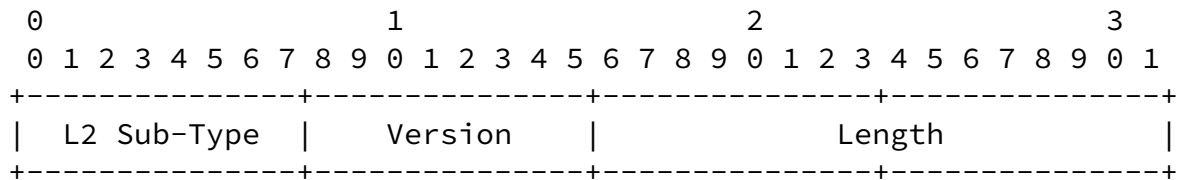
This document defines the MultiAccess Reachability Protocol (MARP), which allows for the fast notification of loss of connectivity to devices attached to a shared media (broadcast) subnet.

3. MARP Packet Format

MARP runs directly over layer 2. The data portion of the packet consists of a header and TLVs as described in this section.

3.1. The MARP Header

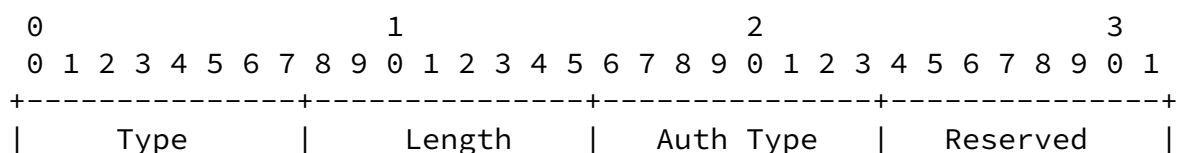
The header, as well as all the other components, is simplified as much as possible to keep the protocol light weight.



- o L2 Sub-Type (1 octet): reserved field for use if the underlying layer 2 media requires it. Otherwise, it SHOULD be sent as 0 and ignored by the receiver.
- o Version (1 octet): the version of the protocol; current value is 1.
- o Length (2 octets): total length of the MARP packet in octets.

3.2. The Authentication TLV

The Authentication TLV is used to optionally provide authentication information to the receiver.



```

+-----+-----+-----+-----+
|                                     Authentication String...                                     |
+-----+-----+-----+-----+

```

- o Type (1 octet): the type of the TLV. The Authentication TLV has a type of 1.
- o Length (1 octet): the total length of the TLV in octets.
- o Authentication Type (1 octet): an unsigned integer indicating the type of authentication present (described below).

- o Reserved (1 octet): reserved for future use; SHOULD be sent as 0 and ignored by the receiver.
- o Authentication String (variable length): contains the authentication information.

The Authentication Type field serves to indicate what type of authentication is present, as well as its length.

- 0 Reserved, it MUST NOT be used.
- 1 Plain text authentication included (authentication string is 16 octets).
- 2 MD5 [[MD5](#)] authentication included (authentication string is 16 octets).

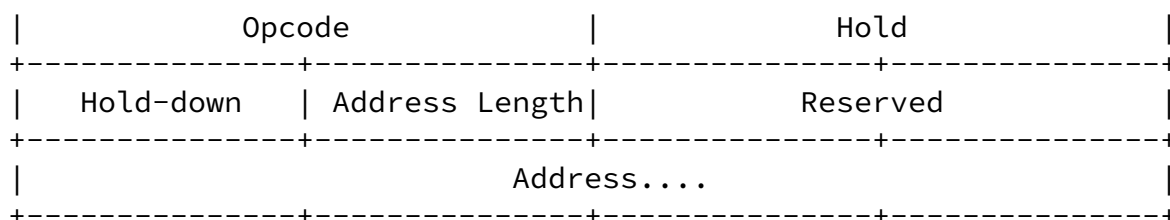
[3.3](#). The Reachability Notification TLV

The Reachability Notification TLV is used to provide information about the need for monitoring and the reachability of an address. Details are provided in the "MARP Operation" section.

```

      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
+-----+-----+-----+-----+
|      Type      |      Length      |      Reserved      |
+-----+-----+-----+-----+

```



- o Type (1 octet): the type of the TLV. The Reachability Notification TLV has a type of 2.
- o Length (1 octet): the total length of the TLV in octets.
- o Opcode (2 octets): A bit field containing information about how the packet should be handled (described below).
- o Hold (2 octets): the number of minutes the receiving device should track the list of addresses included in the packet; note that the hold time of any given entry need not match the hold time of any other entry on the network.

- o Hold-down (1 octet): the time in seconds a port which loses connectivity to the addresses listed in the packet should be held in the down state. The default value is 5 sec.
- o Address Length (1 octet): length in octets of each address included in this TLV.
- o Address (variable length - more than one field may be present in a packet): each field contains one address. The format of the address depends on the underlying media.
- o Reserved: reserved for future use; SHOULD be sent as 0 and ignored by the receiver.

The Opcode field is used to determine how the TLV should be processed when it is received.

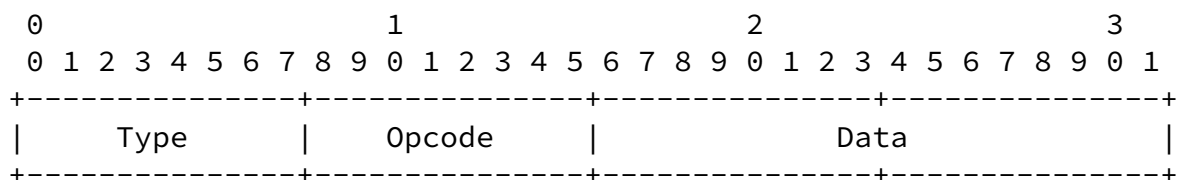
- o If the high order bit of this field is set, then the remaining 15 bits are vendor implementation specific.
- o If the high order bit of this field is not set, then the two low order bits indicate the message type:

- 00 UPDATE: MARP servers SHOULD provide notification when reachability to the address(es) listed fails. A MARP server may have an upper limit to the number of addresses it can track, but this limit SHOULD NOT be lower than 100 per broadcast domain.
- 01 NOTIFY_HARD: Reachability to the address(es) listed has failed.
- 10 NOTIFY_SOFT: Reachability to the address(es) listed may have failed.
- 11 NACK: The address(es) listed cannot be tracked by a MARP server at this time.

3.4. The Fast Reachability Verification TLV

As its name suggests, the Fast Reachability Verification TLV is used to verify the reachability of a node. Details are provided in the "MARP Operation" section.

In order to guarantee a fast response, this TLV SHOULD be the only one present in the MARP packet.



- o Type (1 octet): the type of the TLV. The Fast Reachability Verification TLV has a type of 3.
- o Opcode (1 octet): A bit field containing information about how the packet should be handled (described below).
- o Data (2 octets): User defined data.

The Opcode field is used to determine how the TLV should be processed

when it is received.

- o If the high order bit of this field is set, then the remaining 7 bits are vendor implementation specific.
- o If the high order bit of this field is not set, then the low order bit indicates the message type:
 - 0 Fast Reachability Verification Detection (FRD): the receiver MUST send the message back to the sender, indicating that it is now a Fast Reachability Verification Reply message and including a logical NOT of the information in the Data field.
 - 1 Fast Reachability Verification Reply (FRR): used to reply when an FRD is received.

[4. MARP Operation](#)

As described in this document, MARP can provide two basic services: reachability notification and fast reachability verification. These services are described in the following sections.

[4.1. Reachability Notification](#)

Reachability notification represents MARP's core service. In general, the service consists in a device (MARP server) notifying a group of other devices (MARP clients) about the loss in reachability of another device (identified by an "interesting" address).

Two sub-sections follow to discuss the operation within a MARP

client, and then a MARP server. Note that a single device MAY be a MARP server and a MARP client at the same time.

[4.1.1. MARP Client Operation](#)

A MARP client is a network device that wants to receive a notification when a peer (such as a routing protocol neighbor, for example)

is no longer reachable. The operation is as follows:

- o The MARP client compiles a list of "interesting" addresses (these addresses MUST be significant to the underlying media) that correspond to its peers. The MARP client's own address MAY be part of the list.
- o The list of "interesting" addresses is advertised using the Reachability Notification TLV with an UPDATE Opcode.
- o If a NACK message is received, the MARP client MAY use the process defined in the "Fast Reachability Verification" section to temporarily verify the reachability of any address(es) that the MARP server cannot service at the time.
- o An UPDATE message MUST be resent before the Hold Time expires. If a received UPDATE message includes some (or all) of the locally "interesting" addresses, then the Hold Time should be locally reset to prevent the transmission of unnecessary UPDATES. On the other hand, to avoid the possible effects of a lost UPDATE, they SHOULD be resent at least twice within the Hold Time.
- o A MARP client that receives a NOTIFY_HARD or NOTIFY_SOFT message MAY use this information to reset known adjacencies, check adjacency status, or take other action as deemed appropriate locally.
- o If a NOTIFY_SOFT message is received, the MARP client MAY want to verify the reachability of its peer before taking an action. To do so, the process defined in the "Fast Reachability Verification" section MAY be followed.

All the messages described in this section MUST be sent to a well-known multicast address specific to the underlying media.

A MARP server is a network device capable of tracking the reachability of devices (including itself) on the same broadcast domain. The operation is as follows:

- o If an UPDATE message is received, and the request cannot be serviced at the time (because the MARP server reached its internal limit to the number of addresses it can track, for example), then a NACK MUST be sent immediately in response. If the request can be serviced, then for each address a MARP server MUST determine whether it has reachability to it.
- o If the address is found to not be reachable, then it should be silently ignored.
- o If the address is found to be reachable, then the Hold Time MUST be set to the maximum of the current value or the time specified in the message. The Hold-down Time MUST be set to the maximum of the current value or the time specified in the message.
- o A MARP server MUST stop tracking any layer 2 addresses listed in a NOTIFY_HARD packet.
- o A MARP server SHOULD ignore any NOTIFY_SOFT packets.
- o If a MARP server detects loss of connectivity to an address it is tracking (and the Hold Time has not expired), it MUST send a notification message (NOTIFY_HARD or NOTIFY_SOFT according to the local configuration). If the loss of connectivity was due to a port failure (physical or logical), then the corresponding port SHOULD be maintained in the down state for the length of the corresponding Hold-down Time.

In conjunction with processing the messages as described, the MARP server SHOULD, if applicable, also forward them according to the local multicast forwarding rules.

[4.2.](#) Fast Reachability Verification

Fast reachability verification is an optional MARP service that uses the Fast Reachability Verification TLV. It can be used by a MARP client to verify the reachability of a peer after a NOTIFY_SOFT message is received or as a general mechanism by any network device.

For the purpose of describing the operation of this service, two

devices are considered: the requesting node and the target. In general, the requesting node wants to verify the reachability of the target. The operation is as follows:

- o The requesting node sends an FRD message to the target.
- o The target sends an FRR message that includes a logical NOT of the information in the Data field of the FRD message.

The FRD message MAY be sent directly to the target or to a well-known multicast address specific to the underlying media. If a multicast destination is used, then several targets MAY reply. The FRR message MUST always be sent to the requesting node.

[4.3.](#) An Example of MARP Operation

This section presents an example of MARP being used to provide the reachability notification service.

Given the following network:

R1----(port1)S1(port2)----(port3)S2(port4)----R2

In the figure, R1 and R2 are MARP clients, while S1 and S2 are MARP servers.

- 1 R1 sends an UPDATE message that includes R2's address in it.
- 2 S1 determines that R2's address is reachable via port2, and would thus mark port2 with enough information to note that the failure of this port would be an "interesting" event. The UPDATE message is also forwarded by S1 out all ports on the same broadcast domain, including port2.
- 3 S2 receives the UPDATE message on port3, and finds R2's address available through port4, so it marks port4 as "interesting". The UPDATE message is also forwarded by S2 out all ports on the same broadcast domain, including port4.
- 4 Two independent failure scenarios may occur.
 - 4a The link between S1 and S2 fails. S1 will now send a notification message (NOTIFY_HARD or NOTIFY_SOFT according to the local configuration), with R2's address in it, out

all ports on the same broadcast domain as port2, including port1.

- 4b The link between S2 and R2 fails. S2 will now send a notification message (NOTIFY_HARD or NOTIFY_SOFT according to the local configuration), with R2's address in it, out all ports on the same broadcast domain as port4, including port3. On receiving this notification message, S1 must forward it out all links on the same broadcast domain (except the one it was received on), including port1.
- 5 R1 receives the notification message indicating that R2's address is no longer reachable.
 - 5a If a NOTIFY_SOFT message was received, then R1 may send an FRD message to verify R2's reachability.
 - 5b R1 may take a locally defined action.

5. Security Considerations

This document presents a new protocol which provides a mechanism for a device to notify another device that a particular destination is no longer reachable within a given broadcast domain. While the threat zone is limited to only the local broadcast domain, it is recommended that authentication be used to minimize the threat of false (or spoofed) notifications of lost connectivity.

6. IANA Considerations

The section "MARP Packet Format" defines the fields that make up a MARP packet and it defines meaning to some of the values in them. IANA is expected to maintain the registry for these values as follows.

L2 Sub-Type Field:

- o This field is to be used by the underlying layer 2 media. If not specifically needed by the underlying transport, then it MUST be treated as a Reserved field (described below).

Reserved Fields: These fields, or parts of them, MUST be assigned using the "IETF Consensus" policy defined in [RFC2434](#) [[RFC2434](#)].

Version Number Field:

- o Version number 0 is reserved.
- o Version number 1 is assigned to the current version specified in

this document.

- o Version numbers 2 through 127 MUST be assigned using the "IETF Consensus" policy defined in [RFC2434](#) [[RFC2434](#)].
- o Version numbers 128 through 191 SHOULD be assigned using the "Specification Required" policy defined in [RFC2434](#) [[RFC2434](#)].
- o Version numbers 192 through 255 are for "Private Use" as defined in [RFC2434](#) [[RFC2434](#)].

TLV Type Field:

- o Type code 0 is reserved.
- o Type codes 1, 2 and 3 are assigned in this document.
- o Type codes 4 through 127 MUST be assigned using the "IETF Consensus" policy defined in [RFC2434](#) [[RFC2434](#)].
- o Type codes 128 through 191 SHOULD be assigned using the "Specification Required" policy defined in [RFC2434](#) [[RFC2434](#)].
- o Type codes 192 through 255 are for "Private Use" as defined in [RFC2434](#) [[RFC2434](#)].

Authentication Type Field:

- o Types 0 through 2 are explicitly defined in this document.
- o Authentication Type values 3 thru 63 MUST be assigned using the "IETF Consensus" policy defined in [RFC2434](#) [[RFC2434](#)].

- o Authentication Type values 64 thru 127 SHOULD be assigned using the "Specification Required" policy defined in [RFC2434](#) [[RFC2434](#)].
- o Authentication Type values 128 thru 255 are for "Private Use" as defined in [RFC2434](#) [[RFC2434](#)].

Opcode Field (Reachability Notification TLV):

- o Bit 15 (high order bit) is reserved to indicate if the remaining bits are vendor specific or not.
- o Bits 0 and 1 (two low order bits) are reserved to indicate the message type.

- o Bits 2 through 4 (and its combinations with bits 0 and 1) are to be used for additional message types and SHOULD be assigned using the "IETF Consensus" policy defined in [RFC2434](#) [[RFC2434](#)].
- o Bits 5 through 9 MUST be assigned using the "IETF Consensus" policy defined in [RFC243](#) [[RFC2434](#)].
- o Bits 10 through 14 SHOULD be assigned using the "Specification Required" policy defined in [RFC2434](#) [[RFC2434](#)].

Opcode Field (Fast Reachability Verification TLV)

- o Bit 7 (high order bit) is reserved to indicate if the remaining bits are vendor specific or not.
- o Bit 0 (low order bit) is reserved to indicate the message type.
- o Bits 1 through 4 MUST be assigned using the "IETF Consensus" policy defined in [RFC243](#) [[RFC2434](#)].
- o Bits 5 through 6 SHOULD be assigned using the "Specification Required" policy defined in [RFC2434](#) [[RFC2434](#)].

The IETF has been notified of intellectual property rights claimed in regard to some or all of the specification contained in this document. For more information consult the online list of claimed rights.

8. Acknowledgements

We want to acknowledge David Oran, who had the original idea from which MARP grew. We would like to thank all the people (too many to list individually) who have shown interest in MARP for their valuable input.

9. References

[RFC2119]

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," [RFC 2119](#), March 1997.

[ISIS_SHORT]

Parker, J., McPherson, D., and Alaettinoglu, Cengiz, "Short Adjacency Hold Times in IS-IS", Work In Progress ([draft-parker-short-isis-hold-times-01.txt](#)), July 2001.

[LSP_PING]

Kompella, K., Pan, P., Sheth, N., Cooper, D., Swallow, G., Wadhwa, S., and Bonica, R., "Detecting Data Plane Liveliness in MPLS", Work In Progress ([draft-ietf-mpls-lsp-ping-00.txt](#)), March 2002.

[FLIP]

Sandick, H., Squire, M., Cain, B., Duncan, I., Haberman, B., "Fast

Liveness Protocol (FLIP)", Work In Progress ([draft-sandiick-flip-00.txt](#)), February 2000.

[PLP]

Kompella, K., "Protocol Liveness Protocol", Work In Progress ([draft-kompella-rag-plp-00.txt](#)), October 2002.

[MD5]

Rivest, R., "The MD5 Message-Digest Algorithm", [RFC1321](#), April 1992.

[RFC2434]

Narten, T., Alvestrand, H., "Guidelines for Writing an IANA Considerations Section in RFCs", [RFC 2434](#), [BCP 26](#), October, 1998.

[10](#). Authors' Addresses

Alvaro Retana
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
EMail: aretana@cisco.com

Russ White
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
EMail: riw@cisco.com