

Network Working Group
Internet Draft
Expiration Date: November 2001

Jed Kaplan
P.J. Singh
Allegro Networks, Inc.

Mike O'Dell

John Hayes
Archduke Design, Inc.

Ted Schroeder
Alteon WebSystems, Inc.

Daemon Morrell
Juniper Networks, Inc.

Jennifer Hsu

Extended Ethernet Frame Size Support

[draft-ietf-isis-ext-eth-01.txt](#)

1. 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 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."

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>

2. Abstract

This document presents an extension to current Ethernet Frame specifications for hardware and frame format to support payloads greater than 1500 Bytes for Type interpretation and Length interpretation frames. This is useful for Gigabit Ethernet

technology, providing a means to carry large MTU packets without fragmentation over a high-speed broadcast network.

3. Overview

There are two fundamental frame types defined for Ethernet: Type interpretation [IEEE-ISO] [RFC894] and Length interpretation [IEEE-ISO]. Length interpretation headers may be followed by a Logical Link Control header, 802.2 [IEEE802.2]. Both types of encapsulations can co-exist on the same media at the same time. Encodings for Type interpretation and Length interpretation frames evolved such that, as long as payloads were less than 1500 bytes, Type interpretation frames could always be distinguished from Length interpretation frames.

However, when the payload is greater than 1500 bytes frames may not be uniquely distinguishable as conforming to Type interpretation or Length interpretation formats. This document extends the Ethernet frame format to allow Type interpretation or Length interpretation frame payloads larger than 1500 bytes to be uniquely distinguished.

4. Ethernet Frame Formats

A. Type interpretation

```
+-----+-----+-----+-----+-----+
| DA | SA | Type | Data | FCS |
+-----+-----+-----+-----+-----+
```

DA	Destination MAC Address	(6 bytes)
SA	Source MAC Address	(6 bytes)
Type	Protocol Type	(2 bytes)
Data	Protocol Data	(46 - 1500 bytes)
FCS	Frame Checksum	(4 bytes)

B. Length interpretation and derivatives

```
+-----+-----+-----+-----+-----+
| DA | SA | Len | Data | FCS |
+-----+-----+-----+-----+-----+
```

DA	Destination MAC Address	(6 bytes)
SA	Source MAC Address	(6 bytes)
Len	Length of Data field	(2 bytes)
Data	Protocol Data	(46 - 1500 bytes)
FCS	Frame Checksum	(4 bytes)

The derivatives include LLC (802.2) and SNAP which prefix the data field with an LLC header. In these instances the Len field then corresponds to the combined size of both the data portion of the frame and the LLC header.

On reception, the two formats are differentiated based on the magnitude of the Type/Length field, as follows:

> 1536 bytes: value corresponds to a type field. The frame is an Ethernet II frame, with type values starting at 1536 (600 hex).

<= maxValidFrame bytes: value corresponds to a length field. The frame is an IEEE 802.3 format (or derivative) with a maximum data length of 1500 bytes.

5. Problem with Large Length interpretation Frames in the presence of Type Interpretation Frames

Some protocols commonly used in the Internet have no reserved Ethertype. An example is the set of ISO Network layer protocols, of which ISIS is a member. Such protocols are only defined to use the IEEE 802.3/802.2 encoding, and so their packets are limited in length to 1500 bytes.

Type Interpretation frames have no length field. Protocols encapsulated in Type interpretation frames, such as IP, are not limited in length to 1500 bytes by framing.

6. Proposed Ethernet Frame Extension

Large Length interpretation and Type interpretation frames can be supported by the following:

- + Define an Ethertype for 802.3, 0x8870, and encode large frames (where the data field is greater than 1500 bytes), exclusive of the Destination MAC address, Source MAC address, and Data length fields, within Type interpretation frames.

Large 802.3/802.2 frames would have the following fields:

```
+-----+-----+-----+-----+-----+-----+-----+-----+
| DA | SA | Type | DSAP | SSAP | Ctrl | Data | FCS |
+-----+-----+-----+-----+-----+-----+-----+-----+
=== 802.2 Header ===
```

DA	Destination MAC Address	(6 bytes)
SA	Source MAC Address	(6 bytes)
Type	0x8870 (Ethertype)	(2 bytes)
DSAP	802.2 Destination Service Access Point	(1 byte)
SSAP	802.2 Source Service Access Point	(1 byte)
Ctrl	802.2 Control Field	(1 byte)
Data	Protocol Data	(46 bytes)
FCS	Frame Checksum	(4 bytes)

- + Allow Type interpretation frames to have payloads greater than 1500 bytes.

There is no loss of information from 802.3/802.2 frames. Although the 802.3 length field is missing, the frame length is known by virtue of the frame being accepted by the network interface.

In this manner, all Type interpretation frames, including large Length interpretation frames, can be longer than 1500 bytes, yet are uniquely identified.

7. Default MTU

The motivation for this draft was support of FDDI MTU, 4470, without fragmentation.

Supporting MPLS over Ethernet without fragmentation requires 4 or 8 additional bytes beyond Ethernet MTU.

Some servers support MTU of 9180. Additionally [RFC 1626](#) prescribes an MTU of 9180 for ATM AAL5 supporting IP over ATM. Some equipment providers claim significant cost to support this MTU.

Assuming that there isn't a great need for an MTU of 9180, an MTU of 4470 is recommended, to support the expected predominate applications of extended Ethernet frames, without causing undue burden on equipment providers.

8. References

[IEEE-ISO] IEEE Std. 802.3 [2000 Edition], ISO/IEC 8802-3:2000.

[RFC894] IETF [RFC 894](#)

[IEEE802] IEEE Std 802

[IEEE802.3Z] IEEE Std 802.3z

[802.3X] IEEE Std. 802.3X [March 20, 1997], sec 4.2.7.1.

[EXT.FRAME] "Use of Extended Frame Sizes in Ethernet Networks", draft 2.1, Alteon Networks, Inc.

9. Author's Addresses

Jed Kaplan
Allegro Networks, Inc.
[12700](#) Failr Lakes Circle
Suite 100
Fairfax, Va 22033

email: jkaplan@allegronetworks.com

P.J. Singh
Allegro Networks, Inc.
6399 San Ignacio Blvd.
San Jose, Ca. 95119
408-281-5500
email: pjsingh@allegronetworks.com

Mike O'Dell
email: mo@ccr.net

John Hayes
Archduke Design, Inc.
24700 Skyland Road
Los Gatos, CA 95033
(408) 691-6891
email: hayes@archdukedesign.com

Ted Schroeder
email: ted@alteon.com

Daemon Morrell
Juniper Networks, Inc.
12343-D Sunrise Valley Drive
Reston, VA 20191
email: dmorrell@juniper.net

Jennifer Hsu
jhsu@mur.com

10. Appendix 1. Following is the response from the IEEE to [draft-kaplan-isis-ext-eth-02.txt](#).

From: Geoff Thompson, Chair, IEEE 802.3
To: Scott O. Bradner, IETF
Re: 802.3 Position on Extended Ethernet Frame Size Support

Scott-

This is in response to your query for a position regarding the publication of Extended Ethernet Frame Size Support - [draft-kaplan-isis-ext-eth-02.txt](#) - as an informational RFC. This response was approved in concept and draft by 802.3 during its closing plenary at Hilton Head on March 15. The final form was drafted by myself and reviewed by an ad hoc that was formed during our closing plenary. It should be considered the position of the 802.3 Working Group.

The response is composed of two parts, specific comments on the draft and general comments on the use of jumbo frames in Ethernet

networks, however, virtually all traffic uses the type/length field as a type field. It seems unlikely that the implementations using the length format would take advantage of longer packets. Therefore, the draft conveys a very limited value.

Specific comments on: Extended Ethernet Frame Size Support - [draft-kaplan-isis-ext-eth-02.txt](#)

The draft makes no mention that extended frames are not likely to be successfully handled by Ethernet equipment unless the network is composed entirely of equipment that is specifically designed, beyond the specifications of the Ethernet Standard, to relay extended size frames.

In [section 2](#), Abstract, the document asserts that it presents an extension to the "current Ethernet Frame Standards to support payloads greater than 1500 bytes..." Neither the original Ethernet specification (it was not a "Standard") nor IEEE Std. 802.3 is a "frame standard". They are, rather, complete specifications for hardware and frame format with the expectation that parameters from one portion of the standard can be taken as a given in other portions of the Standard. Moreover, this draft is not an "extension" to those documents but rather a proposal to violate specific provisions of those documents.

In [section 3](#), the draft refers to "Ethernet II [ETH] and points to the reference [ETH] The reference, as cited, is incorrect or incomplete.

Ethernet II would seem to point to Ethernet Version 2.0. That would specifically not be "version 1.0...September 1980". The citation in fact points to 2 different documents and fails to note that the November 1982 edition is in fact Version 2.0. Further, both of these

are obsolete references and have been superseded by IEEE Std. 802.3 and ISO/IEC 8802-3. The current version of these Standards is IEEE Std. 802.3 [2000 Edition] and ISO/IEC 8802-3 : 2000.

The details of [section 4](#) are badly out of date. IEEE Std. 802.3 has included both Type and Length encoded packets ever since the adoption of IEEE Std. 802.3x on March 20, 1997. The current text of the 802.3 text covering this reads:

=====

3.2.6 Length/Type Field

This two-octet field takes one of two meanings, depending on its numeric value. For numerical evaluation, the first octet is the most significant octet of this field.

- a) If the value of this field is less than or equal to the value of `maxValidFrame` (as specified in 4.2.7.1), then the Length/Type field indicates the number of MAC client data octets contained in the subsequent data field of the frame (Length

interpretation).

- b) If the value of this field is greater than or equal to 1536 decimal (equal to 0600 hexadecimal), then the Length/Type field indicates the nature of the MAC client protocol (Type interpretation). The Length and Type interpretations of this field are mutually exclusive.

=====
Please note that any value over "the value of maxValidFrame" is NOT a valid value for encoding length. Additionally, the values between maxValidFrame and "1536 decimal" are undefined in the Ethernet standard. The behavior of equipment at these values is not specified and can not be depended on. The draft implies that these values are valid type fields. This is not true. These values are not valid for either Type or Length.

[Section 4](#) Re: "...are not limited in length to 1500 bytes by framing." While this seems to be true, it is not necessarily true for a number of sometimes subtle reasons, some of which are noted in the "General" section below.

[Section 5](#): Regarding the statement "Although the 802.3 length field is missing, the frame length is missing, the frame length is known by virtue of the frame being accepted by the network interface." This statement is not correct. Many Ethernet interfaces, particularly those of relay equipment, accept frames without regard for packet type or content. There is no reasonable expectation that standards based Ethernet/802.3 equipment will reject the proposed frames. They may very well accept the frame and corrupt it before passing it on. This corruption may consist of truncation or alteration of the data within the packet.

General comments on the use of jumbo frames in Ethernet networks:

Consideration #1: The expectation of no more than 15-1600 bytes between frames and an interpacket gap before the next frame is deeply ingrained throughout the design and implementation of standardized Ethernet/802.3 hardware. This shows up in buffer allocation schemes, clock skew and tolerance compensation and fifo design.

Consideration #2: For some Ethernet/802.3 hardware (repeaters are one specific example) it is not possible to design compliant equipment which meets all of the requirements and will still pass extra long frames. Further, since clock frequency may vary with time and temperature, equipment may successfully pass long frames at times and corrupt them at other times. Therefore, attempts to verify the ability to send long frames over a path may produce inaccurate results.

Consideration #3: The error checking mechanism embodied in the 4 byte checksum has not been well characterized at greater frame

lengths, but is known to degrade. Therefore the data reliability of transfers in long frame transfers will have a greater rate of undetected frame errors.

Consideration #4: The length of frames proposed by this draft can not be assured to pass through standards conformant hardware. The huge value of Ethernet/802.3 systems in the data networking universe is their standardization and the resulting assurance that systems will all interoperate. No such assurance can be provided for oversize frames with both the current broadly accepted standard and the large installed base of standards based equipment.

In summary with regard to greatly longer frames for Ethernet, much of the gear produced today would be intolerant of greatly longer frames. There is no way proposed to distinguish between frame types in the network as they arrive from the media. Bridges might and repeaters would drop or truncate frames (and cause errors doing so) right and left for uncharacterized reasons. It would be a mess. What might seem okay for small carefully characterized networks would be enormously difficult or impossible to do across the Standard.

The choice of frame size for Ethernet packets is really the domain of 802.3 (CSMA/CD) and 802.1 (Bridging, VLANs). The only time the frame size has been modified over the history of the Standard was in order to increase maximum length by four bytes in order to accommodate VLANs, 802.1 initiated this work and 802.3 also modified the Ethernet standard to include these few extra bytes. The people with the experience dealing with this sort of thing attend IEEE 802. It's easy to define a new ethertype, but it's not too easy to figure out what happens when these non-standard frames are given to standardized transmission equipment e.g. bridges. We would expect discussions of this type to take place in both 802.3 & 802.1.

The giant frame issue has been mentioned several times over the years in 802.3, discussed in the back halls and considered each time we move to a higher speed. It has never had consensus support in that context. It has never been brought forward as a separate proposal. Backward compatibility has always been more important than ease of performance improvement. The problem is that the change is very easy to do in the standard and hard to do in the world. It is just like changing the gauge on railroad tracks. All you have to do is change one line in the standard, never mind all of the rails you have to move.

The Kaplan draft is just meant for carrying IS-IS routing protocol frames (the IS-IS working group is the intended sponsor of this draft). We expect those vendors supporting the larger frame will support this will show up and support this proposal. Those vendors not supporting the larger frame as well as those protecting the

installed base will not support this activity nor having this sort of item standardized outside IEEE 802.3.

With best regards,

Geoff Thompson, Chair, IEEE 802.3

11. Appendix 2. Comments from the draft's authors.

With respect to Consideration #2, paragraph 1:

With the advent of full duplex ethernet and higher speed ethernet phys, transcievers have gone from transmitting when they have data ready send to transmitting constantly, sending IDLEs when they have nothing to send. Any clock drift due to time and temperature will affect the system without regard to the frame lengths being used.

With respect to Consideration #3, paragraph 1:

The error checking mechanism of ethernet (CRC-32) was characterized by R. Jain, "Error Characteristics of Fiber Distributed Data Interface (FDDI)," IEEE Transactions on Communications, Vol. 38, No. 8, August 1990, pp. 1244-1252. <http://www.cis.ohio-state.edu/~jain/papers/xie1.htm>

FDDI and Ethernet use the same error checking mechanism; CRC-32. The probability of undetected errors remains constant for frame sizes between 3007 and 91639 bits (approximately 376 to 11455 bytes). Setting the maximum size of jumbo frames to 9018 bytes falls well within this range. There is no increase in undetected errors when using jumbo frames and the existing CRC-32 as the error detection mechanism.

With respect to Consideration #4, paragraph 1:

This proposal provides a workable mechanism to identify and handle jumbo frames through systems that do support jumbo frames.