Network Working Group   J. Postel
Request for Comments: 1042   J. Reynolds
                                                                     ISI
Obsoletes: RFC-948   February 1988

A Standard for the Transmission of IP Datagrams over IEEE 802 Networks

Status of this Memo

This RFC specifies a standard method of encapsulating the Internet Protocol (IP) [1] datagrams and Address Resolution Protocol (ARP) [2] requests and replies on IEEE 802 Networks. This RFC specifies a protocol standard for the Internet community. Distribution of this memo is unlimited.

Acknowledgment

This memo would not exist with out the very significant contributions of Drew Perkins of Carnegie Mellon University, Jacob Rekhter of the T.J. Watson Research Center, IBM Corporation, and Joseph Cimmino of the University of Maryland.

Introduction

The goal of this specification is to allow compatible and interoperable implementations for transmitting IP datagrams and ARP requests and replies. To achieve this it may be necessary in a few cases to limit the use that IP and ARP make of the capabilities of a particular IEEE 802 standard.

The IEEE 802 specifications define a family of standards for Local Area Networks (LANs) that deal with the Physical and Data Link Layers as defined by the ISO Open System Interconnection Reference Model (ISO/OSI). Several Physical Layer standards (802.3, 802.4, and 802.5) [3,4,5] and one Data Link Layer Standard (802.2) [6] have been defined. The IEEE Physical Layer standards specify the ISO/OSI Physical Layer and the Media Access Control Sublayer of the ISO/OSI Data Link Layer. The 802.2 Data Link Layer standard specifies the Logical Link Control Sublayer of the ISO/OSI Data Link Layer.

This memo describes the use of IP and ARP on the three types of networks. At this time, it is not necessary that the use of IP and ARP be consistent across all three types of networks, only that it be consistent within each type. This may change in the future as new IEEE 802 standards are defined and the existing standards are revised allowing for interoperability at the Data Link Layer.

It is the goal of this memo to specify enough about the use of IP and ARP on each type of network to ensure that:

(1) all equipment using IP or ARP on 802.3 networks will interoperate,

(2) all equipment using IP or ARP on 802.4 networks will interoperate,

(3) all equipment using IP or ARP on 802.5 networks will interoperate.

Of course, the goal of IP is interoperability between computers attached to different networks, when those networks are interconnected via an IP gateway [8]. The use of IEEE 802.1 compatible Transparent Bridges to allow interoperability across different networks is not fully described pending completion of that standard.

Description

IEEE 802 networks may be used as IP networks of any class (A, B, or C). These systems use two Link Service Access Point (LSAP) fields of the LLC header in much the same way the ARPANET uses the "link" field. Further, there is an extension of the LLC header called the Sub-Network Access Protocol (SNAP).

IP datagrams are sent on IEEE 802 networks encapsulated within the 802.2 LLC and SNAP data link layers, and the 802.3, 802.4, or 802.5 physical networks layers. The SNAP is used with an Organization Code indicating that the following 16 bits specify the EtherType code (as listed in Assigned Numbers [7]).

Normally, all communication is performed using 802.2 type 1 communication. Consenting systems on the same IEEE 802 network may use 802.2 type 2 communication after verifying that it is supported by both nodes. This is accomplished using the 802.2 XID mechanism. However, type 1 communication is the recommended method at this time and must be supported by all implementations. The rest of this specification assumes the use of type 1 communication.

The IEEE 802 networks may have 16-bit or 48-bit physical addresses. This specification allows the use of either size of address within a given IEEE 802 network.

Note that the 802.3 standard specifies a transmission rate of from 1 to 20 megabit/second, the 802.4 standard specifies 1, 5, and 10 megabit/second, and the 802.5 standard specifies 1 and 4 megabit/second. The typical transmission rates used are 10 megabit/second for 802.3, 10 megabit/second for 802.4, and 4 megabit/second for 802.5. However, this specification for the transmission of IP Datagrams does not depend on the transmission rate.

Header Format

```
                                                                Header

   ...--------+--------+--------+
              MAC Header        |                       802.{3/4/5} MAC
   ...--------+--------+--------+

   +--------+--------+--------+
   | DSAP=K1| SSAP=K1| Control|                                 802.2 LLC
   +--------+--------+--------+

   +--------+--------+---------+--------+--------+
   |Protocol Id or Org Code =K2|    EtherType    |              802.2 SNAP
   +--------+--------+---------+--------+--------+
```

The total length of the LLC Header and the SNAP header is 8-octets, making the 802.2 protocol overhead come out on an nice boundary.

The K1 value is 170 (decimal).

The K2 value is 0 (zero).

The control value is 3 (Unnumbered Information).

Address Mappings

The mapping of 32-bit Internet addresses to 16-bit or 48-bit IEEE 802 addresses must be done via the dynamic discovery procedure of the Address Resolution Protocol (ARP) [2].

Internet addresses are assigned arbitrarily on Internet networks. Each host's implementation must know its own Internet address and respond to Address Resolution requests appropriately. It must also use ARP to translate Internet addresses to IEEE 802 addresses when needed.

The ARP Details

The ARP protocol has several fields that parameterize its use in any specific context [2]. These fields are:

hrd	16 - bits	The Hardware Type Code
pro	16 - bits	The Protocol Type Code
hln	8 - bits	Octets in each hardware address
pln	8 - bits	Octets in each protocol address
op	16 - bits	Operation Code

The hardware type code assigned for the IEEE 802 networks (of all kinds) is 6 (see [7] page 16).

The protocol type code for IP is 2048 (see [7] page 14).

The hardware address length is 2 for 16-bit IEEE 802 addresses, or 6 for 48-bit IEEE 802 addresses.

The protocol address length (for IP) is 4.

The operation code is 1 for request and 2 for reply.

Broadcast Address

The broadcast Internet address (the address on that network with a host part of all binary ones) should be mapped to the broadcast IEEE 802 address (of all binary ones) (see [8] page 14).

Trailer Formats

Some versions of Unix 4.x bsd use a different encapsulation method in order to get better network performance with the VAX virtual memory architecture. Consenting systems on the same IEEE 802 network may use this format between themselves. Details of the trailer encapsulation method may be found in [9]. However, all hosts must be able to communicate using the standard (non-trailer) method.

Byte Order

As described in Appendix B of the Internet Protocol specification [1], the IP datagram is transmitted over IEEE 802 networks as a series of 8-bit bytes. This byte transmission order has been called "big-endian" [11].

Maximum Transmission Unit

The Maximum Transmission Unit (MTU) differs on the different types of IEEE 802 networks. In the following there are comments on the MTU for each type of IEEE 802 network. However, on any particular network all hosts must use the same MTU. In the following, the terms "maximum packet size" and "maximum transmission unit" are equivalent.

Frame Format and MAC Level Issues

For all hardware types

IP datagrams and ARP requests and replies are transmitted in standard 802.2 LLC Type 1 Unnumbered Information format, control code 3, with the DSAP and the SSAP fields of the 802.2 header set to 170, the assigned global SAP value for SNAP [6]. The 24-bit Organization Code in the SNAP is zero, and the remaining 16 bits are the EtherType from Assigned Numbers [7] (IP = 2048, ARP = 2054).

IEEE 802 packets may have a minimum size restriction. When necessary, the data field should be padded (with octets of zero) to meet the IEEE 802 minimum frame size requirements. This padding is not part of the IP datagram and is not included in the total length field of the IP header.

For compatibility (and common sense) the minimum packet size used with IP datagrams is 28 octets, which is 20 (minimum IP header) + 8 (LLC+SNAP header) = 28 octets (not including the MAC header).

The minimum packet size used with ARP is 24 octets, which is 20 (ARP with 2 octet hardware addresses and 4 octet protocol addresses) + 8 (LLC+SNAP header) = 24 octets (not including the MAC header).

In typical situations, the packet size used with ARP is 32 octets, which is 28 (ARP with 6 octet hardware addresses and 4 octet protocol addresses) + 8 (LLC+SNAP header) = 32 octets (not including the MAC header).

IEEE 802 packets may have a maximum size restriction. Implementations are encouraged to support full-length packets.

For compatibility purposes, the maximum packet size used with IP datagrams or ARP requests and replies must be consistent on a particular network.

Gateway implementations must be prepared to accept full-length packets and fragment them when necessary.

Host implementations should be prepared to accept full-length packets, however hosts must not send datagrams longer than 576 octets unless they have explicit knowledge that the destination is prepared to accept them. A host may communicate its size preference in TCP based applications via the TCP Maximum Segment Size option [10].

Datagrams on IEEE 802 networks may be longer than the general Internet default maximum packet size of 576 octets. Hosts connected to an IEEE 802 network should keep this in mind when sending datagrams to hosts not on the same IEEE 802 network. It may be appropriate to send smaller datagrams to avoid unnecessary fragmentation at intermediate gateways. Please see [10] for further information.

IEEE 802.2 Details

While not necessary for supporting IP and ARP, all implementations are required to support IEEE 802.2 standard Class I service. This requires supporting Unnumbered Information (UI) Commands, eXchange IDentification (XID) Commands and Responses, and TEST link (TEST) Commands and Responses.

When either an XID or a TEST command is received a response must be returned; with the Destination and Source addresses, and the DSAP and SSAP swapped.

When responding to an XID or a TEST command the sense of the poll/final bit must be preserved. That is, a command received with the poll/final bit reset must have the response returned with the poll/final bit reset and vice versa.

The XID command or response has an LLC control field value of 175 (decimal) if poll is off or 191 (decimal) if poll is on. (See Appendix on Numbers.)

The TEST command or response has an LLC control field value of 227 (decimal) if poll is off or 243 (decimal) if poll is on. (See Appendix on Numbers.)

A command frame is identified with high order bit of the SSAP address reset. Response frames have high order bit of the SSAP address set to one.

XID response frames should include an 802.2 XID Information field of 129.1.0 indicating Class I (connectionless) service. (type 1).

TEST response frames should echo the information field received in the corresponding TEST command frame.

For IEEE 802.3

A particular implementation of an IEEE 802.3 Physical Layer is denoted using a three field notation. The three fields are data rate in megabit/second, medium type, and maximum segment length in hundreds of meters. One combination of of 802.3 parameters is 10BASE5 which specifies a 10 megabit/second transmission rate, baseband medium, and 500 meter segments. This correspondes to the specifications of the familiar "Ethernet" network.

The MAC header contains 6 (2) octets of source address, 6 (2) octets of destination address, and 2 octets of length. The MAC trailer contains 4 octets of Frame Check Sequence (FCS), for a total of 18 (10) octets.

IEEE 802.3 networks have a minimum packet size that depends on the transmission rate. For type 10BASE5 802.3 networks the minimum packet size is 64 octets.

IEEE 802.3 networks have a maximum packet size which depends on the transmission rate. For type 10BASE5 802.3 networks the maximum packet size is 1518 octets including all octets between the destination address and the FCS inclusive.

This allows 1518 - 18 (MAC header+trailer) - 8 (LLC+SNAP header) = 1492 for the IP datagram (including the IP header). Note that 1492 is not equal to 1500 which is the MTU for Ethernet networks.

For IEEE 802.4

The MAC header contains 1 octet of frame control, 6 (2) octets of source address, and 6 (2) octets of destination address. The MAC trailer contains 4 octets of Frame Check Sequence (FCS), for a total of 17 (9) octets.

IEEE 802.4 networks have no minimum packet size.

IEEE 802.4 networks have a maximum packet size of 8191 octets including all octets between the frame control and the FCS inclusive.

This allows 8191 - 17 (MAC header+trailer) - 8 (LLC+SNAP header) = 8166 for the IP datagram (including the IP header).

For IEEE 802.5

The current standard for token ring's, IEEE 802.5-1985, specifies the operation of single ring networks. However, most implementations of 802.5 have added extensions for multi-ring networks using source-routing of packets at the MAC layer. There is now a Draft Addendum to IEEE 802.5, "Enhancement for Multi-Ring Networks" which attempts to standardize these extensions. Unfortunately, the most recent draft (November 10, 1987) is still rapidly evolving. More importantly, it differs significantly from the existing implementations. Therefore, the existing implementations of 802.5 [13] are described but no attempt is made to specify any future standard.

The MAC header contains 1 octet of access control, 1 octet of frame control, 6 (2) octets of source address, 6 (2) octets of destination address, and (for multi-ring networks) 0 to 18 octets of Routing Information Field (RIF). The MAC trailer contains 4 octets of FCS, for a total of 18 (10) to 36 (28) octets. There is one additional octet of frame status after the FCS.

Multi-Ring Extension Details

The presence of a Routing Information Field is indicated by the Most Significant Bit (MSB) of the source address, called the Routing Information Indicator (RII). If the RII equals zero, a RIF is not present. If the RII equals 1, the RIF is present. Although the RII is indicated in the source address, it is not part of a stations MAC layer address. In particular, the MSB of a destination address is the individual/group address indicator, and if set will cause such frames to be interpreted as multicasts. Implementations should be careful to reset the RII to zero before passing source addresses to other protocol layers which may be confused by their presence.

The RIF consists of a two-octet Routing Control (RC) field followed by 0 to 8 two-octet Route-Designator (RD) fields. The RC for all-routes broadcast frames is formatted as follows:

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  B  |    LTH    |D|  LF |  r  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Note that each tick mark represents one bit position.

B - Broadcast Indicators: 3 bits

The Broadcast Indicators are used to indicate the routing desired for a particular frame. A frame may be routed through a single specified route, through every distinct non-repeating route in a multi-ring network, or through a single route determined by a spanning tree algorithm such that the frame appears on every ring exactly once. The values which may be used at this time are (in binary):

000 - Non-broadcast (specific route)
100 - All-routes broadcast (global broadcast)
110 - Single-route broadcast (limited broadcast)

All other values are reserved for future use.

LTH - Length: 5 bits

The Length bits are used to indicate the length or the RI field, including the RC and RD fields. Only even values between 2 and 30 inclusive are allowed.

D - Direction Bit: 1 bit

The D bit specifies the order of the RD fields. If D equals 1, the routing-designator fields are specified in reverse order.

LF - Largest Frame: 3 bits

The LF bits specify the maximum MTU supported by all bridges along a specific route. All multi-ring broadcast frames should be transmitted with a value at least as large as the supported MTU. The values used are:

LF (binary)	MAC MTU	IP MTU
000	552	508
001	1064	1020
010	2088	2044
011	4136	4092
100	8232	8188

All other values are reserved for future use.

The receiver should compare the LF received with the MTU. If the LF is greater than or equal to the MTU then no action is taken; however, if the LF is less than the MTU the frame is rejected.

There are actually three possible actions if LF < MTU. First is the one required for this specification (reject the frame). Second is to reduce the MTU for all hosts to equal the LF. And, third is to keep a separate MTU per communicating host based on the received LFs.

r - reserved: 4 bits

These bits are reserved for future use and must be set to 0 by the transmitter and ignored by the receiver.

It is not necessary for an implementation to interpret routing-designators. Their format is left unspecified. Routing-designators should be transmitted exactly as received.

IEEE 802.5 networks have no minimum packet size.

IEEE 802.5 networks have a maximum packet size based on the maximum time a node may hold the token. This time depends on many factors including the data signalling rate and the number of nodes on the ring. The determination of maximum packet size becomes even more complex when multi-ring networks with bridges are considered.

Given a token-holding time of 9 milliseconds and a 4 megabit/second ring, the maximum packet size possible is 4508 octets including all octets between the access control and the FCS inclusive.

This allows 4508 - 36 (MAC header+trailer with 18 octet RIF) - 8 (LLC+SNAP header) = 4464 for the IP datagram (including the IP header).

However, some current implementations are known to limit packets to 2046 octets (allowing 2002 octets for IP). It is recommended that all implementations support IP packets of at least 2002 octets.

By convention, source routing bridges used in multi-ring 802.5 networks will not support packets larger than 8232 octets. With a MAC header+trailer of 36 octets and the LLC+SNAP header of 8 octets, the IP datagram (including IP header) may not exceed 8188 octets.

A source routing bridge linking two rings may be configured to limit the size of packets forwarded to 552 octets, with a MAC header+trailer of 36 octets and the LLC+SNAP of 8 octets, the IP datagram (including the IP header) may be limited to 508 octets. This is less that the default IP MTU of 576 octets, and may cause significant performance problems due to excessive datagram fragmentation. An implementation is not required to support an MTU of less than 576 octets, although it is suggest that the MTU be a user-configurable parameter to allow for it.

IEEE 802.5 networks support three different types of broadcasts. All-Stations broadcasts are sent with no RIF or with the Broadcast Indicators set to 0 and no Routing Designators, and are copied once by all stations on the local ring. All-Routes broadcasts are sent with the corresponding Broadcast Indicators and result in multiple copies equal to the number of distinct non-repeating routes a packet may follow to a particular ring. Single-Route broadcasts result in exactly one copy of a frame being received by all stations on the multi-ring network.

The dynamic address discovery procedure is to broadcast an ARP request. To limit the number of all rings broadcasts to a minimum, it is desirable (though not required) that an ARP request first be sent as an all-stations broadcast, without a Routing Information Field (RIF). If the all-stations (local ring) broadcast is not supported or if the all-stations broadcast is unsuccessful after some reasonable time has elapsed, then send the ARP request as an all-routes or single-route broadcast with an empty RIF (no routing designators). An all-routes broadcast is preferable since it yields an amount of fault tolerance. In an environment with multiple redundant bridges, all-routes broadcast allows operation in spite of spanning-tree bridge failures. However, single-route broadcasts may be used if IP and ARP must use the same broadcast method.

When an ARP request or reply is received, all implementations are required to understand frames with no RIF (local ring) and frames with an empty RIF (also from the local ring). If the implementation supports multi-ring source routing, then a non-empty RIF is stored for future transmissions to the host originating the ARP request or reply. If source routing is not supported them all packets with non-empty RIFs should be gracefully ignored. This policy will allow all implementations in a single ring environment, to interoperate, whether or not they support the multi-ring extensions.

It is possible that when sending an ARP request via an all-routes broadcast that multiple copies of the request will arrive at the destination as a result of the request being forwarded by several bridges. However, these "copies" will have taken different routes so the contents of the RIF will differ. An implementation of ARP in this context must determine which of these "copies" to use and to ignore the others. There are three obvious and legal strategies: (1) take the first and ignore the rest (that is, once you have an entry in the ARP cache don't change it), (2) take the last, (that is, always up date the ARP cache with the latest ARP message), or (3) take the one with the shortest path, (that is, replace the ARP cache information with the latest ARP message data if it is a shorter route). Since there is no problem of incompatibility for interworking of different implementations if different strategies are chosen, the choice is up to each implementor. The recipient of the ARP request must send an ARP reply as a point to point message using the RIF information.

The RIF information should be kept distinct from the ARP table. That is, there is, in principle, the ARP table to map from IP addresses to 802 48-bit addresses, and the RIF table to map from those to 802.5 source routes, if necessary. In practical implementations it may be convenient to store the ARP and RIF information together.

> Storing the information together may speed up access to the information when it is used. On the other hand, in a generalized implementation for all types of 802 networks a significant amount of memory might be wasted in an ARP cache if space for the RIF information were always reserved.

IP broadcasts (datagrams with a IP broadcast address) must be sent as 802.5 single-route broadcasts. Unlike ARP, all-routes broadcasts are not desirable for IP. Receiving multiple copies of IP broadcasts would have undesirable effects on many protocols using IP. As with ARP, when an IP packet is received, all implementations are required to understand frames with no RIF and frames with an empty RIF.

Since current interface hardware allows only one group address, and since the functional addresses are not globally unique, IP and ARP do not use either of these features. Further, in the IBM style 802.5 networks there are only 31 functional addresses available for user definition.

IP precedence should not be mapped to 802.5 priority. All IP and ARP packets should be sent at the default 802.5 priority. The default priority is 3.

After packet transmission, 802.5 provides frame not copied and address not recognized indicators. Implementations may use these indicators to provide some amount of error detection and correction. If the frame not copied bit is set but the address not recognized bit is reset, receiver congestion has occurred. It is suggested, though not required, that hosts should retransmit the offending packet a small number of times (4) or until congestion no longer occurs. If the address not recognized bit is set, an implementation has 3 options: (1) ignore the error and throw the packet away, (2) return an ICMP destination unreachable message to the source, or (3) delete the ARP entry which was used to send this packet and send a new ARP request to the destination address. The latter option is the preferred approach since it will allow graceful recovery from first hop bridge and router failures and changed hardware addresses.

Interoperation with Ethernet

It is possible to use the Ethernet link level protocol [12] on the same physical cable with the IEEE 802.3 link level protocol. A computer interfaced to a physical cable used in this way could potentially read both Ethernet and 802.3 packets from the network. If a computer does read both types of packets, it must keep track of which link protocol was used with each other computer on the network and use the proper link protocol when sending packets.

One should note that in such an environment, link level broadcast packets will not reach all the computers attached to the network, but only those using the link level protocol used for the broadcast.

Since it must be assumed that most computers will read and send using only one type of link protocol, it is recommended that if such an environment (a network with both link protocols) is necessary, an IP gateway be used as if there were two distinct networks.

Note that the MTU for the Ethernet allows a 1500 octet IP datagram, with the MTU for the 802.3 network allows only a 1492 octet IP datagram.

Appendix on Numbers

The IEEE likes to specify numbers in bit transmission order, or bit-wise little-endian order. The Internet protocols are documented in byte-wise big-endian order. This may cause some confusion about the proper values to use for numbers. Here are the conversions for some numbers of interest.

Number	IEEE HEX	IEEE Binary	Internet Binary	Internet Decimal
UI Op Code	C0	11000000	00000011	3
SAP for SNAP	55	01010101	10101010	170
XID	F5	11110101	10101111	175
XID	FD	11111101	10111111	191
TEST	C7	11000111	11100011	227
TEST	CF	11001111	11110011	243
Info	818000			129.1.0

References

[1] Postel, J., "Internet Protocol", RFC-791, USC/Information Sciences Institute, September 1981.

[2] Plummer, D., "An Ethernet Address Resolution Protocol - or - Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", RFC-826, MIT, November 1982.

[3] IEEE, "IEEE Standards for Local Area Networks: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications", IEEE, New York, New York, 1985.

[4] IEEE, "IEEE Standards for Local Area Networks: Token-Passing Bus Access Method and Physical Layer Specification", IEEE, New York, New York, 1985.

[5] IEEE, "IEEE Standards for Local Area Networks: Token Ring Access Method and Physical Layer Specifications", IEEE, New York, New York, 1985.

[6] IEEE, "IEEE Standards for Local Area Networks: Logical Link Control", IEEE, New York, New York, 1985.

[7] Reynolds, J.K., and J. Postel, "Assigned Numbers", RFC-1010, USC/Information Sciences Institute, May 1987.

[8] Braden, R., and J. Postel, "Requirements for Internet Gateways", RFC-1009, USC/Information Sciences Institute, June 1987.

[9] Leffler, S., and M. Karels, "Trailer Encapsulations", RFC-893, University of California at Berkeley, April 1984.

[10] Postel, J., "The TCP Maximum Segment Size Option and Related

Topics", RFC-879, USC/Information Sciences Institute, November 1983.

[11] Cohen, D., "On Holy Wars and a Plea for Peace", Computer, IEEE, October 1981.

[12] D-I-X, "The Ethernet - A Local Area Network: Data Link Layer and Physical Layer Specifications", Digital, Intel, and Xerox, November 1982.

[13] IBM, "Token-Ring Network Architecture Reference", Second Edition, SC30-3374-01, August 1987.