Network Working Group
Request for Comments: 2625
Category: Standards Track

M. Rajagopal
R. Bhagwat
W. Rickard
Gadzoox Networks

June 1999

IP and ARP over Fibre Channel

Status of this Memo

Copyright Notice

Abstract

Fibre Channel (FC) is a high speed serial interface technology that supports several higher layer protocols including Small Computer System Interface (SCSI) and Internet Protocol(IP). Until now, SCSI has been the only widely used protocol over FC. Existing FC standards [3] do not adequately specify how IP packets may be transported over FC and how IP addresses are resolved to FC addresses. The purpose of this document is to specify a way of encapsulating IP and Address Resolution Protocol(ARP) over Fibre Channel and also to describe a mechanism(s) for IP address resolution.

Table of Contents

1. Introduction

Fibre Channel (FC) is a gigabit speed networking technology primarily used for Storage Area Networking (SAN). FC is standardized under American National Standard for Information Systems of the National Committee for Information Technology Standards (NCITS) and has specified a number of documents describing its protocols, operations, and services.

Need:

Currently, Fibre Channel is predominantly used for communication between storage devices and servers using the SCSI protocol, with most of the servers still communicating with each other over LANs. Although, there exists a Fibre Channel Standard [3] that has architecturally defined support for IP encapsulation and address resolution, it is inadequately specified. ([3] prohibits broadcasts, thus loops are not covered; [10] has no support for Class 3).

This has lead to a nonstandard way of using IP over FC in the past. Once such a standard method is completely specified, servers can directly communicate with each other using IP over FC, possibly boosting performance in Server host-to-host communications. This technique will be especially useful in a Clustering Application.

Objective and Scope:

The major objective of this specification is to promote interoperable implementations of IPv4 over FC. This specification describes a method for encapsulating IPv4 and Address Resolution Protocol (ARP) packets over FC. This specification accommodates any FC topology

(loop, fabric, or point-to-point) and any FC class of service (1, 2 or 3). This specification also describes a FC Address Resolution Protocol(FARP) for associating World Wide Port Names (MAC addresses) and FC Port identifiers.

A secondary objective of this specification is to describe other optional address resolution mechanisms:

- Other FARP mechanisms that directly build IPv4 address and FC Port Identifier (Port_ID) associations.
- Inverse ARP (InARP) that allows learning the IP address of a remote node given its World Wide Port Name (WW_PN) and Port_ID.

"Multicasting" in Fibre Channel is defined as an optional service [11] for FC Classes 3 and 6 only, with no definition for Classes 1 and 2. Currently, there are no vendor implementations of this service for either Class of service. Broadcast service available within Fibre Channel can be used to do multicasting, although less efficiently. Presently, there appears to be no IP applications over Fibre Channel that require support for IP multicasting. This specification therefore does not support IP Multicasting.

Organization:

Section 2 states the problem that is solved in this specification. Section 3 describes the techniques used for encapsulating IP and ARP packets in a FC sequence. Section 4 discusses the ARP protocol(IP address to WW_PN). Section 5 discusses the FARP protocol used in FC Layer mappings (WW_PN to Port_ID). Section 6 describes the "Exchange" Management in FC. Section 7 is a summary section and provides a quick reference to FC header settings, FC Link Service Commands, supported features in ARP, FARP, InARP, FC Sequences, FC Exchanges, and FC Login Parameters. Section 8 discusses security. Section 9 acknowledges the technical contributors of this document. Section 10 provides a list of references, and Section 11 provides the authors' addresses.

Appendix A discusses other optional FARP mechanisms. Appendix B discusses the Inverse ARP protocol(WW_PN to IP address) as an alternate and optional way of building MAC and IP address associations. Appendix C lists some informal mechanisms for FC Layer Mappings. Appendix D provides a discussion on validation of the FC-layer mappings for the different FC topologies. Appendix E provides a brief overview of the FC Protocols and Networks. Appendix F addresses reliability in Class 3 and Sequence Count FC Protocol issues. Appendix G provides a list of acronyms and a glossary of FC Terms used in this specification.

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 RFC 2119 [19].

2. Problem Statement

This specification addresses two problems:

- A format definition and encapsulation mechanism for IPv4 and ARP packets over FC
- Mechanisms for Address Resolution

As noted earlier, the existing FC Standard [3] ([10]) is inadequate to solve the above problems. A solution to both problems was first proposed by the Fibre Channel Association (FCA)[1]. FCA is an industry consortium of FC vendor companies and not a Standards Body. This specification is based on the proposed solution in [1] and builds on it.

Address Resolution is concerned with resolving IP addresses to WW_PN (MAC address) and WW_PN to FC Port Identifiers (Port_ID). ARP provides a solution to the first resolution problem and FARP the second.

An optional FARP mechanism resolves IP address directly to FC Port_IDs. This is useful in some upper layer applications.

InARP is another optional mechanism that resolves WW_PN and Port_ID to an IP address. InARP is useful when a node after performing a PLOGI with another node, knows its WW_PN and Port_ID, but not its IP address.

3. IP and ARP Encapsulation

3.1 FC Frame Format

All FC frames have a standard format much like LAN 802.x protocols. (See Appendix E and F). However, the exact size of each frame varies depending on the size of the variable fields. The size of the variable field ranges from 0 to 2112-bytes as shown in the FC Frame Format in Fig. 1.

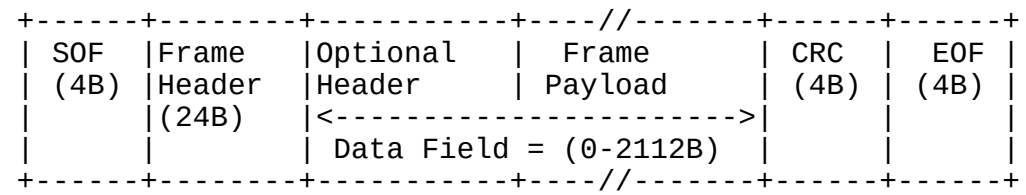


Fig. 1 FC Frame Format

The Start of Frame (SOF) and End of Frame (EOF) are both 4-bytes long and act as frame delimiters.

The CRC is 4-bytes long and uses the same 32-bit polynomial used in FDDI and is specified in ANSI X3.139 Fiber Distributed Data Interface.

The Frame Header is 24-bytes long and has several fields that are associated with the identification and control of the payload. Some of the values and options for this field that are relevant to the IP and ARP payloads are discussed in Section 7.

Current FC Standards allow up to 3 Optional Header fields [11]:

- Network_Header (16-bytes)
- Association_Header (32-bytes)
- Device_Header (up to 64-bytes).

The IP and ARP FC Sequences SHALL carry only the Network_Header field which is 16-bytes long. Other types of optional headers SHALL NOT be used. The Network_Header is REQUIRED in all ARP packets and in the first frame of a logical sequence carrying an IP payload as described below.

An application level payload such as IP is called an Information Unit at the FC-4 Level. Lower FC levels map this to a FC Sequence. (See Appendix E.2 for a description of Sequences and Information Units.) Typically, a Sequence consists of more than one frame. Larger user data is segmented and reassembled using two methods: Sequence Count and Relative Offset [18]. With the use of Sequence Count, data blocks are sent using frames with increasing sequence counts (modulo 65536) and it is quite straightforward to detect the first frame that contains the Network_Header. When Relative Offset is used, as frames arrive, some computation is required to detect the first frame that contains the Network_Header. Sequence Count and Relative Offset field control information, is carried in the FC Header.

In FC, the physical temporal ordering of the frames as it arrives at a destination can be different from that of the order sent because of traversing through a FC Network.

When IP forms the FC Payload then only the first frame of the logical Sequence SHALL include the FC Network_Header. Fig. 2 shows the logical First Frame and logical subsequent frames. Since frames may arrive out of order, detection of the first frame of the logical FC Sequence is necessary.

ARP packets map to a single frame FC Sequence and SHALL always carry the FC Network_Header.

Note the definition of FC Data Field and FC Frame Payload in Fig. 1. FC Data Field includes the FC Frame Payload and the FC Optional Header, that is, Frame Payload definition does not include the FC Optional Header. One or more Frame Payloads together make the FC Sequence Payload as shown in Fig 2 and discussed further in Sections 3.2 and 3.4. FC Sequence Payload includes the mapped IP or ARP packet along with the LLC/SNAP headers.

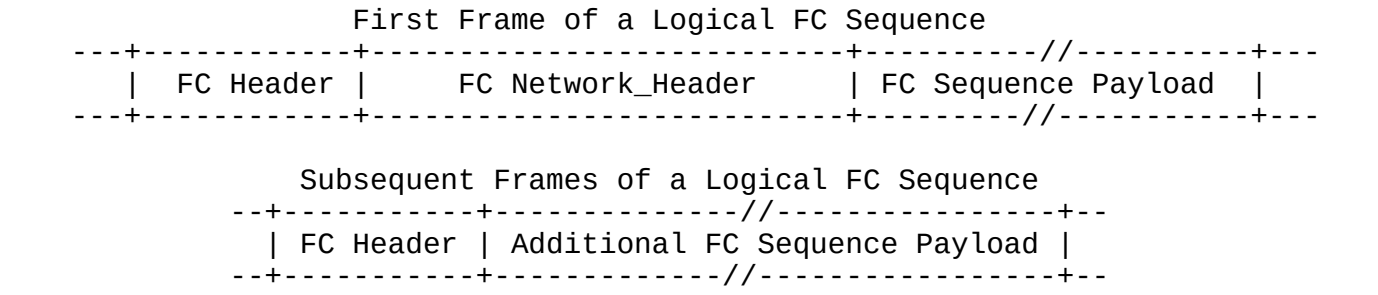


Fig. 2 FC Network_Header in a Frame Sequence

The SOF, CRC, EOF control fields of the FC frame and other optional headers have been omitted in the figure for clarity.

3.2 MTU

3.2.1 IP MTU

An FC Information Unit specific to each protocol such as IP is defined in FC-4. This defines the upper bound on the size of the information that can be transported.

Each IP or ARP Packet is mapped to a single FC Information Unit, which in turn is mapped to a single FC Sequence. There is a one-to-one mapping between an IP or ARP packet and a FC Sequence.

Fibre Channel limits the size of a single Information Unit to 2^32-1, which is very large [2]. However, since the Maximum Transmission Unit (MTU) size of an IPv4 packet does not exceed 65,536-bytes, the mapped IPv4 size is far below the 2^32-1 limit.

IPv4 Packet definition includes the IP Payload and IP Headers - both fixed and optional. The corresponding FC Sequence Payload includes the LLC/SNAP Header and the IPv4 packet.

As noted above, the greatest length allowed for an IPv4 Packet including any optional headers and independent of this standard is 65,536-bytes. However, limiting the IP MTU size to 65,280-bytes helps in buffer resource allocation at N_Ports and also allows for up to 256-bytes of overhead. Since the FC Network_Header requires 16-bytes and the IEEE 802.2 LLC/SNAP header requires 8 bytes, it leaves 232 bytes for future use.

All implementations SHALL restrict the IP MTU size to 65,280 bytes and the corresponding FC Sequence Payload size to 65536-bytes.

3.2.2 Maximally Minimum IPv4 Packet

In order for IP fragmentation and reassembly to work properly it is necessary that every implementation of IP be capable of transporting a maximally minimum size IP packet without fragmentation. A maximally minimum size IP Packet is defined as an IP Packet with an 8-byte payload (the smallest possible non-zero payload size for a fragment) and a 60-byte header (the largest possible header consisting of a 20-byte fixed part and a maximum size option field of 40-bytes) [17].

All implementations SHALL support a FC Data Field of 92-bytes, which is required to support 68-bytes of the maximally minimum sized IP Packet, 16-bytes of the FC Network_Header, and 8-bytes of the LLC/SNAP Header.

3.2.3 ARP MTU

The ARP packet has a fixed size of 28-bytes. All implementations SHALL support a FC Data Field size of 52-bytes, which is required to support 28-bytes of an ARP Packet, 16-bytes of the FC Network_Header, and 8-bytes of the LLC/SNAP Header. Note that the minimum MTU requirement for ARP is already covered by the minimum MTU requirement for IP but it is mentioned here for completeness.

The InARP packet is identical in size to the ARP and the same MTU requirements apply.

3.2.4 FC Data Field containing FARP Packet

The FARP Command is a FC Extended Link Service (ELS) command and maps directly to the FC Data Field without the LLC/SNAP or the FC Network_Header. The FARP Command has a fixed size of 76-bytes. Because FARP operates purely in the FC space, it places no special MTU requirements in this specification.

3.3 FC Port and Node Network Addresses

FC devices are identified by Nodes and their Ports. A Node is a collection of one or more Ports identified by a unique nonvolatile 64-bit World Wide Node name (WW_NN). Each Port in a node, is identified with a unique nonvolatile 64-bit World Wide Port name (WW_PN), and a volatile Port Identifier (Port_ID).

Port_IDs are 24-bits long. A FC frame header carries a Source Port_ID (S_ID) and a Destination Port_ID (D_ID). The Port_ID of a given port is volatile. (The mechanism(s) by which a Port_ID may change in a FC topology is outside the scope of this document. See Appendix D).

The FC Network_Header is normally optional in FC Standards, but REQUIRED in this specification. A FC Network_Header carries source and destination WW_PNs. A WW_PN consists of a 60-bit Network Address and a upper 4-bit Network Address Authority (NAA) field as shown in Fig. 3. The 4-bit NAA field is used to distinguish between the various name registration authorities used to define the Network Address [2].

In this specification, both the Source and Destination 4-bit NAA identifiers SHALL be set to binary ’0001’ indicating that an IEEE 48-bit MAC address is contained in the lower 48 bits of the network address fields. The high order 12 bits in the network address fields SHALL be set to 0x0000. The NAA field value equal to binary ’0001’ allows FC networks to be bridged with other FC networks or traditional LANs.

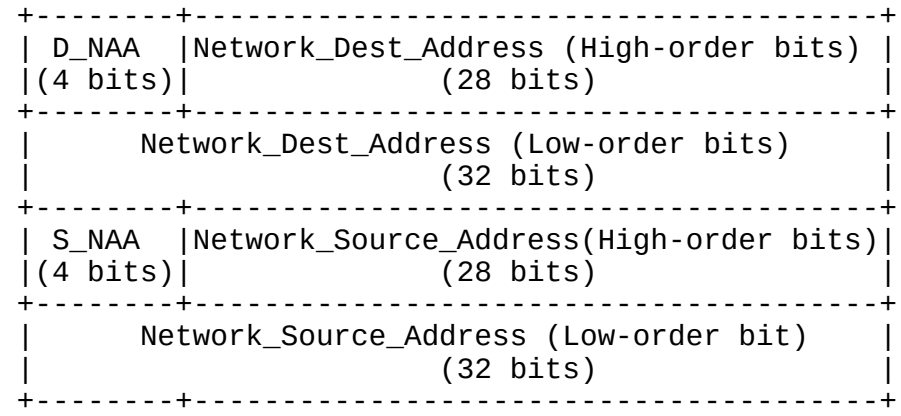


Fig. 3 Format of the Network_Header Field

3.4 FC Sequence Payload Format

FC Payload with IP:

An FC Sequence Payload carrying an IP and ARP packet SHALL use the formats shown in Figs. 4 and 5 respectively. Both formats use the 8-byte LLC/SNAP header.

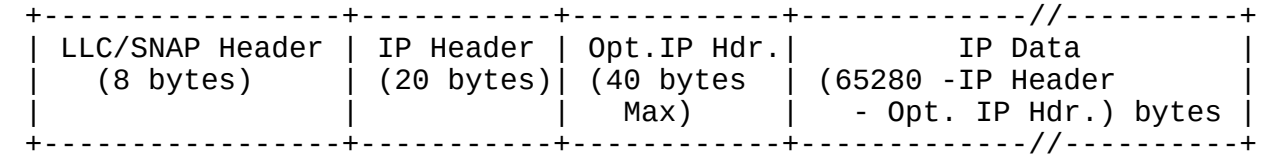


Fig. 4 Format of FC Sequence Payload carrying IP

FC Sequence Payload with ARP:

As noted earlier, FC frames belonging to the same Sequence may be delivered out of order over a Fabric. If the Relative Offset method is used to identify FC Sequence Payload fragments, then the IP Header MUST appear in the frame that has a relative offset of 0.

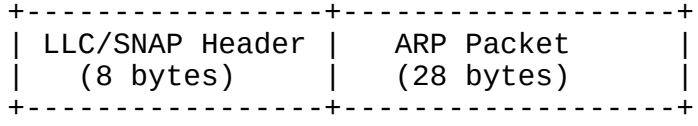


Fig. 5 Format of FC Sequence Payload carrying ARP

FC Sequence Payload with FARP:

FARP Protocol commands are directly mapped to the Frame Sequence Payload and are 76-bytes long. No LLC/SNAP Header or FC Network_Header is used and therefore the FC Data Field size simply consists of the FC Sequence Payload.

LLC:

A Logical Link Control (LLC) field along with a Sub Network Access Protocol (SNAP) field is a method used to identify routed and bridged non-OSI protocol PDUs and is defined by IEEE 802.2 and applied to IP in [8]. In LLC Type 1 operation (i.e., unacknowledged connectionless mode), the LLC header is 3-bytes long and consists of a 1-byte Destination Service Access Point (DSAP)field, a 1-byte Source Service Access Point (SSAP)field, and a 1-byte Control field as shown in Fig. 6.

Fig. 6 LLC Format

The LLC's DSAP and SSAP values of 0xAA indicate that an IEEE 802.2 SNAP header follows. The LLC's CTRL value equal to 0x03 specifies an Unnumbered Information Command PDU. In this specification the LLC Header value SHALL be set to 0xAA-AA-03. Other values of DSAP/SSAP indicate support for other protocols and SHALL NOT be used in this specification.

SNAP:

The SNAP Header is 5-bytes long and consists of a 3-byte Organizationally Unique Identifier (OUI) field and a 2-byte Protocol Identifier (PID) as shown in Fig. 7

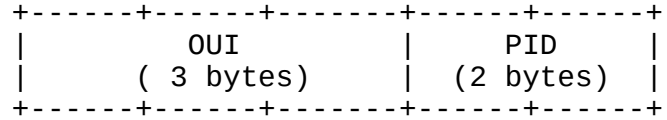


Fig. 7 SNAP Format

SNAP was invented to "encapsulate" LAN frames within the payload. The SNAP OUI value equal to 0x00-00-00 specifies that the PID is an EtherType (i.e., routed non-OSI protocol).

The SNAP OUI value equal to 0x00-80-C2 indicates Bridged Protocols.

With the OUI value set to 0x00-00-00, the SNAP PID value equal to 0x08-00 indicates IP and a PID value equal to 0x08-06 indicates ARP (or InARP).

The complete LLC/SNAP Header is shown in Fig. 8.

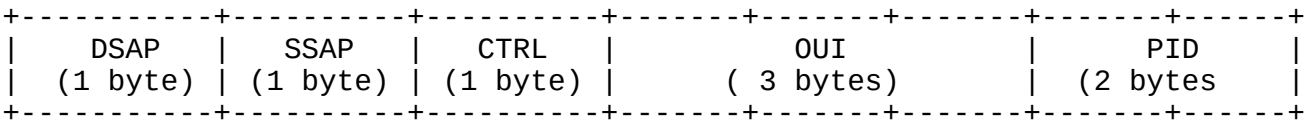


Fig. 8 LLC/SNAP Header

3.5 Bit and Byte Ordering

IP or ARP Packets are mapped to FC-4 Level using the big endian byte ordering, which corresponds to the standard network byte order or canonical form [20]. FC-4 Payload maps with no change in order to the FC-2 Level.

FC-1 Level defines the method used to encode data prior to transmission and subsequently decode the data upon reception. The method encodes 8-bit bytes into 10-bit transmission characters to improve the transmission characteristics of the serial data stream. In Fibre Channel, data fields are aligned on word boundaries. See Appendix E. A word in FC is defined as 4 bytes or 32 bits. The resulting transmission word after the 8-bit to 10-bit encoding consists of 40 bits.

Data words or Ordered Sets (special FC-2 Level control words) from the FC-2 Level map to the FC-1 Level with no change in order and the bytes in the word are transmitted in the Most Significant Byte first to Least Significant Byte order. The transmission order of bits within each byte is the Least Significant Bit to the Most Significant Bit.

4. ARP

4.1 Address Resolution

Address Resolution in this specification is primarily concerned with associating IP addresses with FC Port addresses. As described earlier, FC device ports have two types of addresses:

- a non-volatile unique 64-bit address called World Wide Port_Name (WW_PN)
- a volatile 24-bit address called a Port_ID

The Address Resolution mechanism therefore will need two levels of mapping:

1. A mapping from the IP address to the WW_PN (i.e., IEEE 48-bit MAC address)

2. A mapping from the WW_PN to the Port_ID (see Appendix G for a definition of Port_ID)

The address resolution problem is compounded by the fact that the Port_ID is volatile and the second mapping MUST be valid before use. Moreover, this validation process can be different depending on the network topology used. Appendix D provides a discussion on validation for the different FC topologies.

Architecturally, the first level of mapping and control operation is handled by the Address Resolution Protocol (ARP), and the second level by the FC Address Resolution Protocol (FARP). FARP is described in Section 5.

Other optional mechanisms in FARP that directly map an IP address to a Port_ID, or WW_NN to a Port_ID are described in Appendix A.

The Inverse Address Resolution Protocol (InARP) is yet another optional mechanism that resolves WW_PN and Port_IDs to IP addresses. InARP is described in Appendix B.

4.2 ARP Packet Format

The Address Resolution Protocol (ARP) given in [9] was designed to be a general purpose protocol, and to work with many network technologies, and with many upper layer protocols. Fig 9 shows the ARP packet format based on [9], where the upper layer protocol uses a 4 octet protocol (IP) address and the network technology uses six-octet hardware (MAC) address.

The ARP uses two packet types - Request and Reply - and each type of packet is 28 -bytes long in this specification. The ARP Packet fields are common to both ARP Requests and ARP Replys.

The LLC/SNAP encapsulated ARP Request Packet is mapped to a FC Broadcast Sequence and the exact mechanism used to broadcast a FC Sequence depends on the FC topology. This is discussed later in this section. Compliant ARP Request Broadcasts SHALL include Network_Headers.

The LLC/SNAP encapsulated ARP Reply Packet is mapped to a FC Sequence. Compliant ARP Replys SHALL include Network_Headers.

Note that in all discussions to follow, the WW_PN and the 48-bit MAC address conceptually mean the same thing.

The 'HW Type' field SHALL be set to 0x00-01.

Technically, the correct HW Type value should be set to 0x00-06 according to RFC 1700 indicating IEEE 802 networks. However, as a practical matter a HW Type value of 0x00-06 is known to cause rejections from some Ethernet end stations when FC is bridged to Ethernet. Translational bridges are normally expected to change this field from Type 6 to 1 and vice versa under these configurations, but many do not. It is because of this reason that the Type Code is set to 1 rather than 6. However, both HW Type values of 0x00-01 and 0x00-06 SHALL be accepted.

The 'Protocol' field SHALL be set to 0x08-00 indicating IP protocol.

The 'HW Addr Length' field SHALL be set to 0x06 indicating 6-bytes of HW address.

The 'Protocol Addr Length' field SHALL be set to 0x04 indicating 4-bytes of IPv4 address.

The 'Operation' Code field SHALL be set as follows:

0x00-01 for ARP Request
0x00-02 for ARP Reply

The 'HW Addr of Sender' field SHALL be the 6-byte IEEE MAC address of the sender. It is either the Requester (ARP Request) or the Responder (ARP Reply) address.

The 'Protocol Addr of Sender' field SHALL be the 4-byte IP address of the Requester (ARP Request) or that of the Responder (ARP Reply).

The 'HW Addr of Target' field SHALL be set to zero during an ARP Request and to the 6-byte MAC address of the Requester (ARP Request) in an ARP Reply.

The 'Protocol Addr of Target' field SHALL be set to the 4-byte IP address of the Responder (ARP Reply) in a ARP Request, and to the 4-byte IP address of the Requester (ARP Request) in an ARP Reply.

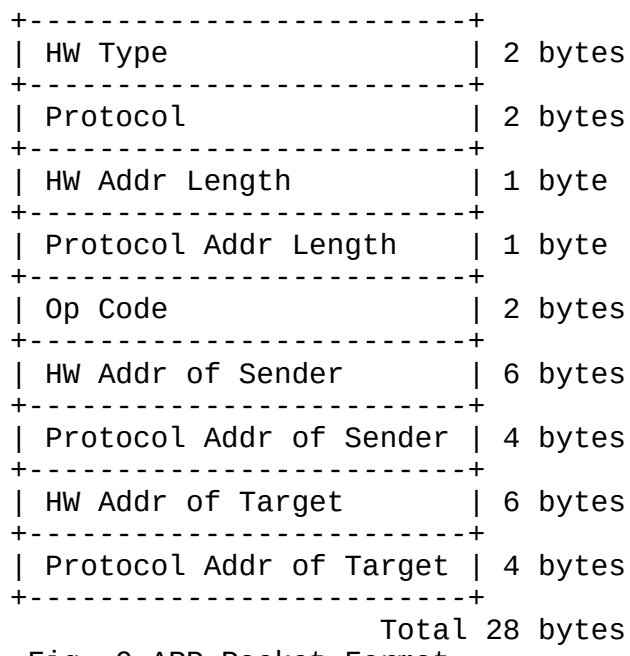


Fig. 9 ARP Packet Format

4.3 ARP Layer Mapping and Operation

Whenever a FC port wishes to send IP data to another FC port, then the following steps are taken:

1. The source port should first consult its local mapping tables to determine the <destination IP address, destination WW_PN>.

2. If such a mapping is found, then the source sends the IP data to the port whose WW_PN address was found in the table.

3. If such a mapping is not found, then the source sends an ARP Request broadcast to its connected FC network in anticipation of getting a reply from the correct destination along with its WW_PN.

4. When an ARP Request Broadcast frame is received by a node with the matching IP address, it generates an ARP Reply. Since the ARP Reply must be addressed to a specific destination Port_ID, the FC layer mapping between the WW_PN and Port_ID (of the ARP Request orginator) MUST be valid before the reply is sent.

5. If no node has the matching IP address, the result is a silent behavior.

4.4 ARP Broadcast in a Point-to-Point Topology

The ARP Request (Broadcast) and Reply mechanism described above still apply, although there is only one node that receives the ARP Request.

4.5 ARP Broadcast in a Private Loop Topology

In a private loop, the ARP Request Broadcast frame is sent using the broadcast method specified in the FC-AL [7]standard.

1. The source port first sends an Open Broadcast Replicate primitive (OPN(fr))Signal forcing all the ports in the loop (except itself), to replicate the frames that they receive while examining the frame header's Destination_ID field.

2. The source port then removes this OPN(fr) signal when it returns to it.

3. The loop is now ready to receive the ARP broadcast. The source now sends the ARP Request as a single-frame Broadcast Sequence in a Class 3 frame with the following FC Header D_ID field and F_CTL bits setting:

Destination ID <Word 0, bit 0:23>: D_ID = 0xFF-FF-FF

Sequence Initiative <Word 2, bit23>: SI=0

Last Sequence <Word 2, bit 20>: LS=1

End Sequence <Word 2, bit 19>: ES=1.

4. A compliant ARP Broadcast Sequence frame SHALL include the Network_Header with destination MAC address set to 0xFF-FF-FF-FF-FF-FF and with NAA = b'0001'

5. The destination port recognizing its IP address in the ARP Request packet SHALL respond with an ARP Reply.

4.6 ARP Broadcast in a Public Loop Topology

The following steps will be followed when a port is configured in a public loop:

1. A public loop device attached to a fabric through a FL_Port MUST NOT use the OPN(fr) signal primitive. Rather, it sends the broadcast sequence to the FL_Port at AL_PA = 0x00.

2. A FC Fabric propagates the broadcast to all other ports including the FL_Port which the broadcast arrived on. This includes all F_Ports, and other FL_Ports.

3. On each FL_Port, the fabric propagates the broadcast by first using the primitive signal OPNfr, in order to prepare the loop to receive the broadcast sequence.

4. A Broadcast Sequence is now sent on all ports (all FL_ports, F_Ports) in Class 3 frame with:

Destination ID <Word 0, bit 23:0>: D_ID = 0xFF-FF-FF

Sequence Initiative <Word 2, bit23>: SI=0

Last Sequence <Word 2, bit 20>: LS=1

End Sequence <Word 2, bit 19>: ES=1.

5. A compliant ARP Broadcast Sequence frame SHALL include the Network_Header with destination MAC address set to 0xFF-FF-FF-FF-FF-FF and with NAA = b'0001'

6. The destination port recognizing its IP address in the ARP Request packet SHALL respond with an ARP Reply.

4.7 ARP Operation in a Fabric Topology

1. Nodes directly attached to fabric do not require the OPN(fr) primitive signal.

2. A Broadcast Sequence is now sent on all ports (all FL_ports, F_Ports) in Class 3 frame with:

    Destination ID <Word 0, bit 23:0>: D_ID = 0xFF-FF-FF

    Sequence Initiative <Word 2, bit23>: SI=0

    Last Sequence <Word 2, bit 20>: LS=1

    End Sequence <Word 2, bit 19>: ES=1.

3. A compliant ARP Broadcast Sequence frame SHALL include the Network_Header with destination MAC address set to 0xFF-FF-FF-FF-FF-FF and with NAA = b'0001'

4. The destination port recognizing its IP address in the ARP packet SHALL respond with an ARP Reply.

5. FARP

5.1 Scope

FC Layer Mapping between the WW_PN and the Port_ID is independent of the ARP mechanism and is more closely associated with the details of the FC protocols. Name Server and FC Address Resolution Protocol (FARP) are two formal mechanisms that can be used to create and maintain WW_PN to Port_ID tables.

FARP is a method using Extended Link Service (ELS) commands that resolves <WW_PN, Port_ID> mappings. The WW_PN to Port_ID address resolution using FARP is especially useful in instances where the Login table entries at a node expire and a Name Server is not available. It is outside the scope of this document to describe Name Server. (See [14].)

Additional address matching mechanisms that resolve <WW_NN, Port_ID> and <IP addr., Port_ID> mapping have been added to FARP. These additional mechanisms are optional and described in Appendix A. Direct IP address to Port_ID mapping is useful in applications where there is no visibility of the MAC address.

Other less formal FC Layer Mapping mechanisms are described in Appendix C.

Since Port_IDs are volatile, all mapped Port_IDs at all times MUST be valid before use. There are many events that can invalidate this mapping. Appendix D discusses conditions when such a validation is required.

5.2 FARP Overview

The FARP protocol uses two ELS commands - FARP-REQ and FARP-REPLY.

Note: In the following discussion 'Requester' means the node issuing the FARP-REQ ELS message; 'Responder' means the node replying to the request by sending the FARP-REPLY command.

The FARP-REQ ELS Broadcast Request command is used to retrieve a specific node's current Port_ID given its unique WW_PN. This Port_ID is sent in a FARP-REPLY unicast command.

The FARP-REQ may indicate that the Responder:

- Perform only a Login with it (Requester) or,
- Send only a FARP-REPLY or,
- Perform a Login and send a FARP-REPLY.

No sequence initiative is transferred with the FARP-REQ and therefore no Reply (ACCEPT or REJECT) follows this command.

Since a Sequence Initiative is transferred with the FARP-REPLY, either a ACCEPT or REJECT follows this command as a response.

Reception of a FARP-REQ requires a higher level entity at the responding node to send a FARP-REPLY or perform a Port Login.

You do not have to be logged in to issue a FARP Request. Also, you do not have to be logged in to the FARP Requester to issue a FARP-REPLY.

The FARP Protocol Steps:

FARP-REQ (ELS broadcast) Request Sequence

(No Reply Sequence)

FARP-REPLY (ELS command) Sequence

Accept/Reject Reply Sequence

The FARP Protocol Format [2] and Size:

FT_1, 76-bytes fixed size

The FARP Protocol Addressing:

- In a FARP-REQ, the S_ID in the FC Header designates the Requester’s Port ID. The D_ID in the FC Header is the broadcast identifier 0xFF-FF-FF.

- In a FARP-REPLY, the S_ID in the FC Header designates the Responder’s Port_ID. The D_ID in the FC Header is the Requester’s Port_ID.

5.3 FARP Command Format

FARP-REQ and FARP-REPLY commands have identical formats (76-bytes fixed size) and fields but use different command codes. See tables below.

FARP-REQ Command		
Field	Size (Bytes)	Remarks
0x54-00-00-00	4	Request Command Code
Match Address Code Points	1	Indicates Address Matching Mechanism
Port_ID of Requester	3	Supplied by Requester = S_ID in FC Header
Responder Flags	1	Response Action to be taken
Port_ID of Responder	3	Set to 0x00-00-00
WW_PN of Requester	8	Supplied by Requester
WW_NN of Requester	8	OPTIONAL; See Appendix A
WW_PN of Responder	8	Supplied by Requester
WW_NN of Responder	8	OPTIONAL; see App. A
IP Address of Requester	16	OPTIONAL; see App. A
IP Address of Responder	16	OPTIONAL; see App. A

FARP-REPLY Command		
Field	Size (Bytes)	Remarks
0x55-00-00-00	4	Reply Command Code
Match Address Code Points	1	Not Used and Unchanged from the FARP-REQ
Port_ID of Requester	3	Extracted from FARP-REQ
Responder Flags	1	Not Used and Unchanged from the FARP-REQ
Port_ID of Responder	3	Supplied by Responder = S_ID in FC Header
WW_PN of Requester	8	Supplied by Requester
WW_NN of Requester	8	OPTIONAL; see App. A
WW_PN of Responder	8	Supplied by Requester
WW_NN of Responder	8	OPTIONAL; see App. A
IP Add. of Requester	16	OPTIONAL; see App. A
IP Address of Responder	16	OPTIONAL; see App. A

Following is a description of the address fields in the FARP Commands.

Port_ID of Requester:

It is the 24-bit Port_ID used in the S_ID field of the FC Header of a FARP-REQ. It is supplied by the Requester in a FARP-REQ and retained in a FARP-REPLY.

Port_ID of Responder:

It is the 24-bit Port_ID used in the S_ID field of the FC Header of a FARP-REPLY. It SHALL be set to 0x00-00-00 in a FARP-REQ. It is supplied by the Responder in a FARP-REPLY.

WW_PN:

This address field is used with the b'001', b'011', b'101, b'111', Match Address Code Points. See Match Address Code Point Table below. The Requester supplies the unique 8-byte WW_PN of the Requester and the Responder. It is retained in a FARP-REPLY.

WW_NN:

The WW_NN address field is used with Match Address Code Points b'010', b'011', b'110', and b'111', which are all optional. Its usage is fully described in Appendix A. When the WW_NN field is not used it SHALL be either set to '0' or a valid non-zero address.

IPv4:

The IPv4 address field is used with the Match Address Code Points b'100', b'101', b'110', and b'111', which are all optional. Its usage is fully described in Appendix A. When the IP Address field is not used it SHALL be either set to '0' or a valid IP address. A valid IP address consists of the 32-bit IPv4 Address with the upper 96 bits set to '0'.

5.4 Match Address Code Points

For each receipt of the FARP-REQ Broadcast ELS, the recipients match one or more addresses based on the encoded bits of the "FARP Match Address Code Points" field shown in the table below. FARP operation with the Match Address Code Point equal to b'001' is described in this section. Other code points are OPTIONAL and are discussed in Appendix A. The upper 5 bits of the Match Address Code Point byte are unused and their use is not currently defined.

Match Address Code Points		
LSBits	Bit name	Action
000	Reserved	
001	MATCH_WW_PN	If 'WW_PN of Responder' = Node's WW_PN then respond
010	MATCH_WW_NN	OPTIONAL; see Appendix A
011	MATCH_WW_PN_NN	OPTIONAL; see Appendix A
100	MATCH_IPv4	OPTIONAL; see Appendix A
101	MATCH_WW_PN_IPv4	OPTIONAL; see Appendix A
110	MATCH_WW_NN_IPv4	OPTIONAL; see Appendix A
111	MATCH_WW_PN_NN_IPv4	OPTIONAL; see Appendix A

When a node receives a FARP-REQ with Code Point b'001', it checks its WW_PN against the one set in 'WW_PN of Responder' field of the FARP-REQ command. If there is a match, then the node issues a response according to the action indicated by the FARP Responder Flag. See table below.

WW_NN and IPv4 address fields are not used with the b'001' Code Point operation. They SHALL be set to '0' or a valid address either by the Requester or the Requester and the Responder.

Note that there can be utmost one FARP-REPLY per FARP-REQ.

5.5 Responder Flags

The Responder Flags define what Responder action to take if the result of the Match Address Code Points is successful. 'Responder Flags' is an 8-bit field (bits 0-7) and is defined in the table below. This field is used only in a FARP-REQ. This field is retained unchanged in a FARP-REPLY. If no bits are set, the Responder will take no action.

	FARP Responder Flag	
Bit Position	Bit Name	Action
0	INIT_P_LOGI	Initiate a P_LOGI to the Requester
1	INIT_REPLY	Send FARP_REPLY to Requester
2 to 7	Reserved	

If INIT_P_LOGI bit is set then, a Login is performed with the port identified by "Port_ID of Requester" field.

If INIT_REPLY is set then, a FARP-REPLY is sent to the Port Identified by "Port_ID of Requester" field.

If both bits are set at the same time, then both Actions are performed.

All other bit patterns are undefined at this time and are reserved for possible future use.

5.6 FARP Support Requirements

Responder action - FARP-REPLY and/or Port Login - for a successful MATCH_WW_PN is always REQUIRED. If there is no address match then a silent behavior is specified.

Support for all other Match Address Code Points is OPTIONAL and a silent behavior from the Responder is valid when it is not supported. Recipients of the FARP-REQ ELS SHALL NOT issue a Service Reject (LS_RJT) if FARP OPTIONAL mechanisms are not supported.

In all cases, if there are no matches, then a silent behavior is specified.

If an implementation issues a FARP-REQ with a Match Address Code Point that is OPTIONAL, and fails to receive a response, and the implementation has not obtained the Port_ID of the Responder’s port by other means (e.g., prior FARP-REQ with other Code Points), then the implementation SHALL reattempt the FARP-REQ with the MATCH_WW_PN Code Point.

Getting multiple FARP Replies corresponding to a single FARP-REQ should normally never occur. It is beyond the scope of this document to specify conditions under which this error may occur or what the corrective action ought to be.

6. Exchange Management

6.1 Exchange Origination

FC Exchanges shall be established to transfer data between ports. Frames on IP exchanges shall not transfer Sequence Initiative. See Appendix E for a discussion on FC Exchanges.

6.2 Exchange Termination

With the exception of the recommendations in Appendix F, Section F.1, "Reliability in Class 3", the mechanism for aging or expiring exchanges based on activity, timeout, or other method is outside the scope of this document.

Exchanges may be terminated by either port. The Exchange Originator may terminate Exchanges by setting the LS bit, following normal FC standard FC-PH [2] rules. This specification prohibits the use of the NOP ELS with LS set for Exchange termination.

Exchanges may be torn down by the Exchange Originator or Exchange Responder by using the ABTS_LS protocol. The use of ABTS_LS for terminating aged Exchanges or error recovery is outside the scope of this document.

The termination of IP Exchanges by Logout is discouraged, since this may terminate active Exchanges on other FC-4s.

7. Summary of Supported Features

Note: 'Settable' means support is as specified in the relevant standard; all other key words are as defined earlier in this document.

7.1 FC-4 Header

Feature	Support	Notes
Type Code (= 5) ISO8802-2 LLC/SNAP	REQUIRED	2
Network_Headers	REQUIRED	3
Other Optional Headers	MUST NOT	

Notes:

1. This table applies only to FC-4 related data, such as IP and ARP packets. This table does not apply to link services and other non-FC-4 sequences (PLOGI, for example) that must occur for normal operation.

2. The TYPE field in the FC Header (Word 2 bits 31-24) MUST indicate ISO 8802-2 LLC/SNAP Encapsulation (Type 5). This revision of the document focuses solely on the issues related to running IP and ARP over FC. All other issues are outside the scope of this document, including full support for IEEE 802.2 LLC.

3. DF_CTL field (Word 3, bits 23-16 of FC-Header) MUST indicate the presence of a Network_Header (0010 0000) on the First logical Frame of FC-4 Sequences. It should not indicate the presence of a Network_Header on any subsequent frames of the Sequence.

7.2 R_CTL

R_CTL in FC-Header: Word 0, bits 31-24

Feature	Support	Notes
Information Category (R_CTL Routing):		
FC-4 Device Data	REQUIRED	1
Extended Link Data	REQUIRED	
FC-4 Link Data	MUST NOT	
Video Data	MUST NOT	
Basic Link Data	REQUIRED	
Link Control	REQUIRED	
R_CTL information :		
Uncategorized	MUST NOT	
Solicited Data	MUST NOT	
Unsolicited Control	REQUIRED	
Solicited Control	REQUIRED	
Unsolicited Data	REQUIRED	1
Data Descriptor	MUST NOT	
Unsolicited Command	MUST NOT	
Command Status	MUST NOT	

Notes:

1. This is REQUIRED for FC-4 (IP and ARP) packets

   - Routing bits of R_CTL field MUST indicate Device Data frames (0000)
   - Information Category of R_CTL field MUST indicate Unsolicited Data (0100)

7.3 F_CTL

F_CTL in FC-Header: Word 2, bits 23-0

Feature	Support	Notes
Exchange Context	Settable	
Sequence Context	Settable	
First / Last / End Sequence (FS/LS/ES)	Settable	
Chained Sequence	MUST NOT	
Sequence Initiative (SI)	Settable	1
X_ID Reassigned / Invalidate	MUST NOT	
Unidirectional Transmit	Settable	
Continue Sequence Condition	REQUIRED	2
Abort Seq. Condition -continue and single Seq.	REQUIRED	3
Relative Offset - Unsolicited Data	Settable	4
Fill Bytes	Settable	

Notes

1. For FC-4 frames, each N_Port shall have a dedicated OX_ID for sending data to each N_Port in the network and a dedicated RX_ID for receiving data from each N_Port as well. Exchanges are used in a unidirectional mode, thus setting Sequence Initiative is not valid for FC-4 frames. Sequence Initiative is valid when using Extended Link Services.

2. This field is required to be 00, no information.

3. Sequence error policy is requested by an exchange originator in the F_CTL Abort Sequence Condition bits in the first data frame of the exchange. For Classes 1 and 2, ACK frame is required to be "continuous sequence".

4. Relative offset prohibited on all other types (Information Category) of frames.

7.4 Sequences

Feature	Support	Notes
Class 2 open Sequences / Exchange	1	1
Length of Seq. not limited by end-to-end credit	REQUIRED	2
IP and ARP Packet and FC Data Field sizes	REQUIRED	3
Capability to receive Sequence of maximum size	OPTIONAL	4
Sequence Streaming	MUST NOT	5
Stop Sequence Protocol	MUST NOT	
ACK_0 support	OPTIONAL	6
ACK_1 support	REQUIRED	6
ACK_N support	MUST NOT	
Class of Service for transmitted Sequences	Class 1, 2, or 3	7
Continuously Increasing Sequence Count	OPTIONAL	8, 9

Notes:

1. Only one active sequence per exchange is optional.

2. A Sequence Initiator shall be capable of transmitting Sequences containing more frames than the available credit indicated by a Sequence recipient at Login. FC-PH [2] end-to-end flow control rules will be followed when transmitting such Sequences.

3. a) IP MTU size is 65280-bytes and resulting FC Sequence Payload size is 65536-bytes.
   b) Maximally Minimum IP Packet size is 68-bytes and resulting FC Data Field size is 92-bytes.
   c) ARP (and InARP) Packet size is 28-bytes and resulting FC Data Field size is 52-bytes.

4. Some OS environments may not handle the max Sequence Payload size of 65536. It is up to the administrator to configure the Max size for all systems.

5. All class 3 sequences are assumed to be non-streamed.

6. Only applies for Class 1 and 2. Use of ACK_1 is default, ACK_0 used if indicated by Sequence recipient at Login.

7. The administrator configured class of service is used, except where otherwise specified (e.g. Broadcasts are always sent in Class 3).

8. Review Appendix F, "Reliability in Class 3".

9. The first frame of the first sequence of a new Exchange must have SEQ_CNT = 0 [2].

7.5 Exchanges

Feature	Support	Notes
X_ID interlock support	OPTIONAL	1
OX_ID=FFFF	MUST NOT	
RX_ID=FFFF	OPTIONAL	2
Action if no exchange resources available	P_RJT	3
Long Lived Exchanges	OPTIONAL	4
Reallocation of Idle Exchanges	OPTIONAL	

Notes:

1. Only applies to Classes 1 and 2, supported by the Exchange Originator. A Port SHALL be capable of interoperating with another Port that requires X_ID interlock. The Exchange Originator facility within the Port shall use the X_ID Interlock protocol in such cases.

2. An Exchange Responder is not required to assign RX_IDs. If a RX_ID of FFFF is assigned, it is identifying Exchanges based on S_ID / D_ID / OX_ID only.

3. In Classes 1 and 2, a Port shall reject a frame that would create a new Exchange with a P_RJT containing reason code "Unable to establish Exchange". In Class 3, the frame would be dropped.

4. When an Exchange is created between 2 Ports for IP/ARP data, it remains active while the ports are logged in with each other. An Exchange SHALL NOT transfer Sequence Initiative (SI). Broadcasts and ELS commands may use short lived Exchanges.

7.6 ARP and InARP

Feature	Support	Notes
ARP Server Support	MUST NOT	1
Response to ARP requests	REQUIRED	2
Class of Service for ARP requests	Class 3	3
Class of Service for ARP replies	Class 1, 2, or 3	4
Response to InARP requests	OPTIONAL	
Class of Service for InARP requests/replies	Class 1, 2 or 3	5

Notes:

1. Well-known Address FFFFFC is not used for ARP requests. Frames from Well-known address FFFFFC are not considered to be ARP frames. Broadcast support is REQUIRED for ARP.

2. The IP Address is mapped to a specific MAC address with ARP.

3. An ARP request is a Broadcast Sequence, therefore Class 3 is always used.

4. An ARP reply is a normal Sequence, thus the administrator configured class of service is used.

5. An InARP Request or Reply is a normal Sequence, thus an administrator configured class of service is used.

7.7 Extended Link Services (ELS)

Feature	Support	Notes
Class of service for ELS commands / responses	Class 1,2 or 3	1
Explicit N-Port Login	REQUIRED	
Explicit F-Port Login	REQUIRED	
FLOGI ELS command	REQUIRED	
PLOGI ELS command	REQUIRED	
ADISC ELS command	REQUIRED	
PDISC ELS command	OPTIONAL	2
FAN ELS command	REQUIRED	5
LOGO ELS command	REQUIRED	
FARP-REQ/FARP-REPLY ELS commands	REQUIRED	3
Other ELS command support	OPTIONAL	4

Notes:

1. The administrator configured class of service is used.

2. PDISC shall not be used as a Requester; ADISC shall be used instead. As a Responder, an implementation may need to respond to both ADISC and PDISC for compatibility with other specifications.

3. Responder Action - FARP-REPLY and/or Port Login - for a successful MATCH_WW_PN is always REQUIRED.
Support for all other match Address Codes Points is a silent behavior from the Responder is valid when it is not supported. Recipients of the FARP-REQ ELS shall not issue a Service Reject (LS_RJT) if FARP is not supported.

4. If other ELS commands are received an LS_RJT may be sent. NOP is not required by this specification, and shall not be used as a mechanism to terminate exchanges.

5. Required for FL_Ports

7.8 Login Parameters

Unless explicitly noted here, a compliant implementation shall use the login parameters as described in [4].

7.8.1 Common Service Parameters - FLOGI

- FC-PH Version, lowest version may be 0x09 to indicate 'minimum 4.3'.
- Can't use BB_Credit=0 for N_Port on a switched Fabric (F_Port).

7.8.2 Common Service Parameters - PLOGI

- FC-PH Version, lowest version may be 0x09 to indicate 'minimum 4.3'.
- Can't use BB_Credit=0 for N_Port in a Point-to-Point configuration

- Random Relative Offset is optional.

- Note that the 'Receive Data Field Size' fields specified in the PLOGI represent both optional headers and payload.

- The MAC Address can therefore be extracted from the 6 lower bytes of the WW_PN field (when the IEEE 48-bit Identifier format is chosen as the NAA) during PLOGI or ACC payload exchanged during Fibre Channel Login [2].

- The MAC Address can also be extracted from the WW_PN field in the Network_Header during ADISC (and ADISC ACC), or PDISC (and PDISC ACC).

7.8.3 Class Service Parameters - PLOGI

- Discard error policy only.

8. Security Considerations

8.1 IP and ARP Related

IP and ARP do not introduce any new security concerns beyond what already exists within the Fibre Channel Protocols and Technology. Therefore IP and ARP related Security does not require special consideration in this document.

8.2 FC Related

FC Standards [11] specify a Security Key Server (independent of IP and ARP) as an optional service. However, there are no known implementations of this server yet. Also, the previously defined [2] use of a Security Header has been discontinued [11].

9. Acknowledgement

This specification is based on FCA IP Profile, Version 3.3. The FCA IP Profile was a joint work of the Fibre Channel Association (FCA) vendor community. The following organizations or individuals have contributed to the creation of the FCA IP Profile: Adaptec, Ancor, Brocade, Clariion, Crossroads, emf Associates, Emulex, Finisar, Gadzoox, Hewlett Packard, Interphase, Jaycor, McData, Migration Associates, Orca Systems, Prisa, Q-Logic, Symbios, Systran, Tektronix, Univ. of Minnesota, Univ. of New Hamshire. Jon Infante from Emulex deserves special mention for his contributions to the FARP Protocol. The authors extend their thanks to all who provided comments and especially to Lansing Sloan from LLNL for his detailed comments.

10. References

[1] FCA IP Profile, Revision 3.3, May 15, 1997

[2] Fibre Channel Physical and Signaling Interface (FC-PH) , ANSI X3.230-1994

[3] Fibre Channel Link Encapsulation (FC-LE), Revision 1.1, June 26, 1996

[4] Fibre Channel Fabric Loop Attachment (FC-FLA), Rev. 2.7, August 12, 1997

[5] Fibre Channel Private Loop SCSI Direct Attach (FC-PLDA), Rev. 2.1, September 22, 1997

[6] Fibre Channel Physical and Signaling Interface-2 (FC-PH-2), Rev. 7.4, ANSI X3.297-1996

[7] Fibre Channel Arbitrated Loop (FC-AL), ANSI X3.272-1996

[8] Postel, J. and J. Reynolds, "A standard for the Transmission of IP Datagrams over IEEE 802 Networks", STD 43, RFC 1042, February 1988.

[9] Plummer, D. "An Ethernet Address Resolution Protocol -or- Converting Network Addresses to 48-bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, November 1982.

[10] FCSI IP Profile, FCSI-202, Revision 2.1, September 8, 1995

[11] Fibre Channel Physical and Signaling Interface -3 (FC-PH-3), Rev. 9.3, ANSI X3.303-199x

[12] Fibre Channel-The Basics, "Gary R. Stephens and Jan V. Dedek", Ancot Corporation

[13] Fibre Channel -Gigabit Communications and I/O for Computers Networks "Alan Benner", McGraw-Hill, 1996, ISBN 0-07-005669-2

[14] Fibre Channel Generic Services -2 (FC-GS-2), Rev. 5.2 X3.288-199x

[15] Bradley, T. and C. Brown, "Inverse Address Resolution Protocol", RFC 1293, January 1992.

[16] Bradley, T., Brown, C. and A. Malis, "Inverse Address Resolution Protocol", RFC 2390, August 1992.

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

[18] The Fibre Channel Consultant: A Comprehensive Introduction, "Robert W. Kembel", Northwest Learning Associates, 1998

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

[20] Narten, T. and C. Burton, "A Caution on The Canonical Ordering of Link-Layer Addresses", RFC 2469, December 1998.

11. Authors' Addresses

Murali Rajagopal
Gadzoox Networks, Inc.
711 Kimberly Avenue, Suite 100
Placentia, CA 92870

Phone: +1 714 577 6805
Fax: +1 714 524 8508
EMail: murali@gadzoox.com

Raj Bhagwat
Gadzoox Networks, Inc.
711 Kimberly Avenue, Suite 100
Placentia, CA 92870

Phone: +1 714 577 6806
Fax: +1 714 524 8508
EMail: raj@gadzoox.com

Wayne Rickard
Gadzoox Networks, Inc.
711 Kimberly Avenue, Suite 100
Placentia, CA 92870

Phone: +1 714 577 6803
Fax: +1 714 524 8508
EMail: wayne@gadzoox.com

Appendix A: Additional Matching Mechanisms in FARP

Section 5 described the FC Layer mapping between the WW_PN and the Port_ID using the FARP Protocol. This appendix describes other optional criteria for address matching and includes:

- WW_NN
- WW_PN & WW_NN at the same time
- IPv4
- IPv4 & WW_PN at the same time
- IPv4 & WW_NN at the same time
- IPv4 & WW_PN & WW_NN at the same time

Depending on the Match Address Code Points, the FARP protocol fundamentally resolves three main types of addresses to Port_IDs and is described in table below.

- For Match Address Code Point b'001': WW_PN Names fields are used to resolve the WW_PN names to Port_IDs. WW_NN and IP address fields are not used with these Code Points and SHALL be set to either '0' or valid addresses by Requester or Requester and Responder.
- For Match Address Code Point b'010': WW_NN Names fields are used to resolve the WW_NN names to Port_IDs. WW_PN and IP address fields are not used with these Code Points and SHALL be set to either '0' or valid addresses by Requester or Requester and Responder.
- For Match Address Code Point b'100': IPv4 fields are used to resolve the IPv4 addresses to Port_IDs. WW_PN and WW_NN fields are not used with these Code Points and SHALL be set to either '0' or valid addresses by Requester or Requester and Responder.
- For all other Match Address Code Points b'011', b'101',b'110', b'111', depending on set bits one or more addresses are jointly resolved to a Port_ID. See table below. If fields are not used, then they are set either to '0' or valid addresses.

The Responder Flags remain the same as before. Note that there can be utmost one FARP-REPLY per FARP-REQ.

Tables showing FARP-REQ and FARP-REPLY and address fields setting are given below:

Match Address Code Points		
LSBits	Bit name	Action
000	Reserved	
001	MATCH_WW_PN	If 'WW_PN of Responder' = Node's WW_PN then respond
010	MATCH_WW_NN	If 'WW_NN of Responder' = Node's WW_NN then respond
011	MATCH_WW_PN_NN	If both 'WW_PN of Responder' & 'WW_NN of Responder' = Node's WW_PN & WW_NN then respond
100	MATCH_IPv4	If 'IPv4 Address of Responder' = Node's IPv4 Address then respond
101	MATCH_WW_PN_IPv4	If 'WW_PN & IPv4 of Responder' = Node's WW_PN and IPv4 then respond
110	MATCH_WW_NN_IPv4	If both 'WW_NN of Responder' & 'IPv4 Address of Responder' = Node's WW_NN & IPv4 then respond
111	MATCH_WW_PN_NN_IPv4	If 'WW_PN of Responder' & 'WW_NN of Responder' & 'IPv4 Address of Responder' = Nodes' WW_PN & WW_NN & IPv4 then respond

FARP-REQ Command		
Field	Size (Bytes)	Remarks
0x54-00-00-00	4	Request Command Code
Match Address Code Points	1	Indicates Address Matching Mechanism
Port_ID of Requester	3	Supplied by Requester
Responder Flags	1	Response Action to be taken
Port_ID of Responder	3	Set to 0x00-00-00
WW_PN of Requester	8	Supplied by Requester
WW_NN of Requester	8	OPTIONAL; Supplied by Requester
WW_PN of Responder	8	Supplied by Requester
WW_NN of Responder	8	OPTIONAL ;Supplied by Requester or Responder
IP Add. of Requester	16	OPTIONAL; Supplied by Requester IPv4 Add.=low 32 bits
IP Address of Responder	16	OPTIONAL; Supplied by Requester or Responder IPv4 Add.=low 32 bits

FARP-REPLY Command		
Field	Size (Bytes)	Remarks
0x55-00-00-00	4	Reply Command Code
Match Address Code Points	1	Not Used and unchanged from the FARP-REQ
Port_ID of Requester	3	Supplied by Requester
Responder Flags	1	Not Used and unchanged from the FARP-REQ
Port_ID of Responder	3	Supplied by Responder
WW_PN of Requester	8	Supplied by Requester
WW_NN of Requester	8	OPTIONAL; Supplied by Requester
WW_PN of Responder	8	Supplied by Requester
WW_NN of Responder	8	OPTIONAL; Supplied by Requester or Responder
IP Add. of Requester	16	OPTIONAL; Supplied by Requester IPv4 Add.=low 32 bits
IP Address of Responder	16	OPTIONAL; Supplied by Requester or Responder IPv4 Add.=low 32 bits

Appendix B: InARP

B.1 General Discussion

Inverse ARP (InARP) is a mechanism described in RFC 1293/2390 [15, 16], which is useful when a node desires to know the protocol address of a target node whose hardware address is known. Situations where this could occur are described in [15, 16]. The motivation for using InARP in FC is to allow a node to learn the IP address of another node with which it has performed a Port Login (PLOGI). PLOGI is a normal FC process that happens between nodes, independent of this standard. PLOGI makes it possible for a node to discover the WW_PN and the Port_ID of the other node but not its IP address. A node in this way may potentially obtain the IP address of all nodes with which it can PLOGI.

Note that the use of the InARP mechanism can result in resolving all WW_PN to IP addresses and ARP may no longer be required. This can be beneficially applied in cases where a particular FC topology makes it inefficient to send out an ARP broadcast.

B.2 InARP Protocol Operation

InARP uses the same ARP Packet format but with different 'Op Codes', one for InARP Request and another for InARP Reply.

The InARP protocol operation is very simple. The requesting node fills the hardware address (WW_PN) of the target device and sets the protocol address to 0x00-00-00-00. Because, the request is sent to a node whose WW_PN and Port_ID are known, there is no need for a broadcast. The target node fills in its Protocol address (IP address in this case) and sends an InARP Reply back to the sender. A node may collect, all such WW_PN and IP addresses pairs in a similar way.

B.3 InARP Packet Format

Since the InARP protocol uses the same packet format as the ARP protocol, much of the discussion on ARP formats given in Section 4 applies here.

The InARP is 28-bytes long in this application and uses two packet types: Request and Reply. Like ARP, the InARP Packet fields are common to both InARP Requests and InARP Replies.

InARP Request and Reply Packets are encapsulated in a single frame FC Sequence much like ARP. Compliant InARP Request and Reply FC Sequences SHALL include Network_Headers.

The 'HW Type' field SHALL be set to 0x00-01.

The 'Protocol' field SHALL be set to 0x08-00 indicating IP protocol.

The 'HW Addr Length' field SHALL be set to 0x06 indicating 6-bytes of HW address.

The 'Protocol Addr Length' field SHALL be set to 0x04 indicating 4-bytes of IP address.

The 'Operation' Code field SHALL be set as follows:

0x00-08 for InARP Request
0x00-09 for InARP Reply

The 'HW Addr of Sender' field SHALL be the 6-byte IEEE MAC address of the Requester (InARP Request) or Responder (InARP Reply).

The 'Protocol Addr of Sender' field SHALL be the 4-byte IP address of the Requester (InARP Request) or Responder (InARP Reply).

The 'HW Addr of Target' field SHALL be set to the 6-byte MAC address of the Responder in an InARP Request and to the 6-byte MAC address of the Requester in an InARP Reply.

The 'Protocol Addr of Target' field SHALL be set to 0x00-00-00-00 in an InARP Request and to the 4-byte IP address of the Requester in an InARP Reply.

B.4 InARP Support Requirements

Support for InARP is OPTIONAL. If a node does not support InARP and it receives an InARP Request message then a silent behavior is specified.

APPENDIX C: Some Informal Mechanisms for FC Layer Mappings

Each method SHALL have some check to ensure PLOGI has completed successfully before data is sent. A related concern in large networks is limiting concurrent logins to only those ports with active IP traffic.

C.1 Login on Cached Mapping Information

This method insulates the level performing Login from the level interpreting ARP. It is more accommodating of non-ARP mechanisms for building the FC-layer mapping table.

1. Broadcast messages that carry a Network_Header contain the S_ID on the FC-header and WW_PN in the Network-Header. Caching this information provides a correlation of Port_ID to WW_PN. If the received Broadcast message is compliant with this specification, the WW_PN will contain the MAC Address.

2. The WW_PN is "available" if Login has been performed to the Port_ID and flagged. If Login has not been performed, the WW_PN is "unavailable".

3. If an outbound packet is destined for a port that is "unavailable", the cached information (from broadcast) is used to look up the Port_ID.

4. After sending an ELS PLOGI command (Port Login) to the Port (from a higher level entity at the host), waiting for an outbound packet before sending this Port Login conserves resources for only those ports which wish to establish communication.

5. After Port Login completes (ACC received), the outbound packet can be forwarded. At this point in time, both ends have the necessary information to complete their <IP address, MAC Address, Port_ID> association.

C.2 Login on ARP Parsing

This method performs Login sooner by parsing ARP before passing it up to higher levels for IP/MAC Address correlation. It requires a low-level awareness of the IP address, and is therefore protocol-specific.

1. When an ARP Broadcast Message is received, the S_ID is extracted from the FC-header and the corresponding Network_Source_Address from the Network_Header.

2. The ARP payload is parsed to determine if
   (a) this host is the target of the ARP request (Target IP Address match), and
   (b) if this host is currently logged in with the port (Port_ID = S_ID) originating the ARP broadcast.

3. The ARP is passed to a higher level for ARP Response generation.

4. If a Port Login is required, an ELS PLOGI command (Port Login) is sent immediately to the Port originating the ARP Broadcast.

5. After Port Login completes, an ARP response can be forwarded. Note that there are two possible scenarios:

   - The ACC to PLOGI returns before the ARP reply is processed and the ARP Reply is immediately forwarded.
   - The ARP reply is delayed, waiting for ACC (successful Login).

6. At this point in time, both ends have the necessary information to complete their <IP address, MAC Address, Port_ID> association.

C.3 Login to Everyone

In Fibre Channel topologies with a limited number of ports, it may be efficient to unconditionally Login to each port. This method is discouraged in fabric and public loop environments.

After Port Login completes, the MAC Address to Port_ID Address tables can be constructed.

C.4 Static Table

In some loop environments with a limited number of ports, a static mapping from a MAC Address to Port_ID (D_ID or AL_PA) may be maintained. The FC layer will always know the destination Port_ID based on the table. The table is typically downloaded into the driver at configuration time. This method scales poorly, and is therefore not recommended.

Appendix D: FC Layer Address Validation

D.1 General Discussion

At all times, the <WW_PN, Port_ID> mapping MUST be valid before use. There are many events that can invalidate this mapping. The following discussion addresses conditions when such a validation is required.

After a FC link interruption occurs, the Port_ID of a port may change. After the interruption, the Port_IDs of all other ports that have previously performed PLOGI (N_Port Login) with this port may have changed, and its own Port_ID may have changed.

Because of this, address validation is required after a LIP in a loop topology [7] or after NOS/OLS in a point-to-point topology [6].

Port_IDs will not change as a result of Link Reset (LR),thus address validation is not required.

In addition to actively validating devices after a link interruption, if a port receives any FC-4 data frames (other than broadcast frames), from a port that is not currently logged in, then it shall send an explicit Extended Link Service (ELS) Request logout (LOGO) command to that port.

ELS commands (Requests and Replies) are used by an N_Port to solicit a destination port (F_Port or N_Port) to perform some link-level function or service.) The LOGO Request is used to request invalidation of the service parameters and Port_ID of the recipient N_Port.

The level of initialization and subsequent validation and recovery reported to the upper (FC-4) layers is implementation-specific.

In general, an explicit Logout (LOGO) SHALL be sent whenever the FC-Layer mapping between the Port_ID and WW_PN of a remote port is removed.

The effect of power-up or re-boot on the mapping tables is outside the scope of this specification.

D.2 FC Layer Address Validation in a Point-to-Point Topology

No validation is required after LR. In a point-to-point topology, NOS/OLS causes implicit Logout of each port and after a NOS/OLS, each port must perform a PLOGI [2].

D.3 FC Layer Address Validation in a Private Loop Topology

After a LIP, a port SHALL not transmit any link data to another port until the address of the other port has been validated. The validation consists of completing either ADISC or PDISC. (See Appendix G.)

ADISC (Address Discovery) is an ELS command for discovering the hard addresses - the 24-bit identifier- of NL_Ports [5], [6].

PDISC (Discover Port) is an ELS command for exchanging service parameters without affecting Login state [5], [6].

As a requester, this specification prohibits PDISC and requires ADISC.

As a responder, an implementation may need to respond to both ADISC and PDISC for compatibility with other FC specifications.

If the three addresses, Port_ID, WW_PN, WW_NN, exactly match the values prior to the LIP, then any active exchanges may continue.

If any of the three addresses have changed, then the node must be explicitly Logged out [4], [5].

If a port’s N_Port ID changes after a LIP, then all active Port-ID to WW_PN mappings at this port must be explicitly Logged out.

D.4 FC Layer Address Validation in a Public Loop Topology

A FAN (Fabric Address Notification) ELS command is sent by the fabric to all known previously logged in ports following an initialization event. Therefore, after a LIP, hosts may wait for this notification to arrive or they may perform a FLOGI.

If the WW_PN and WW_NN of the fabric FL_Port contained in the FAN ELS or FLOGI response exactly match the values before the LIP, and if the AL_PA obtained by the port is the same as the one before the LIP, then the port may resume all exchanges. If not, then FLOGI (Fabric Login) must be performed with the fabric and all nodes must be explicitly Logged out.

A public loop device will have to perform the private loop authentication to any nodes on the local loop which have an Area + Domain Address == 0x00-00-XX

D.5 FC Layer Address Validation in a Fabric Topology

No validation is required after LR (link reset).

After NOS/OLS, a port must perform FLOGI. If, after FLOGI, the S_ID of the port, the WW_PN of the fabric, and the WW_NN of the fabric are the same as before the NOS/OLS, then the port may resume all exchanges. If not, all nodes must be explicitly, Logged out [2].

APPENDIX E: Fibre Channel Overview

E.1 Brief Tutorial

The FC Standard [2] defines 5 "levels" (not layers) for its protocol description: FC-0, FC-1, FC-2, FC-3, and FC-4. The first three levels (FC-0, FC-1, FC-2) are largely concerned with the physical formatting and control aspects of the protocol. FC-3 has been architected to provide a place holder for functions that might need to be performed after the upper layer protocol has requested the transmission of an information unit, but before FC-2 is asked to deliver that piece of information by using a sequence of frames [18]. At this time, no FC-3 functions have been defined. FC-4 is meant for supporting profiles of Upper Layer Protocols (ULP) such as IP and Small Computer System Interface (SCSI), and supports a relatively small set compared to LAN protocols such as IEEE 802.3.

FC devices are called "Nodes", each of which has at least one "Port" to connect to other ports. A Node may be a workstation, a disk drive or disk array, a camera, a display unit, etc. A "Link" is two unidirectional paths flowing in opposite directions and connecting two Ports within adjacent Nodes.

FC Nodes communicate using higher layer protocols such as SCSI and IP and are configured to operate using Point-to-Point, Private Loop, Public Loop (attachment to a Fabric), or Fabric network topologies.

The point-to-point is the simplest of the four topologies, where only two nodes communicate with each other. The private loop may connect a number of devices (max 126) in a logical ring much like Token Ring, and is distinguished from a public loop by the absence of a Fabric Node participating in the loop. The Fabric topology is a switched network where any attached node can communicate with any other. For a detail description of FC topologies refer to [18].

Table below summarizes the usage of port types depending on its location [12]. Note that E-Port is not relevant to any discussion in this specification but is included below for completeness.

Port Type	Location	Topology Associated with
N_Port	Node	Point-to-Point or Fabric
NL_Port	Node	In N_Port mode -Point-to-Point or Fabric In NL_Port mode - Arbitrated Loop
F_Port	Fabric	Fabric
FL_Port	Fabric	In F_Port mode - Fabric In FL_Port mode - Arbitrated Loop
E_Port	Fabric	Internal Fabric Expansion

E.2 Exchange, Information Unit, Sequence, and Frame

The FC 'Exchange' is a mechanism used by two FC ports to identify and manage an operation between them [18]. An Exchange is opened whenever an operation is started between two ports. The Exchange is closed when this operation ends.

The FC-4 Level specifies data units for each type of application level payload called 'Information Unit' (IU). Each protocol carried by FC has a defined size for the IU. Every operation must have at least one IU. Lower FC levels map this to a FC Sequence.

Typically, a Sequence consists of more than one frame. Larger user data is segmented and reassembled using two methods: Sequence Count and Relative Offset [18]. With the use of Sequence Count, data blocks are sent using frames with increasing sequence counts (modulo 65536) and it is quite straightforward to detect the first frame that contains the Network_Header. When Relative Offset is used, as frames arrive, some computation is required to detect the first frame that contains the Network_Header. Sequence Count and Relative Offset field control information, is carried in the FC Header.

The FC-4 Level makes a request to FC-3 Level when it wishes it to be delivered. Currently, there are no FC-3 level defined functions, and this level simply converts the Information Unit delivery request into a 'Sequence' delivery request and passes it on to the FC-2 Level. Therefore, each FC-4 Information Unit corresponds to a FC-2 Level Sequence.

The maximum data carried by a FC frame cannot exceed 2112-bytes [2]. Whenever, the Information Unit exceeds this value, the FC-2 breaks it into multiple frames and sends it in a sequence.

There can be multiple Sequences within an Exchange. Sequences within an Exchange are processed sequentially. Only one Sequence is active at a time. Within an Exchange information may flow in one direction only or both. If bi-directional then one of the ports has the initiative to send the next Sequence for that Exchange. Sequence Initiative can be passed between the ports on the last frame of Sequence when control is transferred. This amounts to half-duplex behavior.

Ports may have more than one Exchange open at a time. Ports can multiplex between Exchanges. Exchanges are uniquely identified by Exchange IDs (X_ID). An Originator OX_ID and a Responder RX_ID uniquely identify an Exchange.

E.3 Fibre Channel Header Fields

The FC header as shown in the diagrams below contains routing and other control information to manage Frames, Sequences, and Exchanges. The Frame-header is sent as 6 transmission words immediately following an SOF delimiter and before the Data Field.

D_ID and S_ID:

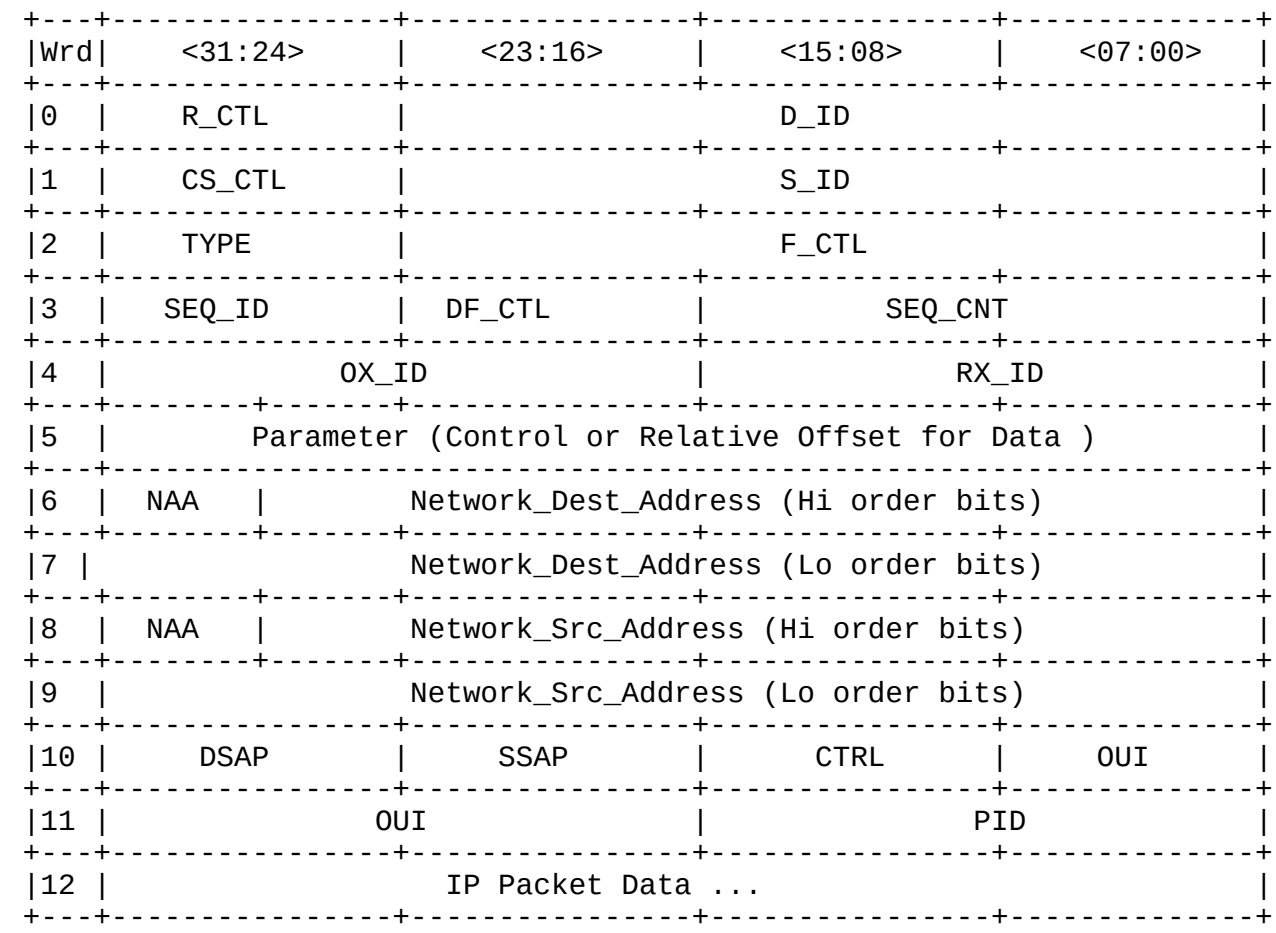
FC uses destination address routing [12], [13]. Frame routing in a point-to-point topology is trivial.

For the Arbitrated Loop topology, with the destination NL_Port on the same AL, the source port must pick the destination port, determine its AL Physical Address, and "Open" the destination port. The frames must pass through other NL_Ports or the FL_Port on the loop between the source and destination, but these ports do not capture the frames. They simply repeat and transmit the frame. Either communicating port may "Close" the circuit.

When the destination port is not on the same AL, the source NL_Port must open the FL_Port attached to a Fabric. Once in the Fabric, the Fabric routes the frames again to the destination.

In a Fabric topology, the Fabric looks into the Frame-header, extracts the destination address (D_ID), searches its own routing tables, and sends the frame to the destination port along the path chosen. The process of choosing a path may be performed at each fabric element or switch until the F_Port attached to the destination N_Port is reached.

Fibre Channel Frame Header, Network_Header, and Payload carrying IP Packet



R_CTL (Routing Control) and TYPE(data structure):

Frames for each FC-4 can be easily distinguished from the others at the receiving port using the R_CTL (Routing Control) and TYPE (data structure) fields in the Frame-header.

The R_CTL has two sub-fields: Routing bits and Information category. The Routing bits sub-field has specific values that mean FC-4 data follows and the Information Category tells the receiver the "Type" of data contained in the frame. The R_CTL and TYPE code points are shown in the diagrams.

Other Header fields:

F_CTL (Frame Control) and SEQ_ID (Sequence Identification), SEQ_CNT (Sequence Count), OX_ID (Originator exchange Identifier), RX_ID (Responder exchange Identifier), and Parameter fields are used to manage the contents of a frame, and mark information exchange boundaries for the destination port.

F_CTL(Frame Control):

The FC_CTL field is a 3-byte field that contains information relating to the frame content. Most of the other Frame-header fields are used for frame identification. Among other things, bits in this field indicate the First Sequence, Last Sequence, or End Sequence. Sequence Initiative bit is used to pass control of the next Sequence in the Exchange to the recipient.

SEQ_ID (Sequence Identifier) and SEQ_CNT (Sequence Count):

This is used to uniquely identify sequences within an Exchange. The <S_ID, D_ID, SEQ_ID> uniquely identifies any active Sequence. SEQ_CNT is used to uniquely identify frames within a Sequence to assure sequentiality of frame reception, and to allow unique correlation of link control frames with their related data frames.

Originator Exchange Identifier (OX_ID) and Responder Exchange Identifier (RX_ID):

The OX_ID value provides association of frames with specific Exchanges originating at a particular N_Port. The RX_ID field provides the same function that the OX_ID provides for the Exchange Originator. The OX_ID is meaningful on the Exchange Originator, and the RX_ID is meaningful on the Responder.

DF_CTL (Data Field Control):

The DF_CTL field specifies the presence or absence of optional headers between the Frame-header and Frame Payload

PARAMETER:

The Parameter field has two meanings, depending on Frame type. For Link Control Frames, the Parameter field indicates the specific type of Link Control frame. For Data frames, this field contains the Relative Offset value. This specifies an offset from an Upper Layer Protocol buffer from a base address.

E.4 Code Points for FC Frame

E.4.1 Code Points with IP and ARP Packets

The Code Points for FC Frames with IP and ARP Packets are very similar with the exception of PID value in Word 11 which is set to 0x08-00 for IP and 0x08-06 for ARP. Also, the Network_Header appears only in the first logical frame of a FC Sequence carrying IP. In the case, where FC frames carry ARP packets it is always present because these are single frame Sequences.

Code Points for FC Frame with IP packet Data

Wrd	<31:24>	<23:16>	<15:08>	<07:00>
0	0x04		D_ID	
1	0x00		S_ID	
2	0x05		F_CTL	
3	SEQ_ID	0x20	SEQ_CNT	
4	OX_ID		RX_ID	
5	0xXX-XX-XX Parameter Relative Offset			
6	0001	0x000	Dest. MAC (Hi order bits)	
7	Dest. MAC (Lo order bits)			
8	0001	0x000	Src. MAC (Hi order bits)	
9	Src. MAC (Lo order bits)			
10	0xAA	0xAA	0x03	0x00
11	0x00-00		0x08-00	
12	IP Packet Data			
13	...			

Code Points for FC Frame with ARP packet Data

Wrd	<31:24>	<23:16>	<15:08>	<07:00>
0	0x04		D_ID	
1	0x00		S_ID	
2	0x05		F_CTL	
3	SEQ_ID	0x20	SEQ_CNT	
4	OX_ID		RX_ID	
5	0xXX-XX-XX-XX Parameter Relative Offset			
6	0001	0x000	Dest. MAC (Hi order bits)	
7	Dest. MAC (Lo order bits)			
8	0001	0x000	Src. MAC (Hi order bits)	
9	Src. MAC (Lo order bits)			
10	0xAA	0xAA	0x03	0x00
11	0x00-00		0x08-06	
12	ARP Packet Data			
13	...			

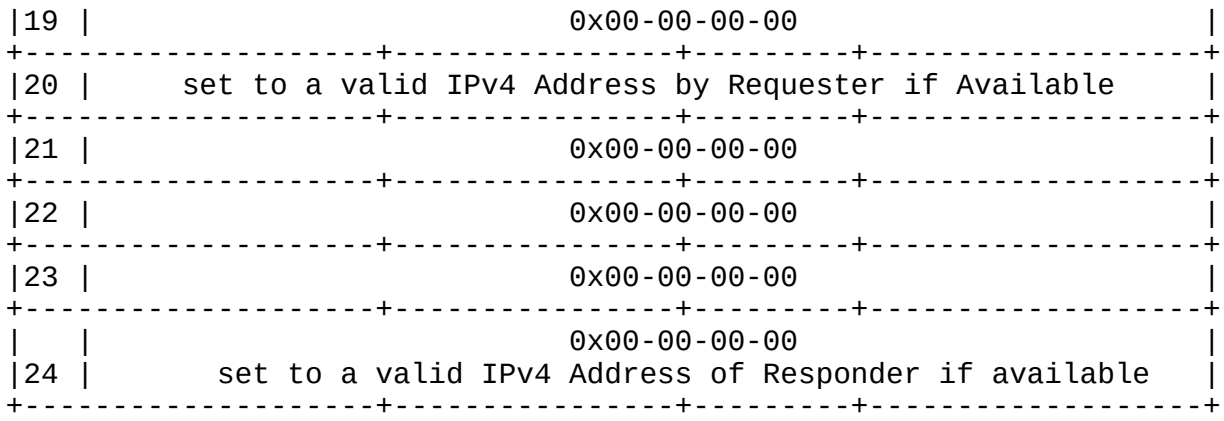
The Code Points for a FARP-REQ for a specific Match Address Code Point MATCH_WW_PN_NN (b'011') is shown below. In particular, note the IP addresses field of the Requester set to a valid address and that of the responder set to '0'. Note also the setting of the D_ID address and the Port_ID of the Responder.

The corresponding code point for a FARP-REPLY is also shown below. In particular, note the setting of the Port_ID of Responder and the IP address setting of the Responder.

E.4.2 Code Points with FARP Command

Code Points for FC Frame with FARP-REQ Command for MATCH_WW_PN_NN

Wrd	<31:24>	<23:16>	<15:08>	<07:00>
0	0x04	0xFF	D_ID = 0xFF	0xFF
1	0x00	S_ID		
2	0x05	F_CTL		
3	SEQ_ID	0x20	SEQ_CNT	
4	OX_ID		RX_ID	
5	0xXX-XX-XX Parameter Relative Offset			
6	0x54	0x00	0x00	0x00
7	Port_ID of Requester = S_ID			Match Addr. Code Points xxxxx011
8	Port_ID of Responder = 0x00	0x00	0x00	Responder Flags
9	0001	0x000	WW_PN Src. MAC(Hi order bits)	
10	WW_PN Src. MAC (Lo order bits)			
11	0001	0x000	WW_NN Src. MAC(Hi order bits)	
12	WW_NN Src. MAC (Lo order bits)			
13	0001	0x000	WW_PN Src. MAC(Hi order bits)	
14	WW_PN Dest. MAC (Lo order bits)			
15	0001	0x000	WW_NN Dest.MAC(Hi order bits)	
16	WW_NN Dest. MAC (Lo order bits)			
17	0x00-00-00-00			
18	0x00-00-00-00			


19	0x00-00-00-00
20	set to a valid IPv4 Address by Requester if Available
21	0x00-00-00-00
22	0x00-00-00-00
23	0x00-00-00-00
	0x00-00-00-00
24	set to a valid IPv4 Address of Responder if available

Code Points for FC Frame with FARP-REPLY Command

Wrd	<31:24>	<23:16>	<15:08>	<07:00>
0	0x04		D_ID	
1	0x00		S_ID	
2	0x05		F_CTL	
3	SEQ_ID	0x20	SEQ_CNT	
4	OX_ID		RX_ID	
5	0xXX-XX-XX Parameter Relative Offset			
6	0x55	0x00	0x00	0x00
7	Port_ID of Requester = D_ID			xxxxx011
8	Port_ID of Responder = S_ID			Resp. Flag
9	0001	0x000	WW_PN Src. MAC(Hi order bits)	
10	WW_PN Src. MAC (Lo order bits)			
11	0001	0x000	WW_NN Src. MAC(Hi order bits)	
12	WW_NN Src. MAC (Lo order bits)			
13	0001	0x000	WW_PN Src. MAC(Hi order bits)	
14	WW_PN Dest. MAC (Lo order bits)			
15	0001	0x000	WW_NN Dest.MAC(Hi order bits)	
16	WW_NN Dest. MAC (Lo order bits)			
17	0x00-00-00-00			
18	0x00-00-00-00			
19	0x00-00-00-00			
20	set to a valid IPv4 Address by Requester			
21	0x00-00-00-00			

22	0x00-00-00-00		
23	0x00-00-00-00		
24	set to a valid IPv4 Address by Responder		

APPENDIX F: Fibre Channel Protocol Considerations

F.1 Reliability In Class 3

Problem: Sequence ID reuse in Class 3 can conceivably result in missing frame aliasing, which could result in delivery of corrupted (incorrectly-assembled) data, with no corresponding detection at the FC level.

Prevention: This specification requires one of the following methods if Class 3 is used.

- Continuously increasing Sequence Count (new Login Bit) - both sides must set When an N_Port sets the PLOGI login bit for continuously increasing SEQ_CNT, it is guaranteeing that it will transmit all frames within an Exchange using a continuously increasing SEQ_CNT (see description in Section B.1 below).
- After using all SEQ_IDs (0-255) once, must start a new Exchange. It is recommended that a minimum of 4 Exchanges be used before an OX_ID can be reused.
- Note: If an implementation is not checking the OX_ID when reassembling Sequences, the problem can still occur. Cycling through some number of SEQ_IDs, then jumping to a new Exchange does not solve the problem. SEQ_IDs must still be unique between two N_Ports, even across Exchanges.
- Use only single-frame Sequences.

F.2 Continuously Increasing SEQ_CNT

This method allows the recipient to check incoming frames, knowing exactly what SEQ_CNT value to expect next. Since the SEQ_CNT will not repeat for 65,536 frames, the aliasing problem is significantly reduced.

A Login bit (PLOGI) is used to indicate that a device always uses a continuously increasing SEQ_CNT, even across transfers of Sequence Initiative. This bit is necessary for interoperability with some devices, and it provides other benefits as well.

In the FC-PH-3 [11], the following is supported:

Word 1, bit 17 - SEQ_CNT (S)
0 = Normal FC-PH rules apply
1 = Continuously increasing SEQ_CNT

Any N_Port that sets Word 1, Bit 17 = 1, is guaranteeing that it will transmit all frames within an Exchange using a continuously increasing SEQ_CNT. Each Exchange SHALL start with SEQ_CNT = 0 in the first frame, and every frame transmitted after that SHALL increment the previous SEQ_CNT by one, even across transfers of Sequence Initiative. Any frames received from the other N_Port in the Exchange shall have no effect on the transmitted SEQ_CNT.

Appendix G: Acronyms and Glossary of FC Terms

It is assumed that the reader is familiar with the terms and acronyms used in the FC protocol specification [2]. The following is provided for easy reference.

First Frame: The frame that contains the SOFi field. This means a logical first and may not necessarily be the first frame temporally received in a sequence.

Code Point: The coded bit pattern associated with control fields in frames or packets.

PDU: Protocol Data Unit

ABTS_LS: Abort Sequence Protocol - Last Sequence. A protocol for aborting an exchange based on the ABTS recipient setting the Last_Sequence bit in the BA_ACC ELS to the ABTS

ADISC: Discover Address. An ELS for discovering the Hard Addresses (the 24 bit NL_Port Identifier) of N_Ports

D_ID: Destination ID

ES: End sequence. This FCTL bit in the FC header indicates this frame is the last frame of the sequence.

FAN: Fabric Address Notification. An ELS sent by the fabric to all known previously Logged in ports following an initialization event.

FLOGI: Fabric Login.

LIP: Loop Initialization. A primitive Sequence used by a port to detect if it is part of a loop or to recover from certain loop errors.

Link: Two unidirectional paths flowing in opposite directions and connecting two Ports within adjacent Nodes.

LOGO: Logout.

LR: Link reset. A primitive sequence transmitted by a port to initiate the link reset protocol or to recover from a link timeout.

LS: Last Sequence of Exchange. This FCTL bit in the FC header indicates the Sequence is the Last Sequence of the Exchange.

Network Address Authority: A 4-bit field specified in Network_Headers that distinguishes between various name registration authorities that may be used to identify the WW_PN and the WW_NN. NAA=b'0001' indicates IEEE-48-bit MAC addresses

Node: A collection of one or more Ports identified by a unique World Wide Node Name (WW_NN).

NOS: Not Operational. A primitive Sequence transmitted to indicate that the port transmitting this Sequence has detected a link failure or is offline, waiting for OLS to be received.

OLS: Off line. A primitive Sequence transmitted to indicate that the port transmitting this Sequence is either initiating the link initialization protocol, receiving and recognizing NOS, or entering the offline state.

PDISC: Discover Port. An ELS for exchanging Service Parameters without affecting Login state.

Primitive Sequence: A primitive Sequence is an Ordered Set that is transmitted repeatedly and continuously.

Private Loop Device: A device that does not attempt Fabric Login (FLOGI) and usually adheres to PLDA. The Area and Domain components of the NL_Port ID must be 0x0000. These devices cannot communicate with any port not in the local loop.

Public Loop Device: A device whose Area and Domain components of the NL_Port ID cannot be 0x0000. Additionally, to be FLA compliant, the device must attempt to open AL_PA 0x00 and attempt FLOGI. These devices communicate with devices on the local loop as well as devices on the other side of a Fabric.

Port: The transmitter, receiver and associated logic at either end of a link within a Node. There may be multiple Ports per Node. Each Port is identified by a unique Port_ID, which is volatile, and a unique World Wide Port Name (WW_PN), which is unchangeable. In this document, the term "port" may be used interchangeably with NL_Port or N_Port.

Port_ID: Fibre Channel ports are addressed by unique 24-bit Port_IDs. In a Fibre Channel frame header, the Port_ID is referred to as S_ID (Source ID) to identify the port originating a frame, and D_ID to identify the destination port. The Port_ID of a given port is volatile (changeable).

PLOGI: Port Login.

SI: Sequence Initiative

World Wide Port_Name (WW_PN): Fibre Channel requires each Port to have an unchangeable WW_PN. Fibre Channel specifies a Network Address Authority (NAA) to distinguish between the various name registration authorities that may be used to identify the WW_PN. A 4-bit NAA identifier, 12-bit field set to 0x0 and an IEEE 48-bit MAC address together make this a 64-bit field.

World Wide Node_Name (WW_NN): Fibre Channel identifies each Node with a unchangeable WW_NN. In a single port Node, the WW_NN and the WW_PN may be identical.

Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.