Network Working Group
Request for Comments: 4338
Obsoletes: 3831, 2625
Category: Standards Track

C. DeSanti
Cisco Systems
C. Carlson
QLogic Corporation
R. Nixon
Emulex

January 2006

Transmission of IPv6, IPv4, and Address Resolution Protocol (ARP) Packets over Fibre Channel

Status of This Memo

Copyright Notice

Abstract

This document specifies the way of encapsulating IPv6, IPv4, and Address Resolution Protocol (ARP) packets over Fibre Channel. This document also specifies the method of forming IPv6 link-local addresses and statelessly autoconfigured IPv6 addresses on Fibre Channel networks, and a mechanism to perform IPv4 address resolution over Fibre Channel networks.

This document obsoletes RFC 2625 and RFC 3831.

Table of Contents

1. Introduction

Fibre Channel (FC) is a high-speed serial interface technology that supports several Upper Layer Protocols including Small Computer System Interface (SCSI), IPv6 [IPv6], and IPv4 [IPv4].

[RFC-2625] defined how to encapsulate IPv4 and Address Resolution Protocol (ARP) packets over Fibre Channel for a subset of Fibre Channel devices. This specification enables the support of IPv4 for a broader category of Fibre Channel devices. In addition, this specification simplifies [RFC-2625] by removing unused options and clarifying current implementations. This document obsoletes [RFC-2625].

Specific [RFC-2625] limitations that this document aims to resolve are the following:

- N_Port_Name format restriction. [RFC-2625] restricts the use of IPv4 to Fibre Channel devices having the format 0x1 N_Port_Name, but many current implementations use other N_Port_Name formats.

- Use of Fibre Channel Address Resolution Protocol (FARP). [RFC-2625] requires the support of FARP to map N_Port_Names to N_Port_IDs, but many current implementations use other methods, such as the Fibre Channel Name Server.

- Missing support for IPv4 multicast. [RFC-2625] does not specify how to transmit IPv4 packets with a multicast destination address over Fibre Channel.

[RFC-3831] defines how to encapsulate IPv6 over Fibre Channel and a method of forming IPv6 link-local addresses [AARCH] and statelessly autoconfigured IPv6 addresses on Fibre Channel networks. [RFC-3831] also describes the content of the Source/Target Link-layer Address option used in Neighbor Discovery [DISC] when the messages are transmitted on a Fibre Channel network. This document obsoletes [RFC-3831].

Warning to readers familiar with Fibre Channel: both Fibre Channel and IETF standards use the same byte transmission order. However, the bit numbering is different. See Appendix C for guidance.

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 [KEYWORDS].

2. Summary of Fibre Channel

2.1. Overview

Fibre Channel (FC) is a gigabit-speed network technology primarily used for storage networking. Fibre Channel is standardized in the T11 Technical Committee of the InterNational Committee for Information Technology Standards (INCITS), an American National Standard Institute (ANSI) accredited standards committee.

Fibre Channel devices are called Nodes. Each Node has one or more Ports that connect to Ports of other devices. Fibre Channel may be implemented using any combination of the following three topologies:

- a point-to-point link between two Ports;
- a set of Ports interconnected by a switching network called a Fabric, as defined in [FC-FS];
- a set of Ports interconnected with a loop topology, as defined in [FC-AL-2].

A Node Port that does not operate in a loop topology is called an N_Port. A Node Port that operates in a loop topology using the loop-specific protocols is designated as an NL_Port. The term Nx_Port is used to indicate a Node Port that is capable of operating in either mode.

A Fabric Port that does not operate in a loop topology is called an F_Port. A Fabric Port that operates in a loop topology using the loop-specific protocols is designated as an FL_Port. The term Fx_Port is used to indicate a Fabric Port that is capable of operating in either mode.

A Fibre Channel network, built with any combination of the FC topologies described above, is a multiaccess network with broadcast capabilities.

From an IPv6 point of view, a Fibre Channel network is an IPv6 Link [IPv6]. IP-capable Nx_Ports are what [IPv6] calls Interfaces.

From an IPv4 point of view, a Fibre Channel network is an IPv4 Local Network [IPv4]. IP-capable Nx_Ports are what [IPv4] calls Local Network Interfaces.

2.2. Identifiers and Login

Fibre Channel entities are identified by non-volatile 64-bit Name_Identifiers. [FC-FS] defines several formats of Name_Identifiers. The value of the most significant 4 bits defines the format of a Name_Identifier. These Name_Identifiers are referred to in a more concise manner as follows:

- an Nx_Port's Name_Identifier is called N_Port_Name;
- an Fx_Port's Name_Identifier is called F_Port_Name;
- a Node's Name_Identifier is called Node_Name;
- a Fabric's Name_Identifier is called Fabric_Name.

An Nx_Port connected to a Fibre Channel network is associated with two identifiers, its non-volatile N_Port_Name and a volatile 24-bit address called N_Port_ID. The N_Port_Name is used to identify the Nx_Port, and the N_Port_ID is used for communications among Nx_Ports.

Each Nx_Port acquires an N_Port_ID from the Fabric by performing a process called Fabric Login, or FLOGI. The FLOGI process is used also to negotiate several communications parameters between the Nx_Port and the Fabric, such as the receive data field size, which determines the maximum size of the Fibre Channel frames that may be transferred between the Nx_Port and the Fabric.

Before effective communication may take place between two Nx_Ports, they must complete a process called Port Login, or PLOGI. The PLOGI process provides each Nx_Port with the other Nx_Port's N_Port_Name, and negotiates several communication parameters, such as the receive data field size, which determines the maximum size of the Fibre Channel frames that may be transferred between the two Nx_Ports.

Both Fabric Login and Port Login may be explicit (i.e., performed using specific FC control messages called Extended Link Services, or ELSes) or implicit (i.e., in which the parameters are specified by configuration or other methods).

2.3. FC Levels and Frame Format

[FC-FS] describes the Fibre Channel protocol using 5 different levels. The FC-2 and FC-4 levels are relevant for this specification. The FC-2 level defines the FC frame format, the transport services, and the control functions necessary for information transfer. The FC-4 level supports Upper Level Protocols, such as IPv6, IPv4, and SCSI. The Fibre Channel frame format is shown in figure 1.

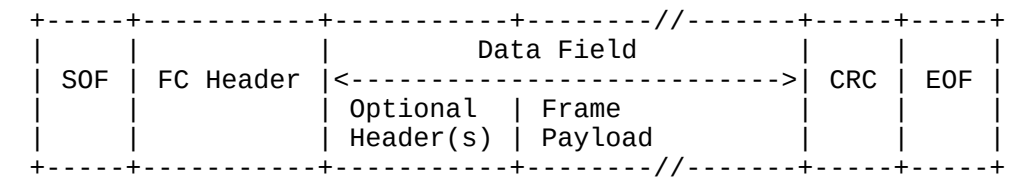

Figure 1: Fibre Channel Frame Format

The Start of Frame (SOF) and End of Frame (EOF) are special FC transmission words that act as frame delimiters. The Cyclic Redundancy Check (CRC) is 4 octets long and is used to verify the integrity of a frame.

The FC Header is 24 octets long and contains several fields associated with the identification and control of the Data Field.

The Data Field is of variable size, ranging from 0 to 2112 octets, and includes the user data in the Frame Payload field and Optional Headers. The currently defined Optional Headers are the following:

- ESP_Header;
- Network_Header;
- Association_Header;
- Device_Header.

The value of the SOF field determines the FC Class of service associated with the frame. Five Classes of service are specified in [FC-FS]. They are distinguished primarily by the method of flow control between the communicating Nx_Ports and by the level of data integrity provided. A given Fabric or Nx_Port may support one or more of the following Classes of service:

- Class 1: Dedicated physical connection with delivery confirmation;
- Class 2: Frame multiplexed service with delivery confirmation;
- Class 3: Datagram service;
- Class 4: Fractional bandwidth;
- Class 6: Reliable multicast via dedicated connections.

Classes 3 and 2 are commonly used for storage networking applications; Classes 1 and 6 are typically used for specialized applications in avionics. Class 3 is recommended for IPv6, IPv4, and ARP (see section 4.3).

2.4. Sequences and Exchanges

An application-level payload such as an IPv6 or IPv4 packet is called an Information Unit at the FC-4 level of Fibre Channel. Each FC-4

Information Unit is mapped to an FC Sequence by the FC-2 level. An FC Sequence consists of one or more FC frames related by the value of the Sequence_ID (SEQ_ID) field of the FC Header.

The architectural maximum data that may be carried by an FC frame is 2112 octets. The maximum usable frame size depends on the Fabric and Nx_Port implementations and is negotiated during the Login process. Whenever an Information Unit to be transmitted exceeds this value, the FC-2 level segments it into multiple FC frames, sent as a single Sequence. The receiving Nx_Port reassembles the Sequence of frames and delivers a reassembled Information Unit to the FC-4 level. The Sequence Count (SEQ_CNT) field of the FC Header may be used to ensure frame ordering.

Multiple Sequences may be grouped together as belonging to the same FC Exchange. The Exchange is a mechanism used by two Nx_Ports to identify and manage an operation between them. The Exchange is opened when the operation is started between the two Nx_Ports, and closed when the operation ends. FC frames belonging to the same Exchange are related by the value of the Exchange_ID fields in the FC Header. An Originator Exchange_ID (OX_ID) and a Responder Exchange_ID (RX_ID) uniquely identify the Exchange between a pair of Nx_Ports.

3. IP-capable Nx_Ports

This specification requires an IP-capable Nx_Port to have the following properties:

- The format of its N_Port_Name MUST be one of 0x1, 0x2, 0x5, 0xC, 0xD, 0xE, 0xF (see section 5.1);
- It MUST support Class 3;
- It MUST support continuously increasing SEQ_CNT [FC-FS];
- It MUST be able to transmit and receive an FC-4 Information Unit at least 1304 octets long (see section 4.1);
- It SHOULD support a receive data field size for Device_Data FC frames of at least 1024 octets (see section 10).

4. IPv6, IPv4, and ARP Encapsulation

4.1. FC Sequence Format for IPv6 and IPv4 Packets

An IPv6 or IPv4 packet is mapped to an Information Unit at the FC-4 level of Fibre Channel, which in turn is mapped to an FC Sequence by the FC-2 level [FC-FS]. An FC Information Unit containing an IP packet MUST carry the FC Network_Header [FC-FS] and the Logical Link Control/SubNetwork Access Protocol (LLC/SNAP) header [IEEE-LLC], resulting in the FC Information Unit format shown in figure 2.

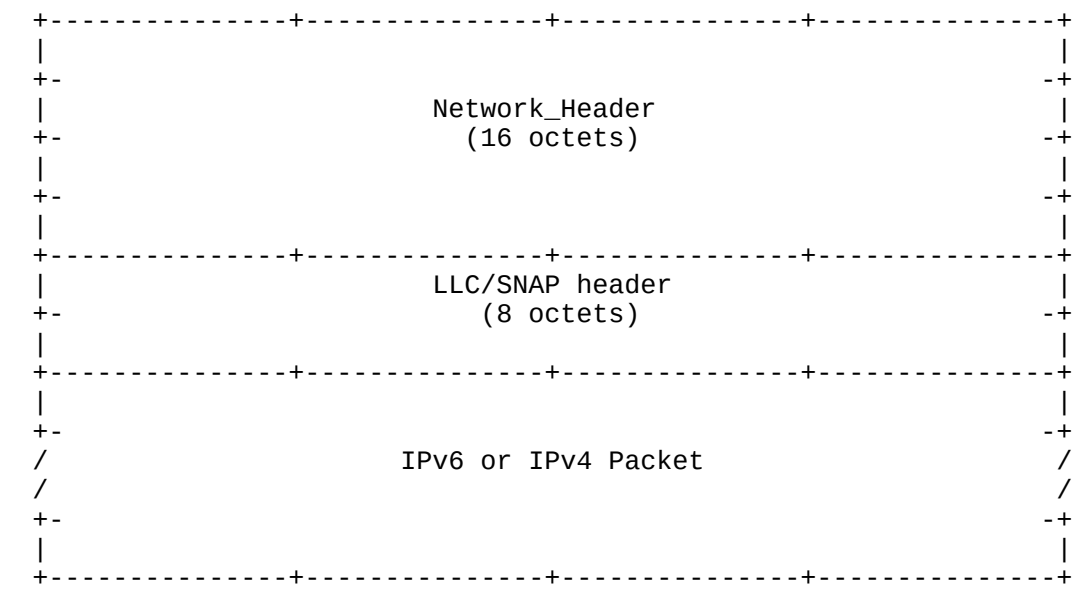

Figure 2: FC Information Unit Mapping an IP Packet

In order to support the minimum IPv6 MTU (i.e., 1280 octets), an Nx_Port supporting IP MUST be able to transmit and receive an FC-4 Information Unit at least 1304 octets long (i.e., 1280 + 8 + 16).

The FC ESP_Header [FC-FS] MAY be used to secure the FC frames composing an IP FC Sequence. Other FC Optional Headers MUST NOT be used in an IP FC Sequence.

An IP FC Sequence often consists of more than one frame, all frames having the same TYPE (see section 4.4). The first frame of the Sequence MUST include the FC Network_Header and the LLC/SNAP header. The other frames MUST NOT include them, as shown in figure 3.

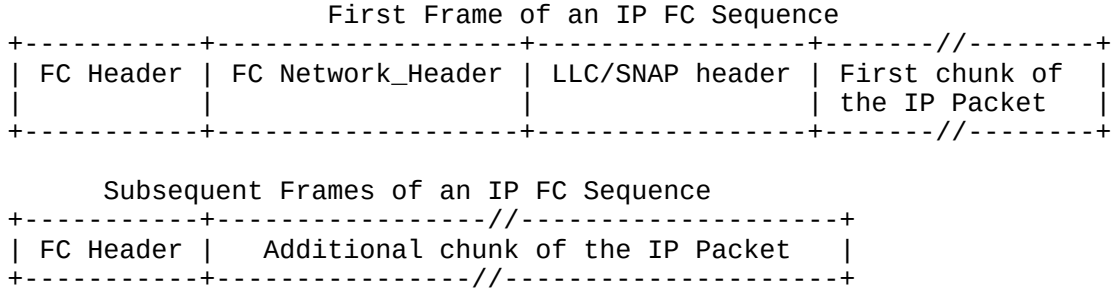

Figure 3: Optional Headers in an IP FC Sequence

4.2. FC Sequence Format for ARP Packets

An ARP packet is mapped to an Information Unit at the FC-4 level of Fibre Channel, which in turn is mapped to an FC Sequence by the FC-2 level. An FC Information Unit containing an ARP packet MUST carry the FC Network_Header [FC-FS] and the LLC/SNAP header [IEEE-LLC], resulting in the FC Information Unit format shown in figure 4.

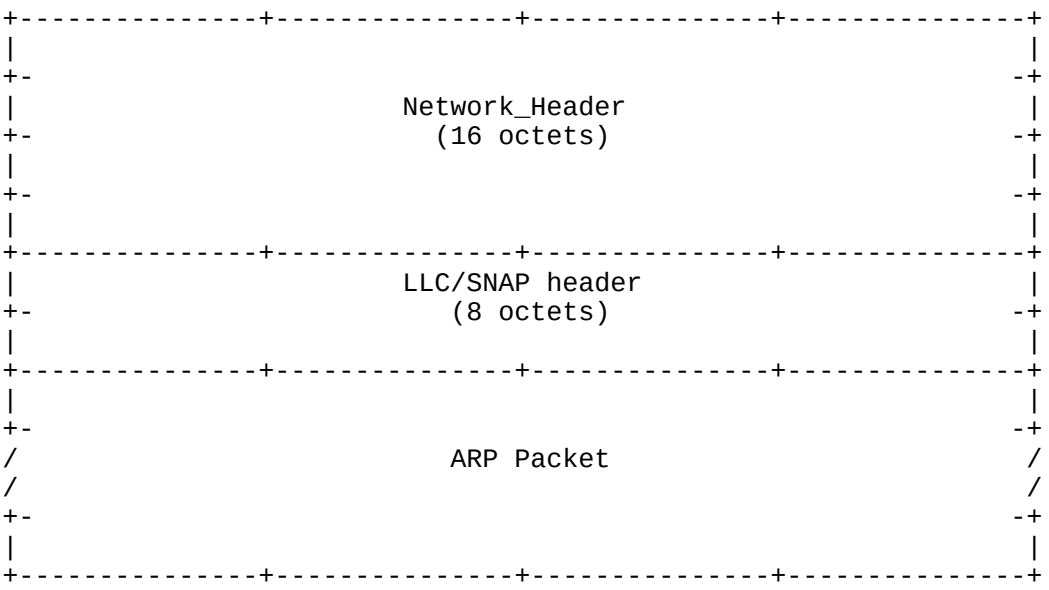

Figure 4: FC Information Unit Mapping an ARP Packet

Given the limited size of an ARP packet (see section 7), an FC Sequence carrying an ARP packet MUST be mapped to a single FC frame that MUST include the FC Network_Header and the LLC/SNAP header.

The FC ESP_Header [FC-FS] MAY be used to secure an FC frame carrying an ARP packet. Other FC Optional Headers MUST NOT be used in an FC frame carrying an ARP packet.

4.3. FC Classes of Service

This specification uses FC Class 3. The following types of packets MUST be mapped in Class 3 FC frames:

- multicast IPv6 packets;
- multicast/broadcast IPv4 packets;
- Control Protocol packets (e.g., ARP packets; IPv6 packets carrying ICMPv6 [ICMPv6], Neighbor Discovery [DISC], or Multicast Listener Discovery [MLDv2] messages; IPv4 packets carrying ICMP [ICMPv4] or IGMP [IGMPv3] messages; IPv6 and IPv4 Routing Protocols packets).

Other IPv6 and IPv4 packets (i.e., unicast IP packets carrying data traffic) SHOULD be mapped in Class 3 FC frames as well. Support for reception of IPv4 or IPv6 packets mapped in FC frames of any Class other than Class 3 is OPTIONAL; receivers MAY ignore them.

4.4. FC Header Code Points

The fields of the Fibre Channel Header are shown in figure 5. The D_ID and S_ID fields contain, respectively, the destination N_Port_ID and the source N_Port_ID.

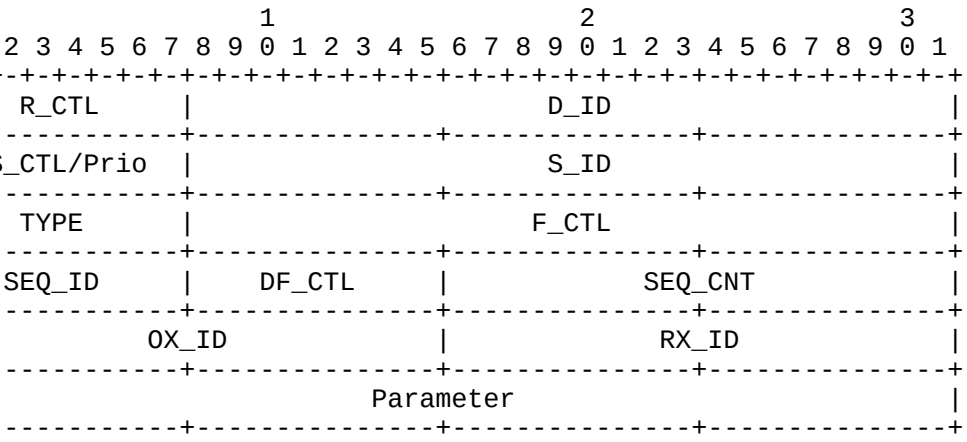

Figure 5: FC Header Format

To encapsulate IPv6 and IPv4 over Fibre Channel, the following code points apply. When a single value is listed without further qualification, that value MUST be used:

- R_CTL: 0x04 (Device_Data frame with Unsolicited Data Information Category [FC-FS]);
- TYPE: 0x05 (IP over Fibre Channel);

- CS_CTL/Prio: 0x00 is the default, see [FC-FS] for other values;
- DF_CTL: 0x20 (Network_Header) for the first FC frame of an IPv6 or IPv4 Sequence, 0x00 for the following FC frames. If the FC ESP_Header is used, then 0x60 for the first FC frame of an IPv6 or IPv4 Sequence, 0x40 for the following FC frames;
- F_CTL, SEQ_ID, SEQ_CNT, OX_ID, RX_ID: see section 11, section 12, and [FC-FS] for additional requirements;
- Parameter: if Relative Offset [FC-FS] is not used, the content of this field MUST be ignored by the receiver, and SHOULD be set to zero by the sender. If Relative Offset is used, see [FC-FS].

To encapsulate ARP over Fibre Channel, the following code points apply. When a single value is listed without further qualification, that value MUST be used:

- R_CTL: 0x04 (Device_Data frame with Unsolicited Data Information Category [FC-FS]);
- TYPE: 0x05 (IP over Fibre Channel);
- CS_CTL/Prio: 0x00 is the default, see [FC-FS] for other values;
- DF_CTL: 0x20 (Network_Header). If the FC ESP_Header is used, then 0x60;
- F_CTL, SEQ_ID, SEQ_CNT, OX_ID, RX_ID: see section 11, section 12, and [FC-FS] for additional requirements;
- Parameter: SHOULD be set to zero.

4.5. FC Network_Header

The fields of the FC Network_Header are shown in figure 6. For use with IPv6, IPv4, and ARP, the N_Port_Names formats MUST be one of 0x1, 0x2, 0x5, 0xC, 0xD, 0xE, 0xF [FC-FS].

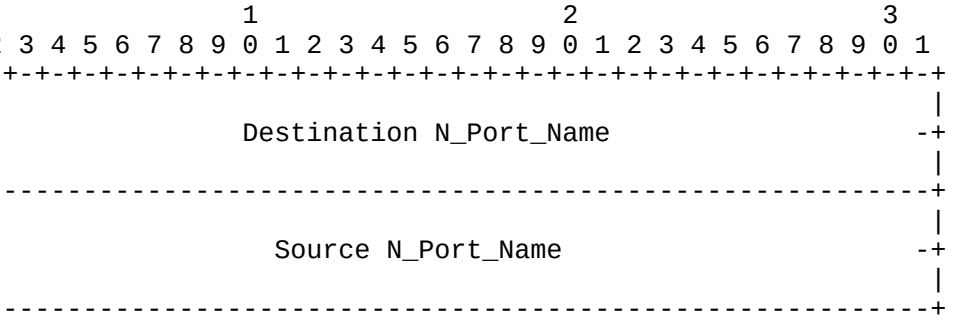

Figure 6: FC Network_Header Format

4.6. LLC/SNAP Header

The fields of the LLC/SNAP header [IEEE-LLC] are shown in figure 7.

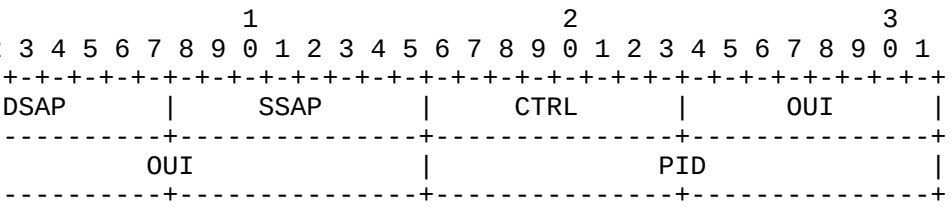

Figure 7: LLC/SNAP Header Format

To encapsulate IPv6, IPv4, and ARP over Fibre Channel, the following code points MUST be used:

- DSAP: 0xAA;
- SSAP: 0xAA;
- CTRL: 0x03;
- OUI: 0x000000;
- PID: 0x86DD for IPv6, 0x0800 for IPv4, 0x0806 for ARP.

4.7. Bit and Byte Ordering

IPv6, IPv4, and ARP packets are mapped to the FC-4 level using the big-endian byte ordering that corresponds to the standard network byte order or canonical form.

4.8. Maximum Transfer Unit

The default MTU size for IPv6 packets over Fibre Channel is 65280 octets. Large IPv6 packets are mapped to a Sequence of FC frames (see section 2.4). This size may be reduced by a Router Advertisement [DISC] containing an MTU option that specifies a smaller MTU, or by manual configuration of each Nx_Port. However, as required by [IPv6], the MTU MUST NOT be lower than 1280 octets. If a Router Advertisement received on an Nx_Port has an MTU option specifying an MTU larger than 65280, or larger than a manually configured value, that MTU option MAY be logged to system management but MUST be otherwise ignored.

As the default MTU size far exceeds the message sizes typically used in the Internet, an IPv6 over FC implementation SHOULD implement Path MTU Discovery [PMTUD6], or at least maintain different MTU values for on-link and off-link destinations.

For correct operation of IPv6 in a routed environment, it is critically important to configure an appropriate MTU option in Router Advertisements.

For correct operation of IPv6 when mixed media (e.g., Ethernet and Fibre Channel) are bridged together, the smallest MTU of all the media must be advertised by routers in an MTU option. If there are no routers present, this MTU must be manually configured in each node that is connected to a medium with a default MTU larger than the smallest MTU.

The default MTU size for IPv4 packets over Fibre Channel is 65280 octets. Large IPv4 packets are mapped to a Sequence of FC frames (see section 2.4). This size may be reduced by manual configuration of each Nx_Port or by the Path MTU Discovery technique [PMTUD4].

5. IPv6 Stateless Address Autoconfiguration

5.1. IPv6 Interface Identifier and Address Prefix

The IPv6 Interface ID [AARCH] for an Nx_Port is based on the EUI-64 address [EUI64] derived from the Nx_Port's N_Port_Name. The IPv6 Interface Identifier is obtained by complementing the Universal/Local (U/L) bit of the OUI field of the derived EUI-64 address. The U/L bit has no function in Fibre Channel; however, it has to be properly handled when a Name_Identifier is converted to an EUI-64 address.

[FC-FS] specifies a method to map format 0x1 (IEEE 48-bit address), 0x2 (IEEE Extended), or 0x5 (IEEE Registered) FC Name_Identifiers in EUI-64 addresses. This allows the usage of these Name_Identifiers to support IPv6. [FC-FS] also defines EUI-64 mapped FC Name_Identifiers (formats 0xC, 0xD, 0xE, and 0xF) that are derived from an EUI-64 address. It is possible to reverse this address mapping to obtain the original EUI-64 address in order to support IPv6.

IPv6 stateless address autoconfiguration MUST be performed as specified in [ACONF]. An IPv6 Address Prefix used for stateless address autoconfiguration of an Nx_Port MUST have a length of 64 bits.

5.2. Generating an Interface ID from a Format 1 N_Port_Name

The Name_Identifier format 0x1 is shown in figure 8.

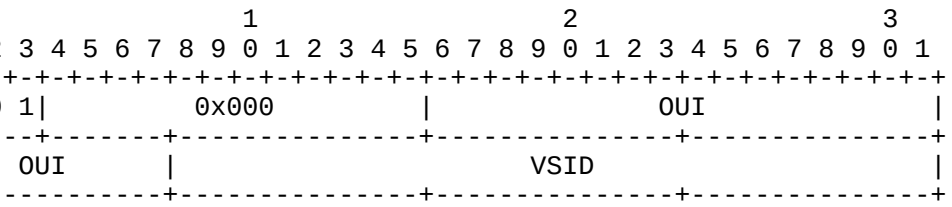

Figure 8: Format 0x1 Name_Identifier

The EUI-64 address derived from this Name_Identifier has the format shown in figure 9 [FC-FS].

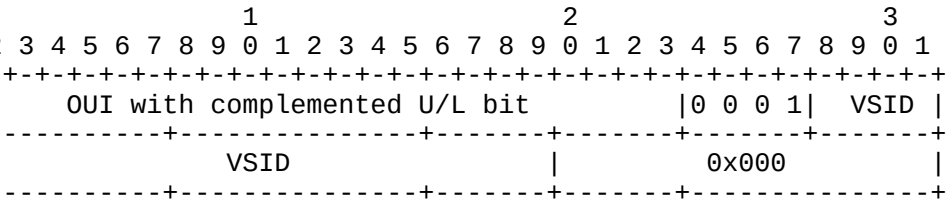

Figure 9: EUI-64 Address from a Format 0x1 Name_Identifier

The IPv6 Interface Identifier is obtained from this EUI-64 address by complementing the U/L bit in the OUI field. Therefore, the OUI in the IPv6 Interface ID is exactly as in the FC Name_Identifier. The resulting IPv6 Interface Identifier has local scope [AARCH] and the format shown in figure 10.

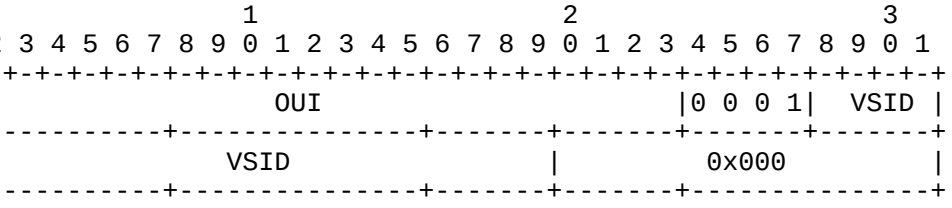

Figure 10: IPv6 Interface ID from a Format 0x1 Name_Identifier

As an example, the FC Name_Identifier 0x10-00-34-63-46-AB-CD-EF generates the IPv6 Interface Identifier 3463:461A:BCDE:F000.

5.3. Generating an Interface ID from a Format 2 N_Port_Name

The Name_Identifier format 0x2 is shown in figure 11.

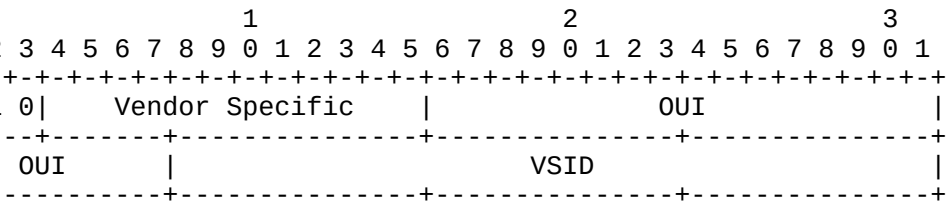

Figure 11: Format 0x2 Name_Identifier

The EUI-64 address derived from this Name_Identifier has the format shown in figure 12 [FC-FS].

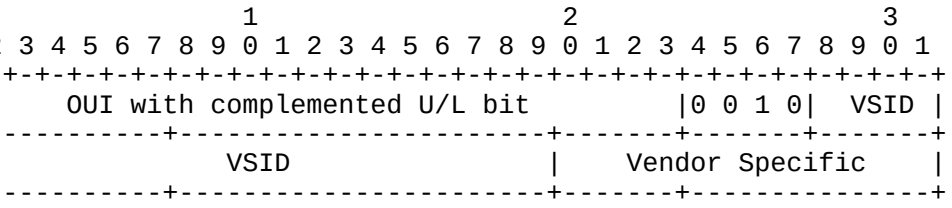

Figure 12: EUI-64 Address from a Format 0x2 Name_Identifier

The IPv6 Interface Identifier is obtained from this EUI-64 address by complementing the U/L bit in the OUI field. Therefore, the OUI in the IPv6 Interface ID is exactly as in the FC Name_Identifier. The resulting IPv6 Interface Identifier has local scope [AARCH] and the format shown in figure 13.

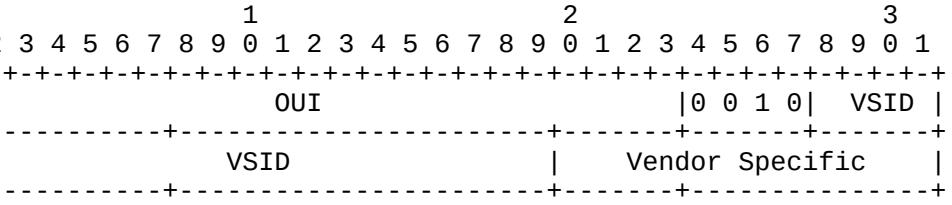

Figure 13: IPv6 Interface ID from a Format 0x2 Name_Identifier

As an example, the FC Name_Identifier 0x27-89-34-63-46-AB-CD-EF generates the IPv6 Interface Identifier 3463:462A:BCDE:F789.

5.4. Generating an Interface ID from a Format 5 N_Port_Name

The Name_Identifier format 0x5 is shown in figure 14.

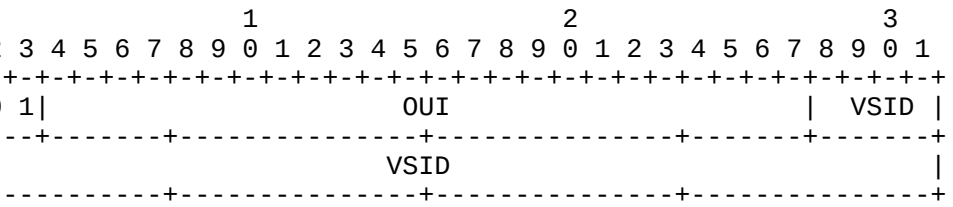

Figure 14: Format 0x5 Name_Identifier

The EUI-64 address derived from this Name_Identifier has the format shown in figure 15 [FC-FS].

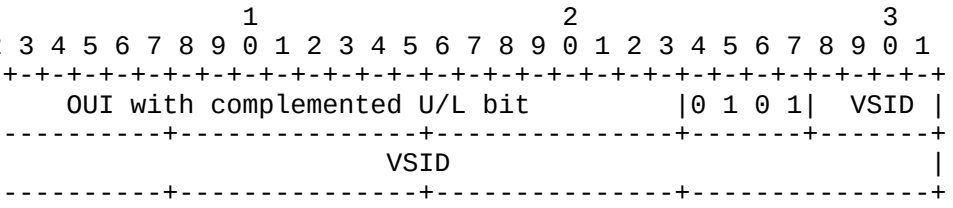

Figure 15: EUI-64 Address from a Format 0x5 Name_Identifier

The IPv6 Interface Identifier is obtained from this EUI-64 address complementing the U/L bit in the OUI field. Therefore, the OUI in the IPv6 Interface ID is exactly as in the FC Name_Identifier. The resulting IPv6 Interface Identifier has local scope [AARCH] and the format shown in figure 16.

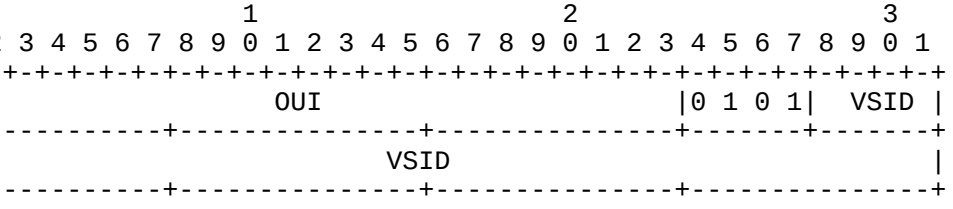

Figure 16: IPv6 Interface ID from a Format 0x5 Name_Identifier

As an example, the FC Name_Identifier 0x53-46-34-6A-BC-DE-F7-89 generates the IPv6 Interface Identifier 3463:465A:BCDE:F789.

5.5. Generating an Interface ID from an EUI-64 Mapped N_Port_Name

The EUI-64 mapped Name_Identifiers formats (formats 0xC through 0xF) are derived from an EUI-64 address by compressing the OUI field of such addresses. The compression is performed by removing the Universal/Local and Individual/Group bits from the OUI, and by putting bits 0 to 5 of the OUI in the first octet of the Name_Identifier, and bits 8 to 23 of the OUI in the second and third octet of the Name_Identifier, as shown in figure 17.

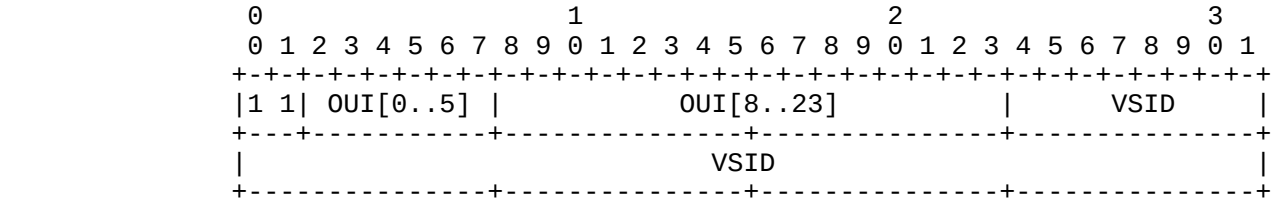

Figure 17: EUI-64 Mapped Name_Identifiers Format

The EUI-64 address used to generate the Name_Identifier shown in figure 17 has the format shown in figure 18.

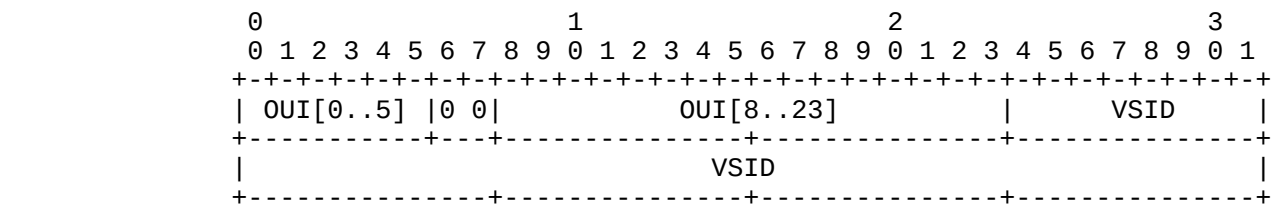

Figure 18: EUI-64 Address from an EUI-64 Mapped Name_Identifier

The IPv6 Interface Identifier is obtained from this EUI-64 address by complementing the U/L bit in the OUI field. The resulting IPv6 Interface Identifier has global scope [AARCH] and the format shown in figure 19.

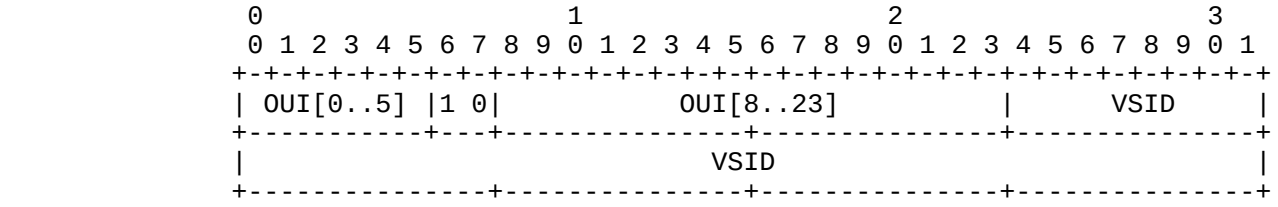

Figure 19: IPv6 Interface ID from an EUI-64 Mapped Name_Identifier

As an example, the FC Name_Identifier 0xCD-63-46-AB-01-25-78-9A generates the IPv6 Interface Identifier 3663:46AB:0125:789A.

6. Link-local Addresses

The IPv6 link-local address [AARCH] for an Nx_Port is formed by appending the Interface Identifier (as defined in section 5) to the prefix FE80::/64. The resulting address is shown in figure 20.

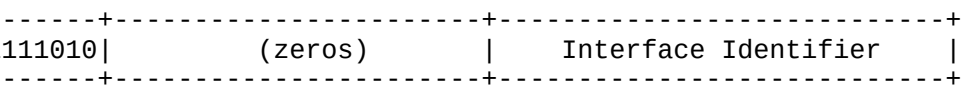

Figure 20: IPv6 Link-local Address Format

7. ARP Packet Format

The Address Resolution Protocol defined in [ARP] is designed to be a general purpose protocol, to accommodate many network technologies and many Upper Layer Protocols.

[RFC-2625] chose to use for Fibre Channel the same ARP packet format used for Ethernet networks. In order to do that, [RFC-2625] restricted the use of IPv4 to Nx_Ports having N_Port_Name format 0x1. Although this may have been a reasonable choice at that time, today there are Nx_Ports with an N_Port_Name format other than 0x1 in widespread use.

This specification accommodates Nx_Ports with N_Port_Names of a format different from 0x1 by defining a Fibre Channel specific version of the ARP protocol (FC ARP), carrying both N_Port_Name and N_Port_ID as Hardware (HW) Address.

IANA has registered the number 18 (decimal) to identify Fibre Channel as ARP HW type. The FC ARP packet format is shown in figure 21. The length of the FC ARP packet is 40 octets.

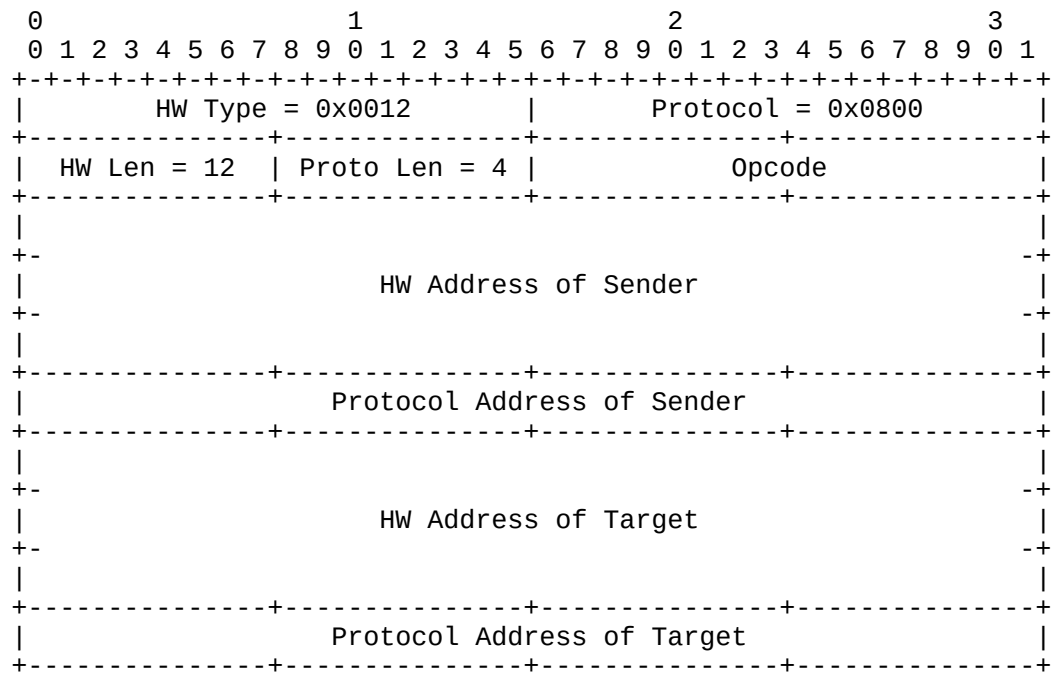

Figure 21: FC ARP Packet Format

The following code points MUST be used with FC ARP:

- HW Type: 0x0012 (Fibre Channel);
- Protocol: 0x0800 (IPv4);
- HW Len: 12 (Length in octets of the HW Address);
- Proto Len: 4 (Length in octets of the Protocol Address);
- Opcode: 0x0001 for ARP Request, 0x0002 for ARP Reply [ARP];
- HW Address of Sender: the HW Address (see section 8) of the Requester in an ARP Request, or the HW Address of the Responder in an ARP Reply;
- Protocol Address of Sender: the IPv4 address of the Requester in an ARP Request, or that of the Responder in an ARP Reply;
- HW Address of Target: set to zero in an ARP Request, and to the HW Address (see section 8) of the Requester in an ARP Reply;
- Protocol Address of Target: the IPv4 address of the Responder in an ARP Request, or that of the Requester in an ARP Reply.

8. Link-layer Address/Hardware Address

The Link-layer Address used in the Source/Target Link-layer Address option (see section 9.2) and the Hardware Address used in FC ARP (see section 7) have the same format, shown in figure 22.

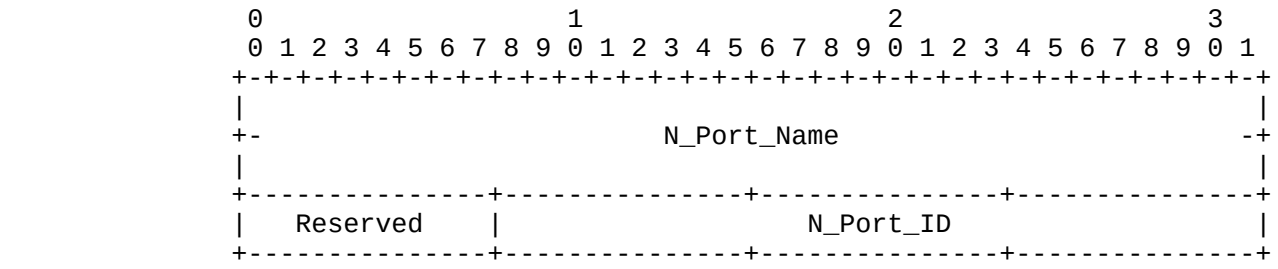

Figure 22: Link-layer Address/HW Address Format

Reserved fields MUST be set to zero when transmitting, and MUST be ignored when receiving.

9. Address Mapping for Unicast

9.1. Overview

An Nx_Port has two kinds of Fibre Channel addresses:

- a non-volatile 64-bit address, called N_Port_Name;
- a volatile 24-bit address, called N_Port_ID.

The N_Port_Name is used to uniquely identify the Nx_Port, and the N_Port_ID is used to route frames to the Nx_Port. Both FC addresses are required to resolve an IPv6 or IPv4 unicast address. The fact that the N_Port_ID is volatile implies that an Nx_Port MUST validate the mapping between its N_Port_Name and N_Port_ID when certain Fibre Channel events occur (see Appendix B).

9.2. IPv6 Address Mapping

The procedure for mapping IPv6 unicast addresses into Fibre Channel link-layer addresses uses the Neighbor Discovery Protocol [DISC]. The Source/Target Link-layer Address option has the format shown in figure 23 when the link layer is Fibre Channel.

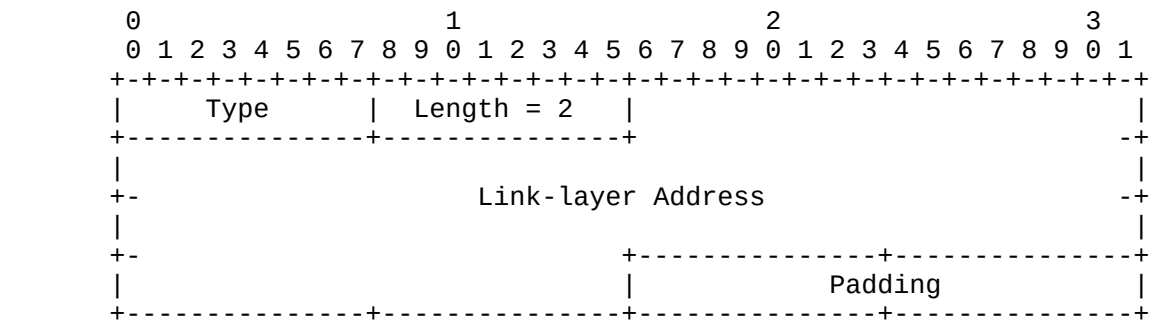

Figure 23: Source/Target Link-layer Address Option for Fibre Channel

Type: 1 for Source Link-layer address.
2 for Target Link-layer address.

Length: 2 (in units of 8 octets).

Padding: MUST be set to zero when transmitting,
MUST be ignored when receiving.

Link-layer Address: the Nx_Port's Link-layer Address (see section 8).

9.3. IPv4 Address Mapping

The procedure for mapping IPv4 unicast addresses into Fibre Channel link-layer addresses uses the FC ARP protocol, as specified in section 7 and [ARP]. A source Nx_Port that has to send IPv4 packets to a destination Nx_Port, known by its IPv4 address, MUST perform the following steps:

1) The source Nx_Port first consults its local mapping tables for a mapping <destination IPv4 address, N_Port_Name, N_Port_ID>.

2) If such a mapping is found, and a valid Port Login is in place with the destination Nx_Port, then the source Nx_Port sends the IPv4 packets to the destination Nx_Port using the retrieved N_Port_ID as D_ID.

3) If such a mapping is not found, or a valid Port Login is not in place with the destination Nx_Port, then the source Nx_Port sends a broadcast FC ARP Request (see section 10) to its connected FC network.

4) When a broadcast FC ARP Request is received by the Nx_Port with the matching IPv4 address, that Nx_Port caches the information carried in the FC ARP Request in its local mapping tables and generates a unicast FC ARP Reply. If a valid Port Login to the Nx_Port that sent the broadcast FC ARP Request does not exist, the Nx_Port MUST perform such a Port Login, and then use it for the unicast reply. The N_Port_ID to which the Port Login is directed is taken from the N_Port_ID field of the Sender HW Address field in the received FC ARP packet.

5) If no Nx_Port has the matching IPv4 address, no unicast FC ARP Reply is returned.

10. Address Mapping for Multicast

IPv6 multicast packets, IPv4 multicast/broadcast packets, and ARP broadcast packets MUST be mapped to FC Sequences addressed to the broadcast N_Port_ID 0xFFFFFF, sent in FC Class 3 in a unidirectional Exchange (see section 12). Appendix A specifies how to transmit a Class 3 broadcast FC Sequence over various Fibre Channel topologies. The Destination N_Port_Name field of the FC Network_Header MUST be set to the value:

- for broadcast ARP and IPv4 packets: 0x10-00-FF-FF-FF-FF-FF-FF;
- for multicast IPv6 packets: 0x10-00-33-33-XX-YY-ZZ-QQ, where XX-YY-ZZ-QQ are the 4 least significant octets of the multicast destination IPv6 address;
- for multicast IPv4 packets: 0x10-00-01-00-5E-XX-YY-ZZ, where the 23 least significant bits of XX-YY-ZZ are the 23 least significant bits of the multicast destination IPv4 address and the most significant bit of XX-YY-ZZ is set to zero.

An Nx_Port supporting IPv6 or IPv4 MUST be able to map a received broadcast Class 3 Device_Data FC frame to an implicit Port Login context in order to handle IPv6 multicast packets, IPv4 multicast or broadcast packets, and ARP broadcast packets. The receive data field size of this implicit Port Login MUST be the same across all the Nx_Ports connected to the same Fabric, otherwise FC broadcast transmission does not work. In order to reduce the need for FC Sequence segmentation, the receive data field size of this implicit Port Login SHOULD be 1024 octets. This receive data field size requirement applies to broadcast Device_Data FC frames, not to ELSes.

Receiving an FC Sequence carrying an IPv6 multicast packet, an IPv4 multicast/broadcast packet, or an FC ARP broadcast packet triggers some additional processing by the Nx_Port when that IPv6, IPv4, or FC ARP packet requires a unicast reply. In this case, if a valid Port Login to the Nx_Port that sent the multicast or broadcast packet

does not exist, the Nx_Port MUST perform such a Port Login, and then use it for the unicast reply. In the case of Neighbor Discovery messages [DISC], the N_Port_ID to which the Port Login is directed is taken from the N_Port_ID field of the Source Link-layer Address in the received Neighbor Discovery message. In the case of FC ARP messages, the N_Port_ID to which the Port Login is directed is taken from the N_Port_ID field of the Sender HW Address field in the received FC ARP packet.

As an example, if a received broadcast FC Sequence carries an IPv6 multicast unsolicited Router Advertisement [DISC], the receiving Nx_Port processes it simply by passing the carried IPv6 packet to the IPv6 layer. Instead, if a received broadcast FC Sequence carries an IPv6 multicast solicitation message [DISC] requiring a unicast reply, and no valid Port Login exists with the Nx_Port sender of the multicast packet, then a Port Login MUST be performed in order to send the unicast reply message. If a received broadcast FC Sequence carries an IPv6 multicast solicitation message [DISC] requiring a multicast reply, the reply is sent to the broadcast N_Port_ID 0xFFFFFF.

11. Sequence Management

FC Sequences carrying IPv6, IPv4, or ARP packets are REQUIRED to be non-streamed [FC-FS]. In order to avoid missing FC frame aliasing by Sequence_ID reuse, an Nx_Port supporting IPv6 or IPv4 is REQUIRED to use continuously increasing SEQ_CNT [FC-FS]. Each Exchange MUST start by setting SEQ_CNT to zero in the first frame; every frame transmitted after that MUST increment the previous SEQ_CNT by one. The Continue Sequence Condition field in the F_CTL field of the FC Header MUST be set to zero [FC-FS].

12. Exchange Management

To transmit IPv6, IPv4, or ARP packets to another Nx_Port or to a multicast/broadcast address, an Nx_Port MUST use dedicated unidirectional Exchanges (i.e., Exchanges dedicated to IPv6, IPv4, or ARP packet transmission and that do not transfer Sequence Initiative). As such, the Sequence Initiative bit in the F_CTL field of the FC Header MUST be set to zero [FC-FS]. The RX_ID field of the FC Header MUST be set to 0xFFFF.

Unicast FC Sequences carrying unicast Control Protocol packets (e.g., ARP packets; IPv6 packets carrying ICMPv6 [ICMPv6], Neighbor Discovery [DISC], or Multicast Listener Discovery [MLDv2] messages; IPv4 packets carrying ICMP [ICMPv4] or IGMP [IGMPv3] messages) SHOULD be sent in short-lived unidirectional Exchanges (i.e., Exchanges containing only one Sequence, in which both the First_Sequence and

Last_Sequence bits in the F_CTL field of the FC Header are set to one [FC-FS]). Unicast FC Sequences carrying other IPv6 and IPv4 packets (i.e., unicast IP packets carrying data traffic) MUST be sent in a long-lived unidirectional Exchange (i.e., an Exchange containing one or more Sequences). IP multicast packets MUST NOT be carried in unicast FC Sequences (see section 10).

Broadcast FC Sequences carrying multicast or broadcast Control Protocol packets (e.g., ARP packets; IPv6 packets carrying ICMPv6 [ICMPv6], Neighbor Discovery [DISC], or Multicast Listener Discovery [MLDv2] messages; IPv4 packets carrying ICMP [ICMPv4] or IGMP [IGMPv3] messages) MUST be sent in short-lived unidirectional Exchanges. Broadcast FC Sequences carrying other IPv6 or IPv4 multicast traffic (i.e., multicast IP packets carrying data traffic) MAY be sent in long-lived unidirectional Exchanges to enable a more efficient multicast distribution.

Reasons to terminate a long-lived Exchange include the termination of Port Login and the completion of the IP communication. A long-lived Exchange MAY be terminated by setting the Last_Sequence bit in the F_CTL field of the FC Header to one, or via the ABTS (Abort Sequence) protocol [FC-FS]. A long-lived Exchange SHOULD NOT be terminated by transmitting the LOGO ELS, since this may terminate active Exchanges on other FC-4s [FC-FS].

13. Interoperability with RFC 2625

The IPv4 encapsulation defined in this document, along with Exchange and Sequence management, are as defined in [RFC-2625]. Implementations following this specification are expected to interoperate with implementations compliant to [RFC-2625] for IPv4 packet transmission and reception.

The main difference between this document and [RFC-2625] is in the address resolution procedure. [RFC-2625] uses the Ethernet format of the ARP protocol and requires all Nx_Ports to have a format 0x1 N_Port_Name. This specification defines a Fibre Channel format for the ARP protocol that supports all commonly used N_Port_Names. In addition, this specification does not use FARP [RFC-2625].

An Nx_Port following this specification, and not having a format 0x1 N_Port_Name, is able to interoperate with an [RFC-2625] implementation by manually configuring the mapping <destination IPv4 address, N_Port_Name, N_Port_ID> on the involved Nx_Ports. Through this manual configuration, the ARP protocol does not need to be performed. However, IPv4 communication is not possible if the [RFC-2625] implementation strictly enforces the requirement for Nx_Ports to use N_Port_Names of format 0x1.

An Nx_Port following this specification, and having a format 0x1 N_Port_Name, is able to interoperate with an [RFC-2625] implementation by manually configuring the mapping <destination IPv4 address, N_Port_Name, N_Port_ID> on the involved Nx_Ports, or by performing the IPv4 address resolution in compatibility mode, as described below:

- When IPv4 address resolution is attempted, the Nx_Port MUST send two ARP Requests, the first one according to the FC ARP format and the second one according to the Ethernet ARP format. If only an Ethernet ARP Reply is received, it provides the N_Port_Name of the Nx_Port having the destination IPv4 address. The N_Port_ID associated with the N_Port_Name received in an Ethernet ARP Reply may be retrieved from the S_ID field of the received ARP Reply, or by querying the Fibre Channel Name Server;
- The Nx_Port MUST respond to a received Ethernet ARP Request with an Ethernet ARP Reply;
- The Nx_Port MAY respond to FARP Requests [RFC-2625].

The reception of a particular format of ARP message does not imply that the sending Nx_Port will continue to use the same format later.

Support of compatibility mode is REQUIRED by each implementation. The use of compatibility mode MUST be administratively configurable.

14. Security Considerations

IPv6, IPv4, and ARP do not introduce any additional security concerns beyond those that already exist within the Fibre Channel protocols. Zoning techniques based on FC Name Server masking (soft zoning) do not work with IPv6 and IPv4, because IPv6 and IPv4 over Fibre Channel do not use the FC Name Server. The FC ESP_Header [FC-FS] may be used to secure the FC frames composing FC Sequences carrying IPv6, IPv4, and ARP packets. All the techniques defined to secure IP traffic at the IP layer may be used in a Fibre Channel environment.

15. IANA Considerations

The directory of ARP parameters has been updated to reference this document for hardware type 18.

16. Acknowledgements

The authors would like to acknowledge the ANSI INCITS T11.3 Task Group members who reviewed this document as well as the authors of [RFC-2625] and [RFC-3831]. The authors also thank the IMSS WG and Brian Haberman for their review and comments.

17. Normative References

[FC-FS] ANSI INCITS 373-2003, "Fibre Channel - Framing and Signaling (FC-FS)".

[FC-AL-2] ANSI INCITS 332-1999, "Fibre Channel - Arbitrated Loop-2 (FC-AL-2)".

[IPv6] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[AARCH] Hinden, R. and S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC 3513, April 2003.

[ACONF] Thomson, S. and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, December 1998.

[DISC] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[PMTUD6] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[IPv4] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[ARP] Plummer, D., "Ethernet Address Resolution Protocol: Or converting network protocol addresses to 48.bit Ethernet address for transmission on Ethernet hardware", STD 37, RFC 826, November 1982.

[IEEE-LLC] IEEE Std 802-2001, "IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture".

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

18. Informative References

[RFC-3831] DeSanti, C., "Transmission of IPv6 Packets over Fibre Channel", RFC 3831, July 2004.

[RFC-2625] Rajagopal, M., Bhagwat, R., and W. Rickard, "IP and ARP over Fibre Channel", RFC 2625, June 1999.

[MLDv2] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

[IGMPv3] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

[PMTUD4] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[ICMPv6] Conta, A. and S. Deering, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 2463, December 1998.

[ICMPv4] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, September 1981.

[EUI64] "Guidelines For 64-bit Global Identifier (EUI-64) Registration Authority", http://standards.ieee.org/regauth/oui/tutorials/EUI64.html

A. Transmission of a Broadcast FC Sequence over FC Topologies (Informative)

A.1. Point-to-Point Topology

No particular mechanisms are required for this case. The Nx_Port connected at the other side of the cable receives the broadcast FC Sequence having D_ID 0xFFFFFF.

A.2. Private Loop Topology

An NL_Port attached to a private loop must transmit a Class 3 broadcast FC Sequence by using the OPN(fr) primitive signal [FC-AL-2].

1) The source NL_Port first sends an Open Broadcast Replicate (OPN(fr)) primitive signal, forcing all the NL_Ports in the loop (except itself) to replicate the frames that they receive while examining the FC Header's D_ID field.

2) The source NL_Port then removes the OPN(fr) signal when it returns to it.

3) The source NL_Port then sends the Class 3 broadcast FC Sequence having D_ID 0xFFFFFF.

A.3. Public Loop Topology

An NL_Port attached to a public loop must not use the OPN(fr) primitive signal. Rather, it must send the Class 3 broadcast FC Sequence having D_ID 0xFFFFFF to the FL_Port at AL_PA = 0x00 [FC-AL-2].

The Fabric propagates the broadcast to all other FC_Ports [FC-FS], including the FL_Port that the broadcast arrives on. This includes all F_Ports, and other FL_Ports.

Each FL_Port propagates the broadcast by using the primitive signal OPN(fr), in order to prepare the loop to receive the broadcast sequence.

A.4. Fabric Topology

An N_Port connected to an F_Port must transmit the Class 3 broadcast FC Sequence having D_ID 0xFFFFFF to the F_Port. The Fabric propagates the broadcast to all other FC_Ports [FC-FS].

B. Validation of the <N_Port_Name, N_Port_ID> Mapping (Informative)

B.1. Overview

At all times, the <N_Port_Name, N_Port_ID> mapping must be valid before use.

After an FC link interruption occurs, the N_Port_ID of an Nx_Port may change, as well as the N_Port_IDs of all other Nx_Ports that have previously performed Port Login with this Nx_Port. Because of this, address validation is required after a Loop Initialization Primitive Sequence (LIP) in a loop topology [FC-AL-2] or after Not_Operational Primitive Sequence / Offline Primitive Sequence (NOS/OLS) in a point-to-point topology [FC-FS].

N_Port_IDs do not change as a result of Link Reset (LR) [FC-FS]; thus, address validation is not required in this case.

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

No validation is required after Link Reset (LR). In a point-to-point topology, NOS/OLS causes implicit Logout of each N_Port and after an NOS/OLS each N_Port must again perform a Port Login [FC-FS].

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

After a LIP [FC-AL-2], an NL_Port must not transmit any data to another NL_Port until the address of the other port has been validated. The validation consists of completing the Address Discovery procedure with the ADISC ELS [FC-FS].

If the three FC addresses (N_Port_ID, N_Port_Name, Node_Name) of a logged remote NL_Port exactly match the values prior to the LIP, then any active Exchange with that NL_Port may continue.

If any of the three FC addresses has changed, then the remote NL_Port must be logged out.

If an NL_Port’s N_Port_ID changes after a LIP, then all active logged-in NL_Ports must be logged out.

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

A Fabric Address Notification (FAN) ELS may be sent by the Fabric to all known previously logged-in NL_Ports following an initialization event. Therefore, after a LIP [FC-AL-2], NL_Ports may wait for this notification to arrive, or they may perform an FLOGI.

If the F_Port_Name and Fabric_Name contained in the FAN ELS or FLOGI response exactly match the values before the LIP and if the AL_PA [FC-AL-2] obtained by the NL_Port is the same as the one before the LIP, then the port may resume all Exchanges. If not, then FLOGI must be performed with the Fabric and all logged-in Nx_Ports must be logged out.

A public loop NL_Port must perform the private loop validation as specified in section B.3 to any NL_Port on the local loop that has an N_Port_ID of the form 0x00-00-XX (i.e., to any private loop NL_Port).

B.5. FC Layer Address Validation in a Fabric Topology

No validation is required after Link Reset (LR).

After NOS/OLS, an N_Port must perform FLOGI. If, after FLOGI, the N_Port's N_Port_ID, the F_Port_Name, and the Fabric_Name are the same as before the NOS/OLS, then the N_Port may resume all Exchanges. If not, all logged-in Nx_Ports must be logged out [FC-FS].

C. Fibre Channel Bit and Byte Numbering Guidance

Both Fibre Channel and IETF standards use the same byte transmission order. However, the bit numbering is different.

Fibre Channel bit numbering can be observed if the data structure heading shown in figure 24 is cut and pasted at the top of the figures present in this document.

```
      3                   2                   1                   0
    1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 24: Fibre Channel Bit Numbering

D. Changes from RFC 2625

- Nx_Ports with N_Port_Name format 0x2, 0x5, 0xC, 0xD, 0xE, and 0xF are supported, in addition to format 0x1;
- An IP-capable Nx_Port MUST support Class 3;
- An IP-capable Nx_Port MUST support continuously increasing SEQ_CNT;
- An IP-capable Nx_Port SHOULD support a receive data field size for Device_Data FC frames of at least 1024 octets;
- The FC ESP_Header MAY be used;
- FC Classes of services other than 3 are not recommended;
- Defined a new FC ARP format;
- Removed support for FARP because some FC implementations do not tolerate receiving broadcast ELSes;
- Added support for IPv4 multicast;
- Clarified the usage of the CS_CTL and Parameter fields of the FC Header;
- Clarified the usage of FC Classes of service;
- Clarified the usage of FC Sequences and Exchanges.

E. Changes from RFC 3831

- Clarified the usage of the CS_CTL and Parameter fields of the FC Header;
- Clarified the usage of FC Classes of service;
- Clarified and updated the mapping of IPv6 multicast on Fibre Channel;
- Clarified the usage of FC Sequences and Exchanges;
- Clarified and updated the format of the Neighbor Discovery Link-layer option for Fibre Channel.

Authors' Addresses

Claudio DeSanti
Cisco Systems, Inc.
170 W. Tasman Dr.
San Jose, CA 95134
USA

Phone: +1 408 853-9172
EMail: cds@cisco.com

Craig W. Carlson
QLogic Corporation
6321 Bury Drive
Eden Prairie, MN 55346
USA

Phone: +1 952 932-4064
EMail: craig.carlson@qlogic.com

Robert Nixon
Emulex
3333 Susan Street
Costa Mesa, CA 92626
USA

Phone: +1 714 885-3525
EMail: bob.nixon@emulex.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).