Internet Fibre Channel Protocol (iFCP) — A Technical Overview

A Storage Networking Industry Association and SNIA IP Storage Forum White Paper

Introduction

This paper outlines the technical details of Internet Fibre Channel Protocol, or iFCP. iFCP is a standards-track document within the Internet Engineering Task Force (IETF) actively pursued in the IETF IP Storage Work Group, http://www.ietf.org/html.charters/ips-charter.html.

The Technical Coordinator within the IETF for iFCP is:
Franco Travostino, Nortel Networks, travos@nortelnetworks.com

For questions or comments on this paper, please email cmonia@NishanSystems.com

What is iFCP?

The iFCP specification defines iFCP as a gateway-to-gateway protocol for the implementation of a Fibre Channel fabric in which TCP/IP switching and routing elements replace Fibre Channel components. The protocol enables the attachment of existing Fibre Channel storage products to an IP network by supporting the fabric services required by such devices.

iFCP supports FCP, the ANSI SCSI serialization standard to transmit SCSI commands, data, and status information between a SCSI initiator and SCSI target on a serial link, such as a Fibre Channel network (FC-2). iFCP replaces the transport layer (FC-2) with an IP network (i.e. Ethernet), but retains the upper layer (FC-4) information, such as FCP. This is accomplished by mapping the existing Fibre Channel transport services to TCP/IP. iFCP, through the use of TCP/IP, can therefore accommodate deployment in environments where the underlying IP network is not reliable.

iFCP's primary advantage as a SAN gateway protocol is the mapping of Fibre Channel transport services over TCP, allowing networked, rather than point-to-point, connections between and among SANs without requiring the use of Fibre Channel fabric elements. Existing FCP-based drivers and storage controllers can safely assume that iFCP, also being Fibre Channel-based, provides the reliable transport of storage data between SAN domains via TCP/IP, without requiring any modification of those products. iFCP is designed to operate in environments that may experience a wide range of latencies.

Why iFCP?

iFCP is designed for customers who may have a wide range of Fibre Channel devices (i.e. Host Bus Adapters, Subsystems, Hubs, Switches, etc.), and want the flexibility to interconnect these devices with IP network. iFCP can interconnect Fibre Channel SANs with IP, as well as allow customers the freedom to use TCP/IP networks in place of Fibre Channel networks for the SAN itself. Through the implementation of iFCP as a gateway-to-gateway protocol, these customers can maintain the benefit of their Fibre Channel devices while leveraging a highly scaleable, manageable and flexible enterprise IP network as the transport medium of choice.

iFCP enables Fibre Channel device-to-device communication over an IP network, providing more flexibility compared to only enabling SAN-to-SAN communication. For example, iFCP has a TCP connection per N_Port to N_Port couple, and such a connection can be set to have its own Quality of Service (QoS) identity. With SAN-to-SAN communication, varying connections cannot be prioritized over one another.

Using a multi-connection model for TCP is important for iFCP as it provides higher aggregate throughput compared to an implementation of a single-connection model. With the single-connection model, a single TCP connection links multiple SAN islands, and therefore multiple N_Port to N_Port sessions. One congestion loss in the connection can disrupt the entire fabric and affect all N_Port to N_Port sessions using that tunnel.

With the multi-connection model of iFCP, congestion loss in one N_Port to N_Port session will only affect the throughput of that session. This model isolates the effects of congestion to specific sessions, without impact to other sessions operating in parallel.

For more information on the user benefits, applications and implementation of iFCP, please see, *The Benefits of Internet Fibre Channel Protocol (iFCP) for Enterprise Storage Networks.*

Overview of an IP Network for Storage

iFCP Network Model

iFCP maps Fibre Channel transport services to an IP fabric as outlined in Figure 1: iFCP Network Model. In this implementation, gateways are used to connect existing Fibre Channel devices to an IP network, and as such will include physical interfaces for both Fibre Channel and IP. iFCP is a TCP/IP protocol that transports encapsulated FC-4 frame images between gateways. iFCP session end points are Fibre Channel N_Ports.

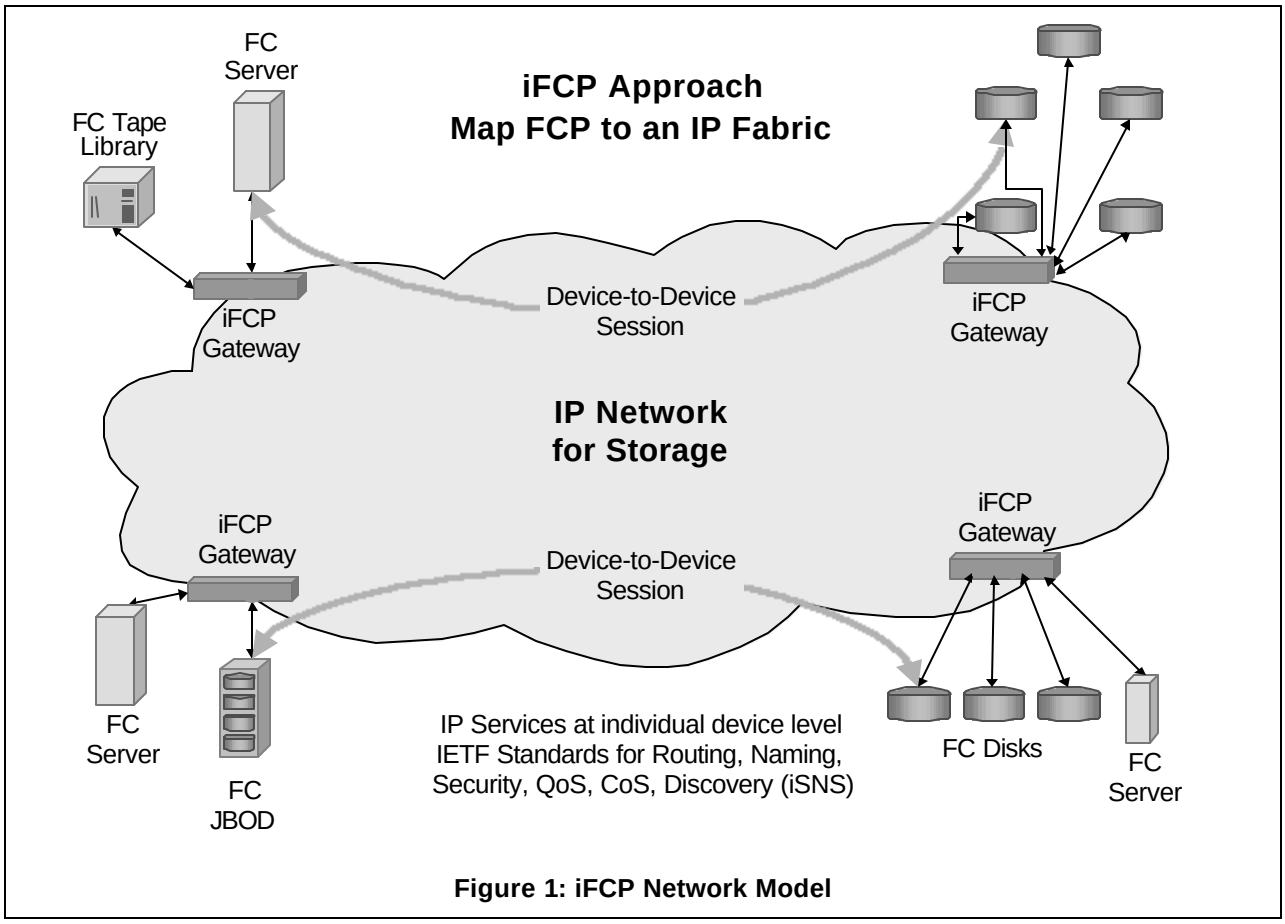


Figure 1: iFCP Network Model

iFCP Implementation Model

Figure 1 outlines the model whereby a storage network is built of Fibre Channel end devices, and TCP/IP switching and routing equipment in the network. Each iFCP gateway maintains its own gateway region as a link between Fibre Channel N_Ports and an IP network. What's more, the gateway region acts as a firewall, isolating the storage fabric from faults within the region.

However, the physical implementation of gateway regions can also include Fibre Channel switches and Fibre Channel Arbitrated Loops.

Figure 2: iFCP Implementation Model, shows an example of two gateway regions that include Fibre Channel switches and arbitrated loops. In each of these cases, the N_Ports in the gateway region (i.e. the actual end node storage devices) are presented to the IP network in the same manner as if there were no Fibre Channel networking equipment in the gateway region.

This opaque view of the gateway region allows iFCP implementations to grandfather existing Fibre Channel installations that include hubs, switches and loops.

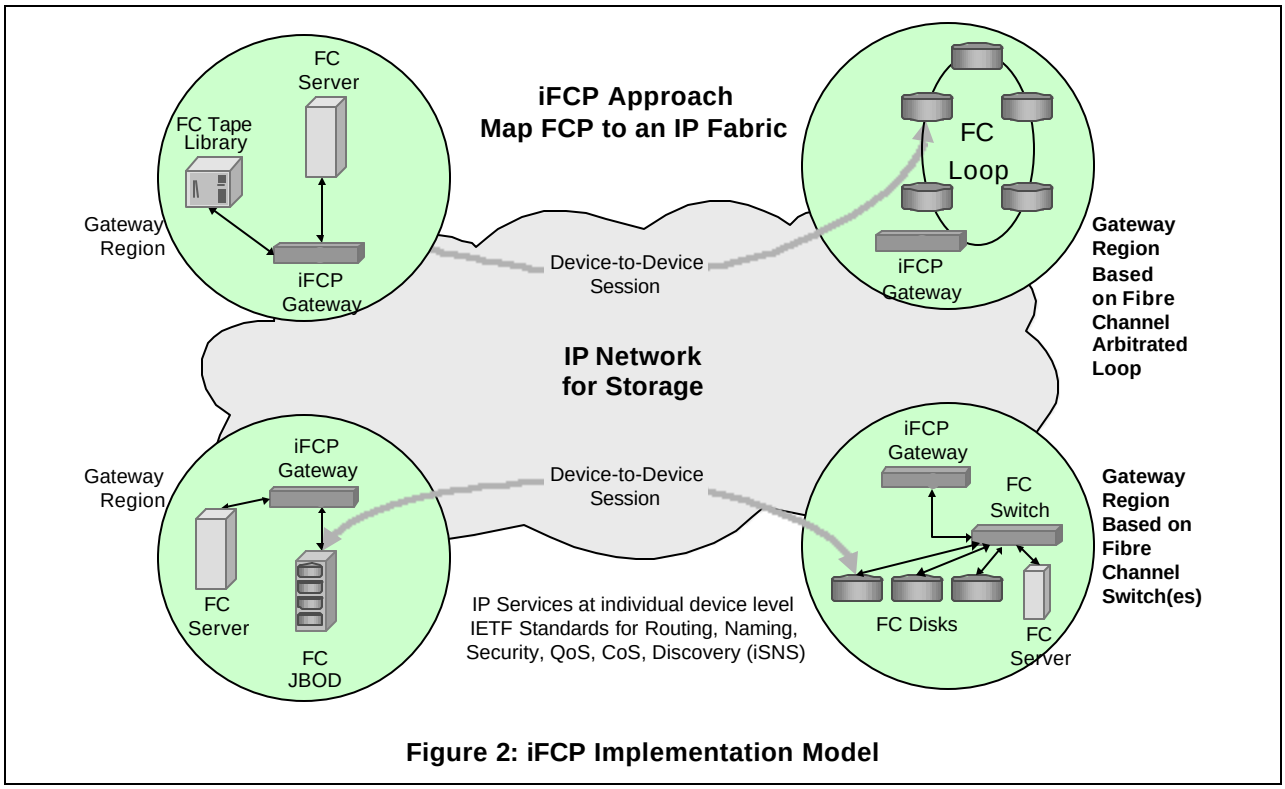


Figure 2: iFCP Implementation Model

Incorporation of storage name services (iSNS)

Because Fibre Channel devices rely on Fibre Channel Generic Services (FC-GS), IP equivalent protocols are provided as companions to iFCP. For example, Internet Storage Name Service (iSNS) is an IP protocol providing FC-compatible storage name services. The iSNS queries are outlined in Figure 3: iSNS for FC-compatible storage name services. iSNS processes FC name service requests, and serves as a generic internet storage name server protocol for the primary IP storage protocols iFCP and iSCSI. The name server repository contains all iFCP storage objects such as discovery domains (i.e. zones), FC devices, N_Ports, gateways, etc. and maintains security and access control information.

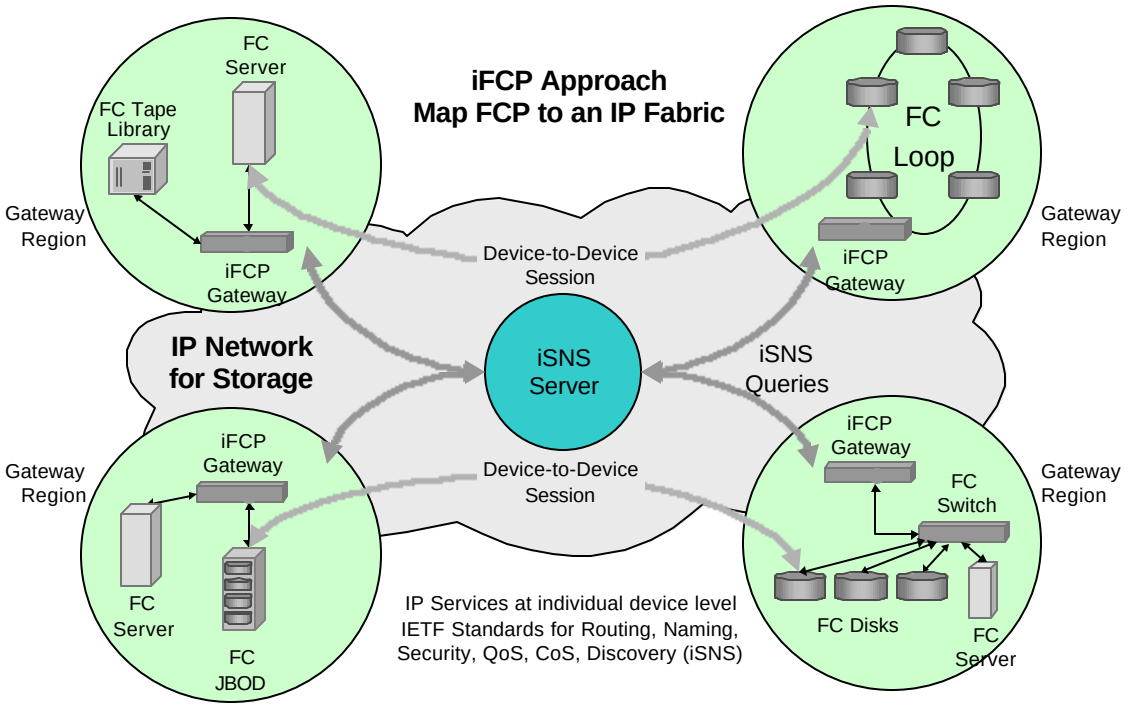


Figure 3: iSNS for FC-compatible storage name services

Key Comparisons between Fibre Channel and iFCP

	Fibre Channel (FC)	Internet Fibre Channel Protocol (iFCP)
Routing	• FSPF	• OSPF or any other IP routing protocol

General Services	• Name services, security key distribution, time services, zone management, fabric configuration services, management services
	• Based on FC-GS2
		• IP based
		• Name services, security key distribution, zoning
		  • iSNS, TLS, etc.
		• Time services
		  • /TBS/

Management	• SNMP	• SNMP
Fabric Services	• Class 1, class 2, class 3 per FC-FS	• Class 2, class 3

Figure 4: Key comparisons between Fibre Channel and iFCP

Role of Gateway Regions

As outlined in Figure 3: iSNS for FC-compatible storage name services, Gateway Regions (GR) are exposed to the IP fabric. Within Gateway Regions, only N_Ports are exposed to the IP fabric,

and other elements such as Fibre Channel hubs or switches are not exposed. This eliminates the need for class F Fibre Channel traffic to flow across the IP network and allows for equivalent class F functionality to be performed using IP protocols.

A Gateway may control the assignment of all N_Port addresses within the Gateway Region , and reconfiguration does not affect the state of other Gateway Regions.

Addressing and Routing

With iFCP, N_Port addressing can be locally assigned by each gateway for a Gateway Region local mode operation. Alternatively, in address-transparent mode, N_Ports can be globally assigned across an interconnected set of gateways.

The routing between Gateway Regions operates with IP only. Routing that takes place within a Gateway Region (if there is any routing within) is opaque to the IP network. For example, Fibre Channel routing and DFS traffic that may be operating within a Gateway Region does not flow between Gateway Regions.

Address Transparent Mode

In address transparent mode, the scope of N_Port addresses is fabric wide. The IP network fabric is defined as a name server object containing a collection of gateways. The iSNS name server acts as a fabric Domain Address Manager, and maintains a pool of Domain IDs for the fabric, assigning FC domain IDs to each gateway within the fabric. Within each Gateway Region, the gateway acts as the *downstream* principal switch.

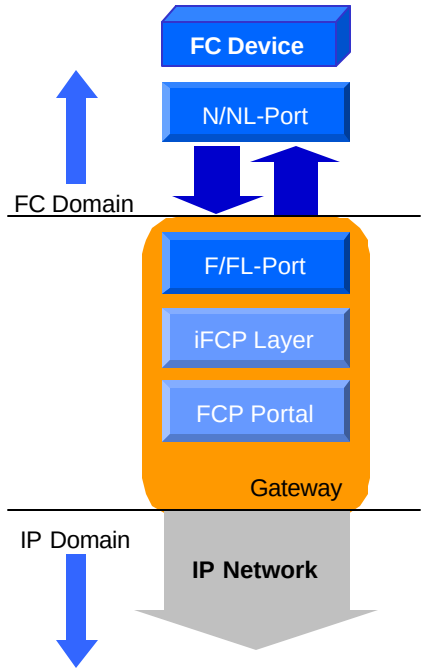
The advantage of address transparent mode is the transparency across the fabric and the resulting simplification of gateway operation. The disadvantage is that each Gateway Region consumes 65K of Node IDs and this is inefficient when the Gateway Region N_Port count is low. Also, Address Transparent Mode is less scaleable as communication among N_Ports is restricted to N_Ports within the fabric.

Gateway Region Local Mode

In Gateway Region Local mode, the scope of the N_Port addresses is local to the Gateway Region. Each gateway maps N_Port network addresses of external devices to N_Port fabric addresses. Normal inter-gateway frame traffic is mapped on the fly.

The advantage of Local Mode is scalability. N_Port connectivity is network-wide, allowing unrestricted addresses within a Gateway Region. Since each gateway is individually responsible for N_Port addresses allocated to its Gateway Region, the fabric becomes more stable as the network scales in size. This is because there is no dependence on a central addressing authority, as is the case with Fibre Channel and iFCP Transparent Mode fabrics.

The disadvantage of local mode is that gateways must be more ELS-aware with special handling required for ELS traffic containing N_Port addresses in the payload. Further, each gateway must maintain and update address translation tables.



FCP Portal Address
- IP Address of Gateway

N_PORT ID
- Gateway-assigned address of a directly attached N_PORT

N_PORT Fabric Address
- Address of an N_PORT in the FC Domain
  - N_PORT ID for directly attached devices
  - Gateway-assigned alias for external devices (GR-local mode)

N_Port Network Address
- Address of an N_PORT in the IP Domain
  - N_PORT ID + FCP Portal Address

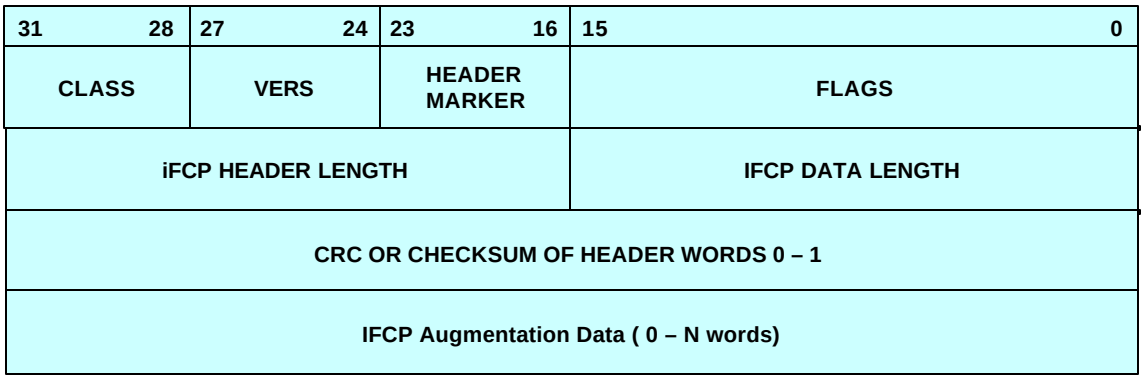
Figure 5: Outline of Gateway Architecture

Explanation of Gateway Architecture

The Gateway architecture includes the FC networking domain and the IP networking domain. Fibre Channel devices are directly connected to the iFCP fabric through F_Ports implemented as part of the iFCP gateway. At the N_Port interface on the Fibre Channel side of the gateway, the network appears as a Fibre Channel fabric. Here, the gateway presents remote N_Ports as directly attached devices. Conversely, on the IP-side, the gateway presents each locally connected N_Port as a logical iFCP device on the IP network. Network addresses from each domain are mapped on to the other through the Gateway architecture.

Within the Fibre Channel device domain, fabric-addressable entities consist of other N_Ports and devices internal to the fabric that perform the fabric services defined in [FGS].

As frames are received from the IP domain, they are mapped to Fibre Channel Port_ID addresses and passed up to the Fibre Channel upper layers supported by the FC device. Similarly, frames passed to the iFCP gateway by Fibre Channel devices are mapped to an N_Port network address, and encapsulated with the proper IP address and N_Port ID before they are delivered to the IP network domain.



Class = FC Class of service (2 or 3)
Version = iFCP Protocol Version Number
Header Marker = 0xAA
iFCP Header length = Size of header in words
(2 + Length of augmented data field)
iFCP Data Length = Length of iFCP frame in words

Augmented data = Present for ELS frames containing N_PORT addresses in payload

FLAGS

15	Sequence End
14	Sequence Start
13	1 = iFCP encapsulation
12 - 3	Reserved
2	TCP ELS
1	Augmentation present
0	Compliance (must be 1)

Figure 6: iFCP Header Detail

Mapping of Fibre Channel to iFCP

Fibre Channel frames ingressing the iFCP gateway are converted to iFCP frames through the process shown in Figure 7. The FC frames may be addressed to remote devices, or to other FC devices attached to the same iFCP gateway. If the latter is the case, no address translation mechanism is needed, and the frame is directly delivered to the local N_Port. If the former is the case, then an address mapping function must occur that maps a key found in the D_ID to the TCP connection addressed to the appropriate remote N_Port network address (N_Port ID and IP address).

The iFCP gateway is responsible for assigning an alias used to look up the N_Port network address for the external device that is not directly attached to the gateway. To perform this function a gateway or edge switch maintains a table that maps frame traffic to the appropriate TCP/IP connection and N_Port ID of all external N_Ports with active sessions. As shown in Figure 7, this alias is assigned using the upper byte DOMAIN_ID set to 0x80.

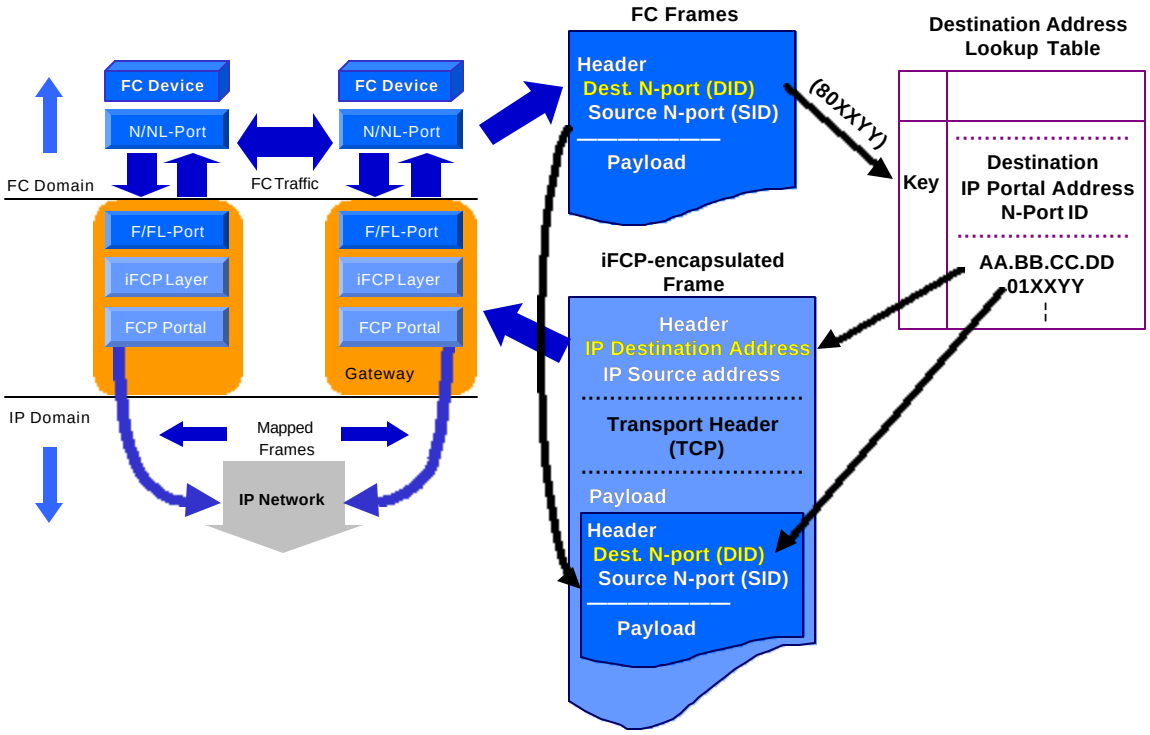


Figure 7: iFCP Mapping Outbound Traffic FC to iFCP

The address translation table mapping N_Port ID's to TCP connections and N_Port addresses is constructed by consulting the iSNS. Each new Port Login command (i.e. PLOGI) forces the iFCP gateway to consult the iSNS, where appropriate N_Port network addresses of external devices can be retrieved. This information is maintained in an internal table within the iFCP gateway that maps the alias found in the D_ID of the incoming Fibre Channel frame to the N_Port network address retrieved from the iSNS. Fibre Channel frames ingressing the F_Port can now be mapped to a TCP connection and N_Port network address on the fly. This process is quite similar to that of ARP (Address Resolution Protocol), used to map IP addresses on to Ethernet MAC addresses.

Mapping of iFCP to Fibre Channel

The opposite process occurs when the FCP Portal receives inbound Fibre Channel frames from a TCP connection. These inbound frames must be mapped on to the gateway-assigned N_Port alias. For inbound frames, the iFCP gateway regenerates the alias from the TCP connection context and destination N_Port ID contained in the encapsulated FC frame. The source N_Port alias found in the lookup is used to replace the Source ID (S_ID) in the received Fibre Channel frame. The translation process is shown below in Figure 8.

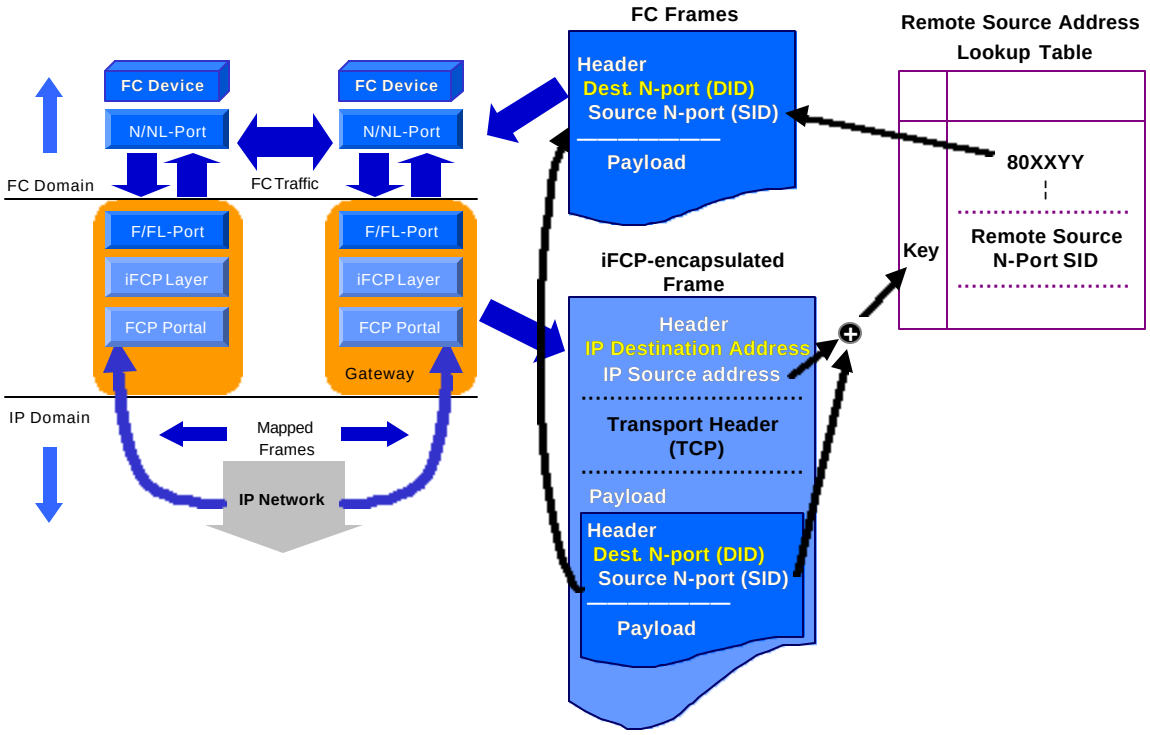


Figure 8: iFCP Mapping Inbound Traffic iFCP to FC

Overview of Extended Link Services

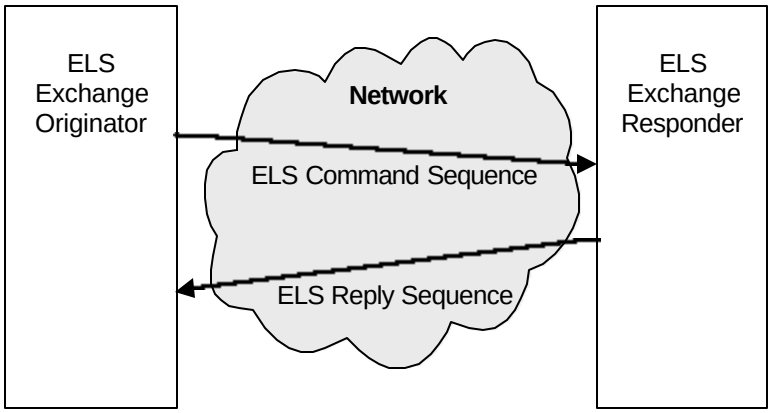


Figure 9: Overview of Extended Link Services

Extended Link Services are a set of Fibre Channel transport services that one port can use to invoke certain functions or services at another port.

A separate exchange takes place for each Extended Link Service operation. Each exchange usually includes a command or request sequence and a reply sequence that ends the exchange. A

sequence consists of one of more frames and operates with the normal rules for sequence identification and management.

Figure 9: Overview of Extended Link Services, provides an outline of the operation.

Processing of Extended Link Services

Most Extended Link Services can be passed transparently by iFCP between the originating and receiving Fibre Channel nodes. However, a subset of ELS commands reference N_Port ID values which must be translated to locally-significant aliases in the receiver's frame of reference. This process of translating nominal Port_ID values is known as ELS Augmentation.

Figure 10 illustrates the process flow whereby outbound ELS commands are examined to see if they reference N_Port values. If the ELS command is addressed to the fabric itself, or to a node attached to the same iFCP gateway, then no ELS Augmentation will be required. If the ELS command is destined to a remote iFCP gateway, then the command itself must be examined to see if it references an N_Port ID value. If augmentation is required, then processing will occur whereby each Port Name (WWPN) corresponding to the remote Port_ID will be added to the payload of the iFCP frame. This process is shown in Figure 10.

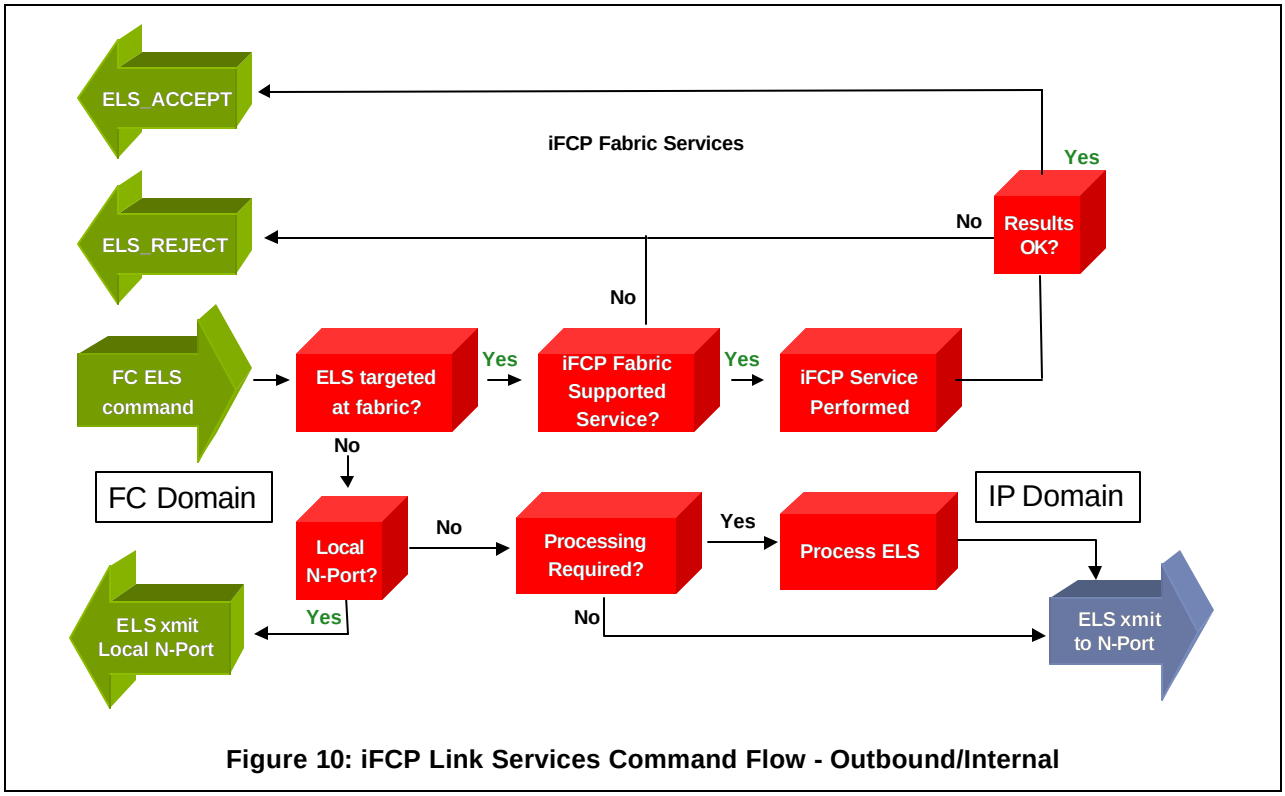


Figure 10: iFCP Link Services Command Flow - Outbound/Internal

Once the augmented ELS command is received by the remote iFCP gateway, the augmented Port Name data is used to construct locally-significant N_Port_ID values that substitute for the original N_Port_ID values in the ELS command, before it is forwarded to the receiving N_Port.

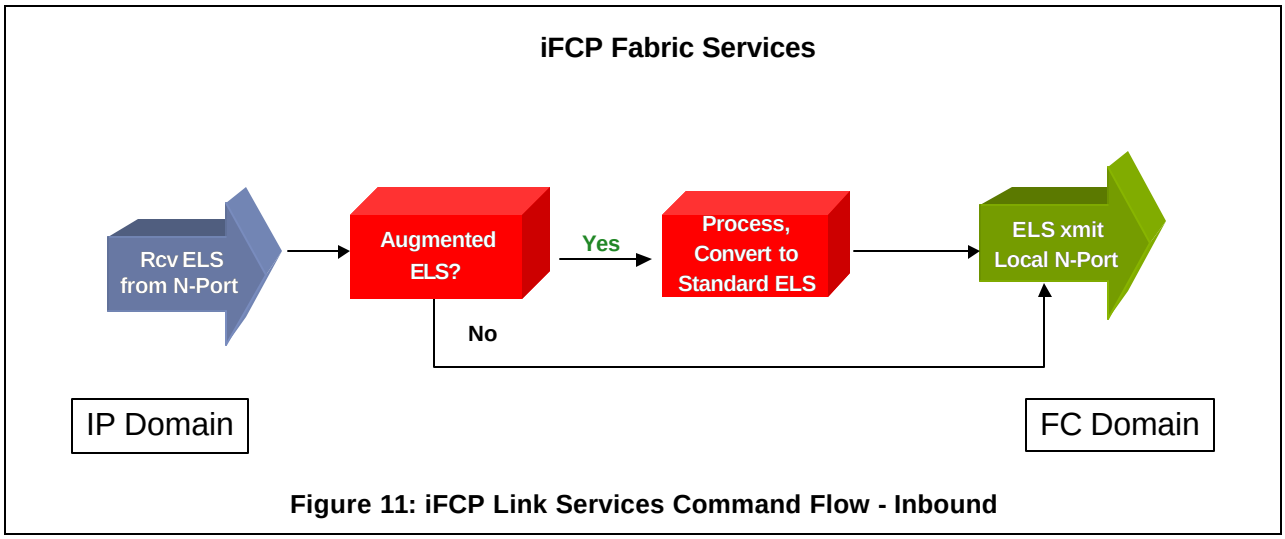


Figure 11: iFCP Link Services Command Flow - Inbound

iSNS Internet Storage Name Service

Each Fibre Channel Name Service message has an equivalent iSNS message. This mapping is transparent, allowing the iFCP fabric with iSNS support to provide the same services that a Fibre Channel fabric can with FC-GS-2.

When an iFCP gateway receives a Name Service ELS, it is directly converted to the equivalent iSNS Name Service message. The iFCP gateway intercepts the response, and maps any addressing information obtained from queries to its internal address translation table before forwarding the Name Service ELS response to the original Fibre Channel requester.

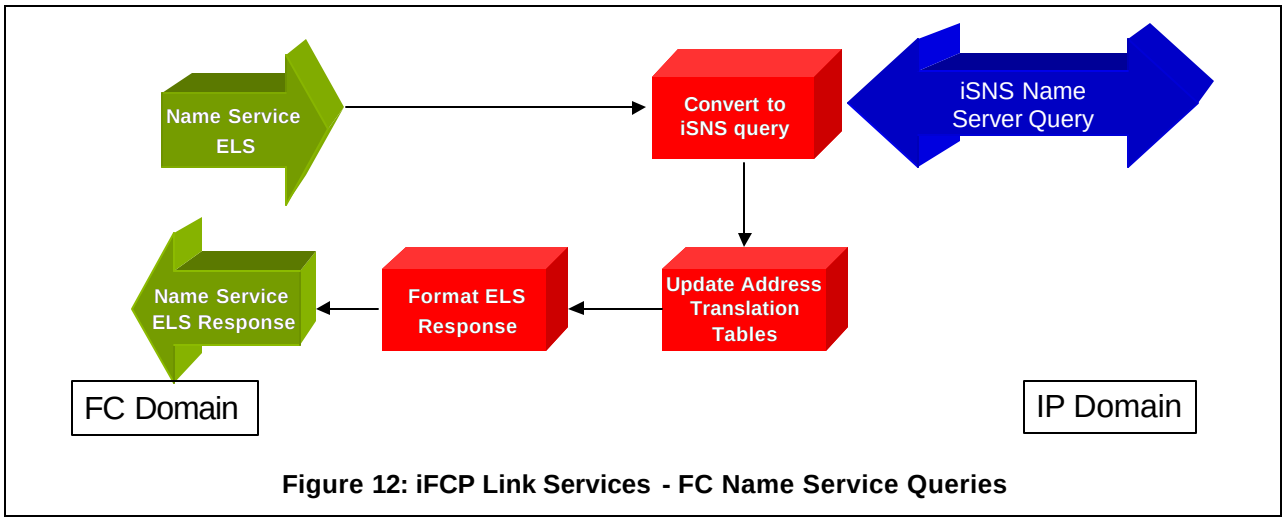


Figure 12: iFCP Link Services - FC Name Service Queries

TCP Connection Setup

TCP connections are established upon detection of the Port Login (PLOGI) ELS command. When a locally attached device attempts a Port Login, the iFCP gateway intercepts the request, and establishes a TCP connection to the appropriate remote iFCP gateway that supports the target of the login request. The Port Login ELS command is then augmented with the Port Name (WWPN) of the target device, and then transmitted over the newly-setup TCP connection.

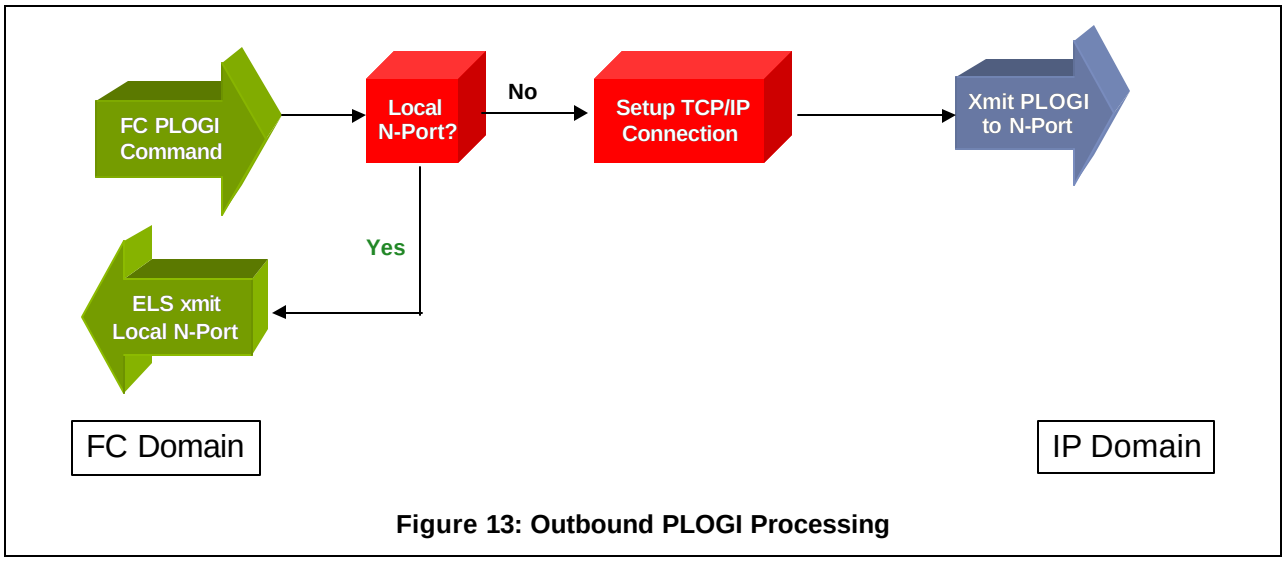


Figure 13: Outbound PLOGI Processing

At the remote iFCP gateway supporting the target of the Port Login request, the PLOGI ELS is received, and mapped to the locally significant N_Port ID using the Port Name (WWPN) found in the augmentation data.

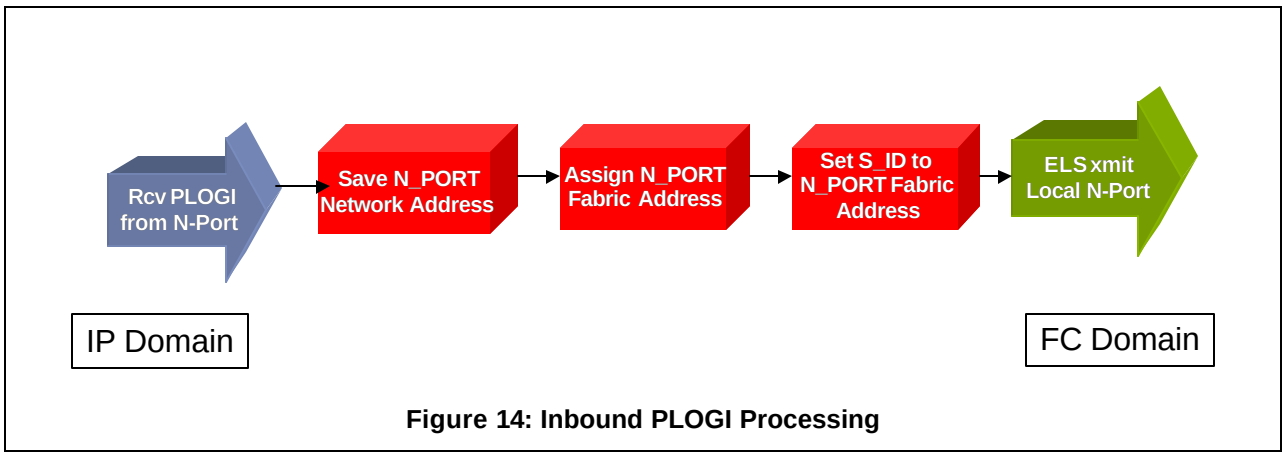


Figure 14: Inbound PLOGI Processing

iFCP Gateway Implementation Example

Fibre Channel
- F-Port
- FL-Port
- FC Auto Port
- E-port

Gigabit Ethernet
- Gigabit Ethernet Port
- Gigabit Ethernet Trunk Port

Management
- IP Java Mgmt Port
- RS232 Console Port

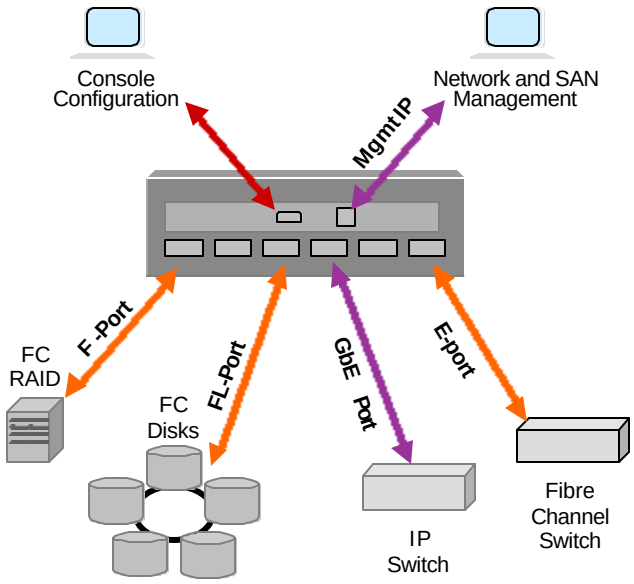


Figure 15: iFCP Gateway Implementation Example

Figure 15: iFCP Gateway Implementation Example, highlights the key functions that might be present in an iFCP product.

Fibre Channel F-port
This port can connect directly to Fibre Channel storage devices such as a Fibre Channel RAID.

Fibre Channel FL-Port
This port can connect directly to Fibre Channel Arbitrated Loops.

Fibre Channel E-port
This port can connect directly to other Fibre Channel switches.

Fibre Channel Auto Port
This port can connect directly to all of the above types of Fibre Channel ports and negotiate between types.

Gigabit Ethernet Port
This port can be connected to IP and Gigabit Ethernet networking equipment.

Gigabit Ethernet Trunk Port
This port can be used to aggregate multiple IP and Gigabit Ethernet links for greater bandwidth.

Management Port (i.e. 10/100 Ethernet)
This port can be used for out of band management such as Java-based device management software.

RS232 Console Port (i.e. serial setup, maintenance)
This port can be used for serial setup such as in assigning and IP address, or for maintenance and diagnostic information of the gateway.

References

Charles Monia, Nishan Systems, cmonia@NishanSystems.com
Josh Tseng, Nishan Systems, jtseng@NishanSystems.com
Franco Travostino, Nortel Networks, travos@NortelNetworks.com
Victor Firoiu, Nortel Networks

Robert W. Kembel, Fibre Channel: A Comprehensive Introduction. © 1998, 2000. Published by Northwest Learning Associates, Inc.

For more information about the iFCP standard, visit www.ietf.org