IPS

Internet Draft
Document: draft-ietf-ips-ifcp-wglcc-00.txt
Category: Informational

C. Monia
F. Travostino
J. Tseng
Nishan Systems

May 2002

Responses to iFCP Rev. 10 Last Call Comments

Status of this Memo

Abstract

This document is a compilation of responses to review comments received for revision 10 if the iFCP specification [IFCP] during the preliminary last call period from 3/4/2002 to 3/18/2002.

Change Log

Revision 0 of draft-monia-ips-ifcplcc-00 to draft-ietf-ips-ifcp-wglcc-00.txt, revision 0.

Comment 49 -- Modified to comply with IESG policy on number of co-authors.

Modified the following responses per feedback from David Black and Mallikarjun Chadalapaka.

Comment 5,
Comment 12,
Comment 94,
Comment 104,
Comment 110.
Comment 120

1. Conventions used in this document

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 [RFC2119].

In the following, [E] designates an editorial comment, [T] a technical comment.

The keywords 'Rejected' or 'Accepted' indicate fundamental agreement or disagreement with the position stated in the comment.

The keyword 'Response' is used when a comment is predicated on a query. The explanatory text should be consulted for details.

2. Comments from David Black

Comment 1. [E} Page 5, Change Log

Remove Change Log in the version after a successful WG Last Call.

Accepted

Comment 2. [E] Section 2.1, page 7, paragraph 1

"Terms needed to clarify the concepts presented in this document are presented here."

I don't like the usage of "clarify". How about "Terms used to describe the concepts presented in this document are defined here." ?

Accepted

The text will be revised as suggested.

Comment 3. [E] Section 2.1, Address Translation Mode Definition

Some tool has helpfully inserted non-ASCII characters. MS Word AutoCorrect is a likely suspect. Hunt all of these down and fix them, then discipline the tool severely ;-).

Accepted.

Comment 4. [T] Section 2.1, Definition of FC-4 Layer

"FC-4 - The fibre channel application layer. This layer is functionally equivalent to the TCP/IP application layer."

I don't understand this. Are you equating FC-4 with OSI layer 7? If so, I'm not sure that is correct, and it might be better to leave out this attempted analogy.

Accepted

The definition will be changed to:

"FC-4 - The fibre channel mapping of an upper level protocol, such as [FCP-2], the fibre channel SCSI mapping."

Comment 5. [T] Section 3.2, page 10

a) "Arbitrated Loop -- A series of N_PORTs connected together in daisy-chain fashion. Data transmission between N_PORTs requires arbitration for control of the loop in a manner similar to a token ring network."

That's not a fabric topology, unless the loop is fabric attached, in which case you're in case c), Mixed Fabric. iFCP can't support an FC-AL loop that isn't fabric-attached.

Accepted in part

The terminology will be changed to "fibre channel network topologies".

In addition, the following definition will be added:

"Fabric -- From [FC-FS]: "The entity which interconnects N_PORTs attached to it and is capable of routing frames by using only the address information in the fibre channel frame."

With regard to FC-AL support, an iFCP gateway implementation can emulate either a public (fabric attached) or private loop environment. The gateway may support a private loop by representing remotely attached devices as if they were resident on a local loop and by emulating the semantics required to support the loop control frames and primitives. Since these functions are implemented internally by the gateway, iFCP protocol support is not required. However, the specification should explicitly prohibit the forwarding of fibre channel loop control frames via iFCP.

A related issue is support for the set of extended link services for remote loop control, such as LINIT (Loop Initialize). These are standard link service messages addressed to the loop fabric address (LFA) of the FL port controlling the loop. A gateway that chooses to expose remote, loop-connected devices as NL_Ports must also expose the LFA. To do so, it should assign the local alias such that the corresponding LFA or the remote loop can be derived by setting the port_id component of the alias to zero in accordance with [FC-FS].

The specification will be modified to discuss these loop topology support issues.

Comment 6. [T] Section 3.2, page 11, para 5

"Depending on the topology, the N_PORT and fabric port variants through which a fibre channel device is attached to the network may be one of the following:

"Fabric Topology	Fabric Port Type	N_PORT Variant
Loop	L_PORT	NL_PORT
Switched	F_PORT	N_PORT
Mixed	FL_PORT	NL_PORT
	F_PORT	N_PORT"

I believe the Loop line in this table does not match the other lines and if so, this is one more reason to leave non-fabric-attached FC-AL out of this description.

Accepted in part

Since the loop topology can be supported, it should remain in the table. However, the terminology should be changed per Comment 5 and the table modified as shown below:

"FC Network Topology	N_PORT Variant	FC Network Interface

Loop	NL_PORT	L_PORT
Switched	N_PORT	F_PORT
Mixed	NL_PORT	FL_PORT via L_PORT"

In the case of a mixed fabric, additional supporting text will be provided.

Comment 7. [E] Section 3.3.1, page 14, para 2

"All switched fabrics must provide the following services:

"Fabric F_PORT server û Services an N_PORT request to access the fabric for communications.

Change "request" to "requests"

Accepted

Replace special character and reword as follows:

"Fabric F_PORT server -- Services N_PORT requests to access the fabric for communications."

Comment 8. [E] Section 4.4, page 21, para 2

"As discussed below, an unbounded iFCP fabric may have any number of switch elements and gateways."

It's not "any", but the limit is a very large number by comparison to 239.

Accepted

The sentence will be changed to:

"As discussed below, an unbounded iFCP fabric is not limited to 239 switches and gateways."

Comment 9. [T] Section 4.4 - iFCP Fabric Properties

At some point the need to reuse 24-bit addresses for outbound traffic from a single FC link behind an iFCP gateway will be a problem. This comment also applies to the second paragraph in Section 4.4.2.

Accepted

A discussion of address re-use issues will be added to the spec.

Comment 10. [E] Section 4.5, page 23, para 2

"In the iFCP protocol, an N_PORT is represented by the following addresses:"

Change "addresses" to "types of addresses" to avoid implying that there's only one alias. Different gateways will assign different aliases to the same N_PORT.

Rejected

The description of an alias will be revised as follows:

b) "A 24-bit N_PORT alias. The fibre channel N_PORT address assigned by each gateway operating in address translation mode to identify a remotely attached N_PORT.

Frame traffic is intercepted by an iFCP gateway and directed to a remotely attached N_PORT by means of the N_PORT alias. The address assigned by each gateway is unique within the scope of the gateway region."

Comment 11. [T] Section 4.5, pp 24, para 14

"The mode of gateway operation is settable in an implementation-specific manner. The implementation MUST NOT allow the mode to be changed after the gateway begins processing fibre channel frame traffic."

Might want to add a MUST that a gateway cannot operate in more than one mode at the same time, and a repeat of the (implied) requirement that all gateways in an iFCP fabric MUST operate in the same mode.

Accepted

Comment 12. [T] Section 4.6. pp 24, para 2

b) "When interoperating with locally attached fibre channel switch elements, each iFCP gateway MUST assume control of DOMAIN_ID assignments in accordance with the appropriate fibre channel standard or vendor-specific protocol specification."

This is ok, but turns up another requirement that needs to be explicitly stated earlier. Any given FC N_PORT MUST NOT be behind more than one iFCP gateway. Address Transparent mode satisfies this because only one gateway can become the principal switch, so the others presumably shut down, but Address Translation mode appears to have the potential for seriously nasty misbehavior unless the "iFCP gateway MUST become the principal switch" requirement is imposed on it also. Need to add a sentence or two on how an iFCP gateway can be assured of becoming the principal switch. Beyond this, the fact that any Fabric Attached FC-AL loop can have only one FL

port completes the picture, ensuring that a loop can't stitch two gateway domains together. Requiring the iFCP gateway to be the principal switch also avoids problems with the gateway being unable to obtain sufficient Domain IDs from the principal switch.

Accepted in Part

An N_PORT can be behind more than one gateway if the following rules are observed:

a) The gateways MUST cooperate in the assignment of N_PORT IDs for locally attached devices and aliases for remotely attached devices such that each local or remotely attached N_PORT has one and only one N_PORT address within the scope of the gateway region.

b) All iFCP frame traffic between any two N_PORTs MUST flow through a single iFCP session. However, each session may traverse a different gateway attached to the region.

In order to meet these constraints, a multi-gateway implementation may require an out of band mechanism for redirecting frame traffic from one physical gateway to another.

The above will be added to the specification.

Comment 13. [T] Section 5.2.2.2, pp 32, para 4

"The gateway SHALL initiate the creation of an iFCP session in response to a PLOGI ELS directed to a remote N_PORT from a locally attached N_PORT as described in the following steps.

a) "Using the D_ID field in the PLOGI frame header, locate the remote N_PORT descriptor. If no descriptor exists, the iFCP gateway SHALL return a response of LS_RJT, with a Reason Code of 'Unable to Perform Command Request' (0x09) and a Reason Code Explanation of 'Invalid N_PORT_ID' (0x1F). An iFCP session SHALL NOT be created."

Need to explain why this is ok.

The answer is that a properly operating FC N_PORT will have previously issued an FC nameserver query that the gateway translated to an iSNS query, and hence when it issues PLOGI to the result of the nameserver query, the iSNS query response created the required descriptor in the gateway before being translated to the FC nameserver result. There's an implication here that remote N_PORT descriptors that result from iSNS queries translated from FC nameserver queries MUST NOT be discarded as long as any N_PORT that has issued a query for that remote N_PORT is logged into the fabric.

Accepted in part

Although a name server query is almost always done in practice prior to a PLOGI, an N_PORT compliant with [FC-FS] is not required to do so. For that reason, the specification should cover the case where a fibre channel device attempts to send frames to an address without having executed a previous name server query.

Also, while the policies for remote N_PORT descriptor retention are implementation-specific, the specification should at least contain recommendations. In that regard, the following added text is proposed:

"Remote N_PORT Descriptors should be reclaimed based on a last in, first out policy.

"An iFCP implementation should have sufficient resources to insure that a newly created descriptor is not reclaimed before the referencing iFCP session is created."

Comment 14. [E] Section 5.2.2.2 - Creating an iFCP Session

e) "If a CBIND response is returned with one of the following statuses, the PLOGI SHALL be terminated with an LS_RJT message. Depending on the CBIND failure status, the Reason Code and Reason Explanation SHALL be set to the following values specified in [FC-FS]."

Add a statement that this plus case f) is a comprehensive list of possible CBIND failure statuses, as specified in Section 6.1.

Accepted

Comment 15. [E] Section 5.2.2.2 - Creating an iFCP Session

f) "A CBIND response with a CBIND STATUS of "N_PORT session already exists" indicates that the remote gateway has concurrently initiated a CBIND request to create an iFCP session between the same pair of N_PORTs. The receiving gateway SHALL terminate this attempt, return the connection to the Unbound state and prepare to respond to an incoming CBIND request as described below."

Add a sentence indicating that the "simultaneous open" race is dealt with by allowing the sender with the numerically larger N_PORT name to succeed in establishing the session.

Accepted

Comment 16. [E] Section 5.2.2.2, pp 34, para 2

"The gateway receiving a CBIND request SHALL respond as follows:

a) "If the receiver has a duplicate iFCP session in the OPEN PENDING state, then the receiving gateway SHALL compare the Source N_PORT Name in the incoming CBIND payload with the Destination N_PORT Name."

b) "If the Source N_PORT Name is greater, the receiver SHALL issue a CBIND response of "Success" and SHALL place the session in the OPEN state."

Add a sentence indicating that in case b), case c) will occur at the other gateway because N_PORT names are globally unique WWNs, and hence this gateway's duplicate session will receive a CBIND STATUS of "N_PORT session already exists" and will be terminated in due course.

Accepted

Comment 17. [T] Section 5.2.2.2 - Creating an iFCP Session

There's no discussion of what to do if a TCP connection closes unexpectedly during this process (e.g., if closing of unbound connections is allowed at arbitrary times for reasons such as reducing the resources consumed by unbound connections). This needs to be added even if the reason in parentheses is not allowed.

Accepted

Comment 18. [T] Section 5.2.2.2, pp 35, para 4

"Upon receiving such a request, the gateway providing the connectivity probe SHALL transmit LTEST messages at the specified interval."

This requires liveness test (LTEST) messages even when the connection is in active use. Was that intended?

Response

The intent is to require LTEST messages at the specified interval regardless of whether or not there is other traffic.

Comment 19. [E] Section 5.2.2.4 - Use of TCP Features and Settings

For Wrapped sequence detection, "Should use" in the table should be "SHOULD use".

Accepted

Comment 20. [T] Section 5.2.3.1, pp 38, para 1

"In response to the Unbind message, either gateway may choose to close the TCP connection or return it to a pool of unbound connections."

This assumes that Unbind is always successful. It can fail, as documented in Section 6.2. Need to specify how to deal with this (e.g., close the TCP connection).

Accepted

The sentence will be modified as follows:

"Upon successful completion of an Unbind operation, either gateway may choose to close the TCP connection or return it to a pool of unbound connections."

The processing for the failure cases will also be specified.

Comment 21. [T] Section 5.2.3.1 - iFCP Session Completion

Can an iFCP gateway reduce the pool of unbound connections (e.g., due to demands for resources for other connections), possibly by closing them? If yes, need to say so.

Accepted

A gateway may close an unbound connection due to resource demands. The spec will be modified appropriately.

Comment 22. [E] Section 5.3 - IANA Considerations

Put this near the end of the document where IANA can more easily find it.

Accepted

Comment 23. [T] Section 5.4.1, pp 40, para 1

"Protocol# IANA-assigned protocol number identifying the protocol using the encapsulation. For iFCP the value is (/TBD/)."

It's 2 - cite the FC Encapsulation draft's IANA Considerations section as the authority for this.

Accepted

Comment 24. [E] Section 5.4.2 - SOF and EOF Delimiter Fields

Need to say that the format is specified in the FC Common Encapsulation document and reproduced here for convenience.

Accepted

Comment 25. [T] Section 5.4.2 - SOF and EOF Delimiter Fields

"SOF (bits 31-24 and bits 23-16 in word 0): iFCP uses the following subset of the SOF fields described in [ENCAP].

This is a problem because these codes are being specified in more than one place. I think the FC Frame Encapsulation document is the right place for the normative specification of these codes (and see my comments against it on the need to get IANA involved). This would be ok as a list of codes that are currently valid, but the specification authority needs to be in one place. Same comment applies to EOF.

Accepted in Part

The specification will be revised in accordance with Comment 24.

Comment 26. [E] Section 6, pp 46

"LS_COMMAND For a special link service ACC response to be processed by iFCP, the LS_COMMAND field SHALL contain bits 31 through 24 of the LS_COMMAND to which the ACC applies. Otherwise the LS_COMMAND field shall be set to zero."

There's an LS_COMMAND field in figure 16 and a second one in the iFCP portion of the FC Common Encapsulation header (from Section 5.4.1).

When a single section discusses both fields, as Section 6 does, this gets confusing fast. Please rename the LS_COMMAND field in the iFCP portion of the FC Common Encapsulation header to something like ACC_LS_COMMAND or LS_COMMAND_ACC.

Accepted

The mnemonic will be changed to LS_COMMAND_ACC.

Comment 27. [T/E] Section 6 - TCP Session Control Messages

Request	LS_COMMAND	Short Name	iFCP Support
Connection Bind	0xE0	CBIND	REQUIRED
Unbind Connection	0xE4	UNBIND	REQUIRED
Test Connection Liveness	0xE5	LTEST	REQUIRED

[T/E] How do we know that those three values (E0, E4, and E5) will not conflict with some future usage by Fibre Channel? I think the answer is that SES=1 in the iFCP flags in the header, and would be 0 in any future use of these values in an ELS, but the use of those three values looks like an attempt to avoid conflict and should be explained.

Accepted

That is correct. These values were chosen as patterns readily distinguishable by a protocol analyzer.

Comment 28. [T] 6.2 - Unbind Connection (UNBIND)

"Unbind Status Description
------------ -----------

0	Successful û No other status
1 - 15	Reserved
16	Failed - Unspecified Reason
18	Failed - Connection ID Invalid
Others	Reserved

"Unbind can fail, but earlier specification of the use of Unbind (e.g., in Section 5.2.3.1) assumes that it cannot fail."

Description of how to deal with Failed status needs to be added there (e.g., close the TCP connection).

Accepted

Comment 29. [E] Section 7.2, pp 56, para 7

"For translation type 3, the receiving gateway SHALL obtain the information needed to fill in the field in the link service frame payload by converting the specified N_PORT worldwide identifier to a gateway IP address and N_PORT ID. This information MUST be obtained through an iSNS name server query."

This requires an iSNS query for every type 3 translation received even if it exists locally in a Remote N_PORT descriptor. It looks like this was intended due to the possibility of the descriptor being stale, but I wanted to check if that was in fact the intention.

Accepted

The intention was to update a potentially stale entry or force the creation of a new descriptor.

Comment 30. [E] Section 7.2, pp 57, para 3

"When the ACC response requires iFCP intervention, the receiving gateway MUST act as a proxy for the originator, retaining the state needed to process the response from the N_PORT to which the request was directed."

That doesn't parse for me. I think the intended meaning was that when an ELS request is sent whose ACC will require iFCP intervention, the ELS also requires intervention to capture the state necessary to process the ACC.

Accepted

The text will be modified as follows:

"When the ACC response requires iFCP intervention, the receiving gateway MUST intervene to process the response from the N_PORT to which the request was directed."

Comment 31. [T] 7.3 - Fibre Channel Link Services Processed by iFCP

"The following Extended and FC-4 Link Service Messages must receive special processing."

Process question - how does this list get coordinated with T11 so that it gets updated when T11 defines a new ELS or FC-4 LS that requires iFCP intervention?

Response

The specification must be revised to track the evolving fibre channel specifications, including, among other things, the addition of new link services that require special processing.

Comment 32. [T] 7.3.1.1 - Abort Exchange (ABTX)

"Fields Requiring Address Translation	Translation Type (see section 7.2)	Supplemental Data (type 3 only)
Exchange Originator S_ID"	1, 2	N/A

Need to specify how to choose the translation type. This comment also applies to RES, RES ACC, RLS, RSS, RRQ, RSI, REC and REC ACC. It may be best resolved by adding additional text in Section 7.2.

Accepted

Comment 33. [E] 7.3.1.3 - Discover Address Accept (ADISC ACC)

Should the Command field be 0x20 or 0x02?

Response

The command field for all ACC response frames is 0x02. No change to the specification is required.

Comment 34. [T] 7.3.1.3 - Discover Address Accept (ADISC ACC)

"Other Special Processing:

The Hard Address of the ELS originator SHALL be set to 0."

Doesn't this also require setting the LS_COMMAND iFCP-specific field (to be renamed) in the FC Common Encapsulation header? This comment also applies to all other ACC's in Section 7.

Accepted

The specification will be modified accordingly.

Comment 35. Section 8.2.1 - Enforcing R_A_TOV Limits

The rules in this section appear to allow forwarding of all frames received while in Unsynchronized mode or with a timestamp of 0,0. This looks like formula for violating R_A_TOV - was this intended?

Response

The intention was to abort all iFCP sessions and not allow the creation of new ones. The specification will be revised accordingly.

Comment 36. [T] Section 9.4.1 - Establishing the Broadcast Configuration

"The broadcast configuration is managed using facilities provided by the iSNS server. Specifically:

a) "An iSNS discovery domain is created and seeded with the network address of the global broadcast server N_PORT. The global server is identified as such by setting the appropriate N_PORT entity attribute."

There are no means for recovery from failure, so loss of the gateway performing the broadcast service results in loss of the broadcast service. This needs to be explained at a minimum, and probably corrected.

Accepted

An implementation may designate a local server as a standby global broadcast server. The local server uses the LTEST message to determine if the global server is functioning and may assume control if not.

The specification will be revised accordingly.

Comment 37. [T] Section 10.2.2, page 82, para 1

"Conformant implementations of the iFCP protocol MAY use such security definitions."

I don't understand this sentence. What was intended?

Accepted

The paragraph will be changed to:

“It is imperative to thwart these attacks, given that an iFCP gateway is the last line of defense for a whole fibre channel island, which may include several hosts and fibre channel switches. To do so, the iFCP gateway must implement and may use confidentiality, data origin authentication, integrity, and replay protection on a per-datagram basis. The iFCP gateway must implement and may use bi-directional authentication of the communication endpoints. Finally, it must implement and may use a scalable approach to key management.”

Comment 38. [T] Section 10.2.3, pp 82, para 1

"Enterprise data center networks are considered mission-critical facilities that must be isolated and protected from all possible security threats. Such networks are usually protected by security gateways, which at a minimum provide a shield against denial of service attacks. The iFCP security architecture is capable of leveraging the protective services of the existing security infrastructure, including firewall protection, NAT and NAPT services, and IPSec VPN services available on existing security gateways."

While this is true of iFCP, iSNS has some serious issues with NAT and NAPT and iFCP cannot be operated without iSNS.

Rejected

iSNS issues with NA(P)Ts are thought to be resolved (see Section 3.6 in the iSNS specification). iSNS has at least two non-exclusive options to cope with NA(P)Ts, a) the use of FQDNs instead of IP addresses, and b) the option to establish a confederation of iSNS servers and have them doctor IP numbers in transit as part of their mutual confederation contract.

Comment 39. [T] Section 10.2.4, pp 82, para 1

"iFCP gateways MUST use Discovery Domain information obtained from the iSNS server [ISNS] to determine whether the initiating fibre channel N_PORT should be allowed access to the target N_PORT. N_PORT identities used in the Port Login (PLOGI) process shall be considered authenticated provided the PLOGI request is received from the remote gateway over a secure, IPSec-protected connection."

Need to say something about the IKE identities (ID payloads) used for the authentication, and how they correspond to information obtained from iSNS - NATs/NAPTs will cause issues here. Just requiring an IPsec-protected connection isn't good enough as it may allow a node not registered with iSNS to get in.

Accepted in part

It would be premature to enumerate ID payloads in section 10.2.4, which describes the scope of the overall security design prior to any IKE/IPsec requirement (to follow in sections 10.3). The requested information will be supplied after the last paragraph in section 10.3.1.

Regarding intervening NA(P)Ts between iSNS clients and servers, it is possible to put a proxy iSNS server at the boundary between addressing domains. Such proxy will terminate the IKE/IPSec so that the ID_IPV4_ADDR identity can be used natively by IKE. It is also possible to use the second method described in the response to comment 38 -- a confederation of iSNS servers where the NAT(P)T mediation now occurs between iSNS servers.

Admission control is performed by the iSNS server, based upon the Discovery Domain (DD) configuration information stored in that iSNS server. Once the authenticity of a gateway is verified (e.g., via a pre-shared key) and IPsec SAs are established, then the gateway is trusted to behave according to the specification, which mandates a handshake with iSNS for admission.

Comment 40. [E] Section 10.2.6 Rekeying

I believe the security draft has changed in this area (small rekeying interval example), please check it.

Response

We appear to be consistent with [SECIPS] version 11 still, end of section 5.4, when Bellare’s results are taken into consideration. Therefore, no change to iFCP is required.

Comment 41. [T] Section 10.2.7 Authorization

"Authorization is outside of the scope of this specification, and is seen as fully orthogonal to the iFCP security design. Such design, however, includes key authorization-enabling features in the form of Identity Payload (e.g., ID_FQDN), certificate-based authentication (e.g., with X509v3 certificates), and discovery domains [ISNS]."

What?? If iSNS doesn't know about an iFCP gateway, that gateway shouldn't be able to talk to any other iFCP gateway. That's access control, which counts as authorization in my book.

Accept

The paragraph will be re-written as follows.

“Basic access control properties stem from the requirement that the communicating iFCP gateways be known to one or more iSNS servers before they can engage in iFCP exchanges. The optional use of Identity Payloads (e.g., ID_FQDNs), certificate-based authentication (e.g., with X509v3 certificates), and discovery domains [ISNS] enables authorization schemas of increasing complexity. The definition of such schemas (e.g., role-based access control) is outside of the scope of this specification."

Comment 42. [E] Section 10.3.2, pp 86, para 8

If an iFCP implementation makes use of unbound TCP connections, and such connections belong to an iFCP Portal with security requirements, then the unbound connections MUST be protected by an SA at all times just like bounded connections.

Change "bounded" to "bound".

Accepted

Comment 43. [T] Section 10.3.2, pp 86, para 9

"Upon receiving an IKE Phase-2 delete message, there is no requirement to terminate the protected TCP connections or delete the associated IKE Phase-1 SA. Since an IKE Phase-2 SA may be associated with multiple TCP connections, terminating such connections might in fact be inappropriate and untimely. An iFCP Portal must instead attempt to create a new Phase-2 SA while there are outstanding iFCP sessions."

That's a problem. If the other side is behaving in accordance with the next paragraph ...:

"To minimize the number of active Phase-2 SAs, IKE Phase-2 delete messages may be sent for Phase-2 SAs whose TCP connections have not handled data traffic for a while. To minimize the use of SA resources while the associated TCP

connections are idle, creation of a new SA may be deferred until new data is to be sent over the connections."

... and is deleting the Phase-2 SAs because it lacks the resources to support them, immediately creating a new Phase-2 SA in response to delete messages risks livelock (massive churn in Phase-2 SA creation/destruction). Creating a new Phase-2 SA in response to a Phase-2 delete message SHOULD be deferred until there is traffic to send over that SA.

Accepted

We shall be removing the misleading sentence “An iFCP Portal must instead attempt to create a new Phase-2 SA while there are outstanding iFCP sessions." and promote from may to SHOULD prior to the word ‘deferred’.

The resulting modified text is shown below:

"Upon receiving an IKE Phase-2 delete message, there is no requirement to terminate the protected TCP connections or delete the associated IKE Phase-1 SA. Since an IKE Phase-2 SA may be associated with multiple TCP connections, terminating such connections might in fact be inappropriate and untimely.

"To minimize the number of active Phase-2 SAs, IKE Phase-2 delete messages may be sent for Phase-2 SAs whose TCP connections have not handled data traffic for a while. To minimize the use of SA resources while the associated TCP connections are idle, creation of a new SA SHOULD be deferred until new data is to be sent over the connections."

Comment 44. [E] Section 13. - Normative References

RFC 2451 reference shows up twice.

Accepted

Comment 45. [T] Section A.2 - Link Services Processed Transparently

"ACC Accept"

Is that right? I thought this was intercepted in some cases, as indicated in Table 6.

Response

The ACC description will be modified to discriminate between the transparent and non-transparent cases.

Comment 46. [T] Section A.2 - Link Services Processed Transparently

FDISC Discover F_Port Service Parameters
FLOGI F_Port Login
RTV Read Timeout Value

Definitely wrong - the iFCP gateway has to implement these itself as specified in Section 9.1.

Accepted in Part

Special link service messages are those which require intervention by an iFCP protocol implementation before they are passed to the destination N_PORT. Transparent link service messages are passed to the destination N_PORT without such intervention. In that regard, the above link services are processed transparently.

The specification will be modified to make the above distinction clearer and the section will be re-titled as: "Link Services Processed Transparently by the iFCP layer".

Comment 47. [T] Section A.2 - Link Services Processed Transparently

LINIT Loop Initialize
LPC Loop Port Control
LSTS Loop Status
SCL Scan Remote Loop

I don't have time to check these, but I'm suspicious about whether anything that has "Loop" as part of its name can/should be forwarded transparently into an FC fabric, although SCL seems plausible. Please verify whether these are transparent.

Response

SCL must be processed as a special link service message. iFCP will be modified accordingly. The remaining link services listed above are transparent.

Comment 48. [T] Section A.2 - Link Services Processed Transparently

RSCN Registered State Change Notification
SCN State Change Notification
SCR State Change Registration

Those can't be transparent, as Section 9.2 requires the iFCP gateway to implement them.

Response

See response to Comment 46.

3. Comments From Elizabeth Rodriguez

Comment 49. [E] Title Page, Number of Authors

Looks like you have 8 authors listed. Rule of thumb I think is 6. I am having difficulty locating the guidelines, but may want to consider how you can move a couple of the listed authors into an acknowledgements section of some sort. With 8, it may or may not get flagged by the IESG...

Accepted

We will reduce the roster of co-authors in accordance with IETF policy.

Comment 50. [E] Capitalize Fibre Channel

I believe "Fibre Channel" should be capitalized throughout document.

Rejected

The specification is consistent with T11 lower case usage.

Comment 51. [E] Acknowledgements

Braces around SECIPS do not match.

Accepted

Comment 52. [E] Section 1.2

"NCITS" should be "INCITS".

Accepted

Comment 53. [E] "About This Document"

There should be a page break before this section.

Accepted

Comment 54. [E] Definitions, iFCP Frame

Technically, the title is of the comment encapsulation specification is "FC Frame Encapsulation".

Accepted

Comment 55. [E] Definitions

In the definitions of "N_PORT Alias" and "N_PORT I/D", two dashes should be used to separate the term from the body of the definition.

Accepted

Comment 56. [E] Section 3, pp 9, para 1

"Fibre channel is a frame-based, serial technology designed for peer-to-peer communication between devices at gigabit speeds and with low overhead and latency."

May want to change to gigabit or greater speeds. Technically, 2, 4, 10 gigabit speeds are still gigabit, but many today interpret gigabit strictly as 1 gigabit.

Rejected

The term 'speeds' implies rates of 1Gb/sec and above.

Comment 57. [E] Section 3.1

a) "N_PORTs -- The end points for fibre channel traffic. In the FC standards, N_PORT interfaces have several variants, depending on the topology of the fabric to which they are attached. As used in this specification, the term applies to any one of the variants."

Suggestion -- sometimes referred to in literature as Nx_PORTs?

Rejected

A parenthetical Nx_PORT digression does not add any value to the iFCP specification, given that the following statement claims that N_PORT is used for any such variants.

Comment 58. [E] Section 3.2, Fabric Topologies

a) "Arbitrated Loop -- A series of N_PORTs connected together in daisy-chain fashion. Data transmission between N_PORTs requires arbitration for control of the loop in a manner similar to a token ring network."

Accepted in part

Rewrite as:

a) “Arbitrated Loop -- A series of N_PORTs connected together in daisy-chain fashion. Loop-connected N_PORTS are referred to as NL_PORTS. Data transmission between NL_PORTS requires...”

Comment 59. [E] Section 3.3, pp 13, para 6

"FC-4 –- Application protocols, such as FCP, the fibre channel SCSI protocol."

Reword to read: "...such as FCP, commonly used abbreviation for "Fibre Channel Protocol for SCSI"

Accepted in part

The sentence will be revised to read: "...such as the fibre channel protocol for SCSI (FCP)."

Comment 60. [E] Section 3.7, pp 16, par 2

"The source and destination N_PORT fabric addresses embedded in the S_ID and D_ID fields represent the physical MAC addresses of originating and receiving N_PORTs."

"I think the term MAC is inappropriate here -- MAC is really an ethernet term. Something like physical world wide unique address, similar to an ethernet MAC address... Or ... represent the physical MAC like address...

Accepted

The text will be changed to:

"The source and destination N_PORT fabric addresses embedded in the S_ID and D_ID fields represent the physical addresses of the originating and receiving N_PORTs."

Comment 61. [E] Section 3.8, Fibre Channel Transport Services

Does class 6 still exist, or has it been made obsolete?

Response

Class 6 is still specified in [FC-FS].

Comment 62. [E] Section 4.5, pp 24, para 5

"The mode of gateway operation is settable in an implementation-specific manner. The implementation MUST NOT allow the mode to be changed after the gateway begins processing fibre channel frame traffic."

Does this need to be qualified -- e.g. MUST NOT allow the mode to be changed after the gateway begins processing Fibre Channel traffic without first terminating all connections to that gateway, or some such -- in other words, really someone can change the mode of operation, but just cannot do so while the gateway is in use.

Rejected

The specification will not be changed. The intent is to latch the operational mode after gateway power is turned on and the gateway begins handling FC frame traffic. A change in operational mode is not intended to be easy or graceful.

Comment 63. [E] Section 5.2.2.1, pp 31, para 9

"When creating a descriptor in response to an incoming CBIND request, the iFCP gateway SHALL perform an iSNS name server query using the worldwide port name of the remote N_PORT in the SOURCE N_PORT NAME field within the CBIND payload. The descriptor SHALL be filled in using the query results."

Need to make sure that iSNS gets through WG last call soon as well, since this is a normative dependency.

Accepted

Comment 64. [E] Section 5.4, pp 39, para 1

"This section describes the iFCP encapsulation of fibre channel frames. The encapsulation is based on the common encapsulation format defined in [ENCAP]."

The reference to "common encapsulation" should be "FC Frame Encapsulation".

Rejected

The reference is appropriate.

Comment 65. [E] Section 6.2, Unbind Connection (UNBIND)

It should be noted that the Unbind status codes listed in this section are decimal values.

Accepted

The rules for numeric representation will be added to the "Conventions" section.

Comment 66. [E] Section 7, pp 53

a) "Transparent – The link service message and reply MUST be transported to the receiving N_PORT by the iFCP gateway without altering the message payload. The link service message and reply are not processed by the iFCP implementation."

Since iFCP has Transparent and Translation modes, use of the term transparent here might get confusing -- Transparent is

referring to the fact that the link service must be propogated across the IP network, correct? As opposed to a link service that is applicable only to transparent mode...

Accepted

The term "transparent" in this context will be changed to "pass-though".

4. Comments from Brian Forbes

Comment 67. [E] Section 2.1, Special Characters

For some reason the file contains a number of occurrences of the character <funny character> instead of a hyphen or dash. Occurs throughout the text.

Accepted

Comment 68. [E] Section 2.1, Definitions

"iFCP Session - An association created when an N_PORT sends a PLOGI request to a remotely attached N_PORT. It is comprised of the N_PORTs and TCP connection that carries traffic between them."

Grammar: “it is comprised of” should be “it comprises”.

Accepted

Comment 69. [E] Section 2.1, Definitions

"N_PORT Alias -- The N_PORT address assigned by a gateway to represent a remote N_PORT accessed via the iFCP protocol. When routing frame traffic in address translation mode, the gateway automatically converts N_PORT aliases to N_PORT network addresses and vice versa."

Consistency: in the list of 2.1 definitions, some entries use a double hyphen and others only a single one (which at least one reader interpreted as a change in level)

Accepted

Comment 70. [E] Section 3.1, The Fibre Channel Network

"Unlike a layered network architecture, a fibre channel network is largely..."

Remove the extra space after the comma.

Accepted

Comment 71. [E] Section 3.3.1, Fabric Supplied Link Services

"Time Server –- Intended for the management of fabric-wide expiration timers or elapsed time values and is not intended for precise time synchronization"

Parallel structure: “and is not intended” seems to read better as “and not intended”

Accepted

Text will be changed to read:

"Time Server –- Intended for the management of fabric-wide expiration timers or elapsed time values and not intended for precise time synchronization"

Comment 72. [E] Section 3.7.1, page 6, para 3

"...The value of the Domain I/D ranges from 1 to 239 (0xEF)."

Both “ID” and “I/D” are used to mean “identifier” within the same paragraph. Common usage suggests “ID” throughout. Also occurs elsewhere, e.g. page 21

Accepted

Comment 73. [E] Section 3.7.1, page 17, para 3

For some reason the file contains a number of occurrences of a non-ascii character instead of an apostrophe. Also occurs on pages 64 for example.

Accepted

Comment 74. [E] Section 3.7.1, page 17, para 4

“FLOGI”: this is the first occurrence of this FC term; it should be spelled out here or a forward reference could be provided

Accepted

Comment 75. [E] Section 3.9, page 18. item a)

a) "Fabric Login (FLOGI) -- An operation whereby the N_PORT registers its presence on the fabric, obtains fabric parameters, such as classes of service supported, and receives its N_PORT address,"

Reads better without the comma after “fabric parameters”.

Accepted

Comment 76. [E] Section 4, page 18, para 3

"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 [FC-GS3]."

“devices” is possibly ambiguous here, could say “FC devices” or iFCP devices” depending on the intent.

Accepted

Text will be changed to:

"Within the fibre channel device domain, fabric-addressable entities consist of other N_PORTs and fibre channel devices internal to the fabric that perform the fabric services defined in [FC-GS3]."

Comment 77. [E] Section 4.6.1, Page 25, Para 1

"As described in section 4.6, each gateway and fibre channel switch in a bounded iFCP fabric MUST have a unique domain I/D. In a gateway region containing fibre channel switch elements, each element obtains a domain I/D by querying the principal switch as described in [FC-SW2] -- in this case the iFCP gateway itself. The gateway in turn MUST obtain domain I/Ds on demand from the iSNS name server acting as the central address allocation authority. In effect, the iSNS server assumes the role of principal switch for the bounded fabric. In that case, the iSNS database contains:"

The fact that a gateway can act as the FC principal switch is mentioned in this section and others, but there seems to be no normative text determining when it must do so. This will be obvious to a knowledgeable reader, or perhaps is covered in an ancillary document, but given the care taken elsewhere to provide normative language for ‘obvious’ functionality it seems to be an oversight

Rejected

Since the paragraph is intended to describe behavior that is normatively specified elsewhere, the use of "MUST" is incorrect. The text will be changed to the following:

"As described in section 4.6, each gateway and fibre channel switch in a bounded iFCP fabric has a unique domain I/D. In a gateway region containing fibre channel switch elements, each element obtains a domain I/D by querying the principal switch as described in [FC-SW2] -- in this case the iFCP gateway itself. The gateway in turn obtains domain I/Ds on demand from

the iSNS name server acting as the central address allocation authority. In effect, the iSNS server assumes the role of principal switch for the bounded fabric. In that case, the iSNS database contains..."

Comment 78. [E] Section 5.2.2.3, page 35, paras 3 and 4

These two paragraphs use the terms ‘heartbeat’ and ‘connectivity probe’ as informal synonyms for LTESTs. Use of the same synonym in both places would keep the reader from wondering whether the two synonyms represent the same concept.

Accepted

Comment 79. [E] Section 5.2.2.4.3, page 37, para 1

"Window scaling, as specified in [RFC1323], allows full utilization of links with large bandwidth - delay products and should be supported by an iFCP implementation."

Is “should” intended to be normative (capitalized)?

Response

The lower case usage is intentional. The goal is to reflect a desirable bias rather than the sort of mandate defined in [RFC2119].

Comment 80. [E] Section 5.2,3, page 32, items c) and d)

a) "For an FC frame received from the IP network, a gateway detects a CRC error in the encapsulation header. The gateway shall abort the session as described in section....

b) "The TCP connection associated with the login session fails for any reason. The gateway detecting the failed connection shall abort the session as described in section...."

“shall” should be capitalized.

Accepted

Comment 81. [E] Section 5.4, page 39, last paragraph

"When operating in Address Translation mode, (see section ...) the iFCP gateway must recalculate the fibre channel CRC."

“must” should be in caps.

Accepted

Comment 82. [E] Section 5.4, page 41, "TRN" mnemonic

It's unfortunate that "TRN" can be read as either "transparent" or "translation" and therefore has less mnemonic value.

Accepted

"TRN", the mnemonic for "transparent mode", will be changed to "TRP".

Comment 83. [E] Section 6.2, page 49, para 1

"UNBIND is used to release a bound TCP connection and. optionally, return it to the pool of unbound TCP connections."

Punctuation: "and. optionally, " should be "and optionally".

Accepted

Comment 84. [E] Section 7.3, page 58, para 2

"The formats of each special link service message, including supplemental data where applicable, are shown in the following sections."

"The formats of each… message are" is awkward, suggest "The format of each… message is".

Accepted

Comment 85. [E] Section 7.3.1.6, page 63, para 2

"This ELS shall always be sent as an augmented ELS regardless of the translation mode in effect."

"Shall" should be capitalized.

Accepted

The text must also be modified to replace "augmented" with "special", as given below.

"This ELS SHALL always be sent as a special ELS regardless of the translation mode in effect."

[E] Section 7.3.1.14, page 71, last paragraph

"The size of each frame to be sent to the destination N_PORT MUST NOT exceed the maximum frame size that the destination N_PORT can accept. The sequence identifier in each frame header SHALL be copied from the augmented ELS and the sequence count shall be monotonically increasing."

"Shall" should be capitalized.

Accepted

Comment 86. [E] Section 10.2.4, page 82, paras 1 and 2

"iFCP is a peer-to-peer protocol. iFCP sessions may be initiated by either or both peer gateways. Consequently, bi-directional authentication of peer gateways MUST be provided.

"iFCP gateways MUST use Discovery Domain information obtained from the iSNS server [ISNS] to determine whether the initiating fibre channel N_PORT should be allowed access to the target N_PORT. N_PORT identities used in the Port Login (PLOGI) process shall be considered authenticated provided the PLOGI request is received from the remote gateway over a secure, IPSec-protected connection."

These paragraphs seem to be statements of required functionality but are too general to use normative language (“MUST”). Later sections contain the normative text necessary to cover these topics.

Accepted

Comment 87. [E] Section 10.2.5, page 82, para 1

See Comment 86.

Accepted

Comment 88. [E] Section 10.2.6, pages 82 and 82, all paragraphs

See Comment 86.

Accepted

5. Comments from Mallikarjun Chadalapaka

Comment 89. [E] Section 1.2

"...standards controlled by NCITS T10 and T11."

"NCITS" should be "INCITS".

Accepted

Comment 90. [E] 2.1 Definitions

"Gateway Region -- The portion of an iFCP fabric accessed through an iFCP gateway. Fibre channel devices in the region

consist of all fibre channel devices locally attached to the gateway."

The first sentence here when interpreted wrt a Nx_port sitting within a given gateway region, implies something that isn't right - viz. the rest of the iFCP fabric. The second sentence makes the intention clear, if "locally attached" includes the entire local fabric. My suggestion would be: "The portion of an iFCP fabric that accesses the rest of the fabric through one iFCP gateway."

Accepted

The definition will be changed to the following:

"Gateway Region -- The portion of an iFCP fabric accessed through an iFCP gateway by a remotely attached N_PORT. Fibre channel devices in the region consist of all those locally attached to the gateway."

Comment 91. [T] Section 3.3.1, pp 14, para 7

"Time Server -- Intended for the management of fabric-wide expiration timers or elapsed time values and is not intended for precise time synchronization."

I am curious about this - is it the conclusion the iFCP authors reached? The reason I ask is that IIRC, FCIP allows using this for time sync.

Response

See Comment 71 for the proposed change to this section.

The characterization is found in the literature and based on the following from the [FC-GS3] specification, section 7, page 161.

"The Time Service is provided to serve time information that is sufficient for managing expiration time."

Comment 92. [T] Section 3.7, pp 14, para 2

"The source and destination N_PORT fabric addresses embedded in the S_ID and D_ID fields represent the physical MAC addresses of originating and receiving N_PORTs."

I am not sure that it is a correct analogy....S_ID and D_ID are actually (potentially transient) addresses assigned by the fabric, Port Names are more akin to the MAC addresses.

Accepted

See Comment 60.

Comment 93. [E] 4. The iFCP Network Model

"The iFCP protocol enables the implementation of fibre channel mixed or switched fabric functionality on an IP network."

I am not sure what is intended by "fibre channel mixed or switched" here, perhaps this could use rewording.

Accepted

The text will be changed to:

"The iFCP protocol enables the implementation of fibre channel fabric functionality on an IP network."

Comment 94. [E] Section 4, pp 20, para 1

"Each iFCP gateway contains two standards-compliant fibre channel ports and an iFCP Portal for attachment to the IP network."

Why are two FC ports required? As far as I can tell, even one E_Port works just as well - is it to be technically called as a "switch"?

Also, is there a reason for limiting to only one IP address (implied by one iFCP Portal)? I see that supporting multiple iFCP Portals would require enahancements to the data structures presented - but can you please comment on any architectural issues here?

Response

The specification will be revised to emphasize that the figure is but one example of a supported implementation. It was intended to parallel the earlier fibre channel fabric example as a way of showing the transition to an equivalent iFCP fabric.

The selected example was chosen because it was easier to depict within the constraints of ASCII text. An E_PORT example could have also been used. In either case, the device incorporating iFCP portal functionality would be called an "iFCP gateway".

The considerations to be addressed when connecting multiple iFCP portals to a gateway region are discussed in Comment 12.

Comment 95. [E] section 4, pp 20, para 2

"... channel switch element. At this interface, remote N_PORTs are presented as fabric-attached devices. Conversely, on the IP

network side, the gateway presents each locally connected N_PORT as a logical fibre channel device."

I am not sure the last sentence is correct - I think "logical fibre channel device" should probably be replaced by "a TCP connection".

Rejected

The logical fibre channel device represents the layer 4 abstraction visible on the IP network.

Comment 96. [E] Section 4.1, pp 20, para 1

"...cases, the gateway may support any standards-compliant fibre channel fabric type by incorporating the functionality required to..."

Can you please comment if really "fabric type" is meant here? Or, is it the "fabric port type"?

Response

More accurately, "fabric type" should be changed to "fibre channel network topology." The specification will be changed accordingly. See Comment 5.

Comment 97. [E] Section 4.1, pp 20, para 1

"...present locally attached N_PORTs as logical iFCP devices."

It may be useful to define "iFCP device" in section 2.1.

Rejected

From section 2.1:

"Logical iFCP Device - The abstraction representing a single fibre channel device as it appears on an iFCP network."

The specification will not be changed.

Comment 98. [T] Section 4.4.1, pp 22, para 2

"...messages, a gateway cannot convert such addresses in the payload of vendor- or user-specific fibre channel frame traffic."

Not being very familiar with today's FC, can you please comment if these proprietary versions of frame formats (with even the D_ID out of place) are legal on regular fabrics? Seems like the entire fabric should be capable of special handling these...

Response

There is one and only one acceptable format for FC frames. That said, the issue is not the frame format but the payload contents.

Besides the addresses in the FC frame header, an iFCP implementation is only cognizant of N_PORT addresses that may be embedded in the payload of standards-compliant link service messages. It cannot remap such addresses if present in the payloads of user-specified or vendor-specific frames.

No change to the specification will be made.

Comment 99. [T] Section 4.4.3, pp 22, para 1

"In an unbounded iFCP fabric, limiting the scope of an N_PORT address to a gateway region reduces the likelihood that reassignment of domain I/Ds caused by a disruption in one gateway region will cascade to others."

"In an unbounded iFCP fabric, limiting the scope of an N_PORT address to a gateway region reduces the likelihood that "

Does it not prevent the likelihood?

Accepted

The text will be changed to:

"In an unbounded iFCP fabric, limiting the scope of an N_PORT address to a gateway region prevents reassignment of domain I/Ds caused when a disruption in one gateway region cascades to others."

Comment 100. [T] Section 4.4.3, pp 22, para 2

"In addition, a bounded iFCP fabric has an increased dependency on..."

Suggest changing "In addition" to "On the other hand".

Accepted

Comment 101. [E] Section 4.4.3, pp 22, para 3

"Finally, adding a gateway to a bounded fabric is more likely to disrupt the operation of all devices in the gateway region along with those already in the fabric as new, fabric-wide N_PORT addresses are assigned."

Isn't the issue in this para the same as that in the first para, albeit from the bounded fabric's perspective? If so, suggest merging them.

Rejected

Adding a new gateway region is distinct from disrupting an existing region and therefore merits its own mention.

Comment 102. [E] Section 4.4.3, pp 23, para 4

...be done non-disruptively and requires only that new gateway's iSNS..."

Change "that" to "that the".

Accepted

Comment 103. [T] Section 4.5, The iFCP N_PORT Address Model

b) "A 24-bit N_PORT alias. A fibre channel N_PORT address assigned by a gateway operating in address translation mode to identify a remotely attached N_PORT. Frame traffic is directed to a remotely attached N_PORT by means of the N_PORT alias."

At any point in time, there can only be 2^24 N_PORTs communicating even in the address translation mode, even though this mode allows the same N_PORT to be mapped to different nports in different gateway regions at different times. If this is a correct interpretation, I suggest that this be made clear in section 4.4.2, which currently simply states that there are no architectural limitations on the number of fibre channel devices in this mode.

Accepted in Part

While the addressability in a given gateway region is constrained by the fibre channel address model, the aggregate addressability of all gateway regions comprising an unbounded iFCP fabric can exceed that limit.

To make this clearer, the text will be changed as follows:

b) "Since N_PORT fibre channel addresses in an unbounded iFCP fabric are not fabric-wide, the number of iFCP gateways, fibre channel devices and switch elements that may be internetworked may exceed the fibre channel fabric limits."

Comment 104. [E] Section 4.6.1, pp 25, para 4

"In its role as principal switch within the gateway region, an iFCP..."

General comment - Change to "...as the Principal Switch...".

Accepted

Comment 105. [T] Section 5.2.1, pp 30, para 4

"...A gateway implementation MAY establish a pool of unbound connections to reduce the session setup time. Such pre-existing TCP connections between iFCP Portals remain unbound and uncommitted until allocated to an iFCP session through a CBIND message"

I wonder if there is a scope for DoS attack here with the possibility of one gateway potentially holding onto several unused TCP connections infinitely...."

Response

No. However, the specification will be modified to point out that a gateway may recover resources at any time by simply closing unbound connections. See Comment 21.

Comment 106. [E] Section 5.2.2.1, pp31, para 6

"If a descriptor does not exist, one SHALL be created in response to an iSNS name server query."

Did you mean "SHALL be created after the response to an iSNS name server query is received"?

Response

The test will be changed to:

"If a descriptor does not exist, one SHALL be created using the information returned by an iSNS name server query."

Comment 107. [E] Section 5.2.2.1.1, pp 31, para 1

"A Remote N_PORT descriptor SHALL only be updated as the result of an iSNS query that returns information for the specified worldwide port name. Following such an update, a new N_PORT alias SHALL NOT be assigned."

I assume you meant "iSNS response" instead of "iSNS query"?

I am a little confused by the SHALL NOT. Here's what I was assuming as the sequence of events

1. Local FC Name Server query.

2. iFCP gateway picks it up.

3. Consults with iSNS server

4. iSNS provides the remote N_PORT for the WWN

5. iFCP gateway assigns a local alias if in translation mode and if the remote N_PORT ID clashes with a pre-existing local NPORT_ID.

I do not see why this sequence should be prohibited. Comments will certainly help.

Accepted in Part

The text will be modified as described in Comment 108.

The 24-bit N_PORT component of the remote N_PORTs address and its local alias can never clash. The gateway transparently converts the alias to a network address, consisting of the TCP connection I/D, TCP Port number and the N_PORT ID assigned by the remote gateway.

Comment 108. [T] Section 5.2.2.1.1, pp 31, para 1

"A Remote N_PORT descriptor SHALL only be updated as the result of an iSNS query that returns information for the specified worldwide port name. Following such an update, a new N_PORT alias SHALL NOT be assigned.

"Until such an update occurs, the contents of a descriptor may become stale as the result of any event that invalidates or triggers a change in the N_PORT network address of the remote device, such as a fabric reconfiguration or the device's removal or replacement."

I assume that generally what is meant by "Until such an update occurs" is "In the absence of an operational iFCP session based on a descriptor". If so, it perhaps requires rewording.

Accepted in Part

Descriptors are only built and updated as a consequence of name server requests or state change notifications. An iFCP session may not necessarily be associated with these activities.

The text will be reworded as shown below to add the state change case and clarify the order of events leading to a stale descriptor.

"A Remote N_PORT descriptor SHALL only be updated as the result of an iSNS query to obtain information for the specified worldwide port name or from information returned by an iSNS state change notification. Following such an update, a new N_PORT alias SHALL NOT be assigned.

"Before such an update, the contents of a descriptor may have become stale as the result of any event that invalidates or triggers a change in the N_PORT network address of the remote device, such as a fabric reconfiguration or the device's removal or replacement."

Comment 109. [E] Section 5.2.2.1.1, pp 31, para 4

"Once the originating N_PORT learns of the reconfiguration, usually through the name server state change notification mechanism, the name server lookup needed to reestablish the iFCP session will automatically purge such stale data from the gateway."

Just a clarification here - it seems to me that the SCN for a remote N_PORT ID needs to delivered via the iFCP gateway anyway, so why not purge the stale mapping then (instead of waiting for the new SNS query from the local N_PORT?

Accepted

The text will be changed to:

"Once the originating N_PORT learns of the reconfiguration, usually through the name server state change notification mechanism, information returned in the notification or the subsequent name server lookup needed to reestablish the iFCP session will automatically purge such stale data from the gateway."

Comment 110. [T] Section 5.2.2.2, pp 33

f) "A CBIND response with a CBIND STATUS of "N_PORT session already exists" indicates that the remote gateway has concurrently..."

I think the document should specify that this status be mapped to the LS_RJT reason code of "Login/command already in progress" - 0x0E. Also, there may be N_PORTs that go down without issuing a LOGO, and attempt a PLOGI once they come back - unbeknownst to the iFCP gateway still with the old session descriptor. It isn't clear to me how this is proposed to be dealt with.

Rejected

As described in Comment 15, the specified behavior is meant to serve as a tie-breaking mechanism for the establishment of the iFCP session. Once the session is established, the PLOGIs from each side are sent and processed by the N_PORTs in accordance with the PLOGI semantics specified in [FC-FS].

A PLOGI after an iFCP session exists is handled in accordance with section 7.3.1.7, paragraph 5, which states:

"As specified in section 5.2.2.2, a PLOGI request addressed to a remotely attached N_PORT MUST cause the creation of an iFCP session if one does not exist. Otherwise, the PLOGI and PLOGI ACC payloads MUST be passed transparently to the destination N_PORT using the existing iFCP session."

Section 5.2.2.2 will be modified to describe the simultaneous PLOGI scenario above and the case of a PLOGI issued when an iFCP session exists.

Comment 111. [T] Section 5.2.2.3, pp 35

b) "An LTEST message is not received within twice the specified interval or the iFCP session has been quiescent for longer than twice the specified interval."

I think "or" above should be "and" - else it implies that the LTEST message must be received periodically even in the presence of other traffic.

Rejected

See Comment 18.

If liveness testing was requested for an iFCP session, an LTEST message must be received within twice the specified interval regardless of whether or not other traffic is present.

Comment 112. [T] Section 5.2.3, pp 37

a) "An LS_RJT response is returned to the gateway that issued the PLOGI ELS. The gateway SHALL forward the LS_RJT to the local N_PORT and complete the session as described in..."

My reading is that the gateway does not "issue" the PLOGI ELS, it merely facilitates the transport of an issued PLOGI ELS. The wording here is a little confusing - I also believe that the forwarding should be to the remote N_PORT, not local.

[Also,] I recommend "terminate"/"close" in all the places "complete" is used.

Accepted

The text will be modified as follows:

a) "An LS_RJT response is returned to the gateway from which the PLOGI ELS originated. That gateway SHALL forward the LS_RJT to the locally attached N_PORT and terminate the session as described in..."

Comment 113. [E] Section 5.2.3.1, pp 37, para 2

"Unbind session control ELS as described in section 6.2."

I am a little confused about the status of Unbind here - is it a FC-FS approved ELS or an iFCP session control message?

Response

Since Unbind is an iFCP session control message, the text will be changed to:

"Unbind session control message as described in section 6.2."

Comment 114. [T] Section 5.2.3.2, pp 38, para 4

"In any event, the TCP connection SHOULD be terminated with a connection reset (RST). If the local N_PORT has logged in to the remote N_PORT, the gateway SHALL send a LOGO to the local N_PORT."

I think the draft should specify both OPEN and OPEN PENDING cases here. For OPEN state, a local LOGO is required as stated, whereas for OPEN PENDING, a local LS_RJT may be appropriate.

Also, it is useful to state that the proxied ELS (in either case) be indistinguishable from the end-to-end ELS in its payload (so any sanity checking done by endnode software would continue to work).

Accepted

Comment 115. [T] Section 5.4.1, pp 40

"Protocol# IANA-assigned protocol number identifying the protocol using the encapsulation. For iFCP the value is (/TBD/)."

Should FCEncap document be referred here instead?

Accepted

See Comment 23.

Comment 116. [E] Section 5.4.3, pp 43, para 2

"Following frame validation, the S_ID and D_ID fields in the frame header SHALL be referenced to lookup the iFCP session descriptor (see section 5.2.2.2). If no iFCP session descriptor exists, the frame SHALL be discarded."

With the exception of PLOGI?

Accepted

The specification will be modified to address the case where a frame triggers the creation of an iFCP session.

Comment 117. [E] Section 5.4.3, pp 43, para 3

"Frames types submitted for encapsulation and forwarding on the IP..."

"Frames" should be "Frame".

Accepted

Comment 118. [E] Section 6, pp 44, para 1

"TCP session control messages are used to create and manage an iFCP session as described in section 5.2.2. They are passed between peer iFCP Portals and are only processed within the iFCP layer.

"The message format is based on the fibre channel extended link service message template shown below...."

It may be useful to state that this message forms the "FC Frame" payload. of the iFCP frame. It may also be useful to state the value of LS_COMMAND in the encap header (0?).

Instead of having two LS_COMMAND fields - one in the header and one in the payload - for these messages, a simpler approach could be to state that LS_COMMAND in the header contains an iFCP-defined value when SES=1 (and remove the one in the payload).

Accepted in Part

In accordance with Comment 26, the mnemonic for the LS_COMMAND field in the encapsulation header will be changed to eliminate confusion as follows:

From section 5.4, Encapsulation of Fibre Channel Frames:

"LS_COMMAND_ACC For a special link service ACC response to be processed by iFCP, the LS_COMMAND_ACC field SHALL contain bits 31 through 24 of the LS_COMMAND to which the ACC applies. Otherwise this field shall be set to zero."

From section 6, TCP Session Control Messages:

LS_COMMAND_ACC 0

With the addition of the new mnemonic, the above text clearly specifies how the field is to be set.

Comment 119. [E] Section 6.1, pp 46, para 2

"The following shows the format of the CBIND request."

I take it that this CBIND structure goes into the Session Control Message starting from word 6? Same question on CBIND response.

Rejected

That is correct. The existing text seems to explain this adequately.

Comment 120. [T] Section 6.2, pp 49, para 1

"UNBIND is used to release a bound TCP connection and. optionally, return it to the pool of unbound TCP connections."

I assume "release" here implies - "remove the binding"?

Is there a way to convey the preference to not terminate the connection on the part of the sender? IOW, where is the optionality selected?

Response

See Comment 21 regarding the disposition of "unbound" TCP connections. The above paragraph will be expanded to clarify the rationale for the unbound connection pool as follows:

"UNBIND is used to terminate an iFCP session and disassociate the TCP connection. To expedite the creation of a new iFCP session, the TCP connection MAY remain open at the discretion of either gateway and kept in a pool of unbound connections.

In order to recover resources, either gateway may spontaneously close the unbound TCP connection at any time."

Comment 121. [E] Section 6.2, pp 50, para 1

"transmitted in the connection that is to be unbound. The time..."

Change "in" to "on".

Accepted

Comment 122. [T] Section 8.2.1, pp 76, para 1

"The R_A_TOV limit on frame lifetimes SHALL be enforced by means of the time stamp in the encapsulation header (see section 5.4.1) as described in this section."

A couple of general questions on this section -

a) Is Unsynchronized operation allowed? If so, how is the R_A_TOV expectation met?

b) If an incorrect configuration causes the timestamp of the incoming frame to be higher than the gateway's time base, it is better if there is a way to detect and perhaps resync both ends with the same SNTP server (as opposed to one out of a list returned by iSNS). As far as I can tell, the current text specifies that it would simply cause the frames to be discarded, but doesn't break the binding nor terminate the TCP connection - perhaps relying on the end nodes to LOGOUT?

Accepted in Part

For item a), see the response to Comment 35.

For item b), iFCP specifies the following behavior:

d) "If the incoming frame has a non-zero time stamp, the receiving gateway SHALL compute the absolute value of the time in flight and SHALL compare it against the value of IP_TOV specified for the IP fabric.

e) "If the result in step (d) exceeds IP_TOV, the encapsulated frame shall be discarded. Otherwise, the frame shall be de-encapsulated as described in section ...."

Since it is impossible to guarantee that one time reference won't be skewed negatively with respect to the other. the propagation delay test is against the absolute value of the time difference.

The iFCP spec will be modified to state that an iFCP gateway implementation MAY terminate an iFCP session if the rate at which stale frames are detected exceeds some administratively-specified threshold.

6. Security Considerations

The applicable security provisions are defined in [IFCP].

7. References

[RFC2026] Bradner, S., "The Internet Standards Process -- Revision

3", BCP 9, RFC 2026, October 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997

[IFCP] Monia, C., et al, ” iFCP - A Protocol for Internet Fibre Channel Storage Networking", draft-ietf-ips-ifcp-10.txt, February 2002

[FC-FS] dpANS INCITS.XXX-200X, "Fibre Channel Framing and Signaling Interface", Revision 1.7, NCITS Project 1331-D, February 2002

[SECIPS] Aboba, B., et-al., "Securing Block Storage Protocols Over IP", revision 11, February 2002

[FC-GS3] dpANS X3.XXX-200X, "Fibre Channel Generic Services -3 (FC-GS3)", revision 7.01, NCITS Project 1356-D, November 2000

8. Author's Addresses

Charles Monia
Josh Tseng

Nishan Systems
3850 North First Street
San Jose, CA 95134
Phone: 408-519-3986
Email:
cmonia@nishansystems.com

Franco Travostino
Director, Content
Internetworking Lab,
Nortel Networks
3 Federal Street
Billerica, MA 01821
Phone: 978-288-7708
Email:
travos@nortelnetworks.com

Full Copyright Statement