

Voting Results on T10 Letter Ballot 01-300r0 on  
Forwarding SRP to first public review

Organization	Name	S Vote	Add'l	Info
Adaptec, Inc.	Ron Roberts	P	Yes	
Amphenol Interconnect	Michael Wingard	P	Yes	
Ancot Corp.	Bart Raudebaugh	P	Yes	
Andiamo Systems, Inc.	Claudio DeSanti	P	Yes	
BREA Technologies, Inc.	Bill Galloway	P	Yes	
Brocade Comm. Systems, Inc.	Brian Forbes	A	YesC	Cmnts
Cisco Systems, Inc.	David Peterson	P	Yes	
Compaq Computer Corp.	Robert C. Elliott	P	YesC	Cmnts
Congruent Software, Inc.	Peter Johansson	P	Yes	
Crossroads Systems, Inc.	Robert Griswold	P	Yes	
Dallas Semiconductor	Ti tkwan Hui	P	Yes	
Dell Computer Corp.	Kevin Marks	P	Yes	
EMC	Gary S. Robinson	P	Yes	
Emul ex			DNV	
ENDL Texas	Ralph O. Weber	P	Yes	
Exabyte Corp.	Joe Breher	P	Yes	
FCI	Douglas Wagner	P	Yes	
Fujitsu	Eugene Lew	P	Yes	
General Dynamics	Nathan Hastad	P	Yes	
Genroco, Inc.	Donald Woelz	P	Yes	
Hewlett Packard Co.	Mr. Randy Haagens	P	YesC	Cmnts
Hitachi Cable Manchester	Randy Wasylak	A	Yes	
IBM / Tivoli Systems	George O. Penokie	P	No	Cmnts
Intel Corp.	Cris Simpson	P	No	Cmnts
Iomega Corp.	Tim Bradshaw	P	Yes	
KnowledgeTek, Inc.	Dennis Moore	P	Yes	
LSI Logic Corp.	John Lohmeyer	P	Yes	
Maxtor Corp.	Mark Evans	P	Yes	
Microsoft Corp.			DNV	
Mol ex Inc.			DNV	
Nishan Systems Inc.	Charles Monia	P	Yes	
Ophi dian Desi gns	Edward A. Gardner	P	No	Cmnts
Panasonic Technologies, Inc	Terence J. Nelson	P	Yes	
Philips Electronics/CD Edge	William P. McFerrin	P	Yes	
Pirus Networks			DNV	
QLogic Corp.	Skip Jones	P	Yes	
Quantum Corp.	Paul Entzel	P	Yes	
Seagate Technology	Gerald Houl der	P	Yes	
Storage Technology Corp.	Erich Oetting	P	Yes	
Sun Microsystems, Inc.	Kenneth Moe	P	Yes	
Texas Instruments	Paul D. Aloisi	P	YesC	Cmnts
Toshiba America Elec. Comp.	Tasuku Kasebayashi	P	Yes	
Troi ka Networks, Inc.	William C. Terrell	P	YesC	Cmnts
TycoElectronics	Charles Brill	P	Yes	
Veritas Software			DNV	
Woven Electronics	Doug Pi per	P	Abs	Cmnts

Ballot totals: (37:3:1:5=46)

37 Yes

3 No

1 Abstain

5 Organization(s) did not vote

46 Total voting organizations

9 Ballot(s) included comments

This 2/3rds majority ballot passed.

37 Yes is at least a majority of the membership [greater than 23] AND

37 Yes is at least 27 (2/3rds of those voting, excluding abstentions [40])

Key:

P Voter is principal member

A Voter is alternate member

YesC Yes with comments vote

Abs Abstain vote

DNV Organization did not vote

Cmnts Comments were included with ballot  
 NoCmnts No comments were included with a vote that requires comments  
 DUP Duplicate ballot (last ballot received from org. is counted)  
 PSWD The password was not correct (vote not counted)  
 ORG? Organization is not voting member of T10 (vote not counted)

\*\*\*\*\*

Comments attached to YesC ballot from Mr. Brian Forbes of Brocade Comm. Systems, Inc.:

Brocade1 001 (E) Page: Many Locator: Many  
 Problem Description:  
 The word 'which' is used inappropriately in many places.  
 Suggested Solution:  
 Do a global search for the word which and replace it with one of the following corrections:  
 A) the word 'that'.  
 B) a new sentence construction that does not require the word.  
 C) nothing. (Which can simply be removed in many cases.)

Brocade1 002 (E) Page: 59 Locator: B.7, figure B.6  
 Problem Description:  
 The word 'must' is used inappropriately.  
 Suggested Solution:  
 The line 'At least one IB I/O controller must be present' should be replaced. I am not sure if this is a requirement that at one or more controllers shall be present. If so, wording like 'At least one IB I/O controller shall be present' is appropriate.

Brocade1 003 (E) Page: vii Locator: Foreword, line 3  
 Problem Description:  
 X3.269 is not the proper name  
 Suggested Solution:  
 This value is not correct and should be marked as TBD or XXX or something like that. In any case, it is an NCITS document, not an X3 document.

Brocade1 004 (E) Page: vii Locator: Foreword, line 8  
 Problem Description:  
 "'by National' s/b 'by the National'"  
 Suggested Solution:  
 Correct as requested.

Brocade1 005 (E) Page: viii Locator: Introduction, line 7  
 Problem Description:  
 "'The working draft SCSI' s/b 'The SCSI'"  
 Suggested Solution:  
 This correction should be made now, even though the document is still a working draft, because it is clearly labeled in lots of places that it is a draft, but the text in it is intended to be the content of the standard.

Brocade1 006 (E) Page: 1 Locator: Title, line 6  
 Problem Description:  
 "'The working draft SCSI' s/b 'The SCSI'"  
 Suggested Solution:  
 This correction should be made now, even though the document is still a working draft, because it is clearly labeled in lots of places that it is a draft, but the text in it is intended to be the content of the standard.

Brocade1 007 (E) Page: 3 Locator: 2.1, lines 32-35  
 Problem Description:  
 Global Engineering should be included here as well, since the drafts are not available from ANSI or NCITS.  
 Suggested Solution:  
 Include Global Engineering as a document source. Include www.t10.org as a document source for standards in development.

Brocade 0 001 (E) Page: Locator:  
 Problem Description:

The draft now seems to equate 'SRP target port' and 'IB service', so an SRP target port is designated by a ServiceID. This implies there can be many ports per IOC. This is a significant change from prior drafts where the target port was equated with an IOC, and there was just a single ServiceID per port. It requires a different model for software (OSs or whatever) to manage which hosts have access to which devices in a multi-host environment. Previously, access control was needed only to the level of IOCs, the draft now implies a need to manage not only who can use which IOCs, but which devices within an IOC.

Suggested Solution:

No solution required if interpretation is correct and implications are understood

Brocade 0 002 (E) Page: 52 Locator: B.3.1.2, lines 22-23

Problem Description:

The definition of 'IB channel adapter GUID' implies it is the Node GUID but doesn't say so; might as well be explicit

Suggested Solution:

'An IB Node GUID that uniquely identifies an IB channel adapter'

Brocade 0 003 (E) Page: 52 Locator: B.3.1.9, lines 39-40

Problem Description:

The definition of 'IB I/O controller GUID' implies it is the IOControllerProfile GUID but doesn't say so; might as well be explicit

Suggested Solution:

'An IB IOControllerProfile GUID that uniquely identifies an IB channel adapter'

Brocade 0 004 (E) Page: 54 Locator: Line 50

Problem Description:

IB GUIDs can have link-local scope and thus may not be 'globally' unique

Suggested Solution:

Change to 'unique within a subnet', or 'either unique within a subnet or globally unique'

Brocade 0 005 (E) Page: 55 Locator: Table B.1, lines 16-17

Problem Description:

IB GUIDs can have link-local scope and thus may not be unique 'worldwide'

Suggested Solution:

Change 'worldwide' to 'IB subnet or worldwide'

Brocade 0 006 (E) Page: 55 Locator: Figure B.3

Problem Description:

Figure B.3's equating of 'SRP Target Ports' with 'IB consumers' is problematic. A 'target port' is a sort of service access point---somewhere where interested parties initially go to obtain service, but without any implication that that's where the service is actually provided. (In IB, it's the Connection Manager that receives the initial connection request, interprets the ServiceID contained therein, and performs some magic that results in the instantiation of a QP bound to some entity that actually provides the target services). This target-services-providing entity fits the definition of 'IB consumer'. But the mapping of ServiceIDs-cum-SRP target ports onto such entities is clearly a matter of implementation, and could be one-to-many, many-to-one, or many-to-many

Suggested Solution:

One possibility: to the left of the IB Consumers show a table/list of service IDs within each IB I/O unit and label these entries as SRP Target Ports; use arrows to show a mapping from the entries to the IB Consumers, with e.g. one Consumer mapped to two IDs and another mapped to one ID to show that the mappings are not always 1 to 1. A further refinement might be to use another set of arrow between the Consumers and the QPs to show that the this mapping is also not 1 to 1

Brocade 0 007 (E) Page: 56 Locator: Line 16

Problem Description:

Missing word

Suggested Solution:

'used by the SRP initiator port'?

Brocade 0 008 (E) Page: 56 Locator: Line 21

Problem Description:  
Names of IB attributes are incomplete  
Suggested Solution:  
'IOUnitInfo, IOControllerProfile, and ServiceEntries'

Brocade 0 009 (E) Page: 56 Locator: Lines 48-49  
Problem Description:  
'I/O' is broken across lines (and pages)  
Suggested Solution:  
Make sure the slash in 'I/O' is non-breaking

Brocade 0 010 (E) Page: 56 Locator: Lines 48-49  
Problem Description:  
The phrase 'processor unit or IB I/O controller' makes an incorrect distinction; target ports can only be found on IB I/O controllers by definition, whether or not the I/O controller embodies a processor unit  
Suggested Solution:  
Omit 'processor unit or'

Brocade 0 011 (E) Page: 57 Locator: Lines 13-14  
Problem Description:  
'IB I/O controllers acting as SRP target ports' could be construed as a 1-to-1 correspondence between controllers and target ports  
Suggested Solution:  
'IB I/O controllers making SRP target ports available' or 'IB I/O controllers hosting SRP target ports'?

Brocade 0 012 (E) Page: 61 Locator: Line 1  
Problem Description:  
'An IB I/O controller acting as an SRP target ports' could be construed as a 1-to-1 correspondence between controllers and target ports  
Suggested Solution:  
'And IB I/O controller making SRP target ports available' or 'An IB I/O controller hosting SRP target ports'?

Brocade 0 013 (E) Page: 61 Locator: Lines 4-5  
Problem Description:  
'IB I/O controllers acting as SRP target ports' could be construed as a 1-to-1 correspondence between controllers and target ports  
Suggested Solution:  
'IB I/O controllers making SRP target ports available' or 'IB I/O controllers hosting SRP target ports'?

\*\*\*\*\*

Comments attached to YesC ballot from Mr. Robert C. Elliott of Compaq Computer Corp.:

CPQ #1 Page a (comment 1 on page)  
Title page  
Remove:  
American National Standard for Information Systems -  
and change "working draft SCSI RDMA Protocol" to "SCSI RDMA Protocol"

CPQ #2 Page a (comment 2 on page)  
General  
Update the PDF properties title and author

CPQ #3 Page ii (comment 1 on page)  
General  
Remove revision history, line numbers, change bars, etc. from final version

CPQ #4 Page 2 (comment 1 on page)  
Section 1  
Delete CAM from figure 1  
Delete these SCSI-2 standards from the example standards list:  
Serial Storage Architecture SCSI-2 Protocol SSA-S2P [ANSI X3.294:1996]  
Common Access Method:

SCSI Common Access Method CAM [ISO/IEC 9316-421]  
[ANSI X3.232:1996]

CPQ #5 Page 2 (comment 2 on page)  
Section 1  
Change Fiber to Fibre

CPQ #6 Page 5 (comment 1 on page)  
Add:  
3.1.8 autosense data: Sense data (see 3.1.49) that is returned in the SRP\_RSP IU payload. See SAM-2.  
3.1.49 sense data: Data returned to an application client as a result of an autosense operation, asynchronous event report, or REQUEST SENSE command. See SPC-2.

CPQ #7 Page 16 (comment 1 on page)  
Section 5.3  
This section should mention the SRP\_CRED\_REQ and SRP\_CRED\_RSP IUs, which are dedicated to flow control service.

CPQ #8 Page 18 (comment 1 on page)  
Section 5.4.2.1  
Table 2  
Remove period from "NO DATA BUFFER DESCRIPTOR PRESENT."

CPQ #9 Page 18 (comment 2 on page)  
Section 5.4.2.1  
Table 2  
There is no reference to note b. It probably needs to be in the 2h row buffer descriptor length cell, where "count" is used

CPQ #10 Page 18 (comment 3 on page)  
Section 5.4.2.1  
Table 2  
Add a period at the end of note c.

CPQ #11 Page 19 (comment 1 on page)  
Section 5.4.2.4  
Add a fairly content-free table showing a direct data buffer containing a memory descriptor so this section has a visual reference like the indirect section.

CPQ #12 Page 20 (comment 1 on page)  
Section 5.4.2.5  
Table 5  
note a  
count should be defined with a note b similar to that in table 2

CPQ #13 Page 20 (comment 2 on page)  
Section 5.4.2.5  
Table 4  
If n is zero in  $16*n+19$ , then the table shows byte 20 followed by byte 19. Remove the 20 and that numbering problem is eluded.

CPQ #14 Page 25 (comment 1 on page)  
Section 6.2  
Change:  
"maximum length"  
to  
"maximum length in bytes"

CPQ #15 Page 25 (comment 2 on page)  
Section 6.2 and elsewhere  
I thought we decided that TAG fields don't have bits labeled (MSB)/(LSB).

CPQ #16 Page 25 (comment 3 on page)  
Section 6.2  
Table 9  
The REQUIRED BUFFER FORMATS cell is missing the horizontal lines present in other multi-byte cells

CPQ #17 Page 26 (comment 1 on page)  
Section 6.2  
Table 10  
Remove period from first Reserved. row

CPQ #18 Page 27 (comment 1 on page)  
Section 6.3  
Change (two places):  
maximum length  
to  
"maximum length in bytes"

CPQ #19 Page 29 (comment 1 on page)  
Section 6.4  
Table 14  
Capitalize Reserved

CPQ #20 Page 29 (comment 2 on page)  
Section 6.4  
Table 13  
The SUPPORTED BUFFER FORMATS cell is missing the horizontal lines present in other multi-byte cells

CPQ #21 Page 31 (comment 1 on page)  
Section 6.6  
Table 17  
Add period after Reserved or remove from other rows

CPQ #22 Page 33 (comment 1 on page)  
Section 6.7  
Rename TASK MANAGEMENT FLAGS to TASK MANAGEMENT FUNCTION. It doesn't really contain flags.

CPQ #23 Page 33 (comment 2 on page)  
Section 6.7  
Table 19  
end each row with a period (or don't)

CPQ #24 Page 33 (comment 3 on page)  
Section 6.7  
Table 19  
Change Codes to Code.

CPQ #25 Page 33 (comment 4 on page)  
Section 6.7  
Table 19  
Remove small caps from TABLE.

CPQ #26 Page 34 (comment 1 on page)  
Section 6.8  
Table 20  
Per Patrick Fitzgerald at JNI, please require that DATA-OUT BUFFER DESCRIPTOR and DATA-IN BUFFER DESCRIPTOR start on 8-byte aligned boundaries. The ADDITIONAL CDB field is only 4 byte aligned.

CPQ #27 Page 34 (comment 2 on page)  
Section 6.8  
Table 20 footnotes  
Change:  
length  
to:  
length in bytes

CPQ #28 Page 35 (comment 1 on page)  
Section 6.8  
Table 21  
SAM-2 rev 20 still requires that untagged tasks be supported by all protocols. 01-318 will remove this requirement and make SRP legal.

CPQ #29 Page 35 (comment 2 on page)  
Section 6.8  
Table 21  
Change a to an in the ACA row

CPQ #30 Page 35 (comment 3 on page)  
Section 6.8  
Table 21  
Remove small caps from TABLE

CPQ #31 Page 38 (comment 1 on page)  
Section 6.9  
After:  
The STATUS field contains the status of a task that completes. See the SAM-2 standard for a list of status codes.  
Add this sentence and a table:  
Some of the status codes defined in SAM-2 are listed in table xx.  
Table xx - Some STATUS codes  
00h GOOD  
02h CHECK CONDITION  
08h BUSY  
18h RESERVATION CONFLICT  
28h TASK SET FULL  
30h ACA ACTIVE  
40h TASK ABORTED  
This helps save the reader a reference to SAM-2 for the most popular fields.

CPQ #32 Page 39 (comment 1 on page)  
Section 6.9  
Remove from 2nd sentence of SENSE DATA paragraph:  
as specified by the SCSI Primary Commands-2 standard.

CPQ #33 Page 39 (comment 2 on page)  
Section 6.9  
Reword the SENSE DATA paragraph to focus on the term autosense which is defined in SAM-2 rather than the REQUEST SENSE command in SPC-2.  
Change:  
The SENSE DATA field contains the information specified by the SCSI Primary Commands-2 standard for presentation by the REQUEST SENSE command. The proper sense data shall be presented when a SCSI status byte of CHECK CONDITION is presented by the SCSI Primary Commands -2 standard.  
to:  
The SENSE DATA field contains the autosense data (see SCSI Architecture Model - 2) when a SCSI STATUS byte of CHECK CONDITION is presented.

CPQ #34 Page 41 (comment 1 on page)  
Section 6.12  
Change  
report an asynchronous event.  
to:  
report an asynchronous event (see SAM-2).

CPQ #35 Page 41 (comment 2 on page)  
Section 6.12  
Add sentence to first paragraph:  
Parameters managing the use of asynchronous event reporting are contained in the Control mode page (see SPC-2).  
This sentence is in SAM-2, but a direct reference from SRP seems helpful.

CPQ #36 Page 42 (comment 1 on page)  
Section 6.13  
Reword the SENSE DATA paragraph like in 6.9, but don't call it autosense here, call it "sense data for the event".

CPQ #37 Page 43 (comment 1 on page)  
Section 7.1

Table 29  
Section 7.3  
LUN should be LU  
(this is broken in SPC too) - the logical unit number is irrelevant here.

CPQ #38 Page 52 (comment 1 on page)  
Annex B  
Change (many places):  
Infiniband  
to:  
InfiniBand

CPQ #39 Page 52 (comment 2 on page)  
Annex B  
There are too many TMs. There only needs to be one per page or one per the whole section.

CPQ #40 Page 62 (comment 1 on page)  
Annex C  
Ralph Weber agreed to put alias formats for each protocol in SPC-3, so this annex can be removed.

\*\*\*\*\*

Comments attached to YesC ballot from Mr. Randy Haagens of Hewlett Packard Co.:

Comment :  
Need a mandatory requirement to persistently report service names (DevMgtGetResp(ServiceEntries)) across IOU/IOC power cycles in order to persistently identify an SRP target port.

Description:  
Table B.8 describes the format of service name as SRP.T10: xxxxxxxxxxxxxxxx. Since the string xxxxxxxxxxxxxxxx in the service name identifies the 64 bit extension identifier value used to construct the SRP target port identifier, it is required that the service name reported by an IOU for a given SRP target port to be persistent across IOU/IOC power cycles. IB boot records contain SRP initiator port identifier, SRP target port identifier and logical unit name to locate an SRP boot LUN and the assumption is that the target port ID is persistent.

Comments on working draft SCSI RDMA Protocol (SRP) T10/1415-D rev.10 dated 3 Oct 2001

These informal comments are the result of a newcomer's first in-depth reading of the SRP specification. I hope they will suggest avenues for further improvement, but they are not formulated at this time as specific requests for changes.

These comments derive from my work on iSCSI, and are in anticipation of development of iWARP, which will be an RDMA protocol for IP networks. iWARP is intended to provide a standard protocol-independent means of doing direct data placement into host memory, without the need for anonymous reassembly buffers. We anticipate that iSCSI and other Internet storage protocols such as CIFS and NFS will be adapted to iWARP. Inclusion of a formalized RDMA transport layer in the IP storage protocol stack places iSCSI on a path to converge with SRP. Each protocol can learn from the other. Today, SRP, while meant to be generally applicable, is demonstrably applicable only to InfiniBand. iSCSI's applicability is similarly limited to IP networks. In the future, we may be able to engineer a single SCSI transport that works both with InfiniBand's RDMA service and with iWARP.

These are my personal comments, and are not meant to reflect an HP consensus. We at HP have not yet taken the time to form an internal consensus on SRP.

Each comment begins with a page-line reference number.

1-11. It is not clear at the outset just what kind of standard SRP is. The



text says that "the SCSI family of standards provides for many different transport protocols?" Is SRP a transport protocol? The text continues, "This standard defines the rules for exchanging information between SCSI devices using an RDMA communication service." So SRP is a mapping from SCSI to an abstract RDMA communication service? What then is the SCSI transport? Is it the combination of SRP and the underlying real RDMA communication service? The standard continues, "Other SCSI transport protocol standards?" So, perhaps SRP is a SCSI transport. A statement along these lines would help a lot: "SRP, in combination with a compatible underlying RDMA communication service, is a SCSI transport. This document defines SRP and the requirements that SRP has for the underlying RDMA communication service."

1-19. "Figure 1 shows the relationship of this standard to the other standards?" But it doesn't. The SRP standard is not identified in the figure. Despite the disclaimer, layering of the blocks does suggest a hierarchy, protocol stack and system architecture. But the figure does not indicate the applicability of SRP to the implementation of a SCSI transport, as far as I can tell.

2-28. SRP is included in a list of transport protocols. So it is a transport protocol. But certainly it is not a complete transport protocol. A discussion of how SRP is used in combination with an underlying RDMA service and its transport protocol to form a SCSI transport protocol would be very instructive to the reader. This would involve a layering diagram-why not?

8-4. It would be useful to say at the beginning of clause 4 that the purpose of clause 4 is to describe an abstract RDMA service that is suitable for supporting SRP. That is, to define SRP's requirements of an underlying RDMA service.

8-17. "This clause describes various functions that may be provided?" Don't you mean to say that this clause describes various functions that must be provided by an RDMA service, in support of SRP? How the function is provided is immaterial, and of course it can be provided through further functional decomposition. Why mention it? Generally, this whole clause 4 seems to be descriptive of RDMA services in general, but not prescriptive in terms of SRP's requirements. It is difficult to separate descriptive information from requirements.

8-20. "Annex B describes the mapping of these functions?" Is it the intention of SRP to work with other RDMA services besides InfiniBand? If so, it might be useful to mention that future revisions of the standard may include other Annexes that define the mapping of SRP to other RDMA services.

10-12. SRP is deficient in not providing a security protocol for client (initiator) authentication. Is the notion of "other parameters required by the RDMA communication service" to be interpreted as suggesting that the RDMA service itself should provide authentication? Given that SCSI port names are conveyed by SRP, this doesn't seem possible. (The RDMA service will have its own names for its end nodes, but they're not related to SCSI/SRP port names.)

11-36. "An RDMA communication service may require?" This sounds to vague and inclusive. What does SRP require of the RDMA service? That's all that should be defined in clause 4. It seems like SRP either will depend on the RDMA service's providing flow control for messages, or it will provide its own flow control. If SRP provides its own flow control, and doesn't depend on flow control from the RDMA service, then there is no reason to discuss flow control except maybe to mention that it is not required.

12-40. 4.5 Ordering and Reliability. Very glad to see this here. Wish it were in SAM-2.

14-24. "Server address" probably should be "server identifier".

15-24. Establishing multiple connections between an I,T port pair is an interesting concept, but may not be very useful, ultimately. The paragraph states that all such RDMA channels are associated with the single I\_T nexus. While there is no ordering assumed between different RDMA channels (15-41), this channel independence cannot be maintained once the tasks are forwarded to the SCSI layer, where the RDMA channel allegiance of the task is forgotten,

and only the I\_T information is retained. Effectively, the tasks will merge from multiple transmission channels into a single queue as they transition from SRP to SCSI, and the original partial order will be replaced by a total order. Correct operation will result, but performance will suffer. Perhaps the only practical use of this construct is for the asynchronous transmission of task management requests, as in the given example.

16-28. A request windowing scheme would be easier to describe than this request limit mechanism. Race conditions would not be an issue.

20-4. Indirect data buffer descriptor. I don't see a good use for this facility in an IO application such as SRP, and I question its inclusion here. The channel adapter local to the memory that is to be read or written (typically the channel adapter of the Initiator) can use a scatter/gather list (SGL) to define an arbitrary virtual memory segment for an I/O buffer, and assign it a unique memory handle. This segment can then be read or written, starting at any offset, and in any order, by the target's RDMA mechanism's simply generating a series of RDMA reads or writes, always referring to the same memory handle, but using different offsets and lengths for each operation. (For example, a series of RDMA writes to increasing offsets, eventually filling the memory segment.) The direct data buffer descriptor format is sufficient for this operation, because the SGL provides for scatter/gather to bufflets that start and end at arbitrary addresses in physical memory (not just page-aligned addresses), just as a traditional DMA controller does.

The only motivation I can find for the indirect model is to reduce the number of SGLs (or mapped memory regions) that the initiator's channel adapter must deal with. Unfortunately, the use of the indirect mechanism means that we must trust the target devices that share a memory region not to step on each other through misoperation or by deliberately generating invalid memory descriptors. While this is the truest form of remote DMA, because it leaves the matter of address generation to the target device, it also leaves the initiator exposed to target device misoperation, or worse.

I am not sufficiently familiar with IB HCA architecture to know whether such HCAs are limited to mapping only regions of contiguous pages, which would necessitate including the indirect data buffer descriptor method to support non-page-oriented IO.

25-1. Login request. The statement that the login request "shall only be sent during RDMA channel establishment" seems to me overly restrictive on the RDMA model. Furthermore, I'm not sure I discern in clause 4 that the RDMA service must transport SRP login information during its own connection establishment, although this requirement is made clear in clause 5, line 14-13. It would seem quite natural to establish an RDMA connection first, and then login SRP using the RDMA connection. (As an example, iSCSI establishes a TCP connection, and then logs it into a new or existing iSCSI session.)

25-1. Login request. Need to resolve how security protocols are handled in the SRP world. The login request does not contain any provision for initiator port authentication to the target.

25-32. So port identifiers are 16 bytes. But SAM-2 rev. 17 allows 8 bytes only, and iSCSI allows 260 bytes or more (still in discussion). These differences need to be rationalized. It would be best if SCSI itself would adopt a naming convention for its ports, rather than delegating this crucial task to its many transports. If SCSI were to name its ports, then SRP would only have to convey the SCSI port identifier passed down the stack by SCSI, and not make provision for conveying an identifier defined by a lower-level transport.

25-32. The port identifier fields, at 16B, are too small to carry identifiers as used by iSCSI. This may prove problematical as we attempt to merge iSCSI and SRP for use with iWARP.

54-1. SRP annex. Are Queue Pairs (QP) in one-to-one correspondence with IB consumers?

54-23. "An IB I/O unit?contains an IB channel adapter." Why restrict it to a

single channel adapter? In Figure B.3 the analogous (but nameless) initiator unit-defined by the dashed lines-is shown with multiple channel adapters. An iSCSI device is conceived as having multiple channel adapters (known informally as channel groups and in the specification as portal groups). OTOH, since an IB I/O unit is not named (it has no GUID associated with it), is there any purpose to the architecture's defining it?

54-28. Figure B.2. Can I/O controllers be virtual objects?

54-28. Figure B.2. What is the purpose of allowing multiple IB consumers per IB I/O controller? Is it so that multiple IB connections can be terminated within an IB I/O controller? (This relates to the question above about correspondence between QPs and IB consumers.)

54-28. Figure B.2. How are shared LUs modeled? Do SRP target ports contain the "task router" function described recently by Penokie? Can two IB I/O controllers have an underlying LU in common, or is this functionality restricted to two IB consumers within the same IB I/O controller?

55-9. Table B.1. IB port GUID is described as "Identifies an IB port within an IB channel adapter". This can be taken to mean that the naming scope for IB port is within a single channel adapter. I doubt that is the intention, since IB port GUIDs are globally unique. Similar comment for IB I/O controller GUID?with the further observation that IB I/O units themselves are not named, and so cannot form a naming scope. It seems to me that the first three lines of this table should read, "Identifies a \_\_\_\_\_", without qualification. It is incidental, isn't it, that an IB port is contained in an IB channel adapter (and an IB I/O controller is contained in an IB I/O unit)? The fact that the discovery process finds IB channel adapters, and then IB I/O controllers, and then IB consumers, utilizing the containment properties, seems irrelevant to describing the naming architecture, when globally unique names are used.

55-25. Figure B.3. What is the object indicated by the dashed lines in the initiator model, analogous to the IB I/O unit in the target model?

55-25. Figure B.3. and 56-1. Table B.2. The rules for constructing initiator ports seem entirely too lax. The text says, "Initiator port identifier should be constructed?" And then the Table indicates that GUID, for example, is the channel adapter GUID. Is there no meaning associated with the initiator port ID? Is the only design goal that the 16B port ID be globally unique? Will any GUID do at all? If so, let's be explicit about this, and let's not make any suggestions about the origin (and possible meaning) of the port name.

But it would be a better model, I think, for the "GUID" used in the initiator port ID to be associated not with the IB channel adapter, but instead with the (unnamed) SRP initiator device. It is the SRP initiator device that is associated with a naming domain such as an operating system image. IB channel adapters will be shared among operating system images, and using them as a naming domain would require that the operating system images cooperate, or that the selection of port identifier extension be delegated to the virtual machine layer, both of which are undesirable.

While we're at it, let's decouple the naming of SRP ports entirely from IB. Although SCSI really should be the layer that names its ports, let's for the moment assume that SCSI continues to delegate port naming to its transport. But let's assume further that SRP accepts the responsibility to name its ports, and doesn't delegate it further to IB. SRP can then generate its own name for SRP initiator device, with an identifier extension to make a unique port name. Analogously, SRP can name the entities identified in the figure as SRP target devices. SRP could adopt a naming scheme that uses 16B "GUIDs" analogously to IB's, and it could draw from the same naming assignment authority that IB uses. But this is not the same as saying that IB defines SRP's port names, and in fact, the description of SRP port naming would be moved from the IB annex to the main SRP text.

This change would require that during the discovery process, the IB I/O unit return the full name of the SRP port from its Service Entries table, in step 3.

This approach to naming ports brings SRP much closer to iSCSI. What is

unresolved is iSCSI's use of long text strings to name iSCSI devices vs. the use of more compact GUID numbers. The two mechanisms could be combined with the introduction of a name service that dereferences string IDs to GUIDs.

\*\*\*\*\*

Comments attached to No ballot from Mr. George O. Penokie of IBM / Tivoli Systems:

#### General

In my comments the notation 'Page xx' refers to all pages in the standard not roman numeral xx. All comments are editorial unless indicated with a '(T)' at the start of the comment.

PDF Page 3

1: Tivoli comment from George Penokie

Page c - The page numbering in first part of the front matter is a,b,c, and d instead of roman numerals. This needs to be corrected.

PDF Page 4

2: Tivoli comment from George Penokie

Page c - d - The Revision list needs to be removed before public review.

3: Tivoli comment from George Penokie

All - All the line numbers need to be removed throughout the document.

PDF Page 6

4: Tivoli comment from George Penokie

All - The printing date information at the bottom of every page needs to be removed.

PDF Page 11

5: Tivoli comment from George Penokie

Page vii - Forward - the BSR number x3.269-199x is not correct for this standard. It should be 'NClTS.xxx-200x' until the actual number is assigned.

PDF Page 12

6: Tivoli comment from George Penokie

Page viii - Line 7 - The statement 'The working draft SCSI RDMA Protocol (SRP) standard is divided into the following clauses:' should be 'The SCSI RDMA Protocol standard is divided into the following clauses:

7: Tivoli comment from George Penokie

All - The acronym SRP should be replaced with 'SCSI RDMA Protocol' in all cases in this document.

PDF Page 13

8: Tivoli comment from George Penokie

Page 1-2 - The following standards should be removed from the list: FC-AL, FC-PH, FC-PH-2, SPI-3, FCP, SPC, and RMC.

PDF Page 16

9: Tivoli comment from George Penokie

Page 4 - section 3.1.1 - The last sentence implies that SRP\_LOGIN\_RSP is the only use for accept data. I believe this is not correct. This should be stated to be an example of accept data.

10: Tivoli comment from George Penokie

All - The full name of a standard should always be used instead of the acronym. This should be change throughout the document.

11: Tivoli comment from George Penokie

Page 4 - line 19 and others - when SRP is used and it is referring to this document then it should be changed to 'this standard'. Line 19 is one case where this appears to be true.

12: Tivoli comment from George Penokie

Page 4 - section 3.1.13 - The statement 'An externally addressable object...' should be 'An addressable object...'. The term externally implies that the addressing is outside the standard.

13: Tivoli comment from George Penokie

Page 4 - section 3.1.15 - The last sentence implies that SRP\_LOGIN\_REQ is the only use for login data. If this is not correct. Then this should be stated to be an example of login data.

14: Tivoli comment from George Penokie

Page 4 - section 3.1.15 - The statement '...server agent or consumer...' should be '...server agent or server consumer...'

15: Tivoli comment from George Penokie

Page 4 - section 3.1 - The terms client consumer, server agent, and server consumer should be definitions in the glossary.

PDF Page 17

16: Tivoli comment from George Penokie  
Page 5 - section 3.1.22 - The statement '...server agent or consumer...' should be '....server agent or server consumer...'

17: Tivoli comment from George Penokie  
Page 5 - section 3.2 - line 34 - The acronym for SRP implies that in almost all cases SRP should be changed to 'this standard'.  
PDF Page 20

18: Tivoli comment from George Penokie  
Page 8 - line 5 - The statement 'by means of' should be change to 'using'.

19: Tivoli comment from George Penokie  
Page 8 - line 44 - The statement 'established and disconnected' should be either 'established and removed' or 'connected and disconnected'. In this case I think the first option is better. The wording in the remaining document must then be made to match this change.

20: Tivoli comment from George Penokie  
Pages 8 - 11 - section 4.2 - This clause should be broken in subclauses and there should be references added between the steps in the figure and the text descriptions of those steps. This will help the reader relate the figures flow to the text.  
PDF Page 21

21: Tivoli comment from George Penokie  
Page 9 - lines 7-9 - The for example text should be change to (e.g., ....).

22: Tivoli comment from George Penokie  
Page 9 - line 2 - The statement '...directed to a server and, if...' is not clear because there is a server agent and a server consumer. Which is this server supposed to be?

23: Tivoli comment from George Penokie  
Page 9 - line 5 - The statement '...identify the server with which...' is not clear because there is a server agent and a server consumer. Which is this server supposed to be?

24: Tivoli comment from George Penokie  
Page 9 - Figure 3 - line 40 - The arrow exiting to the right seems to dead end. Where does the flow go from there. All the other exit points are clear that one is not.  
PDF Page 22

25: Tivoli comment from George Penokie  
(T) Page 10 - line 12 - This states '...the server identifier shall identify one or more SRP target ports, and the login data...'. How is it possible for a single server identifier to identify more than one SRP port.? SCSI requires all target port identifiers be unique within a domain.

26: Tivoli comment from George Penokie  
Page 10 - at least lines 2-15 - The term 'server' is used by itself several times. There needs to be a qualifier on server so the reader does not assume that server equates to server agent and server consumer.

27: Tivoli comment from George Penokie  
Page 10 - lines 28 - 29 - The statement 'With SRP the reject data includes an SRP\_LOGIN\_REJ response (see 6.4).' Is confusing in that it implies the SRP (which is this standard) has additional requirements than what was just specific in the sentence before. That does not compute and needs to be fixed.

28: Tivoli comment from George Penokie  
Page 10 - lines 31 -32 - Is it possible for an RDMA channel to be successfully established and not operational? If not then the statement 'and is operational' should be deleted. If so then it needs to be explained how it is possible.

29: Tivoli comment from George Penokie  
Page 10 - line 34 - The statement '...server agent or consumer...' should be '....server agent or server consumer...'. This needs to be looked for throughout the document and corrected.

30: Tivoli comment from George Penokie  
Page 10 - line 35 - The statement 'With SRP the accept data includes an SRP\_LOGIN\_RSP response (see 6.3).' Is confusing in that it implies the SRP (which is this standard) has additional requirements than what was just specific in the sentence before. That does not compute and needs to be fixed.

31: Tivoli comment from George Penokie  
Page 10 - lines 44-45 - The statement 'With SRP the login data includes an SRP\_LOGIN\_REQ request (see 6.2)...' Is confusing in that it implies the SRP (which is this standard) has additional requirements than what was just specific in the sentence before. That does not compute and needs to be fixed.

32: Tivoli comment from George Penokie  
Page 10 - lines 43 - 44 - The sentence 'The server agent is provided the login

data from the client consumer's request in addition to RDMA communication service specific data.' is awkward. It would be better stated as 'The server agent receives the login data and RDMA communication service specific data from the client consumer's request.'

PDF Page 23

33: Tivoli comment from George Penokie

Page 11 - line 2 - The statement 'With SRP the reject data shall contain an SRP\_LOGIN\_REJ response (see 6.4).' Is confusing in that it implies the SRP (which is this standard) has additional requirements than what was just specific in the sentence before. That does not compute and needs to be fixed.

34: Tivoli comment from George Penokie

Page 11 - lines 5 - 6 - The statement 'With SRP the accept data shall contain an SRP\_LOGIN\_RSP response (see 6.3)...' Is confusing in that it implies the SRP (which is this standard) has additional requirements than what was just specific in the sentence before. That does not compute and needs to be fixed.

35: Tivoli comment from George Penokie

Page 11 - line 11 - The term 'such' should be deleted.

36: Tivoli comment from George Penokie

Page 11 - lines 30-31 - The statement '...to deliver the message to the other consumer associated with the specified RDMA channel (the receiving consumer).' should be changed to '...to deliver the message to the receiving consumer.' There is no need to redefine what a receiving consumer is as that is done in the first paragraph of this section.

37: Tivoli comment from George Penokie

Pages 11 - 12 - section 4.4 - This clause should be broken in subclauses. For example at least an overview, one for read RDMA, and one for write RDMA.

PDF Page 24

38: Tivoli comment from George Penokie

Page 12 - line 5 - The statement 'as well' should be deleted.

39: Tivoli comment from George Penokie

Page 12 - line 14 - The following statement 'Such information may be communicated by an application protocol.' Does not seem relevant to this standard and should be deleted.

40: Tivoli comment from George Penokie

Page 12 - lines 41-43 - This paragraph contains information that is not useful and should be deleted. It essentially states that RDMA communication has characteristics defined here and those not defined here are outside the scope of this standard. That is true but it is also true for every clause in this standard.

41: Tivoli comment from George Penokie

Page 12 - line 45 - The statement 'or else' should be just 'or'.

42: Tivoli comment from George Penokie

Page 12 - line 46 - The term 'exactly' should be deleted. There is no difference between 'exactly once' and 'once'.

43: Tivoli comment from George Penokie

Page many - The terms Write and Read in RDMA Write and RDMA Read should not be capitalized.

PDF Page 25

44: Tivoli comment from George Penokie

Page 13 - line 14 - The term 'satisfy' should be changed to 'meet'.

PDF Page 26

45: Tivoli comment from George Penokie

Page 14 - line 8 - The statement 'I\_T nexus' is correct but there is no reference to where one would find out more about what it is. This needs to be added.

46: Tivoli comment from George Penokie

Page 14 - line 7 - The statement 'for its lifetime' is not clear. It should be stated as 'as long as it is established'. This ties it to the previous section. Note this assumes that the term established in 4.2 is not changed.

47: Tivoli comment from George Penokie

Page 14 - lines 24-28 - This whole paragraph does not look like it belongs here or anywhere and it should be deleted. It appears to be attempting to define things that are either already defined in section 4 or don't need to be defined.

48: Tivoli comment from George Penokie

Page 14 - line 40 - The statement 'that were contained in SRP\_CMD requests (see 6.8)' should be deleted as it is redundant with the statement 'outstanding SCSI tasks'.

49: Tivoli comment from George Penokie

(T) Page 14 - line 43 - The statement '...an SRP target port should send

an...' gives inadequate guidance to a target implementor. This should be required to send the SRP\_T\_LOGOUT or not send it. Or it should be specified when it is required to be sent and when it is not required to be sent.

50: Tivoli comment from George Penokie

PDF Page 27

(T) Page 15 - line 4 - I recommend adding into this list a statement that other SCSI related parameters (e.g., mode pages, logs) not be effected by the disconnect. This should avoid the hole the FC has dug for itself in this area.

51: Tivoli comment from George Penokie

Page 15 - line 18 - The statement '...operation, if accepted, may allow...' should be '...operation may allow...'. The if accepted is redundant with may.

52: Tivoli comment from George Penokie

Page 15 - line 36 - The term 'may' should be deleted.

53: Tivoli comment from George Penokie

Page 15 - lines 36-40 - the format of the e.g is incorrect. It should be...'standards (e.g., ...).'

54: Tivoli comment from George Penokie

Page 15 - line 49 - The statement 'as well as' should be 'or'.

PDF Page 28

55: Tivoli comment from George Penokie

Page 16 - line 3 - The term 'initiation' should be 'start' or 'beginning'.

56: Tivoli comment from George Penokie

Page 16 - line 5 - The term 'all' should be 'the'.

57: Tivoli comment from George Penokie

Page 16 - lines 7-8 - I am not aware of a SCSI command that specifies that status not be returned. If there is such a thing then an e.g., would be helpful. If there is no such thing then this item should be deleted.

58: Tivoli comment from George Penokie

Page 16 - line 18 - What is the 'it' referring to? The 'it' needs to be replaced with whatever 'it' is.

59: Tivoli comment from George Penokie

Page 16 - line 23 - The term 'might' should be 'may'.

60: Tivoli comment from George Penokie

Page 16 - line 23 - What is the 'it' referring to? The 'it' needs to be replaced with whatever 'it' is.

61: Tivoli comment from George Penokie

Page 16 - line 24 - The statement '...to at most one...' seems redundant. It should be '...to one...'

62: Tivoli comment from George Penokie

Page 16 - lines 28-29 - The statement '...present in most information units...' is troublesome. There either needs to be a list of the IUs that have the field or a reference to a place that would tell me which IUs have or do not have the field.

63: Tivoli comment from George Penokie

Page 16 and others? - The when to use small caps rule is not being followed here. The rule is that small caps are only used when the field is being named (e.g., xxx field would have the xxx in small caps). When contents of the field is being called out it is not in small caps (e.g. request limit and request limit delta are both signed...').

64: Tivoli comment from George Penokie

Page 16 - line 49 - The sentence starting with 'An SRP port shall not specify a negative...' should be a separate item in the list.

65: Tivoli comment from George Penokie

(T) Page 16 - section 5.3 - This section on flow control seems overly complex for what appears to be actually needed. The only SRP request that even needs to have multiple outstanding requests in the command. All others should not be streamed but should be interlocked and some should be allowed to occur at any time. This all needs to be looked at to make sure the design point is what we really want.

PDF Page 29

66: Tivoli comment from George Penokie

Page 17 - Figure 4 - The way the arrows are pointing for the virtual address implies that it is not the address of the first byte of the memory segment. It currently implies that it is the space from the memory handle to the beginning of the memory segment which is the memory region. It is also not clear as to what the boundaries are of the memory region. The current drawing implies it is only the area above the memory segment. I do not believe that is correct so it needs to be fixed.

67: Tivoli comment from George Penokie

Page 17 - line 26 - There is no indication as to what kind of value the memory handle is. This would normally not be a problem except that the other two fields to explicitly indicate that they are unsigned integer values. I generally consider all fields to be unsigned integers but in this case there is doubt cast about that assumption.

PDF Page 30

68: Tivoli comment from George Penokie

Page 18 - line 1 - The statement 'A SRP...' should be 'An SRP...' This needs to be checked for throughout the document and corrected.

69: Tivoli comment from George Penokie

Page 18 - line 3 - The statement '...within its memory segment.' should be '...within the memory segment.'

70: Tivoli comment from George Penokie

Page 18 - line 2 - The statement 'SRP target ports shall only issue the appropriate type of RDMA operation for a memory descriptor' appears to be restating what was stated in the previous sentence and therefore should be deleted. The sentence would then read 'SRP target ports shall ensure that each RDMA operation...'

71: Tivoli comment from George Penokie

Page 18 - line 3 - There needs to be a connection between the text above the a.b.c list and the list. Something like 'segment by using the following rules:'.

72: Tivoli comment from George Penokie

Page 18 - lines 15-17 - The sentences 'The format of each data buffer descriptor is specified by a format code value. Some data buffer descriptor format code values use the contents of a count field to further specify the data buffer descriptor format.' should be deleted as the information is a duplicate of what is in table 2.

73: Tivoli comment from George Penokie

Page 18 - table 2 - line 35 - footnote c - There statement 'and and' should be just 'and' and there is not period at the end of the sentence.

74: Tivoli comment from George Penokie

Page 18 - table 2 - line 27 - The equation  $20+16*\text{count}$  should be change to  $20 + 16 \times \text{count}$ . This change from \* to x should be make throughout the document.

75: Tivoli comment from George Penokie

Page 18 - table 2 - footnote b - This should have a reference from the cell with 'count' in it.

76: Tivoli comment from George Penokie

Page 18 - lines 43-45 - The sentence 'An SRP initiator port shall not specify a data buffer descriptor format that was not indicated in the REQUIRED BUFFER FORMATS field value for that RDMA channel.' does not make sense. How can the initiator port be indicating the buffer formats in the REQUIRED BUFFER FORMATS field and at the same time not specifying the buffer formats in the REQUIRED BUFFER FORMATS field that were not indicated in the in the REQUIRED BUFFER FORMATS field. This is circular and needs to be fixed.

77: Tivoli comment from George Penokie

Page 18 - line 41 - There should be a reference to table 2 as follows 'data buffer descriptor formats (see table 2)'

78: Tivoli comment from George Penokie

Page 18 - line 47 - The statement '...RDMA channel and...' should be '...RDMA channel request and...'

79: Tivoli comment from George Penokie

Page 18 - line 49 - The statement '...RDMA channel and...' should be '...RDMA channel request and...'

80: Tivoli comment from George Penokie

Page 18 - line 40 - There should be a reference to table 3 as follows 'The REQUIRED BUFFER FORMATS field (see table 3)...'

PDF Page 31

81: Tivoli comment from George Penokie

Page 19 - line 4 - I believe the 'and' should be an 'or'. I don't believe a target port would do both IU at the same time.

82: Tivoli comment from George Penokie

Page 19 - lines 3-4 - There should be a reference to table 3 as follows 'The SUPPORTED BUFFER FORMATS field (see table 3)...'

83: Tivoli comment from George Penokie

Page 19 - line 8 - The statement '...contents of the REQUIRED BUFFER...' should be '...contents of both the REQUIRED BUFFER...'

84: Tivoli comment from George Penokie

(T) Page 19 - line 18 and line 28 - Why is that when the IDBD bit and the DDBD bit is set to zero it is a should instead of a shall? This should be changed



to a shall unless there is some good reason.

85: Tivoli comment from George Penokie

Page 19 - note 2 - This note should not be a note. It should be part of the main text. It should also be restated as: 'The length of requests sent by an SRP initiator port, as determined by the data buffer descriptor formats, shall be limited to the MAXIMUM INITIATOR TO TARGET IU LENGTH field (see xxx) returned in the SRP\_LOGIN\_RSP response.'

86: Tivoli comment from George Penokie

(T) Page 19 - lines 39 - 40 - The sentence 'SRP target ports are not required to check the contents of the count field.' should be changed to 'SRP target ports shall ignore the contents of the count field.'.

87: Tivoli comment from George Penokie

(T) Page 19 - lines 44 - 45 - The sentence 'SRP target ports are not required to check the contents of the count field.' should be changed to 'SRP target ports shall ignore the contents of the count field.'.

88: Tivoli comment from George Penokie

Page 19 and others - line 39 and others - The term 'count field' is used in many places. First there are two of them so it should be 'count fields'. Second is not clear that these are the count fields in the SRP\_CMD request. I recommend changing 'count field' to 'count fields in the SRP\_CMD request' in all places in the main body text.

PDF Page 32

89: Tivoli comment from George Penokie

Page 20 - line 8 - The statement 'count field' should be 'DATA-OUT BUFFER DESCRIPTOR COUNT field (or DATA-IN BUFFER DESCRIPTOR COUNT field)'.

90: Tivoli comment from George Penokie

Page 20 - line 12 - A reference to table 5 should be added to the end of the paragraph.

91: Tivoli comment from George Penokie

Page 20 - table 4 - line 29 - Footnote a - It's not clear which count field is being referred to. Is it the one in table 2 or the ones in the SRP\_CMD\_ request. This needs to be fixed with the proper terminology and a reference to the correct place.

92: Tivoli comment from George Penokie

Page 20 - line 34 - The statement 'The DATA LENGTH field of the INDIRECT TABLE MEMORY DESCRIPTOR field value contains...' is not correct. It should be 'The DATA LENGTH field of the memory descriptors in the indirect table contains...'.

93: Tivoli comment from George Penokie

Page 20 - line 39 - The sentence 'SRP target port behavior when the TOTAL LENGTH field contains any other value is vendor specific.' should be moved to the end of the paragraph and restated as 'If the TOTAL LENGTH field value is not equal to the sum of the DATA LENGTH field values the SRP target port's behavior shall be vendor specific.'.

94: Tivoli comment from George Penokie

Page 20 - line 42 - It's not clear which count field is being referred to. Is it the one in table 2 or the ones in the SRP\_CMD\_ request. This needs to be fixed with the proper terminology and a reference to the correct place.

95: Tivoli comment from George Penokie

Page 20 - line 47 - This should be the start of a new subclause. Something like 'SRP target port indirect data restrictions'.

PDF Page 33

96: Tivoli comment from George Penokie

Page 21 - line 7 - This paragraph should be the start of a new subclause titled something like 'Examples of Indirect data buffers'.

97: Tivoli comment from George Penokie

(T) Page 20 and 21 - The possibility of having both a data-in and a data-out buffer is not described here. Why not? This needs to be fixed.

98: Tivoli comment from George Penokie

Page 21 - lines 12 and 13 - The term 'might' should be changed to 'may'. This should be done throughout this document.

PDF Page 35

99: Tivoli comment from George Penokie

Page 23 - line 48 - The statement 'A requestor shall provide a TAG value in each SRP request that is unique among all of the requestor's outstanding SRP requests with a particular responder. A responder shall copy the TAG value from each SRP request to the SRP request's SRP response. Responders are not required to check whether the TAG values of outstanding SRP requests are unique.' should be 'Each SRP request shall contain a TAG value that is unique among all of the outstanding SRP requests from a particular SRP initiator'.

port. Each SRP response shall contain a copy of the TAG value from the corresponding SRP request. Responders are not required to check whether the TAG values are unique.'

PDF Page 37

100: Tivoli comment from George Penokie

Page 25 - line 2 - The term 'conveys' should be changed to 'sends'.

101: Tivoli comment from George Penokie

Page 25 - line 42 - The statement '...wishes to send...' should be changed to '...sends...'

102: Tivoli comment from George Penokie

Page 25 - line 42 - The statement '...be 64 or larger.' should be '...be greater than or equal to 64.' or '...be greater than 63.'

PDF Page 38

103: Tivoli comment from George Penokie

Page 26 - lines 1-2 - The statement 'The MULTI-CHANNEL ACTION field identifies how an SRP target port treats any existing RDMA channel associated with the same I\_T nexus. The MULTI-CHANNEL ACTION field is defined in table 10.' should be changed to 'The MULTI-CHANNEL ACTION field (see table 10) indicates how an SRP target port handles existing RDMA channels associated with the same I\_T nexus.'

104: Tivoli comment from George Penokie

Page 26 - table 10 - All the codes except for the 2 that are defined need to be listed as reserved. The row should have '02h - FFh' in the action column and 'reserved' in the description column.

PDF Page 39

105: Tivoli comment from George Penokie

Page 27 - line 2 - The term 'conveys' should be changed to 'sends'.

PDF Page 40

106: Tivoli comment from George Penokie

Page 28 - lines 1-2 - The statement 'MULTI-CHANNEL RESULT identifies how the SRP target port treated existing RDMA channels associated with the same I\_T nexus. Table 12 defines this field.' should be changed to 'The MULTI-CHANNELRESULT field (see table 12) indicates how an SRP target port handles existing RDMA channels associated with the same I\_T nexus.'

107: Tivoli comment from George Penokie

Page 28 - table 12 - All the codes except for the 3 that are defined need to be listed as reserved. The row should have '03h - FFh' in the action column and 'reserved' in the description column.

PDF Page 42

108: Tivoli comment from George Penokie

Page 30 - line 4 - The statement '...failed, rendering it non-operational..' should be changed to '...failed.'

PDF Page 43

109: Tivoli comment from George Penokie

Page 31 - line 4 - The statement '...failed, rendering it non-operational..' should be changed to '...failed.'

PDF Page 44

Page 32 - line 2 - The term 'conveys' should be changed to 'sends'.

110: Tivoli comment from George Penokie

Page 32 - lines 37-38 - The statement '...logical unit component of the nexus for the task management request.' should be changed to '...logical unit to which to send task management request.'

PDF Page 46

111: Tivoli comment from George Penokie

Page 34 - line 2 - The term 'conveys' should be changed to 'sends'.

112: Tivoli comment from George Penokie

Page 34 - 35 - Table 20 - This table splits up a paragon and worse a sentence. This needs to be fixed.

113: Tivoli comment from George Penokie

Page 34 - table 20 - The notation 'do' and 'di' are confusing when placed into a sentence (as in the footnotes). They should be changed to 'x' and 'y'.

PDF Page 48

114: Tivoli comment from George Penokie

Page 36 - line 2 - The term 'conveys' should be changed to 'sends'.

115: Tivoli comment from George Penokie

Page 36 - line 6 - The statement '...message capable of containing...' should be changed to '...message containing...'

PDF Page 49

116: Tivoli comment from George Penokie

Page 37 - The statement 'set to 1' should be 'set to one' and the statement

'set to 0' should be 'set to zero' in all cases throughout this document.

117: Tivoli comment from George Penokie

Page 37 - line 44 - The statement 'are not reliable and' should be deleted as it contains no useful information.

PDF Page 50

118: Tivoli comment from George Penokie

Page 38 - line 3 - Add a reference to the RSP\_CODE values table (table 24) at the end of this paragraph.

119: Tivoli comment from George Penokie

(T) Page 38 - lines 15-17 - The statement 'If DOUNDER is set to 1, a transfer that did not use the entire data-out buffer was performed and the value of DATA-OUT RESIDUAL COUNT shall be equal to: data-out buffer length - highest offset of any data-out byte transferred - 1' needs to be changed to 'If DOUNDER is set to one and a transfer that did not fill the entire data-out buffer was performed the value of DATA-OUT RESIDUAL COUNT is defined as follows:

DATA-OUT RESIDUAL COUNT = (data-out buffer length) - (highest offset of any data-out byte transmitted + 1)'

120: Tivoli comment from George Penokie

(T) Page 38 - lines 22-23 - The statement 'DATA-OUT RESIDUAL COUNT shall be equal to: data-out transfer length required by command - data-out buffer length' needs to be changed to 'The DATA-OUT RESIDUAL COUNT is defined as follows:

DATA-OUT RESIDUAL COUNT = (Transfer length required by command) - (data-out buffer length)'

121: Tivoli comment from George Penokie

(T) Page 38 - lines 34-36 - The statement 'If DIUNDER is set to 1, a transfer that did not fill the entire data-in buffer was performed and the value of DATA-IN RESIDUAL COUNT shall be equal to: data-in buffer length - highest offset of any data-in byte transferred - 1' needs to be changed to 'If DIUNDER is set to one and a transfer that did not fill the entire data-in buffer was performed the value of DATA-IN RESIDUAL COUNT is defined as follows:

DATA-IN RESIDUAL COUNT = (data-in buffer length) - (highest offset of any data-in byte transmitted + 1)'

122: Tivoli comment from George Penokie

(T) Page 38 - lines 41-43 - The statement 'DATA-IN RESIDUAL COUNT shall be equal to: data-in transfer length required by command - data-in buffer length' needs to be changed to 'The DATA-IN RESIDUAL COUNT is defined as follows:

DATA-IN RESIDUAL COUNT = (Transfer length required by command) - (data-in buffer length)'

PDF Page 51

123: Tivoli comment from George Penokie

Page 39 - line 1 - The term 'certain' should be deleted.

124: Tivoli comment from George Penokie

(T) Page 39 - lines 30 - 41 - All this should be deleted and replaced with 'The SENSE DATA field contains the autosense data specified by the SCSI Primary Commands-2 standard. The proper SENSE DATA shall be presented when the SCSI status byte of CHECK CONDITION is presented as specified by the SCSI Primary Commands-2 standard. If no conditions requiring the presentation of SCSI sense data have occurred, the SENSE DATA field shall not be included in the SRP\_RSP response and the RSPVALID bit shall be zero. SRP devices shall perform autosense.'

PDF Page 53

125: Tivoli comment from George Penokie

Page 41 - line 2 - The term 'conveys' should be changed to 'sends'.

PDF Page 54

126: Tivoli comment from George Penokie

(T) Page 42 - lines 3-13 - All this should be deleted and replaced with the following 'The SENSE DATA field contains sense data specified by the SCSI Primary Commands-2 standard.'. This is AER not a check condition they are different things. The only thing that should be stated here is that sense data is returned.

PDF Page 56

127: Tivoli comment from George Penokie

Page 44 and others - line 16 and others - The term 'set to 0' and 'set to 1' should be 'set to one' and 'set to zero'.

128: Tivoli comment from George Penokie

Page 44 - line 19 - The term 'all' should be deleted as it is redundant.

PDF Page 58

129: Tivoli comment from George Penokie  
Page 46 - figure A.2 and A.3 - line 15 and 43 - The statement '(SRP initiator)' should be '(SRP initiator port)'.

130: Tivoli comment from George Penokie  
Page 46 and others- lines 22-26 and others - The 1,2,3 list should not have line spaces between numbered items. This should be fixed in all cases  
PDF Page 64

131: Tivoli comment from George Penokie  
Page 52 - line 20 - The term 'executes' should be changed to 'processes'.

132: Tivoli comment from George Penokie  
Page 52 - line 32 - The statement '...a device or component...' should be 'an IB device or component...'.

PDF Page 65

133: Tivoli comment from George Penokie  
Page 53 - line 20 - There seems to be no definition of what a 'connection manager' is. This should, at least, be added to the glossary.

134: Tivoli comment from George Penokie  
Page 53 - section B.3.2 - The abbreviation IOC needs to be added to the list.  
PDF Page 67

135: Tivoli comment from George Penokie  
Page 55 - lines 1-2 - The sentence 'The IB more IB LIDs and IB GIDs corresponding to an IB port GUID or IB channel adapter GUID.' does not seem to be a complete sentence and is not clear as to what it is trying to state. This needs to be fixed.  
PDF Page 68

136: Tivoli comment from George Penokie  
(T) Page 56 - line 2 - Why is the should not a shall. I believe it should be changed to a shall.

137: Tivoli comment from George Penokie  
Page 56 - line 15 - The statement '...field should an IB GUID...' should be '...field should be an IB GUID...'.

138: Tivoli comment from George Penokie  
Page 56 - line15 - The statement '...port, e.g. the...SRP initiator port.' should be '...port (e.g., the...SRP initiator port)'.

139: Tivoli comment from George Penokie  
Page 56 - lines 15-16 - The statement 'the IB channel adapter GUID for an IB channel adapter used the SRP initiator port.' is not very clear as to what it is. This needs to be fixed.

140: Tivoli comment from George Penokie  
Page 56 - line 20 - There is not clue as what a 'device management agent' is. This could be fixed by replacing 'device management agent' with the more generic term 'entity'.

141: Tivoli comment from George Penokie  
Page 56 - line 22 - The term 'indicated' is confusing in this sentence. A better term would be 'identified'.

142: Tivoli comment from George Penokie  
Page 56 - line 39 - The term 'indicated' is confusing in this sentence. A better term would be 'identified'.

143: Tivoli comment from George Penokie  
Page 56 - lines 41-42 - This sentence seems out of place here. It should be moved to right after figure B.3.

144: Tivoli comment from George Penokie  
Page 56 - line 49 and page 57 - line 1 - The term 'IB I/O' has been split across lines (and in this case across pages) at the /. This needs to be fixed so it will not happen. There is an option in frame that if selected will prevent this. It should be enabled for this document.  
PDF Page 69

145: Tivoli comment from George Penokie  
Page 57 - line 34 - The 'it' at the beginning of the sentence should be replaced with whatever the 'it' is.

146: Tivoli comment from George Penokie  
Page 57 - line 46 and page 58 - line 1 - Why is the receive data-out mapped to RDMA requests and send data-in mapped to RDMA WRITE packets? One is a 'request' the other a 'packet' this seems strange shouldn't they be the same?  
PDF Page 73

147: Tivoli comment from George Penokie  
Page 61 - table B.8 - line 31 - The statement '(binary zeros)' should be '(i.e., binary zeros)'.

148: Tivoli comment from George Penokie  
SRP does not define any format for the 3rd party device identifier for

third party reservations. This needs to be added to comply with requirements in SPC-3.

Comments from IBM on SRP rev 10:

1. IBM Comment  
p50 line 11. "See 4x1" is a typo. I think this should be "See 4.4".
2. IBM Comment  
p50 line 14. "Sever" should be "server".
3. IBM Comment  
p50 line 35. "Sever" should be "server".
4. IBM Comment  
p57 section B.6.5. The descriptions for data-in and data-out are not symmetrical. One is described in terms of an "RDMA READ Request" and the other in terms of "one or more RDMA WRITE packets". I think the rules are the same for both data-in and data-out (please let me know if I'm incorrect in that assumption). Describing them differently implies that they are somehow different, and generates unnecessary confusion. (This is the same as Tivoli comment number 146).

\*\*\*\*\*

Comments attached to No ballot from Mr. Cris Simpson of Intel Corp.:

intel0001 Sect: 1 Pg: 1 Ln: 32  
Transport protocol s/b 'SCSI Protocol'  
Suggest shading box to clarify what we're doing in this spec

intel0002 Sect: 1 Pg: 1 Ln: 37  
Remove 'Physical'

intel0003 Sect: 3.1.11 Pg: 4 Ln: 27  
inconsistent use of 'the' before SRP - suggest no 'the'

intel0004 Sect: 3.1.14 Pg: 4 Ln: 34  
Is it necessary to specify field size in definition?

intel0005 Sect: 3.1.15 Pg: 4 Ln: 35  
'Application protocol' is not defined, thus what constitutes app proto data is unclear

intel0006 Sect: 3.1.16 Pg: 4 Ln: 39  
Key feature is that data placement is under control of receiver

intel0007 Sect: 3.1.17 Pg: 4 Ln: 41  
'path' is a poor term, implies routing

intel0008 Sect: 3.1.17 Pg: 4 Ln: 43 C  
'a transport protocol or service' - which is it?  
There appears to be an abstraction layering problem  
Using 'service' to define a service suggests we don't have a clean definition - we don't

intel0009 Sect: 3.1.23 Pg: 5 Ln: 6  
rewrite as 'specific to an RDMA comm service'

intel0010 Sect: 3.1.27 Pg: 5 Ln: 12 C  
TP ID 'within an RDMA comm service' - another abstraction issue - what is a service?

intel0011 Sect: 3.1.28 Pg: 5 Ln: 15  
Any reason to specify field size?

intel0012 Sect: 3.3.9 Pg: 6 Ln: 12  
reported as AN error

intel0013 Sect: 4 Pg: 8 Ln: 1 C  
Clause 4 alternates between being a generic overview of RDMA, including

discussion of features not used by SRP (e.g., solicited events in 4.3), and being normative (numerous SHALLs), which seems out of place in a clause entitled '...model'

Suggest separating the architectural model from the normative.

intel0014 Sect: 4.2 Pg: 8 Ln: 46  
Seems redundant to Line 10 above.

intel0015 Sect: 4.2 Pg: 9 Ln: 2 C  
Model is unclear:

"A client consumer requests that the RDMA communication service establish an RDMA channel."

But RDMA\_CS is defined as a protocol. The sense should be that the client requests a SERVICE PROVIDER establish a channel.

intel0016 Sect: 4.2 Pg: 9 Ln: 2  
"The request is directed to a server" - Ambiguous

There are several standard meanings for 'server' - a piece of HW, a process, etc.

intel0017 Sect: 4.2 Pg: 9 Ln: 29  
Should we add 'and validate' to 'Determine'?

intel0018 Sect: 4.2 Pg: 9 Ln: 50 C  
We need a similar diagram for channel teardown.

intel0019 Sect: 4.2 Pg: 10 Ln: 17  
(Many places in this clause)  
Some formatting is needed to set off model-specific terms such as "channel establishment failure response" - suggest bold or small caps. This would making parsing and understanding much easier.

intel0020 Sect: 4.2 Pg: 10 Ln: 22  
Given the vague definition of RDMA CS, it's hard to tell what 'internal to the RDMA communication service' does or does not mean.

intel0021 Sect: 4.2 Pg: 10 Ln: 26  
"An RDMA channel rejected response returns reject data"  
s/b "Rejection" data

intel0022 Sect: 4.2 Pg: 10 Ln: 28  
'With SRP the reject data includes' - near duplicate of page 11, Line 2

intel0023 Sect: 4.2 Pg: 10 Ln: 30  
'service specific data' s/b 'service-specific data' (global replace)

intel0024 Sect: 4.2 Pg: 10 Ln: 42  
'requests that are acceptable to the RDMA communication service shall be passed to the server agent.'  
(SHALL in model clause. )  
What does it mean to be acceptable to the service?  
As there is no mapping of 'Server Agent' to any entity, on what is this requirement placed?  
Can this requirement be stated in SRP or Annex B -specific terms?

intel0025 Sect: 4.2 Pg: 11 Ln: 2  
'reject(ion) data shall contain an SRP\_LOGIN\_REJ...' (SHALL)  
Do we need a subclause similar to '4.5 Ordering and Reliability' to capture size issues,  
so we can specify requirements on underlying interconnects? (e.g., Must be able to return  
\_REJ as part of connection establishment protocol.)

intel0026 Sect: 4.2 Pg: 11 Ln: 5  
'accept data' s/b 'acceptance data'

intel0027 Sect: 4.2 Pg: 11 Ln: 9

It is unclear how an RDMA comm svc requests that a channel be disconnected.

intel0028 Sect: 4.2 Pg: 11 Ln: 9

Need to discuss the case of a channel being destroyed due to an error.

intel0029 Sect: 4.2 Pg: 11 Ln: 13

'A disconnect request causes an RDMA channel to become non-operational.'

Is this a request by a consumer to the local CS provider, or to the remote client,

server agent, ... ?

intel0030 Sect: 4.2 Pg: 11 Ln: 16

'may or may not'

Since 'May' and 'May Not' are both defined to be equivalent to 'May or May Not',

there appears to be no reason to include both.

(global)

intel0031 Sect: 4.2 Pg: 11 Ln: 16

Suggest: 'The completion status of operations... is indeterminate.'

intel0032 Sect: 4.2 Pg: 11 Ln: 13

'disconnect request' s/b 'disconnection request'

(global)

intel0033 Sect: 4.3 Pg: 11 Ln: 18

'An RDMA channel may allow its consumers to exchange messages.'

One that did not would be useless for the present case, wouldn't it?

intel0034 Sect: 4.3 Pg: 11 Ln: 21

'may provide normal and solicited message reception notification,'

Since not used by SRP, why included?

intel0035 Sect: 4.3 Pg: 11 Ln: 24

'providing the following to an RDMA communication service'

Again, CS model issue - how do you provide this to a protocol?

intel0036 Sect: 4.4 Pg: 12 Ln: 23 C

'An RDMA communication service is not required to provide a way for a requesting

consumer to determine whether the data has been written into the specified range

of addresses in registered memory.'

If the target does not know whether a write has completed, how does it know when to

send status, and whether status is good or not?

intel0037 Sect: 4.5 Pg: 12 Ln: 45

'or else disconnect the RDMA channel.'

'destroy' is a better term to reflect the error case.

intel0038 Sect: 4.5 Pg: 13 Ln: 15

disconnect s/b destroy

intel0039 Sect: 5.1.1 Pg: 14 Ln: 20

'An SRP target port shall not accept a new RDMA channel unless its SRP target

port identifier matches the value in the SRP\_LOGIN\_REQ request.'

As we have not defined 'match', do we need to explicitly allow wildcards?

intel0040 Sect: 5.1.1 Pg: 14 Ln: 26

Additional - spelling

intel0041 Sect: 5.1.1 Pg: 14 Ln: 31

'Prior to requesting that an RDMA channel be disconnected,

an SRP initiator port may send an SRP\_I\_LOGOUT'

s/b SHALL send

intel0042 Sect: 5.1.1 Pg: 14 Ln: 44

'Prior to requesting that an RDMA channel be disconnected, an SRP target port should send an SRP\_T\_LOGOUT request'  
s/b 'SHALL send'

intel0043 Sect: 5.1.3 Pg: 15 Ln: 16

'Following acceptance of a login specifying single RDMA channel operation that single RDMA channel'  
Add comma after 'operation'

intel0044 Sect: 5.1.3 Pg: 15 Ln: 22

'shall not accept such a login'  
What \_REJ reason code is returned?

intel0045 Sect: 5.1.3 Pg: 15 Ln: 27

identifier

intel0046 Sect: 5.1.3 Pg: 15 Ln: 36

Break E.g. sentence into two or more sentences, or write as a note.

intel0047 Sect: 5.3 Pg: 16 Ln: 24

Suggest creating 5.3.1 Initiator Requests, and 5.3.2 Target Requests, to discuss separately.

'Many\_ reviewers have become confused with 'SRP target ports shall limit...'  
Add pointer to Table 7 and emphasis that these are 'target-initiated SRP requests, \_not\_ RDMA requests.

intel0048 Sect: 5.3 Pg: 16 Ln: 27

'credit based' s/b 'credit-based'

intel0049 Sect: 5.3 Pg: 16 Ln: 40 C

'An SRP initiator port shall not send an SRP request on any RDMA channel whose REQUEST LIMIT has a value less than or equal to zero.'  
What is Target Port response to this?

intel0050 Sect: 5.3 Pg: 16 Ln: 42 C

'To ensure that task management requests may be sent, an SRP initiator port may choose to send commands only when REQUEST LIMIT is greater than one'  
Since TargPort can remove an arbitrary number of credits at any time, Init Port can be prohibited from performing Task Mgmt or sending SRP\_I\_LOGOUT.

intel0051 Sect: 5.3 Pg: 16 Ln: 46 C

'An SRP initiator port shall add...whenever it receives an information unit on that RDMA channel'  
What does 'receive' mean?  
Received at what layer?  
There may be a significant delay between receiving and reading.

intel0052 Sect: 5.3 Pg: 16 Ln: 46 C

Target Port maintains, implicitly or explicitly, a value representing its view of the number of free request contexts (Call this Target Request Limit TRL) When there are no requests outstanding, TRL will be equal to the initiator's REQUEST LIMIT (IRL).

The description in 5.3 only describes IRL, but TRL may differ from IRL, and there is no definition of when IRL is changed. Specifically, when TargPort sends SRP\_CRED\_REQ with a negative value, when does TP update TRL?

It only makes sense to update upon receipt of SRP\_CRED\_RSP, but that is not stated.

Rewrite to describe with state variable at IP and at TP, and rules for



updating.

intel0053 Sect: 5.3 Pg: 16 Ln: 46 C

When can TPort be sure that IPort has seen the REQ\_LIMIT\_DELTA in an SRP\_RSP?  
(Receipt of transport ACK is not enough)

intel0054 Sect: 5.3 Pg: 17 Ln: 1

'An SRP target port shall not specify a negative value of REQUEST LIMIT  
DELTA that

might cause REQUEST LIMIT to drop below  $-2^{31}$ '

Given wrapping, it's impossible to drop below  $-2^{31}$  in 32-bit 2's comp.  
Would  $-2^{16}$  be negative enough?

intel0055 Sect: 5.3 Pg: 17 Ln: 1 C

'An SRP target port shall account for all possible race  
conditions to meet these requirements.'  
Remove this sentence.

intel0056 Sect: 5.4.1 Pg: 17 Ln: 6

'memory segment' and 'memory region' need to be defined before use.

intel0057 Sect: 5.4.1 Pg: 17 Ln: 6

'identifies the byte address'

Isn't the interpretation of a VA up to the particular interconnect/transport?

intel0058 Sect: 5.4.1 Pg: 17 Ln: 6

(Memory Handle) 'The SRP initiator port shall use this value to locate the  
region.'

It doesn't appear to be within our scope to define initiator memory controller  
implementations. Remove this sentence.

intel0059 Sect: 5.4.1 Pg: 17 Ln: 36

Drawing seems to indicate that memory addresses increase moving downward.  
Should be explicit.

intel0060 Sect: 5.4.1 Pg: 18 Ln: 2

'SRP target ports shall only issue the appropriate type of RDMA operation  
for a memory descriptor,'

Add: 'depending on whether the descriptor was a DATA-IN or DATA-OUT  
descriptor'

intel0061 Sect: 5.4.1 Pg: 18 Ln: 4

'a) The RDMA operations VIRTUAL ADDRESS shall be greater'

Should specify STARTING address.

Although VIRTUAL ADDRESS is a field name in Table 1, the field may have a  
different name in a particular interconnect's request format. Should not  
be in CAPS.

intel0062 Sect: 5.4.1 Pg: 18 Ln: 4

'Some data buffer descriptor format code values'  
s/b 'descriptor formats'

intel0063 Sect: 5.4.1 Pg: 18 Ln: 17

'use the contents of a count field to further specify the data buffer  
descriptor format.'  
specify -> describe

intel0064 Sect: 5.4.1 Pg: 18 Ln: 17 C

'use the contents of a count field to further specify the data buffer  
descriptor format.'

'count' is essentially a pointer to another field someplace, but this is far  
from

obvious when reading.

Suggest we define a format for 'virtual fields', e.g. '\*COUNT', or  
'vCOUNT', which

the reader could easily recognize. Clause 3 would contain a table allowing  
\*COUNT

to be looked up as

'SRP\_CMD DATA\_OUT BUFFER DESCRIPTOR COUNT or  
SRP\_CMD DATA\_IN BUFFER DESCRIPTOR COUNT, as appropriate'

intel0065 Sect: 5.4.1 Pg: 18 Ln: 24  
Remove period after PRESENT

intel0066 Sect: 5.4.1 Pg: 18 Ln: 32  
Note 'b' is not referenced above, probably s/b on 'count'

intel0067 Sect: 5.4.1 Pg: 18 Ln: 43  
'initiator port may specify in SRP\_CMD requests (see 6.8) sent on that RDMA channel. An SRP initiator port shall not specify a data buffer descriptor format that was not indicated in the REQUIRED BUFFER FORMATS field value for that RDMA channel.'  
What is target response if it does?

'SRP target ports are not required to check SRP\_CMD requests for data buffer descriptor formats that were not indicated in the REQUIRED BUFFER FORMATS field value.'  
Not clear - are they required to validate that they did a valid format?

intel0068 Sect: 5.4.1 Pg: 18 Ln: 47  
'An SRP target port may accept an RDMA channel and s/b 'channel establishment request'

intel0069 Sect: 5.4.2.2 Pg: 18 Ln: 49  
shall reject the RDMA channel and return after channel, add 'establishment request'

intel0070 Sect: 5.4.2.2 Pg: 19 Ln: 16  
indirect data buffer descriptor (IDBD)  
Use caps or formatting to set off these field names

intel0071 Sect: 5.4.2.2 Pg: 19 Ln: 16  
if the SRP initiator port may specify the INDIRECT s/b 'if the TP will accept...'

intel0072 Sect: 5.4.2.2 Pg: 19 Ln: 18  
does not use  
(Sense is that IP forebears use of indirect) shall not use?

intel0073 Sect: 5.4.2.4 Pg: 19 Ln: 44  
'sixteen bytes'  
Previously defined in Table 2 - eschew multiple definitions

intel0074 Sect: 5.4.2.4 Pg: 19 Ln: 48  
target port shall only issue RDMA Read operations using the memory descriptor contained in the direct data buffer descriptor.  
Statement does not have desired effect - limits what you can read, but does not limit accesses to READS.  
s/b 'shall issue only RDMA Reads when using'

intel0075 Sect: 5.4.2.4 Pg: 20 Ln: 1  
shall issue only RDMA Writes...

intel0076 Sect: 5.4.2.5 Pg: 20 Ln: 6  
format code value  
'value' appears to be superfluous

intel0077 Sect: 5.4.2.5 Pg: 20 Ln: 8  
'The length... sixteen bytes.'  
Drop sentence - redundant to Table 2

intel0078 Sect: 5.4.2.5 Pg: 20 Ln: 10  
'An indirect data buffer is comprised of one or more memory segments'  
Need a real definition.

intel0079 Sect: 5.4.2.5 Pg: 20 Ln: 11  
 segments may or may not be contiguous.  
 s/b 'may be discontinuous'

intel0080 Sect: 5.4.2.5 Pg: 20 Ln: 11  
 remove 'may be in a single memory region'

intel0081 Sect: 5.4.2.5 Pg: 20 Ln: 12  
 of the memory segments (ADD: listed in an IBDB)

intel0082 Sect: 5.4.2.5 Pg: 20 Ln: 13  
 may have any length  
 As the length field is finite, so is the segment length

intel0083 Sect: 5.4.2.5 Pg: 20 Ln: 29  
 value contained in the data buffer descriptor's count field.  
 Implies that the field is contained within the DBD

intel0084 Sect: 5.4.2.5 Pg: 20 Ln: 29  
 'count'  
 Suggest replacing with 'PMDL Length'

intel0085 Sect: 5.4.2.5 Pg: 20 Ln: 31  
 DESCRIPTOR field value is a memory descriptor  
 Suggest: DESCRIPTOR field contains a memory descriptor

intel0086 Sect: 5.4.2.5 Pg: 20 Ln: 33  
 concatenated together  
 'together' is redundant  
 Stamp Out and Abolish Redundancy!

intel0087 Sect: 5.4.2.5 Pg: 20 Ln: 35  
 contains the number of memory descriptors in the indirect table times  
 sixteen.  
 Suggest: contains the length, in bytes, of the indirect table (16 bytes \*  
 number of  
 descriptors in table)

intel0088 Sect: 5.4.2.5 Pg: 20 Ln: 36  
 MEMORY DESCRIPTOR field value contains any other  
 drop 'value'

intel0089 Sect: 5.4.2.5 Pg: 20 Ln: 43  
 list of n memory descriptors  
 Use bold or something to set off n

intel0090 Sect: 5.4.2.5 Pg: 20 Ln: 47  
 shall only issue  
 s/b shall issue only

intel0091 Sect: 5.4.2.5 Pg: 21 Ln: 1  
 shall only issue  
 s/b shall issue only  
 (also Ln 4)

intel0092 Sect: 5.4.2.5 Pg: 21 Ln: 12  
 All four... , each might... , or several might be...  
 Awkward - generalize to: segments may be in different memory regions

intel0093 Sect: 5.4.2.5 Pg: 21 Ln: 44  
 value contains  
 Drop: value  
 ( i.e., )  
 Add: in bytes

intel0094 Sect: 6.1 Pg: 23 Ln: 7  
 Add M/O column, or statement that all are mandatory.

intel0095 Sect: 6.1 Pg: 23 Ln: 24

Add space between Tables 6 and 7 to clarify distinction between I>T and T>I requests.

intel0096 Sect: 6.1 Pg: 23 Ln: 46

shall send SRP\_T\_LOGOUT

What reason code?

intel0097 Sect: 6.1 Pg: 23 Ln: 48

Need to define requestor, responder.

Much reviewer confusion wrt Targ as requestor.

intel0098 Sect: 6.1 Pg: 24 Ln: 2 C

Responders are not required to check whether the TAG values of outstanding

SRP requests are unique.

Since duplicate tags would likely cause a credit leak (one response for two requests),

this could lead to deadlock, as InitReqLimit and TRL would be out of sync.

We either need to require verification of uniqueness, or provide a ReqLimit re-sync mechanism.

intel0099 Sect: 6.2 Pg: 24 Ln: 2

shall only be sent during RDMA

s/b: shall be sent only during RDMA

intel0100 Sect: 6.2 Pg: 24 Ln: 41

maximum length

Add: in bytes

intel0101 Sect: 6.3 Pg: 27 Ln: 4

shall only be sent

s/b: shall be sent only

intel0102 Sect: 6.3 Pg: 27 Ln: 40

maximum length

Add: in bytes

intel0103 Sect: 6.3 Pg: 27 Ln: 45 C

52 or larger

AER\_REQ requires 56

intel0104 Sect: 6.4 Pg: 29 Ln: 3

a(n) SRP target

intel0105 Sect: 6.4 Pg: 29 Ln: 40

too large

Need a way to specify, so that Init does not have to guess

intel0106 Sect: 6.5 Pg: 30 Ln: 20

delay a vendor specific time

Wait for transport ACK or timeout error at least

intel0107 Sect: 6.6 Pg: 31 Ln: 3 C

An SRP\_T\_LOGOUT request may also be used to notify the SRP initiator port that an RDMA channel has failed, rendering it non-operational.

If the channel has failed, it won't be able to carry this IU.

We DO need a way to report failures.

intel0108 Sect: 6.6 Pg: 31 Ln: 30

There are no references in spec to reason codes 2, 3, 6-9.

Do we need some SHALLS pointing to them?

intel0109 Sect: 6.6 Pg: 31 Ln: 45

delay a vendor...

Reference: xport ack or timeout

intel0110 Sect: 6.8 Pg: 34 Ln: 14

COUNT

Change to PMDL Length

intel0111 Sect: 6.8 Pg: 34 Ln: 40  
Add ref a,b to notes below

intel0112 Sect: 6.9 Pg: 36 Ln: 36  
Since SENSE DATA length is 7 bytes + a one-byte length field, at least the top two bytes s/b reserved. We may want to have this field be that one-byte length field, with 7 assumed, as in SPC.

intel0113 Sect: 6.9 Pg: 37 Ln: 9  
length of the...buffer  
Ref 5.4 for length determination

intel0114 Sect: 6.9 Pg: 37 Ln: 26  
indicates (that) the contents ....shall be ignored and (that) the

intel0115 Sect: 6.9 Pg: 37 Ln: 26  
The(value of the) SENSE DATA LIST LENGTH field (be a multiple of four).

intel0116 Sect: 6.9 Pg: 37 Ln: 26 C  
SENSE DATA LIST LENGTH shall contain the length of the truncated SENSE DATA field.  
This is at odds with SPC-2, which returns the total length. How would you know that you had missed some Sense Data?

intel0117 Sect: 6.9 Pg: 37 Ln: 47  
shall contain a length of 4  
Also defined in Table 23 - refer to table instead

intel0118 Sect: 6.9 Pg: 38 Ln: 17  
structure eqn as DOBL - (offset + 1)  
Much easier to understand  
(global change to all similar eqns)  
Formatting - more white space above and below, use bold font

intel0119 Sect: 6.9 Pg: 38 Ln: 24  
may or may not  
not is the more worrisome case  
(more so for Ln 25)

intel0120 Sect: 6.9 Pg: 38 Ln: 31  
Some commands may have a non-zero residual  
Add: e.g., INQUIRY

intel0121 Sect: 6.9 Pg: 38 Ln: 45  
may not

intel0122 Sect: 6.9 Pg: 39 Ln: 1  
certian (SRP) protocol errors

intel0123 Sect: 6.9 Pg: 39 Ln: 18  
Would there ever be a case where a RSP of NO FAILURE was returned?

intel0124 Sect: 6.9 Pg: 39 Ln: 31  
sense data shall be presented  
presented s/b returned  
Also Ln 32,33

intel0125 Sect: 6.9 Pg: 39 Ln: 33  
whose  
Use whose wrt people only

intel0126 Sect: 6.9 Pg: 39 Ln: 30  
SPC-2  
Annex C references SPC-3 - which?

intel0127 Sect: 6.11 Pg: 40 Ln: 43 C  
See comments on 5.3 for CRED\_RSP issues

intel0128 Sect: 6.12 Pg: 41 Ln: 31

Don't need four bytes for SENSE data length (7 + 1 byte)

intel0129 Sect: 6.12 Pg: 41 Ln: 43

The (value of) the SENSE DATA Len field (shall be a multiple of four.)

intel0130 Sect: 6.12 Pg: 41 Ln: 44 C

If no sense data is provided,

What would the point be - to force Init to issue Req Sense Request?

Should this be allowed?

intel0131 Sect: 6.12 Pg: 42 Ln: 1

SENSE DATA LIST LENGTH shall contain the length of the truncated SENSE DATA field.

Appears to violate SPC-2.

intel0132 Sect: 6.12 Pg: 42 Ln: 7

presented s/b (returned in response to)

intel0133 Sect: 7.2 Pg: 43 Ln: 21

The following subclause defines the fields in the disconnect-reconnect mode

Nope - same subclause

intel0134 Sect: 7.2 Pg: 43 Ln: 25

Gray-out or mark as Reserved the fields that are reserved for SRP.

There's a lot of noise for the two fields that are used...

intel0135 Sect: 7.2 Pg: 44 Ln: 1

SRP devices shall only use (the) disconnect-reconnect page parameter fields

Use formatting for disconnect-reconnect

intel0136 Sect: 7.2 Pg: 44 Ln: 1

SRP devices shall only use ... fields defined below.

What about the standard mode page header fields?

intel0137 Sect: 7.2 Pg: 44 Ln: 7

field shall not be implemented by SRP target ports

Define in terms of behavior, not implementation. Appears to have been covered by para above.

intel0138 Sect: 7.2 Pg: 44 Ln: 17

If the EMDP bit is set to 0, the SRP target port shall generate (RDMA requests with)

continuously increasing () addresses for a single SCSI command.

intel0139 Sect: 7.2 Pg: 44 Ln: 19

affect the order of frames within an RDMA.

What's a frame? Within an RDMA what?

intel0140 Sect: 7.2 Pg: 44 Ln: 24

intel0141 Sect: 7.2 Pg: 44 Ln: 24

protocol specific

s/b protocol-specific

(also Ln 27)

intel0142 Sect: 7.2 Pg: 44 Ln: 28

LUN -> PORT

intel0143 Sect: A.1 Pg: 45 Ln: 11

Top right box s/b Device Server?

intel0144 Sect: A.1 Pg: 45 Ln: 29

four step, two step

s/b four-step, two-step

(global)

intel0145 Sect: A.1 Pg: 46 Ln: 16  
Need close paren after initiator

intel0146 Sect: A.1 Pg: 46 Ln: 43  
Close paren

intel0147 Sect: A.3 Pg: 47 Ln: 11  
See table A.1 for the definitions of the names used within  
Don't see names there - objects?

intel0148 Sect: A.4.1 Pg: 48 Ln: 44  
Use bold for EXECUTE COMMAND

intel0149 Sect: B.3.1.7 Pg: 52 Ln: 35  
IBTA uses caps for G S I

intel0150 Sect: B.3.1.2 Pg: 52 Ln: 23  
Do we need to define, spell out GUID?

intel0151 Sect: B.3.1.14 Pg: 53 Ln: 1  
Ports also present on switches.

intel0152 Sect: B.3.1.16 Pg: 53 Ln: 5  
Spell out QP, use IBTA definition.

intel0153 Sect: B.3.2 Pg: 53 Ln: 20  
IBTA uses caps for R T U

intel0154 Sect: B.4 Pg: 54 Ln: 50 C  
Each IB GUID is globally unique,  
Not true - see IBA Vol 1 4.1.1

intel0155 Sect: B.4 Pg: 55 Ln: 17  
worldwide  
Varies - see IBA Vol 1, 4.1.1

intel0156 Sect: B.5 Pg: 56 Ln: 2  
An SRP initiator device is one or more IB consumers  
may consist of

intel0157 Sect: B.5 Pg: 56 Ln: 15  
The GUID field should (be) an IB GUID available to the SRP initiator  
port,  
Must it be a GUID, an IB GUID, ....?

intel0158 Sect: B.5 Pg: 56 Ln: 17  
The IDENTIFIER EXTENSION field shall be chosen by the SRP initiator  
port to ensure that all SRP initiator port identifiers are unique.  
Over what domain?

intel0159 Sect: B.5 Pg: 56 Ln: 36  
[containing] the SRP target port.  
providing?

intel0160 Sect: B.5 Pg: 56 Ln: 41  
The service delivery subsystem contains queue pairs, IB channel  
adapters, IB ports, and the InfiniBand TM Architecture fabric.  
Contains exclusively?  
How does this map to Clause 4 RDMA Comm Service?

intel0161 Sect: B.5 Pg: 56 Ln: 47  
general service interface  
IBTA uses caps

intel0162 Sect: B.5 Pg: 56 Ln: 48  
I/ (breaks across page) 0  
Remove slash from FRAME list of characters for line breaks.

intel0163 Sect: B.6.2 Pg: 57 Ln: 13

open IBA connections  
use establish instead

intel0164 Sect: B. 6. 3 Pg: 57 Ln: 25  
Port and CM Redirection or Port Redirection.  
Very hard to parse - use bold or underscores inside the names

intel0165 Sect: B. 6. 4 Pg: 57 Ln: 38  
SRP\_LOGOUT IU  
list as T\_LOGOUT, I\_LOGOUT or define as a virtual field

intel0166 Sect: B. 6. 4 Pg: 57 Ln: 38  
CM disconnect request  
use caps -it's not generic

intel0167 Sect: B. 6. 4 Pg: 57 Ln: 38  
The sender may disconnect if its send queue has transitioned to (THE)  
error state.  
What do you mean by disconnect here - local action?

intel0168 Sect: B. 6. 4 Pg: 57 Ln: 42  
The receiver of an SRP\_LOGOUT IU shall respond with an InfiniBand TM  
Architecture transport acknowledgement and disconnect.  
Destroy QP, send DREQ, ...?

intel0169 Sect: B. 6. 5 Pg: 57 Ln: 46  
to an ... RDMA READ Request.  
One or more requests.

intel0170 Sect: B. 6. 5 Pg: 58 Ln: 1  
WRITE packets  
WRITE requests

intel0171 Sect: B. 7 Pg: 58 Ln: 37  
outcommands

intel0172 Sect: B. 7 Pg: 59 Ln: 7  
Why list ChangelD and OptionROM to say we don't care about them?

intel0173 Sect: B. 7 Pg: 60 Ln: 23 C  
Send Message Depth  
Reserved -> Maximum Initiator Request Limit  
This allows initiators to efficiently allocate buffers

intel0174 Sect: B. 7 Pg: 60 Ln: 24 C  
RDMA Read Depth  
reserved -> Maximum IOC-issued RDMA depth  
Allows inits to efficiently allocate RDMA resources

intel0175 Sect: B. 7 Pg: 60 Ln: 26 C  
Send Message Size  
rsvd -> MAXIMUM INITIATOR TO TARGET IU SIZE  
Eliminates need to guess this value

intel0176 Sect: B. 7 Pg: 60 Ln: 46 C  
This field is expected to be marked obsolete in future versions of the  
InfiniBand TM Architecture  
Not for T10/ANSI to say

intel0177 Sect: B. 7 Pg: 61 Ln: 13 C  
Is :reserved a literal?  
If not, express as :zzzz, explain below that it is reserved.

intel0178 Sect: B. 7 Pg: 61 Ln: 16  
No references to Table notes.

intel0179 Sect: B. 7 Pg: 61 Ln: 16  
padded  
s/b extended



Comment from IBTA by William Futral (Intel):

The IBTA Application Working Group understands that the SRP document is out for review and would like to offer the following comment.

The value assigned to I/O Class field in Table B.7 of the SRP document needs to be changed as a result of a change made to the format of this component in the latest InfiniBand(TM) Identifiers Annex, which is a supplement to InfiniBand(TM) Architecture Specification Volume 1.

Attached is a PDF document that contains the new wording in the IBTA Annex (see T10/01-319).

A Class Category needs to be selected for the SRP protocol and inserted in the I/O class field in place of the 0xFF value currently stated. For example, if the Storage Class was selected, the value for I/O class in your Table b.7 would become 0x0100.

Bill Futral  
Application Working Group Co Chair  
InfiniBand Trade Association

\*\*\*\*\*

Comments attached to No ballot from Mr. Edward A. Gardner of Ophi dian Design s:

see 01-325

\*\*\*\*\*

Comments attached to YesC ballot from Mr. Paul D. Aloisi of Texas Instruments:

This has the appearance of a draft copy, not a final review copy. Change bars and line numbers should not be on a letter ballot document.

\*\*\*\*\*

Comments attached to YesC ballot from Mr. William C. Terrell of Troika Networks, Inc.:

1. The TYPE code value of 80h in tble 13 is incorrect according to table 6 and should be value C2h.

\*\*\*\*\*

Comments attached to Abs ballot from Mr. Doug Piper of Woven Electronics:

Can not Contri bute

\*\*\*\*\* End of Ballot Report \*\*\*\*\*