Working Draft American National Standard

Project T10/BSR INCITS xxx

Revision 01
12 July 2016

Information technology - SCSI RDMA Protocol-2 (SRP-2)

T10 Technical Editor: David Landsman
Western Digital
951 SanDisk Drive
Milpitas, CA 95035-7933
USA

Telephone: (206) 275-4385
Email: dave.landsman@sandisk.com

Reference number
ISO/IEC xxxx-xxx:200x
BSR INCITS *-201x**

Printed 8:50 AM Tuesday 12 July 2016

Points of Contact

InterNational Committee for Information Technology Standards (INCITS) T10 Technical Committee

T10 Chair
Ralph O. Weber
Western Digital Corp.
18484 Preston Road, Suite 102, PMB 178
Dallas, TX 75252
USA

Telephone: (214) 912-1373
Email: Ralph.Weber@WDC.com

T10 Vice-Chair
William Martin
Samsung Semiconductor, Inc
7213 Marblethorpe Drive
Roseville, CA 95747-5925
USA

Telephone: (916) 765-6875
Email: bill.martin@ssi.samsung.com

T10 Web Site: http://www.t10.org

IINCITS Secretariat
Suite 610
1101 K Street, NW
Washington, DC 20005-7031
USA

Telephone: (202) 737-8888
Web site: http://www.incits.org
Email: incits@itic.org

Information Technology Industry Council
Web site: http://www.itic.org

Purchase INCITS Standards
Web site: http://www.incits.org/standards-information/purchase-standards-or-download-dpans

American National Standard
for Information Technology

SCSI RDMA Protocol-2 (SRP-2)

Secretariat
Information Technology Industry Council

Approved mm.dd.yy

American National Standards Institute, Inc.

ABSTRACT

This standard describes the message format and protocol definitions required to transfer commands and data between a SCSI (Small Computer System Interface) initiator port and a SCSI target port using an RDMA communication service.

American National Standard

Approval of an American National Standard requires verification by ANSI that the requirements for due process, consensus, and other criteria for approval have been met by the standards developer. Consensus is established when, in the judgment of the ANSI Board of Standards Review, substantial agreement has been reached by directly and materially affected interests. Substantial agreement means much more than a simple majority, but not necessarily unanimity. Consensus requires that all views and objections be considered, and that effort be made towards their resolution.

The use of American National Standards is completely voluntary; their existence does not in any respect preclude anyone, whether he has approved the standards or not, from manufacturing, marketing, purchasing, or using products, processes, or procedures not conforming to the standards.

The American National Standards Institute does not develop standards and will in no circumstances give interpretation on any American National Standard. Moreover, no person shall have the right or authority to issue an interpretation of an American National Standard in the name of the American National Standards Institute. Requests for interpretations should be addressed to the secretariat or sponsor whose name appears on the title page of this standard.

CAUTION NOTICE: This American National Standard may be revised or withdrawn at any time. The procedures of the American National Standards Institute require that action be taken periodically to reaffirm, revise, or withdraw this standard. Purchasers of American National Standards may receive current information on all standards by calling or writing the American National Standards Institute.

CAUTION: The developers of this standard have requested that holders of patents that may be required for the implementation of the standard, disclose such patents to the publisher. However, neither the developers nor the publisher have undertaken a patent search in order to identify which, if any, patents may apply to this standard. As of the date of publication of this standard, following calls for the identification of patents that may be required for the implementation of the standard, no such claims have been made. No further patent search is conducted by the developer or the publisher in respect to any standard it processes. No representation is made or implied that licenses are not required to avoid infringement in the use of this standard.

Published by

American National Standards Institute
25 West 43rd Street 4th floor, New York, New York 10036-7422

Printed in the United States of America

Revision Information

R.1 Revision SRP-2r00 (23 December 2015)

This is revision 0 of this working draft. The project proposal was 16-xxxr0.

This revision incorporated these:

a) the published ANSI INCITS 365-2002 (SRP);
b) 03-057r1 - SRP-2 Immediate Data Proposal [Simpson] (see Minutes of T10 Plenary Meeting #53 - January 16, 2003 (03-048r0) for approval vote; and
c) removed SRP_AER_REQ request and SRP_AER_RSP response information units after the removal of asynchronous event reporting from SAM-3 (see T10/02-457r0).

Contents

Tables

Figures

FOREWORD (This foreword is not part of this standard)

Requests for interpretation, suggestions for improvement and addenda, or defect reports are welcome. They should be sent to the INCITS Secretariat, International Committee for Information Technology Standards, Information Technology Industry Council, 1101 K Street, Suite 610, NW, Washington, DC 20005-7031.

This standard was processed and approved for submittal to ANSI by the International Committee for Information Technology Standards (INCITS). Committee approval of the standard does not necessarily imply that all committee members voted for approval. At the time it approved this standard, INCITS had the following members:

INCITS Technical Committee T10 on SCSI Storage Interfaces, which developed and reviewed this standard, had the following members:

Ralph O. Weber, Chair
William Martin, Vice-Chair
John Geldman, Secretary
~~John B. Lohmeyer, Chair~~
~~William Martin, Vice-Chair~~
~~Ralph O. Weber, Secretary~~

INTRODUCTION

General

The Small Computer System Interface (SCSI) command set is widely used and applicable to a wide variety of device types. The transmission of SCSI command set information across an RDMA communication service allows the large body of SCSI application and driver software to be successfully used on the InfiniBand™ [1.] Architecture, the VI Architecture and other interfaces that support RDMA communication service semantics.

The this standard is divided into the following clauses:

Clause 1 is the scope.
Clause 2 enumerates the normative references that apply to this standard.
Clause 3 describes the definitions, symbols, abbreviations, and conventions used in this standard.
Clause 4 describes the RDMA communication service model.
Clause 5 describes significant concepts of SRP.
Clause 6 describes the information units used by SRP.
Clause 7 defines the SCSI management features for SRP, including the SRP mode pages.Annex

Annex A and Annex B form an integral part of this standard.[1]

1.InfiniBand is a trademark and service mark of the InfiniBand Trade Association

American National Standard
for Information Technology -

SCSI RDMA Protocol-2 (SRP-2)

1 Scope

The SCSI family of standards provides for many different transport protocols that define the rules for exchanging information between different SCSI devices. This standard defines the rules for exchanging information between SCSI devices using a serial interconnect. Other SCSI transport protocol standards define the rules for exchanging information between SCSI devices using other interconnects.

Figure 1 shows the relationship of this standard to the other standards and related projects in the SCSI family of standards as of the publication of this standard.

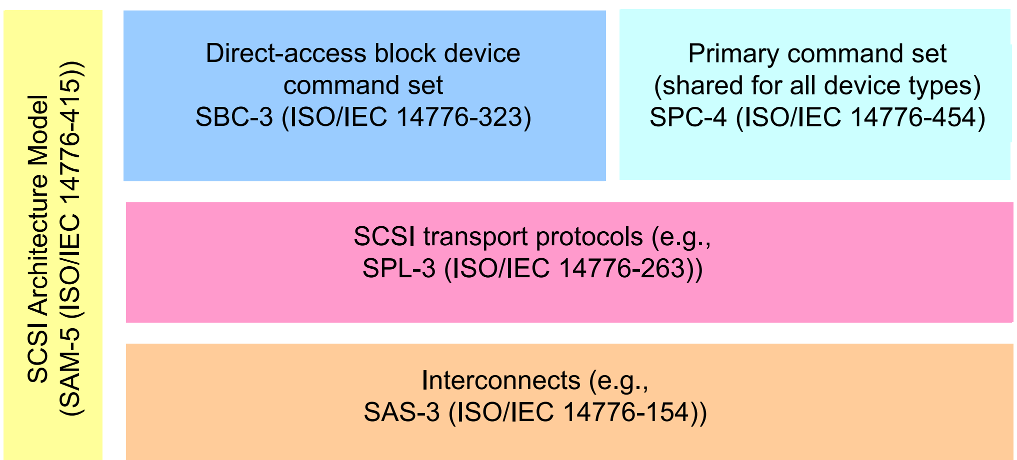


Figure 1 — SCSI document structure

Figure 1 shows the general relationship of the documents to one another, and do not imply any hierarchy, protocol stack, or system architecture relationship.

These standards specify the interfaces, functions and operations necessary to ensure interoperability between conforming implementations. This standard is a functional description. Conforming implementations may employ any design technique that does not violate interoperability.

This standard has made obsolete the following:

a) SRP_AER_REQ request;
b) SRP_AER_RSP response; and
c) asynchronous event solicited notification (AESOLNT) bit.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 14776-454, *SCSI Primary Commands - 4 (SPC-4)*

IISO/IEC 14776-415, *SCSI Architecture Model - 5 (SAM-5)*

3 Definitions, symbols, abbreviations, and conventions

3.1 Definitions

3.1.1 acceptance data
application protocol data communicated from a server consumer to the client consumer when a new RDMA channel is accepted (see 4.2)

Note 1 to entry: This protocol uses acceptance data to communicate the SRP_LOGIN_RSP response (see 6.3).

3.1.2 application client
object that is the source of SCSI commands and task management function requests (see SAM-5)

3.1.3 byte
sequence of eight contiguous bits considered as a unit

3.1.4 channel attributes
information provided during RDMA channel establishment that identifies the type and characteristics of the desired RDMA channel (see 4.2)

Note 1 to entry: The format and interpretation of channel attributes are specific to a particular RDMA communication service.

3.1.5 command
request describing a unit of work to be performed by a device server (see SAM-5).

3.1.6 command descriptor block (CDB)
structure used to communicate commands from an application client to a device server

Note 1 to entry: See SAM-5.

3.1.7 consumer
object that communicates with other consumers using an RDMA communication service (see 4.1)

Note 1 to entry: In this protocol, a consumer is either an SRP target port or an SRP initiator port.

3.1.8 data-in buffer
buffer identified by the application client to receive data from the device server during the processing of a command

Note 1 to entry: See SAM-5.

3.1.9 data-out buffer
buffer identified by the application client to supply data that is sent from the application client to the device server during the processing of a command

Note 1 to entry: See SAM-5.

3.1.10 device server
object that processes SCSI commands

Note 1 to entry: See SAM-5.

3.1.11 expected
behavior of the hardware or software in the design models assumed by this standard

Note 1 to entry: Other hardware and software design models may also be implemented.

3.1.12 ignored
unused bit, byte, word, field, or code value

Note 1 to entry: The contents or value of an ignored bit, byte, word, field, or code value shall not be examined by the receiving SCSI device and may be set to any value by the transmitting SCSI device.

3.1.13 information unit
organized collection of data specified by this protocol to be transferred as login data, rejection data, acceptance data, or a message on an RDMA channel

3.1.14 initiator port identifier
value by which a SCSI initiator port is referenced within a domain

Note 1 to entry: See SAM-5.

3.1.15 logical unit
object that implements a device model and processes SCSI commands sent by an application client

Note 1 to entry: See SAM-5.

3.1.16 logical unit number (LUN)
identifier for a logical unit

Note 1 to entry: See SAM-5.

3.1.17 login data
data communicated from a client consumer to a server agent or server consumer during RDMA channel establishment (see 4.2) that is meaningful within the client/server application protocol

Note 1 to entry: This protocol uses login data to communicate the SRP_LOGIN_REQ request (see 6.2).

3.1.18 message
communication sent by one consumer to another using an RDMA channel

Note 1 to entry: See 4.3.

3.1.19 RDMA channel
communication path between two consumers of an RDMA communication service (see 4.1)

3.1.20 RDMA communication service
software, protocols, and interconnect that provides message and RDMA operations between pairs of consumers (see clause 4)

3.1.21 RDMA operation
either an RDMA Read operation or an RDMA Write operation

3.1.22 RDMA Read operation
operation by which a requesting consumer may fetch data from memory registered by the other consumer associated with an RDMA channel (see 4.2)

Note 1 to entry: See 4.4.

3.1.23 RDMA Write operation
operation by which a requesting consumer may store data into memory registered by the other consumer associated with an RDMA channel (see 4.2)

Note 1 to entry: See 4.4.

3.1.24 rejection data
application protocol data communicated from a server agent or server consumer to the client consumer when a new RDMA channel is rejected (see 4.2)

Note 1 to entry: This protocol uses rejection data to communicate the SRP_LOGIN_REJ response (see 6.4).

3.1.25 sense data
data describing command completion information that a device server delivers to an application client in the same I_T nexus transaction as the status or as parameter data in response to a REQUEST SENSE command

Note 1 to entry: In this protocol sense data is returned to an application client in the SENSE DATA field of an SRP_RSP response ~~or an SRP_AER_REQ request~~.

Note 2 to entry: See SAM-5.

3.1.26 server agent
entity that provides services on behalf of a server consumer

Note 1 to entry: An example of a service is connection management.

3.1.27 server identifier
information provided to an RDMA communication service by a client consumer that allows the RDMA communication service to locate the desired server consumer

Note 1 to entry: The format and interpretation of a server identifier are specific to the RDMA communication service.

3.1.28 SRP initiator port
SCSI initiator port that uses this protocol to communicate with an SRP target port

3.1.29 SRP initiator port identifier
value by which an SRP initiator port is identified to an SRP target port

3.1.30 SRP target port
SCSI target port that uses this protocol to communicate with an SRP initiator port

3.1.31 SRP target port identifier
value by which an SRP target port is identified within an SRP domain

3.1.32 status
response information sent from a device server to an application client upon completion of each command

Note 1 to entry: See SAM-5.

3.1.33 target port identifier
value by which a SCSI target port is referenced within a domain

Note 1 to entry: See SAM-5.

3.2 Symbols and abbreviations

3.2.1 Abbreviations

Abbreviations used in this standard:

Abbreviation	Meaning
CDB	Command Descriptor Block
LSB	Least significant bit
LUN	Logical unit number
MSB	Most significant bit
RDMA	Remote Direct Memory Access
SAM-5	SCSI Architecture Model-5 (see clause 2)
SCSI	Small Computer System Interface family of standards
SPC-4	SCSI Primary Commands-4 (see clause 2)

3.2.2 Mathematical operators

Mathematical operators used in this standard:

Mathematical Operators	Meaning
×	multiplication
+	add
-	subtract

3.3 Keywords

3.3.1 invalid
keyword used to describe an illegal or unsupported bit, byte, word, field, or code value

Note 1 to entry: Receipt by a device server of an invalid bit, byte, word, field, or code value shall be reported as an error.

3.3.2 mandatory
keyword indicating an item that is required to be implemented as defined in this standard

3.3.3 may
keyword that indicates flexibility of choice with no implied preference

3.3.4 may not
keyword that indicates flexibility of choice with no implied preference

3.3.5 obsolete
keyword indicating that an item was defined in prior SCSI standards but has been removed from this standard

3.3.6 option, optional
keywords that describe features that are not required to be implemented by this standard

Note 1 to entry: If any optional feature defined by this standard is implemented, then it shall be implemented as defined in this standard.

3.3.7 prohibited
keyword used to describe a feature, function, or coded value that is defined in a non-SCSI standard (i.e., a standard that is not a member of the SCSI family of standards) to which this standard makes a normative reference where the use of said feature, function, or coded value is not allowed for implementations of this standard

3.3.8 reserved
keyword referring to bits, bytes, words, fields, and code values that are set aside for future standardization

Note 1 to entry: A reserved bit, byte, word, or field shall be set to zero, or in accordance with a future extension to this standard.

Note 2 to entry: Recipients are not required to check reserved bits, bytes, words, or fields for zero values.

Note 3 to entry: Receipt of reserved code values in defined fields shall be reported as error.

3.3.9 restricted
keyword referring to bits, bytes, words, and fields that are set aside for other identified standardization purposes

Note 1 to entry: A restricted bit, byte, word, or field shall be treated as a reserved bit, byte, word, or field in the context where the restricted designation appears.

3.3.10 shall
keyword indicating a mandatory requirement

Note 1 to entry: Designers are required to implement all such mandatory requirements to ensure interoperability with other products that conform to this standard.

3.3.11 should
keyword indicating flexibility of choice with a strongly preferred alternative

3.3.12 vendor specific
something (e.g., a bit, field, code value) that is not defined by this standard

Note 1 to entry: Specification of the referenced item is determined by the SCSI device vendor and may be used differently in various implementations.

3.4 Editorial conventions

Certain words and terms used in this standard have a specific meaning beyond the normal English meaning. These words and terms are defined either in the glossary or in the text where they first appear.

Upper case is used when referring to the name of a numeric value defined in this specification or a formal attribute possessed by an entity. When necessary for clarity, names of objects, procedure calls, arguments or discrete states are capitalized or set in bold type. Names of fields are identified using small capital letters (e.g., NACA bit).

Names of procedure calls are identified by a name in bold type (e.g., **Execute Command**). Names of arguments are denoted by capitalizing each word in the name (e.g., Sense Data is the name of an argument in the **Execute Command** procedure call). For more information on procedure calls see 3.7.

Quantities having a defined numeric value are identified by large capital letters (e.g., CHECK CONDITION). Quantities having a discrete but unspecified value are identified using small capital letters. (e.g., TASK COMPLETE, indicates a quantity returned by the **Execute Command** procedure call). Such quantities are associated with an event or indication whose observable behavior or value is specific to a given implementation standard.

Lists sequenced by lowercase or uppercase letters show no ordering relationship between the listed items.

EXAMPLE 1 - The following list shows no relationship between the named items:

a) various forms of red such as:
    A) crimson; or
    B) amber;
b) blue; or
c) green.

Lists sequenced by numbers show an ordering relationship between the listed items.

EXAMPLE 2 -The following list shows an ordered relationship between the named items:

1) top;
2) middle; and
3) bottom.

Lists are associated with an introductory paragraph or phrase, and are numbered relative to that paragraph or phrase (i.e., all lists begin with an a) or 1) entry).

If a conflict arises between text, tables, or figures, then the order of precedence to resolve the conflicts is:

1) text;
2) tables; and
3) figures.

Not all tables or figures are fully described in the text.

Notes and examples do not constitute any requirements.

Notes are numbered consecutively throughout this standard.

3.5 Numeric and character conventions

3.5.1 Numeric conventions

A binary number is represented in this standard by any sequence of digits comprised of only the Arabic numerals 0 and 1 immediately followed by a lower-case b (e.g., 0101b). Underscores or spaces may be included in binary number representations to increase readability or delineate organizational boundaries (e.g., 00010101 11001110b, 00010101_11001110b, 0 0101 1010b or 0_0101_1010b).

A hexadecimal number is represented in this standard by any sequence of digits comprised of only the Arabic numerals 0 to 9 and/or the upper-case English letters A to F immediately followed by a lower-case h (e.g., FA23h). Underscores or spaces may be included in hexadecimal number representations to increase readability or delineate organizational boundaries (e.g., 3456FDCA 84BD5E7Ah, 3456FDCA_84BD5E7Ah, B FD8C FA23h, or B_FD8C_FA23h).

A decimal number is represented in this standard by any sequence of digits comprised of only the Arabic numerals 0 to 9 not immediately followed by a lower-case b or lower-case h (e.g., 25).

This standard uses the following conventions for representing decimal numbers:

a) the decimal separator (i.e., separating the integer and fractional portions of the number) is a period;
b) the thousands separator (i.e., separating groups of three digits in a portion of the number) is a space;
c) the thousands separator is used in both the integer portion and the fraction portion of a number; and
d) the decimal representation for a year is 1999 not 1 999.

Table 1 shows some examples of decimal numbers using various conventions.

Table 1 — Numbering conventions

French	English	This standard
0,6	0.6	0.6
3,141 592 65	3.14159265	3.141 592 65
1 000	1,000	1 000
1 323 462,95	1,323,462.95	1 323 462.95

A decimal number represented in this standard with an overline over one or more digits following the decimal point is a number where the overlined digits are infinitely repeating (e.g., $666.\overline{6}$ means 666.666 666... or 666 2/3, and $12.\overline{142\ 857}$ means 12.142 857 142 857... or 12 1/7).

A range of numeric values is represented in this standard in the form “a to z”, where a is the first value included in the range, all values between a and z are included in the range, and z is the last value included in the range (e.g., the representation “0h to 3h” includes the values 0h, 1h, 2h, and 3h).

3.5.2 Units of measure

This standard represents values using both decimal units of measure and binary units of measure. Values are represented by the following formats:

a) for values based on decimal units of measure:
   1) numerical value (e.g., 100);
   2) space; and
   3) prefix symbol and unit:
      1) decimal prefix symbol (e.g., M) (see table 2); and
      2) unit abbreviation (e.g., B);

   and

b) for values based on binary units of measure:
   1) numerical value (e.g., 1 024);
   2) space; and
   3) prefix symbol and unit:
      1) binary prefix symbol (e.g., Gi) (see table 2); and
      2) unit abbreviation (e.g., b).

Table 2 compares the prefix, symbols, and power of the binary and decimal units.

Table 2 — Comparison of decimal prefixes and binary prefixes

Decimal			Binary		
Prefix name	Prefix symbol	Power (base-10)	Prefix name	Prefix symbol	Power (base-2)
kilo	k	10^3	kibi	Ki	2^{10}
mega	M	10^6	mebi	Mi	2^{20}
giga	G	10^9	gibi	Gi	2^{30}
tera	T	10^{12}	tebi	Ti	2^{40}
peta	P	10^{15}	pebi	Pi	2^{50}
exa	E	10^{18}	exbi	Ei	2^{60}
zetta	Z	10^{21}	zebi	Zi	2^{70}
yotta	Y	10^{24}	yobi	Yi	2^{80}

3.5.3 Byte encoded character strings conventions

When this standard requires one or more bytes to contain specific encoded characters, the specific characters are enclosed in single quotation marks. The single quotation marks identify the start and end of the characters that are required to be encoded but are not themselves to be encoded. The characters that are to be encoded are shown in the case that is to be encoded.

An ASCII space character (i.e., 20h) may be represented in a string by the character '¬' (e.g., 'SCSI¬device').

The encoded characters and the single quotation marks that enclose them are preceded by text that specifies the character encoding methodology and the number of characters required to be encoded.

EXAMPLE - Using the notation described in this subclause, stating that eleven ASCII characters 'SCSI device' are to be encoded is the same as writing out the following sequence of byte values: 53h 43h 53h 49h 20h 64h 65h 76h 69h 63h 65h.

3.6 Bit and byte ordering

In this standard, data structures may be defined by a table. A table defines a complete ordering of elements (i.e., bits, bytes, fields, and dwords) within the structure. The ordering of elements within a table does not in itself constrain the order of storage or transmission of the data structure, but in combination with other normative text in this standard, the ordering of elements within a table may constrain the order of storage or transmission of the structure.

In a table, any element that is presented in a row above another element in a lower row is more significant than the lower element, and any element presented to the left of another element in the same row is more significant than the element to the right.

If a table shows bit numbering (see table 3), the least significant bit (LSB) is numbered 0 and each more significant bit has the next greater number than the immediately less significant bit. If a table shows numbering of bytes or characters (see table 4), the most significant byte or character is represented at the lowest number and each less significant byte or character has the next greater number than the immediately more significant byte.

In a field in a table consisting of more than one bit that contains a single value (e.g., a number), the least significant bit (LSB) is shown on the right and the most significant bit (MSB) is shown on the left (e.g., in a byte, bit 7 is the MSB and is shown on the left, bit 0 is the LSB and is shown on the right). The MSB and LSB are not labeled if the field consists of eight or fewer bits. The MSB and LSB are labeled if the field consists of more than eight bits and has no internal structure defined.

In a field in a table consisting of more than one byte that contains multiple fields each with their own values (e.g., a descriptor), there is no MSB and LSB of the field itself and thus there are no MSB and LSB labels. Each individual field has an MSB and LSB, but they are not labeled.

In a field containing a text string (e.g., ASCII or UTF-8), only the MSB of the first character and the LSB of the last character are labeled.

Multiple byte fields are represented with only two rows, with the non-sequentially increasing byte number denoting the presence of additional bytes.

A data dword consists of 32 bits. Table 3 shows a data dword containing a single value, where the MSB is on the upper left in bit 31 and the LSB is on the lower right in bit 0.

Table 3 — Example of ordering of bits and bytes within a data dword

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB) Bit 31	Bit 30	Bit 29	Bit 28	Bit 27	Bit 26	Bit 25	Bit 24
1	Bit 23	Bit 22	Bit 21	Bit 20	Bit 19	Bit 18	Bit 17	Bit 16
2	Bit 15	Bit 14	Bit 13	Bit 12	Bit 11	Bit 10	Bit 9	Bit 8
3	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0 (LSB)
Note - The Bit x labels in the individual table cells are for reference only and should not appear within tables that use this element format.								

Table 4 shows a data dword containing four one-byte elements, where byte 0 (the first byte) is on the top and byte 3 (the fourth byte) is on the bottom. Each byte has an MSB on the left and an LSB on the right.

Table 4 — Example of ordering of bits and bytes within a data dword element

Bit Byte	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
0	First byte							
	Bit 7 (MSB)	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0 (LSB)
1	Second byte							
	Bit 7 (MSB)	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0 (LSB)
2	Third byte							
	Bit 7 (MSB)	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0 (LSB)
3	Fourth byte							
	Bit 7 (MSB)	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0 (LSB)
Note - The Bit x labels in the individual table cells and the xx byte labels in the individual bytes are for reference only and should not appear within tables that use these element formats. In this example the MSB and LSB labels are for reference only. However, they may appear in multi-byte fields as described in this subclause.								

3.7 Notation for procedure calls

In this standard, the model for functional interfaces between entities is a procedure call. Such interfaces are specified using the following notation:

[Result =] Procedure Name (IN ([input-1] [,input-2] …), OUT ([output-1] [,output-2] …))

Where:

Result	A single value representing the outcome of the procedure or function.
Procedure Name	A descriptive name for the function to be performed.
IN (Input-1, Input-2, …)	A comma-separated list of names identifying caller-supplied input data objects.
OUT (Output-1, Output-2, …)	A comma-separated list of names identifying output data objects to be returned by the procedure.
[…]	Brackets enclose optional or conditional parameters and arguments.

This notation allows arguments to be specified as inputs and outputs. An interface between entities may require only inputs. If a procedure call has no output arguments, the word OUT, preceding comma, and associated pair of balanced parentheses are omitted.

The following is an example of a procedure call specification:

Found = **Search (IN (Pattern, Item List), OUT ([Item Found]))**

Input arguments:

Pattern: Argument containing the search pattern.

Item List: **Item<NN>** contains the items to be searched for a match.

Output arguments:

Item Found: Item located by the search procedure call. This argument is only returned if the search succeeds.

4 RDMA communication service model

4.1 Overview

This protocol is designed to operate using an RDMA communication service. An RDMA communication service provides communication between pairs of consumers using messages for control information and RDMA operations for data transfers. This clause describes an abstract RDMA communication service suitable for supporting this protocol. Annex B describes the mapping of these functions to those provided by the InfiniBand™ Architecture.

Figure 3 shows an example system that uses an RDMA communication service. Communication is provided by RDMA channels. An RDMA channel provides communication between two consumers. A single pair of consumers may communicate using many RDMA channels if sufficient resources are available. Some environments may use multiple special purpose RDMA channels between a single pair of consumers (e.g., a pair of consumers may use certain RDMA channels for messages and other RDMA channels for RDMA operations).

The RDMA communication service in figure 3 is comprised of adapters and other unspecified components (e.g. wires, fabric switches). The components of an RDMA communication service are vendor specific.

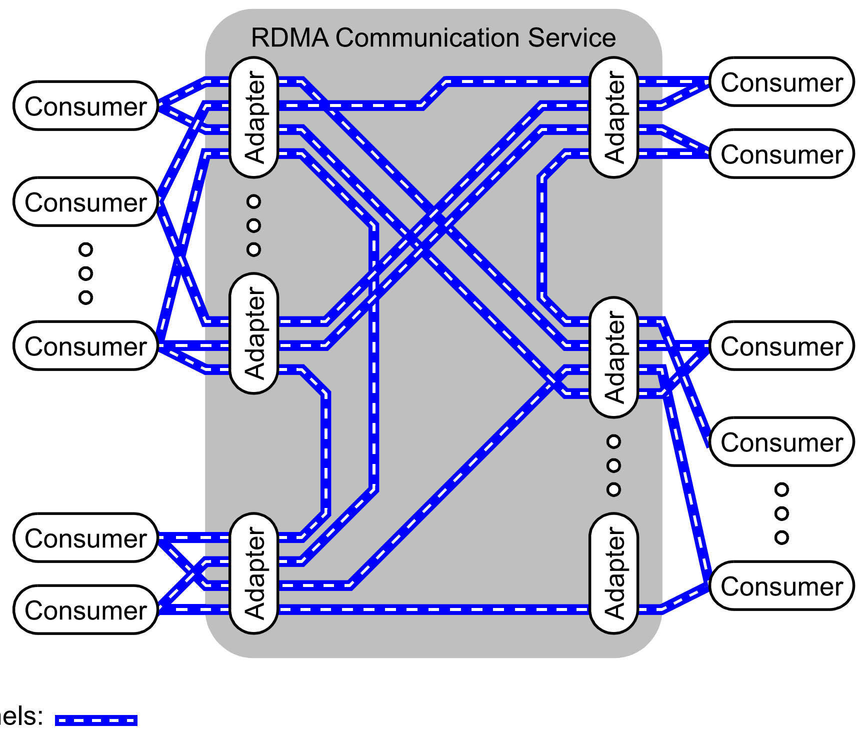


Figure 2 — RDMA communication service example

4.2 RDMA Channels

4.2.1 Introduction

An RDMA channel provides communication between a pair of consumers using messages, RDMA operations, or both. An RDMA channel is a dynamic association, established and destroyed upon request. Establishing an RDMA channel may require obtaining resources to support the RDMA channel, either within the RDMA channel's consumers or within the RDMA communication service or both. The resources associated with an RDMA channel may be released after the RDMA channel is disconnected.

4.2.2 Establishment

Figure 4 shows an example of the process by which an RDMA channel is established. A client consumer requests that the RDMA communication service establish an RDMA channel. The request is directed to a server agent and, if successful, resolved to a server consumer. The resulting RDMA channel provides communication between the client consumer and the server consumer.

A client consumer provides a server identifier to establish an RDMA channel. The format and interpretation of a server identifier are specific to the RDMA communication service. A server identifier may specify an individual server consumer or multiple server consumers (e.g., a server identifier may identify an adapter as shown in figure 2, specifying all consumers that implement a specific application protocol and are accessible through that adapter).

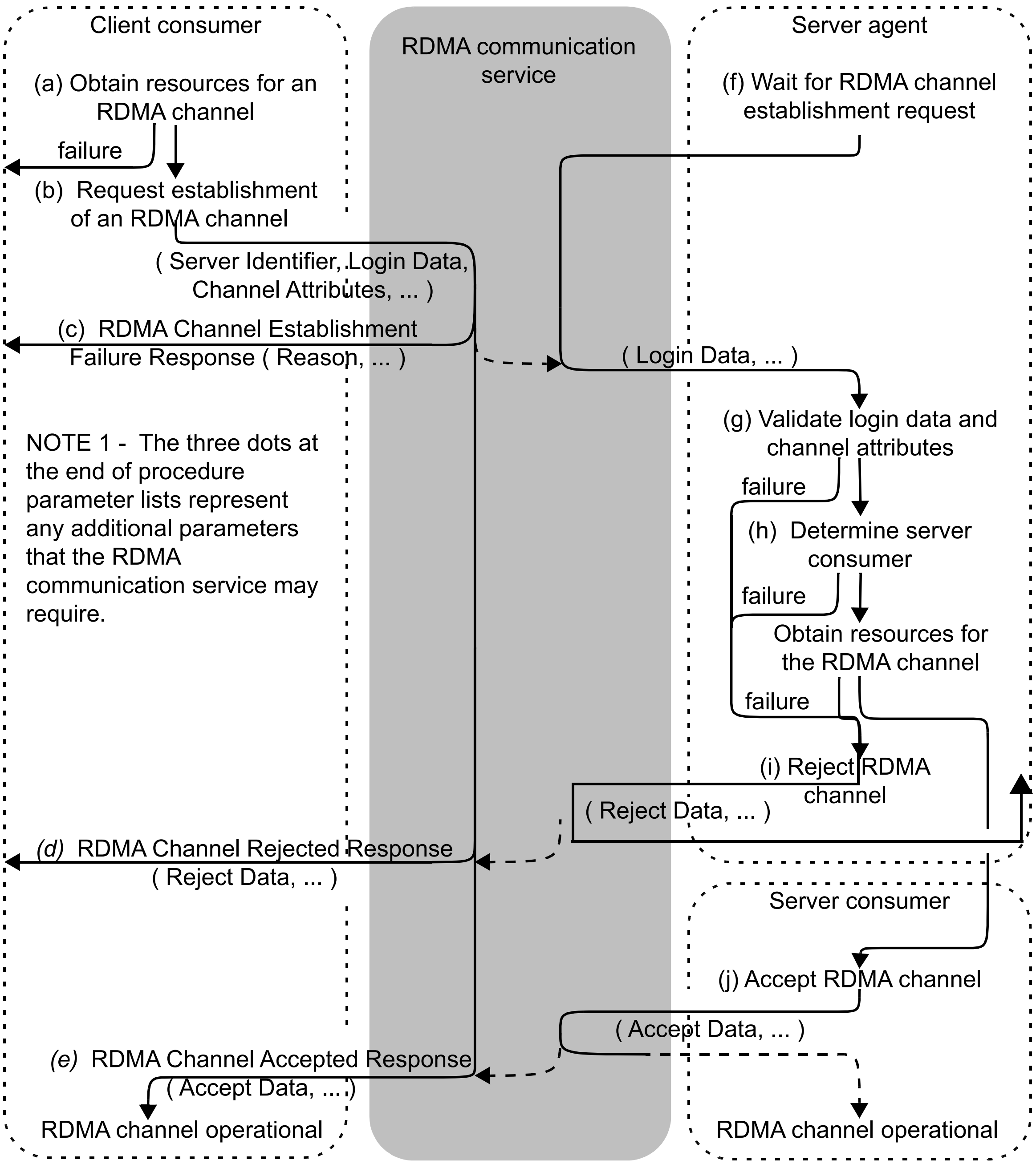


Figure 3 — Example RDMA channel establishment

In the example shown in figure 3, the recipient of an RDMA channel establishment request, identified by a server identifier, is called a server agent. The server agent determines whether an RDMA channel establishment request may be accepted and the server consumer to which it shall be assigned. A server agent may not be a distinct object. Some or all of the actions that figure 3 shows being performed by a server agent may be performed by a server consumer or by the RDMA communication service.

An RDMA communication service may require that the client consumer obtain resources (see figure 3, step a) before requesting that an RDMA channel be established. After obtaining those resources, the client consumer may request (see figure 3, step b) that the RDMA communication service establish an RDMA channel. The request includes a server identifier, login data, channel attributes, and any other parameters required by the RDMA communication service. Under this protocol, the client consumer is an SRP initiator port, the server identifier is an interconnect specific value that enables the RDMA communication service to locate one or more SRP target ports, and the login data contains an SRP_LOGIN_REQ request (see 6.2).

The RDMA communication service returns one of three responses to the client consumer for an RDMA channel establishment request:

a) an RDMA Channel Establishment failure response (see figure 3, step c);
b) an RDMA Channel Rejected response (see figure 3, step d); or
c) an RDMA Channel Accepted response (see figure 3, step e).

An RDMA Channel Establishment Failure response (see figure 3, step c) indicates that the RDMA channel was not established for some reason internal to the RDMA communication service. An RDMA Channel Establishment Failure response may return an RDMA communication service-specific reason code to identify the cause of the failure as well as other RDMA communication service-specific data.

An RDMA Channel Rejected response (see figure 3, step d) indicates that the request was rejected by the server agent or server consumer. An RDMA Channel Rejected response may return rejection data provided by the server agent or server consumer. Rejection data may include a reason for rejecting the request. In this protocol, the rejection data includes an SRP_LOGIN_REJ response (see 6.4). An RDMA Channel Rejected response may also return RDMA communication service specific data.

An RDMA Channel Accepted response (see figure 3, step e) indicates that the RDMA channel has been successfully established. The client consumer may use the RDMA channel in accordance with the application protocol. An RDMA Channel Accepted response returns acceptance data provided by the server agent or server consumer. In this protocol, the acceptance data includes an SRP_LOGIN_RSP response (see 6.3). An RDMA Channel Accepted Response may also return data specific to the RDMA communication service.

An RDMA communication service may require that a server agent register (see figure 3, step f) itself prior to receiving connection establishment requests. In see figure 4 this is shown as a registration request (e.g., subroutine call) that returns control to the server agent when an RDMA Channel Establishment request is received. The way that a server agent registers with an RDMA communication service is specific to that service or the server.

RDMA Channel Establishment requests that are acceptable to the RDMA communication service are passed (see figure 3, step g) to the server agent. The server agent receives the login data and RDMA communication service-specific data from the client consumer's request. In this protocol, the login data includes an SRP_LOGIN_REQ request (see 6.2).

The server agent determines whether the RDMA Channel Establishment request may be accepted and determines (see figure 3, step i) the server consumer to be associated with the RDMA channel. If the request is not accepted, then the server agent or server consumer instructs the RDMA communication service to reject (see figure 3, step j) the RDMA channel. The server agent or server consumer provides rejection data and any RDMA communication service-specific data that is required. In this protocol, the rejection data shall contain an SRP_LOGIN_REJ response (see 6.4).

If the RDMA Channel Establishment request is accepted, then the server agent or server consumer instructs the RDMA communication service to accept (see figure 3, step k) the RDMA channel. The server agent or server consumer provides acceptance data and any RDMA communication service-specific data that is required. In this protocol, the acceptance data shall contain an SRP_LOGIN_RSP response (see 6.3).

4.2.3 Disestablishment

An RDMA channel may be disconnected by the RDMA communication service due to an error or by request from either of the RDMA channel's consumers. The consumers may each be notified that the RDMA channel has been disconnected, allowing the consumers to recover any resources associated with the RDMA channel. The time to deliver a notification may vary depending upon the RDMA communication service, the consumer being notified, and the specific circumstances of the disconnection request.

A disconnection request or error causes an RDMA channel to become non-operational. Operations in progress on an RDMA channel at the time it becomes non-operational and operations requested subsequently may not complete, and the status of those operations may be indeterminate.

4.3 Messages

A message is sent by one consumer associated with an RDMA channel (i.e., the sending consumer) to the other consumer associated with the RDMA channel (i.e., the receiving consumer). A message contains a payload of some number of data bytes. An RDMA communication service may provide a facility known as solicited message reception notification, by which some messages may be marked by the sender to indicate to the recipient that the marked message is more urgent than a message that is not marked. How such a marked message is handled by the recipient is outside the scope of this standard.

A sending consumer requests that a message be sent by providing the following to an RDMA communication service:

a) the message's payload length;
a) the message's payload data;
a) the RDMA channel to use; and
a) whether to use normal message or solicited message reception notification.

The RDMA communication service attempts to deliver the message to the receiving consumer. If delivery succeeds, then the RDMA communication service notifies the receiving consumer that a message has been received, providing the message's length, payload, and the RDMA channel on which the message was received. The RDMA communication service may also provide an indication of whether the sending consumer specified normal message or solicited message reception notification.

Sending a message on an RDMA channel when no receive buffer has been provided or when the provided receive buffer is too small for the message, may result in behavior that is not specified by this standard.

As an example this behavior may include disconnecting the RDMA channel, discarding or truncating the message, or delaying delivery of the message until a suitable message receive buffer becomes available. The RDMA communication service may not provide an error indication.

An RDMA communication service may not provide a way for a sending consumer to determine whether a message has been delivered to the receiving consumer (e.g., an acknowledgement may indicate only that a message was received without error).

4.4 RDMA operations

4.4.1 Overview

An RDMA channel may provide RDMA Write operations, RDMA Read operations, or both between its consumers.

A consumer may allow RDMA access by registering some or all of its memory with an RDMA communication service. The RDMA communication service returns a memory handle to identify the registered memory. The consumer may specify that the memory handle is usable for memory access on only a specified RDMA channel or on a group of RDMA channels. The consumer may impose other access restrictions allowed by the RDMA communication service (e.g. read-only access).

A consumer that has registered memory and obtained a memory handle may communicate the memory handle to another consumer. This may be done using an application protocol contained in message payloads. The other consumer may then use the memory handle to request RDMA operations that access the memory registered by the first consumer.

The registered memory identified by a memory handle is represented as a memory address space. Accessible locations are identified by addresses. An RDMA communication service is not required to provide a way to determine, from a message handle, which memory locations are accessible, the number of locations that are accessible, or the type of access allowed.

4.4.2 RDMA Write

An RDMA Write operation allows a requesting consumer to store data into memory registered by another consumer. A requesting consumer provides the following to an RDMA communication service when it requests an RDMA Write operation:

a) an RDMA channel to use for the operation;
b) a memory handle that is usable for access on that RDMA channel;
c) a range of addresses within the memory address space identified by the memory handle; and
d) data to be written into the specified range of addresses.

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 (e.g., an acknowledgment may only signify error-free reception of the data). An RDMA communication service is not required to provide a way for the consumer that registered the memory to determine whether an RDMA Write operation is in progress or has completed.

4.4.3 RDMA Read

An RDMA Read operation allows a requesting consumer to fetch data from memory registered by another consumer. A requesting consumer provides the following to an RDMA communication service when it requests an RDMA Read operation:

a) an RDMA channel to use for the operation;
b) a memory handle that is usable for access on that RDMA channel;
c) a range of addresses within the memory address space identified by the memory handle; and
d) a buffer into which to place the data read from the specified range of addresses.

The RDMA communication service notifies the requesting consumer after data has been successfully obtained from the specified range of addresses and placed in the requestor’s buffer. An RDMA communication service is not required to provide a way for the consumer that registered the memory to determine whether an RDMA Read operation is in progress or has completed.

4.5 Ordering and Reliability

4.5.1 Ordering and reliability overview

This protocol operates using an RDMA communication service having the characteristics described in 4.5.2 and 4.5.3. Use of this protocol with an RDMA communication service having different characteristics is outside the scope of this standard.

4.5.2 Reliability

An RDMA communication service shall deliver each message sent on an RDMA channel to the receiving consumer or destroy the RDMA channel. Each delivered message shall be delivered to the receiving consumer once, without duplication; the RDMA communication service shall discard any duplicates that may result from retransmission or other mechanisms. Each delivered message shall be delivered to the receiving consumer complete and error free.

The RDMA communication service shall provide to the sending consumer an indication of the completion status of each RDMA communication service request. This status shall be one of the following:

a) **successful**: The request completed without error;
b) **error**: The request was not completed due to an error. The RDMA communication service may provide additional information about the error. This status should be returned immediately when the RDMA channel does not exist or has experienced an error; or
c) **timeout**: No indication was received, completion status of request is unknown, RDMA communication service has experienced an error. The length of time after which a timeout indication is returned is specific to the RDMA communication service.

4.5.3 Ordering

Messages sent on an RDMA channel shall be delivered to the receiving consumer in the order they were sent. The data for all RDMA Write operations requested on an RDMA channel by a consumer prior to that same consumer sending a message on the same RDMA channel shall be available to the receiving consumer (e.g. stored into registered memory) before the message is delivered to the receiving consumer. The order in which multiple RDMA Writes, without an intervening message, are applied to a particular memory location is specific to the RDMA communication service.

Messages sent on different RDMA channels may be delivered in any order. The data for RDMA Write operations may be stored into registered memory in any order relative to the delivery of messages sent on other RDMA channels. RDMA Write operations requested on different RDMA channels may store data into the same registered memory location in any order.

RDMA Read operations may be processed in any order.

If an RDMA communication service fails to meet the ordering requirements of this subclause on an RDMA channel, then it shall destroy the RDMA channel.

5 Structure and concepts

5.1 Overview of SRP operation

5.1.1 RDMA channel establishment and login

SRP initiator ports login with SRP target ports when a new RDMA channel is established for use with this protocol. The login process associates an RDMA channel with a specific SRP initiator port and SRP target port (i.e., an I_T nexus (see SAM-5)) and negotiates parameters that govern the use of that RDMA channel.

SRP initiator ports and SRP target ports shall be determined by their role during RDMA channel establishment. An object that requests RDMA channel establishment as a client consumer (see 4.2) shall be an SRP initiator port. An object that accepts RDMA channel establishment as a server consumer (see 4.2) shall be an SRP target port.

Login occurs during RDMA channel establishment. An SRP initiator port shall provide an SRP_LOGIN_REQ request (see 6.2) as the login data when establishing a new RDMA channel. If an SRP target port accepts a new RDMA channel, then it shall provide an SRP_LOGIN_RSP response (see 6.3) as the acceptance data. If an SRP target port does not accept a new RDMA channel, then it shall provide an SRP_LOGIN_REJ response (see 6.4) as the rejection data when rejecting the new RDMA channel.

The SRP_LOGIN_REQ request (see 6.2) contains an SRP initiator port identifier and an SRP target port identifier. An SRP target port shall not accept a new RDMA channel unless its SRP target port identifier is identical to the value in the SRP_LOGIN_REQ request. If an SRP target port accepts a new RDMA channel, then it shall treat all communication on that RDMA channel as being with the SRP initiator port identified by the SRP initiator port identifier specified in the SRP_LOGIN_REQ request.

5.1.2 Single RDMA channel operation

An SRP initiator port may specify single RDMA channel operation during login. If an SRP target port accepts such a login, then it shall:

a) attempt to send an SRP_T_LOGOUT request (see 6.6) on any established RDMA channel that specified the same SRP initiator port identifier. The reason code shall indicate that the RDMA channel was disconnected due to a MULTI-CHANNEL ACTION code in a new SRP_LOGIN_REQ request (see 6.2);
b) request disconnection of any established RDMA channel (see 5.1.4) that specified the same SRP initiator port identifier; and
c) reject any other RDMA channel establishment requests it has received that specified the same SRP initiator port identifier and that the SRP target port has not yet accepted.

Following acceptance of a login specifying single RDMA channel operation, that single RDMA channel shall be used for all communication between the specified SRP initiator port and SRP target port. Subsequent logins specifying other modes of operation may allow communication using multiple RDMA channels.

When an I_T nexus is using this protocol in the single RDMA channel mode, the following events are I_T nexus loss notification events (see SAM-5 and SPC-4):

a) sending or receiving an SRP_I_LOGOUT request or an SRP_T_LOGOUT request;
b) requesting that an RDMA channel be disconnected; and
c) receiving notification that an RDMA channel has been disconnected.

5.1.3 Multiple independent RDMA channel operation

An SRP initiator port may specify multiple independent RDMA channel operation during login. An SRP target port shall not accept such a login if doing so requires disconnecting an established RDMA channel with the same SRP initiator port and shall return the SRP_T_LOGOUT request reason code RDMA CHANNEL LIMIT REACHED FOR THIS INITIATOR.

Following acceptance of a login specifying multiple independent RDMA channel operation, one or more RDMA channels may be used for communication between the SRP initiator port and the SRP target port. All such RDMA channels are associated with the single I_T nexus defined by the SRP initiator port identifier and the SRP target port identifier.

If multiple independent RDMA channels are used, then operation of each SRP request is confined to a single RDMA channel. The sender of an SRP request chooses an RDMA channel to use for sending the SRP request. The sender of an SRP response shall use the same RDMA channel as the SRP request for sending the SRP response. All RDMA operations associated with the SRP request shall also use the same RDMA channel as the SRP request.

While each SRP request is confined to a single RDMA channel, SCSI tasks and task management functions may be conveyed on independent RDMA channels associated with the same I_T nexus. SCSI tasks and task management functions interact as specified by SAM-5, SPC-4, and other SCSI command standards (e.g., within an I_T nexus, a SCSI task sent on one RDMA channel may be aborted by an ABORT TASK sent on a different RDMA channel).

An RDMA communication service may not provide any ordering relationship between SRP requests, SRP responses, and RDMA operations that use different RDMA channels. If ordering is important for a sequence of SRP requests, then they should be sent using the same RDMA channel.

If an I_T nexus is using this protocol in the multiple independent RDMA channel mode, then the following events are I_T nexus loss notification events (see SAM-5 and SPC-4) when they occur with respect to the last, or only, channel associated with the I_T nexus:

a) sending or receiving an SRP_I_LOGOUT request or an SRP_T_LOGOUT request;
b) requesting that an RDMA channel be disconnected; and
c) receiving notification that an RDMA channel has been disconnected.

5.1.4 RDMA channel disconnection

RDMA channel disconnection may cause (see 5.1.2 and 5.1.3) an I_T nexus loss notification event as described in SAM-5 and SPC-4.

An SRP initiator port should send an SRP_I_LOGOUT request (see 6.5) and wait for the RDMA communication service status indication (see 4.5.2) before requesting that an RDMA channel be disconnected.

After requesting that an RDMA channel be disconnected, after being notified that an RDMA channel has been disconnected, or upon receiving an SRP_T_LOGOUT request (see 6.6), an SRP initiator port shall:

a) discard any outstanding request received from an SRP target port on that RDMA channel, without returning a response;
b) not send any further messages on that RDMA channel;
c) discard any subsequent messages received on that RDMA channel; and
d) for any outstanding SCSI tasks sent on that RDMA channel, indicate to the application client that the task has terminated with a service delivery system failure.

An SRP target port should send an SRP_T_LOGOUT request (see 6.6) and wait for the RDMA communication service status indication (see 4.5.2) before requesting that an RDMA channel be disconnected.

After requesting that an RDMA channel be disconnected, after being notified that an RDMA channel has been disconnected, or upon receiving an SRP_I_LOGOUT request (see 6.5), an SRP target port shall:

a) abort all outstanding SCSI tasks that were contained in SRP_CMD requests (see 6.8) received on that RDMA channel, without returning a response;
b) discard any other outstanding requests received from an SRP initiator port on that RDMA channel, without returning a response;
c) not send any further messages on that RDMA channel;
d) discard any subsequent messages received on that RDMA channel; and

e) not alter previously established conditions (e.g., MODE SELECT parameters, reservations, ACA) as a result of the disconnection.

5.2 Identifiers

Initiator port identifiers and target port identifiers (see SAM-5) for this protocol are 16 bytes in length.

5.3 Alias associations

There are no events specific to this protocol that clear alias associations (see SPC-4).

5.4 Information unit classes

Each SRP information unit is classified as an SRP request (see table 9 and table 11) or an SRP response (see table 10 and table 12). SRP requests convey SCSI commands, task management requests and RDMA channel management requests. SRP responses convey SCSI command and task management service responses and RDMA channel management responses. RDMA channel management requests may be issued by SRP target ports or SRP initiator ports.

In normal operation, SRP requests and SRP responses occur in pairs. Each SRP request elicits a single corresponding SRP response from the SRP device receiving the SRP request. An SRP request communicates the start of a remote procedure call; the corresponding SRP response communicates the remote procedure call's completion.

An SRP response shall not be returned for:

a) an SRP_CMD request if the associated task is aborted;
b) an SRP_T_LOGOUT request (see 6.6);
c) an SRP_I_LOGOUT request (see 6.5); and
d) outstanding SRP requests received on an RDMA channel when an SRP device becomes aware of a failure preventing further communication on that RDMA channel. In this case, the device shall abort all outstanding SRP requests received on that RDMA channel.

In all other cases an SRP device shall return a single SRP response for each SRP request it receives.

SRP responses shall be sent on the RDMA channel on which the corresponding SRP request was received.

5.5 SRP target port buffer management

5.5.1 Buffer management overview

SRP target port buffer management allows an SRP target device to limit the number of SRP requests that may be sent on an RDMA channel. SRP devices may use SRP target port buffer management to manage internal and RDMA channel-related resources.

SRP responses are not subject to buffer management, they may be sent at any time. An SRP device may limit the number of SRP responses it may receive by limiting the number of SRP requests it has outstanding.

5.5.2 SRP requests issued by target port

SRP target ports shall limit themselves to one outstanding SRP request (see table 11) per RDMA channel. Upon sending an SRP request, an SRP target port shall not send another SRP request on the same RDMA channel until after it receives the SRP response (see table 12) for the previous SRP request.

5.5.3 Requests issued by initiator port

This protocol uses a credit-based buffer management algorithm to limit the number of SRP requests (see table 5) that an SRP initiator port may send to an SRP target port. The algorithm uses:

a) the REQUEST LIMIT DELTA field present in information units sent by an SRP target port to an SRP initiator port except in the SRP_LOGIN_REJ or SRP_T_LOGOUT; and
b) the REQUEST LIMIT state variable.

Most information units containing a REQUEST LIMIT DELTA field do not generate a confirmation that the SRP initiator port has received the information unit and processed the contents of the REQUEST LIMIT DELTA field. The SRP_CRED_REQ (see 6.10) request ~~and SRP_AER_REQ (see 6.12) request~~ generates a confirmation through the SRP_CRED_RSP (see 6.11) response ~~and SRP_AER_RSP (see 6.13) response~~. An SRP initiator port shall process the REQUEST LIMIT DELTA fields of information units received on an RDMA channel in the order that they are received. An SRP initiator port shall process the REQUEST LIMIT DELTA field of a request before sending that request's response (e.g., an SRP initiator port shall process the REQUEST LIMIT DELTA field of an SRP_CRED_REQ request before sending the SRP_CRED_RSP response). The following rules specify the buffer management algorithm for SRP requests sent by SRP initiator ports:

a) the Request Limit variable and Request Limit Delta variable are both signed, two's complement 32-bit integers. SRP initiator ports shall implement a separate copy of the request limit variable for each RDMA channel;
b) upon successful completion of RDMA channel establishment an SRP initiator port shall initialize the RDMA channel's Request Limit variable to the value of the REQUEST LIMIT DELTA field received in the SRP_LOGIN_RSP response (see 6.3). Except for providing an SRP_LOGIN_REQ request (see 6.2) when requesting RDMA channel establishment, the SRP initiator port shall not send any SRP information units on the RDMA channel prior to initializing the Request Limit variable;
c) an SRP initiator port may send an SRP request on an RDMA channel when the value of the RDMA channel's Request Limit variable is greater than zero. An SRP initiator port shall not send an SRP request on any RDMA channel whose Request Limit variable has a value less than or equal to zero. The results of doing so are vendor specific. To ensure that task management requests may be sent, an SRP initiator port may choose to send commands only when the value of the Request Limit variable is greater than one;
d) an SRP initiator port shall decrement an RDMA channel's Request Limit variable by one whenever it sends an SRP request on that RDMA channel;
e) an SRP initiator port shall add using two's complement addition the value of the REQUEST LIMIT DELTA field to an RDMA channel's Request Limit variable whenever the SRP initiator port receives an information unit on that RDMA channel;
f) an SRP target port shall not specify a positive value of the REQUEST LIMIT DELTA field that may cause the SRP initiator port's Request Limit variable to exceed 2^{30}; and
g) an SRP target port shall not specify a negative value for the REQUEST LIMIT DELTA field in an information unit that may cause the SRP initiator port's Request Limit variable to drop below -2^{31}.

5.6 Data buffers

5.6.1 Memory descriptors

A memory descriptor is a 16-byte structure that identifies a memory segment (see table 5). Figure 4 shows the mapping of a memory descriptor to a memory segment.

Table 5 — Memory descriptor

Byte \ Bit	7	6	5	4	3	2	1	0
0	(MSB)							
•••	VIRTUAL ADDRESS							
7								(LSB)
8	(MSB)							
•••	MEMORY HANDLE							
11								(LSB)
12	(MSB)							
•••	DATA LENGTH							
18								(LSB)

The VIRTUAL ADDRESS field contains an unsigned integer value that identifies the byte address within the memory region of the first byte of the memory segment.

The MEMORY HANDLE field contains an SRP initiator port-specific value that identifies the region that contains the memory segment. The SRP target port shall supply this value with any RDMA operation that accesses the memory segment.

The DATA LENGTH field contains an unsigned integer value that identifies the length of the memory segment in bytes. The interpretation of a memory descriptor where the sum of the VIRTUAL ADDRESS field and DATA LENGTH field exceeds 2^{64} is vendor specific.

An SRP target port may use a memory descriptor for either RDMA Read operations or RDMA Write operations but not both. SRP target ports shall issue only the appropriate type of RDMA operation for a memory descriptor, depending on whether the descriptor was a data-in or data-out descriptor, and shall ensure that each RDMA operation is wholly contained within the memory segment by using the following rules:

a) the RDMA operation’s starting virtual address shall be greater than or equal to the memory descriptor’s virtual address and less than the sum of the memory descriptor’s virtual address and data length; and

b) the sum of the RDMA operation's virtual address and data length shall be greater than the memory descriptor's virtual address and less than or equal to the sum of the memory descriptor's virtual address and data length.

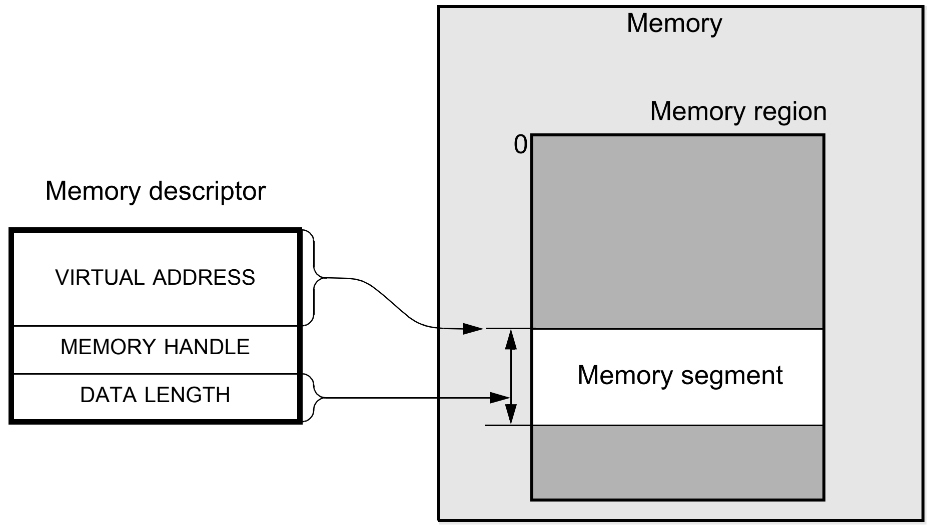


Figure 4 — Memory descriptor mapping

5.6.2 Data buffer descriptors

5.6.2.1 Overview

An SRP_CMD request (see 6.8) may contain a data-out buffer descriptor, a data-in buffer descriptor, both, or neither, depending upon the data transfers requested by the SCSI command. The format of each data buffer descriptor is specified by a format code value. In an SRP_CMD request with both data-in and data-out buffer descriptors, there is no requirement that both buffer descriptors be of the same format. Some data buffer descriptor formats use the contents of a count field (i.e., SRP_CMD request DATA-OUT BUFFER DESCRIPTOR COUNT field or DATA-IN BUFFER DESCRIPTOR COUNT field) to further describe the data buffer descriptor format. Table 2 defines data buffer descriptor format code values.

Table 6 — Status codes

Data buffer descriptor format code	Reference	format code value [a]	buffer descriptor length (bytes) [c]
NO DATA BUFFER DESCRIPTOR PRESENT	5.6.2.3	0h	0
DIRECT DATA BUFFER DESCRIPTOR	5.6.2.4	1h	16
INDIRECT DATA BUFFER DESCRIPTOR	5.6.2.5	2h	20 + 16 × count [b]
IMMEDIATE DATA BUFFER		3h	4

[a] The format code value for a data-out buffer descriptor is specified by the DATA-OUT BUFFER DESCRIPTOR FORMAT field of an SRP_CMD request (see 6.8). The format code value for a data-in buffer descriptor is specified by the DATA-IN BUFFER DESCRIPTOR FORMAT field of an SRP_CMD request (see 6.8).

[b] The count field for a data-out buffer descriptor is the DATA-OUT BUFFER DESCRIPTOR COUNT field of an SRP_CMD request (see 6.8). The count field for a data-in buffer descriptor is the DATA-IN BUFFER DESCRIPTOR COUNT field of an SRP_CMD request (see 6.8).

[c] The length of a data buffer descriptor is determined from its format code value and the contents of its count field.

5.6.2.2 Supported data buffer descriptor formats

The REQUIRED BUFFER FORMATS field (see table 7) of the SRP_LOGIN_REQ request (see 6.2) indicates the data buffer descriptor formats (see table 6) that an SRP initiator port may specify in requests sent on an RDMA channel. An SRP initiator port shall set the REQUIRED BUFFER FORMATS field to indicate all data buffer descriptor formats that the SRP initiator port may specify in SRP_CMD requests (see 6.8) sent on that RDMA channel. An SRP initiator port shall not issue an SRP_CMD request indicating a data buffer descriptor format that was not indicated in the REQUIRED BUFFER FORMATS field value for that RDMA channel. 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. If a target port does detect that an initiator has specified a descriptor format not indicated in the REQUIRED BUFFER FORMATS field, then the target port shall send an SRP_T_LOGOUT request (see 6.6) with the reason code UNSUPPORTED FORMAT CODE VALUE SPECIFIED IN DATA-OUT BUFFER DESCRIPTOR FORMAT FIELD or the reason code UNSUPPORTED FORMAT CODE VALUE SPECIFIED IN DATA-IN BUFFER DESCRIPTOR FORMAT FIELD, as appropriate.

An SRP target port may accept an RDMA channel establishment request and return an SRP_LOGIN_RSP response (see 6.3) if the SRP target port is able to support all of the data buffer descriptor formats indicated in the REQUIRED BUFFER FORMATS field on that RDMA channel. An SRP target port shall reject the RDMA channel establishment request and return an SRP_LOGIN_REJ response (see 6.4) with reason ONE OR MORE REQUESTED DATA BUFFER DESCRIPTOR FORMATS ARE NOT SUPPORTED if the SRP target port is unable to support one or more of the data buffer descriptor formats indicated in the REQUIRED BUFFER FORMATS field on that RDMA channel.

An SRP target port shall indicate the data buffer descriptor formats that it supports in the SUPPORTED BUFFER FORMATS field (see table 7) of the SRP_LOGIN_RSP response (see 6.3) or the SRP_LOGIN_REJ response (see 6.4). All SRP target ports shall support the DIRECT DATA BUFFER DESCRIPTOR format and may support other data buffer descriptor formats.

Table 7 defines the contents of both the REQUIRED BUFFER FORMATS field and the SUPPORTED BUFFER FORMATS field.

Table 7 — Supported data buffer descriptor formats

Bit / Byte	7	6	5	4	3	2	1	0
0	Reserved							
1	Reserved				IMM	IDBD	DDBD	Reserved

An SRP initiator port sets the indirect data buffer descriptor (IDBD) bit to one in a SRP_LOGIN_REQ request (see 6.2) if the SRP initiator port requires that the target port support the INDIRECT DATA BUFFER DESCRIPTOR format.

The target port shall set the IDBD bit to one in an SRP_LOGIN_RSP response (see 6.3) or in an SRP_LOGIN_REJ response (see 6.4) if the SRP target port supports the INDIRECT DATA BUFFER DESCRIPTOR format. The IDBD bit shall be set to zero in an SRP_LOGIN_RSP response or in an SRP_LOGIN_REJ response if the SRP target port does not support the INDIRECT DATA BUFFER DESCRIPTOR format.

An SRP initiator port sets the direct data buffer descriptor (DDBD) bit to one in a SRP_LOGIN_REQ request (see 6.2) if it requires that the target port support the DIRECT DATA BUFFER DESCRIPTOR format.

The target port shall set the DDBD bit to one in an SRP_LOGIN_RSP response (see 6.3) or in an SRP_LOGIN_REJ response (see 6.4).

An SRP initiator port sets the IMM (immediate data) bit to one in an SRP_LOGIN_REQ request (see 6.2) if it requests that the target port supports the IMMEDIATE DATA BUFFER DESCRIPTOR format.

An SRP target port shall set the IMM bit to one in an SRP_LOGIN_RSP response (see 6.3) or in an SRP_LOGIN_REJ response (see 6.4) if the SRP target port supports the IMMEDIATE DATA BUFFER DESCRIPTOR format.

The IMM bit shall be set to zero in an SRP_LOGIN_RSP response or in an SRP_LOGIN_REJ response if the SRP target port does not support the IMMEDIATE DATA BUFFER DESCRIPTOR format.

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 returned in the SRP_LOGIN_RSP response (see 6.3).

5.6.2.3 No data buffer descriptor present

The NO DATA BUFFER DESCRIPTOR PRESENT format code value specifies that the corresponding data buffer descriptor field is not present. The contents of the count field in the SRP_CMD request (i.e., DATA-OUT BUFFER DESCRIPTOR COUNT field or DATA-IN BUFFER DESCRIPTOR COUNT field) are reserved. SRP target ports shall ignore the contents of the count field.

5.6.2.4 Direct data buffer descriptor format

The DIRECT DATA BUFFER DESCRIPTOR format code value specifies that the corresponding data buffer descriptor field is as defined in table 6 and contains a direct data buffer descriptor. The contents of the count field in the SRP_CMD request are reserved. SRP target ports shall ignore the contents of the count field.

A direct data buffer descriptor contains a single memory descriptor (see table 5). The memory descriptor identifies the data buffer, which is a single memory segment within a memory region's virtual address space. If a direct data buffer descriptor defines a data-out buffer, then the SRP target port shall issue only RDMA Read operations using the memory descriptor contained in the direct data buffer descriptor. If a direct data buffer descriptor defines a data-in buffer, then the SRP target port shall issue only RDMA Write operations using the memory descriptor contained in the direct data buffer descriptor. The SRP target port shall use the contents of the DATA LENGTH field of the memory descriptor as the length of the data-out buffer or data-in buffer.

5.6.2.5 Indirect data buffer descriptor format

The INDIRECT DATA BUFFER DESCRIPTOR format code value specifies that the corresponding data buffer descriptor field contains an indirect data buffer descriptor (see table 6).

An indirect data buffer is comprised of one or more memory segments. The memory segments may be discontiguous. The memory segments may be spread among several memory regions. The indirect data buffer is the concatenation of the memory segments listed in the indirect data buffer descriptor. Each memory segment may have any length supported by the RDMA communication service, including a length of zero bytes (see figure 5).

Table 8 shows the format of an indirect data buffer descriptor.

Table 8 — Indirect data buffer descriptor

Bit / Byte	7	6	5	4	3	2	1	0
0	Indirect table memory descriptor							
•••								
15								
16	(MSB)	TOTAL LENGTH						
•••								
19								(LSB)
20	PARTIAL MEMORY DESCRIPTOR LIST							
•••								
19+16×n								

[a] The value n is the value contained in the data buffer descriptor's count field. The COUNT field for a data-out buffer descriptor is the DATA-OUT BUFFER DESCRIPTOR COUNT field of an SRP_CMD request (see 6.8). The COUNT field for a data-in buffer descriptor is the DATA-IN BUFFER DESCRIPTOR COUNT field of an SRP_CMD request.

The indirect table memory descriptor is a memory descriptor (see table 5) that specifies a memory segment containing an indirect table. An indirect table is a list of one or more memory descriptors. The memory segments specified by the memory descriptors in the indirect table form the indirect data buffer. The value of the DATA LENGTH field of the indirect table memory descriptor represents the length, in bytes, of the indirect table, and is the number of memory descriptors in the indirect table multiplied by sixteen (i.e., the length, in bytes, of a memory descriptor). SRP target port behavior when the DATA LENGTH field of the indirect table memory descriptor contains any other value is vendor specific.

The TOTAL LENGTH field value is the sum of the DATA LENGTH field values of the memory descriptors in the indirect table. The SRP target port shall use either the TOTAL LENGTH field value or the sum of the DATA LENGTH field values as the length of the data-out buffer or data-in buffer. If the value of the TOTAL LENGTH field is not equal to the sum of the values of the DATA LENGTH fields, then the SRP target port behavior is vendor specific.

The PARTIAL MEMORY DESCRIPTOR LIST field is only present when the SRP_CMD information unit's data buffer descriptor's count field (i.e., DATA-OUT BUFFER DESCRIPTOR COUNT field or DATA-IN BUFFER DESCRIPTOR COUNT field) contains a non-zero value. The PARTIAL MEMORY DESCRIPTOR LIST field contains a list of n memory descriptors that are copies of the first n memory descriptors in the indirect table. The value n is the value contained in the associated count field. The SRP target port behavior when the PARTIAL MEMORY DESCRIPTOR LIST field contains any other value is vendor specific.

5.6.2.5.1 SRP target port indirect data restrictions

An SRP target port shall issue only RDMA Read operations to the indirect table.

If an indirect data buffer descriptor specifies a data-out buffer, then the SRP target port shall issue only RDMA Read operations using the memory descriptors contained in the indirect table or the PARTIAL MEMORY DESCRIPTOR LIST field value.

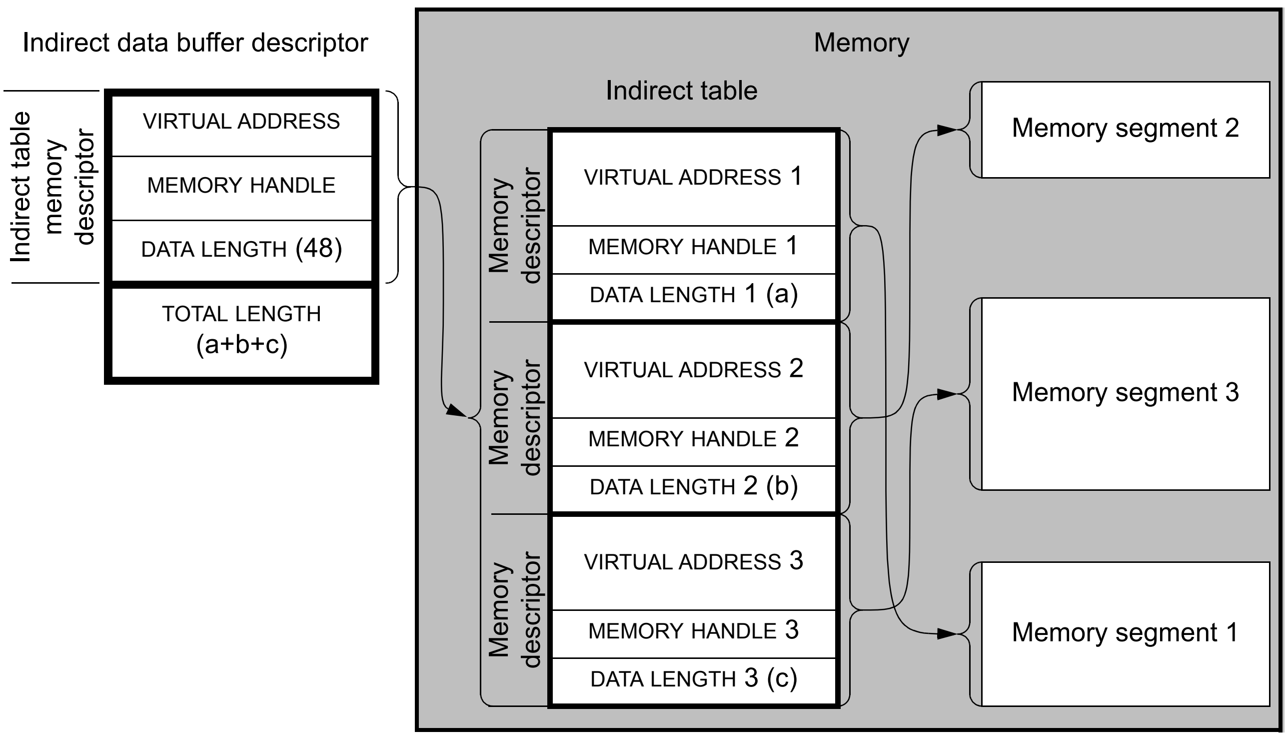
If an indirect data buffer descriptor specifies a data-in buffer, then the SRP target port shall only issue RDMA Write operations using the memory descriptors contained in the indirect table or the PARTIAL MEMORY DESCRIPTOR LIST field value.

5.6.2.5.2 Examples of indirect data buffers

Figure 5 shows an indirect data buffer descriptor that does not contain a PARTIAL MEMORY DESCRIPTOR LIST field. Memory is shown containing four memory segments:

a) the indirect table;
b) memory segment 1;
c) memory segment 2; and
d) memory segment 3.

The mapping of each memory descriptor to its memory segment has been shown as a single arrow. For details of this mapping see 5.6.1 and figure 4. Figure 5 does not show the memory regions in which the memory segments reside. All four segments may be in a single memory region or may be in different memory regions.



NOTE 1 - Format code is 2h.
NOTE 2 - Count field contains 0.
NOTE 3 - Data buffer descriptor length is 20 bytes.

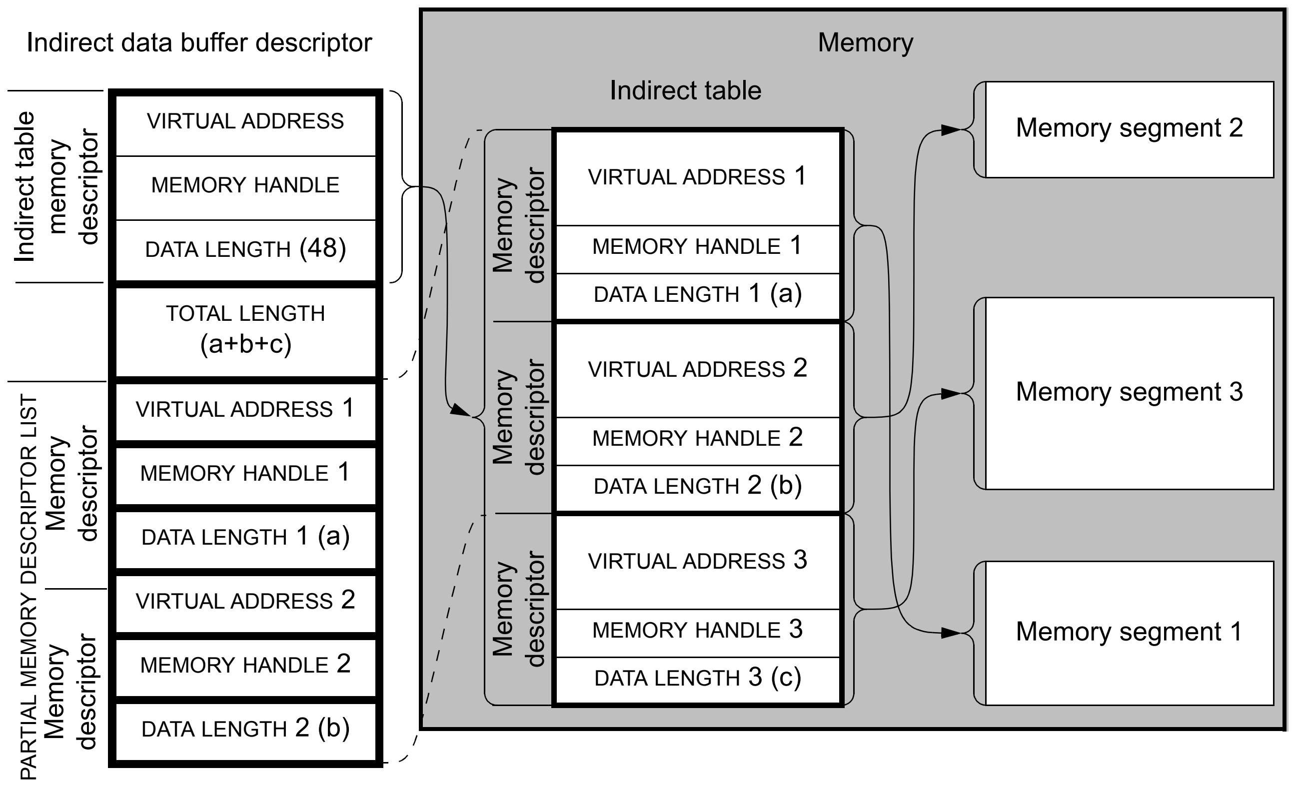
Figure 5 — Example indirect data buffer descriptor with no PARTIAL MEMORY DESCRIPTOR LIST field

In the example shown in figure 5 the data buffer descriptor format code value is 2h and the count field contains zero. The indirect data buffer descriptor is 20 bytes long. The data buffer is comprised of three memory segments:

a) memory segment 1;
b) memory segment 2; and
c) memory segment 3.

A separate memory segment contains the indirect table, a list of three memory descriptors specifying memory segments 1 to 3. The INDIRECT TABLE MEMORY DESCRIPTOR field value of the indirect data buffer descriptor specifies the memory segment containing the indirect table. The DATA LENGTH field of the INDIRECT TABLE MEMORY DESCRIPTOR field value contains 48 (i.e. the length of the indirect table). The TOTAL LENGTH field of the data buffer descriptor contains the sum of the DATA LENGTH field values of the memory descriptors in the indirect table (i.e. the sum of the contents of data length 1 + data length 2 + data length 3). This sum is the total length of the data buffer.

Figure 6 shows the same example as in figure 5 except with a PARTIAL MEMORY DESCRIPTOR LIST field. The data buffer, indirect table, INDIRECT TABLE MEMORY DESCRIPTOR field value, and TOTAL LENGTH field value are all identical to the example in figure 5. The data buffer descriptor format code is 2h, the same as in figure 5. However the COUNT field contains the value two, indicating that the PARTIAL MEMORY DESCRIPTOR LIST field is present and contains two memory descriptors. Those two memory descriptors are copies of the first two memory descriptors in the indirect table. The third memory descriptor is only present in the indirect table. The indirect data buffer descriptor is 52 bytes long.



NOTE 1 - Format code is 2h.
NOTE 2 - Count field contains 2.
NOTE 3 - Data buffer descriptor length is 52 bytes.

Figure 6 — Example indirect data buffer descriptor with a PARTIAL MEMORY DESCRIPTOR LIST field

5.6.2.6 Immediate data buffer descriptor format

The IMMEDIATE DATA BUFFER DESCRIPTOR format code value specifies that the corresponding data buffer descriptor field is as defined in table 2 and contains an immediate data buffer descriptor. An immediate data buffer descriptor contains a single 32-bit value, representing the length of the immediate data in bytes. The data-out buffer may be an immediate data buffer. The data-in buffer shall not be an immediate data buffer.

The contents of the count field in the SRP_CMD request are reserved. SRP target ports shall ignore the contents of the count field.

6 SRP Information Units

6.1 Summary

The information units used by SRP and their characteristics are shown in table 9, table 10, table 11, and table 12.

All SRP initiator ports shall support sending the information units listed in table 9 and table 12, and shall support receiving the information units listed in table 10 and table 11.

All SRP target ports shall support sending the information units listed in table 10 and table 11, and shall support receiving the information units listed in table 9 and table 12.

Table 9 — SRP requests sent from SRP initiator ports to SRP target ports

Information Unit	Reference	TYPE field (code)	Length (bytes)	Description
SRP_LOGIN_REQ	6.2	00h	64	Login request
SRP_TSK_MGMT	6.7	01h	64	SCSI task management function
SRP_CMD	6.8	02h	48 minimum	SCSI command
SRP_I_LOGOUT	6.5	03h	16	SRP initiator port logout notification

Table 10 — SRP responses sent from SRP target ports to SRP initiator ports

Information Unit	Reference	TYPE field (code)	Length (bytes)	Description
SRP_LOGIN_RSP	6.3	C0h	52	Login successful response
SRP_RSP	6.9	C1h	36 [a] minimum	SCSI status or service response
SRP_LOGIN_REJ	6.4	C2h	32	Login failure response
[a] 36 bytes is not sufficient to return any sense data.				

Table 11 — SRP requests sent from SRP target ports to SRP initiator ports

Information Unit	Reference	TYPE field (code)	Length (bytes)	Description
SRP_T_LOGOUT	6.6	80h	16	SRP target port logout
SRP_CRED_REQ	6.10	81h	52	SRP target port credit adjustment request
~~SRP_AER_REQ~~Obsolete	~~6.12~~	82h	~~36 [a] minimum~~	~~Asynchronous event report request~~

[a] 36 bytes is not sufficient to return any sense data.

Table 12 — SRP responses sent from SRP initiator ports to SRP target ports

Information Unit	Reference	TYPE field (code)	Length (bytes)	Description
SRP_CRED_RSP	6.11	41h	64	Response to SRP target port credit adjustment request
~~SRP_AER_RSP~~Obsolete	~~6.13~~	42h	~~16~~	~~Asynchronous event report response~~

Byte 0 of each SRP information unit contains a type code. The type code value uniquely identifies the information unit and its format. The length of an information unit is indicated by its type code and selected fields within the information unit. If an SRP target port receives an SRP information unit with an invalid type code or whose length is incorrect for the information unit's type code, then the SRP target port shall send an SRP_T_LOGOUT request (see 6.6) and disconnect the RDMA channel.

Bytes 8 to 15 of each information unit contain a tag value, which provides a mechanism for matching SRP requests with their corresponding SRP responses. Each SRP request information unit (see table 9 and table 11) shall contain a tag value that is unique among all the outstanding SRP requests from a particular initiator port or target 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 of outstanding SRP requests are unique. If tag values are not unique, then responder behavior is unpredictable.

6.2 SRP_LOGIN_REQ request

An SRP_LOGIN_REQ request (see table 13) conveys SRP login parameters from an SRP initiator port to an SRP target port. The SRP_LOGIN_REQ request shall be sent as login data during RDMA channel establishment.

Table 13 — SRP_LOGIN_REQ request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (00h)							
1 ••• 7	Reserved							
8 ••• 15	(MSB) TAG (LSB)							
16 ••• 23	Reserved							
24 ••• 25	REQUIRED BUFFER FORMATS							
26	FIRSTBURST	AESOLNTObsolete	CRSOLNT	LOSOLNT	Reserved		MULTICHANNEL ACTION	
027	FIRST BURST DATA OFFSET IMMEDIATE DATA OFFSET							
278	IMMEDIATE DATA OFFSETReserved							
••• 31								
32 ••• 47	INITIATOR PORT IDENTIFIER							
48 ••• 63	TARGET PORT IDENTIFIER							

The TYPE field shall be set as shown in table 13 for the SRP_LOGIN_REQ request.

The TAG field is defined in 6.1.

The REQUESTED MAXIMUM INITIATOR TO TARGET IU LENGTH field specifies the maximum length in bytes of any information unit that the SRP initiator port sends on this RDMA channel. This value shall be greater than 63 and shall include the length of any first burst data immediate data.

The REQUIRED BUFFER FORMATS field is defined in 5.6.2.2.

The asynchronous event solicited notification (AESOLNT) bit specifies whether an SRP_AER_REQ request (see 6.12) should use normal message or solicited message reception notification. The AESOLNT bit shall be set to one to request solicited notification or set to zero to request normal notification.

The first burst (FIRSTBURST) bit specifies whether the initiator port is requesting that the target port support first burst data. If FIRSTBURST is set to one, then the initiator port is requesting that the target port accept first burst data within the SRP_CMD request information unit (see 6.8) beginning at the offset specified by FIRST BURST DATA OFFSET field. If FIRSTBURST is set to zero, then there is no first burst data and the device server shall ignore the FIRST BURST DATA OFFSET field.

The immediate data field (IMMEDIATE DATA OFFSET) specifies where the immediate data starts in the SRP_CMD request information unit (see 6.8). The IMMEDIATE DATA OFFSET is measured from the start of the SRP_CMD request. The value of the IMMEDIATE DATA OFFSET shall be such that the immediate data follows the data descriptors in the SRP_CMD request information unit.

The credit request solicited notification (CRSOLNT) bit specifies whether an SRP_CRED_REQ request (see 6.10) should use normal message or solicited message reception notification. The CRSOLNT bit shall be set to one to request solicited notification or set to zero to request normal notification.

The logout solicited notification (LOSOLNT) bit specifies whether an SRP_T_LOGOUT request (see 6.6) should use normal message or solicited message reception notification. The LOSOLNT bit shall be set to one to request solicited notification or set to zero to request normal notification.

The MULTICHANNEL ACTION field (see table 14) indicates how an SRP target port handles existing RDMA channels associated with the same I_T nexus.

Table 14 — MULTICHANNEL ACTION field

Code	Description
00b	Single RDMA channel operation (see 5.1.2)
01b	Multiple independent RDMA channel operation (see 5.1.3)
10b to 11b	Reserved

The INITIATOR PORT IDENTIFIER field and the TARGET PORT IDENTIFIER field specify the I_T nexus that shall be associated with this RDMA channel.

6.3 SRP_LOGIN_RSP response

An SRP_LOGIN_RSP response (see table 15) indicates successful RDMA channel establishment and conveys SRP login parameters from an SRP target port to an SRP initiator port. An SRP_LOGIN_RSP response shall be sent as acceptance data during RDMA channel establishment (see 4.2).

Table 15 — SRP_LOGIN_RSP response

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (C0h)							
1 ••• 3	Reserved							
4	(MSB)							
•••	REQUEST LIMIT DELTA							
7								(LSB)
8	(MSB)							
•••	TAG							
15								(LSB)
16	(MSB)							
•••	MAXIMUM INITIATOR TO TARGET IU LENGTH							
19								(LSB)
20 ••• 23	MAXIMUM TARGET TO INITIATOR IU LENGTH							
24 25	SUPPORTED BUFFER FORMATS							
26	Reserved			SOLNTSUP	Reserved		MULTICHANNEL ACTION	
27 ••• 51	Reserved							

The TYPE field shall be set as shown in table 15 for the SRP_LOGIN_RSP response.

The REQUEST LIMIT DELTA field is defined in 5.5.

The TAG field is defined in 6.1.

The MAXIMUM INITIATOR TO TARGET IU LENGTH field specifies the maximum length in bytes of any information unit that the SRP target port is able to receive on this RDMA channel. This value shall be 64 or larger and greater than or equal to the value of REQUESTED MAXIMUM INITIATOR TO TARGET IU LENGTH field specified in the SRP_LOGIN_REQ request (see 6.2). The SRP initiator port shall not send any information unit on this RDMA channel longer than this value.

The MAXIMUM TARGET TO INITIATOR IU LENGTH field specifies the maximum length in bytes of any information unit that the SRP target port may send on this RDMA channel. This value shall be 52 or larger. The SRP target port shall not send any information unit on this RDMA channel longer than this value.

The SUPPORTED BUFFER FORMATS field is defined in 5.6.2.2.

The solicited notification supported (SOLNTSUP) bit indicates whether the SRP target port supports solicited message reception notification for messages sent from the SRP target port to an SRP initiator port (see 4.3). If the SOLNTSUP bit is one, then the SRP target port supports solicited message reception notification. If the SOLNTSUP bit is zero, then the SRP target port only supports normal message reception notification.

The MULTICHANNEL ACTION field (see table 16) indicates how an SRP target port handles existing RDMA channels associated with the same I_T nexus.

Table 16 — MULTICHANNEL ACTION field

Code	Description
00b	No existing RDMA channels were associated with the same I_T nexus.
01b	One or more existing RDMA channels were terminated.
10b	One or more existing RDMA channels continue to operate independently.
11b	Reserved

6.4 SRP_LOGIN_REJ response

An SRP_LOGIN_REJ response (see table 17) indicates that an RDMA channel is not able to be established. An SRP_LOGIN_REJ response shall be sent as rejection data (see 4.2)

Table 17 — SRP_LOGIN_REJ response

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (C2h)							
1 ••• 3	Reserved							
4	(MSB)							
•••	REASON							
7								(LSB)
8	(MSB)							
•••	TAG							
15								(LSB)
16 ••• 23	Reserved							
24 ••• 25	SUPPORTED BUFFER FORMATS							
26 ••• 31	Reserved							

The TYPE field shall be set as shown in table 17 for the SRP_LOGIN_REJ response.

The REASON field (see table 18) indicates the reason that the RDMA channel was not able to be established.

Table 18 — REASON field

Code	Description
0000 0000h to 0000 FFFFh	Reserved
0001 0000h	Unable to establish RDMA channel, no reason specified
0001 0001h	Insufficient RDMA channel resources
0001 0002h	The contents of the REQUESTED MAXIMUM INITIATOR TO TARGET IU LENGTH field is too large
0001 0003h	Unable to associate RDMA channel with specified I_T nexus
0001 0004h	One or more requested data buffer descriptor formats are not supported
0001 0005h	SRP target port does not support multiple RDMA channels per I_T nexus
0001 0006h	RDMA channel limit reached for this initiator
0001 0007h to FFFF FFFFh	Reserved

The TAG field shall be set to the same value as the TAG field in the SRP_LOGIN_REQ request (see 6.2).

The SUPPORTED BUFFER FORMATS field is defined in 5.6.2.2.

6.5 SRP_I_LOGOUT request

An SRP_I_LOGOUT request (see table 19) is sent by an SRP initiator port to notify the SRP target port that the SRP initiator port is disconnecting the RDMA channel. An SRP_I_LOGOUT request shall be sent as a 16-byte message with normal message reception notification (see 4.3).

After sending an SRP_I_LOGOUT request, an SRP initiator port shall:

1) wait for the RDMA communication service status indication (see 4.5.2); and
2) request that the RDMA channel be disconnected and perform the actions specified in 5.1.4.

Upon receiving an SRP_I_LOGOUT request an SRP target port shall perform the actions specified in 5.1.4. The SRP target port shall not send an SRP response to an SRP_I_LOGOUT request.

Table 19 — SRP_I_LOGOUT request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (03h)							
1 ••• 7	Reserved							
8 ••• 15	(MSB) TAG (LSB)							

The TYPE field shall be set as shown in table 19 for the SRP_I_LOGOUT request.

The TAG field is defined in 6.1.

6.6 SRP_T_LOGOUT request

An SRP_T_LOGOUT request (see table 20) is sent by a SRP target port to notify the SRP initiator port that the SRP target port is disconnecting the RDMA channel. An SRP_T_LOGOUT request shall be sent as a 16-byte message.

After sending an SRP_T_LOGOUT request, an SRP target port shall:

1) wait for the RDMA communication service status indication (see 4.5.2); and
2) request that the RDMA channel be disconnected and perform the actions specified in 5.1.4.

Upon receiving an SRP_T_LOGOUT request an SRP initiator port shall perform the actions specified in 5.1.4. The SRP initiator port shall not send an SRP response to an SRP_T_LOGOUT request.

Table 20 — SRP_T_LOGOUT request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (80h)							
1	Reserved							SOLNT
2	Reserved							
3	Reserved							
4	(MSB)							
•••	REASON							
7								(LSB)
8	(MSB)							
•••	TAG							
15								(LSB)

The TYPE field shall be set as shown in table 20 for the SRP_T_LOGOUT request.

The solicited notification (SOLNT) bit indicates whether the SRP initiator port specified normal message or solicited message reception notification for SRP_T_LOGOUT requests during login (see 6.2). The SOLNT bit shall contain the value specified in the LOSOLNT bit of the SRP_LOGIN_REQ request.

If the SOLNT bit is set to one and the SRP target port supports solicited message reception notification (see 6.3), then the SRP target port shall send the SRP_T_LOGOUT response with solicited message reception notification (see 4.3). If the SOLNT bit is zero, then the SRP target port should send the SRP_T_LOGOUT response with normal message reception notification. An SRP initiator port shall not validate the SOLNT bit against whether an SRP_RSP response was received with normal message or solicited message reception notification.

The REASON field (see table 21) indicates the reason for disconnecting the RDMA channel.

Table 21 — REASON field

Code	Description
0000 0000h	No reason specified
0000 0001h	Inactive RDMA channel (reclaiming resources)
0000 0002h	Invalid information unit TYPE code received by SRP target port
0000 0003h	SRP initiator port sent response (see table 12) with no corresponding SRP target port request (see table 11) outstanding
0000 0004h	RDMA channel disconnected due to multichannel action code in new SRP_LOGIN_REQ
0000 0005h	Reserved
0000 0006h	Unsupported format code specified in DATA-OUT BUFFER DESCRIPTOR FORMAT field
0000 0007h	Unsupported format code specified in DATA-IN BUFFER DESCRIPTOR FORMAT field
0000 0008h	Invalid length for IU type
0000 0009h to FFFF FFFFh	Reserved

The TAG field is defined in 6.1.

6.7 SRP_TSK_MGMT request

An SRP_TSK_MGMT request conveys a SCSI task management request (see table 22). An SRP_TSK_MGMT request shall be sent with normal message reception notification (see 4.3).

Table 22 — SRP_TSK_MGMT request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (01h)							
1	Reserved					UCSOLNT	SCSOLNT	Reserved
2 ••• 7	Reserved							
8	(MSB)							
•••	TAG							
15								(LSB)
16 ••• 19	Reserved							
20 ••• 27	LOGICAL UNIT NUMBER							
28	Reserved							
29	Reserved							
30	TASK MANAGEMENT FUNCTION							
31	Reserved							
31	(MSB)							
•••	TAG OF TASK TO BE MANAGED							
39								(LSB)
40 ••• 47	Reserved							

The TYPE field shall be set as shown in (see table 22) for the SRP_TSK_MGMT request.

The unsuccessful completion solicited notification (UCSOLNT) bit specifies whether an SRP_RSP response (see 6.9) reporting unsuccessful completion of the task management request should use normal message or solicited message reception notification. The UCSOLNT bit shall be set to one to request solicited notification. The UCSOLNT bit shall be set to zero to request normal notification.

The successful completion solicited notification (SCSOLNT) bit specifies whether an SRP_RSP response (see 6.9) reporting successful completion of the task management request should use normal message or solicited message reception notification. The SCSOLNT bit shall be set to one to request solicited notification. The SCSOLNT bit shall be set to zero to request normal notification.

The TAG field is defined in 6.1.

The LOGICAL UNIT NUMBER field identifies the logical unit to which the task management request is directed. The structure of the LOGICAL UNIT NUMBER field shall be as defined in the SAM-5 standard. The LOGICAL UNIT NUMBER field is reserved if the task management request is not directed to either an I_T_L or I_T_L_Q nexus.

The TASK MANAGEMENT FUNCTION field is defined in table 23. If TASK MANAGEMENT FUNCTION field contains a reserved or restricted value, then the task manager shall return an SRP_RSP response (see 6.9) containing GOOD status. The RSP_CODE field shall be set to TASK MANAGEMENT FUNCTION NOT SUPPORTED.

Table 23 — TASK MANAGEMENT FUNCTION field

Code	Description
00h	Reserved
01h	The task manager shall perform an ABORT TASK function (see SAM-5).
02h	The task manager shall perform an ABORT TASK SET function (see SAM-5).
03h	Reserved
04h	The task manager shall perform a CLEAR TASK SET function (see SAM-5).
05h to 07h	Reserved
08h	The task manager shall perform a LOGICAL UNIT RESET function (see SAM-5).
09h to 1Fh	Reserved
20h	Reserved
21h-3Fh	Reserved
40h	The task manager shall perform a CLEAR ACA function (see SAM-5).
41 to FFh	Reserved

If TASK MANAGEMENT FLAGS specifies that an ABORT TASK function shall be performed, then the TAG OF TASK TO BE MANAGED field specifies the tag from the SRP_CMD request (see 6.8) of the task to be aborted. The TAG OF TASK TO BE MANAGED field shall be ignored if TASK MANAGEMENT FLAGS specifies any other function.

6.8 SRP_CMD request

An SRP_CMD request conveys a SCSI command (see table 24). An SRP_CMD request shall be sent as a message whose length is 48 bytes plus the lengths of the ADDITIONAL CDB field, DATA-OUT BUFFER DESCRIPTOR field, and DATA-IN BUFFER DESCRIPTOR field. An SRP_CMD request shall be sent with normal message reception notification (see 4.3).

Table 24 — SRP_TSK_MGMT request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (02h)							
1	Reserved					UCSOLNT	SCSOLNT	Reserved
2 ••• 4	Reserved							
5	DATA-OUT BUFFER DESCRIPTOR FORMAT				DATA-IN BUFFER DESCRIPTOR FORMAT			
6	DATA-OUT BUFFER DESCRIPTOR COUNT							
7	DATA-IN BUFFER DESCRIPTOR COUNT							
8 ••• 15	(MSB) TAG (LSB)							
16 ••• 19	Reserved							
20 ••• 27	LOGICAL UNIT NUMBER							
28	Reserved							
29	Reserved					TASK ATTRIBUTE		
30	Reserved							
31	ADDITIONAL CDB LENGTH (n dwords)						Reserved	
32 ••• 47	CDB							
48 ••• 47+4×**n**	ADDITIONAL CDB							
48+4×**n** ••• 47+4×**n**+do	DATA-OUT BUFFER DESCRIPTOR							
48+4×**n**+do ••• 47+4×**n**+do+di	DATA-OUT BUFFER DESCRIPTOR							

Key:

do = The length in bytes of the DATA-OUT BUFFER DESCRIPTOR field, determined from the format code value contained in the DATA-OUT BUFFER DESCRIPTOR FORMAT field and the count value contained in the DATA-OUT BUFFER DESCRIPTOR COUNT field (see 5.6.2).
di = The length in bytes of the DATA-IN BUFFER DESCRIPTOR field, determined from the format code value contained in the DATA-IN BUFFER DESCRIPTOR FORMAT field and the count value contained in the DATA-IN BUFFER DESCRIPTOR COUNT field (see 5.6.2).

The TYPE field shall be set as shown in table 24 for the SRP_CMD request.

The unsuccessful completion solicited notification (UCSOLNT) bit specifies whether an SRP_RSP response (see 6.9) reporting unsuccessful completion of the task management request should use normal message or solicited message reception notification. The UCSOLNT bit shall be set to one to request solicited notification. The UCSOLNT bit shall be set to zero to request normal notification.

The successful completion solicited notification (SCSOLNT) bit specifies whether an SRP_RSP response (see 6.9) reporting successful completion of the task management request should use normal message or solicited message reception notification. The SCSOLNT bit shall be set to one to request solicited notification. The SCSOLNT bit shall be set to zero to request normal notification.

The DATA-OUT BUFFER DESCRIPTOR FORMAT field specifies the format of the DATA-OUT BUFFER DESCRIPTOR field (see 5.6.2).

The DATA-IN BUFFER DESCRIPTOR FORMAT field specifies the format of the DATA-IN BUFFER DESCRIPTOR field (see 5.6.2).

The DATA-OUT BUFFER DESCRIPTOR COUNT field provides additional information to specify the format of the DATAOUT BUFFER DESCRIPTOR field (see 5.6.2).

The DATA-IN BUFFER DESCRIPTOR COUNT field provides additional information to specify the format of the DATA-IN BUFFER DESCRIPTOR field (see 5.6.2).

The TAG field is defined in 6.1.

The LOGICAL UNIT NUMBER field specifies the logical unit number of the logical unit to which the task router shall route the command. The structure of the LOGICAL UNIT NUMBER field shall be as defined in SAM-5. If the addressed logical unit does not exist, then the task router shall follow the rules for selection of incorrect logical units defined in SAM-5.

The TASK ATTRIBUTE field is defined in table 25.

Table 25 — TASK ATTRIBUTE field

Code	Task attribute	Description
000b	SIMPLE	Specifies that the command be managed according to the rules for a SIMPLE task attribute (see SAM-5).
001b	HEAD OF QUEUE	Specifies that the command be managed according to the rules for a HEAD OF QUEUE task attribute (see SAM-5).
010b	ORDERED	Specifies that the command be managed according to the rules for an ORDERED task attribute (see SAM-5).
011b	Reserved	
100b	ACA	Specifies that the command be managed according to the rules for an ACA task attribute (see SAM-5).
101b to 111b	Reserved	

The ADDITIONAL CDB LENGTH field contains the length in dwords (i.e., four bytes) of the ADDITIONAL CDB field.

The CDB field and ADDITIONAL CDB BYTES field together contain the CDB to be interpreted by the addressed logical unit. Any bytes after the end of the actual CDB within the two fields shall be ignored (e.g., a six-byte CDB occupies the first six bytes of the CDB field, the remaining ten bytes of the CDB field are ignored, and the ADDITIONAL CDB BYTES field is not present).

The contents of the CDB are defined in the SCSI command standards (e.g., SPC-4).

The DATA-OUT BUFFER DESCRIPTOR field specifies the buffer that shall be used for data-out transfers (see 5.6.2).

The DATA-IN BUFFER DESCRIPTOR field specifies the buffer that shall be used for data-in transfers (see 5.6.2).

6.9 SRP_RSP response

An SRP_RSP response (see table 26) conveys an SRP response to an SRP_TSK_MGMT request (see 6.7) or an SRP_CMD request (see 6.8) received by a SRP target port. SRP_RSP responses that contain neither RESPONSE DATA or SENSE DATA shall be sent as a 36 byte message. SRP_RSP responses that contain either RESPONSE DATA or SENSE DATA shall be sent as the minimum length message containing those fields.

Table 26 — SRP_RSP response

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (C1h)							
1	Reserved							SOLNT
2	Reserved							
3	Reserved							
4	(MSB)							
•••	REQUEST LIMIT DELTA							
7								(LSB)
8	(MSB)							
•••	TAG							
15								(LSB)
16	Reserved							
17	Reserved							
18	Reserved		DIUNDER	DIOVER	DOUNDER	DOOVER	SNSVALID	RSPVALID
19	STATUS							
20	(MSB)							
•••	DATA-OUT RESIDUAL COUNT							
23								(LSB)
24	(MSB)							
•••	DATA-IN RESIDUAL COUNT							
27								(LSB)
28	(MSB)							
•••	SENSE DATA LIST LENGTH = n							
31								(LSB)
32	(MSB)							
•••	RESPONSE DATA LIST LENGTH = m							
35								(LSB)
36	(MSB)							
•••	RESPONSE DATA (m bytes long)							
35+m								(LSB)
36+m	(MSB)							
•••	SENSE DATA (n bytes long)							
35+m+n								(LSB)

The TYPE field shall be set as shown in table 26 for the SRP_RSP response.

The solicited notification (SOLNT) bit indicates whether the SRP initiator port specified normal message or solicited message reception notification for this response. If the STATUS field is set to a non-zero value or if the RSP_CODE field is present and set to a non-zero value, then the SOLNT bit shall contain the value specified in the UCSOLNT bit of the corresponding SRP_CMD request or SRP_TSK_MGMT request, otherwise the SOLNT bit shall contain the value specified in the SCSOLNT bit of the corresponding SRP_CMD request or SRP_TSK_MGMT request.

If the SOLNT bit is set to one and the SRP target port supports solicited message reception notification (see 6.3), then the SRP target port shall send the SRP_RSP response with solicited message reception notification (see 4.3), otherwise the SRP target port should send the SRP_RSP response with normal message reception notification. An SRP initiator port shall not validate the SOLNT bit against whether an SRP_RSP response was received with normal message or solicited message reception notification.

The REQUEST LIMIT DELTA field is defined in 5.5.

The TAG field shall contain the same value as the TAG field in the SRP_TSK_MGMT request (see 6.7) or the SRP_CMD request (see 6.8) for which this SRP_RSP response is a response.

A data-out under (DOUNDER) bit set to one indicates that the DATA-OUT RESIDUAL COUNT field is valid and contains the count of data bytes that were expected to be transferred from the data-out buffer that were not transferred. The application client should examine the DATA-OUT RESIDUAL COUNT field in the context of the command to determine whether or not an error condition occurred.

A data-out over (DOOVER) bit set to one indicates that the DATA-OUT RESIDUAL COUNT field is valid and contains the count of data bytes that were not able to be transferred from the data-out buffer as a result of the length of the data-out buffer not being sufficient. The application client should examine the DATA-OUT RESIDUAL COUNT field in the context of the command to determine whether or not an error condition occurred.

If the DOUNDER bit is set to zero and the DOOVER bit is set to zero, then the DATA-OUT RESIDUAL COUNT field is not valid and the SRP initiator port shall ignore the DATA-OUT RESIDUAL COUNT field. The SRP target port shall not set the DOUNDER bit and DOOVER bit to one.

A data-in under (DIUNDER) bit set to one, indicates that the DATA-IN RESIDUAL COUNT field is valid and contains the count of data bytes that were expected to be transferred to the data-in buffer that were not transferred. The application client should examine the DATA-IN RESIDUAL COUNT field in the context of the command to determine whether or not an error condition occurred.

A data-in over (DIOVER) bit set to one indicates that the DATA-IN RESIDUAL COUNT field is valid and contains the count of data bytes that were not able to be transferred to the data-in buffer as a result of the length of the data-in buffer not being sufficient. The application client should examine the DATA-IN RESIDUAL COUNT field in the context of the command to determine whether or not an error condition occurred.

If the DIUNDER bit is set to zero and the DIOVER is set to zero, then the DATA-IN RESIDUAL COUNT field is not valid and the SRP initiator port shall ignore the DATA-IN RESIDUAL COUNT field. The SRP target port shall not set the DIUNDER bit and the DIOVER bit to one.

A sense data valid (SNSVALID) bit set to zero, indicates that the contents of the SENSE DATA LIST LENGTH field shall be ignored and that the SENSE DATA field is not present. A SNSVALID bit set to one indicates that the contents of the SENSE DATA LIST LENGTH field specify the number of bytes in the SENSE DATA field.

If sense data is provided, then the SNSVALID bit shall be set to one and the SENSE DATA LIST LENGTH field shall specify the number of bytes in the SENSE DATA field.

If returning all the sense data provided causes the SRP_RSP response (see 6.9) to be longer than the value of the MAXIMUM TARGET TO INITIATOR IU LENGTH field indicated in the SRP_LOGIN_RSP response (see 6.3) when the RDMA channel was established, then the SRP target port shall return an SRP_RSP response whose length is the value from the MAXIMUM TARGET TO INITIATOR IU LENGTH field. The SENSE DATA field shall be truncated as needed to achieve this length. The SENSE DATA LIST LENGTH field shall contain the length of the truncated SENSE DATA field.

If no sense data is provided, then:

a) THE SNSVALID bit shall be set to zero;
b) the SRP initiator port shall ignore the SENSE DATA LIST LENGTH field; and

c) the SRP initiator port shall assume a sense data length of zero.

A response data valid (RSPVALID) bit set to zero indicates that the contents of the RESPONSE DATA LIST LENGTH field shall be ignored and the RESPONSE DATA field is not present. An RSPVALID bit set to one indicates that:

a) the contents of the RESPONSE DATA LIST LENGTH field specify the number of bytes in the RESPONSE DATA field; and
b) the contents of the STATUS field shall be ignored by the SRP initiator port.

If response data is provided, then the RSPVALID bit shall be set to one and the RESPONSE DATA LIST LENGTH field shall specify the number of bytes in the RESPONSE DATA field (see table 27). The RESPONSE DATA LIST LENGTH field shall be set to four. Other lengths are reserved for future standardization.

If no response data is provided, then:

a) RSPVALID bit shall be set to zero;
b) the SRP initiator port shall ignore the RESPONSE DATA LIST LENGTH field; and
c) the SRP initiator port shall assume a length of zero.

Response data shall be provided in any SRP_RSP response that is sent in response to an SRP_TSK_MGMT request (see 6.7). The information in the RSP_CODE field (see table 28) shall indicate the completion status of the task management function.

Response data shall not be provided in any SRP_RSP response that returns a non-zero status code in the STATUS field.

The STATUS field contains the status of a task that completes. See the SAM-5 for a list of status codes.

If the DOUNDER bit is set to one or the DOOVER bit is set to one, then the DATA-OUT RESIDUAL COUNT field contains a count of the number of residual data bytes that were not transferred from the data-out buffer for this SCSI command. Upon successful completion of an SRP I/O operation, the residual data-out byte count should be zero and the contents of the DATA-OUT RESIDUAL COUNT field should not be valid. Some commands may have a non-zero residual data-out byte count that is not an error. SRP target ports are not required to check the data-out length implied by the contents of the CDB for overrun or underrun before processing a SCSI command.

If the DOUNDER bit is set to one and a transfer that did not fill the entire data-out buffer was performed, then 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)

A DOUNDER bit set to one may not be an error for some devices and some commands.

If the DOOVER bit is set to one, then the transfer was truncated as a result of the data-out transfer required by the SCSI command being longer than the data-out buffer (see 5.6.2). Those bytes that were not able to be transferred without exceeding the length of the data-out buffer shall not be transferred. The data-out residual count is defined as follows:

data-out residual count = (Transfer length required by command) - (data-out buffer length)

If the DOOVER bit is set to one, then the termination state of the SRP I/O operation is not certain. Data may not have been transferred from the data-out buffer and the SCSI status byte may not provide correct command completion information.

If either the DIUNDER is set to one or the DIOVER bit is set to one, then the DATA-IN RESIDUAL COUNT field contains a count of the number of residual data bytes that were not transferred to the data-in buffer for this SCSI command. Upon successful completion of an SRP I/O operation, the residual data-in byte count should be zero and the contents of the DATA-IN RESIDUAL COUNT field should not be valid. Some commands (e.g., INQUIRY) may have a non-zero residual data-in byte count that is not an error. SRP target ports are not required to check the data-in length implied by the contents of the CDB for overrun or underrun before processing a SCSI command.

If the DIUNDER bit is set to one and a transfer that did not fill the entire data-in buffer was performed, then 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)

A DIUNDER bit set to one may not be an error for some devices and some commands.

If the DIOVER bit is set to one, then the transfer was truncated as a result of the data-in transfer required by the SCSI command was longer than the data-in buffer (see 5.6.2). Those bytes that were not able to be transferred without exceeding the length of the data-in buffer shall not be transferred. The data-in residual count is defined as follows:

data-in residual count = (Transfer length required by command) - (data-in buffer length)

If the DIOVER bit is set to one, then the termination state of the SRP I/O operation is not certain. Data may not have been transferred to the data-in buffer and the SCSI status byte may not provide correct command completion information.

The DATA-OUT RESIDUAL COUNT field, DATA-IN RESIDUAL COUNT field, SENSE DATA LIST LENGTH field, and RESPONSE DATA LIST LENGTH field shall always be present in the SRP_RSP response, regardless of whether their contents are valid.

The RESPONSE DATA field (see table 27) contains information describing protocol failures detected during processing of an SRP request received by the SRP target port. The RESPONSE DATA field shall be present if the SRP target port detects any of the conditions described by a non-zero response code (see table 28).

Table 27 — RESPONSE DATA field

Bit / Byte	7	6	5	4	3	2	1	0
0	Reserved							
1	Reserved							
2	Reserved							
3	RSP_CODE							

The RSP_CODE field is defined in table 28.

Table 28 — RSP_CODE field

Code	Description
00h	TASK MANAGEMENT FUNCTION COMPLETE
01h	Reserved
02h	REQUEST FIELDS INVALID
03h	Reserved
04h	TASK MANAGEMENT FUNCTION NOT SUPPORTED
05h	TASK MANAGEMENT FUNCTION FAILED
06h to FFh	Reserved

If the SNSVALID bit is set to one, then the SENSE DATA field shall be set to the sense data (see SAM-5). If no conditions requiring the presentation of SCSI sense data have occurred, then the SENSE DATA field shall not be included in the SRP_RSP response and the SNSVALID bit shall be set to zero.

6.10 SRP_CRED_REQ request

An SRP target port may use SRP_CRED_REQ requests (see table 29) to adjust an SRP initiator port's request limit (see 5.5). An SRP_CRED_REQ request shall be sent as a 16 byte message.

Table 29 — SRP_CRED_REQ request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (81h)							
1	Reserved							SOLNT
2	Reserved							
3	Reserved							
4	(MSB)							
•••	REQUEST LIMIT DELTA							
7								(LSB)
8	(MSB)							
•••	TAG							
15								(LSB)

The TYPE field shall be set as shown in table 29 for the SRP_CRED_REQ request.

The solicited notification (SOLNT) bit indicates whether the SRP initiator port specified normal message or solicited message reception notification during login (see 6.2) for SRP_CRED_REQ requests. The SOLNT bit shall contain the value specified in the CRSOLNT bit of the SRP_LOGIN_REQ request.

If the solicited notification (SOLNT) bit is one and the SRP target port supports solicited message reception notification (see 6.3), then the SRP target port shall send the SRP_CRED_REQ request with solicited message reception notification (see 4.3), otherwise the SRP target port should send the SRP_CRED_REQ request with normal message reception notification. An SRP initiator port shall not validate the SOLNT bit against whether an SRP_CRED_REQ request was actually received with normal message or solicited message reception notification.

The REQUEST LIMIT DELTA field is defined in 5.5.

The TAG field is defined in 6.1.

6.11 SRP_CRED_RSP response

An SRP_CRED_RSP response (see table 30) is the response to an SRP_CRED_REQ request (see 6.10) received by an SRP initiator port. All SRP initiator ports shall support generating SRP_CRED_RSP responses. An SRP_CRED_RSP response shall be sent as a 16-byte message with normal message reception notification (see 4.3).

Table 30 — SRP_CRED_RSP response

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (41h)							
1 ••• 7	Reserved							
8 ••• 15	(MSB) TAG (LSB)							

The TYPE field shall be set as shown in table 30 for the SRP_CRED_RSP response.

The TAG field shall contain the same value as the TAG field in the SRP_CRED_REQ request (see 6.10) for which this SRP_CRED_RSP response is a response.

6.12 SRP_AER_REQ request

A target port sends an SRP_AER_REQ request (see table 31) to report an asynchronous event. An SRP_AER_REQ request shall be sent as the minimum length message capable of carrying the fields. See SPC-4 for parameters for managing the use of asynchronous event reporting.

Table 31 — SRP_AER_REQ request

Byte \ Bit	7	6	5	4	3	2	1	0
0	TYPE (82h)							
1	Reserved							SOLNT
2	Reserved							
3	Reserved							
4	(MSB)							
•••	REQUEST LIMIT DELTA							
7								(LSB)
8	(MSB)							
•••	TAG							
15								(LSB)
16								
•••	Reserved							
19								
20								
•••	LOGICAL UNIT NUMBER							
27								
28	(MSB)							
•••	SENSE DATA LIST LENGTH = n							
31								(LSB)
32								
•••	Reserved							
35								
36	(MSB)							
•••	SENSE DATA (n bytes long)							
35+n								(LSB)

The TYPE field shall be set as shown in table 31 for the SRP_AER_REQ request.

The solicited notification (SOLNT) bit indicates whether the SRP initiator port specified normal message or solicited message reception notification during login (see 6.2) for SRP_CRED_REQ requests. The SOLNT bit shall contain the value specified in the CRSOLNT bit of the SRP_LOGIN_REQ request.

If the solicited notification (SOLNT) bit is one and the SRP target port supports solicited message reception notification (see 6.3), then the SRP target port shall send the SRP_CRED_REQ request with solicited message reception notification (see 4.3), otherwise the SRP target port should send the SRP_CRED_REQ request with normal message reception notification. An SRP initiator port shall not validate the SOLNT bit against whether an SRP_CRED_REQ request was actually received with normal message or solicited message reception notification.

The REQUEST LIMIT DELTA field is defined in 5.5.

The TAG field is defined in 6.1.

The SENSE DATA LIST LENGTH field shall specify the number of bytes in the SENSE DATA field.

If returning all the sense data provided causes the SRP_AER_REQ request to be longer than the value of the MAXIMUM TARGET TO INITIATOR IU LENGTH field indicated in the SRP_LOGIN_RSP response (see 6.3) when the RDMA channel was established, then the SRP target port shall return an SRP_AER_REQ request whose length is the value from the MAXIMUM TARGET TO INITIATOR IU LENGTH field. The SENSE DATA field shall be truncated as needed to achieve this length. The SENSE DATA LIST LENGTH field shall contain the length of the truncated SENSE DATA field.

The SENSE DATA field shall be set to the sense data (see SAM-5).

6.13 SRP_AER_RSP response

An SRP_AER_RSP response (see table 32) conveys an SRP initiator port's SRP response to an SRP_AER_REQ request (see 6.12). An SRP_AER_RSP response shall be sent as a 16-byte message with normal message reception notification (see 4.3).

Table 32 — SRP_AER_RSP response

Bit / Byte	7	6	5	4	3	2	1	0
0	TYPE (41h)							
1								
•••	Reserved							
7								
8	(MSB)							
•••	TAG							
15								(LSB)

The TYPE field shall be set as shown in table 32 for the SRP_AER_RSP response.

The TAG field shall contain the same value as the TAG field in the SRP_AER_REQ request (see 6.12) for which this SRP_AER_RSP response is a response.

7 SCSI mode parameters

7.1 SCSI mode parameter overview and codes

This subclause describes the block descriptors and the pages used with MODE SELECT command and MODE SENSE command that influence, control, and report the behavior of the SRP target port. All mode parameters not defined in this standard shall influence the behavior of the SCSI devices as specified in the appropriate command set standard. The mode pages are addressed to the device server of a logical unit. The mode pages associated with this protocol are listed in table 33.

Table 33 — SRP mode page codes

Mode Page Code	Subpage code	Description	Reference
02h	00h	Disconnect-reconnect page	7.2
18h	00h	Protocol specific LUN page	7.3
	01h to DFh	Reserved	
	E0h to FEh	Vendor specific	
	FFh	Return all subpages for this mode page code	SPC-4
19h	00h	Protocol specific port page	7.4
	01h to DFh	Reserved	
	E0h to FEh	Vendor specific	
	FFh	Return all subpages for this mode page code	SPC-4

If any field in an implemented mode page is not implemented, then the value of the field shall be assumed to be zero (i.e., as if the field is set to zero) (see SPC-4).

If a mode page defined by this standard is not implemented, then the value of each field in that mode page that is:

a) allowed by this standard to be changeable (e.g., not defined as a read only field); and
b) not used solely to define the mode page structure or coordinate access to the mode page,

shall be assumed to be zero (i.e., as if the mode page is implemented and the field is set to zero).

7.2 Disconnect-reconnect mode page

The disconnect-reconnect mode page (see table 34) provides the application client the means to tune the performance of the service delivery subsystem. Table 34 defines the parameters that are applicable to SRP.

The SCSI application client sends the values in the fields to be used by the SCSI device server to control SRP by means of a MODE SELECT command. The SCSI device server shall then communicate the field values to the SRP target port. The field values are communicated from the SCSI device server to the SRP target port in a vendor specific manner.

SRP devices shall only use the parameter fields defined in table 34. If any other fields within the Disconnect-Reconnect mode page of the MODE SELECT command contain a nonzero value, then the SCSI device server shall terminate the MODE SELECT command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN PARAMETER LIST.

Table 34 — Disconnect-Reconnect mode page for SRP

Byte\Bit	7	6	5	4	3	2	1	0
0	PS	SPF (0b)	PAGE CODE (02h)					
1	PAGE LENGTH (0Eh)							
2	Reserved							
3	Reserved							
4	Reserved							
5								
6	Reserved							
7								
8	Reserved							
9								
10	(MSB)	MAXIMUM BURST SIZE						
11								(LSB)
12	EMDP	Reserved						
13	Reserved							
14	(MSB)	~~Reserved~~FIRST BURST SIZE						
15								(LSB)

The PARAMETERS SAVEABLE (PS) bit is defined in SPC-4.

The SUBPAGE FORMAT (SPF) bit is defined in SPC-4 and shall be set as shown in table 34 for the Disconnect-Reconnect mode page for SAS SSP.

The PAGE CODE field is defined in SPC-4 and shall be set as shown in table 34 for the Disconnect-Reconnect mode page for SRP.

The PAGE LENGTH field is defined in SPC-4 and shall be set as shown in table 34 for the Disconnect-Reconnect mode page for SRP.

The MAXIMUM BURST SIZE field indicates the maximum size of an RDMA Read or RDMA Write operation that the device server shall perform. This value is expressed in increments of 512 bytes (e.g., a value of 0001h in this field means that the number of bytes transferred to the SRP initiator port for the nexus is less than or equal to 512, and a value of 0002h in this field means that the number of bytes transferred to the SRP initiator port for the nexus is less than or equal to 1 024). The device server may round this value down as defined in SPC-4. A value of zero indicates that the maximum transfer size is limited only to that of the underlying interconnect. The value zero shall be implemented by all SRP devices. The application client and device server may use the value of this parameter to adjust internal maximum buffering requirements. A router between an SRP device and another protocol device may intercept and adjust this value to reflect its own maximum buffering capabilities.

The enable modify data pointers (EMDP) bit indicates whether the SRP target port may use the random buffer access capability to order RDMA requests for a single SCSI command. An EMDP bit set to zero specifies the SRP target port shall generate RDMA requests with continuously increasing addresses for a single SCSI command. An EMDP bit is set to one specifies the SRP target port may issue RDMAs for a single SCSI command in any order.The EMDP function shall be implemented by SRP devices.

~~The FIRST BURST SIZE field indicates the maximum amount of write data that the target port supports receiving along with the command (i.e., the first burst). This value is expressed in increments of 512 bytes (e.g., a value of 0001h in this field means that the number of bytes transferred by the initiator port is less than or equal to 512 and a value of 0002h in this field means that the number of bytes transferred by the initiator port is less than or equal to 1 024). A value of zero indicates that the target port does not support first burst functionality. The maximum first burst size value for a connection is set at login (see 6.2). This value shall not be changeable (see SPC-4).~~

7.3 Protocol specific LUN mode page

The Protocol Specific LUN mode page shall not be implemented by SRP target ports.

7.4 Protocol specific port mode page

The Protocol Specific Port mode page shall not be implemented by SRP target ports.

Annex A
(Normative)

SRP interface protocol and services

A.1 Service interface protocol

This standard describes a SCSI device’s behavior in terms of functional levels, service interfaces between levels and peer-to-peer protocols. For a full description of the model used in this standard see SAM-2. Figure A.1 shows the model as it appears from the point of view of this standard.

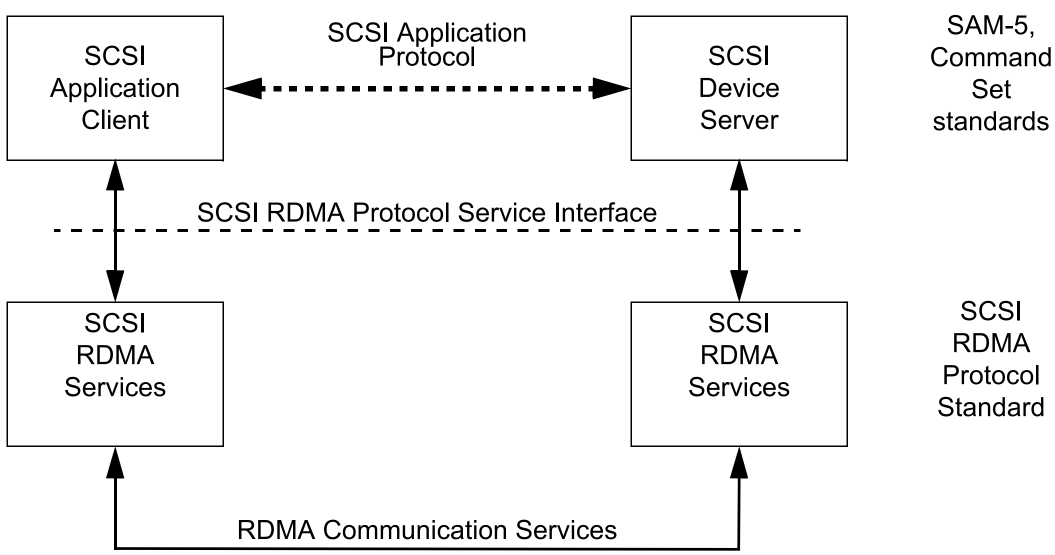


Figure A.1 — SRP reference model

Services between service levels are either four-step confirmed services or two-step confirmed services. A four-step confirmed service consists of a service request, indication, response, and confirmation (see figure A.2). A two-step confirmed service consists of a service request and confirmation.

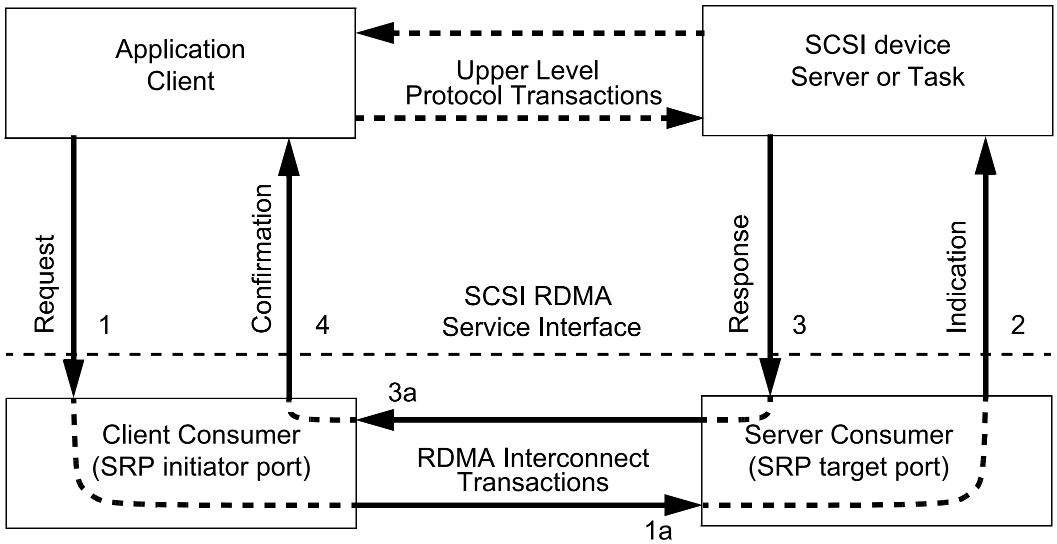


Figure A.2 — Model for a four-step confirmed service

The SCSI RDMA four-step confirmed service protocol consists of the following interactions:

1) a request to the client consumer to invoke a service;
2) an indication from the server consumer notifying the SCSI device server or task manager of an event;
3) a response from the SCSI device server or task manager in reply to an indication; and
4) a confirmation from the client consumer upon service completion.

Only application clients shall request a four-step confirmed service be invoked.

Figure A.3 shows the service and protocol interactions for a two-step confirmed service.

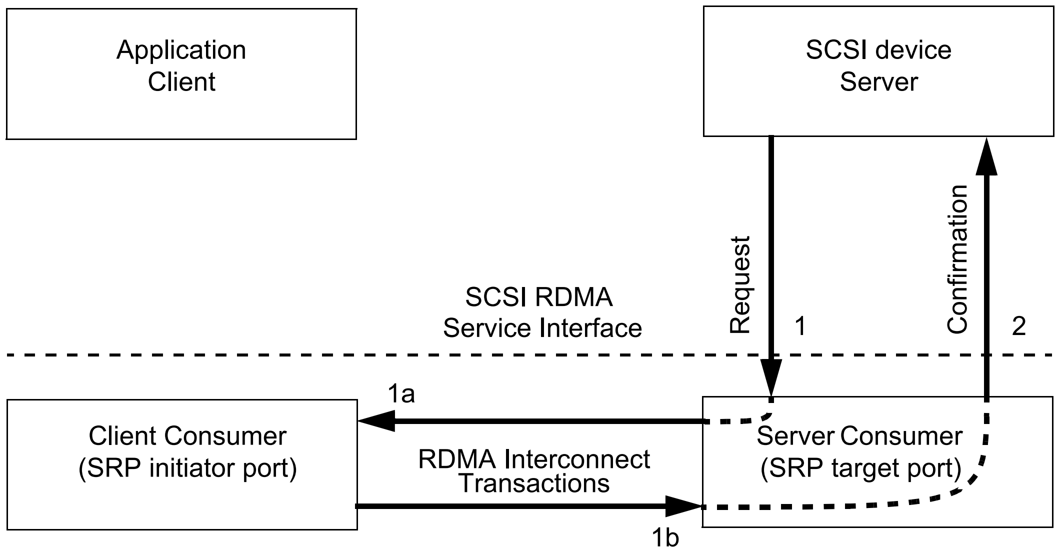


Figure A.3 — Model for a two-step confirmed service

The SCSI RDMA two-step confirmed service interface consists of the following interactions:

1) a request to the server consumer to invoke a service; and
2) a confirmation from the server consumer upon service completion.

Only SCSI device servers shall request a two-step confirmed service be invoked.

A.2 SRP services

SRP provides services to enable an application client to request and manage commands (see SAM-5) and to enable a device server to receive commands and move data to and from an application client. The SRP services are described in terms of the services the SRP initiator port and SRP target port provide.

A.3 Terminology mapping to SAM-5

See table A.1 for the SRP terminology mapping to SAM-5 identifiers and names.

Table A.1 — Terminology mapping to SAM-5

Term used in SAM-5	Term used in this standard
initiator port identifier	initiator port identifier
initiator port name	
target port identifier	target port identifier
target port name	

A.4 Procedure arguments

See table A.2 for a list of the procedure arguments used within this standard, the name of the standard where the arguments are defined, the standard where the binary contents of the arguments are defined, and the routing of the arguments. The routing shows:

a) the source of the argument;
b) the final destination of the argument; and

c) the routing of the argument.

Table A.2 — Procedure arguments

Procedure argument	Standard where argument format is defined	Argument routing
application client buffer offset	SAM-5	DS → targ → init
data-out buffer size	SAM-5	AC → init
data-in buffer size	SAM-5	AC → init
command descriptor block	SAM-5/cmd [a]	AC → init → targ → DS
data-in buffer	cmd [b]	DS → targ → init → AC
data-out buffer	cmd [b]	AC → init → targ → DS
device server buffer	cmd [b]	DS → targ → init
first burst enabled	this standard	AC → init → targ → DS
I_T_L_x nexus	this standard	AC → init → targ → DS **or** AC → init → targ → TM **or** DS → targ → init
request byte count	SAM-5	DS → targ
service response	this standard [c]	DS → targ → init → AC **or** targ → DS
autosense request	SAM-5	AC → init → targ → DS
sense data	SPC-2	DS → targ → init → AC
status	SAM-5	DS → targ → init → AC
task attribute	this standard	AC → init → targ → DS

Key:

AC=application client
DS=device server
initiator port
targ=SRP target port

cmd=SCSI command standards
init=SRP
TM=task manager

[a] The portions not defined in SAM-5 are defined in the SCSI command standards (e.g., SPC-4).
[b] Parameter lists are defined within one of the SCSI command standards (e.g., SPC-4). SCSI standards do not define nonparameter list information.
[c] The SERVICE DELIVERY OR TARGET FAILURE value of the service response is not defined in SCSI.

A.5 Application client SCSI command services

A.5.1 Application client SCSI command services overview

The SCSI command services shall be requested by the application client using a procedure call defined as:

Service response = Execute Command (IN (I_T_L_x Nexus, CDB, Task Attribute, [Data-In Buffer Size], [Data-Out Buffer], [Data-Out Buffer Size], [Command Priority]), OUT ([Data-In Buffer], [Sense Data], [Sense Data Length], Status, [Status Qualifier]))

A.5.2 Send SCSI command service

The send SCSI command service is a four-step confirmed service (see figure A.2) that provides the means to transfer a command data block to a device server.

Processing the execute command procedure call for a send SCSI command service shall be composed of the four-step confirmed service shown in table A.3.

Table A.3 — Processing of execute command procedure call for a send SCSI command service

Step (step number) [a]	Source to Destination	Protocol service name	SCSI Protocol Service Interface procedure calls
request (1)	application client to client consumer	send SCSI command request	**Send SCSI command (IN (I_T_L_x nexus, CDB, [task attribute], [data-in buffer size], [data-out buffer], [data-out buffer size], [First Burst Enabled], autosense request))**
information unit transfer (1a)	client consumer to server consumer	SRP_CMD request or SRP_TSK_MGMT request	See 6.7 and 6.8
indication (2)	server consumer to device server	send SCSI command indication	**SCSI command received (IN (I_T_L_x nexus, command descriptor block, [task attribute], [First Burst Enabled], autosense request))**
If the send SCSI command requires a data transfer, then see A.6.2 for data-out delivery services and A.6.3 for data-in delivery services			
response (3)	device server to server consumer	send SCSI command response	**Send command complete (IN (I_T_L_x nexus, [sense data], status, service response))**
information unit transfer (3a)	server consumer to client consumer	SRP_RSP response	See 6.9
confirmation (4)	client consumer to application client	send SCSI command confirmation	**Command complete received (IN (I_T_L_x nexus, [data-in buffer], [sense data], status, service response))**
[a] See figure A.2 for step number			

A.6 Device server SCSI command services

A.6.1 Device server SCSI command services overview

The SCSI data buffer movement services shall be requested from the device server using a procedure call defined as:

Receive Data-Out (IN (I_T_L_x nexus, device server buffer, application client buffer offset, request byte count))

Either data-in delivery, data-out delivery, both data-in and data-out delivery, or neither data delivery may be used while processing one command. If both are used, then the device server shall combine the data-in and data-out service responses into one service response.

A.6.2 Data-out delivery service

The data-out delivery service is a two-step confirmed service (see figure A.3) that provides the means to transfer a parameter list or data from an SRP initiator port to a device server.

Processing the execute command procedure call for a data-out delivery service shall be composed of the two-step confirmed service shown in table A.4.

Table A.4 — Processing of execute command procedure call for a data-out delivery service

Step (step number) [a]	Source to Destination	Protocol service name	SCSI Protocol Service Interface procedure calls
request (1)	device server to server consumer	data-out delivery request	**Receive data-out (IN (I_T_L_x nexus, application client buffer offset, request byte count, device server buffer))**
data-out transfer (1a and 1b)	server consumer to client consumer [b]	RDMA data-out transfer	See 4.4.3.
confirmation (2)	server consumer to device server	data-out delivery confirmation	**Data-out received (IN (I_T_L_x nexus))**

[a] See figure A.3 for step number
[b] RDMA transfers are typically performed by hardware without the intervention of the client consumer

A.6.3 Data-in delivery service

The data-in delivery service is a two-step confirmed service (see figure A.3) that provides the means to transfer a parameter list or data from a device server to an SRP initiator port.

Processing the execute command procedure call for a data-in delivery service shall be composed of the two-step confirmed service shown in table A.5.

Table A.5 — Processing of execute command procedure call for a data-in delivery service

Step (step number) [a]	Source to Destination	Protocol service name	SCSI Protocol Service Interface procedure calls
request (1)	device server to server consumer	data-in delivery request	**Send data-in (IN (I_T_L_x nexus, device server buffer, application client buffer offset, request byte count))**
data-in transfer (1a and 1b)	server consumer to client consumer [b]	RDMA data-in transfer	See 4.4.
confirmation (2)	server consumer to device server	data-in delivery confirmation	**Data-In delivered (IN (I_T_L_x nexus))**

[a] See figure A.3 for step number
[b] RDMA transfers are typically performed by hardware without the intervention of the client consumer

A.7 Task management services

A.7.1 Task management functions overview

The task management services shall be requested from the application client using a procedure call defined as:

Function name (IN (nexus), service response)

A.7.2 Task management functions

This standard handles task management functions as a four-step confirmed service that provides the means to transfer task management functions to a task manager.

The task management functions are defined in the SAM-2. This standard defines the actions taken by the SRP services to carry out the requested task management functions.

A.7.3 ABORT TASK

The SRP services request the SRP initiator port issue an SRP_TSK_MGMT request (see 6.7) with a TASK MANAGEMENT FLAGS field set to indicate an ABORT TASK function to be sent to the selected SCSI device.

A.7.4 ABORT TASK SET

The SRP services request the SRP initiator port issue an SRP_TSK_MGMT request (see 6.7) with a TASK MANAGEMENT FLAGS field set to indicate an ABORT TASK SET function to be sent to the selected SCSI device.

A.7.5 CLEAR ACA

The SRP services request the SRP initiator port issue an SRP_TSK_MGMT request (see 6.7) with a TASK MANAGEMENT FLAGS field set to indicate a CLEAR ACA function to be sent to the selected SCSI device.

A.7.6 CLEAR TASK SET

The SRP services request the SRP initiator port issue an SRP_TSK_MGMT request (see 6.7) with a TASK MANAGEMENT FLAGS field set to indicate a CLEAR TASK SET function to be sent to the selected SCSI device.

A.7.7 LOGICAL UNIT RESET

The SRP services request the SRP initiator port issue an SRP_TSK_MGMT request (see 6.7) with a TASK MANAGEMENT FLAGS field set to indicate a LOGICAL UNIT RESET function to be sent to the selected SCSI device.

A.7.8 TARGET RESET

This protocol does not support use of the TARGET RESET task management function.

A.7.9 WAKEUP

This protocol does not support use of the WAKEUP task management function.

Annex B
(Normative)

SRP for the InfiniBand™ Architecture

B.1 Overview

This annex specifies requirements for mapping SRP onto the IB, a transport that implements a superset of the RDMA communication service (see clause 4). See IBAS for a description of the IB.

B.2 Normative references

Infiniband™ Architecture Specification Volume 1 Release 1.0.a[1]

IETF RFC 2373, IP Version 6 Addressing Architecture[2]

B.3 Definitions

B.3.1 Introduction to definitions

The definitions in B.3 are incomplete without reference to IBAS.

B.3.2 Definitions

B.3.3 IB channel adapter
device that terminates an IB link and processes transport level functions

B.3.4 IB channel adapter GUID
IB GUID that uniquely identifies an IB channel adapter

B.3.5 IB communication manager
software, hardware, or combination of the two that supports the IB communication management mechanisms and protocols

B.3.6 IB consumer
object that communicates with other IB consumers using the IB

B.3.7 IB GID
128-bit value that conforms to the IPv6 address format

B.3.8 IB GUID
globally unique value that identifies an IB device or component

B.3.9 IB General Service Interface
interface providing management services other than IB subnet management

B.3.10 IB I/O controller
part of an IB I/O unit that provides I/O services

1.Copies of the Infiniband™ Architecture Specification may be obtained from the Infiniband™ Trade Association at www.infinibandta.org.
2.Copies of the IETF RFCs may be obtained at http://www.ietf.org/.

B.3.11 I/O controller GUID
IB GUID that uniquely identifies an IB I/O controller

Note 1 to entry: This value is present as the GUID field of the IOControllerProfile attribute.

Note 2 to entry: See table B.7.

B.3.12 IB I/O unit
one or more IB I/O controllers attached to the IB fabric through a single IB channel adapter

B.3.13 IB LID
port address used for directing IB packets within an IB subnet

B.3.14 IB MAD
IB packet used to manage an IB network

B.3.15 IB packet
indivisible unit of IB data transfer and routing, consisting of one or more headers, a packet payload, and one or two CRCs

B.3.16 IB port
location on an IB channel adapter, switch, or router to which a link connects

B.3.17 IB port GUID
IB GUID that uniquely identifies an IB port

B.3.18 IB Queue Pair
interface used for communication, consisting of a Send work queue and a Receive work queue

B.3.19 IB service ID
value that allows an IB communication manager to associate an incoming connection request with the entity providing the service

B.3.20 IB subnet
set of IB ports connected via IB switches that have a common IB subnet ID and are managed by a common IB subnet manager

B.3.21 IB subnet manager
entity that configures and controls an IB subnet

Note 3 to entry: See IBAS.

B.3.22 IPv6 address
28-bit address constructed in accordance with IETF RFC 2373 for Internet Protocol version 6

Note 4 to entry: See IETF RFC 2373.

B.4 Abbreviations

The abbreviations in this subclause are incomplete without reference to IBAS.

Abbreviations used in this annex:

Abbreviation	Meaning
CM:Ready To Use	IB communication manager Ready to Use message
CM:Reject	IB communication manager Reject message
CM:Request	IB communication manager Request message
CM:Response	IB communication manager Response message
CRC	Cyclic Redundancy Check
GID	IB Global ID
GUID	Globally unique identifier
IB	InfiniBand™ Architecture
IBAS	Infiniband™ Architecture Specification Volume 1 Release 1.0.a
IOC	IB IO Controller
IPv6	Internet Protocol version 6
LID	IB Local ID
MAD	IB Management datagram
QP	IB Queue pair

B.5 IB overview

IB devices contain IB consumers and one or more IB channel adapters. Each IB channel adapter contains one or more IB ports. Associated with each IB channel adapter are IB QPs that interface between IB consumers and the IB channel adapter. Figure B.1 shows an example IB device.

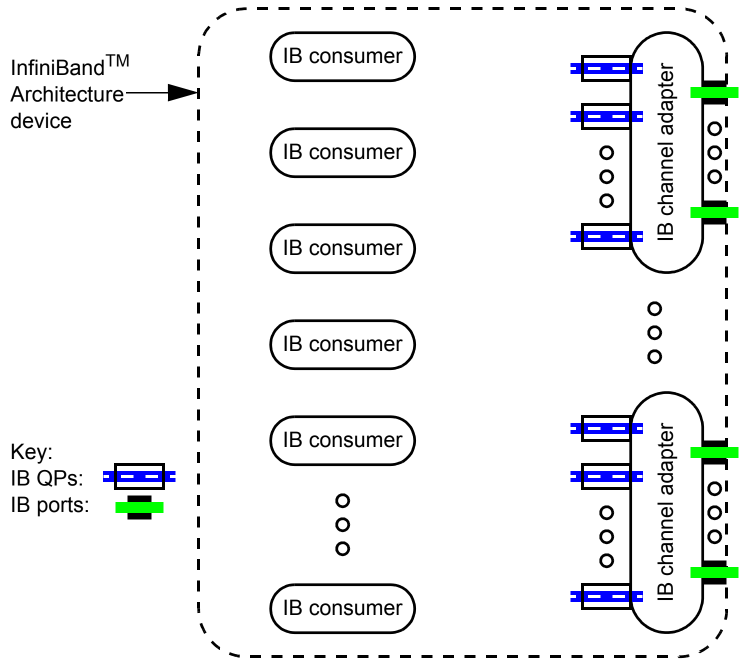


Figure B.1 — IB device example

An IB I/O unit is an IB device that contains an IB channel adapter with one or more IB ports, IB QPs, and one or more IB I/O controllers. Figure B.2 shows an example IB I/O unit.

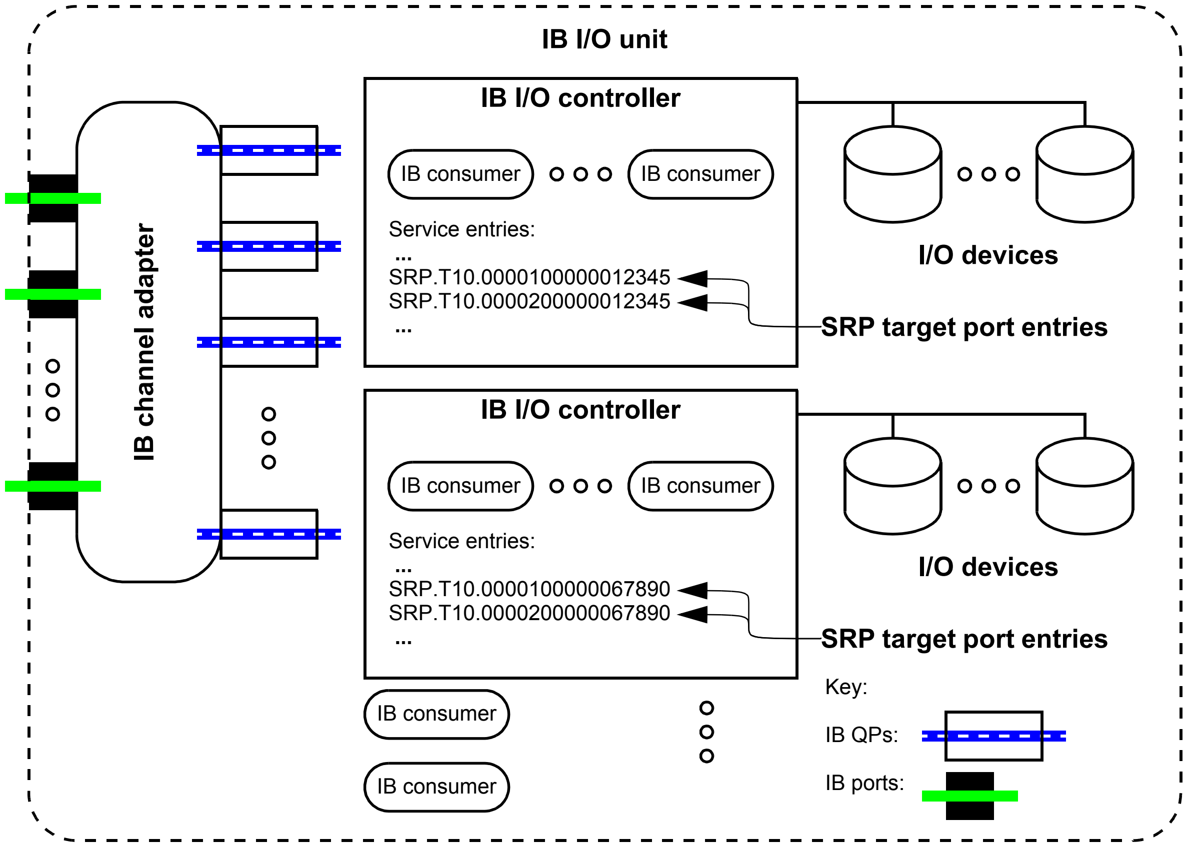


Figure B.2 — IB I/O unit example

Each IB port has a 64-bit globally unique identifier called an IB port GUID. Each IB channel adapter has a IB channel adapter GUID, which is shared by all IB ports on the IB channel adapter.

Each IB I/O controller has an IB I/O controller GUID.

The IB subnet manager assigns one or more IB LIDs and one or more IB GIDs to each IB port. Table B.1 summarizes the IB names (IB GUIDs) and addresses (IDs) relevant to this protocol.

Table B.1 — IB names and addresses

Name	Scope of uniqueness	Size	Description
IB port GUID	worldwide	64 bits	Identifies an IB port
IB channel adapter GUID	worldwide	64 bits	Identifies a IB channel adapter
IB I/O controller GUID	worldwide	64 bits	Identifies an IB I/O controller
IB LID	IB subnet	16 bits	Local routing address assigned to each IB port by the IB subnet manager
IB GID	varies [a]	128 bits	Address assigned by the IB subnet manager; (e.g., IB subnet prefix plus the IB port GUID)
[a] Refer to IBAS			

B.6 SCSI architecture mapping

Figure B.3 shows how SCSI initiator devices, SRP initiator ports, SRP target ports, and SCSI target devices map to InfiniBand^{TM} Architecture objects.

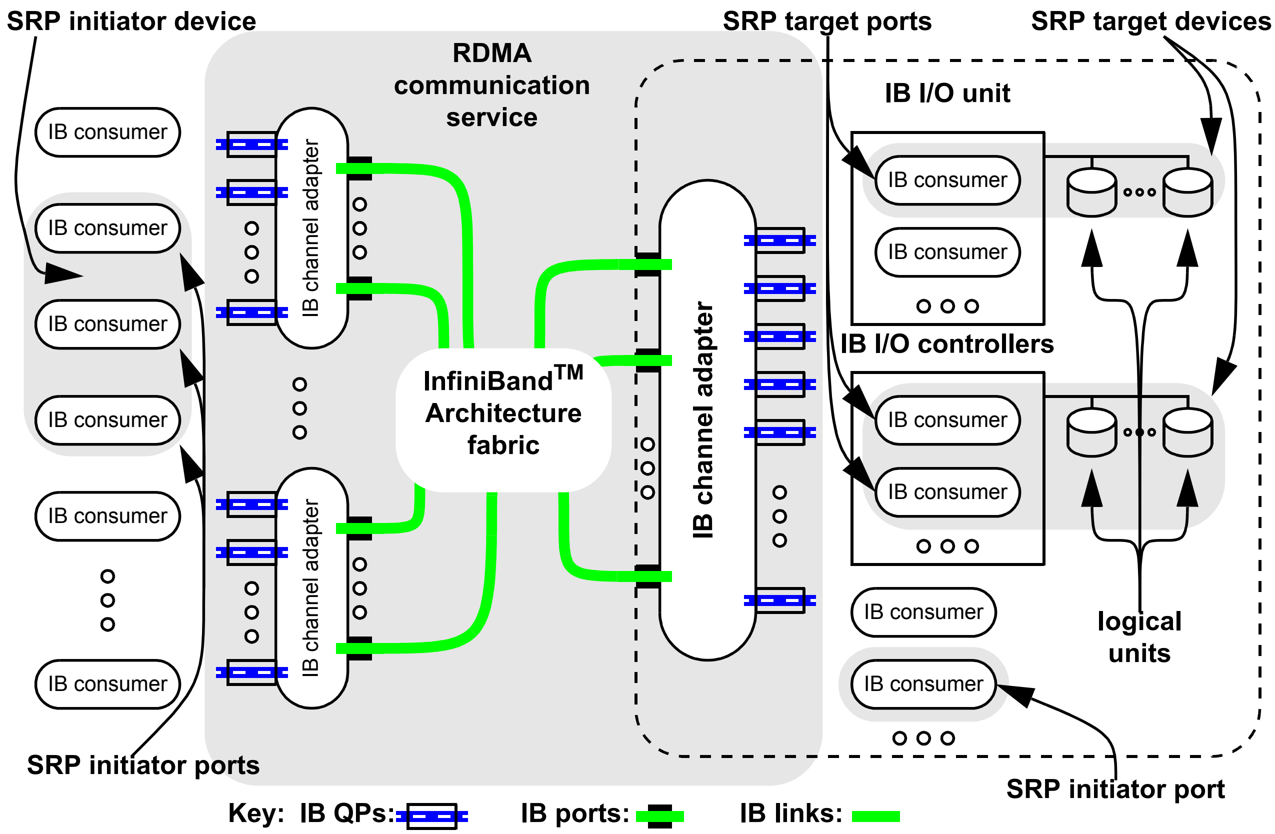


Figure B.3 — SCSI architecture mapping

The RDMA communication service (see clause 4) includes IB queue pairs, IB channel adapters, IB ports, software, and the IB fabric.

An IB consumer in any IB device may be an SRP initiator port. An SRP initiator device may consist of one or more IB consumers. The SRP initiator port identifier shall be constructed as shown in table B.2.

Table B.2 — IB SRP initiator port identifier

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB)							
•••				IDENTIFIER EXTENSION				
7								(LSB)
8	(MSB)							
•••				GUID (e.g., IB channel adapter GUID)				
15								(LSB)

The IDENTIFIER EXTENSION field shall be chosen by the SRP initiator port to ensure that all SRP initiator port identifiers are unique.

The GUID field should an IB GUID available to the SRP initiator port.

SRP target ports shall be implemented by IB IO Controllers in IB I/O units. The IB I/O unit shall include an IB device management agent to provide the IOUnitInfo, IOControllerProfile, and ServiceEntries attributes.

An SRP target port is identified by a ServiceEntries attribute of an IB I/O controller. The SRP target port identifier shall be constructed as shown in table B.3.

Table B.3 — IB SRP target port identifier

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
•••	IDENTIFIER EXTENSION							
7								(LSB)
8	(MSB)							
•••	IO CONTROLLER GUID							
15								(LSB)

The IDENTIFIER EXTENSION field shall be the value from the ServiceEntries attribute that identifies the SRP target port (see table B.8).

The IO CONTROLLER GUID field shall be the IB I/O controller GUID of the IB I/O controller providing the SRP target port.

B.7 Communication management

B.7.1 Communication management overview

IB communications managers on each IB device manage IB connections using IB MADs transported over the IB General Service Interface. SRP initiator ports and SRP target ports shall use the active/passive (client/server) connection establishment protocol. The processor unit or IB I/O controller containing the SRP target port shall act as the server and the processor unit or IB I/O controller containing the SRP initiator port shall act as the client.

B.7.2 Discovering SRP target ports

To discover the IB service ID of an SRP target port in an IB I/O unit, an SRP initiator port may use this sequence:

1) retrieve the IOUnitInfo attribute from an IB I/O unit using a DevMgtGet IB MAD to determine the presence and slot number of each IB I/O controller attached to the IB I/O unit;
2) retrieve the IOControllerProfile attributes from each IB I/O controller, each of which includes a ServiceEntries table; and
3) search the ServiceEntries table for service names matching the rules described in table B.8.

The IB service ID associated with each matching service name may be used in the communication management process to establish IB connections to IB I/O controllers providing SRP target ports. The SRP target port identifier for each SRP target port is constructed as described in table B.3.

B.7.3 Establishing a connection

To establish an IB connection, the client places the IB service ID in an IB communication management CM:Request message. The server associates the request with the appropriate SRP target port. The PrivateData field of the CM:Request message shall include an SRP_LOGIN_REQ request (see 6.2).

The SRP target port may choose to refuse the connection based on the SRP_LOGIN_REQ request content by returning a CM:Reject message with the reason code set to Consumer Reject. The PrivateData field of the CM:Reject message shall include an SRP_LOGIN_REJ response (see 6.4).

The SRP target port may choose to redirect the connection to a different endpoint (e.g. another IB port) by returning a CM:Reject message with the reason code set to either PORT AND CM REDIRECTION or PORT REDIRECTION. The SRP initiator port should retry the connection establishment using the new endpoint. See IB Specification Volume 1 Release 1.0.a.

If the server accepts the connection request and SRP login, then the server returns a CM:Response message. The PrivateData field of the CM:Response message shall include an SRP_LOGIN_RSP response (see 6.3). The SRP initiator port may choose to refuse the connection based on the SRP_LOGIN_RSP response content by returning a CM:Reject message with a Reason code set to Consumer Reject. In this case, the PrivateData field of the CM:Reject message is reserved.

If the client accepts the connection reply and the SRP login response, then it replies with a CM:Ready To Use message indicating both an IB and an SRP connection are open. The client may then start using the connection for communication.

B.7.4 Releasing a connection

The SRP initiator port may send an SRP_I_LOGOUT request or the SRP target port may send an SRP_T_LOGOUT request with a SEND operation. The sender shall send a CM Disconnect Request as described in IBAS upon receipt of an IB transport level acknowledgement to the SRP_I_LOGOUT request or SRP_T_LOGOUT request information unit. The receiver of an SRP_I_LOGOUT request or SRP_T_LOGOUT request information unit shall respond with an IB transport acknowledgement and may send a CM Disconnect Request as described in IBAS, or may wait to receive a CM Disconnect Request.

B.7.5 Errors

Some errors cause an IB queue pair to enter the Error state, which destroys the connection. The IB communication manager for the queue pair consumer should send a CM Disconnect Request as described in IBAS.

B.7.6 Data-out and data-in operations

An SRP target port shall map a Receive Data-out SCSI protocol service interface procedure call to one or more IB RDMA READ requests. An SRP target port shall map a Send Data-in SCSI protocol service interface procedure call to one or more IB RDMA WRITE requests. Table B.4 specifies the value of the IB RDMA header fields.

Table B.4 — IB RDMA header fields

IB RDMA Extended Transport Header field	Value
Virtual Address	VIRTUAL ADDRESS [a] + application client buffer offset [b]
Remote Key	MEMORY HANDLE [c]
DMA Length	request byte count [d]

[a] The contents of the VIRTUAL ADDRESS field in the memory descriptor (see table 5).
[b] The application client buffer offset parameter to the receive data-out (see table A.4) or send data-in (see table A.5) SCSI protocol service interface procedure call.
[c] The contents of the MEMORY HANDLE field in the memory descriptor (see table 5).
[d] The request byte count parameter to the receive data-out (see table A.4) or send data-in (see table A.5) SCSI protocol service interface procedure call.

B.8 InfiniBand™ Architecture protocol requirements

SRP target ports and SRP initiator ports shall support the Reliable Connection transport service type.

SRP target ports shall implement the device management class of general management services.

SRP initiator ports and SRP target ports shall support the transport functions described in table B.5.

Table B.5 — Transport operation support requirements

Transport functions	SRP initiator port	SRP target port
Send to	Mandatory	Mandatory
Send from	Mandatory	Mandatory
RDMA write to	Mandatory	Not used
RDMA write from	Not used	Mandatory
RDMA read to	Mandatory for data-out commands	Not used
RDMA read from	Not used	Mandatory for data-out commands
RDMA Write with immediate data (to or from)	Not used	Not used
ATOMIC (to or from)	Not used	Not used

IB I/O units containing an IB I/O controller acting as an SRP target port shall report the device management IOUnit attributes defined in IBAS as described in table B.6.

Table B.6 — IOUnit attributes for SRP target ports

Field	SRP requirement
Max Controllers	At least one
Controller List	At least one IB I/O controller acting as an SRP target port shall be present
[a] This protocol does not change or override IB requirements on the values of fields not listed.	

IB I/O controllers acting as SRP target ports shall report the device management IOControllerProfile attributes defined in IBAS as described in table B.7.

Table B.7 — IOControllerProfile attributes for SRP target ports

Field	SRP requirement
I/O Class	0100h
I/O Subclass	609Eh
Protocol	0108h
Protocol Version	0001h
Service Connections	At least one
Initiators Supported	At least one
Send Message Depth	Reserved
RDMA Read Depth	Maximum IOC-issued RDMA depth [a]
Send Message Size	MAXIMUM INITIATOR TO TARGET IU SIZE [b]
RDMA Transfer Size	Reserved
Controller Operations Capability Mask:	
0: ST; Send Messages To IOCs	Shall be set to one.
1: SF; Send Messages From IOCs	Shall be set to one.
2: RT; RDMA Read Requests To IOCs	No requirement
3: RF; RDMA Read Requests From IOCs	Shall be set to one if an SRP target port supports data-out commands. No requirement otherwise.
4: WT; RDMA Write Requests To IOCs	No requirement
5: WF; RDMA Write Requests From IOCs	Shall be set to one.
6: AT; Atomic Operations To IOCs	No requirement
7: AF; Atomic Operations From IOCs	No requirement
Controller Services Capability Mask	Reserved [a]
Service Entries	At least one

[a] This protocol does not change or override InfiniBand Architecture requirements on the values of fields not listed, or for those marked as having 'no requirement'.

[b] The largest number of RDMA Read requests that this IO Controller may have outstanding on one channel.

[c] This value shall be no less than the largest value, in bytes, of MAXIMUM INITIATOR TO TARGET IU SIZE that this IO Controller shall return in the SRP_LOGIN_RSP information unit.

IB I/O controllers providing SRP target ports shall include at least one ServiceName/ServiceID pair in the device management ServiceEntries attribute pair (see IBAS) as described in table B.8.

Table B.8 — ServiceEntries attribute pair for SRP target ports

Field	Length (bits)	SRP requirement
ServiceName_n	320	’SRP.T10:xxxxxxxxxxxxxxxx’ or ’SRP.T10:xxxxxxxxxxxxxxxx:reserved’
ServiceID_n	64	Assigned by the IB I/O controller

[a] A service name that identifies an SRP target port shall meet the rules described in this table.
[b] The string ’SRP.T10’ and the colons shall appear exactly as shown (e.g. capital letters only).
[c] The string ’xxxxxxxxxxxxxxxx’ in the service name shall be sixteen hexadecimal digits. Only the characters 0 to 9 and A to F (capital letters only) are permitted. If any other character appears, then the service name shall not be recognized as identifying an SRP target port.
[d] The string ’xxxxxxxxxxxxxxxx’ in the service name identifies the 64-bit extension identifier value used to construct the SRP target port identifier (see table B.3)
[e] The literal string ’reserved’ shall either be ignored by SRP initiator ports or treated in accordance with a future revision of this standard.
[f] If the service name does not completely fill ServiceName_n field (i.e. it is less than 40 bytes), then it shall be extended with null characters (i.e., binary zeros).

Bibliography

ISO/IEC 14776-153, *Serial Attached SCSI-2 (SAS-3)*

ISO/IEC 14776-323, *SCSI Block Command-3 (SBC-3)*

ISO/IEC 14776-263, *SAS Protocol Layer-3 (SPL-3)*