

To: T10 Technical Committee
 From: Rob Elliott, Compaq Computer Corporation (Robert.Elliott@compaq.com)
 Date: 21 May 2001
 Subject: SRP bidirectional residuals

Revision History

Revision 0 (18 May 2001) first revision

Revision 1 (21 May 2001) incorporated comments from George Penokie - changed execute to process, changed some whichs to thats, and reworted the "devices having indeterminate data lengths" phrases.

Related Documents

srp-r04 - SCSI RDMA Protocol revision 4

fcp2r07 - Fibre Channel Protocol - 2 revision 7

01-092r1 - Minutes of the SRP WG, March 8-9, 2001

Overview

SRP revision 4 does not explain how residuals for bidirectional transfers are handled. FCP-2 did not change the meaning of the OVER and UNDER bits and the RESIDUAL_COUNT field for unidirectional transfers; they can apply to either read or write direction depending on the command. For bidirectional commands, new BIDI_OVER, BIDI_UNDER, and BIDI_RESIDUAL_COUNT fields were added to cover the read direction; the historical fields are used for the write direction. iSCSI follows that mold, but it implements bidirectional commands using an optional header in the CDB. (Although it saves the bytes in the CMD IU, it always includes the extra residual field in the RSP IU.)

Since SRP natively supports bidirectional commands, it makes sense to support both direction's residuals as was done for scatter gather lists. A unidirectional read command uses a DATA IN RESIDUAL COUNT; a unidirectional write command uses a DATA OUT RESIDUAL COUNT; a bidirectional command uses both.

The disadvantage with this approach is it makes SRP_RSP 36 bytes instead of 32 bytes when there is no response or sense data. There are enough reserved bytes (bytes 2, 3, 4, 16, and 17) that the format could be completely restructured to fit in the same size, but that would destroy commonality with other IUs. The extra residual could be defined as only present if the associated over/under bit is set, but that creates a variable length structure that is difficult to handle. Or, the whole residual feature could be dropped.

The March SRP working group "agreed that each descriptor will define what constitutes the 'data length' used for under/overrun calculations." Text doing this is also proposed.

Suggested Changes

4.4.2.1 Direct data buffer descriptors

The presence of a direct data buffer descriptor is indicated by DOIND or DIIND containing 0 and DOCOUNT or DICOUNT containing 1. A direct data buffer descriptor contains a single memory descriptor (table 1). The memory descriptor identifies the data buffer, which is a single contiguous memory segment within a memory region's virtual address space. If the direct data buffer defines a data out buffer, the SRP target shall only issue RDMA Read operations to the memory segment. If the direct data buffer defines a data in buffer, the SRP target shall only issue RDMA Write operations to the memory segment. The target shall use the DATA LENGTH field as the data buffer length for residual calculations.

4.4.2.2 Indirect data buffer descriptors

INDIRECT TABLE MEMORY DESCRIPTOR identifies a memory segment that contains the indirect table. The indirect table is a list of one or more memory descriptors. The memory segments identified

by the memory descriptors in the indirect table, concatenated together, comprise the indirect data buffer. The DATA LENGTH field of INDIRECT TABLE MEMORY DESCRIPTOR ~~is expected to shall~~ be the number of memory descriptors in the indirect table times 64. ~~The consequences of any other value are vendor specific.~~

TOTAL LENGTH ~~is expected to shall~~ contain the sum of the DATA LENGTH fields of the memory descriptors in the indirect table. ~~The consequences of any other value are vendor specific. The target shall use the TOTAL LENGTH field as the data buffer length for residual calculations.~~

CACHED MEMORY DESCRIPTOR LIST is only present when DOCOUNT or DICOUNT contain a value larger than 1. It is a list of m-1 memory descriptors, where m is the value contained in DOCOUNT or DICOUNT. CACHED MEMORY DESCRIPTOR LIST ~~is expected to shall~~ contain identical copies of the first m-1 memory descriptors in the indirect table. ~~The consequences of any other value are vendor specific. The consequences of CACHED MEMORY DESCRIPTOR LIST being longer than the indirect table are vendor specific.~~

...

5.7 SRP RSP response

[In Table 19, change the fields as described:

Byte	Bit	Current	Proposed
18	2	RIDOVER	DO_OVER
18	3	RIDUNDER	DO_UNDER
18	4	reserved	DI_OVER
18	5	reserved	DI_UNDER
20-23		RESIDUAL COUNT	DATA OUT RESIDUAL COUNT
24-27		insert new field DATA IN RESIDUAL COUNT	
28-end		shift 24-end down to 28-n	

Change the 4 byte RESIDUAL COUNT field to DATA OUT RESIDUAL COUNT.

Add a 4 byte DATA IN RESIDUAL COUNT right after it.

[Editor's note: Consider creating subsections for each field description as was done for FCP-2.]
Change the text as follows:

DO RIDUNDER, when set to 1, 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, but 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.

DO RIDOVER, when set to 1, indicates that the DATA OUT RESIDUAL COUNT field is valid and contains the count of data bytes that could not be transferred from the data out buffer because the length of the data out buffer was not 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.

DO RIDUNDER and DO RIDOVER, when both set to 0, indicate that the DATA OUT RESIDUAL COUNT field is not valid; the initiator shall ignore its contents. The target shall not set both DO RIDUNDER and DO RIDOVER to 1.

DI RIDUNDER, when set to 1, 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, but 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.

DI_RIDOVER, when set to 1, indicates that the DATA IN RESIDUAL COUNT field is valid and contains the count of data bytes that could not be transferred to the data in buffer because the length of the data in buffer was not 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.

DI_RIDUNDER and DI_RIDOVER, when both set to 0, indicate that the DATA IN RESIDUAL COUNT field is not valid; the initiator shall ignore its contents. The target shall not set both DI_RIDUNDER and DI_RIDOVER to 1.

...

If either DO_RIDUNDER or DO_RIDOVER is set to 1, the DATA OUT RESIDUAL COUNT field contains a count of the number of residual data bytes ~~which could not be~~ 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 is normally zero and the DATA OUT RESIDUAL COUNT value is not valid. ~~Devices having indeterminate data lengths~~ Some commands may have a non-zero residual data out byte count ~~after completing valid operations~~ that is not an error. Targets are not required to ~~verify that~~ check the data out length implied by the contents of the CDB ~~will create a data transfer for~~ overruns or underruns before ~~beginning execution~~ of processing a SCSI command.

If DO_RIDUNDER is set to 1, a transfer that did not ~~fill~~ use the entire data out buffer was performed and the value of DATA OUT RESIDUAL COUNT shall be equal to:
data out buffer length - (highest offset of any data out byte transferred + 1)

A condition of DO_RIDUNDER set to 1 may not be an error for some devices and some commands.

If DO_RIDOVER is set to 1, the transfer was truncated because the data out transfer required by the SCSI command was longer than the data out buffer. Those bytes that could ~~not~~ be transferred without exceeding the length of the data out buffer ~~shall not~~ may be transferred. DATA OUT RESIDUAL COUNT shall be equal to:

data out transfer length required by command - data out buffer length

If DO_RIDOVER is set to 1, the termination state of the SRP I/O operation is not certain. Data may or may not have been transferred from the data out buffer and the SCSI status byte may or may not provide correct command completion information.

If either DI_RIDUNDER or DI_RIDOVER is set to 1, 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 is normally zero and the DATA IN RESIDUAL COUNT value is not valid. Some commands may result in a non-zero residual data in byte count that is not an error. Targets are not required to check the data in length implied by the contents of the CDB for overruns or underruns before processing a SCSI command.

If DI_RIDUNDER is set to 1, a transfer that did not use the entire data in buffer was performed and the value of DATA IN RESIDUAL COUNT shall be equal to:
_____ data in buffer length - (highest offset of any data in byte transferred + 1)

A condition of DI_RIDUNDER set to 1 may not be an error for some devices and some commands.

If DI_RIDOVER is set to 1, the transfer was truncated because the data in transfer required by the SCSI command was longer than the data in buffer. Those bytes that could be transferred

without exceeding the length of the data in buffer may be transferred. DATA IN RESIDUAL COUNT shall be equal to:

data in transfer length required by command - data in buffer length

If DI_RIDOVER is set to 1, the termination state of the SRP I/O operation is not certain. Data may or may not have been transferred to the data in buffer and the SCSI status byte may or may not provide correct command completion information.

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