To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 9 October 2002
Subject: T10/02-404r0 SAM-3 Data In Size and Sense Data Size for Execute Command

Revision history
Revision 0 (9 October 2002) first revision

Related documents
sam3r03 - SCSI Architecture Model - 3 revision 3
spc3r09 - SCSI Primary Commands - 3 revision 9
sas-r02a - Serial Attached SCSI revision 2a

Overview
Thanks to Ralph Weber for pointing out this issue.

Issue #1. SCSI read commands like INQUIRY can return variable amounts of data. However, the Execute Command () RPC in SAM-3 does not return the number of bytes returned in the data-in buffer. A Data-In Size output argument should be added to Execute Command

In some protocols (e.g. SRP) where the initiator is not aware of data transfers or doesn't trust the interconnect not to drop frames, the Data-In Size is calculated from the residual carried in a RESPONSE IU. In other protocols (e.g. SAS), the initiator maintains a data count and returns the Data-In Size directly to the application client. In other protocols (e.g. FCP), both mechanisms can be used.

Example: INQUIRY is run with a Data-In Buffer of 256 bytes. The logical unit transfers 56 bytes of data and GOOD status.

SRP: The logical unit reports a Data-In residual of (256 - 56)=200 along with GOOD status. The initiator returns Data-In Size=(request size - residual)=(256 - 56)=200, to the application client with a TASK COMPLETE service response.

SAS: The logical unit reports GOOD status. The initiator counted 56 bytes and returns Data In Size=56 to the application client with a TASK COMPLETE service response.

FCP: The logical unit reports a Data-In residual of (256 - 56)=200 along with GOOD status. The initiator counted 56 bytes. The initiator returns Data In Size=56 calculated either way to the application client. If they don't match, it knows a frame was lost and can report a SERVICE DELIVERY OR TARGET FAILURE service response.

Issue #2. SCSI commands return variable length sense data, especially with the new long sense data format. However, the Execute Command () RPC just returns [Sense Data] without indicating its length. The caller does not provide a sense data buffer length (like it does for a data-in buffer), and the RPC does not return the sense data length delivered. These arguments should be patterned after the Data-In arguments.

Suggested changes to SAM-3

5.1 The Execute Command remote procedure

An application client requests the processing of a SCSI command by invoking the SCSI transport protocol services described in 5.4, the collective operation of which is conceptually modeled in the following remote procedure:

Service response =Execute Command (
IN (I_T_L_x Nexus, CDB, [Task Attribute], [Data-In Buffer Size], [Data-Out Buffer], [Data-Out Buffer Size], [Sense Data Buffer Size], [Autosense Request], [Command Reference Number]),
OUT ([Data-In Buffer], Data-In Size, [Sense Data Buffer], Sense Data Size, Status))

Input Arguments:

I_T_L_x Nexus: Either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).
CDB: Command descriptor block (see 5.2).
Task Attribute: A value specifying one of the task attributes defined in 8.5. This argument shall not be specified for an untagged command or the second and subsequent commands in a sequence of linked commands. Untagged tasks shall implicitly have the SIMPLE attribute. The attribute of a task that processes linked commands shall be set according to the Task Attribute argument specified for the first command in the sequence.
Data-In Buffer Size: The number of bytes available for data transfers to the Data-In Buffer (see 5.4.3).
Data-Out Buffer: A buffer containing command specific information to be sent to the logical unit, such as data or parameter lists needed to service the command. The buffer size is indicated by the Data-Out Buffer Size argument. The content of the Data-Out Buffer shall not change during the lifetime of the command (see 5.5) as viewed by the application client.
Data-Out Buffer Size: The number of bytes available for data transfers from the Data-Out Buffer (see 5.4.3).
Sense Data Buffer Size: The number of bytes available for transfers to the Sense Data Buffer.
Autosense Request: An argument requesting the automatic return of sense data by means of the autosense mechanism specified in 5.9.4.3. It is not an error for the application client to provide this argument when autosense is not supported by the SCSI transport protocol or logical unit. SCSI transport protocols may require that the Autosense Request argument always request automatic return of the sense data.
Command Reference Number (CRN): When this argument is used, all sequential commands of an I_T_L nexus shall include a CRN argument that is incremented by one. The CRN shall be set to one for each I_T_L nexus involving the SCSI port after the SCSI port receives a hard reset or detects I_T nexus loss. The CRN shall be set to one after it reaches the maximum CRN value supported by the protocol. The CRN value zero shall be reserved for use as defined by the SCSI transport protocol. It is not an error for the application client to provide this argument when CRN is not supported by the SCSI transport protocol or logical unit.

Output Arguments:

Data-In Buffer: A buffer to contain command specific information returned by the logical unit by the time of command completion. The buffer size is indicated by the Data-In Buffer Size argument. The application client shall not assume that the buffer contents are valid unless the command completes with a status of GOOD, CONDITION MET, INTERMEDIATE, or INTERMEDIATE-CONDITION MET. While some valid data may be present for other values of status, the application client should obtain additional information from the logical unit, such as sense data, to determine the state of the buffer contents. If the command ends with a service response of SERVICE DELIVERY OR TARGET FAILURE, the application client shall consider this parameter to be undefined.
Data-In Size: The number of bytes delivered to the Data-In Buffer.
Sense Data Buffer: A buffer to contain sense data returned by means of the autosense mechanism (see 5.9.4.3). The buffer size is indicated by the Sense Data Buffer Size argument. If the command ends with a service response of SERVICE DELIVERY OR TARGET FAILURE, the application client shall consider this parameter to be undefined.
Sense Data Size: The number of bytes delivered to the Sense Data Buffer.
Status: A one-byte field containing command completion status (see 5.3). If the command ends with a service response of SERVICE DELIVERY OR TARGET FAILURE, the application client shall consider this parameter to be undefined.

Service Response assumes one of the following values:

TASK COMPLETE: A logical unit response indicating that the task has ended. The status parameter shall have one of the values specified in 5.3 other than INTERMEDIATE or INTERMEDIATE-CONDITION MET.

LINKED COMMAND COMPLETE: Logical unit responses indicating that the task has not ended and that a linked command has completed successfully. As specified in 5.3, the status parameter shall have a value of INTERMEDIATE or INTERMEDIATE-CONDITION MET.

SERVICE DELIVERY OR TARGET FAILURE: The command has been ended due to a service delivery failure (see 3.1.115) or SCSI target device malfunction. All output parameters are invalid.

The actual SCSI transport protocol events corresponding to a response of TASK COMPLETE, LINKED COMMAND COMPLETE or SERVICE DELIVERY OR TARGET FAILURE shall be specified in each SCSI transport protocol standard.

An application client requests processing of a linked command by setting the LINK bit to one in the CDB CONTROL byte as specified in 5.2.3. The task attribute is determined by the Task Attribute argument specified for the first command in the sequence. Upon receiving a response of LINKED COMMAND COMPLETE, an application client may issue the next command in the series through an Execute Command remote procedure call having the same I_T_L_x nexus and omitting the Task Attribute argument. If the logical unit receives the next command in a series of linked commands before completing the current command in that linked command series, the overlapped command condition described in 5.9.2 shall result.

Figure xx shows the Data-Out Buffer, Data-In Buffer, and Sense Data Buffer provided by the application client for the lifetime of the command. The buffers shall not overlap.

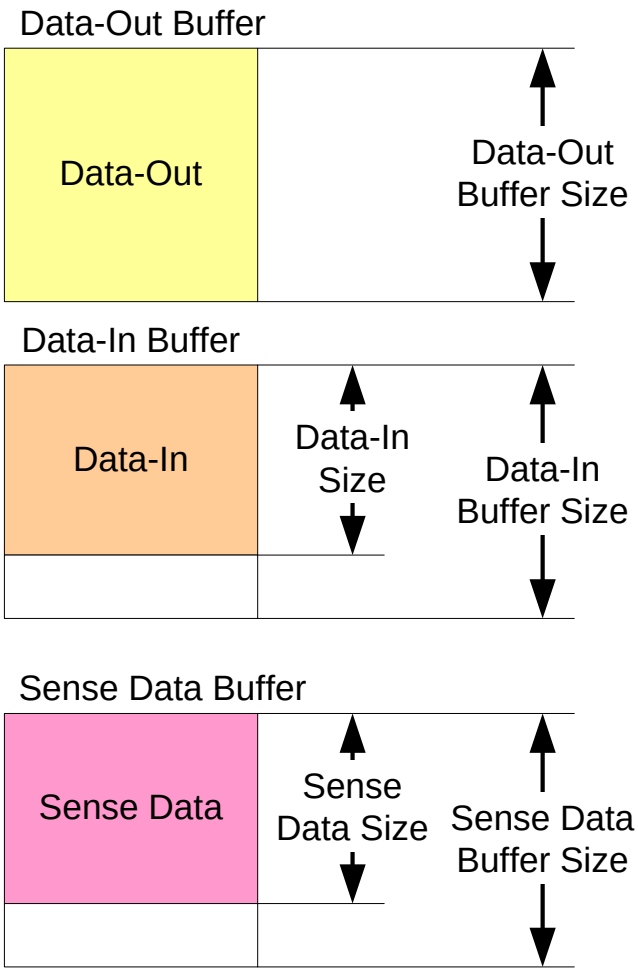

Figure xx. Application client buffers

5.4.1 Overview

The SCSI transport protocol services that support the **Execute Command** remote procedure call are described in 5.4. Two groups of SCSI transport protocol services are described. The SCSI transport protocol services that support the request and confirmation for the **Execute Command** remote procedure call are described in 5.4.2. The SCSI transport protocol services that support the data transfers associated with processing a SCSI command are described in 5.4.3.

5.4.2 Execute Command request/confirmation SCSI transport protocol services

All SCSI transport protocol standards shall define the SCSI transport protocol specific requirements for implementing the **Send SCSI Command** SCSI transport protocol service request and the **Command Complete Received** confirmation. Support for the **SCSI Command Received** indication and **Send Command Complete** response by a SCSI transport protocol standard is optional. All SCSI I/O systems shall implement these SCSI transport protocols as defined in the applicable SCSI transport protocol specification.

SCSI Transport Protocol Service Request:

Send SCSI Command (IN (I_T_L_x Nexus, CDB, [Task Attribute], [Data-In Buffer Size], [Data-Out Buffer], [Data-Out Buffer Size], [Autosense Request], [Command Reference Number]))

Input Arguments:

I_T_L_x Nexus: Either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).

CDB: Command descriptor block (see 5.2).
Task Attribute: A value specifying one of the task attributes defined in 8.5. For specific requirements on the Task Attribute argument see 5.1.
Data-In Buffer Size: The number of bytes available for data transfers to the Data-In Buffer (see 5.4.3).
Data-Out Buffer: A buffer containing command specific information to be sent to the logical unit, such as data or parameter lists needed to service the command (see 5.1). The content of the Data-Out Buffer shall not change during the lifetime of the command (see 5.5) as viewed by the application client.
Data-Out Buffer Size: The number of bytes available for data transfers from the Data-Out Buffer (see 5.4.3).
Autosense Request: An argument (see 5.1) requesting the automatic return of sense data by means of the autosense mechanism specified in 5.9.4.3.
Command Reference Number (CRN): When this argument is used, all sequential commands of an I_T_L nexus shall include a CRN argument that is incremented by one (see 5.1).

SCSI Transport Protocol Service Indication:

SCSI Command Received (IN (I_T_L_x Nexus, CDB, [Task Attribute], [Autosense Request], [Command Reference Number]))

Input Arguments:

I_T_L_x Nexus: Either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).
CDB: Command descriptor block (see 5.2).
Task Attribute: A value specifying one of the task attributes defined in 8.5. For specific requirements on the Task Attribute argument see 5.1.
Autosense Request: This parameter is only present if the **Autosense Request** parameter was specified in the **Send SCSI Command** call and autosense delivery is supported by the SCSI transport protocol and logical unit.
Command Reference Number (CRN): When this argument is used, all sequential commands of an I_T_L nexus shall include a CRN argument that is incremented by one (see 5.1).

SCSI Transport Protocol Service Response (from device server)**:**

Send Command Complete (IN (I_T_L_x Nexus, [Sense Data], Sense Data Size, Status, Service Response))

Input Arguments:

I_T_L_x Nexus: Either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).
Sense Data: If present, this argument instructs the SCSI target port to return sense information to the SCSI initiator port automatically (see 5.9.4.3).
Status: Command completion status (see 5.1).
Service Response: Possible service response information for the command (see 5.1).

SCSI Transport Protocol Service Confirmation:

Command Complete Received (IN (I_T_L_x Nexus, [Data-In Buffer], Data-In Size, [Sense Data Buffer], Sense Data Size, Status, Service Response))

Input Arguments:

I_T_L_x Nexus: Either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).
Data-In Buffer: A buffer containing command specific information returned by the logical unit on command completion (see 5.1).
Data-In Size: The number of bytes consumed in the Data-In Buffer.
Sense Data Buffer: Autosense data (see 5.9.4.3).
Sense Data Size: The number of bytes used in the Sense Data Buffer.
Status: Command completion status (see 5.1).

Service Response: Service response for the command (see 5.1).

...

5.4.3 Data transfer SCSI transport protocol services

5.4.3.1 Introduction

The data transfer services described in 5.4.3 provide mechanisms for moving data to and from the SCSI initiator port in response to commands transmitted using the **Execute Command** remote procedure call. All SCSI transport protocol standards shall define the protocols required to implement these services.

The application client's Data-In Buffer and/or Data-Out Buffer each appear~~s~~ to the device server as a single, logically contiguous block of memory large enough to hold all the data required by the command (see figure 28). The model allows either unidirectional or bidirectional data transfer. The processing of a SCSI command may require the transfer of data from the application client using the Data-Out Buffer, or to the application client using the Data-In Buffer, or both to and from the application client using both the Data-In Buffer and the Data-Out Buffer.

It is assumed that the buffering resources available to the logical unit are limited and may be less than the amount of data that is capable of being transferred in one SCSI command. Such data needs to be moved between the application client and the media in segments that are smaller than the transfer size specified in the SCSI command. The amount of data moved per segment is usually a function of the buffering resources available to the logical unit. Figure 28 shows the model for such incremental data transfers.

The movement of data between the application client and device server is controlled by the following arguments:

Application Client Buffer Size: The total number of bytes in the application client's buffer (Data-In or Data-Out).

Application Client Buffer Offset: Offset in bytes from the beginning of the application client's buffer (Data-In or Data-Out) to the first byte of transferred data.

Byte Count Requested by Device Server: Number of bytes to be moved by the data transfer request.

For any specific data transfer SCSI transport protocol service request, the **Byte Count Requested by Device Server** is less than or equal to the combination of **Application Client Buffer Size** minus the **Application Client Buffer Offset**.

If a SCSI transport protocol supports random buffer access, the offset and byte count specified for each data segment to be transferred may overlap. In this case the total number of bytes moved for a command is not a reliable indicator of highest byte transferred and shall not be used by a SCSI initiator device or SCSI target device implementation to determine whether all data has been transferred.

All SCSI transport protocol standards shall define support for a resolution of one byte for the above arguments. A SCSI initiator device shall support a resolution of one byte. A SCSI target device may support any resolution.

Random buffer access occurs when the device server requests data transfers to or from segments of the application client's buffer that have an arbitrary offset and byte count. Buffer access is sequential when successive transfers access a series of monotonically increasing, adjoining buffer segments. Support for random buffer access by a SCSI transport protocol standard is optional. A device server implementation designed for any SCSI transport protocol implementation should be prepared to use sequential buffer access when necessary.

The LLP confirmed services specified in 5.4.3.2 and 5.4.3.3 are used by the device server to request the transfer of command data to or from the application client. The SCSI initiator device SCSI transport protocol service interactions are unspecified.

The model provides only for the transfer phases to be sequential. Provision for overlapping transfer phases is outside the scope of this standard.

...

5.4.3.3 Data-Out delivery service

Request:

Receive Data-Out (IN (I_T_L_x Nexus, Application Client Buffer Offset, Request Byte Count, Device Server Buffer))

Argument descriptions:

I_T_L_x Nexus: either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).
Device Server Buffer: Buffer from which data is to be transferred.
Application Client Buffer Offset: Offset in bytes from the beginning of the application client's buffer to the first byte of transferred data.
Request Byte Count: Number of bytes to be moved by this request.

Confirmation:

Data-Out Received (IN (I_T_L_x Nexus))

This confirmation notifies the device server that the requested data has been successfully delivered to its buffer.

Argument descriptions:

I_T_L_x Nexus: either an I_T_L nexus or an I_T_L_Q nexus (see 4.11).