To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 20 January 2003
Subject: T10/02-403r2 SAM-3 SPC-3 Maximum and first burst size and protocol services

Revision history
Revision 0 (9 October 2002) first revision
Revision 1 (6 November 2002) incorporated comments from November CAP WG.
Revision 2 (20 January 2003) incorporated comments from January CAP WG.

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

Overview
In developing the Serial Attached SCSI state machines, it became apparent that the Disconnect-Reconnect mode page MAXIMUM BURST SIZE and FIRST BURST SIZE fields are best enforced at the SCSI application layer rather than by the SCSI transport protocol layer.

MAXIMUM BURST SIZE can be described solely in SPC-3 as a restriction on how the device server chooses the byte counts for the Receive Data-Out () or Send Data-In () protocol services.

FIRST BURST SIZE can be described in SAM-3 as including data that will automatically be sent after Send SCSI Command (). It is not received with SCSI Command Received(), however; the data just shows up in a first burst data buffer. When Receive Data-Out() is called, it pulls data from that buffer.

Suggested changes to SPC-3
The MAXIMUM BURST SIZE field indicates the maximum amount of data that the SCSI target port shall transfer during a single data transfer operation. This value is expressed in increments of 512 bytes (e.g., a value of one means 512 bytes, two means 1024 bytes, etc.). The relationship (if any) between data transfer operations and interconnect tenancies is specified in the individual SCSI protocol standards. A value of zero indicates there is no limit on the amount of data transferred per data transfer operation.

In terms of the SCSI protocol services (see SAM-3), the device server shall limit the Request Byte Count argument to the Receive Data-Out protocol service and the Send Data-In protocol service to the amount specified in the MAXIMUM BURST SIZE field.

The FIRST BURST SIZE field indicates the maximum amount of data that may be transferred to the SCSI target port for a command along with the command (i.e., the first burst). This value is expressed in increments of 512 bytes; a value of one means 512 bytes, two means 1024 bytes, etc. A value of zero indicates that there is no first burst size limit. The meaning of a value of zero is transport protocol specific. Transport protocols supporting this field shall provide an additional mechanism to enable and disable the first burst function.

In terms of the SCSI protocol services (see SAM-3), the Receive Data-Out protocol service shall retrieve the first FIRST BURST SIZE amount of data from the first burst.

Suggested changes to SAM-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], [First Burst Enabled]))

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).

      First Burst Enabled: An argument specifying that a transport protocol specific number of bytes from the Data Out Buffer shall be delivered to the logical unit without waiting for the device server to invoke the Receive Data-Out protocol service,

SCSI Transport Protocol Service Indication:

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

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).

      First Burst Enabled: An argument indicating that a transport protocol specific number of bytes from the Data Out Buffer are being delivered to the logical unit without waiting for the device server to invoke the Receive Data-Out protocol service.

...
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 appears 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.

Transport protocols may allow logical units to accept the initial portion of the write data along with the command without waiting for the device server to invoke the Receive Data-Out protocol service. This is modeled using Receive Data-Out protocol service calls for which the transport protocol may have moved the first burst data prior to the call. Logical units with this capability shall implement the FIRST BURST SIZE field in the Disconnect-Reconnect mode page (see SPC-3).

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.

If the SCSI Command Received protocol service included a First Burst Enabled argument and
random buffer access is not supported, the device server shall use the first burst data.

If the SCSI Command Received protocol service included a First Burst Enabled argument and
random buffer access is supported, the device server may re-fetch the data transferred as first
burst data.

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).