

To: T10 Technical Committee
From: Rob Elliott (elliott@hp.com) and Mallikarjun Chadalapaka, HP (mallikarjun.chadalapaka@hp.com)
Date: 11 July 2006
Subject: 06-341r0 SAM-4 Response Fence for protocol services

Revision history

Revision 0 (11 July 2006) First revision

Related documents

sam4r04 - SCSI Architecture Model - 4 (SAM-4) revision 4
04-072r0 SAM-3 SPC-3 Unit attention for status delivery problems (Rob Elliott, HP)
draft-ietf-ips-iscsi-impl-guide-02.txt - iSCSI Implementor's Guide revision 2 (by Mallikarjun Chadalapaka, HP).
Available as <http://www.ietf.org/internet-drafts/draft-ietf-ips-iscsi-impl-guide-02.txt> through September 2006.

Overview

iSCSI supports multiple connections per session (i.e., multiple connections between an initiator port and a target port) and Serial Attached SCSI (SAS) supports multiple connections between a wide initiator port and a wide target port.

A device server requests that a target port transmit a RESPONSE frame (including the SCSI status, sense data, or task management service response) by invoking the Send Command Complete () or Task Management Function Executed () protocol service. SAM-4 does not define whether these calls complete immediately (e.g., before the RESPONSE frame is sent) or after it has been sent. There is no response value for these calls (e.g., there is no "Result = Send Command Complete ()").

Sometimes the device server needs to ensure that the target port has sent certain previously requested RESPONSE frames before it sends a new RESPONSE frame, so the initiator port doesn't receive them out of order. For example, when the device server requests a RESPONSE frame be sent for a CLEAR TASK SET task management function, the task management function's RESPONSE frame must not arrive at the initiator port before some of the RESPONSE frames for the tasks that were supposed to be cleared. Another example is a PERSISTENT RESERVE OUT command with a PREEMPT AND ABORT service action; that command must not complete until all the affected commands are aborted.

The device server could do this by single-threading its calls to Send Command Complete () and Task Management Function Executed (), but that does not reflect the reality that they can and will send them concurrently in most cases.

A Response Fence argument is proposed to be added to the Send Command Complete () and Task Management Function Executed () protocol services. If set, the target port is required to drain all previous responses before sending this one, and must finish sending this one before sending any subsequent responses. If not set, the target port is allowed to send them in any order.

04-072r0 proposed a Delivery Result output argument be added to Send Command Complete (), indicating whether or not the RESPONSE frame was successfully delivered. With that also added to Task Management Function Executed (), the device server could simply wait for the previous protocol services of interest to complete before issuing the one of interest, and not issue subsequent protocol services until the one of interest completes. The CAP WG didn't like that proposal, however (its main goal was to allow the device server to create a unit attention condition and log the event if one of its RESPONSE frames could not be delivered).

Suggested changes

5.4 SCSI transport protocol services in support of Execute Command

5.4.1 Overview

The SCSI transport protocol services that support the Execute Command 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 procedure call are described in 5.4.2. The

SCSI transport protocol services that support the data transfers associated with processing a 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 request, the SCSI Command Received indication, the Send Command Complete response, and the Command Complete Received confirmation SCSI transport protocol services.

All SCSI initiator devices shall implement the Send SCSI Command request and the Command Complete Received confirmation SCSI transport protocol services as defined in the applicable SCSI transport protocol standards. All SCSI target devices shall implement the SCSI Command Received indication and the Send Command Complete response SCSI transport protocol services as defined in the applicable SCSI transport protocol standards.

SCSI Transport Protocol Service Request:

Send SCSI Command (IN (I_T_L_Q Nexus, CDB, Task Attribute, [Data-In Buffer Size], [Data-Out Buffer], [Data-Out Buffer Size], [Command Reference Number], [Task Priority], [First Burst Enabled]))

Input Arguments:

I_T_L_Q Nexus: The I_T_L_Q nexus identifying the task (see 4.12).

CDB: Command descriptor block (see 5.2).

Task Attribute: A value specifying one of the task attributes defined in 8.6. 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). SCSI transport protocols may interpret this argument to include both the size and the location of the Data-In Buffer.

Data-Out Buffer: A buffer containing command specific information to be sent to the logical unit (e.g., data or parameter lists needed to process 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).

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).

Task Priority: The priority assigned to the task (see 8.7).

First Burst Enabled: An argument specifying that a SCSI 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 SCSI transport protocol service.

SCSI Transport Protocol Service Indication:

SCSI Command Received (IN (I_T_L_Q Nexus, CDB, Task Attribute, [Command Reference Number], [Task Priority], [First Burst Enabled]))

Input Arguments:

I_T_L_Q Nexus: The I_T_L_Q nexus identifying the task (see 4.12).

CDB: Command descriptor block (see 5.2).

Task Attribute: A value specifying one of the task attributes defined in 8.6. For specific requirements on the Task Attribute argument see 5.1.

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).

Task Priority: The priority assigned to the task (see 8.7).

First Burst Enabled: An argument specifying that a SCSI 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 SCSI transport protocol service.

SCSI Transport Protocol Service Response (from device server):

Send Command Complete (IN (I_T_L_Q Nexus, [Sense Data], [Sense Data Length], Status, Service Response, [Retry Delay Timer], [\[Response Fence\]](#)))

Input Arguments:

I_T_L_Q Nexus: The I_T_L_Q nexus identifying the task (see 4.12).

Sense Data: If present, this argument instructs the SCSI target port to return sense data to the SCSI initiator port (see 5.8.6).

Sense Data Length: The length in bytes of the sense data to be returned to the SCSI initiator port.

Status: Command completion status (see 5.1).

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

Retry Delay Timer: The Retry Delay Timer code for the command (see 5.3.2).

[Response Fence:](#) The target port shall:

a) complete sending all previous Command Complete responses and Task Management Function Executed responses (see 7.10) before sending this response; and

b) complete sending this response before sending any subsequent Command Complete responses and Task Management Function Executed responses.

SCSI Transport Protocol Service Confirmation:

Command Complete Received (IN (I_T_L_Q Nexus, [Data-In Buffer], [Sense Data], [Sense Data Length], Status, Service Response, [Retry Delay Timer]))

Input Arguments:

I_T_L_Q Nexus: The I_T_L_Q nexus identifying the task (see 4.12).

Data-In Buffer: A buffer containing command specific information returned by the logical unit on command completion (see 5.1).

Sense Data: Sense data returned in the same I_T_L_Q nexus transaction (see 3.1.47) as a CHECK CONDITION status (see 5.8.6).

Sense Data Length: The length in bytes of the received sense data.

Status: Command completion status (see 5.1).

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

Retry Delay Timer: The Retry Delay Timer code for the command (see 5.3.2).

7.10 Task management SCSI transport protocol services

The SCSI transport protocol services described in this subclause are used by a SCSI initiator device and SCSI target device to process a task management procedure call. The following arguments are passed:

Nexus: An I_T Nexus, I_T_L Nexus, or I_T_L_Q Nexus (see 4.12).

Function Identifier: Argument encoding the task management function to be performed.

All SCSI transport protocol standards shall define the SCSI transport protocol specific requirements for implementing the **Send Task Management Request**, the **Task Management Request Received** indication, the **Task Management Function Executed** response, and the **Received Task Management Function Executed** confirmation SCSI transport protocol services described in this subclause.

A SCSI transport protocol standard may specify different implementation requirements for the **Send Task Management Request** SCSI transport protocol service for different values of the Function Identifier argument.

All SCSI initiator devices shall implement the **Send Task Management Request** and the **Received Task Management Function Executed** confirmation SCSI transport protocol services as defined in the applicable SCSI transport protocol standards.

All SCSI target devices shall implement the **Task Management Request Received** indication and the **Task Management Function Executed** response SCSI transport protocol services as defined in the applicable SCSI transport protocol standards.

Request sent by an application client:

Send Task Management Request (IN (Nexus, Function Identifier))

Argument descriptions:

Nexus: An I_T Nexus, I_T_L Nexus, or I_T_L_Q Nexus (see 4.12).

Function Identifier: Argument encoding the task management function to be performed.

Indication received by the task manager:

Task Management Request Received (IN (Nexus, Function Identifier))

Argument descriptions:

Nexus: An I_T Nexus, I_T_L Nexus, or I_T_L_Q Nexus (see 4.12).

Function Identifier: Argument encoding the task management function to be performed.

Response from task manager:

Task Management Function Executed (IN (Nexus, Service Response, [\[Response Fence\]](#)))

Argument descriptions:

Nexus: An I_T Nexus, I_T_L Nexus, or I_T_L_Q Nexus (see 4.12).

Service Response: An encoded value representing one of the following:

FUNCTION COMPLETE: The requested function has been completed.

FUNCTION SUCCEEDED: The requested function is supported and completed successfully.

FUNCTION REJECTED: The task manager does not implement the requested function.

INCORRECT LOGICAL UNIT NUMBER: An optional task router response indicating that the function requested processing for an incorrect logical unit number.

SERVICE DELIVERY OR TARGET FAILURE: The request was terminated due to a service delivery failure (see 3.1.112) or SCSI target device malfunction. The task manager may or may not have successfully performed the specified function.

Response Fence: The target port shall:

a) complete sending all previous Command Complete responses (see 5.4) and Task Management Function Executed responses before sending this response; and

b) complete sending this response before sending any subsequent Command Complete responses and Task Management Function Executed responses.

Confirmation received by application client:

Received Task Management Function Executed (IN (Nexus, Service Response))

Argument descriptions:

Nexus: An I_T Nexus, I_T_L Nexus, or I_T_L_Q Nexus (see 4.12).

Service Response: An encoded value representing one of the following:

FUNCTION COMPLETE: The requested function has been completed.

FUNCTION SUCCEEDED: The requested function is supported and completed successfully.

FUNCTION REJECTED: The task manager does not implement the requested function.

INCORRECT LOGICAL UNIT NUMBER: An optional task router response indicating that the function requested processing for an incorrect logical unit number.

SERVICE DELIVERY OR TARGET FAILURE: The request was terminated due to a service delivery failure (see 3.1.112) or SCSI target device malfunction. The task manager may or may not have successfully performed the specified function.

Each SCSI transport protocol shall allow a **Received Task Management Function Executed** confirming completion of the requested task to be associated with the corresponding **Send Task Management Request**.