

Date: September 29, 2005

To: T10 Committee (SCSI)

From: George Penokie (IBM/Tivoli)

Subject: SAS-2, SAM-4, SPC-4, NOTIFY (POWER FAILURE EXPECTED) Fixes

Overview

Some holes were found in the 05-167r5 that need to be fixed. None of the changes constitute a technical change from the intent of 05-167r5.

SAS-2 changes

The changes marked in this proposal for SAS-2 represent changes from the combination of 05-176r5 and 05-307r0.

0.0.0.1 NOTIFY

NOTIFY may be transmitted in place of any ALIGN (see 7.2.5.2) being transmitted for clock skew management (see 7.3), rate matching (see 7.13), or STP initiator phy throttling (see 7.17.2). Substitution of a NOTIFY may or may not affect the ALIGN rotation (i.e., the NOTIFY may take the place of one of the ALIGNs in the rotation through ALIGN (0), ALIGN (1), ALIGN (2), or ALIGN (3) or it may delay the rotation). A specific NOTIFY shall not be transmitted in more than three consecutive dwords before at least three other dwords have been transmitted.

NOTIFY shall not be forwarded through expander devices. Expander devices shall substitute an ALIGN for a NOTIFY if necessary.

SAS target devices are not required to detect every transmitted NOTIFY.

The versions of NOTIFY representing different reasons are defined in table 1.

Table 1 — NOTIFY primitives

Primitive	Description
NOTIFY (ENABLE SPINUP)	Specify to an SAS target device that it may temporarily consume additional power while transitioning into the active or idle power condition state.
NOTIFY (POWER FAILURE EXPECTED)	Specify to an SAS target device that power may be lost within a time specified by the power fail timeout timer (see 0.0.2.1.1).
NOTIFY (RESERVED 1)	Reserved.
NOTIFY (RESERVED 2)	Reserved.

NOTIFY (ENABLE SPINUP) is transmitted by a SAS initiator port or expander port and is used to specify to an SAS target device that it may temporarily consume additional power (e.g., while spinning-up rotating media) while transitioning into the active or idle power condition state. The length of time the SAS target device consumes additional power and the amount of additional power is vendor specific. NOTIFY (ENABLE SPINUP) shall interact with the device's power condition state transitions, controlled by the Power Conditions mode page (see SPC-3) and/or the START STOP UNIT command (see SBC-2), as described in 10.2.10.

SAS initiator devices and expander devices shall use NOTIFY (ENABLE SPINUP) while attached to SAS target devices (i.e., devices that report SSP target support in their IDENTIFY address frames). They shall transmit one NOTIFY (ENABLE SPINUP) after power on when the enclosure is ready for initial spin-up. After the initial NOTIFY (ENABLE SPINUP), they shall transmit NOTIFY (ENABLE SPINUP) periodically. Otherwise, the selection of when and how often to transmit NOTIFY (ENABLE SPINUP) is outside the scope of this standard.

NOTE 1 - The SAS initiator device or expander device uses NOTIFY (ENABLE SPINUP) to avoid exceeding enclosure power supply capabilities during spin-up of multiple SAS target devices. It may choose to rotate transmitting NOTIFY (ENABLE SPINUP) across all of its ports, distributing it to N ports at a time if the enclosure power supply is capable of powering N SAS target devices spinning up at a time. An expander device may allow this timing to be configured by a NVROM programming with enclosure-specific sequencing patterns, or may employ more complex, dynamic interaction with the enclosure power supply.

NOTE 2 - NOTIFY (ENABLE SPINUP) should be transmitted as frequently as possible to avoid incurring application layer timeouts.

I_T nexus loss, logical unit reset, and hard reset shall not cause a SAS target device to spin-up automatically on receipt of NOTIFY (ENABLE SPINUP).

SAS target devices with multiple SAS target ports shall honor NOTIFY (ENABLE SPINUP) from all SAS target ports equivalently (e.g., NOTIFY (ENABLE SPINUP) received on SAS target port A serves as a wakeup for a START STOP UNIT command received through SAS target port B).

NOTIFY (POWER FAILURE EXPECTED) is transmitted by a SAS initiator port or expander port and is used to specify to a SAS target device that a power failure may occur within a time defined in the power failure timeout field of the Protocol-Specific Logical Unit mode page (see 0.0.2.1.1). The NOTIFY (POWER FAILURE EXPECTED) shall be transmitted at least three times by the SAS initiator port or expander port.

If a target device supports NOTIFY (POWER FAILURE EXPECTED) and a NOTIFY (POWER FAILURE EXPECTED) is received, then the device server for each logical unit accessed by the SAS target port that received the NOTIFY (POWER FAILURE EXPECTED) shall:

- 1) stop writing data to the media on a block boundary (e.g., all write activity shall continue until a block boundary is reached then all writing shall stop); ~~and~~
- 2) clear ~~the all~~ task sets (i.e., the device server acts as if it has received a CLEAR TASK SET task management function (see SAM-4) for each task set); and
- 3) establish a unit attention condition for the initiator port associated with every I_T nexus with the additional sense code set to ~~POWER FAIL EVENT COMMANDS CLEARED~~ COMMANDS CLEARED BY POWER FAILURE NOTIFICATION.

If a target device supports NOTIFY (POWER FAILURE EXPECTED) and a NOTIFY (POWER FAILURE EXPECTED) is received, then within 1 ms after the NOTIFY (POWER FAILURE EXPECTED) is received, each SAS phy within the target device shall:

- a) if there is a connection, then transmit a BREAK on that connection; and
- b) transmit an OPEN_REJECT (RETRY)s in response to all OPEN address frames received until the power fail timeout timer expires (see 0.0.2.1.1) or power is lost.

NOTE 3 - A SAS target device may transmit a BREAK on an open connection and begin transmitting OPEN_REJECT (RETRY)s as soon as a NOTIFY (POWER FAILURE EXPECTED) is received.

If any frames are received by a SAS target device after a NOTIFY (POWER FAILURE EXPECTED) is received and before a connection is closed (i.e., by transmitting a BREAK in response to the NOTIFY (POWER FAILURE EXPECTED)), then the SAS target device shall discard the received frames.

~~If the power fail timeout timer expires, then each device server shall establish a unit attention condition for the initiator port associated with every I_T nexus with the additional sense code set to ~~POWER FAIL EVENT COMMANDS CLEARED~~ COMMANDS CLEARED BY POWER FAILURE NOTIFICATION. After power on the power fail timeout timer shall be initialized and stopped until a NOTIFY (POWER FAILURE EXPECTED) is received.~~

NOTIFY (RESERVED 1), and NOTIFY (RESERVED 2) shall be ignored by all devices.

SPC-4 changes

Add in a new ASC/ASCQ of 2Fh 01h COMMANDS CLEARED BY POWER FAILURE NOTIFICATION.

SAM-4 changes

0.1 Task and command lifetimes

This subclause specifies the events delimiting the beginning and end (i.e., lifetime) of a task or pending command from the viewpoint of the device server and application client.

The device server shall create a task upon receiving a **SCSI Command Received** indication unless the command represents a continuation of a linked command as described in 5.1.

The task shall exist until:

- a) The device server sends a SCSI transport protocol service response for the task of TASK COMPLETE; or
- b) The task is aborted as described in 0.2.

The application client maintains an application client task to interact with the task from the time the **Send SCSI Command** SCSI transport protocol service request is invoked until it receives one of the following SCSI target device responses:

- a) A service response of TASK COMPLETE for that task;
- b) Notification of a unit attention condition with one of the following additional sense codes;
 - A) ~~COMMANDS CLEARED BY ANOTHER INITIATOR, if in reference to the task set containing the task~~ Any additional sense code whose ADDITIONAL SENSE CODE field contains 2Fh (e.g., CLEARED BY ANOTHER INITIATOR, or POWER FAIL EVENT COMMANDS CLEARED), if in reference to the task set containing the task; or
 - B) Any additional sense code whose ADDITIONAL SENSE CODE field contains 29h (e.g., POWER ON, RESET, OR BUS DEVICE RESET OCCURRED; POWER ON OCCURRED; SCSI BUS RESET OCCURRED; BUS DEVICE RESET FUNCTION OCCURRED; DEVICE INTERNAL RESET; or I_T NEXUS LOSS OCCURRED);
- c) Notification that the task manager has detected the use of a duplicate I_T_L_Q nexus (see 5.8.3);
- d) A service response of FUNCTION COMPLETE following an ABORT TASK task management function directed to the specified task;
- e) A service response of FUNCTION COMPLETE following an ABORT TASK SET or a CLEAR TASK SET task management function directed to the task set containing that task;
- f) A service response of FUNCTION COMPLETE following an I_T NEXUS RESET task management function delivered on the I_T nexus used to deliver that task; or
- g) A service response of FUNCTION COMPLETE in response to a LOGICAL UNIT RESET task management function directed to the logical unit.

0.2 Aborting tasks

0.2.1 Mechanisms that cause tasks to be aborted

A task is aborted when an event or SCSI initiator device action causes termination of the task prior to its successful completion.

The following events cause a task or several tasks to be aborted:

- a) The return of an **Execute Command** service response of SERVICE DELIVERY OR TARGET FAILURE as described in 5.1;
- b) An I_T nexus loss (see 6.3.4);
- c) A logical unit reset (see 6.3.3);
- d) A hard reset (see 6.3.2); ~~or~~
- e) A power on condition (see 6.3.1); or
- f) SCSI transport protocol specific conditions.

An action transmitted via one I_T nexus may abort task(s) received on that I_T nexus and/or task(s) received on other I_T nexuses.

The following actions affect only the task(s) received on the I_T nexus on which the action is transmitted:

- a) Completion of an ABORT TASK task management function directed to the specified task;
- b) Completion of an ABORT TASK SET task management function under the conditions specified in 7.3;
- c) Completion of an I_T NEXUS RESET task management function; or
- d) Completion of a command with a CHECK CONDITION status, without establishing an ACA condition (see 5.8.1.3) or establishing an ACA condition (see 5.8.2.2), while the Control mode page (see SPC-3) contains fields that are set as follows:
 - A) The QERR field set to 01b and the TST field set to 001b; or
 - B) The QERR field set to 11b.

The actions shown in table 2 affect the task(s) received on the I_T nexus on which the action is transmitted and/or task(s) received on other I_T nexuses.

Table 2 — Actions that affect task(s) received on this or other I_T nexuses

Action	Unit attention additional sense code, if any (see 5.6.3)
Completion of a CLEAR TASK SET task management function referencing the task set containing the specified task	COMMANDS CLEARED BY ANOTHER INITIATOR
Completion of a command with a CHECK CONDITION status, with or without establishing an ACA condition, and the QERR field was set to 01b and the TST field was set to 000b in the Control mode page (see SPC-3)	COMMANDS CLEARED BY ANOTHER INITIATOR
Completion of a PERSISTENT RESERVE OUT command with a PREEMPT AND ABORT service action with a reservation key that is associated with the I_T nexus on which the task was received (see SPC-3)	COMMANDS CLEARED BY ANOTHER INITIATOR
Completion of a LOGICAL UNIT RESET task management function (see 7.7) directed to the logical unit	BUS DEVICE RESET FUNCTION OCCURRED
SCSI transport protocol specific conditions	As defined by the applicable SCSI transport protocol standard

If one or more tasks are cleared or aborted, the affected tasks are also cleared from the initiator ports in a manner that is outside the scope of this standard.