To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 2 May 2008
Subject: 08-126r1 SPC-4 SBC-3 SAS-2.1 Power condition enhancements

Revision history

Revision 0 (6 March 2008) First revision
Revision 1 (2 May 2008) Incorporated feedback from March CAP WG.

Related documents

sbc3r13 - SCSI Block Commands - 3 (SBC-3) revision 13
sas2r14 - Serial Attached SCSI - 2 (SAS-2) revision 14
spc4r12 - SCSI Primary Commands - 4 (SPC-4) revision 12
mmc6r01 - Multimedia Commands - 6 (MMC-6) revision 1
rbc-r10a - Reduced Block Commands (RBC) revision 10a
T13/ata8-acs-r4c - ATA Command Set (ATA8-ACS) revision 4c
01-134r2 - SAM-2 SPC-3 SPI-4 SBC-2 WAKEUP and reset cleanup (Rob Elliott, Compaq)
07-485r2 - SAT-2 Additional power management methods (Frederick Knight, NetApp)

Overview

With renewed interest in power management, and plans to add interface power management in SAS-2.1, the device level power management features (e.g., active/idle/standby power conditions) in SCSI should be enhanced.

After discussion in the March 2008 CAP WG, a new model, new commands, and new VPD page are proposed. This seems better than enhancing the current SPC-4 power conditions model and Power Conditions mode page, and the START STOP UNIT command available for some but not all peripheral device types.

1. MANAGE POWER OUT command
    a) Enter a power management state.
    b) Set up timer to transition between states automatically.
2. MANAGE POWER IN command
    a) Report the current power management state.
3. Power Management VPD page
    a) Return a list of power management states and their attributes:
        A) State numeric value (one byte). 00h means “active”; all others are vendor-specific.
        B) ASCII string describing the state
        C) Commands/TMFs processing behavior during the state
        D) Power consumption during the state
        E) Recovery method
        F) Recovery power consumption to transition to state 00h
        G) Recovery time to transition to state 00h

Sleep power condition background

Some SCSI standards currently define or previously defined a **sleep power condition** to send a target into an “off” mode, ready for power to be removed. This mode:

a) consumes the least amount of power possible, short of actually removing power from the device.
    A) all I/O context can be lost (unlike Standby, where commands resume)
    B) write cache DRAM can be turned off (no need to preserve contents)
    C) embedded processor can go to sleep mode (no need to restore context later)
    D) target port interface prepared to be turned off
b) requires a hard reset to wake up
    A) once software selects this mode, there is no backing out
    B) target port interface might remain on for a while (waiting on other logical units)

C) in a serial interface, target port interface must accept some sort of wake up signal (e.g. COMINIT in SAS) so a hard reset can be detected

The SCSI parallel interface (SPI-5) was the last T10 transport protocol to support the sleep power condition. It defined a WAKEUP task management function to wakeup from Sleep, defined as as a “bus reset condition,” not a special signal. Among non-T10 SCSI protocols, ATAPI supports it.

Three SCSI command sets have defined the sleep power condition. Entry is via the START STOP UNIT command (which is defined in individual command sets, not in SPC), not the Power Condition mode page (in SPC):

a) MMC-n (Multi-Media Commands);
b) RBC (Reduced Block Commands); and
c) SBC-n (SCSI Block Commands).

NOTE 1 - SBC-1 defined the POWER CONDITION field set to 5h as “Place device into Sleep” condition. SBC-2 revisions 0, 1, and 2 mistakenly relocated that to 6h. Revision 3 restored it to 5h. 02-464r3 changed that value to “obsolete” in SBC-2 revision 9.

ATA8-ACS also defines a sleep power condition, and a SCSI equivalent is needed so SCSI to ATA translation layers (SATLs) can map all the capabilities of an ATA device into SCSI.

The new model needs to support the sleep mode.

New power conditions

Some idle mode enhancements have also been deployed in the industry:

a) low-rpm idle. The disk keeps spinning at a reduced rate, providing a shorter recovery time than the full spin-down required by the standby power condition. Not as much power savings as standby.
b) park/retract heads. Reduce friction on the spinning disk during the idle power condition. Very fast recovery time. Not as much power savings as low-rpm idle.

The new model needs to support these, and still allow flexibility for additional modes to be deployed.

Editor’s Note 1: MMC-6 changes will be needed as well; not covered yet

Suggested changes to SPC-4

5.99 Power management model [all new]

5.99.1 Power management model overview

Editor’s Note 2: Need to write the model section

One suggestion in March: use a well-known logical unit to control the SCSI target device state, and don’t let individual logical units have their own states. This simplifies the RAID controller implementation, but makes other implementations impossible (like a tape library where individual logical units represent unique tape drives). Also, software often only knows about logical units, and loses the SCSI target device relationship between them.

The SCSI target device shall make inactive each SCSI target port through which the logical unit is accessible if all other logical units in the SCSI target device that are accessible through that SCSI target port are also in this state. It shall not make a SCSI target port inactive while any logical unit that is accessible through that SCSI target port is not in this state; and

5.99.2 Timers

Editor's Note 3: describe idle timers here

5.19 MANAGE POWER OUT command [all new]

The MANAGE POWER OUT command (see table 2) requests that the device server change the power management state of the logical unit or configure the timers for automatic entry into one or more power management states (see 4.16).

Logical units that contain cache shall write all cached logical blocks to the medium (e.g., as they would do in response to a SYNCHRONIZE CACHE command (see 5.20 and 5.21) with the SYNC_NV bit set to zero, the LOGICAL BLOCK ADDRESS field set to zero, and the NUMBER OF LOGICAL BLOCKS field set to zero) prior to entering into any power condition that prevents accessing the medium (e.g., before the rotating media spindle motor is stopped during transition to the stopped power condition).

Editor's Note 4: Interaction with deferred download microcode needs to be considered.

Table 2 — MANAGE POWER OUT command

Byte\Bit	7	6	5	4	3	2	1	0
0	OPERATION CODE (A4h)							
1	Reserved				SERVICE ACTION (nnh)			
2	Reserved							
5								
6	(MSB)	PARAMETER LIST LENGTH						
9								(LSB)
10	Reserved							
11	CONTROL							

The OPERATION CODE field and SERVICE ACTION field are defined in 4.3.2 and shall be set to the values defined in table 2.

The PARAMETER LIST LENGTH field is defined in 4.3.5.5.

If the parameter list length results in the truncation of the header or any timer configuration descriptor list entry, then the device server shall make no changes to the timers and shall terminate the command with CHECK CONDITION status, with the sense key set to ILLEGAL REQUEST, and the additional sense code set to PARAMETER LIST LENGTH ERROR.

Table 3 defines the power management parameter list.

Table 3 — Power management parameter list

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)	PAGE LENGTH (n - 2)						
1								(LSB)
2	Reserved							
3								
4	REQUESTED PM STATE							
5	Reserved							
6								
7	NUMBER OF TIMER CONFIGURATION DESCRIPTORS							
Timer configuration descriptor list								
8	Timer configuration descriptor (first)(see table 4)							
11								
...	...							
n - 3	Timer configuration descriptor (last)(see table 4)							
n								

The PAGE LENGTH field specifies the number of bytes in the power management parameter list that follow and is set to the value defined in table 3.

The REQUESTED PM STATE field specifies the power management state into which the logical unit shall transition, and shall be set to one of the values reported in the Power Management VPD page (see table 8 in 6.x). A REQUESTED PM STATE field set to 00h specifies that no change be made to the current power management state. If the task set contains any commands that are not able to be processed in the specified power management state, the device server shall terminate the command with CHECK CONDITION status with a sense key set to ABORTED COMMAND and an additional sense code set to INVALID FIELD IN PARAMETER LIST and shall not change any of the timers.

Editor's Note 5: Could set a marker that lets the task set drain, blocking any new commands from entering the task set while waiting, then enter the requested state. Could add a "force" bit to enter even if there are commands in the task set (possibly losing them if entry is to a state like sleep; probably causing timeouts if entry is to a lighter weight state)

The NUMBER OF TIMER CONFIGURATION DESCRIPTORS field specifies the number of timer configuration descriptors in the timer configuration descriptor list.

The timer configuration descriptor list contains timer configuration descriptors sorted in ascending order by the STATE IDENTIFIER field in the descriptor.

If any timer configuration descriptor specifies a timer for a state that does not have its IDLE TIMER SUPPORTED bit set to one in the Power Management VPD page (see 6.x), the device server shall terminate the command with CHECK CONDITION status with a sense key set to ABORTED COMMAND and an additional sense code set to INVALID FIELD IN PARAMETER LIST and shall not change any of the timers.

Table 4 defines the timer configuration descriptor.

Table 4 — Timer configuration descriptor

Byte\Bit	7	6	5	4	3	2	1	0
0	STATE IDENTIFIER							
1	Reserved							
2	(MSB)	IDLE TIME						
3								(LSB)

The STATE IDENTIFIER field specifies the power management state of the logical unit for which the timer is being configured, and is set to one of the values reported in the Power Management VPD page (see table 8 in 6.x).

The IDLE TIME field specifies the amount of idle time in 100 ms increments that the logical unit shall detect before entering the power management state. The IDLE TIME field set to 0000h specifies that the timer for the power management state shall be disabled.

Editor's Note 6: FFFFh = 65,535. In 100 ms increments, the max is 655.35 seconds.

Editor's Note 7: Define "idle time". Other "idle time" detection algorithms may be desired. For example, only consider media access commands. Step down and remain in a slower operational mode (e.g. active to silent) if the workload is light (not using one specific idle time, just detect a longer average idle time); don't go back to fully active just because a command is received.

5.19 MANAGE POWER IN command [all new]

The MANAGE POWER IN command (see table 5) reports the power management state of the logical unit (see 4.16)..

Table 5 — MANAGE POWER IN command

Byte\Bit	7	6	5	4	3	2	1	0
0	OPERATION CODE (A3h)							
1	Reserved				SERVICE ACTION (nnh)			
2		Reserved						
5								
6	(MSB)	ALLOCATION LENGTH						
9								(LSB)
10	Reserved							
11	CONTROL							

The OPERATION CODE field and SERVICE ACTION field are defined in 4.3.2 and shall be set to the values defined in table 5.

The ALLOCATION LENGTH field is defined in 4.3.5.6.

Table 6 defines the power management parameter data.

Table 6 — Power management parameter data

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)	PAGE LENGTH (06h)						
1								(LSB)
2	Reserved							
3								
4	CURRENT PM STATE							
5	REASON							
6	Reserved							
7								

The PAGE LENGTH field indicates the number of bytes in the power management data that follow and is set to the value defined in table 5.

The CURRENT PM STATE field indicates the current power management state of the logical unit, and is set to one of the values reported in the Power Management VPD page (see table 8 in 6.x).

The REASON field indicates the method of entry into the power management state and is defined in table 7.

Table 7 — REASON field

Code	Description
00h	Power on
01h	MANAGE POWER OUT command
02h	Timer
All others	Reserved

7.7.4 Power Management VPD page [all new]

The Power Management VPD page (see table 8) provides the application client with a means to obtain information about the power management capabilities of the logical unit.

Table 8 — Power Management VPD page

Byte\Bit	7	6	5	4	3	2	1	0
0	PERIPHERAL QUALIFIER			PERIPHERAL DEVICE TYPE				
1	PAGE CODE (nnh)							
2	(MSB) PAGE LENGTH (nnh)							
3								
4	Reserved							
6								
7	NUMBER OF POWER MANAGEMENT STATES							
Power management state descriptor list								
8	Power management state descriptor (first)(see table 10)							
39								
...	...							
n - 31	Power management state descriptor (last)(see table 10)							
n								

The NUMBER OF POWER MANAGEMENT STATES field indicates the number of power management state descriptors in the power management state descriptor list.

Editor's Note 8: Report the power-up mode(s). ATA devices can power up in standby mode; should that be allowed here, so SAT mapping is legal?

The power management state descriptor list contains power management state descriptors sorted in ascending order by the STATE IDENTIFIER field in the descriptor. The power management state descriptor list shall start with the descriptor for the active power management state defined in table 9.

Table 9 — Mandatory power management state descriptor

STATE IDENTIFIER field	STATE NAME field	ABILITY field	RECOVERY METHOD field	POWER CONSUMPTION field	RECOVERY POWER CONSUMPTION field	RECOVERY TIME field
01h	"active"	00h	00h	vendor specific	vendor specific	0 s

Editor's Note 9: That belongs in the model section

Table 10 defines the power management state descriptor.

Table 10 — Power management state descriptor

Byte\Bit	7	6	5	4	3	2	1	0
0	STATE IDENTIFIER							
1	ABILITY							
2	RECOVERY METHOD							
3	Reserved						POWER ON STATE	IDLE TIMER SUPPORTED
4	Reserved							
7								
8	(MSB)	BUDGETARY POWER CONSUMPTION						
9								(LSB)
10	(MSB)	BUDGETARY RECOVERY POWER CONSUMPTION						
11								(LSB)
12	(MSB)	RECOVERY TIME						
13								(LSB)
14	Reserved							
15								
16	STATE NAME							
31								

The STATE IDENTIFIER field (see table 322) indicates the binary identifier for the power management state described by the power management state descriptor. A STATE IDENTIFIER field set to 00h is reserved. A STATE IDENTIFIER field set to 01h indicates the active power management state.

The ABILITY field (see table 11) indicates the commands and task management functions that are allowed during the power management state.

Table 11 — ABILITY field

Code	Description
00h	All task management functions are allowed. All commands are allowed.
01h	All task management functions are allowed. Non-media access commands are allowed. Media access commands are allowed, but may incur an extra delay. The logical unit remains in the same power management state.
02h	All task management functions are allowed. Non-media access commands are allowed. Media access commands are terminated with CHECK CONDITION status with an additional sense code set to POWER MANAGEMENT STATE TRANSITION UNDERWAY. The logical unit transitions to a different power management state capable of processing the command.
03h	All task management functions are allowed. Non-media access commands are allowed. Media access commands are terminated with CHECK CONDITION status with an additional sense code set to INCORRECT POWER MANAGEMENT STATE. The logical unit remains in the same power management state.
FFh	No task management functions are allowed. No commands are allowed. Commands are terminated with CHECK CONDITION status with an additional sense code set to POWER MANAGEMENT STATE DOES NOT ACCEPT COMMANDS. Task management functions are terminated with a function response of FUNCTION REJECTED or SERVICE DELIVERY OR TARGET FAILURE.
All others	Reserved

Editor’s Note 10: The target device may contain multiple target ports and multiple logical units. A target port is only turned off when all the logical units are in the sleep power condition, meaning that some could continue to receive commands while waiting for their siblings to go to sleep; thus the error reporting in b) (just letting the commands and TMFs timeout would be another viable approach). If there are multiple target ports, they are sent to sleep at the same time (when all the logical unit reach the sleep power condition).

The RECOVERY METHOD field (see table 12) indicates the method required for exiting the power management state.

Table 12 — RECOVERY METHOD field

Code	Description
00h	Any commands
01h	Media access commands
02h	MANAGE POWER OUT command requesting a different power management state
03h	Hard reset. The SCSI target port may be unavailable during this state for anything except a hard reset event.
All others	Reserved

Editor's Note 11: automatic link reset after the drive goes down after entering SLEEP cannot cause a wakeup, or else we defeat the purpose. Requiring a hard reset solves that, but requires software to send a hard reset. If the initiator/expander loses power but the drive doesn't, the normal power up sequence will find the drive unresponsive.

A POWER ON STATE bit set to one indicates that the logical unit enters this power management state after power on. An IDLE TIMER SUPPORTED bit set to zero indicates that the logical unit does not enter this power management state after power on.

An IDLE TIMER SUPPORTED bit set to one indicates that the logical unit supports an idle timer for the power management state. An IDLE TIMER SUPPORTED bit set to zero indicates that the logical unit does not support an idle timer for the power management state.

The BUDGETARY POWER CONSUMPTION field indicates the budgetary maximum power consumption in 100 mW increments that the logical unit should be expected to consume while in the power management state.

Editor's Note 12: FFFFh = 65,535. In 100 mW increments, the max is 655.35 W.

The BUDGETARY RECOVERY POWER CONSUMPTION field indicates the budgetary maximum power consumption in 100 mW increments that the logical unit should be expected to consume while transitioning from the power management state to power management state 01h (i.e., active) (e.g., because spin-up of rotating media consumes extra power). If the logical unit consumes no additional power than the active state, then the value is the same as the power consumption in the active state.

The RECOVERY TIME field indicates the maximum time in 100 ms increments that the logical unit takes to return from the power management state to power management state 01h (i.e., active). Transport protocol-specific spinup control signals (e.g., NOTIFY (ENABLE SPINUP) in SAS) may increase the recovery time beyond this maximum.

Editor's Note 13: FFFFh = 65,535. In 100 ms increments, the max is 655.35 seconds.

The STATE NAME field contains a left-aligned ASCII string describing the power management state. A STATE NAME field set to “active” indicates the power management state 00h (i.e., active).

Table 13 lists an example set of power management states.

Table 13 — Power management states example

STATE IDENTIFIER field	STATE NAME field	ABILITY field	RECOVERY METHOD field	BUDGETARY POWER CONSUMPTION field	BUDGETARY RECOVERY POWER CONSUMPTION field	RECOVERY TIME field
01h	“active”	00h	00h	13.6 W	13.6 W	0 s
02h	“silent”	00h	00h	11.5 W	13.6 W	0 s
10h	“idle”	01h	00h	9.6 W	13.6 W	0 s
11h	“idle unload”	01h	01h	7.5 W	27 W	1 s
12h	“idle low rpm”	01h	01h	5.1 W	27 W	7 s
80h	“standby”	02h	01h	2 W	27 W	15 s
90h	“stopped”	03h	02h	2.1 W	27 W	20 s
FFh	“sleep”	FFh	03h	1.7 W	30 W	31 s

Editor’s Note 14: This example is based on the Hitachi Deskstar 7K500 Hard Disk Drive specification revision 1.3 available on http://www.hitachigst.com, and is just used to illustrate how the power states could be implemented. Actual numbers for actual products will vary. The goal is to let software know the absolute numbers so it can decide whether it is worth the hassle to select a power management state.

Editor’s Note 15: This is conceptually the same as the PCI Power Budgeting Data Register, which reports the power a PCI card consumes. It indicates whether the value is for D0, D1, D2, and/or D3. It indicates whether the value is for auxilary, idle, sustained, maximum. It indicates whether the value is for 12V, 3.3V, or 1.8V.

Editor’s Note 16: Interface power management numbers belong elsewhere. PCIe 2.0 reports these latencies to its software: L0s exit latency (64 ns to 4 us), L1 exit latency (1 us to 64 us), L0s Acceptable Latency (64 ns to 4 us to no limit), L1 Acceptable Latency (1 us to 64 us to no limit)

Editor’s Note 17: PCIe 2.0 messages and DLLPs: PM_Enter_L1, PM_Enter_L23, PM_Active_State_Request_L1, RM_Request_Ack, PM_Active_State_Nak, PM_PME, PME_Turn_Off, PME_TO_Ack

~~**3.1.5 active power condition**: When a device server is capable of responding to all of its supported commands, including media access requests, without delay. See 5.9.~~

3.1.51 idle power condition: When a device server is capable of responding to all of its supported commands, including media access requests, but commands may take longer to complete than when in the active power condition. See 5.9.

3.1.150 standby power condition: When a device server is capable of accepting commands, but not capable of processing media access commands. See 5.9.

5.9 Power conditions

5.9.1 Power conditions overview

The optional Power Condition mode page (see 7.4.12) allows an application client to control the power condition of a logical unit in a manner that may reduce power consumption of the SCSI target device. This control is invoked by enabling and setting the idle condition timer and/or, the standby condition timer using the mode page. A change in the power condition of any logical unit in a SCSI target device may result in a change in the SCSI target device's power consumption.

In addition to the Power Condition mode page, the power condition of a logical unit may be controlled by the START STOP UNIT command (see SBC-3 or RBC). If both the Power Condition mode page and the START STOP UNIT command methods are being used to control the power condition of the same logical unit, then any START STOP UNIT command's power condition specification shall override the Power Condition mode page's power control and may disable the idle condition and standby condition timers.

There shall be no notification to the application client that a logical unit has transitioned from one power condition to another. The REQUEST SENSE command (see 6.28) indicates if a logical unit is in the idle power condition or the standby power condition.

Command standards (see 3.1.18) may define for their peripheral device types additional power conditions (e.g., the stopped power condition defined by SBC-3 for direct-access block devices) and extensions to the REQUEST SENSE command for reporting power conditions.

No power condition shall affect the supply of any power required for proper operation of a service delivery subsystem.

Logical units that contain cache memory shall write all cached data to the medium for the logical unit (e.g., as a logical unit would do in response to a SYNCHRONIZE CACHE command as described in SBC-2) prior to entering into any power condition that prevents accessing the media (e.g., before a hard drive stops its spindle motor during transition to the standby power condition).

~~The power conditions are described in table 42.~~

Table 42 — Power Conditions

Power Condition	Description
~~active~~	~~While in the active power condition (see 3.1.5):~~ ~~a) A device server is capable of responding to all of its supported commands including media access requests;~~ ~~b) A logical unit completes processing of operations in the shortest time when compared to the time required for completion while in the idle or standby power conditions; and~~ ~~c) The SCSI target device may consume more power than when the logical unit is in the idle power condition (e.g., a disk drive's spindle motor may be active).~~
~~idle~~	~~While in the idle power condition (see 3.1.51):~~ ~~a) A device server is capable of responding to all of its supported commands including media access requests;~~ ~~b) A logical unit may take longer to complete processing a command than it would while in the active power condition (e.g., the device may have to activate some circuitry before processing a command); and~~ ~~c) The power consumed by the SCSI target device should be less than or equal to the power consumed when the logical unit is in the active power condition and may be greater than the power consumed when the logical unit is in the standby power condition.~~
~~standby~~	~~While in the standby power condition (see 3.1.150):~~ ~~a) A device server is not capable of processing media access commands; and~~ ~~b) The power consumed by the SCSI target device should be less than or equal to the power consumed when the logical unit is in the idle power condition (e.g., a disk drive's spindle motor is stopped).~~

~~...~~

~~5.9.2 Power condition state machine~~

~~5.9.2.1 Power condition state machine overview~~

~~The PC (power condition) state machine describes the logical unit power states and transitions resulting from Power Condition mode page settings.~~

~~The PC states are as follows:~~

a) ~~PC0:Powered_on (see 5.9.2.2) (initial state);~~
b) ~~PC1:Active (see 5.9.2.3);~~
c) ~~PC2:Idle (see 5.9.2.4); and~~
d) ~~PC3:Standby (see 5.9.2.5).~~

~~The PC state machine stall start in the PC0:Powered_on state after power on.~~

~~Figure 5 describes the PC state machine.~~

~~Figure 5 — Pow er condition state machine~~

~~5.9.2.2 PC0:Powered_on state~~

~~5.9.2.2.1 PC0:Powered_on state description~~

~~The logical unit shall enter this state upon power on. This state consumes zero time.~~

~~5.9.2.2.2 Transition PC0:Powered_on to PC1:Active~~

~~This transition shall occur after the logical unit is ready to begin its power on initialization.~~

~~5.9.2.3 PC1:Active state~~

~~5.9.2.3.1 PC1:Active state description~~

~~While in this state, if power on initialization is not complete, then the logical unit shall complete its power on initialization.~~

~~While in this state, if power on initialization is complete, then:~~

~~a) The logical unit is in the active power condition (see table 42);~~
~~b) If the idle condition timer is active, then the idle condition timer is running; and~~
~~c) If the standby condition timer is active, then the standby condition timer is running.~~

~~5.9.2.3.2 Transition PC1:Active to PC2:Idle~~

~~This transition shall occur after:~~

~~a) The idle condition timer is active; and~~
~~b) The idle condition timer is zero.~~

~~5.9.2.3.3 Transition PC1:Active to PC3:Standby~~

~~This transition shall occur after:~~

~~a) The standby condition timer is active; and~~
~~b) The standby condition timer is zero.~~

~~5.9.2.4 PC2:Idle state~~

~~5.9.2.4.1 PC2:Idle state description~~

~~While in this state:~~

~~a) The logical unit is in the idle power condition (see table 42);~~
~~b) The device server processes the REQUEST SENSE command as described in 6.28; and~~
~~c) If the standby condition timer is active, then the standby condition timer is running.~~

~~5.9.2.4.2 Transition PC2:Idle to PC1:Active~~

~~This transition shall occur after the device server processes a command that requires the logical unit to be in the PC1:Active state to process the command.~~

~~5.9.2.4.3 Transition PC2:Idle to PC3:Standby~~

~~This transition shall occur after:~~

~~a) The standby condition timer is active; and~~
~~b) The standby condition timer is zero.~~

~~5.9.2.5 PC3:Standby state~~

~~5.9.2.5.1 PC3:Standby state description~~

~~While in this state:~~

~~a) The logical unit is in the standby power condition (see table 42); and~~
~~b) The device server processes the REQUEST SENSE command as described in 6.28.~~

5.9.2.5.2 Transition PC3:Standby to PC1:Active

This transition shall occur after the device server processes a command that requires the logical unit to be in the PC1:Active state to process the command.

5.9.2.5.3 Transition PC3:Standby to PC2:Idle

This transition shall occur after the device server processes a command that requires the logical unit to be in the PC2:Idle state to process the command.

7.4.12 Power Condition mode page

The Power Condition mode page provides an application client with methods to control the power condition of a logical unit (see 5.9). These methods include:

a) Specifying that the logical unit transition to a power condition without delay; and
b) Activating and setting of idle condition and standby condition timers to specify that the logical unit wait for a period of inactivity before transitioning to a specified power condition.

The mode page policy (see 6.9) for this mode page shall be shared.

When a device server receives a command while in a power condition based on a setting in the Power Condition mode page, the logical unit shall transition to the power condition that allows the command to be processed. If either the idle condition timer or the standby condition timer has been set, then they shall be reset on receipt of the command. On completion of the command, the timer(s) shall be started.

Logical units that contain cache memory shall write all cached data to the medium for the logical unit (e.g., as a logical unit does in response to a SYNCHRONIZE CACHE command as described in SBC-2) prior to entering into any power condition that prevents accessing the media (e.g., before a hard drive stops its spindle motor during transition to the standby power condition).

The logical unit shall use the values in the Power Condition mode page to control its power condition after a power on or a hard reset until a START STOP UNIT command setting a power condition is received.

Table 319 defines the Power Condition mode page.

Table 319 — Power Condition mode page

Byte\Bit	7	6	5	4	3	2	1	0
0	PS	SPF (0b)	PAGE CODE (1Ah)					
1	PAGE LENGTH (0Ah)							
2	Reserved							
3	Reserved						IDLE	STANDBY
4	(MSB)	IDLE CONDITION TIMER						
7								(LSB)
8	(MSB)	STANDBY CONDITION TIMER						
11								(LSB)

The PS bit, SPF bit, PAGE CODE field, and PAGE LENGTH field are described in 7.4.5.

The IDLE and STANDBY bits specify which timers are active.

If the IDLE bit is set to one and the STANDBY bit is set to zero, then the idle condition timer is active and the device server shall transition to the idle power condition when the idle condition timer is zero.

If the IDLE bit is set to zero, then the device server shall ignore the idle condition timer.

If the STANDBY bit is set to one and the IDLE bit is set to zero, then the standby condition timer is active and the device server shall transition to the standby power condition when the standby condition timer is zero.

~~If the STANDBY bit is set to zero, then the device server shall ignore the standby condition timer.~~

~~If both the IDLE and STANDBY bits are set to one, then both timers are active and run concurrently. When the idle condition timer is zero the device server shall transition to the idle power condition. When the standby condition timer is zero the device server shall transition to the standby power condition. If the standby condition timer is zero before the idle condition timer is zero, then the logical unit shall transition to the standby power condition.~~

~~The value in the IDLE CONDITION TIMER field specifies the inactivity time in 100 millisecond increments that the logical unit shall wait before transitioning to the idle power condition when the IDLE bit is set to one. The idle condition timer is expired when:~~

~~a) The IDLE CONDITION TIMER field is set to zero; or~~
~~b) The number of milliseconds specified by the value in the IDLE CONDITION TIMER field times 100 milliseconds has elapsed since the last activity (e.g., processing a command that requires the active power condition or performing a self test).~~

~~The value in the STANDBY CONDITION TIMER field specifies the inactivity time in 100 millisecond increments that the logical unit shall wait before transitioning to the standby power condition when the STANDBY bit is set to one. The standby condition timer is expired when:~~

~~a) The STANDBY CONDITION TIMER field is set to zero; or~~
~~b) The number of milliseconds specified by the value in the STANDBY CONDITION TIMER field times 100 milliseconds has elapsed since the last activity (e.g., processing any command or performing a self test).~~

Suggested changes to SBC-4

Editor's Note 18: These changes are targeted for SBC-4, not SBC-3

Editor's Note 19: Keep START STOP UNIT for removable media (load/eject bit) but obsolete using it for changing power conditions (idle, standby, stopped)

4.2.2 Rotating media

The typical application of a direct-access block device is a magnetic disk device. The medium is a spinning disk with a magnetic material that allows flux changes to be induced and recorded. An actuator positions a read-write head radially across the spinning disk, allowing the device to randomly read or write the information at any radial position. Data is stored by using the write portion of the head to record flux changes and is read by using the read portion of the head to read the recorded data.

The circular path followed by the read-write head at a particular radius is called a track. The track is divided into sectors each containing blocks of stored data. If there are more than one disk spinning on a single axis and the actuator has one or more read-write heads to access the disk surfaces, the collection of tracks at a particular radius is called a cylinder.

A logical block is stored in one or more sectors, or a sector may store more than one logical block. Sectors may also contain information for accessing, synchronizing, and protecting the integrity of the logical blocks.

A rotating media-based direct-access block device is ready when the disks are rotating at the correct speed and the read-write circuitry is powered and ready to access the data, and may require a ~~START STOP UNIT command (see 5.19)~~ MANAGE POWER OUT (see SPC-4) to bring the logical unit to the ready state.

Rotating media-based direct-access block device are usually non-volatile.

The defect management scheme of a disk device may not be discernible through this command set, though some aspects (see 4.9) may be accessible to the application client with the READ LONG commands and the WRITE LONG commands (see 5.16, 5.17, 5.35, and 5.36).

4.2.3 Memory media

Memory media is based on solid state random access memories (RAMs) (e.g., static RAM (SRAM), dynamic RAM (DRAM), magnetoresistive RAM (MRAM), ferroelectric RAM (FeRAM), or flash memory). Memory media-based direct-access block devices may be used for fast-access storage.

A memory media-based direct-access block device is ready after power on, and does not require a ~~START STOP UNIT command (see 5.19)~~ MANAGE POWER OUT (see SPC-4) to bring the logical unit to a ready state.

These logical units may be non-mechanical, and therefore logical blocks may be accessed with similar access times regardless of their location on the medium. Memory media-based direct-access block devices may store less data than disks or tapes, and may be volatile.

The defect management scheme (e.g., ECC bytes) (see 4.9) may be accessible to the application client with the READ LONG commands and the WRITE LONG commands (see 5.16, 5.17, 5.35, and 5.36).

Memory media may be volatile (e.g., SRAM or DRAM) or non-volatile (e.g., SRAM or DRAM with battery backup, MRAM, FeRAM, or flash memory).

4.3 Removable medium

4.3.1 Removable medium overview

The medium may be removable or non-removable. The removable medium may be contained within a cartridge or jacket to prevent damage to the recording surfaces.

A removable medium has an attribute of being mounted or unmounted on a suitable transport mechanism in a direct-access block device. A removable medium is mounted when the direct-access block device is capable of performing write, read, and verify operations to the medium. A removable medium is unmounted at any other time (e.g., during loading, unloading, or storage).

An application client may check whether a removable medium is mounted by issuing a TEST UNIT READY command (see SPC-4). A direct-access block device containing a removable medium may not be accessible for write, read, and verify operations until it receives a START STOP UNIT command (see 5.19).

If the direct-access block device implements cache, either volatile or non-volatile, it ensures that all logical blocks of the medium contain the most recent user data and protection information, if any, prior to permitting unmounting of the removable medium.

If the medium in a direct-access block device is removable, and the medium is removed, then the device server shall establish a unit attention condition with the additional sense code set to the appropriate value (e.g., NOT READY TO READY CHANGE, MEDIUM MAY HAVE CHANGED).

The PREVENT ALLOW MEDIUM REMOVAL command (see 5.6) allows an application client to restrict the unmounting of the removable medium. This is useful in maintaining system integrity.

If the application client issues a START STOP UNIT command to eject the removable medium, and the direct-access block device is prevented from unmounting by the PREVENT ALLOW MEDIUM REMOVAL command, the START STOP UNIT command is rejected by the device server.

~~4.16 START STOP UNIT and power conditions~~

~~4.16.1 START STOP UNIT and power conditions overview~~

~~The START STOP UNIT command (see 5.19) allows an application client to control the power condition of a logical unit. This method includes specifying that the logical unit transition to a power condition.~~

~~In addition to the START STOP UNIT command, the power condition of a logical unit may be controlled by the Power Condition mode page (see SPC-4). If both the START STOP UNIT command and the Power Condition mode page methods are being used to control the power condition of the same logical unit, then the power condition specified by any START STOP UNIT command shall override the Power Condition mode page's power control.~~

~~There shall be no notification to the application client that a logical unit has transitioned from one power condition to another. The REQUEST SENSE command (see SPC-4) indicates if a logical unit is in the idle power condition or the standby power condition and may indicate if a logical unit is in the stopped power condition.~~

~~If the logical unit is in the idle power condition, the device server shall process a REQUEST SENSE command by:~~

1) ~~returning parameter data containing sense data with the sense key set to NO SENSE and the additional sense code set to:~~
   A) ~~LOW POWER CONDITION ON if the reason for entry into the idle power condition is unknown;~~
   B) ~~IDLE CONDITION ACTIVATED BY TIMER if the logical unit entered the idle power condition due to the idle condition timer (see SPC-4); and~~
   C) ~~IDLE CONDITION ACTIVATED BY COMMAND if the logical unit entered the idle power condition due to a START STOP UNIT command or receipt of a command requiring the idle power condition while it was in the standby power condition;~~

   ~~and~~
2) ~~returning GOOD status for the REQUEST SENSE command.~~

~~If the logical unit is in the standby power condition, the device server shall process a REQUEST SENSE command by:~~

1) ~~returning parameter data containing sense data with the sense key set to NO SENSE and the additional sense code set to:~~

A) LOW POWER CONDITION ON if the reason for entry into the standby power condition is unknown;
B) STANDBY CONDITION ACTIVATED BY TIMER if the logical unit entered the standby power condition due to the standby condition timer (see SPC-4); and
C) STANDBY CONDITION ACTIVATED BY COMMAND if the logical unit entered the idle power condition due to a START STOP UNIT command;

and

2) returning GOOD status for the REQUEST SENSE command.

If the logical unit is in the stopped power condition, the device server shall process a REQUEST SENSE command by:

1) returning parameter data containing sense data with:
   A) the sense key set to NO SENSE and the additional sense code set to NO ADDITIONAL SENSE INFORMATION; or
   B) the sense key set to NOT READY and the additional sense code set to LOGICAL UNIT NOT READY, INITIALIZING COMMAND REQUIRED;

   and

2) returning GOOD status for the REQUEST SENSE command.

No power condition shall affect the supply of any power required for proper operation of a service delivery subsystem.

4.16.2 START STOP UNIT and power conditions state machine

4.16.2.1 START STOP UNIT and power conditions state machine overview

The SSU_PC (start stop unit power condition) state machine for logical units implementing the START STOP UNIT command describes the logical unit power states and transitions resulting from settings by the START STOP UNIT command and settings in the Power Condition mode page (see SPC-4).

The SSU_PC states are as follows:

a) SSU_PC0:Powered_on (see 4.16.2.2) (initial state);
b) SSU_PC1:Active (see 4.16.2.3);
c) SSU_PC2:Idle (see 4.16.2.4);
d) SSU_PC3:Standby (see 4.16.2.5); and
e) SSU_PC4:Stopped (see 4.16.2.6)..

The SSU_PC state machine shall start in the SSU_PC0:Powered_on state after power on.

NOTE 2 - NOTE 6 - The SSU_PC state machine is an enhanced version of the Power Condition state machine described in SPC-4.

Figure 4 describes the SSU_PC state machine.

Figure 4 — Power condition state machine for logical units implementing the START STOP UNIT command

4.16.2.2 SSU_PC0:Powered_on state

4.16.2.2.1 SSU_PC0:Powered_on state description

The logical unit shall enter this state upon power on. This state consumes zero time.

4.16.2.2.2 Transition SSU_PC0:Powered_on to SSU_PC1:Active

This transition shall occur if:

a) the logical unit has been configured to transition to the SSU_PC1:Active state.

4.16.2.2.3 Transition SSU_PC0:Powered_on to SSU_PC4:Stopped

~~This transition shall occur if:~~

a) ~~the logical unit has been configured to transition to the SSU_PC4:Stopped state.~~

~~4.16.2.3 SSU_PC1:Active state~~

~~4.16.2.3.1 SSU_PC1:Active state description~~

~~While in this state, if power on initialization is not complete, then the logical unit completes its power on initialization.~~

~~While in this state, after power on initialization is complete, then:~~

a) ~~the logical unit is in the active power condition (see SPC-4);~~
b) ~~if the idle condition timer is active (see SPC-4) and not disabled (see 5.19), then the idle condition timer is running; and~~
c) ~~if the standby condition timer is active (see SPC-4) and not disabled (see 5.19), then the standby condition timer is running.~~

~~4.16.2.3.2 Transition SSU_PC1:Active to SSU_PC2:Idle~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to IDLE;~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0; or~~
c) ~~the idle condition timer is active (see SPC-4), enabled (see 5.19), and zero.~~

~~4.16.2.3.3 Transition SSU_PC1:Active to SSU_PC3:Standby~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to STANDBY;~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0; or~~
c) ~~the standby condition timer is active (see SPC-4), enabled (see 5.19), and zero.~~

~~4.16.2.3.4 Transition SSU_PC1:Active to SSU_PC4:Stopped~~

~~This transition shall occur after the device server processes a START STOP UNIT command with the START bit set to zero and the POWER CONDITION field set to START_VALID.~~

~~4.16.2.4 SSU_PC2:Idle state~~

~~4.16.2.4.1 SSU_PC2:Idle state description~~

~~While in this state:~~

a) ~~the logical unit is in the idle power condition (see SPC-4);~~
b) ~~the device server processes the REQUEST SENSE command as described in 4.16.1; and~~
c) ~~if the standby condition timer is active (see SPC-4) and not disabled (see 5.19), then the standby condition timer is running.~~

~~4.16.2.4.2 Transition SSU_PC2:Idle to SSU_PC1:Active~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the START bit set to one;~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to ACTIVE; or~~
c) ~~the device server processes a command that requires the logical unit to be in the SSU_PC1:Active state to process the command.~~

~~4.16.2.4.3 Transition SSU_PC2:Idle to SSU_PC3:Standby~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to STANDBY;~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0; or~~
c) ~~the standby condition timer is active (see SPC-4), enabled (see 5.19), and zero.~~

~~4.16.2.4.4 Transition SSU_PC2:Idle to SSU_PC4:Stopped~~

~~This transition shall occur after the device server processes a START STOP UNIT command with the START bit set to zero.~~

~~4.16.2.5 SSU_PC3:Standby state~~

~~4.16.2.5.1 SSU_PC3:Standby state description~~

~~While in this state:~~

a) ~~the logical unit is in the standby power condition (see SPC-4); and~~
b) ~~the device server processes the REQUEST SENSE command as described in 4.16.1.~~

~~4.16.2.5.2 Transition SSU_PC3:Standby to SSU_PC1:Active~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the START bit set to one;~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to ACTIVE; or~~
c) ~~the device server processes a command that requires the logical unit to be in the SSU_PC1:Active state to process the command.~~

~~4.16.2.5.3 Transition SSU_PC3:Standby to SSU_PC2:Idle~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to IDLE;~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0; or~~
c) ~~the device server processes a command that requires the logical unit to be in the SSU_PC2:Idle state to process the command.~~

~~4.16.2.5.4 Transition SSU_PC3:Standby to SSU_PC4:Stopped~~

~~This transition shall occur after the device server processes a START STOP UNIT command with the START bit set to zero.~~

~~4.16.2.6 SSU_PC4:Stopped state~~

~~4.16.2.6.1 SSU_PC4:Stopped state description~~

~~While in this state:~~

a) ~~the logical unit is in the stopped power condition;~~
b) ~~the device server is not capable of processing medium access commands. The device server shallterminate each medium access command or TEST UNIT READY command processed while in this state with CHECK CONDITION status with the sense key set to NOT READY and the additional sense code set to LOGICAL UNIT NOT READY, INITIALIZING COMMAND REQUIRED;~~
c) ~~the device server processes the REQUEST SENSE command as described in 4.16.1; and~~
d) ~~the power consumed by the SCSI target device should be less than or equal to that consumed than when the logical unit is in the SSU_PC1:Active, SSU_PC2:Idle, or SSU_PC3:Standby states.~~

~~4.16.2.6.2 Transition SSU_PC4:Stopped to SSU_PC1:Active~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the START bit set to one; or~~

b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to ACTIVE.~~

~~4.16.2.6.3 Transition SSU_PC4:Stopped to SSU_PC2:Idle~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to IDLE; or~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0.~~

~~4.16.2.6.4 Transition SSU_PC4:Stopped to SSU_PC3:Standby~~

~~This transition shall occur after:~~

a) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to STANDBY; or~~
b) ~~the device server processes a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0.~~

5.19 START STOP UNIT command

The START STOP UNIT command (see table 60) requests that the device server ~~change the power condition of the logical unit (see 4.16) or~~ load or eject the medium. ~~This includes specifying that the device server enable or disable the direct-access block device for medium access operations by controlling power conditions and timers.~~

~~Logical units that contain cache shall write all cached logical blocks to the medium (e.g., as they would do in response to a SYNCHRONIZE CACHE command (see 5.20 and 5.21) with the SYNC_NV bit set to zero, the LOGICAL BLOCK ADDRESS field set to zero, and the NUMBER OF LOGICAL BLOCKS field set to zero) prior to entering into any power condition that prevents accessing the medium (e.g., before the rotating media spindle motor is stopped during transition to the stopped power condition).~~

If any deferred downloaded code has been received as a result of a WRITE BUFFER command (see SPC-4), then that deferred downloaded code shall replace the current operational code.

Table 60 — START STOP UNIT command

Byte\Bit	7	6	5	4	3	2	1	0
0	OPERATION CODE (1Bh)							
1	Reserved							IMMED
5	Reserved							
5	Reserved							
4	~~POWER CONDITION~~ Obsolete				Reserved		LOEJ	START
5	CONTROL							

The OPERATION CODE field is defined in SPC-4 and shall be set to the value defined in table 60.

If the immediate (IMMED) bit is set to zero, then the device server shall return status after the operation is completed. If the IMMED bit set to one, then the device server shall return status as soon as the CDB has been validated.

~~The POWER CONDITION field is used to specify that the logical unit be placed into a power condition or to adjust a timer as defined in table 61. If this field is supported and is set to a value other than 0h, then the START and LOEJ bits shall be ignored.~~

Table 61 — POWER CONDITION field

Code	Name	Description
~~0h~~	~~START_VALID~~	~~Process the START and LOEJ bits.~~
~~1h~~	~~ACTIVE~~	~~Place the deviceinto the active power condition.~~
~~2h~~	~~IDLE~~	~~Place the device into the idle power condition.~~
~~3h~~	~~STANDBY~~	~~Place the device into the standby power condition.~~
~~4h - 6h~~	~~Reserved~~	
~~7h~~	~~LU_CONTROL~~	~~Transfer control of power conditions to the logical unit.~~
~~8h - 9h~~	~~Reserved~~	
~~Ah~~	~~FORCE_IDLE_0~~	~~Force the idle condition timer to zero.~~
~~Bh~~	~~FORCE_STANDBY_0~~	~~Force the standby condition timer to zero.~~
~~Ch - Fh~~	~~Reserved~~	

The contents of the CONTROL byte are defined in SAM-4.

~~If the START STOP UNIT command is processed with the POWER CONDITION field set to ACTIVE, IDLE, or STANDBY, then:~~

a) ~~the logical unit shall transition to the specified power condition;~~
b) ~~the logical unit shall change power conditions only after receipt of another START STOP UNIT command or a logical unit reset; and~~
c) ~~the device server shall disable the idle condition timer if it is active (see SPC-4) and disable the standby condition timer if it is active (see SPC-4) until another START STOP UNIT command is processed that returns control of the power condition to the logical unit, or a logical unit reset occurs.~~

~~If the START STOP UNIT command is processed with the POWER CONDITION field set to LU_CONTROL, then the device server shall enable the idle condition timer if it is active (see SPC-4) and disable the standby condition timer if it is active (see SPC-4).~~

~~If the START STOP UNIT command is processed with the POWER CONDITION field set to FORCE_IDLE_0 or FORCE_STANDBY_0, then the device server shall:~~

a) ~~force the specified timer to zero, cause the logical unit to transition to the specified power condition, and return control of the power condition to the device server; or~~
b) ~~terminate a START STOP UNIT command that selects a timer that is not supported by the device server or a timer that is not active. The command shall be terminated with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN CDB.~~

~~It is not an error to specify that the logical unit transition to its current power condition.~~

If the load eject (LOEJ) bit is set to zero, then the START bit is obsolete and the logical unit shall take no action regarding loading or ejecting the medium. If the LOEJ bit is set to one, then:

a) the logical unit shall unload the medium if the START bit is set to zero~~. If the LOEJ bit is set to one, then~~ ; or
b) the logical unit shall load the medium if the START bit is set to one.

~~If the START bit is set to zero, then the logical unit shall transition to the stopped power condition, disable the idle condition timer if it is active (see SPC-4), and disable the standby condition timer if it is active (see~~

SPC-4). If the START bit set to one, then the logical unit shall transition to the active power condition, enable the idle condition timer if it is active, and enable the standby condition timer if it is active.

Suggested changes to SAS-2.1 (please ignore in this revision of the proposal)

Editor's Note 20: **Please ignore this section of the proposal**, which was only valid in 08-126r0. Eventually, changes will be needed to support NOTIFY (ENABLE SPINUP) behavior in the new model. There won't be a simple state machine with 4 states to modify, since there will be possibly 256 power management states.

2.2 Approved references

At the time of publication, the following referenced standards or technical reports were approved:

ISO/IEC 14776-326, *Reduced Block Commands (RBC)* (ANSI INCITS 330-2000 with ANSI INCITS 330-2003/AM1)

ANSI INCITS TR-35-2004, *Methodologies for Jitter and Signal Quality Specification (MJSQ)*. When MJSQ is referenced from this standard, the FC Port terminology used within MJSQ should be substituted with SAS phy terminology.

2.3 References under development

At the time of publication, the following referenced standards were still under development. For information on the current status of the document, or regarding availability, contact the relevant standards body or other organization as indicated.

ISO/IEC xxxxx-xxx, *ATA Attachment-8 ATA/ATAPI Architecture Model (ATA8-AAM)* (T13/1700-D)

ISO/IEC xxxxx-xxx, *ATA Attachment-8 ATA/ATAPI Command Set (ATA8-ACS)* (T13/1699-D)

ISO/IEC 14776-414, *SCSI Architecture Model-4 (SAM-4)* (T10/1683-D)

ISO/IEC 14776-454, *SCSI Primary Commands-4 (SPC-4)* (T10/1731-D)

ISO/IEC 14776-323, *SCSI Block Commands-3 (SBC-3)* (T10/1799-D)

ISO/IEC 14776-372, *SCSI Enclosure Services-2 (SES-2)* (T10/1559-D)

ISO/IEC 14776-366, *Multi-Media Commands-6 (MMC-6)* (T10/1836-D)

NOTE 3 - For more information on the current status of these documents, contact the INCITS Secretariat at 202-737-8888 (phone), 202-638-4922 (fax) or via Email at incits@itic.org. To obtain copies of these documents, contact Global Engineering at 15 Inverness Way, East Englewood, CO 80112-5704 at 303-792-2181 (phone), 800-854-7179 (phone), or 303-792-2192 (fax) or see http://www.incits.org.

2.4 Symbols and abbreviations

See 2.1 for abbreviations of standards bodies (e.g., ISO). Units and abbreviations used in this standard:

Abbreviation	Meaning
ATA8-AAM	AT Attachment - 8 ATA/ATAPI Architecture Model standard (see 2.3)
ATA8-ACS	AT Attachment - 8 ATA/ATAPI Command Set standard (see 2.3)
MMC-6	Multi-Media Commands standard (see 2.3)
RBC	Reduced Block Commands standard (see 2.2)
SAM-4	SCSI Architecture Model - 4 standard (see 2.3)
SBC-3	SCSI Block Commands - 3 standard (see 2.3)
SPC-4	SCSI Primary Commands - 4 standard (see 2.3)

7.2.5.3.2 NOTIFY (ENABLE SPINUP)

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-4) and/or the START STOP UNIT command (see SBC-3, MMC-6, or RBC), as described in 10.2.10.

7.10 Power management

SATA interface power management is not supported in STP.

STP initiator ports shall not generate SATA_PMREQ_P, SATA_PMREQ_S, or SATA_PMACK. If an STP initiator port receives SATA_PMREQ_P or SATA_PMREQ_S, it shall reply with SATA_PMNAK.

If an expander device receives SATA_PMREQ_P or SATA_PMREQ_S from a SATA device while an STP connection is not open, it shall not forward it to any STP initiator port and shall reply with SATA_PMNAK. If one of these primitives arrives while an STP connection is open, it may forward the primitive to the STP initiator port.

SCSI idle ~~and~~, standby, and sleep power conditions, implemented with the START STOP UNIT command (see SBC-3, MMC-6, or RBC) and the Power Condition mode page (see SPC-4), may be supported by SSP initiator ports and SSP target ports as described in 10.2.10.

ATA idle ~~and~~, standby, and sleep power modes, implemented with the IDLE, IDLE IMMEDIATE, STANDBY, STANDBY IMMEDIATE, SLEEP, and CHECK POWER MODE commands (see ATA8-ACS), may be supported by STP initiator ports. ~~The ATA sleep power mode, implemented with the SLEEP command, shall not be used.~~

10.2.6.5 START STOP UNIT command

The power condition states controlled by the START STOP UNIT command (see SBC-3, MMC-6, or RBC) for a SAS device are described in 10.2.10.

10.2.10 SCSI power conditions

10.2.10.1 SCSI power conditions overview

The logical unit power condition states from the Power Condition mode page (see SPC-4) and START STOP UNIT command (see SBC-3, MMC-6, or RBC), if implemented, shall interact with the NOTIFY (ENABLE SPINUP) primitive (see 7.2.5.3) to control temporary consumption of additional power (e.g., spin-up of rotating media) as described in this subclause.

The logical unit uses NOTIFY (ENABLE SPINUP) to:

a) initiate spin-up after power on; and
b) delay spin-ups requested by START STOP UNIT commands.

Table 23 defines the power conditions supported by this standard.

Table 62 — Power conditions

Power condition	Target port(s)	Maximum recovery time [b]	References
Active	on [a]	See SPC-4	10.2.10.2.3 10.2.10.2.8 SPC-4
Idle	on [a]	See SPC-4	10.2.10.2.4 10.2.10.2.9 SPC-4
Standby	on [a]	See SPC-4	10.2.10.2.5 SPC-4
Stopped	on [a]	15 s	10.2.10.2.6 SBC-3, MMC-6, RBC
Sleep	off	30 s	10.2.10.2.7 SBC-3, MMC-6, RBC

[a] Interface power management (see TBD) may used to temporarily turn off a target port, but recovery is automatic.
[b] NOTIFY (ENABLE SPINUP) may increase the recovery time beyond this maximum.

Editor's Note 21: Not sure what contents belong in this standard. Don't want to double-specify anything.

10.2.10.2 SA_PC (SCSI application layer power condition) state machine

10.2.10.2.1 SA_PC state machine overview

The SA_PC (SCSI application layer power condition) state machine describes how the SAS target device processes logical unit power condition state change requests and NOTIFY (ENABLE SPINUP) if it is a SCSI target device.

NOTE 4 - This state machine is an enhanced version of the logical unit power condition state machines described in SPC-4 and SBC-3.

This state machine consists of the following states:

a) SA_PC_0:Powered_On (see 10.2.10.2.2)(initial state);
b) SA_PC_1:Active (see 10.2.10.2.3);
c) SA_PC_2:Idle (see 10.2.10.2.4);
d) SA_PC_3:Standby (see 10.2.10.2.5);
e) SA_PC_4:Stopped (see 10.2.10.2.6)(specific to ~~SBC-3~~ logical units with certain peripheral device types);
f) SA_PC_5:Sleep (see 10.2.10.2.6)(specific to logical units with certain peripheral device types);
g) SA_PC_~~5~~6:Active_Wait (see 10.2.10.2.8)(specific to SAS devices); and
h) SA_PC_~~6~~7:Idle_Wait (see 10.2.10.2.9)(specific to SAS devices).

This state machine shall start in the SA_PC_0:Powered_On state after power on.

If the device server processes a START STOP UNIT command (see SBC-3, MMC-6, or RBC) with the IMMED bit set to one, it may complete the command before completing the transition, if any, specified by the POWER CONDITION field and the START bit.

Figure 220 describes the SA_PC state machine.

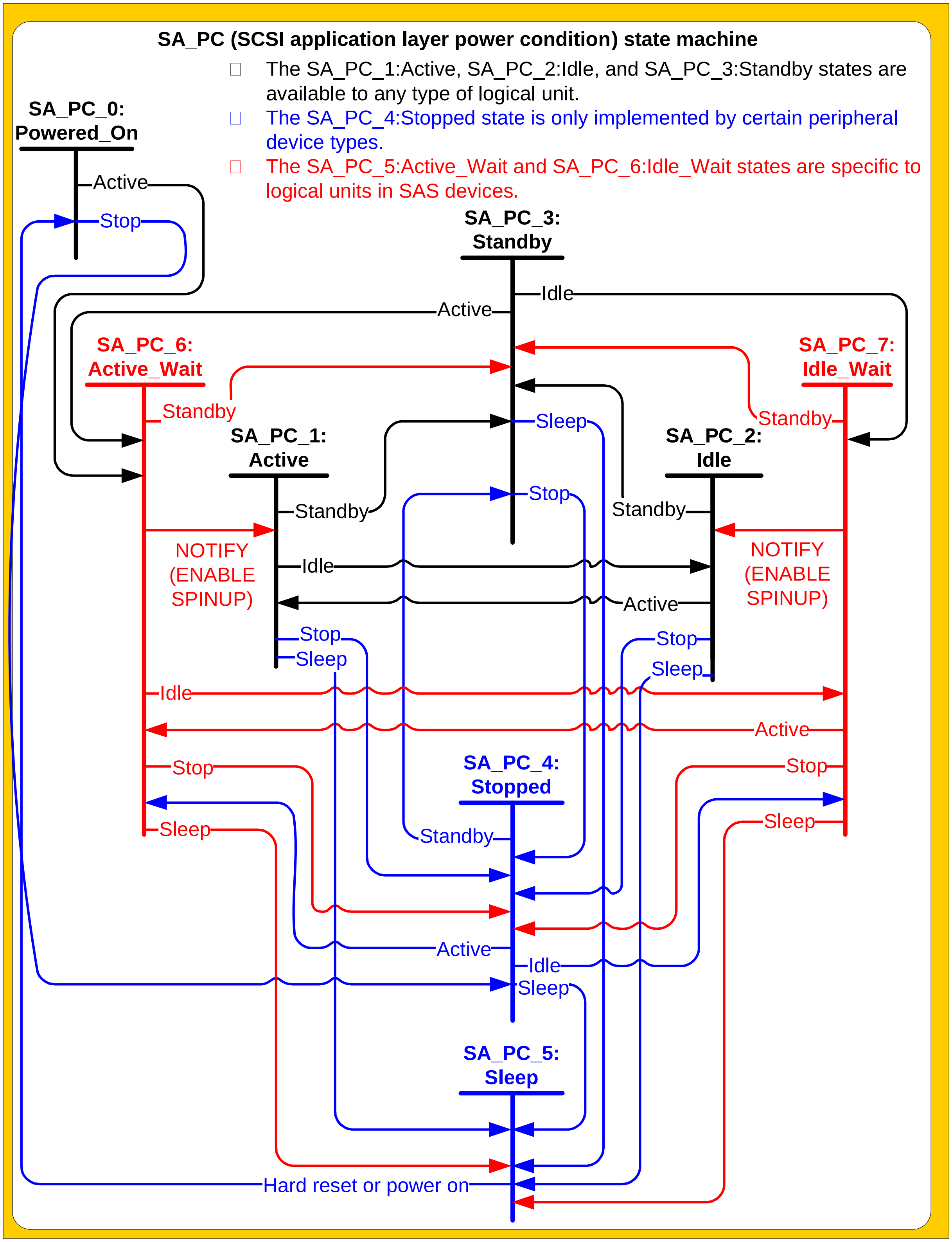

Figure 220 — SA_PC (SCSI application layer power condition) state machine for SAS [updated]

10.2.10.2.2 SA_PC_0:Powered_On state

10.2.10.2.2.1 State description

This state shall be entered upon power on. This state consumes zero time.

10.2.10.2.2.2 Transition SA_PC_0:Powered_On to SA_PC_4:Stopped

This transition shall occur if the SAS device has been configured to start in the SA_PC_4:Stopped state.

10.2.10.2.2.3 Transition SA_PC_0:Powered_On to SA_PC_~~5~~6:Active_Wait

This transition shall occur if the SAS device has been configured to start in the SA_PC_~~5~~6:Active_Wait state.

10.2.10.2.3 SA_PC_1:Active state

10.2.10.2.3.1 State description

While in this state, rotating media in block devices shall be active (i.e., rotating or spinning).

See SPC-4 for more details about this state.

10.2.10.2.3.2 Transition SA_PC_1:Active to SA_PC_2:Idle

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to IDLE is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0 is processed; or
c) the Power Condition mode page idle condition timer expires.

10.2.10.2.3.3 Transition SA_PC_1:Active to SA_PC_3:Standby

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to STANDBY is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0 is processed; or
c) the Power Condition mode page standby condition timer expires.

10.2.10.2.3.4 Transition SA_PC_1:Active to SA_PC_4:Stopped

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to zero is processed.

10.2.10.2.3.5 Transition SA_PC_1:Active to SA_PC_5:Sleep

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to SLEEP is processed.

10.2.10.2.4 SA_PC_2:Idle state

10.2.10.2.4.1 State description

While in this state, rotating media in block devices shall be active (i.e., rotating or spinning).

See SPC-4 for more details about this state.

10.2.10.2.4.2 Transition SA_PC_2:Idle to SA_PC_1:Active

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to one is processed;

b) a START STOP UNIT command with the POWER CONDITION field set to ACTIVE is processed; or
c) a command that requires the active power condition is processed.

10.2.10.2.4.3 Transition SA_PC_2:Idle to SA_PC_3:Standby

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to STANDBY is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0 is processed; or
c) the Power Condition mode page standby condition timer expires.

10.2.10.2.4.4 Transition SA_PC_2:Idle to SA_PC_4:Stopped

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to zero is processed.

10.2.10.2.4.5 Transition SA_PC_2:Idle to SA_PC_5:Sleep

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to SLEEP is processed.

10.2.10.2.5 SA_PC_3:Standby state

10.2.10.2.5.1 State description

While in this state, rotating media in block devices shall be stopped.

See SPC-4 for more details about this state.

10.2.10.2.5.2 Transition SA_PC_3:Standby to SA_PC_4:Stopped

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to zero is processed.

10.2.10.2.5.3 Transition SA_PC_3:Standby to SA_PC_5:Sleep

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to SLEEP is processed.

10.2.10.2.5.4 Transition SA_PC_3:Standby to SA_PC_~~5~~6:Active_Wait

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to one is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to ACTIVE is processed; or
c) a command that requires the active power condition is processed.

If the transition is based on a START STOP UNIT command with the IMMED bit set to zero, the device server shall not complete the command until this state machine reaches the SA_PC_1:Active state.

10.2.10.2.5.5 Transition SA_PC_3:Standby to SA_PC_~~6~~7:Idle_Wait

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to IDLE is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0 is processed; or
c) a command that requires the idle power condition is processed.

If the transition is based on a START STOP UNIT command with the IMMED bit set to zero, the device server shall not complete the command until this state machine reaches the SA_PC_2:Idle state.

10.2.10.2.6 SA_PC_4:Stopped state

10.2.10.2.6.1 State description

This state is only implemented in ~~block devices~~certain peripheral device types (see SBC-3, MMC-6, or RBC).

While in this state, rotating media shall be stopped.

See ~~SBC-3~~the command standard for the peripheral device type for more details about this state.

10.2.10.2.6.2 Transition SA_PC_4:Stopped to SA_PC_3:Standby

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to STANDBY is processed; or
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0 is processed.

10.2.10.2.6.3 Transition SA_PC_3:Stopped to SA_PC_5:Sleep

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to SLEEP is processed.

10.2.10.2.6.4 Transition SA_PC_4:Stopped to SA_PC_~~5~~6:Active_Wait

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to one is processed; or
b) a START STOP UNIT command with the POWER CONDITION field set to ACTIVE is processed.

If the transition is based on a START STOP UNIT command with the IMMED bit set to zero, the device server shall not complete the command until this state machine reaches the SA_PC_1:Active state.

10.2.10.2.6.5 Transition SA_PC_4:Stopped to SA_PC_~~6~~7:Idle_Wait

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to IDLE is processed; or
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0 is processed.

If the transition is based on a START STOP UNIT command with the IMMED bit set to zero, the device server shall not complete the command until this state machine reaches the SA_PC_2:Idle state.

10.2.10.2.7 SA_PC_5:Sleep state

10.2.10.2.7.1 State description

This state is only implemented in certain peripheral device types (see SBC-3, MMC-6, or RBC).

When a phy is inactive because the logical unit is in this state:

a) the phy transmitter shall transmit D.C. idle; and
b) the phy receiver shall be capable of detecting COMINIT.

See the command standard for the peripheral device type for more details about this state.

Editor's Note 22: Is D.C. idle necessary or is completely tristate better?

4.16.2.6.2 Transition SA_PC_5:Sleep to SA_PC_0:Powered_On

This transition shall occur after:

a) hard reset or power on.

10.2.10.2.8 SA_PC_56:Active_Wait state

10.2.10.2.8.1 State description

This state shall only be implemented in SAS devices.

While in this state, rotating media in block devices shall be stopped. The device server shall be capable of processing commands and shall terminate each media access command or TEST UNIT READY command with CHECK CONDITION status with the sense key set to NOT READY and the additional sense code set to LOGICAL UNIT NOT READY, NOTIFY (ENABLE SPINUP) REQUIRED.

In response to a REQUEST SENSE command processed in this state, the device server shall return parameter data containing sense data with the sense key set to NOT READY and the additional sense code set to LOGICAL UNIT NOT READY, NOTIFY (ENABLE SPINUP) REQUIRED and return GOOD status for the command.

10.2.10.2.8.2 Transition SA_PC_56:Active_Wait to SA_PC_1:Active

This transition shall occur if:

a) a NOTIFY (ENABLE SPINUP) is detected; or
b) the SAS device does not consume additional power as a result of the transition to SA_PC_1:Active.

10.2.10.2.8.3 Transition SA_PC_56:Active_Wait to SA_PC_3:Standby

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to STANDBY is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0 is processed; or
c) the Power Condition mode page standby condition timer expires.

10.2.10.2.8.4 Transition SA_PC_56:Active_Wait to SA_PC_4:Stopped

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to zero is processed.

10.2.10.2.8.5 Transition SA_PC_6:Active_Wait to SA_PC_5:Sleep

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to SLEEP is processed.

10.2.10.2.8.6 Transition SA_PC_56:Active_Wait to SA_PC_67:Idle_Wait

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to IDLE is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_IDLE_0 is processed; or
c) the Power Condition mode page idle condition timer expires.

If the transition is based on a START STOP UNIT command with the IMMED bit set to zero, the device server shall not complete the command until this state machine reaches the SA_PC_2:Idle state.

10.2.10.2.9 SA_PC_67:Idle_Wait state

10.2.10.2.9.1 State description

This state shall only be implemented in SAS devices.

While in this state, rotating media in block devices shall be stopped. The device server shall be capable of processing commands and shall terminate each media access command or TEST UNIT READY command with CHECK CONDITION status with the sense key set to NOT READY and the additional sense code set to LOGICAL UNIT NOT READY, NOTIFY (ENABLE SPINUP) REQUIRED.

In response to a REQUEST SENSE command processed in this state, the device server shall return parameter data containing sense data with the sense key set to NOT READY and the additional sense code set to LOGICAL UNIT NOT READY, NOTIFY (ENABLE SPINUP) REQUIRED and return GOOD status for the command.

10.2.10.2.9.2 Transition SA_PC_~~6~~7:Idle_Wait to SA_PC_2:Idle

This transition shall occur if:

a) a NOTIFY (ENABLE SPINUP) is detected; or
b) the SAS device does not consume additional power as a result of the transition to SA_PC_2:Idle.

10.2.10.2.9.3 Transition SA_PC_~~6~~7:Idle_Wait to SA_PC_3:Standby

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to STANDBY is processed;
b) a START STOP UNIT command with the POWER CONDITION field set to FORCE_STANDBY_0 is processed; or
c) the Power Condition mode page standby condition timer expires.

10.2.10.2.9.4 Transition SA_PC_~~6~~7:Idle_Wait to SA_PC_4:Stopped

This transition shall occur if:

a) a START STOP UNIT command with the START bit set to zero is processed.

10.2.10.2.9.5 Transition SA_PC_7:Idle_Wait to SA_PC_5:Sleep

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to SLEEP is processed.

10.2.10.2.9.6 Transition SA_PC_~~6~~7:Idle_Wait to SA_PC_~~5~~6:Active_Wait

This transition shall occur if:

a) a START STOP UNIT command with the POWER CONDITION field set to ACTIVE is processed; or
b) a command that requires the active power condition is processed.

If the transition is based on a START STOP UNIT command with the IMMED bit set to zero, the device server shall not complete the command until this state machine reaches the SA_PC_1:Active state.

10.4.3.28 PHY CONTROL function

The PHY CONTROL function requests actions by the specified phy. This SMP function may be implemented by any management device server. In zoning expander devices, if zoning is enabled then this function shall only be processed from SMP initiator ports that have access to zone group 2 or the zone group of the specified phy (see 4.9.3.2).

Table 63 defines the request format.

Table 63 — PHY CONTROL request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (91h)							
2	ALLOCATED RESPONSE LENGTH							
3	REQUEST LENGTH (00h or 09h)							
4	(MSB)	EXPECTED EXPANDER CHANGE COUNT						
5								(LSB)
6	Reserved							
8								
9	PHY IDENTIFIER							
10	PHY OPERATION							
11	Reserved							UPDATE PARTIAL PATHWAY TIMEOUT VALUE
12	Reserved							
23								
24	ATTACHED DEVICE NAME							
31								
32	PROGRAMMED MINIMUM PHYSICAL LINK RATE				Reserved			
33	PROGRAMMED MAXIMUM PHYSICAL LINK RATE				Reserved			
34	Reserved							
35								
36	Reserved				PARTIAL PATHWAY TIMEOUT VALUE			
37	Reserved							
39								
40	(MSB)	CRC						
43								(LSB)

The SMP FRAME TYPE field is defined in 10.4.3.2.2 and shall be set to the value defined in table 63.

The FUNCTION field is defined in 10.4.3.2.3 and shall be set to the value defined in table 63.

The ALLOCATED RESPONSE LENGTH field is defined in 10.4.3.2.4.

If the ALLOCATED RESPONSE LENGTH field is set to 00h, then the management device server shall:

a) set the RESPONSE LENGTH field to 00h in the response frame; and
b) return the first 4 bytes defined in table 66 plus the CRC field as the response frame.

If the ALLOCATED RESPONSE LENGTH field is not set to 00h, then the management device server shall:

a) set the RESPONSE LENGTH field in the response frame to the value defined in table 66 (i.e., 00h); and
b) return the response frame as specified by the ALLOCATED RESPONSE LENGTH field.

NOTE 5 - Future versions of this standard may change the value defined in table 66.

The REQUEST LENGTH field is defined in 10.4.3.2.5 and shall be set to one of the values defined in table 63 based on the LONG RESPONSE bit in the REPORT GENERAL response (see 10.4.3.4). A REQUEST LENGTH field set to 00h specifies that there are 9 dwords before the CRC field.

The EXPECTED EXPANDER CHANGE COUNT field is defined in the SMP CONFIGURE GENERAL request (see 10.4.3.18).

The PHY IDENTIFIER field specifies the phy (see 4.2.10) to which the SMP PHY CONTROL request applies.

Table 64 defines the PHY OPERATION field.

Table 64 — PHY OPERATION field (part 1 of 3)

Code	Operation	Description
00h	NOP	No operation.

[a] Phys compliant with previous versions of this standard did not reject this phy operation due to affiliations.
[b] Phys compliant with previous versions of this standard returned SMP FUNCTION REJECTED.

Table 64 — PHY OPERATION field (part 2 of 3)

Code	Operation	Description
01h	LINK RESET	If: a) a SAS phy is attached; b) a SATA phy is attached and there is no affiliation; or c) a SATA phy is attached and an affiliation exists for the STP initiator port with the same SAS address as the SMP initiator port that opened this SMP connection, then: a) if the specified phy is a physical phy, perform a link reset sequence (see 4.4) on the specified phy and enable the specified phy; and b) if the specified phy is a virtual phy, perform an internal reset and enable the specified phy. If a SATA phy is attached and an affiliation does not exist for the STP initiator port with the same SAS address as the SMP initiator port that opened this SMP connection, then the management device server shall return a function result of AFFILIATION VIOLATION in the response frame (see table 245 in 10.4.3.3). [a] See 7.11 for Broadcast (Change) requirements related to this phy operation in an expander device. Any affiliation (see 7.17.4) shall continue to be present. The phy shall bypass the SATA spinup hold state, if implemented (see 6.8.3.9). The management device server shall return the PHY CONTROL response without waiting for the LINK RESET phy operation to complete. While the LINK RESET phy operation is in progress, the management device server sets the NEGOTIATED PHYSICAL LINK RATE field and the NEGOTIATED PHYSICAL LINK RATE field to RESET_IN_PROGRESS in the SMP DISCOVER response (see 10.4.3.10).
02h	HARD RESET	If the specified phy is a physical phy, perform a link reset sequence (see 4.4) on the specified phy and enable the specified phy. If the attached phy is a SAS phy or an expander phy, the link reset sequence shall include a hard reset sequence (see 4.4.2). If the attached phy is a SATA phy, the phy shall bypass the SATA spinup hold state. See 7.11 for Broadcast (Change) requirements related to this phy operation in an expander device. If the specified phy is a virtual phy, perform an internal reset and enable the specified phy. Any affiliation (see 7.17.4) shall be cleared. The management device server shall return the PHY CONTROL response without waiting for the HARD RESET phy operation to complete. While the HARD RESET phy operation is in progress, the management device server sets the NEGOTIATED PHYSICAL LINK RATE field and the NEGOTIATED PHYSICAL LINK RATE field to RESET_IN_PROGRESS in the SMP DISCOVER response (see 10.4.3.10).

[a] Phys compliant with previous versions of this standard did not reject this phy operation due to affiliations.
[b] Phys compliant with previous versions of this standard returned SMP FUNCTION REJECTED.

Table 64 — PHY OPERATION field (part 3 of 3)

Code	Operation	Description
03h	DISABLE	Disable the specified phy (i.e., stop transmitting valid dwords and receiving dwords on the specified phy). The LINK RESET and HARD RESET operations may be used to enable the phy. See 7.11 for Broadcast (Change) requirements related to this phy operation in an expander device.
04h	~~Reserved~~ SLEEP	Place the specified phy in sleep mode, where it does not transmit anything until its receiver detects COMINIT. The LINK RESET and HARD RESET operations may be used to enable the phy. See 7.11 for Broadcast (Change) requirements related to this phy operation in an expander device.
05h	CLEAR ERROR LOG	Clear the error log counters reported in the REPORT PHY ERROR LOG function (see 10.4.3.11) for the specified phy.
06h	CLEAR AFFILIATION	Clear an affiliation (see 7.17.4) from the STP initiator port with the same SAS address as the SMP initiator port that opened this SMP connection. If there is no such affiliation, the management device server shall return a function result of AFFILIATION VIOLATION [b] in the response frame (see table 245 in 10.4.3.3).
07h	TRANSMIT SATA PORT SELECTION SIGNAL	This function shall only be supported by phys in an expander device. If the expander phy incorporates an STP/SATA bridge and supports SATA port selectors, the phy shall transmit the SATA port selection signal (see 6.6) which causes the SATA port selector to select the attached phy as the active host phy and make its other host phy inactive. See 7.11 for Broadcast (Change) requirements related to this phy operation in an expander device. Any affiliation (see 7.17.4) shall be cleared. If the expander phy does not support SATA port selectors, then the management device server shall return a function result of PHY DOES NOT SUPPORT SATA. If the expander phy supports SATA port selectors but is attached to a SAS phy or an expander phy, the management device server shall return a function result of SMP FUNCTION FAILED in the response frame (see table 245 in 10.4.3.3).
08h	CLEAR STP I_T NEXUS LOSS	The STP I_T NEXUS LOSS OCCURRED bit in the REPORT PHY SATA function (see 10.4.3.12) shall be set to zero.
09h	SET ATTACHED DEVICE NAME	If the expander phy is attached to a SATA phy, set the ATTACHED DEVICE NAME field reported in the DISCOVER response (see 10.4.3.10) to the value of the ATTACHED DEVICE NAME field in the PHY CONTROL request.
All others	Reserved	

[a] Phys compliant with previous versions of this standard did not reject this phy operation due to affiliations.
[b] Phys compliant with previous versions of this standard returned SMP FUNCTION REJECTED.

Editor's Note 23: Interface power management will be invisible to the SCSI layers, but a device going to sleep should be visible (since there is no automatic wakeup). Add 04 SLEEP to 7.11 and elsewhere.

If the operation specified by the PHY OPERATION field is unknown, the management device sever shall return a function result of SMP FUNCTION FAILED in the response frame (see table 245 in 10.4.3.3) and not process any other fields in the request.

If the PHY IDENTIFIER field specifies the phy which is being used for the SMP connection and a phy operation of LINK RESET, HARD RESET, or DISABLE is requested, the management device server shall not perform the requested operation and shall return a function result of SMP FUNCTION FAILED in the response frame (see table 245 in 10.4.3.3).

An UPDATE PARTIAL PATHWAY TIMEOUT VALUE bit set to one specifies that the PARTIAL PATHWAY TIMEOUT VALUE field shall be honored. An UPDATE PARTIAL PATHWAY TIMEOUT VALUE bit set to zero specifies that the PARTIAL PATHWAY TIMEOUT VALUE field shall be ignored.

The ATTACHED DEVICE NAME field is used by the SET ATTACHED DEVICE NAME phy operation and is reserved for all other phy operations. If a management application client detects the ATTACHED DEVICE NAME field set to zero in the DISCOVER response when a SATA device is attached, it shall set the ATTACHED DEVICE NAME field based on the IDENTIFY (PACKET) DEVICE data retrieved by an ATA application client in the same SAS initiator device as follows:

a) if IDENTIFY (PACKET) DEVICE data word 255 (i.e., the Integrity word) is correct and words 108-111 (i.e., the World Wide Name field) are not set to zero, set this field to the world wide name indicated by words 108-111 according to table 13 in 4.2.7;
b) if IDENTIFY (PACKET) DEVICE data word 255 (i.e., the Integrity word) is correct and words 108-111 (i.e., the World Wide Name) are set to zero, set this field to 00000000 00000000h; or
c) if IDENTIFY (PACKET) DEVICE data word 255 (i.e., the Integrity word) is not correct, set this field to 00000000 00000000h.

The PROGRAMMED MINIMUM PHYSICAL LINK RATE field specifies the minimum physical link rate the phy shall support during a link reset sequence (see 4.4.1). Table 65 defines the values for this field. This value is reported in the DISCOVER response (see 10.4.3.10). If this field is changed along with a phy operation of LINK RESET or HARD RESET, that phy operation shall utilize the new value for this field.

The PROGRAMMED MAXIMUM PHYSICAL LINK RATE field specifies the maximum physical link rates the phy shall support during a link reset sequence (see 4.4.1).Table 65 defines the values for this field. This value is reported in the DISCOVER response (see 10.4.3.10). If this field is changed along with a phy operation of LINK RESET or HARD RESET, that phy operation shall utilize the new value for this field.

Table 65 — PROGRAMMED MINIMUM PHYSICAL LINK RATE and PROGRAMMED MAXIMUM PHYSICAL LINK RATE fields

Code	Description
0h	Do not change current value
1h - 7h	Reserved
8h	1.5 Gbps
9h	3 Gbps
Ah	6 Gbps
Bh - Fh	Reserved for future physical link rates

If the PROGRAMMED MINIMUM PHYSICAL LINK RATE field or the PROGRAMMED MAXIMUM PHYSICAL LINK RATE field is set to an unsupported or reserved value, or the PROGRAMMED MINIMUM PHYSICAL LINK RATE field and PROGRAMMED MAXIMUM PHYSICAL LINK RATE field are set to an invalid combination of values (e.g., the minimum is greater than the maximum), the management device server shall not change either of their values and may return a function result of SMP FUNCTION FAILED in the response frame (see table 245 in 10.4.3.3). If it returns a function result of SMP FUNCTION FAILED, it shall not perform the requested phy operation.

The PARTIAL PATHWAY TIMEOUT VALUE field specifies the amount of time in microseconds the expander phy shall wait after receiving an Arbitrating (Blocked On Partial) confirmation from the ECM before requesting that the ECM resolve pathway blockage (see 7.12.4.5). A PARTIAL PATHWAY TIMEOUT VALUE field value of zero (i.e., 0 µs) specifies that partial pathway resolution shall be requested by the expander phy immediately upon reception of an Arbitrating (Blocked On Partial) confirmation from the ECM. This value is reported in the DISCOVER response (see 10.4.3.10). The PARTIAL PATHWAY TIMEOUT VALUE field is only honored when the UPDATE PARTIAL PATHWAY TIMEOUT VALUE bit is set to one.

The CRC field is defined in 10.4.3.2.8.

Table 66 defines the response format.

Table 66 — PHY CONTROL response

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (91h)							
2	FUNCTION RESULT							
3	RESPONSE LENGTH (00h)							
4	(MSB)	CRC						
7								(LSB)

The SMP FRAME TYPE field is defined in 10.4.3.3.2 and shall be set to the value defined in table 66.

The FUNCTION field is defined in 10.4.3.3.3 and shall be set to the value defined in table 66.

The FUNCTION RESULT field is defined in 10.4.3.3.4.

The RESPONSE LENGTH field is defined in 10.4.3.3.5 and shall be set to the value defined in table 66.

The CRC field is defined in 10.4.3.3.8.

Background - Sleep power condition definition in MMC-6, RBC, SBC-2 revision 7, and ATA8-ACS

Background: MMC-6 excerpts about sleep power condition

4.1.10 Power Management

Table 12 — POWER CONDITION Power Conditions

Mode	Description
Sleep	The lowest power consumption, with power applied, occurs in the Sleep condition. When in the Sleep condition a MM Drive requires a WAKEUP task management function to be activated
Standby	In the Standby condition a MM Drive is capable of accepting commands, but media is not immediately accessible (e.g., the spindle is stopped).
Idle	In the Idle condition a MM Drive is capable of responding quickly to media access requests. However, a MM Drive in the idle condition may take longer, than in the active condition, to complete the execution of a command because it may have to activate some circuitry.
Active	In the Active condition a MM Drive is capable of responding immediately to media access requests, and operations complete execution in the shortest time compared to the other power conditions.

6.39.2.4 Power conditions

Table 634 — POWER CONDITION field

Code	Description
5h	Place Drive into Sleep State. Immed has no meaning when sleep state is requested. Before entering the sleep state, all buffers shall be successfully flushed by the Drive. If the sleep command is successful, the Host should not issue new commands after receiving the successful completion status. The Device shall de-power and disable the interface only after all Drives have successfully completed the sleep operation.

6.39.3.2 Power Condition Changes

When the Drive enters the sleep state, any queued GET EVENT/STATUS NOTIFICATION commands shall be removed from the command queue without command completion.

If any commands other then GET EVENT STATUS NOTIFICATION are in the command queue when a the sleep requested, the command shall be terminated with CHECK CONDITION status and sense bytes SK/ASC/ASCQ shall be set to ILLEGAL REQUEST/COMMAND SEQUENCE ERROR.

If a request to go to a power state fails, the Drive shall remain in the current power state and shall generate power management class event with the Power Event Field set to PwrChg-Fail.

All power state change requests, except sleep, that complete successfully shall generate a power management class event with the Power Event field set to PwrChg-Successful.

Notification of power states shall occur upon entering a new power state.

B.3.5 Power Management and Device Reset in SCSI

When a SCSI Device is in the Power Managed Sleep state, a reset through the service delivery subsystem are used to wake the device.

H.1 Power Management States

... Sleep defines the “Off” state. ...

Table 635 — H.1 Power Management Model States

State	Power Consumption	Drive Context Retained	Restore time
Active	As needed for operation	All	None
Idle	Less than Active	All	The Drive should be restored to active state within 1 second on any request to enter active state, independent of the de-powering process.
Standby	Less than idle	All buffers are empty before entering Standby state.	Vendor specific: Greater than or equal to Idle to Active
Sleep	Less than Standby	None, Buffer & All of command queues are empty before entering Sleep state.	Greater than or equal to Standby to Active. Vendor Specific. May Need full initialization. The Initiator may remove Vcc.

H.2 Power State Transitions

Sleep State:

Maximum power saving state. Buffers and all command queues, including GET EVENT STATUS NOTIFICATION commands, should be emptied before entering into the Sleep state. When the Drive enters the sleep state, any GET EVENT STATUS NOTIFICATION commands present in the command queue, should be removed from the command queue, without command completion. In this Sleep state, all functions are stopped and no commands, except for reset may be received. The unit is consuming less power than when in the Standby state. The Drive context is invalid in the Sleep state.

The Initiator software should fully initialize the Drive after exiting Sleep state, as all context may be lost in the Sleep state. Most devices provide a manual eject mechanism for removing/inserting a disc independent of any lock/unlock mechanism employed. Given this possibility, when the Logical Unit is unable to determine if media has been changed while the Drive was in the sleep state, the Drive should report NEW MEDIA on the next GET EVENT STATUS NOTIFICATION (Media Status) command.

In the Sleep state, the Initiator may completely remove power from the device by turning off Vcc.

...

A power-on or hard reset always returns the Power State to the Standby State. A Device Reset does not alter the current Power State, unless the current Power State is Sleep. A Device Reset received while in sleep state returns the Power State to Standby.

The Sleep state is entered when the Drive has been commanded to go to Sleep but Vcc is still applied to the device. Removing Vcc always takes the device to the Power Off State. Removing Vcc is recommended only when all Drives on a given bus are in Sleep State.

Background: RBC revision 10a excerpts about the sleep power condition

4.3.1 START STOP UNIT command state restrictions

A removable medium device shall be in either PREVENT state 00b or 10b in order to successfully execute a START STOP UNIT command with thePOWER CONDITIONS field set to the Sleep state (5).

If a removable medium device, in either PREVENT state 01b or 11b, receives a START STOP UNIT command with the POWER CONDITIONS field set to the Sleep state (5), the device shall respond with status set to CHECK CONDITION (02h), the sense key to ILLEGAL REQUEST (05h), and the ASC/ASCQ to ILLEGAL POWER CONDITION REQUEST (2Ch/05h).

A removable medium device in the SLEEP State shall eject the media without causing the media to spin up in accordance with the PREVENT/ALLOW MEDIUM REMOVAL command requirements.

Refer to clause 0 for a description of the POWER CONDITIONS field values.

5.4.1 Power Conditions

Table 8 — Power Conditions

Code		Description
5	M	Place device in Sleep condition

Sleep (condition 5): Devices in the Sleep state are at a lower power consumption level than when in the Standby condition and have very little of the drive circuitry consuming power. A device reset may be required before access to the device is allowed. Prior to entering the Sleep state the device shall ensure that logical blocks in cache have their most recent data value recorded on the physical medium.

7.1.1 Power condition change notification

RBC devices shall notify an initiator of the intent to change power conditions via asynchronous event notification. The status value shall be set to CHECK CONDITION (02h), the sense key to UNIT ATTENTION (06h), and the ASC to POWER CONDITION CHANGE NOTIFICATION (5Eh). The ASCQ shall be set to the value of the new power condition plus 40h as shown in Table 22.

Table 22 — Power condition sense code and qualifier values

ASC	ASCQ	Description
5Eh	45	POWER CONDITION CHANGE TO SLEEP

Background: SBC-2 revision 7 excerpts about the sleep power condition

If the START STOP UNIT command is issued with the POWER CONDITION field set to 5h the device server shall:

a) suspend any Power Condition timers that are active on receipt of the START STOP UNIT command until a wakeup;
b) not respond to commands and task management functions until a wakeup.

On receipt of a wakeup any previously active power condition timers shall be restored to those values indicated by the saved Power Condition mode page parameters. Before returning a function complete response the target port shall place itself into a condition capable of receiving commands and task management functions and shall create a unit attention condition for all initiators. The sense key shall be set to UNIT ATTENTION with the additional sense code set to LOW POWER CONDITION ACTIVE.

...

In the sleep power condition the device server shall only respond to a wakeup. When a target port has multiple logical units attached it shall enter the sleep power condition only after all the logical units have been placed into a sleep power condition.

Background: ATA8-ACS revision 4c excerpts about the sleep power condition

4.18.2 Power management commands

...

The SLEEP command moves a device to Sleep mode. The device's interface becomes inactive after the device reports command completion for the SLEEP command. A device only transitions from Sleep mode after processing a hardware reset, a software reset, or a DEVICE RESET command.

4.18.4 Power modes

...

Transition PM0:PM3: When a SLEEP command is received, the device shall make a transition to the PM3:Sleep state.

Transition PM1:PM3: When a SLEEP command is received, the device shall make a transition to the PM3:Sleep state.

Transition PM2:PM3: When a SLEEP command is received, the device shall make a transition to the PM3:Sleep state.

PM3: Sleep: This state shall be entered when the device receives a SLEEP command.

A device transitions from Sleep mode only after processing a hardware reset, a software reset, or a DEVICE RESET command. The time to respond may be as long as 30 s. Sleep state provides the lowest power consumption of any state.

In Sleep state, see the applicable transport standard for a description of the device's interface behavior.
Transition PM3:PM2:, A device shall transition to the PM2:Standby state after processing a hardware reset, software reset, or DEVICE RESET command.

7.52 SLEEP - E6h, Non-Data

7.52.1 Feature Set

This 28-bit command is mandatory for devices implementing the Power Management feature set.

7.52.2 Description

This command is the only way to cause the device to enter Sleep mode. The device shall exit Sleep (i.e., State PM3) only after processing a hardware reset, a software reset, or a DEVICE RESET command.

A device shall not power-on in Sleep state.