Date: December 14, 2007

To: T10 Committee (SCSI)

From: George Penokie (IBM)

Subject: SAS-2: Limiting SAS Target response to OPEN_REJECT (RETRY)

1 Overview

In a SAS topology with a large imbalance between the number of initiators and targets (e.g., one or two initiators with 50 or more targets) with a heavy workload it is not unusual for the minority SAS device to become bandwidth limited. When this happens the bandwidth limited SAS device will start issuing OPEN_REJECT (RETRY) to most open attempts. In response to the OPEN_REJECT (RETRY) many SAS devices will immediately issue another OPEN. This causes a flood of OPEN/OPEN_REJECT (RETRY)s which accomplishes nothing except to add congestion to an already congested topology.

Revision 1 - Moves the timer to the Shared Port Control mode page for SSP and the CONFIGURE GENERAL function for SMP and STP. Adds the timer description into the port layer specifically the PL_OC state machine.

Revision 2 - Moves the timer to the Protocol-Specific Port mode page for SSP and the REPORT GENERAL function for STP. Removed the option from SMP.

2 Proposed SAS-2 changes

8.2 PL (port layer) state machines

8.2.1 PL state machines overview

The PL (port layer) consists of state machines that run in parallel and perform the following functions:

   a) receive requests from the SSP, SMP, and STP transport layer state machines for connection
       management (e.g., requests to open or close connections) and frame transmission;
   b) send requests to the SAS link layer state machines for connection management and frame
       transmission;
   c) receive confirmation from the SAS link layer state machines; and
   d) send confirmations to the SSP, SMP, and STP transport layer state machines.

The port layer state machines are as follows:

   a) PL_OC (port layer overall control) state machines (see 8.2.2); and
   b) PL_PM (port layer phy manager) state machines (see 8.2.3).

There is one PL_OC state machine per port (see 4.1.4). There is one PL_PM state machine for each phy contained in the port. Phys are assigned to ports by the management application layer. More than one port in a SAS device may have the same SAS address.

Figure 1 shows examples of the port layer state machines and their interaction with the transport and link layers.
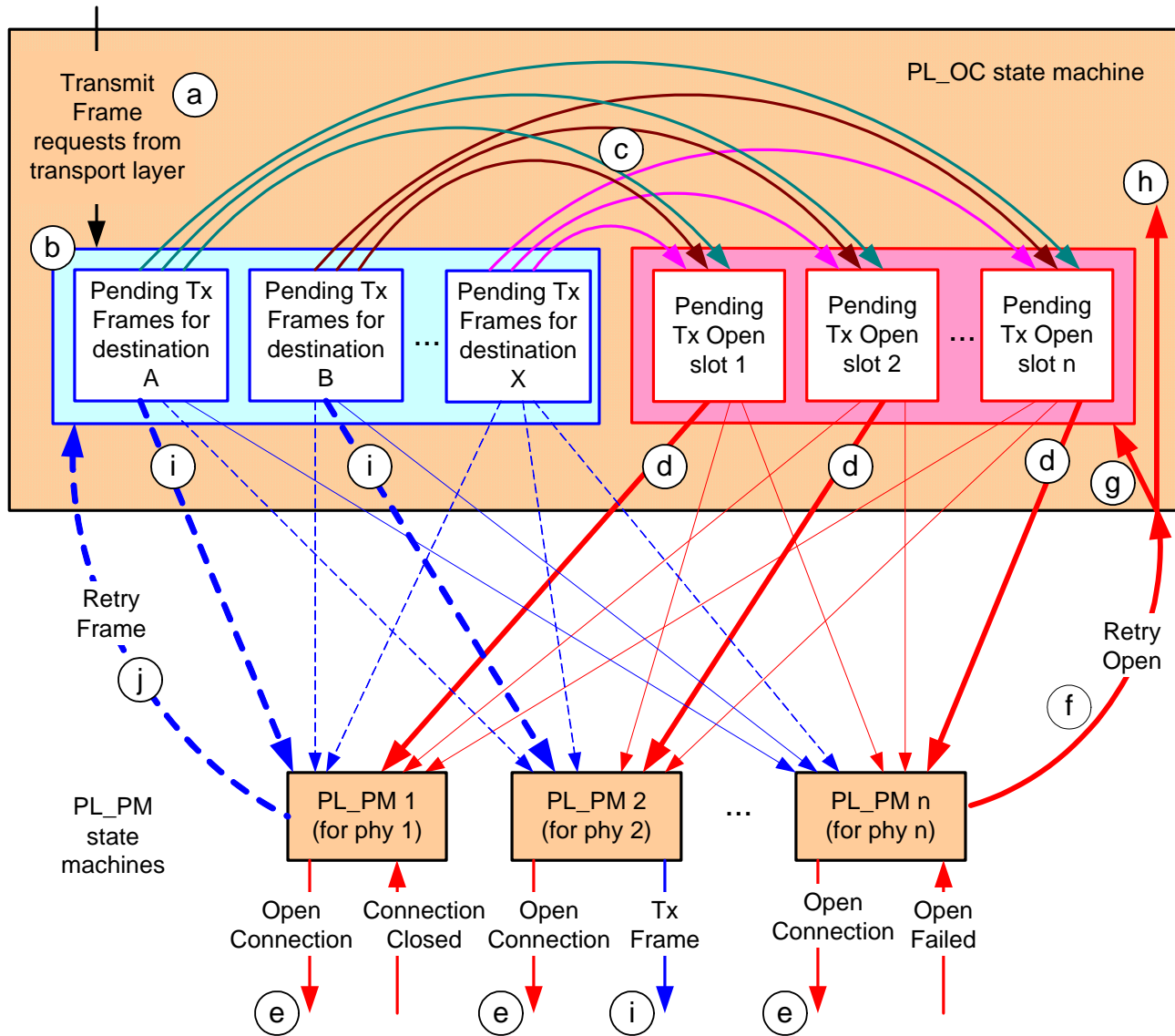

Figure 1 — Port layer examples

The following is a description of the example processes in figure 1. These example processes do not describe all of the possible condition or actions.

a) Transmit Frame requests are received by the PL_OC state machine;

b) the PL_OC state machine converts Transmit Frame requests into pending Tx Frame messages associated with the destination SAS address;

c) the PL_OC state machine generates a pending Tx Open message for a pending Tx Frame message when there is a pending Tx Open slot available (i.e., the number of pending Tx Open messages is less than or equal to the number of phys);

d) the PL_OC state machine sends a pending Tx Open message as a Tx Open message to a PL_PM state machine when a PL_PM machine is available; a slot is then available for a new pending Tx Open message;

e) when a PL_PM state machine receives a Tx Open message, the PL_PM state machine attempts to establish a connection with the destination SAS address through the link layer;

f) if a PL_PM state machine is unable to establish a connection with the destination SAS address, then the PL_PM state machine sends a Retry Open message to the PL_OC state machine;

g) the PL_OC state machine converts a Retry Open message to a pending Tx Open message if there is a pending Tx Open slot available; if the PL_OC state machine converts a Retry Open message into a pending Tx Open message, then the pathway blocked count and arbitration wait time context from the Retry Open message are applied to the pending Tx Open message;

h) after the reject to open limit timer, if any, has expired, and if there is a pending Tx Open slot available, then the PL_OC state machine converts a Retry Open message to a pending Tx Open message with the pathway blocked count and arbitration wait time context from the Retry Open message applied to the pending Tx Open message;

i) if the PL_OC state machine does not convert a Retry Open to a pending Tx Open frame, then the PL_OC discards the Retry Open message. The PL_OC state machine may create a new Tx Open message for the same pending Tx Frame at a later time. If the PL_OC state machine discards a Retry Open message, then the pathway blocked count and arbitration wait time context from the Retry Open message are also discarded;

j) after a PL_PM state machine establishes a connection with a destination SAS address, the PL_OC state machine sends pending Tx Frame messages for the destination to the PL_PM state machine as Tx Frame messages;

k) if a PL_PM state machine is unable to send a Tx Frame message to the link layer as a Tx Frame request (e.g., due to a credit timeout), then the PL_PM state machine sends a Retry Frame message to the PL_OC state machine, and the PL_OC state machine converts the Retry Frame message into a pending Tx Frame message; and

l) if the PL_PM state machine is able to send a Tx Frame message as a Tx Frame request to the link layer, then the PL_PM state machine sends a Transmission Status confirmation to the transport layer.

The Transmission Status confirmation from either the PL_OC state machine or a PL_PM state machine shall include the following as arguments:

a) tag;
b) destination SAS address; and
c) source SAS address.

8.2.2 PL_OC (port layer overall control) state machine

8.2.2.1 PL_OC state machine overview

A PL_OC state machine:

a) receives requests from the SSP, SMP, and STP transport layers;
b) sends messages to the PL_PM state machine;
c) receives messages from the PL_PM state machine;
d) selects frames to transmit;
e) selects phys on which to transmit frames;
f) receives confirmations from the link layer;
g) sends confirmations to the transport layer;
h) has Arbitration Wait Time timers; and
i) has I_T Nexus Loss timers; and
j) may have Reject To Open Limit timers.

This state machine consists of the following states:

a) PL_OC1:Idle (see 8.2.2.2) (initial state); and
b) PL_OC2:Overall_Control (see 8.2.2.3).

After power on this state machine shall start in the PL_OC1:Idle state.

The PL_OC state machine shall maintain:

a) a pool of pending Tx Frame messages for each destination SAS address; and
b) as many pending Tx Open message slots as there are phys in the port.

The PL_OC state machine shall maintain the timers listed in table 1.

Table 1 — PL_OC state machine timers

| Timer	Maximum number of timers	Initial value
I_T Nexus Loss timer	One per destination SAS address	Depending on the protocol used by the port: a) for SSP target ports, the value in the I_T NEXUS LOSS TIME field in the Protocol-Specific Port mode page (see 10.2.7.4); b) for SSP initiator ports, the value in the I_T NEXUS LOSS TIME field in the Protocol-Specific Port mode page for the SSP target port with that destination SAS address (see 10.2.7.4); c) for STP target ports, the value in the STP SMP I_T NEXUS LOSS TIME field in the SMP CONFIGURE GENERAL function (see 10.4.3.18); d) for STP initiator ports, the value in the STP SMP I_T NEXUS LOSS TIME field in the SMP REPORT ~~CONFIGURE~~ GENERAL function (see 10.4.3.18) for the STP target port with that destination SAS address; or e) for SMP initiator ports, the value in the STP SMP I_T NEXUS LOSS TIME field in the SMP REPORT ~~CONFIGURE~~ GENERAL function (see 10.4.3.18).
Arbitration Wait Time timer	One per pending Tx Open message	0000h, a vendor-specific value less than 8000h (see 7.12.3), or the value received with a Retry Open message.
Reject To Open Delay timer	One per Retry Open message	Depending on the protocol used by the port: a) for SSP target ports, the value in the REJECT TO OPEN DELAY field in the Shared Port Control mode page (see 10.2.7.6); b) for SSP initiator ports, a vendor specific value; c) for STP target ports, the value in the STP REJECT TO OPEN DELAY field in the SMP CONFIGURE GENERAL function (see 10.4.3.18); or d) for STP initiator ports, a vendor specific value.

Figure 2 shows the PL_OC state machine.
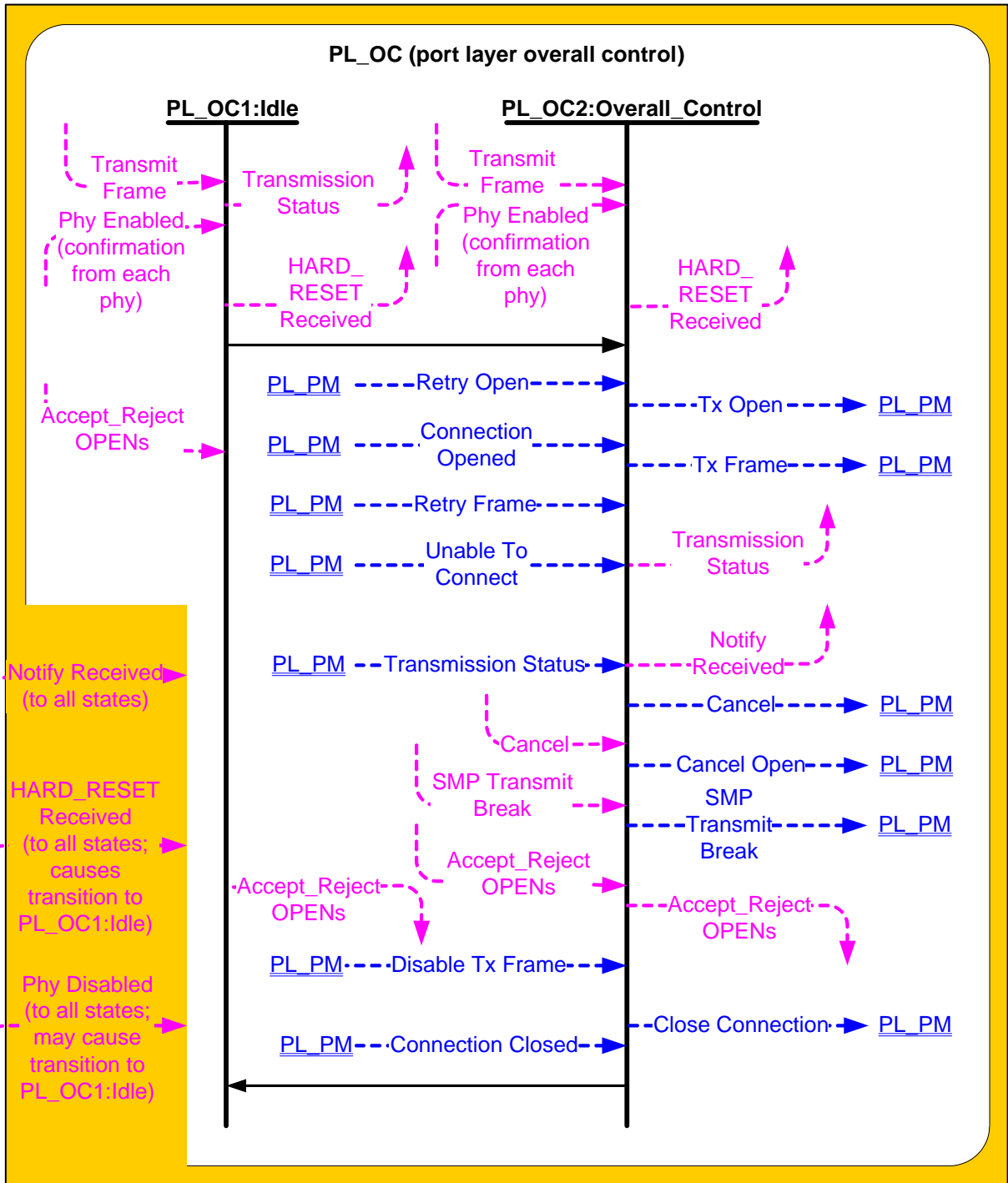
Figure 2 — PL_OC (port layer overall control) state machine

8.2.2.2 PL_OC1:Idle state

8.2.2.2.1 PL_OC1:Idle state description

This state is the initial state of the PL_OC state machine.

If this state receives a HARD_RESET Received confirmation, then this state shall send a HARD_RESET Received confirmation to the transport layer.

If this state receives a Notify Received (Power Loss Expected) confirmation, then this state shall send a Notify Received (Power Loss Expected) confirmation to the transport layer.

If this state receives an Accept_Reject Opens request, then this state shall send an Accept_Reject Opens request to all link layers in the port.

If this state receives a Transmit Frame request, then this state shall send a Transmission Status (No Phys In Port) confirmation to the transport layer.

If an I_T Nexus Loss timer expires for a destination SAS address, this state shall perform the following:

a) delete the I_T Nexus Loss timer for the SAS address;
b) send a Transmission Status (I_T Nexus Loss) confirmation for each pending Tx Frame message for the SAS address; and
c) discard each pending Tx Frame message for the SAS address and any corresponding pending Tx Open messages.

If the port is an STP target port or an STP initiator port, the port shall handle all pending commands as described in 4.5.

8.2.2.2.2 Transition PL_OC1:Idle to PL_OC2:Overall_Control

This transition shall occur after a Phy Enabled confirmation is received for at least one phy assigned to the port.

8.2.2.3 PL_OC2:Overall_Control state

8.2.2.3.1 PL_OC2:Overall_Control state overview

This state may receive Transmit Frame requests from the transport layers (i.e., SSP and SMP) and Retry frame messages from PL_PM state machines. This state shall create a pending Tx Frame message for each received Transmit Frame request and Retry Frame message. There may be more than one pending Tx Frame message at a time for each SSP transport layer. There shall be only one pending Tx Frame message at a time for each SMP transport layer.

This state selects PL_PM state machines through which connections are established. This state shall only attempt to establish connections through PL_PM state machines whose phys are enabled. In a vendor-specific manner, this state selects PL_PM state machines on which connections are established to transmit frames. This state shall receive a response to a message from a PL_PM state machine before sending another message to that PL_PM state machine.

This state also:

a) receives connection management requests from the transport layers;
b) sends connection management messages to PL_PM state machines;
c) receives connection management messages from PL_PM state machines; and
d) sends connection management confirmations to the transport layers.

After receiving a Transmit Frame request for a destination SAS address for which there is no connection established and for which no I_T Nexus Loss timer has been created, this state shall create an I_T Nexus Loss timer for that SAS address if:

a) the protocol is SSP, the port is an SSP target port, the Protocol-Specific Port mode page is implemented, and the I_T NEXUS LOSS TIME field in the Protocol-Specific Port mode page (see 10.2.7.4) is not set to 0000h;
b) the protocol is STP, the port is an STP target port, and the STP SMP I_T NEXUS LOSS TIME field in the SMP CONFIGURE GENERAL function is not set to 0000h; or
c) the protocol is SMP, the port is an SMP initiator port, and the STP SMP I_T NEXUS LOSS TIME field in the SMP CONFIGURE GENERAL function is not set to 0000h.

This state may create an I_T Nexus Loss timer for that SAS address if:

a) the protocol is SSP and the port is an SSP initiator port; or;
b) the protocol is STP and the port is an STP initiator port.

When this state creates an I_T Nexus Loss timer it shall:

1) initialize the I_T Nexus Loss timer as specified in table 1 (see 8.2.2.1); and
2) not start the I_T Nexus Loss timer.

If this state machine is in an SSP initiator port, then this state may create an I_T Nexus Loss timer for the SAS address. If a state machine in an SSP initiator port and creates an I_T Nexus Loss timer, then the state machine should use the value in the I_T NEXUS LOSS TIME field in the Protocol-Specific Port mode page for the SSP target port (see 10.2.7.4) as the initial value for its I_T Nexus Loss timer.

If there are no pending Tx Frame messages for a destination SAS address and an I_T Nexus Loss timer has been created for that destination SAS address, then this state shall delete the I_T Nexus Loss timer for that destination SAS address.

If this state receives a HARD_RESET Received confirmation, then this state shall discard all pending Tx Frame messages and delete all I_T Nexus Loss timers and send a HARD_RESET Received confirmation to the transport layer.

If this state receives a Notify Received (Power Loss Expected) confirmation, then this state shall:

a) discard all pending Tx Frame messages, if any;
b) delete all I_T Nexus Loss timers, if any;
c) send a Close Connection message to all the PL_PM state machines;
d) send a Cancel Open message to all the PL_PM state machines; and
e) send a Notify Received (Power Loss Expected) confirmation to the transport layer.

8.2.2.3.2 PL_OC2:Overall_Control state establishing connections

This state receives Phy Enabled confirmations indicating when a phy is available.

This state receives Retry Open messages from a PL_PM state machine.

This state creates pending Tx Open messages based on pending Tx Frame messages and Retry Open messages. Pending Tx Open messages are sent to a PL_PM state machine as Tx Open messages.

If this state receives a Retry Open (Retry) message, then this state shall process the Retry Open message.

If this state receives a Retry Open (No Destination) or a Retry Open (Open Timeout Occurred) message and an I_T Nexus Loss timer has not been created for the destination SAS address (e.g., an SSP target port does not support the I_T NEXUS LOSS TIME field in the Protocol-Specific Port mode page or the field is set to 0000h), then this state shall process the Retry Open message as either a Retry Open message or an Unable To Connect message. This selection is vendor-specific.

If this state receives a Retry Open (Pathway Blocked) message and an I_T Nexus Loss timer has not been created for the destination SAS address, then this state shall process the Retry Open message.

If this state receives a Retry Open (No Destination), Retry Open (Open Timeout Occurred), or Retry Open (Pathway Blocked) message, and an I_T Nexus Loss timer has been created for the destination SAS address with an initial value of FFFFh, then this state shall process the Retry Open message (i.e., the Retry Open message is never processed as an Unable to Connect message).

If this state receives a Retry Open (No Destination) or a Retry Open (Open Timeout Occurred) message, an I_T Nexus Loss timer has been created for the destination SAS address, and there is no connection established with the destination SAS address, then this state shall check the I_T Nexus Loss timer, and:

a) if the I_T Nexus Loss timer is not running, the I_T nexus loss time is not set to FFFFh, and the CONFIGURING bit is set to zero in the REPORT GENERAL response (see 10.4.3.3) for each expander device between this port and the destination port that is two or more levels away from this port, then this state shall start the timer;
b) if the I_T Nexus Loss timer is not running and the I_T nexus loss time is not set to FFFFh, then this state shall start the timer;
c) if the I_T Nexus Loss timer is running, then this state shall not stop the timer; and
d) if the I_T Nexus Loss timer has expired, then this state shall process the Retry Open message as if it were an Unable To Connect message (see 8.2.2.3.4).

If this state receives a Retry Open (Pathway Blocked) message, an I_T Nexus Loss timer has been created for the destination SAS address, and there is no connection established with the destination SAS address, then this state shall check the I_T Nexus Loss timer, and:

a) if the I_T Nexus Loss timer is running, then this state shall not stop the timer; and
b) if the I_T Nexus Loss timer has expired, then this state shall process the Retry Open message as if it were an Unable To Connect message (see 8.2.2.3.4).

If this state receives a Retry Open (Retry) and an I_T Nexus Loss timer is running for the destination SAS address, then this state shall:

a) stop the I_T Nexus Loss timer (if the timer has been running); and
b) initialize the I_T Nexus Loss timer.

This state shall create a pending Tx Open message if:

a) this state has a pending Tx Frame message or has received a Retry Open message;
b) this state has fewer pending Tx Open messages than the number of PL_PM state machines (i.e., the number of phys in the port);
c) there is no pending Tx Open message for the destination SAS address; and
d) there is no connection established with the destination SAS address.

This state may create a pending Tx Open message if:

a) this state has a pending Tx Frame message, or this state has received a Retry Open message and has not processed the message by sending a confirmation; and
b) this state has fewer pending Tx Open messages than the number of PL_PM state machines.

This state shall have no more pending Tx Open messages than the number of PL_PM state machines.

If this state receives a Retry Open message and there are pending Tx Frame messages for which pending Tx Open messages have not been created, then this state should create a pending Tx Open message from the Retry Open message.

If this state does not create a pending Tx Open message from a Retry Open message (e.g., the current number of pending Tx Open messages equals the number of phys), then this state shall discard the Retry Open message. This state may create a new pending Tx Open message at a later time for the pending Tx Frame message that resulted in the Retry Open message.

If this state receives a Retry Open (Opened By Destination) message and the initiator port bit and protocol arguments match those in the Tx Open messages that resulted in the Retry Open message, then this state may discard the Retry Open message and use the established connection to send pending Tx Frame messages as Tx Frame messages to the destination SAS address. If this state receives a Retry Open (Opened By Destination) message, then, if this state has a pending Tx Open slot available, this state may create a pending Tx Open message from the Retry Open message.

> NOTE 1 - If a connection is established by another port as indicated by a Retry Open (Opened By Destination) message, credit may not be granted for frame transmission. In this case this state may create a pending Tx Open message from a Retry Open message in order to establish a connection where credit is granted.

This state shall send a pending Tx Open message as a Tx Open message to a PL_PM state machine that has an enabled phy and does not have a connection established. If there is more than one pending Tx Open message, this state should send a Tx Open message for the pending Tx Open message that has been pending for the longest time first.

If this state creates a pending Tx Open message from one of the following messages:

a) a Retry Open (Opened By Destination);
b) a Retry Open (Opened By Other);
c) a Retry Open (Collided); or
d) a Retry Open (Pathway Blocked),

then this state shall:

1) create an Arbitration Wait Time timer for the pending Tx Open message;
2) set the Arbitration Wait Time timer for the pending Tx Open message to the arbitration wait time argument from the Retry Open message; and
3) start the Arbitration Wait Time timer for the pending Tx Open message.

When a pending Tx Open message is sent to a PL_PM state machine as a Tx Open message, the Tx Open message shall contain the following arguments to be used in an OPEN address frame:

a) initiator port bit from the Transmit Frame request;
b) protocol from the Transmit Frame request;
c) connection rate from the Transmit Frame request;
d) initiator connection tag from the Transmit Frame request;
e) destination SAS address from the Transmit Frame request;
f) source SAS address from the Transmit Frame request;
g) pathway blocked count; and
h) arbitration wait time.

If this state creates a pending Tx Open message from one of the following:

a) a Transmit Frame request;
b) a Retry Open (No Destination) message;
c) a Retry Open (Open Timeout Occurred) message; or
d) a Retry Open (Retry) message,

then this state shall:

a) set the pathway blocked count argument in the Tx Open message to zero; and
b) set the arbitration wait time argument in the Tx Open message to zero or a vendor-specific value less than 8000h.

If a pending Tx Open message was created as the result this state receiving a Retry Open (Retry) message and;

a) the protocol for the connection is SSP, the Shared Port Control mode page is implemented, and REJECT TO OPEN DELAY field in the Shared Port Control mode page (see 10.2.7.6) is not set to zero; or
b) the protocol for the connection is STP and the STP REJECT TO OPEN DELAY field is not set to zero in the SMP REPORT GENERAL response (see 10.4.3.3),

then, this state shall:

1) create a Reject To Open Limit timer associated with the pending Tx Open message that received the Retry Open (Retry) message;
2) initialize the Reject To Open Limit timer as specified in table 1;
3) start the Reject To Open Limit timer; and
4) wait at least until the Reject To Open Limit timer expires before sending a Tx Open message.

If a pending Tx Open message was created as the result this state receiving a Retry Open (Pathway Blocked) message, then this state shall set the pathway blocked count argument in the Tx Open message to the value of the pathway blocked count argument received with the message plus one, unless the pathway blocked count received with the argument is FFh.

If a pending Tx Open message was created as the result of this state receiving one of the following:

a) a Retry Open (Opened By Destination) message;
b) a Retry Open (Opened By Other) message;
c) a Retry Open (Collided) message; or
d) a Retry Open (Pathway Blocked) message;

then this state shall set the arbitration wait time argument in the Tx Open message to be the value from the Arbitration Wait Time timer created as a result of the Retry Open message.

After this state sends a Tx Open message, this state shall discard the pending Tx Open message from which the Tx Open messages was created. After this state discards a pending Tx Open message, this state may create a new pending Tx Open message.

If this state receives a Connection Opened message and the initiator port bit and protocol arguments match those in a pending Tx Open message, then:

   a) if credit is granted, then any pending Tx Open message associated with the established connection may be discarded; and
   b) any Reject To Open Limit timer associated with that pending Tx Open message shall be discarded.

If this state receives a Connection Opened message and the initiator port bit and protocol arguments match those in any pending Tx Frame messages, then this state may use the established connection to send pending Tx Frame messages as Tx Frame messages to the destination SAS address.

10.2.7.4 Protocol-Specific Port mode page

The Protocol-Specific Port mode page (see SPC-4) contains parameters that affect SSP target port operation. If the mode page is implemented by one logical unit in a SCSI target device, it shall be implemented by all logical units in the SCSI target device that support the MODE SELECT or MODE SENSE commands.

The mode page policy (see SPC-4) for this mode page shall be either shared or per target port. If a SAS target device has multiple SSP target ports, the mode page policy should be per target port.

Parameters in this mode page:

   a) shall affect all phys in the SSP target port if the mode page policy is per target port; and
   b) shall affect all SSP target ports in the SAS target device if the mode page policy is shared.

Table 2 defines the format of the page for SAS SSP.

Table 2 — Protocol-Specific Port mode page for SAS SSP

Byte\Bit	7	6	5	4	3	2	1	0
0	PS	SPF (0b)	PAGE CODE (19h)					
1	PAGE LENGTH (~~06h~~0Eh)							
2	Reserved	CONTINUE AWT	BROADCAST ASYNCHRONOUS EVENT	READY LED MEANING	PROTOCOL IDENTIFIER (6h)			
3	Reserved							
4	(MSB)	I_T NEXUS LOSS TIME						
5								(LSB)
6	(MSB)	INITIATOR RESPONSE TIMEOUT						
7								(LSB)
8		Reserved						
9								
10	(MSB)	REJECT TO OPEN LIMIT						
11								(LSB)
12		Reserved						
15								

The PARAMETERS SAVEABLE (PS) bit is defined in SPC-4.

The SUBPAGE FORMAT (SPF) bit shall be set to zero for access to this mode page.

The PAGE CODE field shall be set to 19h.

The PAGE LENGTH field shall be set to the number of bytes in the page after the PAGE LENGTH field (i.e., 06h).

A CONTINUE AWT bit set to one specifies that the a SAS port shall not stop the Arbitration Wait Time timer and set the Arbitration Wait Time timer to zero when the SAS port receives an OPEN_REJECT (RETRY). A CONTINUE AWT bit set to zero specifies that the a SAS port shall stop the Arbitration Wait Time timer and set the Arbitration Wait Time timer to zero when it receives an OPEN_REJECT (RETRY).

A BROADCAST ASYNCHRONOUS EVENT bit set to one specifies that the device server shall enable origination of Broadcast (Asynchronous Event) (see 4.1.13). A BROADCAST ASYNCHRONOUS EVENT bit set to zero specifies that the device server shall disable origination of Broadcast (Asynchronous Event).

The READY LED MEANING bit specifies the READY LED signal behavior (see 10.4.1). Regardless of the mode page policy (see SPC-4) for this mode page, the shared mode page policy shall be applied to the READY LED MEANING bit.

The PROTOCOL IDENTIFIER field shall be set to 6h indicating this is a SAS SSP specific mode page.

The I_T NEXUS LOSS TIME field contains the time that the SSP target port shall retry connection requests to an SSP initiator port that are rejected with certain responses indicating that the SSP initiator port may no longer be present (see 8.2.2) before recognizing an I_T nexus loss (see 4.5). Table 3 defines the values of the I_T

NEXUS LOSS TIME field. If this mode page is not implemented, the I_T nexus loss time is vendor specific. This value is enforced by the port layer (see 8.2.2).

Table 3 — I_T NEXUS LOSS TIME **field**

Code	Description
0000h	Vendor-specific amount of time.
0001h to FFFEh	Time in milliseconds.
FFFFh	The SSP target port shall never recognize an I_T nexus loss (i.e., it shall retry the connection requests forever).

NOTE 2 - If this mode page is implemented, the default value of the I_T NEXUS LOSS TIME field should be non-zero. It is recommended that this value be 2 000 ms.

NOTE 3 - An SSP initiator port should retry connection requests for the time indicated by the I_T NEXUS LOSS TIME field in the Protocol-Specific Port mode page for the SSP target port to which it is trying to establish a connection (see 4.5).

The INITIATOR RESPONSE TIMEOUT field contains the time in milliseconds that the SSP target port shall wait for the receipt of a frame (e.g., a write DATA frame) before aborting the command associated with that frame. An INITIATOR RESPONSE TIMEOUT field value of zero indicates that the SSP target port shall disable the initiator response timeout timer. If this mode page is not implemented, the logical unit shall not implement an initiator response timeout timer. This value is enforced by the transport layer (see 9.2.6.3).

The REJECT TO OPEN LIMIT field specifies the minimum time in 10 us increments that the target port shall wait to establish a connection request with an initiator port on an I_T nexus after receiving an OPEN_REJECT (RETRY), an OPEN_REJECT (CONTINUE 0), or an OPEN_REJECT (CONTINUE 1) (see table 113xx). This value may be rounded as defined in SPC-4. A value of 0000h specifies that the time limit is vendor specific. This time limit is enforced by the port layer (see 8.2.3).

10.4.3.3 REPORT GENERAL function

The REPORT GENERAL function returns general information about the SAS device (e.g., a SAS device contained in an expander device). This SMP function shall be implemented by all management device servers.

Table 4 defines the request format.

Table 4 — REPORT GENERAL request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (00h)							
2	Reserved							
3	REQUEST LENGTH (00h)							
4	(MSB)							
7	CRC							(LSB)

The SMP FRAME TYPE field shall be set to 40h.

The FUNCTION field shall be set to 00h.

The REQUEST LENGTH field shall be set to 00h.

The CRC field is defined in 10.4.3.1.

Table 5 defines the response format.

Table 5 — REPORT GENERAL response (part 1 of 3)

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (00h)							
2	FUNCTION RESULT							
3	RESPONSE LENGTH (110h)							
4	(MSB)				EXPANDER CHANGE COUNT			
5								(LSB)
6	(MSB)				EXPANDER ROUTE INDEXES			
7								(LSB)
8	Reserved							
9	NUMBER OF PHYS							
10	TABLE TO TABLE SUPPORTED	Reserved				CONFIGURES OTHERS	CONFIGURING	EXTERNALLY CONFIGURABLE ROUTE TABLE
11	Reserved							
12	ENCLOSURE LOGICAL IDENTIFIER							
19								
20	Reserved							
29								
30	(MSB)				STP BUS INACTIVITY TIME LIMIT			
31								(LSB)
32	(MSB)				STP MAXIMUM CONNECT TIME LIMIT			
33								(LSB)
34	(MSB)				STP SMP I_T NEXUS LOSS TIME			
35								(LSB)
36	NUMBER OF ZONE GROUPS	Reserved		ZONE LOCKED	PHYSICAL PRESENCE SUPPORTED	PHYSICAL PRESENCE ASSERTED	ZONING SUPPORTED	ZONING ENABLED
37	Reserved			SAVING	SAVING ZONE MANAGER PASSWORD SUPPORTED	SAVING ZONE PHY INFORMATION SUPPORTED	SAVING ZONE PERMISSION TABLE SUPPORTED	SAVING ZONING ENABLED SUPPORTED

Table 5 — REPORT GENERAL response (part 2 of 3)

Byte\Bit	7	6	5	4	3	2	1	0
38	(MSB)							
39				MAXIMUM NUMBER OF ROUTED SAS ADDRESSES				(LSB)
40								
47				ACTIVE ZONE MANAGER SAS ADDRESS				
48	(MSB)							
49				ZONE LOCK INACTIVITY TIME LIMIT				(LSB)
50								
51				Reserved				
52				Reserved				
53				FIRST ENCLOSURE CONNECTOR ELEMENT INDEX				
54				NUMBER OF ENCLOSURE CONNECTOR ELEMENT INDEXES				
55				Reserved				
56	REDUCED FUNCTIONALITY				Reserved			
57				TIME TO REDUCED FUNCTIONALITY				
58				INITIAL TIME TO REDUCED FUNCTIONALITY				
59				MAXIMUM REDUCED FUNCTIONALITY TIME				
60	(MSB)							
61				LAST SELF-CONFIGURATION STATUS DESCRIPTOR INDEX				(LSB)
62	(MSB)							
63				MAXIMUM NUMBER OF STORED SELF-CONFIGURATION STATUS DESCRIPTORS				(LSB)
64	(MSB)							
65				LAST PHY EVENT LIST DESCRIPTOR INDEX				(LSB)
66	(MSB)							
67				MAXIMUM NUMBER OF STORED PHY EVENT LIST DESCRIPTORS				(LSB)

Table 5 — REPORT GENERAL response (part 3 of 3)

Byte\Bit	7	6	5	4	3	2	1	0
68					Reserved			
69								
70	(MSB)				STP REJECT TO OPEN LIMIT			
71								(LSB)
72	(MSB)				CRC			
75								(LSB)

The SMP FRAME TYPE field shall be set to 41h.

The FUNCTION field shall be set to 00h.

The FUNCTION RESULT field is defined in 10.4.3.2.

An STP CONTINUE AWT bit set to one indicates that the a STP port is not stopping the Arbitration Wait Time timer or setting the Arbitration Wait Time timer to zero when it receives an OPEN_REJECT (RETRY). An STP CONTINUE AWT bit set to zero indicates that the a STP port is stopping the Arbitration Wait Time timer and setting the Arbitration Wait Time timer to zero when the STP port receives an OPEN_REJECT (RETRY).

> Editor's Note 1: SAS editor to pick the location of the STP CONTINUE AWT bit

The RESPONSE LENGTH field shall be set to 11~~0~~h. For compatibility with previous versions of this standard, a RESPONSE LENGTH field set to 00h indicates that there are 6 dwords before the CRC field.

The EXPANDER CHANGE COUNT field counts the number of Broadcast (Change)s originated by an expander device (see 7.11). Management device servers in expander devices shall support this field. Management device servers in other device types (e.g., end devices) shall set this field to 0000h. This field shall be set to at least 0001h at power on. If the expander device has originated Broadcast (Change) for any reason described in 7.11 since transmitting any SMP response frame containing an EXPANDER CHANGE COUNT field, it shall increment this field at least once from the value in the previous REPORT GENERAL response. It shall not increment this field when forwarding a Broadcast (Change). This field shall wrap to at least 0001h after the maximum value (i.e., FFFFh) has been reached.

> NOTE 4 - Application clients that use the EXPANDER CHANGE COUNT field should read it often enough to ensure that it does not increment a multiple of 65 535 times between reading the field in an expander device compliant with this standard or a multiple of 65 536 times between reading the field in an expander device compliant with previous versions of this standard.

> NOTE 5 - Management device servers in expander devices compliant with previous versions of this standard may return an EXPANDER CHANGE COUNT field set to 0000h.

> NOTE 6 - The originated Broadcast (Change) count is also reported in the REPORT BROADCAST response (see 10.4.3.8).

The EXPANDER ROUTE INDEXES field indicates the maximum number of expander route indexes per phy for the expander device (see 4.6.7.3). Management device servers in externally configurable expander devices containing phy-based expander route tables shall support this field. Management device servers in other device types (e.g., end devices, externally configurable expander devices with expander-based expander route tables, and self-configuring expander devices) shall set the EXPANDER ROUTE INDEXES field to zero. Not all phys in an externally configurable expander device are required to support the maximum number indicated by this field.

The NUMBER OF PHYS field indicates the number of phys in the device, including any virtual phys and any vacant phys.

A TABLE TO TABLE SUPPORTED bit set to one indicates the expander device is a self-configuring expander device that supports its table routing phys being attached to table routing phys in other expander devices. The TABLE TO TABLE SUPPORTED bit shall only be set to one if the EXTERNALLY CONFIGURABLE ROUTE TABLE bit is set to zero. A TABLE TO TABLE SUPPORTED bit set to zero indicates the expander device is not a self-configuring expander device that supports its table routing phys being attached to table routing phys in other expander devices.

A CONFIGURES OTHERS bit set to one indicates that the expander device is a self-configuring expander device that performs the configuration subprocess defined in 4.8. A CONFIGURES OTHERS bit set to zero indicates the expander device may or may not perform the configuration subprocess. Self-configuring expander devices compliant with this standard shall set the CONFIGURES OTHERS bit to one.

> NOTE 7 - If the CONFIGURES OTHERS bit is set to zero, the expander device may configure all externally configurable expander devices in the SAS domain.

A CONFIGURING bit set to one indicates that either:

a) the management device server is in a self-configuring expander device, the self-configuring expander device's management application client is currently performing the discover process (see 4.7), and it has identified at least one change to its expander routing table; or
b) the zoning expander device is locked and the zoning expander shadow values differ from the zoning expander active values.

A CONFIGURING bit set to zero indicates that the management device server is not in a self-configuring expander device currently performing the discover process and changing its expander routing table. Changes in this bit from one to zero result in a Broadcast (Change) being originated (see 7.11). Management device servers in self-configuring expander devices shall support this bit. Management device servers in externally configurable expander devices and in other device types shall set the CONFIGURING bit to zero.

An EXTERNALLY CONFIGURABLE ROUTE TABLE bit set to one indicates that the management device server is in an externally configurable expander device that has a phy-based expander route table that is required to be configured with the SMP CONFIGURE ROUTE INFORMATION function (see 4.6.7.3). An EXTERNALLY CONFIGURABLE ROUTE TABLE bit set to zero indicates that the management device server is not in an externally configurable expander device (e.g., it is in an end device, in a self-configuring expander device, or in an expander device with no phys with table routing attributes).

The ENCLOSURE LOGICAL IDENTIFIER field identifies the enclosure, if any, in which the device is located, and is defined in SES-2. The ENCLOSURE LOGICAL IDENTIFIER field shall be set to the same value reported by the enclosure services process, if any, for the enclosure. An ENCLOSURE LOGICAL IDENTIFIER field set to zero indicates no enclosure information is available.

The STP BUS INACTIVITY TIME LIMIT field indicates the bus inactivity time limit for STP connections, which is set by the CONFIGURE GENERAL function (see 10.4.3.18).

The STP MAXIMUM CONNECT TIME LIMIT field indicates the maximum connect time limit for STP connections, which is set by the CONFIGURE GENERAL function (see 10.4.3.18).

The STP SMP I_T NEXUS LOSS TIME field indicates the time that an STP target port and an SMP initiator port retry certain connection requests which is set by the CONFIGURE GENERAL function (see 10.4.3.18).

The NUMBER OF ZONE GROUPS field indicates the number of zone groups (e.g., the number of entries in the zone group permission table) supported by the expander device and is defined in table 6.

Table 6 — NUMBER OF ZONE GROUPS field

Code	Description
00b	128 zone groups
01b	256 zone groups
All others	Reserved

A ZONE LOCKED bit set to one indicates that the zoning expander device is locked (see 4.9.6.2). A ZONE LOCKED bit set to zero indicates that the zoning expander device is not locked.

A PHYSICAL PRESENCE SUPPORTED bit set to one indicates that the expander device supports physical presence as a mechanism for allowing locking from phys in zone groups without access to zone group 2. A PHYSICAL PRESENCE SUPPORTED bit set to zero indicates that the expander device does not support physical presence as a mechanism for allowing locking.

A PHYSICAL PRESENCE ASSERTED bit set to one indicates that the expander device is currently detecting physical presence. A PHYSICAL PRESENCE ASSERTED bit set to zero indicates that the expander device is not currently detecting physical presence. The PHYSICAL PRESENCE ASSERTED bit shall be set to zero if the PHYSICAL PRESENCE SUPPORTED bit is set to zero.

A ZONING SUPPORTED bit set to one indicates that zoning is supported by the expander device (i.e., it is a zoning expander device). A ZONING SUPPORTED bit set to zero indicates that zoning is not supported by the expander device.

A ZONING ENABLED bit set to one indicates that zoning is enabled in the expander device. A ZONING ENABLED bit set to zero indicates that zoning is disabled in the expander device. The ZONING ENABLED bit shall be set to zero if the ZONING SUPPORTED bit is set to zero.

A SAVING bit set to one indicates that the management device server is currently saving zoning values to non-volatile storage and may return a function result of BUSY for zone management functions that access saved zoning values. A SAVING bit set to zero indicates that the management device server is not currently saving zoning values to non-volatile storage.

A SAVING ZONE MANAGER PASSWORD SUPPORTED bit set to one indicates that saving the zone manager password is supported. A SAVING ZONE MANAGER PASSWORD SUPPORTED bit set to zero indicates that saving the zone manager password is not supported.

A SAVING ZONE PHY INFORMATION SUPPORTED bit set to one indicates that saving the zone phy information is supported. A SAVING ZONE PHY INFORMATION SUPPORTED bit set to zero indicates that saving the zone phy information is not supported.

A SAVING ZONE PERMISSION TABLE SUPPORTED bit set to one indicates that saving the zone permission table is supported. A SAVING ZONE PERMISSION TABLE SUPPORTED bit set to zero indicates that saving the zone permission table is not supported.

A SAVING ZONING ENABLE SUPPORTED bit set to one indicates that saving the ZONING ENABLED bit is supported. A SAVING ZONING ENABLE SUPPORTED bit set to zero indicates that saving the ZONING ENABLED bit is not supported.

The MAXIMUM NUMBER OF ROUTED SAS ADDRESSES field indicates the number of routed SAS addresses in an expander-based expander route table (see 4.6.7.3 and 4.9.3.4). Management device servers in expander devices containing expander-based expander route tables shall support this field. Management device servers in other device types (e.g., end devices and expander devices with phy-based expander route tables) shall set this field to 0000h.

The ACTIVE ZONE MANAGER SAS ADDRESS field indicates the SAS address of the zone manager that last locked the zoning expander device. If the zoning expander device is currently being configured by a vendor-specific

sideband method then the ACTIVE ZONE MANAGER SAS ADDRESS field shall be set to zero. This field shall be set to zero at power on.

The ZONE LOCK INACTIVITY TIME LIMIT field indicates the minimum time between any SMP ZONE LOCK requests, SMP zone configuration function requests, or SMP ZONE ACTIVATE requests from the active zone manager that the locked expander device allows and is set in the SMP ZONE LOCK request (see 10.4.3.21).

The FIRST ENCLOSURE CONNECTOR ELEMENT INDEX field indicates the lowest CONNECTOR ELEMENT INDEX field of all the expander phys in all the expander devices in the enclosure that have CONNECTOR TYPE fields set to 20h through 2Fh (i.e., an internal connector to an end device) in their SMP DISCOVER responses.

The NUMBER OF ENCLOSURE CONNECTOR ELEMENT INDEXES field indicates the number of expander phys in all the expander devices in the enclosure that have CONNECTOR TYPE fields set to 20h through 2Fh (i.e., an internal connector to an end device) in their SMP DISCOVER responses.

> NOTE 8 - The NUMBER OF ENCLOSURE CONNECTOR ELEMENT INDEXES field assumes that all internal connectors to end devices are assigned to a contiguous range of CONNECTOR ELEMENT INDEX field values.

A REDUCED FUNCTIONALITY bit set to one indicates that:

a) the expander device is scheduled to reduce its functionality (see 4.6.8) in the time indicated in the TIME TO REDUCED FUNCTIONALITY field; or
b) that the expander device is currently operating with reduced functionality (see 4.6.8).

A REDUCED FUNCTIONALITY bit set to zero indicates the expander device is not scheduled to reduce functionality and that the contents of the TIME TO REDUCED FUNCTIONALITY field shall be ignored.

If the REDUCED FUNCTIONALITY bit set to one, then the TIME TO REDUCED FUNCTIONALITY field indicates the time, in 100 ms increments, remaining until the expander device is scheduled to reduce functionality. The expander device starts the reduced functionality delay timer after originating a Broadcast (Expander) (see 4.6.8).

The INITIAL TIME TO REDUCED FUNCTIONALITY field indicates the minimum period of time, in 100 ms increments, that an expander device waits from originating a Broadcast (Expander) to reducing functionality. The expander device should set the default value for the INITIAL TIME TO REDUCED FUNCTIONALITY field to at least 2 000 ms (i.e., 14h).

The MAXIMUM REDUCED FUNCTIONALITY TIME field indicates the maximum time, in seconds, that the expander device responds with OPEN_REJECT (RETRY) to connection requests that map to an expander phy or an SMP target port that is not accessible during expander device reduced functionality. This timer starts after the reduced functionality delay timer expires.

The LAST SELF-CONFIGURATION STATUS DESCRIPTOR INDEX field is defined in the REPORT SELF-CONFIGURATION STATUS response (see 10.4.3.5).

The MAXIMUM NUMBER OF STORED SELF-CONFIGURATION STATUS DESCRIPTORS field indicates the maximum number of self-configuration status descriptors (see 10.4.3.5.4) that the management device server supports.

The LAST PHY EVENT LIST DESCRIPTOR INDEX field is defined in the REPORT PHY EVENT LIST response (see 10.4.3.16).

The MAXIMUM NUMBER OF STORED PHY EVENT LIST DESCRIPTORS field indicates the maximum number of phy event list descriptors (see 10.4.3.13.4) that the management device server supports.

The STP REJECT TO OPEN LIMIT field indicates the minimum time in 10 us increments that an STP port waits to establish a connection request with an initiator port on an I_T nexus after receiving an OPEN_REJECT (RETRY), an OPEN_REJECT (CONTINUE 0), or an OPEN_REJECT (CONTINUE 1) (see table 113xx). A value of 0000h indicates that the time limit is vendor specific.

The CRC field is defined in 10.4.3.2.

10.4.3.18 CONFIGURE GENERAL function

The CONFIGURE GENERAL function requests actions by the device containing the management device server. 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 (see 4.9.3.2).

Table 7 defines the request format.

Table 7 — CONFIGURE GENERAL request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (80h)							
2	Reserved							
3	REQUEST LENGTH (04h)							
4	(MSB)			EXPECTED EXPANDER CHANGE COUNT				
5								(LSB)
6				Reserved				
7								
8	Reserved			UPDATE STP REJECT TO OPEN LIMIT	UPDATE INITIAL TIME TO REDUCED FUNCTIONALITY	UPDATE STP SMP I_T NEXUS LOSS TIME	UPDATE STP MAXIMUM CONNECT TIME LIMIT	UPDATE STP BUS INACTIVITY TIME LIMIT
9	Reserved							
10	(MSB)			STP BUS INACTIVITY TIME LIMIT				
11								(LSB)
12	(MSB)			STP MAXIMUM CONNECT TIME LIMIT				
13								(LSB)
14	(MSB)			STP SMP I_T NEXUS LOSS TIME				
15								(LSB)
16	INITIAL TIME TO REDUCED FUNCTIONALITY							
17				STP REJECT TO OPEN LIMIT				
18								
19	Reserved							
20	(MSB)			CRC				
23								(LSB)

The SMP FRAME TYPE field shall be set to 40h.

The FUNCTION field shall be set to 80h.

The REQUEST LENGTH field shall be set to 04h.

If the management device server is not in an expander device or the EXPECTED EXPANDER CHANGE COUNT field is set to 0000h, the EXPECTED EXPANDER CHANGE COUNT field shall be ignored. If the management device server is in an expander device and the EXPECTED EXPANDER CHANGE COUNT field is not set to 0000h, then:

    a)   if the EXPECTED EXPANDER CHANGE COUNT field contains the current expander change count (i.e., the value of the EXPANDER CHANGE COUNT field that would be returned by an SMP REPORT GENERAL response at this time), the management device server shall process the function; and

    b)   If the EXPECTED EXPANDER CHANGE COUNT field does not contain the current expander change count, the management device server shall return a function result of INVALID EXPANDER CHANGE COUNT in the response frame (see table 315 in 10.4.3.2).

An UPDATE STP REJECT TO OPEN LIMIT bit set to one specifies that the STP REJECT TO OPEN LIMIT field shall be honored. An UPDATE STP REJECT TO OPEN LIMIT bit set to zero specifies that the STP REJECT TO OPEN LIMIT field shall be ignored.

An UPDATE INITIAL TIME TO REDUCED FUNCTIONALITY bit set to one specifies that the INITIAL TIME TO REDUCED FUNCTIONALITY field shall be honored. An UPDATE INITIAL TIME TO REDUCED FUNCTIONALITY bit set to zero specifies that the INITIAL TIME TO REDUCED FUNCTIONALITY field shall be ignored.

An UPDATE STP BUS INACTIVITY TIME LIMIT bit set to one specifies that the STP BUS INACTIVITY TIME LIMIT field shall be honored. An UPDATE STP BUS INACTIVITY TIME LIMIT bit set to zero specifies that the STP BUS INACTIVITY TIME LIMIT field shall be ignored.

An UPDATE STP MAXIMUM CONNECT TIME LIMIT bit set to one specifies that the STP MAXIMUM CONNECT TIME LIMIT field shall be honored. An UPDATE STP MAXIMUM CONNECT TIME LIMIT bit set to zero specifies that the STP MAXIMUM CONNECT TIME LIMIT field shall be ignored.

An UPDATE STP SMP I_T NEXUS LOSS TIME bit set to one specifies that the STP SMP I_T NEXUS LOSS TIME field shall be honored. An UPDATE STP SMP I_T NEXUS LOSS TIME bit set to zero specifies that the STP SMP I_T NEXUS LOSS TIME field shall be ignored.

The STP BUS INACTIVITY TIME LIMIT field specifies the maximum time in 100 μs increments that an STP target port is permitted to maintain a connection (see 4.1.12) while transmitting and receiving SATA_SYNC. When this time is exceeded, the STP target port shall close the connection. A value of 0000h in this field specifies that there is no bus inactivity time limit. This value is reported in the STP BUS INACTIVITY TIME LIMIT field in the SMP REPORT GENERAL response (see 10.4.3.3). The bus inactivity time limit is enforced by the port layer (see 8.2.3).

The STP MAXIMUM CONNECT TIME LIMIT field specifies the maximum duration of a connection (see 4.1.12) in 100 μs increments (e.g., a value of 0001h in this field means that the time is less than or equal to 100 μs and a value of 0002h in this field means that the time is less than or equal to 200 μs). When this time is exceeded, the STP target port shall close the connection at the next opportunity. If the STP target port is transferring a frame when the maximum connection time limit is exceeded, the STP target port shall complete transfer of the frame before closing the connection. A value of 0000h in this field specifies that there is no maximum connection time limit. This value is reported in the STP MAXIMUM CONNECT TIME LIMIT field in the SMP REPORT GENERAL response (see 10.4.3.3). The maximum connection time limit is enforced by the port layer (see 8.2.3).

The STP SMP I_T NEXUS LOSS TIME field specifies the time that an STP target port or SMP initiator port shall retry connection requests that are rejected with responses indicating the destination port may no longer be present (see 8.2.2) before recognizing an I_T nexus loss (see 4.5). Table 8 defines the values of the STP SMP I_T NEXUS LOSS TIME field. This value is enforced by the port layer (see 8.2.2).

Table 8 — STP SMP I_T NEXUS LOSS TIME field

Code	Description
0000h	Vendor-specific amount of time.
0001h to FFFEh	Time in milliseconds.
FFFFh	The port shall never recognize an I_T nexus loss (i.e., it shall retry the connection requests forever).

NOTE 9 - The default value of the STP SMP I_T NEXUS LOSS TIME field should be non-zero. It is recommended that this value be 2 000 ms.

NOTE 10 - An STP initiator port should retry connection requests for the time indicated by the STP SMP I_T NEXUS LOSS TIME field in the SMP REPORT GENERAL response for the STP target port to which it is trying to establish a connection.

The INITIAL TIME TO REDUCED FUNCTIONALITY field specifies the minimum period of time, in 100 ms increments, that an expander device shall wait from originating a Broadcast (Expander) to reducing functionality (see 4.6.8). This value is reported in the INITIAL TIME TO REDUCED FUNCTIONALITY field in the SMP REPORT GENERAL response (see 10.4.3.3).

The STP REJECT TO OPEN LIMIT field specifies the minimum time in 10 us increments that an STP port shall wait to establish a connection request with an initiator port on an I_T nexus after receiving an OPEN_REJECT (RETRY), an OPEN_REJECT (CONTINUE 0), or an OPEN_REJECT (CONTINUE 1) (see table 113xx). This value may be rounded as defined in SPC-4. A value of 0000h specifies that the time limit is vendor specific. This time limit is enforced by the port layer (see 8.2.3).

The CRC field is defined in 10.4.3.1.