To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 12 December 2006
Subject: 06-466r1 SAS-2 OPEN_REJECT RETRY during self-configuration changes

Revision history
Revision 0 (26 October 2006) First revision
Revision 1 (12 December 2006) Incorporated comments from November SAS protocol WG.

Related documents
sas2r07 - Serial Attached SCSI - 2 (SAS-2) revision 7
04-340r2 - OPEN_REJECT (BAD DESTINATION) handling (Rob Elliott, HP) - incorporated into SAS-1.1
06-384r1 - OPEN_REJECT (RETRY) during zoning changes (Rob Elliott, HP) - incoporated into SAS-2

Overview
1. After an initiator performs the discover process and discovers a target, it may start opening connections to that target. If a self-configuring expander along the pathway has not yet put the target's SAS address into its downstream routing tables, the initiator will get OPEN_REJECT (NO DESTINATION). This is not a problem; the initiator also discovers the expanders and should have noticed that one of them has its CONFIGURING bit set to one, indicating that it should extend the I_T nexus loss time indefinitely.

If the downstream route tables are all complete, the initiator could open the target successfully, send a read command (like INQUIRY), and close the connection. However, when the target tries to open the initiator, there is no guarantee that all the expanders have finished self-configuration in the upstream direction, so the target could get OPEN_REJECT (NO DESTINATION). Since targets don't perform the discover process, it doesn't know an expander still has its CONFIGURING bit set to one and that it should extend its I_T nexus loss time. If self-configuration takes too long, it will declare an I_T nexus loss (unless the initiator happens to open it in the meantime).

Even on the initiator side, reliance on the CONFIGURING bit is imperfect. The CONFIGURING bit could change in the time between the initiator discovering the expander and the initiator attempting to open the target. BROADCAST (CHANGE)s don't propagate immediately, and the discover process in the expander doesn't set the CONFIGURING bit to one until it detects that a change is needed.

Zoning expanders return OPEN_REJECT (RETRY) rather than OPEN_REJECT (ZONE VIOLATION) while zoning configuration changes are in progress. This proposal has self-configuring expanders do the same, and return OPEN_REJECT (RETRY) rather than OPEN_REJECT (NO DESTINATION) if self-configuring changes are in progress. This will take effect as soon as BROADCAST (CHANGE) is received (rather than wait until the discover process determines a change is needed).

See figure 1.

2. SAS-1.1 was (purposely) imprecise about whether an expander device returns OPEN_REJECT (BAD DESTINATION) vs. OPEN_REJECT (NO DESTINATION) if the connection request loops back on a table routing phy, saying NO DESTINATION "should" be returned. This proposal upgrades that to a "shall" - OPEN_REJECT (NO DESTINATION) shall be returned if the port is using the table or subtractive routing method.

Since OPEN_REJECT (BAD DESTINATION) is clearly a fatal error based on those rules, returning OPEN_REJECT (RETRY) during self-configuration changes is not going to help anything. Therefore, it is excluded from the new treatment; OPEN_REJECT (BAD DESTINATION) shall be returned if the port is using the direct routing method. (Without this, there would be no defined reason for OPEN_REJECT (BAD DESTINATION any more.)

.

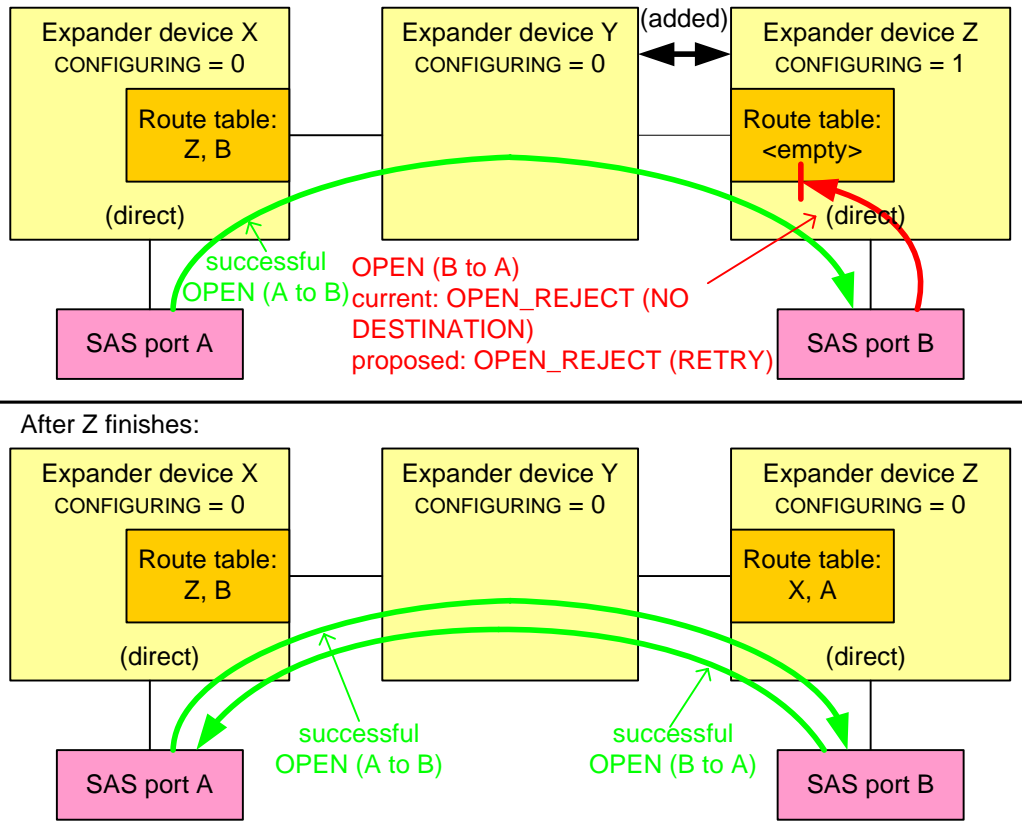After expander Y and Z are attached, assume X and Y complete self-configuration quickly while Z is slow:

Figure 1 — Self-configuation not complete example

Suggested changes

4.6.6.3 ECM interface

> Editor's Note 1: Note: a strikethrough/underlined combo is not existing text; it was temporarily proposed during development of this proposal

Table 12 describes the requests from an expander logical phy to the ECM. The XL state machine (see 7.15) defines specifically when each request is sent.

Table 12 — Expander logical phy to ECM requests

Message	Description
Request Path (arguments)	Request for a connection (i.e., the expander logical phy is in the XL1:Request_Path state (see 7.15.4.1)).
Partial Pathway Timeout Timer Expired	The Partial Pathway Timeout Timer expired (i.e., the expander logical phy is in the XL1:Request_Path state (see 7.15.4.1) and the timer expires).

Table 13 describes the responses from an expander logical phy to the ECM. The XL state machine (see 7.15) defines specifically when each response is sent.

Table 13 — Expander logical phy to ECM responses

Message	Description
~~Phy Status (Partial Pathway)~~	~~Response meaning that an expander logical phy:~~ ~~a) is being used for an unblocked partial pathway (i.e., the expander logical phy is in the XL3:Open_Confirm_Wait state (see 7.15.6.1) and the last AIP transmitted was not AIP (WAITING ON PARTIAL), or the expander logical phy is in the XL6:Open_Response_Wait state (see 7.15.9.1) and the last AIP received was not AIP (WAITING ON PARTIAL)); or~~ ~~b) has sent a Request Path request to the ECM (i.e., the expander logical phy is in the XL1:Request Path state (see 7.15.6.1)) and is receiving Arbitrating (Waiting On Partial) from the ECM.~~
Phy Status (Partial Pathway)	Response meaning that an expander logical phy: a) is being used for an unblocked partial pathway; or b) is waiting on another phy being used for a partial pathway.
~~Phy Status (Blocked Partial Pathway)~~	~~Response meaning that an expander logical phy:~~ ~~a) is being used for a blocked partial pathway (i.e., the expander logical phy is in the XL3:Open_Confirm_Wait state (see 7.15.6.1) and the last AIP transmitted was AIP (WAITING ON PARTIAL), or the expander logical phy is in the XL6:Open_Response_Wait state (see 7.15.9.1) and the last AIP received was AIP (WAITING ON PARTIAL)); or~~ ~~b) has sent a Request Path request to the ECM (i.e., the expander logical phy is in the XL1:Request Path state (see 7.15.6.1)) and is receiving Arbitrating (Blocked On Partial) from the ECM.~~
Phy Status (Blocked Partial Pathway)	Response meaning that an expander logical phy: a) is being used for a blocked partial pathway; or b) is waiting on another phy being used for blocked partial pathway.
~~Phy Status (Connection)~~	~~Response meaning that an expander logical phy:~~ ~~a) is being used for a connection (i.e., the expander logical phy is in the XL7:Connected state (see 7.15.10.1) or the XL8:Close_Wait state state (see 7.15.11.1)); or~~ ~~b) has sent a Request Path request to the ECM (i.e., the expander logical phy is in the XL1:Request Path state (see 7.15.6.1)) and is receiving Arbitrating (Waiting On Connection) from the ECM.~~
Phy Status (Connection)	Response meaning that an expander logical phy: a) is being used for a connection; or b) is waiting on another phy being used for a connection.

Table 14 describes the confirmations from the ECM to an expander phy. These confirmations are sent in confirmation of a Request Path request. See 7.12.4 for specific definitions about when each confirmation is sent.

Table 14 — ECM to expander phy confirmations (part 1 of 2)

Message	Description
Arbitrating (Normal)	Confirmation that the ECM has received the Request Path request.
Arbitrating (Waiting On Partial)	~~Confirmation that the ECM has determined that:~~ ~~a) there is a destination port capable of routing to the requested destination SAS address;~~ ~~b) at least one phy within the destination port supports the requested connection rate;~~ ~~c) each of the phys within the destination port is returning a Phy Status (Partial Pathway) or Phy Status (Blocked Partial Pathway) response; and~~ ~~d) at least one of the phys within the destination port is returning a Phy Status (Partial Pathway) response.~~ Confirmation that the ECM is waiting on a partial pathway (see 4.1.11).
Arbitrating (Blocked On Partial)	~~Confirmation that the ECM has determined that:~~ ~~a) there is a destination port capable of routing to the requested destination SAS address;~~ ~~b) at least one phy within the destination port supports the requested connection rate; and~~ ~~each of the phys within the destination port is returning a Phy Status (Blocked Partial Pathway) response.~~ Confirmation that the ECM is waiting on a blocked partial pathway (see 4.1.11).
Arbitrating (Waiting On Connection)	~~Confirmation that the ECM has determined that the connection request is blocked due to one of the following reasons:~~ ~~a) the connection request is blocked by an active connection; or~~ ~~b) there are insufficient routing resources within the expander to complete the connection request.~~ ~~A connection request shall be considered blocked by an active connection when:~~ ~~a) there is a destination port capable of routing to the requested destination SAS address;~~ ~~b) at least one phy within the destination port supports the requested connection rate;~~ ~~c) each of the phys within the destination port is returning a Phy Status (Partial Pathway), Phy Status (Blocked Partial Pathway), or Phy Status (Connection) response; and~~ ~~at least one of the phys within the destination port is returning a Phy Status (Connection) response.~~ Confirmation that the ECM is waiting for a connection to complete.
Arb Won	Confirmation that an expander logical phy has won path arbitration.
Arb Lost	Confirmation that an expander logical phy has lost path arbitration.
~~Arb Reject (No Destination)~~	~~Confirmation that~~ the expander device is not configuring (see 4.7.3) and~~:~~ ~~a) there is no operational expander logical phy capable of routing to the requested destination SAS address; or~~ ~~b) the requested destination SAS address maps back to the requesting port~~ and the requesting port (i.e., the destination port) is using the table routing method or the subtractive routing method (see 7.12.4.3).

Table 14 — ECM to expander phy confirmations (part 2 of 2)

Message	Description
Arb Reject (No Destination)	Confirmation that the expander device is not configuring (see 4.7.3) and there is no path to the destination.
~~Arb Reject (Bad Destination)~~	~~Confirmation that:~~ ~~a)~~ ~~the requested destination SAS address maps back to the requesting port; and the requesting port (i.e., the destination port) is using the direct routing method~~ *~~or the table routing method; and~~* ~~b)~~ *~~the ECM has not chosen to return Arb Reject (No Destination) (see 7.12.4.3).~~*
Arb Reject (Bad Destination)	Confirmation that the path to the destination maps back to the requesting expander port.
Arb Reject (Connection Rate Not Supported)	Confirmation that the ECM has determined that there is a destination port capable of routing to the requested destination SAS address but no phys within the destination port are configured to support the requested connection rate.
Arb Reject (Zone Violation)	Confirmation that the the expander device is not locked and the ECM has determined that there is a zoning violation (see 4.9.3).
Arb Reject (Pathway Blocked)	Confirmation that the ECM has determined that the requesting expander logical phy shall back off according to SAS pathway recovery rules.
Arb Reject (Retry)	Confirmation that: a) ~~the ECM has determined that there is a zoning violation (see 4.9.3) while the zoning expander device is locked.;~~ b) the expander device is locked (see 4.7.3) and the ECM would otherwise have returned Arb Reject (Zone Violation); or c) the expander device is configuring (see 4.7.3) and the ECM would otherwise have returned Arb Reject (No Destination).

4.6.6.4 ECR interface

Table 15 describes the requests from an expander logical phy to the ECR and the corresponding indications from the ECR to another expander logical phy. The XL state machine (see 7.15) defines specifically when each request is sent.

Table 15 — Expander logical phy to ECR to expander logical phy requests and indications

Message	Description
Forward Open (arguments)	Request/indication to forward an OPEN address frame.
Forward Close	Request/indication to forward a CLOSE.
Forward Break	Request/indication to forward a BREAK.
Forward Dword	Request/indication to forward a dword.

Table 16 describes the responses from an expander logical phy to the ECR and the corresponding confirmations from the ECR to another expander logical phy. These responses are sent in response to a Forward Open indication. The XL state machine (see 7.11) defines specifically when each response is sent.

Table 16 — Expander logical phy to ECR to expander logical phy responses and confirmations

Message	Description
Arb Status (Normal)	Confirmation/response that AIP (NORMAL) has been received.
Arb Status (Waiting On Partial)	Confirmation/response that AIP (WAITING ON PARTIAL) has been received.
Arb Status (Waiting On Connection)	Confirmation/response that AIP (WAITING ON CONNECTION) has been received.
Arb Status (Waiting On Device)	Confirmation/response that: a) AIP (WAITING ON DEVICE) has been received; or b) the expander logical phy has completed the forwarding of an OPEN address frame and has entered the XL6:Open_Response_Wait state.
Open Accept	Confirmation/response that OPEN_ACCEPT has been received.
Open Reject	Confirmation/response that OPEN_REJECT has been received.
Backoff Retry	Confirmation/response that: a) a higher priority OPEN address frame has been received (see 7.12.3); and b) the source SAS address and connection rate of the received OPEN address frame are not equal to the destination SAS address and connection rate of the transmitted OPEN address frame.
Backoff Reverse Path	Confirmation/response that: a) a higher priority OPEN address frame has been received (see 7.12.3); and b) the source SAS address and connection rate of the received OPEN address frame are equal to the destination SAS address and connection rate of the transmitted OPEN address frame.

4.6.6.5 BPP interface

Table 17 describes the requests from an expander logical phy to the BPP. Requests from the management device server about SMP ZONED BROADCAST requests received from the SMP target port in zoning

expander devices with zoning enabled are not described. See 4.9.5 for more information on how zoning expander devices with zoning enabled handle Broadcasts.

Table 17 — Expander logical phy to BPP requests

Message	Description
Broadcast Event Notify (Phy Not Ready)	Request to originate a Broadcast (Change) because the expander logical phy's SP state machine transitioned from the SP15:SAS_PHY_Ready or SP22:SATA_PHY_Ready state to the SP0:OOB_COMINIT state (see 6.8).
Broadcast Event Notify (SATA Spinup Hold)	Request to originate a Broadcast (Change) because the SATA spinup hold state has been reached (see 6.8 and 6.11) by the expander phy.
Broadcast Event Notify (Identification Sequence Complete)	Request to originate a Broadcast (Change) because the expander logical phy has completed the identification sequence (see 7.9).
Broadcast Event Notify (SATA Port Selector Change)	Request to originate a Broadcast (Change) because the expander phy detected that a SATA port selector appeared or disappeared.
Broadcast Event Notify (Change Received)	Request to forward a Broadcast (Change) because the expander logical phy received a Broadcast (Change). See 7.11 and 7.15.
Broadcast Event Notify (Reserved Change 0 Received)	Request to forward a Broadcast (Reserved Change 0) because the expander logical phy received a Broadcast (Reserved Change 0). See 7.11 and 7.15.
Broadcast Event Notify (Reserved Change 1 Received)	Request to forward a Broadcast (Reserved Change 1) because the expander logical phy received a Broadcast (Reserved Change 1). See 7.11 and 7.15.
Broadcast Event Notify (SES Received)	Request to forward a Broadcast (SES) because the expander logical phy received a Broadcast (SES). See 7.15.
Broadcast Event Notify (Expander Received)	Request to forward a Broadcast (Expander) because the expander logical phy received a Broadcast (Expander). 7.11 See 7.15.
Broadcast Event Notify (Asynchronous Event Received)	Request to forward a Broadcast (Asynchronous Event) because the expander logical phy received a Broadcast (Asynchronous Event). See 7.15.
Broadcast Event Notify (Reserved 3 Received)	Request to forward a Broadcast (Reserved 3) because the expander logical phy received a Broadcast (Reserved 3). See 7.15.
Broadcast Event Notify (Reserved 4 Received)	Request to forward a Broadcast (Reserved 4) because the expander logical phy received a Broadcast (Reserved 4). See 7.15.

Table 18 describes the indications from the BPP to an expander logical phy. Indications to the management application client to generate SMP ZONED BROADCAST functions from the SMP initiator port in a zoning expander device with zoning enabled are not described. See 4.9.5 for more information on how zoning expander devices with zoning enabled handle Broadcasts.

Table 18 — BPP to expander logical phy indications

Message	Description
Transmit Broadcast (type)	Indication to transmit a BROADCAST with the specified type.

4.6.7 Expander device routing

4.6.7.1 Routing attributes and routing methods

Each expander phy in an expander device shall support one of the following routing attributes:

    a)   direct routing attribute;
    b)   table routing attribute; or
    c)   subtractive routing attribute.

The routing attributes allow the ECM to determine which routing method to use when routing connection requests to the expander logical phys in the expander phy:

    a)   the table routing method routes connection requests to attached expander devices using an expander route table;
    b)   the subtractive routing method routes unresolved connection requests to an attached expander device; or
    c)   the direct routing method routes connection requests to attached end devices, the SMP port of an attached expander device, or SAS devices contained in the expander device.

Table 19 describes the routing methods that the ECM uses based on the routing attributes of an expander phy.

Table 19 — Routing attributes and routing methods

Routing attribute of an expander phy	Routing method used by ECM for the expander phy
Direct	Direct [a]
Table	Direct, if attached to an end device
	Direct, if attached to an expander device, for the SAS address of the expander device
	Table, if attached to an expander device, for SAS addresses beyond the expander device
Subtractive	Direct, if attached to an end device
	Subtractive, if attached to an expander device
[a] If attached to an expander device, the ECM is only able to route to the expander device itself through a phy with the direct routing attribute	

An expander device may have zero or more phys with the table routing attribute.

An expander device shall have at most one defined port containing phys with the subtractive routing attribute.

An expander device that does not support table-to-table attachment shall only use phys with the table routing attribute to attach to phys with the subtractive routing attribute in other expander devices.

If multiple phys within an expander device have subtractive routing attributes and are attached to expander devices, they shall attach to phys with identical SAS addresses (i.e., the same expander port).

If multiple phys within an expander device have subtractive routing attributes and are attached to expander devices that do not have identical SAS addresses, the application client that is performing the discover process (see 4.7) shall report an error in a vendor-specific manner.

4.6.7.2 Connection request routing

The ECM shall determine how to route a connection request from a source expander logical phy to a destination expander logical phy in a different expander port if the destination expander logical phy is enabled

and operating at a valid logical link rate (e.g., the SMP DISCOVER response reports a NEGOTIATED LOGICAL LINK RATE field set to G1 (i.e., 8h), G2 (i.e., 9h), or G3 (i.e., Ah)) using the following precedence:

1) route to an expander logical phy with the direct routing attribute or table routing attribute when the destination SAS address matches the attached SAS address;
2) route to an expander logical phy with the table routing attribute when the destination SAS address matches an enabled SAS address in the expander route table;
3) route to an expander logical phy with the subtractive routing attribute; or
4) return an Arb Reject confirmation (see 4.6.6.3) to the source expander logical phy.

Editor's Note 2: The list should including zoning considerations as well

If the destination expander logical phy only matches an expander logical phy in the same expander port from which the connection request originated, then the ECM shall return an Arb Reject confirmation.

If the destination SAS address of a connection request matches a disabled SAS address in an expander route table, then the ECM shall ignore the match.

4.7 Discover process

4.7.2 Discover process overview

...

~~If~~While a self-configuring expander device's CONFIGURING bit is set to one in the SMP REPORT GENERAL response, connection requests for destination ports two or more levels beyond the self-configuring expander device ~~may result in responses such as OPEN_REJECT (NO DESTINATION) (see 7.2.5.12). If a management application client (e.g., in an end device) detects a self-configuring expander device with the configuring bit set to one, the SAS port shall stop the I_T Nexus Loss timer for I_T nexuses involving such destination ports as long as the CONFIGURING bit is set to one (see 8.2.2.1)~~that would otherwise have returned OPEN_REJECT (NO DESTINATION) return OPEN_REJECT (RETRY) instead (see 4.6.6.3 and 4.7.3).

...

4.7.3 Discover process in a self-configuring expander device

The management application client of a self-configuring expander device shall configure:

a) the expander routing table in that expander device; and
b) the expander routing table in each externally configurable expander device in the SAS domain that is not located behind another self-configuring expander device.

When a self-configuring expander device receives a Broadcast (Change) the management application client shall start the discover process using the expander port that received the Broadcast (Change). If a change to the expander route table is identified then the management device server shall set its CONFIGURING bit to one in the SMP REPORT GENERAL response (see 10.4.3.3).

The management application client shall set the CONFIGURING bit to zero when the discover process is complete. When the CONFIGURING bit changes from one to zero:

a) a zoning expander device with zoning enabled shall originate a Broadcast (Change) on each expander port that has access to the expander port through which the discover process was performed based on the zone permission table; and
b) an expander device with zoning disabled shall originate a Broadcast (Change) on each expander port other than the one through which the discover process was performed.

When a Broadcast (Change) has been received by a self-configuring expander device, all previously valid SAS addresses shall continue to be routable until they are determined to be no longer valid. When a change has been identified, all unaffected SAS addresses shall continue to be routable.

While the CONFIGURING bit is set to one, the expander device shall return OPEN_REJECT (RETRY) for any connection requests that would otherwise have resulted in OPEN_REJECT (NO DESTINATION) (see 4.6.6.3).

7.2.5.12 OPEN_REJECT

OPEN_REJECT specifies that a connection request has been rejected and specifies the reason for the rejection. The result of some OPEN_REJECTs is to abandon (i.e., not retry) the connection request and the result of other OPEN_REJECTs is to retry the connection request.

All of the OPEN_REJECT versions defined in table 20 shall result in the originating port abandoning the connection request.

Table 20 — OPEN_REJECT abandon primitives

Primitive	Originator	Description
OPEN_REJECT (BAD DESTINATION)	Expander phy	~~A connection request arrives through an expander phy using the direct routing or table routing method and the expander device determines the connection request would have to be routed to the same expander port as the expander port through which the connection request arrived (e.g., the destination SAS address equals the source SAS address), and the expander device has not chosen to return OPEN_REJECT (NO DESTINATION) (see 7.12.4.4).~~ <u>The connection request routes to a destination expander phy in the same expander port as the source expander phy and the expander port is using the direct routing method (see x.x.x).</u>
OPEN_REJECT (CONNECTION RATE NOT SUPPORTED)	Any phy	The requested connection rate is not supported on some physical link on the pathway between the source phy and destination phy. When a SAS initiator phy is directly attached to a SAS target phy, the requested connection rate is not supported by the destination phy. The connection request may be modified and reattempted as described in 7.12.2.2.
OPEN_REJECT (PROTOCOL NOT SUPPORTED)	Destination phy	Phy with destination SAS address exists but the destination phy does not support the requested initiator/target role, protocol, initiator connection tag, or features (i.e., the values in the INITIATOR PORT bit, the PROTOCOL field, the INITIATOR CONNECTION TAG field, and/or the FEATURES field in the OPEN address frame are not supported).
OPEN_REJECT (RESERVED ABANDON 1)	Unknown	Reserved. Process the same as OPEN_REJECT (WRONG DESTINATION).
OPEN_REJECT (RESERVED ABANDON 2)	Unknown	Reserved. Process the same as OPEN_REJECT (WRONG DESTINATION).
OPEN_REJECT (RESERVED ABANDON 3)	Unknown	Reserved. Process the same as OPEN_REJECT (WRONG DESTINATION).
OPEN_REJECT (STP RESOURCES BUSY)	Destination phy	STP target port with destination SAS address exists but the STP target port has an affiliation with another STP initiator port or all of the available task file registers have been allocated to other STP initiator ports (see 7.17.5). Process the same as OPEN_REJECT (WRONG DESTINATION) for non-STP connection requests.
OPEN_REJECT (WRONG DESTINATION)	Destination phy	The destination SAS address does not match the SAS address of the SAS port to which the connection request was delivered.
OPEN_REJECT (ZONE VIOLATION)	Zoning expander phy	The connection request is from a zone group that does not have permission to access the zone group that contains the destination phy according to the zone permission table of an unlocked zoning expander device.

All of the OPEN_REJECT versions defined in table 21 shall result in the originating port retrying the connection request.

Table 21 — OPEN_REJECT retry primitives (part 1 of 2)

Primitive	Originator	Description
OPEN_REJECT (NO DESTINATION) [a]	Expander phy	An expander in the pathway is not configuring and determines that~~Either~~: a) there is no ~~No~~ such destination phy; b) ~~the expander device determines the connection request would have to be routed to the same expander port as the expander port through which the connection request arrived (e.g., the destination SAS address equals the source SAS address) and the expander device has not chosen to return OPEN_REJECT (BAD DESTINATION) (see 7.12.4.4)~~ the connection request routes to a destination expander phy in the same expander port as the source expander phy and the expander port is using the direct routing method (see x.x.x); or c) the SAS address is valid for an STP target port in an STP/SATA bridge, but the initial Register - Device to Host FIS has not been successfully received (see 10.4.3.9).
OPEN_REJECT (PATHWAY BLOCKED) [b]	Expander phy	An expander device determined the pathway was blocked by higher priority connection requests.
OPEN_REJECT (RESERVED CONTINUE 0) [c]	Unknown	Reserved. Process the same as OPEN_REJECT (RETRY).
OPEN_REJECT (RESERVED CONTINUE 1) [c]	Unknown	Reserved. Process the same as OPEN_REJECT (RETRY).
OPEN_REJECT (RESERVED INITIALIZE 0) [a]	Unknown	Reserved. Process the same as OPEN_REJECT (NO DESTINATION).
OPEN_REJECT (RESERVED INITIALIZE 1) [a]	Unknown	Reserved. Process the same as OPEN_REJECT (NO DESTINATION).
OPEN_REJECT (RESERVED STOP 0) [b]	Unknown	Reserved. Process the same as OPEN_REJECT (PATHWAY BLOCKED).

Table 21 — OPEN_REJECT retry primitives (part 2 of 2)

Primitive	Originator	Description
OPEN_REJECT (RESERVED STOP 1) b	Unknown	Reserved. Process the same as OPEN_REJECT (PATHWAY BLOCKED).
OPEN_REJECT (RETRY) c	Destination phy or zoning expander phy	Either: a) Phy with destination SAS address exists but is not able to accept connections~~,~~: b) an expander in the pathway is configuring and it would otherwise have returned OPEN_REJECT (NO DESTINATION); or c) an expander in the pathway is locked and it would otherwise have returned OPEN_REJECT (ZONE VIOLATION)(see table 15)~~the connection request is from a zone group that does not have permission to access the zone group that contains the destination phy according to the zone permission table of a locked zoning expander device~~.

a If the I_T Nexus Loss timer is already running, it continues running; if it is not already running, it is initialized and started. Stop retrying the connection request if the I_T Nexus Loss timer expires.
b If the I_T Nexus Loss timer is already running, it continues running. Stop retrying the connection request if the I_T Nexus Loss timer expires.
c If the I_T Nexus Loss timer (see 8.2.2) is already running, it is stopped.

7.12 Connections

7.12.1 Connections overview

7.12.2 Opening a connection

7.12.3 Arbitration fairness

7.12.4 ~~Arbitration and resource management in an expander device~~Arbitration inside an expander device

7.12.4.1 Expander phy arbitration requirements

[moved from below]

An expander phy shall set the Retry Priority Status argument to IGNORE AWT when it requests a path after:

a) it has forwarded an OPEN address frame to the physical link;
b) an OPEN address frame is received with higher arbitration priority (see 7.12.3); and
c) the destination SAS address and connection rate of the received OPEN address frame are not equal to the source SAS address and connection rate of the transmitted OPEN address frame (see 7.15.4 and 7.15.9).

Otherwise, the expander phy shall set the Retry Priority Status argument to NORMAL.

See the XL state machine (see 7.15) for detailed expander phy requirements.

7.12.4.2 ~~Arbitration and resource management in an expander device overview~~ECM arbitration requirements

7.12.4.2.1 ECM arbitration requirements overview

The ECM shall arbitrate and assign or deny path resources for Request Path requests from each expander phy.

Arbitration includes adherence to the SAS arbitration fairness algorithm and path recovery. Path recovery is used to avoid potential deadlock scenarios within the SAS topology by deterministically choosing which partial pathway(s) to tear down to allow at least one connection to complete.

~~The ECM responds to each Request Path request by returning an Arb Won, Arb Lost, or Arb Reject confirmation to the requesting expander phy.~~

Several of the Request Path arguments are used for arbitration. The Arbitration Wait Time, Source SAS Address, and Connection Rate arguments are filled in from the received OPEN address frame and are used to by the ECM to compare Request Path requests. The Retry Priority Status argument is used to prevent the Arbitration Wait Time argument from being considered during an arbitration which occurs after a Backoff Retry response is sent by an expander phy (see 7.15.4).

[moved from 7.12.4.7]

When the ECM in an expander device receives a connection request:

1) if the destination SAS address is that of the expander device itself, the ECM shall arbitrate for access to its SMP target port ~~and forward the connection request~~;
2) if the destination SAS address matches the SAS address ~~of a device to which~~attached to one or more of the expander phys ~~is attached~~, the ECM shall arbitrate for access to ~~one of the matching~~those expander phys ~~and forward the connection request~~;
3) if the destination SAS address matches an enabled SAS address in the expander route table for one or more expander phys ~~which~~that is using the table routing method ~~has a table routing attribute (see 4.6.7.1) and is attached to an expander device~~, the ECM shall arbitrate for access to ~~one of the matching~~those expander phys ~~and forward the connection request~~; and
4) if at least one expander phy is using the subtractive routing method ~~has the subtractive routing attribute and is attached to an expander device (i.e., the phy is using the subtractive routing method)~~, and the request did not come from ~~that expander device~~one of those expander phys, the ECM shall arbitrate for access to ~~one of the subtractive routing~~those expander phys ~~and forward the connection request~~.

The ECM shall respond to each Request Path request by returning the following confirmations to the requesting expander logical phy while processing the Request Path request:

a) Arbitrating (Normal) (see 7.12.4.2.2);
b) Arbitrating (Waiting On Partial) (see 7.12.4.2.2);
c) Arbitrating (Blocked On Partial) (see 7.12.4.2.2);
d) Arbitrating (Waiting On Connection) (see 7.12.4.2.2).

The ECM shall complete responding to each Request Path request by returning one of the following confirmations to the requesting expander logical phy:

a) Arb Won (see 7.12.4.2.3);
b) Arb Lost (see 7.12.4.2.4); or
c) Arb Reject (see 7.12.4.2.5).

[moved rest of text in this section into those subsections]

7.12.4.2.2 Arbitrating confirmations [reworded from ECM to expander logical phy table entries]

The ECM shall send an Arbitrating (Normal) confirmation after it has received a Request Path request.

The ECM shall send an Arbitrating (Waiting On Partial) confirmation if it is waiting on a partial pathway (see 4.1.11). The ECM is waiting on a partial pathway if:

a) there is a destination port capable of routing to the requested destination SAS address;
b) at least one phy within the destination port supports the requested connection rate;
c) each of the phys within the destination port is returning a Phy Status (Partial Pathway) or Phy Status (Blocked Partial Pathway) response; and
d) at least one of the phys within the destination port is returning a Phy Status (Partial Pathway) response.

The ECM shall send an Arbitrating (Blocked On Partial) confirmation if it is waiting on a blocked partial pathway (see 4.1.11). The ECM is waiting on a blocked partial pathway if:

     a)   there is a destination port capable of routing to the requested destination SAS address;
     b)   at least one phy within the destination port supports the requested connection rate; and
     c)   each of the phys within the destination port is returning a Phy Status (Blocked Partial Pathway) response.

The ECM shall send an Arbitrating (Waiting On Connection) confirmation if it is waiting on a connection to complete. The ECM is waiting on a connection to complete if:

     a)   the connection request is blocked by an active connection; or
     b)   there are insufficient routing resources within the expander to complete the connection request.

A connection request shall be considered blocked by an active connection when:

     a)   there is a destination port capable of routing to the requested destination SAS address;
     b)   at least one phy within the destination port supports the requested connection rate;
     c)   each of the phys within the destination port is returning a Phy Status (Partial Pathway), Phy Status (Blocked Partial Pathway), or Phy Status (Connection) response; and
     d)   at least one of the phys within the destination port is returning a Phy Status (Connection) response.

7.12.4.2.3 Arb Won confirmation [text is unchanged, but it is placed into a subsection]

The ECM shall generate the Arb Won confirmation when all of the following conditions are met:

     a)   the Request Path request maps to a destination expander phy that:
         A)   supports the connection rate; and
         B)   is not reporting a Phy Status (Partial Pathway), Phy Status (Blocked Partial Pathway), or Phy Status (Connection) response, unless that expander phy is arbitrating for the requesting expander phy;
     b)   there are sufficient routing resources to complete the connection request;
     c)   no higher priority Request Path requests are present with the requesting expander phy as the destination; and
     d)   the Request Path request is the highest priority Request Path request (see table 123 and table 124) mapping to the destination expander phy (i.e., only send one Arb Won confirmation for Request Path requests to the same destination phy).

If two or more Request Path requests contend and all of the Request Path requests include a Retry Priority Status argument set to NORMAL, the ECM shall select the winner by comparing the OPEN address frame contents described in table 123.

Table 123 — Arbitration priority for contending Request Path requests in the ECM when all requests have Retry Priority Status arguments of NORMAL

Bits 83-68 (83 is MSB)	Bits 67-4	Bits 3-0 (0 is LSB)
ARBITRATION WAIT TIME field value	SOURCE SAS ADDRESS field value	CONNECTION RATE field value

If two or more Request Path requests contend and one or more of the Request Path requests include a Retry Priority Status argument set to IGNORE AWT, the ECM shall select the winner from the set of Request Path requests with Retry Priority Status arguments set to IGNORE AWT by comparing the OPEN address frame contents described in table 124.

Table 124 — Arbitration priority for contending Request Path requests in the ECM among requests with Retry Priority Status arguments of IGNORE AWT

Bits 67-4 (67 is MSB)	Bits 3-0 (0 is LSB)
SOURCE SAS ADDRESS field value	CONNECTION RATE field value

7.12.4.2.4 Arb Lost confirmation [text is unchanged, but it is placed into a subsection]

The ECM shall generate the Arb Lost confirmation when all of the following conditions are met:

 a) the Request Path request maps to a destination expander phy that:
   A) supports the connection rate; and
   B) is not reporting a Phy Status (Partial Pathway), Phy Status (Blocked Partial Pathway), or Phy Status (Connection) response unless that expander phy is arbitrating for the requesting expander phy;
 b) there are sufficient routing resources to complete the connection request; and
 c) one of the following conditions are met:
   A) the destination expander phy is making a Request Path request with the requesting expander phy as its destination (i.e., when two expander phys both receive an OPEN address frame destined for each other, the ECM provides the Arb Lost confirmation to the expander phy that received the lowest priority OPEN address frame); or
   B) the ECM is sending an Arb Won confirmation to another expander phy that is using the requesting expander phy as the destination.

7.12.4.2.5 Arb Reject confirmation [text is changed and placed into a subsection]

The ECM shall generate the following Arb Reject confirmation (see 4.6.6.3) when any of the following conditions are met and all the Arb Won conditions (see 7.12.4.3) are not met:

 1) ~~Arb Reject (No Destination), or Arb Reject (Bad Destination) if the Request Path request does not map to a destination expander phy that is not part of the same expander port as the requesting expander phy (i.e., there is no direct routing or table routing match and there is no subtractive phy)(see 7.12.4.3);~~
 2) ~~Arb Reject (Connection Rate Not Supported) if the Request Path request does not map to any expander phy that supports the connection rate (i.e., none of the prospective physical links support the requested connection rate);~~
 3) ~~Arb Reject (Zone Violation) if the Request Path request maps to expander phys with a zoning violation (see 4.9.3); or~~
 4) ~~Arb Reject (Pathway Blocked) if the Request Path request maps to that all contain blocked partial pathways (i.e., are all returning Phy Status (Blocked Partial Pathway)) and pathway recovery rules require this Request Path request to release path resources (see 7.12.4.5).~~
 1) Arb Reject (Bad Destination) if the source expander phy and destination expander phy(s) are in the same expander port and are using the direct routing method;
 2) if the source expander phy and destination expander phy(s) are in the same expander port and are using the table routing method or the subtractive routing method:
   A) Arb Reject (No Destination) if the expander device is not configuring; and
   B) Arb Reject (Retry) if the expander device is configuring;
 3) if there are no destination expander phys (i.e., there is no direct routing or table routing match and there is no subtractive phy):
   A) Arb Reject (No Destination) if the expander device is not configuring; and
   B) Arb Reject (Retry) if the expander device is configuring;
 4) Arb Reject (Connection Rate Not Supported) if none of the destination expander phys supports the connection rate;
 5) if access to the destination expander phy(s) is prohibited by zoning (see 4.9.3):
   A) Arb Reject (Zone Violation) if the zoning expander device is unlocked; and
   B) Arb Reject (Retry) if the zoning expander device is locked;

   and

 6) Arb Reject (Pathway Blocked) if all the destination expander phys that support the connection rate contain blocked partial pathways (i.e., are all returning Phy Status (Blocked Partial Pathway)) and pathway recovery rules require this Request Path request be rejected to release path resources (see 7.12.4.5).

[moved from 7.12.4.7 but then deleted]

16

If it does not find a match and no phy using the subtractive routing method exists, the ECM shall reply with Arb Reject (No Destination).

If the destination phy is in the same expander port as the source phy and is using the subtractive routing method, the ECM shall reply with Arb Reject (No Destination).

If the destination phy is in the same expander port as the source phy and is using the direct routing method or the table routing method, the ECM shall reply with either Arb Reject (No Destination) or Arb Reject (Bad Destination); it should reply with Arb Reject (No Destination).

> NOTE 43 - ECMs in expander devices compliant with previous versions of this standard were required to reply with Arb Reject (Bad Destination).

7.12.4.3 Arbitration status

Arbitration status shall be conveyed between expander devices and by expander devices to SAS endpoints using AIP (see 7.2.5.1). This status is used to monitor the progress of connection attempts and to facilitate pathway recovery as part of deadlock recovery.

The arbitration status of an expander phy is set to the last type of AIP received.

Before an expander device transmits AIP, it may have transmitted an OPEN address frame on the same physical link. Arbitration fairness dictates which OPEN address frame wins (see 7.12.3).

After an expander device transmits an AIP, it shall not transmit an OPEN address frame unless it has higher arbitration priority than the incoming connection request.

After transmitting an AIP primitive sequence, and expander device shall transmit at least one other dword (e.g., an idle dword) before transmitting another AIP primitive sequence.

Expander devices shall transmit at least one AIP every 128 dwords while transmitting AIP (NORMAL), AIP (WAITING ON PARTIAL), or AIP (WAITING ON CONNECTION).

> NOTE 44 - Expander devices compliant with previous versions of this standard were not required to transmit three consecutive AIP primitives, as AIP was defined as a single primitive sequence (see 7.2.4.2) rather than as an extended primitive sequence (see 7.2.4.5).

Expander devices shall transmit an AIP (e.g., an AIP (NORMAL)) within 128 dwords of receiving an OPEN address frame.

7.12.4.4 Expander devices

[entire section moved above and modified]