

21 August 2006

06-384r0 SAS-2 OPEN_REJECT RETRY during zoning changes

To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 21 August 2006
Subject: 06-384r0 SAS-2 OPEN_REJECT RETRY during zoning changes

Revision history

Revision 0 (21 August 2006) First revision

Related documents

sas2r05a - Serial Attached SCSI - 2 (SAS-2) revision 5a
06-358 SAS-2 Zone Configuration model (Tim Symons, PMC-Sierra)

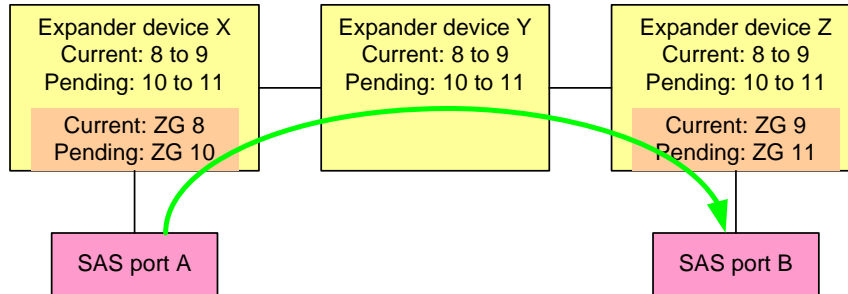
Overview

OPEN_REJECT (ZONE VIOLATION) is an abandon-class OPEN_REJECT. If a SAS port receives it, it is supposed to abandon the connection request and act as if an I_T nexus loss has occurred.

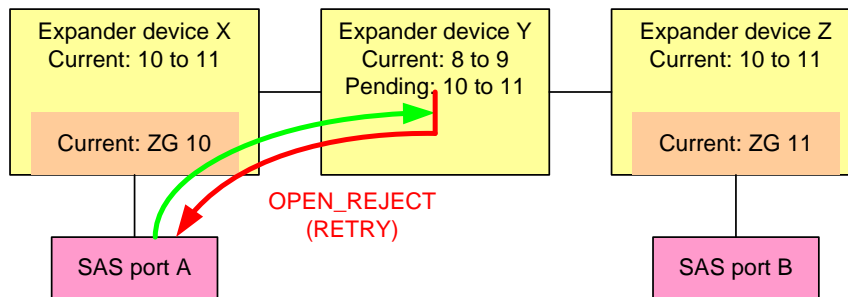
While zoning changes are being activated (per the four-step Lock, Configure, Activate, Unlock process defined in 06-358), the expander devices inside a ZPSDS temporarily but inevitably will have different zone permission table contents. During this time, an expander device that is not activated yet could respond to a connection request with OPEN_REJECT (ZONE VIOLATION) even though the connection request would succeed after that expander device is activated.

For example, assume a phy in zone group 8 has permission to access a phy in zone group 9. The zone manager is reassigning zone group numbers so those phys will change to zone groups 10 and 11. The expander devices that assigns the phys their zone groups is updated (both the zone phy information and the new zone permission table), then one of the phys transmits a connection request. The OPEN address frame is forwarded to another expander that is not yet activated, claiming a source zone group of 10. Since 10 did not have permission to access 11 in the old table, it gets rejected.

Before expander devices X, Y, and Z are activated:



After expander devices X and Z are activated (but Y is not yet):



After expander devices X, Y and Z are activated:

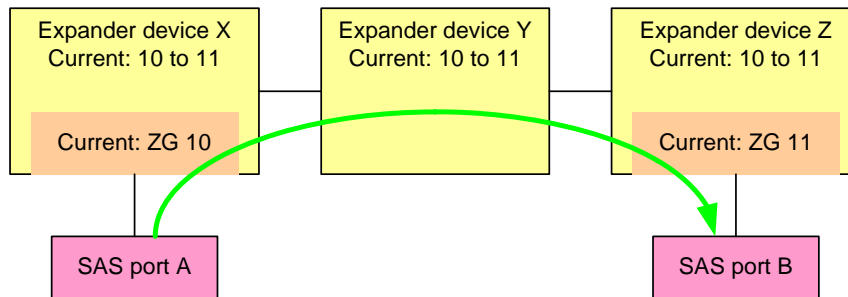


Figure 1 — ZG change example

To avoid this, any locked expander device should return OPEN_REJECT (RETRY) rather than OPEN_REJECT (ZONE VIOLATION). A SAS port that receives OPEN_REJECT (RETRY) will try the connection request again (and not even run its I_T nexus loss timer) and not report failure to the application layer. This way, the I/Os will continue to operate during zoning changes.

There are four cases:

1. If access from A to B is being changed from allowed to allowed (e.g., with different zone group assignments), accesses need to be allowed throughout the zoning change. Issuing some OPEN_REJECT (RETRY)s is OK (as long as zoning updates don't take longer than the I/O timeouts).
2. If access from A to B is being changed from allowed to disallowed, it doesn't matter exactly when the OPEN_REJECT (ZONE VIOLATION) shows up.

3. If access from A to B is being changed from disallowed to allowed, it doesn't matter exactly when the connection requests start going through. If the source port is a SAS initiator port, it can deal with OPEN_REJECT (ZONE VIOLATION) and reissue a similar request later. If the source port is a SAS target port, it shouldn't be trying to access a phy yet.

NOTE 1 - This could still be a problem if A can talk to B but B cannot talk to A. A might send a command to B and then B is stuck. Zone permission tables are required to be symmetric, however, which should prevent this from occurring.

4. If access from A to B is being changed from disallowed to disallowed (e.g. with different zone group assignments), accesses should be blocked throughout the zoning change. Upgrading from OPEN_REJECT (ZONE VIOLATION) to OPEN_REJECT (RETRY) temporarily is OK.

Suggested changes

4.6.6.3 ECM interface

...

Table 1 describes the confirmations from the ECM to an expander phy. These confirmations are sent in confirmation of a Request Path request.

Table 1 — 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: <ul style="list-style-type: none"> 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.
Arbitrating (Blocked On Partial)	Confirmation that the ECM has determined that: <ul style="list-style-type: none"> 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.
Arbitrating (Waiting On Connection)	Confirmation that the ECM has determined that the connection request is blocked due to one of the following reasons: <ul style="list-style-type: none"> 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: <ul style="list-style-type: none"> 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.
Arb Won	Confirmation that an expander phy has won path arbitration.
Arb Lost	Confirmation that an expander phy has lost path arbitration.

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

Message	Description
Arb Reject (Bad Destination)	Confirmation that: a) the requested destination SAS address maps back to the requesting port; b) the requesting port is using the direct routing method or the table routing method; and c) the ECM has not chosen to return Arb Reject (No Destination) (see 7.12.4.3).
Arb Reject (Bad Connection Rate)	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 (No Destination)	Confirmation that: a) there is no operational expander phy capable of routing to the requested destination SAS address; or b) the requested destination SAS address maps back to the requesting port (see 7.12.4.3).
Arb Reject (Pathway Blocked)	Confirmation that the ECM has determined that the requesting expander phy shall back off according to SAS pathway recovery rules.
Arb Reject (Retry)	Confirmation that the ECM has determined that there is a zoning violation (see 4.9.3) while the zoning expander device is locked .
Arb Reject (Zone Violation)	Confirmation that the ECM has determined that there is a zoning violation (see 4.9.3) while the zoning expander device is not locked .

4.9.9.3 Zone permission table

The zone permission table specifies access permission between zone groups. If a bit in the zone permission table is set to one then connection requests shall be permitted between phys in the zone groups. If a bit in the zone permission table is set to zero then connection requests between phys in the zone groups shall be rejected with OPEN_REJECT (ZONE VIOLATION) [or OPEN_REJECT \(RETRY\)](#) as described in 4.9.3.5.

4.9.3.5 Source zone group and destination zone group determination

When a zoning expander device with zoning enabled receives an OPEN address frame (see 7.8.3):

- a) the zone group of the source port (i.e., s) is identified as defined in table 26; and
- b) the zone group of the destination port (i.e., d) is identified as defined in table 27.

If the ZP[s, d] bit is set to one then access between the phys shall be permitted and the zoning expander device shall perform the ECM arbitration procedure. If the ZP[s, d] bit is set to zero then access between the phys is not permitted and the zoning expander device shall transmit [an OPEN_REJECT](#) ~~OPEN_REJECT (ZONE VIOLATION)~~ in response to the connection request: [as follows](#):

- a) [OPEN_REJECT \(RETRY\) if the zoning expander device is locked; and](#)
- b) [OPEN_REJECT \(ZONE VIOLATION\) if the zoning expander device is not locked.](#)

7.2.2 Primitive summary

Table 2 defines the primitives not specific to the type of connection.

Table 2 — Primitives not specific to type of connection

Primitive	Use ^a	From ^b			To ^b			Primitive sequence type ^c
		I	E	T	I	E	T	
OPEN_ACCEPT	NoConn	I		T	I		T	Single
OPEN_REJECT (BAD DESTINATION)	NoConn		E		I		T	Single
OPEN_REJECT (CONNECTION RATE NOT SUPPORTED)		I	E	T				
OPEN_REJECT (NO DESTINATION)			E					
OPEN_REJECT (PATHWAY BLOCKED)			E					
OPEN_REJECT (PROTOCOL NOT SUPPORTED)		I		T				
OPEN_REJECT (ZONE VIOLATION)			E					
OPEN_REJECT (RESERVED ABANDON 1)								
OPEN_REJECT (RESERVED ABANDON 2)								
OPEN_REJECT (RESERVED ABANDON 3)								
OPEN_REJECT (RESERVED CONTINUE 0)								
OPEN_REJECT (RESERVED CONTINUE 1)								
OPEN_REJECT (RESERVED INITIALIZE 0)								
OPEN_REJECT (RESERVED INITIALIZE 1)								
OPEN_REJECT (RESERVED STOP 0)								
OPEN_REJECT (RESERVED STOP 1)								
OPEN_REJECT (RETRY)			I	E				
OPEN_REJECT (STP RESOURCES BUSY)			E	T				
OPEN_REJECT (WRONG DESTINATION)		I		T			T	
SOAF	NoConn	I	E	T	I	E	T	Single

^a The Use column indicates when the primitive is used:
 a) NoConn: SAS physical links, outside connections;
 b) Conn: SAS physical links, inside connections;
 c) All: SAS physical links, both outside connections or inside any type of connection; or
 d) STP: SAS physical links, inside STP connections.

^b The From and To columns indicate the type of ports that originate each primitive or are the intended destinations of each primitive:
 a) I for SAS initiator ports;
 b) E for expander ports; and
 c) T for SAS target ports.
 Expander ports are not considered originators of primitives that are passing through from expander port to expander port.

^c The Primitive sequence type columns indicate whether the primitive is sent as a single primitive sequence, a repeated primitive sequence, a continued primitive sequence, an extended primitive sequence, a triple primitive sequence, or a redundant primitive sequence (see 7.2.4).

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 3 shall result in the originating port abandoning the connection request.

Table 3 — 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 7.12.4.3).
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 4 shall result in the originating port retrying the connection request.

Table 4 — OPEN_REJECT retry primitives

Primitive	Originator	Description
OPEN_REJECT (NO DESTINATION) ^a	Expander phy	Either: a) 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 7.12.4.3); 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 10.4.3.8).
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).
OPEN_REJECT (RESERVED STOP 1) ^b	Unknown	Reserved. Process the same as OPEN_REJECT (PATHWAY BLOCKED).
OPEN_REJECT (RETRY) ^c	Destination phy <u>or zoning expander phy</u>	Phy with destination SAS address exists but is not able to accept connections, <u>or 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.</u>
<p>^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.</p> <p>^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.</p> <p>^c If the I_T Nexus Loss timer (see 8.2.2) is already running, it is stopped.</p>		

7.12.2.2 Results of a connection request

...

A SAS port shall stop the Arbitration Wait Time timer and set it to zero when it receives one of the following connection responses:

- a) OPEN_ACCEPT;
- b) OPEN_REJECT (PROTOCOL NOT SUPPORTED);
- c) OPEN_REJECT (RESERVED ABANDON 1);
- d) OPEN_REJECT (RESERVED ABANDON 2);
- e) OPEN_REJECT (RESERVED ABANDON 3);
- f) OPEN_REJECT (STP RESOURCES BUSY);
- g) OPEN_REJECT (WRONG DESTINATION);
- h) OPEN_REJECT (RESERVED CONTINUE 0);
- i) OPEN_REJECT (RESERVED CONTINUE 1); or
- j) OPEN_REJECT (RETRY).

NOTE 2 - Connection responses that are conclusively from the destination phy (see table 3 and table 4 in 7.2.5.12) are included in the list. Except for OPEN_REJECT (RETRY), c Connection responses that are only from expander phys or may be from expander phys are not included.