

July 1516, 2002

To: T10 Technical Committee
From: Timothy Hoglund (LSI Logic)
Subject: SAS Pathway Recovery modification

The presently specified method for avoiding deadlock and reclaiming blocked pathway resources within a SAS topology requires SAS expanders to perform a two-pass algorithm to insure that deadlock is avoided. This method therefore incurs additional complexity in expander implementations and will cause additional time (and associated performance degradation) to resolve pathway conflicts. The two-pass algorithm is necessary due to the present dependency on comparing Arbitration Wait Time values which in a multiple expander topology introduces uncertainty to the comparison, i.e. a second pass is required should the first pass not resolve the conflict.

This proposal calls for a single-pass method to resolve blocked pathways through removal of Arbitration Wait Time comparisons from the Pathway Recovery algorithm and by introducing a new value to be used. This value, Pathway Blocked Count, is contained in all Open Address frames and specifies the number of times a port has received a OPEN_REJECT(PATHWAY BLOCKED) response to a connection request. Expanders use the Pathway Blocked Count value to determine which connection requests to abort as specified below in 4.8.4.4.3.

Port Layer Management of Pathway Blocked Count (to be added to Port Layer state machines)

The SAS port layer shall initialize the Pathway Blocked Count value to zero when any of the following conditions are met:

- a) upon first connection request to a destination;
- b) an OPEN_REJECT with a reason code not equal to PATHWAY BLOCKED is received;
- c) an OPEN_ACCEPT is received; or
- d) a BREAK is received

The SAS port layer shall increment the Pathway Blocked Count value when OPEN_REJECT(PATHWAY BLOCKED) is received in response to a connection request. The port shall not increment the Pathway Blocked Count value past FFh.

Revisions to current Pathway Recovery text**4.8.4.4.3 Pathway Recovery**

Pathway recovery provides a means to abort connection requests in order to prevent deadlock-
using Pathway Recovery Priority comparisons. Pathway Recovery Priority is defined as the concatenation of PATHWAY BLOCKED COUNT and the SOURCE SAS ADDRESS fields within the OPEN address frame of a blocked connection request. Figure 1 illustrates the components of Pathway Recovery Priority.

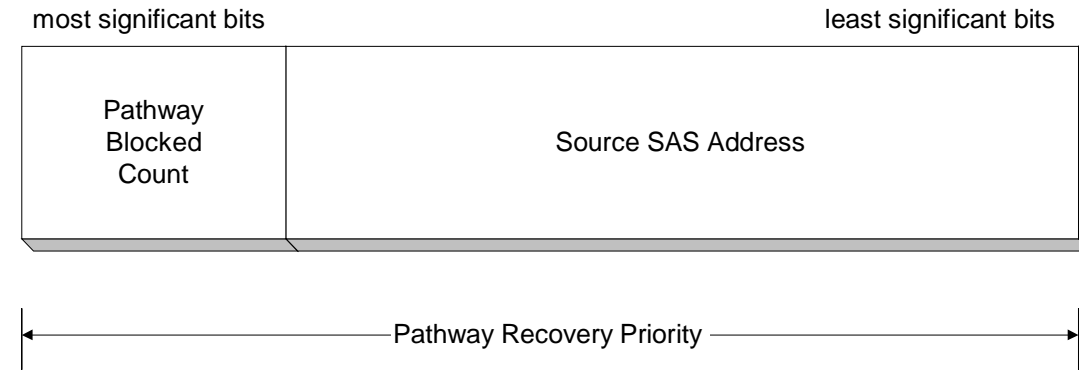


Figure 1: Pathway Recovery Priority

When the PPT timer for an arbitrating link expires (i.e., reaches a value of zero), the Expander Function shall determine whether to continue the connection request or to abort the connection request as follows:

The Expander Function shall instruct the arbitrating link to reject the connection request by transmitting OPEN_REJECT(PATHWAY_BLOCKED) when the PPT timer expires and ~~one of the following conditions are met: the Pathway Recovery Priority of the arbitrating link (link requesting the connection) is less than the Pathway Recovery Priority of all links within the destination port with an arbitration status of WAITING_ON_PARTIAL.~~

~~the Pathway Blocked Count value of the arbitrating link (link requesting the connection) is less than the Pathway Blocked Count value of all links within the destination port with an arbitration status of WAITING_ON_PARTIAL; or~~

- ~~b)a) the Pathway Blocked Count value of the arbitrating link is equal to the Pathway Blocked Count value of all links within the destination port with an arbitration status of WAITING_ON_PARTIAL and the source SAS address of the arbitrating link is less than the source SAS address of all links within the destination port with an arbitration status of WAITING_ON_PARTIAL.~~

Revisions to OPEN address frame

7.5.3 OPEN address frame

[Table 1](#) defines the OPEN address frame format used for connection requests.

Table 1. OPEN address frame format

Byte	7	6	5	4	3	2	1	0
0	INITIATOR	PROTOCOL			ADDRESS FRAME TYPE (1h)			
1	FEATURES				LINK RATE			
2	(MSB)	INITIATOR CONNECTION TAG						(LSB)
3								
4	DESTINATION DEVICE NAME							
11								
12	SOURCE DEVICE NAME							
19								
20	Reserved							
21	PATHWAY BLOCKED COUNT							
22	SCALE	(MSB)	ARBITRATION WAIT TIME					(LSB)
23								
24	(MSB)	Reserved						(LSB)
27								
28	(MSB)	CRC						(LSB)
31								

The PROTOCOL field indicates the protocol for the connection being requested and is defined in [Table 3](#).

Table 32. Protocol

PROTOCOL	Protocol
000b	SMP
001b	SSP
010b	STP
All others	Reserved

The destination device shall reject the connection request with OPEN_REJECT(PROTOCOL NOT SUPPORTED) if the PROTOCOL field is set to a value it does not support.

An INITIATOR bit set to one indicates the source device is acting as an initiator device. An INITIATOR bit set to zero indicates the source device is acting as a target device. If a target/initiator device sets the INITIATOR bit to one, it shall operate only in its initiator role during the connection. If a target/initiator device sets the INITIATOR bit to zero, it shall operate only in its target role during the connection.

If a target/initiator device accepts an OPEN address frame with the INITIATOR bit to one, it shall operate only in its target role during the connection. If a target/initiator device accepts an OPEN address frame with the INITIATOR bit to zero, it shall operate only in its initiator role during the connection.

fixfix

An initiator-only device shall reject connection requests from an initiator device with OPEN_REJECT(PROTOCOL NOT SUPPORTED). A target-only device shall reject connection requests from a target device with OPEN_REJECT(PROTOCOL NOT SUPPORTED).

The LINK RATE field indicates the rate of the link on which the OPEN address frame originated, and the rate at which all links on the pathway must support, and is defined in [Table 5](#).

Table 53. Link rate

LINK RATE	Link rate
0h	1,5 Gbps
1h	3,0 Gbps
2h - Fh	Reserved

The FEATURES field shall be set to zero. The destination device shall reject the connection request with OPEN_REJECT(PROTOCOL NOT SUPPORTED) if the FEATURES field is set to a value it does not support.

The INITIATOR CONNECTION TAG field is used for SSP and STP connection requests to provide an initiator port an easier context lookup when the target port originates a connection request. SSP or STP initiator ports shall set the INITIATOR CONNECTION TAG field to FFFFh if they do not require the field be provided by the target port. Otherwise, an SSP or STP initiator port should set the INITIATOR CONNECTION TAG field to a unique value per target port. When requesting a connection to an initiator port, a target port shall set the INITIATOR CONNECTION TAG field to the value received in connection requests from the initiator port. An initiator port shall use the same tag for all connection requests to the same target port, and may only change the tag when it has no commands outstanding to that target port. Targets are not required to check consistency of the tags in different connection requests from the same initiator port. SMP initiator ports shall set the INITIATOR CONNECTION TAG field to FFFFh.

The destination device shall ignore the contents of reserved fields in the OPEN address frame.

The PATHWAY BLOCKED COUNT field indicates the number of times the port has retried a connection request due to receiving OPEN_REJECT (PATHWAY BLOCKED). The port shall not increment the Pathway Blocked Count value past FFh.

The SCALE bit and the ARBITRATION WAIT TIME field indicate how long the port sending the OPEN address frame has been waiting for a connection request to be accepted, and are defined in [Table 7](#)~~Table 4~~. See **Error! Reference source not found.** for details on arbitration fairness.

Table 7~~Table 4~~. Arbitration wait time

SCALE	ARBITRATION WAIT TIME		
	Description	Minimum value (0000h) meaning	Maximum value (7FFFh) meaning
0b	Microseconds since the connection request was first made	0 μ s	32 767 μ s
1b	Milliseconds since the connection request was first made minus 32 768 microseconds	32 768 μ s	32 768 ms + 32 768 μ s

The DESTINATION DEVICE NAME field indicates the device name of the port with which a connection is being requested.

The SOURCE DEVICE NAME field indicates the device name of the port sending the OPEN address frame.