To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 4 April 2006
Subject: 06-164r1 SAS-2 Require expanders transmit three AIPs

Revision history
Revision 0 (28 March 2006) First revision
Revision 1 (4 April 2006) Corrected error noticed by Andrew Roy (LeCroy) - in option B, the Extended
    primitive sequence only requires one dword between sequences, not three.

Related documents
sas2r03 - Serial Attached SCSI - 2 (SAS-2) revision 3
04-246r1 SAS-1.1 Send AIP multiple times (Rob Elliott, HP)

Overview
Based on 04-246r1, SAS-1.1 added a statement that expander devices *may* transmit three AIPs rather than
one to ensure they get through even if a bit error occurs. This prevents problems in arbitration, where AIP
affects the determination of the winner when OPEN address frames occur in both directions. SAS-1.1
included a NOTE that this behavior might be required (a *shall*) in future versions of the standard.

SAS-2 is such a future version of the standard, so this recommendation should be upgraded to a requirement.

Two approaches are discussed:

    a)   keeping AIP as a Single primitive sequence, and require it be transmitted 3 times in 7.12.4.2
    b)   defining AIP as a new Extended primitive sequence with send-3 receive-1 rules

Suggested changes: Option A (AIP remains a single primitive sequence)

7.12.4.2 Arbitration status

Arbitration status shall be conveyed between expander devices and by expander devices to SAS endpoints
using AIP primitives. 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.

Expander devices shall transmit no more than three consecutive AIPs without transmitting an idle dword.

Expander devices ~~may~~shall transmit three consecutive AIPs to provide better tolerance of errors. Expander
devices shall transmit at least one set of three consecutive AIP every 128 dwords while transmitting AIP
(NORMAL), AIP (WAITING ON PARTIAL), or AIP (WAITING ON CONNECTION).

> NOTE 31 - ~~Future versions of this standard may require thatt e~~Expander devices compliant with previous
> versions of this standard were not required to transmit three consecutive AIPs.

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

Suggested changes: Option B (AIP becomes a new Extended primitive sequence)

7.2.2 Primitive summary

Table 1 defines the primitives not specific to the type of connection.

Table 1 — Primitives not specific to type of connection

Primitive	Use [a]	From [b]			To [b]			Primitive sequence type [c]
		I	**E**	**T**	**I**	**E**	**T**	
AIP (NORMAL)	NoConn		E		I	E	T	~~Single~~ Extended
AIP (RESERVED 0)	NoConn				I	E	T	~~Single~~ Extended
AIP (RESERVED 1)	NoConn				I	E	T	~~Single~~ Extended
AIP (RESERVED 2)	NoConn				I	E	T	~~Single~~ Extended
AIP (RESERVED WAITING ON PARTIAL)	NoConn				I	E	T	~~Single~~ Extended
AIP (WAITING ON CONNECTION)	NoConn		E		I	E	T	~~Single~~ Extended
AIP (WAITING ON DEVICE)	NoConn		E		I	E	T	~~Single~~ Extended
AIP (WAITING ON PARTIAL)	NoConn		E		I	E	T	~~Single~~ Extended
CLOSE (CLEAR AFFILIATION)	STP	I					T	Triple
CLOSE (NORMAL)	Conn	I		T	I		T	Triple
CLOSE (RESERVED 0)	Conn				I		T	Triple
CLOSE (RESERVED 1)	Conn				I		T	Triple
...	...							...

[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, a triple primitive sequence, an extended primitive sequence, or a redundant primitive sequence (see 7.2.4).

Editor's Note 1: Change footnotes of the other primitive tables in the same way

7.2.4.1 Primitive sequences overview

Table 2 summarizes the types of primitive sequences.

Table 2 — Primitive sequences

Primitive sequence type	Number of times the transmitter transmits the primitive to transmit the primitive sequence	Number of times the receiver receives the primitive to detect the primitive sequence	Reference
Single	1	1	7.2.4.2
Repeated	1 or more	1	7.2.4.3
Continued	2 followed by SATA_CONT	1	7.2.4.4
Extended	3	1	7.2.4.x
Triple	3	3	7.2.4.5
Redundant	6	3	7.2.4.6

Any number of ALIGNs and NOTIFYs may be sent inside primitive sequences without affecting the count or breaking the consecutiveness requirements. Rate matching ALIGNs and NOTIFYs shall be sent inside primitive sequences inside of connections if rate matching is enabled (see 7.13).

7.2.4.x Extended primitive sequence

Primitives that form extended primitive sequences (e.g., AIP) shall be transmitted three times consecutively. ALIGNs and NOTIFYs may be sent inside primitive sequences as described in .

A receiver shall detect an extended primitive sequence after the primitive is received one time. After receiving an extended primitive sequence, a receiver shall not detect a second instance of the same extended primitive sequence until it has received one dword that is not any of the following:

    a)   the original primitive; or
    b)   an ALIGN or NOTIFY.

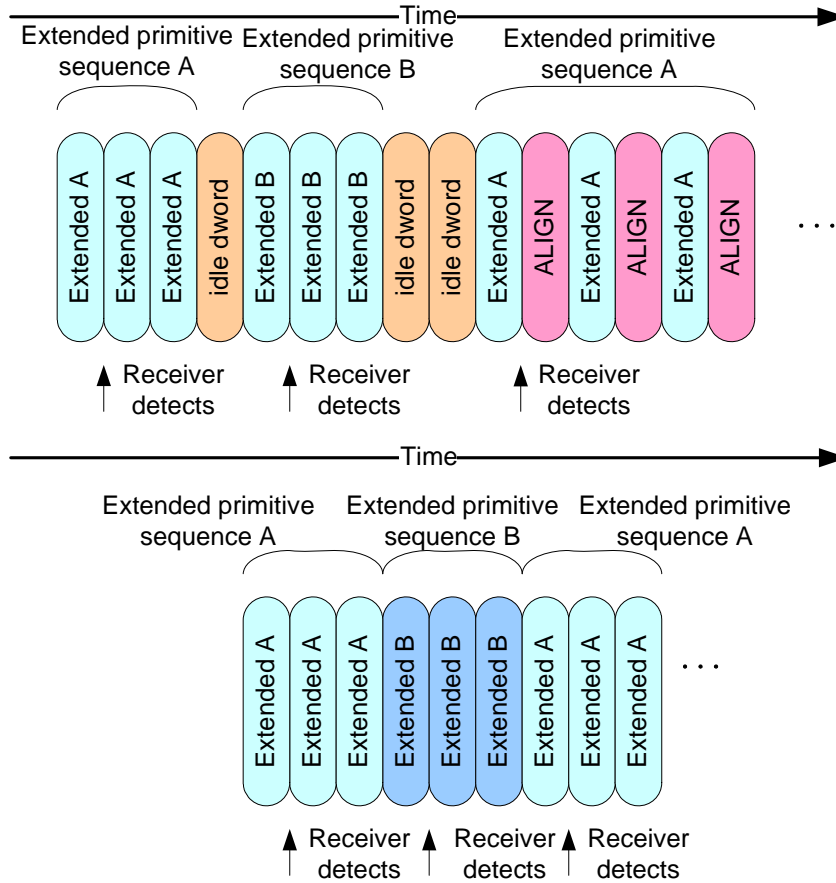Figure 1 shows examples of extended primitive sequences.

Figure 1 — Extended primitive sequence

7.2.5.1 AIP (Arbitration in progress)

AIP is sent by an expander device after a connection request to specify that the connection request is being processed and specify the status of the connection request.

The versions of AIP representing different statuses are defined in table 3.

Table 3 — AIP primitives

Primitive	Description
AIP (NORMAL)	Expander device has accepted the connection request. This may be sent multiple times (see 7.12.4.2).
AIP (RESERVED 0)	Reserved. Processed the same as AIP (NORMAL).
AIP (RESERVED 1)	Reserved. Processed the same as AIP (NORMAL).
AIP (RESERVED 2)	Reserved. Processed the same as AIP (NORMAL).
AIP (WAITING ON CONNECTION)	Expander device has determined the routing for the connection request, but either the destination phys are all being used for connections or there are insufficient routing resources to complete the connection request. This may be sent multiple times (see 7.12.4.2).
AIP (WAITING ON DEVICE)	Expander device has determined the routing for the connection request and forwarded it to the output physical link. This is sent one time (see 7.12.4.2).
AIP (WAITING ON PARTIAL)	Expander device has determined the routing for the connection request, but the destination phys are all busy with other partial pathways. This may be sent multiple times (see 7.12.4.2).
AIP (RESERVED WAITING ON PARTIAL)	Reserved. Processed the same as AIP (WAITING ON PARTIAL).

See 7.12 for details on connections.

7.12.4.2 Arbitration status

Arbitration status shall be conveyed between expander devices and by expander devices to SAS endpoints using AIP ~~primitives~~ (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.

Expander devices shall ~~transmit no more than three~~ not transmit consecutive AIPs without transmitting at least one other dword (e.g., an idle dword) in between the primitive sequences.

~~Expander devices may transmit three consecutive AIPs to provide better tolerance of errors.~~ 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 31 - ~~Future versions of this standard may require thatt e~~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.

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