To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 13 August 2007
Subject: 07-090r2 SAS-2 Transmit IDENTIFY three times

Revision history

Revision 0 (26 April 2007) First revision
Revision 1 (7 May 2007) Incorporated comments from May 2007 SAS protocol WG.
Revision 2 (13 August 2007) Changed SL_IR_IRC to notify the link layer that it is ready to go after transmitting the first IDENTIFY address frame rather than wait until it has transmitted the third.

Related documents

sas2r09a - Serial Attached SCSI - 2 (SAS-2) revision 9a
05-086r0 SAS-1.1 Link layer timeout race conditions (Rob Elliott, HP)
05-094r0 SAS Protocol WG minutes 7 March 2005 (Ralph Weber, ENDL, and John Lohmeyer, LSI)
07-334 SAS-2 Add minimum number of dwords after IDENTIFY frame (Gerry Houlder, Seagate)

Overview

In most cases, SAS recovers fairly gracefully from single-bit errors. If a single bit error occurs in an IDENTIFY address frame, however, the only recovery mechanism available is to rerun the entire link reset sequence (OOB + speed negotiation). Although the IDENTIFY address frame includes a CRC, there is no NAK to trigger retransmission on failed reception.

To better tolerate single-bit errors, the IDENTIFY address frame should be sent 3 times rather than just 1 time. Transmitting 3 times rather than 2 allows for DFE receiver error expansion of a single-bit error on the wire to a short burst of errors in the receiver and allows the frames to be sent back-to-back (SOAF ... EOAF, SOAF ... EOAF, SOAF ... EOAF).

This should not confuse a SAS-1.1 receiver, which is supposed to only honor the first SOAF it sees. As noted by Jeff Gauvin (LSI), the SL_IR state machine in SAS-1.1 (which handles IDENTIFY address frames) differs from the SL_RA state machine (which handles OPEN address frames) in its handling of unexpected address frames and SOAFs before EOAFs. SL_IR only honors the first SOAF it sees, while SL_CC restarts on any SOAF. It is likely that some SL_IR implementations don't make that distinction and restart on each SOAF, to share logic with SL_RA. If they do make the distinction, then SAS-2 to SAS-1.1 just falls back to the SAS-1.1 case of rerunning the link reset sequence if an error occurs.

This proposal:

a) Changes SL_IR_RIF to restart reception of an address frame on any SOAF
b) Changes SL_IR_RIF to keep receiving address frames until it finds a correct IDENTIFY address frame. It does not quit if the CRC is bad.
c) Allows SL_IR_TIF to transmit either one or three IDENTIFY address frames.

Two complaints were lodged against revision 1 of this proposal in the May T10 plenary (where the vote failed):

1. Frames too fast. 07-334 asks that 3 idle dwords be included after each IDENTIFY address frame to avoid triggering a bug in existing SAS-1.1 designs that would be exacerbated by sending three IDENTIFYs back-to-back.

2. Steve Finch complained that there is no maximum time between frames; the transmitter has to understand the receiver timeout and send the frames fast enough to meet it, but there is no advice in the transmitter section about how long to take. However, this is no different than the rest of the standard; on 3/7/2005, discussing proposal 05-086, the SAS protocol WG voted 9-3 to not define any transmitter time limits in SSP state machines and leave the standard vague. Unless this is changed everywhere, then just including time limits in SL_IR could cause confusion (e.g., why are there in SL_IR but not SSP?).

Suggested changes to SAS-2

7.8.2 IDENTIFY address frame

Table 116 defines the IDENTIFY address frame format used for the identification sequence. The IDENTIFY address frame is sent by each logical phy after the phy reset sequence completes if the physical link is a SAS physical link. The IDENTIFY address frames sent by each logical phy in a physical phy shall be identical.

...

7.9 Identification and hard reset sequence

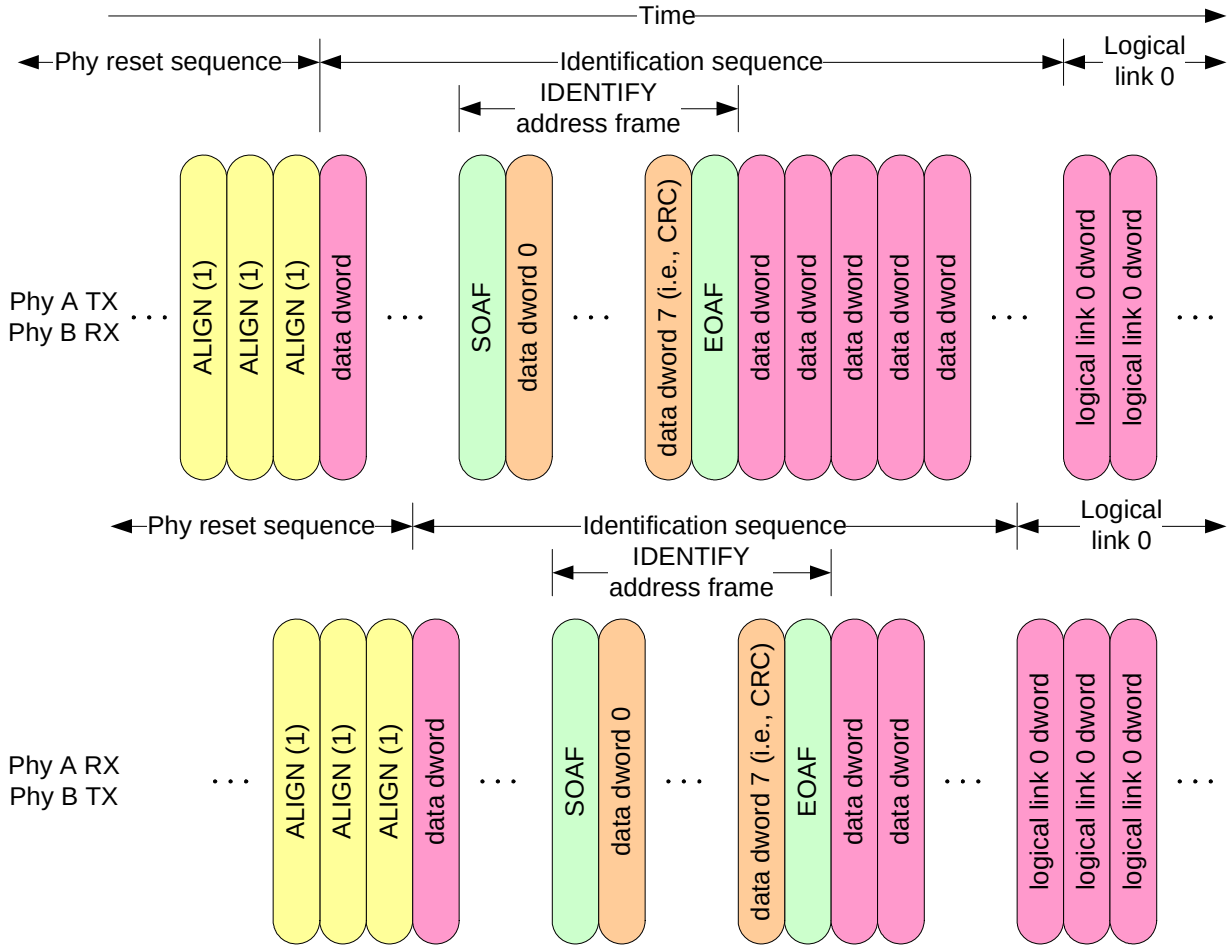
7.9.1 Identification and hard reset sequence overview

After the phy reset sequence (see 6.7) has been completed indicating the physical link is using SAS rather than SATA, each logical phy either:

a) performs and identification sequence by transmitting ~~an~~one or three IDENTIFY address frames (see 7.8.2); or
b) performs a hard reset sequence by transmitting a HARD_RESET primitive sequence (see 7.2.5.9).

If multiplexing (see 6.10) is enabled (see table 86 in 6.7.4.2.3.3), the identification sequence and hard reset sequence begin with a multiplexing sequence (see 6.7.4.3).

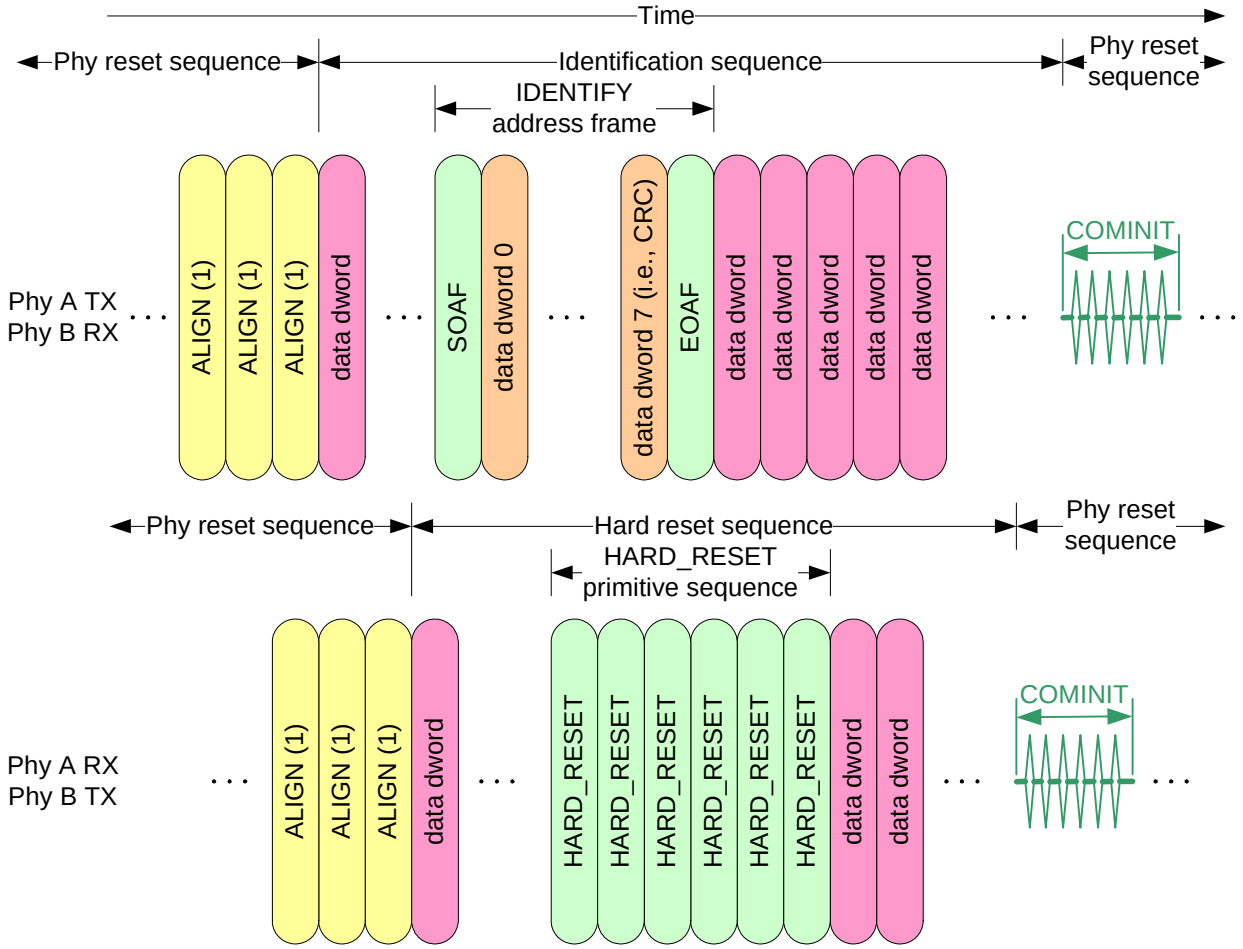
Figure 1 shows two phys with multiplexing disabled performing the identification sequence. Only one IDENTIFY address frame is shown in this example.



NOTE: Phys transmit deletable primitives for clock skew management after the phy reset sequence.

Figure 1 — Identification sequence

Figure 2 shows phy A performing the identification sequence and phy B performing the hard reset sequence. Multiplexing is disabled and only one IDENTIFY address frame is shown in this example.



NOTE: Phys transmit deletable primitives for clock skew management after the phy reset sequence.

Figure 2 — Hard reset sequence

Each logical phy receives an IDENTIFY address frame or a HARD_RESET primitive sequence from the logical phy to which it is attached. The combination of a phy reset sequence, an optional hard reset sequence followed by another phy reset sequence, and an identification sequence is called a link reset sequence (see 4.4.1).

If a phy receives a valid IDENTIFY address frame within 1 ms of phy reset sequence completion, the SAS address in the outgoing IDENTIFY address frame(s) and the SAS address in the incoming IDENTIFY address frame determine the port to which a phy belongs (see 4.1.4). The phy ignores subsequent IDENTIFY address frames and HARD_RESET primitives until another phy reset sequence occurs.

If a phy receives a HARD_RESET primitive sequence within 1 ms of phy reset sequence completion, it shall be considered a reset event and cause a hard reset (see 4.4.2) of the port containing that phy.

If a phy does not receive a HARD_RESET primitive sequence or a valid IDENTIFY address frame within 1 ms of phy reset sequence completion, it shall restart the phy reset sequence.

7.9.2 SAS initiator device rules

After a link reset sequence, or after receiving a Broadcast (Change), a management application client behind an SMP initiator port should perform a discover process (see 4.7).

When a discover process is performed after a link reset sequence, the management application client discovers all the devices in the SAS domain. When a discover process is performed after a Broadcast (Change), the management application client determines which devices have been added to or removed from the SAS domain.

The discover information may be used to select connection rates for connection requests (see 7.8.3).

After receiving a Broadcast (Expander), a management application client behind an SMP initiator port should issue a REPORT GENERAL function (see 10.4.3.3) to all expander devices to determine:

a) the expander devices, if any, that are reducing their functionality (i.e., the REDUCED FUNCTIONALITY bit is set to one in the REPORT GENERAL response)(see 4.6.8); and
b) the amount of time remaining until the reduced functionality occurs (i.e., the contents of the TIME TO REDUCED FUNCTIONALITY field in the REPORT GENERAL response).

7.9.3 Expander device rules

After completing the link reset sequence on a phy and completing internal initialization, the ECM within an expander device shall be capable of routing connection requests through that phy. The expander device may return OPEN_REJECT (NO DESTINATION) until it is ready to process connection requests.

After a link reset sequence, or after receiving a Broadcast (Change), the management application client behind an SMP initiator port in a self-configuring expander device shall follow the SAS initiator device rules (see 7.9.2) to perform a discover process (see 4.7).

The ECM of an externally configurable expander device is dependent on the completion of the discover process (see 4.7) for routing connection requests using the table routing method.

7.9.4 SL_IR (link layer identification and hard reset) state machines

7.9.4.1 SL_IR state machines overview

The SL_IR (link layer identification and hard reset) state machines control the flow of dwords on the physical link that are associated with the identification and hard reset sequences. The state machines are as follows:

a) SL_IR_TIR (transmit IDENTIFY or HARD_RESET) state machine (see 7.9.4.3);
b) SL_IR_RIF (receive IDENTIFY address frame) state machine (see 7.9.4.4); and
c) SL_IR_IRC (identification and hard reset control) state machine (see 7.9.4.5).

The SL_IR state machines send the following messages to the SL state machines (see 7.14) in SAS devices or the XL (see 7.15) state machine in expander devices:

a) Enable Disable SAS Link (Enable); and
b) Enable Disable SAS Link (Disable).

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

Table 1 — SL_IR_IRC timers

Timer	Initial value
Receive Identify Timeout timer	1 ms

Figure 3 shows the SL_IR state machines.

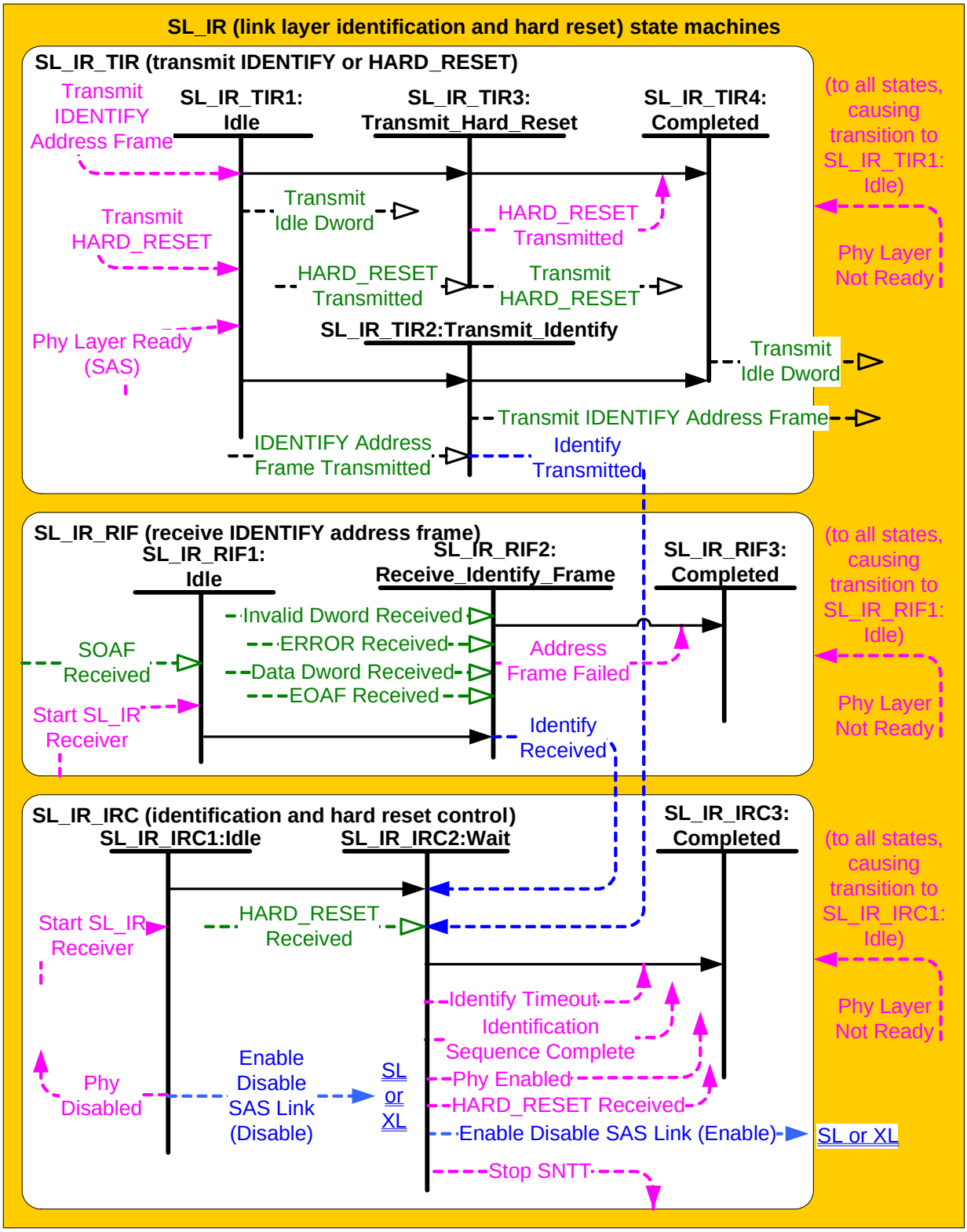


Figure 3 — SL_IR (link layer identification and hard reset) state machines

7.9.4.2 SL_IR transmitter and receiver

The SL_IR transmitter receives the following messages from the SL_IR state machines indicating primitive sequences, frames, and dwords to transmit:

a) Transmit IDENTIFY Address Frame;
b) Transmit HARD_RESET; and
c) Transmit Idle Dword.

The SL_IR transmitter sends the following messages to the SL_IR state machines:

a) HARD_RESET Transmitted; and
b) IDENTIFY Address Frame Transmitted.

The SL_IR receiver sends the following messages to the SL_IR state machines indicating primitive sequences and dwords received from the SP_DWS receiver (see 6.9.2):

a) SOAF Received;
b) Data Dword Received;
c) EOAF Received;
d) ERROR Received;
e) Invalid Dword Received; and
f) HARD_RESET Received.

The SL_IR receiver shall ignore all other dwords.

7.9.4.3 SL_IR_TIR (transmit IDENTIFY or HARD_RESET) state machine

7.9.4.3.1 SL_IR_TIR state machine overview

The SL_IR_TIR state machine's function is to transmit ~~a single IDENTIFY address frame~~one or three IDENTIFY address frames or a HARD_RESET primitive after the phy layer enables the link layer. This state machine consists of the following states:

a) SL_IR_TIR1:Idle (see 7.9.4.3.2)(initial state);
b) SL_IR_TIR2:Transmit_Identify (see 7.9.4.3.3);
c) SL_IR_TIR3:Transmit_Hard_Reset (see 7.9.4.3.4); and
d) SL_IR_TIR4:Completed (see 7.9.4.3.5).

This state machine shall start in the SL_IR_TIR1:Idle state. This state machine shall transition to the SL_IR_TIR1:Idle state from any other state after receiving a Phy Layer Not Ready confirmation.

7.9.4.3.2 SL_IR_TIR1:Idle state

7.9.4.3.2.1 State description

This state shall request idle dwords be transmitted by repeatedly sending Transmit Idle Dword messages to the SL_IR transmitter.

7.9.4.3.2.2 Transition SL_IR_TIR1:Idle to SL_IR_TIR2:Transmit_Identify

This transition shall occur after both:

a) a Phy Layer Ready (SAS) confirmation is received; and
b) a Transmit IDENTIFY Address Frame request is received.

7.9.4.3.2.3 Transition SL_IR_TIR1:Idle to SL_IR_TIR3:Transmit_Hard_Reset

This transition shall occur after both:

a) a Phy Layer Ready (SAS) confirmation is received; and
b) a Transmit HARD_RESET request is received.

7.9.4.3.3 SL_IR_TIR2:Transmit_Identify state

7.9.4.3.3.1 State description

Upon entry into this state, this state shall send ~~a~~either one or three Transmit IDENTIFY Address Frame message~~s~~ to the SL_IR transmitter.

NOTE 1 - Phys compliant with previous versions of this standard only transmitted one Transmit IDENTIFY Address Frame message.

After this state receives an IDENTIFY Address Frame Transmitted message in response to its first Transmit IDENTIFY Address Frame message, this state shall send an Identify Transmitted message to the SL_IR_IRC state machine.

7.9.4.3.3.2 Transition SL_IR_TIR2:Transmit_Identify to SL_IR_TIR4:Completed

If this state sends one Transmit IDENTIFY Address Frame message, this ~~This~~ transition shall occur after sending an Identify Transmitted message to the SL_IR_IRC state machine.

If this state sends three Transmit IDENTIFY Address Frame messages, this transition shall occur after receiving three Identify Transmitted messsages.

7.9.4.3.4 SL_IR_TIR3:Transmit_Hard_Reset state

7.9.4.3.4.1 State description

Upon entry into this state, this state shall send a Transmit HARD_RESET message to the SL_IR transmitter.

After this state receives a HARD_RESET Transmitted message, this state shall send a HARD_RESET Transmitted confirmation to the management application layer.

7.9.4.3.4.2 Transition SL_IR_TIR3:Transmit_Hard_Reset to SL_IR_TIR4:Completed

This transition shall occur after sending a HARD_RESET Transmitted confirmation to the management application layer.

7.9.4.3.5 SL_IR_TIR4:Completed state

This state shall request idle dwords be transmitted by repeatedly sending Transmit Idle Dword messages to the SL_IR transmitter.

7.9.4.4 SL_IR_RIF (receive IDENTIFY address frame) state machine

7.9.4.4.1 SL_IR_RIF state machine overview

The SL_IR_RIF state machine receives an IDENTIFY address frame and checks the IDENTIFY address frame to determine if the frame should be accepted or discarded by the link layer.

This state machine consists of the following states:

a) SL_IR_RIF1:Idle (see 7.9.4.4.2)(initial state);
b) SL_IR_RIF2:Receive_Identify_Frame (see 7.9.4.4.3); and
c) SL_IR_RIF3:Completed (see 7.9.4.4.4).

This state machine shall start in the SL_IR_RIF1:Idle state. This state machine shall transition to the SL_IR_RIF1:Idle state from any other state after receiving a Phy Layer Not Ready confirmation.

7.9.4.4.2 SL_IR_RIF1:Idle state

7.9.4.4.2.1 State description

This state waits for an SOAF to be received from the physical link, indicating an address frame is arriving.

7.9.4.4.2.2 Transition SL_IR_RIF1:Idle to SL_IR_RIF2:Receive_Identify_Frame

This transition shall occur after both:

a) a Start SL_IR Receiver confirmation is received; and
b) an SOAF Received message is received.

7.9.4.4.3 SL_IR_RIF2:Receive_Identify_Frame state

7.9.4.4.3.1 State description

This state receives the dwords of an address frame and the EOAF.

If this state receives an SOAF Received message, then this state shall ~~discard the address frame (i.e., the subsequent Data Dword Received and EOAF Received messages) and~~ discard any previously received dwords for the address frame, send an Address Frame Failed confirmation to the management application layer to indicate that an invalid IDENTIFY address frame was received, and start receiving the new address frame.

If this state receives more than eight Data Dword Received messages after an SOAF Received message and before an EOAF Received message, then this state shall discard the address frame and send an Address Frame Failed confirmation to the management application layer to indicate that an invalid IDENTIFY address frame was received.

If this state receives an Invalid Dword Received message or an ERROR Received message after an SOAF Received message and before an EOAF Received message, then this state shall:

a) ignore the invalid dword or ERROR; or
b) discard the address frame and send an Address Frame Failed confirmation to the management application layer to indicate that an invalid IDENTIFY address frame was received.

After receiving an EOAF Received message, this state shall check if it the received frame is a valid IDENTIFY address frame.

This state shall accept an IDENTIFY address frame and send an Identify Received message to the SL_IR_IRC state machine if:

a) the ADDRESS FRAME TYPE field is set to Identify;
b) the number of bytes between the SOAF and EOAF is 32; and
c) the CRC field contains a good CRC.

Otherwise, this state shall discard the ~~IDENTIFY~~ address frame and send an Address Frame Failed confirmation to the management application layer to indicate that an invalid ~~IDENTIFY~~ address frame was received.

7.9.4.4.3.2 Transition SL_IR_RIF2:Receive_Identify_Frame to SL_IR_RIF3:Completed

This transition shall occur after sending an Identify Received message ~~or Address Frame Failed~~ confirmation.

7.9.4.4.4 SL_IR_RIF3:Completed state

This state waits for a Phy Layer Not Ready confirmation.

7.9.4.5 SL_IR_IRC (identification and hard reset control) state machine

7.9.4.5.1 SL_IR_IRC state machine overview

The SL_IR_IRC state machine ensures that IDENTIFY address frames have been both received and transmitted before enabling the rest of the link layer, and notifies the link layer if a HARD_RESET primitive sequence is received before an IDENTIFY address frame has been received.

This state machine consists of the following states:

a) SL_IR_IRC1:Idle (see 7.9.4.5.2)(initial state);
b) SL_IR_IRC2:Wait (see 7.9.4.5.3); and

c) SL_IR_IRC3:Completed (see 7.9.4.5.4).

This state machine shall start in the SL_IR_IRC1:Idle state. This state machine shall transition to the SL_IR_IRC1:Idle state from any other state after receiving a Phy Layer Not Ready confirmation.

7.9.4.5.2 SL_IR_IRC1:Idle state

7.9.4.5.2.1 State description

This state waits for the link layer to be enabled. Upon entry into this state, this state shall:

a) send an Enable Disable SAS Link (Disable) message to SL state machines (see 7.14) or XL state machine (see 7.15) halting any link layer activity; and
b) send a Phy Disabled confirmation to the port layer and the management application layer indicating that the phy is not ready for use.

7.9.4.5.2.2 Transition SL_IR_IRC1:Idle to SL_IR_IRC2:Wait

This transition shall occur after a Start SL_IR Receiver confirmation is received.

7.9.4.5.3 SL_IR_IRC2:Wait state

7.9.4.5.3.1 State description

This state ensures that an IDENTIFY address frame has been received by the SL_IR_RIF state machine and that an IDENTIFY address frame has been transmitted by the SL_IR_TIR state machine before enabling the rest of the link layer. The IDENTIFY address frames may be transmitted and received on the physical link in any order.

After this state receives an Identify Received message, it shall send a Stop SNTT request to the phy layer.

After this state receives an Identify Transmitted message, it shall initialize and start the Receive Identify Timeout timer. If an Identify Received message is received before the Receive Identify Timeout timer expires, this state shall:

a) send an Identification Sequence Complete confirmation to the management application layer, with arguments carrying the contents of the incoming IDENTIFY address frame;
b) send an Enable Disable SAS Link (Enable) message to the SL state machines (see 7.14) in a SAS logical phy or the XL state machine (see 7.15) in an expander logical phy indicating that the rest of the link layer may start operation; and
c) send a Phy Enabled confirmation to the port layer and the management application layer indicating that the phy is ready for use.

If the Receive Identify Timeout timer expires before an Identify Received message is received, this state shall send an Identify Timeout confirmation to the management application layer to indicate that an identify timeout occurred.

If this state receives a HARD_RESET Received message before an Identify Received message is received, this state shall send a HARD_RESET Received confirmation to the port layer and the management application layer and a Stop SNTT request to the phy layer.

If this state receives a HARD_RESET Received message after an Identify Received message is received, the HARD_RESET Received message shall be ignored.

7.9.4.5.3.2 Transition SL_IR_IRC2:Wait to SL_IR_IRC3:Completed

This transition shall occur after sending a HARD_RESET Received confirmation, Identify Timeout confirmation, or an Identification Sequence Complete and an Phy Enabled confirmation.

7.9.4.5.4 SL_IR_IRC3:Completed state

This state waits for a Phy Layer Not Ready confirmation.

7.14.3 SL_RA (receive OPEN address frame) state machine

The SL_RA state machine's function is to receive address frames and determine if the received address frame is an OPEN address frame and whether or not it was received successfully. This state machine consists of one state.

This state machine receives SOAFs, dwords of an OPEN address frames, and EOAFs.

This state machine shall ignore all messages except SOAF Received, Data Dword Received, and EOAF Received.

If this state machine receives a subsequent SOAF Received message after receiving an SOAF Received message but before receiving an EOAF Received message, then this state machine shall discard the Data Dword Received messages received before the subsequent SOAF Received message.

If this state machine receives more than eight Data Dword Received messages after an SOAF Received message and before an EOAF Received message, then this state machine shall discard the address frame.

If this state machine receives an Invalid Dword Received message or an ERROR Received message after an SOAF Received message and before an EOAF Received message, then this state machine shall:

a) ignore the invalid dword or ERROR; or
b) discard the address frame.

After receiving an EOAF Received message, this state machine shall check if the address frame is a valid OPEN address frame.

This state machine shall accept an address frame if:

a) the ADDRESS FRAME TYPE field is set to Open;
b) the number of data dwords between the SOAF and EOAF is 8; and
c) the CRC field contains a good CRC.

Otherwise, this state machine shall discard the address frame. If the frame is not discarded then this state machine shall send a OPEN Address Frame Received message to the SL_CC0:Idle state and the SL_CC1:ArbSel state with an argument that contains all the data dwords received in the OPEN address frame.