

To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 27 June 2006
Subject: 06-263r3 SAS-2 Spread-spectrum clocking

Revision history

Revision 0 (31 May 2006) First revision

Revision 1 (14 June 2006) SAS disk drives and HBAs only need to choose between to transmit with down-spreading SSC or no SSC; there is no need for them ever to transmit faster than that. Expander phys are the only phys that must transmit with up-spreading/center-spreading SSC or no SSC to ensure they are always at least as fast as a non-SSC phy originating dwords. All SAS phys must be able to receive up-spreading/center-spreading SSC, down-spreading SSC, and no SSC. HBAs can get by with two clock trees, disk drives with one clock tree, and only expanders must suffer with three clock trees.

Revision 2 (22 June 2006) Incorporated comments from 15 June 2006 SAS physical WG teleconference.

Revision 3 (27 June 2006) Incorporated comments from 23 June 2006 SAS physical WG teleconference.

Related documents

sas2r04 - Serial Attached SCSI - 2 (SAS-2) revision 4

05-381 SAS-2 Multiplexing (Rob Elliott, HP)

06-064 SAS-2 SSC Investigation (Barry Olawsky, HP)

06-129 Spread Spectrum Clocking (SSC) consideration list (Alvin Cox, Seagate)

06-192 Spread spectrum clocking considerations (Harvey Newman, Infineon)

06-193 Symmetrical SSC in SAS-2 physical interface (Yuriy Greshishchev, PMC-Sierra)

06-246 SAS-2 Physical WG minutes 9 May 2006 (Alvin Cox, Seagate)

06-267 SAS-2 Spread spectrum clocking options (Kevin Witt, Greg Tabor, and Adrian Robinson, Vitesse)

06-275 SAS-2 ALIGN insertion rate during STP connections (Rob Elliott, HP)

06-295 SAS-2 Speed negotiation (Amr Wassal and Robert Watson, PMC-Sierra)

06-301 SAS-2 PHYSICAL address frame (Rob Elliott, HP)

Background reading

Spread Spectrum Clock Generation for the Reduction of Radiated Emissions by Keith Hardin, John Fessler, and Donald Bush of Lexmark International. IEEE International Symposium on Electromagnetic Compatibility, August 1994. Available at <http://www.lexmark.com/vgn/images/portal/emc1994.pdf>.

Notes on SSC and Its Timing Impacts by Michael Zhang, Intel, February 1998.

Why Spread Spectrum Clocking of Computing Devices is Not Cheating by Harry Skinner and Kevin Slattery, Intel. 2001 IEEE International Symposium on Electromagnetic Compatibility, August 2001.

Design for EMI. Intel Application Note AP-589, February 1999.

A Spread Spectrum Clock Generator for SATA-II by Wei-Ta Chen, Jen-Chien Hsu, Hong-Wen Line, and Chau-chin Su. IEEE International Symposium on Circuits and Systems (ISCAS 2005), May 2005.

Spread Spectrum Clock Generator with Delay Cell Array to Reduce Electromagnetic Interference by Jonghoon Kim, Dong Gun Kam, Pil Jung Jun, and Joungho Kim. IEEE Transactions on Electromagnetic Compatibility, November 2005.

See <http://www.lexmark.com/sscg> for a variety of papers and licensing information for Lexmark (a spin-off of IBM) SSCG patents including:

Electronic apparatus having low radio frequency interference from system clock signal. US patent 4,507,796 by David Stumfall, Printronix, granted 26 March 1985.

Spread Spectrum Clock Generator and associated method. US patents 5,872,802 and 5,867,524 (continuations of 5,488,627 and 5,631,920) by Keith Hardin and others, Lexmark International, granted 2 February 1999.

Overview

Serial ATA has always supported spread-spectrum clocking (SSC) with a range of +0/-5000 ppm (down-spreading) added to the clock frequency tolerance of +350/-350 ppm, resulting in an overall clock frequency tolerance of +350/-5350 ppm. This is why SATA requires phys to ensure that 2/256 dwords (0.7812%) are ALIGNs.

NOTE 1 - Some SATA devices and SATA hosts do not support SSC; these are technically non-compliant and there is no standard way to identify them. SATA disk vendors who have released SSC drives seem to have all backpedaled and default to non-SSC as of mid 2006. One SATA disk vendor provides a utility to enable/disable SSC that must be run on a system that supports SSC to disable SSC, and may be run on any system to enable it. Other vendors include jumpers on their disk drives.

Serial Attached SCSI, on the other hand, does not support SSC and has a clock frequency tolerance of +100/-100 ppm. This is why SAS only requires phys originating data to ensure 1/2048 dwords (0.0488%) are ALIGNS. Expanders forwarding data do not ensure that number is met on their transmitting phys; the receiving phy deletes all incoming ALIGNs and the transmitting phy adds them back as needed (i.e., when it underflows).

SAS devices never transmit with SSC, nor do they transmit enough ALIGNs to be carried over a physical link beyond an expander device using down-spreading SSC.

SAS HBAs and expanders usually support receiving with down-spreading SSC since they support attachment to SATA devices. SAS disk drives and tape drives do not usually support receiving with SSC, since they do not support attachment to SATA hosts.

Figure 1 shows how SAS-1.1 phys supports both SATA phys that transmit with down-spreading SSC and SAS-1.1 phys that transmit with no SSC.

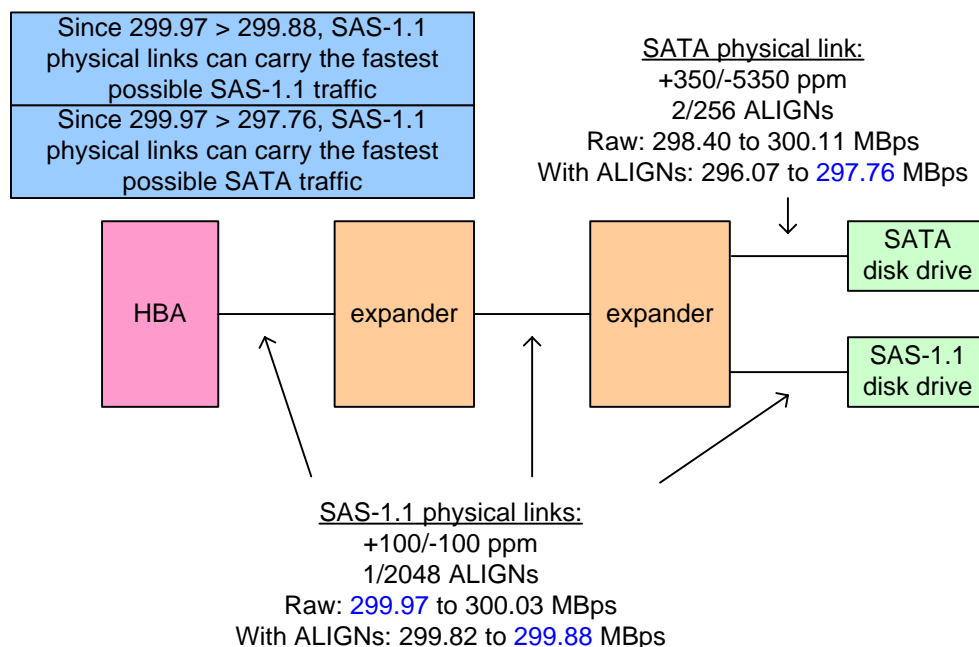


Figure 1 — SAS-1.1 support of SATA SSC

As discussed in 06-064, SSC for SAS would be helpful in reducing emissions at 3 Gbps and even more helpful (possibly necessary) at 6 Gbps. The main problems are:

- SAS-1 and SAS-1.1 phys in disk drives are not required to receive with SSC;
- SAS-1.1 and SAS-1.1 phys do not insert enough ALIGNs in SSP and SMP connections to travel on physical links with down-spreading SSC.

Table 1 shows 3 Gbps unit interval (UI), the raw bandwidth, and the remaining bandwidth after certain numbers of ALIGNs are inserted for different spreading options. For simplicity, only the 3 Gbps physical link rate is discussed.

Table 1 — 3 Gbps UI, raw bandwidth, and bandwidth after ALIGNs

Phy	UI	Raw bandwidth	Bandwidth after ALIGNs
SAS 3 Gbps +100/-100 ppm 1/2048 ALIGNs	333.30 to 333.37 ps	299.97 to 300.03 MBps	299.82 to 299.88 MBps
SATA 3 Gbps +100/-5100 ppm 2/256 ALIGNs	333.30 to 335.03 ps	298.48 to 300.03 MBps	296.15 to 297.69 MBps
SATA 3 Gbps +350/-5350 ppm 2/256 ALIGNs	333.22 to 335.12 ps	298.40 to 300.11 MBps	296.07 to 297.76 MBps
SAS phy +5100/-100 ppm 2/256 ALIGNs	331.63 to 333.37 ps	299.97 to 301.54 MBps	297.63 to 299.18 MBps
SAS phy +5100/-100 ppm 4/256 ALIGNs			295.28 to 296.83 MBps

Down-spreading

Figure 2 shows the problem with adding down-spreading SSC (e.g., +0/-5000 ppm) to SAS-2. This works fine for SATA phys, which also down-spread, but fails if there is a SAS-1.1 phy in the SAS domain that does not support SSC. If such a phy is attached anywhere in the SAS domain, its traffic overflows any SAS physical link (either expander-to-expander or expander-to-end device) implementing down-spreading SSC. Down-spreading SSC would have to be disabled everywhere if any non-SSC device were added to the SAS domain.

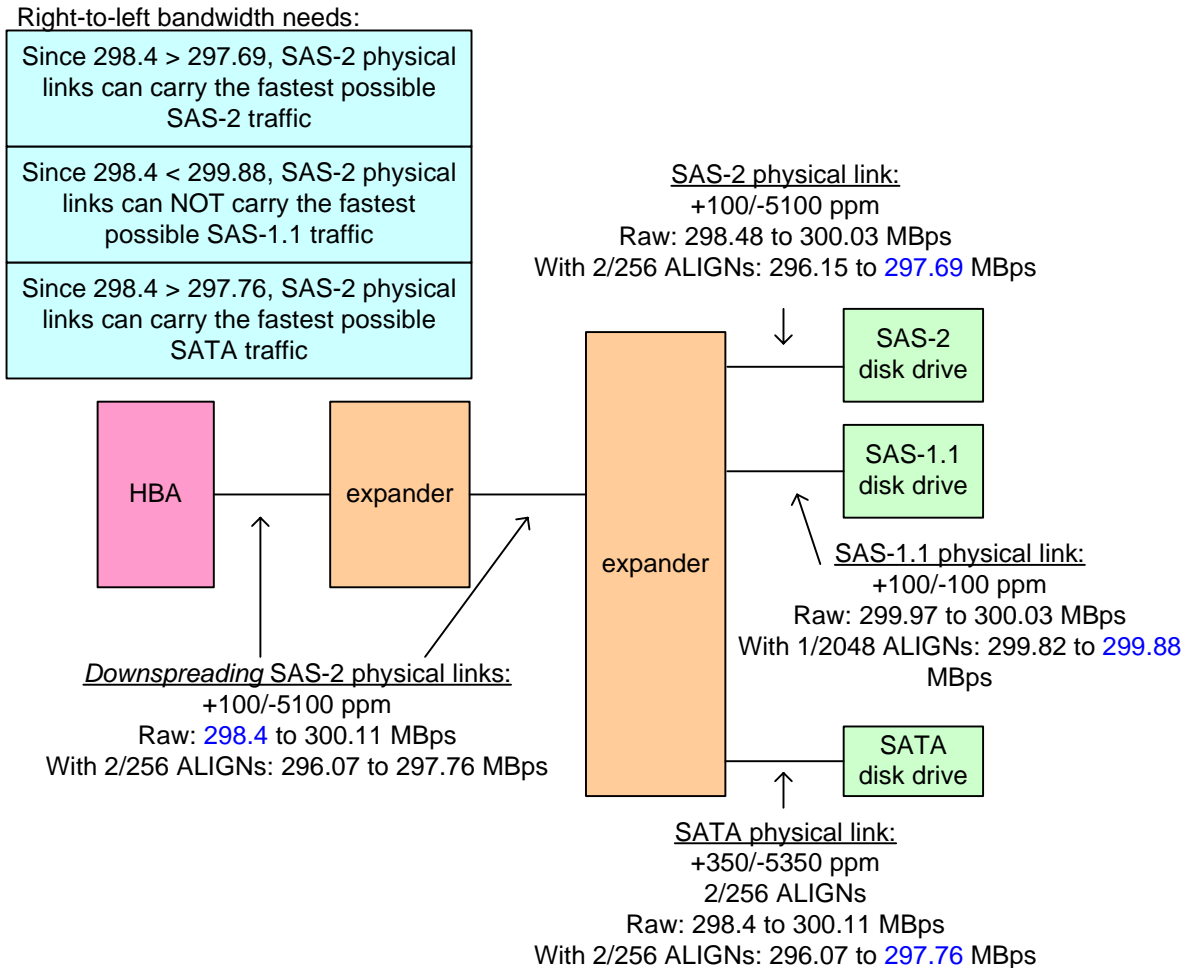


Figure 2 — Down-spreading SSC in SAS-2

Center-spreading

Center-spreading (e.g., +5000/-5000 ppm) has some of the same problems as down-spreading, since the SAS-2 physical link runs for periods of time slower than a SAS-1.1 physical link. Although it also runs for equivalent periods of time faster than a SAS-1.1 physical link, averaging out to the same rate, an expander must buffer dwords that accumulate during the slow period.

At 3 Gbps, there are $0.5 * (3,000,000,000 \text{ bits per second} / 30,000 \text{ per second}) = 50,000 \text{ bits} = 1250 \text{ dwords}$ during the slow period and 1250 dwords during the fast period.

In 06-267, Kevin Witt (Vitesse) calculated this means about 7 dwords will accumulate, worst case (assuming a triangular modulation pattern), which is a tolerable number to buffer in an expander. Expanders already must have ~ 32 dword buffers available for STP flow control; this expands those buffer sizes a bit.

One advantage of center-spreading over up-spreading is more widespread availability of clock generators. Down-spreading and center-spreading clock generators are common; up-spreading clock generators are not.

Up-spreading

Up-spreading SSC (e.g., +5000/-0 ppm) ensures that SAS-2 physical links will always be equal or faster than any SAS-1.1, SAS-2, or SATA physical links. Figure 3 shows up-spreading SSC.

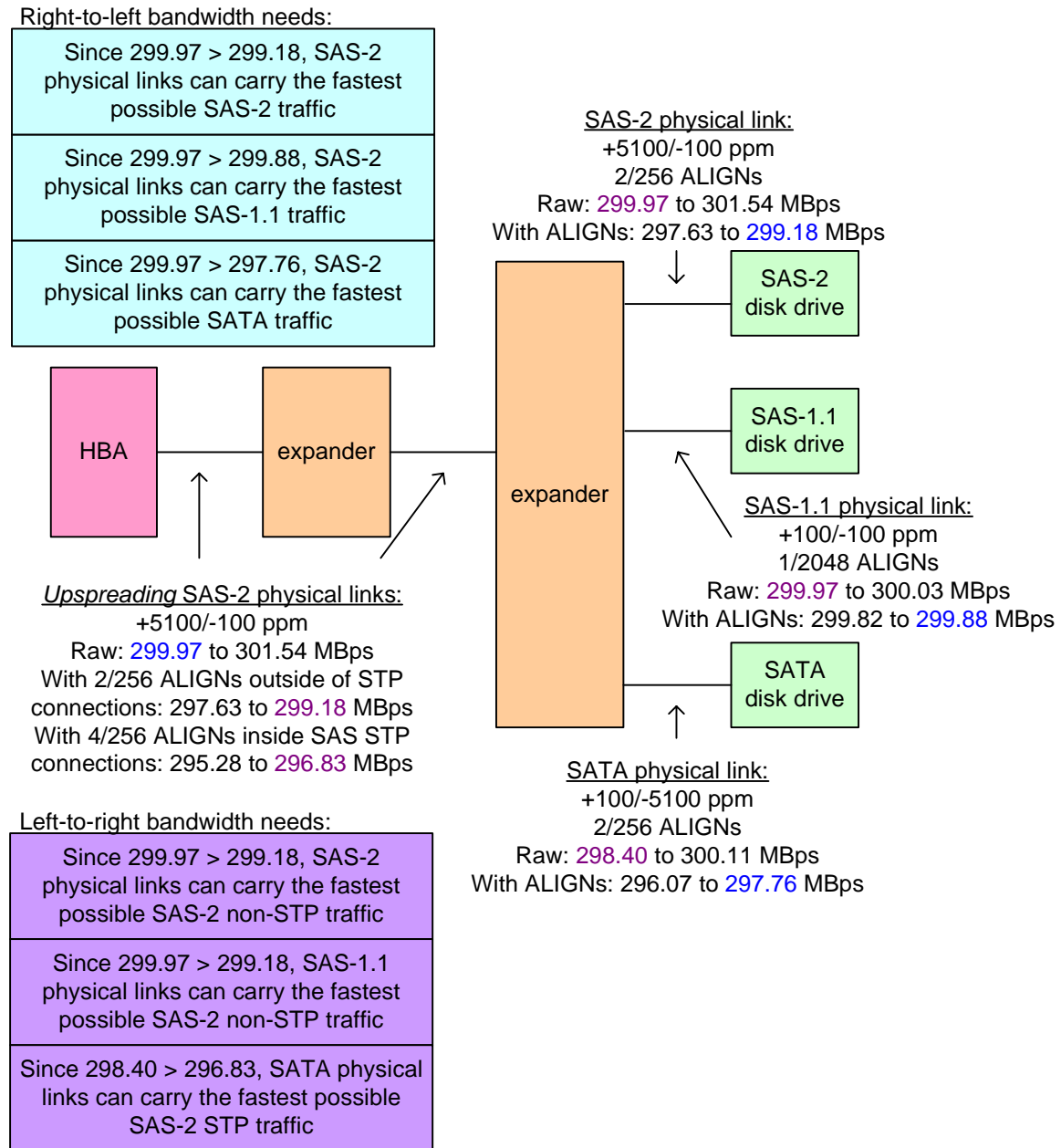


Figure 3 — Up-spreading SSC in SAS-2

Extra ALIGNs are needed to accomodate SSC.

Summary of proposed rules

Overall summary:

- Center-spreading and down-spreading SSC for SAS physical links is supported at 6 Gbps, 3 Gbps, or 1.5 Gbps
- Down-spreading SSC for SATA physical links is supported at any physical link rate

Receiver rules summary:

- a) SAS phys and expander phys shall support receiving with:
 - A) +2600/-2600 ppm SSC (i.e., center-spreading) from expander phys supporting SSC;
 - B) +100/-100 ppm (i.e., none) from SAS phys and expander phys not supporting SSC;
 - C) +350/-350 ppm (i.e., none) from SATA phys not supporting SSC, for non-disk drive SAS phys and expander phys;
 - D) +100/-5100 ppm SSC (i.e., down-spreading) from SAS phys and expander phys supporting SSC, for disk drive SAS phys; and
 - E) +350/-5350 ppm SSC (i.e., down-spreading) from SATA phys, for non-disk drive SAS phys and expander phys.
- b) SAS phys and expander phys shall support receiving OOB signals with SSC
- c) SAS phys and expander phys shall support receiving with SSC during both SAS and SATA speed negotiation

Table 2 summarizes the receiver rules.

Table 2 — Receiver phy rules summary

Input clock	SAS initiator phy or expander phy attached to SAS	SAS initiator phy or expander phy attached to SATA	SAS disk drive phy
+2600/-2600 ppm SSC (i.e., center-spreading)	yes	no	yes
+100/-100 ppm (i.e., none)	yes	no	yes
+350/-350 ppm (i.e., none)	no	yes	no
+100/-5100 ppm SSC (i.e., down-spreading)	yes	no	yes
+350/-5350 ppm SSC (i.e., down-spreading)	no	yes	no

Transmitter rules summary:

- a) Non-expander phys (i.e., end devices) shall transmit with:
 - A) +100/-5100 SSC (i.e., down-spreading) to SAS phys supporting SSC or SATA phys; and
 - B) +100/-100 (i.e., none) otherwise.

NOTE 2 - There is no reason for an end device to ever transmit fast. It is originating all the dwords it transmits, so can throttle them as it wants.

- b) Expander phys shall transmit with:
 - A) +2600/-2600 SSC (i.e., center-spreading) to SAS phys and expander phys supporting SSC; or
 - B) +100/-100 ppm (i.e., none) to SAS phys and expander phys not supporting SSC; and
 - C) +100/-5100 SSC (i.e., down-spreading) to SATA phys;

NOTE 3 - This ensures they can always carry traffic from non-SSC end devices.

Table 3 summarizes the transmitter rules.

Table 3 — Transmitter phy rules summary

Output clock	SAS phy	Expander phy
+2600/-2600 ppm SSC (i.e., center-spreading)	no	yes
+100/-100 ppm (i.e., none)	yes	yes
+100/-5100 ppm SSC (i.e., down-spreading)	yes	yes

Transmitter ALIGN (deletable primitive) insertion rules when SSC is enabled summary:

- a) phys originating dwords shall insert 4/256 ALIGNs (i.e., 1/64) inside SSP connections;
- b) phys originating dwords shall insert 4/256 ALIGNs (i.e., 1/64) inside SMP connections;
- c) phys originating dwords shall insert 4/256 ALIGNs (i.e., 1/64) outside of connections;
- d) phys originating dwords shall insert 4/256 ALIGNs (i.e., 1/64) inside STP connections;

NOTE 4 - Rules are kept equal for simplicity. Phys outside connections could use 2/256, since they only worry about the difference to nominal. SAS phys in STP connections could use 2/256 since they transmit down-spreading or nominal, the same as the SATA device. Center-spreading expander phys originating dwords in STP connections (e.g., during HOLD/HOLDA flow control) need to use 4/256 (+2600 to -5350).

Editor's Note 1: 4/256 was chosen to keep everything a power of two. See note in proposed text below for further discussion.

- e) **Implementation tip:** phys originating dwords should include separate programmable ALIGN insertion rates for each of these cases, in case anything else changes:
 - A) outside connections;
 - B) inside SSP connections;
 - C) inside SMP connections; and
 - D) inside STP connections;

NOTE 5 - Combining SSP/SMP is probably safe.

- f) expander phys forwarding (i.e., not originating) dwords shall insert ALIGNs only as necessary (i.e., whenever the dword stream underflows);
- g) expander phys forwarding dwords with rate matching shall insert ALIGNs only when necessary, not blindly inserting them every 1 of 2 (e.g., for 6 Gbps to 3 Gbps) or 3 of 4 dwords (e.g., for 6 Gbps to 1.5 Gbps).

Phy reset sequence rules summary:

- a) Transmit OOB signals without SSC
- b) Transmit SATA speed negotiation with SSC
- c) Transmit SAS speed negotiation windows (SNW) G1, G2, and G3 without SSC
- d) G3 window is not just for 6 Gbps any more
- e) G3 window consists of a 1.5 Gbps speed negotiation followed by phy feature support information
 - A) rates supported: 1.5, 3, 6, etc.
 - B) whether the phy supports receiving with SSC;
- f) Run the final SNW at the highest rate supported by both phys
- g) If both phys support SSC, enable transmission with SSC before starting the Final SNW
 - A) continue transmitting with SSC enabled until transmitting another OOB signal (e.g. after loss of dword synchronization, or start of any new link reset sequence)

Complications

This requires an end device ASIC to have 2 clock trees, with each phy being able to select from +100/-5100 ppm SSC and +100/-100 ppm based on the capabilities of the attached phy.

This requires an expander ASIC to have 3 clock trees. Figure 4 shows the problem.

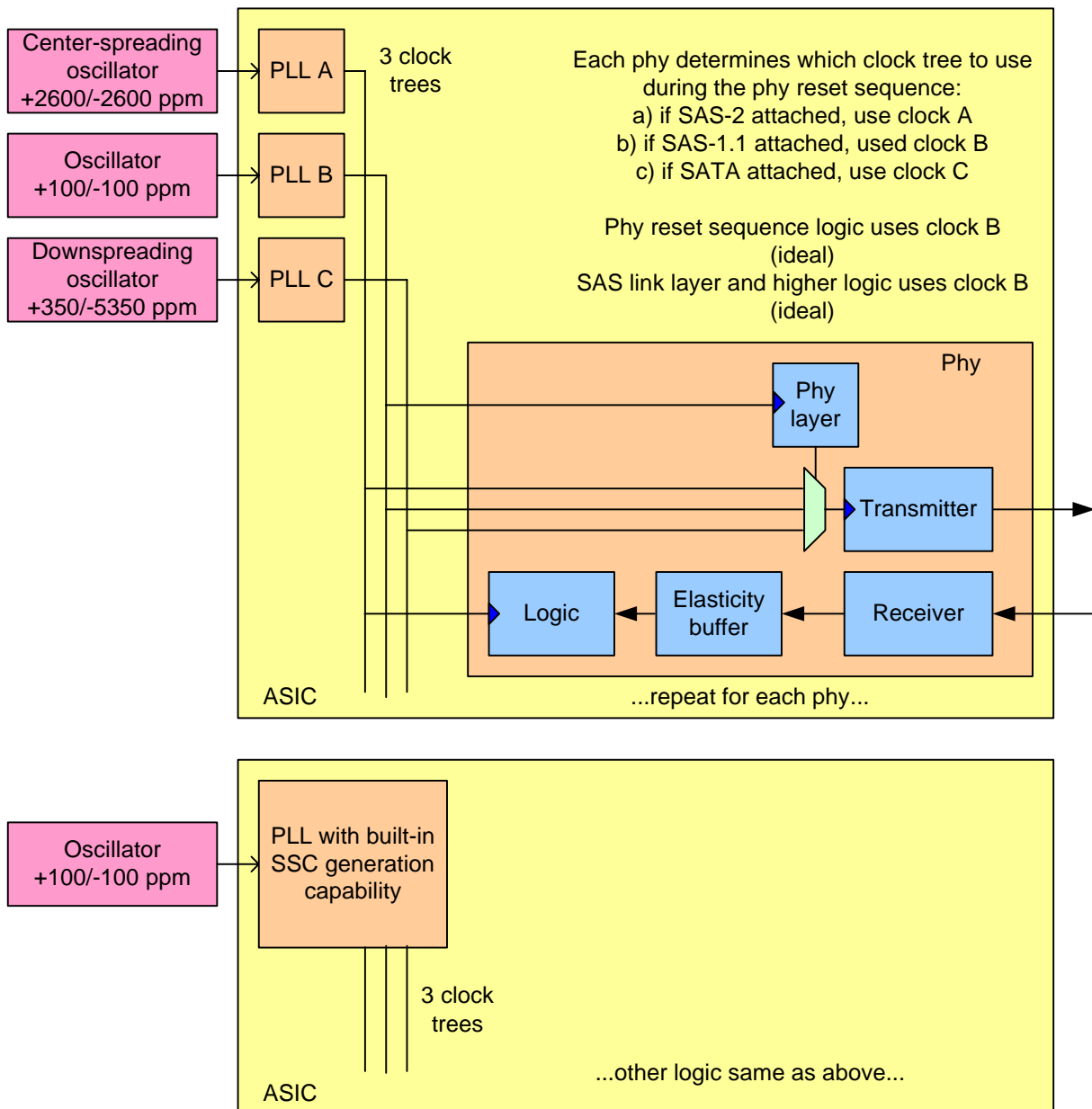


Figure 4 — Three clock trees

Disk drive phys can always transmit with down-spreading SSC. Since SATA devices have that luxury, SAS devices are not treading new ground by doing so. SAS initiator phys and expander phys supporting SATA attachment are supposed to be prepared for this already.

Suggested changes to SAS-2**Suggested changes to chapter 3 (Definitions)**

3.1.xxx electromagnetic interference (EMI): Any electromagnetic disturbance that interrupts, obstructs, toher otherwise degrades or limits the effective performance of electronics/electrical equipment.

3.1.220 spread spectrum clocking (SSC): The technique of modulating the operating frequency of a transmitted signal (i.e., the physical link rate) to reduce the measured peak amplitude of radiated emissions. See 5.3.8.

3.2 Symbols and abbreviations

...

SSC spread spectrum clocking (see 3.1.220)
EMI electromagnetic interference (see 3.1.xxx)

...

Suggested changes to chapter 4 (Model)**4.3.2 Transmit data path**

Figure 32 shows the transmit data path in a SAS phy.

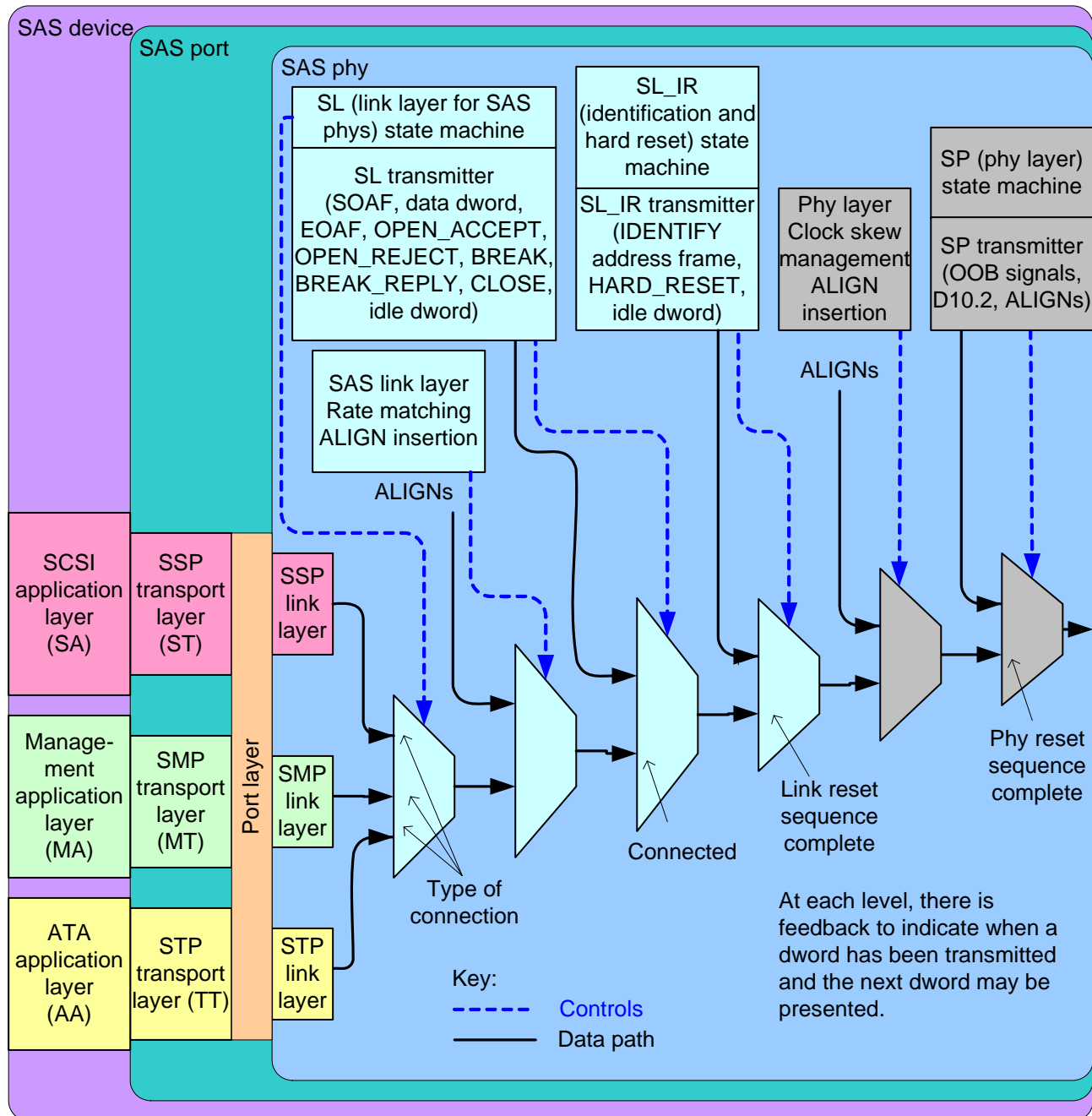


Figure 32 — Transmit data path in a SAS phy

Editor's Note 2: In figure 32, change "Clock skew management" to "Clock frequency difference tolerance" (or whatever term matches section 7.3). Note that 05-381(Multiplexing) also makes changes to this figure.

...

Figure 33 shows the transmit data path for the STP link layer and communication to the port layer, STP transport layer, and ATA application layer. Only the STP link layer (i.e., not the port, transport, or application layer) transmits dwords.

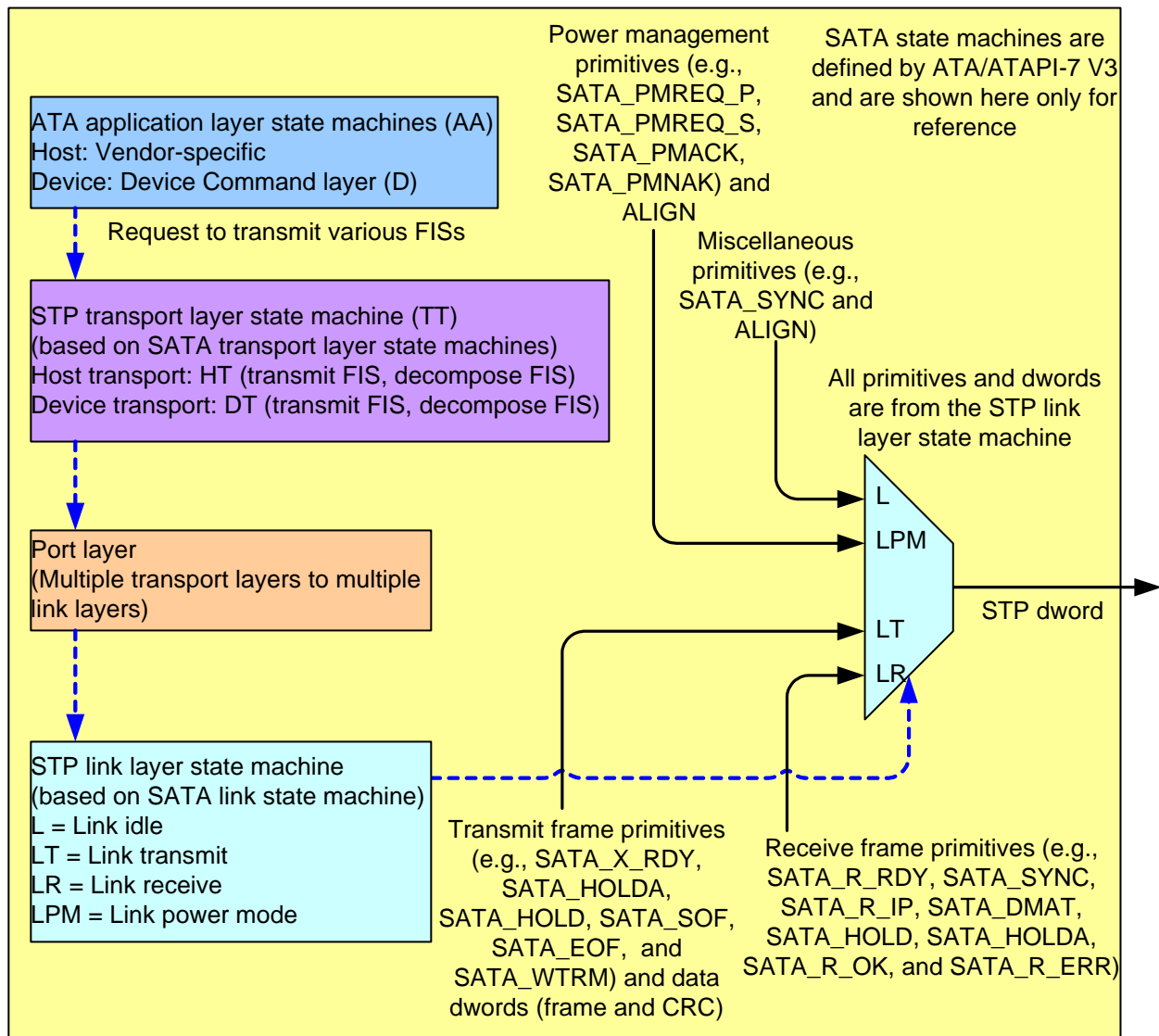


Figure 33 — STP link, port, STP transport, and ATA application layer state machines

Editor's Note 3: Figure 33 should have included STP initiator phy throttling ALIGN insertion. If the fixed 1/64 ratio for all types of connections is used instead per this proposal, that is no longer a special case.

Figure 36 shows the transmit data path in an expander phy.

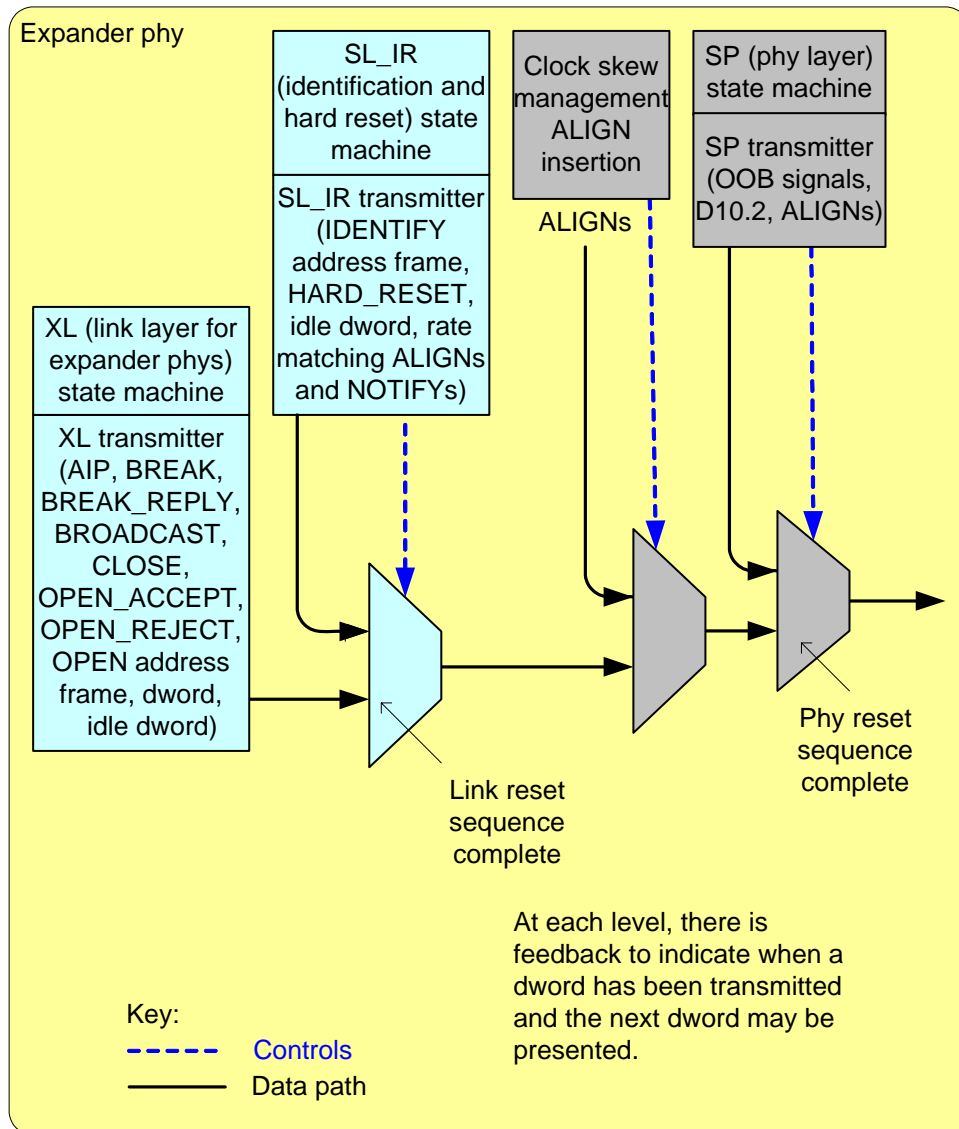


Figure 36 — Transmit data path and state machines in an expander phy

Editor's Note 4: In figure 36, change "Clock skew management" to "Clock frequency difference tolerance" (or whatever term matches section 7.3). Note that 05-381(Multiplexing) also makes changes to this figure.

Suggested changes to chapter 5 (Physical layer)

5.3.3 General electrical characteristics

...

Editor's Note 5: It has not yet been determined how 6 Gbps will be integrated into chapter 5. This material will probably move several times before it settles down. 6 Gbps is only important to this proposal for the center-spreading tolerance buffer size; none of the SSC parameters are based on link rate (ppm is a percentage of the unit interval, so scales automatically).

Table 49 defines the general electrical characteristics.

Table 49 — General electrical characteristics

Characteristic	Units	1,5 Gbps (i.e., G1)	3,0 Gbps (i.e., G2)
Physical link rate (nominal)	MBps	150	300
Bit rate (nominal)	Mbaud	1 500	3 000
Unit interval (UI)(nominal)	ps	666,6	333,3
Differential TxRx connection impedance (nominal)	ohm	100	
Maximum A.C. coupling capacitor ^a	nF	12	
Maximum noise during OOB idle time ^b	mV(P-P)	120	
^a The coupling capacitor value for A.C. coupled transmit and receive pairs. A.C. coupling requirements for transmitter devices are described in 5.3.6.1. A.C. coupling requirements for receiver devices are described in 5.3.7.1.			
^b With a measurement bandwidth of 1,5 times the highest supported baud rate (e.g., 4,5 GHz for 3,0 Gbps), no signal level during the idle time shall exceed the specified maximum differential amplitude.			

Table 50 defines the transmitter device general electrical characteristics.

Table 50 — General transmitter device electrical characteristics

Characteristic	Units	1,5 Gbps	3,0 Gbps
Physical link rate long-term stability at IT and CT	ppm	± 100	
Physical link rate tolerance SSC modulation at IT and CT	ppm	± 100 See table 52 in 5.3.8.1 and table 53 and table 54 in 5.3.8.2	
...	
a ...			

Table 51 defines the receiver device general electrical characteristics.

Table 51 — Receiver device general electrical characteristics

Characteristic	Units	1,5 Gbps	3,0 Gbps
Physical link rate long-term tolerance at IR if SATA is not supported	ppm	+100 / -100	
Physical link rate long-term tolerance at IR if SATA is supported	ppm	+350 / -350	
Physical link rate SSC modulation tolerance at IR if SATA is supported ^a	ppm	+350 / -350 See table 52 in 5.3.8.1 and table 55 in 5.3.8.3	
Physical link rate tolerance at IR if SATA is not supported and at CR	ppm	±100	
Physical link rate tolerance at IT and CT	ppm	±100	
...	
^a Allows support for SATA devices with spread spectrum clocking (see ATA/ATAPI-7 V3 and SATAII-PHY).			
^b ...			

5.3.8 Spread spectrum clocking (SSC)

5.3.8.1 SSC overview

~~Transmitter devices shall not transmit with spread spectrum clocking.~~

~~Receiver devices that support SATA shall support receiving with spread spectrum clocking (see ATA/ATAPI-7 V3 and SATAII-PHY). Receiver devices that do not support SATA are not required to support receiving with spread spectrum clocking.~~

~~An expander device shall retime data received from a SATA phy with an internal clock before forwarding to the rest of the SAS domain.~~

[\[All new text follows - blue underlines not included\]](#)

Spread spectrum clocking (SSC) is the technique of modulating the operating frequency of a transmitted signal to reduce the measured peak amplitude of radiated emissions.

Phys shall transmit with SSC as defined in 5.3.8.2 and receive with SSC as defined in 5.3.8.3.

NOTE 6 - Phys compliant with previous versions of this standard do not transmit with SSC. Phys compliant with previous versions of this standard that do not support being attached to SATA devices were not required to receive with SSC.

Table 52 defines the SSC modulation types.

Table 52 — SSC modulation types

SSC modulation type	SSC frequency deviation ^a
Center-spreading	+2 500 / -2 500 ppm
None	+0 / -0 ppm
Down-spreading	+0 / -5 000 ppm
^a This is in addition to the physical link rate long-term stability and tolerance defined in table 50 and table 51 (see 5.3.3).	

A phy may be transmitting with a different SSC modulation type than it is receiving (e.g., a phy is transmitting with center-spreading while it is receiving with down-spreading).

The phy shall transmit within the specified SSC frequency deviation with an SSC modulation frequency that is a minimum of 30 kHz and a maximum of 33 kHz.

The SSC modulation profile (e.g., triangular, exponential) is vendor-specific, but is intended to provide the maximum amount of electromagnetic interference (EMI) reduction. For center-spreading, the average amount of up-spreading (i.e., > 0 ppm) in the SSC modulation profile shall be the same as the average amount of down-spreading (i.e., < 0 ppm).

Editor's Note 6: the average will not be exactly the same; a real phy will skew one way or the other. This would lead to an infinite center-spreading tolerance buffer size, except for the fact that ALIGNs can be deleted if needed. The amount of slop must be less than 288 ppm, which is the legacy ALIGN rate of 1/2048 (488 ppm) minus the +/- 100 ppm long-term stability rate (range of 200 ppm), so an ALIGN can occasionally be deleted to account for this slop.

SSC-induced high-frequency jitter is included in the deterministic jitter (DJ) and consequently in total jitter (TJ) at the transmitter output. The jitter is measured after the application of a single pole high-pass frequency-weighting function that progressively attenuates jitter at 20 dB/decade below a frequency of ((bit rate) / 1 667).

5.3.8.2 Transmitter SSC modulation

A SAS phy shall transmit with the SSC modulation types defined in table 53.

Table 53 — SAS phy transmitter SSC modulation

SSC modulation type	Condition
None	When attached to a phy that does not support SSC
Down-spreading	When attached to a phy that supports SSC

NOTE 7 - A SAS device (e.g., disk drive) that attach to a SAS domain in a location (e.g., a drive bay) that also supports a SATA device may always transmit with down-spreading on all its phys, since phys that attach to SAS have always been required to tolerate the SATA clock frequency range of +350/-5 350 ppm (see table 55 in 5.3.8.3).

NOTE 8 - SAS phys compliant with previous versions of this standard were only allowed to transmit with an SSC modulation type of none.

An expander phy shall transmit with the SSC modulation types defined in table 54.

Table 54 — Expander phy transmitter SSC modulation types

SSC modulation type	Condition
Center-spreading	When attached to a SAS phy or expander phy that supports SSC
None	When attached to a SAS phy or expander phy that does not support SSC
Down-spreading	When attached to a SATA phy

NOTE 9 - Expander phys compliant with previous versions of this standard were only allowed to transmit with an SSC modulation type of none.

5.3.8.3 Receiver SSC modulation tolerance

SAS phys and expander phys shall support (i.e., tolerate) receiving with SSC modulation types defined in table 55.

Table 55 — Receiver SSC modulation tolerance

SSC modulation type ^a	Phys that are required to tolerate
Center-spreading	All phys
None	All phys
Down-spreading	All phys
^a This is in addition to the physical link rate long-term tolerance defined in table 51 (see 5.3.3).	

NOTE 10 - Phys compliant with previous versions of this standard that do not support being attached to SATA devices were only required to tolerate an SSC modulation type of none. Phys compliant with previous versions of this standard that support being attached to SATA devices were only required to tolerate SSC modulation types of none and down-spreading.

5.3.8.4 Expander device forwarding dwords while center-spreading

Expander devices forwarding dwords shall support a center-spreading tolerance buffer for each connection with the buffer size defined in table 56 to hold any dwords it receives during the up-spreading portion(s) of the SSC modulation period that it is unable to forward because the ECR and/or the transmitting expander phy is slower than the receiving expander phy and because the dword stream does not include enough deletable primitives. The expander device unloads the center-spreading tolerance buffer during the down-spreading portion(s) of the SSC modulation period when the receiving expander phy is slower than the ECR and the transmitting expander phy.

Table 56 — Expander center-spreading tolerance buffer

Physical link rate	Minimum buffer size
6 Gbps	7 dwords
3 Gbps	4 dwords
1,5 Gbps	2 dwords

NOTE 11 - The minimum buffer size is based on the number of dwords that may be transmitted during half of the longest allowed SSC modulation period (i.e., half of the period indicated by 30 kHz) at the maximum physical link rate (i.e., +2 600 ppm) minus the number that may be transmitted at the nominal physical link rate rate (i.e., 0 ppm). This accounts for forwarding dwords in a connection that originated from a phy compliant with

previous versions of this standard (i.e., a phy with an SSC modulation type of none and inserting ALIGNs and/or NOTIFYs at a rate supporting the long-term frequency stability).

Editor's Note 7: See last page of this proposal for calculations behind those numbers

Figure 37 shows an example of center-spreading tolerance buffer usage.

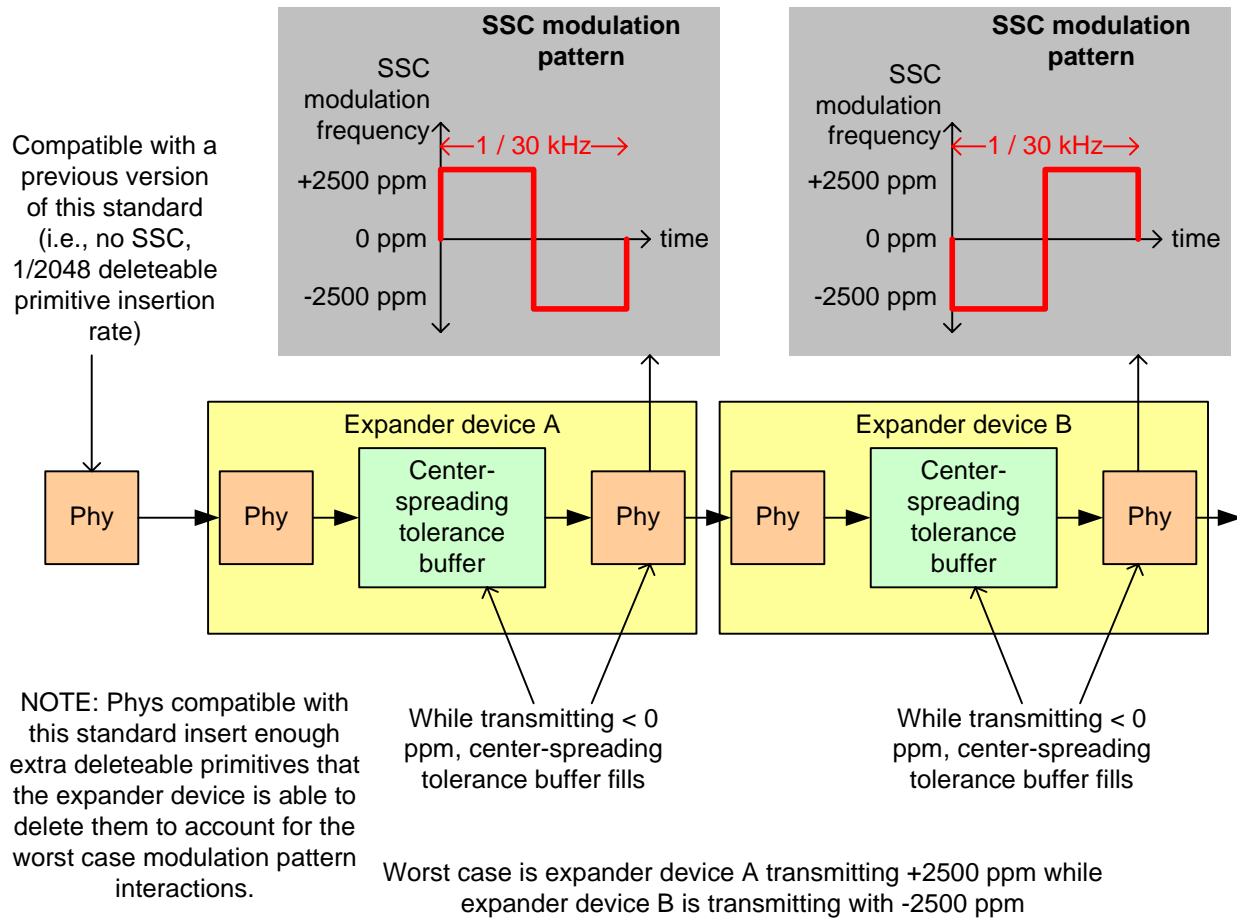


Figure 37 — Center-spreading tolerance buffer

[\[End of all-new text; change marks resume\]](#)

5.3.9 Non-tracking clock architecture

Transceivers shall be designed with a non-tracking clock architecture (i.e., the receive clock derived from the bit stream received by the receiver device shall not be used as the transmit clock by the transmitter device).

Receiver devices that support SATA shall tolerate clock tracking by the SATA device. Receiver devices that do not support SATA are not required to tolerate clock tracking by the SATA device.

Suggested changes to chapter 7 (Link layer)

7.2.5.2 ALIGN

ALIGNs are used for:

- OOB signals;
- character and dword alignment during the speed negotiation sequence;
- clock ~~skew management~~[frequency difference tolerance](#) after the phy reset sequence (see 7.3); [and](#)

- d) rate matching during connections (see 7.13); ~~and~~
- e) ~~STP initiator phy throttling during STP connections (see 7.16.2).~~

Table 57 defines the different versions of ALIGN primitives.

Table 57 — ALIGN primitives

Primitive	Description
ALIGN (0)	Used for OOB signals, the speed negotiation sequence, clock skew-management <u>frequency difference tolerance</u> , and rate matching, and STP initiator phy throttling .
ALIGN (1)	Used for the speed negotiation sequence, clock skew-management <u>frequency difference tolerance</u> , and rate matching, and STP initiator phy throttling .
ALIGN (2)	Used for clock skew-management <u>frequency difference tolerance</u> , and rate matching, and STP initiator phy throttling .
ALIGN (3)	Used for clock skew-management <u>frequency difference tolerance</u> , and rate matching, and STP initiator phy throttling .

Phys shall use ALIGN (0) to construct OOB signals as described in 6.6. Phys shall use ALIGN (0) and ALIGN (1) during the speed negotiation sequence as described in 6.7.4.2. Phys shall rotate through ALIGN (0), ALIGN (1), ALIGN (2), and ALIGN (3) for all ALIGNs sent after the phy reset sequence.

Phys receiving ALIGNs after the phy reset sequence shall not verify the rotation and shall accept any of the ALIGNs at any time.

Phys shall only detect an ALIGN after decoding all four characters in the primitive.

NOTE 12 - SATA devices are allowed to decode every dword starting with a K28.5 as an ALIGN, since ALIGN is the only primitive defined starting with K28.5.

For clock ~~skew-management~~frequency difference tolerance, ~~and~~ rate matching, ~~and STP initiator phy throttling~~, ALIGNs may be replaced by NOTIFYs (see 7.2.5.10). ALIGNs shall not be replaced by NOTIFYs during OOB signals and speed negotiation.

7.2.5.10 NOTIFY

7.2.5.10.1 NOTIFY overview

NOTIFY may be transmitted in place of any ALIGN (see 7.2.5.2) being transmitted for clock ~~skew-management~~frequency difference tolerance (see 7.3), ~~or~~ rate matching (see 7.13), ~~or STP initiator phy throttling (see 7.16.2)~~. Substitution of a NOTIFY may or may not affect the ALIGN rotation (i.e., the NOTIFY may take the place of one of the ALIGNs in the rotation through ALIGN (0), ALIGN (1), ALIGN (2), or ALIGN (3) or it may delay the rotation). A specific NOTIFY shall not be transmitted in more than three consecutive dwords until at least three other dwords have been transmitted.

...

7.3 Clock ~~skew-management~~frequency difference tolerance

Editor's Note 8: 05-381 (Multiplexing) and 06-275 (ALIGNs during STP connections) also make changes to this section.

Editor's Note 9: "Clock skew management" is a poor name, since it's not really related to skew between clocks, it's related to slight frequency differences (that do cause the clock edges to be

skewed differently). “Clock frequency difference tolerance” is better. Would “Physical link rate tolerance” be better standardese?

7.3.1 Clock frequency difference tolerance overview

The internal clock for a device is typically based on a PLL with its own clock generator and is used when transmitting dwords on the physical link. When receiving, however, dwords need to be latched based on a clock derived from the input bit stream itself. Although the input clock is nominally a fixed frequency, it may differ slightly from the internal clock frequency up to the physical link rate tolerance defined in table 50 (see 5.3.3). Over time, if the input clock is faster than the internal clock, the ~~device~~phy receiver may receive a dword and not be able to forward it to an internal buffer; this is called an overrun. If the input clock is slower than the internal clock, the ~~device~~phy receiver may not have a dword when needed in an internal buffer; this is called an underrun.

To solve this problem, ~~transmitting devices~~phy transmitters insert ALIGNs or NOTIFYs in the dword stream. ~~Receivers~~Phy receivers may pass ALIGNs and NOTIFYs through to their internal buffers, or may strip them out when an overrun occurs. ~~Receivers~~Phy receivers add ALIGNs or NOTIFYs when an underrun occurs. The internal logic shall ignore all ALIGNs and NOTIFYs that arrive in the internal buffers.

Elasticity buffer circuitry, as shown in figure 38, is required to absorb the slight differences in frequencies between the SAS initiator phy, SAS target phy, and expander phys. The frequency tolerance for a phy is specified in 5.3.3. The depth of the elasticity buffer is vendor-specific but shall accommodate the clock ~~skew-~~managementfrequency difference tolerance ALIGN insertion requirements in table 58.

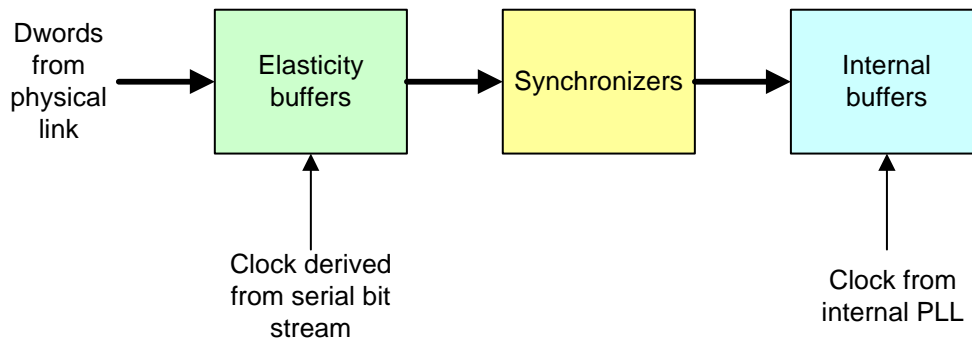


Figure 38 — Elasticity buffers

7.3.2 Phys originating dwords

A phy that is the original source for the dword stream (i.e., a phy that is not an expander phy forwarding dwords from another expander phy) shall insert one ALIGN or NOTIFY for clock ~~skew-management~~frequency difference tolerance as described in table 58.

Table 58 — Clock ~~skew-management~~frequency difference tolerance ALIGN insertion requirement

Physical link rate	Requirement
1,5 Gbps	One ALIGN or NOTIFY within every 2-048 64 dwords
3,0 Gbps	Two ALIGNs or NOTIFYs within every 4-096 128 dwords
<u>6.0 Gbps</u>	<u>Four ALIGNs or NOTIFYs within every 256 dwords</u>

NOTE 13 - These numbers account for the worst case clock frequency differences between the fastest phy transmitter and the slowest phy receiver (e.g., an expander phy originating dwords in an STP connection at +2 600 ppm to a SATA phy receiving dwords with an internal clock of -5 350 ppm). The difference of 7 950 ppm (i.e., 0.795 %) is less than the ALIGN insertion rate of 1/64 (i.e., 1.5625 %), ensuring there are enough ALIGNs or NOTIFYs for the phy receiver to delete in that situation.

Editor's Note 10: One number is defined for all cases: SSP connections, STP connections, SMP connections, and outside connections. This gets rid of the need for the STP initiator phy throttling ALIGN type altogether, but means more ALIGNs are inserted than necessary in some cases.

If a phy transmitter has SSC disabled or is using down-spreading SSC, it only needs to insert 1/128 ALIGNs (the SATA ratio), which covers +100 to -5350 ppm. It will never overrun a center-spreading or no-SSC receiver.

If a phy transmitter has SSC disabled or is using down-spreading SSC only, it only needs to insert 1/128 ALIGNs (the SATA ratio), which covers +100 to -5350 ppm.

Is that complication worth a $1/128 = 0.78\%$ performance improvement? (e.g. at 6 Gbps, this works out to be 4.77 MBytes/sec).

Editor's Note 11: It may not be necessary that the ratio be 1 over a power-of-two. The actual ratio needed is 1/126 (+2600 to -5350 ppm), just barely more than 1/128. With some slop, a ratio like 1/100 might be a better choice. Again, implementors are advised to make this programmable, no matter what the working drafts of the standard choose.

ALIGNs and NOTIFYs inserted for clock ~~skew management~~[frequency difference tolerance](#) are in addition to ALIGNs and NOTIFYs inserted for rate matching (see 7.13) ~~and STP initiator phy throttling (see 7.16.2)~~. See Annex H for a summary of their combined requirements.

See 7.2.5.2 for details on rotating through ALIGN (0), ALIGN (1), ALIGN (2), and ALIGN (3). NOTIFYs may also be used in place of ALIGNs (see 7.2.5.10) on SAS physical links.

7.3.3 Expander phys forwarding dwords

An expander device that is forwarding dwords (i.e., is not the original source) is allowed to insert or delete as many ALIGNs and/or NOTIFYs as required to match the transmit and receive connection rates. It is not required to transmit the number of ALIGNs and/or NOTIFYs for clock ~~skew management~~[frequency difference tolerance](#) described in table 58 when forwarding to a SAS physical link. It may increase or reduce that number based on clock frequency differences between the phy transmitting the dwords to the expander device and the expander device's receiving phy.

NOTE 14 - One possible implementation for expander devices forwarding dwords is for the expander device to delete all ALIGNs and NOTIFYs received and to insert ALIGNs and/or NOTIFYs at the transmit port whenever its elasticity buffer is empty.

The STP target port of an STP/SATA bridge is allowed to insert or delete as many ALIGNs and/or NOTIFYs as required to match the transmit and receive connection rates. It is not required to transmit any particular number of ALIGNs and/or NOTIFYs for clock ~~skew management~~[frequency difference tolerance](#) when forwarding to a SAS physical link and is not required to ensure that any ALIGNs and/or NOTIFYs it transmits are in pairs.

NOTE 15 - Due to clock ~~skew~~[frequency difference tolerance](#) ALIGN and NOTIFY removal, the STP target port may not receive a pair of ALIGNs and/or NOTIFYs every 256 dwords, even though the STP initiator port transmitted at least one pair. However, the rate of the dword stream allows for ALIGN or NOTIFY insertion by the STP/SATA bridge. One possible implementation is for the STP/SATA bridge to delete all ALIGNs and NOTIFYs received by the STP target port and to insert two consecutive ALIGNs at the SATA host port when its elasticity buffer is empty or when 254 non-ALIGN dwords have been transmitted. It may need to buffer up to 2 dwords concurrently being received by the STP target port while it does so.

Editor's Note 12: work on the above note to reflect that special throttling ALIGNs are now just

7.13 Rate matching

Each successful connection request contains the connection rate (see 4.1.10) of the pathway.

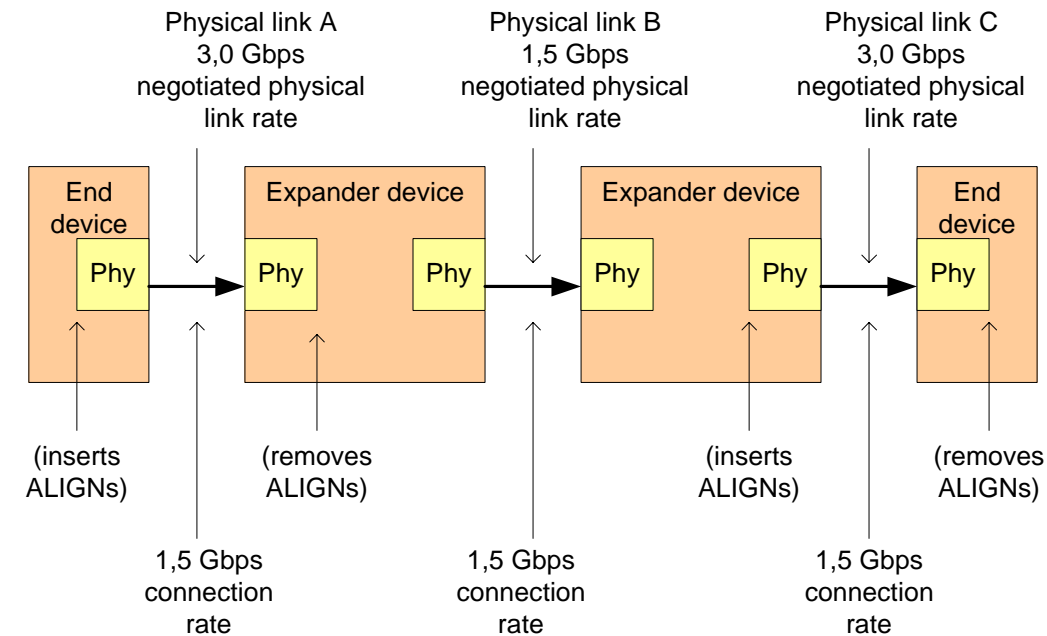
Each phy in the pathway shall insert ALIGNs and/or NOTIFYs between dwords if its physical link rate is faster than the connection rate as described in table 59.

Table 59 — Rate matching ALIGN and/or NOTIFY insertion requirements

Physical link rate	Connection rate	Requirement
1,5 Gbps	1,5 Gbps	None
3,0 Gbps	1,5 Gbps	One ALIGN or NOTIFY within every 2 dwords that are not clock skew-management frequency difference tolerance ALIGNs or NOTIFYs (i.e., every overlapping window of 2 dwords)(e.g., a repeating pattern of an ALIGN or NOTIFY followed by a dword or a repeating pattern of a dword followed by an ALIGN or NOTIFY)
	3,0 Gbps	None

ALIGNs and NOTIFYs inserted for rate matching are in addition to ALIGNs and NOTIFYs inserted for clock ~~skew-management~~frequency difference tolerance (see 7.3) and ~~STP initiator phy throttling (see 7.16.2)~~. See Annex H for a summary of their combined requirements.

Figure 39 shows an example of rate matching between a 3,0 Gbps source phy and a 3,0 Gbps destination phy, with an intermediate 1,5 Gbps physical link in between them.



Sample dwords on physical links (from left to right) during a 1,5 Gbps connection:

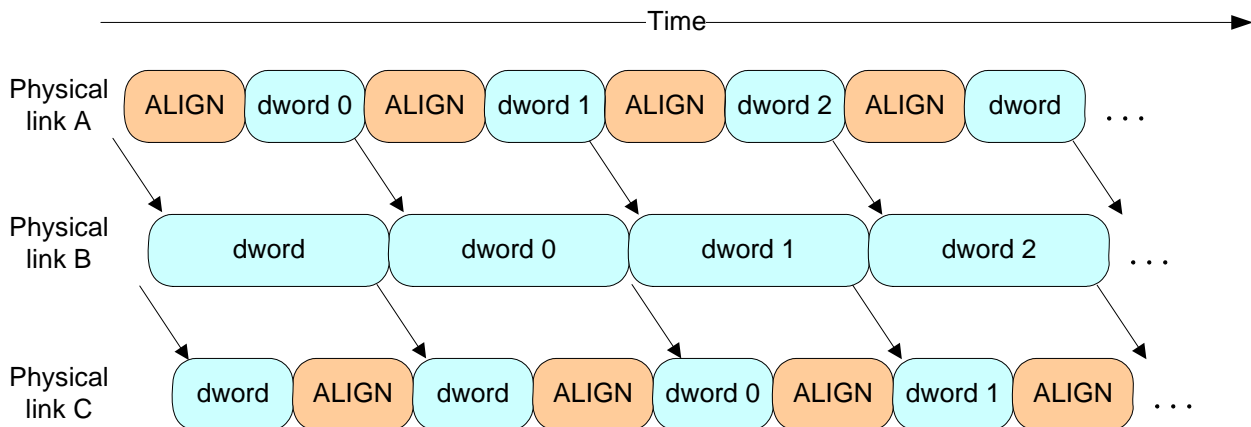


Figure 39 — Rate matching example

A phy originating dwords shall start rate matching at the selected connection rate starting with the first dword that is not an ALIGN or NOTIFY inserted for clock ~~skew management~~frequency difference tolerance following:

- transmitting the EOAF for an OPEN address frame; or
- transmitting an OPEN_ACCEPT.

An expander phy forwarding dwords shall not insert ALIGNs and NOTIFYs for rate matching based on counting dwords transmitted. It shall insert ALIGNs and NOTIFYs whenever it underflows.

The source phy transmits idle dwords including ALIGNs and NOTIFYs at the selected connection rate while waiting for the connection response. This enables each expander device to start forwarding dwords from the source phy to the destination phy after forwarding an OPEN_ACCEPT.

A phy shall stop inserting ALIGNs and/or NOTIFYs for rate matching after:

- transmitting the first dword in a CLOSE;
- transmitting the first dword in a BREAK;
- transmitting the first dword in a BREAK_REPLY;

- d) receiving an OPEN_REJECT for a connection request; or
- e) losing arbitration to a received OPEN address frame.

If an expander phy attached to a SATA phy is using a physical link rate greater than the maximum connection rate supported by the pathway from an STP initiator port, a management application client should use the SMP PHY CONTROL function (see 10.4.3.16) to set the PROGRAMMED MAXIMUM PHYSICAL LINK RATE field of the expander phy to the maximum connection rate supported by the pathway from that STP initiator port.

7.15 XL state machine

7.15.2 XL transmitter and receiver

...

The XL transmitter shall ensure clock ~~skew management~~frequency difference tolerance requirements are met (see 7.3) while originating dwords.

The XL transmitter shall ensure clock ~~skew management~~frequency difference tolerance requirements are met (see 7.3) during and after switching from forwarding dwords to originating dwords, including, for example:

- a) when transmitting BREAK;
- b) when transmitting BREAK_REPLY;
- c) when transmitting CLOSE;
- d) when transmitting an idle dword after closing a connection (i.e., after receiving BREAK, BREAK_REPLY, or CLOSE);
- e) while transmitting a SATA frame to a SAS physical link, when transmitting the first SATA_HOLD in response to detection of SATA_HOLD; and
- f) while receiving dwords of a SATA frame from a SAS physical link, when transmitting SATA_HOLD.

NOTE 16 - The XL transmitter may always insert an ALIGN or NOTIFY before transmitting a BREAK, BREAK_REPLY, CLOSE, or SATA_HOLD in response to detection of SATA_HOLD to meet clock ~~skew management~~frequency difference tolerance requirements.

The XL transmitter shall insert an ALIGN or NOTIFY before switching from originating dwords to forwarding dwords, including, for example:

- a) when transmitting OPEN_ACCEPT;
- b) when transmitting the last idle dword before a connection is established (i.e., after receiving OPEN_ACCEPT);
- c) while transmitting a SATA frame to a SAS physical link, when transmitting the last dword from the SATA flow control buffer in response to release of SATA_HOLD;
- d) while transmitting a SATA frame to a SAS physical link, when transmitting the last SATA_HOLD in response to release of SATA_HOLD (e.g., if the SATA flow control buffer is empty); and
- e) while receiving dwords of a SATA frame from a SAS physical link, when transmitting the last SATA_HOLD.

NOTE 17 - This ensures that clock ~~skew management~~frequency difference tolerance requirements are met, even if the forwarded dword stream does not include an ALIGN or NOTIFY until the last possible dword.

The XL transmitter shall ensure rate matching requirements are met during a connection (see 7.13).

~~The XL transmitter shall ensure STP initiator phy throttling requirements are met (see 7.16.2) when:~~

- ~~a) transmitting dwords in the direction of an STP target port while originating dwords (e.g., while transmitting SATA_HOLD, SATA_HOLD in response to release of SATA_HOLD, or unloading the SATA flow control buffer);~~
- ~~b) switching from forwarding dwords to originating dwords; and~~
- ~~c) switching from originating dwords to forwarding dwords.~~

...

7.16 STP link layer

7.16.1 STP frame transmission and reception

STP frame transmission is defined by SATA (see ATA/ATAPI-7 V3). During an STP connection, frames are preceded by SATA_SOF and followed by SATA_EOF as shown in figure 40.

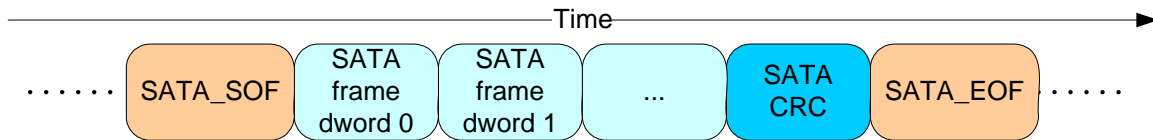


Figure 40 — STP frame transmission

The last data dword after the SOF prior to the EOF always contains a CRC (see 7.5).

Other primitives may be interspersed during the connection as defined by SATA.

STP encapsulates SATA with connection management. Table 60 summarizes STP link layer differences from the SATA link layer (see ATA/ATAPI-7 V3) that affect behavior during an STP connection.

Table 60 — STP link layer differences from SATA link layer during an STP connection

Feature	Description	Reference
STP initiator phy throttling	Limit the number of dwords transmitted to make room for more ALIGN primitives that may be deleted by the SATA device. This meets the ALIGN requirements for the physical link from the STP/SATA bridge to the attached SATA device.	7.16.2
STP flow control	Flow control through an STP connection is point-to-point, not end-to-end. Expander devices accept dwords in a temporary holding buffer after transmitting SATA HOLD to avoid losing data en-route before the transmitting phy acknowledges the SATA_HOLD with SATA_HOLD_A.	7.17.3
Continued primitive sequence	Sustain the continued primitive sequence if a SATA_CONT appears after the continued primitive sequence has begun.	7.17.4

7.16.2 ~~STP initiator phy throttling~~

Editor's Note 13: This whole concept goes away if the number of ALIGNs for clock frequency difference tolerance purposes are increased for all protocols. All that is unique about STP is the discussion about the ALIGNs being back-to-back in SATA but not necessarily so in SAS. There is a note in 7.3 discussing that already.

~~On a SATA physical link, phys are required to transmit two consecutive ALIGN (0) primitives within every 256 dwords. To ensure an STP/SATA bridge is able to meet this requirement, an STP initiator phy has to reduce (i.e., throttle) the rate at which it is sourcing dwords by the same amount.~~

~~During an STP connection, an STP initiator phy shall insert two ALIGNs or NOTIFYs within every 256 dwords (i.e., within every overlapping window of 256 dwords) that are not ALIGNs or NOTIFYs for clock skew management or rate matching. They are not required to be inserted consecutively, because a phy in the pathway may delete one of them for clock skew management since STP initiator phy throttling ALIGNs and NOTIFYs are indistinguishable from clock skew management ALIGNs and NOTIFYs.~~

STP target phys are not required to insert extra ALIGNs and/or NOTIFYs, because SATA hosts are not supported by SAS domains. STP initiator phys, the only recipients of data from STP target phys, do not require extra ALIGNs or NOTIFYs.

ALIGNs and NOTIFYs inserted for STP initiator phy throttling are in addition to ALIGNs and NOTIFYs inserted for clock skew management (see 7.3) and rate matching (see 7.13). See Annex H for a summary of their combined requirements.

A phy shall start inserting ALIGNs and NOTIFYs for STP initiator phy throttling after:

- a) transmitting an OPEN_ACCEPT; or
- b) sending the first SATA primitive after receiving an OPEN_ACCEPT.

A phy shall stop inserting ALIGNs and NOTIFYs for STP initiator phy throttling after:

- a) transmitting the first dword in a CLOSE;
- b) transmitting the first dword in a BREAK; or
- c) transmitting the first dword in a BREAK_REPLY.

Suggested changes to annexes

Annex H (informative)

ALIGN and/or NOTIFY insertion rate summary

Table H.1 shows all the possible combinations of ALIGN and/or NOTIFY insertion rates for clock ~~skew management~~frequency difference tolerance (see 7.3), and rate matching (~~see 7.13~~), and STP initiator phy throttling (~~see 7.17.2~~).

Table H.1 — ALIGN and/or NOTIFY insertion rate examples

Physical link rate	Connection rate	Type of dword stream	ALIGN and/or NOTIFY insertion rate (per specified number of dwords)
6 Gbps	6 Gbps		4 per 256 (clock frequency difference tolerance)
	3 Gbps		4 per 256 (clock frequency difference tolerance) + 1 per 2 (rate matching)
	1.5 Gbps		4 per 256 (clock frequency difference tolerance) + 3 per 4 (rate matching)
3,0 Gbps	3,0 Gbps	all but to STP target	2 per 4 096 128 (clock skew management <u>frequency difference tolerance</u>)
		to STP target	2 per 4 096 (clock skew management) + 2 per 256 (STP initiator phy throttling)
	1,5 Gbps	all but to STP target	2 per 4 096 128 (clock skew management <u>frequency difference tolerance</u>) + 1 per 2 (rate matching)
		to STP target	2 per 4 096 (clock skew management) + 1 per 2 (rate matching) + 2 per 256 (STP initiator phy throttling)
1,5 Gbps	1,5 Gbps	all but to STP target	1 per 2 048 64 (clock skew management <u>frequency difference tolerance</u>)
		to STP target	1 per 2 048 (clock skew management) + 2 per 256 (STP initiator phy throttling)

Background information (for this proposal, not for the standard)

The number of dwords that must be buffered by an expander when center-spreading is occurring is determined by the area under the SSC modulation profile curve. A worst case pattern would be a square wave where the phy receives at +2600 ppm for the first half-period (1/30 kHz) and -2600 ppm for the second half-period. As long as the area under (towards 0 ppm) each half-period is equal, the expander device can empty out the buffer during the second half-period.

Table 61 lists some ratios used in the calculations for the center-spreading tolerance buffer size and the ALIGN insertion rates.

Table 61 — Interesting ratios

ppm	Ratio	Decimal	Notes
15 625 ppm	1/64	0.015 625	SAS-2 ALIGN rate
7 950 ppm	1/126	0.007 95	SAS-2 widest range (+2 600 to -5 350 ppm)
7 812 ppm	1/128	0.007 812 5	SATA ALIGN rate
5 700 ppm	1/175	0.0057	SATA down-spreading range (+5 350 to -350 ppm)
5 350 ppm	1/186	0.00535	SATA down-spreading total
5 200 ppm	1/192	0.0052	SAS down-spreading range (+100 to -5 100 ppm)
-5 100 ppm	1/196	0.0051	SAS down-spreading total
-5 000 ppm	1/200	0.005	Down-spreading deviation
±2 600 ppm	1/384	0.002 6	Center-spreading total
±2 500 ppm	1/400	0.002 5	Center-spreading deviation
488 ppm	1/2 048	0.000 488 281 25	SAS-1.1 ALIGN rate
±350 ppm	1/2 857	0.000 35	SATA base
200 ppm	1/5 000	0.000 2	SAS base range (+100 to -100 ppm)
±100 ppm	1/10 000	0.000 1	SAS base

UI = 166.666 ps/bit at 6 Gbps, 333.333 ps/bit at 3 Gbps, and 666.666 ps/bit at 6 Gbps

What is the dword rate at + 2600 ppm?

166.666 ps/bit * (1 - .0026) = 166.233 sec/bit = 6 015 640 665 bits/sec

333.333 ps/bit * (1 - .0026) = 332.466 sec/bit = 3 007 820 332 bits/sec

666.666 ps/bit * (1 - .0026) = 664.933 sec/bit = 1 503 910 166 bits/sec

Note that 30 kHz = 30000 / sec.

How many dwords are transmitted in a 30 kHz half-period at nominal rates?

0.5 * (6 000 000 000 bits/sec) * (sec / 30 000) * (1 dword / 40 bits) = 2500 dwords at 6 Gbps

0.5 * (3 000 000 000 bits/sec) * (sec / 30 000) * (1 dword / 40 bits) = 1250 dwords at 3 Gbps

0.5 * (1 500 000 000 bits/sec) * (sec / 30 000) * (1 dword / 40 bits) = 625 dwords at 1.5 Gbps

How many dwords are transmitted in a 30 kHz half-period at +2600 ppm (compared to dwords at nominal rates)?

0.5 * (6 015 640 665 bits/sec) * (sec / 30 000) * (1 dword / 40 bits) = 2506.517 dwords at 6 Gbps (7 more)

0.5 * (3 007 820 332 bits/sec) * (sec / 30 000) * (1 dword / 40 bits) = 1253.258 dwords at 3 Gbps (4 more)

0.5 * (1 503 910 166 bits/sec) * (sec / 30 000) * (1 dword / 40 bits) = 626.629 dwords at 1.5 Gbps (2 more)

(looking forward, at 12 Gbps the numbers are 5013.0338 vs. 5000 dwords (14 more))

(this assumes that SSC is generated based on actual time, not by counting nominal unit intervals)