To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 14 June 2006
Subject: 06-263r1 SAS-2 Spread-spectrum clocking

Revision history

Revision 0 (31 May 2006) First revision
Revision 1 (10 June 2006) SAS disk drives and HBAs only need to choose between to transmit with
    downspreading 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 upspreading/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
    upspreading/center-spreading SSC, downspreading 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.

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)

Overview

Serial ATA has always supported spread-spectrum clocking (SSC) with a range of +0/-5000 ppm
(downspreading) 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
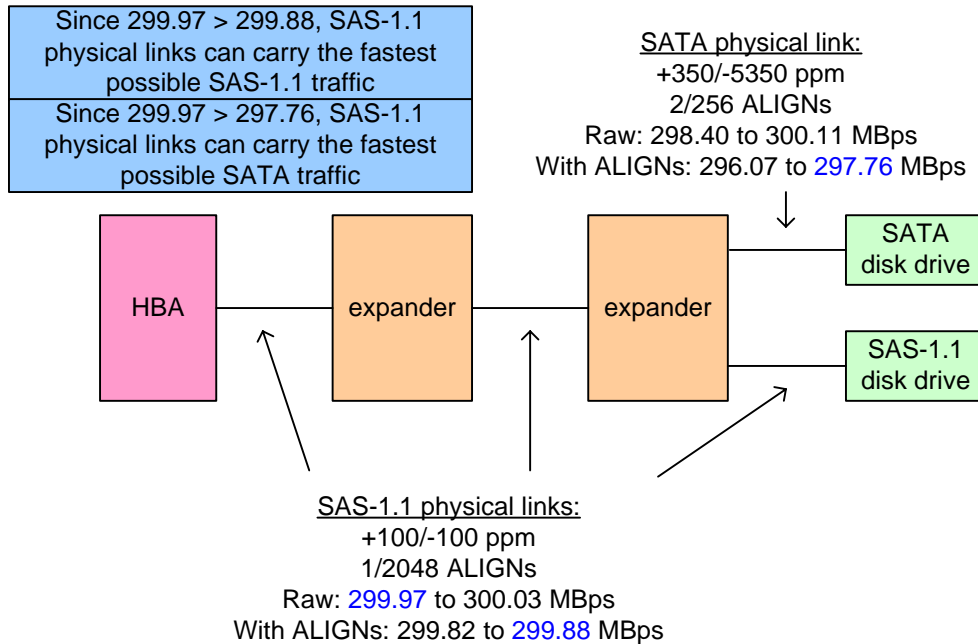(0.7812%) are ALIGNs.

> 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 downspreading SSC.

SAS HBAs and expanders usually support receiving with downspreading 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 downspreading SSC and SAS-1.1 phys that transmit with no SSC.

Since 299.97 > 299.88, SAS-1.1 physical links can carry the fastest possible SAS-1.1 traffic

Since 299.97 > 297.76, SAS-1.1 physical links can carry the fastest possible SATA traffic

SATA physical link:
+350/-5350 ppm
2/256 ALIGNs
Raw: 298.40 to 300.11 MBps
With ALIGNs: 296.07 to 297.76 MBps

HBA — expander — expander — SATA disk drive / SAS-1.1 disk drive

SAS-1.1 physical links:
+100/-100 ppm
1/2048 ALIGNs
Raw: 299.97 to 300.03 MBps
With ALIGNs: 299.82 to 299.88 MBps

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:

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

Table 1 shows 3 Gbps unit interval (IU), 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

Downspreading

Figure 2 shows the problem with adding downspreading SSC (e.g., +0/-5000 ppm) to SAS-2. This works fine for SATA phys, which also downspread, 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 downspreading SSC. Downspreading SSC would have to be disabled everywhere if any non-SSC device were added to the SAS domain.
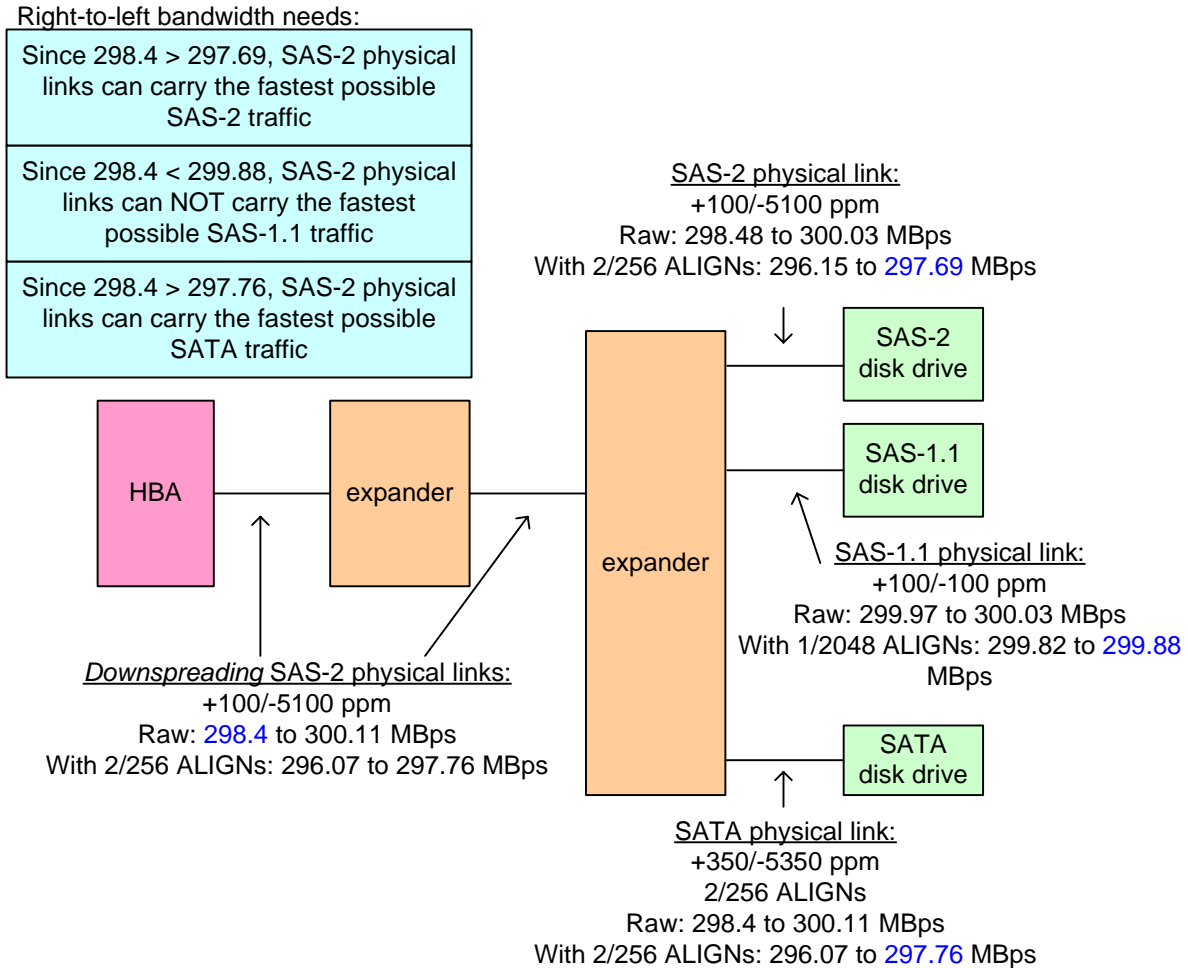
Right-to-left bandwidth needs:

Since 298.4 > 297.69, SAS-2 physical links can carry the fastest possible SAS-2 traffic

Since 298.4 < 299.88, SAS-2 physical links can NOT carry the fastest possible SAS-1.1 traffic

Since 298.4 > 297.76, SAS-2 physical links can carry the fastest possible SATA traffic

SAS-2 physical link:
+100/-5100 ppm
Raw: 298.48 to 300.03 MBps
With 2/256 ALIGNs: 296.15 to 297.69 MBps

SAS-2 disk drive

SAS-1.1 disk drive

SAS-1.1 physical link:
+100/-100 ppm
Raw: 299.97 to 300.03 MBps
With 1/2048 ALIGNs: 299.82 to 299.88 MBps

HBA

expander

expander

Downspreading SAS-2 physical links:
+100/-5100 ppm
Raw: 298.4 to 300.11 MBps
With 2/256 ALIGNs: 296.07 to 297.76 MBps

SATA disk drive

SATA physical link:
+350/-5350 ppm
2/256 ALIGNs
Raw: 298.4 to 300.11 MBps
With 2/256 ALIGNs: 296.07 to 297.76 MBps

Figure 2 — Downspreading SSC in SAS-2

Center-spreading

Center-spreading (e.g., +5000/-5000 ppm) has some of the same problems as downspreading, 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.
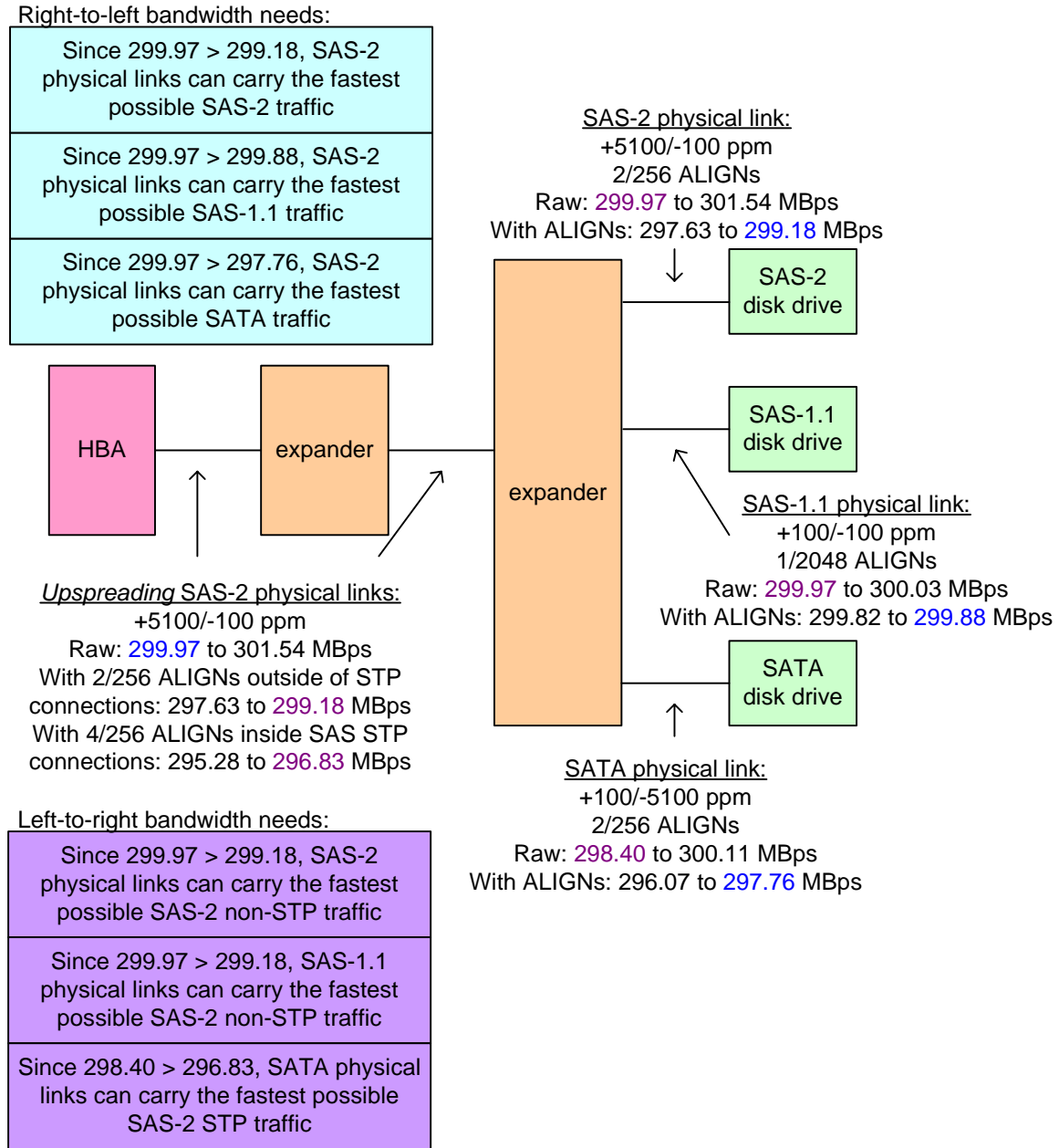
At 3 Gbps, there are 0.5 * (3,000,000,000 bits per second/ 30,000 per second)) = 50,000 bits = 1250 dwords during the slow period and 1250 dwords during the fast period.

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

One advantage of center-spreading over upspreading is more widespread availabilty of clock generators. Downspreading and center-spreading clock generators are common; upspreading clock generators are not.

Upspreading

Upspreading 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 upspreading SSC.

Right-to-left bandwidth needs:

Since 299.97 > 299.18, SAS-2 physical links can carry the fastest possible SAS-2 traffic

Since 299.97 > 299.88, SAS-2 physical links can carry the fastest possible SAS-1.1 traffic

Since 299.97 > 297.76, SAS-2 physical links can carry the fastest possible SATA traffic

SAS-2 physical link:
+5100/-100 ppm
2/256 ALIGNs
Raw: 299.97 to 301.54 MBps
With ALIGNs: 297.63 to 299.18 MBps

SAS-2 disk drive

SAS-1.1 disk drive

HBA

expander

expander

Upspreading SAS-2 physical links:
+5100/-100 ppm
Raw: 299.97 to 301.54 MBps
With 2/256 ALIGNs outside of STP connections: 297.63 to 299.18 MBps
With 4/256 ALIGNs inside SAS STP connections: 295.28 to 296.83 MBps

SAS-1.1 physical link:
+100/-100 ppm
1/2048 ALIGNs
Raw: 299.97 to 300.03 MBps
With ALIGNs: 299.82 to 299.88 MBps

SATA disk drive

SATA physical link:
+100/-5100 ppm
2/256 ALIGNs
Raw: 298.40 to 300.11 MBps
With ALIGNs: 296.07 to 297.76 MBps

Left-to-right bandwidth needs:

Since 299.97 > 299.18, SAS-2 physical links can carry the fastest possible SAS-2 non-STP traffic

Since 299.97 > 299.18, SAS-1.1 physical links can carry the fastest possible SAS-2 non-STP traffic

Since 298.40 > 296.83, SATA physical links can carry the fastest possible SAS-2 STP traffic

Figure 3 — Upspreading SSC in SAS-2

Extra ALIGNs are needed to accomodate SSC.

Summary of proposed rules

Overall summary:

a) Center-spreading and downspreading SSC for SAS physical links is supported at 6 Gbps, 3 Gbps, or 1.5 Gbps
b) Downspreading SSC for SATA physical links is supported at any physical link rate

Receiver rules:

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., downspreading) from SAS phys and expander phys supporting SSC, for disk drive SAS phys; and
    E) +350/-5350 ppm SSC (i.e., downspreading) 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

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., downspreading)	yes	no	yes
+350/-5350 ppm SSC (i.e., downspreading)	no	yes	no

Transmitter rules:

a) Non-expander phys (i.e., end devices) shall transmit with:
    A) +100/-5100 SSC (i.e., downspreading) 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., downspreading) 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

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., downspreading)	yes	yes

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

    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 downspreading 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.

    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:

    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 requries 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 require an expander ASIC to have 3 clock trees. Figure 4 shows the problem.
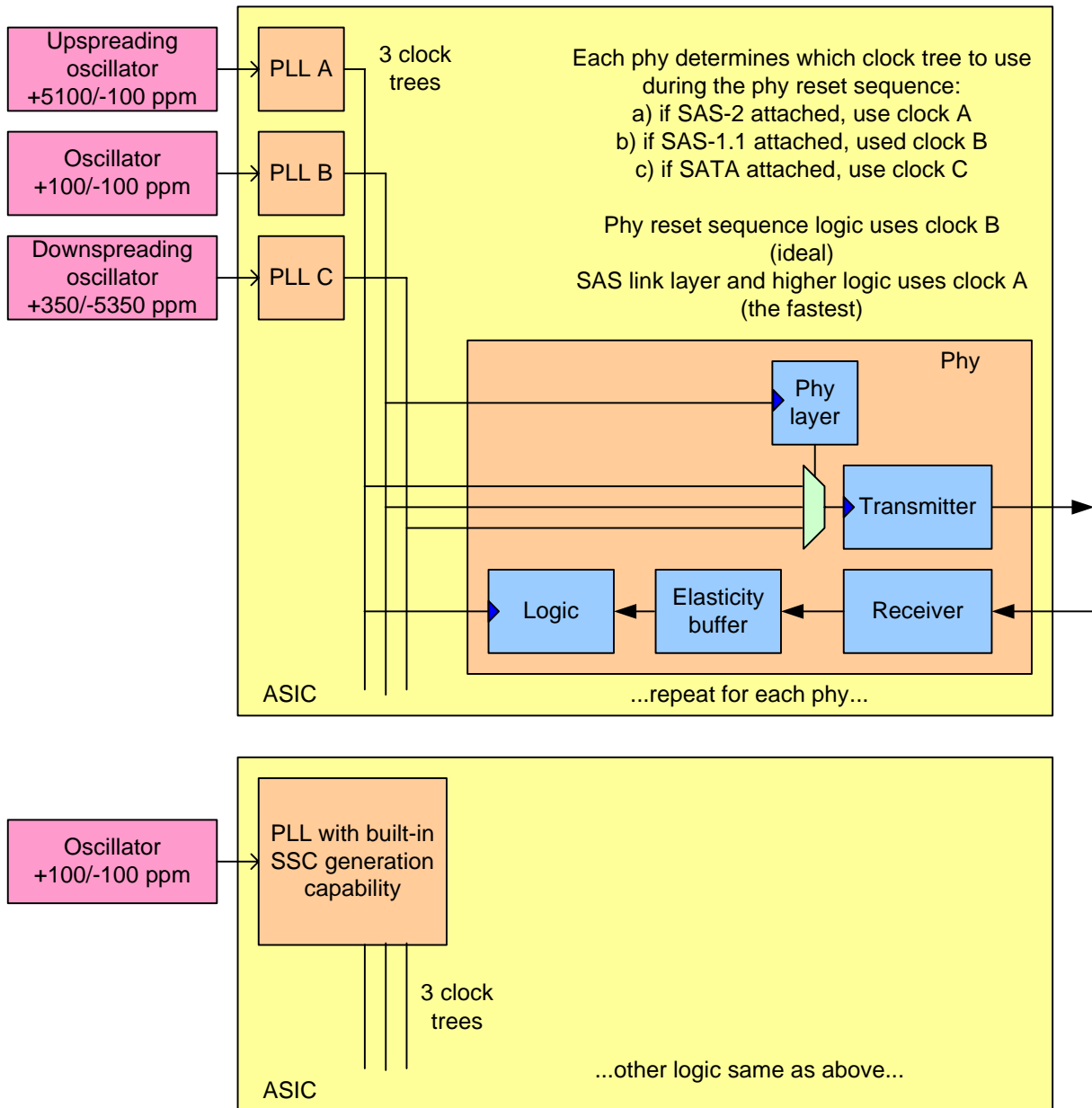
Figure 4 — Three clock trees

Disk drive phys can always transmit with downspreading 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 chapter 3 (Definitions)

3.1.220 spread spectrum clocking (SSC): The technique of modulating the operating frequency of a transmitted signal 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)

...

> Editor's Note 1: The Chapter 4 transmit data path figures may need updates to clarify ALIGN
> insertion responsibilities. 05-381 SAS-2 Multiplexing is also making changes to those figures.

Suggested changes to chapter 5 (Physical layer)

5.3.3 General electrical characteristics

...

> Editor's Note 2: 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.

Table 4 defines the transmitter device general electrical characteristics.

Table 4 — General transmitter device electrical characteristics

Characteristic	Units	1,5 Gbps	3,0 Gbps
Physical link rate tolerance at IT and CT	ppm	± 100 +100 / -5 100, +100/-100, or +2600 / -2 600	
...	...	...	
a ...			

Table 5 defines the receiver device general electrical characteristics.

Table 5 — Receiver device general electrical characteristics

Characteristic	Units	1,5 Gbps	3,0 Gbps
Physical link rate tolerance at IR if SATA is supported a	ppm	+350 / -5 350, +350 / -350 and +2600 / -2 600	
Physical link rate tolerance at IR if SATA is not supported and at CR	ppm	± 100 +100 / -5 100, +100/-100, and +2600 / -2 600	
Physical link rate tolerance at IT and CT	ppm	± 100	
...	...	...	
a Allows support for SATA devices, SAS devices, and expander devices with spread spectrum clocking (see ATA/ATAPI-7 V3, and SATAII-PHY, and 5.3.8).			
b Allows support for SAS devices and expander devices with spread spectrum clocking (see 5.3.8).			
c ...			

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.~~

Editor's Note 3: 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.

The transmitter and receiver phy may be operating with different SSC ranges.

5.3.8.2 Transmitter device clock frequency modulation

When SSC is enabled, the phy shall modulate the transmitter clock within a defined frequency range at a modulation rate greater than or equal to 30 kHz and less than or equal to 33 kHz.

SAS phys shall transmit with a clock frequency range defined in table 6.

Table 6 — SAS phy transmitter clock frequency modulation

Clock frequency range	Description
+100 / -100 ppm	When attached to a phy that does not support SSC
+100 / -5 100 ppm	When attached to a phy that supports SSC [a]
[a] Phys compliant with previous versions of this standard were only allowed to transmit with a clock frequency range of +100 / -100 ppm.	

NOTE 6 - 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 a clock frequency range of +100/-5100 ppm on all its phys, since phys that attach to SAS have always been required to tolerate the SATA clock frequence range of +350/-5 350 ppm (see table 8 in 5.3.8.3).

Expander phys shall transmit with a clock frequency range defined in table 7.

Table 7 — Expander phy transmitter clock frequency modulation

Clock frequency range	Description
+2 600 / -2 600 ppm	When attached to a SAS phy or expander phy that supports SSC [a]
+100 / -100 ppm	When attached to a phy that does not support SSC
+100 / -5 100 ppm	When attached to: [a] a) a SAS phy that supports SSC; or b) a SATA phy.
[a] Phys compliant with previous versions of this standard were only allowed to transmit with a clock frequency range of +100 / -100 ppm.	

5.3.8.3 Receiver device clock frequency tolerance

SAS phys and expander phys shall support receiving with clock frequency tolerances defined in table 8.

Table 8 — Receiver clock frequency tolerance

Clock frequency tolerance	Phys that are required to tolerate
+2 600 / -2 600 ppm	All phys [a]
+100 / -100 ppm	All phys
+100 / -5 100 ppm	All phys [a]
+350 / -350 ppm	Phys that support being attached to SATA devices
+350 / -5 350 ppm	Phys that support being attached to SATA devices
[a] Phys compliant with previous versions of this standard that did not support being attached to SATA devices were only required to tolerate +100 / -100 ppm.	

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:

a) OOB signals;
b) character and dword alignment during the speed negotiation sequence;
c) clock skew management 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 9 defines the different versions of ALIGN primitives.

Table 9 — ALIGN primitives

Primitive	Description
ALIGN (0)	Used for OOB signals, the speed negotiation sequence, clock skew management, and rate matching, and STP initiator phy throttling.
ALIGN (1)	Used for the speed negotiation sequence, clock skew management and rate matching, and STP initiator phy throttling.
ALIGN (2)	Used for clock skew management, and rate matching, and STP initiator phy throttling.
ALIGN (3)	Used for clock skew management, 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 7 - 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, 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 (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

> Editor's Note 5: Clock skew management is misnamed; it should be called "clock frequency difference tolerance"

> Editor's Note 6: 05-381 (Multiplexing) and 06-275 (ALIGNs during STP connections) also make changes to this section.

7.3.1 Clock skew management 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 49 (see 5.3.3). Over time, if the input clock is faster than the internal clock, the devicephy 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 devicephy receiver may not have a dword when needed in an internal buffer; this is called an underrun.

To solve this problem, transmitting devicesphy transmitters insert ALIGNs or NOTIFYs in the dword stream. ReceiversPhy receivers may pass ALIGNs and NOTIFYs through to their internal buffers, or may strip them out when an overrun occurs. ReceiversPhy 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 5, 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 management ALIGN insertion requirements in table 10.
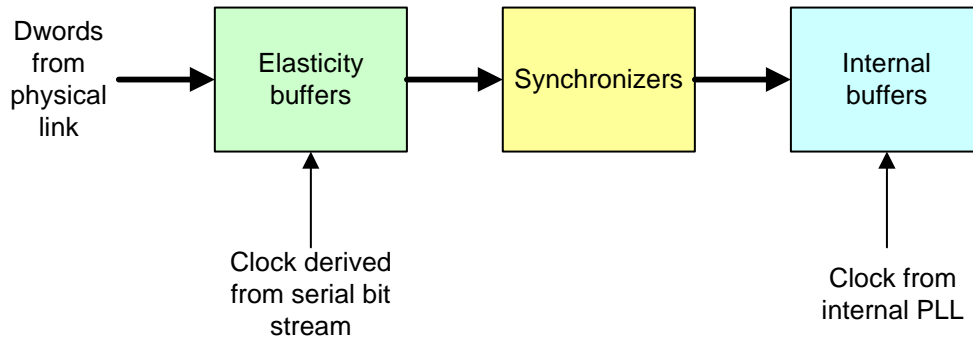

Figure 5 — Elasticity buffers

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 as described in table 10.

Table 10 — Clock skew management ALIGN insertion requirement

Physical link rate	Requirement
1,5 Gbps	One ALIGN or NOTIFY within every 2 04864 dwords
3,0 Gbps	Two ALIGNs or NOTIFYs within every 4 096128 dwords
6,0 Gbps	Four ALIGNs or NOTIFYs within every 256 dwords

NOTE 8 - These numbers account for the worst case clock frequency differences between the fastest phy transmitter and the slowest phy receiver (e.g., +2 600 ppm to -5 350 ppm is a difference of 7 950 ppm (i.e., 0,795 %). 1/64 is 1,5625 %, ensuring there are enough ALIGNs or NOTIFYs for the phy receiver to delete in that situation.

Editor's Note 7: If a phy transmitter has SSC disabled or is using downspread SSC only, it could get away with inserting fewer ALIGNs - 1/128 (the SATA ratio) would cover sending to either SAS or SATA phys with downspread SSC. Is that complication worth a 0.8% performance improvement? (e.g. at 6 Gbps, this is 4.77 MBps).

ALIGNs and NOTIFYs inserted for clock skew management 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 described in table 10 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 9 - 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 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 10 - Due to clock skew 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 8: work on the above note to reflect that special throttling ALIGNs are now just clock normal skew management ALIGNs

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 11.

Table 11 — 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 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 (see 7.3) and STP initiator phy throttling (see 7.16.2). See Annex H for a summary of their combined requirements.

Figure 6 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.
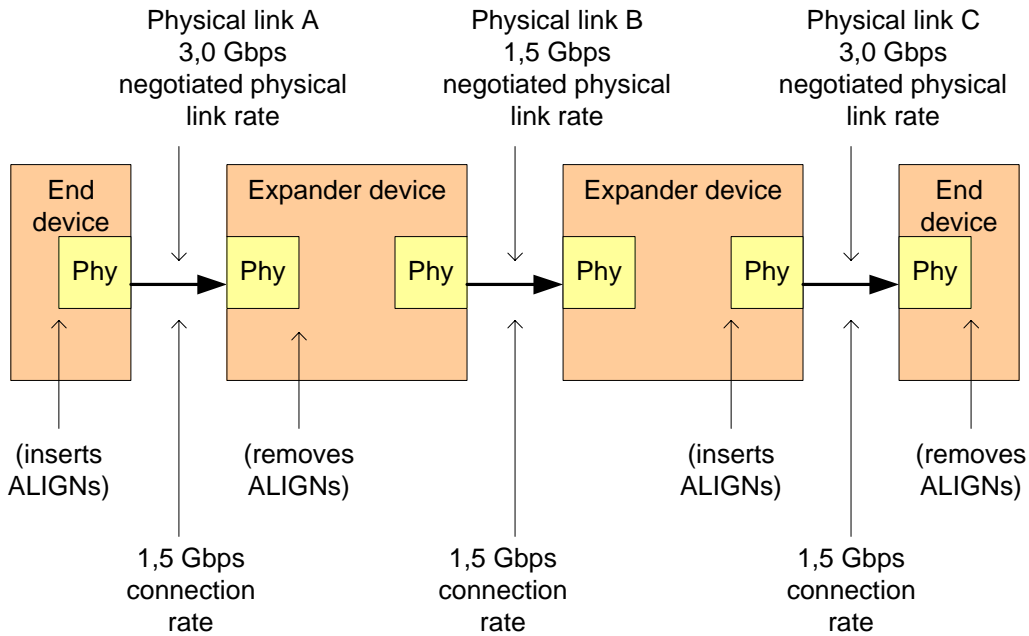
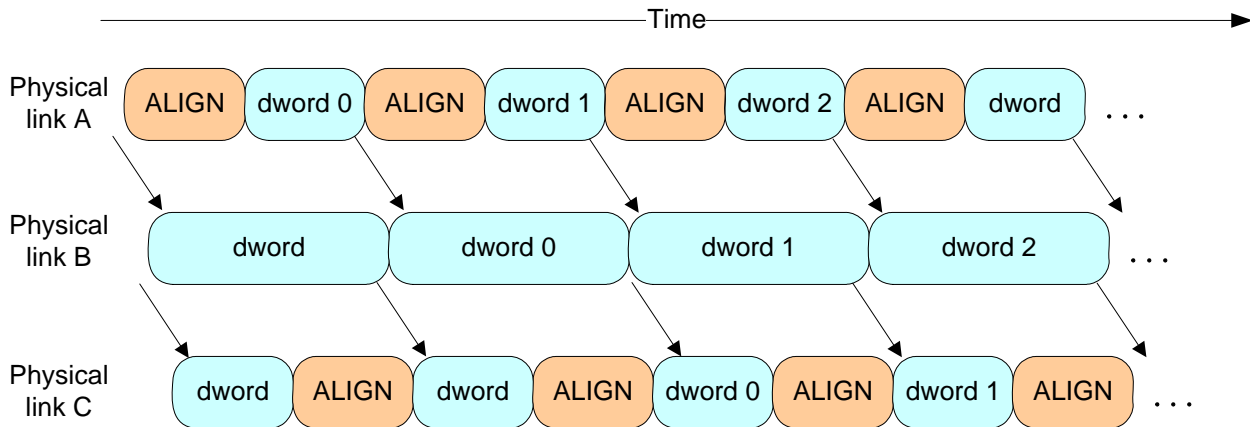Figure 6 — Rate matching example

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 following:

    a)   transmitting the EOAF for an OPEN address frame; or
    b)   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:

    a)   transmitting the first dword in a CLOSE;
    b)   transmitting the first dword in a BREAK;
    c)   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 10.4.3.14) 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 requirements are met (see 7.3) while originating dwords.

The XL transmitter shall ensure clock skew management 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_HOLDA 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 11 - The XL transmitter may always insert an ALIGN or NOTIFY before transmitting a BREAK, BREAK_REPLY, CLOSE, or SATA_HOLDA to meet clock skew management 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_HOLDA 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 12 - This ensures that clock skew management 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_HOLDA, 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 7.
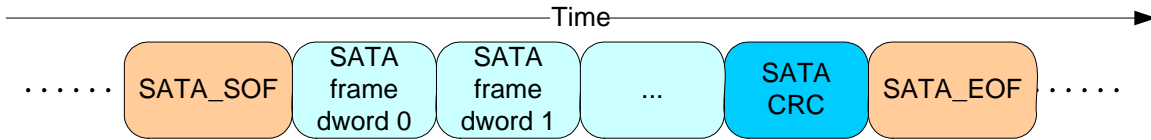

Figure 7 — STP frame transmission

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

Table 12 — 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_HOLDA.	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 9: This whole concept goes away as the number of ALIGNs for clock skew management purposes are increased. 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.