To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 31 May 2006
Subject: 06-263r0 SAS-2 Spread-spectrum clocking with upspreading

Revision history
Revision 0 (31 May 2006) First revision

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

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.

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.

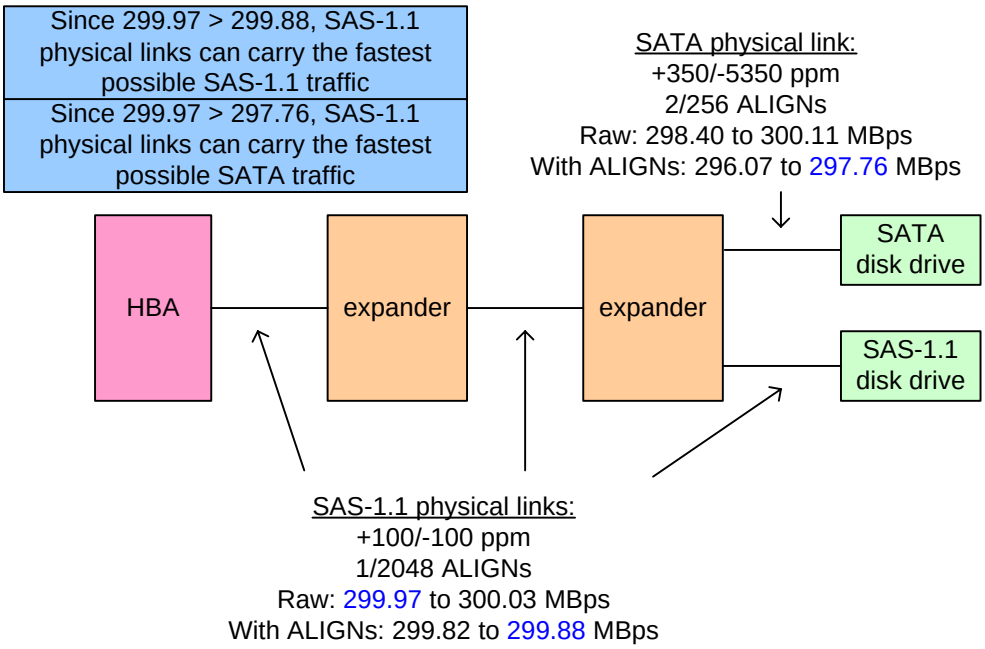

Figure 1 — SAS-1.1 support of SATA SSC

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 downspread 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 for SAS-1.1 phy which does not support SSC. If such a phy is attached anywhere in the SAS domain, its traffic overflows any SAS physical link performing SSC. SSC would have to be disabled everywhere if any non-SSC device were added to the SAS domain.

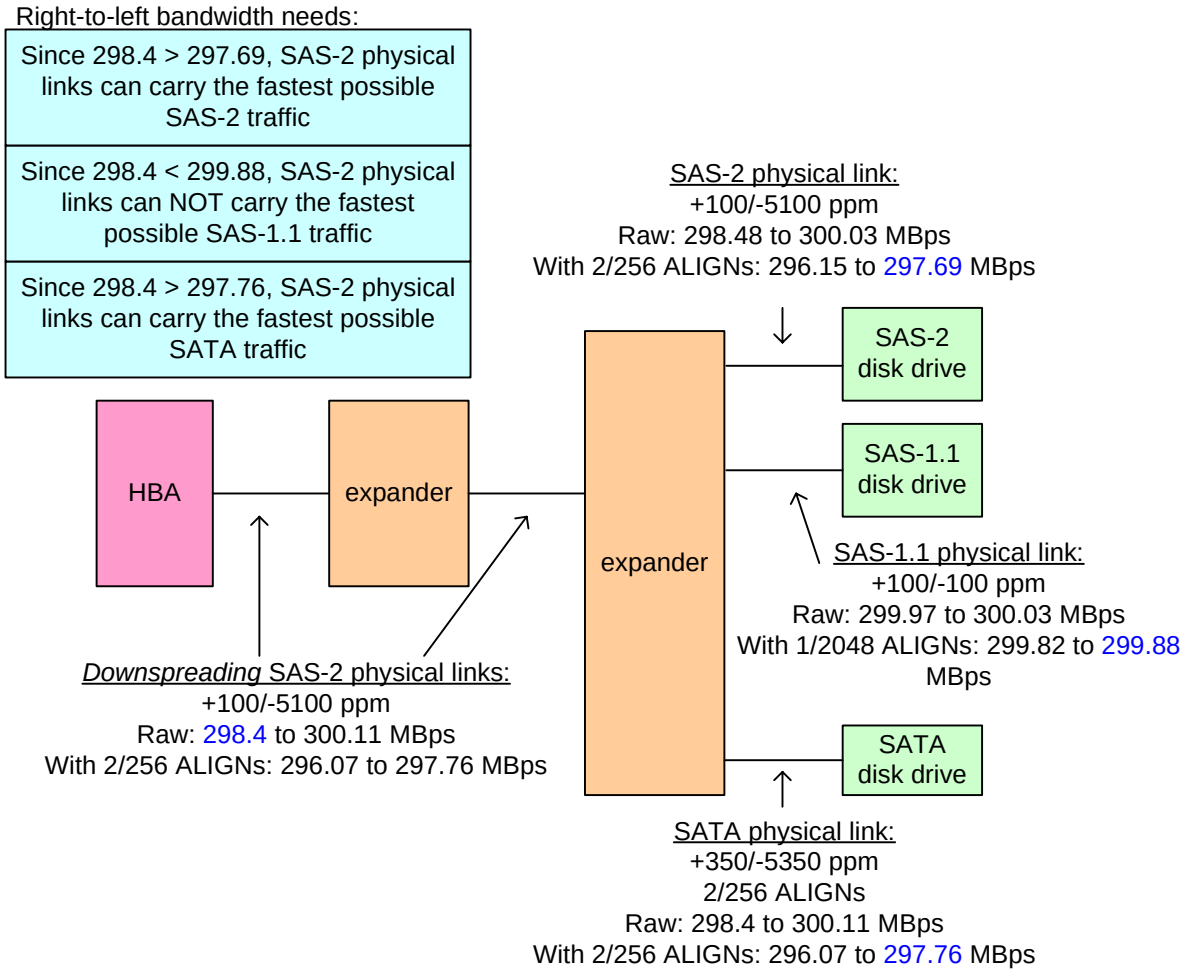

Figure 2 — Downspreading SSC in SAS-2

Center spreading

Center spreading (e.g., +5000/-5000 ppm) has the same problem as downspreading, since the SAS-2 physical link runs for long periods of time slower than a SAS-1.1 physical link. Although it also runs for long periods of time faster than a SAS-1.1 physical link, averaging out to the same rate, there are far too many dwords accumulated during the slow period to practically buffer for release in the fast period (at 3 Gbps, 0.5 * (3,000,000,000 bits per second/ 30,000 per second)) = 50,000 bits = 1250 dwords.

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.

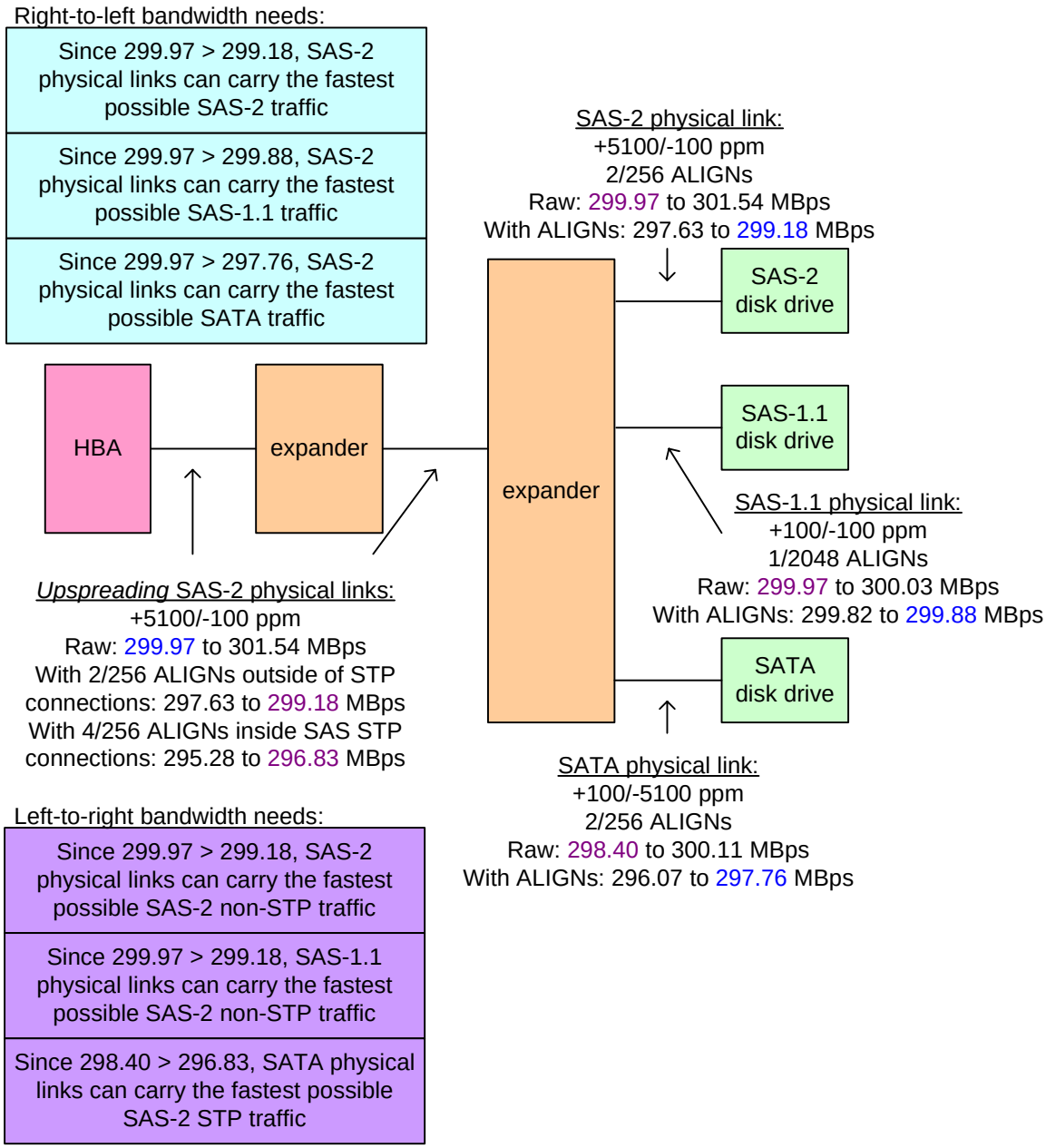

Figure 3 — Upspreading SSC in SAS-2

To reduce the number of options necessary for phys attached to disk drive bays, the initiator or expander phys attached to the disk drive bays could transmit with either +100/-100 or +100/-5100 ppm but forgo using +5100/-100 ppm. They still must support receiving with +5100/-100 ppm in this mode (so transmit and receive are different), since a SAS disk drive might be fixed at that rate. As long as the HBA to expander and expander to expander physical links use +5000/-0 or +0/-0, they will always be fast enough to carry traffic from the disk drive. This would require initiator and expander phys to insert 4/256 ALIGNs in SSP connections (to account for the +5100 to -5100 ppm total difference); the disk drive is fine with 2/256.

Rules

Overall:

a) SSC for SAS physical links may be supported at 6 Gbps
b) SSC for SAS physical links shall not be supported at 3 Gbps or 1.5 Gbps
c) SSC for SATA physical links is supported at any physical link rate

Receiver rules:

a) Phy that support being attached to SATA phys (e.g., initiator and expander phys) shall support receiving with +5100/-100 ppm (e.g., from SAS-2 phys), +100/-100 ppm (e.g., from SAS-1.1 phys), and +350/-5350 ppm (e.g., from SATA phys) SSC;
b) Phys that do not support being attached to SATA phys (e.g. target phys) shall support receiving with +5100/-100 ppm (e.g., from SAS-2 phys) and +100/-100 ppm (e.g., from SAS-1.1 phys).

Transmitter rules:

a) SAS phys and expander phys attached to SAS phys or expander phys shall transmit with +5100/-100 or +100/-100 ppm SSC;
b) SAS phys and expander phys attached to SAS-1.1 phys shall transmit with +100/-100 ppm SSC;
c) SAS phys and expander phys attached to SATA phys shall transmit with +350/-5350 or +350/-350 ppm SSC.

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

a) phys originating dwords shall insert 4/256 ALIGNs outside of STP connections (i.e., inside SSP connections, inside SMP connections, and outside of connections altogether);
b) phys originating dwords shall insert 4/256 ALIGNS in STP connections;
c) phys originating dwords should include programmable ALIGN insertion rates for each case, in case anything else changes:
    A) outside connections;
    B) inside SSP connections;
    C) inside SMP connections;
    D) inside STP connections;
d) expander phys not originating dwords shall insert ALIGNs only as necessary (i.e., whenever the dword stream underflows);
e) 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 speed negotiation windows (SNW) G1, G2, and G3 without SSC
c) G3 window consists of a 1.5 Gbps speed negotiation followed by phy feature support information such as SSC capability
    A) the phy transmits whether or not it supports receiving with SSC;
    B) the phy receives indication about whether or not the other phy supports SSC;
d) 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

Complications

This requires HBA and expander ASICs to have 3 clock trees, which may not be viable. Figure 4 shows the problem.

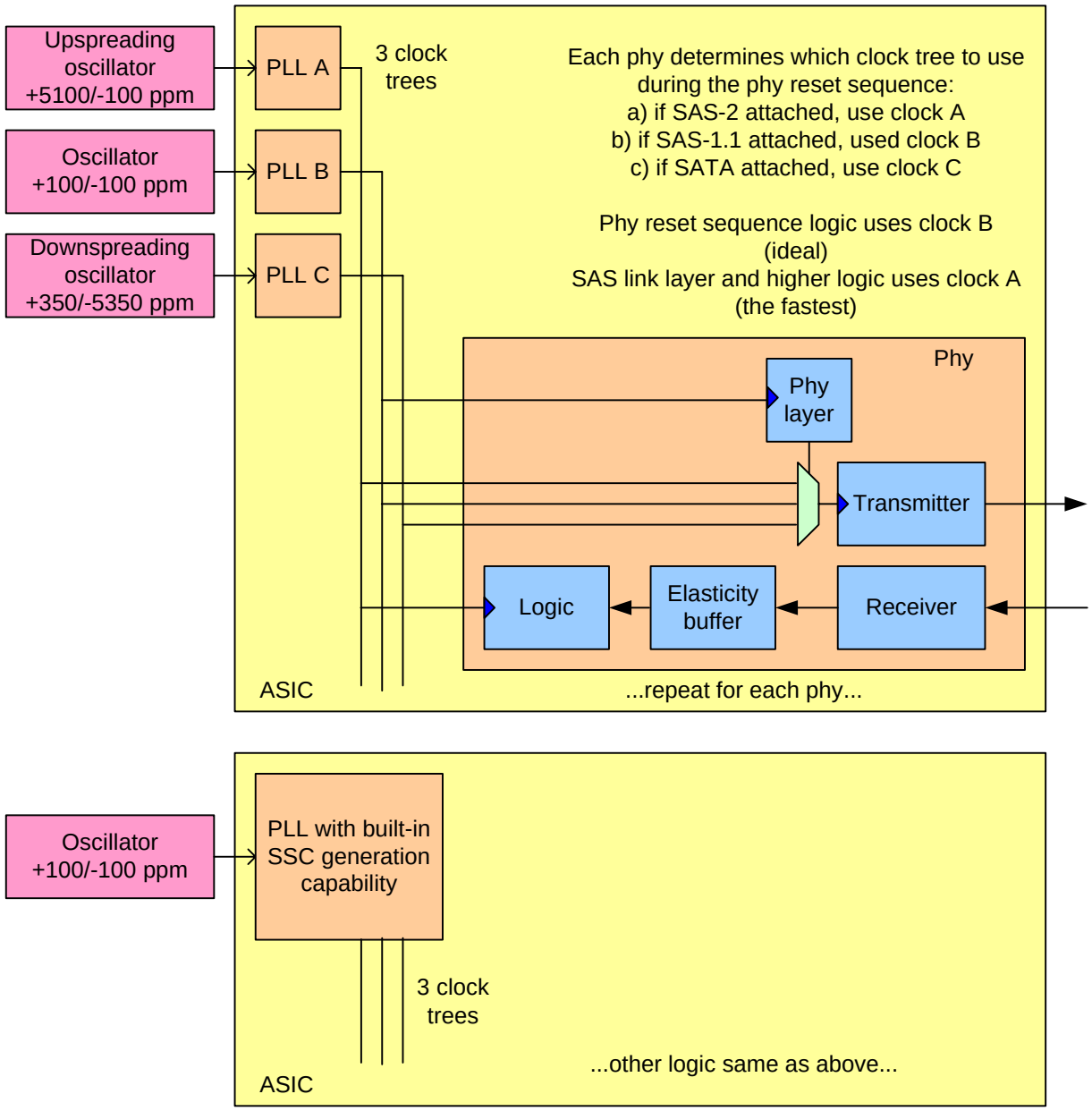

Figure 4 — Three clock trees

Some relief can be granted if phys are not universal. If a phy is known to attach only to disk drive bays, it could be restricted to choosing between clock B and clock C. If a phy is known to attach only SAS phys but not SATA phys, it could be restricted to choosing between clock A and clock B. If all devices are known to support SSC (a closed environment), then clock C could be used everywhere. Groups of phys (e.g. 4 phys) often share properties since SAS cable assemblies are 4-wide.

A dual-ported disk drive requires 2 clock trees, since even if its phys are attached to expander/initiator phys that have the same level of SSC support, they will not both be running the phy reset sequence and transmitting OOB signals at exactly the same times.

Downspreading is common in clock generators for the PC market, as it is used by PCI, PCI Express, Serial ATA, FB-DIMM, and other interface standards. Center-spreading is also common. Upspreading is uncommon.

A clock generator might have to be programmed at a base rate of <core + 5000 ppm> and then downspread from that, or programmed to a base rate of <core + 2500 ppm> and center-spread from that.

The choice of +5000 ppm mirrors SATA's use of -5000 ppm. According to 06-193, a receiver designed for -5000 ppm is probably able to handle +5000 ppm without much extra effort.

Other options to consider

Skew the outputs between transmitters in an ASIC with multiple phys so they don't all transition at the same time (e.g., phy 0 is nominal, phy 1 is 10 ps later, phy 2 is 20 ps later, etc.).

Implement SSC as a variable skew implemented at each transmitter inside the ASIC. Route only the ideal clock to them rather than three clocks.

Run devices on a board at slightly different but fixed frequencies (e.g., one nominally at +0 ppm, another at +2000 ppm, another at +4000 ppm). Around that base, they transmit with +100/-100 ppm variation.