The basics of synchronized Ethernet

By Slobodan Milijevic
Senior Applications Engineer
Zarlink Semiconductor Inc.
E-mail: slobodan.milijevic@zarlink.com

Over the past two decades, Ethernet has become the dominant technology for data transmission—in particular, with telecom and wireless providers—due to its simplicity and low cost. However, the asynchronous nature of Ethernet provides certain transmission challenges.

For example, time division multiplexing (TDM) services such as T1/E1 and SONET/SDH require synchronized clocks at both the source and destination nodes. Similarly, wireless base stations require synchronization to a common clock to ensure a smooth call hand-off between adjacent cells.

While there are several ways to achieve synchronization over Ethernet, one gaining momentum is Synchronous Ethernet (SyncE). SyncE uses the physical layer interface to pass timing from node to node in the same way timing is passed in SONET/SDH or T1/E1. This gives telecom and wireless providers confidence that networks based on SyncE will be not only cost-effective, but also as highly reliable as SONET/SDH and T1/E1 based networks.

As interest from carriers and service providers grows, many Ethernet equipment vendors are developing SyncE-enabled equipment targeting this lucrative new market. However, Ethernet equipment designers often lack in-depth understanding of synchronization and may underestimate the complexity of the issue.

A common assumption is synchronization over Ethernet can be achieved merely by replacing the free-running crystal oscillator used for Ethernet Physical Layer Device (PHY) with a general purpose synchronization device (PLL).

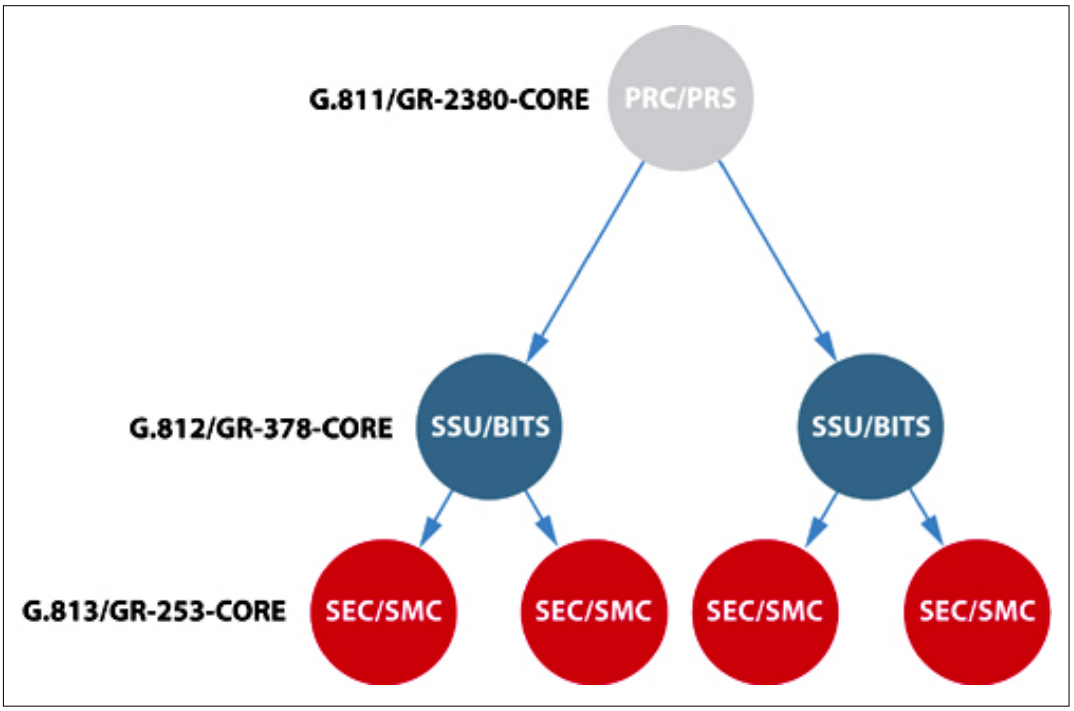

Figure 1: Network synchronization in telecom systems is based on clock hierarchy, with the highest accuracy clock at the top.

Certainly, this is not the case and designs based on such an assumption are destined to fail. This article discusses the basics of SyncE performance, reviews synchronization concepts and requirements and highlights some common design hurdles faced by board designers to help ensure right-first-time SyncE designs.

Network synchronization

Timing is critical for telecom system performance. Telecom systems are based on TDM technologies (T1/E1 and SONET/SDH), which are suited for transmission of constant bit rate traffic such as digitized voice and video. TDM technologies support small transmission delay and small transmission delay variation—two key parameters for voice quality and video transmission.

Small transmission delay can only be achieved if data buffering at each node is minimal. This implies that all nodes in a TDM-based network must be tightly synchronized to a common clock to prevent data loss. If one node has a slightly different frequency, even for a short period, the buffer at that node will either overflow or underflow and the data sample(s) will be lost or repeated to keep the bit rate constant.

Network synchronization in telecom systems is based on clock hierarchy, with the highest accuracy clock at the top (**Figure 1**).

At the top of the hierarchy is the Primary Reference Clock (PRC) or Primary Reference Source (PRS) with clock accuracy of 10-11. This means that a clock with this accuracy will have one extra or one less pulse for every 1011 pulse relative to the ideal clock. A wristwatch timed with such a clock would be off one second every 1011sec. (3,172 years).

PRC/PRS can be generated from a cesium (atomic) clock or from cesium clock-controlled radio signals, such as GPS, Global Orbiting Navigation Satellite System and Long Range Navigation System Version C.

At the next level of hierarchy is Synchronization Supply Unit (SSU) or Building Integrated Timing Supply (BITS). SSU/BITS includes holdover, a feature that allows it to generate a clock with higher accuracy than its intrinsic free-running accuracy for a short period after it loses synchronization with PRC/PRS. SSU/BITS is usually implemented with a Digital PLL (DPLL) driven by a rubidium clock.

The third level is the SDH Equipment Clock (SEC) or SONET Minimum Clock (SMC). SEC/SMC also features holdover, but its holdover and free-run accuracy performance is lower than what is required for SSU/BITS. SEC/SMC is usually implemented with a DPLL driven by an ovenized crystal oscillator (OCXO) or temperature-controlled crystal oscillator (TCXO). It should be noted that the second and lower levels of hierarchy will have clock accuracy equal to PRC/PRS, so long as their path to the

PRC/PRS is not broken.

For reliability reasons, it is unrealistic to expect all global telecommunication networks to be synchronized to a single PRC/PRS. Real networks use a flatter timing distribution structure with a number of PRC/PRS running independently. Each telecom provider usually has its own PRC/PRS, which means that the worldwide telecommunication network consists of synchronized islands connected with plesiochronous links.

While PRC/PRS and SSU/BITS are usually implemented as standalone products with timing-only functionality (no data transmission), SEC/SMC are almost exclusively implemented as a part of networking product such as an add-drop multiplexer.

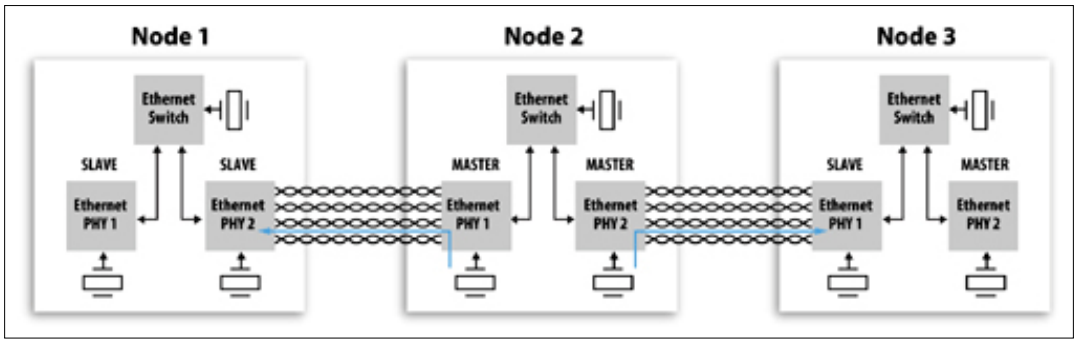

Figure 2: Synchronization does exist in Ethernet on each hop between two adjacent nodes, but it is not passed from hop to hop.

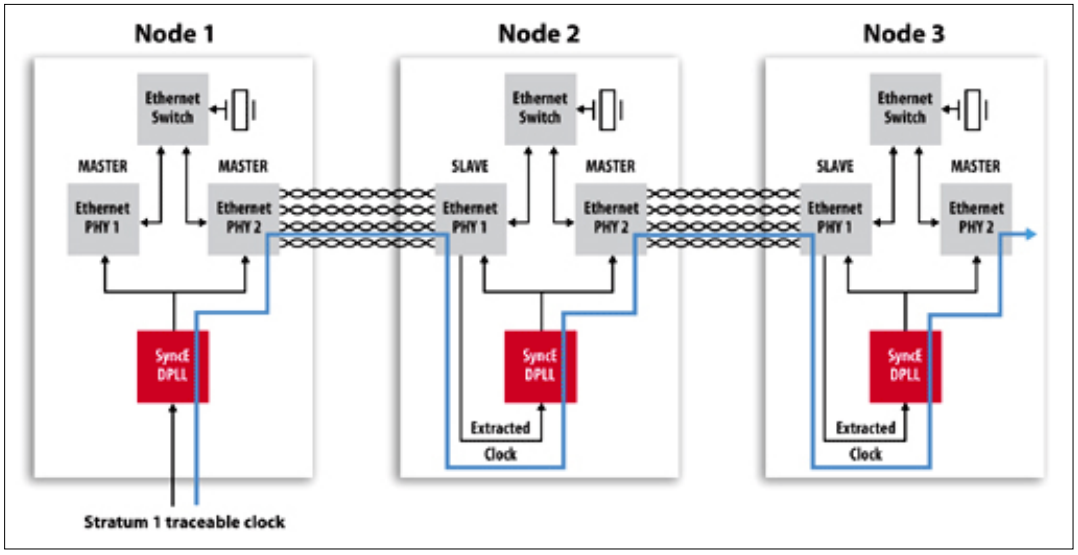

Figure 3: Passing synchronization is relatively simple—take the recovered clock from the node receiving synchronization and with this clock, feed all nodes that are transmitting synchronization.

Traditional vs. synchronized

Traditional Ethernet was originally intended for transmission of asynchronous data traffic, meaning there was no requirement to pass a synchronization signal from the source to destination. In fact, the old 10Mbit/s (10Base-T) Ethernet is not even capable of synchronization signal transmission over the physical layer interface because a 10Base-T transmitter stops sending pulses during idle periods.

A 10Base-T transmitter simply sends a single pulse ("I am alive" pulse) every 16ms to notify its presence to the receiving end. Of course, such infrequent pulses are not sufficient for clock recovery at the receiver.

Idle periods in faster Ethernet flavors (100Mbit/s, 1Gbit/s and 10Gbit/s) are continuously filed with pulse transitions, allowing continuous high-quality clock recovery at the receiver—good candidates for synchronized Ethernet.

Figure 1 highlights Gigabit Ethernet over copper (1000Base-T). To reduce clutter, each node has only two ports, although typically each node has multiple ports. Gigabit Ethernet over copper provides an additional challenge for SyncE implementation, which does not exist in Ethernet over fiber.

Gigabit Ethernet over copper uses line coding and transmission over all four pairs of CAT-5 cable to compensate for limited bandwidth of twisted pairs used in CAT-5 cables. The transmission is done in both directions simultaneously, similar to ISDN and xDSL where DSP algorithms have to be used for echo cancellation.

The echo cancellation is greatly simplified if the symbol rate (frequency at which data is transmitted) is identical in both directions. This is accomplished with a GbE master/slave concept.

The master generates a transmit clock locally from free-running crystal oscillator and the slave recovers the master clock from the received data and uses this recovered clock to transmit its own data. Master and slave are determined during the auto-negotiation process. The master is generally assigned randomly using a seed value but it can also be set manually.

Figure 2 illustrates that synchronization does exist in Ethernet on each hop between two adjacent nodes, but it is not passed from hop to hop. Passing synchronization is relatively simple—take the recovered clock from the node receiving synchronization and with this clock, feed all nodes that are transmitting synchronization (**Figure 3**).

Of course, the recovered signal needs to be cleaned with a PLL to remove jitter generated from the clock recovery circuit before being fed to the transmitting device. Ports need to be manually set in the clock path to alternate the master and slave function (only for 1000Base-T).

This is not an issue for Gigabit Ethernet over fiber (1000Base-X) or for 10GbE (10GBASE) because one fiber is always used for transmission and the other for reception—there is no bi-directional transmission on a single fiber. Thus, there is no need for master and slave functions.

Any Gigabit or 10GbE PHY device should be able to support synchronized Ethernet, so long as it provides a recovered clock on one of its output pins. The recovered clock is cleaned by the PLL and fed to the 25MHz crystal oscillator input pin on the PHY device. Some new Ethernet PHY devices provide a dedicated pin for the synchronization input. The advantage of this approach is that frequency input can be higher than 25MHz—higher clock frequencies usually have lower jitter. In addition, this approach avoids any potential timing loop problems within the PHY device.

Clean jitter

From the discussion so far, it appears that the only requirement for a PLL used in SyncE is to clean jitter from the recovered clock,

which can be accomplished with general purpose PLLs. However, the PLL used in SyncE must provide additional functions beyond jitter cleaning.

For example, if the receiving PHY device (Node 2, PHY 1 in Figure 3) gets disconnected from the line, the recovered clock frequency will either stop or start to drift depending on the implementation of the clock recovery circuit. The general purpose PLL will pass this big frequency change to the transmitting PHY device (Node 2, PHY 2 in Figure 3). As a result, not only is the transmission of synchronization signal going to fail, but the data transmission could fail as well.

The PLL used in SyncE must be able to detect failure of the recovered clock and switch the PLL to either another good reference in the system or into holdover mode. Requirements for SyncE are outlined in the timing characteristics of synchronous Ethernet equipment clock (ITU G.8262/Y1362) specifications. These specifications are based on ITU-T G.813 specification for SDH clocks. The major requirements of ITU-T G.8262/Y1362 are the following:

- *Free-run accuracy*—The accuracy of PLL output when it is not driven by a reference should be equal or better than ±4.6ppm over a time period of one year. This is a very accurate clock relative to the clock accuracy for traditional Ethernet (±100ppm).
- *Holdover*—The PLL constantly calculates the average frequency of the locked reference. If the reference fails and no other references are available, the PLL goes into holdover mode and generates an output clock based on a calculated average value. Holdover stability depends on the resolution of the PLL averaging algorithm and the frequency stability of the oscillator used as the PLL master clock.
- *Reference monitoring*—The PLL needs to constantly monitor the quality of its input references. If the reference deteriorates (disappears or drifts in frequency), the PLL raises an alarm (interrupt) and switches to another valid reference.
- *Hitless reference switching*—If the PLL's reference fails, it will lock to another available reference without phase disturbances at its output.
- Jitter and wander filtering—The PLL can be viewed as a jitter and wander filter. The narrower the loop bandwidth, the better the jitter and wander attenuation.
- *Jitter and wander tolerance*—The PLL should tolerate large jitter and wander at its input, and still maintain synchronization without raising any alarms.

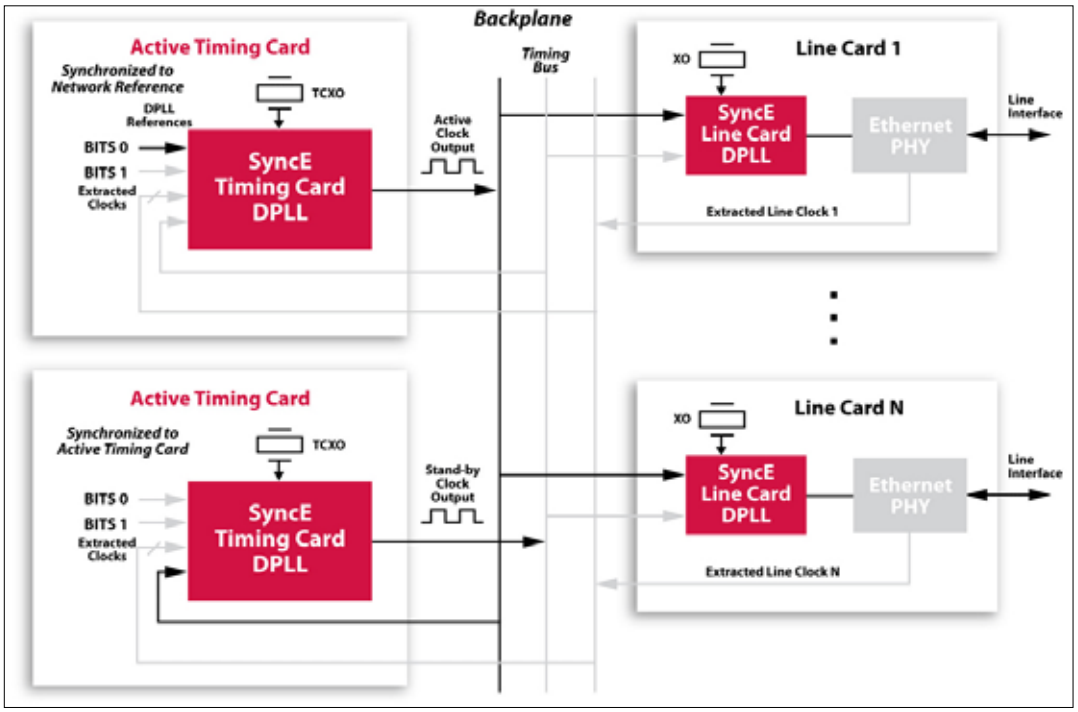

Figure 4: Timing in a carrier-grade SyncE system is composed of two timing cards that feed clocks to multiple line cards via a common backplane.

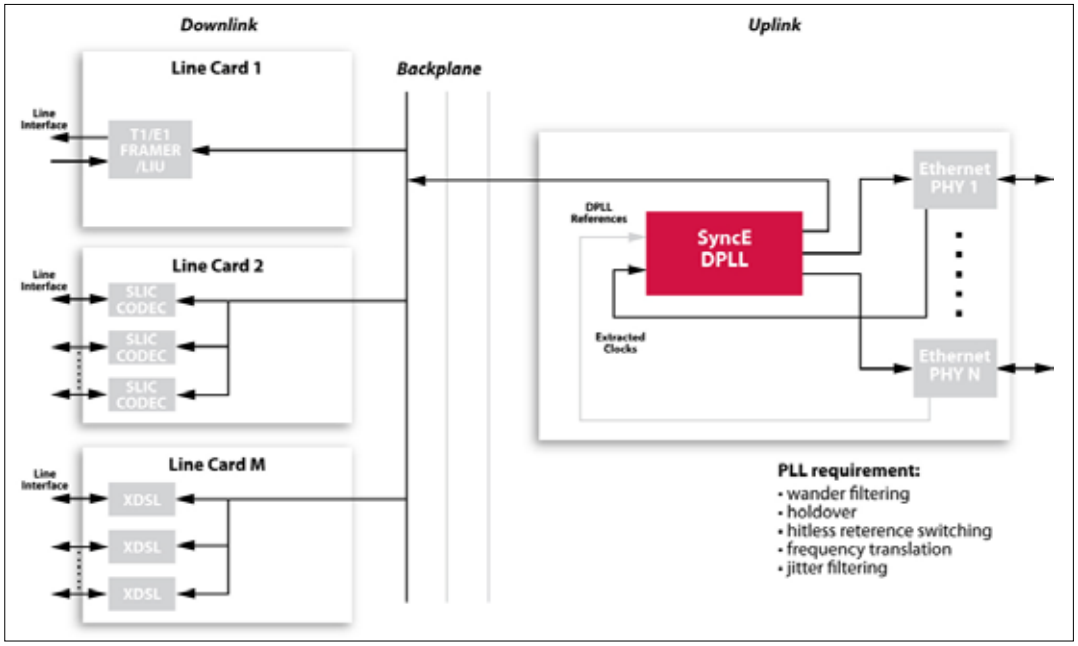

Figure 5: A next-generation DLC requires Ethernet and telecom clock frequencies.

These stringent requirements can be met only with a DPLL similar to DPLLs used for SONET/SDH clocks. The major difference is that a SyncE DPLL needs to be able to lock and generate clock frequencies used in Ethernet (25MHz, 125MHz and 156.25MHz), as opposed to telecom clocks (19.44MHz, 155.52MHz) used in SONET/SDH.

Carrier-grade SyncE systems must provide highly reliable operation under all network conditions. To do this, the most critical components within the system are made redundant, including timing.

Timing in a carrier-grade SyncE system is composed of two timing cards that feed clocks to multiple line cards via a common backplane (**Figure 4**). All line cards synchronize to the clock coming from an active timing card. If the active timing card clock fails (i.e. the card is unplugged), line cards will synchronize to the clock coming from the redundant timing card. Switching from one timing card to the other should not cause any interruption or failure in the system.

Having two timing cards protects against an internal failure if one of the cards fails. As seen in Figure 4, to protect from external clock reference failures, the timing cards are designed to be able to synchronize to more than one reference. A timing card accepts references from multiple sources, selects one, cleans it from phase noise with a DPLL, and distributes it to the line cards via the backplane. The DPLL is the most important part of the timing card. Timing card DPLL references can come externally from a SSU/BITS, internally from line cards, or from the other timing card in the system. Timing card DPLLs should meet all ITU-T G.8262/Y1362 requirements.

As seen in Figure 4, the line cards each have a DPLL that is used for jitter reduction and frequency translation. For example, frequency translation is required to convert from the 25MHz backplane clock to one or more clocks required by the Ethernet PHY, such as 125MHz, 156.25MHz, 155.52MHz or any other.

The line card DPLL must also provide hitless switching between the active and redundant clocks and provide clock continuity for a short period, such as when the active clock unexpectedly disappears before the system detects active reference failure and switches the line card DPLL to lock to the redundant reference. Like any DPLL, a line card DPLL requires a crystal oscillator.

However, this can be a low-cost oscillator as the line card DPLL is not required to go into holdover (except for short time periods when switching between active and redundant clocks). For long-term holdover, the system relies on a timing card DPLL. Therefore, a timing card DPLL requires higher quality crystal oscillators (TCXO, OCXO).

Smaller SyncE systems that do not need timing redundancy will generally have only one DPLL. This DPLL should meet all requirement of the timing card DPLL and the line card DPLL combined. This DPLL should have narrow loop bandwidth, good holdover (TCXO or OCXO required), hitless reference switching and very low intrinsic jitter. Depending on the application, this DPLL might also need to generate telecom frequencies such as 8KHz, 2.048MHz, 1.544MHz, 34.368MHz, 44.736MHz and many others.

Figure 5 illustrates a next-generation Digital Loop Carrier (DLC) requiring Ethernet and telecom clock frequencies. DLCs are installed in the neighborhood to aggregate traffic from multiple POTS, xDSL and T1/E1 to minimize the number of the lines going to the Central Office (CO) and increase xDSL data rates by shortening the length of the copper lines.

Aggregated traffic is carried to the CO via a fiber cable or several copper lines. Traditionally, DLCs have used SONET/SDH or T3/E3 to transmit data between the DLC and CO. However, these links are being replaced by Ethernet because of its lower capital and operational costs.