April 2022

TRANSPORT OF 800GE TRAFFIC OVER TWO 400GE FRAMES

Gilberto Loprieno

Emanuele Giacometti

Luca Della Chiesa

Jim Duda

TRANSPORT OF 800GE TRAFFIC OVER TWO 400GE FRAMES

AUTHORS:
Gilberto Loprieno
Emanuele Giacometti
Luca Della Chiesa
Jim Duda

ABSTRACT

The 800 gigabit per second (Gb/s) Ethernet (GE) technology is a new protocol that has been standardized by the Ethernet Technology Consortium. Techniques are presented herein that split an 800GE stream into two 400GE frames (e.g., 400GE-A and 400GE-B) without requiring any mapping. Under aspects of the presented techniques the 32 virtual lanes of the 800GE stream may be grouped into two groups of 16 virtual lanes; both the 400GE-A and 400GE-B links are protected by the Forward Error Correction (FEC) of the 800GE stream; alignment markers (AMs) are inserted according to the 800GE standard; the AM field CM2 is replaced every M AMs with two new markers, which may be used to identify the first 16 virtual lanes (0-15) and the second 16 virtual lanes (16-31) in the 800GE stream; and during transmission the new markers are inserted at the same time on all of the virtual lanes. In the event that a digital signal processing (DSP) facility does not support AM transparency, aspects of the presented techniques uniquely transcode AMs into special rate compensation markers that transport the CM2 and unique markers. At a receiving side, the 400GE-A and 400GE-B frames may be received, aligned by checking the CM2 field (with an extended deskew value), and then reassembled to recover the 800GE traffic.

DETAILED DESCRIPTION

The 800 gigabit per second (Gb/s) Ethernet (GE) technology is a new protocol that has been standardized by the Ethernet Technology Consortium. 800GE is defined through the reuse of 400GE building blocks. For example, it employs two 400GE Physical Coding Sublayers (PCSs) with the included Forward Error Correction (FEC). 400GE is a well-defined protocol that already is in mass production.

Techniques are presented herein that support splitting an 800GE stream over two 400GE streams by grouping the 32 virtual lanes of the 800GE stream into two groups of 16 virtual lanes each, generating two 400GE streams that may be forwarded to a coherent optics digital signal processing (DSP) facility. Figure 1, below, provides a high-level illustration of the described approach.

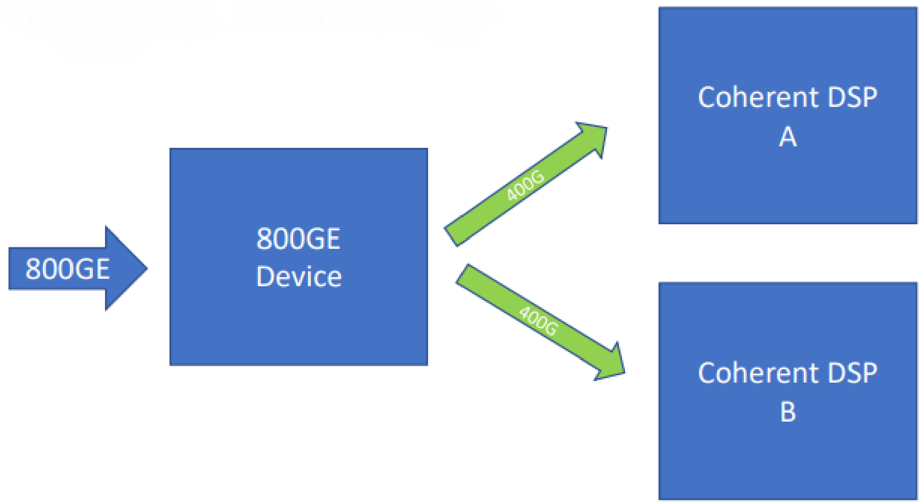

Figure 1: Illustrative Splitting of an 800GE Stream

It is important to note that in the arrangement that is depicted in Figure 1, above, the coherent DSPs must prevent the removal of AMs so that it is possible to uniquely mark AMs (as will be described below) in support of a simple splitting and reassembly of an 800GE stream.

According to aspects of the techniques presented herein, periodically two special markers (one for the first 400GE traffic stream and a different one for the second 400GE traffic stream) may be inserted, thus increasing the deskew capability of a receiver (as will be described and illustrated in the below narrative). In such a way, an 800GE stream may be split and transported as two 400GE traffic streams.

An 800GE stream may be transported over two 400GE streams (or on smaller containers) through available Flexible Ethernet (FlexE) or Flexible Optical Transport Network (FlexO)-XE technology but in a more complex way. For example, FlexE requires the insertion and removal of idles and the mapping of an 800GE stream over a FlexE frame.

Similarly, FlexO-XE requires a generic mapping procedure (GMP) mapping over a FlexO frame.

Importantly, the techniques presented herein do not require any mapping and they are timing transparent. Additionally, they can be easily implemented in a digital design (such as, for example, in an application-specific integrated circuit (ASIC)).

As noted previously, 800GE is a new protocol that has been defined by the Ethernet Technology Consortium. However, it has not yet been specified by the Institute of Electrical and Electronics Engineers (IEEE). Employing the techniques presented herein allows for the transport of 800GE traffic over a mature and already consolidated technology, such as 400GE.

800GE traffic is transmitted through 32 virtual lanes. Each virtual lane has its own unique alignment marker (AM) to allow for a reordering of the lanes. AMs are periodically inserted (every 163,840 x 257b blocks, which is every 102 microseconds) over all 32 lanes. 400GE employs a similar structure where most of the AMs are common with 800GE but support is for only 16 virtual lanes.

Aspects of the techniques presented herein split the 32 virtual lanes of an 800GE stream into two groups of 16 lanes, from which two 400GE split blocks may be constructed. Figure 2, below, illustrates elements of such an approach.

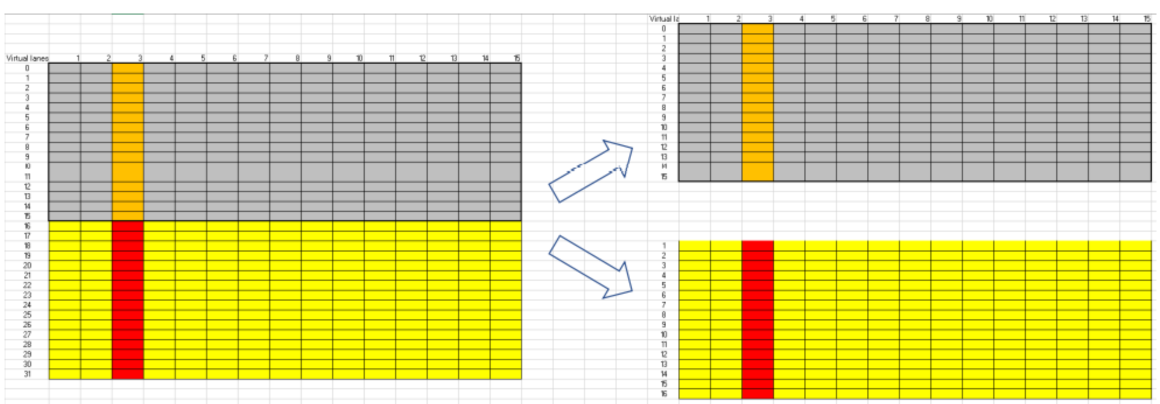

Figure 2: Illustrative Splitting of an 800GE Stream

Such a split, as described and illustrated above, is very simple and reliable since its group is protected by its own FEC. In this way, the 800GE stream may segmented and split over two different DSPs and then forwarded over different lambdas (but also different fibers).

Further aspects of the techniques presented herein periodically replace one AM field with a new marker. For simplicity of exposition, in the balance of the instant narrative the CM2 field will be employed. As defined in the specifications, the CM2 field is filled with a fixed pattern (0x26) and it is the same on all 32 of the virtual lanes. The CM2 value of 0x26 is based on a sequence of three ones and five zeros. Under the presented techniques, the same number of ones are employed, but in a transposed order. Such an ordering results in the new marker 0x91 for virtual lanes 0 through 15 and the new marker 0xC8 for virtual lanes 16 through 31.

Figure 3, below, depicts elements of an exemplary CM2 replacement approach according to aspects of the techniques presented herein and reflective of the above discussion.

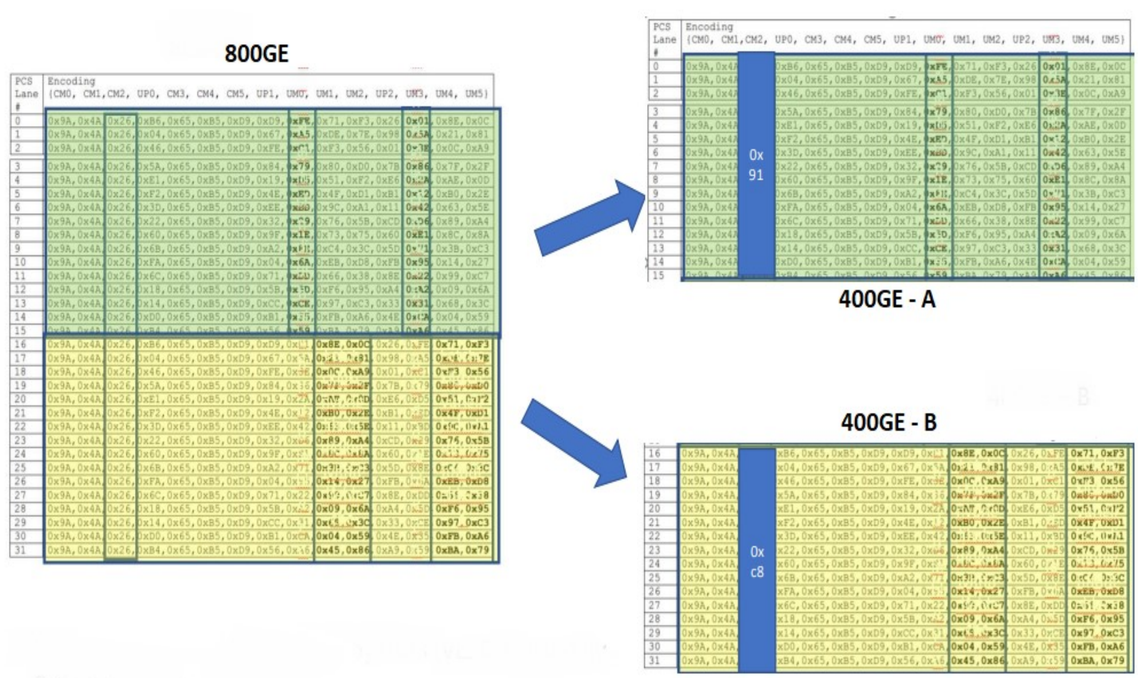

Figure 3: Exemplary CM2 Replacements

As depicted in Figure 3, above, the exemplary 800GE stream (comprising 32 virtual lanes each having a CM2 value of 0x26) is split into two 400GE streams. A first 400GE stream (which is identified in the figure as 400GE – A) comprises 16 virtual lanes each having a CM2 value of 0x91. A second 400GE stream (which is identified in the figure as 400GE – B) comprises 16 virtual lanes each having a CM2 value of 0xC8.

As described and illustrated in the above narrative, under aspects of the techniques presented herein an 800GE traffic stream may be segmented into two different groups, with each group transmitted along a 400GE path. Consequently, a receiver must be able to compensate for different skews. Figure 4, below, illustrates how virtual lanes may arrive at a receiving (Rx) side with different skews.

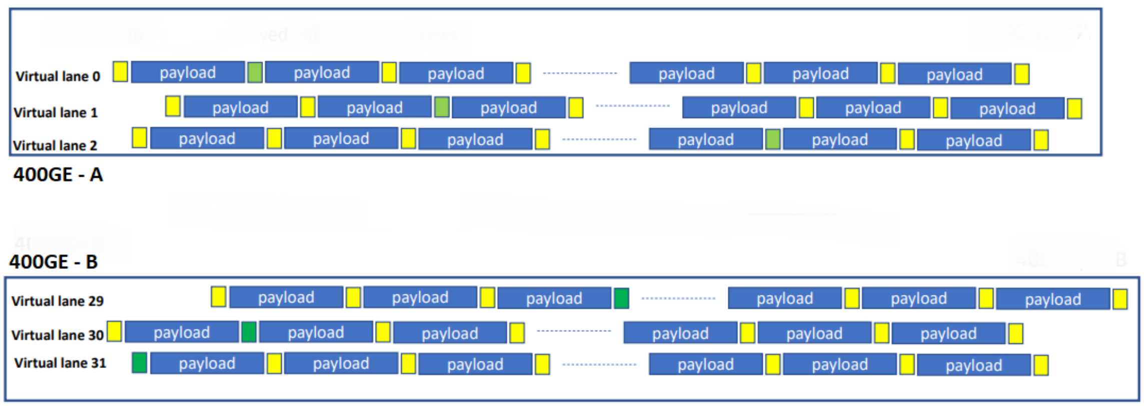

Figure 4: Illustrative Skewing

Accordingly, aspects of the techniques presented herein support replacing one CM every M AM insertions (i.e., M x 163,840 x 257b blocks) where M may be any number. At the same time, as described previously a CM2 value will be replaced with the value 0x91 for virtual lanes 0-15 and with the value 0xC8 for virtual lanes 16-31.

It is important to note that deskew in not limited by a transmitter but rather from the memory that is available at a receiver where an 800GE stream is reassembled. Increasing the value of M can compensate for a higher skew (e.g., assuming M is equal to 16 then the maximum deskew is 1.6 milliseconds).

After replacing a CM2 value with either 0x91 or 0xC8, as described above, the 32 virtual lanes may be grouped into two 400GE blocks. The first block transports virtual lanes 0 through 15 and the second block transports virtual lanes 16 through 31. Both of the blocks may then be sent to a coherent optical DSP. The DSP will terminate FEC and any correcting errors and then map the constant bit rate (CBR) traffic to a far-end side.

Importantly, according to the techniques presented herein AMs must be transparently transported. However, a DSP will usually remove AMs and then replace them with rate compensation (RC) markers (e.g., RC0 and RC1).

In light of this, aspects of the techniques presented herein support replacing IEEE standard 802.3 AMs with special RC markers or, alternatively, transcoding AMs into RC markers. Such new RC markers employ the leading 0xFF value to allow for a simple 265 - 257 transcoding (as the 0xFF value is acceptable as the first byte in the instant process); transparently pass CM0, CM1, and CM2 along with their inverts CM3, CM4, and CM5; use CM2 for the deskew feature; and transparently pass UM0 and UM1 (which may be used to differentiate between a 400GE-A block and a 400GE-B block).

Figure 5, below, illustrates elements of the process that was described above.

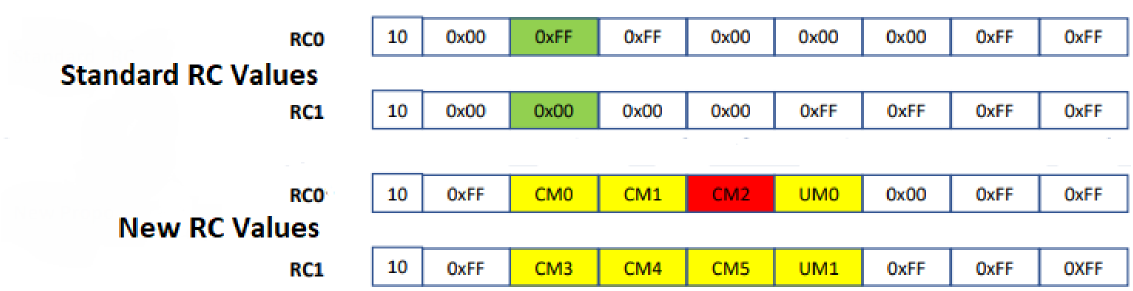

Figure 5: Exemplary RC Values

As depicted in Figure 5, above, the new RC values are transcoding AMs and are transparently transporting most of the common and unique markers, thus supporting the transparent transport of AMs.

At an endpoint, the UP0, UP1, UP2, UM2, UM3, UM4, and UM5 values may (according to aspects of the techniques presented herein) be replaced with their proper values from the IEEE 802.3 standard. In this way, the synchronization message that is transported through the new CM2 approach, as described and illustrated above, is preserved.

A far-end DSP may transcode the RC values into AMs and the device that is in charge of the reassembly of the 800GE stream from the two 400GE blocks (i.e., 400GE-A and 400GE-B, as depicted in Figure 3, above) may deskew all of the virtual lanes. Figure 6, below, illustrates elements of the process that was described above.

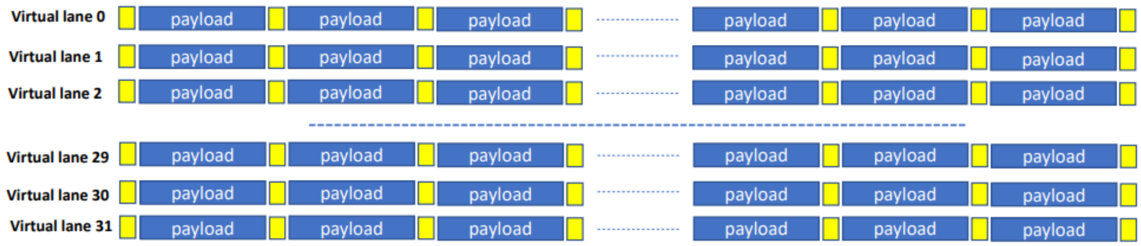

Figure 6: Exemplary Reassembled 800GE Stream

As described and illustrated in the above narrative, the techniques presented herein allow for the segmentation of an 800GE stream through the simple customization of AMs and improve the deskew capability when the 800GE stream is reassembled. Importantly, the presented techniques also support the transparency of AMs that usually are removed (under, for example, the Optical Networking Forum (OIF) standard 400ZR and the International Telecommunication Union (ITU) Telecommunication Standardization Sector (ITU-T) Recommendation G.709).

Importantly, splitting an 800GE stream into two 400GE blocks (according to the techniques presented herein) allows for the reuse of 800GE and 400GE designs with minimal changes (e.g., only to guarantee AM transparency). Further, the presented techniques are compatible with current AM state machines. Additionally, a network equipment vendor may implement support for aspects of the techniques presented herein in a line card.

It is important to note that under the techniques presented herein, it is not necessary to modify AMs in any application which requires less than 102 microseconds of skew. In the event that is necessary to improve deskew, then aspects of the presented techniques support periodically changing one CM. For example, by replacing an AM every 64 insertions, it is possible to deskew up to 6.5 milliseconds.

In summary, techniques have been presented herein that split an 800GE stream into two 400GE frames (e.g., 400GE-A and 400GE-B) without requiring any mapping. Under aspects of the presented techniques the 32 virtual lanes of the 800GE stream may be grouped into two groups of 16 virtual lanes; both the 400GE-A and 400GE-B links are protected by the FEC of the 800GE stream; AMs are inserted according to the 800GE standard; the AM field CM2 is replaced every M AMs with two new markers which may be used to identify the first 16 virtual lanes (0-15) and the second 16 virtual lanes (16-31) in the 800GE stream; and during transmission the new markers are inserted at the same time on all of the virtual lanes. In the event that a DSP facility does not support AM transparency, aspects of the presented techniques uniquely transcode AMs into special rate compensation markers that transport the CM2 and unique markers. At a receiving side, the 400GE-A and 400GE-B frames may be received, aligned by checking the CM2 field (with an extended deskew value), and then reassembled to recover the 800GE traffic.