Using Native 40 Gigabit Ethernet

What You Will Learn

The increasing need for high-bandwidth data transfer rates in the data center challenges organizations to increase the available bandwidth to both the servers and the backplane. One approach has been to use link aggregation and PortChannels consisting of multiple 10-Gbps links. However, this approach may not achieve the desired throughput because of hashing mechanism imbalances and the need to support higher bandwidth flows.
The solution is faster links using native 40 Gigabit Ethernet.

The IEEE 802.3ba task force was formed to develop a 40 Gigabit Ethernet standard to support transmission transfer rates of 40 Gbps. The standard was ratified in June 2010. Technology that meets this standard uses four lanes of 10 Gigabit Ethernet, as with link aggregation, but it can support true 40-Gbps flows, unlike link aggregation.

This document discusses the advantages of the IEEE 802.3ba standard over simple link aggregation and explains how native 40 Gigabit Ethernet based on IEEE 802.3ba works.

64/66-Bit Line Code

To understand native 40 Gigabit Ethernet, first consider a 10 Gigabit Ethernet network and its use of 64/66-bit line code to transfer data. The 64/66-bit line code is used in 10 Gigabit Ethernet to provide efficient bandwidth with about 3 percent overhead. With this encoding, 64 bits of data are transmitted as a 66-bit word after 2 bits are added to signify the kind of data that is being transferred. Figure 1 shows the process when a single 10 Gigabit Ethernet physical coding sublayer (PCS) lane is used.

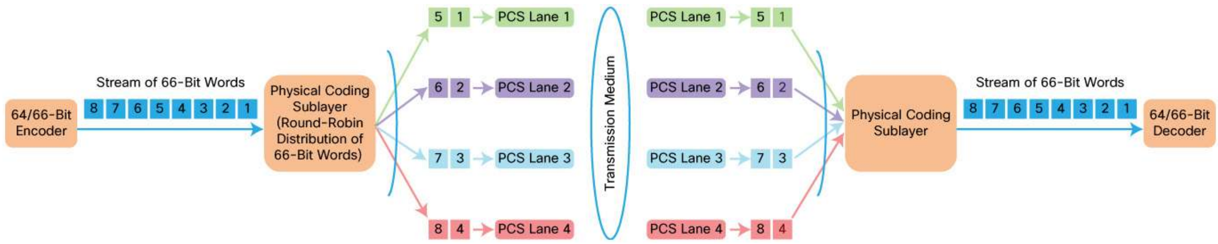
Figure 1. 64/66-Bit Line Code Process for 10 Gigabit Ethernet

Multilane Distribution

Using the 64/66-bit line code, the IEEE 802.3ba task force developed multilane distribution, which transmits each 66-bit word in a round-robin fashion across the four 10 Gigabit Ethernet lanes. This approach allows a single flow to use the aggregate bandwidth of the lanes.

Multilane distribution uses the physical coding sublayer to distribute the 66-bit words across the four lanes using a simple round-robin mechanism. On the receiving side, the physical coding sublayer rearranges the 66-bit words into their proper sequence and sends them to the 64/66-bit encoder. The multilane distribution unit and multiple SerDes (serializer deserializer) reassemble the stream and put it back into order. Figure 2 shows the multilane distribution process on the transmit side.

Figure 2. Multistream Distribution Process: Transmit Side



With native 40 Gigabit Ethernet based on the IEEE 802.3ba standard, line-rate bandwidth is achieved using all the physical lanes simultaneously and without the flow-based constraints seen with a 4 x 10 Gigabit Ethernet PortChannel and link-aggregation bonded link.

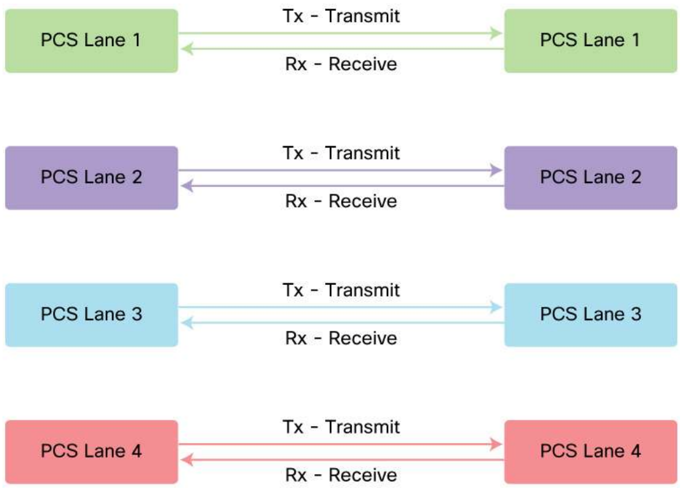
Native 40 Gigabit Ethernet's reuse of the 10 Gigabit Ethernet 64/66-bit encoder helps protect investments in existing technology. 40 Gigabit Ethernet's compatibility with the 10 Gigabit Ethernet standard helps reduce hardware development costs.

Transmission Medium

There are two options for transmitting the data across an optical fiber: using four fiber pairs or a single fiber pair.

You can use four pairs of fiber, transmitting each lane on a separate fiber pair. In Figure 3, each arrow represents one fiber-optic strand. A fiber pair is needed to create a single 10 Gigabit Ethernet link: one fiber for transmission and one for receipt. This is the approach traditionally used for typical 10 Gigabit Ethernet links. Thus, four pairs of fiber are needed to create a single 40 Gigabit Ethernet link.

Figure 3. Transmission Using Four Fiber Pairs

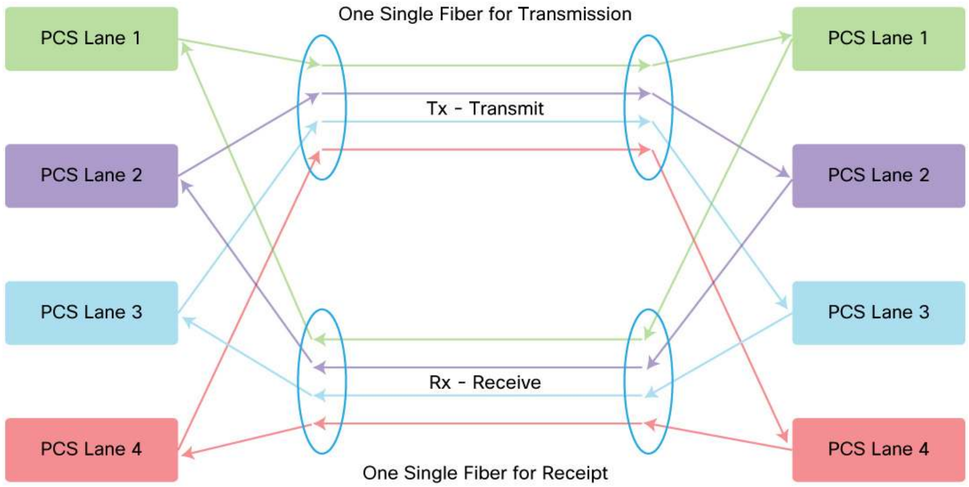


This type of transmission is often referred to as parallel optics because four pairs of lasers and receivers are used. Because there are four different pairs, the same optical light frequency (lamda or wavelength) can be used for each fiber pair.

You can also use different lambdas for each pair. When four different lambdas are used, you can combine all the lambdas onto a single fiber. With everything combined on a single fiber, coarse wave-division multiplexing (CWDM) technology is used.

In this method, a single fiber pair is used for both receive and transmit traffic. Four PCS lanes of traffic are still used. These lanes are separated on the wire by the use of different lambdas for each lane of traffic. Figure 4 shows this approach.

Figure 4. Transmission Using One Fiber Pair



This method supports link lengths of up to 10km using standard pair of G.652 single-mode fiber with duplex LC connectors.

10 Gigabit Ethernet Mode

40 Gigabit Ethernet inherently also works in 10 Gigabit Ethernet mode because it is essentially a stack of 10 Gigabit Ethernet lanes. 10 Gigabit Ethernet mode is easily implemented with parallel optics when the lambdas used match those of the 10 Gigabit Ethernet optics on the other end of the link. Fiber-optic or Twinax breakout cables are used to provide the final physical connectivity. In addition, the Enhanced Quad Small Form-Factor Pluggable (QSFP+) optics used for 40 Gigabit Ethernet fiber connections can work in both native 40 Gigabit Ethernet mode and 4 x 10 Gigabit Ethernet mode.

Conclusion

Native implementation of 40 Gigabit Ethernet is different from link aggregation, even though both look alike physically. Because of hashing issues and because only one link can be used per flow, full-bandwidth utilization is difficult to achieve with link aggregation. With native 40 Gigabit Ethernet, because the traffic is split across multiple lanes at a very low level, full bandwidth is utilized.

For More Information

http://www.cisco.com/en/US/prod/collateral/modules/ps5455/data_sheet_c78-660083_ps11541_Products_Data_Sheet.html.

http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps12806/guide_c07-726164.html.

http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps12806/guide_c07-726163.pdf.

http://www.ieee802.org/3/ba/.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-726674-00 03/13