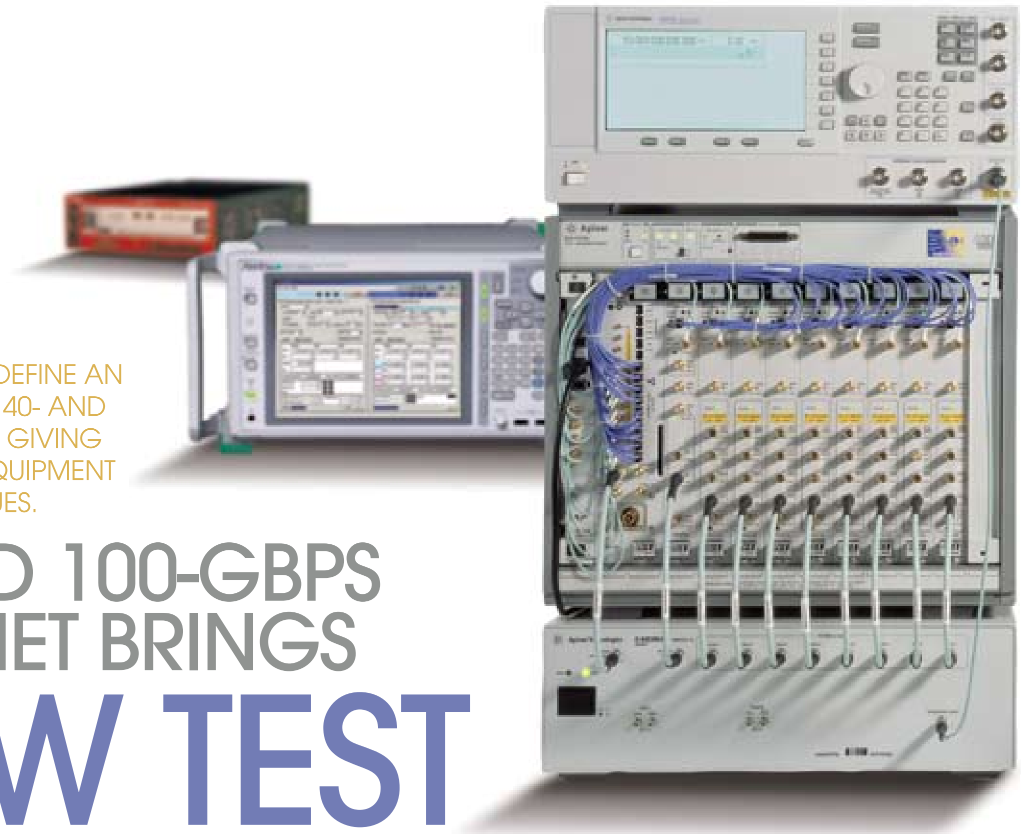

IEEE P802.3BA WILL DEFINE AN ARCHITECTURE FOR 40- AND 100-GBPS ETHERNET, GIVING RISE TO NEW TEST EQUIPMENT AND TEST TECHNIQUES.

40- AND 100-GBPS ETHERNET BRINGS NEW TEST CHALLENGES

BY MARTIN ROWE • SENIOR TECHNICAL EDITOR, *TEST & MEASUREMENT WORLD*

Internet and data-center users always demand higher bandwidth to carry voice, data, and especially video. Because of that demand, today's 10-Gbps optical and electrical links are running out of capacity. Data centers and core networks need faster links, according to participants at the OFCNFOEC (Optical Fiber Communication Conference and Exposition/National Fiber Optic Engineers Conference) in March 2008 (**Reference 1**). At the time, the IEEE P802.3ba Ethernet task force was developing a standard that will define an architecture for 40- and 100-Gbps Ethernet (**Reference 2**). Although IEEE P802.3ba is still in the works, engineers worldwide are beginning to develop products that employ these links, and those products will need testing. Much of the testing for IEEE P802.3ba will use 10-Gbps technology, but some tests will require new equipment and new techniques.

FLEXIBLE ARCHITECTURE

John D'Ambrosia, senior scientist at Force10 Networks and chairman of the IEEE P802.3ba task force, leads a team of engineers that is working out the details of IEEE P802.3ba. Although details are still forthcoming, the IEEE P802.3ba task force has defined a general architecture, the architecture's protocol sublayers, and how those sublayers function (**Reference 3**). **Figure 1** shows a simplified diagram of the architectures for 40- and 100-Gbps optical networks. Although the architectures differ slightly, IC manufacturers could implement both in the same part.

For optical networks, IEEE P802.3ba defines the PCS (physical-coding sublayer), the PMA (physical-medium-attachment) sublayer, and the PMD (physical-medium-dependent) sublayer. Both 40- and 100-Gbps implementations can use 10-Gbps fiber PHY (physical) links. In the future, the same architecture will support 100-Gbps links using four 25-Gbps lanes. IEEE P802.3ba also defines two additional sublayers for copper connections.

The PCS performs 64b/66b encoding on an aggregate 40- or 100-Gbps data stream to produce 66-byte blocks. It then sends the blocks across four lanes for 40-Gbps transmission or 20 lanes for 100-Gbps transmission. "We chose 20 lanes because [20] is divisible by one, two, four, five, and 10," says D'Ambrosia. So far, it appears that PCS will operate with four or 10 lanes. Each of the 20 PCS lanes will include a lane marker that identifies a lane and provides timing information for each data block (**Reference 4**). **Figure 2** shows how the 66-bit blocks will travel within the PCS lanes. Distribution of the blocks occurs in a round-robin fashion across the PCS lanes. A striping process divides the blocks across the PCS lanes.

The PMA sublayer matches the number of PCS lanes to the number of lanes that a physical layer requires. For 40-Gbps transmission, the PMA sublayer maintains four lanes. For 100-Gbps transmission, the PMA sublayer converts 20 XLAUI (40-Gbps-attachment-unit-interface) or CAUI (100-Gbps-attachment-unit-interface) lanes to 10 XLAUI or CAUI lanes. The PMD sublayer provides the final interface to a physical medium. **Table 1** describes these physical media and their respective transmission distances. In general, 40-Gbps links will find use in data centers, and 100-Gbps links will operate in core networks.

IEEE P802.3ba's flexible architecture supports long-reach and extended-reach optical links, which use WDM (wavelength-division multiplexing) across SMF (single-mode fiber). Short-reach links use multiple fibers, each carrying a different lane. Because optical and electrical components that send and receive 10-Gbps signals are now available, initial implementations of 40- and 100-Gbps transmission systems will use four and 10 10-Gbps lanes, respectively. Four-lane implementations using 25-Gbps links—actually, 25.78125 Gbps due to encoding—will take some time to appear, but communications carriers plan to use WDM, which has four wavelengths, on single-mode fiber for 40-km transmissions (**Reference 5**).

TESTING P802.3BA

Because 10-Gbps lanes will appear first, you'll be able to use your 10-Gbps optical test equipment to measure parameters such as timing jitter, amplitude, and BER (bit-error rate). You can test only one transmission path at a time, however, until equipment that can test

AT A GLANCE

- Today's 10-Gbps optical and electrical links are running out of capacity.
- The IEEE P802.3ba standard defines the PCS (physical-coding sublayer), the PMA (physical-medium-attachment) sublayer, and the PMD (physical-medium-dependent) sublayer.
- Initial implementations of 40- and 100-Gbps transmission systems will use four and 10 10-Gbps lanes, respectively.
- When testing an implementation that uses four or 10 fibers or wires in each direction, you must individually test each path.
- You'll be able to use your 10-Gbps optical test equipment to measure parameters such as timing jitter, amplitude, and BER (bit-error rate).

TEST EQUIPMENT FOR 40- AND 100-GBPS ETHERNET

Because testing the 10-Gbps physical lanes that comprise 40 and 100 links is similar to testing 10-GbE (gigabit Ethernet), you can test the physical layer of 40 and 100 GbE with oscilloscopes, BER (bit-error-rate) testers, and optical spectrum analyzers. Currently, little test equipment is available for testing at 100 Gbps, but some is starting to reach the market. For example, Ixia used its 100GE development-accelerator system to perform a proof-of-concept demonstration of 100-Gbps Ethernet in June 2008 at NXTcomm08 (Figure A and Reference A). The system, which Ixia formally announced to the public in September 2008 and began shipping in February 2009, generates and analyzes Layer 2 Ethernet traffic at 100 Gbps.

Among other recently introduced test equipment for Ethernet are products from Agilent, Anritsu, and Centellax. For example, Agilent Technologies' E4899A serial-BER tester for research labs and standards bodies lets you perform BER tests at 40 and 100 Gbps (Figure B and Reference B). In March, Agilent introduced the N4931A optical-modulation analyzer, which analyzes 40- and 100-Gbps optical signals using microwave techniques. Anritsu recently announced I/O cards for its MP1800A signal analyzer, which lets you modulate and analyze 40- and 100-Gbps optical signals (Figure C and Reference C). Further, Centellax recently introduced a 40-Gbps clock and data multiplexer. You can use the instrument to perform BER measurements on physical-layer links. In addition, PicoSolve, which Exfo acquired in February, manufactures PC-based optical sampling oscilloscopes that can characterize and monitor high-speed transmissions at 40 Gbps.

Although 10-GbE test equipment will get things started, IEEE P802.3ba will likely give rise to new test equipment. For example, testing with a four- or 10-lane BER system takes less time than using a single-channel instrument. Protocol analyzers that decode the data blocks into Ethernet packets will also help. In addition, testers that inject unexpected alignment blocks and remove expected alignment blocks will help

Figure A Ixia used its 100GE development-accelerator system in a demonstration of 100-Gbps Ethernet.

Figure B Agilent Technologies' 40/100-Gbps BER tester lets you test physical layers for bit errors.

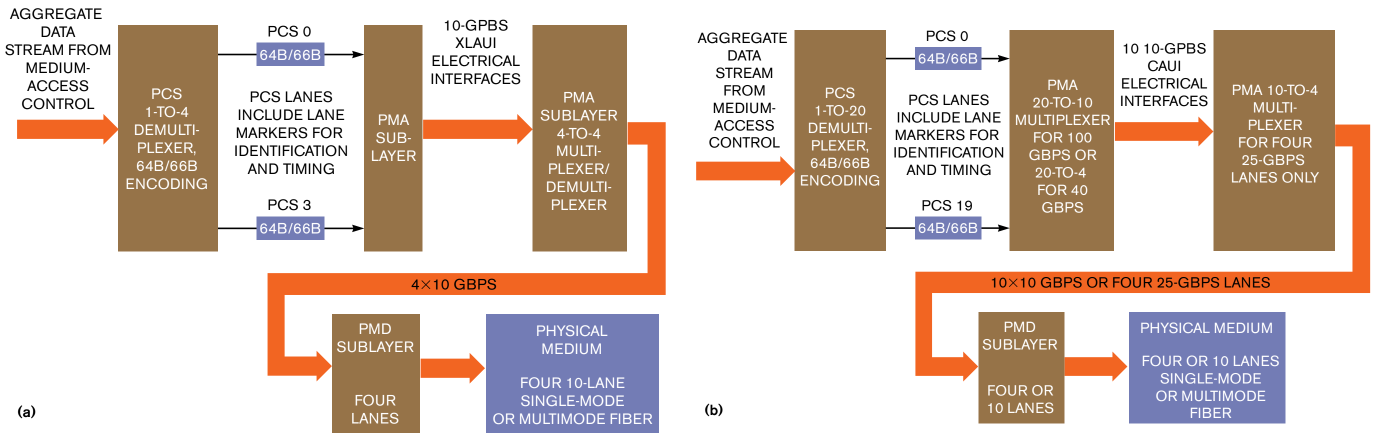

Figure 1 The IEEE P802.3ba architecture defines three sublayers (brown boxes) for optical transmission that support four data lanes to achieve 40 Gbps (a) or four to 10 data lanes to achieve 100 Gbps (b).

four or 10 lanes becomes available. "Tests for the individual optical lanes will be very similar to tests for existing 10-Gbps technology," says Edward Nakamoto, director of hardware for Spirent Communications. Testing multilane Ethernet will require you to start by individually testing each lane for signal integrity and BER (see **sidebar** "Test equipment for 40- and 100-Gbps Ethernet"). When testing four or 10 10-Gbps lanes, you must illuminate all lanes, preferably with data, and perform BER measurements on each lane. You then must modulate the laser with the data. **Reference 4** also describes modulation techniques that are in development for use with single-mode fibers for long-haul transmissions.

you test your network link. IC manufacturers are currently working with test-equipment makers to develop the components that will carry the protocol sublayers. One example is Altera, which has announced that its high-end FPGAs are available with 24 transceivers that can communicate at 11.3 Gbps. Thus, one device can handle 10 or 11 10-Gbps lanes with additional transceivers available for generating errors and monitoring traffic. The FPGA can connect directly to optical modules, and it can hold the PCS (physical-coding sublayer) and PMA (physical-medium-attachment) sublayer in the device. If designers use the Altera device, however, test engineers won't have access to the interface between the PCS and PMA sublayers for testing. It's an engineering trade-off that you may have to make.

The IEEE P802.3ba task force also plans to define an architecture that provides for implementations that use four 25-Gbps physical lanes to achieve 100-Gbps throughput. These implementations will need new technology in the form of ICs and optical components that will run at those speeds. Modulation methods such as DP-QPSK (dual-polarization quadrature-phase-shift keying) and DQPSK (dual quadrature-phase-shift keying) that can handle those speeds are under development. The PSK (phase-shift-keying) methods could solve problems due to polarization-mode and chromatic dispersion (Reference D).

The higher data rate will require engineers to develop electrical and optical components that can reach that speed, and test equipment will need to keep up. For example, real-time oscilloscopes will need even more bandwidth than the current 30-GHz, state-of-the-art technology available in LeCroy's WaveMaster 8 Zi series (Reference E). You also need BER testers, clock-recovery units, optical spectrum analyzers, and other equipment that can work with signals at those speeds. Stressed-eye testers will also let you test optical receivers for the added signal distortion that will occur at the higher bit rate.

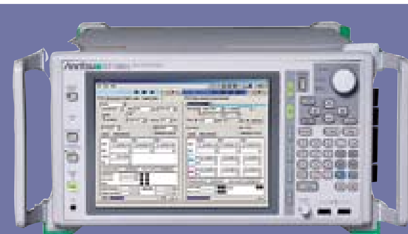

Figure C I/O cards for Anritsu's MP1800A signal analyzer let you modulate and analyze 40- and 100-Gbps optical signals.

REFERENCES

A "High Speed Ethernet (HSE) 100 GE Proof of Concept Demonstration," www.ixiacom.com/hse/100GE-POC.

B "40 Gb/s-100 Gb/s Serial BERT (E4899A)," Agilent Technologies, www.home.agilent.com/agilent/redirector.jspx?action=ref&cname=PRODUCT&ckey=747185&lc=eng&cc=US.

C "Flexible, Highly Accurate 40G/100G Test Solution Introduced by Anritsu Company," Anritsu, www.us.anritsu.com/news/2008/Flexible,-Highly-Accurate-40G/100G-Test-Solution-Introduced-by-Anritsu-Company_ARNid880.aspx.

D "Trends and Issues in Ultra-High-Speed Transmission Technologies," Anritsu, November 2008, www.us.anritsu.com/downloads/files/MP1800A_40_100GbE_EE1200.pdf.

E "LeCroy oscilloscopes reach 30 GHz," *Test & Measurement World*, Jan 7, 2009, www.tmworld.com/article/CA6626932.

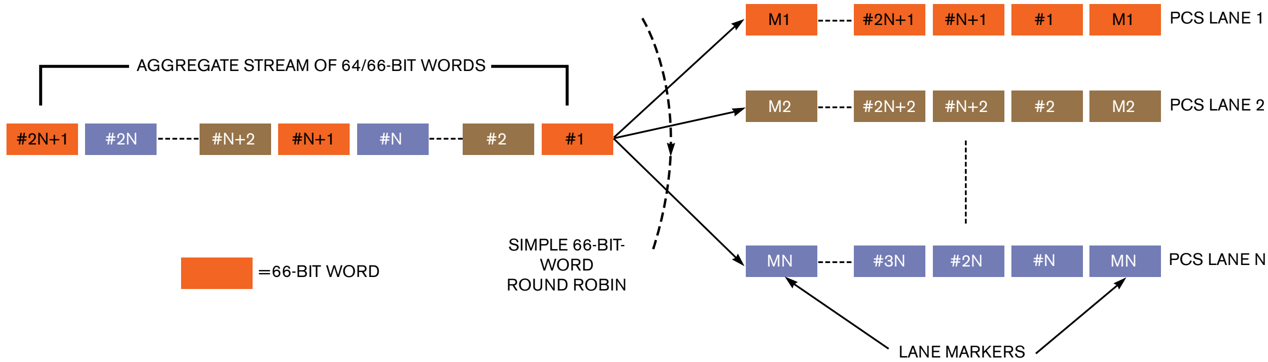

Figure 2 The PCS places 66-bit blocks into 20 lanes that include lane markers (courtesy Ethernet Alliance).

Michael Fleischer-Reumann, strategic-product planner at Agilent Technologies, notes that you can test an optical medium by generating PRBS (pseudo-random-bit-sequence) test patterns with a BER tester. When testing an implementation that uses four or 10 fibers or wires in each direction, you must individually test each path. "When testing a WDM-multimode fiber with four or 10 wavelengths, you'll need a tunable laser," he says. Testing individual lanes is a good start, but you must also test lane timing and skew. When you are using WDM fiber, each data stream uses a unique wavelength, but each wavelength has a different propagation speed. Optical receivers must compensate for timing differences in transmissions.

To achieve that compensation, IEEE P802.3ba will define alignment blocks at the PCS that convey timing information. Alignment blocks appear once every 16,384 blocks in a data stream (**Reference 5**). Receivers use those blocks to realign the lane blocks before reconstructing the data stream. "The standard is very insensitive to skew," says Jeff Lapak, 10 Gigabit Ethernet Consortium manager at the UNH-IOL (University of New Hampshire Interoperability Lab). "It will be difficult to break alignment as part of our testing, but we'll do it." Lapak intends to test for skew-induced errors by removing alignment blocks from where they belong and inserting them where they don't belong in a data stream. Unfortunately, detecting errors is complicated because testers may not have access to each sublayer. IC designers will try to integrate the sublayers into as few devices as possible, so it's likely that the PCS, PMA sublayer, and PMD sublayer won't reside in three ICs. At least two—and perhaps all three—may reside in a single IC.

Lapak is waiting to see how IC designers package the sublayers before UNH-IOL engineers and students develop test tools. "IOL wants to test where sublayer interfaces will be the most consistent among manufacturers," he says. Companies such as Sarance Technologies have begun implementing 40- and 100-Gbps Ethernet cores into Xilinx FPGAs (**Reference 6**). This implementation combines the PCS and the PMA sublayer, and it adds MAC (media-access control), a layer above PCS. Because the sublayers reside in ICs, the interfaces between the devices will be electrical regardless of the physical-transport medium. Electrical crosstalk will be a significant challenge. "Interference and crosstalk may occur on the electrical side," says Toshihiro Suzuki, assistant manager at the product-planning center of the marketing division of Anritsu and a member of the IEEE P802.3ba task force. "At the design and verification stage, engineers must test for crosstalk and interference by using a multichannel pattern generator," he says. You'll then need an oscilloscope to look at adjacent lanes—if you can get access to them.

At the system level, you must test the multiplexing and demultiplexing functions. "Tests [must] be developed [for] the breaking up of the traffic into multiple lanes," says Spirent's Nakamoto.

Lapak of the UNH-IOL goes a step further, saying that he expects to test block encoding and decoding by reversing the order of the data on PCS lanes. "You don't know which lane will carry a block," he says. "Every lane has to be able to carry blocks from any other lane." Thus, he expects to develop test tools to reverse the order of the lanes and test whether a receiver will receive alignment information and reconstruct the data. EDN

TABLE 1 TRANSMISSION DISTANCES FOR PHYSICAL MEDIA

Minimum distance	Medium	40-Gbps Ethernet	100-Gbps Ethernet
1 m	Backplane	40GBaseKR4	
10m	Copper cable	40GBaseCR4	100GBaseCR10
100m	OM3 multimode fiber	40GBaseSR4	100GBaseSR10
10 km	Single-mode fiber	40GBaseLR4	100GBaseLR4
40 km	Single-mode fiber		100GBaseER4

A version of this article appears on the *Test & Measurement World* Web site, www.tmworld.com/article/CA6640178. There you will find a list of test equipment introduced at OFCNFOEC 2009.

+ For a list of references cited in this article, as well as more information on the companies it mentions, go to www.edn.com/090423df.