PennWell®

LIGHTWAVE®

EDITORIAL GUIDE

40G/100G Component Technology Advances

Requirements for data rates greater than 10 Gbps have emerged in both carrier and enterprise networks. A new generation of technology has emerged to meet these requirements, including building blocks for new classes of transceivers/ transponders. This Lightwave Editorial Guide looks at both component advances and module work being done for 40G/100G applications.

Next-gen 40/100 Gigabit Ethernet transceiver technologies and designs

With the first generation of modules beginning to ship, work is already underway to reduce transceiver power and form factor.

By **JON ANDERSON** and **ROB HANNAH**

WITH THE COMPLETION of 40G/100 Gigabit Ethernet (GbE) optical interface standards (IEEE 802.3ba-2010) and pluggable optical transceiver module specifications (CFP-MSA Rev 1.4), and with the production shipment of first-generation 40GbE/100GbE CFP products underway, optical module vendors are focusing on developing technologies and proving design-ins for their next-generation 40/100GbE pluggable optical transceivers. Key objectives include significant reductions in module power dissipation and size, which are critical to increasing system port density and reducing overall optical port cost for system vendors and their customers.

The 100GbE CFP module provides the highest faceplate density (in terms of Gbps per faceplate aperature-module pitch area) for MSA-specified pluggable optical modules to date. However, from a systems point of view, the CFP port density most likely will be limited by thermal constraints on power dissipation, which may typically be greater than 25 W for first-generation 100GbE CFP modules. This article discusses some of the key challenges facing optical module vendors considering these design objectives and outlines some of the more promising technical approaches to tackle and overcome these challenges.

First-generation 40/100GbE pluggble optical transceivers

Let's first review the technologies and designs of choice in the first generation of 40GbE/100GbE pluggable optical transceivers. The 40GbE and 100GbE optical interfaces specified in IEEE 802.3ba-2010 are summarized in Table 1.

Table 1. IEEE Std 802.3ba-2010 optical interfaces.

Physical Medium	40GbE	100GbE
Multimode Fiber (at least 100 m OM3)	40GBase-SR4	100GBase-SR10
Singlemode Fiber (at least 10 km)	40GBase-LR4	100GBase-LR4
Singlemode Fiber (at least 40 km)	Not Defined	100GBase-ER4

From a market application perspective, the 40GbE and 100GbE LR4 singlemode fiber optical interfaces are high priority and present the greatest technological challenges when it comes to significant reductions in transceiver power dissipation and form factor size.

CFP 40GBase-LR4 design. The first-generation 40GBase-LR4 optical transceiver is based on a 4x10G architecture that comprises the following discrete components:

- 10G CWDM 1310-nm uncooled directly modulated distributed feedback (DM-DFB) transmit optical subassemblies (TOSAs)
- 10G PIN photodiode (PD) with integrated transimpedance amplifier (TIA)
- receive optical subassemblies (ROSAs)
- four-channel optical multiplexing/de-multiplexing filters
- quad dual-channel clock/data recovery (CDR) IC. The quad CDR IC provides the XLAUI 4x10G electrical interface defined in the IEEE Std 802.3ba-2010 specification.

These components are packaged into the pluggable CFP module; the module's mechanical, electrical, and management interface specifications are given in the recently completed CFP MSA Rev. 1.4, as shown in Figure 1. The CFP module-host system management interface is based on the IEEE Std MDIO/MDC interface and includes several new features, such as programmable controls and alarms, module state transitions, and error rate calculations. The first-generation design leverages existing 10G optoelectronic device technology and uses innovations in packaging to realize high-performance, low-cost modules for high-volume production.

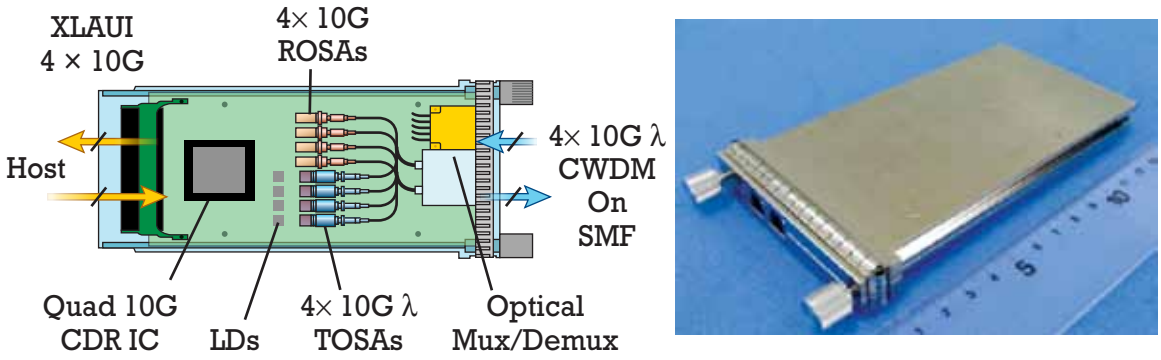


Figure 1. *First-generation 40GBase-LR4 optical transceiver design and CFP pluggable module.*

The CFP 40GBase-LR4 module power dissipation is typically in the range of 6 W, which fits well within the CFP module's 32-W power maximum. Thus, there is considerable interest in reducing the 40G-LR4 module form factor in next-generation designs for increased 40GbE port density. This will be addressed later in this article.

CFP 100GBase-LR4/ER4 design. The first-generation 100GBase-LR4/ER4 optical transceiver architecture is similar to that of the 40GBase-LR4, but with the speed of the active optoelectronic components increased to 28 Gbps for realizing a 4x28G optical interface. Additionally, the CAUI electrical interface defined in IEEE 802.3ba-2010 is widened from 4x10G lanes to 10x10G lanes. A 10:4/4:10 "gearbox" serializer/deserializer IC is used to implement the electrical interface between the 10-lane host data path and the four-lane optical data path.

The optical interface defined in IEEE 802.3ba-2010 uses a four-wavelength LAN-WDM 800-GHz wavelength grid in the 1310-nm band and optical multiplexing/de-multiplexing on singlemode fiber. The transmitter optical specifications for LR4 and ER4 are based on cooled electro-absorption modulation with integrated DFB (EA-DFB) laser technology, but were written to allow eventual implementation with directly modulated DFB lasers for smaller size, lower power consumption, and lower cost TOSAs.

The receiver optical specifications for LR4 and ER4 are based upon PIN-PD detector technology with integrated TIA. The receiver specification also includes optical amplification, such as from a semiconductor optical amplifier, to compensate for optical fiber attenuation loss in the ER4 40-km application.

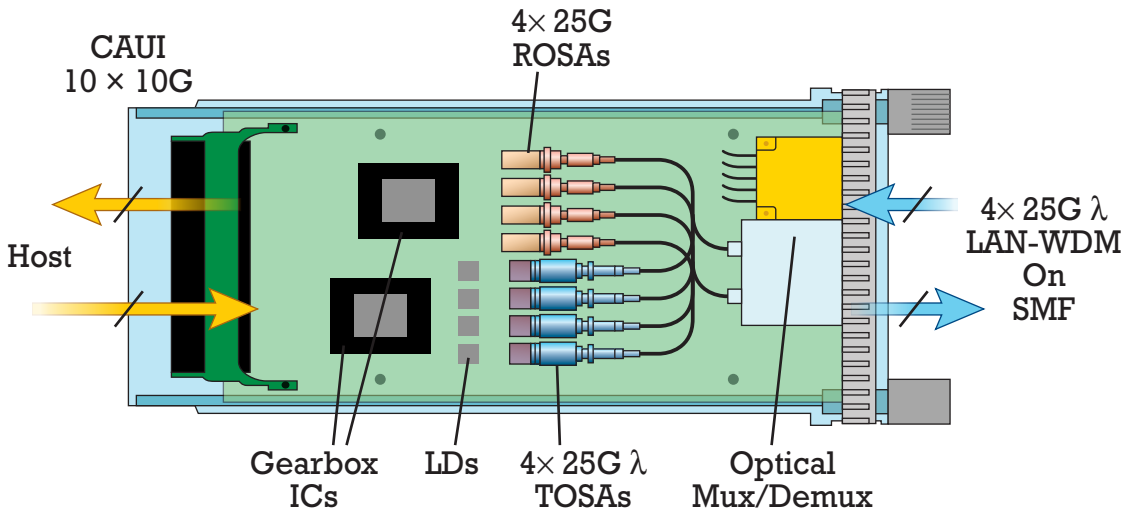


Figure 2. *First-generation 100GBase-LR4 optical transceiver design.*

These components are packaged into the CFP pluggable module (previously shown in Figure 1) with non-coaxial, 28-Gbps electrical connections between the discrete component TOSAs, ROSAs, and gearbox IC, as illustrated in Figure 2.

The first-generation 100GBase-LR4 module power dissipation is typically in the range of 24 W, which poses significant thermal management challenges for system designers, particularly as they seek to increase 100GbE optical port density. Thus, there is strong motivation to significantly reduce the 100GBase-LR4 optical transceiver module power dissipation in the next-generation design.

Next-generation 40GbE/100GbE optical module design targets

For the next-generation 40GbE/100GbE optical transceiver modules, system designers want significant reductions in power dissipation and form factor size. These objectives are particularly critical as system houses work to scale their core switching and routing input/output capacities and reduce constraints on port densities due to thermal management limits.

For 40GBase-LR4, the priority target is module form factor reduction. In terms of faceplate density, the current CFP form factor for 40GbE ports is 2.5X less efficient compared to the 100GbE CFP.

One approach the CFP design enables is to double or possibly triple the number of 40GbE ports within a single CFP module. While this approach increases port

density, it suffers from reduced port provisioning modularity.

A more feasible approach is to make use of the existing QSFP+ form factor (SFF-8436) and the non-retimed XLPPI electrical interface specified in IEEE 802.3ba-2010. This approach increases faceplate density by more than 60 percent over the CFP while retaining port modularity. To make this switch, however, optical module vendors need to not only reduce the physical size of their optical components, but they must reduce component power dissipation by over 50 percent so as to fit into the 3.5-W maximum power envelope of the QSFP+.

For 100GBase-LR4/ER4, the priority target is power dissipation reduction. System makers are looking for power consumption reduction on the order of 50 percent or more. This will ease system thermal management and enable 100GbE port count scaling in the short term. For the longer term, however, system designers seek 100GbE transceiver roadmaps with significant reductions in both power dissipation and form factor size.

Next-generation 40GbE/100GbE technologies

Several promising technological advances in progress could be used by optical module designers to achieve their next-gen 40GbE/100GbE module design targets. These include:

:: laser array/planar lightwave circuitry (PLC) hybrid integrated TOSA

:: PD/TIA array/PLC hybrid integrated ROSA

:: low-power BiCMOS ICs and CMOS gearbox IC

:: Narrow 4x28G-VSR electrical interface, electrical connector, and 28G CDR IC

:: CFP2 electro-mechanical module development.

Hybrid integration of DFB discrete or array lasers with optical multiplexing PLCs has been investigated intensively across the industry. Some of the key challenges to using this technology in TOSA development include laser/PLC device alignment and optical coupling loss minimization. Use of a laser array is desired, as it minimizes the number of process steps in active alignment with the PLC device. However, a DFB laser array is particularly challenging to realize with sufficient gain across all channels for a wide temperature range. Nevertheless, four-channel devices appear to be feasible for realizing an optical hybrid integrated TOSA.

Similarly, hybrid integration of PIN-PD and TIA arrays with optical de-multiplexing PLC has also been investigated. Early progress using this type of ROSA was made with 10GBase-LX4 module designs, so it appears feasible to scale ROSA rates up to 4x10G and 4x28G. Challenges still remain in PD/PLC device passive alignment, control of attenuation and polarization mode dispersion losses, and temperature stability for realizing an optical hybrid integrated ROSA.

Significant reduction of 40GbE/100GbE optical transceiver power dissipation will come from improvements in the component ICs. For next-generation 40GBase-LR4, use of the non-retimed XLPPI electrical interface enables elimination of the quad CDR device, which results in more than 30 percent power consumption savings. Further process and design improvements in laser drivers and TIAs will assist with an overall module power consumption reduction of over 50 percent.

For next-gen 100GBase-LR4, EA-DFB driver IC process improvement and CMOS gearbox ICs will be a major factor in module power consumption reduction. With these factors, plus improvements in TIA and module DC-DC power conversion, 50 percent overall CFP module power dissipation looks feasible in the near term.

For the longer term, it is desirable to narrow the module electrical interface to four parallel lanes operating each at 28 Gbps. This would enable replacement of the gearbox IC with a quad 28G CDR IC, thus reducing electrical interface IC complexity and power consumption. Work is underway in the Optical Internetworking Forum (OIF) to specify a host chip to optical module electrical interface, called CEI 28G-VSR. Electrical connector suppliers, physical layer IC suppliers, host system vendors, and module vendors are working together in the OIF to confirm application requirements and specify a 28G channel model and electrical interface characteristics.

Even with all of these developments, it still appears that 100GBase-LR4 power dissipation will be on the order of 10 W, which is still too high to fit into the existing QSFP+ form factor power envelope. To reduce module form factor size, consideration of a next-generation CFP module, “CFP2,” is underway that would be compactly sized for sub-10-W power dissipation and support a narrow 4x28G electrical interface.

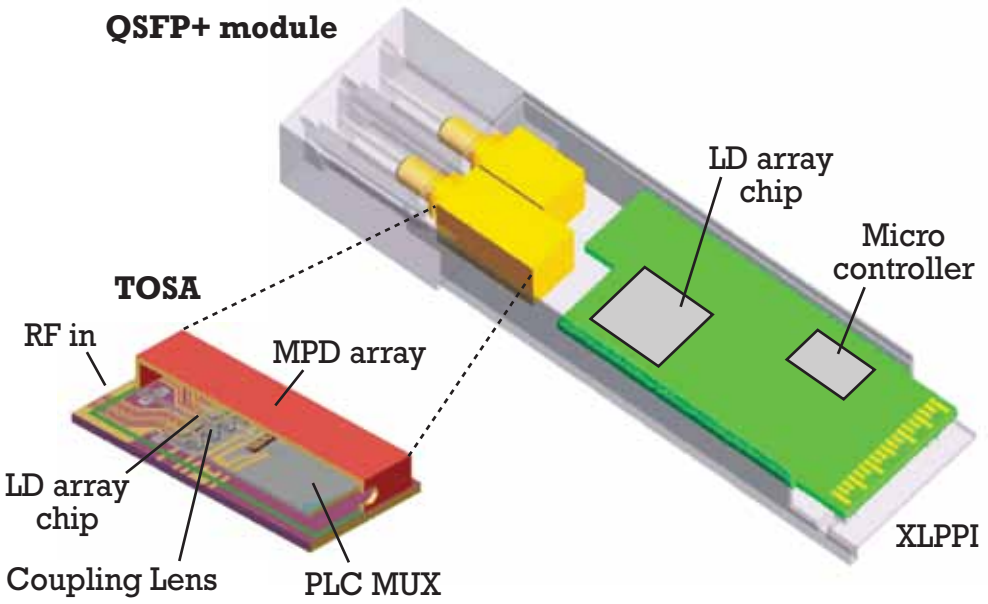


Figure 3. *Next-generation 40GBase-LR4 optical transceiver design.*

With the above-noted technology advances, the next-generation 40GBase-LR4 QSFP+ conceptually could be realized as illustrated in Figure 3. A future 100GBase-LR4 CFP2 would look similar architecturally, with operation at 4 x28G and inclusion of a quad 28G CDR electrical interface. The CFP2 module dimension specifications are an open point of study at this time, but past design experience suggests the CFP2 may look mechanically similar to the existing X2 form factor.

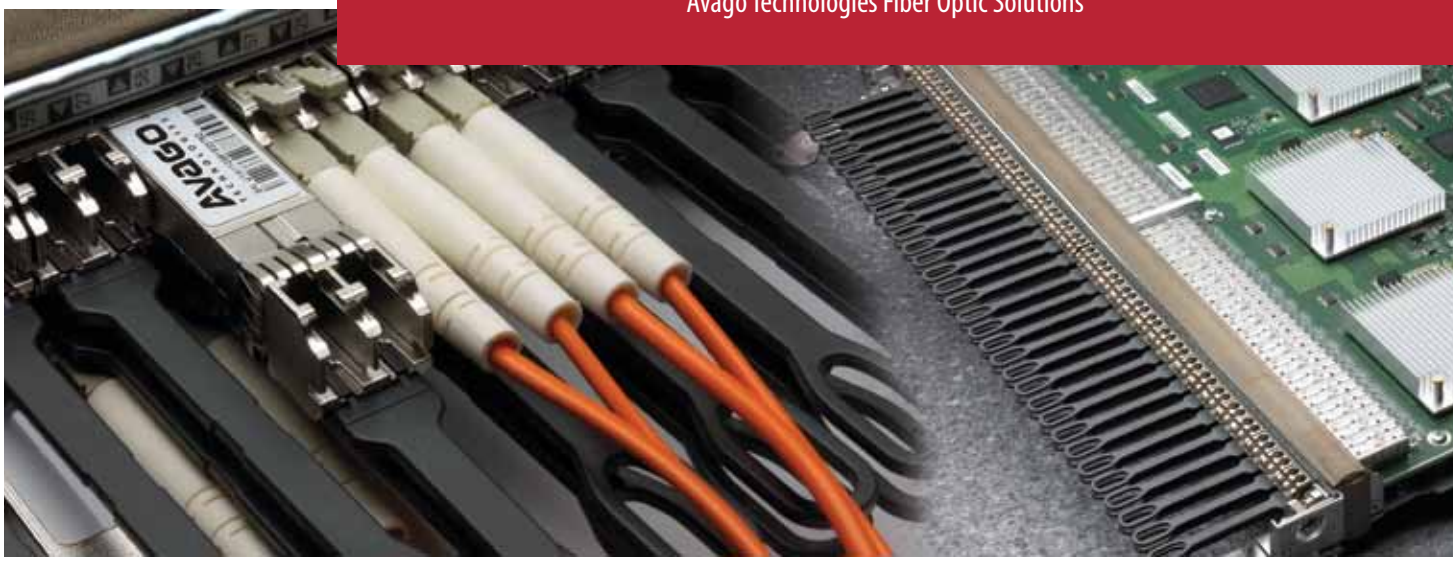
Future trends

First-generation 40GbE/100GbE CFP optical transceivers are now completing customer qualification and shipping in production. Key design targets for next-generation optical transceivers are: significant reduction of module power dissipation and form factor size.

Critical technologies for tackling these design targets include 4x10G and 4x28G hybrid integrated TOSAs/ROSAs and process improvements in 28G gearbox and CDR ICs. There also may be consideration of uncooled CWDM 28G laser technology for realizing 100GbE optical transceivers in a QSFP+-like form factor for short singlemode fiber (<2 km) applications.

JON ANDERSON, PH.D., is director of technology programs at Opnext.

Avago Technologies Fiber Optic Solutions



Mini SFP+ increases system port density using existing cable infrastructure

- Smaller form factor enables 64 ports of 10Gb in a 19" 1RU blade compared to standard SFP+ that can only support 48 ports
- Compatible with standard LC optical cable installed base (cables locked to 5.25mm instead of 6.25mm)
- Re-use standard SFP+ electrical connector and pin-out
- New "push/pull tab" de-latch mechanism to allow extraction in high density cabled applications

To learn more or to request a sample please go to:

www.avagoresponsecenter.com/27

AVAGO TECHNOLOGIES

Your Imagination, Our Innovation
Sense · Illuminate · Connect

Is 100G a DIY proposition?

By **MEGHAN FULLER HANNA**

SYSTEMS VENDORS TRADITIONALLY have relied on MSA-based modules to help reduce costs. Now, control of the complex electronics required for coherent detection at 40G, 100G, and beyond may make a "do it yourself" approach more appealing.

In July 2010, Cisco acquired optical module and electronics manufacturer CoreOptics, signaling a sea change in the relationship between systems houses and transponder vendors, say some analysts. In the past, equipment vendors relied on MSA-based modules to help reduce their costs, making Cisco's move a bit puzzling, particularly since it led the MSA charge at 10G. But CoreOptics' expertise is really in the digital signal processing (DSP) required for coherent detection at 40/100G.

Complex electronics—including DSP, high-speed, analog-to-digital conversion (ADC), and forward error correction (FEC)—are the gating factor for coherent approaches to 40/100G. Several systems vendors, Cisco included, have seemingly decided it's in their best interest to retain this functionality in house. If this is true, where does that leave the module suppliers?

Systems houses control DSP, but for how long?

With the acquisition of CoreOptics, Cisco can add its name to a short but growing list of systems houses who have already introduced internally developed 100G coherent systems or announced their intention to do so. The vendors include Ciena/Nortel, Alcatel-Lucent, Huawei, and Infinera.

Those systems houses will bypass the optical module vendors and the IC suppliers and use their own expertise for coherent detection. In a research note entitled "Optical system and module vendors rush to grab 100G expertise," author Karen Liu, vice president of components at Ovum, poses two key questions: "Who is best able to take ownership of the DSP? And for whom is it most critical to take ownership?"

Nortel, now part of Ciena, installed the first commercially available 100G system for Verizon in 2009. It used homegrown electronics—and marked the start of a trend in systems design. (Photo courtesy of Ciena)

The answers to these questions are different, she says. Ovum believes the systems houses are best able to take ownership of the DSP, but there is more incentive for the optical module vendors to do so.

As Liu sees it, the systems houses already control the complex electronics required for coherent detection; the more salient question is how long will they continue to do so? Historically, OEMs have developed critical technology in house because they wanted to make sure they get it right or because the technology was not yet available from a third party. Then, to reduce the cost of future generations of equipment, they relinquished control of that technology to the merchant market.

"In my mind," says Liu, "it's not a question of whether they will take ownership; I think they already have, by and large. The question is, is that a permanent move or is it a temporary time-to-market [move]?"

For Niall Robinson, vice president of marketing at Mintera Corp., the situation unfolding in the coherent market seems familiar. In the mid to late 1990s, when 10G was still in its infancy, "we had this new thing come into terrestrial transport that the submarine market had been using for a while," he recalls. "It was called forward error correction [FEC]."

Nortel led that charge as well, developing a "soft FEC" that leveraged unused overhead in the SONET frame to add FEC capability. Though it provided only a few decibels of optical signal-to-noise ratio (OSNR) improvement initially, the industry recognized how important FEC could be for terrestrial applications.

"And what did everyone do?" posits Robinson. "They did exactly what is playing

out again with coherent. They went away and developed their own FEC because there were no chips to buy on the market. So all the products to hit the market, day one, were home-grown, proprietary solutions," he explains. "But ask any systems house today, 'Do you make your own 10G stuff?' Everybody buys it externally."

Alcatel-Lucent added 100G to its 1830 PSS. (Photo courtesy of Ciena)

Robinson expects the same scenario to play out with coherent detection. The systems houses will continue to develop complex electronics in house, probably for the next three or four years, he says. But eventually, the technology will become "fairly commonplace" and offerings will begin to appear from third-party vendors. "And people will start shifting over to those modules because they'll be lower cost than building their own cards with very little performance differentiation."

One indication that Robinson may be right: ClariPhy Communications' cofounder and CEO Paul Voois, in speaking with *Lightwave*'s editorial director Stephen Hardy about his company's latest round of funding, said it will be used to develop 40/100G DSP ICs. In addition to Oclaro, which contributed $7.5 million, other investors included "multiple telecom OEMs." While Voois declined to name names, he did reveal that those investors are, in fact, systems houses.

Alain Couder, CEO of Oclaro, agrees that the DSP capability will not remain the sole domain of the systems houses. "The number of companies that can afford to do their own transponder will be going down," he said in a conversation with Lightwave shortly after the partnership with ClariPhy was announced. "And you see that right now. Most of the Tier 2 companies are buying modules, while the Tier 1 companies are buying modules at 10G — not all of them, but some at 10G — but most of them do their own 40G and 100G. All of that is going to evolve. I think we'll see less and less people doing their own module — if [optical module manufacturers can deliver] the right module with the right performance."

Liu, by contrast, is not entirely convinced this scenario will play out. Thanks

in part to the Optical Internetworking Forum's efforts, there seems to be "remarkable clarity and agreement" about how to implement 100G, but "the kicker in my mind is that it's not a straightforward turn of the crank going forward from 100G," she says. "Discussions about 400G and a terabit are technically all over the map." Because of this technical uncertainty in the road ahead, systems houses may want to retain ownership of the electronics, she says.

"Let's say 100G matures," Liu continues. "They wouldn't turn that [expertise] over because they are actually going to need that effort to help them with 400[G and beyond]. That would be the thing that changes this from a purely economic [decision]; 'Okay, I've got this design working. Now I can send it out.' It's the fact that [the DSP expertise] has a tie-in to the whole roadmap and the need to have that in house."

Liu wonders whether this is a lesson the industry learned from Nortel. "When Nortel came out with [coherent detection at 40G], everyone suddenly realized that they had kept a capability, and suddenly they could do it. I think that's potentially a wake-up call. You can't wait for vendors to come out with this stuff; you have to control it."

Optical module vendors optimistic

For their part, Oclaro and Mintera remain optimistic about the potential market opportunity for them and other module vendors. They cite the high cost of ASIC development as a key reason why systems houses will eventually turn to the merchant market.

"Honestly, I'm not sure there's a real positive business case on developing your own ASIC," Robinson notes. "If these things take $15 [million] to $20 million at best to develop, you'd have to sell an awful lot of systems to really start to recoup that. And I think some of the ASICs that are coming to market now have taken more than one go at getting the ASIC right, so my guess is that their development costs are significantly higher than that."

Oclaro agrees that developing high-speed electronics in house may prove cost-prohibitive for systems houses. "We think that over time, the price of developing very high-speed electronics will increase," Couder asserted in the conversation mentioned previously. He argued that "a chipset that has been designed to be

pre-tested, pre-designed with optical components [like the one Oclaro is currently developing in partnership with ClariPhy] can save time to market for some of the people who currently are doing their own chips."

Still, Liu believes the optical module vendors have a tough road ahead. In her research note, she maintains that "their product must address the optical link requirements of multiple OEM designs, which themselves must be tailored for different carrier networks. In addition, they face harsher density and thermal requirements than OEMs designing linecards from discrete components."

But where one person sees challenges, another sees opportunity. Vladimir Kozlov, founder and chief analyst at LightCounting, shares the optical module vendors' glass-half-full perspective, arguing that systems houses may become less reliant on standard modules now.

"If system vendors decide on different approaches to 100G technologies and they design their DSP for a specific protocol, then they may require a specific kind of transponder to go with it," Kovlov muses. "So instead of shipping a standard module to all the customers, transceiver vendors would have to make custom-made devices, which is actually going to be very good news for transceiver vendors because they will be able to charge much more for custom-made modules than standard ones."

Regardless of which suppliers end up controlling the high-speed electronics market, Kozlov believes Cisco's acquisition of CoreOptics is a positive development for the industry overall, as it signals Cisco's commitment to the 100G market.

"If you look back 10 years," he says, "40G was supposed to be right around the corner. In my opinion, it is still just emerging from around the corner. So the question is, are we going to have to wait another 10 years for 100G to come, or is it going to happen sooner than that? With Cisco's [acquisition of CoreOptics], I think there's a good chance we won't have to wait another 10 years for this to really happen."

MEGHAN FULLER HANNA is a senior editor at Lightwave.

Pluggable 40G and 100G Parallel Optic Transceivers for Next Generation HPC and Data Center Networks

Short Reach Solutions for High Density Pluggable Interconnects (QSFP+ and CXP)

- >120Gbps aggregate bandwidth
- 850nm VCSEL array design
- Proven performance and reliability

Best in Class Performance

- Excellent Signal Integrity
- Comprehensive Digital Diagnostic Monitoring

Flexible Cable Management with Removable Cables

- Enables solutions for patch panel installations
- Supports 100m on OM3 cable, 150m on OM4 cable
- Ease of installation for cable lengths longer than 30 meters
- Fully MSA and Standards Compliant

Enabling Next Generation Applications

- 40Gb Ethernet (802.3ba)
- Next Generation Infiniband (QDR/FDR)
- 16x Fibre Channel (14Gbps) links

To learn more or to request a sample please go to:
www.avagoresponsecenter.com/27

AVAGO TECHNOLOGIES

Your Imagination, Our Innovation
Sense · Illuminate · Connect

Carriers drive serial 40-GbE standard for optical compatibility

The IEEE 802.3bg Task Force is defining a serial 40-Gigabit Ethernet technology that will be optically compatible with the carriers' installed base of 40-Gbps client interfaces and enable the development of a common tri-rate module.

By **MEGHAN FULLER HANNA**

THE DEVELOPMENT OF an Ethernet standard at 40 Gbps was somewhat controversial when it was originally added to the IEEE 802.3ba Ethernet standards activities. Conventional wisdom stated that its application would be limited to rack-to-rack and shelf-to-shelf applications in data centers, and the carrier community, bullish on 100-Gigabit Ethernet, was reluctant to embrace the slower speed.

Fast forward several years, and the carriers are looking at 40-Gigabit Ethernet (40GbE) with

Carriers would like to see serial 40-Gigabit Ethernet capabilities added to existing 40-Gbps 300-pin modules such as this one. (Photo courtesy of Opnext)

renewed interest as an interconnection between their transport networks or as an uplink interconnection from enterprise networks, data centers, and other network operators into their transport networks. In fact, the new, swiftly moving IEEE 802.3bg 40-Gbps Ethernet singlemode fiber PMD task force is hard at work defining a serial 40GbE standard tailored specifically for the carriers' needs.

"We're seeing [40GbE] as sort of coming after 100 Gig, more as a backfill," reports Mark Nowell, senior director of engineering at Cisco and chair of the IEEE 802.3bg task force. "The data centers and the enterprises are starting to look for uplinks into [carrier] networks. The carriers are very aware of 40-Gigabit Ethernet, and they want to be prepared for it when it comes. But they have also been very consistent over the last number of years that their main interest and focus has been 100 Gig."

Nevertheless, the members of the IEEE 802.3bg task force are committed to a serial 40GbE implementation and working diligently to get the standard ratified in record time. Barring any major hurdles, the standard should be completed in June 2011, a mere 18 months after the original call for interest (CFI) in November 2009.

Driven by backwards compatibility

Among other objectives, the 802.3bg standard aims to:

:: Support a bit error rate (BER) better than or equal to 10^{-12}

:: Support 40GbE operation over at least 2 km on singlemode fiber

:: Enable optical compatibility with existing carrier 40-Gbps client interfaces.

It is this third objective—optical compatibility—that ultimately led to the formation of the IEEE 802.3bg project, which will be published as an amendment to the 802.3ba Ethernet standard.

Because 40GbE was originally perceived as a data center/enterprise technology, the 802.3ba Ethernet standard reflects the particular needs of that market, which is and always has been very cost sensitive. Though the group ultimately settled on a 4x10-Gbps CWDM implementation (known as 40GBase-LR4) that leverages lower-cost 10-Gbps components, a contingent of members, Cisco among them, pushed for a serial implementation as well. Their rationale was simple:

For 40-Gigabit Ethernet to penetrate carrier transport networks, it must easily interface with their existing 40-Gbps SONET/SDH and OTN infrastructures.

"There was a lot of passionate discussion around price curves and how a single laser is always going to be cheaper than four lasers, and while CWDM may be cheaper, there could be some crossover point in the future," Nowell recalls. "But what happened in the end was the [802.3ba] 40GbE group chose to go with CWDM as their main proposal, and that's what came to be."

As it is currently defined in the IEEE 802.3ba Ethernet standard, 40GBase-LR4 is optically incompatible with the carriers' installed base of 40-Gbps client interfaces, which includes the OC-768/STM-256, 40-Gbps packet-over-SONET (POS), and Optical Transport Network OTU3 protocols.

To minimize costs, carriers have deployed dual-rate or dual-protocol OTU3/OC-768 modules that are software selectable, enabling them to use one module for multiple applications. Both OTU3 and OC-768 are based on a 40-Gbps NRZ modulation format, and their optical specifications are nearly equivalent. Now carriers would like to add 40GbE to the mix. But the CWDM variance defined by the 802.3ba standard prevents the development of a so-called common tri-rate module.

In a presentation entitled "Economic Feasibility," delivered in January 2010 during the group's interim meeting, representatives from AT&T, Verizon, NTT, et al., noted, "Carrier opex is proportional to the number of module types that have to be supported." The development of a common tri-rate module would reduce sparring and simplify inventory management, thus lowering overall operating expenses.

Existing modules for 40G are housed in the 300-pin form factor and support the ITU specifications for SONET/SDH and OTN. Explains Jon Anderson, director of technology programs at Opnext, "For an Ethernet rate, which is 41.25 Gig, basically, the idea is to have that rate supported in the existing 300-pin 40G transceiver and have it be [software] selectable so the operator can select the [SONET/]SDH rate, the OTN rate, or the Ethernet rate using the same piece of hardware."

The tri-rate module, then, would support SONET/SDH (OC-768/STM-256) with a bit rate of 39.81 Gbps as defined by the ITU; OTU3 with a bit rate of 43 Gbps as defined by the ITU; and 40GbE with a bit rate of 41.256 as defined by the IEEE 802.3ba standard.

Anderson believes the first tri-rate modules should appear in late 2011/early 2012. The hardware effectively exists today, but testing the devices remains an obstacle. "It's a question of when the specs will be completely solid and made available on test equipment, and then made available to module vendors like ourselves to use for testing the devices," he says. "I suspect that will probably occur sometime mid-next year."

Completing the standard

By all accounts, the 802.3bg task force is moving swiftly to completion, in part because they are able to leverage work already done by the ITU and the 802.3ba task force. "We delayed the start of this project waiting for the 802.3ba group to finish what it was doing," says Nowell, "and that gave us a period of time to do a lot of incubation work, to get the proposals hashed out."

Of course, that doesn't mean all decisions have been unanimous. In a task force meeting held May 26–27, 2010, in Geneva, the group resolved the key technical challenge of selecting a wavelength operating range for the transmitter (1550 nm), but not without a fair amount of debate.

According to Anderson, who was also a presenter at the task force meeting, the discussion often seemed to fall along geographic lines, with several North American representatives favoring the 1550-nm operating range, while several Asian representatives lobbied for a 1310-nm option.

The key advantage of using the 1310-nm wavelength, says Anderson, is reach. "1550 [nm] at 43 Gig is dispersion-limited to 2 km," he reports. "At 1310 [nm], it can fan out to a 10-km reach with effectively no cost difference between the two."

Meanwhile, the key advantage of using the 1550-nm wavelength is one of backwards compatibility; transceivers based on the 1550-nm wavelength would be backwards compatible with legacy transceivers, while the working group

found the 1310-nm variant to be incompatible with an estimated 25% of legacy transceivers.

In the end, the group decided backwards compatibility was its most important criteria. But, says Anderson, "It was an exciting time. Some of the Asian suppliers and some of the Asian carriers—China Telecom in particular—who were looking for a 1310-nm solution for their networks may still have some interest in pursuing a standard in this direction."

So it seems the conventional wisdom—that 40GbE technology would be limited to computing and network storage applications—will not hold after all. With the ratification of the IEEE 802.3bg amendment, suppliers will be able to develop a common tri-rate module that should drive 40GbE penetration in carrier transport networks.

"As with anything," says John d'Ambrosia, director of Ethernet-based standards at Force10 Networks and chair of the IEEE 802.3ba Ethernet task force, "the market will decide where it's going to use [40GbE] and where it makes sense to use it. Do I believe that 40 Gig is going to have a home in servers? Yes. Do I believe that 40 Gig will have a home in networking? Yes."

MEGHAN FULLER HANNA is a senior editor at Lightwave.

SFP+ optical transceivers require engineering tradeoffs

The lower cost of SFP+ modules comes at the cost of increased engineering complexity for systems designers.

By **RUDY PRATER** and **ROB HANNAH**

TRADITIONALLY, 10-GBPS TRANSCEIVER implementations have used two approaches to ease host board design constraints. In the first, the data is retimed to re-establish a clean clock in the data stream with minimal jitter. In the second, the data stream is converted from a serial data path at the full bit rate to multiple data paths with lower clock rates. These techniques, however, increase power dissipation and system complexity. Traditionally for 10-Gbps transceivers, this extra complexity was built into the module, increasing unit costs significantly.

The SFP+ platform modifies the allocation of these elements between the transceiver and the host to minimize overall costs. However, this architecture change puts a greater burden on the host system designer, who must design the electrical transmission channels with greater care and discipline than has been necessary for systems operating at 4 Gbps and below.

CDRs and SerDes

The major change between SFP+ applications and those using SFP modules is the increased difficulty in designing the systems' electrical signal path. Even when using the enhanced grade of FR4 PCB material, the rate of signal degradation during propagation down even well-designed differential pair traces is significant.

As with previous evolutions of transceiver form factors, the decision about how and where to implement the necessary protocol functional blocks has been changed for SFP+.

The additional functions that are frequently used at data rates greater than 4.25 Gbps between the optical channel and the basic functions accomplished by the host system (i.e., the retiming and the conversion to lower-rate parallel bit streams) are accomplished through the clock and data recovery (CDR) and through serializer/deserializer (SerDes) functional blocks (see Figure 1).

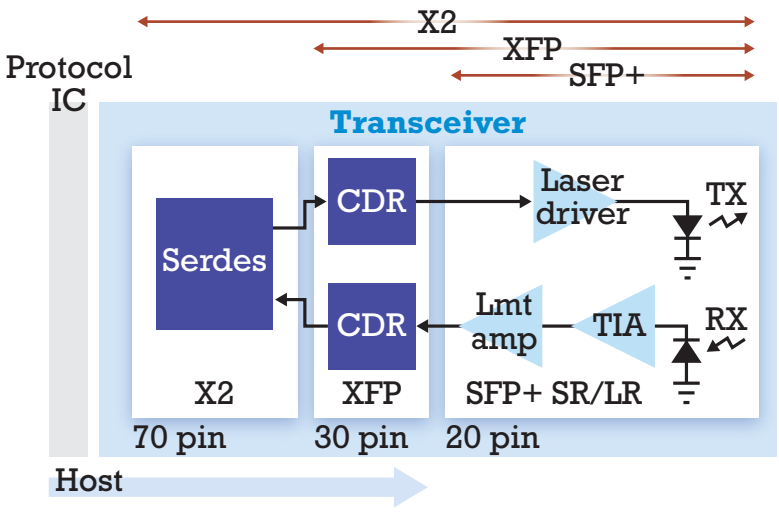


Figure 1. *System function allocations with different transceiver package sizes.*

The first generation of 10 Gbps transceivers included these functions inside the transceiver, as indicated in Figure 1. Unfortunately, this implementation, while technically feasible, required that the transceivers have:

1. a wide electrical connector to accommodate the parallel bit streams (limiting the achievable port/count density)
2. relatively high power dissipation (creating thermal challenges)
3. perhaps most importantly, relatively high cost (which limited the adoption of 10G interfaces).

The second generation of 10-Gbps transceivers moved the SerDes function outside the module onto the host board. Here, it could be integrated with other system functions more efficiently, decreasing the electrical power required by the transceiver and enabling a smaller connector interface and overall physical footprint. The internal CDR ensured that the high-speed electrical signal at 10 Gbps was retimed inside the transceiver, eliminating some of the accumulated signal degradation before it reached the host board.

The SFP+ platform progresses to the next level, moving the CDR function to the host, as well. This change enables a significant reduction in the size of

the transceiver electrical connector/housing, as well as a further reduction in power dissipation. The attendant overall cost reduction also has promoted SFP+ adoption.

Yet, as mentioned previously, these architecture changes mean a tradeoff between leveraging the benefits of the SFP+ and the considerably greater level of care required in designing the host board signal paths.

LRM applications and the SFP+ linear interface

Designs for 10GBase-LRM applications face an additional twist. To support 300-m reach over legacy multimode fiber, 10GBase-LRM calls for the use of electronic dispersion compensation (EDC). The EDC function generally is realized in PHY integrated circuits that reside on the host board. Since the EDC has been engineered to correct signal degradation with very specific characteristics that are expected from the propagation of the optical signal in the fiber, the EDC circuit needs to receive an accurate representation of the optical signal incident on the receive detector.

Specifically, the amplification of the optical signal representation from the photodetector must be very linear so that a faithful representation of the received signal is processed by the digital filter. While some additional degradation from the transmission of the electrical signal through the RF traces and the module connector will also be corrected, the lower the amount of additional degradation present, the more easily the filter can adapt to correct the optical path degradation.

The end result is: the better the design of the signal path in the transceiver—but especially the design of the generally longer traces on the host board—the better the performance of the entire link. While this statement is true for SR and LR links as well, the quality of the PCB RF design is especially important for LRM links.

New challenges in testing

Systems implementing SFP+ at 10-Gbps data rates are vulnerable to poor link performance that results when inadequate care is taken in the design of the electrical channel on the host PCB.

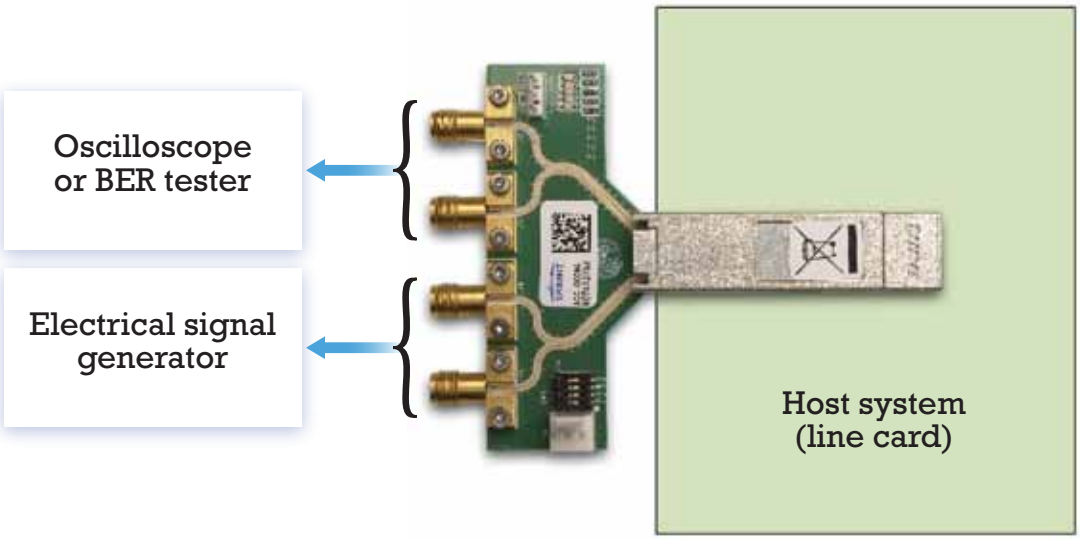


Figure 2. *Insert a host compliance board into a host board to provide access to test equipment for standard measurements.*

The SFP+ MSA identifies limits for 10-Gbps trace lengths at 200 mm (8 inches) using enhanced-grade FR4, and 150 mm (6 inches) using standard-grade FR4. While these limits could be longer in 8.5-Gbps designs, pushing this limit too far can increase project costs through added debug time and PCB changes, if performance does not meet expectations.

Whereas trace length is an easily quantifiable metric, many other geometric factors also influence signal integrity. It is a significant challenge to keep the paths between the SFP+ connectors and the PHY IC within spec for all ports in a high-density host design, as well as to find the optimum geometry for each port.

The interface between the transceiver and the host system needs more careful definition at these data rates and, when a design is not working, it is important to be able to isolate clearly which part of the system is causing a particular issue.

The primary method for quantifying the host channel design performance is the measurement of the system's S-parameters. To this end, the MSA has defined some standard test conditions and test fixtures (see Figure 2) so that a meaningful, well-defined test can reasonably compare one system to another, not to mention assist in isolating the weakest element in the system when performance is sub-par.

S-parameters are fundamentally a measure of the effective impedance of a circuit as a function of frequency. The MSA, therefore, specifies a maximum value as a function of frequency. The measurements can seem abstract and difficult to those unaccustomed to them, but they are actually straight forward with the appropriate equipment, as long as proper care is taken to calibrate the measurement according to the instructions of the equipment manufacturer.

Figure 3 shows the results of such measurements. Note that the S-parameters specified are differential S-parameters, characteristics of the energy transmitted along the differential pair of traces on the PCB. The differential S-parameters can be constructed from measurements of the single-ended S-parameters made on each trace separately.

The other major change in testing SFP+ systems arises from the need to use more advanced signal processing on the electrical signals in the host to compensate for propagation degradation while traversing the system's electrical channel.

For example, on the optical transmit side, the host system component may employ pre-emphasis when launching energy into the electrical channel that feeds the optical transmitter in the module. On the receive side of the transceiver, compensation is useful as well. Some transceivers have an optional pre-emphasis built in to the receive channel output, but many do not. For these transceivers, post compensation or equalization can be applied by the next stage in the system, the PHY integrated circuit that performs the SerDes function as well as protocol

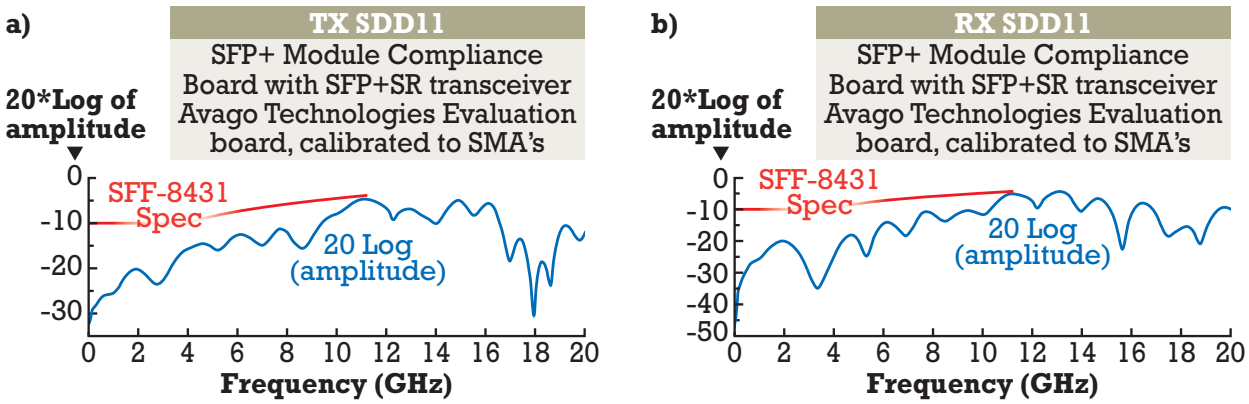


Figure 3. *(a) Comparison of typical S-parameter measurements with standard requirements for the transmit path; (b) comparison of typical S-parameter measurements with standard requirements for the receive path.*

coding. Equalization compensates for the static characteristics of the copper traces on the PCB and the effects of the SFP+ connector. Equalization uses a subset of the same technology as the EDC required for the LRM standards. The 10G systems, thus, have the additional design issue of switching the operation of the PHY circuit, depending on whether an LRM module has been inserted as opposed to an SR or LR transceiver.

The primary disadvantage of this new reliance on signal processing techniques, beyond increased system complexity, is the impact on testing and debugging the system. The extreme case of this effect is that the LRM transceiver has specifications to define performance when operated as a standalone device that can seem abstract and unintuitive. The standard quality benchmark for modules—receiver sensitivity at 8 Gbps and below and stressed receiver sensitivity for 10GBase modules—holds no meaning for a standalone SFP+ LRM transceiver. This benchmark can only be used to characterize a link that includes the EDC function.

For a meaningful measure, three parts of the link are critical:

1. the transceiver
2. the PHY IC with EDC
3. the exact layout and configuration of the copper traces in the transceiver but, more importantly, on the host PCB.

While the SFP+ LRM receiver offers an extreme example of the effects of electrical signal processing on testing and debugging, the same issues pertain to the receivers in SFP+ SR and SFP+ LR modules, if less severely.

More troubling can be the influence of the application of preemphasis to the electrical input of the transmitter. Qualitative judgments are often made about transmit performance, and these can be strongly influenced by the pre-emphasis of the signal applied to the transceiver input pins. The shape of the eye can change significantly, depending on the signal characteristics at the SFP+ connector. Attempts to quantify these effects are necessary, using either mask margin or a variety of jitter measurements. These also will vary significantly with different pre-emphasis settings, if only because the shape of the eye affects the choices for the 1 level, 0 level, or the level on the edge where the jitter is

measured. These measurement points are often implicitly determined by the test equipment, but it is important that the test engineer be aware of such issues and use appropriate skepticism and care when drawing conclusions from the data.

From the system design perspective, choices have to be made that are often somewhat arbitrary. These include whether a single TX emphasis setting is sufficient for all ports and all modules or whether some variation in the emphasis settings should be made to depend on each particular port's layout geometry. Another possible refinement might be to vary the pre-emphasis settings according to either transceiver type, transceiver manufacturer, or both.

It is important to realize, however, that the correlation between small variations in mask margin or jitter and performance over a worst case link is not obvious. While it is clear that the extreme values for mask margin and jitter lead to poor link performance, the sensitivity of system BER to commonly seen tolerances for these values is not obvious and, unfortunately, not easy to test.

In summary, the effects of preemphasis and equalization are significant and complicate system testing, qualification, and system debugging for SFP+ links in ways not generally experienced with prior data rates. The interaction between the various components of the system, especially including the design of the RF copper connections, are quite likely to be the source of more performance issues than the performance of any isolated component. Thus, more care will be needed for disciplined, quantitative approaches to testing and debugging.

RUDY PRATER is a former senior applications engineer in the Fiber Optic Products Division of Avago Technologies. **ROB HANNAH** is applications engineering manager at Avago.

About Avago Technologies

Avago Technologies is a leading supplier of analog interface components for communications, industrial and consumer applications. The company's analog, mixed signal and optoelectronics products serve four diverse end markets: wireless communications, wired infrastructure, industrial and automotive electronics, and consumer and computing peripherals. Further information is available at www.avagotech.com.

LINKS:

- To learn more or to request a free sample, please go to: http://www.avagoresponsecenter.com/597
- Fiber Optics Product Information Pages on Avagotech.com http://www.avagotech.com/pages/fiber_optics/
- Flash Video: Fiber Optic's Optical Transceiver Flash Demo ("SFP Optical Transceivers With Bulk Optics") http://www.avagotech.com/pages/fiber_optic/
- White Paper: Fiber Optic Solutions for Base Station Applications http://www.avagotech.com/docs/AV02-2123EN