

100 GIGABIT ETHERNET AND BEYOND

JOHN D'AMBROSIA

Introduction

At the November 2009 IEEE 802 Plenary meeting, the IEEE P802.3ba 40 Gb/s and 100 Gb/s Ethernet project reached the next milestone on its journey to being ratified as a standard, as the draft amendment to the IEEE 802.3 standard was approved to move to the second and final stage of balloting, known as Sponsor Ballot. The success of meeting this milestone keeps the project on schedule for ratification in June 2010. The ratification of the P802.3ba standard will be a major accomplishment for the Task Force, the end result of hard work and diligent efforts, and a major milestone in the long running history of Ethernet as a protocol.

The IEEE P802.3ba standard introduces a family of physical layer (PHY) specifications for 100 Gigabit Ethernet (100GbE) defined in Table 1.

Below is a description of each of the different PHYs in the 100GbE family:

- The 100GBASE-CR10 PHY supports transmission of 100GbE over 7 m of twin axial copper cable across 10 differential pairs in each direction. The PHY leverages past work from the Backplane Ethernet project by utilizing the 10GBASE-KR architecture, channel budget, and physical medium-dependent sublayer (PMD).
- The 100GBASE-SR10 PHY is based on 850 nm multimode fiber (MMF) optical technology and supports transmission of 100GbE across 10 parallel fibers in each direction. The effective data rate per lane is 10 Gb/s. Optical Multimode 3 (OM3) grade fiber, which has an effective modal bandwidth of 2000 MHz/km, can support reaches up to at least 100 m, while Optical Multimode 4 (OM4) grade fiber, which has an effective modal bandwidth of 4700 MHz/km, can support reaches up to at least 125 m.
- The 100GBASE-LR4 PHY is based on dense wavelength-division multiplexing (DWDM) technology and supports transmission of at least 10 km over a pair of single-mode fibers (SMFs). The four center wavelengths are 1295 nm, 1300 nm, 1305 nm, and 1310 nm. The center frequencies are spaced at 800 GHz, and are members of the frequency grid for 100 GHz spacing and above defined in ITU-T G.694.1. The effective data rate per lambda is 25 Gb/s. Therefore, the 100GBASE-LR4 PMD supports transmission of 100GbE over four wavelengths on a single SMF in each direction.
- The 100GBASE-ER4 PHY is also based on DWDM technology and supports transmission of at least 40 km over a pair of single-mode fibers. The four center wavelengths are 1295 nm, 1300 nm, 1305 nm, and 1310 nm. The center frequencies are spaced at 800 GHz, and are members of the frequency grid for 100 GHz spacing and above defined in ITU-T G.694.1. The effective data rate per lambda is 25 Gbit/s. Therefore, the 100GBASE-LR4 PMD supports transmission of 100GbE over four wavelengths on a single SMF in each direction. To achieve the 40 km reaches called for, it is anticipated that implementations may need to include semiconductor optical amplifier (SOA) technology.

Additionally, optional retimed and non-retimed electrical interfaces based on 10 lanes of 10 Gb/s in each direction are defined. The retimed interface, known as CAUI (100 Gigabit Attachment Unit Interface), is specified for chip-to-chip and chip-to-module applications. The optional non-retimed elec-

trical interface, CPPI (100 Gigabit Parallel Physical Interface), has been designed to support chip-to-module applications, and is optimized based on the 100GBASE-SR physical layer specification.

Reviewing the description of the PHY members of the 100GbE family, it can be observed that existing electrical and optical signaling technologies are leveraged heavily. The real innovation in the standard is the development of an architecture that is flexible and scalable to support both 40GbE and 100GbE, and the multiple PHYs being developed as part of the IEEE P802.3ba standard, as well as PHY specifications that may be developed by future task forces. The keys to this flexible architecture reside in the lane distribution scheme of the physical coding sublayer (PCS) and the multiplex/demultiplex functionality of the physical medium attachment (PMA) sublayer. Further insight into the IEEE P802.3ba architecture may be found in "40 Gigabit Ethernet and 100 Gigabit Ethernet: The Development of a Flexible Architecture," which was written by the author for the March 2009 IEEE Optical Communications Supplement to *IEEE Communications Magazine*.

As early adopters look to the initial deployment of systems supporting 100GbE to provide relief for congested networks, others are looking to the development of the next generation of electrical and optical signaling technologies that will enable reductions in 100GbE port cost and power, while simultaneously maximizing the usable port densities per system. These same technologies will also provide the building blocks for the next rate of Ethernet. As noted, the conception of the IEEE P802.3ba architecture was driven by the need to be flexible in order to support currently available and future electrical and optical signaling technologies, but it was also conceived as an architecture that could scale to support future speeds of Ethernet [1].

While the upcoming ratification of the IEEE P802.3ba standard will be a major accomplishment, it is anticipated that the 100GbE family will undergo further development. The development of 100GbE was justified based on the observation that the bandwidth requirements of core networking applications are doubling every 18 months. This observation correlates to core networking applications needing Terabit Ethernet in 2015. Therefore, one must consider the future of electrical and optical signaling technologies for 100GbE in order to consider the next rate of Ethernet.

From an electrical signaling perspective, 100GbE leverages 10 Gb/s signaling for its chip-to-chip, chip-to-module, and 100GBASE-CR10 electrical specifications. From an optical signaling perspective, 100GBASE-SR10 leverages 10 Gb/s 850 nm technology, while 100GBASE-LR4 and 100GBASE-ER4 leverage 25 Gb/s WDM technology. Given the multilane approach employed electrically and optically, the most obvious direction in development activity would be to reduce the number of lanes or wavelengths by increasing the data transmission rate across a given lane.

Development Activity for Electrical Signaling

Recent communications between the IEEE 802.3 and the Optical Internetworking Forum (OIF) indicate an industry interest in the development of 25 Gb/s electrical signaling. The OIF, which has been working on two signaling specifica-

(Continued on page S8)

Port Type	Reach	Description
100GBASE-CR10	At least 7 m cu cable	10 × 10 Gb/s
100GBASE-SR10	At least 100 m OM3 MMF (at least 125 m OM4 MMF)	10 × 10 Gb/s (use of Parallel Fiber)
100GBASE-LR4	At least 10 km SMF	4 × 25 Gb/s
100GBASE-ER4	At least 40 km SMF	4 × 25 Gb/s

Table 1. 100GbE physical layer specifications.

(Continued from page S6)

tions related to 25 Gb/s signaling — CEI-25G-LR (intended for backplane applications) and CEI-28G-SR (intended for chip-to-chip and chip-to-module applications) — requested feedback from IEEE 802.3. In its response IEEE 802.3 noted that it was the opinion of participants within IEEE 802.3 that the priority for industry development is on a four-lane by 25 Gb/s electrical interface that targets chip-to-module applications.

The development of a 4 × 25 Gb/s electrical signaling scheme is most applicable to the 100GBASE-LR4 and 100GBASE-ER4 PHYs, as they are based on a 4 × 25 Gb/s per wavelength scheme. Figure 1 illustrates the architecture being implemented in first-generation modules. The electrical interface to the module utilizes CAUI, which is a 10 × 10 Gb/s in each direction. The use of this interface requires the use of a 10:4 serializer on the transmit side of the module and a 4:10 deserializer on the receive side of the module, which will yield higher power consumption due to their non-integer ratio [2]. Thus, the width of the CAUI interface and the higher power

consumption of the implementation will drive the size of a module footprint upward. For example, one currently proposed industry standard 100GbE optical module, the CFP, is specified at 144.75 mm deep by 82 mm wide [3]. The width of the CFP is over two times that of the XENPAK module, which was one of the earlier form factors for 10 Gigabit Ethernet (10GbE).

Figure 2 illustrates how the development of a retimed 25 Gb/s electrical interface could be applied to a next-generation module implementation of 100GBASE-LR4 [4]. Reducing the width and number of pins of the electrical interface, and reducing the power by eliminating the serial-izer/deserializer in the module could help enable

a smaller module size, potentially toward the width of the XENPAK. This smaller module would enable higher port counts, thus enabling integration to help drive per port costs down.

While it is hopeful that cooperation and collaboration between IEEE 802.3 and OIF will enable faster implementation of 25 Gb/s electrical signaling, one must also consider that 10 Gb/s electrical signaling underwent several years of refinement. Therefore, it should be expected that the fastest electrical signaling that will be available for the next speed of Ethernet will most likely be based on 25 Gb/s.

Development Activity for Optical Signaling

Optical signaling is being explored in a multitude of areas — from chip-to-chip and board-to-board interconnects to distances that will cover thousands of kilometers. There are three optical PHYs in the 100GbE family: 100GBASE-SR10, which supports up to 100 m over OM3 fiber; 100GBASE-LR4, which supports up to 10 km over SMF; and 100GBASE-

(Continued on page S10)

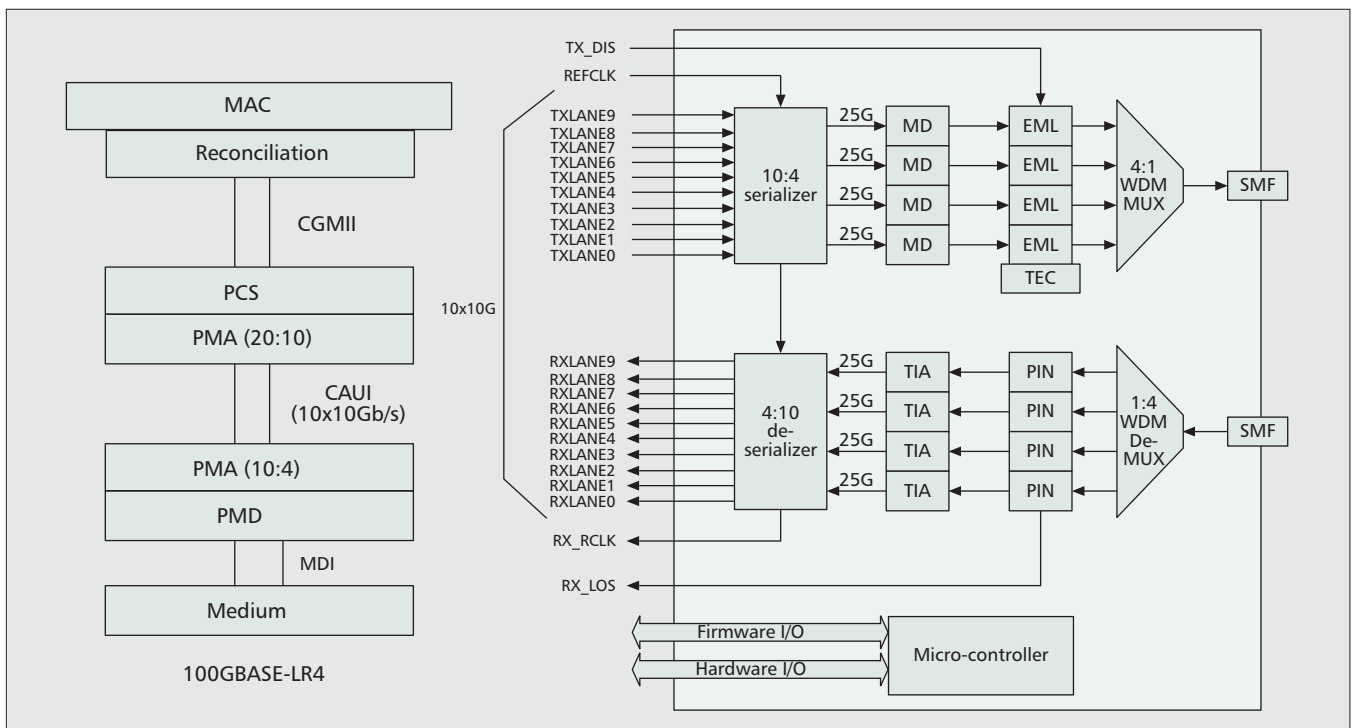


Figure 1. Generation 1 100GBASE-LR4 implementation concept.

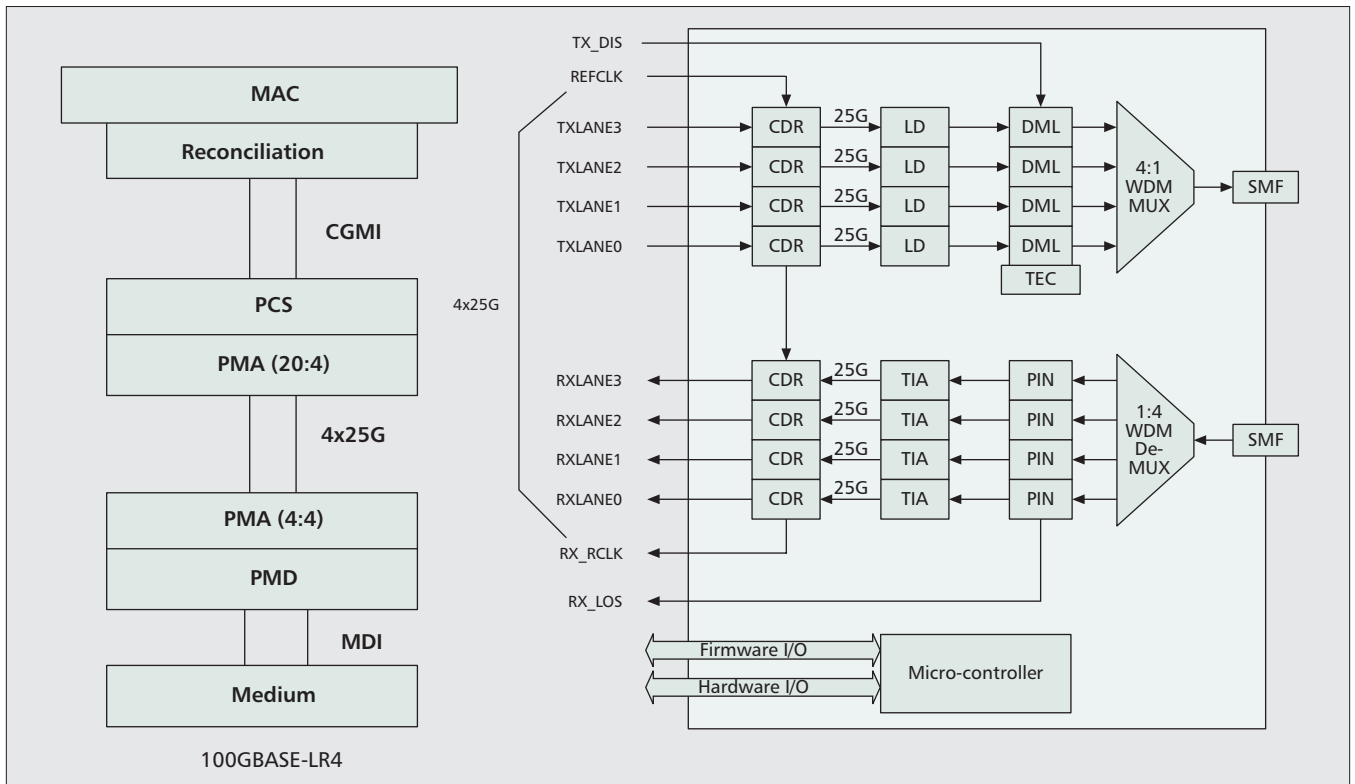


Figure 2. Generation 2 100GBASE-LR4 implementation concept.

(Continued from page S8)

ER4, which supports up to 40 km over SMF. While the selected PHYs for 100GbE were described earlier, it is worthwhile to consider other proposals that were considered for the respective PHYs, in order to assess the potential for future development efforts.

Potential development efforts for the 100GBASE-SR family were suggested during the proposal selection process of the

project. While the IEEE P802.3ba Task Force selected a solution based on parallel multimode fibers (MMFs), there were a number of discussions regarding the development of a solution that could be delivered across a duplex MMF solution. One proposal suggested using a WDM-based architecture, similar to the one adopted for 100GBASE-LR4 shown in Fig. 1, that utilized 4 wavelengths on a 7 nm channel spacing

(Continued on page S12)

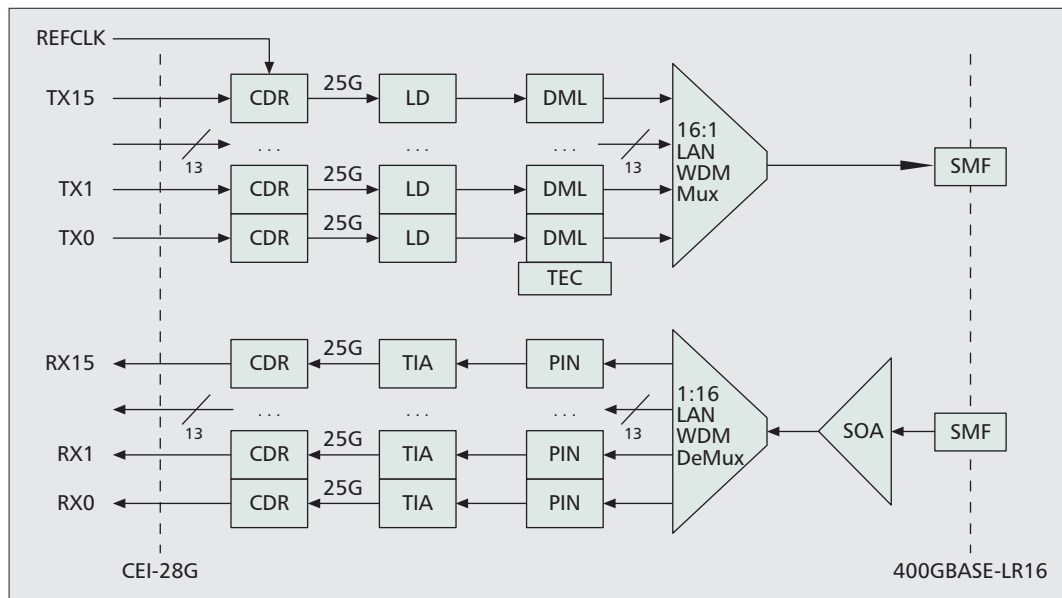


Figure 3. Potential concept for 400GbE.

(Continued from page S10)

around 850 nm at 25 Gb/s per wavelength [5]. Also, during the proposal selection process for the 40GBASE-SR PHY, one presentation suggested that with OM4 fiber 40 Gb/s serial rates could be achieved [6].

For 100GbE over SMF the Task Force quickly moved toward the adopted WDM approach, as no competing proposals were provided. It was realized, however, from early work during the Study Group phase of the project that work exploring alternate modulation schemes, such as differential quadrature phase shift keying (DQPSK), was underway, especially for long-haul transmission. Industry consensus on a single modulation scheme has not yet been achieved; therefore, it is unclear whether such an optical signaling technology would be ready to be utilized for the next speed of Ethernet. So, while a serial or single wavelength solution was not adopted, the knowledge of exploratory work in this space helped influence the desire to develop an architecture that would ultimately be able to utilize such a PMD solution when it became available.

At the time of writing this column, IEEE 802.3 has formed a new study group that will examine the need for the development of a 40 Gb/s Ethernet SMF is optimized for client applications in the carrier environment. Given that carrier deployment of 40G technology started in 2004 with modules deployed that could support either OTU3 or OC-768, which are both based on non-return-to-zero (NRZ) technology [7], it is anticipated that an Ethernet solution for this space would

need to coexist with these protocols, and hence will also be a serial solution.

At this time it appears that the 100GbE family of optical PHYs could grow, but it would seem that the obvious candidate would be a new PHY focused on delivering 100GbE across a duplex MMF solution using some sort of WDM solution based on 25 Gb/s per optical wavelength. At this time it is uncertain when a new project targeting adding a serial or single-wavelength PHY to the 100GbE family might occur.

Looking Ahead

The introduction of 100GbE will drive the development of the next generation of electrical and optical signaling technologies to drive the reduction in cost and power per 100GbE port and maximize the usable port densities per system.

As noted earlier, however, based on the observation that bandwidth requirements for core networking applications are doubling approximately every 18 months, core networking applications are forecasted to need Terabit Ethernet in 2015. Thus, discussions regarding the next speed of Ethernet have already begun.

The IEEE P802.3ba architecture has been developed to be a flexible and scalable architecture. The future development efforts in electrical and optical signaling will be the gating items for developing the next speed. The next generation of electrical signaling will target interfaces based on 25 Gb/s. For optical transmission over multi-mode fiber two possible paths exist — development of a 4 × 25 Gb/s per optical wavelength

The TechMate™ Series

Easy, Effective, Cost Efficient



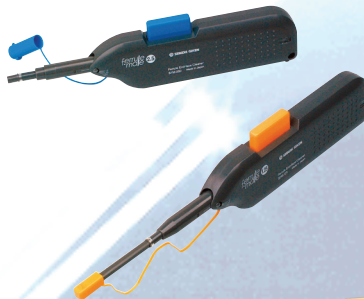
HandiMate -

with it's unique Slide-Guide application, offeres a low-cost, handheld alternative to bulky, expensive cartridge cleaners.



FerruleMate -

the only commercially available bulkhead connector cleaner approved by all major Telco operators for MDU, Pedestal and CO sites.



RepairMate -

provides all the tools necessary to repair and restore scratched connectors in the field without the high cost of replacement.



Please come visit us at Booth # 2939.

Tel: 770-279-6602 • Sales@sg-usa.com • www.SeikohGiken.com

PHY over duplex fiber for 100GbE or the development of a serial 40 Gb/s solution for 40GbE. For optical transmission over single-mode fiber the path forward is unclear — a single lambda solution using a modulation technique, such as DQPSK, is a possible answer, but the actual timing for standardization is unclear.

The debate regarding the next speed of Ethernet has also been influenced by the decision of the IEEE P802.3ba Task Force to break with the tradition of only doing 10x leaps in speed. This has fueled speculation that the next leap in speed for Ethernet will be 400 Gb/s, which could leverage electrical and optical technologies being developed for the next generation of 100GbE. For example, one proposal for 400GbE suggested a solution based on a 16-lane interface based on 25 Gb/s electrical signaling and 16 optical wavelengths [8]. By leveraging these technologies, the costs associated with developing such a solution could be minimized. Conversely, based on the potential technologies that will be developed over the next few years in support of 100GbE, it is not clear how Terabit Ethernet would even be accomplished. For example, the electrical interface by itself would be 40 lanes wide in each direction with each lane operating at 25 Gb/s.

Conclusions

During the course of the study group phase of the IEEE P802.3ba project, a number of individuals with data center backgrounds got up and expressed interest in 100GbE, based on the aggregation of Gigabit Ethernet servers in computer clusters. Meanwhile, network engineers in data centers look ahead to deployment of server clusters built on 10GbE-based servers, with cross-sectional bandwidths of fabrics reaching into multi-terabits of capacity. Similar needs have also been discussed by those in the carrier community and Internet exchanges. So while technical feasibility, which is one of the five criteria on which new projects are judged within the IEEE, is a key issue, another is broad market potential. As noted, the forecasted need in 2015 is for Terabit Ethernet. Therefore, while 400GbE is technically feasible, the question is will there be customer interest in it? This will be the key question that will be discussed within the industry in the months ahead. The next rates for networks will ultimately be discussed and decided within standards bodies such as the IEEE and ITU-T.

The discussion regarding the next rate of Ethernet, which will be decided by the next IEEE 802.3 Higher Speed Study Group, has begun!

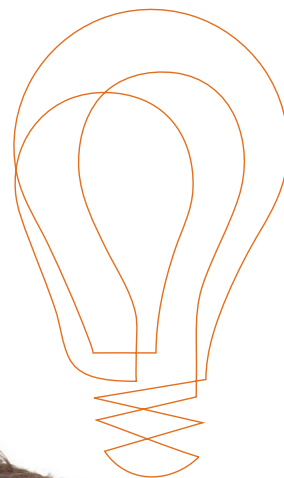
REFERENCES

- [1] J. D'Ambrosia, "40 Gigabit Ethernet and 100 Gigabit Ethernet: The Development of a Flexible Architecture," *IEEE Commun. Mag.*, Mar. 2009.
- [2] Cole et al., "100GBE Optical LAN Technologies," *IEEE Applications & Practice*, Dec. 2007.
- [3] CFP MSA Draft1.0, <http://www.cfp-msa.org/Documents/CFP-MSA-DRAFT-rev-1-0.pdf>, Mar. 23, 2009.
- [4] Cole et al., "100GBE Optical LAN Technologies," *IEEE Applications & Practice*, Dec. 2007.
- [5] Kropp et al., "40Gbit/s and 100Gbit/s Transmission over OM3 Duplex Fiber," *IEEE P802.3ba May 2008 Interim Meeting*, http://www.ieee802.org/3/ba/public/may08/kropp_01_0508.pdf
- [6] Kropp et al., "1x40 Gbit/s and 4x25 Gbit/s Transmission at 850nm on Multimode Fiber," *IEEE 802*, Mar. 2008 Plenary Meeting, http://www.ieee802.org/3/ba/public/mar08/kropp_01_0308.pdf
- [7] 40GE SMF PMD Call for Interest, *IEEE Nov 2009 802 Plenary*, http://www.ieee802.org/3/40GSMF/public/nov09/40GSMF_CFI_1109.zip
- [8] Cole, "Life Beyond IEEE P802.3ba," *Ethernet Alliance — Technology Exploration Forum*, Sept. 2009.
- [9] Lee, "Life Beyond 100G," *OIDA*, 18th Annual Forum, Dec. 2009.

Innovation Meets Brilliance

Gennum solutions bring innovation and integration to high speed mixed signal IC's.

Solve your next design's signal integrity issues with Gennum's high performance end-to-end solutions.



Visit us at OFC, Booth # 203
or online at www.gennum.com



 gennum