Technical Feasibility of New 200 Gb/s and 400 Gb/s Links for Data Centers

Y. Sun[(1)], A. Swartz[(1)], R. Lingle, Jr.[(1)], J. Kamino[(1)], R. Shubochkin[(2)], D. Braganza[(2)]

(1) OFS, 2000 Northeast Expressway, Norcross, GA 30071, USA

(2) OFS, 50 Hall Road, Sturbridge, MA 01566, USA

Email: ysun@ofsoptics.com

Abstract

The low cost and low power advantages of the VCSEL-MMF paradigm will continue through future generations of Datacom networks, at speeds of 200 Gb/s, 400 Gb/s and beyond, meeting the need for higher bandwidth in cloud and large enterprise data centers. Two baseline proposals for 400 Gb/s modules operating over multimode fiber– 400GBASE-SR8 and 400GBASE-SR4.2 – were adopted by the IEEE P802.3cm task force [1, 2, 3]. In this paper, we will review the literature [4] and provide new experimental evidence that demonstrated the technical feasibility of new 200 Gb/s and 400 Gb/s PHYs over fewer MMF fiber pairs.

Keywords: 400 Gb/s; Datacenter; Datacom; Multimode fiber; OM4; PAM4; Pulse Amplitude Modulation; Short Wavelength Division Multiplexing; SWDM; TIA; Wideband

1. Introduction

The market shows strong internet traffic growth generated by video streaming, social networking, cloud computing and Internet-of-Things (IoT). The VCSEL-MMF advantage of low cost and low power will continue through future generations of short reach links, at speeds of 200 Gb/s, 400 Gb/s and beyond. IEEE has defined 50, 100, 200, and 400 Gb/s links in the P802.3bs [5] and P802.3cd [6] Ethernet standards. The PMDs over multimode optical fiber (MMF) in those standards are based on 850 nm VCSEL operation. IEEE P802.3cd 50GBASE-SR, 100GBASE-SR2 and 200GBASE-SR4 use 50 Gb/s PAM4 modulation. 50 Gb/s PAM4 lanes support the traditional reach of 100 m reach over OM4 MMF. IEEE 802.3bs 400GBASE-SR16 uses 25 Gb/s Non-Return to Zero (NRZ) and requires 16 fiber pairs, which is not compatible with current cabling practices of either duplex fibers or 4-fiber pairs.

In November 2017, the IEEE 802.3 Ethernet Working Group authorized the formation of a study group to consider "Next-generation 200 Gb/s and 400 Gb/s PHYs over fewer MMF pairs than in existing Ethernet projects and standards". At the IEEE 802.3 Next-generation 200 Gb/s and 400 Gb/s MMF PHYs Study Group Interim meeting in Geneva [1] several proposals for new 200 G and 400 G PMDs were presented:

- 400GBASE-SR4.2 would use 4 fiber pairs, 2 wavelengths per fiber and 50 Gb/s PAM4 per wavelength
- 400GBASE-SR8 would use 8 fiber pairs, 1 wavelength and 50 Gb/s PAM4 per wavelength
- 400GBASE-SR1.8 would use 1 fiber pairs, 8 wavelengths per fiber and 50 Gb/s PAM4 per wavelength
- 400GBASE-SR4 would use 4 fiber pairs, 1 wavelength per fiber and 100 Gb/s PAM4 per wavelength
- 200GBASE-SR1.4 would use 1 fiber pair, 4 wavelengths per fiber and 50 Gb/s PAM4 per wavelength.

The IEEE P802.3cm 400 Gb/s over Multimode Fiber Task Force was created in May 2018. Baseline proposals for 400GBASE-SR8 and 400GBASE-SR4.2 were adopted in May and July [2,3], respectively, and draft documents are expected in the second half of 2018. 400GBASE-SR8 baseline proposal defines eight 26.5625 GBd parallel lanes, with KP4 FEC, which was defined in IEEE P802.3bs standard. It also reuses 850 nm sources defined for 200GBASE-SR4. Two connector types defined are a 24-fiber (two rows of 12) MPO-24 and single row MPO-16. Form factors are not specified by Ethernet, but would include QSFP-DD, COBO 8-lane and OSFP. The 400GBASE-SR4.2 baseline proposal is based on bi-directional (BiDi) two wavelengths WDM (850 and 910 nm), 26.5625 GBd PAM4 and KP4 FEC. It will allow break out to a 100G-PAM4-BiDi module. A 200 Gb/s PMD will not be standardized at this time, but may be offered as a MSA or proprietary product in the future, based on market demand.

In this paper, we will review past literature [4] and provide new experimental evidence that demonstrates the technical feasibility of new 200 and 400 Gb/s PHYs over fewer MMF fiber pairs.

2. Technical Feasibility of New 200/400 G MMF Links from the Literature

Previous papers have documented the technical feasibility of 2, 4 and 8 wavelengths, each carrying 50 Gb/s, as well as a research demonstration of 100 Gb/s PAM4 over OM3/4/5 [4]. Two/four wavelength transmission over MMF and 50 Gb/s PAM4 modulation on VCSELs at 850 nm are established technologies viable in commercial 100 Gb/s products, e.g. 100G-PAM4-BiDi and 100G-SWDM4, supporting 100 m link distances on OM4 and 150 m on OM5 MMF. Both 100G-SWDM4 over 300 m OM5 using a commercial switch, and 100G-PAM4-BiDi over 400 m OM5 were demonstrated at OFC and recaptured in references [7-8]. These commercial offerings serve as clear indications that SR4.2 and SR8 variants of 400 Gb/s modules are technically feasible, with a reach of 100 m over OM4 MMF.

Table 1. The literature demonstrating 50 Gb/s or 100 Gb/s PAM4 SWDM over OM3/4/5 (details in [4]).

Data Rate (Gbit/s)	Length @ 850 nm	Length @ 880 nm	Length @ 905 nm	Length @ 910 nm	Length @ 940 nm	Length @ 980 nm	Length @ 1060 nm	Fiber Type	Publication	Author
53.125	300m	300m	-	300m	300m	300m	-	OM5	OFC 2017	F. Chang *et al.*
50	200m/400m	-	200m/400m	-	-	-	-	OM4/5	IWCS 2017	Y. Sun *et al.*
53.125	300m	300m	-	300m	300m	300m	200m	OM5	JLT 2017	Y. Sun *et al.*
50	200m	-	-	-	-	200m	-	OM5	OFC 2016	J. Castro *et al.*
51.56	-	-	100m	-	-	-	-	OM4	ECOC 2014	S. Pavan *et al.*
51.56	100m/100m	-	-	-	-	-	200m/100m	OM3/4	JLT 2017	S. Pavan *et al.*
100	105m	105m	-	105m	105m	-	-	OM5	OFC 2017	J. Lavrencik *et al.*
100	100m	-	-	-	-	-	-	OM4	ECOC 2017	J. Lavrencik et al.

It is significant that the proposals for 400 Gb/s over 4-pair and 8-pair MMF cabling were authored or co-authored by major cloud companies in the US and China. Cloud data centers are increasingly important in the data center market. From those contributions [1], it is clear that VCSEL-MMF based links continue to be the low-cost choice for short reach applications. Historically VCSEL-MMF links have held a cost and power advantage for short-reach interconnects because 1) relaxed alignment tolerances of MMF (several microns vs. sub-micron) allow passive alignment in modules and offer better cost/loss trade-offs for connectors; 2) MMF connectors are more resilient to dirt, while cleaning SMF connectors is common issue in datacenters; 3) VCSELs need lower drive current than a DFB (5-10 mA vs. 50-60 mA), and 4) testing of VCSELs is done on-wafer. It is significant that adding wavelengths & PAM4 to MMF modules preserves the historical cost and power advantages over SMF modules. First, tolerances for mux/demux are significantly more relaxed in the case of MMF than SMF. Second, reduction of laser RIN for PAM4 is at least as, if not more, difficult for DFBs than for VCSELs. Also, packaging for VCSEL sources at 50 Gb/s PAM4 is based on known technology, whereas packaging for 1310 nm sources at 100 Gb/s per lane PAM4 has required significant development, which will result in yield and cost penalties – at least during the ramp up of 100 G/lane devices.

200GBASE-SR1.4 has been proposed as a solution for large enterprise datacenters, which often re-use duplex cabling originally installed for Gigabit Ethernet or 10GBASE-SR. Enterprise customers typically prefer duplex cabling due to embedded space, lower cost, familiarity with duplex connectors, and space constrains in patch panels or cable trays. A 100 m reach over OM4 MMF, and longer reach over OM5 MMF, appears possible for these links based on detailed contributions on technical feasibility [4]. The Chang paper at OFC 2017 validated the feasibility of a 265.625 Gb/s (53.125 Gb/s × 5λ) aggregate speed over a single OM5 fiber, while the Castro OFC 2016 paper also demonstrated 50 Gb/s/λ transmission over a 5 λ grid, using OM4 MMF. These papers confirmed the strong potential for 100 m reach on OM4 and 150 m reach on OM5 MMF using a 200GBASE-SR1.4 module.

Parameter	OM4 MMF	WideBand (OM5) MMF
Effective modal bandwidth at 850 nm, min (MHz.km)	4700	4700
Effective modal bandwidth at 953 nm, min (MHz.km)	Not spec'd	2470
Chromatic dispersion at 840 nm, max (ps/nm*km)	108.4	103
Chromatic dispersion at 953 nm, max (ps/nm*km)	65	61.7
Zero dispersion wavelength (nm)	$1295 \le \lambda_0 \le 1340$ nm	$1297 \le \lambda_0 \le 1328$ nm
Zero dispersion slope (ps/nm^2.km)	$S_0 \le 0.105$ for $1295 \le \lambda0 \le 1310$ nm, and $\le 0.000375(1590-\lambda_0)$ for $1310 \le \lambda0 \le 1340$ nm	$S_0 \le 4(-103)/(840(1-(\lambda_0/840)^4))$
Fiber attenuation at 850 nm per TIA-492AAAD&E, max (dB/km)	2.5	2.5
Fiber attenuation at 953 nm per TIA-492AAAD&E, max (dB/km)	Not spec'd	1.8

Table 2. Performance Spec Comparison of OM4 & OM5.

It is also technically feasible to build a shorter reach (possibly 30 m, to be determined), but potentially very useful, 400 G transceiver for duplex MMF using eight VCSEL wavelengths. The Sun JLT 2017 and Pavan JLT 2017 papers demonstrated 50 Gb/s PAM4 transmission over a 8-λ range from 850 nm to 1060 nm indicating the feasibility of 400 Gb/s over a single pair of OM3/4/5 MMF. The reach number in this case may drop from 100 m to perhaps ~30 m, with further study needed.

The Lavrencik work published at OFC'17 and ECOC'17 showed that VCSELs with modestly higher bandwidth might transmit 100 Gbps per λ over 100 m OM4 MMF at 850 nm and 105 m over OM5 MMF from 850 to 940 nm, with strong equalization. Practical 100G VCSELs may exist by 2021, enabling an 800GBASE-SR8 module to breakout to a middle-of-row switch for 100 G server-switch interconnects when copper can no longer meet reach requirements.

Table 2 compares the fiber characteristics of OM4 and OM5 fiber. OM5 has less chromatic dispersion and more bandwidth than OM4 at 953 nm. Note that according to IEC [9] the worst case effective modal bandwidth of OM4 is less than the specified 2470 MHz.km for OM5 as shown in Figure 1. Additionally, a co-doped OM5 design used in Sun JLT'17 [4, 10] could have EMB larger than TIA OM5 specification at much wider wavelength range. Also, these proposals do not require new OM5 fiber to be applied at the traditional 100 m reach targets that have been common in Ethernet standards. However, OM5 fiber will be useful in extending the supported link distances for multi-wavelength solution.

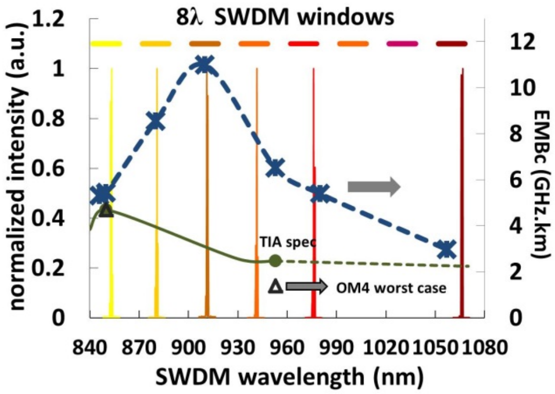


Figure 1. EMBc of a OM5 sample, TIA specification of EMBc for OM5, worst case OM4 EMBc at 953 nm [4,9].

3. New Experimental Evidence

At OFC 2018, we demonstrated 100 Gb/s Ethernet traffic transmission over a simulated two tiers data center network using a commercial Dell Z9100-ON 100 G switch without special cooling. The simulated spine-leaf link transmits 100 Gb/s Ethernet traffic over a 400 m duplex OM5 cable between two Finisar 100G-SWDM4 transceivers. The simulated leaf-server link transmits a 100GBASE-SR4 signal over an OM4 MPO cable concatenated by three 100 m OM4 cables, three patch cords and six MPO connections, then break out to four 25 GBASE-SR transceivers through a 4-pair MPO to 1-pair LC breakout cable. The demo ran three days without generating any frame error or bit error.

We use the same simulated two-tiers data transmission link to validate duplex transmission of 100G-PAM4-BiDi signals by replacing the two 100G-SWDM4 transceivers in Figure 2. Three

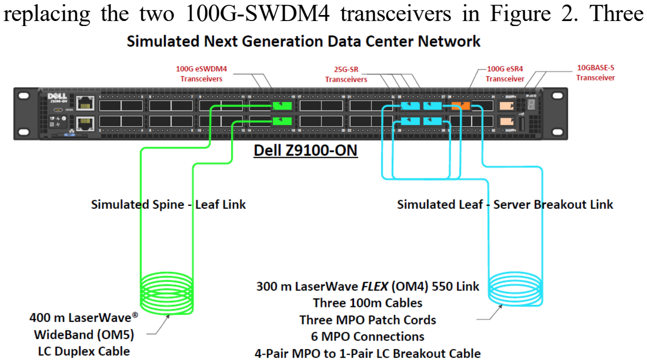


Figure 2. Simulated two tiers 100 Gb/s data center network: the simulated spine-leaf link transmits 100G-SWDM4 Ethernet signal over 400 m duplex OM5 cable; the simulated leaf-server link transmits 100GBASE-SR4 signals over a OM4 MPO cable concatenated by three 100 m cables, three patch cords and 6 MPO connections, then break out to four 25 GBASE-SR transceivers through a 4-pair MPO to 1-pair LC breakout cable.

OM5 fibers at various lengths from 100 m to 400 m and one marginal OM4 fiber from 100 to 200 m were evaluated. To assess the link performance, a variable optical attenuator (VOA) was added in the optical link. The attenuation of the VOA was adjusted to find the threshold for error-free transmission after FEC. (Note: the FEC is terminated in the 100G-BiDi module. It is preferable to do these experiments with pre-FEC BER and plot waterfall curves to extract Rx sensitivity and penalties, but we were not able to turn FEC off in these experiments.) The received power at the error-free threshold is considered an "effective Rx sensitivity," and the difference in effective Rx sensitivity with & without fiber in the link is denoted as a dispersion power penalty (DPP). As a rough guide to interpreting results, we have found in the past that DPPs extracted from pre-FEC BER data should be less than 4-5 dB for stable link performance.

The received power at the threshold at 855 and 907 nm for uplink and downlink was recorded. The effective Rx sensitivity was measured by an Agilent power meter at the two ends of the optical link before the optical signals entering the receiver at 855 or 907 nm. Note that the disconnection of the optical links to the transmitter and receiver may introduce an error bar of +/-0.5 dB. The measurement accuracy is expected to be improved by directly monitoring Tx/Rx

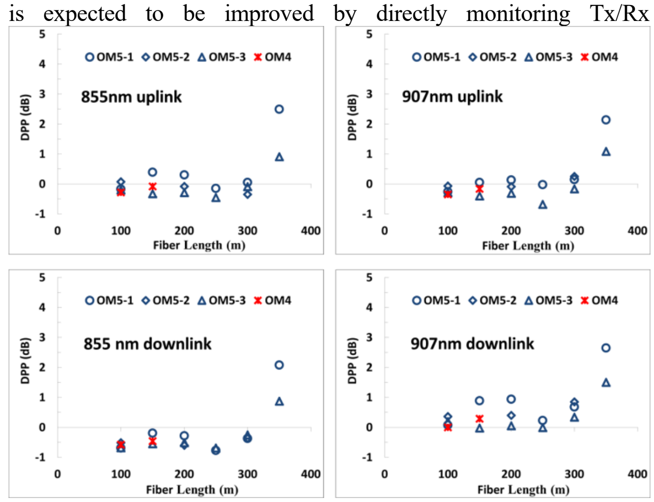


Figure 3. Dispersion power penalty (DPP) of 100G-PAM4_BiD transmission over three OM5 and one OM4 links using a Dell Z9100-ON switch.

power in switch diagnostic tool in the future. Since it could not be determined which wavelength failed first, the DPP of one wavelength may be better than shown in Figure 3. The DPP is less than 3 dB for all three OM5 fibers up to 350 m at both 855 and 907 nm for uplink and downlink, indicating that there is more than ample margin to support robust link performance over the design link distance of 150 m over OM5 MMF, including environmental perturbations such as temperature variation. At 400 m, the link is susceptible to failure with environmental perturbation such as shaking of the jumper. For OM4, the DPP is almost negligible at the design reach of 100 m as well as at the extended reach of 150 m, at both wavelengths. However, OM4 fiber could not support at 200 m transmission.

To assess the critical impact of temperature variation on the 100G-PAM4-BiDi link performance, we used a thermal electric cooler (TEC) to vary the temperature of a 100G-PAM4-BiDi transceiver plugged into a MCB board. Full 100 Gb/s Ethernet traffic was generated by a 100G Viavi ONT. Samples of OM5-1 at 100, 150, 200, 300 and 400 m are tested at 12 ^{0}C, 25 ^{0}C, 50 ^{0}C and 75 ^{0}C. The RMS spectral width of the 100G-PAM4-BiDi transceiver was less than 0.45 nm at both wavelengths across the temperature range. Effective Rx sensitivity was tested and DPPs were extracted. DPP was less than 2 dB up to 300 m at 12 ^{0}C, 25 ^{0}C, 50 ^{0}C and 75 ^{0}C. The DPP exceeded the acceptable range at 5.21 dB for 400 m at 12 ^{0}C and 907 nm. This experiment demonstrates the feasibility of 100G-PAM4-BiDi transmission over extended OM5 link lengths. In the future, temperature control at a wider range from -4 to 75 ^{0}C at finer incremental step and more transceivers and fibers will be evaluated.

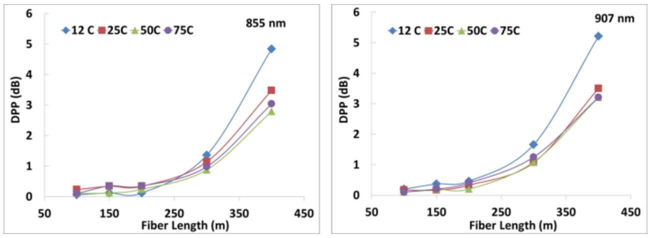


Figure 4. DPP for OM5-1 of 100, 150, 200, 300 and 400m at 12 ^{0}C, 25 ^{0}C, 50 ^{0}C and 75 ^{0}C. Left is 855 nm and right is 907 nm.

4. Conclusions

VCSEL-MMF short reach optical interconnects will continue to have a cost and power advantage in next generation 200, 400 Gb/s and higher speed networks. SWDM and 50G-PAM4 technologies have enabled commercial 100G duplex MMF solutions. 100G-PAM4 transceivers and OM4/OM5 MMF provides an upgrade path to a standards-based 400 Gb/s four pair SR4.2 solution. 50G-PAM4 transmission using 850 nm sources provide a building block for 400GBASE-SR8. The IEEE P802.3cm task force for 400 Gb/s over MMF has adopted 400GBASE-SR8 and 400GBASE-SR4.2 baseline proposals. Several publications have demonstrated the technical feasibility of 200 Gb/s and 400 Gb/s duplex MMF transmission over OM4 and OM5 fiber. New experiments using a 100G commercial switch and Viavi Optical Network Tester demonstrate 100G-PAM4-BiDi transmission over duplex OM4 and OM5 MMF. The advantage of OM5 in supporting longer link lengths is shown. Transmission under a wide temperature range is documented, demonstrating the robust performance of the link.

Acknowledgement: We acknowledge Foxconn Interconnect Technology, Ltd (FOIT) and Finisar for providing the 100G transceiver modules.

5. References

[1] http://www.ieee802.org/3/NGMMF/public/index_Jan_2018.html

[2] J. King, P. Kolesar, http://www.ieee802.org/3/cm/public/May18/king_3cm_01b_0518.pdf, "400 Gb/s 100-m 8-pair MMF objective baseline proposal" , IEEE P802.3cm, Pittsburg, PA, May 2018

[3] J. Ingham, http://www.ieee802.org/3/cm/public/July18/ingham_3cm_01b_0718.pdf, "Baseline proposal for a 400 Gb/s optical PMD supporting four MMF pairs," IEEE P802.3cm 400 Gb/s over Multimode Fiber Task Force, San Diego, CA, July 2018

[4] E. Parsons, Y. Sun, F. Chang, and A. Amezcua, "Technical Feasibility of 50 Gb/s PM4 using VCSELs from 850 to 1060nm", http://www.ieee802.org/3/NGMMF/public/parsons_NGMMF_01_jan18.pdf, Geneva, Jan 2018

[5] "SMF PMD Nominal Specifications," http://www.ieee802.org/3/bs/public/adhoc/smf/14_10_14/cole_01a_1014.smf.pdf, Oct. 2014

[6] IEEE802.3cd, http://www.ieee802.org/3/cd/

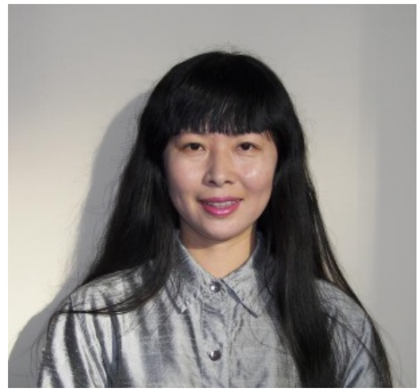
[7] Y. Sun, K. Scott, R. Shubochkin, A. McCurdy, J. Kamino, A. Swarts, R. Lingle, D. Braganza, D.S. Vaidya, M. Chow, J Wynia, T. Gray, WJ. Fan, S. Koelher, K. Wade, D. Gazula, G. Landry, J. Tatum, "High speed short wavelength division multiplexing transmission over next generation wideband multimode optical fiber," Proceedings of the 65th IWCS conference, pp476-484, the International Wire & Cable Symposium, 2016

[8] Y. Sun, J. Kamino, R. Shubochkin, A. Swartz, R. Lingle, D. Braganza, "High speed short reach optical interconnect over OM4 and OM5 multimode optical fiber," Proceedings of the 66th IWCS conference, p791-796, the International Wire & Cable Symposium, 2017

[9] IEC 60793-2-10 ed.7 "Optical fibres - Part 2-10: Product specifications - Sectional specification for category A1 multimode fibres", Committee Draft for voting

[10] K. Balemarthy, R. Shubochkin, Y. Sun, "Next-Generation wideband multimode fibers for data centers", Proc. SPIE977504, March 7, 2016

6. Pictures of Authors



Yi Sun has a BS/MS in Astronomy from Peking University, a MS in EE and a Ph.D. in Physics from Northwestern University. Dr. Yi Sun joined Bell Laboratories in 2001 and currently is a Senior R&D Engineer and Distinguished Member of Technical Staff in OFS Fitel, LLC in USA. Dr. Sun's research work covers various novel optical fibers and their applications in Datacom and SDM. Dr. Sun is the holder of 12 patents and has authored more than 100 publications. Dr. Sun receives several outstanding contribution awards by OFS and the 2011 Spirit Endeavor Award by TAG. Dr. Yi Sun is a senior member of IEEE.

.

Alex Swartz is a co-op student working in the Fiber Design and Systems Research Department at OFS and is pursuing a B.S. degree in Electrical Engineering at the Georgia Institute of Technology in Atlanta. He has worked on the design and construction of robotic and autonomous vehicles with IEEE and now focuses on the testing, measurement, and development of single mode and advanced multimode optical fibers for long-haul and short reach interconnector applications, respectively. His academic and research involvement focuses on optical communication networks as well as signal processing for optical systems.

Robert Lingle, Jr. is Director of Systems & Technology Strategy at OFS in Norcross, GA, as well as Adjunct Professor of Electrical and Computer Engineering at the Georgia Institute of Technology. He has a research background in short pulse lasers and their application to fundamental processes in liquids and interfaces, with a Ph.D. in physics from LSU and a postdoc in surface physics at UC Berkeley. At Bell Labs and now OFS, he first worked in sol-gel materials chemistry and then managed the development and commercialization of many new optical fiber products. His current research is in the confluence of optical and electronic methods for mitigating impairments in optical communications systems.

John Kamino is a Senior Manager – Product Management for OFS (www.ofsoptics.com). A 27-year veteran of the company, he has also held positions in optical connectivity product management, offer management, product marketing, sales, and engineering. He has a B.S. in Chemical Engineering from the University of Nebraska-Lincoln, and a M.B.A. from Mercer University.

Roman Shubochkin is a Senior R&D Engineer and a Distinguished Member of Technical Staff at OFS. He received BS/MS in Optical Engineering from Moscow Power Engineering Institute and MS/PhD in Electrical Engineering from Brown University, Providence RI. He worked as a research fellow at the Fiber Optics and Solid State Physics Divisions of the General Physics Institute in Moscow, Russia, as well as a senior research associate at Boston University. His research background includes design and fabrication of specialty optical fibers, research on fiber lasers, fiber optic sensors, pulsed solid state lasers, and rare-earth doped glasses. He joined the Optical Fiber R&D group at OFS in 2012 and works on the design, fabrication, and testing of novel transmission optical fibers. Dr. Shubochkin serves as a Vice Chair for the TR42.12 TIA Subcommittee on Optical Fibers and Cables and represents OFS at the Fibre Channel INCITS/T11 Technical Committee. Dr. Shubochkin is a member of OSA and Sigma Xi.

David Braganza is Technical Manager for Fiber R&D at OFS. He received the B.E. degree in electrical engineering from Pune University, Pune, India, and the M.S. and Ph.D. in Electrical Engineering from Clemson University, Clemson S.C., USA. In 2007 he joined the optical fiber R&D group at OFS where his early work focused on the development of novel processes for optical fiber manufacturing. Dr. Braganza is a senior member of IEEE.