multiple servers in a SAN. This allows heterogeneous SANs to be constructed which are still able to take advantage of policy-based management, workload balancing, and similar functions (the cost of network management can be 3 to 10 times greater than the SAN hardware cost). The switched FCP fabric is also designed for massive scalability without disruption to the existing traffic, and for high reliability, availability, and serviceability with no single points of failure. An important alternative to FCP-based SANs is the Fibre Connection (FICON)™ protocol,[2] originally developed for enterprise servers and currently being standardized by the American National Standards Institute (ANSI). FICON is implemented in essentially the same way as FCP below layer 3 protocols, and includes enhancements above layer 4 which improve data integrity. Through various design features, FICON channels maintain their maximum data throughput over distances of at least 100 km.

A complimentary approach to SANs known as Network Attached Storage (NAS) relies on IP protocols like Gigabit Ethernet to interconnect application servers with storage appliances. NAS does not segregate storage traffic from other types of IP data, and is typically optimized for smaller data packet transfers, using file protocols rather than block protocols (the maximum data block size for Gigabit Ethernet is 9 kilobytes using jumbo frames). The principle benefit of NAS is its flexibility and ease of installation; storage appliances are designed to "plug and play" on existing local area networks (LANs) with a minimum of effort. The difference in complexity is apparent; while you install a NAS device, you implement a SAN infrastructure. Thus, NAS architectures can vary from shared nothing (each server accesses storage independently with no clustering, load balancing, or failover capabilities) to shared everything (any server can concurrently access any storage device). NAS devices can be installed as either a multi-node cluster or distributed file servers; in the latter case, NAS management is non-centralized. The two approaches are not mutually exclusive, and various types of SAN/NAS gateways are available to merge the two approaches and groom storage data traffic. For example, the IBM SANergy™ gateway and related offerings from IBM Global Services provide heterogeneous file sharing between SAN or NAS servers and file storage.

Storage applications over IP are generating interest because of their ability to use existing IP network infrastructures over metropolitan or wide are network (MAN/WAN) distances. Although transport of Gigabit Ethernet over DWDM has been demonstrated to distances over 1000 km,[3] performance at such long latencies remains to be determined. There are currently three major ways to transport storage data over IP. First, iSCSI (SCSI over IP) uses a software agent on the server to encapsulate SCSI data into an IP packet, and tunnel that encapsulated packet through the TCP/IP stack; the data is unwrapped at the other end of the link. Second, storage over IP (SoIP) converts Fibre Channel to IP outside the server and bypasses the TCP/IP stack in the host, eliminating stack overhead in the process. Finally, the Fibre Channel over IP specification (FCoIP) is a proposed standard which allows routers to connect a Fibre Channel SAN to an IP network (note that similar options for Fibre Channel over ATM are also available). Of these options, iSCSI has attracted attention recently because of its handling of block I/O protocols and ready integration with embedded storage (iSCSI appliances) or gateways to an existing SAN or NAS. As a block transfer based network accessible over an IP network, iSCSI has the potential to offer the advantages of both SAN and NAS environments.

The fundamental building blocks of a SAN include the server, optical transport fabric (gateways, hubs, wavelength multiplexers), switches or routers, storage devices (disk control units or tape libraries), management software, and deployment service offerings. Steadily increasing server performance has driven the requirement for increased network bandwidth and storage capacity.[2] In the past decade alone, the cost per megabyte of disk storage has fallen by an order of magnitude to less than 50 cents. Data density on a storage medium has been doubling annually since 1997, growing faster than Moore's Law. Although optical storage is becoming more common, magnetic storage continues to grow due to recent advances in material science[4] which are expected to enable 100 gigabits/square inch by 2003, a level previously thought impossible (this would enable a laptop PC to hold 200 Gbytes, enough to accommodate 200,000 books or 300 CDs). There is also a significant opportunity to improve the optical transport fabric technology. For example, DWDM has emerged as a cost-effective way to extend storage networks over the MAN and WAN, and is expected to form the backbone for future storage service provider offerings. SONET backbones will likely co-exist with Ethernet encapsulation, especially in the MAN where new approaches such as resilient packet ring are expected to find applications. This in turn has driven interest in very high speed (nanosecond) all-optical switches and cross-connects. These technologies are among the elements of the future optical data center.

There are various organizations driving the next generation of industry standards for SANs, including trade associations for SCSI, Fibre Channel, and Infiniband. Today's storage environment is a combination of multiple SAN and NAS environments, typically with little attempt at global optimization. In the future, storage networks will evolve through the use of gateways, modular storage, and hybrid approaches such as iSCSI into a shared pool of intelligent storage devices accessible from any server, anywhere. Efforts are under way towards storage virtualization, which separates a single, logical view of storage resources from the physical device configuration; however, a standardized approach has yet to be established. Intelligent storage devices are expected to automate storage administration, making it simpler, less expensive, and continuously available; example include the IBM StorageTank™ SAN file system. With the growth of DWDM as a cost effective means to implement disaster recovery over extended distances, storage networks are spreading into the MAN and WAN. Today, clustered servers in a Parallel Sysplex™ support synchronous peer-to-peer remote copying over distances up to 40 km, and asynchronous remote copy using FICON is available up to 100 km.[2] This trend towards geographically dispersed SANs (GDSANs) is expected to continue, and represents a significant market for storage application providers. With consolidation of SAN, NAS, and iSCSI networks and extension of storage into the WAN, some have even proposed that in the future there will be only one SAN in the world. Whether or not this is the case, the potential of SANs is yet to be fully realized—application neutral, distance independent, infinitely scalable, user-centric networks that catalyze a host of new computing applications.

ESCON, FICON, SANergy, StorageTank, and Parallel Sysplex are trademarks of IBM Corporation; GeoMax is a servicemark of Quest; Ultravailable Broadband Network is a servicemark of AT&T.

References

1. A.R. Newton, et. al., "Future of Storage Area Networks," University of California study (1997).
2. C. DeCusatis (Editor in Chief), D. Clement, E. Maass, and R. Lasky (Associate Editors), *Handbook of Fiber Optic Data Communications volume 1 and 2* (Academic Press, New York, 2001).
3. IEEE: W. Xin, G.K. Chang, and T.T. Gibbons, "Transport of Gigabit Ethernet directly over WDM for 1062 km in the MONET Washington, D.C. network" paper WC1.4 in *Proceedings of IEEE summer topical meeting, Boca Raton, Florida,* (Institute of Electrical and Electronics Engineers, New York, 2000), pp. WC1-WC-4.
4. J. Lohau, A. Moser, D.T. Margulies, E.F> Fullerton, M.E. Schabes, "Dynamic coercivity measurements of antiferromagnetically coupled magnetic media layers," App. Phys. Lett. 78, 2748–2752 (2001).

ThH6 **10:00 am**

Design, transport performance study and engineering of a 11 Tb/s US mesh metro network

N. Madamopoulos, N. Antoniades, I. Roudas, *Applied Research, Photonic Research and Test Center, Corning Incorporated, 2200 Cottontail Lane, Somerset, NJ 08873, Email: madamopon@corning.com*

M. D. Vaughn, R. E. Wagner, *Optical Network Architecture, Corning Incorporated, Corning, NY 14831*

1. Introduction

There have been several experimental and theoretical studies focusing on WDM metro network optical layer engineering.[1–3] Most of these have been limited to either small[1,2] or rather symmetrical networks.[3] Furthermore, the above sample networks were not designed based on actual or projected traffic demands in the metro environment, but rather physical layer designs that were intended to demonstrate specific technical virtues. In this paper, we present for the first time an effective top-down simulation approach that enables the design of a 29-node, 11-Tb/s US mesh metro network in a transparent way based on realistic network requirements. Projected traffic demands based on current known demands and existing Central Offices (CO) are used to derive network architectures for a number of different US carriers for year 2005. A physical layer topology, based on network dimensioning, is then superimposed on the network architectures and the network is engineered for 100% of the required optical connections. This involves engineering paths with the longest length as well as the ones with the maximum number of optical hops. It was shown in the derived mesh network topology that transparent paths with maximum lengths of 465 km and paths with maximum number of hops equal to twelve were possible, thus enabling all other paths in a 29-node mesh WDM metro optical network.

2. Traffic Projection—Network Dimensioning

The optical network architecture under study is derived from traffic modeling of a major US metropolitan area and different market penetrations (low, moderate and high demand scenarios) of existing (ILEC) and new (CLEC) carriers presented.[4] The traffic projections are based on projected voice, transaction data and Internet traffic demands taking into account knowledge of current demands, the change of the population, the change of the non-production employees, and the Internet hosts.[5] Three different carriers were considered: one ILEC with seventeen COs, one CLEC (CLEC1) with ten COs, and nine other new carriers (CLEC2) with only one or two COs each. There are also 12 peripheral nodes that represent the points where traffic from the greater metropolitan area is aggregated. The projected traffic demands and CO locations are used to derive individual traffic matrices for the different carriers. The traffic demand of the core network (which excludes the peripheral nodes) is the driver of the assumed connectivity (selection of links connecting metro nodes—Fig. 1(a)). After having selected a specific connectivity for the core, the projected traffic from the peripheral nodes is also considered and the connectivity is expanded further to include all possible connections between any two nodes in the network. Fig. 1(a) shows the resulting mesh ILEC network topology. The nodes of all other carriers are left unconnected because as will be shown they have less demanding network performance requirements. A commercial network dimensioning and routing tool (WDMNetDesign [6]) is then used to generate required network parameters, such as connection lengths, add/drop percentages, and number of fibers at each node, and to assign wavelengths for each connection. For the network case study of Fig. 1(a) 40 wavelengths per fiber and 1 + 1 protection were considered.

The peripheral nodes in Fig. 1(a) (e.g., Z4) that have connecting fibers coming from two directions (connectivity 2) are Optical Add/Drop Multiplexers (OADMs). All nodes with connectivity larger than two (e.g., O, K, etc) have Optical Crossconnect (OXC) equipment. Note that each fiber link of Fig. 1(a) may consist of more than one fiber and as a result the size of the OXC module as well as the number of ports of the switching fabric will vary. The design details of the OADMs and OXCs are beyond the scope of this article and thus in trying to maintain the generality of our methodology we used a "black box" approach for these network elements. Insertion losses, filter, crosstalk and noise characteristics from commercial components are used for the transport layer analysis below.

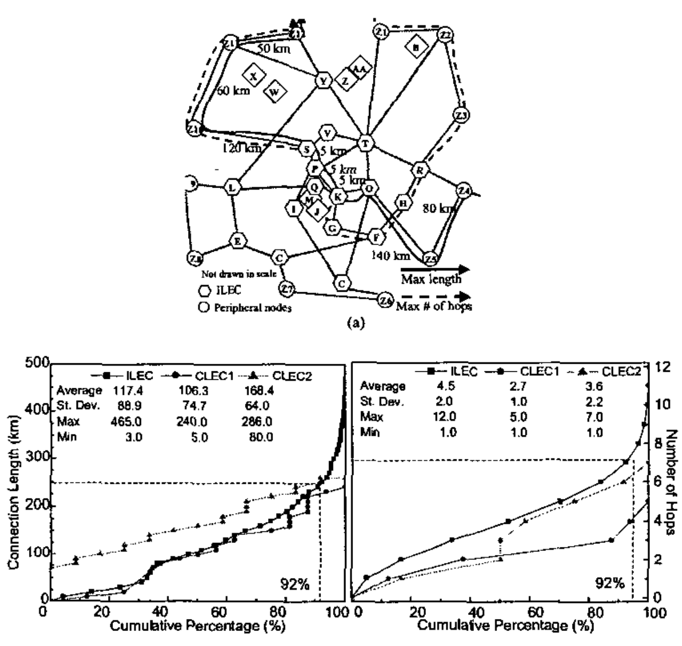
After studying the different market share scenarios[4] for 2.5 Gbps and 10 Gbps, we concentrated our transport layer performance modeling study on 10 Gbps since this will be typical of metro-area networks in the near-term. Fig. 1(b) shows the cumulative percentage of all the possible number of connections that are within a specific connection length and traverse a specific number of hops for all three carrier scenarios. *Hops* are number of nodes that the signal has to traverse in an end-to-end connection (excluding the originating node). In particular, Fig. 1(b) shows that 92% of the required ILEC connections have connection lengths of less than 250 km, while they traverse no more than seven hops. Fig. 1(b) also shows the statistics for the connection lengths and number of hops required for the three carriers. It must be noted that the results of Fig. 1(b) include both working and protection paths. It is obvious that the ILEC network contains the longest connection lengths, as well as the maximum number of hops compared to the CLECs. Hence, as mentioned earlier, we chose this network as our case study in Fig. 1(a). Longest length- and maximum number of hops-paths are modeled since they represent the worst-case engineering scenarios. The former will suffer from more fiber nonlinearities and chromatic dispersion whereas the latter will accumulate the worst crosstalk, filtering and ASE noise effects. We model Z4-Z5-O-K-P-S-Z10-Z11-Z12 (longest length) and Z1-Z2-Z3-R-H-F-G-K-P-S-Z10-Z11-Z12 (maximum number of hops) paths in Fig. 1(a). Engineering for the above paths assures better performance for the smaller size CLEC networks.

3. Network Engineering

The remaining part of our simulation methodology is network engineering based on the superimposed physical layer topology presented above. In trying to engineer the identified worst-case paths of Fig. 1(a), different transport layer impairments that will degrade system performance need to be examined. These include amplified spontaneous emission (ASE) noise, power ripple of components, chromatic dispersion, optical crosstalk, waveform distortion due to filter concatenation, fiber nonlinearities and polarization effects such as polarization mode dispersion (PMD), and polarization dependent loss (PDL). Complete time/frequency-domain simulation of such a large network is a very time consuming process. As a result, a three-step computationally efficient simulation methodology is derived. The first step involves performing wavelength-domain simulation on the entire network,[7] followed by conventional time/frequency-domain simulations on the identified worst-case paths. Finally, a budgeting approach based on accurate impairment models and information obtained from the previous two simulation steps is used to estimate the Q-performance on these worst-case paths in the network.

Metro-optimized EDFAs are used as pre- and post-amplifiers at each OADM and in-line EDFAs are used in long links (>80 km). EDFAs are engineered to gains of 10 to 21 dB, with noise figures 8.7 to 5.4 dB, respectively. Standard single-mode fiber and per-link chromatic dispersion compensation is used. Typical through insertion loss used is 26 dB for the OXC and 12 dB for the OADMs.

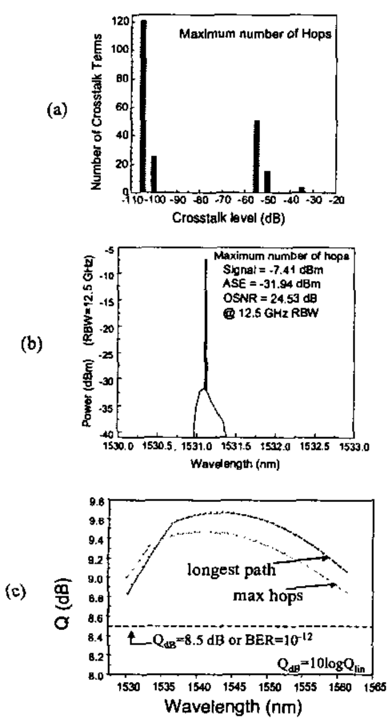
During the *first simulation step* optical crosstalk terms generated in the network are calculated for all possible paths. For the purposes of this work only common-channel crosstalk is considered (leakage terms at the same nominal wavelength) since it is the most detrimental. Furthermore, as a worst case scenario, it is assumed that the signal and the crosstalk terms have identical



ThH6 Fig. 1. (a) Metro network case study with longest-length path (solid) and maximum number of hops path (dotted). Examined network is that of an ILEC. Unconnected nodes are part of CLEC carrier networks; (b) Connection lengths and number of hops required for all connections in the three carrier networks.

polarization but uncorrelated phase noises (incoherent homodyne crosstalk case). Typical sources of such crosstalk are MUX/DMUXs, switch fabrics, and optical filters. Only terms with crosstalk level above −110 dB are considered in the simulations. We then calculate the crosstalk-induced Q-penalty which is defined as the difference in Q (dB, defined as $10\log(Q_{lin})$) at the optical receiver that is observed at a given error probability (P_e) for the case of crosstalk-free system and the one that contains the effects of crosstalk. For the evaluation of the P_e of an optically amplified direct-detection receiver in the presence of N interferers at the same nominal wavelength as the signal, we use the model in Gimlett et al.[8] Fig. 2(a) shows the histogram of the crosstalk terms generated for the path with the maximum number of hops. About 210 terms are collected from which the dominant ones (two at −35 dB) are generated at the OADMs. Fourteen crosstalk terms are at −50 dB and forty-nine terms at −55 dB below the signal originate from the OXCs. Similar numbers of crosstalk terms were observed for the longest connection length case too. Crosstalk-induced Q penalty of 1dB is obtained for both examined paths. Fig. 2(b) shows the spectrum, for one possible wavelength (at 1530.13 nm), at the end of the path, before the preamplified receiver. A 24.5 dB OSNR is achieved and the effect of optical filtering is obvious on the ASE noise.

The *second simulation step* consists of time-domain simulations on the two worst-case paths.



ThH6 Fig. 2. (a) Histogram of the crosstalk terms generated for the path having the maximum number of hops; (b) spectrum of one channel as calculated using wavelength domain simulation for the same path as 2(a); (c) Q-performance results obtained from Q-budgeting simulation approach for both the longest path as well as the maximum hops path of Fig. 1.

Accurate error probability calculations are performed for each path taking into account filter concatenation, and ASE noise accumulation.[9] A 0.5 dB Q-penalty due to filter concatenation was calculated assuming tight ripple requirements on the optical filters. Nonlinearities and polarization effects are further budgeted with 0.5 dB Q-penalty.[10] Based on the above, a Q-budgeting approach is used as a *third and final simulation step* in trying to engineer the above network design.[3] This is a budgeting approach where Q-penalties for the various effects obtained from the previous steps of the simulation methodology are used. The Q-performance of all possible wavelengths to be considered for this channel, applied on the worst-case paths, can then be calculated by keeping track of the signal and ASE noise powers through the network. Channel power divergence, dispersion/chirp, ASE noise, receiver noise terms and budgeting for all other impairments in the form of dBQ is considered. Note that this is different from a simplistic Q-budget since the margin allocated for each impairment is not fixed, but it is calculated from a corresponding impairment model. Fig. 2(c) presents the final results of the three-step transport layer simulation methodology. The Q-performance of all possible wavelengths for both of the paths examined in Fig. 1 are above the required 8.5 dB (Bit Error Rate $<10^{-12}$). The shape of the curves is due to the dispersion map of the system. It is important to note that Fig. 2 demonstrates that engineering for the above worst-case paths assures that 100% of the paths of the mesh metro network for the typical ILEC can be transparently engineered.

4. Summary

A new top-down simulation methodology for successfully enabling transparent optical networks in the metro environment is presented. A 11 Tb/s ILEC network was engineered based on projected traffic demands for a typical medium size US metropolitan area. The approach combines traffic modeling, network dimensioning, and transport layer performance modeling. The latter is further composed of three main simulation steps: efficient wavelength-domain simulations for network crosstalk analysis, time-domain simulation for determining Q-penalties for pulse propagation effects and finally a Q-budgeting approach that combines the above results for flexible network design and engineering. As a result metro players like ILECs or CLECs can effectively and quickly engineer transparent networks in the metro environment using simulation methodologies like the above. Details on the actual traffic and transport layer models used as well as experimental validation results for the models will be presented in more detail at the conference.

References

1. W.T. Anderson, "The MONET-project a final report," *J. of Lightwave Techn.*, 18, 1988–2009, (2000).
2. D. Stoll, et al., "Metropolitan DWDM: a dynamically configurable ring for the KomNet field trial in Berlin," *IEEE Comm. Mag.*, 39, 106–113, (2001).
3. N. Antoniades et al, "Performance engineering and topological design of metro WDM optical networks using computer simulation," *IEEE J. of Sel. Areas in Comm., Special Issue on WDM-based network architectures*, To appear Feb 2002.
4. N. Madamopoulos, et al., "Metro network architecture scenarios, equipment requirements and implications for carriers," in Proc. OFC 2001, WL2, Anaheim, CA, March 17–22, 2001.
5. M.D. Vaughn and R.E. Wagner, "Metropolitan Network Traffic Demand Study," in Proc. LEOS 2000, MK4, Rio Grande, PR, November 13–16, 2000.
6. P. Arijs, et al., "Design of ring and mesh based WDM transport networks", *Optical Networks Magazine*, Vol. 1, No. 3, pp. 25–40, July 2000.
7. I. Roudas, et al., "Wavelength-domain simulation of multiwavelength optical networks," *IEEE J. of Sel. Top. In Quant. Electr*, 6, 348–362 (2000).
8. J.L. Gimlett, et al. "Effects of phase-to-intensity noise conversion by multiple reflections on gigabit-per-second. . . . ," *J. Lightwave Techn.*, 7, 888–895, (1989).
9. M. Kuznetsov, et al. "Power penalty for optical signals due to dispersion slope in WDM filter cascades," *IEEE Phot. Technol. Lett.*, 11, 1411–1413, (1999).
10. N. Antoniades et al., "Computer simulation of a Metro WDM interconnected ring network", *IEEE Phot. Technol. Lett.*, 12, 1576–1578, (2000).

ThH7 **10:15 am**

An All-Optical Multi-hop (Cascaded) High Bit Rate Wireless Communication Field Trial

Uri Adar, Tal Kuzniz, Ilan D. Haber, Nahum Budin, *Oraccess Inc., ISRAEL, Email: uria@oraccess.com*

Moshe Tur, *Faculty of Engineering, Tel-Aviv University, Tel-Aviv, ISRAEL 69978, Email: tur@eng.tau.ac.il*

1. Introduction

Traditional optical wireless systems use electrical-to-optical and optical-to-electrical conversions on its transmitting and receiving terminals, respectively.[1] Commercial systems offer up to 1.5 Gbps rates. Few companies announced 2.5 Gbps systems. 2.5 Gbps seems to be the maximum practical rate for systems based on regeneration due to receiver performance limitations. However, higher bit-rates and DWDM traffic are needed for metro networks and disaster recovery applications as bandwidth requirements increase. Some experimental works presented higher bandwidth transmission[2,3,4] but none of them was based on a full operational system that exhibits long-term stability and enables cascading of links. All-Optical wireless communication is considered to be the next generation of optical wireless communication. Such systems will carry DWDM traffic with 10 Gbps rate per channel and will be a true fiber replacement in terms of capacity and reliability for short distances.

The unique characteristics of such systems are: (i) Transparency to wavelengths, bit-rates, and protocols for seamless integration with fiber optic networks; (ii) Wide wavelength range for DWDM payloads; (iii) Wide dynamic range that is needed to overcome diverse changes in atmospheric conditions; (iv) Scalability.

The unique components for such a system are: an all-Optical Automatic Gain Control (OAGC) to overcome turbulence effects, optical ampli-