IEEE APPLICATIONS & PRACTICE:
TOPICS IN OPTICAL COMMUNICATIONS

Storage Area Network Extension over Passive Optical Networks (S-PONs)

Si Yin, Yuanqiu Luo, Lei Zong, Stephen Rago, Jianjun Yu, Nirwan Ansari, and Ting Wang

ABSTRACT

After 9/11 and the accidental failure of the power grid in North America in 2003, storage area network (SAN) extension has emerged as critical to ensuring business continuity. However, SAN extension encounters challenges in the access network, including scalability problems, cost challenges, bandwidth bottlenecks and low throughput. In this article, we propose a new solution to address these problems: SAN extension over passive optical networks (S-PONs). To tackle the scalability problems and cost challenges, we designed the S-PON architecture based on the existing point-to-multiple-point (P2MP) PON infrastructure. To address the bandwidth bottlenecks in SAN extension, we propose three solutions for carrying storage signals with gigabit-level transmission. We also introduce a new device, XtenOLT, for implementing buffer pools by a new buffer- management scheme to improve SAN extension throughput and utility. Our experimental results show that, in the physical layer, the proposed S-PON transmission technologies successfully deliver SAN traffic to the long-haul at the rate of 2.5 Gb/s; in the network layer, S-PON with XtenOLT dramatically enhances deliverable throughput and utility over long-distance transmission.

INTRODUCTION

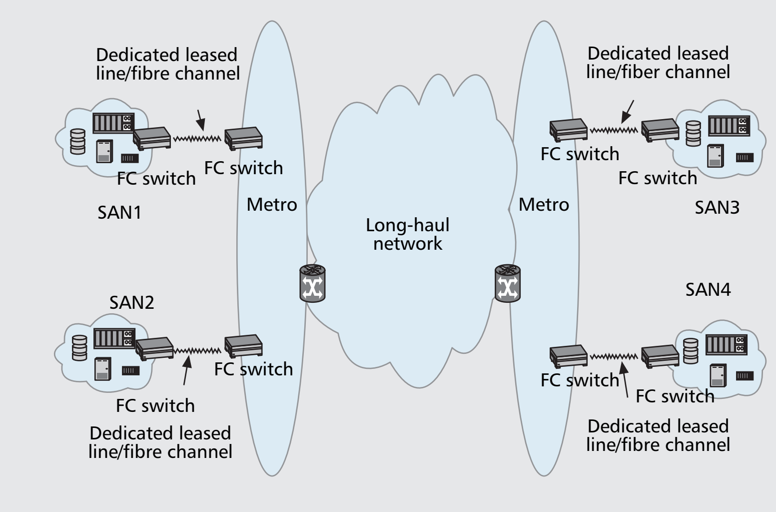
The rapid growth of data-intensive applications, including multimedia, e-business, e-learning, and Internet protocol television (IPTV), is driving the demand for higher data-storage capacity. Organizations want their huge amounts of data to be stored so it can be easily accessible and manageable. Furthermore, they require the critical data to be securely transported, stored, and consolidated at high speeds. The storage area network (SAN) is emerging as the data-storage technology of choice because of its significant performance advantages, such as better scalability and higher availability over the traditional storage architectures [1]. The SAN is a high-speed and special-purpose network that interconnects a set of storage devices with their servers. SAN architectures have won attention from large enterprises such as Google, Yahoo, and Amazon that have tremendous amounts of data to back up and consolidate, as well as replicate among different locations.

After 9/11 and the power grid failure in North America in 2003, SANs were widely deployed as the major data disaster recovery system infrastructure. To avoid severe damage from widespread power outages, earthquakes, fire, and terrorist attack, the storage sites must be physically separated by up to hundreds or even thousands of miles so that only one site will be affected in a disaster [2–4]. United States federal regulators, such as the Office of Management and Budget and the General Services Administration, have adopted a similar disaster recovery strategy to insure the continuity of operations plan (COOP), which is applicable to all federal agencies, airports, and financial institutions [5]. Recently, one 860-km-long testbed was set up in Europe to demonstrate the new services over SAN extension [6].

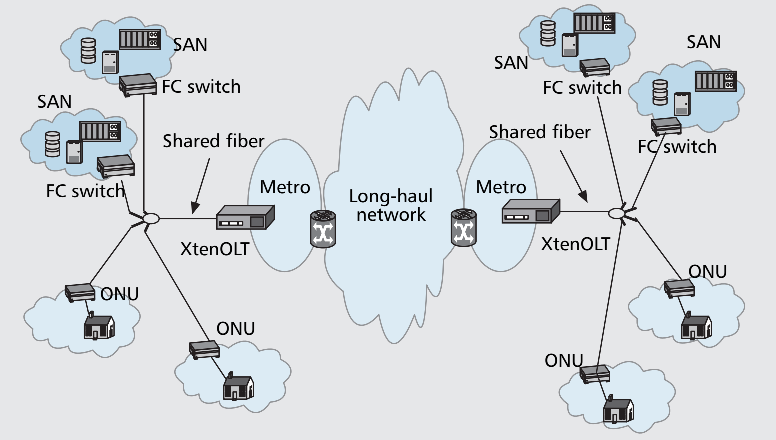
The existing literature covering SAN extension is mainly about long-haul overlay. The proposed solutions include optical-based extension solutions and IP-based extension solutions. The optical-based extension solutions include extending SAN over synchronous optical network (SONET) and over wavelength division multiplexing (WDM). SONET-based extension essentially assigns a dedicated SONET channel with fixed bandwidth to each SAN connection [7]. On the other hand, WDM-based extension divides bandwidth on a fiber into several non-overlapping channels (i.e., wavelengths) and conducts simultaneous message transmission on different wavelengths in the core network [8]. Finally, IP-based extension solutions encapsulate data units of SAN traffic into standard IP frames to be transported over core networks [9]. Several protocols, including Internet small computer system interface (iSCSI) [10], fibre channel over TCP/IP (FCIP) [11], and Internet fibre-channel protocol (iFCP) [12] have been introduced to transport the SCSI commands and responses, either by major vendors or the IP Storage Working Group of the Internet Engineering Task Force (IETF).

Few papers, however, address the challenges presented by extending SANs over the access network. The first challenge comes from the

0163-6804/08/$25.00 © 2008 IEEE



(a) Conventional SAN extension topology



(b) S-PON topology

Figure 1. *Comparison of conventional SAN extension and S-PON.*

conventional SAN extension solution, such as the fibre channel (FC), which uses point-to-point dedicated "dark fiber" to connect the SAN into the metro network. This approach lacks scalability and is cost-prohibitive because maintaining the storage service requires dedicated delivery and extra manpower. The second challenge is the bandwidth bottleneck of current access technologies. For example, DSL, cable and T1/E1 can only provide low megabit-level bandwidth, which is far below the gigabit-level bandwidth requirement of SANs [1]. The third challenge comes from the buffering constraints of SAN switches. The legacy SAN switch was historically designed with limited amounts of buffers for local transmission of up to 10 km. This leads to low throughput when the SAN extends to hundreds of miles. The aforementioned three-folded challenges motivated us to develop a new approach to extend SAN over the access network to improve scalability, lower cost and increase the speed of transmission, as well as improve throughput.

In the rest of this article, we introduce a new architecture, S-PON, to tackle the scalability problems and cost challenges and we present three different transmission techniques to solve the bandwidth bottleneck of current SAN extension methods. A new device and a new buffer management scheme in S-PON are discussed, respectively. Simulation results are summarized, and we provide an S-PON implementation example and draw conclusions.

S-PON: ARCHITECTURE

To overcome the scalability problem and cost challenges of dedicated FC, we propose to extend the SAN over the passive optical network (PON). The resulting architecture is called S-PON. Instead of using point-to-point "dark fiber" (Fig. 1a), S-PON employs the point-to-multiple-point (P2MP) architecture of PON, illustrated in Fig. 1b.

The PON infrastructure has been widely deployed in recent years. For example, Verizon's FiOS service, facilitated by PON technologies, has been deployed in 16 different states in the U.S. and is targeted to reach 50 percent of households by 2010 [13]. Since PON is leading the trend to next-generation broadband access, S-PON naturally solves the FC scalability problem by building on the growth of PON coverage. Furthermore, the P2MP architecture of S-PON allows SAN to share a single feeder fiber up to 20 km long in the access network with other optical network units (ONUs), thus greatly reducing the cost of SAN extension.

S-PON: TRANSMISSION TECHNOLOGIES

To solve the bandwidth bottleneck of current SAN extension techniques, we propose three different transmission technologies for the physical layer: in-band transmission, out-of-band transmission, and out-of-wavelength transmission.

With in-band transmission, the SAN shares the upstream channel with other ONUs through time division multiple access (TDMA). In this way, the SAN is regarded as a special ONU, sharing the 1 Gb/s bandwidth with other ONUs. The TDMA in-band transmission technique is illustrated in Fig. 2a.

For more critical SAN applications, S-PON fulfills the bandwidth requirements with out-of-band transmission technology. This is facilitated by sub-carrier multiple access (SCMA), as shown in Fig. 2b. The basedband carrier f_0 is for LAN traffic transmission, while two sub-carriers, f_1 and f_2, are used to transmit the storage data from SAN1 and SAN2, respectively. Either SAN can transmit gigabit-level traffic by using the allocated sub-carrier through the proposed communication infrastructure.

Out-of-wavelength techniques are employed for the most critical storage data transmission requiring high quality. The practical method takes advantage of wavelength division multiple access (WDMA). As shown in Fig. 2c, LANs are assigned wavelength λ_1 for data transmission, and SAN1 and SAN2 are assigned two other wavelengths, λ_2 and λ_3 respectively, for storage data transmission.

The remote node is responsible for multiplexing wavelengths in the upstream direction and

Transmission techniques	Media Access	Bandwidth	Security	Cost
In-band techniques	TDMA	Low	Low	Low
Out-of-band transmission	SCMA	Medium	Medium	Medium
Out-of-wavelength transmission	WDMA	High	High	High

Table 1. *Comparison of technologies for physical layer transmission in S-PON.*

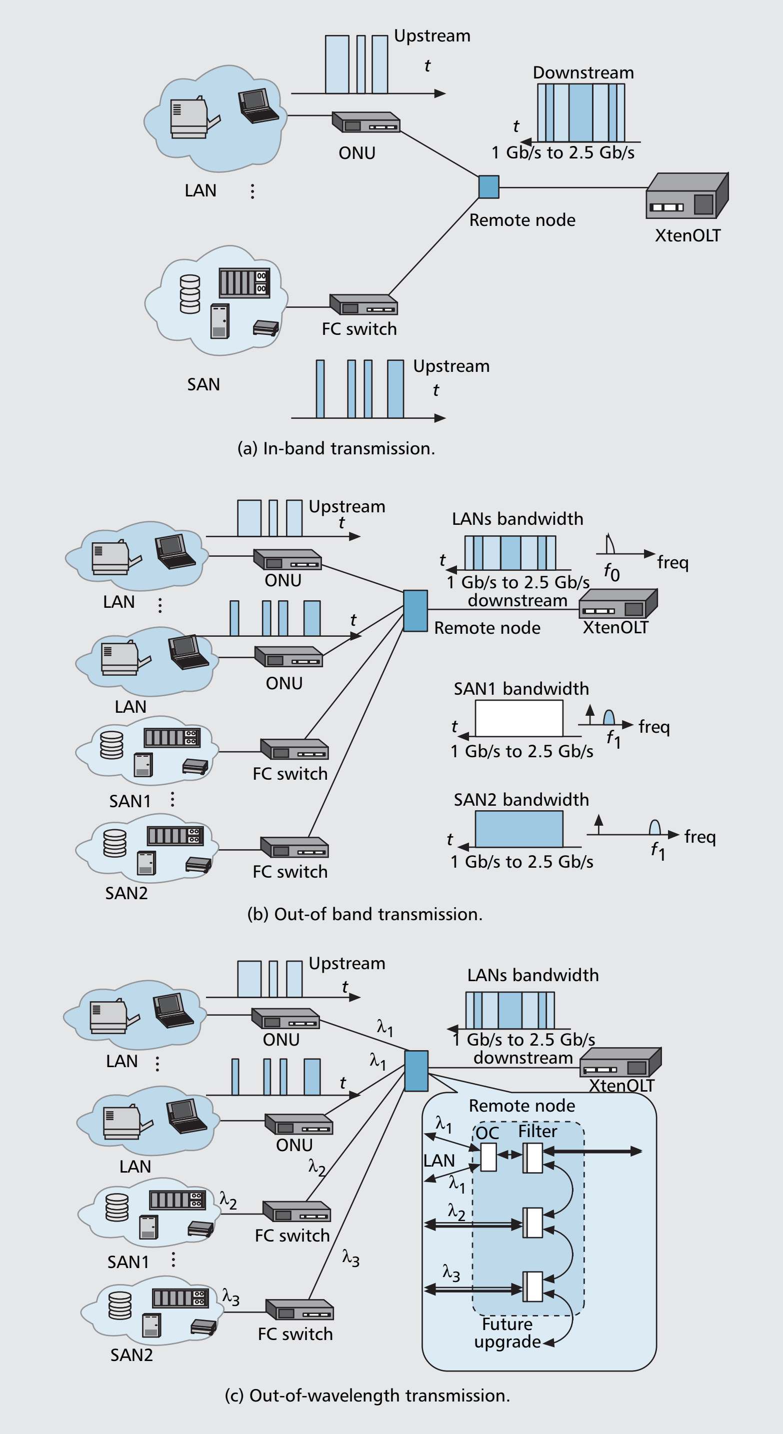


Figure 2. *Transmission technologies.*

demultiplexing in the downstream direction. Regular PON remote nodes, such as optical splitters, can still be used in the TDMA and SCMA scenarios. In the WDMA scenario, however, a modified remote node needs to be implemented, which is shown in the inset of Fig. 2c. The modified remote node separates the downstream LAN data and SAN data with a set of optical filters or an arrayed-waveguide-grating (AWG) to achieve high security and enhanced transmission rate. For the upstream data, the modified remote node multiplexes the LAN and SAN data and sends them to XtenOLT.

In addition, the optical transmitters and receivers in the FC switches and the OLT need to be upgraded to support sub-carrier and WDM transmission for the SCMA and WDMA scenarios, respectively. In the TDMA scenario, transmitters and receivers for a regular PON can still be used.

Table 1 summarizes the pros and cons of the three transmission techniques in term of media access, bandwidth, security and cost.

XTENOLT: A NEW DEVICE IN S-PON

The conventional SAN switch node was designed with few buffers for short-distance transmission. The storage flow control mechanism was implemented with buffers to hold the incoming FC frames before receiving acknowledgments [1]. When transmitting over hundreds of miles, these insufficient buffers lead to low throughput because of the storage flow control sensitivity to the long distance round-trip time.

To solve this problem, we propose a new device, XtenOLT, in the S-PON architecture. XtenOLT is an enhanced optical line terminal (OLT) in PON with storage service provisioning. The internal architecture of XtenOLT is illustrated in Fig. 3. Two buffer pools are constructed for buffering the incoming FC frames from local SANs and remote sites, respectively. The flow control and switch module is composed of a buffer-to-buffer (BTB) flow-control sub-module, an end-to-end (ETE) flow-control sub-module, and a switch interface, which are responsible for the BTB and ETE flow control and the packet switch [1]. An OLT module is also included, which is responsible for OLT arbitration. The transmission module is responsible for physical layer transmission through the TDMA, SCMA or WDMA sub-modules. The dynamic resource management (DRM) module is responsible for efficiently managing buffer pools. Lastly, the service differentiation (SD) module is responsible for differentiating services among the SANs.

TETRIS: A NEW BUFFER MANAGEMENT SCHEME

Among the various functional modules in XtenOLT, the DRM is the core module for buffer management. Various buffer management schemes have been proposed in the literature. The conventional fixed scheme simply allocates a constant number of buffers to each SAN regard-

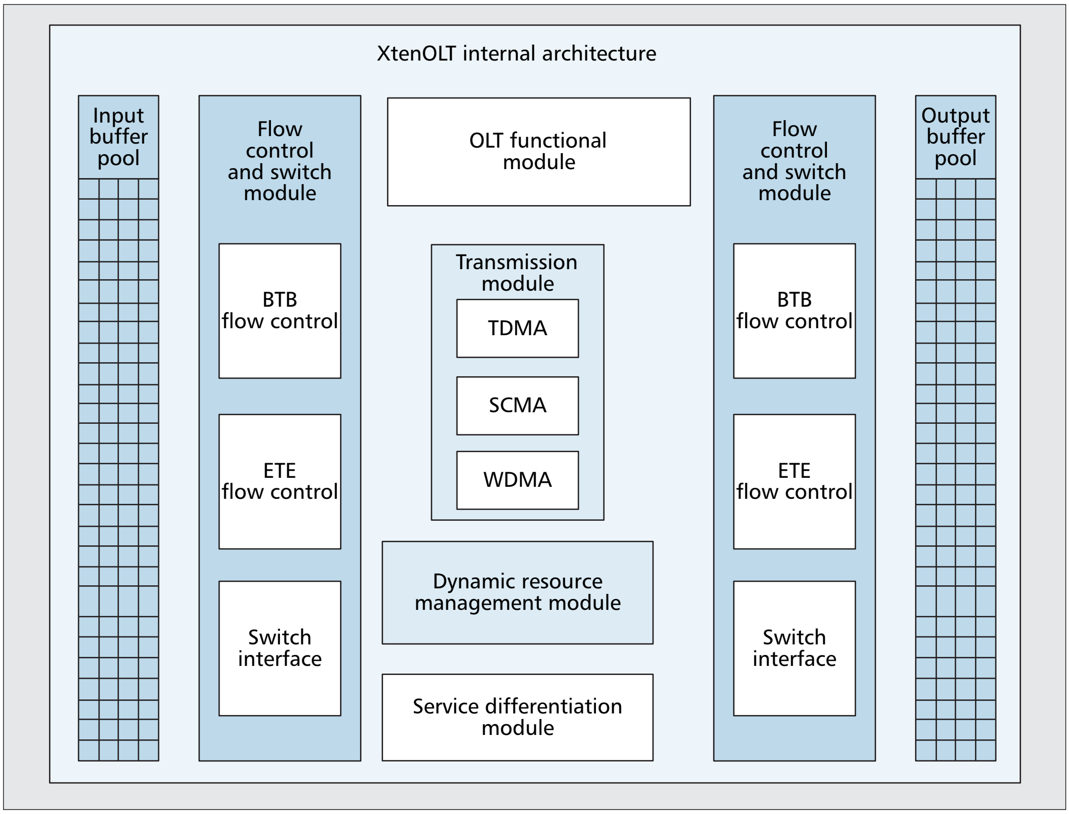


Figure 3. *Internal architecture of XtenOLT.*

less of the traffic. Under such a scheme, a fixed threshold is set for each SAN and the arriving packets are discarded if the queue length is beyond the prescribed threshold. It has been shown that the fixed buffering scheme leads to poor performance [14]. Furthermore, the SAN traffic follows a self-similar pattern with bursty characteristics [15–17]. The fixed buffering scheme also ignores the bursty nature of SAN traffic by preventing the heavily loaded traffic from accessing the free space in the shared buffer pool, thus leading to overall inefficiency.

The linear proportional scheme is another commonly used buffer management technique [18]. Under such a scheme, the number of buffers granted to each SAN is linearly proportional to the request in the previous time interval. Because the linear proportionality scheme favors SANs with large buffer requirements, it causes problems of unfairness and low utility [19].

To overcome the problems of existing buffer management schemes, we propose an algorithm called Tetris, which allocates the buffers to the SANs dynamically. The basic idea of the Tetris algorithm is to grant each SAN the number of buffers equivalent to the minimal request among the SANs. In each time cycle, the Tetris algorithm may take several rounds to complete until all available buffers are successfully granted (see Appendix 1 for details).

Figure 4 shows a simple illustration of this algorithm. Assume there are four SANs requesting buffers in time interval *n* and their requests are represented by four columns. In round 1, the granted buffers to each SAN are equal to the minimal request, which is request 2. Thereafter, the granted buffers (i.e., request 2) are then chopped from each request, as illustrated by the dashed line in round 1. Request 2 is therefore 100 percent fulfilled in round 1. In round 2, there are only three requests, requests 1, 3 and 4. Similarly, the granted buffers to each SAN are equal to the minimal request, which is request 3.

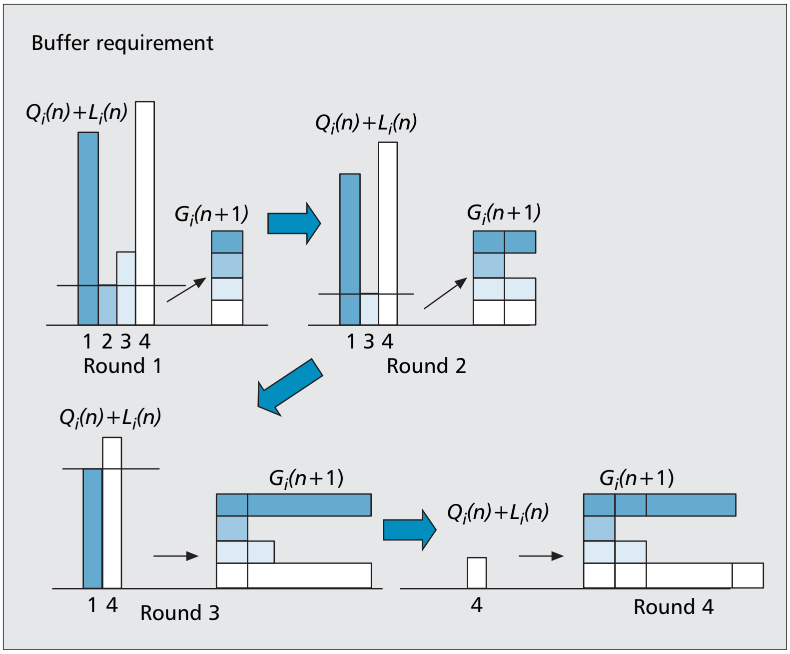
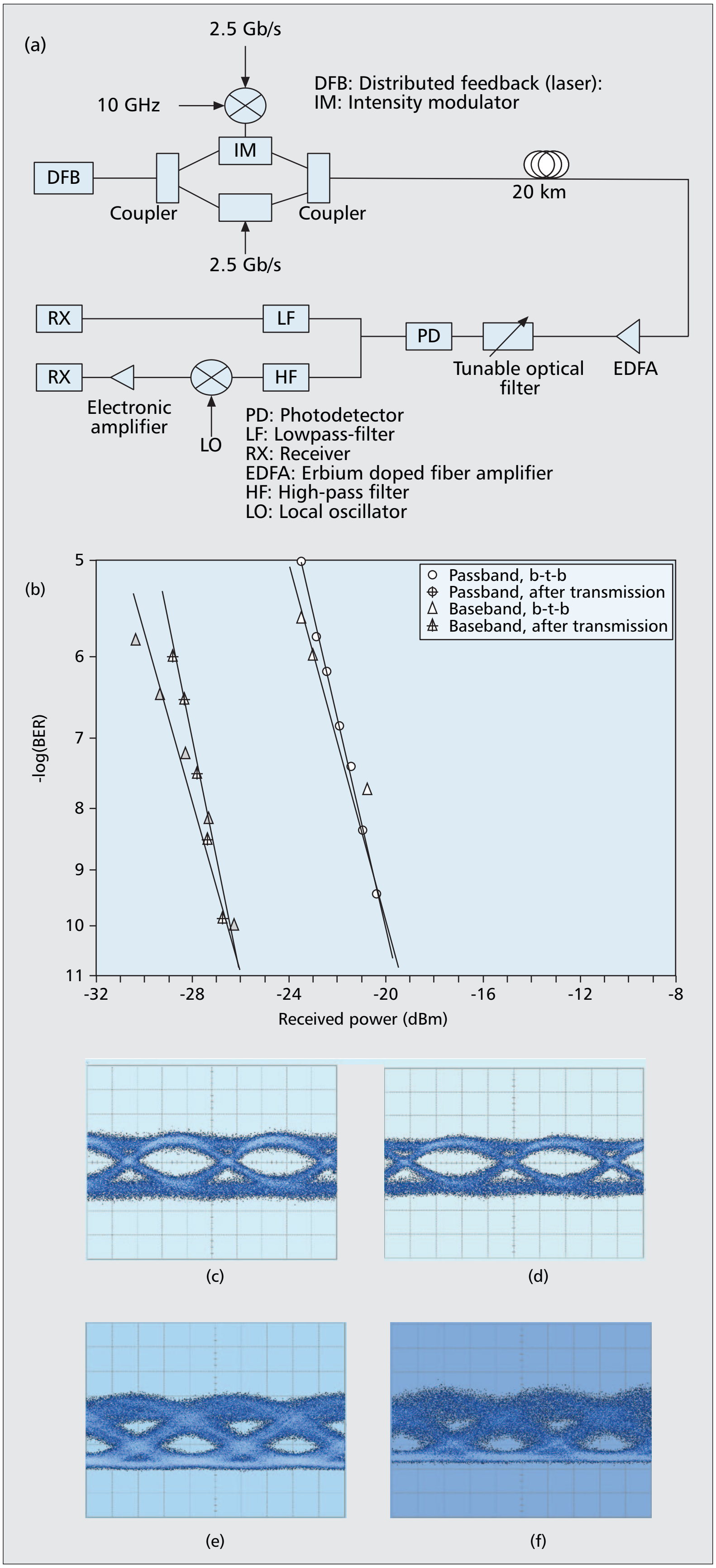


Figure 4. *The Tetris scheme.*

The granted buffers (i.e., request 3) are then chopped from each request. Therefore, request 3 is fulfilled. By following the same process, round 3 fulfills request 1, and round 4 fulfills request 4.



■ **Figure 5.** *The S-PON experiment: a) setup, b) bit error rate (BER) curves, c) eye diagram for 2.5 Gb/s data signal before transmission, d) eye diagram for 2.5 Gb/s data signal after 20-km transmission, e) eye diagram for 2.5 Gb/s storage signal before transmission, f) eye diagram for 2.5-Gb/s storage signal after 20-km transmission.*

EXPERIMENTS AND SIMULATIONS

We performed experiments to evaluate the various physical layer transmission techniques. We also simulated several buffer management schemes to evaluate their performance. This section summarizes our results.

PHYSICAL LAYER SIMULATION

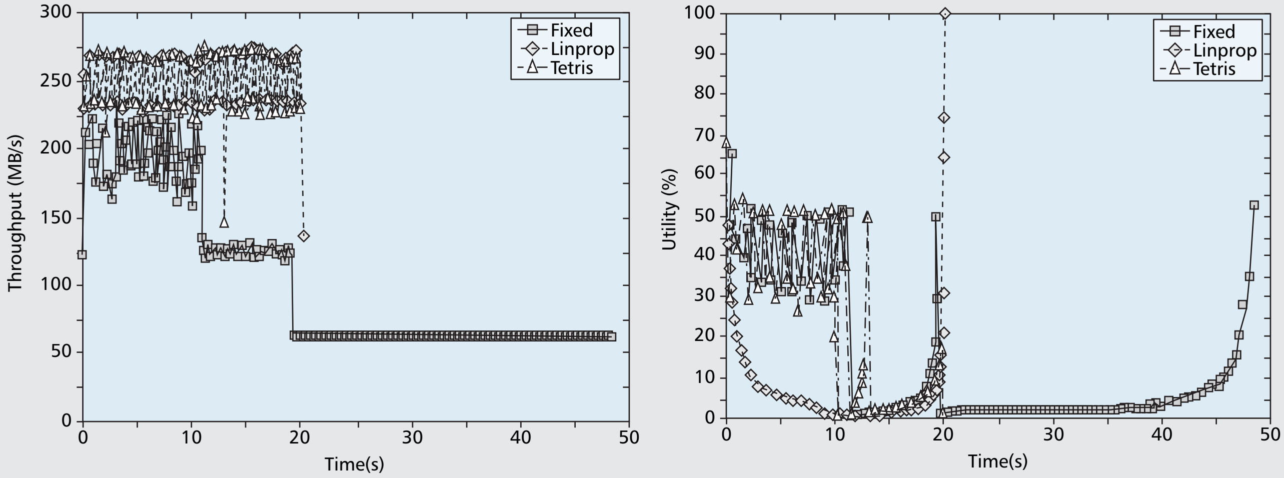
Our S-PON experimental setup is shown in Fig. 5a. One wavelength is employed to carry the 2.5-Gb/s data signal (to emulate LAN traffic) and the 2.5-Gb/s storage signal (to emulate SAN traffic) in the upstream direction. The storage signal is mixed with a 10-GHz carrier before they are used to drive the modulator to generate the sub-carrier multiplexing signal. One photodetector (PD) is employed after an erbium doped fiber amplifier (EDFA) and a tunable optical filter to receive both data and storage signals. A low-pass filter is used for data signal receiving. To receive the storage signal, a high-pass filter, a 10-GHz mixer and an electronic amplifier are employed.

The BER measurement results are shown in Fig. 5b. We observe that both data and storage signals show small power penalties after a 20-km transmission. Figure 5c–5f show the eye diagrams measured for the signals. The experimental results demonstrate that by employing the low-cost electric filters, the baseband data signal and the modulated storage signal are correctly detected simultaneously at the OLT side, and, thus, the extended storage service can be provided by using the widely deployed PON access network architecture.

BUFFER MANAGEMENT SIMULATIONS

We simulated our Tetris buffer management algorithm to evaluate its performance. The experimental S-PON connects two sites about 5000 km apart, one in New York City and one in San Francisco. Each site consists of four SANs, which are connected to the XtenOLT node through the PON architecture. In the simulation, each SAN carries its own local traffic, which are 100 Mb/s, 500 Mb/s, 1 Gb/s and 2.5 Gb/s, respectively. All the traffic patterns are simulated by using a self-similar traffic generator, with the Hurst parameter H set to 0.8. This parameter, with a range of 0.5-1, is a measure of the self-similarity of a time series of traffic. The generated traffic exhibits higher self-similarity when H is closer to the value of 1, and lower self-similarity when H is closer to 0.5 [15]. The long-distance link capacity is set to be 2.5 Gb/s (i.e., 320 MB/s) and 4800 buffers are configured in XtenOLT. We also compared the performance of the Tetris scheme with two other buffer-management schemes, namely, the fixed and linear proportional schemes. The simulation results are shown in Fig. 6a and Fig. 6b.

Figure 6a shows the instantaneous throughput comparison of the three algorithms. It shows that both the Tetris and the linear proportional



(a) Comparison of instantaneous throughput of three algorithms.

(b) Comparison of overall utility of three algorithms.

Figure 6. *The performance comparison of fixed, proportional and Tetris schemes.*

scheme achieve around 250 MB/s, which is 78 percent of the link capacity. Since both schemes make full use of free buffer space, the instantaneous aggregated throughput of the two algorithms overlap most of the time, so it is difficult to distinguish the difference between the two in the throughput graph (one color obscures the other). On the other hand, the fixed scheme achieves an average throughput of 100 MB/s, which is 31 percent of the link capacity. The fixed scheme achieves low throughput because it ignores the bursty nature of SAN traffic and prevents the heavily loaded traffic from accessing the free space in the shared buffer pool. Fig. 6a also shows that the fixed scheme may cause severe congestion when the queue length reaches a certain threshold level (i.e., the throughput of the fixed scheme in the12th and the 19th seconds), which also explains why the fixed scheme takes longer than Tetris and linear proportional schemes to transmit the SAN traffic in Fig. 6a.

The instantaneous measurement of the overall utility of the three algorithms in the simulation are compared in Fig. 6b. Here, the overall utility is defined as the request-to-grant ratio in each time cycle, which is a measure of the degree of customer satisfaction. (see Appendix 2 for a detailed definition of overall utility.) Figure 6b shows that the Tetris, fixed and linear proportional schemes achieve 23 percent, 20 percent and 5 percent average overall utility, respectively. The linear proportional scheme has the lowest overall utility because when heavily-loaded SANs constantly request large numbers of buffers, the linear proportional scheme has no way to prevent the heavy traffic from monopolizing the buffer pool. Consequently, the lightly loaded traffic begins to starve, leading to low utility. The Tetris scheme, on the other hand, always satisfies the SAN with the minimal request, and thus prevents the heavily loaded traffic from monopolizing the buffer pool. In this way, the overall utility is greatly enhanced, as shown in Fig. 6b.

In the simulation, the fixed scheme provides higher utility than the linear proportional scheme, because the low-traffic SAN requests are always fully satisfied by the buffers allotted to the SAN. On the other hand, the linear proportional scheme provides better throughput than the fixed scheme, because underutilized buffers do not remain idle, and instead are used to satisfy requests from other SANs. The Tetris scheme exhibits the higher throughput of the linear proportional scheme while also exhibiting the higher utility of the fixed scheme.

S-PON IMPLEMENTATION

The implementation of S-PON is based on the regular PON infrastructure with minor modifications, such as the aforementioned changes in the remote node for the WDMA scenario, extra transceivers in the SCMA/WDMA scenario, and the upgraded OLT device (i.e., XtenOLT). To further demonstrate the possible practical application of the S-PON architecture, we show an S-PON implementation exemplified in this section.

In Fig. 7, SAN1 and SAN2 are the two primary sites in San Francisco and Dallas, respectively, while SAN3 is the storage center in Atlanta for routine backup and disaster recovery. S-PON enables the provisioning of storage service nationwide, separating the primary site and backup center by hundreds or thousands of miles. All three sites are deployed with NEC "S" serial SAN products and SpectralWave GPON products. In addition, the XtenOLT is implemented in each site as an enhanced OLT with storage provisioning capacity.

Different transmission technologies are employed in the S-PON architecture of each site. TDMA is utilized at SAN1 in San Francisco, which enables SAN1 to share the 1 Gb/s transmission bandwidth with other LAN traffic through its own time slots. Therefore, the remote node and transceivers in this S-PON remain the same as in the regular PONs. On the other hand, the storage data of SAN2 in Dallas is modulated by sub-carrier signals to XtenOLT, which provides SAN2 with a 2.5-Gb/s transmission rate

The implementation of S-PON is based on the regular PON infrastructure with minor modifications, such as the aforementioned changes in the remote node for the WDMA scenario, extra transceivers in the SCMA/WDMA scenario, and the upgraded OLT device.

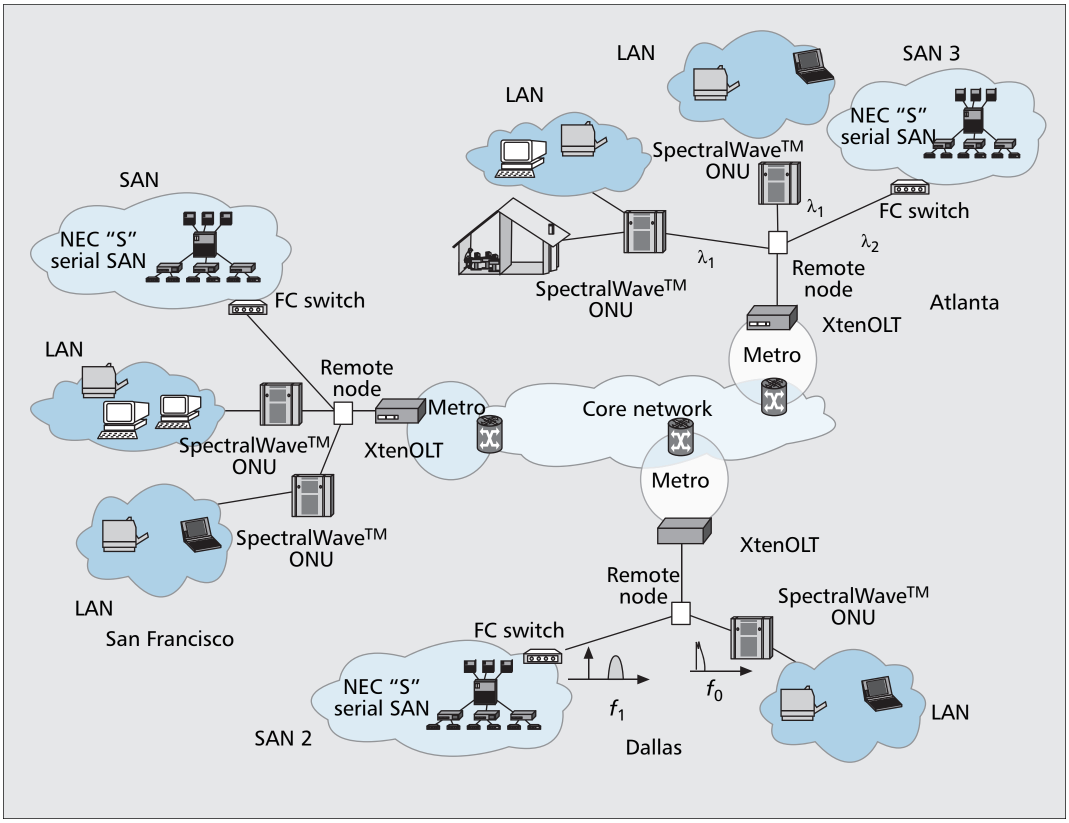


Figure 7. *The implementation of S-PON.*

through its sub-band f_1. To support the SCMA transmission, a pair of transceivers that support SCM are configured in the XtenOLT and the FC switch of SAN2. Lastly, as the data backup center, SAN3 in Atlanta has the highest bandwidth and security requirements, and is thus allocated one dedicated wavelength, λ_2, for storage data transmission, which provides SAN3 with up to a 10-Gb/s transmission rate. To fulfill the transmission requirements under this scenario, the remote node of the PON is upgraded, as shown in Fig. 2c, to support WDMA, and WDM transceivers are implemented in XtenOLT and in the FC switch of SAN3.

With S-PON, businesses in San Francisco and Dallas are able to locate their remote mirroring sites in Atlanta, which is far away from their primary sites, to divert the risk of all sites being hit by the same disaster. S-PON offers a high transfer speed comparable to conventional long-haul technologies, thereby enabling businesses in San Francisco and Dallas to recover all their critical data in a short time. In addition, S-PON can transfer data without critical loss and hence improves the efficiency of data consolidation. This would help businesses consolidate their data effectively, and make it readily available to their subunits in real time.

CONCLUSIONS

In this article, we have proposed a new solution, S-PON, to tackle the challenges of extending the SAN into the long-haul network. S-PON adopts the P2MP architecture and leverages the existing PON infrastructure to solve the key issues of scalability and cost. Furthermore, three transmission technologies, TDMA, SCMA and WDMA, were investigated to tackle the legacy transmission bottleneck. We have also proposed a new device to deliver storage service over PON and to solve the low throughput of conventional SAN extension. A new buffer management scheme called Tetris is implemented in XtenOLT. Our experiments and simulations have shown that, in the physical layer, the proposed S-PON transmission technologies successfully deliver SAN traffic to the long-haul at the rate of 2.5 Gb/s; in the network layer, XtenOLT with the Tetris buffer-management scheme dramatically enhances the deliverable throughput and overall utility over a 5000-km distance.

APPENDIX 1: THE MATHEMATICAL DERIVATIONS OF THE TETRIS ALGORITHM

Let's define the request $R_i(n)$ the sum of the queued length $Q_i(n)$ at the beginning of time interval n and the arrived data length $L_i(n)$ at the end of time interval n. In the linear proportional scheme, the granted buffers to SANi is calculated by

$$G_i(n+1) = M \times \frac{Q_i(n)+L_i(n)}{\sum_{i=1}^{k}[Q_i(n)+L_i(n)]}, i=1,\ldots,k \qquad (1)$$

where M is the total available buffers in time interval $n+1$ and k is the number of SANs connected to the switch. Since the linear proportional scheme favors SANs with large buffer requirements, it causes the problems of unfairness and low utility [19].

To overcome the problems of existing buffer management schemes, the Tetris algorithm is proposed to allocate the buffers to the SANs more efficiently. The Tetris algorithm is described as follows: assume there are k SANs requesting buffers in time interval n, and their requests $R_i(n)$ are also defined as the sum of the queued length $Q_i(n)$ at the beginning of time interval n and the arrived data length $L_i(n)$ at the end of time interval n. In round 1, grant $G_i(n+1)$ to each SAN is equal to the minimal request of k SANs, say $R_1^{\min}$. Without loss of generality, we assume SANi has the minimal requirements in round i, and thus,

$$\begin{aligned} G_1(n+1) &= G_2(n+1) = \ldots = G_k(n+1) \\ &= \min\{Q_i(n)+L_i(n), \\ &\qquad i = 1, 2, \ldots, k\} = R_1^{\min} \end{aligned} \tag{2}$$

After round 1, there are $k - 1$ requests left with the value of $Q_i(n+1) + L_i(n) - R_1^{\min}$, $i = 1, 2, \ldots, k - 1$. Assume the minimal value of the left request is $R_2^{\min}$, we then have the grants in round 2 as,

$$\begin{cases} G_1(n+1) = R_1^{\min} \\ G_2(n+1) = G_3(n+1) = \ldots = G_k(n+1) \\ \quad = R_1^{\min} + \min\left\{Q_i(n) + L_i(n) - R_1^{\min}, i = 2, 3, \ldots k\right\} \\ \quad = R_1^{\min} + R_2^{\min} \end{cases} \tag{3}$$

As long as the available buffer M is larger than $k \times \min\{Qi + L_i(n), i = 1, 2, \ldots k$, Tetris continues to allocate buffer until the last request is granted, i.e.,

$$\begin{cases} G_1(n+1) = R_1^{\min} \\ G_2(n+1) = G_3(n+1) = R_1^{\min} + R_2^{\min} \\ \qquad \ldots \\ G_k(n+1) = R_1^{\min} + R_2^{\min} + \ldots + R_{k-1}^{\min} \\ \quad + \left[Q_k(n) + L_k(n) - R_1^{\min} - R_2^{\min} - \ldots - R_{k-1}^{\min}\right] \\ \quad = Q_k(n) + L_k(n) \end{cases} \tag{4}$$

A critical condition for deploying the Tetris algorithm is to ensure that $M > k \times \min\{Q_i(n) + L_i(n)$, i = 1, 2, … $k\}$ always holds. However, it is possible that the available buffers are not larger than $k \times \min\{Q_i(n) + L_i(n), i = 1, 2, \ldots k\}$ after several rounds. In this case, the leftover available buffers will be distributed to each SAN following a certain remainder distribution policy (RDP).

APPENDIX 2: THE DEFINITION OF OVERALL UTILITY

The utility of SANi in time interval n is defined as

$$u_i(n) = \frac{G_i(n)}{R_i(n)} \tag{5}$$

where $G_i(n)$ is the granted buffer, and $R_i(n)$ is the sum of the queued length $Q_i(n)$ at the beginning of time interval n and the arrived data length $L_i(n)$ at the end of time interval n. $u_i(n+1)$ essentially represents how much of the ratio of the requests are granted in each time interval. Assuming that there are k SANs, the overall utility in the time interval n is defined as

$$u_i^{overall}(n) = \frac{\sum_{i=1}^{k} u_i(n)}{k} \tag{6}$$

S-PON can transfer data without critical loss and hence improves the efficiency of data consolidation. This would help businesses consolidate their data effectively, and make it readily available to their subunits in real time.

REFERENCES

[1] T. Clark, *Designing Storage Area Networks: A Practical Reference for Implementing Fibre Channel and IP SANs*, Addison-Wesley), 2003.

[2] R. Telikepalli, T. Drwiega, and J. Yan. "Storage Area Network Extension Solutions and Their Performance Assessment," *IEEE Commun. Mag.*, Apr. 2004, pp. 56–63.

[3] "Improving Business Resilience with Data Replication and Storage Area Network Extension Technology," Apr. 2007, Hitachi white paper, http://www.hds.com/assets/pdf/wp-improving-business-resilience-with-data-replication-and-storage-area-network-extension-technology.pdf

[4] X. Qiu *et al.*, "Reliability and Availability Assessment of Storage Area Network Extension Solutions," *IEEE Commun. Mag.*, Mar. 2005, pp. 80–85.

[5] "Thinking Storage Network Extension," Ciena Inc. white paper, http://www.fcw.com/vendorsolutions/THINK_Ciena-2-final2.pdf

[6] T. Kessler, A. Ehrhardt, and A. Weber, "SAN Extension Testbed Within the Global Seamless Networks Demonstrator," *Proc. 2nd Int'l. Conf. Testbeds and Research Infrastructures for the Development of Networks and Communities*, Mar. 1–3, 2006, pp. 209–15.

[7] A. Helland, "The Economics of Large-Scale Data Transfer: the Benefits of Fibre Channel and SONET for High-Performance, Cost-Effective Data Transport Through WAN Storage Networking," *Computer Technology Review*, Dec. 2002.

[8] "EMC and Cisco Metro Optical Storage Networking Solutions," Cisco Systems Inc. White paper, 2001. http://whitepapers.zdnet.co.uk/0,39025945,60031308p-39000566q,00.htm

[9] T. Clark, *IP SANs: A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks*, (Pearson Education), 2002.

[10] J. Satran *et al.*, "Internet Small Computer Systems Interface (iSCSI)," IETF draft, 2004, http://www.ietf.org/rfc/rfc3720.txt

[11] M. Rajagopal, E. Rodriguez, and R. Weber, "Fiber Channel Over TCP/IP," IETF draft, 2004, http://tools.ietf.org/html/rfc3821

[12] C. Monia *et al.*, "iFCP -A Protocol for Internet Fibre Channel Storage Networking," IETF draft, 2005, http://www.ietf.org/rfc/rfc4172.txt

[13] V. Ruesterholz, "Verizon FiOS: Executing Strategy and Revenue Opportunities," http://investor.verizon.com/news/20060927/20060927.pdf

[14] A. K. Choudhury and E. L. Hahne, "Dynamic Queue Length Threshold for Shared-Memory Packet Switches," *IEEE/ACM Trans. Net.*, vol. 6, no. 2, Apr. 1998, pp. 130–40.

[15] X. Molero, F. Silla, V. Santonja and J. Duato, "Modeling and Simulation of Storage Area Networks," *Proc. 8th Int'l. Symp. Modeling, Analysis and Simulation of Computer and Telecommun. Systems*, Aug. 29–Sept. 1, 2000, pp. 307–14.

[16] M.E. Gomez and V. Santonja, "Analysis of Self-Similarity in I/O Workload Using Structural Modeling," *Proc. 7th Int'l. Symp. IEEE Analysis and Simulation of Computer and Telecommun. Systems*, Oct. 24–28, 1999.

[17] M. E. Gomez and V. Santonja, "Self-Similarity in I/O Workload: Analysis and Modeling," *Proc. IEEE Workload Characterization: Methodology and Case Studies*, Nov. 1998, pp. 97–104.

[18] Y. Afek *et al.*, "Dynamic Bandwidth Allocation Policies," *Proc. IEEE INFOCOM'96*, San Francisco.

[19] S. Biswas and R. Izmailov, “Design of a Fair Bandwidth Allocation Policy for VBR Traffic in ATM Networks,” *IEEE/ACM Trans. Net.*, vol. 8, Issue 2, April 2000, pp. 212–23.

BIOGRAPHIES

SI YIN [S] (sy44@njit.edu) received the B.E. degree from the University of Electronic Science and Technology of China, Chengdu, in 2002, and the M.E. (by research) degree from National University of Singapore, in 2005, both in electrical engineering. He is working toward to his Ph.D. degree in the department of electrical and computer engineering, New Jersey Institute of Technology (NJIT), Newark. His research interests are in the areas of broadband access networks (xPON, WiMAX, and optical wireless integration), control theory in network resource management, and network storage management. He was the recipient of NEC Laboratories America SEEDS Award (2007), first place in the IEEE North Jersey Student Presentation Contest (2007), and Singapore A*STAR Scholarship Award (2003 and 2004).

YUANQIU LUO [M] (yluo@nec-labs.edu) received her Ph.D. degree in electrical engineering from New Jersey Institute of Technology (NJIT). She is currently with NEC Laboratories America in Princeton, NJ. Her research interests are in the areas of broadband access networks, network survivability, network modeling, and integrated optical/wireless networks. Her awards and recognitions include the NEC Laboratories America SEEDS Award (2007), nomination for the New Faces of Engineering in conjunction with National Engineers Week 2006, New Jersey Inventors Hall of Fame Graduate Student Award (2005), NJIT Hashimoto Fellowship (2005), and first place in the IEEE North Jersey Section Student Presentation Contest (2004). She received both her B.S. degree in electronics and information Systems and her M.E. degree in electrical engineering from Shandong University, China.

LEI ZONG [M] (lanezong@nec-labs.com) is a research staff member in the optical networking department of NEC Laboratories America. He received bachelor’s, master’s and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, in 1991, 1998 and 2001, respectively. He joined NEC Laboratories America in September 2001. Dr. Zong has published more than 30 papers in peer-reviewed journals and conference proceedings. He also holds 6 U.S. patents and pending patents. His research interests include optical switching, high-speed transmission and internetworking.

STEPHEN RAGO [M] (sar@nec-labs.com) is a research staff member in the Robust and Secure Systems Group at NEC Laboratories America. He received a bachelor’s degree in electrical engineering and a master’s degree in computer science, both from Steven’s Institute of Technology, in 1984. He holds two patents and is the author of UNIX System V Network Programming (Addison-Wesley, 1993) and co-author of the second edition of Advanced Programming in the UNIX Environment (Addison-Wesley, 2005). His research interests include networking, operating systems and storage systems.

JIANJUN YU [SM] received a B.S. degree in optics from Xiangtan University, China, in September 1990, and M.E. and Ph.D. degrees in optical communications from the Beijing University of Posts and Telecommunications in 1996 and 1999, respectively. From June 1999 to January 2001, he worked at the Research Center COM, Technical University of Denmark, in Lyngby, as an assistant research professor. From February 2001 to December 2002, he was a member of the technical staff at Bell Labs’ Lucent Technologies and Agere Systems, in New Jersey. He joined the Georgia Institute of Technology, Atlanta, in January 2003, where he was on the research faculty and served as director of the Optical Network Laboratory. He is currently a member of technical staff of NEC Laboratories America. He is also an adjunct professor at the Georgia Institute of Technology and the Beijing University of Posts and Telecommunications. His current research interests include 100-Gb/s high-speed transmission systems, new modulation-format techniques, radio-over-fiber systems and networks, wavelength-division-multiplexing passive optical networks and optical-label switching in optical networks. He has co-authored more than 250 peer-reviewed journal and conference papers, and is the first author of more than 100 of them. He holds two U.S. patents and five pending patents. He is a Senior Member of the IEEE Lasers and Electro-Optics Society (LEOS). He is serving as a guest editor for a special issue, “Convergence of Optical and Wireless Networks,” for IEEE/OSA Journal of Lightwave Technology. He serves as a technical committee member of the IEEE LEOS annual meetings from 2005-7, as well as a reviewer of IEEE Photonics Technology Letters, Journal of Lightwave Technology, IEEE J. of Quantum Electronics, Optics Letters, Electronics Letters, Optics Express, Optical Engineering, and Optics Communications.

NIRWAN ANSARI [SM] received a B.S.E.E. degree (summa cum laude) from New Jersey Institute of Technology in 1982, an M.S.E.E. degree from the University of Michigan, Ann Arbor, IN, in 1983, and a Ph.D. degree from Purdue University, West Lafayette, IN, in 1988. He joined NJIT’s department of electrical and computer engineering as an assistant professor in 1988, and has been a full professor since 1997. He has also assumed various administrative positions. He authored Computational Intelligence for Optimization (Springer, 1997, translated into Chinese in 2000) with E.S.H. Hou, and edited Neural Networks in Telecommunications (Springer, 1994) with B. Yuhas. His current research focuses on various aspects of broadband networks and multimedia communications. He has also written over 300 technical papers, over one third of which are published in refereed journals and magazines. He is a senior technical editor of IEEE Communications Magazine, and also serves on the editorial board of Computer Communications, the ETRI Journal, and the Journal of Computing and Information Technology. He was the founding general chair of the First IEEE International Conference on Information Technology: Research and Education (ITRE2003), was instrumental, while serving as its chapter chair, in rejuvenating the North Jersey Chapter of the IEEE Communications Society, which received the 1996 Chapter of the Year Award and a 2003 Chapter Achievement Award, served as chair of the IEEE North Jersey Section and on the IEEE Region 1 Board of Governors during 2001–2002, and has been serving on various IEEE committees, including as chair of the IEEE COMSOC Technical Committee on Ad Hoc and Sensor Networks, and general/technical program chair or vice-chair of several conferences/symposia. His awards and recognitions include the NJIT Excellence in Teaching Award in Graduate Instruction (1998), IEEE Region 1 Award (1999), IEEE Leadership Award (2007, from the IEEE Princeton/Central Jersey Section), and designation as an IEEE Communications Society Distinguished Lecturer.