Encyclopedia of Internet Technologies and Applications

Mario Freire
University of Beira Interior, Portugal

Manuela Pereira
University of Beira Interior, Portugal

Information Science
REFERENCE

INFORMATION SCIENCE REFERENCE
Hershey · New York

Acquisitions Editor: Kristin Klinger
Development Editor: Kristin Roth
Senior Managing Editor: Jennifer Neidig
Managing Editor: Sara Reed
Copy Editor: Larissa Vinci and Mike Goldberg
Typesetter: Amanda Appicello and Jeffrey Ash
Cover Design: Lisa Tosheff
Printed at: Yurchak Printing Inc.

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com/reference

and in the United Kingdom by
Information Science Reference (an imprint of IGI Global)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: http://www.eurospanonline.com

Library of Congress Cataloging-in-Publication Data

Encyclopedia of Internet technologies and applications / Mario Freire and Manuela Pereira, editors.
p. cm.
Summary: "This book is the single source for information on the world's greatest network, and provides a wealth of information for the average Internet consumer, as well as for experts in the field of networking and Internet technologies. It provides the most thorough examination of Internet technologies and applications for researchers in a variety of related fields"--Provided by publisher.
Includes bibliographical references and index.
ISBN 978-1-59140-993-9 (hardcover) -- ISBN 978-1-59140-994-6 (ebook)
1. Internet--Encyclopedias. I. Freire, Mário Marques, 1969- II. Pereira, Manuela.
TK5105.875.I57E476 2007
004.67'803--dc22
2007024949

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Optical Burst Switching

Kyriakos Vlachos
University of Patras, Greece

INTRODUCTION

Switching in core optical networks is currently being performed using high-speed electronic or all-optical circuit switches. Switching with high-speed electronics requires optical-to-electronic (O/E) conversion of the data stream, making the switch a potential bottleneck of the network: any effort (including parallelization) for electronics to approach the optical speeds seems to be already reaching its practical limits. Furthermore, the store-and-forward approach of packet-switching does not seem suitable for all-optical implementation due to the lack of practical optical random-access-memories to buffer and resolve contentions. Circuit switching on the other hand, involves a pre-transmission delay for call setup and requires the aggregation of microflows into circuits, sacrificing the granularity and the control over individual flows, and is inefficient for bursty traffic. Optical burst switching (OBS) has been proposed by Qiao and Yoo (1999) to combine the advantages of both packet and circuit switching and is considered a promising technology for the next generation optical internet.

BACKGROUND

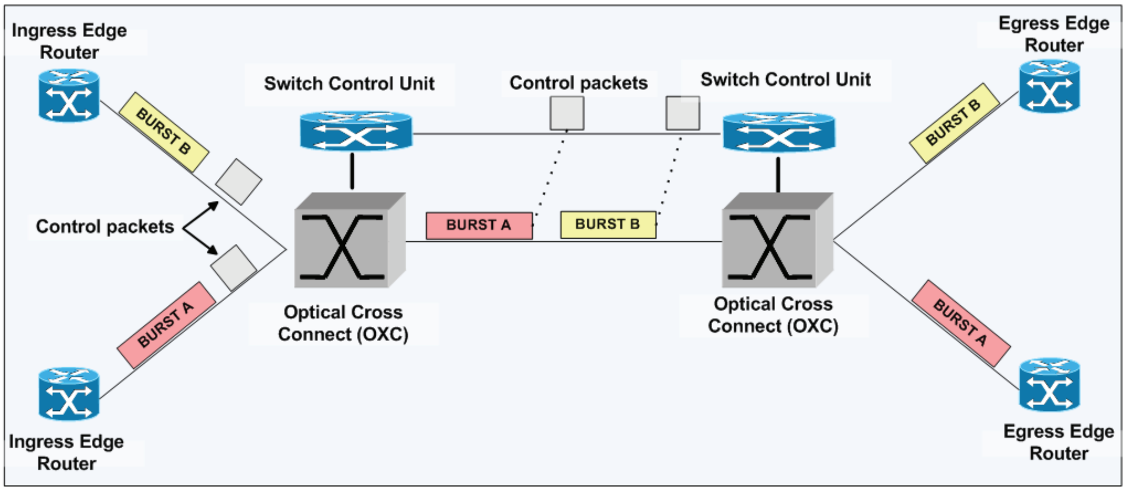
An OBS network consists of a set of optical core routers and edge routers. The basic idea is to amortize switching and protocol processing overhead over a larger amount of payload data, and thus enable affordable and less intelligent switches to be employed. An optical burst is constructed at the network edge by aggregating a number of variable size packets of different protocols (IP packets, ATM cells…). The burst is then transmitted in the network and is forwarded transparently (all-optically) to its end destination router. OBS builds upon the tell-and-go protocol (TAG) class developed by Hudek and Muder (1995). According to this class of protocols, a virtual circuit is set up on the fly for a burst of packets that can go through intermediate switches without buffering and without waiting for acknowledgement of the allocation of the circuit. To this end, a control packet that carries routing and overhead information is transmitted prior to burst transmission. It must be noted here that other signaling schemes exists as well, but OBS was initially designed based on the "*one-way*" reservation concept.

Unlike packet switching, an optical burst can size from a few bytes to multi-giga byte packets, while unlike circuit switching reservation duration is known in advance via the communication to all nodes of the control packet. Thus, reservation is so called *delayed reservation* in the sense that bandwidth is reserved only for time it is actually needed; that is for a time equal to the burst size. Figure 1 illustrates the transmission of two bursts and their associated control packets.

In particular, Figure 1 shows two edge routers that transmit two bursts of data heading for different egress routers. The control packets carry overhead information for the bursts (signaling and routing overhead) and these are communicated to all the *switch control units* of the core routers. The latter process the control packet and signal the optical cross connects (OXCs) to configure their states. In the case of contention, one of the burst is dropped. The control packet precedes the data burst by a time offset to compensate for its processing delay and thus avert the bursts to surpass it. Upon the reception of the burst at the egress router, the latter disassembles the burst and forwards its contents to their end-users.

The success of OBS technology relies on its small control overhead for a large amount of payload data. Data are optically switched in the core and thus there is no need for high-speed electronics, while control packet can be at a significant lower rate. OBS shifts

Figure 1. Optical burst switching network architecture. An edge router is commissioned to assemble the bursts, while the core routers transparently (all-optically) forward the bursts to their end destinations. Control packets carry signaling and routing overhead information.



its complexity at the network edge, where the bursts have to be constructed. Figure 2 illustrates a top level architecture of an OBS edge router. The edge router maintains a separate queue per destination, where a separate burst scheduler is responsible for constructing the bursts from packets coming from the access network. Upon completion of the so called *burst assembly* process, a *link scheduler* schedules the burst for transmission. The link scheduler is responsible for wavelength assignment and routing table look up.

RESEARCH AND DEVELOPMENT ISSUES OF OPTICAL BURST SWITCHING

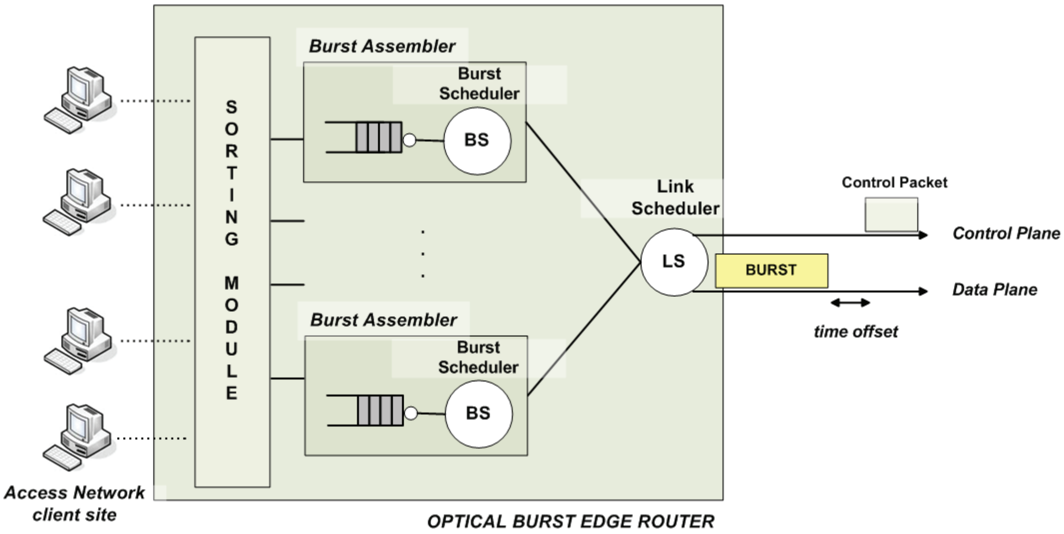
OBS became a hot research topic soon after its introduction. Several issues are under study and some other have been successfully addressed. In particular, issues of pivotal importance include the development of burst assembly algorithms, efficient signaling protocols, contention resolution schemes as well as quality of service provisioning mechanisms. In what follows, we provide an overview of current development and research trends in the aforementioned key areas.

Burst Assembly Algorithms

Burst assembly defines how packets are assembled to form a burst. The burst assembly process starts with the arrival of a packet from a high layer application and continues until a predefined criterion is met. The criterion defines when the burst assembly process stops and the newly generated burst is sent into the network. The assembly process affects, through the assembly criterion, the burst size, the burst duration, as well as the packet delay at the edge router. The packet delay is defined as the time that the packets must wait before burst transmission.

Two distinct burst assembly algorithms have been proposed in the literature: the *timer-based* and the *threshold-based* method. In the timer-based method, also denoted as T_{MAX} in the literature, (Callegati & Tamil, 2000), a time counter starts any time a packet arrives and when the timer reaches a time threshold (T_{MAX}), a burst is created; the timer is then reset to 0 and it remains so until the next packet arrival at the queue. Hence, the ingress router generates periodically bursts, every T_{MAX} time, independently of the yielding burst size. In the second scheme, (Vokkrane, Haridoss, & Jue, 2002), a threshold is used to determine the end of the assembly process. In most cases the threshold used is the burst length denoted in the literature as B_{MAX}.

Figure 2. Optical burst switching edge router architecture. A burst scheduler is commissioned to assemble burst from packets with the same end destination, while a link scheduler is commissioned to assign wavelengths and schedule bursts for transmission



In that case, bursts are thought as containers of a fixed size B_{MAX}, and as soon as the container is completely filled with data, the burst is transmitted.

Apart from the aforementioned assembly schemes, other more complex schemes have been also proposed, which are usually a combination of the *timer -based*, and the *threshold-based* methods. For example the *min-burst length-max-assembly-period (MBMAP)* algorithm, (Cao, Chen, & Qiao, 2002), sends out a burst when its size exceeds a minimum burst length (MBL) or when the assembly period times out. However, all of the above burst assembly criteria do not take into account traffic situation so as to adapt the burst assembly process accordingly. This is very important for higher layer protocols such as TCP, because it limits its effective throughput. To this end, adaptive burst assembly schemes have been also proposed as for example the *adaptive-assembly-period* (AAP) algorithm proposed by Cao, Chen, and Qiao (2002). The AAP algorithm dynamically changes the assembly time at the ingress node according to the length of the burst recently sent. Nevertheless, the proper selection of the timer or threshold parameter is important and is still an open issue. For example, the use of a burst-length threshold may result in long assembly times under light loads while the use of a timer-based method may result to diverse sizes of bursts, making scheduling in the core a difficult task. What is important is to minimize loss ratio in the core while predicting the assembly expiration time. This will allow transport protocols to predict future round trip times (RTT) and thus minimize time-outs.

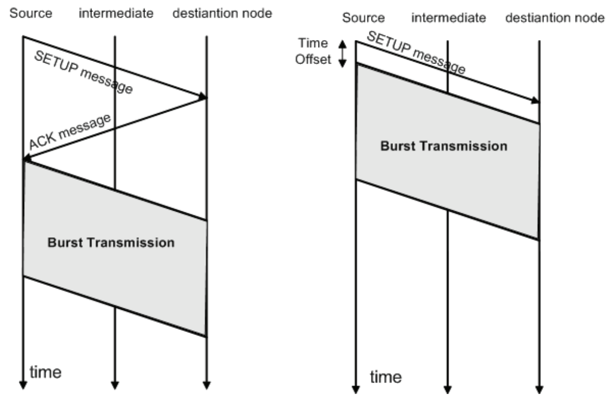
Signaling Protocols

A signaling scheme is required for reserving resources and configuring switches in OBS networks. All the signaling schemes developed can be categorized in two main classes: In two-way reservation schemes (also called *tell-and-wait, TAW*) and one-way reservation schemes (also called *tell-and-go*, *TAG*). Figure 3(a) and (b) illustrate the timing considerations of these two schemes.

In two-way reservation schemes, end-to-end connections are fully established before burst transmission, while resources at intermediate nodes are reserved immediately upon the arrival of the SETUP packet at these nodes. Recent research efforts like the WR-OBS (Dueser & Bayvel, 2002), have shown that such reservation schemes can enable the implementation of a bufferless core network with limited node wavelength conversion capability by moving the processing and buffering functions at the edge.

In one-way reservation schemes, a setup packet is sent in advance over the path, preceding the arrival of the burst by a minor offset. This minimizes the pre-transmission delay, but can result in high burst dropping

Figure 3. Timing considerations of (a) two-way and (b) one-way OBS reservation schemes



(a) *(b)*

probability. A number of one-way reservation schemes have been proposed for OBS networks, including the just-enough-time (JET) (Qiao & Yoo, 1997), the horizon (Turner 1999), the just-in-time (JIT) (Wei & MacFarland Jr., 2000), and the ready-to-go virtual circuit protocol (Varvarigos & Sharma, 1997). The differences among these variances lie mainly in the time instances that determine the allocation and the release of the resources. These can be *implicit* where capacity is freed immediately after burst traversing the OBS node (burst length information is stored in the preceding control packet) or *explicit* with a separate *release* message. Furthermore, for the one-way schemes that employ delayed reservations, sophisticated channel scheduling and void filling algorithms have been proposed to resolve contentions and efficiently utilize the available bandwidth (Xiong, Vandenhoute, & Cankaya, 2000). The use of one-way reservation schemes has introduced a new era for OBS networking and opened new research lines. One-way schemes can guarantee minimum delays at the edge node, on-demand use of bandwidth resources and very low switch setup times. New research lines have emerged on how to avoid burst dropping when contention occurs or how to provision quality of service for bursts carrying packets of a higher priority.

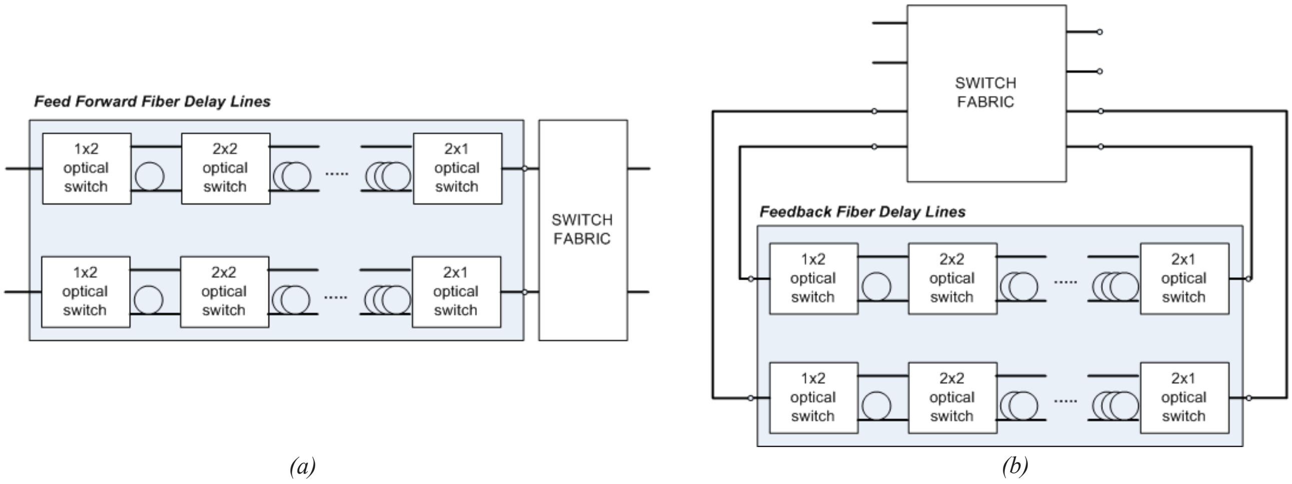
Contention Resolution Schemes

Contention resolution takes place at any of the intermediate nodes upon the reception of two bursts that request the same outgoing link. Since capacity is not reserved the switch has to resolve contention or otherwise drops one of the bursts. Contention resolution can be performed in one of the following domains:

- In the time domain, employing a fiber-delay-line (FDL) structure for buffering/delaying a burst until the contention situation is resolved. In contrast to electronic RAM based buffers, optical FDLs only provide a fixed delay, which must be long enough to temporary store contending bursts. Under the FDL-based buffering scheme, two prime architectures exist; namely the feed forward and the feedback one. Figure 4 illustrates these two FDL architectures. In the feed-forward method, bursts are fed into fiber delay lines of different lengths and when a burst reaches the output it has to be switched out. In the feedback scheme, a burst may re-circulate from the buffer output to the buffer input until contention is resolved and the requested outgoing link is free.

- In the wavelength domain by means of wavelength conversion, where a burst can be sent on a different wavelength channel of the designated output line.
- In the space domain, where a burst is deflected to a different output line of the OBS switch and follows an alternative route than the predetermined one.

Figure 4. (a) Feed-forward and (b) feedback fiber delay lines schemes for contention resolution in OBS networks



(a) *(b)*

Another significant contention resolution strategy relies on the burst segmentation. According to this technique, only the contending part of the two bursts is dropped or routed on another wavelength or deflected (Vokkrane & Jue, 2003).

Most studies on contention resolution in OBS networks focus in the wavelength domain and assume full wavelength conversion at all nodes. For a low to medium load, such an assumption provides a low burst loss ratio. However, for high loads, and in order to achieve a loss ratio of less than 10^{-6}, the number of wavelengths has to be very large and in particular more than >100, making this impractical for a potential deployment.

Similarly, the use of Fiber delay lines is also impractical because of the huge length of fiber needed for an efficient resolution. For example, buffering of a few Mbytes of data requires more than 150km of optical fibre at 10Gb/s. Thus, feedback FDL structures are more attractive. However, the infinite recirculation of the burst data may impair signal quality due to noise accumulation. Further, these schemes increase complexity and size of an OBS node. In particular for an (NxN) switch with L delay lines for burst buffering, instead of (NxN), an $(N+L)$ x $(N+L)$ space switch is required.

To this end, it is not still clear which OBS contention resolution strategy to follow, since all exhibit major drawbacks that lag their commercial implementation. To overcome this drawback, research efforts are focusing to other mechanisms in order to carry high loads on a per wavelength channel. These techniques include intelligent burst scheduling for load balancing as well as hybrid signaling schemes such as the INI scheme (Karanam et al., 2003).

QoS Provision and Service Differentiation Schemes

Quality of service (QoS) in OBS networks has emerged as an extremely important issue in order to guarantee services to end users. The use of one-way reservation schemes and the absence of an efficient contention resolution mechanism urge the QoS support of optical burst switching. In OBS networks, QoS can be provisioned by introducing service differentiation at any point of the network including for example the burst assembly process, the contention resolution process as well as the burst scheduling process. There are two basic models for QoS provision in optical burst switching networks: relative QoS and absolute QoS. In the relative QoS model, the performance of one class is defined with respect to the other classes. For example it is guaranteed that high priority bursts will exhibit a lower edge delay or lower loss ratio as against other classes. However, its absolute performance still depends on the traffic characteristics of the rest classes. On the other hand, absolute QoS model provides an absolute performance metric of quality as for example defining a worst-case loss ratio for bursts belonging to the same class.

Typical schemes of the first category are the *offset-based* approach, (Qiao & Yoo, 2000), the composite-burst assembly (Vokkrane & Jue, 2003), and the preemptive wavelength reservation (Liao & Loi, 2004). According to the first scheme, an extra offset time is given to higher priority bursts to overcome contention in the core, while the *composite-burst assembly* scheme mixes traffic classes during burst assembly and provides QoS via prioritized burst segmentation. Finally, the third scheme associates each class with a predefined usage limit. Bursts that comply with their usage limits preempt others that do not.

Under the second category of absolute QoS provision, we find two techniques; the *probabilistic preemptive* approach (Yang, Jiang, & Jiang, 2003), according to which high-priority bursts, may preempt lower classes in a probabilistic way and the *early dropping* approach that randomly drops bursts depending on their class (Zhang et al., 2004). This is done on purpose, in order to maintain the required loss ratios of higher priority bursts.

The issue of QoS provision is still under study and is being investigated together with the effect on TCP traffic. The prime target is to provide a guaranteed TCP throughput for end-users, minimizing variance of its performance.

FUTURE RESEARCH LINES

Optical burst switching offers significant advantages when compared to traditional circuit and packet switching. Research effort is now focusing on how to utilize this technology for applications that really depend on huge data exchanges as for examples GRID computing. Currently, GRID networks are using an optical network infrastructure, which is dedicated to a small number of well known organizations with extremely large jobs (e.g., large data file transfers between known users or destinations). OBS has the potential of meeting several important objectives of GRIDS as for example: high bandwidth, low latency as well as transparency (bitrate, protocol, and service) in the transmission of huge data bursts.

Another significant area of research in OBS networks concerns burst scheduling. It is clear that the loss of a large burst that may contain packets from numerous users and applications will have an imminent effect in service delivery. Toward this, research efforts are focusing on multi-constrain burst scheduling techniques that take into account burst size, burst destination, and instant traffic situation.

Finally, research effort is still devoted on assessing the effect of OBS in higher layer protocols and in particular in TCP. In a typical IP network, packet loss probability of each packet is independent of other packets and is largely due to overflow of buffers at the routers. In OBS networks this does not apply. When a burst is lost due to contention, numerous TCP agents will time-out, since numerous clients may have packets in that burst and will not receive an acknowledgement. All these sources will enter a slow start phase, where the congestion window will be set to one and TCP throughput will reset. It is therefore clear that such a situation must be avoided.

To this end, research is focusing towards two directions. The first direction aims to investigate the effect of the burst drop probability in TCP throughput and in particular how this probability depends upon the network load and the level of burst contentions in the network. The second direction involves the investigating of the *burstification* effect and particularly how the burst assembly processes affect TCP throughput. Burst assembly introduces an unpredictable delay that prohibits TCP to predict future round trip times by sampling the behavior of packets sent over a connection and averaging these measurements into a "*smoothed*" round-trip time estimate.

CONCLUSION

Optical burst switching has been proposed as an alternative switching paradigm to combine the strengths of both optical packet and circuit (wavelength-routing) switching. Several issues are still under investigation and even more will emerge as technology continues to mature. Within this context, we have analyzed current research trends in burst assembly algorithms, signaling and contention resolution schemes as well as provided future guidelines on topics like QoS provision and multi-cost burst scheduling, which will be important for a successful deployment of this technology. Two important factors that argue for the success of this technology are that a suitable application has been identified, namely GRID computing, and that OBS can be implemented within the framework of generalized multi-protocol label switching (GMPLS).

REFERENCES

Callegati, F., & Tamil, L., (2000). IEEE Communications Letters, 4(3), 98-100.

Cao, X. Y., Chen, J., Li, & Qiao, C. (2002). IEEE Globecom 2002 (Vol. 3, pp. 2808-2812).

Dueser, M., & Bayvel, P. (2002). IEEE/OSA Journal of Lightwave Technology, 20(4), 574-585.

Hudek, G. C., & Muder, D. J. (1995). IEEE Proceedings of International Conference on Communications (pp. 1206 -1210).

Karanam, R., Vokkarane, V. M., & Jue, J. P. (2003). In *Proceedings, IEEE/OSA Optical Fiber Communication Conference* 2003.

Liao, W., & Loi, C. (2004). IEEE/OSA Journal of Lightwave Technology, 22(7), 1651-1660.

Vokkarane, V.M., Haridoss, K., and Jue, J.P., (2002). In *Proceedings of the Optical Communications Conference* (Vol. 1, pp. 125-136).

Qiao, C. & Yoo, M. (1997). IEEE/LEOS Technology Global Information Infrastructure, 26–27.

Qiao, C. & Yoo, M. (1999). Journal of High Speed Networks, 8, 69-84.

Qiao, C. & Yoo, M. (2000). IEEE Journal on Selected Areas in Communications, 18(10), 2062-2071.

Turner, J. S. (1999). Journal of High Speed Networks, 8(1), 3-16.

Varvarigos, E. A., & Sharma, V. (1997). IEEE/ACM Transactions on Networking, 5 (5), 705-718.

Vokkarane, V., & Jue, J. P. (2003). IEEE Journal on Selected Areas in Communications, 21(7), 1198-1209.

Wei, J. Y., & MacFarland Jr, R. I. (2000). IEEE/OSA Journal of Lightwave Technology, 18(12), 2019-2037.

Xiong, Y., Vandenhoute, M., & Cankaya, H. (2000). IEEE Journal on Selected Areas in Communications, 18(10), 1838-1851.

Yang, L., Jiang, Y., & Jiang, S. (2003). In *Proceedings of the IEEE Global Telecommunication Communication Conference* (pp. 2689-2673).

Zhang, Q., Vokkarane, V. M., Jue, J. P., & Chen, B. (2004). IEEE Journal on Selected Areas in Communications, 22(9), 1781-1795.

KEY TERMS

FDL: Fiber delay line is a fixed length of fibre that is used to induce a given delay in the passing through optical signal.

GMPLS: Generalized multi-protocol label switching protocol allows traffic paths to be set up through a switched network automatically. This involves the configuration of core switches for the transparent forwarding data from a given start to given end point.

GRID: GRID is an emerging computing model that provides the ability to execute complex processing tasks in a number of distributed, inter-networked computers.

Optical Burst Switching (OBS): A new switching concept which lies between optical circuit switching and optical packet switching. In optical burst switching, the switching payload is the aggregation of numerous packets, usually call burst of packets.

OXC: Optical cross-connect is network device (switch fabric) used by network operators to switch high-speed optical signals. It is capable of switching multiple high-speed signals that are not multiplexed together.

Quality of Service (QoS): Refers to the capability of a telecommunication network to meet a requested quality or traffic contract. In many cases quality of service is refered to the probability of a packet succeeding in propagating through a certain link or path in the network, within its delay bounds.