**Universität Stuttgart**

INSTITUT FÜR
KOMMUNIKATIONSNETZE
UND RECHNERSYSTEME
Prof. Dr.-Ing. Andreas Kirstädter

Institute of Communication Networks and Computer Engineering
University of Stuttgart
Pfaffenwaldring 47, D-70569 Stuttgart, Germany
Phone: ++49-711-685-68026, Fax: ++49-711-685-67983
Email: mail@ikr.uni-stuttgart.de, http://www.ikr.uni-stuttgart.de

# Deterministic Delay Guarantee in OBS Edge Node for Premium Services

Guoqiang Hu, Carla Raffaelli, and Andrea Perin

*Abstract*—In optical burst switched (OBS) networks, the queueing delay in ingress edge node is an important performance measure. In this paper, we propose a deterministic delay model and derive the upper bound of the burst queueing delay in an edge node. On this basis, the edge-to-edge delay guarantee framework for premium services is further outlined.

*Index Terms*—optical burst switching, deterministic QoS, delay analysis

## I. INTRODUCTION

**O**PTICAL burst switching (OBS) [1], [2] is one of the representative architectures proposed for future all-optical networks. In OBS core nodes, optical data frames called *bursts* are switched transparently via optical switching fabric. The header information of each burst for routing/switching decision is encapsulated in a *burst header packet* (BHP). BHPs are transmitted on separate control wavelengths and processed by an electronic switching control unit (SCU) in each switching node. To alleviate per-hop BHP processing workload of the SCU, OBS bursts have large sizes. Furthermore, a BHP needs to arrive at a switching node earlier than the correspondent optical burst by an *offset time* for the compensation of the BHP buffering and processing latency in the SCU.

To support these features, ingress edge nodes have the tasks in classifying/assembling client traffic into OBS bursts, generation of BHPs and scheduling of BHP/burst transmission with the insertion of offset time. Here, the delay becomes a crucial performance issue for time-critical services. Previous work was concentrated on the statistics of random delay components of single node either by means of simulation [2]–[4] or by queueing analysis [5]–[7]. These statistical evaluations, however, rely very much on the accurate modeling of network traffic. In practical deployment, robustness problem can arise due to the inherent high variability and dynamics of the contemporary network traffic.

On the other hand, the network-wide quality of service (QoS) framework with deterministic guarantee has quite mature theories and technologies [8]–[10]. Despite of the overestimation in resource engineering, the deterministic QoS paradigm provides robust absolute end-to-end (E2E) performance guarantee, which is an important feature desired by many premium services.

In this paper, a novel edge-node delay model is proposed following the deterministic QoS paradigm for premium services in OBS networks. After introducing the system model and parameters in Section II, we derive in Section III the upper bound for the burst queueing delay in the edge node with consideration of different offset time requirements among multiple flows. Section IV shows how the delay model is incorporated into an edge-to-edge delay guarantee framework. Conclusions are drawn in Section V.

## II. SYSTEM MODEL AND PARAMETERS

We look at an ingress edge node that has $m$ input ports connected to client networks. Each port has the same channel rate $C_{\text{in}}$. The traffic coming through the input ports is classified into $n$ forwarding equivalence classes (FEC) according to their targeted egress node and QoS class. For each FEC, the client traffic is collected in an assembly buffer. Timer-based assembly is used for the assembly control with the timeout period $T_i$ for FEC $i$. The timer is set upon the arrival of the first packet in an empty assembly buffer. As timeout occurs, all data in the assembly buffer are encapsulated into a burst and immediately forwarded to a transmission buffer. The transmission buffer is shared by all $n$ FECs. Unlimited buffer size is assumed for the delay analysis. From the delay bound derived later, the practical buffer size can be easily calculated. The burst transmission buffer is equipped with a single data wavelength with transmission rate $C_{\text{WL}}$.

For each assembled burst, a BHP is generated. To insert the offset time between the BHP and the burst, the burst transmission is artificially delayed. Let $\Delta_i$ denote the required offset time of an arbitrary FEC $i$. If a burst of FEC $i$ arrives in the transmission buffer at time $t$, the earliest time allowed for its transmission, which is generally called *eligible time*, equals to $t + \Delta_i$.

The transmission channel is reserved for incoming bursts in a first-come-first-serve (FCFS) manner. Once the reservation is made, the transmission of the correspondent BHP is planned by considering the offset time. Further changes in this reservation are not allowed. The transmission time window reserved for a newly arriving burst cannot overlap or precede those of already existing reservations, which is similar to the Horizon reservation scheme [2] proposed for OBS core nodes. An example is illustrated in Fig. 1. The two upward arrows stand for two burst arrivals in the transmission buffer. The first burst is from FEC 1. Since the channel is idle, the time window of its channel reservation (the shaded block labeled with 1) starts exactly at the eligible time of the burst. The second burst is from FEC 2 and has a shorter offset time $\Delta_2$.

G. Hu was with the Institute of Communication Networks and Computer Engineering in the University of Stuttgart when this work was initiated. He is currently a research member in the Centre for Quantifiable Quality of Service in Communication Systems in the Norwegian University of Science and Technology, (e-mail: guoqiang.hu@q2s.ntnu.no).

C. Raffaelli and A. Perin are with the Department of Electronics, Computer Science and Systems in the University of Bologna, (email: carla.raffaelli@unibo.it, perin.andrea@gmail.com).

Fig. 1: Burst scheduling in the edge node



Fig. 2: Queueing delay for a test burst of FEC $i$
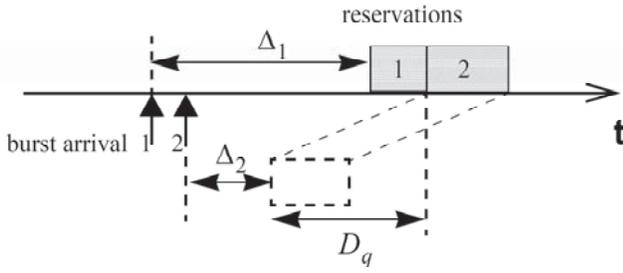
Although the channel can accommodate the burst transmission starting at the eligible time, the reservation of the 2nd burst (the shaded block labeled with 2) must be placed subsequently after the channel reservation of the first burst. Consequently, an additional queueing delay occurs at the 2nd burst and is denoted by $D_q$. Note that the burst queueing delay in the transmission buffer does not include the constant burst delay due to the offset time. In the following, we use $D_{q,i}$ denote the burst queueing delay of FEC $i$. This paper concentrates on the delay bound analysis for $D_{q,i}$.

For the analysis, the incoming client traffic for each FEC is modeled as an aggregation of regulated micro-flows from many users. Regulated micro-flows are suitable to model, e.g., audio/video flows with source shaping. Each micro-flow is characterized by a *constraint function* $A(t)$ which stands for the maximal amount (in bits, bytes or cells) of traffic arrival in an arbitrary time interval of $t$ [8], [10]. A typical constraint function is: $A(t) = \min[p \cdot t, \sigma + r \cdot t]$ where $p$ denotes the peak traffic rate and $r$ is the sustainable rate. $\sigma$ is the maximal batch size defined as the maximal amount of traffic volume that arrives instantaneously without considering the channel limitation. The constraint function has the superposition property. Let $A_i(t)$ represent the constraint function for FEC $i$. Then, $A_i(t)$ equals to the sum of the constraint functions of individual micro-flows.

## III. BURST QUEUEING DELAY

### A. Delay Bound Analysis

The states of the data wavelength channel are classified into "idle" and "occupied". The channel is occupied at a specific time instant if it is transmitting a burst or it is reserved for the transmission of a burst. Otherwise, it is idle. This is exemplified in Fig. 2. The time period in which the channel is continuously occupied is called *busy period*.

An arbitrary burst arrival in the transmission buffer is selected for the analysis. This burst is referred to as the *test burst*. It is from FEC $i$ and the offset time is $\Delta_i$. The channel occupancy at the arrival of the test burst is shown by the shaded blocks in Fig. 2. For the analysis, we are interested in the situation that the eligible time (i.e., the arrival time plus the offset time) of this burst is earlier than the tail of the latest reservation window on the channel. In this case, the
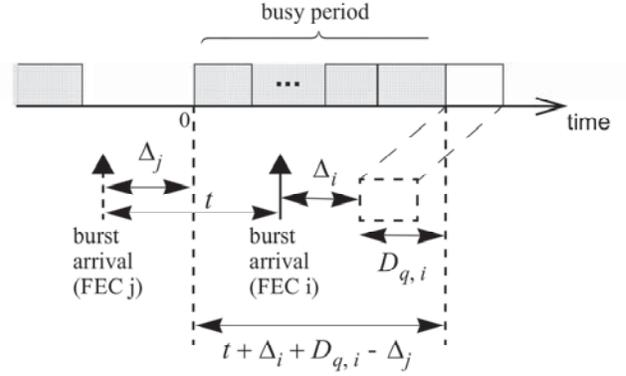
burst is subject to a non-zero queueing delay $D_{q,i}$, as depicted in Fig. 2.

Look at the current busy period to which the test burst attaches its reservation window. The *starting burst* of this busy period must have a zero queueing delay. So the busy period begins actually at the eligible time of the burst. Suppose that this burst comes from FEC $j$, its arrival time at the transmission buffer precedes the busy period by $\Delta_j$. Without loss of generality, we set the arrival time of the starting burst as the "0" point of the time axis. Let $t : t \geq 0$ denote the arrival time of the test burst. The existing channel reservations (shaded blocks) in the current busy period correspond to all bursts arriving after the starting burst (included) and before the test burst (excluded). The conservation law leads to:

$$C_{\mathrm{WL}} \cdot (t + \Delta_i + D_{q,i} - \Delta_j) = \sum_{k=1}^{n} W_k(0, t). \qquad (1)$$

Recall that $C_{\mathrm{WL}}$ is the wavelength channel rate. $W_k(0, t)$ here represents the workload arriving between the time instant 0 and $t$ from FEC $k$. Note that bursts arriving exactly at 0 or $t$ are also accounted into $W_k(0, t)$ except for $k = i$. $W_i(0, t)$ does not include the test burst itself. An implicit assumption here is that the test burst is the last to be scheduled if there are multiple burst arrivals at time instant $t$. In the following, $W_k(0, t)$ is analyzed for different FECs respectively.

1) FEC $j$ of the Starting Burst ($j \neq i$): As the minimal burst inter-arrival time from FEC $j$ equals to the timeout period $T_j$ of the assembler, it is straightforward that the maximal number of burst arrivals between 0 and $t$ is $1 + \lfloor t/T_j \rfloor$ including the starting burst. Since the assembly process is lossless, the total workload of these bursts equals to the traffic amount for FEC $j$ in the time interval of $T_j + T_j \cdot \lfloor t/T_j \rfloor$. According to the definition of the constraint function $A_j(t)$ for FEC $j$, an upper bound for $W_j(0, t)$ is derived:

$$W_j(0, t) \leq A_j(T_j + T_j \cdot \lfloor \frac{t}{T_j} \rfloor). \qquad (2)$$

2) FEC $k$ ($k \neq j$ and $k \neq i$): For an FEC $k$ of neither the starting burst nor the test burst, the number of burst arrivals between 0 and $t$ reaches the maximum when the burst inter-arrival time is minimal $T_k$ and one burst of

FEC $k$ arrives exactly at time instant 0 or at time instant $t$. In either case, it yields:

$$W_k(0,t) \leq A_k(T_k + T_k \cdot \lfloor \frac{t}{T_k} \rfloor). \quad (3)$$

3) FEC $i$ of the Test Burst: Since the test burst itself is not counted, the maximal number of burst arrivals between 0 and $t$ equals to $\lfloor t/T_i \rfloor$. That is:

$$W_i(0,t) \leq A_i(T_i \cdot \lfloor \frac{t}{T_i} \rfloor). \quad (4)$$

By inserting Equ. (2)-(4) into Equ. (1), we obtain:

$$C_{\text{WL}} \cdot (t + D_{q,i}) = \sum_{k=1}^{n} W_k(0,t) + C_{\text{WL}} \cdot (\Delta_j - \Delta_i) \quad (5)$$

$$\leq \sum_{k \neq i} A_k(T_k + T_k \cdot \lfloor \frac{t}{T_k} \rfloor) + \quad (6)$$

$$A_i(T_i \cdot \lfloor \frac{t}{T_i} \rfloor) + C_{\text{WL}} \cdot (\Delta_{\max} - \Delta_i)$$

where $\Delta_{\max}$ is the maximal offset time among all $n$ FECs. Note that Equ. (6) is derived with the assumption that the starting burst and the test burst are from different FECs ($j \neq i$). In case they are of the same FEC ($j = i$), it can be proved that Equ. (6) is valid as well. On this basis, if there exists a value $d$ such that for all $t \geq 0$:

$$C_{\text{WL}} \cdot (t + d) \geq \sum_{k \neq i} A_k(T_k + T_k \cdot \lfloor \frac{t}{T_k} \rfloor) + \quad (7)$$

$$A_i(T_i \cdot \lfloor \frac{t}{T_i} \rfloor) + C_{\text{WL}} \cdot (\Delta_{\max} - \Delta_i),$$

then $D_{q,i} \leq d$. To get a tight upper bound for $D_{q,i}$, the minimal $d$ satisfying Equ. (7) is searched. Concluding, the burst queueing delay for FEC $i$ is bounded by:

$$D_{q,i} \leq \min_{d \geq 0} \{ d : \forall t \geq 0, C_{\text{WL}} \cdot (t + d) \geq \quad (8)$$

$$\sum_{k \neq i} A_k(T_k + T_k \cdot \lfloor \frac{t}{T_k} \rfloor) + A_i(T_i \cdot \lfloor \frac{t}{T_i} \rfloor) +$$

$$C_{\text{WL}} \cdot (\Delta_{\max} - \Delta_i) \}.$$

*Example A:* To further outline the derivation of the queueing delay bound, an example scenario is studied for the edge node. For simplicity, it is assumed that all FECs have the same offset time. In this case, the burst queueing delay in the transmission buffer resembles that of an unbounded FIFO queue. With $\Delta_{\max} = \Delta_i$, Equ. (8) leads to:

$$D_{q,i} \leq \min_{d \geq 0} \{ d : \forall t \geq 0, C_{\text{WL}} \cdot (t + d) \geq \quad (9)$$

$$\sum_{k \neq i} A_k(T_k + T_k \cdot \lfloor \frac{t}{T_k} \rfloor) + A_i(T_i \cdot \lfloor \frac{t}{T_i} \rfloor) \}.$$

The offset time has no influence on the queueing delay any more.

Totally $n = 10$ FECs are considered. The client traffic for each FEC is aggregated from 150 micro-flows of video conference services. The constraint function of each micro-flow has the setting borrowed from Table I in [11]: peak rate $p = 10$ Mbps, sustainable rate $r = 0.5$ Mbps and maximal batch size $\sigma = 0.08$ Mbits. By superposition, the constraint function of FEC $k$ is $A_k(t) = 150 \min[10t, 0.08 + 0.5t] = \min[1500t, 12 + 75t]$ in Mbits for $1 \leq k \leq 10$. Notice that this represents a high traffic peakness of $p/r = 20$. The assembly timeout $T_k = 0.8$ ms for all $k : 1 \leq k \leq 10$. The wavelength channel rate $C_{\text{WL}} = 10$ Gbps.

Since the FEC flows are homogeneous, all FECs have the same delay bound. A schematic determination of the delay bound is shown in Fig. 3. The stepwise increasing curve corresponds to the traffic workload on the right hand side of the condition in Equ. (9). This stands for the maximal amount of traffic volume that is sent before the transmission of the test burst within the current busy period. The function $C_{\text{WL}} \cdot (t+d)$ on the left hand side is sketched in lines with circle markers which are referred to as *service curves*. It is seen that with $d = 0$ the delay bound condition is not satisfied. To meet the condition, $d$ is increased by shifting the line $C_{\text{WL}} \cdot t$ leftwards until the workload curve is completely covered under the line of $C_{\text{WL}} \cdot (t+d)$. As shown, this becomes true with $d \geq 0.00468$. So, the smallest upper bound of the burst queueing delay is 4.68 ms.

### B. A Refinement

The preceding analysis does not take into account the restriction of the input ports on the shape of the client traffic. In this subsection, we aim to exploit the total rates $m \cdot C_{\text{in}}$ of the input ports as a hard limitation on the peak rate summed up from all FECs. For the consistency in the presentation, $A_k(t)$ with $1 \leq k \leq n$ still refers to the constraint function derived from the superposition of regulated micro-flows without considering the influence of the input ports.

Summarizing Equ. (2)-(4), we find that in order to evaluate the workload process $W_k(0,t)$ in the transmission buffer for any FEC $k : 1 \leq k \leq n$, the relevant time window size of the client traffic process amounts to $t + T_k$ at most. Therefore, $\sum_{k=1}^{n} W_k(0,t)$ is bounded by $m \cdot C_{\text{in}} \cdot (t + T_{\max})$ where $T_{\max}$ is the maximal timeout period among all FECs. Corresponding to Equ. (6), we get:

$$C_{\text{WL}} \cdot (t + D_{q,i}) \leq m \cdot C_{\text{in}} \cdot (t + T_{\max}) + \quad (10)$$

$$C_{\text{WL}} \cdot (\Delta_{\max} - \Delta_i).$$

With consideration of both Equ. (6) and (10), the delay bound in Equ. (8) is refined to:

$$D_{q,i} \leq \min_{d \geq 0} \{ d : \forall t \geq 0, C_{\text{WL}} \cdot (t + d) \geq \quad (11)$$

$$\min[\sum_{k \neq i} A_k(T_k + T_k \cdot \lfloor \frac{t}{T_k} \rfloor) + A_i(T_i \cdot \lfloor \frac{t}{T_i} \rfloor),$$

$$m \cdot C_{\text{in}} \cdot (t + T_{\max})] + C_{\text{WL}} \cdot (\Delta_{\max} - \Delta_i) \}.$$

*Example B:* A schematic determination of the delay bound with the refined analysis is plotted in Fig. 4 with respect to the same scenario of Example A. Again, the influence of the offset time is excluded due to the homogeneous setting of the FECs. In addition, the total rate of the input ports is set to $m \cdot C_{\text{in}} = 12$ Gbps, which is larger than the wavelength channel rate $C_{\text{WL}} = 10$ Gbps. Notice that the refined workload curve is obtained from the min-operator on the original workload curve of Example A and the port limitation $m \cdot C_{\text{in}} \cdot (t +$
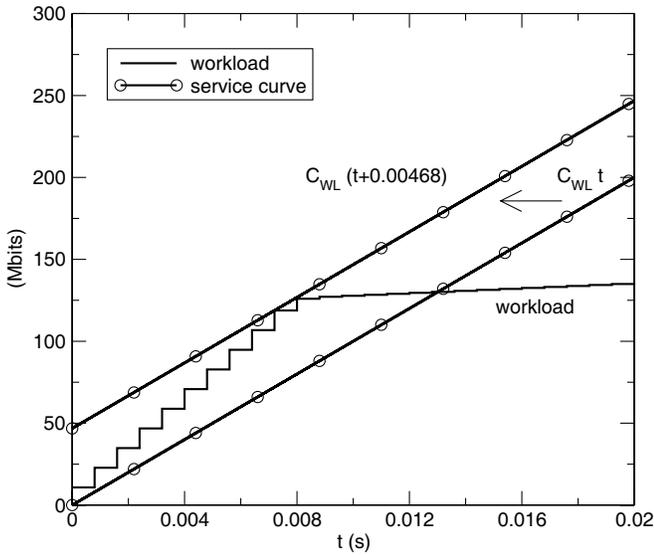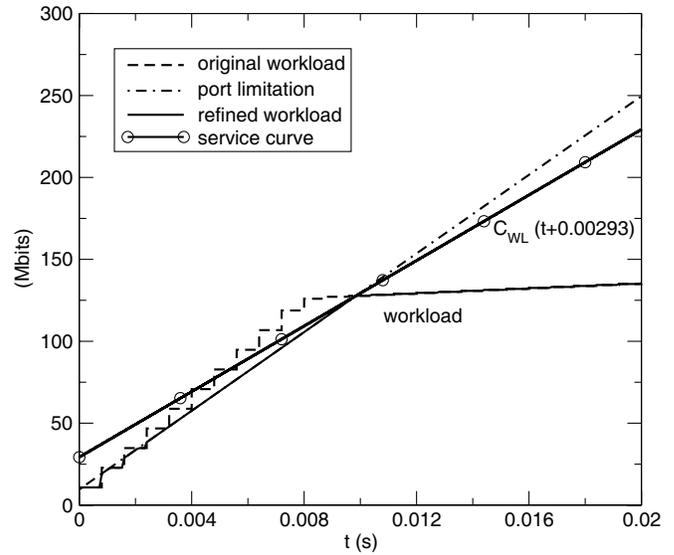
Fig. 3: Determination of the delay bound



Fig. 4: Refinement on the delay bound

$T_{\max}) = 12000 \cdot (t + 0.0008)$ Mbits. By shifting the service curve $C_{WL} \cdot (t + d)$ leftwards (starting with $d = 0$) until it covers the refined workload curve completely, a tighter delay bound of 2.93 ms is obtained.

## IV. EXTENSION FOR AN EDGE-TO-EDGE GUARANTEE

In Section III, the assembly timeout period $T_i$ and the offset time $\Delta_i$ are treated as the known parameters in a given system configuration. In an edge-to-edge delay guarantee, however, they themselves are important delay components and need to be configured with care.

For a sufficient delay compensation in core nodes, the offset time should at least equal to the edge-to-edge BHP latency in SCUs. Let $U_{q,i}$ denote the upper bound of the queueing delay $D_{q,i}$ analyzed in Section III. It can be shown that the BHP departure flow of FEC $i$ from an ingress edge node follows the traffic specification of a generic cell rate algorithm (GCRA) [12] with a constant inter-packet time $T_i$ and a delay variation $U_{q,i}$, i.e., GCRA($T_i$, $U_{q,i}$). On this basis, the edge-to-edge BHP latency in SCUs can be bounded by means of per-hop packet scheduling [8], which further guides the setting of $\Delta_i$.

Through the above delay bound analysis, the burst queueing delay $D_{q,i}$ and the offset time $\Delta_i$ are both related to the assembly timeout period $T_i$, the configuration of which is finally determined by the edge-to-edge delay budget specification. This underlies a deterministic edge-to-edge delay guarantee in OBS networks.

## V. CONCLUSION

In this paper, we analyze the upper bound of the burst queueing delay in the OBS ingress edge node following the deterministic QoS paradigm. The analytical result is further related to the configuration of the offset time and the assembly

timeout period under an edge-to-edge guarantee framework. Our work provides a solution to the deterministic delay guarantee for premium services in OBS networks.

## REFERENCES

[1] C. Qiao and M. Yoo, "Optical burst switching (OBS)—a new paradigm for an optical Internet," *J. High Speed Networks*, vol. 8, no. 1, 1999.
[2] Y. Xiong, M. Vanderhoute, and H. C. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, pp. 1838-1851, Oct. 2000.
[3] F. Callegati and C. Raffaelli, "End-to-end performance of an optical transparent packet network," in *IEEE ICC*, Atlanta, June 1998, pp. 508-513.
[4] R. Rajaduray, S. Ovadia, and D. Blumenthal, "Analysis of an edge router for span-constrained optical burst switched (OBS) networks," *IEEE J. Lighwave Technology*, vol. 22, no. 11, 2004.
[5] J. Choi, J. Choi, and M. Kang, "Dimensioning burst assembly process in optical burst switching networks," *IEICE - Trans. Commun.*, vol. E88-B, no. 10, pp. 3855-3863, 2005.
[6] G. Hu, "Performance model for a lossless edge node of OBS networks," in *Proc. IEEE GLOBECOM*, November 2006.
[7] H. Li and I. L.-J. Thng, "Edge node buffer usage in optical burst switching networks," *Photonic Network Commun.*, vol. 13, no. 1, pp. 31-51, Jan. 2007.
[8] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," in *Proc. IEEE*, vol. 83, no. 10, pp. 1374-1396, Oct. 1995.
[9] S. Shenker, C. Partridge, and R. Guerin, "Specification of Guaranteed Quality of Service," IETF, RFC 2212, Sept. 1997.
[10] V. Firoiu, J.-Y. L. Boudec, D. Towsley, and Z.-L. Zhang, "Theories and models for Internet quality of service," in *Proc. IEEE*, vol. 90, no. 9, pp. 1565-1590, Sept. 2002.
[11] V. Sivaraman and F. Chiussi, "Providing end-to-end statistical delay guarantees with earliest deadline first scheduling and per-hop traffic shaping," in *Proc. IEEE INFOCOM*, 2000.
[12] W. Stallings, *High-Speed Networks and Internets: Performance and Quality of Service*, Prentice Hall, 2002.