

# A framework for deterministic delay guarantee in OBS networks

Guoqiang Hu · Carla Raffaelli

Received: 3 September 2009 / Accepted: 24 July 2010 / Published online: 11 August 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** In OBS networks, the delay of control packets in the switch control unit (SCU) of core nodes influences burst loss performance in the optical switching and should be constrained. Furthermore, the end-to-end (E2E) delay requirements of premium services need queueing delay guarantee in network nodes throughout the transmission path. For this purpose, a framework for deterministic delay guarantee is proposed in this article. It incorporates the deterministic delay model in the ingress edge node as well as in the SCUs of core nodes. On this basis, the configuration of the assembler and the offset time is addressed by means of an optimization problem under the delay constraints. Scenario studies are carried out with reference to realistic transport network topologies. Compared to statistical delay models in the literature, the deterministic model has advantages in rendering robust absolute delay guarantee for individual FEC flows, which is especially appreciated in the provisioning of premium services. By performance evaluation in comparison with the statistical models, it is shown that the adopted deterministic delay models lead to practical delay bounds in a magnitude that is close to the delay estimations by stochastic analysis.

**Keywords** Optical burst switching · Burst assembler · Deterministic quality of service · Delay bound

---

G. Hu (✉)  
Centre for Quantifiable Quality of Service in Communication  
Systems, Norwegian University of Science and Technology,  
Trondheim, Norway  
e-mail: guoqiang.hu@q2s.ntnu.no

C. Raffaelli  
Department of Electronics, Computer Science and Systems,  
University of Bologna, Bologna, Italy  
e-mail: carla.raffaelli@unibo.it

## 1 Introduction

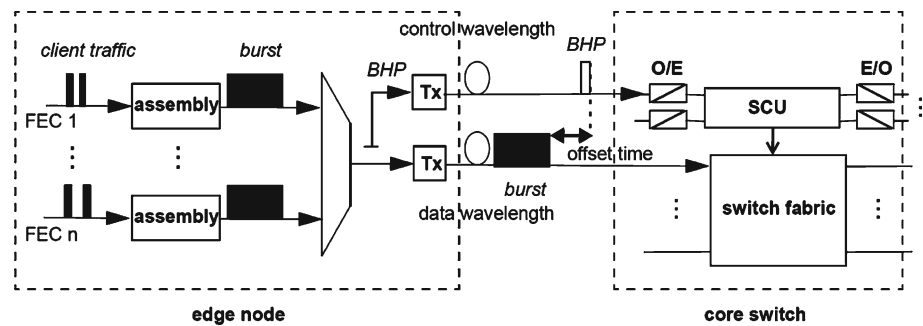
### 1.1 OBS network architecture

One of the main challenges in the development of future packet-switched optical transport networks is the implementation of packet header processing that matches the ultra-high data throughput of the optical switching fabric. While the relatively low processing speed of electronic devices cannot keep up with the increasing switch capacity requirements, optical packet switching is not feasible in the near future due to the difficulties in implementing optical buffering and signal processing.

Optical burst switching (OBS) was proposed [1–3] as an intermediate solution between all optical circuit and packet switching. Electronic header processing is adopted in the core switching node so that a good technological feasibility can be assured. Furthermore, OBS adopts large data frames called *bursts* in the core network to reduce the header processing overhead in switching nodes.

In Fig. 1, schematic OBS edge node and core switch are plotted to illustrate the OBS network architecture. The edge node is the gateway between the OBS core network and client networks. In the graph, only the burst transmitting part of an ingress edge node is shown. Here, the packets of the incoming client traffic (e.g., IP packet) are classified into different forwarding equivalence classes (FEC) according to their targeted egress node and quality-of-service (QoS) class. Packets of the same FEC are collected and assembled into OBS bursts following some assembly strategy [3–5]. For each burst, a burst header packet (BHP) is generated at the same time. The BHP contains all the essential information for the routing/switching of the corresponding burst and is transmitted through a separate control wavelength. The transmission of the burst onto the data wavelength is delayed on purpose for

**Fig. 1** OBS network architecture



a certain amount of time after the BHP transmission. In this way, an *offset time* separates in time the BHP and the data burst. This offset time is to compensate the buffering and processing latency of the BHP in the core nodes throughout the OBS network.

An OBS core switching node is basically composed of electronic switch control unit (SCU) and optical switching fabric. The SCU terminates the incoming control wavelengths and processes the BHP to determine the switching scheme for the corresponding optical burst. On this basis, switching resources are reserved for the upcoming burst. The BHP itself is updated after the local processing and immediately forwarded to the next hop. According to the reservation, the optical switching fabric will be configured to timely set up an internal path and switch the optical burst “on-the-fly”.

The resource reservation by the BHP in a core node can fail in case multiple incoming bursts are to be forwarded to the same output link at the same time and compete in the channel reservation. An unsuccessful reservation results in data burst discard. Since the optical switching fabric is short of buffering capacity in comparison to the conventional electronic packet switch, relatively high data loss rate can occur due to the channel competition. In last years, the burst loss probability has been intensively studied as a major OBS performance measure [6–8]. Different contention resolution schemes are available to improve the loss performance of the optical switching [9–13]. To enable the QoS provisioning in OBS networks, plenty of service differentiation schemes in terms of the loss performance [14–21] was proposed. The framework for absolute end-to-end (E2E) loss guarantee was studied in [18, 20–23].

## 1.2 Essential delay guarantee

Besides the burst loss caused by the competition on the data channels, the BHP latency in the SCU is another important performance metric. The “on-the-fly” nature of burst switching requires that the switching fabric is configured on time before the arrival of the data burst. In other words, the processing of a BHP in the SCU must be finished within a limited period. In fact the overall BHP latency in SCUs along

the route needs to be bounded by the offset time set by the edge node. Otherwise, the burst is blocked when it arrives at a switching node that does not finish its switching configuration yet.

Apparently, the BHP latency in an SCU depends on the workload of the SCU, which is further decided by the traffic assembly in the ingress edge node. A burst assembler reduces the traffic intensity of the BHP flow at the expense of additional assembly delay. In order to provide premium services which have generally a latency requirement, the delay in the edge node must be constrained.

Therefore, OBS networks must essentially guarantee:

1. The BHP latency in the SCUs on the OBS route in order to avoid the burst loss attributed to the overdue BHP processing;
2. The delay of the data traffic in the ingress edge node in order to meet the E2E delay requirement of premium services.

The above two design goals are related with each other so that a comprehensive QoS model is necessary to achieve delay guarantee. This motivates the contributions of this article.

## 1.3 Related work

Simulation study on burst delay performance in the edge node was reported in [24–26] as well as for the BHP delay in the SCU [3]. Analytical queueing models for the edge node were proposed in [27, 28]. Choi et al. [29] used an approximate model to study the BHP delay distribution in the SCU and applied the result to guide the configuration of the burst assembler. In [30], the burst loss performance due to the BHP delay/loss in the SCU was studied.

The aforementioned analytical models are all based on statistical queueing analysis and aimed at deriving the tail probability of the queueing delay. The assumption of first-in-first-out (FIFO) buffers shared by all traffic flows in the edge node or in the SCU is generally applied. The resulting delay estimation is actually the averaged value of all FECs. Per-class delay analysis is difficult to develop. This limits the

application of the statistical models in the differentiated QoS provisioning. Furthermore, the statistical queueing analysis relies very much on the accurate characterization of the traffic. In the practical deployment, robustness problem arises when the traffic dynamic is beyond the assumption adopted by the analysis.

In [31], a delay model was proposed for OBS edge nodes following the deterministic QoS paradigm [32]. This avoids the above problems in the statistical QoS guarantee. However, deterministic guarantee is generally regarded to be much more conservative than the statistical guarantee. At the same time, a comprehensive view with respect to the network-wide delay guarantee is still missing.

#### 1.4 Contributions

In this article, we construct a framework for the deterministic E2E delay guarantee for premium services in OBS networks. It assures delay bounds in both the ingress edge node and the SCU. The client traffic is modeled by the constraint functions that describe the deterministic traffic envelopes for individual flows. The ingress edge node applies timer-based assembly scheme to generate data bursts so that the assembly delay is bounded by the timeout period. The resulting traffic intensity of the BHP flow is thus bounded irrespective of the data throughput. The overall delay bound in the edge node can be analyzed according to [31]. The constraint function of each BHP flow is further derived. In SCUs, conventional scheduling disciplines [33] based on the generalized processor sharing (GPS) are deployed to schedule the BHP processing of different FEC flows. This assures a good service differentiation/separation between the FECs. As a result, the E2E BHP latency in the SCUs of core nodes is guaranteed for each FEC. By a comprehensive delay evaluation with respect to the edge node and the BHP path in the core OBS network, the guideline for the assembler/offset time configuration is outlined.

In comparison to statistical models proposed in the literature, the deterministic delay bounds in this framework firstly stand for the strongest QoS guarantee. The burst loss attributed to the overdue BHP processing is completely eliminated. The buffer in OBS edge nodes and SCUs can be dimensioned free of overflow according to the delay bounds. Secondly, the proposed model provides the delay bounds for each individual FEC and meets the service differentiation requirement. Thirdly, the deterministic QoS model characterizes the traffic by the constraint function enforced by traffic shapers/regulators at the service access point or by the bandwidth of the feeding channels. Therefore, the model is robust to the traffic dynamic. On the other hand, deterministic queueing analysis serves as the worst case analysis and is generally supposed to be too conservative. In this regard, the deterministic delay bounds are evaluated in comparison to the statistical queueing

models. We show that in the representative OBS network scenario the deterministic delay models lead to practical delay bounds in a magnitude that is close to the delay estimations by stochastic analysis.

In the remainder of this article, Sect. 2 deals with the delay guarantee in the edge node for data bursts. BHP delay in the SCUs of core nodes is studied in Sect. 3. In both sections, the relevant system model and parameters are introduced first to specialize the QoS problem. The traffic model is specified and the deterministic delay bound is given. An example study is carried out for the comparison with statistical queueing models. In Sect. 4, a comprehensive delay analysis is performed based on the delay guarantee in preceding sections. The configuration of the burst assembler and the offset time is addressed by means of an optimization problem under the delay constraints. Scenario studies are carried out with reference to realistic transport network topologies. In Sect. 5, the conclusion and outlook are given.

## 2 Delay guarantee in edge node

### 2.1 System model and parameters

Figure 2 illustrates an ingress edge node that has  $m$  input ports connected to client networks. Without loss of generality, each port is assumed to have the same channel rate  $C_{in}$ . The traffic coming through the input ports is classified into  $n$  FECs and collected in the corresponding 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 the transmission buffer. Unlimited buffer size is assumed in order to focus on the delay analysis. The burst transmission buffer is equipped with a single data wavelength with transmission rate  $C_{WL}$ . This corresponds to the practical system implementation by assigning a dedicated transmission buffer for each high-speed data channel to reduce the bandwidth requirement of the memory device.

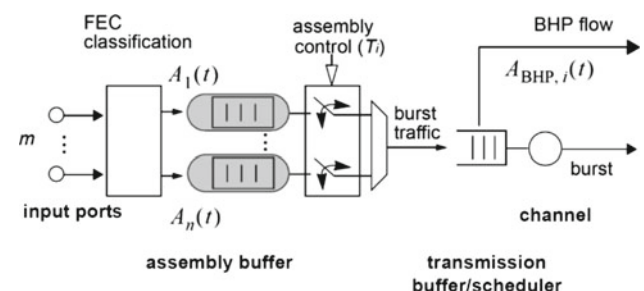
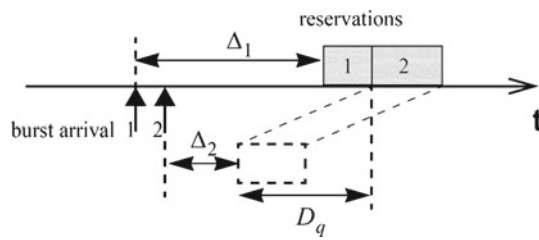


Fig. 2 OBS edge node



**Fig. 3** Burst scheduling in the edge node

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 scheduler reserves the transmission channel for the incoming bursts in a first-come-first-serve (FCFS) manner. Once the reservation is made, the transmission of the corresponding BHP is planned by considering the offset time. Further changes in this reservation (e.g., re-scheduling or preemption) 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 [1,3] proposed for OBS core nodes.

An example is illustrated in Fig. 3. 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 (i.e., 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  $\Delta_2$ . Although the channel is able to accommodate the burst transmission starting at the eligible time, the reservation of the second burst (i.e., 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 second burst and is denoted by  $D_q$ .

### 2.2 Traffic model

In the QoS provisioning for premium services, a service request is subject to the call admission control (CAC). If the request is accepted, its traffic flow is further enforced by the usage parameter control (UPC) to conform to the traffic specifications agreed in the CAC [34]. Such a shaped/regulated traffic 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  $t$  [32,33]. Obviously, the constraint function is non-decreasing with  $t$ . A typical example of the constraint function is:

$$A(t) = \min[p \cdot t, \sigma + r \cdot t] \tag{1}$$

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 operator  $\min[a, b]$  is to choose the smaller one between  $a$  and  $b$ .

After the FEC classification in the ingress edge node, each FEC flow turns out to be the aggregation of regulated micro-flows from many users. Let  $A_i(t)$  represent the constraint function for FEC  $i$  as marked in Fig. 2. Since the constraint function has the superposition property,  $A_i(t)$  equals to the sum of the constraint functions of individual micro-flows.

### 2.3 Bound on the queueing delay

Let  $D_{\text{edge},i}$  denotes the delay in the ingress edge node from the perspective of FEC  $i$ . It mainly consists of three components: the assembly delay, the burst delay for the offset time, and the burst queueing delay. The timer-based assembly scheme has the assembly delay of FEC  $i$  bounded by  $T_i$ . The offset time delay is constant  $\Delta_i$ . Use  $D_{q,i}$  to represent the burst queueing delay of FEC  $i$  in the transmission buffer. The delay in the edge node is then:

$$D_{\text{edge},i} = T_i + \Delta_i + D_{q,i}. \tag{2}$$

$D_{q,i}$  is further evaluated by [31]:

$$D_{q,i} \leq \min_{d \geq 0} \{d : \forall t \geq 0, C_{\text{WL}} \cdot (t + d) \geq \min \left[ \sum_{k \neq i} A_k \left( T_k + T_k \cdot \left\lfloor \frac{t}{T_k} \right\rfloor \right) + A_i \left( T_i \cdot \left\lfloor \frac{t}{T_i} \right\rfloor \right), m \cdot C_{\text{in}} \cdot (t + T_{\text{max}}) \right] + C_{\text{WL}} \cdot (\Delta_{\text{max}} - \Delta_i) \}, \tag{3}$$

where  $T_{\text{max}}$  is the maximal timeout period and  $\Delta_{\text{max}}$  is the maximal offset time among all  $n$  FECs.

The delay upper bound in the edge node is obtained by inserting (3) into (2).

### 2.4 Comparison with the statistical model

In this subsection, the bound of the burst queueing delay in (3) is compared to popular stochastic queueing model.

Due to the superposition of multiple FEC flows in the transmission buffer, an exact statistical queueing analysis is difficult for a large number of FECs. The offset time differences between FECs further increase the hardness. To the best of our knowledge, only approximate queueing models were used in [27,28] without considering the different offset times between FECs. Among them, the  $M/G/1$  model in [28] is the most suited to our system model. This model is justified with the assumption that the incoming traffic is

not long range dependent (LRD). In this case, the aggregated traffic converges to the Poisson process when the number of FECs grows large.

To specify the burst size distribution for the use of the  $M/G/1$  model, we further assume that the burst size is constant and equals to the maximal burst size allowed by the constraint function of the correspondent FEC  $k$ , i.e.,  $A_k(T_k)$ . This is then equivalent to the  $M/MTU/1$  model in [35]. The  $M/MTU/1$  here is interpreted as an  $M/D/1$  model with each data frame having the size of the maximum transmission unit (MTU). In [35], it is heuristically shown that the  $M/MTU/1$  model serves as a worst case model for the  $M/G/1$  system. The delay distribution is calculated according to Sect. III-B1 in [35]. Considering the tolerable burst loss rate in the range of  $10^{-6}$  for OBS networks [36], the  $1 - 10^{-6}$  quantile of the delay distribution is selected as a stochastic delay bound for the comparison with the deterministic bound of (3).

To fit to the statistical model, the evaluation scenario assumes that all FECs have the same offset time so that the burst queue is simplified to an FIFO queue. The client traffic of each FEC is aggregated from a large number of homogeneous regulated micro-flows. The constraint function of each micro-flow is in the form of (1) and has the typical setting of a video conference session: sustainable rate  $r = 0.5$  Mbps and maximal batch size  $\sigma = 0.08$  Mbits. Different peakness  $p/r$  is checked to reflect the impact of the source shaping [37]. All FECs have the same traffic intensity and each FEC flow is aggregated from 300 micro-flows. By superposition, the constraint function of FEC  $k$  is  $A_k(t) = 300 \min[p \cdot t, 0.08 + 0.5t]$  in Mbits for  $1 \leq k \leq n$ . Also, the same assembly timeout  $T_k = 0.8$  ms for  $1 \leq k \leq n$ . The wavelength channel rate  $C_{WL} = 10$  Gbps. To estimate the delay bound under different load situations, we tune the system load  $\rho$  by changing the number of FECs.

A comparison between the deterministic delay bound and the statistical bound is shown in Fig. 4 for the case  $m \cdot C_{in} = 15$  Gbps. The peakness  $p/r$  is set to 4 and 8, respectively. It is seen that in the range of light loads, the results of the deterministic analysis and the statistical analysis are in the same magnitude. At medium and heavy loads, the deterministic bound is much more conservative than the statistical delay bound. It is further discovered that the load point by which the deterministic bound begins to grow dramatically corresponds to the point at which the peak rate summed up from all micro-flows injecting the system exceeds the output channel rate  $C_{WL}$ . A smaller peakness helps to shift the location of this critical points to higher loads. Note also that the deterministic bounds in the case of  $p/r = 4$  and  $p/r = 8$  overlap with each other at heavy loads because the constraint functions in these two cases turn to be the same in the range of large  $t$ . The peak rate  $p$  has the impact only for small values of  $t$ .

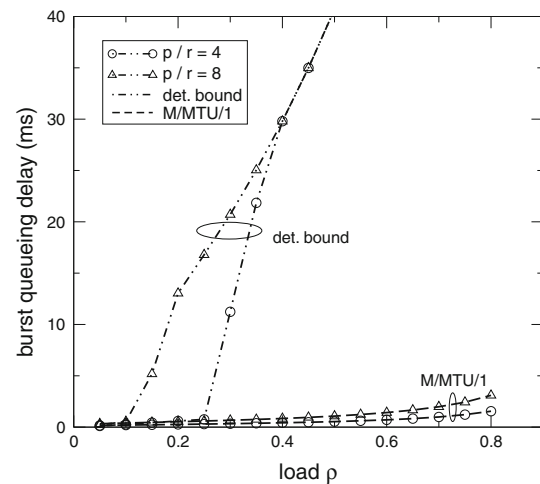


Fig. 4 Burst queuing delay ( $m \cdot C_{in} = 15$  Gbps)

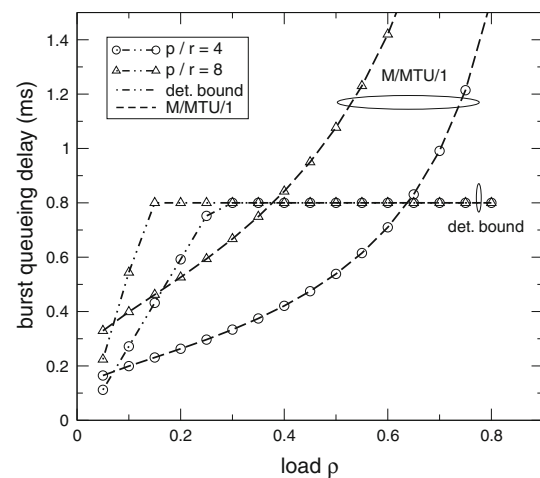


Fig. 5 Burst queuing delay ( $m \cdot C_{in} = 10$  Gbps)

In OBS networks, the practical load on the data wavelength is supposed to be low in order to keep a tolerable burst loss performance [6,22,36] in core switching nodes. In this case, the disadvantage of the deterministic bound is marginal.

An interesting case is to have the total channel rate of the input ports equal to the output channel rate:  $m \cdot C_{in} = C_{WL}$ . This is a very common configuration in the practical system implementation. The results of the delay analysis are plotted in Fig. 5. The input port limitation now suppresses the deterministic queuing delay under the timeout period  $T_k = 0.8$  ms, irrespective of the system load. In contrast, the statistical bounds grow continuously with the increasing load. Note that the statistical bounds here are the same as those in Fig. 4 because the  $M/MTU/1$  model cannot seize the change in  $m \cdot C_{in}$  from 15 Gbps to 10 Gbps. The advantage of the deterministic model is apparent in this scenario.

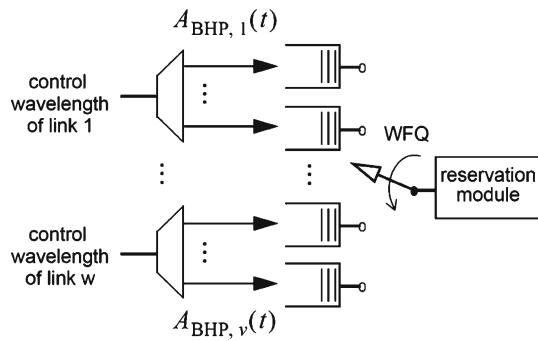


Fig. 6 System model for the SCU

### 3 Delay guarantee in SCUs

#### 3.1 System model and parameters

In the BHP processing by an SCU in the core node, the channel reservation for the upcoming optical burst is the most time-consuming task that constrains the system throughput [38]. In this section, we analyze the performance of the reservation module of an SCU. The system is modeled by a multi-queue single server system depicted in Fig. 6. The BHP traffic from input links is classified into multiple queues according to their FEC identifiers. Weighted fair queuing (WFQ) [33,39], a realization of the GPS policy, is applied to schedule the BHPs of different FECs to the reservation processing. WFQ here assures a good per-node QoS separation and guarantees the E2E BHP delay along the OBS network path.

An appropriate realization of the reservation module can assure a constant BHP processing time that is independent of the burst size [38]. We denote the reservation processing time by  $h$ . For the queuing delay analysis in the SCU, the BHP workload is measured in terms of the number of BHP arrivals. The size of individual BHPs is not an issue here.  $A_{BHP,i}$  is defined as the constraint function for the BHP flow of an arbitrary FEC  $i$ .

#### 3.2 Traffic characterization for BHP flow

We consider a BHP flow of FEC  $i$  that comes directly from an edge node. Since the channel capacity of the control wavelength is not regarded as a performance constraint, the BHP traffic profile is not subjected to deformation in the transmission phase. The characteristic of the incoming BHP flow is equivalent to that departs from the edge node (cf. Fig. 2).

In the edge node, a BHP of FEC  $i$  is always sent earlier than the corresponding optical burst by the fixed offset time  $\Delta_i$ . Therefore, the BHP departure process is equivalent to the burst departure process if the traffic is measured in the number of BHP/burst departures. Note that the maximal burst rate is reached when bursts are assembled constantly at

the interval  $T_i$ . As a result, the bursts of FEC  $i$  are delivered to the transmission buffer in the constant rate. As the bursts experience queuing delay in the transmission buffer, a delay variation equal to the queuing delay bound is brought into the periodic burst flow. This is similar to a flow controlled by the generic cell rate algorithm (GCRA) [34].

In asynchronous transfer mode (ATM) networks, it was known that a constant-cell-rate flow imposed by an additional delay variation is characterized by the traffic descriptor GCRA( $I, \tau$ ) [34]. Here,  $I$  denotes the constant cell inter-arrival time if there were no delay variation.  $\tau$  stands for the delay variation. Apparently, the burst departures from the edge node follow the same traffic pattern. Let  $U_{q,i}$  denotes the upper bound of the queuing delay  $D_{q,i}$ . The burst departures are thus characterized by the descriptor GCRA( $T_i, U_{q,i}$ ). This GCRA descriptor can be further converted to the form of a constraint function [40]:

$$A_{BHP,i}(t) = \frac{t}{T_i} + \left( \frac{U_{q,i}}{T_i} + 1 \right). \tag{4}$$

The term  $1/T_i$  corresponds to the sustainable rate and  $(U_{q,i}/T_i + 1)$  specifies the maximal batch size.

#### 3.3 Bound on the BHP delay in SCUs

Let  $z_i$  denotes the number of switching hops of FEC  $i$  throughout the OBS core network. All switching nodes have the same BHP processing rate of  $1/h$ . In the WFQ configuration of each SCU, FEC  $i$  is allocated with a processing rate of  $\epsilon_i : 1/T_i \leq \epsilon_i \leq 1/h$ . Use  $U_{BHP,i}$  to represent the upper bound on the E2E BHP delay in the SCUs along the OBS network path. According to the theory of the WFQ scheduling [33,39], the following bound is obtained for the BHP flow characterized by (4):

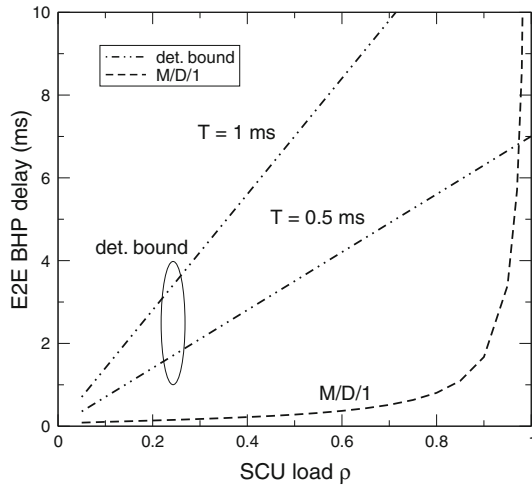
$$U_{BHP,i} = \frac{U_{q,i}}{T_i \cdot \epsilon_i} + \frac{1}{\epsilon_i} + z_i \cdot \frac{1}{\epsilon_i} + z_i \cdot h. \tag{5}$$

The BHP delay bound in SCUs turns out dependent on the burst queuing delay in the edge node.

#### 3.4 Comparison with statistical model

An example scenario is studied for the performance comparison between the deterministic delay model and the statistical model. The FEC  $i$  is assumed to traverse 10 hops in an OBS network. In each SCU, one reservation module is equipped for each output link [38]. The BHP processing time is  $h = 500$  ns. We assume that 20% of the data wavelength channels on an output link are assigned for premium services by means of channel partitioning or its variations [18,21] to guarantee the burst loss performance.

A comparable statistical queuing model for the SCU is the  $M/D/1$  model [29,38]. Here, it is assumed that the BHP



**Fig. 7** Comparison of E2E BHP delay in SCUs

flows of premium services are multiplexed into an FIFO buffer. The aggregated traffic is approximated by a Poisson process. Corresponding to the data channel allocation, the WFQ scheduler of the reservation module is configured to guarantee 20% of the processing capacity for the premium services. This is equivalent to a service rate of  $4 \times 10^5$  BHPs per second in the  $M/D/1$  model. Similar to Sect. 2, the queueing delay distribution is calculated according to [35] and the  $1 - 10^{-6}$  quantile of the delay distribution is taken as the stochastic delay bound. For the evaluation, the load  $\rho$  of the SCU reservation module is the same at each hop. The E2E BHP delay in SCUs is measured by summing up the statistical BHP delay bound of each hop.

For the deterministic analysis, we look at homogeneous BHP flows. For each FEC,  $T_i$  is set to 0.5 and 1 ms, respectively.  $U_{q,i}$  is fixed to  $3T_i$ . The total BHP processing rate ( $4 \times 10^5$  BHPs per second) for the premium services is equally allocated to BHP flows by the WFQ. The same SCU load  $\rho$  is assumed in each hop. The deterministic BHP delay bound is calculated by (5). We change the SCU load by tuning the number of BHP flows.

The analytical results are plotted in Fig. 7. It shows that the deterministic bound is larger than the statistical bound at light and medium loads. However, in the heavy load situation the statistical bound grows dramatically. In contrast, the deterministic model assures an upper bound even at a 100% load. In any case, the deterministic upper bound is in the magnitude from millisecond to tens of millisecond, meeting well the E2E delay requirements of premium services. For example, the E2E delay budget for interactive real time services can amount to 100 ms in practice [41].

#### 4 Edge node configuration with delay budget

In Sects. 2 and 3, the delay guarantee for the burst queueing in the edge node and the BHP latency in the SCUs is stud-

ied separately. Also, the system parameters like the timeout period  $T_i$  and the offset time  $\Delta_i$  are used as already configured. In this section, the practical configuration problem in the edge node is solved by the joint application of the delay bounds in previous sections.

##### 4.1 Scenario

Referring to the edge node model in Fig. 2, we specify that all the  $n$  FECs are of the same premium service class. They are differentiated from each other according to the destined egress edge node of the traffic. The number of switching hops in the OBS network is  $z_i$  for FEC  $i : 1 \leq i \leq n$ . An edge node delay budget  $\phi_i$  is specified for each FEC  $i$  that covers the transit delay of the data between the input port and the output burst transmitter. Note that  $\phi_i$  can be directly derived from the E2E delay budget by subtracting the constant burst propagation delay in the core OBS network. As a representative design in the device implementation, the total channel rate of input ports is set equal to the output channel rate:  $m \cdot C_{in} = C_{WL}$ . Each FEC  $i$  needs to configure its parameter  $T_i, \Delta_i$  and  $\epsilon_i$  so as to meet the delay budget and consume the SCU processing capacity efficiently.

##### 4.2 Problem formulation

The key to the configuration problem is the tradeoff between the SCU load and the delay performance. A larger timeout period  $T_i$  helps to reduce the traffic rate of the BHP flow, but increases the burst queueing delay as well as the BHP delay. From the perspective of the network operator, it suffices to limit the delay within the delay budget. A further delay reduction does not bring added value. Therefore, we take the delay budget as a constraint on the configuration and aim to minimize the demand on the processing capacity in the SCUs. Hence, the guaranteed SCU processing rate  $\epsilon_i$  for FEC  $i$  is assigned with the minimal value that equals to the BHP sustainable rate. With  $\epsilon_i = 1/T_i$ , the configuration is formulated by an optimization problem with the objective function:

$$\min : \sum_{i=1}^n z_i \cdot \epsilon_i = \min : \sum_{i=1}^n z_i \cdot \frac{1}{T_i} \tag{6}$$

The function represents the sum of the allocated SCU processing capacity in every hops for all  $n$  FECs of the edge node.

The delay constraint for individual FEC  $i$  is formulated from (2):

$$T_i + \Delta_i + U_{q,i} \leq \phi_i \quad \text{for } 1 \leq i \leq n. \tag{7}$$

Recall that  $U_{q,i}$  is the bound of the burst queueing delay. With  $m \cdot C_{in} = C_{WL}$ , a slightly loosened  $U_{q,i}$  can be derived

from (3) that is independent of the specific constraint function per FEC:

$$U_{q,i} = T_{\max} + \Delta_{\max} - \Delta_i. \tag{8}$$

Insertion of (8) into (7) leads to:

$$T_i + T_{\max} + \Delta_{\max} \leq \phi_i \quad \text{for } 1 \leq i \leq n. \tag{9}$$

The offset time  $\Delta_i$  must at least cover the bound of the E2E BHP delay in SCUs. Plug  $\epsilon_i = 1/T_i$  and (8) into (5). It is derived:

$$\Delta_i \geq U_{\text{BHP},i} = \frac{1}{2}(\Delta_{\max} + T_{\max} + T_i + z_i \cdot (T_i + h)) \tag{10}$$

for  $1 \leq i \leq n$ .

Equations (9) and (10) here form the basic constraints in configuring  $T_i$  and  $\Delta_i$ . Furthermore, the facts that  $T_{\max} = \max_{1 \leq i \leq n} \{T_i\}$  and  $\Delta_{\max} = \max_{1 \leq i \leq n} \{\Delta_i\}$  are loosely formulated by the constraints  $T_i \leq T_{\max}$  and  $\Delta_i \leq \Delta_{\max}$ . Therefore, all the constraints are linear. Standard methods are available to solve such a convex programming problem [42].

When the optimization solution is obtained,  $T_k = T_{\max}$  and  $\Delta_w = \Delta_{\max}$  must hold for some FEC  $k, w$ . Otherwise, according to (9)  $T_{\max}$  and  $\Delta_{\max}$  can be further decreased to increase the timeout periods of FECs. Therefore, in the optimization solution  $T_{\max}$  and  $\Delta_{\max}$  become exactly the respective maximum of  $T_i$  and  $\Delta_i$  for  $1 \leq i \leq n$ .

### 4.3 Example solutions

For the evaluation, two reference transport network scenarios are selected [43]: the 28-nodes pan-European network and the 14-nodes US network. In both networks, one node is taken as the local ingress edge node that is to be configured. For each FEC, the routing to the egress node follows the shortest path (in the number of hops). The BHP processing time in SCUs is  $h = 500$  ns. The propagation delay of optical bursts is calculated according to the physical distance matrix. We assume an E2E delay budget of 30 ms for each FEC. The edge node delay budget  $\phi_i$  is determined by subtracting the constant propagation delay from the E2E delay budget.

In the example, Dublin is chosen in the pan-European network. The flows destined to the other 27 nodes are classified into 27 FECs. The resulting hop count  $z_i$  and delay budget  $\phi_i$  are summarized in the vector  $\mathbf{d}$  and  $\boldsymbol{\phi}$ :  $\mathbf{d} = \{1, 1, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7\}$ ;  $\boldsymbol{\phi} = \{27.7, 26.6, 24.0, 23.9, 22.6, 21.6, 21.1, 21.0, 20.2, 20.5, 20.2, 19.2, 17.0, 16.1, 18.9, 17.9, 17.1, 16.5, 15.3, 15.3, 15.2, 13.7, 12.9, 9.2, 13.2, 11.4, 7.8\}$  ms. In the US network, NY is taken as the local edge node. The correspondent parameters are:  $\boldsymbol{\phi} : \mathbf{d} = \{1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3\}$ ;  $\boldsymbol{\phi} = \{27.8, 27.7, 26.4, 26.1, 23.6, 22.8, 15.8, 12.3, 19.3, 8.9, 6.5, 6.5, 3.2\}$  ms. It is noticeable that the European network has a large network diameter in the number of hops. The US network has a small diameter but the propagation latency on the long-haul links consumes a large delay budget.

For both networks, the optimization problem is solved numerically by the MATLAB optimization toolbox. The resulting setting of the timeout period and the offset time is plotted in Figs. 8 and 9. The x-axis is labeled with the abbreviations of destination nodes [43] that correspond to individual FECs. They are arranged in the same order as that of the vector  $\mathbf{d}$  and  $\boldsymbol{\phi}$ , i.e., the hop count ascends from left to right. In general, the following observations are made:

1. The offset time is several times larger than the timeout period. This is because the minimal BHP processing rate  $\epsilon_i = 1/T_i$  is allocated in SCUs for FEC  $i$ . The BHP latency at each hop equals at least to the timeout period  $T_i$ .
2. FECs with longer paths are generally assigned with smaller edge node delay budgets due to their larger propagation delay. Whereas, a larger offset time is necessary for the larger hop count. This leads to the relatively small timeout periods granted for these FECs.
3. Among the different FECs, the difference in the setting of the timeout period is not very large. This is attributed to the influence of the common term  $T_{\max}, \Delta_{\max}$  in (9) and (10).

By comparing Figs. 8 and 9, it is found that in average the allowable timeout period in the pan-European network

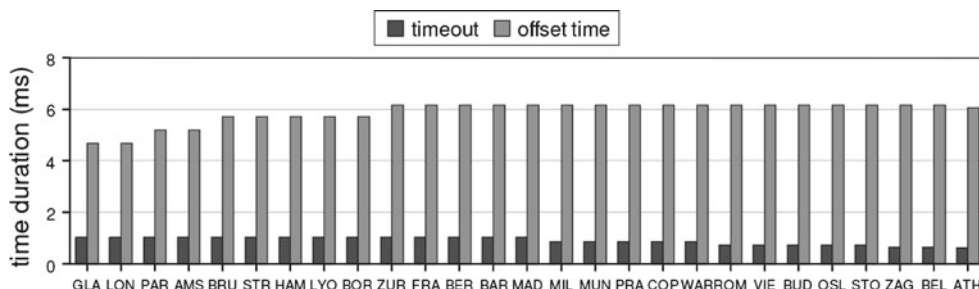


Fig. 8 Configuration for the pan-European network



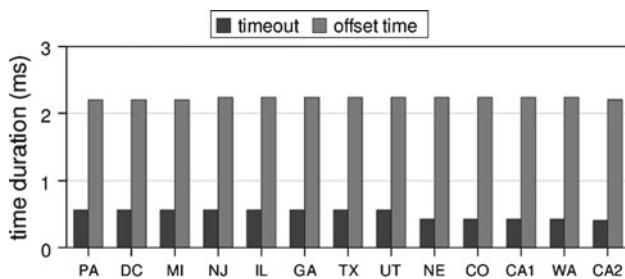


Fig. 9 Configuration for the US network

is larger than that of the US network, because all FECs in the US network suffer from the relatively small delay budget  $\phi_{13}3.2$  ms (egress node CA2) through the common term  $T_{\max}$  and  $\Delta_{\max}$  in the formula. On the other hand, the small diameter of the US network leads to a small ratio of the offset time and the timeout period.

## 5 Conclusion and outlook

This article presented a framework for the deterministic delay guarantee for premium services in OBS networks. It enables a robust configuration of the burst assembler and the offset time so as to: (1) eliminate the burst loss problem in relation to the overdue BHP processing; (2) meet the E2E delay requirement of services. With reference to the representative network scenario, the configuration problem is solved by means of the optimization procedure. Example studies have been carried out for realistic transport network topologies.

Compared to the statistical delay models in the literature, the proposed deterministic model has advantages in rendering the robust absolute QoS guarantee for individual FEC flows, which is especially appreciated in premium services provisioning. A drawback of the deterministic analysis could be its conservative estimation with respect to the worst case. In this regard, the deterministic delay bounds have been compared with the statistical delay bounds obtained from the widely used Markov models in the literature. It was seen that the deterministic models lead to practical delay bounds in a magnitude that is close to the results of stochastic analysis. In some cases, the deterministic model is able to capture the inherent hardware limitation on the traffic profile and leads to better performance estimations than the statistical model. In the GPS-based traffic scheduling, the deterministic model can assure a 100% resource utilization while confining the delay in the practical magnitude for premium services.

Future study will be related to the design of advanced burst scheduling schemes for the edge node that are more efficient in the bandwidth utilization and provide tight delay bounds at the same time. Furthermore, a comprehensive network evaluation will be carried out by integrating the deterministic delay guarantee into the global QoS architecture of OBS networks.

**Acknowledgments** The study described in this article was carried out with the support of the BONE-project (“Building the Future Optical Network in Europe”), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] Turner, J.S.: Terabit burst switching. *J. High Speed Netw.* **8**(1), 3–16 (1999)
- [2] Qiao, C., Yoo, M.: Optical burst switching (OBS)—a new paradigm for an optical Internet. *J. High Speed Netw.* **8**(1), 69–84 (1999)
- [3] Xiong, Y., Vanderhoute, M., Cankaya, H.C.: Control architecture in optical burst-switched WDM networks. *IEEE J. Sel. Areas Commun.* **18**(10), 1838–1851 (2000)
- [4] Raffaelli, C., Zaffoni, P.: Packet assembly at optical packet network access and its effects on TCP performance. In: *Workshop on High Performance Switching and Routing (HPSR)*, Torino, pp. 141–146, June 2003
- [5] de Vega Rodrigo, M., Goetz, J.: An analytical study of optical burst switching aggregation strategies. In: *Proceedings of the Third International Workshop on Optical Burst Switching (WOBS)*, San Jose, pp. 1–13, October 2004
- [6] Dolzer, K., Gauger, C.M.: On burst assembly in optical burst switching networks—a performance evaluation of just-enough-time. In: *Proceedings of the 17th International Teletraffic Congress (ITC 17)*, pp. 149–160, September 2001
- [7] Izal, M., Aracil, J.: On the influence of self-similarity on optical burst switching traffic. In: *IEEE Global Telecommunications Conference (Globecom)*, vol. 3, pp. 2308–2312 (2002)
- [8] Yu, X., Li, J., Cao, X., Chen, Y., Qiao, C.: Traffic statistics and performance evaluation in optical burst switched networks. *J. Lightwave Technol.* **22**(12), 2722–2738 (2004)
- [9] Hunter, D., Chia, M., Andonovic, I.: Buffering in optical packet switches. *IEEE/OSA J. Lightwave Technol.* **16**(12), 2081–2094 (1998)
- [10] Eramo, V., Listanti, M.: Packet loss in a bufferless optical WDM switch employing shared tunable wavelength converters. *IEEE/OSA J. Lightwave Technol.* **18**(12), 1818–1833 (2000)
- [11] Tancevski, L., Yegnanarayanan, S., Castanon, G., Tamil, L., Masetti, F., McDermott, T.: Optical routing of asynchronous, variable length packets. *IEEE J. Sel. Areas Commun.* **18**(10), 2084–2093 (2000)
- [12] Chich, T., Cohen, J., Fraigniaud, P.: Unslotted deflection routing: a practical and efficient protocol for multihop optical networks. *IEEE/ACM Trans. Netw.* **9**(1), 47–59 (2001)
- [13] Gauger, C.M.: Optimized combination of converter pools and FDL buffers for contention resolution in optical burst switching. *Photonic Netw. Commun.* **8**(2), 139–148 (2004)
- [14] Yoo, M., Qiao, C., Dixit, S.: QoS performance of optical burst switching in IP over WDM networks. *IEEE J. Sel. Areas Commun.* **18**, 2062–2071 (2000)
- [15] Vokkarane, V.M., Jue, J.P.: Prioritized burst segmentation and composite burst-assembly techniques for QoS support in optical burst-switched networks. *IEEE J. Sel. Areas Commun.* **21**(7), 1198–1209 (2003)
- [16] Dolzer, K.: Mechanisms for quality of service differentiation in optical burst switched networks—86th report on studies in congestion theory. Ph.D. dissertation, University of Stuttgart (2004)

- [17] Kaheel, A., Alnuweiri, H.: Quantitative QoS guarantees in labeled optical burst switching networks. *IEEE Globecom* **3**, 1747–1753 (2004)
- [18] Zhang, Q., Vokkarane, V., Jue, J., Chen, B.: Absolute QoS differentiation in optical burst-switched networks. *IEEE J. Sel. Areas Commun.* **22**(9), 1781–1795 (2004)
- [19] Guan, X., Thng, I., Jiang, Y., Li, H.: Providing absolute QoS through virtual channel reservation in optical burst switching networks. *Comput. Commun.* **28**, 967–986 (2005)
- [20] Overby, H., Stol, N.: Providing absolute QoS in asynchronous bufferless optical packet/burst switched networks with the adaptive preemptive drop policy. *Comput. Commun.* **28**(9), 1038–1049 (2005)
- [21] Yang, L., Rouskas, G.: A framework for absolute QoS guarantees in optical burst switched networks. In: *Proceedings of Broadnets*, pp. 1–10 (2006)
- [22] Rosberg, Z., Vu, H.L., Zukerman, M., White, J.: Blocking probabilities of optical burst switching networks based on reduced load fixed point approximation. In: *Proceedings of IEEE INFOCOM*, vol. 3, pp. 2008–2018, March 2003
- [23] Phung, M., Chua, K., Mohan, G., Motani, M., Wong, D.: An absolute QoS framework for loss guarantees in optical burst-switched networks. *IEEE Trans. Commun.* **55**(6), 1191–1201 (2007)
- [24] Callegati, F., Raffaelli, C.: End-to-end performance of an optical transparent packet network. In: *Proceedings of the IEEE International Conference on Communications (ICC)*, Atlanta, pp. 508–513, June 1998
- [25] Rajaduray, R., Ovadia, S., Blumenthal, D.: Analysis of an edge router for span-constrained optical burst switched (OBS) networks. *IEEE J. Lightwave Technol.* **22**(11), 2693–2705 (2004)
- [26] Hu, G., Koehn, M.: Evaluation of packet delay in OBS edge nodes. In: *Proceedings of the 8th International Conference on Transparent Optical Networks (ICTON)*, vol. 3, Nottingham, pp. 66–69 (2006)
- [27] Hu, G.: Performance model for a lossless edge node of OBS networks. In: *Proceedings of the 49th IEEE GLOBECOM Technical Conference*, pp. 1–6, November 2006
- [28] Li, H., Thng, I.L.-J.: Edge node buffer usage in optical burst switching networks. *Photonic Netw. Commun.* **13**(1), 31–51 (2007)
- [29] Choi, J., Choi, J., Kang, M.: Dimensioning burst assembly process in optical burst switching networks. *IEICE Trans. Commun.* **E88-B**(10), 3855–3863 (2005)
- [30] Pantaleo, A., Tornatore, M., Pattavina, A., Raffaelli, C., Callegati, F.: Dimensioning for in-band and out-of-band signalling protocols in OBS networks. *IET Commun.* **3**(3), 418–427 (2009)
- [31] Hu, G., Raffaelli, C., Perin, A.: Deterministic delay guarantee in OBS edge node for premium services. *Trans. Lett. IEEE Trans. Commun.* **57**(6), 1618–1621 (2009)
- [32] Firoiu, V., Boudec, J.-Y.L., Towsley, D., Zhang, Z.-L.: Theories and models for internet quality of service. *Proc. IEEE* **90**(9), 1565–1590 (2002)
- [33] Zhang, H.: Service disciplines for guaranteed performance service in packet-switching networks. *Proc. IEEE* **83**(10), 1374–1396 (1995)
- [34] Stallings, W.: *High-Speed Networks and Internets: Performance and Quality of Service*, 2nd ed. Prentice Hall, Upper Saddle River, NJ (2002)
- [35] Bonald, T., Proutiere, A., Roberts, J.: Statistical performance guarantees for streaming flows using expedited forwarding. In: *Proceedings of IEEE INFOCOM*, vol. 2, pp. 1104–1112 (2001)
- [36] Gauger, C.M.: Trends in optical burst switching. In: *Proceedings of SPIE ITCOM*, pp. 765–775, September 2003
- [37] Lakshman, T., Ortega, A., Reibman, A.R.: VBR video: tradeoffs and potentials. In: *Proceedings of the IEEE*, vol. 86, no. 5, pp. 952fb973, May 1998
- [38] Junghans, S., Gauger, C.M.: Architectures for resource reservation modules for optical burst switching core nodes. In: *Proceedings of the 4th ITG Symposium on Photonic Networks*, Leipzig, pp. 109–117, May 2003
- [39] Kumar, A., Manjunath, D., Kuri, J.: *Communication Networking: an Analytical Approach*. Morgan Kaufmann, Burlington (2004)
- [40] Boudec, J.-Y.L.: Application of network calculus to guaranteed service networks. *IEEE Trans. Inf. Theory* **44**(3), 1087–1096 (1998)
- [41] ITU-T.: *Network Performance Objective for IP-based Service*. International Telecommunication Union, ITU-T, Geneva (2002)
- [42] Pioro, M., Medhi, D.: *Routing, Flow, and Capacity Design in Communication and Computer Networks*. Morgan Kaufmann, Burlington (2004)
- [43] Betker, A., Gerlach, C., Hlsermann, R., Jäger, M., Barry, M., Bodamer, S., Späth, J., Gauger, C.M., Köhn, M.: Reference transport network scenarios. [http://www.pt-it-dlr.de/\\_media/MTN\\_Referenz\\_Netze.pdf](http://www.pt-it-dlr.de/_media/MTN_Referenz_Netze.pdf)

### Author Biographies



**Guoqiang Hu** received his B.S. Degree in Computer Science from the Shanghai Jiao-Tong University, China, in 1997 and the M.S. Degree in Information Technology from the University of Stuttgart, Germany, in 2002. From 2002 to 2007, he was with the Institute of Communication Networks and Computer Engineering, Stuttgart, Germany as a research staff and Ph.D. candidate. Since 2008, he has been a research fellow in the

Centre for Quantifiable Quality of Service in Communication Systems, Trondheim, Norway. His research interests include network design and architectures, teletraffic theory, performance evaluation, admission control, and resource allocation.



**Carla Raffaelli** is an associate professor in Telecommunications at the University of Bologna, Italy. She received the M.S. and the Ph.D degrees in electrical engineering and computer science from the University of Bologna, Italy, in 1985 and 1990, respectively. Since 1985 she has been with the Department of Electronics, Computer Science and Systems of the University of Bologna, Italy, where she was Research Associate

in 1990. Her research interests include performance analysis of telecommunication networks, switching architectures, protocols, and broadband communication. Since 1993, she participated in European funded projects on optical packet-switched networks, the RACE-ATMOS, the ACTS-KEOPS, and the IST-DAVID projects. She was active in the EU funded e-photon/One network of excellence and, at present, in its follow-up, BONE. She also participated in many national research projects on multi-service telecommunication networks. She is the author of many technical articles on broadband switching and network modelling and regularly acts as a reviewer for top international conferences and journals. She is author or co-author of more than 100 conference and journal articles mainly in the field of optical networking and network performance evaluation. She is IEEE member.