# Performance analysis of the CSMA/CA MAC protocol in the DBORN optical MAN network architecture

Guoqiang Hu, Christoph M. Gauger and Sascha Junghans

University of Stuttgart, Institute of Communication Networks and Computer Engineering
Pfaffenwaldring 47, D-70569 Stuttgart, Germany
e-mail: {hu, gauger, junghans}@ikr.uni-stuttgart.de

**Abstract:** In this paper, we present a detailed performance analysis of the MAC protocol of the DBORN optical burst-mode metro network architecture. We introduce an exact analytical performance model for the slotted operation mode and models for the unslotted operation mode which yield good upper and lower bounds. Also, we introduce a new moment analysis approach to derive the mean waiting time of preemptive repeat identical priority queueing systems. Then, we validate the analytical model by simulation and assess important architectural options regarding delay and maximum achievable network load. In order to consider realistic burst traffic characteristics, we extend the evaluation towards general arrival processes and finally towards a burst assembly module with self-similar IP traffic.

## 1. INTRODUCTION

Demands for higher capacity metropolitan area networks (MAN) are continuously rising due to several strong trends in information technology both for private life and business. The huge success of broadband Internet access deployment not only leads to an ever increasing number of Internet users but also to the roll-out of new multimedia services which today are provided over dedicated networks, e.g., cable or satellite. Also, virtualization of business processes and even entire companies is enabled by virtual private LAN services for voice as well as for bandwidth demanding services like video, enterprise resource planning, or shared storage.

After the fast growth in core network capacity over the past years, transport and Internet service providers today ask for equipment with higher bandwidth but lower cost. This motivates the application of optical solutions with wavelength division multiplexing (WDM) technology. However, as active optical switching elements are still costly, an optical network architecture without active optical switching elements is desirable. The DBORN (Dual Bus Optical Ring Network) architecture is a MAN technology suitable for these constraints [4,16].

DBORN connects metro edge nodes, in which IP packets are aggregated into bursts, to the core network via a hub node. As cost tradeoffs lead to some constraints in the area of medium access, a new medium access control (MAC) protocol and interface card design was required [4,16] which we survey in more detail in Section 2.

In this paper, we present for the first time our detailed teletraffic theoretical analysis [8] as well as a comprehensive performance evaluation based on both analysis and simulations. First, we describe the exact mean waiting time model for the slotted mode as well as upper and lower bound models for unslotted mode based on the preemptive repeat identical (PRI) priority

queueing system. Second, a new moment analysis approach is derived for PRI systems for Poisson arrivals. This new method not only simplifies the derivation of the mean waiting time by only using first and second moments of busy period and completion time, but also explains its exact relation to the busy period of high priority customers. Third, we validate the analytical models and bounds by simulation for a Poisson arrival process. The applicability of the Poisson traffic model is then further validated by comparing the results with general independent (GI) traffic models of different variability. We then extend our work to realistic aggregate burst traffic that is assembled from self-similar IP traffic. Following this systematic approach, we show that under most practical circumstances Poisson traffic can be a good approximation or act as a useful worse case traffic model.

Regarding DBORN, only individual aspects have been evaluated and reported in literature so far. Key technology solutions, e.g., burst-mode transceivers, and implementation issues in the context of DBORN are presented in [6]. A protocol which is designed to provide access fairness among all the network nodes and assure QoS for premier traffic is reported in [2,4] and further evaluated in [11]. Work dedicated to system performance of DBORN is available in [3,8,10,11]. While [3] and [8] both present a lower bound analysis, which was independently found in parallel, our presentation in [8] also contains the exact analysis for slotted operation mode, the analytical model for the upper bound based on the new mean value analysis, and a detailed performance evaluation. In [10], we presented simulation results for basic MAC performance, a back-pressure mechanism and the buffer allocation in the transceiver card.

Regarding PRI priority queueing systems, standard analytical models following Jaiswal [14] and Takagi [17] are commonly applied to obtain the Laplace-Stieltjes transformation of the waiting time and then derive its moments. In contrast, we derive the mean waiting time more directly by setting up its relation to the moments of busy period and completion time.

The remainder of this paper is structured as follows. In Section 2, we will survey the DBORN network architecture and its MAC protocol, Section 3 describes the analytical models of the MAC protocol and Section 4 shows the results of the model validation and the performance evaluation studies. In Section 5, conclusions are drawn and further work is outlined.

## 2. NETWORK ARCHITECTURE AND MAC PROTOCOL

DBORN is a high speed network solution for metropolitan areas [16]. On the basis of advances in the optical transmitter and receiver technology [6], the carrier sense multiple access with collision avoidance (CSMA/CA) is realized in DBORN.

### 2.1 Network architecture

DBORN is an optical metro ring architecture connecting several edge nodes, e. g., metro clients like enterprise, campus or local area networks (LAN), to a regional or core network. The ring consists of two parallel fibers called working and protection fiber (Figure 1 left) in order to provide resilience in case of single link failures. Each ring employs WDM and carries a set of wavelengths which are further classified into downstream and upstream wavelength channels (Figure 1 right). While downstream wavelength channels start from the transmitters in the hub, upstream wavelength channels are terminated by the receivers in the hub.

Several edge nodes share upstream and downstream channels respectively in asynchronous time division multiplexing. For load balancing purposes, an edge node can be attached to more than one upstream or downstream channel. In order to keep the edge node interface cards

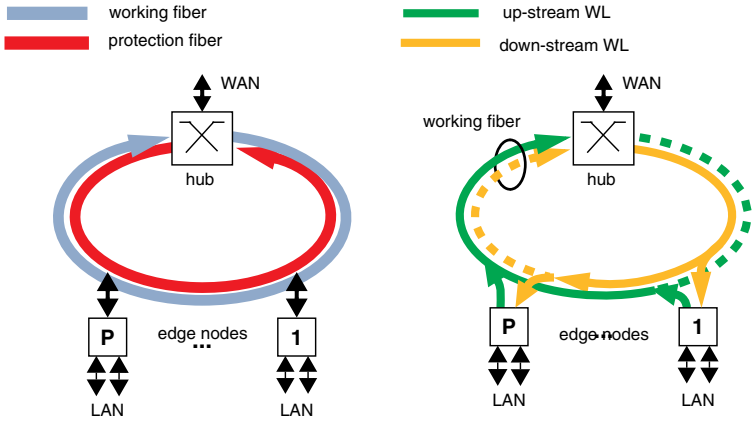


**Figure 1** DBORN architecture

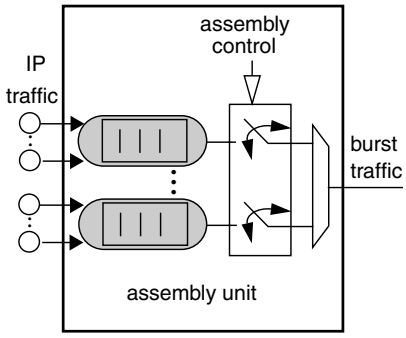


**Figure 2** Burst assembly in the edge node

as simple as possible, all traffic has to pass the hub. Specifically, no edge node receives or even removes traffic on upstream channels or inserts traffic on downstream channels. Thus, both upstream and downstream channels can be modelled as shared unidirectional buses.

As the hub node exclusively transmits on the downstream channel, traditional scheduling mechanisms can be applied here. However, medium access of edge nodes has to be controlled on the upstream channel which will be analysed in depth in the rest of the paper.

## 2.2 Burst size and burst assembly

In order to provide for safe transmitting and receiving on the ring a guard time has to be inserted between consecutive optical transmission units. Typical guard times with current technologies are $50ns$ [6], which results in transmission times of about 63 B on a 10 Gbps link.

DBORN targets transaction data and Internet traffic which is commonly transported over Ethernet, i. e. client layer packet sizes are in the range of 40 to 1500 B [20] bounded by the Ethernet maximum transmission unit (MTU). As transmission of individual client layer packets/frames would lead to a significant overhead due to guard times, all client layer traffic is assembled into larger units called *bursts* for transmission. A considerable amount of literature on burst assembly is available in the context of optical burst switching (e.g. [7][9][19]). The small MTU value only allows a limited degree of assembly gain. In future versions, this could be improved by segmentation of client layer packets [1] or by selection of a different optical layer burst format, e. g. ITU-T's G.709 frame format with a size of about 16 KB [13].

In case the G.709 burst format is applied, a traffic assembler needs to be deployed in the edge node (Figure 2). Incoming IP packets are classified based on the address of the destination node and stored in correspondent assembly queues accordingly. A general assembly algorithm is so-called MaxSize-TimeOut assembly, where the MaxSize $S_{\max}$ refers to the maximal size of the burst and TimeOut $\tau$ refers to a predefined timeout parameter. The timer is set upon arrival of the first IP packet to an empty assembly queue. A burst becomes ready for transmission whenever the accumulated data volume in the assembly queue is about to exceed the MaxSize or the timeout occurs. Therefore, the MaxSize sets the upper bound for the burst size and the timeout serves as an upper bound for the assembly duration of one burst. Burst assembly has an impact on the traffic characteristic and therefore also influences the network performance [5]. This issue will be studied in Section 4.

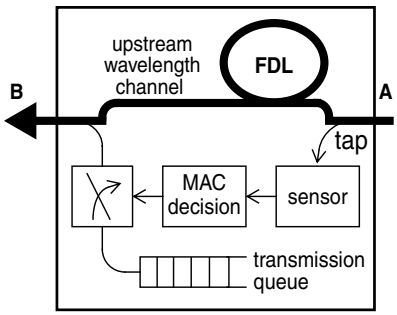


**Figure 3** Functional model of transmitter interface in edge nodes

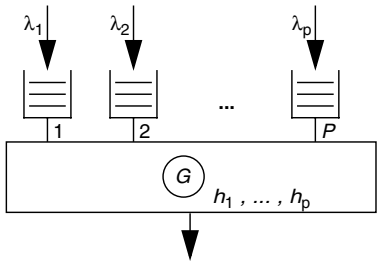


**Figure 4** Priority queueing model

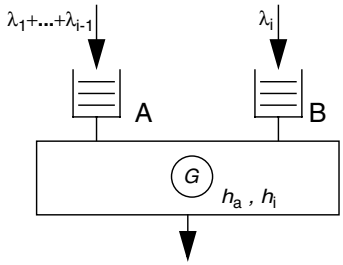


**Figure 5** Approximate resulting model

### 2.3 MAC protocol

As DBORN targets a cost efficient optical ring solution no active optical components, e. g. switches, are used on the interface cards and transmitting and receiving part are strictly separated. Figure 3 depicts a functional model of the transmitter interface, which was designed to allow a collision-free medium access.

Between the input (point A) and the output (point B) of the edge node a fiber delay line (FDL) is inserted into the ring. The length of the FDL should correspond to a delay equal to or greater than the transmission time of the maximum burst size. At the input of the edge node, a simple sensor taps the upstream channel and constantly monitors the channel status—busy or idle. On the other side of the FDL a laser is coupled into the same channel and controlled by the decision unit to send bursts safely. Due to the delay introduced by the FDL, the edge node can determine the duration of voids on the channel up to the FDL delay before they pass the coupling point of the laser and thus decide on the medium access avoiding collisions.

There are two possible operation modes for DBORN: *slotted* and *unslotted*. In the slotted mode, the channel is divided into constant duration slots and the transmission is allowed if the edge node finds an idle slot on the upstream channel. On the one hand, this requires some basic synchronization between network nodes, on the other hand edge nodes only have to check whether a slot is idle or used. In the unslotted mode, no synchronisation is required and bursts can have an arbitrary transmission time up to the FDL delay. By comparing the duration of an available void on the channel and the transmission time of the first burst in the transmission queue the edge node can decide when to transmit a burst.

## 3. ANALYSIS FOR A SINGLE UPSTREAM CHANNEL

In this section, we present our performance model for a single upstream channel applying the CSMA/CA MAC protocol described in Section 2.3. The performance metric is the mean waiting time of an already assembled burst before transmission. In Section 3.1, we deduce the exact mean burst waiting time for slotted mode. Bounds on the mean burst waiting time are given by approximate solutions for unslotted mode in Sections 3.2 and 3.3.

A DBORN edge node can only make use of bandwidth (voids) on the channel which was left over by other nodes located further upstream. This behavior can be modelled by a priority queueing system as illustrated in Figure 4 in which $p$ queues compete for a single server and a queue is only allowed to transmit if all queues of higher priority are empty [8]. Nodes are indexed in ascending order following the traffic flow direction and are abstracted by their

transmission queue. The class priority of the queue is defined by the node index $i$, i.e., a smaller node/queue index corresponds to a higher priority class. Note that the distance between the edge nodes only affects the propagation delay but does not impact the mean waiting time analysis in our scenario and consequently is modelled to be zero here.

For the analysis, we assume that each class injects traffic following a Poisson arrival process with rate $\lambda_i$ and that traffic streams of different nodes are independent of each other. The service time of bursts is independently and identically distributed with mean $h_i$.

### 3.1 Exact analysis of mean waiting time for slotted mode

In slotted operation mode, a fixed slot time is assumed which is equal to the fixed burst transmission time $h$. Thus, an edge node which has a burst to transmit decides on medium access at the slot boundary based on whether the slot is busy or idle. In the corresponding priority system this means that the queues compete for the right of transmission only at slot boundaries, which is described by a slotted priority system without preemption.From a mean value analysis [15] it follows that the mean waiting time $W_i$ of a class-$i$ customer equals

$$E[W_i] = \frac{h}{2} + \sum_{k=1}^{i} E[X_k]h + \sum_{k=1}^{i-1} (E[W_i]\lambda_k)h \tag{1}$$

where $X_i$ denotes the queue length of class $i$. This equation expresses that the mean waiting time experienced by a test customer[1] consists of three parts: (i) the average residual lifetime of a time slot, (ii) the workload of those customers of higher or equal priority who have been in the system upon arrival and will be served prior to the test customer, and (iii) the workload of those customers of higher priority who will arrive after the test customer and will be served prior to the test customer.

From Little's Theorem $E[X_i] = \lambda_i E[W_i]$ an Equ. (1), the mean waiting time follows:

$$E[W_i] = \frac{h/2}{(1-\rho_{\le i})(1-\rho_{\le i-1})} \tag{2}$$

where $\rho_{\le i} = \sum_{k=1}^{i} \lambda_k h$ denotes the total offered traffic of classes $1, \ldots, i$. From Equ. (2) it can be seen that the mean waiting time $E[W_i]$ is always finite as long as the $\rho_{\le i} < 1$, which corresponds to the work-conserving property of the slotted operation mode.

### 3.2 Upper bound on mean waiting time for unslotted mode

In unslotted operation mode, no slot synchronization is available and bursts can have an arbitrary size up to the maximum burst size. An edge node only sends a burst if it can find a void on the channel which is large enough. As a consequence, there may be voids becoming too small to be filled, so called channel fragmentation, and burst transmission is no longer strictly in the priority order of node location but also depends on void and burst sizes.

Thus, the unslotted operation mode does not lend itself to the straightforward analysis used for the slotted mode. Also, arrivals of busy/idle periods on the channel observed by a downstream edge node now follow a correlated random process and thus renewal theory does not apply any more. Still, a *preemptive repeat identical* (PRI) queueing system as alternative, approximate model yields an upper and a lower bound on the mean waiting time.

In a PRI system, a low priority customer in service can be preempted by a newly arriving high priority customer. The turn comes to the preempted customer again after all higher prior-

[1]Arrival customer and outside observer waiting time are identical according to the PASTA theorem.

ity customers are served and service starts from the beginning which is called preemptive repeat. In each repetition a customer's required service time remains identical, i.e., there is no resampling after preemption. A low priority customer can only be completely served if it finds a service time interval without arrivals of higher-priority customers.

This is an essential analogy to the unslotted DBORN MAC protocol. The difference lies only in the fact that an edge node will not start transmission of its burst if the void is too small, so these voids can still be utilized by edge nodes located further downstream, while in the PRI system the customer occupies the server no matter whether he will be preempted or not. As the server capacity wasted by the unfinished service in case of preemption leads to a performance degradation, i. e., increased mean waiting time, for all lower priority classes the PRI system yields an upper bound for the mean burst waiting time.

We use $D_{i,\,\mathrm{PRI}}$ to denote the duration of the interval for a class-$i$ customer in the PRI system between his arrival to the system and the beginning of his effective service period, i.e., the period in which he gets fully served without interruptions. The upper bound condition of the mean burst waiting time in node $i$ can thus be expressed in terms of $E[D_{i,\,\mathrm{PRI}}]$ by

$$E[W_i] \le E[D_{i,\,\mathrm{PRI}}]\,. \tag{3}$$

Note that in Equ. (3) the equality holds for $i = 1, 2$. Only classes with $i \ge 3$, suffer from the non-effective consumption of server capacity due to the preemption.

For the derivation of $E[D_{i,\,\mathrm{PRI}}]$ several intermediate parameters are needed. Completion time $C_i$ denotes the interval between the instant at which a class-$i$ customer enters service for the first time and the final completion of his service. Note that we assume here that the customer leaves the queue when he first starts service, i.e., he does not return to the queue $i$ when being preempted. In this way the completion time can be regarded as "virtual service time" of a customer. The waiting time $W_{i,\,\mathrm{PRI}}$ of a class-$k$ customer refers only to the time he waits in the queue. Busy period $B_i$ denotes the duration of an interval in which there is at least one customer of class $i$ or of higher priority in the system. This corresponds to the time the server is continuously occupied by traffic of class $i$ or of higher priority. Both completion time process and busy period process are renewal processes. $E[D_{i,\,\mathrm{PRI}}]$ can thus be expressed as

$$E[D_{i,\,\mathrm{PRI}}] = E[W_{i,\,\mathrm{PRI}}] + E[C_i] - h_i \tag{4}$$

The solution of $E[C_i]$ is available in [14]. In [17] a derivation of $E[W_{i,\,\mathrm{PRI}}]$ is given by applying complex transformation techniques. Here, we present a novel, more direct but still exact derivation of $E[W_{i,\,\mathrm{PRI}}]$ by using the mean value analysis method. $E[W_{i,\,\mathrm{PRI}}]$ can be further decomposed into two parts

$$E[W_{i,\,\mathrm{PRI}}] = E[X_i]E[C_i] + P_{\mathrm{busy},i}E[\Gamma_i] \tag{5}$$

where $X_i$ denotes the queue length of class $i$ and the first term on the right hand side denotes the completion time of all customers in queue $i$ which arrived earlier and have not yet started service. $\Gamma_i$ represents the residual sojourn time of the class-$i$ customer at the head of the queue before he leaves the queue and will be looked at below in more detail. $P_{\mathrm{busy},i}$ is the probability that the server is in a busy period in terms of class-$i$ customers with

$$P_{\mathrm{busy},i} = E[B_i]/(E[B_i] + 1/\lambda_{\le i})\,. \tag{6}$$

Here, $\lambda_{\le i}$ denotes the total rate of traffic of class $i$ and of higher priority and represents the termination rate of the idle period between two busy periods $B_i$. Note that as arrivals are Poisson and an idle period is terminated by any arrival of classes 1, …, $i$, its duration is negative exponentially distributed with its expected value $1/\lambda_{\le i}$.

Now, we concentrate on $E[\Gamma_i]$ by studying following situations:

1. With probability $\lambda_i/\lambda_{\le i}$ the current busy period $B_i$ starts with the service of a class-$i$ customer. In this case it can be proven[1] that presently there is definitely a class-$i$ customer in the system who has left the queue but not yet finished his service. Therefore, $E[\Gamma_i]$ equals the residual completion time and it yields $E[\Gamma_i] = E[C_i^2]/(2E[C_i])$.
2. If the busy period $B_i$ starts with a higher priority customer or equivalently starts with a busy period of $B_{i-1}$, two possibilities exist:
   2.1. With probability $E[B_{i-1}]/E[B_i]$ the present time falls in this first busy period $B_{i-1}$ contained in the current $B_i$. Then $\Gamma_i$ equals the residual time of the busy period $B_{i-1}$: $E[\Gamma_i] = E[B_{i-1}^2]/(2E[B_{i-1}])$,
   2.2. Otherwise, with the same argument as in case 1, there is $E[\Gamma_i] = E[C_i^2]/(2E[C_i])$.

Based on those arguments, $E[\Gamma_i]$ can finally be expressed as

$$\frac{\lambda_i}{\lambda_{\le i}} \cdot \frac{E[C_i^2]}{2E[C_i]} + \frac{\lambda_{\le i-1}}{\lambda_{\le i}} \cdot \left( \frac{E[B_{i-1}]}{E[B_i]} \cdot \frac{E[B_{i-1}^2]}{2E[B_{i-1}]} + \left(1 - \frac{E[B_{i-1}]}{E[B_i]}\right) \cdot \frac{E[C_i^2]}{2E[C_i]} \right). \tag{7}$$

The first and secondary ordinary moment of $C_i$ and $B_i$ can be calculated for $1 \le i \le p$ according to the iterative formulas in [14], which are also presented in the Appendix. Then, Equ. (6) and (7) can be computed directly. Using Little's Theorem $E[X_i] = \lambda_i E[W_{i,\mathrm{PRI}}]$ and inserting Equ. (6) and (7) into Equ. (5) we obtain $E[W_{i,\mathrm{PRI}}]$. At last, the exact solution for $E[D_{i,\mathrm{PRI}}]$ can be derived from Equ. (4).

### 3.3 Lower bound on mean waiting time for unslotted mode

Based on the case of equality in Equ. (3), we found an alternative approximate resulting model as illustrated in Figure 5. From the point of view of edge node $i$, channel traffic generated by the $i-1$ upstream edge nodes is approximately the same as the traffic generated by one upstream edge node but with same total traffic intensity. This system can be abstracted by a two-class PRI system. Queue A models all upstream nodes and has a traffic arrival rate $\sum_{k=1}^{i-1} \lambda_k$. Queue B represents the observed edge node $i$ with traffic arrival rate $\lambda_i$. The mean waiting time of a class-$B$ customer can be computed as described in Section 3.2.

However, modelling edge nodes $1, \ldots, i-1$ with one queue entirely removes the effect of channel fragmentation introduced by the MAC protocol, i. e., the fact that all edge nodes $2, \ldots, i-1$ experience an additional waiting time due to bursts arriving on the ring and too small voids in between. Thus, this approximation leads to an optimistic estimation of the performance and constitutes a lower bound on the mean waiting time for unslotted operation.

## 4. PERFORMANCE EVALUATION

In this section, the mean waiting time analysis of Section 3 will be validated under the modeling assumptions using simulation. Then, slotted and unslotted mode will be compared, still with the assumption that burst arrivals follow a Poisson process. To analyze MAC protocol performance in a more real network environment, we also look at the cases when burst arrivals follow general independent (GI) process for which we can study different levels of

[1] A hint for the proof: The counter-example arises if and only if a busy period $B_{i-1}$ starts exactly at the instant when a completion time $C_i$ ends. However, this occurs with probability 0.

variation. Finally, we also include a burst assembly unit into the scenario to generate burst traffic from self-similar IP traffic and present the impact on performance.

The evaluation scenario considers a DBORN ring, to which 10 edge nodes are attached on a single 10 Gbps upstream channel. Traffic is homogenous across all nodes regarding demand and burst characteristics. For unslotted mode both fixed burst size and variable burst size are considered, for slotted mode only the fixed size is used. We use the 16 KB maximal burst size (c.f. Section 2.2) which has a transmission duration of 12.8 μs. The term load always refers to the ratio of average traffic bitrate and channel capacity. In all graphs, mean waiting time is normalized by the FDL delay which here equals to 12.8 μs, i.e., the transmission duration of a burst of 16 KB.

### 4.1 Principle behavior and validation of the analytical models

For the slotted mode, the mean waiting time analysis is exact and was found to be in perfect consistency with the simulation results. For the unslotted mode, the upper bound and lower bound calculated according to Section 3.2 and Section 3.3 are depicted in Figure 6 for fixed burst size and in Figure 7 for variable burst size over node index. In the variable size case, we use independent discrete uniform distributions between 5058 B and 16 KB to cover a broad spectrum of burst size variability in the presence of a fixed upper bound of 16 KB.

First of all, the curves for the two approximations bound the simulation results very well. The bounds are tighter for upstream nodes and scenarios with lower load, which can be explained by the smaller channel fragmentation in both cases. It can be observed that for a total network load up to 0.7 the mean waiting time is less than 20 FDL delay time which is about 0.25 ms. Downstream nodes experience a larger delay due to the intrinsic location priority property of the DBORN MAC protocol. However, at small and medium load levels, the mean waiting time different between the edge nodes is not really prominent.

Comparing the figures, it can be found that the mean waiting time is higher in the fixed size than in the variable size case which is due to the smaller mean burst size. This effect can already be explained by standard queueing models, but also by the fact that the smaller burst size also increases the probability of fitting a burst into a void.

### 4.2 Maximal network load

A key techno-economic metric for a network operator is the maximal network load for given quality of service (QoS) constraints. As DBORN does not incur burst loss and as the last edge node has the worst performance, we define the critical QoS constraint in DBORN as the delay of the last node. For a given mean waiting time in the last node of $w$, the bound on the network load can be exactly calculated from Equ. (2) for slotted mode and by numerically solving the upper bound in Equ. (4) for unslotted mode.

In Figure 8, the maximal network load is drawn with respect to the number of total edge nodes in the network for slotted and unslotted mode and a delay constraint $w$ of 5 and 10 times the FDL delay. Again, in unslotted mode we distinguish the fixed and variable burst size case and use the parameters as in Section 4.1. In general, it can be observed that the admissible network load gets smaller with large number of network nodes. In a priority system, more priority levels generally yield a larger absolute performance differentiation at a given total load. Correspondingly, the last edge node is more discriminated when the node number increases. However, it can be also seen in Figure 8 that such a dependence between the load bound and the number of edge nodes decreases when the network is large.

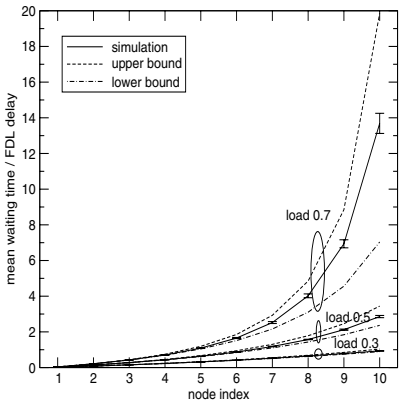


**Figure 6** Mean waiting time wrt. node position for fixed burst size

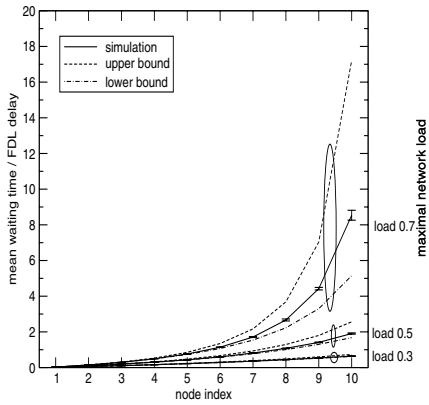


**Figure 7** Mean waiting time wrt. node position for variable burst size

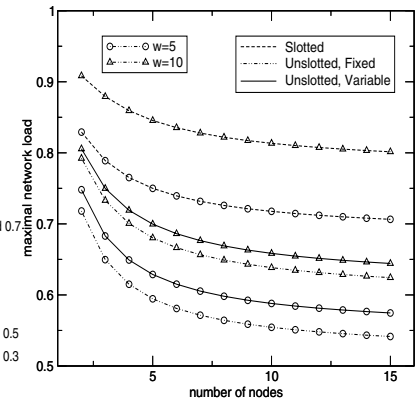


**Figure 8** Maximal network load

Slotted mode outperforms the unslotted mode by more than approx. 10% in the maximal operational network load as it causes no bandwidth fragmentation which we also showed by extensive simulation results in [10]. In unslotted mode, the variable burst size case is a little better than the fixed burst size case due to the smaller mean transmission time of one burst.

In the following, we focus on the performance of the unslotted mode which is preferred for DBORN due to its flexibility in channel access as well as in implementation.

### 4.3 Impact of burst interarrival time distribution

To study the impact of burst interarrival time variability, we use a GI arrival process based on a phase type model for the probability distribution. Through proper parameter settings, the coefficient of variation (CoV) $c$ is tuned to 0.5, 0.8, 1 and 2 respectively. For a load of 0.7 and a fixed burst size of 16 KB, Figure 9 shows the mean waiting time over node index. Variable burst size leads to very similar system behavior and is thus not shown here. For comparison, the analytical upper bound and lower bounds for the Poisson arrival process are also sketched. Note $c = 1$ corresponds exactly to the Poisson arrival process.

It can be seen that the performance is very sensitive to the traffic variability in the interarrival time. A large value of CoV ($c = 2$) leads a much worse fairness performance which cannot be estimated from the analytical performance bounds any more. As long as $c < 1$ the Poisson burst traffic can act as a worst case assumption. Especially, when the CoV is close to 1 ($c = 0.8$), the delay performance in each edge node is also similar to that of Poisson traffic.

### 4.4 Assembled burst traffic

Now, the burst assembler described in Section 2.2 is included in each of the 10 edge nodes. IP packets are classified based on the node which serves as gateway (9 other edge nodes and the hub node) according to their destination address. For this study, we restrict ourselves to a uniform traffic pattern. We apply an M/Pareto model to generate self-similar IP traffic [18]. Data *sessions* (we avoid the term *burst* here) arrive according to a Poisson process. The size of a session follows the Pareto distribution with mean value 10 KB, minimum value 3750 B and a shape parameter of 1.6 which corresponds to a Hurst parameter of 0.7. Data sessions are segmented into packets with maximal packet size of 1000 B and are then sent at a constant rate of 100 Mbps which corresponds to the access link rate of an end user.

The burst assembler parameters are $S_{\max} = 16$ KB and $\tau = \alpha S_{\max} / R_{\text{in}}$. Here $R_{\text{in}}$ denotes the average IP traffic arrival rate observed by each assembly queue. $\alpha$ is a scalar to

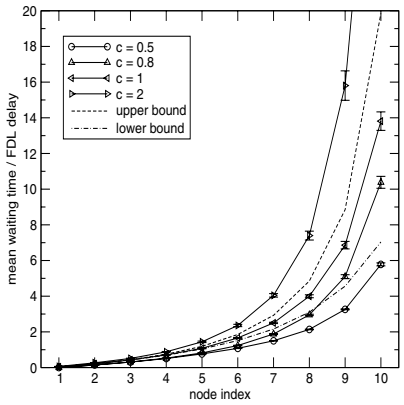


**Figure 9** Mean waiting time with GI burst arrvial and fixed burst size

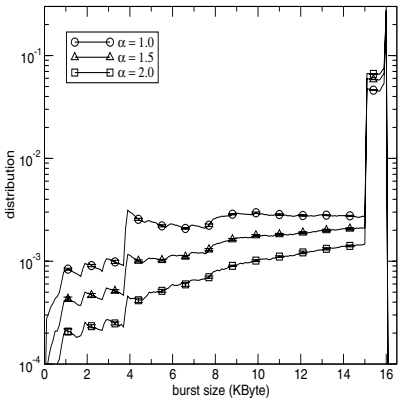


**Figure 10** Distribution of burst size with combined burst assembly

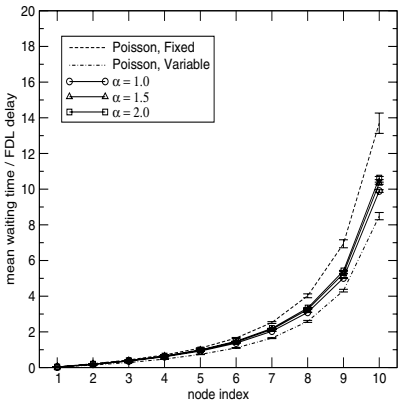


**Figure 11** Mean waiting time with assembled burst traffic

tune the values of the timeout $\tau$. As we assume a well-designed or adaptive burst assembler we do not consider small values of $\alpha$ and only use 1.0, 1.5, and 2.0, respectively.

First, the distribution of the burst size is plotted in Figure 10 for a network load 0.7. It can be seen that the density is mainly located in the range larger than the minimum session size of 3750 B. As the relatively large timeout values (about 1.8, 2.7, and 3.7 ms for $\alpha = 1.0, 1.5, 2.0$) are greater than the minimum (0.3 ms) and even mean (0.8 ms) session duration, it is very unlikely that less data than the minimal session size is collected during $\tau$. On the other side, $S_{\max}$ =16 KB bounds the distribution. In general, the distribution looks like a combination of uniform distribution (in the range of 3750 and 15 KB) and deterministic probability distribution (in the area close to 16 KB). Larger values of $\alpha$ shift the weight of the distribution to the right end, i.e., the distribution approaches the fixed burst size case.

Figure 11 depicts the normalized mean waiting time for bursts arriving from the assembler and from a Poisson process with fixed (16 KB) and variable burst size (uniform distribution between 3750 and 16 KB). We observe that the performance difference caused by the assumption of Poisson burst arrivals is not prominent. This can be explained by the fact that the aggregate burst traffic in each node is multiplexed from 10 burst assembly queues. Although the burst traffic from each assembly queue is rather smooth [19] multiplexing increases the traffic variability again and finally makes the system performance approaching that under Poisson traffic [12]. Consequently, the Poisson assumption and fixed burst size can together serve as a worst case scenario: With a large number of assembly queues the CoV of the multiplexed burst interarrival time is very close to the 1 (about 0.9 in the scenarios here) and the Poisson process models burst arrivals well. This can be seen in Figure 11 as the mean waiting time with assembled bursts is located in between the case of Poisson arrivals with fixed burst size and Poisson arrival with variable burst size.

## 5. CONCLUSION AND OUTLOOK

This paper models and studies the performance of the MAC protocol of the DBORN optical burst-mode metro network architecture. We describe and successfully validate our analytical performance models for Poisson arrivals. The model for slotted operation mode is exact while the approximate models for unslotted operation lead to upper and lower bounds. Also, a new analysis approach for the mean waiting time of preemptive repeat identical priority queueing systems based on moment analysis is given and applied to the models for unslotted mode.

Our performance evaluations show that unslotted operation with variable burst size performs better than with fixed size and almost as good as slotted operation with fixed size. Simulations with more realistic traffic models like general independent arrival distributions or real burst traffic assembled from self-similar IP traffic show that the analytical Poisson models with fixed and with uniformly distributed burst length are very useful as performance bounds.

Future work should extend the models towards network dimensioning, evaluate the performance to a wider set of scenarios, e.g., non-uniform traffic, and include QoS differentiation.

# 6. APPENDIX

Iterative formulas [14] for the solutions of the first and secondary ordinary moment of the completion time $C_i$ and busy period $B_i$:

$$E[C_i] = \left(\frac{1}{\lambda_{\le i-1}} + E[B_{i-1}]\right)(E[e^{\lambda_{\le i-1} T_{H,i}}] - 1) \tag{8}$$

$$\begin{aligned} E[C_i^2] = {} & 2\left(\frac{1}{\lambda_{\le i-1}} + E[B_{i-1}]\right)^2 E[(e^{\lambda_{\le i-1} T_{H,i}} - 1)^2] \\ & + \left(E[B_{i-1}^2] + \frac{2E[B_{i-1}]}{\lambda_{\le i-1}} + \frac{2}{\lambda_{\le i-1}^2}\right)(E[e^{\lambda_{\le i-1} T_{Hi}}] - 1) \\ & - 2\left(E[B_{i-1}] + \frac{1}{\lambda_{\le i-1}}\right)E[T_{Hi} e^{\lambda_{\le i-1} T_{Hi}}] \end{aligned} \tag{9}$$

$$E[B_i] = \frac{\lambda_i}{\lambda_{\le i}} \cdot \frac{E[C_i]}{1 - \lambda_i E[C_i]} + \frac{\lambda_{\le i-1}}{\lambda_{\le i}} \cdot \frac{E[B_{i-1}]}{1 - \lambda_i E[C_i]} \tag{10}$$

$$E[B_i^2] = \frac{\lambda_i}{\lambda_{\le i}} \cdot \frac{E[C_i^2]}{(1 - \lambda_i E[C_i])^3} + \frac{\lambda_{\le i-1}}{\lambda_{\le i}}\left(\frac{E[B_{i-1}^2]}{(1 - \lambda_i E[C_i])^2} + \frac{\lambda_i E[B_{i-1}] E[C_i^2]}{(1 - \lambda_i E[C_i])^3}\right) \tag{11}$$

where $1 < i \le p$ and $T_{H,i}$ is the service time random variable of class-$i$ with mean $E[T_{H,i}] = h_i$. For class-1, it is exactly a M/G/1 queue, i.e., the iterative computation is initiated by $E[C_1] = h_1$, $E[B_1] = h_1/(1 - \lambda_1 h_1)$ and $E[B_1^2] = E[T_{H,i}^2]/(1 - \lambda_1 h_1)^3$ as in [15].

# 7. REFERENCES


[1] T. Atmaca, D. Popa: "Logical performance of the optical packet metropolitan ring architecture." *Proceedings of 17th International Symposium on Computer and Information Sciences (ISCIS)*, Orlando, Oct. 2002.

[2] N. Bouabdallah, L. Ciavaglia, E. Dotaro, N.L. Sauze: "Matching fairness and performance by preventive traffic control in optical multiple access networks." *Proc. OptiComm 2003*, Dallas, Oct. 13-17, 2003.

[3] N. Bouabdallah, A. Beylot, G. Pujolle: "Fairness issues in burst-based optical access networks." *NETWORKING 2004*, Athens, May 9-14, 2004, pp. 914-925

[4] N. Bouabdallah, E. Dotaro, L. Ciavaglia, N.L. Sauze, G. Pujolle: "Resolving the fairness issue in bus-based optical access networks." *IEEE Optical Communications*, Nov. 2004, pp. S12-S18

[5] K. Dolzer, C. M. Gauger, On burst assembly in optical burst switching networks - a performance evaluation of Just-Enough-Time Proceedings of the 17th International Teletraffic Congress (ITC 17), 2001.

[6] G. Eilenberger, L. Dembeck, W. Lautenschläger, J. Wolde: "Optische Burst-Techniken in Metronetzen." *ITG-Fachtagung Photonische Netze 2002*, pp. 13-18

[7] C.M. Gauger: "Trends in Optical Burst Switching." *Proc. SPIE ITCom 2003*, Orlando/FL, Sept. 2003.

[8] G. Hu, C.M. Gauger, S. Junghans: “Analysis of the CSMA/CA MAC Protocol in a New Optical MAN Network Architecture“, University of Stuttgart, IKR, Prof. P. J. Kühn, Internal Report No. 48, Febr. 2004.
[9] G. Hu, K. Dolzer, C.M. Gauger: “Does burst assembly really reduce the self-similarity.” *Optical Fiber Communication Conference (OFC 2003)*, Atlanta, Mar. 2003, pp. 124-126.
[10] G. Hu, C.M. Gauger, S. Junghans: “Performance Evaluation of the DBORN MAC Layer.” *Beiträge zur 5. ITG-Fachtagung Photonische Netze*, Leipzig, May 2004
[11] G. Hu, C.M. Gauger, S. Junghans: “Performance of MAC Layer and Fairness Protocol for the Dual Bus Optical Ring Network (DBORN).” *Proceedings of ONDM 2005*, Milan, 2005, pp. 467-476
[12] M. Izal, J. Aracil: “On the influence of self-similarity on optical burst switching traffic.” *Proceedings of IEEE GLOBECOM 2002,* Taipei, Nov. 2002
[13] ITU-T Recommendations on the OTN Transport Plane, G.709
[14] N.K. Jaiswal: “Priority queues.” Academic Press, New York and London, 1968
[15] L. Kleinrock: “Queueing systems.” Volume I/II, John Wiley & Sons, 1975
[16] N.L. Sauze, A. Dupas, E. Dotaro, L. Ciavaglia, M.H.M Nizam, A. Ge, L. Dembeck: “A novel, low cost optical packet metropolitan ring architecture.” *Proceedings of ECOC 2001*, Amsterdam, Oct. 2001.
[17] H. Takagi: “Queueing analysis." Volume I, NORTH-HOLLAND, 1991, pp. 365-373
[18] R. G. Addie, M. Zukerman and T. D. Neame, "Broadband traffic modeling: simple solutions to hard problems", IEEE Communication Magazine, August 1998, pp. 88-95.
[19] X. Yu, Y. Chen, C. Qiao: “A study of traffic statistics of assembled burst traffic in optical burst switched networks.” *Opticomm 2002*, pp. 149-159
[20] http://www.caida.org/analysis/AIX/plen_hist/