QoS Provision with EDF Scheduling, Stochastic Burstiness and Stochastic Guarantees

Erez Biton and Ariel Orda *

Many of the typical network input processes can be characterized only stochastically. Consequently, only stochastic QoS guarantees can be provided to such traffic profiles. Such guarantees are usually sufficient, in view of the elasticity of typical (QoS demanding) applications, which translates into looser QoS requirements. Moreover, with stochastic guarantees, better network utilization can be achieved. Accordingly, this study considers QoS provision schemes for connections with stochastic traffic profiles and stochastic QoS requirements. We concentrate on the class of Rate-Controlled Earliest Deadline First (RC-EDF) scheduling disciplines, which have several well known advantages, in particular simplicity of implementation and flexibility. Assuming the Exponentially Bounded Burstiness (EBB) traffic model, we establish results that extend the deterministic study of RC-EDF, both for a single server in isolation and for networks of servers. For a single traffic shaper followed by an EDF scheduler, we establish stochastic bounds on the distribution of the delay for each session. In the general (multi-hop) setting, we first establish stochastic bounds on the distribution of the end-to-end delay for traffic shaper elements in series; then, we establish stochastic bounds for RC-EDF networks. The application of these bounds is illustrated through an example, in which our bounds practically outperform the corresponding bounds for RPPS-PGS networks with EBB traffic.

1. Introduction

Emerging Broadband networks are expected to provide real time and multimedia applications with various Quality of Service (QoS) guarantees. One of the major problems in the provision of QoS guarantees is to identify a feasible path that can meet the QoS requirements. Schedulability conditions, as well as worst case bounds on the delay, constitute major tools for quantifying the ability of a path to meet the QoS requirements.

The ability to support QoS requirements depends on the scheduling policy employed in the nodes. In this paper, we consider the Rate-Controlled Earliest-Deadline-First (RC-EDF) scheduling discipline [1]. In the rate-controlled class of service disciplines, the traffic of each connection is reshaped at every node to ensure that the traffic offered to the scheduler conforms to specific characteristics. Reshaping makes the traffic at each node more predictable and, therefore, simplifies the task of guaranteeing performance to individual connections. When used with a particular scheduling policy, it allows the specification of worst case delay bounds at each node. End-to-end bounds can then be computed as

*Authors are with the Department of Electrical Engineering, Technion, Haifa, Israel. E-mail: {berez@tx,ariel@ee}.technion.ac.il
the sum of worst-case delay bounds at each node along the path. Some important advantages of a rate-controlled service discipline, especially when compared with Generalized Processor Sharing (GPS) [2], are simplicity of implementation and flexibility [3]. Thus, a rate-controlled service discipline is often a better solution in terms of scalability.

The EDF scheduling policy associates a per-hop deadline with each packet and schedules packets in the order of their assigned deadlines. Some important properties have been established for this scheduling policy, in the context of a deterministic setting. In [4], exact schedulability conditions have been established; these conditions detect violations of delay guarantees in a network of EDF switches. In [5], EDF was proven to be an optimal scheduling discipline; that is, if a set of tasks is schedulable under any scheduling discipline, then this set is also schedulable under EDF. Also, RC-EDF was proven to outperform GPS in providing end-to-end delay guarantees in a network [3].

As mentioned, the above results have been established solely within a deterministic setting. Under such a setting, the input traffic bursts are assumed to be of bounded length. This is not the case with many typical input processes, e.g., traffic aggregates. Hence, a setting that considers the stochastic nature of the traffic is desired. Under a stochastic setting, only stochastic QoS can be guaranteed, i.e., it is guaranteed that the end-to-end delay experienced by a high percentage of the packets does not significantly exceed the required delay. Such guarantees are appropriate for many applications, in particular multimedia applications, which can tolerate a certain amount of loss due to either late arrival or buffer overflow. Furthermore, schemes that guarantee no loss have a low connection-carrying capacity for bursty traffic. In other words, with stochastic guarantees, better network utilization can be achieved. Consequently, several studies have investigated the provision of stochastic QoS guarantees for stochastic traffic profiles, e.g., [6–8]; however, these studies were carried only in the context of the GPS scheduling discipline.

In view of the significant practical advantages of the RC-EDF service discipline, in this study we investigate QoS provision through RC-EDF scheduling for connections with stochastic traffic profiles and stochastic QoS requirements. Specifically, we consider Exponentially Bounded Burstiness (EBB) processes, which appropriately model typical input processes [9]. First, we study the single node case. We derive schedulability conditions for the EDF scheduling discipline as well as stochastic bounds on the delay experienced by a packet entering an EDF scheduling element. Furthermore, we introduce the concept of EBB traffic shapers, and derive stochastic bounds on the delay experienced by any packet entering such a traffic shaper. Next, we study the multiple node (i.e., multi-hop path) case, and derive a stochastic bound on the end-to-end delay. This bound, for a network of EDF schedulers, packetized EBB traffic and stochastic QoS guarantees, is the main contribution of this study.

We illustrate the application of our derived end-to-end bound through an example. In particular, we demonstrate that this bound, for RC-EDF networks and EBB traffic, practically outperforms the corresponding bound for RPPS-GPS networks and EBB traffic [7].

Some previous studies also considered QoS provisioning under a stochastic setting [10–12]. However, those studies dealt with deterministic input traffic and deterministic traffic shaping (dual-leaky-bucket). Our stochastic model and framework are different. In our
case, the stochasticity of the end-to-end guarantees is (solely) due to the stochastic nature of the session input traffic, whereas in [10–12], the exploitation of statistical multiplexing results in a provision of (only) stochastic guarantees even for deterministically bounded input traffic.

The rest of the paper is structured as follows. In Section 2, we formulate the model. Next, in Section 3, we study the provisioning of QoS with EDF scheduling and stochastic guarantees: first, we consider a service element in isolation; then, we study the multiple node case and derive a stochastic bound on the end-to-end delay; finally, we present some numerical and simulation results, carried on an example network. In Section 4, we conclude the paper and discuss possible future work.

Due to space constraints, all proofs and some technical details are omitted, and can be found in [13].

2. Model Formulation

We consider a store-and-forward network comprising of packet switches in which a packet scheduler is available at each output link. Packetized traffic from a particular connection entering the switch passes through a traffic shaper before being delivered to the scheduler. The traffic shaper regulates traffic, so that the output of the shaper satisfies certain pre-specified traffic characteristics. We consider the traffic model introduced in [9], of Exponentially Bounded Burstiness (EBB) processes. The traffic shaper reshapes the incoming traffic by delaying the packets so that the output is EBB, and then delivers them to the scheduler. We focus on the Earliest Deadline First (EDF) scheduling discipline. The EDF scheduler associates a deadline $t + \delta j$ with each packet of a session $j$ that arrives at time $t$. The packets are served in the order of their assigned deadlines.

The network is represented by a directed Graph $G(V,E)$, in which nodes represent switches and arcs represent links. $V$ is the set of nodes and $E$ is the set of links interconnecting them. Each link $l \in E$ is characterized by a service rate $r_l$. We assume that link propagation delays are negligible. Let $L^j$ be the maximal packet size of a session $j$ and $L_{\text{max}}$ be the maximal packet size in the network.

We assume a discrete time domain, in which the amount of information transmitted on a link with capacity $r = 1$ during one time slot is regarded as a unit of data. In this context, the concept of traffic stream $A(t)$, $t \in \mathbb{R}$, reduces to a sequence $\{A(t)\}_{t \in \mathbb{N}}$ of random variables.

Exponentially Bounded (EB) and Exponentially Bounded Burstiness processes are defined as follows. A stochastic process $A(t)$ is Exponentially Bounded (EB) with parameters $(\Lambda, \alpha)$ if, for any $t$ and any $\sigma \geq 0$, $\Pr\{A(t) \geq \sigma\} \leq \Lambda e^{-\alpha \sigma}$. Let $A(t)$ be the instantaneous traffic rate. $A(t)$ has Exponentially Bounded Burstiness (EBB) with parameters $(\rho, \Lambda, \alpha)$, if for any $s, \tau$ and any $\sigma \geq 0$, $\Pr\{A[s, \tau] \geq \rho (\tau - s) + \sigma\} \leq \Lambda e^{-\alpha \sigma}$, where $A[s, \tau] = \sum_{n=s+1}^{\tau} A(n)$.

3. QoS Provisioning with Stochastic Guarantees

As mentioned, we consider a network with EDF schedulers and traffic streams with EBB profiles. Our goal is to derive a stochastic bound on the end-to-end delay. For that,
we need to formulate a stochastic version of the class of rate-controlled service disciplines introduced in [1].

We start by considering the basic single node case, namely, a (single) traffic shaper followed by a (single) EDF scheduler. Then, we consider the multiple node case, namely, a path of interconnected such elements.

3.1. The single node case

Here, we consider a service element in isolation, which consists of an EBB traffic shaper followed by an EDF scheduler. First, we generalize the notion of EDF-schedulability in order to apply to the considered stochastic setting, and we derive the corresponding schedulability conditions. With that at hand, we establish bounds on the delay experienced by a packet entering a single EDF scheduler. Next, we introduce the notion of an EBB traffic shaper. Here, we establish the rule for delaying packets in the traffic shaper so that the traffic delivered to the scheduler conforms with the required EBB characteristics. This rule is presented as an upper bound on the delay distribution.

3.1.1. EDF Scheduler

An Earliest-Deadline-First (EDF) scheduler assigns each arriving packet a time stamp corresponding to its deadline, i.e., a packet from connection \( j \) with a deadline \( \delta^j \) that arrives at the scheduler at time \( t \) is assigned a time stamp of \( t + \delta^j \). The EDF scheduler always selects the packet with the lowest deadline for transmission. Under the deterministic setting, if a packet is not transmitted by its time stamp then a deadline violation has occurred. A set of connections is said to be schedulable if deadline violations never occur. Obviously, the above definition cannot apply to the stochastic setting.

Consider a set \( \mathcal{N} \) of connections where each connection’s traffic rate is \( \{A^j(t)\}_{j \in \mathcal{N}} \). Each connection \( j \) requires some stochastic delay guarantees, denoted by \( \Pr\{D^j(t) \geq \delta^j + d\} \leq f^j(d) \) for any \( d \geq 0 \). Then, the stochastic schedulability of a set of connections entering an EDF server is defined as follows.

**Definition 1** Given are a scheduler and a set \( \mathcal{N} \) of connections, where each connection \( j \in \mathcal{N} \) is characterized by an EBB traffic profile \( (A^j, \delta^j) \) and by stochastic delay requirements \( \Pr\{D^j(t) \geq \delta^j + d\} \leq f^j(d) \forall j \in \mathcal{N} \). The set of connections is said to be EDF-schedulable if, for all \( t > 0 \), the required stochastic delay guarantees are provided for each connection.

**Proposition 1** A set \( \mathcal{N} \) of connections is EDF-schedulable if for all \( k \in \mathcal{N} \), for all \( t \) and for any \( d \geq 0 \):

\[
\sum_{\tau=0}^{t} \Pr \left\{ \sum_{j \in \mathcal{N}} A^j \left[ t - \hat{\tau}, t + \delta^k - \delta^j \right] + \max_{\delta^j > \tau + \delta^k} L^j > r \left( \hat{\tau} + \delta^k + d \right) \right\} \leq f^k(d) \tag{1}
\]

With these schedulability conditions at hand, we show that the delay experienced by a packet entering an EDF scheduler is exponentially bounded (i.e., EB).
Proposition 2 Suppose that \( \{A_j\}_{j \in \mathcal{N}} \) are \( |\mathcal{N}| \) independent \( (\rho^j, \Lambda^j, \alpha^j) \)-EBB processes sharing an EDF server with delay assignment \( \{\delta^j\}_{j \in \mathcal{N}} \). Then, at any time \( t \), for any \( d \geq \delta^k \) and for all \( k \in \mathcal{N} \),

\[
\Pr \{ D^k(t) \geq d \} \leq \tilde{\Lambda} e^{-\tilde{\alpha} d} \tag{2}
\]

where \( \frac{1}{\tilde{\alpha}} = \sum_{j \in \mathcal{N}} \frac{1}{\alpha^j}, \quad \tilde{\Lambda} = \frac{1}{1-e^{-\sum_{j \in \mathcal{N}} \alpha^j}} \left( e^{-\tilde{\alpha} \left( \sum_{j \in \mathcal{N}} \rho^j \delta^j - \left( \sum_{j \in \mathcal{N}} \rho^j \right) \delta^k - L_{\max} \right)} \right). \]

3.1.2. EBB Traffic Shaper

The \((\rho, \Lambda, \alpha)\)-shaper has one input link and one output link with equal rate. The shaper receives an arbitrary stream on the input link and it buffers data, if necessary, so that the output stream transmitted in the output link has EBB with parameters \((\rho, \Lambda, \alpha)\).

Suppose that the rate of the traffic input to a shaper \( A_1 \) is represented by \( A_0 \). Let \( s_i \) be the time at which the \( i \)th packet starts to arrive on the input link, and \( L_i \) is the length in bits of the packet. Suppose that the packet exits on the output link at time \( t_i \) and let \( A_1 \) represent the rate of traffic exiting the shaper. The shaper transmits packets on the output link in an FCFS order such that

\[
\Pr \{ W_\rho (A_1) (t) \geq \sigma \} \leq \Lambda e^{-\alpha \sigma} \tag{3}
\]

where \( W_\rho (A_1) (t) = \max_{s \leq t} \{ A_1 (s, t) - \rho (t - s) \} \).

Proposition 3 Let the delay suffered by the \( i \)th packet, \( D^{(A_0,A_1)}_i = t_i - s_i \), be as small as possible subject to the constraint (3). It holds that\(^2\)

\[
\Pr \left\{ D^{(A_0,A_1)}_i \leq d \right\} = \begin{cases} 0 & d < t_{i-1} - s_i \\ \leq \Lambda e^{-\alpha \left( W_\rho (A_0) (s_i) - \rho d \right)} & t_{i-1} - s_i \leq d < \frac{W_\rho (A_0) (s_i)}{\rho} \\ 1 & d \geq \frac{W_\rho (A_0) (s_i)}{\rho} \end{cases} \tag{4}
\]

3.2. The multiple node case

We proceed to consider a sequence of service elements, which constitute a path in the network.

3.2.1. Traffic shaper elements in series

First, we study the effect of connecting a sequence of some \( n \) \((\rho, \Lambda, \alpha)\)-shaper elements, say \( \{(\rho, \Lambda_1, \alpha_1), (\rho, \Lambda_2, \alpha_2), \ldots, (\rho, \Lambda_n, \alpha_n)\} \). Let \( (\rho, \Lambda_k, \alpha_k) \) be the parameters of the \( k \)th shaper \( A_k \). \( A_k \) represents the rate of the traffic output of the \( k \)th shaper and \( A_0 \) represents the rate of the traffic input to the system. Let \( D^{(0,k)}_i \) be the difference between the time at which the \( i \)th packet begins to exit the \( k \)th shaper \((t^k_i)\) and the time at which it begins to arrive to the system \((s_i)\).

\(^2\)Note that \( W_\rho (A_0) (s_i) \) is the size of the backlog at time \( s_i \) in a virtual work-conserving system, which accepts traffic at the rate \( A_0 \) and transmits at the rate \( \rho \). This value is known at the time \( s_i \) that the \( i \)th packet arrives.
Proposition 4 For all $k = 1, 2, \ldots, n$ it holds that:

$$\Pr \left\{ D^{(0,k)}_i \leq d \right\} = \begin{cases} 
0 & d < t^{k}_{i-1} - s_i \\
\min_{m=1,2,\ldots,k} \Lambda_m e^{-\alpha_m(W_\rho(A_0)(s_i) - \rho d)} & t^{k}_{i-1} - s_i \leq d < \frac{W_\rho(A_0)(s_i)}{\rho} \\
1 & d \geq \frac{W_\rho(A_0)(s_i)}{\rho} 
\end{cases} \tag{5}$$

Consider an EBB stream $A_0$ with the parameters $(\rho, \Lambda_0, \alpha_0)$ entering a sequence of EBB traffic shapers.

We have $\Pr \left\{ W_\rho(A_0)(s_i) \geq \sigma \right\} \leq \Lambda_0 e^{-\alpha_0 \sigma}$ and, from (5), $D^{(0,k)}_i \leq \frac{W_\rho(A_0)(s_i)}{\rho}$. Therefore,

$$\Pr \left\{ D^{(0,k)}_i \geq d \right\} \leq \Lambda_0 e^{-\alpha_0 \rho d} \tag{6}$$

We note that the upper bound in (6) is loose. This looseness results from taking $D^{(0,k)}_i = \frac{W_\rho(A_0)(s_i)}{\rho}$. That is, assuming that the sequence of shaper elements delays all packets until the backlog $W_\rho(A_k)(t^k_i)$ is cleared.

3.2.2. EDF-Scheduler and EBB-Shaper in Cascade

Our goal is to derive an upper bound on the tail distribution of the delay experience by a packet entering a system of a (single) EDF-scheduler followed by a (single) EBB-shaper. To that end, we start by considering two systems, $S_1$ and $S_2$, where system $S_1$ consists of a $(\rho, \Lambda, \alpha)$-shaper and system $S_2$ consists of a “delay subsystem” and an identical $(\rho, \Lambda, \alpha)$-shaper connected in series as depicted in Fig. 1. The “delay subsystem” delays the $i$th arriving packet by $\theta_i \geq 0$ and then delivers it to the shaper. The following lemma relates the delays experienced by a packet in the two systems $S_1$ and $S_2$. More precisely, the lemma states that the delay distribution in system $S_2$ is upper-bounded by the same function that upper bounds the distribution of the delay in system $S_1$.

Lemma 1 Assume that packets arrive to systems $S_1, S_2$ according to the same arrival process $A_0$. If $D^{(1)}_i$ and $D^{(2)}_i$ are the delays of packet $i$ in the traffic shaper in systems $S_1$

Figure 1. The Systems $S_1$ and $S_2$  
Figure 2. A scheduler followed by a shaper
and $S_2$ respectively, then, for all $i = 1, 2, \ldots$,

$$
\Pr \left\{ D_i^{(2)} + \theta_i \leq d \right\} = \begin{cases} 
0 & d < t^{(1)}_{i-1} - s^{(1)}_i \\
\Lambda e^{-\alpha \left(W_{\rho}(A_0) (s^{(1)}_i - \rho d) \right)} & t^{(1)}_{i-1} - s^{(1)}_i \leq d < \frac{W_{\rho}(A_0) (s^{(1)}_i)}{\rho}
end{cases}
$$

(7)

The following proposition provides the required bound, namely on the delay distribution of an EDF-scheduler followed by an EBB-shaper.

**Proposition 5** Assume that the output of a $(\rho, \Lambda_0, \alpha_0)$-shaper $A_0$ enters a system $S$, for which it is known that the delay experienced by a packet $i$ is exponentially bounded, i.e.,

$$
\Pr \{ D_i(t) \geq d \} \leq \Lambda e^{-\bar{\alpha}rd} \quad \forall d \geq \delta. 
$$

The output of system $S$ enters $(\rho, \Lambda_1, \alpha_1)$-shaper $A_1$ (see Fig. 2).

The total delay, $\hat{D}_i$, experienced by a packet $i$, from the time it enters the scheduler till the time it exits $A_1$, is exponentially bounded as follows:

$$
\Pr \{ \hat{D}_i \geq d \} \leq \left( \Lambda_0 + \Lambda \right) e^{-\frac{1}{\rho \alpha_0} \bar{\alpha}rd} \quad \forall d \geq \delta.
$$

(8)

### 3.2.3. End-to-end Delay

Finally, consider a connection $k$ with $(\rho_0^k, \Lambda_0^k, \alpha_0^k)$-EBB input traffic. The connection is routed through a path $p$ in which all nodes employ the RC-EDF service discipline. Then, the stochastic bound on the end-to-end delay is given in the following theorem.

**Theorem 1** For any session $k$ in an RC-EDF network with $(\rho_0^k, \Lambda_0^k, \alpha_0^k)$-EBB traffic, the end-to-end delay $D^k(p)$ is stochastically upper bounded as follows:

$$
\Pr \{ D^k(p) \geq d \} \leq \left( \Lambda_0^k + \sum_{l \in p} \bar{\Lambda}_l \right) e^{-\frac{1}{\rho_0^k + \sum_{l \in p} \alpha_l} \bar{\alpha}rd}
$$

(9)

where

$$
\frac{1}{\bar{\alpha}_l} = \sum_{j \in N_l} \frac{1}{\alpha_j^l}, \quad \bar{\Lambda}_l = \frac{\left( \sum_{j \in N_l} \lambda_j^l \right)}{1 - e^{-\bar{\alpha}_l \left( \sum_{j \in N_l} \rho_j^l - \left( \sum_{j \in N_l} \rho_j^l \right) \delta_j^l - \delta_j^l \right)}} \cdot e^{-\bar{\alpha}_l \left( \sum_{j \in N_l} \rho_j^l \delta_j^l - \left( \sum_{j \in N_l} \rho_j^l \right) \delta_j^l - \delta_j^l \right)}.
$$

$(\rho_j^l, \lambda_j^l, \alpha_j^l)$ are the traffic parameters of session $j$ in node $l$, and $\delta_j^l$ is the EDF deadline of session $j$ in node $l$.

### 3.2.4. An Example

We now illustrate the results we have obtained for RC-EDF networks by way of an example. Specifically, following [7,9], we consider the three-node tree structured network depicted in Fig. 3. The rates of the servers are assumed to be 1. Suppose that there are four sessions in the network, which are routed as depicted in Fig. 3. The source traffic for
each session is modelled by a (mutually) independent Bernoulli process. For each session $i$ ($i = 1, 2, 3, 4$), the Bernoulli parameter is $P^i$. In [9], it is proven that the Bernoulli process is EBB and it is shown how to choose appropriate EBB parameters. We use this result to derive the EBB parameters ($\rho^i_0, \Lambda^i_0, \alpha^i_0$). The Bernoulli parameters and the corresponding EBB parameters are presented in Table 1. For reasons that shall become clear in the sequel, at the output of the shapers we enforce EBB parameters that are different than those of the input; more precisely, for each session $i$, we enforce there the EBB parameters ($\rho^i_0, \frac{1}{10}\Lambda^i_0, 10\alpha^i_0$). Choosing the same EDF deadlines for all sessions at all nodes, say, $\delta^i = 1$ for $i = 1, 2, 3, 4$, and applying Theorem 1, we obtain bounds on the tail distribution of the end-to-end delay for the four sessions. These bounds are presented by the solid lines in Fig. 4(a). In that figure, we also plot the simulation results for the actual Bernoulli processes (dotted lines). Evidently, Fig. 4(a) validates the established end-to-end delay bound for RC-EDF networks and EBB traffic. Note that, since Bernoulli processes do not satisfy Cruz’ Burstiness Constraint model [14], no worst-case deterministic bounds can be derived in this case by the methods proposed in [3].

<table>
<thead>
<tr>
<th>session</th>
<th>$P^i$</th>
<th>$\rho^i_0$</th>
<th>$\Lambda^i_0$</th>
<th>$\alpha^i_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.2</td>
<td>1</td>
<td>1.51</td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>0.25</td>
<td>1</td>
<td>1.21</td>
</tr>
<tr>
<td>3</td>
<td>0.1</td>
<td>0.2</td>
<td>1</td>
<td>1.51</td>
</tr>
<tr>
<td>4</td>
<td>0.15</td>
<td>0.25</td>
<td>1</td>
<td>1.21</td>
</tr>
</tbody>
</table>

Table 1. EBB characterization

Figure 3. An example network

Figure 4. End-to-end delay distribution in logscale
In Fig. 4(b) we compare the bounds of the RC-EDF service discipline with those presented in [7] for the RPPS-GPS service discipline and EBB traffic. The three non-solid lines represent the bounds of Theorem 1 for three sets of EBB shaping parameters, namely, $(\rho_0, \Lambda_0^i, \alpha_0^i)$, $(\rho_0, \frac{1}{10} \Lambda_0^i, 100\alpha_0^i)$, $(\rho_0, \frac{1}{100} \Lambda_0^i, 1000\alpha_0^i)$. The loosest RC-EDF bound in the figure (the “- - -” line) is achieved when the traffic shapers have the same traffic parameters as the input traffic. Tighter bounds (the “- -” and “- - -” lines) are achieved by employing EBB parameters with lower $\Lambda$’s and higher decay factors $\alpha$’s. The solid lines represent the corresponding bounds for RPPS-GPS networks. Specifically, assuming RPPS-GPS servers and applying Theorem 5 of [7], we obtain bounds on the tail distribution of the end-to-end delay for the four sessions in the three-node network of Fig. 3. One can see that, by enforcing traffic parameters at internal nodes that are different than those of the input traffic, the bounds of the RC-EDF service discipline practically outperform those of the RPPS-GPS service discipline. More specifically, better bounds are obtained for the “probable” delay values; for example, for $(\rho_0, \frac{1}{100} \Lambda_0^i, 1000\alpha_0^i)$ and for probabilities $p \geq 10^{-9}$, the delay value $d$ for which $\text{Prob} \{ D_i(t) > d \} = p$ is smaller with the EDF bound.

4. Conclusion

This study considers end-to-end QoS provisioning in RC-EDF networks with stochastic traffic profiles and stochastic guarantees. Previous studies have dealt with either the generalized processor sharing scheduling discipline (GPS), under both deterministic and stochastic settings (e.g., [7,9]), or with the rate-controlled earliest deadline first discipline (RC-EDF), under a deterministic setting (e.g., [3–5]). The present study is the first to provide end-to-end bounds for exponentially bounded burstiness (EBB) traffic and systems of RC-EDF schedulers.

After having established the bounds analytically, we tested them through some simulations and numerical calculations, carried on a three-node example. In the example, the obtained bound for the RC-EDF service discipline practically outperformed the bound presented in [7] for the RPPS-GPS service discipline under EBB traffic. Hence, an interesting direction for future research is to conduct a comprehensive performance comparison between the RC-EDF scheduling discipline and the GPS discipline under stochastic settings. In particular, in [3] it was shown that any end-to-end delay bounds that can be guaranteed by the GPS discipline, can also be achieved by an RC-EDF discipline, by using a simple algorithm to determine how to reshape the traffic; whether a similar result can be obtained also under the stochastic setting is a challenging open question.

This study has provided the required foundations for several related network control problems. In [15], we present some initial results on the application of the proposed bounds in the context of call admission control and routing. Due to the complexity of the end-to-end bound under the stochastic setting, these schemes are quite complex. The establishment of more efficient call admission control and routing schemes is an important direction for future research.

In order to achieve tighter bounds, a more general framework, of Stochastically Bounded Burstiness (SBB), was developed in [16]. The SBB calculus is known to be a powerful tool for obtaining much tighter bounds than the EBB calculus for multiple time-scale processes. Accordingly, establishing bounds on the end-to-end tail distribution for RC-
EDF networks and SBB traffic, as was done in [8] for RPPS-GPS networks, is yet another important direction for future research.

REFERENCES