Statistical Analysis of Generalized Processor Sharing Scheduling Discipline

Z. ZHANG, D. TOWSLEY and J. KUROSE

CMPSCI Technical Report 95-10

February, 1995

Statistical Analysis of Generalized Processor Sharing Scheduling Discipline*

Zhi-Li Zhang, Don Towsley and Jim Kurose
Computer Science Department
University of Massachusetts
Amherst, MA 01003, USA

Abstract

In this paper, we develop bounds on the individual session backlog and delay distribution under the Generalized Processor Sharing (GPS) scheduling discipline . This work is motivated by, and is an extension of, Parekh and Gallager's deterministic study of the GPS scheduling discipline with leaky-bucket token controlled sessions [PG93a,b, Parekh92]. Using the exponentially bounded burstiness (E.B.B.) process model introduced in [YaSi93] as a source traffic characterization, we establish results that extend the deterministic study of GPS: for a single GPS server in isolation, we present statistical bounds on the distributions of backlog and delay for each session. In the network setting, we show that networks belonging to a broad class of GPS assignments, the so-called Consistent Relative Session Treatment (CRST) GPS assignments, are stable in a stochastic sense. In particular, we establish simple bounds on the distribution of backlog and delay for each session in a Rate Proportional Processor Sharing (RPPS) GPS network with arbitrary topology.

1 Introduction

The provision of *Quality-of-Service* (QOS) guarantees has become an increasingly important and challenging topic in the design of high-speed networks. One important issue in the provision of QOS guarantees is the study of the scheduling disciplines to be employed at network switches. Ideally these scheduling disciplines should, on the one hand, provide isolation between sessions so that the misbehavior of one session will not affect other sessions, and on the other hand, exploit statistical multiplexing gain [CSZ92]. They should also, preferably, be amenable

*This work supported in part by NSF grant NCR-9116183.

to analysis so that theoretical bounds can be derived. The provable bounds can be either deterministic (i.e., *hard guarantees*) or statistical (i.e., *soft guarantees*) depending on the nature of applications.

Perhaps the most widely studied non-FCFS scheduling is the *Generalized Processor Sharing* (GPS) scheduling discipline (also known as *Weighted Fair Queueing (WFQ)* in the literature [e.g. DKS89, CSZ92]). In [PG93a,b, Parekh92], Parekh and Gallager presented a thorough examination of GPS with leaky-bucket token controlled incoming traffic (hence traffic from each session conforms to Cruz's *Linear Bounded Arrival Process* (LBAP) traffic model [Cruz91a, Cruz91b]). Given that each session is leaky-bucket controlled and that the total rate of the arrivals is smaller than the service rate, it was shown, in the case of a single GPS server in isolation, that the backlog and delay of each session are bounded from above; and in the case of a network of GPS servers, that under a broad class of GPS assignments known as *Consistent Relative Session Treatment* (CRST) GPS assignments, the network is stable. These bounds are actually attainable in the worst-case scenario. Of particular interest is a subclass of CRST GPS networks, the so-called *Rate Proportional Processor Sharing* (RPPS) GPS networks. For RPPS GPS networks, bounds on backlog and delay for each session do not depend on the length of the route the session traverses but only on the bottleneck node on the route. In this case, simple closed form expressions can be derived for each session. These nice features of GPS make it an attractive choice when QOS guarantees must be made.

The work of Parekh and Gallager makes it possible to provide worst-case deterministic bounds (i.e., hard guarantees) for networks employing the GPS scheduling. As many applications, especially multimedia applications, can tolerate a certain amount of loss due either to late arrival or buffer overflow, it is more desirable if statistical bounds (i.e., soft guarantees) can also be provided. Moreover, simulation results [YKTH93] show that deterministic upper bounds are usually very conservative. Hence, if these bounds are used as admission control criteria, low utilization of network bandwidth will result.

This motivates us to investigate the behavior of the GPS scheduling discipline in a stochastic setting. In this paper, we model the source session traffic as an E.B.B. process, a notion introduced in [YaSi93]. Using techniques similar to those of Parekh and Gallager's, we examine the sample path behavior of the sessions served by a single GPS server in isolation and derive several useful relations on a sample path. Based on these relations, we derive statistical bounds on the backlog, delay and departure processes for each session. We then apply our result for an isolated GPS server to networks of GPS servers. We introduce the notion of a feasible partition, which captures the essence of the notion of feasible orderings introduced by Parekh and Gallager. With the help of this notion, we are able to obtain tighter bounds for each session. Parallel to the deterministic case studied by Parekh and Gallager, we show that a CRST GPS network with arbitrary topology is stable in a statistical sense. In particular, for RPPS GPS networks, the upper bounds on the backlog and delay distributions for each session are shown not to depend on the length of the route the session traverses but only on the bottleneck node on the route and have simple closed form expressions.

In [YaSi94], Yaron and Sidi also studied GPS scheduling discipline using E.B.B. processes as source traffic models for sessions. Their focus was mostly on establishing the input-output relation for a single GPS server and the stability of CRST GPS network with E.B.B. arrival processes. Their analysis was based on direct investigation of the output processes of the sessions, and the E.B.B. characterizations for the output processes were expressed in a recursive fashion. In contrast, we take a decomposition approach in the analysis of the sample path behavior of a GPS server which allows us to decouple the sessions. As a consequence, we are able to obtain bounds on the backlog and delay distribution for each session and derive simpler E.B.B. characterizations for the output processes. Moreover, in the network setting, we define CRST GPS assignments in terms of the feasible partition at each node. This yields a larger set of CRST GPS assignments than Parekh and Gallager's original definition and hence we are able to show stability results for a broader class of GPS networks than [YaSi94]. In particular, we obtain simple but important results for RPPS GPS networks which parallel those for the deterministic case.

A number of papers have also been concerned with developing performance bounds. For example, several studies [Cruz91a, Cruz91b, Chang94] have obtained performance bounds in the case that sources are characterized by deterministic traffic models (Cruz's LBAP model and Chang's *Envelope Process* model). Stochastic models have been considered for policies other than GPS in [Kurose91, YaSi93, Chang94]. The second of these, [YaSi93], introduced the *Exponentially Bounded Burstiness* (E.B.) process model which we use in our work. These studies derive stochastic bounds on some interested performance metrics under arbitrary work-conserving service disciplines. Other work includes studies of the *Stop-and-go* scheduling policy, [Go90, Go91], *Hierarchical Round Robin* [KKK90], and some rate-control or jitter-control based algorithms (e.g., [ZF93, VZF91]).

The rest of the paper is organized as follows. In Section 2, we introduce the necessary notation. In Section 3, we examine the sample path behavior of sessions for a single GPS server. In Section 4, we present bounds on backlog and delay distribution and establish input-output relations. In Section 5, we introduce the concept of a feasible partition and refine the bounds obtained in the previous section. In Section 6, we analyze a network of GPS servers, show that CRST GPS networks are stable and study RPPS GPS networks in particular. We also present a numerical example for a simple RPPS GPS network. Finally in Section 7, we conclude the paper and discuss possible future work.

2 Preliminaries

Generalized Processor Sharing (GPS) is a work-conserving scheduling discipline defined under the assumption that sources are described by fluid models, (i.e., packets are infinitely divisible[1]) and is work-conserving. Extensions to account for the packet nature of communication are not difficult [PG93a, YaSi93b], but will not be considered here. Consider a GPS server with rate r serving N sessions. Following Parekh and Gallager's definition, each session i is assigned a fixed real-valued positive parameter ϕ_i, where $\{\phi_i\}_{1\leq i\leq N}$ is called a GPS assignment. The N sessions share the server in the following way:

For $1 \leq i \leq N$, let $S_i(\tau,t)$ be the amount of session i traffic served in a time interval $[\tau, t]$. Then

$$\frac{S_i(\tau,t)}{S_j(\tau,t)} \geq \frac{\phi_i}{\phi_j}, \quad j = 1, 2, \ldots, N \tag{1}$$

for any session i that is *backlogged* throughout the interval $[\tau, t]$. A session is backlogged throughout an interval if there is always traffic queued for that session at all times in the interval [PG93a, Parekh92]. From (1), one can derive that when session i is backlogged, it is guaranteed a backlog clearing rate (or equivalently, a guaranteed amount of service per unit time) of $g_i = \frac{\phi_i}{\sum_{j=1}^{N}\phi_j}r$. For simplicity, we will assume $r = 1$ throughout the paper, except in Section 6 where we consider a network of GPS queues.

We use the exponentially bounded burstiness (E.B.B.) process introduced in [YaSi93] to model traffic generated by a source. We say a session i arrival process, A_i, is a $(\rho_i, \Lambda_i, \alpha_i)$-E.B.B. process, if for any τ and t such that $\tau \leq t$ and for any $x \geq 0$,

$$Pr\{A_i(\tau,t) \geq \rho_i(t-\tau) + x\} \leq \Lambda_i e^{-\alpha_i x} \tag{2}$$

where ρ_i is called the long term *upper rate* of the arrival process, Λ_i the prefactor and α_i the decay rate of the exponential decay function. As a necessary stability condition, we require that $\sum_{i=1}^{N} \rho_i < 1$.

A corresponding concept of an E.B.B. process is the concept of an *exponentially bounded* (E.B.) process. We say a stochastic process $X(t)$ is an (α, Λ)-E.B. process if for any t and any $x \geq 0$,

$$Pr\{X(t) \geq t\} \leq \Lambda e^{-\alpha x}. \tag{3}$$

As in the case of E.B.B. process, α is the decay rate and Λ the prefactor of the E.B. process. Relationships between E.B.B. and E.B. processes are found in [YaSi93].

[1]Namely, we are regarding traffic from each session as a bit flow. There is no notion of "packet" in the fluid model [PG93a, Parekh92].

In [PG93a,b, Parekh92], Parekh and Gallager showed[2] that, given $\sum_{i=1}^{N} \rho_i < 1$, there exists an ordering among the sessions such that, after relabeling of the sessions,

$$\rho_i < \frac{\phi_i}{\sum_{j=i}^{N} \phi_j}(1 - \sum_{j=1}^{i-1} \rho_j), \;\; 1 \leq i \leq N. \tag{4}$$

Such an ordering is called a *feasible ordering* (with respect to $\{\rho_i\}_{1\leq i\leq N}$ and $\{\phi_i\}_{1\leq i\leq N}$). In general, there are many feasible orderings associated with a given set of $\{\rho_i\}_{1\leq i\leq N}$ and $\{\phi_i\}_{1\leq i\leq N}$.

We introduce several basic notations for the statistical analysis of a single GPS server in isolation. Notation for the network analysis is deferred until Section 6.

For $1 \leq i \leq N$, A_i denotes the arrival process for session i and S_i the corresponding departure process, where for any $\tau \leq t$, $A_i(\tau, t)$ is the amount of traffic from session i during the time interval $[\tau, t]$, and $S_i(\tau, t)$ the amount of service session i received during the same period. The session i backlog at time t, denoted as $Q_i(t)$, is given by $Q_i(t) = \sup_{\tau\leq t}\{A_i(\tau, t) - S_i(\tau, t)\}$. The delay experienced by a session i traffic arriving at time t is denoted as $D_i(t)$. If the traffic from a session is serviced by the FCFS scheduling discipline (which is assumed to be so in this paper), then $D_i(t)$ is the time that the session i backlog at time t, $Q_i(t)$, gets cleared. Furthermore, we say a time interval, B_i, is a session i busy period if it is the maximal interval such that session i is backlogged throughout the interval. For any time $t \in B_i$, we say session i is busy at time t. A time interval B is a system busy period if it is the maximal interval such that, at any time in the interval, at least one session is busy.

3 Sample Path Behavior of a Single GPS Server

We first study the sample path behavior of the sessions served by a single GPS server. For convenience of exposition, we assume that each session has its own infinite capacity queue. Recall that traffic for each session, A_i, is characterized by an E.B.B. process with long term upper rate ρ_i such that $\sum_{i=1}^{N} \rho_i < r = 1$. By abuse of notation, let A_i also denote a sample path (or a realization) of a random arrival process A_i, so $A_i(\tau, t)$ is the amount of traffic from session i during the time interval $[\tau, t]$ on this sample path. In general, we assume that A_i is a right continuous function with left limit, therefore $A_i(t, t) = 0$ for any t. Similarly, we will use S_i, Q_i and D_i to denote the corresponding sample paths of the corresponding random processes S_i, Q_i and D_i.

The analysis of a stochastic system such as a GPS system where several queues share a server is generally very difficult, as the amount of service a session receives at any moment depends

[2] Actually the strict inequality here and in (4) can be replaced by "$\leq$" in the deterministic case they considered.

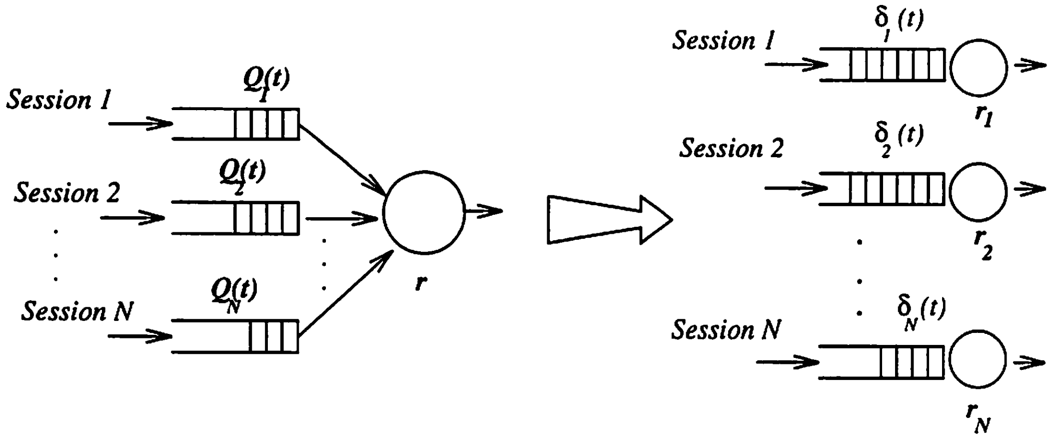


Figure 1: **Decomposition of a GPS system.**

not only on its arrival process and the content of its own queue at the moment but also on the arrival processes to the other sessions and the contents of their queue. To circumvent this difficulty, we would like to decompose the GPS system in such a way that the sessions are decoupled. Imagine that we divide a GPS server of rate $r = 1$ into a set of N (fictitious) servers with rates $r_1, \ldots, r_N$, so that, instead of having a GPS system with N sessions sharing a server, we have a decomposed system consisting of N separate queues, each of which has a dedicated server (see Figure 1). For each session i in the *decomposed system*, let $\delta_i(t)$ denote its queue length (or backlog) at time t. We hope to bound the (actual) session i backlog $Q_i(t)$ of the *real* GPS system in terms of the (fictitious) session i backlog $\delta_i(t)$'s of the *imaginary* decomposed system. Now the crucial question is how to decompose the GPS system so that our goal can be achieved, i.e., how shall we choose the r_i's? Obviously we must have $\sum_{i=1}^{N} r_i \leq 1$ and $\rho_i < r_i$ for $\delta_i(t)$ to be well-defined. We need to impose an additional relation on the r_i's in order to reflect the GPS scheduling discipline, corresponding to relation (4). We choose a set of r_i's such that, after some relabeling of the sessions,

$$r_i \leq \frac{\phi_i}{\sum_{j=i}^{N} \phi_j}(1 - \sum_{j=1}^{i-1} r_j), \ 1 \leq i \leq N. \tag{5}$$

As long as $\sum_{i=1}^{N} r_i \leq 1$, such a feasible ordering of the sessions always exists. Note that as $\rho_i < r_i$, we have $\epsilon_i = r_i - \rho_i > 0$. *Without loss of generality, we assume that* $1, 2, \ldots, N$ *is a feasible ordering of the N sessions with respect to* $\{r_i\}_{1 \leq i \leq N}$ *throughout the rest of this section and the next section.*

From the study of G/G/1 queue, it is well known that $\delta_i(t)$ can be expressed in the form,

$$\delta_i(t) = \sup_{s \leq t}\{A_i(s,t) - r_i(t-s)\}. \tag{6}$$

Clearly, $\delta_i(t) \geq 0$ for any t. For $\tau \leq t$, applying (6) to $\delta_i(\tau)$ and $\delta_i(t)$, we can easily derive the following useful inequality:

$$A_i(\tau, t) \leq r_i(t-\tau) + \delta_i(t) - \delta_i(\tau) \leq r_i(t-\tau) + \delta_i(t). \tag{7}$$

For any t, let $\eta_i(t) = Q_i(t) - \delta_i(t)$. As $S_i(\tau, t) = A_i(\tau, t) + Q_i(\tau) - Q_i(t)$, from (7), we have

$$S_i(\tau, t) \leq r_i(t-\tau) + \eta_i(\tau) - \eta_i(t) \leq r_i(t-\tau) + \eta_i(\tau) + \delta_i(t). \tag{8}$$

where the last inequality holds as $Q_i(t) \geq 0$.

Now we state an important fact, the proof of which is relegated to the appendix.

Lemma 1 *For each session i, $1 \leq i \leq N$, at any time t,*

$$\sum_{j=1}^{i} Q_j(t) \leq \sum_{j=1}^{i} \delta_j(t). \tag{9}$$

This lemma says that on a sample path basis, the sum of the (actual) backlogs of the first i sessions of a feasible ordering in the real GPS system is bounded from above by the sum of the (fictitious) backlogs of the corresponding sessions in the imaginary decomposed system. We can also bound the (actual) backlog, $Q_i(t)$, and delay, $D_i(t)$, of each individual session of the real GPS system in terms of the $\delta_i(t)$'s, but first we need to establish a lower bound on the session i service function S_i when it is in a busy period. This is stated in Lemma 2 below. The bounds on individual sessions are then given in Lemma 3.

Lemma 2 *For any t, let τ be the beginning of a session i busy period that contains t. Then*

$$S_i(\tau, t) \geq r_i(t-\tau) - \frac{\phi_i}{\sum_{j=i}^{N} \phi_j} \sum_{j=1}^{i-1} \delta_j(t). \tag{10}$$

Proof: As session i is busy in $[\tau, t]$, from the definition of GPS, we can easily derive that

$$S_i(\tau, t) \geq \frac{\phi_i}{\sum_{j=i}^{N} \phi_j}\left(t - \tau - \sum_{j=1}^{i-1} S_j(\tau, t)\right). \tag{11}$$

By (8) and Lemma 1, we have

$$\begin{aligned}
\sum_{j=1}^{i-1} S_j(\tau, t) &\leq \sum_{j=1}^{i-1} (\eta_j(\tau) + r_j(t-\tau) + \delta_j(t)) \\
&\leq \sum_{j=1}^{i-1} (r_j(t-\tau) + \delta_j(t)).
\end{aligned}$$

Substituting the above inequality into (11) and noticing that $r_i \leq \frac{\phi_i}{\sum_{j=i}^N \phi_j}(1 - \sum_{j=1}^{i-1} r_j)$, we obtain that

$$S_i(\tau, t) \geq r_i(t-\tau) - \frac{\phi_i}{\sum_{j=i}^N \phi_j} \sum_{j=1}^{i-1} \delta_j(t). \tag{12}$$

■

Lemma 3 *For any t,*

$$Q_i(t) \leq \delta_i(t) + \frac{\phi_i}{\sum_{j=i}^N \phi_j} \sum_{j=1}^{i-1} \delta_j(t) \tag{13}$$

and

$$D_i(t) \leq \frac{1}{g_i}\left(\delta_i(t) + \frac{\phi_i}{\sum_{j=i}^N \phi_j} \sum_{j=1}^{i-1} \delta_j(t)\right) \tag{14}$$

where $g_i = \frac{\phi_i}{\sum_{j=1}^N \phi_j}$ *is the guaranteed backlog clearing rate for session* i.

Proof: Let τ be the beginning of a session i busy period that contains t. As $Q_i(t) = A_i(\tau, t) - S_i(\tau, t)$, from (7) and Lemma 2, we obtain (13). (14) follows immediately as session i is guaranteed a backlog clearing rate of g_i. ■

For $\tau \leq t$, as $S_i(\tau, t) \leq A_i(\tau, t) + Q_i(\tau)$, applying Lemma 3 to $Q_i(\tau)$, we have

Lemma 4 *Over any time interval* $[\tau, t]$,

$$S_i(\tau, t) \leq A_i(\tau, t) + \delta_i(\tau) + \frac{\phi_i}{\sum_{j=i}^N \phi_j} \sum_{j=1}^{i-1} \delta_j(\tau). \tag{15}$$

Before we leave this section, we point out the connection between our notations and those of Parekh and Gallager. In [PG93a, Parekh92], Parekh and Gallager defined

$$l_i(t) = \inf_{s \leq t}\{r_i(t-s) + \sigma_i - A_i(s,t)\} \tag{16}$$

and

$$\sigma_i(t) = Q_i(t) + l_i(t). \tag{17}$$

In terms of leaky buckets, $l_i(t)$ is the number of tokens left in the bucket at time t, and denotes the maximal allowable burst at the next moment, t^+. $\sigma_i(t)$ is the potential discrepancy between the arrival function and the service function at this next moment. We observe that $l_i(t) = \sigma_i - \delta_i(t)$ and $\sigma_i(t) = \eta_i(t) + \sigma_i$. Therefore, in this context, $\delta_i(t)$ can be interpreted as the excess of session i traffic over the tokens generated up to time t (excluding the full token bucket at the very beginning and the tokens lost due to bucket overflow), and $\eta_i(t)$ the discrepancy between the session i arrival function (with this excess, $\delta_i(t)$, excluded) and the service function at time t.

Our notations, however, are more general. They can be better used to describe another token control scheme: for each session i, tokens are generated as a continuous bit flow at a constant rate of r_i. They are consumed immediately by the incoming session i traffic arriving at the same moment. If not all tokens generated are consumed, they are discarded (i.e., the token bucket size is zero[3]). The arriving traffic in excess of the tokens generated at the moment is marked and let into the network (instead of being lost or buffered as in the typical leaky bucket control scheme). Under this scheme, $\delta_i(t)$ is the amount of marked session i traffic at time t, and $\eta_i(t)$ is the backlog of unmarked session i traffic at time t, whereas $Q_i(t) = \eta_i(t) + \delta_i(t)$ is the total backlog of session i traffic (both unmarked and marked) at time t. Under this scheme, ρ_i is the long term upper rate of session i which reflects the true characteristic of session i whereas r_i is the token generating rate for session i. The methods in the next section can be used to study the distribution of the amount of marked traffic of each session at any time and the impact the marked traffic has on the unmarked traffic in terms of backlog and delay distribution.

4 Statistical Analysis of a Single GPS Server

In the previous section we have established some useful relations on backlog and delay on a sample path basis. It is interesting to note that they are independent of whatever stochastic model we use to characterize each session source traffic. In this section we will use an E.B.B. process characterization of a session to establish upper bounds on the tail distributions of backlog and delay for each session. For simplicity of exposition, at first we will assume that the arrival processes are all independent. The results for the case where the arrival processes are dependent will be mentioned at the end of this session.

From Lemma 3, we see that in order to bound the tail distributions of $Q_i(t)$ and $D_i(t)$ we only need to bound $\delta_i(t)$ for any t. Under the E.B.B. process model, such a bound can be easily obtained. For example, the proof of Theorem 1 in [YaSi93] provides an upper bound on $\delta_i(t)$ which is paraphrased below.

Lemma 5 *For any $x > 0$,*

$$Pr\{\delta_i(t) \geq x\} \leq \frac{\Lambda_i e^{\alpha_i \rho_i \xi}}{1 - e^{-\alpha_i \epsilon_i \xi}} e^{-\alpha_i x} \tag{18}$$

where ξ satisfies $0 < \xi \leq \frac{\ln(\Lambda_i + 1)}{\alpha_i \epsilon_i}$.

[3]We can modify our notations to describe the case where the token buckets are not of zero size but of size σ_i for each session i: $\delta_i^{\sigma_i}(t) = \sup_{s \leq t}\{A_i(s,t) - r_i(t-s) - \sigma_i\}$.

Remark: In [YaSi93a], the condition on ξ is such that for some fixed x_0, $0 < \xi < \frac{x_0}{r_i}$ (actually $0 < \xi \le \frac{x_0}{r_i}$ is sufficient) and $\frac{\Lambda_i e^{\alpha_i \rho_i \xi}}{1-e^{-\alpha_i \epsilon_i \xi}} \ge e^{\alpha_i x_0}$. Note that the condition we impose on ξ ensures that $\frac{\Lambda_i e^{\alpha_i \rho_i \xi}}{1-e^{-\alpha_i \epsilon_i \xi}} \ge e^{-\alpha_i r_i \xi}$. Hence we can always pick an x_0 such that $\frac{\Lambda_i e^{\alpha_i \rho_i \xi}}{1-e^{-\alpha_i \epsilon_i \xi}} \ge e^{\alpha_i x_0} \ge e^{\alpha_i r_i \xi}$.

For reasons that will become clearer, we are actually interested in bounding the moment generating function of $\delta_i(t)$, i.e., $Ee^{\theta \delta_i(t)}$ for some θ. We first observe that for $\tau \le t$ and for $0 < \theta < \alpha_i$,

$$Ee^{\theta A_i(\tau,t)} \le e^{\theta(\rho_i(t-\tau)+\hat{\sigma}_i(\theta))} \tag{19}$$

where $\tau \le t$ and $\hat{\sigma}_i(\theta) = \frac{1}{\theta}\ln(1 + \frac{\theta \Lambda_i}{\alpha_i - \theta})$. To see why this is true, for $0 < \theta < \alpha_i$ and $\tau \le t$, let $y = e^{\theta \rho_i (t-\tau)}$, then

$$\begin{aligned} Ee^{\theta A_i(\tau,t)} &= \int_0^\infty Pr\{e^{\theta A_i(\tau,t)} \ge x\}dx \\ &= \int_0^y Pr\{e^{\theta A_i(\tau,t)} \ge x\}dx + \int_y^\infty Pr\{e^{\theta A_i(\tau,t)} > x\}dx. \end{aligned}$$

As $Pr\{e^{\theta A_i(\tau,t)} \ge x\} \le 1$, the first integral is at most $y = e^{\theta \rho_i (t-\tau)}$ whereas by a change of variable $x = e^{\theta[\rho_i(t-\tau)+\sigma_i]}$, the second integral becomes

$$\begin{aligned} &\theta e^{\theta\rho_i(t-\tau)} \int_0^\infty Pr\{A_i(\tau,t) \ge \rho_i(t-\tau) + \sigma\} e^{\theta\sigma} d\sigma \\ &\le \theta e^{\theta\rho_i(t-\tau)} \int_0^\infty \Lambda_i e^{-(\alpha_i - \theta)\sigma} d\sigma \\ &= \frac{\Lambda_i \theta}{\alpha_i - \theta} e^{\theta \rho_i (t-\tau)}. \end{aligned}$$

Therefore, (19) follows.

Now we show that

Lemma 6 *For* $0 < \theta < \alpha_i$,

$$Ee^{\theta\delta_i(t)} \le \frac{e^{\theta(\hat{\sigma}_i(\theta) + \rho_i \xi)}}{1 - e^{-\theta \epsilon_i \xi}} \tag{20}$$

where $\xi > 0$ *is an arbitrary discretization parameter.*

Proof: Recall that

$$\delta_i(t) = \sup_{\tau \le t}\{A_i(\tau,t) - r_i(t-\tau)\}. \tag{21}$$

Let s, $s \le t$, be such that $\delta_i(t) = A_i(s,t) - r_i(t-s)$. For any $\xi > 0$, there is a $k \ge 1$, such that $t - k\xi < s \le t - k\xi + \xi$. Then,

$$\delta_i(t) = A_i(s,t) - r_i(t-s) \le A_i(t-k\xi, t) - k\xi r_i + r_i \xi. \tag{22}$$

Therefore, for $0 < \theta < \alpha_i$,

$$\begin{aligned} e^{\theta\delta_i(t)} &\leq \max_{k\geq 1}\{\exp[\theta(A_i(t-k\xi,t)-k\xi r_i+\xi r_i)]\} \\ &\leq e^{\theta r_i \xi}\sum_{k=1}^{\infty}\exp[\theta(A_i(t-k\xi,t)-k\xi r_i)]. \end{aligned}$$

Taking the expectation on both sides yields

$$\begin{aligned} Ee^{\theta\delta_i(t)} &\leq e^{\theta r_i\xi}\sum_{k=1}^{\infty}E\exp[\theta(A_i(t-k\xi,t)-k\xi r_i)] \\ &\leq e^{\theta r_i\xi}\sum_{k=1}^{\infty}e^{\theta(k\xi\rho_i+\hat{\sigma}_i(\theta))}e^{-\theta k\xi r_i} \\ &= e^{\theta(r_i\xi+\hat{\sigma}_i(\theta))}\sum_{k=1}^{\infty}e^{-k\theta\epsilon_i\xi} \\ &= e^{\theta(r_i\xi+\hat{\sigma}_i(\theta))}\frac{e^{-\theta\epsilon_i\xi}}{1-e^{-\theta\epsilon_i\xi}} \\ &= \frac{e^{\theta(\hat{\sigma}_i(\theta)+\rho_i\xi)}}{1-e^{-\theta\epsilon_i\xi}} \end{aligned}$$

where the second inequality follows from (19). ■

Remarks:

(1) One can show that $f(\xi) = \frac{e^{\theta\rho_i\xi}}{1-e^{-\theta\epsilon_i\xi}}$ is a convex function that attains its minimum at $\xi_0 = \frac{\ln r_i - \ln\rho_i}{\epsilon_i\theta}$ and $f(\xi_0) = \frac{r_i^2}{\epsilon_i\rho_i}e^{\frac{\rho_i}{\epsilon_i}}$. Hence the righthand side of (20) attains its minimum value $e^{\theta\hat{\sigma}_i(\theta)}f(\xi_0) = (1+\frac{\Lambda_i}{\alpha_i-\theta})\frac{r_i^2}{\epsilon_i\rho_i}e^{\frac{\rho_i}{\epsilon_i}}$. Similarly, the prefactor $\frac{\Lambda_i e^{\alpha_i\rho_i\xi}}{1-e^{-\alpha_i\epsilon_i\xi}}$ of the right hand side of (18) attains its mimimun value at $\xi = \min\{\frac{\ln(\Lambda_i+1)}{\alpha_i\epsilon_i}, \frac{\ln r_i - \ln\rho_i}{\alpha_i\epsilon_i}\}$. In other words, the minimum value is $(\Lambda_i+1)^2 e^{\frac{\rho_i}{\epsilon_i}}$ if $\Lambda_i \leq \frac{\epsilon_i}{\rho_i}$, and $\frac{\Lambda_i r_i^2}{\epsilon_i\rho_i}e^{\frac{\rho_i}{\epsilon_i}}$ otherwise. In the rest of the paper, however, when using (20), we will choose $\xi = 1$ for simplicity of notation.

(2) In the discrete time case, similar bounds on $\delta_i(t)$ as (18) and (20) can also be obtained (see [YaSi93] and [Chang94]).

(3) In general, if we know the more specific structure of the arrival process, better bounds for $\delta_i(t)$ are usually available. As $\delta_i(t)$ is the backlog of session i when it is serviced by a server with rate r_i, known results for the single server queue can be applied. For example, in the case that arrivals are described by Markov modulated processes, results of the studies reported in [LNT94, BD94] can be applied. We note that by using better bounds on $\delta_i(t)$ in the proofs of the following theorems, better bounds than those stated in the theorems can be obtained. We will not elaborate on this here. In section 6.3, we will present an example to illustrate this point.

For each i, let $\psi_i = \frac{\phi_i}{\sum_{j=i}^{N}\phi_j}$. The following is the main theorem of this section.

Theorem 7 *Suppose that* $\{A_i\}_{1\le i\le N}$ *are* N *independent* $(\rho_i, \Lambda_i, \alpha_i)$*-E.B.B processes sharing a GPS server with assignment* $\{\phi_i\}_{1\le i\le N}$. *Assume* $\sum_{i=1}^N \rho_i < 1$ *and that* $1, 2, \ldots, N$ *is a feasible ordering with respect to* $\{\phi_i\}_{1\le i\le N}$ *and* $\{r_i = \rho_i + \epsilon_i\}_{1\le i\le N}$, *where* $\epsilon_i > 0$ *and* $\sum_{i=1}^N \epsilon_i \le 1 - \sum_{i=1}^N \rho_i$. *Then at any time t, for any* $q > 0$,

$$Pr\{Q_i(t) \ge q\} \le \Lambda_i^{out} e^{-\theta q}, \tag{23}$$

and for any $d > 0$,

$$Pr\{D_i(t) \ge d\} \le \Lambda_i^{out} e^{-\theta g_i d}, \tag{24}$$

moreover, for any fixed time interval $[\tau, t]$ *and any* $x > 0$,

$$Pr\{S_i(\tau, t) \ge \rho_i(t-\tau) + x\} \le \Lambda_i^{out} e^{-\theta x} \tag{25}$$

where

$$\Lambda_i^{out} = \frac{\exp(\theta[\hat{\sigma}_i(\theta) + \rho_i + \psi_i \sum_{j=1}^{i-1}(\hat{\sigma}_j(\psi_i\theta) + \rho_j)])}{(1 - e^{-\theta\epsilon_i})\prod_{j=1}^{i-1}(1 - e^{-\psi_i\theta\epsilon_j})} \tag{26}$$

and $0 < \theta < \min_{1\le j\le i} \alpha_j$.

Proof: From (13), we have

$$Pr\{Q_i(t) \ge q\} \le Pr\{\delta_i(t) + \psi_i \sum_{j=1}^{i-1} \delta_j(t) \ge q\}. \tag{27}$$

As $\delta_i(t)$ only depends on A_i up to time t and $A_1, A_2, \ldots, A_N$ are independent, $\delta_1(t)$, $\delta_2(t)$, $\ldots$, $\delta_N(t)$ are also independent for any fixed t. Using Chernoff's Bound, for any θ such that $0 < \theta < \min_{1\le j\le i} \alpha_j$, we have

$$Pr\{\delta_i(t) + \psi_i \sum_{j=1}^{i-1} \delta_j(t) \ge q\} \le E[e^{\theta\delta_i(t)}] \prod_{j=1}^{i-1} E[e^{\psi_i\theta\delta_j(t)}] e^{-\theta q}. \tag{28}$$

For $1 \le k \le i$, choosing an appropriate θ', applying (20) and substituting into (28) yields (23).

Equation (24) then follows as the session i bits are assumed to be serviced in a first-come-first-serve manner, and the session i is guaranteed a minimum service rate of g_i whenever its queue is not empty.

To prove (25), for $\tau \le t$, we define

$$\delta_i^\tau(t) = \delta_i(\tau) + A_i(\tau, t) - \rho_i(t - \tau). \tag{29}$$

Then from Lemma 4, we have

$$S_i(\tau, t) \le \rho_i(t-\tau) + \delta_i^\tau(t) + \frac{\phi_i}{\sum_{j=i}^N \phi_j} \sum_{j=1}^{i-1} \delta_j(\tau). \tag{30}$$

By an argument similar to the one used in the proof of Lemma 6, we can show that

$$Ee^{\theta\delta_i^\tau(t)} \le \frac{e^{\theta(\hat{\sigma}_i(\theta)+\rho_i\xi)}}{1-e^{-\theta\epsilon_i\xi}} \tag{31}$$

where $\xi > 0$. As before, we take $\xi = 1$ for simplicity.

Clearly, given that $A_j, 1 \le j \le i$, are independent, so are $\delta_j(\tau), 1 \le j \le i-1$, and $\delta_i^\tau(t)$. Now the same argument as used above for proving (23) yields (25). ∎

Theorem 7 states that the backlog and the delay for each session decay exponentially with rate θ as q and d increase. Stochastic processes with such property are called *exponentially bounded* (E.B.) processes in [YaSi93a]. Furthermore, (25) says that for $0 < \theta < \min_{1\le j\le i}\alpha_j$, the departure process for each session i is an E.B.B. process. This yields a simple input-output relation under the assumption that the arrival processes are independent. To obtain an input-output relation that will be useful in the analysis of networks of GPS servers, we need to do away with this independence assumption. Fortunately, the case where the arrival processes are dependent can be resolved with the help of Hölder's inequality. For $1 \le i \le N$, let $\{p_j\}_{1\le j\le i}$ be such that $p_j > 1$ and $\sum_{j=1}^{i}\frac{1}{p_j} = 1$ (e.g. $p_j = i$). Then

$$E[\exp(\theta(\delta_i(t) + \psi_i\sum_{j=1}^{i}\delta_j(t)))] \le E[\exp(p_i\theta\delta_i(t))]^{1/p_i}\prod_{j=1}^{i-1}(E[\exp(p_j\psi_i\theta\delta_j(t))])^{1/p_j}. \tag{32}$$

Using this fact, it is easy to prove the following

Theorem 8 *Suppose that $\{A_i\}_{1\le i\le N}$ are N $(\rho_i, \Lambda_i, \alpha_i)$-E.B.B processes sharing a GPS server with assignment $\{\phi_i\}_{1\le i\le N}$. Assume $\sum_{i=1}^{N}\rho_i < 1$ and that $1, 2, \ldots, N$ is a feasible ordering with respect to $\{\phi_i\}_{1\le i\le N}$ and $\{r_i = \rho_i + \epsilon_i\}_{1\le i\le N}$, where $\epsilon_i > 0$ and $\sum_{i=1}^{N}\epsilon_i \le 1 - \sum_{i=1}^{N}\rho_i$. For $1 \le i \le N$, let $\{p_j\}_{1\le j\le i}$ be such that $p_j > 1$ and $\sum_{j=1}^{i}\frac{1}{p_j} = 1$. Then at any time t, for any $q > 0$,*

$$Pr\{Q_i(t) \ge q\} \le \Lambda_i^{out'}e^{-\theta q}, \tag{33}$$

and for any $d > 0$,

$$Pr\{D_i(t) \ge d\} \le \Lambda_i^{out'}e^{-\theta g_i d}, \tag{34}$$

moreover, for any fixed time interval $[\tau, t]$ and any $x > 0$,

$$Pr\{S_i(\tau, t) \ge \rho_i(t-\tau) + x\} \le \Lambda_i^{out'}e^{-\theta x} \tag{35}$$

where

$$\Lambda_i^{out'} = \frac{\exp(\theta[\hat{\sigma}_i(p_i\theta) + \rho_i + \psi_i\sum_{j=1}^{i-1}(\hat{\sigma}_j(p_j\psi_i\theta) + \rho_j)])}{(1-e^{-p_i\theta\epsilon_i})\prod_{j=1}^{i-1}(1-e^{-p_j\psi_i\theta\epsilon_j})} \tag{36}$$

and $0 < \theta < \min_{1\le j\le i}\alpha_j/p_j$.

We see that without the independence assumption of the arrival processes, the prefactor $\Lambda_i^{out'}$ will be generally larger. Moreover, the range of the decay rate θ is small. Note that $\min_{1\leq j\leq i} \alpha_j/p_j$ is maximized by choosing p_j such that $\alpha_j/p_j = \alpha_i/p_i$ for $1 \leq j \leq i$. The maximum value is $(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \cdots + \frac{1}{\alpha_i})^{-1}$.

In [YaSi93b], a result similar to (35) is also established. However, the E.B.B. decay parameters for their output processes are obtained in a recursive manner such that they depend on the E.B.B. characterization of the output processes of the previous sessions in a feasible ordering. Furthermore, as their argument is based directly on the analysis of the output processes, they cannot take advantage of the independence assumption of the input processes to derive better bounds as we obtained in Theorem 7.

5 Feasible Partition and Rate Proportional Processor Sharing

In the analysis of the last section, we observe that the backlog and delay bounds of the ith session in a feasible ordering with respect to $\{r_i\}_{1\leq i\leq N}$ depend only on sessions $1, \ldots, i-1$ and not on sessions $i+1, \ldots, N$. Therefore the position of a session in a feasible ordering affects the bounds we obtain. Note that the feasible ordering is defined with respect to the parameters $\{r_i\}_{1\leq i\leq N}$ that we chose. There are many different ways to choose $\{r_i\}_{1\leq i\leq N}$ that satisfies (5). Furthermore, even for a given $\{r_i\}_{1\leq i\leq N}$, there are possibly many different feasible orderings associated with the sessions. It is natural to wonder whether there is something inherent in these orderings, something uniquely determined by the system parameters $\{\phi_i\}_{1\leq i\leq N}$ and $\{\rho_i\}_{1\leq i\leq N}$. Close scrutiny shows these orderings are essentially determined by the ratios, $\{\rho_i/\phi_i\}_{1\leq i\leq N}$. This leads to a partition of the sessions $\mathcal{H} = \{H_l\}_{1\leq l\leq L}$, $H_1 \cup \cdots \cup H_L = \{1, 2, \ldots, N\}$, where H_l is defined recursively as follows:

$$i \in H_1 \text{ if } \frac{\rho_i}{\phi_i} < \frac{1}{\sum_{j=1}^{N} \phi_j}, \tag{37}$$

and for $k \geq 1$, if $H^k := H_1 \cup \cdots \cup H_k \neq \{1, 2, \ldots, N\}$, then H_{k+1} is defined such that

$$i \in H_{k+1} \text{ if } \frac{\rho_i}{\phi_i} < \frac{1}{\sum_{j\notin H^k} \phi_j}(1 - \sum_{j\in H^k} \rho_j). \tag{38}$$

We call such a partition $\mathcal{H}$ the *feasible partition* induced by $\{\phi_i\}_{1\leq i\leq N}$ and $\{\rho_i\}_{1\leq i\leq N}$. From the definition, we see that for any $i \in H_{k+1}$,

$$\frac{1}{\sum_{j\notin H^{k-1}} \phi_j}(1 - \sum_{j\in H^{k-1}} \rho_j) \leq \frac{\rho_i}{\phi_i} < \frac{1}{\sum_{j\notin H^k} \phi_j}(1 - \sum_{j\in H^k} \rho_j). \tag{39}$$

In particular, a session is in H_1 if its long term upper rate, ρ_i, is smaller than its guaranteed backlog clearing rate, g_i.

The feasible partition induces a partial ordering among the sessions. Intuitively, for a given set of system parameters, this partial ordering captures the inherent "priority" among the sessions so that the decay rate of the backlog of a session in the partition class H_k will be determined only by the sessions in the partition classes $H_l, 1 \leq l < k$, but not by the other sessions in H_k or in $H_l, l > k$; moreover, the backlog of this session cannot decrease faster than the backlogs of those sessions in the lower-numbered partition classes.

For $1 \leq k \leq L$, let us lump together the sessions in H_k into a single new session, called the *aggregate session* k. For $0 < \theta < \min_{i \in H_k} \alpha_i$, the associated arrival process for the aggregate session k is a $(\bar{\rho}_i, e^{\theta \bar{\sigma}_k(\theta)}, \theta)$-E.B.B. process with $\bar{\rho}_k = \sum_{i \in H_k} \rho_i$ and $\bar{\sigma}_k(\theta) = \sum_{i \in H_k} \hat{\sigma}_i(\theta)$. We have an induced single server GPS system with L sessions and a GPS assignment $\{\bar{\phi}_k\}_{1 \leq k \leq L}$ where $\bar{\phi}_k = \sum_{i \in H_k} \phi_i$. By the definition of feasible partition, we see that with respect to $\{\bar{\rho}_k\}_{1 \leq k \leq N}$, there is a *unique* feasible ordering on these L aggregate sessions such that

$$\frac{\bar{\rho}_1}{\bar{\phi}_1} < \frac{1}{\sum_{l=1}^{L} \bar{\phi}_l} \leq \frac{\bar{\rho}_2}{\bar{\phi}_2} < \ldots \leq \frac{\bar{\rho}_k}{\bar{\phi}_k} < \frac{1}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l) \leq \ldots \leq \frac{\bar{\rho}_L}{\bar{\phi}_N}. \tag{40}$$

More interestingly, this unique feasible ordering is preserved in the decomposition of the induced GPS system as stated in the following lemma.

Lemma 9 *For $1 \leq k \leq L$, let $\bar{r}_k = \bar{\rho}_{i_k} + \bar{\epsilon}_k$ where $\bar{\epsilon}_k > 0$. If $\sum_{k=1}^{L} \bar{r}_k \leq 1$, then $1, 2, \ldots, L$ is also a feasible ordering with respect to $\{\bar{r}_k\}_{1 \leq k \leq L}$, i.e., for $1 \leq k \leq L$,*

$$\bar{r}_k \leq \frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{r}_l). \tag{41}$$

Proof: As $\bar{r}_l = \bar{\rho}_l + \bar{\epsilon}_l, 1 \leq l \leq L$, then for any k, $1 \leq k \leq L$, we have

$$\begin{aligned} \frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{r}_l) &= \frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l - \sum_{l=1}^{k-1} \bar{\epsilon}_l) \\ &\geq \frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l) - \sum_{l=1}^{k-1} \bar{\epsilon}_l. \end{aligned} \tag{42}$$

Hence, if we can prove that

$$\frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l) - \bar{\rho}_k \geq \sum_{l=1}^{k} \bar{\epsilon}_l, \tag{43}$$

then (41) follows.

Let $\epsilon = 1 - \sum_{l=1}^{L} \bar{\rho}_l$, then $\epsilon \geq \sum_{l=1}^{L} \bar{\epsilon}_l \geq \sum_{l=1}^{k} \bar{\epsilon}_l$. We prove that the left hand side of (43) is bigger than ϵ. By the definition of the feasible partition (see (40)), for any l, $k < l \leq L$,

$$\bar{\rho}_l > \frac{\bar{\phi}_l}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l). \tag{44}$$

Summing both sides from $k+1$ to L yields

$$\sum_{l=k+1}^{L} \bar{\rho}_l > \frac{\sum_{l=k+1}^{L} \bar{\phi}_l}{\sum_{l=k}^{L} \bar{\phi}_{h^i l}}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l). \tag{45}$$

Therefore,

$$1 - \sum_{l=k+1}^{L} \bar{\rho}_l < \sum_{l=1}^{k-1} \bar{\rho}_l + \frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l). \tag{46}$$

Thus we have

$$\frac{\bar{\phi}_k}{\sum_{l=k}^{L} \bar{\phi}_l}(1 - \sum_{l=1}^{k-1} \bar{\rho}_l) - \bar{\rho}_k > \epsilon. \tag{47}$$

■

Using the notion of feasible partition, we can derive better bounds than those obtained in the previous section. For example, for sessions in H_1, even without assuming that $A_1, \ldots, A_N$ are independent, we have the following bounds.

Theorem 10 *Let session i be a session in H_1, then at any time t, for any $q > 0$,*

$$Pr\{Q_i(t) \geq q\} \leq \Lambda_i^* e^{-\alpha_i q} \tag{48}$$

and for any $d > 0$,

$$Pr\{D_i(t) \geq d\} \leq \Lambda_i^* e^{-\alpha_i g_i d} \tag{49}$$

where

$$\Lambda_i^* = \frac{\Lambda_i e^{\alpha_i \rho_i \xi}}{1 - e^{-\alpha_i (g_i - \rho_i)\xi}} \tag{50}$$

and $0 < \xi \leq \frac{\ln(\Lambda_i + 1)}{\alpha_i (g_i - \rho_i)}$.

Proof: For any session i in H_1, consider a sample path of the session i arrival process A_i. Take $r_i = g_i = \frac{\phi_i}{\sum_{j=1}^{N} \phi_j}$. For any $t \geq 0$, let τ be the beginning of a session i busy period that contains t. Then $S_i(\tau, t) \geq g_i(t - \tau)$. As $A_i(\tau, t) \leq g_i(t - \tau) + \delta_i(t)$, we have $Q_i(t) \leq \delta_i(t)$. Thus $Pr\{Q_i(t) \geq q\} \leq Pr\{\delta_i(t) \geq q\}$ and $Pr\{D_i(t) \geq d\} \leq Pr\{\delta_i(t) \geq g_i d\}$. The theorem now follows from Lemma 5. ■

More generally, for sessions in the partition classes other than H_1, better bounds can also be obtained. The general result is stated below, the proof is slightly more complicated. In the following, for any session i in H_k, $1 \leq k \leq L$, $\psi_i = \frac{\phi_i}{\sum_{j \notin H_1^{k-1}} \phi_j}$.

Theorem 11 *Suppose that $\{A_i\}_{1 \leq i \leq N}$ are N independent $(\rho_i, \Lambda_i, \alpha_i)$-E.B.B processes sharing a GPS server with assignment $\{\phi_i\}_{1 \leq i \leq N}$. Assume that $\sum_{i=1}^{N} \rho_i < 1$ and $\mathcal{H} = \{H_k\}_{1 \leq k \leq L}$ is the feasible partition with respect to $\{\rho_i\}_{1 \leq i \leq N}$ and $\{\phi_i\}_{1 \leq i \leq N}$. Let i be any session H_k where $1 \leq k \leq L$. Then at any time t, for any $q > 0$,*

$$Pr\{Q_i(t) \geq q\} \leq \tilde{\Lambda}_i^{out} e^{-\theta q}, \tag{51}$$

and for any $d > 0$,

$$Pr\{D_i(t) \geq d\} \leq \tilde{\Lambda}_i^{out} e^{-\theta g_i d}, \tag{52}$$

moreover, for any fixed time interval $[\tau, t]$ and any $x > 0$,

$$Pr\{S_i(\tau, t) \geq \rho_i(t - \tau) + x\} \leq \tilde{\Lambda}_i^{out} e^{-\theta x} \tag{53}$$

where

$$\tilde{\Lambda}_i^{out} = \frac{\exp(\theta[\hat{\sigma}_i(\theta) + \rho_i + \psi_i \sum_{j \in H_1^{k-1}} (\hat{\sigma}_j(\psi_i \theta) + \rho_j)])}{(1 - e^{-\theta(g_i - \rho_i)/k})^k}. \tag{54}$$

and $0 < \theta < \min\{\alpha_i, \min_{j \in H_1^{k-1}} \alpha_j\}$.

Proof: For a session i in H_k, $1 \leq k \leq L$, we would like to find the tightest bound possible on the session i backlog (and delay) distribution. From the result in the previous session, we see that the position session i lies in a feasible ordering will affect the bound we derive. A feasible ordering where session i has the lowest possible position is the most desirable.Let $\{r_i\}_{1 \leq i \leq N}$ be a set of numbers such that $\sum_{j=1}^{N} r_j \leq 1$. Since the bounds on the session i backlog and delay only depend on the sessions ahead of i and not on the sessions after i in a feasible ordering with respect to $\{r_i\}_{1 \leq i \leq N}$, it is only essential to require that $\rho_j < r_j$ for $j \leq i$. As long as $\sum_{j=1}^{N} r_j \leq 1$, a feasible ordering with respect to $\{r_i\}_{1 \leq i \leq N}$ always exists. Furthermore, recall that we can regard the sessions in a partition class as an aggregate session. Now consider the GPS system consisting of the first $k-1$ aggregate sessions and the rest of the sessions $j \in H_l$, $l \geq k$. We will construct a feasible ordering on these (mixed) sessions such that session i of class H_k will be the kth session in a feasible ordering. For $1 \leq j \leq i$, let $\epsilon_j = r_j - \rho_j$, and for $1 \leq l \leq k-1$, let $\tilde{r}_l = \sum_{j \in H_l} r_j$ and $\tilde{\epsilon}_l = \sum_{j \in H_l} \epsilon_j$. As long as

$$\epsilon_i + \psi_i \sum_{l=1}^{k-1} \tilde{\epsilon}_l \leq g_i - \rho_i \tag{55}$$

one can show, using the property of feasible partition, that there is a feasible ordering such that the aggregate sessions $1, 2, \ldots, k-1$ and session i are the first k sessions in the ordering. Hence we can appy Theorem 7 to session i with respect to this ordering to derive bounds on backlog and delay distribution for session i. For example, $\epsilon_i = \psi_i \bar{\epsilon}_1 = \cdots = \psi_i \bar{\epsilon}_{k-1} = \frac{g_i - \rho_i}{k}$ satisfies (55). Moreover, as $\sum_{l=1}^{k-1} \bar{r}_l + r_i \leq \sum_{l=1}^{k-1} \bar{\rho}_l + \rho_i + \psi_i^{-1}(g_i - \rho_i) = 1 - (\psi_i^{-1} - 1)\rho_i \leq 1$, we can choose $r_j, j \in H_l, l \geq k$ and $j \neq i$, such that $\sum_{l=1}^{k-1} \bar{r}_l + \sum_{j \in H_l, l \geq k} r_j \leq 1$. With respect to this set of numbers, there is a feasible ordering such that the aggregate sessions $1, 2, \ldots, k-1$ and session i are the first k sessions in the ordering. Hence the theorem follows. ∎

In the case that the arrival processess are not independent, using Hölder's inequality and the same argument as above, we have

Theorem 12 *Suppose that* $\{A_i\}_{1 \leq i \leq N}$ *are* N $(\rho_i, \Lambda_i, \alpha_i)$*-E.B.B processes sharing a GPS server with assignment* $\{\phi_i\}_{1 \leq i \leq N}$. *Assume that* $\sum_{i=1}^N \rho_i < 1$ *and* $\mathcal{H} = \{H_k\}_{1 \leq k \leq L}$ *is the feasible partition with respect to* $\{\rho_i\}_{1 \leq i \leq N}$ *and* $\{\phi_i\}_{1 \leq i \leq N}$. *Let* i *be any session* H_k *where* $1 \leq k \leq L$ *and let* $\{p_l\}_{1 \leq l \leq k}$ *be such that* $p_l > 1$ *and* $\sum_{l=1}^k \frac{1}{p_l} = 1$. *Then at any time* t, *for any* $q > 0$,

$$Pr\{Q_i(t) \geq q\} \leq \tilde{\Lambda}_i^{out'} e^{-\theta q}, \tag{56}$$

and for any $d > 0$,

$$Pr\{D_i(t) \geq d\} \leq \tilde{\Lambda}_i^{out'} e^{-\theta g_i d}, \tag{57}$$

moreover, for any fixed time interval $[\tau, t]$ *and any* $x > 0$,

$$Pr\{S_i(\tau, t) \geq \rho_i(t - \tau) + x\} \leq \tilde{\Lambda}_i^{out'} e^{-\theta x} \tag{58}$$

where

$$\tilde{\Lambda}_i^{out'} = \frac{\exp(\theta[\hat{\sigma}_i(p_k\theta) + \rho_i + \psi_i \sum_{l=1}^{k-1} \sum_{j \in H_l}(\hat{\sigma}_j(p_l\psi_i\theta) + \rho_j)])}{\prod_{l=1}^k (1 - e^{-p_l\theta(g_i - \rho_i)/k})}. \tag{59}$$

and $0 < \theta < \min\{\alpha_i, \min_{1 \leq l \leq k-1} \alpha_j : j \in H_l\}$.

Note that $\min\{\alpha_i, \min_{1 \leq l \leq k-1} \alpha_j : j \in H_l\}$ is maximized by choosing p_l such that $\bar{al}_l/p_l = \alpha_k/p_k$ where $\bar{\alpha}_l = \sum_{j \in H_l} \alpha_j$ for $1 \leq l < k$. The maximum value is $(\frac{1}{\bar{\alpha}_1} + \cdots + \frac{1}{\bar{\alpha}_{k-1}} + \frac{1}{\alpha_k})^{-1}$.

A special but important GPS assignment is the so-called *Rate Proportional Processor Sharing* (RPPS) GPS assignment where $\phi_i = \rho_i, 1 \leq i \leq N$. It is easy to show that the feasible partition $\mathcal{H}$ comprises of only one partition class, $H_1 = \{1, \ldots, N\}$. Therefore Theorem 10 applies to all sessions, yielding simple stochastic bounds. These bounds only depend on the source traffic characterization of the session, not on those of other sessions. Hence, under the RPPS GPS assignment, from a bounding standpoint, each session appears to behave "independently" even though it may be correlated with other sessions.

6 Statistical Analysis of GPS Networks

Consider a network consisting of M nodes labelled $m = 1, \ldots, M$ that each uses GPS. The traffic source for session i is modeled as a $(\rho_i, \Lambda_i, \alpha_i)$ E.B.B. process as before. At node m, the service rate is r^m and the GPS assignment for a session i is ϕ_i^m. The set of sessions present at node m is denoted by $I(m)$. For $1 \leq i \leq N$, let A_i be the random process describing the session i traffic entering the network. The route traversed by session i is denoted by $P(i)$ and the kth node in $P(i)$ by $P(i,k)$. K_i is the total number of nodes in $P(i)$. For $1 \leq k \leq K_i$, denote the session i arrival process at the kth node of its route by $A_i^{(k)}$, and the session i departure by $S_i^{(k)}$. Note that we have $A_i^{(1)} = A_i$ and $A_i^{(k+1)} = S_i^{(k)}$, i.e., the session i arrival process at node $P(i, k+1)$ is the session i departure process at node $P(i,k)$. $S_i^{(K_i)}$ describes the session i traffic that leaves the network. Similarly, denote the session i backlog (resp. delay) at node $P(i,k)$ at time t by $Q_i^{(k)}(t)$ (resp. $D_i^{(k)}(t)$), the total amount of session i traffic queued in the network at time t by $Q_i^{net}(t)$ and the session i end-to-end delay at time t by $D_i^{net}(t)$. A network system (resp. session i) busy period is defined to be the maximal interval B (resp. B_i) such that for every $\tau \in B$ (resp. $\tau \in B_i$), there is at least one server in the network that is in a system (resp. session i) busy period at time τ.

We say a GPS network is *stable* if, for each session i in the network, $\lim_{q\to\infty} Pr\{Q_i^{net}(t) > q\} = 0$, or equivalently, $\lim_{d\to\infty} Pr\{D_i^{net}(t) > d\} = 0$. If at every node of the network, the backlog (or the delay) process for each session at the node is an exponentially bounded (E.B.) process, then clearly the network is stable. If we have an acyclic GPS network with E.B.B. arrival processes, then by the input-output relation established in Section 4, the network can be easily shown to be stable. However for a cyclic network, stability is generally much harder to establish. In this section, we will show that under a broad class of GPS assignments a GPS network with arbitrary topology is stable if $\sum_{i\in I(m)} \rho_i < r^m$ for all nodes m in the network. This class of GPS assignment is called the class of *Consistent Relative Session Treatment* (CRST) networks. We will then look at a special case of CRST networks, the so-called class of *Rate Proportional Processor Sharing* networks. Finally, we present a simple numerical example to illustrate the bounds we derive for the RPPS GPS networks.

6.1 Stability of CRST GPS Networks

The class of the Consistent Relative Session Treatment (CRST) networks is defined as follows. For each node m, $m = 1, \ldots, M$, let $\{\phi_i^m\}_{i\in I(m)}$ be the GPS assignment for the sessions at node m and $\mathcal{H}^m = \{H_l^m\}_{1\leq l\leq L_m}$ be the induced feasible partition of the sessions at node m. We say this collection of GPS assignments $\{\phi_i^m : i \in I(m), 1 \leq m \leq M\}$ for the network is a Consistent Relative Session Treatment (CRST) GPS assignment if there is a partition $\mathcal{H} = \{H_l\}_{1\leq l\leq L}$ of the N sessions in the network such that $\mathcal{H}$ is *consistent* with $\mathcal{H}^m$ at each node m, i.e., for any $i, j \in I(m)$, if $i \in H_l$ and $j \in H_k$ such that $l < k$, then $i \in H_{l'}^m$ and

$j \in H_{k'}^m$ such that $l' \leq k'$. We call $\mathcal{H}$ the *CRST partition* of the sessions in the network. Networks with CRST GPS assignment are called CRST GPS networks. It is worth pointing out that CRST is a condition imposed on the global GPS assignments, not on the topology of the network.

The stability of CRST GPS networks can be established by recursively applying the results obtained for the single node case. Let $\mathcal{H} = \{H_l\}_{1 \leq l \leq L}$ be the CRST partition of the N sessions in the network. From the statistical analysis of a single GPS server, we see that for a session $i \in H_l$, at any node m along its route, the bounds on the distribution of session i backlog $Q_i^m(t)$ and delay $D_i^m(t)$, $t \geq 0$, $1 \leq m \leq N$ depend only on the sessions at node m that are in $H_k, k < l$, and not on the sessions at node m that are in $H_k, k \geq l$. The same statement also holds for the characterization of the session i output process at node m, S_i^m, i.e. the decay coefficient $\Lambda_i^{m,out}$ and the decay factor $\alpha_i^{m,out}$ are functions of only those parameters of the arrival E.B.B. processes A_j^m for sessions j at node m where $j \in H_k, k < l$. This suggests a recursive procedure for computing backlog and delay bounds and characterizing the output process for each session at any node along its route. For sessions in H_1, the interested performance metrics can be obtained independently along their routes by applying the input-output relations and the bounds established for the single node case. For any k, $2 \leq k \leq N$, once the output process characterization has been derived for each session in $H_l, l < k$, at every node of its route, the stochastic bounds on backlog and delay distribution and the output process characterization for sessions in H_k can then be derived at any node along their routes. Finally the end-to-end performance metrics such as the stochastic bound on the end-to-end delay can be computed by convolving the per-node bounds along the session routes. Therefore, we have

Theorem 13 *Given that each session i in a CRST GPS network is a $(\rho_i, \Lambda_i, \alpha_i)$-E.B.B. process, then it is stable if $\sum_{i \in I(m)} \rho_i < r^m$ at each node m where r^m is the service rate at node m.*

Remark Our definition of CRST is slightly weaker than the one introduced by Parekh and Gallager using the notion of "impede". Under their definition, if at node m session i impedes session j, i.e., $\frac{\rho_i^m}{\phi_i^m} < \frac{\rho_j^m}{\phi_j^m}$, but node m', session j impedes session i, then this GPS assignment is not CRST. However as long as session i and session j belong to the same class of the feasible partition at every node they share, this GPS assignment will still be CRST under our definition. On the other hand, a CRST GPS assignment under their definition is clearly also a CRST GPS assignment under our definition.

6.2 Rate Proportional Processor Sharing GPS Networks

A Rate Proportional Processor Sharing (RPPS) GPS network is a network where, at every node m of the network and for every session $i \in I(m)$, $\phi_i^m = \rho_i$. As $\rho_i / \phi_i^m = 1$ for any session i and at each node m, the CRST partition $\mathcal{H}$ consists of only one class H_1 encompassing all sessions. For RPPS networks, closed-form expressions exist for the backlog and delay distribution bounds as will be proved below.

Let g_i^m denote the backlog clearing rate for session i at node m, i.e.,

$$g_i^m = \frac{\phi_i^m}{\sum_{j\in I(m)} \phi_j^m} r^m \tag{60}$$

where r^m is the service rate at node m. Define $g_i^{net} = \min_{m\in P(i)} g_i^m$. Then session i is guaranteed a backlog clearing rate of g_i^{net} at every node along its route. Furthermore, if $\sum_{i\in I(m)} \rho_i < r^m$, then the RPPS GPS assignment implies that $g_i^{net} > \rho_i$.

The following important observation is a restatement of Lemma 3.2 in [Parekh92] which can be shown to hold also in the case we are considering here.

Lemma 14 *For every interval $[\tau, t]$ that is contained in a single session i network busy period, we have*

$$S_i^{(K_i)}(\tau, t) \geq g_i^{net}(t-\tau). \tag{61}$$

With the help of this lemma, we can prove the following interesting result.

Theorem 15 *If every session i in a RPPS GPS network is a $(\rho_i, \Lambda_i, \alpha_i)$-E.B.B. process, and at every node m, $\sum_{j\in I(m)} \rho_j < r^m$ where r^m is the rate of the server at node m. Then at any time t, for any $q > 0$,*

$$Pr\{Q_i^{net}(t) \geq q\} \leq \Lambda_i^{net} e^{-\alpha_i q} \tag{62}$$

and for any $d > 0$,

$$Pr\{D_i^{net}(t) \geq d\} \leq \Lambda_i^{net} e^{-\alpha_i g_i^{net} d} \tag{63}$$

where

$$\Lambda_i^{net} = \frac{\Lambda_i e^{\alpha_i \rho_i \xi}}{1 - e^{-\alpha_i (g_i^{net} - \rho_i)\xi}} \tag{64}$$

and $0 < \xi < \frac{\ln(\Lambda_i + 1)}{\alpha_i (g_i^{net} - \rho_i)}$.

Proof: For $1 \leq i \leq N$, consider a sample path A_i of A_i. Define $r_i = \rho_i + \epsilon_i$ where $\epsilon_i = g_i^{net} - \rho_i$. For any $t \geq 0$, let τ be the first time before t when there is no session i traffic backlogged in the network. From (7) and Lemma 14, we have

$$\begin{aligned} Q_i^{net}(t) &= A_i(\tau, t) - S_i^{(K_i)}(\tau, t) \\ &\leq r_i(t-\tau) + \delta_i(t) - g_i^{net}(t-\tau) \\ &\leq \delta_i(t). \end{aligned}$$

As session i is guaranteed a backlog clearing rate of g_i^{net}, hence

$$D_i^{net}(t) \leq \frac{Q_i^{net}(t)}{g_i^{net}} \leq \frac{\delta_i(t)}{g_i^{net}}. \tag{65}$$

For $0 < \theta < \alpha_i$, applying Lemma 5, we have that for any $q \geq 0$,
$Pr\{Q_i^{net}(t) \geq q\} \leq Pr\{\delta_i(t) \geq q\} \leq \Lambda_i^{net} e^{-\theta q}$
and similarly for any $d \geq 0$, $Pr\{D_i^{net}(t) \geq d\} \leq Pr\{\delta_i(t) \geq g_i^{net} d\} \leq \Lambda_i^{net} e^{-\theta g_i^{net} q}$. ■

This theorem reveals that under the RPPS GPS assignment, the backlog and delay bounds are independent of the route length and the topology of the network, a result analogous to Parekh and Gallager's deterministic result. It essentially reduces the analysis of a network of queues to that of a single queue, more precisely, that of the bottleneck node (the node with the minimum g_i) for each session. Since the stochastic bounds come from the bound on $\delta_i(t)$, when applying Theorem 15 to more specific arrival processes, we can often get around the E.B.B. characterization and directly use results regarding $\delta_i(t)$ to obtain tighter bounds. For example, if arrivals are described by a Markov modulated arrival process, then results in [LNT94] can be used to obtain tighter bounds than those given by Theorem 15. However, results like those in [LNT94] are not always available for other types of traffic sources.

Last, we remark that Theorem 15 actually applies to any session i that is guaranteed a backlog clearing rate of $g_i^{net} > \rho_i$ at all nodes along its route, regardless of what GPS assignment is used.

6.3 A Numerical Example

In this section we present a simple numerical example to illustrate the results we obtained for RPPS GPS networks. We consider a simple three-node tree structured network (Figure 2) as used in [YaSi93]. The rate of the servers and the capacity of the links are all assumed to be 1. Suppose there are only 4 sessions in the network, two sessions at node 1 and two sessions at node 2; all sessions congregate at node 3. The source traffic for each session is modeled by a discrete time two-state on-off Markov process. For $1 \leq i \leq 4$, the transition probability from the off-state to the on-state is p_i, and from on-state to off-state is q_i. The traffic rate in the on-state is λ_i and in the off-state zero. The average traffic rate $\bar{\lambda}_i$ is $\frac{p_i \lambda_i}{p_i + q_i}$. The values for the parameters of the four arrival processes are listed in Table 1. To obtain E.B.B. characterizations for the arrival processes, we can choose any $\rho_i, i = 1, 2, 3, 4$ such that $\rho_i > \bar{\lambda}_i$ and $\sum_{i=1}^{4} \rho_i < 1$. Two sets of ρ_i's are shown in Table 2, the α_i's and Λ_i's are obtained using the results for discrete time two-state on-off Markov processes in [LNT94].

Under the RPPS GPS assignment, $\phi_i^m = \rho_i$ for $i = 1, 2, 3, 4$ and for $m = 1, 2, 3$. Using the discrete version of the bound (18) and then applying Theorem 15, we have that for $q > 0$,

$$Pr\{Q_i^{net}(t) \geq q\} \leq \frac{\Lambda_i}{1 - e^{-\alpha_i(g_i - \rho_i)}} e^{-\alpha_i q} \tag{66}$$

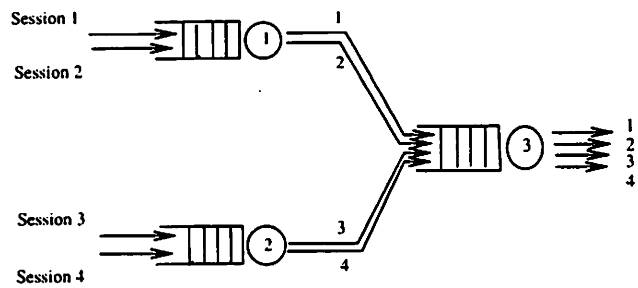


Figure 2: **An Example Network.**

session	p_i	q_i	λ_i	λ_i
1	0.3	0.7	0.5	0.15
2	0.4	0.4	0.4	0.2
3	0.3	0.3	0.3	0.15
4	0.4	0.6	0.5	0.2

Table 1: Parameters for the Arrival Processes

and for $d > 0$,

$$Pr\{D_i^{net}(t) \geq d\} \leq \frac{\Lambda_i}{1 - e^{-\alpha_i(g_i - \rho_i)}} e^{-\alpha_i g_i d} \tag{67}$$

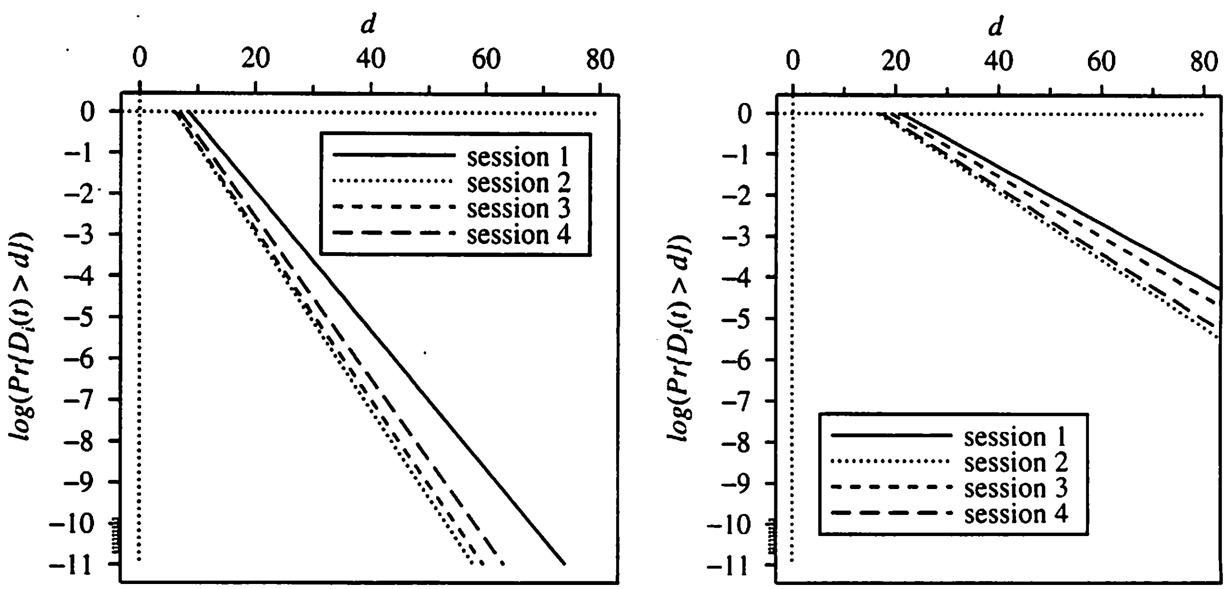
The bounds on the end-to-end delay distributioon for the four sessions using the two sets of E.B.B. parameters are shown in Figure 3(a) and Figure 3(b), respectively. The bounds on the backlog distributions for the four sessions looks similar are not included here.

From Table 2 and Figure 3, we see that the smaller ρ_i we choose for a session, the larger Λ_i is and the smaller α_i is. This in turn leads to a slower decay rate of the delay bound. On the other hand, the smaller ρ_i for each session, the smaller the sum of ρ_i's. From the perspective of call admission control, this means that the number of calls admitted into a system can be increased. What ρ_i to choose is closely related to what QoS requirement a session demands and what call admission control mechanism is employed in the network, we will address these issues in a later paper. It is worth pointing out here that in this context it is more appropriate to use C. S. Chang's *Envelope Processes* model as the source characterization. Although the two are the same in essense, C. S. Chang's model allows one to use the elegant notion of *effective badnwidth* in addressing the issues of QoS requirement and call admission control.

This example also helps to illustrate another disadvantage of E.B.B. process model. From

	Set 1			Set 2		
session	ρ_i	Λ_i	α_i	ρ_i	Λ_i	α_i
1	0.2	1.0	1.74	0.17	1.0	0.729
2	0.25	0.92	1.76	0.22	0.968	0.672
3	0.2	0.84	2.13	0.17	0.929	0.775
4	0.25	1.0	1.62	0.22	1.0	0.655

Table 2: Two Sets of E.B.B. Characterizations for the Arrival Processes

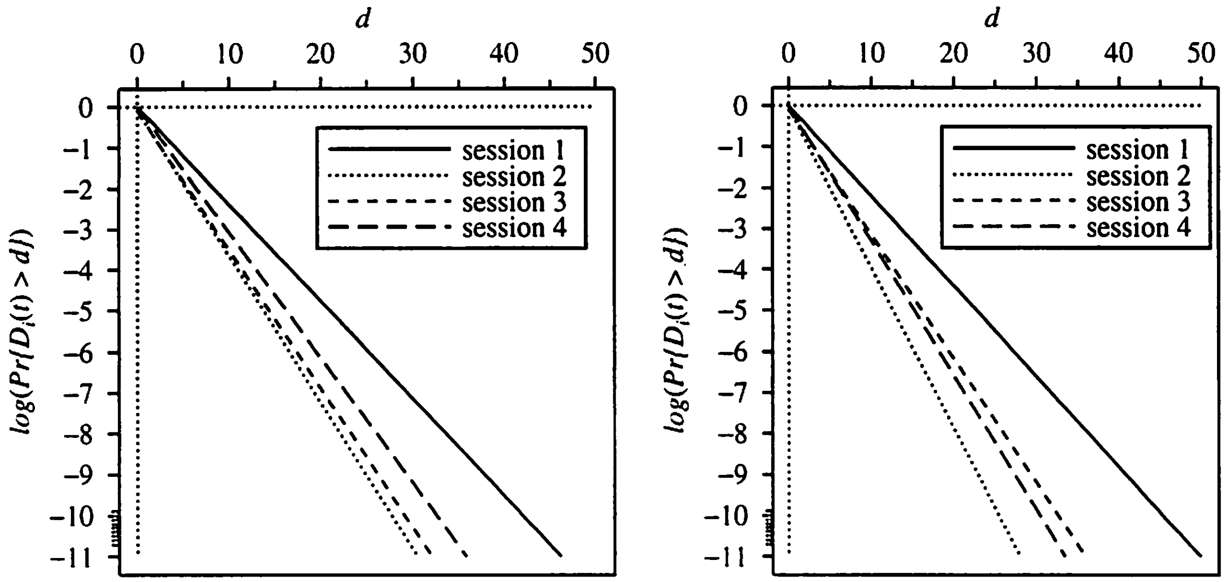


(a) Set 1

(b) Set 2

Figure 3: **Bounds on the End-to-End Delay Distributions in Logscale.**

Table 2, we see that a decrease of ρ_i from Set 1 to Set 2 leads to drastic decrease in the value of α_i, thus the *delay bounds* in Figure 3(b) decay much slower than those of Figure 3(a) . However, if one looks at the guaranteed bandwidths g_i's, for session 1 and session 3, both g_1 and g_3 decrease from 0.22 to approximately 0.218, whereas both g_2 and g_4 actually increase from 0.28 to approximately 0.282 for session 2 and sesion 4. Under RPPS, this should imply that the *real* end-to-end delays of sessions 1 and 3 should decay slightly slower and those of sessions 2 and 4 should decay slightly faster. However, this fact is not reflected in Figure 3(b), where the decay rates of the delay bound for all sessions become much smaller. The problem lies in the E.B.B. model. As we choose ρ_i close to the mean rate $\bar{\rho}_i$ of a session, α_i decreases rapidly. This drastic decrease in α_i offsets the slight increase in the guaranteed bandwith sharing for sessions 2 and 4. When using the E.B.B. source characterization, since the decay



(a) Set 1 (b) Set 2

Figure 4: **Improved Bounds on the End-to-End Delay Distributions in Logscale.**

rate of the delay bound is no smaller than α_i, the decay rate thus obtained will be limited by the value of α_i. Fortunately, if we know the specific structure of the arrival processes, this problem can, in general, be circumvented. By applying a better bound on $\delta_i(t)$ (such a bound is usually readily available) than the bound obtained from the E.B.B. characterization in Lemma 5 or in Lemma 6, the theorems in the earlier sections, for instant, Theorem 15, can be easily modified to produce tighter bounds. In this example, we can use the results in [LNT94] to directly bound $\delta_i(t)$, yielding tighter bounds as shown in Figure 4.

Lastly, we note that since the two-state on-off Markov processes do not satisfy Cruz's LBAP model, no worst-case deterministic bounds can be derived by Parekh and Gallager's method.

7 Conclusions and Future Work

In this paper we studied the statistical behavior of the generalized processor sharing (GPS) scheduling discipline using exponentially bounded burstiness (E.B.B.) processes as source session traffic models and derived upper bounds on the tail distributions of session backlog and delay, both for a single GPS server in isolation and RPPS GPS networks with arbitrary topology. We also established the stability of CRST GPS network with E.B.B. arrival processes. Although our main focus is on GPS, the results can be easily extended to packetized version of GPS —PGPS (cf., [PG92a, YaSi93b]). Clearly, simulation needs to be conducted to verify how good the theoretical bounds we derived in this paper are.

Our study places no restriction on the arrival process for each session other than that it has to be an E.B.B process, which appears to include most processes. Although fine-tuning the bounds we obtained is still possible, an unavoidable disadvantage associated with the use of a general traffic model is that the bounds based on such a model will generally be loose. One may take advantage of the specific structure of the arrival processes in question to derive better bounds. For example, several papers (e.g., [LNT94, BD94]) have appeared that assume a more specific traffic model and improve those results of [Chang93, YaSi93a] obtained for a general traffic model in dealing with a general service discipline (in particular FCFS). These bounds can be used directly to bound $\delta_i(t)$ instead of applying Lemma 6, yielding better prefactors for the derived bounds on backlog and delay tail distributions.

Throughout the analysis, we have assumed that the E.B.B. characterizations of the arrival processes are given. In practice, how to obtain these characterizations and how good the characterizations are is a concern, especially given that the E.B.B. parameters chosen will affect the tightness of the performance bounds. There are tradeoffs in the choices of ρ's, α's and Λ's. In general, the closer ρ to the average rate of the arrival rate, the smaller α will be and the bigger Λ will be. It seems that the issue of source traffic characterization and its impact on the performance bounds can be more adequately addressed in the context of C. S. Chang's envelope process model. From the viewpoint of provision of QOS guarantees under GPS scheduling discipline, the issue of source traffic characterization is apparently closely related to the theory of effective bandwidth [e.g., EM93, GAN91, KWC93] established mostly for FCFS scheduling. Note that our analysis of GPS yields an upper bound on the backlog decay rate of each session under GPS. To complete the picture, it will be interesting to provide a lower bound on the backlog decay rate of each session, yielding an analogous theory of effective bandwidth that will be useful for call admission control under GPS scheduling. We believe that the notion of feasible partition will play a role in the determination of the individual session backlog decay rates.

Another issue of importance in practice is how to choose the GPS assignment $\{\phi_i\}_{1\leq i\leq N}$. Recall that in the analysis we have assumed that the number of total sessions in the system, N, is fixed. What happens if a new session wants to join the system. This question is of particular importance from the aspect of call admission control. Clearly, dynamically adding, deleting or updating ϕ's will incur considerable overhead in maintaining the system state and computing performance bounds on-line for the purpose of call admission control. This problem suggests that we might consider combining or integrating GPS with other scheduling policies. In [CSZ92], Clark, et al. discusses the relative merits of GPS and FCFS scheduling disciplines. They argued that GPS is a good scheduling discipline to provide isolation among sessions, but can be too strict to allow sessions, especially sessions with similar characteristics, to maximally exploit the multiplexing gains. The notion of feasible partition of sessions introduced in this paper provides some ideas as how to combine GPS with other scheduling policies to alleviate this problem. We observe that bounds for a session in a feasible partition class H_k only depend on previous classes H_l, $l < k$. If we use, say, FCFS, to schedule sessions within a class, then the bounds for the aggregate sessions can be used as the worst-case statistical bounds for each session within a class, but sessions within a class can exploit the multiplexing gains due to FCFS. One approach one can take is to categorize the traffic in a network into several traffic classes such that traffic with identical or similar characteristics will be grouped into one class, say, voice class, video classes with different resolutions. One reason one may want to do this is that traffic with identical or similar characteristics will likely have identical or similar QOS requirement and thus will be treated similarly by the network. So grouping them together could make resource management such as bandwidth allocation simpler or more efficient. From the analytical standpoint, the advantage of this approach is that traffic with similar characteristics as reflected by the value of ρ_i/ϕ_i will probably fall into the same partition classes defined by the feasible partition. This will greatly simplify the computation of performance bounds. Let us look at an example where we have three traffic classes. For the first class, the sessions are assigned "peak rate" ρ_i, i.e., $\rho_i/\phi_i = 1$; for the second class, the sessions are assigned around 75% of the "peak rate" ρ_i , i.e., $\rho_i/\phi_i \approx 1/0.75 = 4/3$; and for the third class, the sessions are assigned around half of the "peak rate" ρ_i, i.e., $\rho_i/\phi_i \approx 1/0.5 = 2$ for i in class 3. The GPS scheduling discipline is used to provide protection between the classes. For the sessions within a class, FCFS or other scheduling disciplines can be used to exploit multiplexing gains. When computing performance bounds for traffic classes 2 or 3, only the aggregate effect of previous classes need to be taken into account. Moreover, the well-studied notion of effective bandwidth can be applied directly to FCFS sessions within a class for call admission control.

Acknowledgement The authors would like to thank Zhen Liu and Philippe Nain of INRIA for many helpful discussions.

References

[BD94] E. Buffet and N. G. Duffield, Exponential Upper Bounds via Martingales for Multiplexers with Markovian Arrivals, *J. Appl. Probability.*, 1994. To appear.

[Chang94] C. S. Chang, Stability, Queue Length and Delay of Deterministic and Stochastic Queueing Networks, *IEEE Transaction on Automatic Control*, May 1994.

[Cruz91a] R. L. Cruz, A Calculus for Network Delay, Part I: Network Elements in Isolation, *IEEE Transaction on Information Theory*, Vol. 37, No. 1, Jan. 1991, pp. 114-131.

[Cruz91b] R. L. Cruz, A Calculus for Network Delay, Part II: Network Analysis, *IEEE Transaction on Information Theory*, Vol. 37, No. 1, Jan. 1991, pp. 132-141.

[CSZ92] D. Clark, S. Shenker and L. Zhang, Supporting Real-time Applications in an Integrated Service Packet Network: Architecture and Mechanism. In *Proceedings of ACM SIGCOMM'92*, pp. 14-26, 1992.

[DKS89] A. Demers, S. Keshav and S. Shenker, Analysis and Simulation of a Fair Queueing Algorithm, *Journal of Internetworking: Research and Experience*, 1, pp. 3-26, 1990. Also in *Proceedings of ACM SIGCOMM '89*, pp. 3-12.

[EM92] A. Elwalid and D. Mitra, Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks, *IEEE/ACM Trans. on Networking*, Vol. 1, No. 3, June 1993, pp. 329-357.

[GAN91] R. Guérin, H. Ahmadi and M. Naghshineh, Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks, *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 7, Sept.,1991, pp. 968-981.

[Go90] S. J. Golestani, A Stop and Go Queueing Framework for Congestion Management, In *Proceedings of SIGCOMM'90*, pp. 8-18, 1990.

[Go91] S. J. Golestani, Duration-Limited Statistical Multiplexing of Delay Sensitive Traffic in Packet Networks, In *Proceedings of INFOCOMM '91*, 1991.

[KKK90] C. Kalmanek, H. Kanakia and S. Keshav, Rate Controlled Servers for Very High-Speed Networks, In *Proceedings of GlobalComm '90*, pp. 300.3.1-300.3.9, 1990.

[KWC93] G. Kesidis, J. Walrand and C. S. Chang, Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources, *IEEE/ACM Trans. Networking*, Aug. 1993.

[Kurose91] J. Kurose, On Computing Per-Session Performance Bounds in High-Speed Multi-Hop Computer Networks, *Performance Evaluation Review*, Vol. 20, No. 1, June, 1992, pp. 128-139.

[LNT94] Z. Liu, P. Nain and D. Towsley, Exponential Bounds for a Class of Stochastic Processes with Application to Call Admission Control in Networks, to appear in the Proc. of the 33rd *Conference on Decision and Control (CDC'93)*, February, 1994.

[Parekh92] A. K. Parekh, A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks, Ph.D Thesis, Department of Electrical Engineering and Computer Science, MIT, February 1992.

[PG93a] A. K. Parekh and R. G. Gallager, A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case, em IEEE/ACM Transaction on Networking, Vol. 1, No. 3, June 1993, pp. 344-357.

[PG93b] A. K. Parekh, R. G. Gallager, A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case, *Proceedings of IEEE INFOCOM '93*, pp.521-530. To appear in *IEEE/ACM Transaction on Networking*.

[VZF91] D. Verma, H. Zhang and D. Ferrari, Delay Jitter Control for Real-Time Traffic, In *Proceedings of TriCom '91*, pp. 35-43, 1991.

[YaSi93] O. Yaron , M. Sidi, Performance and Stability of Communication Networks via Robust Exponential Bounds, *IEEE/ACM Trans. on Networking*, Vol. 1, No. 3, pp. 372-385, 1993.

[YaSi94] O. Yaron , M. Sidi, Generalized Processor Sharing Networks with Exponentially Bounded Burstiness Arrivals, In *Proceedings of IEEE INFOCOM '94*, pp.??-??, June 1994.

[YKTH93] D. Yates, J. Kurose, D. Towsley and M. Hluchyj, On Per-Session End-to-End Delay and Call Admission Problem for Real-Time Applications with QOS Requirements , *SIGCOMM '93*, 1993.

[ZF93] H. Zhang and D. Ferrari, Rate-Controlled Static-Priority Queueing, In *Proceedings of IEEE INFOCOM '93*, pp. 227-236, April 1993.

A Appendix

Proof of Lemma 1: (9) is equivalent to $\sum_{j=1}^{i} \eta_j(t) \leq 0$. We prove this inequality by induction on i. When $i = 1$, let τ be the beginning of session 1 busy period that contains t, then $Q_1(\tau) = 0$, and $\eta_1(\tau) = -\delta_1(\tau) \leq 0$. By the definition of GPS and the fact that $r_1 < \frac{\phi_1}{\sum_{j=1}^{N} \phi_j}$, we have

$$S_1(\tau, t) \geq \frac{\phi_1}{\sum_{j=1}^{N} \phi_j}(t-\tau) > r_1(t-\tau). \tag{68}$$

Then from (8),

$$\eta_1(t) \leq \eta_1(\tau) + r_1(t-\tau) - S_1(\tau, t) \leq 0. \tag{69}$$

Now we assume the lemma is true for $1, 2, \ldots, i-1$, and show that it is also true for i. First, if $\eta_i(t) \leq 0$, then by the induction hypothesis, the claim follows easily. The case where $\eta_i(t) > 0$ is a bit harder. Note that $\eta_i(t) > 0$ implies that $Q_i(t) > \delta_i(t) \geq 0$. Let τ be the beginning of a session i busy period that contains t, thus $Q_i(\tau) = 0$, and $\eta_i(\tau) = -\delta_i(\tau) \leq 0$. From (8),

$$S_i(\tau, t) \leq r_i(t-\tau) - \eta_i(t) < r_i(t-\tau). \tag{70}$$

Let $x > 0$ be such that

$$S_i(\tau, t) = r_i(t - \tau) - x. \tag{71}$$

As

$$r_i < \frac{\phi_i}{\sum_{j=i}^{N} \phi_j}(1 - \sum_{j=1}^{i-1} r_j), \tag{72}$$

from (71), we have

$$S_i(\tau, t) < \frac{\phi_i}{\sum_{j=i}^{N} \phi_j}(1 - \sum_{j=1}^{i-1} r_j)(t - \tau) - x. \tag{73}$$

Moreover, by the definition of GPS, for any j,

$$S_i(\tau, t) \geq \frac{\phi_i}{\phi_j} S_j(\tau, t). \tag{74}$$

Thus

$$\sum_{j=i}^{N} S_j(\tau, t) \leq (\sum_{j=i}^{N} \frac{\phi_j}{\phi_i}) S_i(\tau, t). \tag{75}$$

Using (73), we have

$$\begin{aligned} \sum_{j=i}^{N} S_j(\tau, t) &< (t - \tau)(1 - \sum_{j=1}^{i-1} r_j) - x \sum_{j=i}^{N} \frac{\phi_j}{\phi_i} \\ &\leq (t - \tau)(1 - \sum_{j=1}^{i-1} r_j) - x. \end{aligned} \tag{76}$$

On the other hand, since the system is in a system busy period,

$$\sum_{j=1}^{i-1} S_j(\tau, t) = t - \tau - \sum_{j=i}^{N} S_j(\tau, t). \tag{77}$$

From (76)

$$\sum_{j=1}^{i-1} S_j(\tau, t) > (t - \tau) \sum_{j=1}^{i-1} r_j + x. \tag{78}$$

Adding (71) to (78) yields

$$\sum_{j=1}^{i} S_j(\tau, t) > (t - \tau) \sum_{j=1}^{i} r_j. \tag{79}$$

Now from (8), for $1 \leq j \leq i$,

$$\eta_j(t) \leq \eta_j(\tau) + r_j(t - \tau) - S_j(\tau, t). \tag{80}$$

Summing over j and using (79), we have

$$\sum_{j=1}^{i} \eta_j(t) \leq \sum_{j=1}^{i} \eta_j(\tau) = \sum_{j=1}^{i-1} \eta_j(\tau) + \eta_i(\tau) \leq 0 \tag{81}$$

where the last inequality follows from the induction hypothesis and the fact that $\eta_i(\tau) \leq 0$. This concludes the proof for the lemma. ■