Journal of High Speed Networks 3 (1994), 375–387
IOS Press

GENERALIZED PROCESSOR SHARING NETWORKS WITH EXPONENTIALLY BOUNDED BURSTINESS ARRIVALS

Opher Yaron* and Moshe Sidi†
Department of Electrical Engineering
Technion — Israel Institute of Technology
Haifa 32000, Israel

Abstract
We consider virtual circuit packet switching communication networks that employ processor sharing type service disciplines and study their stability when the offered traffic to the network has exponentially bounded burstiness (EBB). The advantages of processor sharing switching techniques both in terms of service flexibility and the potential capability to guarantee certain grades of service, were emphasized in [7]. The study presented there assumes that the traffic offered to the network is flow-controlled by the leaky-bucket mechanism, and the performance of both generalized processor sharing (GPS) and packet-by-packet generalized processor sharing (PGPS) is analyzed.

In this paper we employ an exponential characterization (EBB) introduced in [11] to analyze this type of systems in a stochastic setting. We first examine the GPS and PGPS servers in isolation, and demonstrate their superiority to a general server in the stochastic environment. We then show that a network of servers, that are all either GPS or PGPS, is stable whenever the service rate of each node is larger than the total arrival rate to it. In addition, we provide exponential upper bounds to the traffic flows within the network links, and to the backlogs of each session in the various nodes on its path.
Keywords: Communication Networks, Generalized Processor Sharing, Exponential Bounds, Stability, Packet Switching, Burstiness

1. Introduction

The issue of estimating the performance of a given communication network, when it is loaded with some known (at least stochastically) user traffic, is crucial in many

* O. Yaron is currently with PowerSpectrum Technology Ltd., Kiryat Bialik 27000, Israel, a subsidiary of Geotek Communications Inc., Montvale, NJ 07645 (Tel.: (972) 4-792910, Fax: (972) 4-792882, e-mail: opher@pst.co.il).
† This research was supported by The Israel Science Foundation administrated by The Israel Academy of Siences and Humanities.

aspects of the discipline of high-speed networks. Such estimates are unavoidable, for instance, when there is a need to decide upon the amount of resources necessary in the various components of the network to provide acceptable service. Another example is when developing an appropriate admittance policy. Such policy must decide whether the network can meet the requirements of an applying new user, without violating the grades of service guaranteed to existing users. A preliminary condition one must check is that the service rate of each node in the network is larger than the total arrival rate to it. To see why this condition, which we call *the throughput condition*, must hold, consult [11].

Indeed, performance estimation has been addressed in a variety of settings, and in some cases simple solutions are available [6] for various important probability distributions such as queue length, delay, etc. One major case which is of great interest in the area of high-speed networks is that of virtual circuit packet switching. Many studies of recent years, both theoretical and practical, have established this switching technique as a favorable compromise between flexibility of service and performance guarantees. Unfortunately, the problem we are interested in is usually intractable for exact analysis in the context of virtual circuit switching, even in very simple scenarios. Consequently, an alternative approach to exact analysis is needed to handle this class of networks and investigate their performance.

Such an alternative is the calculation of bounds, and several models have been proposed that give rise to this modus operandi. One novel approach, introduced in [1], is to assume that each input traffic stream to the network satisfies some deterministic characterization, called *a burstiness constraint*, and to derive deterministic upper bounds to the backlogs and delays that develop in the nodes of the network. This approach is particularly significant, since high-speed networks are usually considered in conjunction with appropriate flow-control mechanisms that are designed to protect the network from being saturated by ill-behaved customers. One mechanism, which is most popular in this respect, is the Leaky-Bucket [10], that produces a burstiness constrained output. This model has been employed in [1]–[3] to analyze isolated network elements of a general work-conserving service discipline, as well as some specific disciplines, such as FCFS and locally FCFS. These results are easily extended to feed-forward networks. However, general topology networks are proved to be stable only when they are lightly loaded.

Further progress within the model proposed in [1] has been presented in [7]–[9]. In this study two new service disciplines are introduced — General Processor Sharing (GPS), and its packet-by-packet version, Packet-based GPS (PGPS). Apparently, GPS and PGPS are identical to the generalized service disciplines introduced in the Appendix of [5]. These generalized disciplines are the non-uniformly weighted versions of bit-by-bit Round Robin and of Fair Queueing, which are presented and studied in [4], and compared to each other analytically in [5].

GPS and PGPS are analyzed in isolation, and in a network setting, when the parameter assignment belongs to a class called Consistent Relative Session Treatment (CRST), which is broad enough to support a wide variety of service requirements. For leaky-bucket controlled arrivals, it is shown that such a network is stable whenever the throughput condition holds, and bounds to total delays and backlogs within the network are derived.

In a recent study [11] we introduced a different model, designed to take into account the stochastic nature of the input traffic streams to communication networks. We proposed the Exponentially Bounded Burstiness (EBB) characterization, and claimed that most commonly used stochastic models for traffic flows satisfy this characterization. Moreover, we analyzed isolated network elements, as well as feed-forward and cyclic networks assuming that the servers at the nodes are general, and derived sufficient conditions for these systems to be stable. We showed that when these conditions hold, the distributions of backlogs within the network are upper bounded by decaying exponentials.

In this paper we apply the techniques developed in [11] to the system model introduced in [7]. We show that if a general topology network of GPS or PGPS servers, that is assigned with CRST service parameters, is loaded with EBB user traffic streams, then the system is stable whenever the throughput condition holds. This result is based on extending the ideas of [7] in a straightforward manner to the EBB model, and using analytical techniques that are specific to this case, many of them by applying material developed and presented in [11].

Similar results have been recently presented in [12]. The treatment of isolated servers is based on a sample path analysis, which exploits more of the structural properties of GPS, and allows for backlog bounds which are generally tighter than the ones presented here, though still fairly loose. The analysis of GPS networks, however, is parallel to ours.

The notations we use in the analysis will naturally be a combination of those used in [11] and in [7], and they will be briefly reviewed in the next section.

2. Concepts and Notations

2.1. *Network Topology and Routing*

A communication network is modeled by a directed graph, where vertices represent switches and edges represent links. Since we are interested in virtual circuit switching, the route taken by the traffic of each of the users (occasionally called sessions) is a predetermined path in the graph. We will denote the path taken by user i by $P(i) = (P(i,1), P(i,2), \ldots, P(i,K_i))$, where $P(i,k)$ is the kth node in the path $P(i)$, K_i is its length.

2.2. *Traffic Flow Characterization*

A traffic flow in the network is a stochastic process, specified by an instantaneous rate $R(t)$. The amount of data carried by this traffic flow during the time interval $[s,t]$ will be denoted by $R^{s,t} = \int_s^t R(u)\,du$. Subscripts will be used to denote the session. Thus, R_i is the session i traffic flow into the network, $R_{i,k}$ is the session i traffic flow into $P(i,k)$, and R_{i,K_i+1} is the session i traffic that leaves the network. We will also use the notation $R_{i,k,\text{out}}$ for session i traffic that leaves node $P(i,k)$ after being served there. Thus, the session i traffic that leaves the network can also be referred to by $R_{i,K_i,\text{out}}$.

Remark. When we refer to a particular switch v in the network we will use a somewhat different notation. $I(v)$ will be the set of sessions that are served by v, $R_{i,v}$ the session i traffic that enters v, $R_{i,v,\mathrm{out}}$ the session i traffic that leaves v.

The offered traffic processes we will consider are all EBB. The EBB characterization of session i traffic flow into the network consists of a long term average rate η_i (occasionally called just *rate*), a decay factor α_i, and a constant A_i. Consult [11] for an introduction to the concepts of EBB processes and EB processes, and for the notations we use. Especially, note that if the session i traffic flow into the network is indeed characterized as EBB with the three parameters η_i, α_i, and A_i, then we say that it is (η_i, A_i, α_i)-EBB. Recall that we are interested in user traffic flows that satisfy the throughput condition, i.e.,

$$\sum_{i \in I(v)} \eta_i < C^v$$

for every node v, where C^v is its service rate. Consequently, we can choose a small enough positive ϵ such that

$$\sum_{i \in I(v)} (\eta_i + \epsilon) < C^v \tag{1}$$

holds for every node v. We associate each session i with a virtual rate $\rho_i = \eta_i + \epsilon$, and obtain a new collection of rates, which, by (1), also satisfy the throughput condition. For clarity of exposition we will use only the set of ρs, and write $\rho_i - \epsilon$ wherever η_i is to be expected. Therefore, we use the following EBB characterization for the amount of session i traffic entering the network during any time interval $[s, t]$,

$$\Pr\left\{ \int_s^t R_i(u)\,\mathrm{d}u \geq (\rho_i - \epsilon)(t - s) + \sigma \right\} \leq A_i \mathrm{e}^{-\alpha_i \sigma} \tag{2}$$

for all $\sigma \geq 0$. The additional ϵ will be needed to control the effect of the bursts of $R_i(t)$, which are unbounded in nature, and can be arbitrarily large.

2.3. *GPS Multiplexing and CRST Assignments*

Each of the switches serves the traffic destined to any of its outgoing links according to a predetermined service discipline. In what follows we assume that every switch of the network is either a GPS multiplexer with one outgoing link, or a demultiplexer. Combinations of switches of these two types allow modeling of various server architectures, as illustrated by the server architecture of [7, p. 97], which may be considered a variant of a two-input two-output server with output-queueing. However, in the analysis we herein present, only the GPS switches are considered. This is due to the fact that a demultiplexer does not affect the instantaneous traffic rates of the sessions that traverse it. It merely directs the traffic flow of each session to the next switch in its path. We will also relate the results to the case when the multiplexers employ the packet-based version of the GPS discipline, the PGPS.

A GPS server that serves N sessions on an outgoing link of capacity C is specified by N positive real numbers $\phi_1, \phi_2, \ldots, \phi_N$. It is defined (see [8]) by the following two rules:

1. It is work conserving.
2. If session i is backloged in the server during the whole interval $[s, t]$ then

$$\frac{R_{i,\text{out}}^{s,t}}{R_{j,\text{out}}^{s,t}} \geq \frac{\phi_i}{\phi_j} \tag{3}$$

   for all $j = 1, 2, \ldots, N$.

Note that whenever session i is backloged, it is guaranteed a service rate of at least

$$g_i = \frac{\phi_i}{\sum_{j=1}^{N} \phi_j} C \tag{4}$$

Remark. When we refer to a particular switch v in the network we add its index as a superscript. Thus, C^v is the capacity of the outgoing link of v. Similarly, ϕ_i^v is the service parameter of session i at v, g_i^v is the service rate session i is guaranteed at v.

Following [7] we say that session i *impedes* session j at a switch v if

$$\frac{\phi_i^v}{\phi_j^v} > \frac{\rho_i}{\rho_j}$$

When no confusion may result, we drop the superscript v and say that i impedes j.

One evident disadvantage of the GPS discipline is that it serves multiple sessions simultaneously. An attempt to overcome this drawback is the PGPS, which approximates GPS in the more practical case, when the server can only handle one packet at a time. A PGPS server considers a packet as *arrived* after its last bit has arrived, and it uses a simulation of GPS to determine the order at which arrived packets receive service. Whenever a PGPS server becomes free, it selects for service the arrived packet that would be the first to complete service in the GPS simulation, if no additional packet were to arrive. The extent of this approximation is considered in the following restatement of Theorem 1.2 of [7].

Theorem 1. *Consider a GPS server and a PGPS server that are loaded with the same user traffic. Assume the input traffic of each session is a packet arrival process, modeled by a series of impulses, and that the size of each packet is at most L. Let $R_{i,\text{out}}^{0,t}$, $\widehat{R}_{i,\text{out}}^{0,t}$ be the amount of session i traffic served in the two servers, respectively, from 0 to t. Then for all times t*

$$R_{i,\text{out}}^{0,t} - \widehat{R}_{i,\text{out}}^{0,t} \leq L$$

An assignment of GPS (or PGPS) service parameters is called *Consistent Relative Session Treatment (CRST)* if there exists an ordering π of the sessions such that if $\pi^{-1}(i) < \pi^{-1}(j)$ (i.e., i precedes j in π) then session j does not impede session i at any of the nodes of the network. In this notation an ordering π of a set S of sessions is a bijection π: $\{1, 2, \ldots, |S|\} \mapsto S$. As discussed in [7], the sessions of a network with a CRST assignment can be partitioned into non-empty classes $H_1, H_2, \ldots, H_M$, such that every session, i, in H_m does not impede any of the sessions in $H_1, H_2, \ldots, H_m$.

3. Isolated Servers

3.1. *Preliminaries*

Consider an isolated server, either GPS or PGPS, that serves N sessions on an outgoing link of capacity C, and assume its service parameters are $\phi_1, \phi_2, \ldots, \phi_N$. We begin with two definitions, the first of which is due to [7].

Definition 1. *An ordering π is* a feasible ordering *of the sessions if it satisfies*

$$\rho_{\pi(k)} < \frac{\left(C - \sum_{i=1}^{k-1} \rho_{\pi(i)}\right)\phi_{\pi(k)}}{\sum_{i=1}^{N} \phi_i - \sum_{i=1}^{k-1} \phi_{\pi(i)}}$$

for all $k = 1, 2, \ldots, N$.

Definition 2. *An ordering π is* a consistent ordering *of the sessions if it satisfies $\pi^{-1}(i) < \pi^{-1}(j)$ for any two sessions i and j such that i impedes j.*

Thus an ordering is consistent if i precedes j in it whenever i impedes j.

Notice that both the concepts of feasibility and consistency are related to a specific server, and depend only on the service parameters of this server, and on the rates of the sessions it serves. When we wish to examine an ordering of the sessions of *a network*, we can only relate to the ordering it *induces* on the sessions $I(v)$ that are served by a specific server v of the network. We can then determine whether this induced ordering is a feasible ordering or a consistent ordering for v. If so, we will say that the induced ordering is a feasible ordering or a consistent ordering of the sessions $I(v)$, respectively.

Proposition 1. *Let π be a consistent ordering. Then π is a feasible ordering.*

Proof. We prove by induction that for every $k = 1, 2, \ldots, N$ session $\pi(k)$ can be chosen as the kth session in a feasible ordering that starts with $\pi(1), \pi(2), \ldots, \pi(k-1)$. The basis step and the inductive step are similar, and will be performed together. Assume the claim is true for $k - 1$ and prove it for k (k equals 1 for the basis step). By the proof to Lemma 2.8 of [7] any partial feasible ordering can be completed into a full feasible ordering. Consequently, there exists at least one new session j such that

$$\rho_j < \frac{\left(C - \sum_{i=1}^{k-1} \rho_{\pi(i)}\right)\phi_j}{\sum_{i=1}^{N} \phi_i - \sum_{i=1}^{k-1} \phi_{\pi(i)}}$$

Since π is a consistent ordering, session j does not impede session $\pi(k)$ and hence

$$\frac{\phi_{\pi(k)}}{\phi_j} \geq \frac{\rho_{\pi(k)}}{\rho_j}$$

Therefore

$$\rho_{\pi(k)} \leq \frac{\phi_{\pi(k)}}{\phi_j}\rho_j < \frac{\left(C - \sum_{i=1}^{k-1} \rho_{\pi(i)}\right)\phi_{\pi(k)}}{\sum_{i=1}^{N} \phi_i - \sum_{i=1}^{k-1} \phi_{\pi(i)}}$$

and the claim is proved. □

A fundamental lemma, which is a simple modification of Lemma 2.11 of [7] is the following.

Lemma 1. *Let π be a feasible ordering of the sessions of a GPS server, and suppose that a session $\pi(k)$ is backloged during the interval $[s, t]$. Assume further that*

$$R^{s,t}_{\pi(k),\text{out}} < \rho_{\pi(k)}(t - s) - \sigma$$

for some positive σ. Then

$$\sum_{i=1}^{k-1} R^{s,t}_{\pi(i),\text{out}} > \sum_{i=1}^{k-1}(t - s)\rho_{\pi(i)} + \sigma$$

3.2. *An Isolated GPS Server*

Theorem 2. *Let π be a feasible ordering, and consider a session $i = \pi(k)$, $1 \leq k \leq N$. If the input traffic flows $R_{\pi(l)}$ to a GPS server are*

$$(\rho_{\pi(l)} - \epsilon, A_{\pi(l)}, \alpha_{\pi(l)})\text{-EBB}$$

for all $l \leq k$ then the output traffic $R_{i,\text{out}}$ is also EBB with rate $\rho_i - \epsilon$.

Notice the significant advantage of this result over the one presented in [11] for a general server. In the GPS case, for the session i traffic that leaves the server to be EBB, it is only required that the entering traffic of session i and of the sessions that precede i be EBB. In the general server case, on the other hand, it was required that *all* the input traffics be EBB.

Proof. By induction over k. Without loss of generality we may assume that π is the identity, and consider session 1 first. Moreover, since this session is first in a feasible ordering, we have $\rho_1 < g_1$.

Fix some positive σ_0 and choose a small enough δ such that $0 < \delta < \sigma_0/\rho_1$ and

$$\frac{A_1 e^{\alpha_1(\rho_1 - \epsilon)\delta}}{1 - e^{-\alpha_1 \epsilon \delta}} \geq e^{\alpha_1 \sigma_0} \tag{5}$$

Notice that such a choice is always possible since the left hand side of (5) tends to infinity when δ approaches 0. Let $d(s)$ be the r.v. defined by

$$d(s) = \min \{u: \; Q_1(s-u) = 0\}$$

where $Q_1(t)$ is the backlog of session 1 traffic at time t. Thus $d(s)$ is the time that has passed since the last time prior to s when session 1 was not backloged. Two direct consequences of this definition are the following.

$$R_{1,\text{out}}^{s-d(s),s} \geq g_1 d(s) > \rho_1 d(s) \tag{6}$$

$$R_{1,\text{out}}^{s-d(s),t} \leq R_1^{s-d(s),t} \tag{7}$$

Combining (6) and (7) we get

$$R_1^{s-d(s),t} \geq R_{1,\text{out}}^{s-d(s),t} = R_{1,\text{out}}^{s-d(s),s} + R_{1,\text{out}}^{s,t} \geq \rho_1 d(s) + R_{1,\text{out}}^{s,t} \tag{8}$$

Suppose that $\sigma \geq \sigma_0$. We have

$$\begin{aligned} &\Pr\left\{R_{1,\text{out}}^{s,t} \geq (\rho_1 - \epsilon)(t-s) + \sigma\right\} \\ &\quad = \sum_{m=0}^{\infty} \Pr\Big(\left\{R_{1,\text{out}}^{s,t} \geq (\rho_1 - \epsilon)(t-s) + \sigma\right\} \cap \{m\delta \leq d(s) < (m+1)\delta\}\Big) \end{aligned} \tag{9}$$

and by (8) we get

$$\begin{aligned} &\left\{R_{1,\text{out}}^{s,t} \geq (\rho_1 - \epsilon)(t-s) + \sigma\right\} \cap \{m\delta \leq d(s) < (m+1)\delta\} \\ &\quad \subset \left\{R_1^{s-d(s),t} \geq (\rho_1 - \epsilon)(t-s) + \sigma + \rho_1 d(s)\right\} \cap \{m\delta \leq d(s) < (m+1)\delta\} \\ &\quad \subset \left\{R_1^{s-(m+1)\delta,t} \geq (\rho_1 - \epsilon)(t-s) + \sigma + \rho_1 m\delta\right\} \\ &\quad = \left\{R_1^{s-(m+1)\delta,t} \geq (\rho_1 - \epsilon)(t-s+(m+1)\delta) + \epsilon(m+1)\delta + \sigma - \rho_1\delta\right\} \end{aligned} \tag{10}$$

Recall that $\sigma \geq \sigma_0$ and $\rho_1\delta < \sigma_0$. Therefore, the event (10) is of the form $\{R_1^{t_1,t_2} \geq (\rho_1 - \epsilon)(t_2 - t_1) + \sigma_1\}$ where $\sigma_1 > 0$, and we can use the assumption that R_1 is EBB. Substituting into (9) we get

$$\begin{aligned} &\Pr\left\{R_{1,\text{out}}^{s,t} \geq (\rho_1 - \epsilon)(t-s) + \sigma\right\} \\ &\quad \leq \sum_{m=0}^{\infty} A_1 \mathrm{e}^{-\alpha_1(\epsilon(m+1)\delta + \sigma - \rho_1\delta)} = \frac{A_1 \mathrm{e}^{\alpha_1(\rho_1-\epsilon)\delta}}{1 - \mathrm{e}^{-\alpha_1\epsilon\delta}} \mathrm{e}^{-\alpha_1\sigma} \end{aligned} \tag{11}$$

It is here that we employ the additional ϵ we provided in ρ_1, which compensates for the unbounded nature of $d(s)$. The same technique will be used again to complete the induction step.

This concludes the proof that $R_{1,\text{out}}$ is EBB with a constant

$$A_{1,\text{out}} = \frac{A_1 e^{\alpha_1(\rho_1-\epsilon)\delta}}{1-e^{-\alpha_1\epsilon\delta}}$$

since (11) also holds for $\sigma < \sigma_0$ by the choice (5). Notice that the decay factor we get for session 1 is $\alpha_{1,\text{out}} = \alpha_1$.

Consider next a session i, where $i > 1$. By the assumptions of the theorem the input traffic flow R_j is $(\rho_j - \epsilon, A_j, \alpha_j)$-EBB for all $j \le i$. By the induction hypothesis the output traffic flow $R_{j,\text{out}}$ is $(\rho_j - \epsilon, A_{j,\text{out}}, \alpha_{j,\text{out}})$-EBB for all $j < i$.

Let $\{p_j\}_{j\le i}$ be a set of positive numbers that sum to one, such that $\alpha_{j,\text{out}} p_j = \alpha_i p_i$ for all $j < i$, and choose $\alpha_{i,\text{out}}$ to be

$$\alpha_{i,\text{out}} = \alpha_i p_i = \frac{1}{\frac{1}{\alpha_i} + \sum_{j<i} \frac{1}{\alpha_{j,\text{out}}}}$$

Fix some positive σ_0 and let δ be small enough, such that $0 < \delta < p_j\sigma_0/\rho_j$ for all $j \le i$ and

$$A_{i,\text{out}} = \frac{A_i e^{\alpha_i(\rho_i-\epsilon)\delta}}{1-e^{-\alpha_i\epsilon\delta}} + \sum_{j<i} \frac{A_{j,\text{out}} e^{\alpha_{j,\text{out}}(\rho_j-\epsilon)\delta}}{1-e^{-\alpha_{j,\text{out}}\epsilon\delta}} \ge e^{\alpha_{i,\text{out}}\sigma_0}$$

Such a choice of δ is always possible for a reason similar to the one mentioned following (5). Notice that the last requirement defines the choice of $A_{i,\text{out}}$, and establishes the theorem for $\sigma < \sigma_0$. Therefore we only have to consider the case $\sigma \ge \sigma_0$. We have

$$\begin{aligned} &\Pr\left\{R_{i,\text{out}}^{s,t} \ge (\rho_i - \epsilon)(t-s) + \sigma\right\} \\ &= \sum_{m=0}^{\infty} \Pr\Big(\left\{R_{i,\text{out}}^{s,t} \ge (\rho_i - \epsilon)(t-s) + \sigma\right\} \cap \{m\delta \le d(s) < (m+1)\delta\}\Big) \end{aligned} \tag{12}$$

where $d(s)$ is defined, similarly to before, by $d(s) = \min\{u\colon\ Q_i(s-u) = 0\}$. To upper bound each of the summands of (12), it is sufficient to intersect each event in this sum with $E = \{R_{i,\text{out}}^{s-d(s),s} \ge \rho_i d(s) - (1-p_i)\sigma\}$ and with its complement E^c, and to add the probabilities of the two resulting events. Since we assumed $1, 2, \ldots, N$ is a feasible ordering, we can handle E^c by means of Lemma 1.

$$\left\{R_{i,\text{out}}^{s-d(s),s} < \rho_i d(s) - (1-p_i)\sigma\right\} \subset \left\{\sum_{j<i} R_{j,\text{out}}^{s-d(s),s} \ge \sum_{j<i} \rho_j d(s) + (1-p_i)\sigma\right\} \tag{13}$$

Applying (13), and using relations similar to (6) and (7) for session i, we get

$$\begin{aligned} &\left\{R_{i,\text{out}}^{s,t} \ge (\rho_i - \epsilon)(t-s) + \sigma\right\} \cap \{m\delta \le d(s) < (m+1)\delta\} \\ &\subset \left\{R_i^{s-(m+1)\delta,t} \ge (\rho_i - \epsilon)(t - s + (m+1)\delta) + \epsilon(m+1)\delta + p_i\sigma - \rho_i\delta\right\} && (14) \\ &\cup \bigcup_{j<i} \left\{R_{j,\text{out}}^{s-(m+1)\delta,s} \ge (\rho_j - \epsilon)(m+1)\delta + \epsilon(m+1)\delta + p_j\sigma - \rho_j\delta\right\} && (15) \end{aligned}$$

Now we can apply the assumptions of the theorem to (14), and the induction hypothesis to each of the events in (15). Substituting into (12), we can write

$$\Pr\left\{R_{i,\text{out}}^{s,t} \geq (\rho_i - \epsilon)(t-s) + \sigma\right\}$$
$$\leq \sum_{m=0}^{\infty} \left(A_i e^{\alpha_i(\rho_i-\epsilon)\delta} e^{-\alpha_i \epsilon m \delta} e^{-\alpha_i p_i \sigma} + \sum_{j<i} A_{j,\text{out}} e^{\alpha_{j,\text{out}}(\rho_j-\epsilon)\delta} e^{-\alpha_{j,\text{out}}\epsilon m\delta} e^{-\alpha_{j,\text{out}} p_j \sigma} \right)$$
$$\leq \left(\frac{A_i e^{\alpha_i(\rho_i-\epsilon)\delta}}{1-e^{-\alpha_i\epsilon\delta}} + \sum_{j<i} \frac{A_{j,\text{out}} e^{\alpha_{j,\text{out}}(\rho_j-\epsilon)\delta}}{1-e^{-\alpha_{j,\text{out}}\epsilon\delta}} \right) e^{-\alpha_{i,\text{out}}\sigma}$$

where the last inequality is due to the choice of the parameters p_j. □

Note that a similar proof will show that when the assumptions of Theorem 2 hold, Q_i is $(A_{i,\text{out}}, \alpha_{i,\text{out}})$-EB.

3.3. *An Isolated PGPS Server*

Our goal in this section is to extend Theorem 2 to the case of an isolated PGPS server. We assume the traffic flow each session i introduces to the server is divided into packets, and that the maximal packet size is L.

Besides the fact that the PGPS service discipline is different from GPS, we recall that, unlike the GPS server, a PGPS server does not start serving a packet if it is only partially received. In other words, the PGPS is a Store and Forward technique, while GPS is Cut Through. To cope with this difficulty, we separate the PGPS server into two imaginary parts — a regulator and a PGPS core (see also the discussion in Chapter 4 of [7]). The function of the regulator is to perform the Store and Forward part. It collects the received data into packets, and passes only complete packets to the PGPS core. The output of this regulator, which is the input to the PGPS core, is a series of impulses, whose heights represent the sizes of the packets. Let R_i be the session i input traffic to the PGPS server, which is also the input to the regulator part, $\widehat{R}_i$ is the output traffic from the regulator, which is the input traffic to the PGPS core. We have

$$\widehat{R}_i^{s,t} = l_i(s) + R_i^{s,t} - l_i(t)$$

where $l_i(u)$ is the amount of data belonging to a session i packet that is partially received in the regulator by time u. Since $0 \leq l(u) \leq L$, we conclude that

$$\widehat{R}_i^{s,t} \leq R_i^{s,t} + L \tag{16}$$

Considering a PGPS server with EBB inputs, we can employ (16) and get

$$\begin{aligned} &\Pr\left\{\widehat{R}_i^{s,t} \geq (\rho_i - \epsilon)(t-s) + \sigma\right\} \\ &\leq \Pr\left\{R_i^{s,t} \geq (\rho_i - \epsilon)(t-s) + \sigma - L\right\} \\ &\leq \left(A_i e^{\alpha_i L}\right) e^{-\alpha_i \sigma} \end{aligned} \tag{17}$$

which holds for all values of $\sigma > 0$. Notice that the application of (16) is valid only for $\sigma \geq L$, but in case $\sigma < L$ the right hand side of (17) is at least 1, since the constant of any EBB characterization is at least 1, and therefore $A_i \geq 1$ for all i. Thus the inputs to the PGPS core are also EBB, with the same rates and decay factors, and with the constants changed from A_i to $A_i e^{\alpha_i L}$.

Since we wish to apply Theorem 2, we resort to Theorem 1. We denote the session i output traffic from the PGPS core by $\widehat{R}_{i,\text{out}}$, and consider a GPS server that replaces the PGPS core, with output traffic denoted by $R_{i,\text{out}}$.

Consider some two time instances $t > s$. By Theorem 1 we have

$$\sum_{j \neq i} R_{j,\text{out}}^{0,t} - \sum_{j \neq i} \widehat{R}_{j,\text{out}}^{0,t} \leq (N-1)L \tag{18}$$

Since both the GPS and the PGPS service disciplines are work-conserving, we have

$$\sum_{j=1}^{N} R_{j,\text{out}}^{0,t} = \sum_{j=1}^{N} \widehat{R}_{j,\text{out}}^{0,t} \tag{19}$$

Subtracting (18) from (19) we get

$$R_{i,\text{out}}^{0,t} - \widehat{R}_{i,\text{out}}^{0,t} \geq -(N-1)L$$

and thus, by applying Theorem 1 again

$$R_{i,\text{out}}^{s,t} - \widehat{R}_{i,\text{out}}^{s,t} = \left(R_{i,\text{out}}^{0,t} - \widehat{R}_{i,\text{out}}^{0,t}\right) - \left(R_{i,\text{out}}^{0,s} - \widehat{R}_{i,\text{out}}^{0,s}\right) \geq -(N-1)L - L = -NL \tag{20}$$

We are now in a position to bound the distribution of $\widehat{R}_{i,\text{out}}^{s,t}$, since by (20)

$$\begin{aligned}
&\Pr\left\{\widehat{R}_{i,\text{out}}^{s,t} \geq (\rho_i - \epsilon)(t-s) + \sigma\right\} \\
&\leq \Pr\left\{R_{i,\text{out}}^{s,t} \geq (\rho_i - \epsilon)(t-s) + \sigma - NL\right\}
\end{aligned}$$

Moreover, we already know by (17) that the input traffic flows to the PGPS core are EBB. Therefore we can apply Theorem 2, and conclude its peer for the PGPS case.

Theorem 3. *Let π be a feasible ordering, and consider a session $i = \pi(k)$, $1 \leq k \leq N$. If the input traffic flows $R_{\pi(l)}$ to a PGPS server are*

$(\rho_{\pi(l)} - \epsilon, A_{\pi(l)}, \alpha_{\pi(l)})$*-EBB*

for all $l \leq k$ then the output traffic $R_{i,\text{out}}$ is also EBB with rate $\rho_i - \epsilon$.

The appropriate constants $\alpha_{i,\text{out}}$ and $A_{i,\text{out}}$ can be determined by Theorem 2, taking (17) and (20) into account.

4. The Stability of CRST Networks

We are now ready to examine the traffic flows of the various sessions within a network with a CRST assignment, where each of the nodes is either a GPS or a PGPS server (which we jointly call GPS type servers). Our goal is to show that the traffic flow of each session, in each of the links it traverses, is EBB. This fact will guarantee that the backlog of each session in each of the nodes it visits is EB, and the network is stable. As mentioned earlier, we partition the sessions of the network into non-empty disjoint classes $H_1, H_2, \ldots, H_M$.

Definition 3. *An ordering* π *is* a staggered ordering *of the sessions of the network if* i *precedes* j *whenever* $i \in H_l$, $j \in H_k$, *and* $l < k$.

Thus an ordering is staggered if it is such that all the sessions of H_1 come first, all the sessions of H_2 come next, and so forth. Notice that a staggered ordering induces a consistent ordering over the set $I(v)$ of sessions that are served by v, for every switch v.

Theorem 4. *A network of GPS type servers with CRST assignment, which is loaded with N EBB sessions, is stable whenever the throughput condition holds.*

Proof. To show that the traffic flow of each session, in each of the links it traverses, is EBB, we use a double induction. The first is on the class index $1 \leq m \leq M$, and the second, which considers a specific session i, $i \in H_m$, is on the path index $1 \leq k \leq K_i$.

By the first induction hypothesis we assume that the traffic flow of each session i, $i \in H_1 \cup H_2 \cup \ldots \cup H_m$, in each of the links it traverses, is EBB. Notice that for $m = 0$ this induction hypothesis is trivial, hence we only have to establish the inductive step.

Consider a specific session i, $i \in H_{m+1}$, and let $v = P(i,k)$. By the second induction hypothesis we assume the session i traffic that enters v is EBB. Again, the case $k = 1$ is due to the EBB nature of the session i input traffic to the network, and we only have to establish the inductive step.

Let π be a staggered ordering in which $i \in H_{m+1}$ is first among the sessions of H_{m+1}. The ordering π induces on $I(v)$ is a consistent ordering, in which session i is only preceded by sessions of classes with indices m or less. Moreover, by Proposition 1 it is also a feasible ordering. Considering this induced ordering, we can apply the appropriate theorem (either Theorem 2 if v is a GPS server, or Theorem 3 if it is a PGPS server), and deduce that the session i traffic leaving $P(i,k)$ is EBB. Since we are dealing with virtual circuit routing, this EBB traffic is the session i traffic that enters $P(i, k + 1)$, and the proof is complete. □

5. Conclusions

We demonstrated the power of the characterization presented in [11] by analyzing the GPS type systems presented in [7]. The network model considered employs the virtual circuit packet switching technique and it consists of two kinds of processor sharing type servers, the GPS and the PGPS. We first analyzed an isolated server, and showed that when it is loaded with EBB user traffic flows, its output traffic streams are also EBB. Moreover, we emphasized the improvement of this analysis over the one presented in [11] for the general work conserving service discipline. We then considered the network model, assuming each of the various nodes of the network employs either the GPS or the PGPS service disciplines. We showed that if the network is loaded with EBB user traffic, then it is stable whenever the throughput condition holds. The exponential bounds to the distributions of the backlogs and the output streams, derived for the isolated server case, can be easily employed to deduce similar bounds to both the backlogs of each session in each node, and the flows of each session in each link. Referring to the applicability of the EBB characterization, established in [11], we can conclude that from the queueing-theory perspective the above network model is stable when it is loaded with common stochastic processes, such as Poisson, Bernoulli, or AMS.

References

[1] R. L. Cruz, *A Calculus for Network Delay and a Note on Topologies of Interconnection Networks*, Ph.D. Thesis, University of Illinois, Urbana-Champaign, (1987).

[2] R. L. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation", *IEEE Trans. Inform. Th.* **37** (1), (Jan. 1991), 114–131.

[3] R. L. Cruz, "A Calculus for Network Delay, Part II: Network Analysis", *IEEE Trans. Inform. Th.* **37** (1), (Jan. 1991), 132–141.

[4] A. Demers, S. Keshav and S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm", pp. 1–12 in *Proc. SIGCOMM '89*, 1989.

[5] A. G. Greenberg and N. Madras, "How Fair is Fair Queueing?", *J. ACM* **39** (3), (Jul. 1992), 568–598.

[6] F. P. Kelly, *Reversibility and Stochastic Networks*, John Wiley & Sons, 1979.

[7] 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, (1992).

[8] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case", *IEEE/ACM Trans. Networking* **1** (3), (Jun. 1993), 344–357.

[9] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case", *IEEE/ACM Trans. Networking* **2** (2), (Apr. 1994), 137–150.

[10] J. Turner, "New Directions in Communications (or Which Way to the Information Age?)", *IEEE Comm. Mag.* **24** (10), (Oct. 1986), 8–15.

[11] O. Yaron and M. Sidi, "Performance and Stability of Communication Networks via Robust Exponential Bounds", *IEEE/ACM Trans. Networking* **1** (3), (Jun. 1993), 372–385.

[12] Z. Zhang, D. Towsley and J. Kurose, "Statistical Analysis of Generalized Processor Sharing Scheduling Discipline", submitted to *SIGCOMM '94*, Jan. 1994.