DPS queues with stationary ergodic service times and the performance of TCP in overload

Eitan Altman*, Tania Jiménez†, and Daniel Kofman‡
*INRIA Sophia Antipolis, 06902 Sophia Antipolis, France
altman@sophia.inria.fr
† CESIMO, Facultad de Ingeniería, Univ. de Los Andes, Mérida, Venezuela
tania@ula.ve
‡ Network Department, E.N.S.T., 46 rue Barrault, 75634 Paris Cedex, France
kofman@res.enst.fr

Abstract— **In a recent paper, Bonald and Roberts [1] studied non-persistent TCP connections in transient overload conditions, under the assumption that all connections have the same round-trip times. In this paper our goal is to develop theoretical tools that will enable us to relax this assumption and obtain explicit expressions for the rate of growth of the number of connections at the system, the rate at which TCP connections leave the system, as well as the time needed for the completion of a connection. To that end, we model the system as a DPS (Discriminatory Processor Sharing) system which we analyze under very mild assumptions on the probability distributions related to different classes of arrivals: we only assume that the arrival rates of connections exist, and that the amount of information transmitted during a connection of a given type forms a stationary ergodic sequence. We then proceed to obtain explicit expressions for the growth rate of the number of connections at the DPS system for several specific probability distributions. We check through simulations the applicability of our queueing results for modeling TCP connections sharing a bottleneck.**

Methods Keywords: Stochastic processes/Queueing theory.

I. INTRODUCTION

In this paper we provide a new queueing analysis of DPS (Discriminate Processor Sharing) queueing model (see e.g. [8]) with stationary ergodic service times (possibly correlated). We apply this model to the performance evaluation of multiplexing of heterogeneous TCP connections (heterogeneity is taken with respect to both round trip delay as well as with respect to the probability distribution of the amount of data to be transmitted by a connection). We focus on the overload regime of the queue in which the amount of workload that arrives exceeds the system capacity. Our mathematical analysis is based on techniques from reference [7], who presented a fluid limit for a single queue operating under the processor sharing regime. A remarkable property of these fluid limits is that the throughput, and the rate of growth of the queue size was shown to depend not only on expectation of interarrival and service times, but on the whole distribution of the service times. In addition to the extension of the model of [7] to the case of several classes (and to the DPS regime), we consider in this paper a more general arrival processes that need not be stationary nor ergodic[1]. This allows us to extend our results to some networks of DPS queues. For some distributions (exponential and hyper-exponential) of the service times, we are able to obtain exact expressions that describe the system's performance.

We then apply our mathematical results to study the performance of non-persistent TCP connections at transient overload periods. We use the well known fact that at a session level, TCP can be analyzed using a processor sharing approach [1], [4]. We consider K types of TCP sessions where each type is characterized by its average round trip time and the distribution of the amount of information it has to transfer. Within each type, there may be very short connections ("mice") i.e. connections with small amount of information to transfer, as well as long connections ("elephants"). We focus on the case that the sessions share a bottleneck link. We then extend the result to some more general networks.

As already shown in [1] for the case of equal round trip times (RTTs) of connections that share a single

[1]Our framework is related to that of [2] that studied stability of non-stationary systems; in our case we obtain "stability" properties of the "instability" of an overloaded system

bottleneck node, during transient overload periods there is a *linear* growth in the number of ongoing TCP sessions and their waiting times in the system, and the average throughput of a session decreases to zero. Yet the rate at which sessions of a given class leave the system converges to a constant which our DPS model can predict.

Our model shows that even in overload periods, eventually all sessions manage to complete (including ”elephants”). Other surprising feature of the analysis is that the departure rate of sessions need not be monotone increasing in the arrival rates[2]. Yet we show that the number of ongoing sessions is monotone in the arrival rates.

The structure of the paper is as follows. We begin in the next section by presenting the discriminatory processor sharing regime and analyze its performance under overload condition. We then discuss the stability conditions in Section III for the case of a single common bottleneck queue. We extend the analysis to the case of several overloaded links (with possibly different routes for different TCP sessions) in Section IV. Explicit expressions are obtained in Section V for the system's performance for some special marginal distributions of the service times. In Section VI we apply our model to the analysis of TCP connections with different round trip time sharing a common bottleneck node; using simulations performed with ns simulator [9], we show that the DPS model is well adapted to the way TCP connections share the bandwidth at overload.

II. Discriminatory Processor Sharing: model and analysis

We analyze a single server queue, fed by K independent arrival processes under the discriminatory processor sharing discipline. This discipline is defined as follows. A coefficient g_i is associated to the arrival process i, $i = 1 \dots K$. We denote L_t^i the number of customers of class i in the queue at time t. Under the DPS discipline each class i customer is served with a rate

$$\frac{g_i}{\sum_{j=1}^{K} g_j L_t^j}.$$

The total server rate is assumed equal to 1.

We introduce the following assumptions and notation. All the processes introduced below are assumed to be defined on a common probability space, on which a probability measure P is defined. Let $N^i[a,b]$ be the number of class i arrivals in the interval $[a,b]$. We assume that the limits

$$\lambda_i := \lim_{t\to\infty} \frac{N^i[0,t]}{t} \tag{1}$$

exist $P-a.s.$ Let t_n^i denote the instant at which the nth customer of class i arrives. Let σ_n^i denote the amount of service required by the nth customer of class i arriving to the system. (In the case of TCP connection, this corresponds to the size of the nth file divided by the bottleneck link throughput.) We shall assume that for each class i, the service times σ_n^i are stationary ergodic under the shift in n. (This setting, in which the inter-arrival times are not required to be stationary ergodic, is a generalization of the setting in [7] even for one class of customers. It will become necessary for us to consider this setting, as we shall consider networks where the arrival into one queue are related to the output of another queue. The inter-departure times from the first queue are, in general, nonstationary.) Denote

- $\hat{P}^i$ the Palm probability[3] associated to the process N^i
- $\hat{E}^i$ the expectation with respect to this measure.

The following hypothesis implies the transient behaviour of the queue under study :

$$\sum_{i=1}^{K} \lambda_i \hat{E}^i(\sigma_0^i) > 1 \tag{2}$$

We will use the following lemma which is an adaptation of the one in [7] to our setting (with weaker assumptions):

Lemma 1: Let $N[0,t]$ be a point process. Denote

$$\overline{\lambda} \doteq \limsup_{t\to\infty} \frac{1}{t} N[0,t], \qquad \underline{\lambda} \doteq \liminf_{t\to\infty} \frac{1}{t} N[0,t],$$

and suppose that $\overline{\lambda} < \infty$ and $\underline{\lambda} > 0$. To each point t_n of the point process we associate a mark σ_n. If the sequence $\{\sigma_n\}$ is stationary and ergodic (under the shift in n), for any measurable functions $f, g : \mathbb{R}_+ \to [0,1]$ we have, almost surely

$$i) \limsup_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq f(\sigma_n)t\}} \leq \overline{\lambda} E((fg)(\sigma_0)),$$

$$ii) \liminf_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq f(\sigma_n)t\}} \geq \underline{\lambda} E((fg)(\sigma_0)).$$

[2]Examples have been obtained in [6] for the single class case

[3]for a stationary ergodic marked point process N^i, the term ”Palm probability” is the probability describing the process embeded at times T_n^i; we use this term with some abuse of terminology as N^i need not be stationary. It will however only be used for the service times σ_n^i, which are indeed assumed to be stationarity ergodic.

In particular, if $\underline{\lambda} = \overline{\lambda} = \lambda$ then

$$\lim_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq f(\sigma_n)t\}} = \lambda E((fg)(\sigma_0)).$$

Proof: We start showing that i) holds when $A = f(\mathbb{R}_+)$ if finite. In this case

$$\begin{aligned}
&\limsup_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq f(\sigma_n)t\}} \\
&\leq \sum_{a\in A} \limsup_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq at\}} \mathbb{1}_{\{f(\sigma_n)=a\}} \\
&\leq \sum_{a\in A} \limsup_{t\to\infty} \frac{1}{t} \sum_{n\geq 0}^{N[0,at]} g(\sigma_n) \mathbb{1}_{\{f(\sigma_n)=a\}} \\
&\leq \sum_{a\in A} \limsup_{t\to\infty} \frac{N[0,at]}{at} \limsup_{t\to\infty} \\
&\quad \times \frac{1}{N[0,at]} \sum_{n=0}^{N[0,at]} g(\sigma_n) a \mathbb{1}_{\{f(\sigma_n)=a\}} \\
&\leq \sum_{a\in A} \overline{\lambda} E[g(\sigma_0) a \mathbb{1}_{\{f(\sigma_0)=a\}} = \overline{\lambda} E[gf(\sigma_0)]
\end{aligned}$$

The last inequality is based in the fact that $\lim_{t\to\infty} N[0,at] = \infty$ since $\underline{\lambda} > 0$.

Now, let us define

$$f_m(x) = k/m \;\forall x \;\text{ such that } \lfloor mf(x) \rfloor = k-1 \quad (3)$$

Here $\lfloor y \rfloor$ is the greater integer smaller than y.
We have

$$\begin{aligned}
&\limsup_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq f(\sigma_n)t\}} \\
&\leq \limsup_{t\to\infty} \frac{1}{t} \sum_{n\geq 0} g(\sigma_n) \mathbb{1}_{\{t_n \leq f_m(\sigma_n)t\}} \\
&\leq \overline{\lambda} E((f_m g)(\sigma_0))
\end{aligned}$$

We can now obtain i) by using the Lebesgue theorem, that allows us to write

$$\lim_{m\to\infty} \overline{\lambda} E((f_m g)(\sigma_0)) = \overline{\lambda} E((fg)(\sigma_0)) \quad (4)$$

The proof of ii) can be done in the same manner. ∎

We will denote T_n^i the time of departure of the n-th class i customer (i.e. of the customer that arrived at time t_n^i).

The next proposition shows that a unique limit exists to the rates of growth of number of customers of different classes, and that this limit can be computed as a solution of a set of K equations. We have recently learnt that this result has been obtained in parallel in [5] in a Markovian setting, i.e. for the special case that for each i, the sequence $\{\sigma_n^i\}_n$ are i.i.d. and the arrivals are independent Poisson processes.

For any random variable X, we denote by

$$ess \inf(X) := \sup(a : P(X > a) = 1),$$

$$ess \sup(X) := \inf(a : P(X < a) = 1).$$

Proposition 2: Let L_t^i be the number of customers of class i in the queue at time t. Then

$$\lim_{t\to\infty} \frac{L_t^i}{t} = \alpha^i, \quad a.s. \quad (5)$$

where $\alpha = (\alpha^i)_{i=1}^K$ is the unique positive solution of

$$x^i = \lambda_i \left[1 - \hat{E}^i \exp\left(-g_i^{-1} \sigma_0^i \sum_{j=1}^K x^j g_j \right) \right], \quad (6)$$

$i = 1 \dots K$. Let W_n^i be the sojourn time of the n-th customer of class i, then

$$\lim_{n\to\infty} \frac{W_n^i}{n} - \frac{(e^{\alpha^i \sigma_n^i} - 1)}{\lambda_i} = 0 \quad (7)$$

in distribution.

Proof: Denote

$$\overline{\alpha}^i = ess \sup\left(\limsup_{t\to\infty} \frac{L_t^i}{t} \right),$$
$$\underline{\alpha}^i = ess \inf\left(\liminf_{t\to\infty} \frac{L_t^i}{t} \right).$$

Observe that $\overline{\alpha}^i \leq \lambda_i$ since the number of customers of class i in the buffer at time t is less than or equal to the total number of customers of class i that have arrived until t. Let $\overline{\alpha}$ and $\underline{\alpha}$ be the vectors whose ith components are $\overline{\alpha}^i$ and $\underline{\alpha}^i$, respectively.

The proof follows 4 steps.

Step 1 : **A bound for $\overline{\alpha}$.** Let $\beta = (\beta^i)_{i=1}^K$ be a real vector such that $\beta^i > \overline{\alpha}^i$ i=1…K. Then, there exist C_0 such that $L_t^i \leq \beta^i t \;\; \forall t > C_0,\; i = 1 \dots K$. We have

$$T_n^i \leq t_n^i e^{\gamma^i \sigma_n^i} \quad \forall\; t_n^i > C_0 \quad (8)$$

where $\gamma^i = g_i^{-1} \sum_{j=1}^K g_j \beta^j$. Indeed, because of the DPS service discipline,

$$\sigma_n^i = \int_{t_n^i}^{T_n^i} \frac{g_i}{\sum_{j=1}^K g_j L_u^j} du, \quad n \geq 0 \quad (9)$$

If $t > C_0$, then $L_t^i \leq \beta^i t$. Then we have

$$\sigma_n^i \geq \int_{t_n^i}^{T_n^i} \frac{g_i}{\sum_{j=1}^K g_j \beta^j u} du = \frac{g_i}{\sum_{j=1}^K g_j \beta^j} Log \frac{T_n^i}{t_n^i}$$

for $n \geq 0$. This implies (8). On the other hand, eq. (8) and the fact that

$$\sum_{0 \leq t_n^i < C_0} \mathbb{I}_{\{t_n^i \leq t < T_n^i\}} \leq N^i[0, C_0]$$

implies that

$$L_t^i = \sum_{n \geq 0} \mathbb{I}_{\{t_n^i \leq t < T_n^i\}} \leq N^i[0, C_0] + \sum_{t_n^i \geq C_0} \mathbb{I}_{\{t \exp(-\gamma^i \sigma_n^i) < t_n^i \leq t\}} \quad (10)$$

Using Lemma (1), and writing $\mathbb{I}_{\{t \exp(-\gamma^i \sigma_n^i) < t_n^i \leq t\}} = \mathbb{I}_{\{t_n^i \leq t\}} - \mathbb{I}_{\{t_n^i \leq t \exp(-\gamma^i \sigma_n^i)\}}$ we have

$$\limsup_{t \to \infty} \frac{L_t^i}{t} \leq \lambda_i (1 - \hat{E}^i (\exp(-\gamma^i \sigma_0^i))) \quad (11)$$

and letting β converging to $\overline{\alpha}$, we have

$$\limsup_{t \to \infty} \frac{L_t^i}{t} \leq \lambda_i (1 - \hat{E}^i (\exp\left(-\overline{\delta}^i \sigma_0^i\right))) \quad (12)$$

where $\overline{\delta}^i = g_i^{-1} \sum_{j=1}^K g_j \overline{\alpha}^j$.

Step 2 : **A bound for $\underline{\alpha}$.** Let $\beta = (\beta^i)_{i=1}^K$ be a real vector such that $\beta^i < \underline{\alpha}^i$ i=1…K. Then, there exist C_0 such that $L_t^i \geq \beta^i t \ \ \forall t > C_0, \ i = 1 \ldots K$. We have

$$T_n^i \geq t_n^i \exp\left(\gamma^i \sigma_n^i\right) \quad \forall \ t_n^i > C_0 \quad (13)$$

where $\gamma^i = g_i^{-1} \sum_{j=1}^K g_j \beta^j$. This can be shown using the same arguments as in Step 1. This implies that

$$L_t^i \geq \sum_{t_n^i \geq C_0} \mathbb{I}_{\{t \exp(-\gamma^i \sigma_n^i) < t_n^i \leq t\}} \quad (14)$$

and then it can be shown as in Step 1 (letting β converging to $\underline{\alpha}$) that

$$\liminf_{t \to \infty} \frac{L_t^i}{t} \geq \lambda_i (1 - \hat{E}^i (\exp\left(-\underline{\delta}^i \sigma_0^i\right))) \quad (15)$$

where $\underline{\delta}^i = g_i^{-1} \sum_{j=1}^K g_j \underline{\alpha}^j$. At this point we cannot yet conclude from (12) and (15) that L_t/t has a limit as $t \to \infty$ since we have not showed that the right hand side of (15) is larger than the right hand side of (12). We shall establish the existence of the limit only in step 4, by using properties derived in the step 3 below.

Step 3 : **Computing the unique positive solution.** We show that there exists a unique non null solution of the set of equations

$$x^i = \lambda_i \left[1 - \hat{E}^i \exp\left(-g_i^{-1} \sigma_0^i \sum_{j=1}^K x^j g_j \right) \right], \quad (16)$$

$i = 1 \ldots K$. Let $\tilde{x}^i = x^i g_i, \ i = 1 \ldots K$ and $\tilde{\lambda}_i = \lambda_i g_i, \ i = 1 \ldots K$. In terms of this transformation, we look for a solution of

$$\tilde{x}^i = \tilde{\lambda}_i \left[1 - \hat{E}^i \exp\left(-\frac{\sum_{j=1}^K \tilde{x}^j}{g_i} \sigma_0^i \right) \right], \quad (17)$$

$i = 1 \ldots K$. Let $x = \sum_{j=1}^K \tilde{x}^j$, the function

$$Z(x) := x - \sum_{j=1}^K \tilde{\lambda}_i \left[1 - \hat{E}^i \exp\left(-g_i^{-1} x \sigma_0^i\right) \right] \quad (18)$$

has a unique strictly positive solution. In fact, it is convex, converges toward $+\infty$ at infinity and, due to the inequality (2), has a strictly negative derivative at 0. Denote ν this solution (note that ν is the smallest solution of $Z(x) \geq 0, x > 0$, and the largest solution of $Z(x) \leq 0$) and define

$$\alpha^i = \lambda_i (1 - \hat{E}^i (\exp\left(-g_i^{-1} \nu \sigma_0^i\right))), \ i = 1 \ldots K.$$

Let $\alpha = (\alpha^i)_{i=1}^K$. It is easy to see that α is the unique strictly positive solution of (16). Indeed, once the sum of $\tilde{x}^j$ are determined uniquely as the positive solution of $Z(x) = 0$, the value of each $\tilde{x}^j$ is uniquely defined as seen from eq. (17).

Step 4 : **Relation between the solution of (16) and the increase rate of the number of sessions.**

Eq. (12) implies that

$$\sum_{i=1}^K g_j \overline{\alpha}^j \leq \sum_{i=1}^K g_j \lambda_j \left(1 - \hat{E}^i (\exp(-\overline{\delta}^i \sigma_0^i)) \right).$$

Then, due to the characteristics of the function (18), $\sum_{i=1}^K g_j \overline{\alpha}^j \leq \nu$.

Eq. (15) implies that

$$\sum_{i=1}^K g_j \underline{\alpha}^j \geq \sum_{i=1}^K g_j \lambda_j \left(1 - \hat{E}^i (\exp(-\underline{\delta}^i \sigma_0^i)) \right).$$

Then, in order to prove that $\sum_{i=1}^K g_j \underline{\alpha}^j \geq \nu$ it is enough to prove that $\sum_{i=1}^K g_j \underline{\alpha}^j \neq 0$ and this can be done in the same way as in [7].

Then
$$\sum_{i=1}^{K} g_j \underline{\alpha}^j = \sum_{i=1}^{K} g_j \overline{\alpha}^j.$$
Using (12) and (15) we have $\underline{\alpha} = \overline{\alpha} = \alpha$. This prove the first part of the proposition.
To obtain the second part, we note that (8) and (13), together with the fact that $\underline{\alpha} = \overline{\alpha} = \alpha$, imply that
$$\lim_{n\to\infty} T_n^i/t_n^i - \exp(\gamma^i \sigma_n^i) = 0,$$
where
$$\gamma^i = g_i^{-1} \sum_{j=1}^{K} g_j \alpha^j.$$
We then use the relation $W_n^i = T_n - t_n$ together with (1). ∎

Remark 3: Note that it follows from (8) that even in permanent overload conditions, eventually all sessions manage to complete. [1].

Next, we present some continuity and monotonicity properties of the rate of growth of the queue as a function of the input rate.

Proposition 4: Under the same hypothesis of Proposition 2, the explosion rate of class i customers is a continuous function of $\lambda_j,\ j = 1, \ldots, K$.
Proof: It follows from the Implicit Function Theorem that the strictly positive solution x of the equation $Z(x) = 0$ (see (18)) is continuous in all λ_j's. From (17) we now see that the strictly positive solution of (16) is continuous in that x and hence in the λ_j's. The result now follows from Proposition 2. ∎

In the following proposition we use the notations introduced in Proposition 2.

Proposition 5: (Monotonicity) Under the same hypothesis as in Proposition 2, the explosion rate of class i customers is an increasing function of each input rate $\lambda_j,\ j = 1, \ldots, K$.
Proof: From the definition of ν and α^i (see Proposition 2), it is enough to prove that ν is an increasing function of $\lambda_j,\ j = 1, \ldots, K$. Let us recall that ν is the unique positive root of
$$G_{\tilde{\lambda}_1,\ldots,\tilde{\lambda}_K}(x) = x - \sum_{j=1}^{K} \tilde{\lambda}_i \left[1 - \hat{E}^i\left(e^{-g_i^{-1} x \sigma_0^i}\right)\right] \quad (19)$$
The proposition is then a consequence of the fact that $G_{\tilde{\lambda}_1,\ldots,\tilde{\lambda}_K}(x)$ is convex, has another 0 at $\nu = 0$ and is decreasing with respect to $\lambda_j,\ j = 1, \ldots, K$. ∎

III. SUFFICIENT AND NECESSARY CONDITIONS FOR STABILITY

Since we are not dealing with an ergodic stationary setting, we shall say that the system is stable if $\alpha^j = 0\ \forall j$. Proposition 2 shows that
$$\sum_{i=1}^{K} \lambda_i \hat{E}^i(\sigma_0^i) \leq 1 \quad (20)$$
is a necessary condition for stability. We show now that this is also a sufficient condition. Indeed, under condition (20), the function (18) is strictly increasing and has a root in zero. Then it has no positive roots. By using the same arguments as in Step 4 of Proposition 2, one can easily prove that $\bar{\alpha}^j = 0\ \forall j$. This implies that $\alpha^j = 0\ \forall j$ so that the system is stable.

IV. A NETWORK OF DPS QUEUES

The fact that the results of the previous Section did not require stationarity of interarrival times but just the existence of rates, was sufficient to show that the output processes also have rates, which are uniquely defined as the positive solution of a set of implicit equations. This implies that the results above holds in fact to any network of DPS queues, that can be described as a directed graph provided that there are no feed-backs. The only requirement is that in each queue l and class i, the sequence of service times $\sigma_0^i(l), \sigma_1^i(l), \sigma_2^i(l), ...$ are stationary ergodic where $\sigma_k^i(l)$ are the amount of service required by the kth arrival of class i at queue l.

The next question is whether this can be generalized to networks with feedback. We first note that the results for a single queue may generalize to a single queue with feedback under the following assumptions:

- The arrival rates $\lambda_i := \lim_{t\to\infty} \frac{N^i[0,t]}{t}$ exist $P - a.s.$
- The nth customer of class i gets r_n^i services of durations $\sigma_n^i(1), \sigma_n^i(2), ..., \sigma_n^i(r_n^i)$. The jth service of the nth customer of class i starts immediately after the $j-1$st service ends. Denote $\overline{\sigma}_n = \sum_{j=1}^{r_i^n} \sigma_n^i(j)$. Each one of the K sequences $\{\overline{\sigma}_n^i\}$ is assumed to be stationary (under the shift in n) under P.

Under the above assumptions, it is clear that the results of the previous section hold, when we replace σ_n^i by $\overline{\sigma}_n^i$. Hence, under this setting, we may still apply all previous results to network of DPS queues, provided that there are no feedback.

When considering networks with feedback, the analysis becomes much more difficult and will be pursued in the future.

V. Examples of distributions

Consider the case where for each class $k = 1, ..., K$, the marginal distribution of the service times is exponential with parameter μ_i (Recall that the service times need not be i.i.d.). We now compute explicitly the solution of (6). The expression

$$\hat{E}^i \exp\left(-g_i^{-1}\sigma_0^i \sum_{j=1}^{K} x^j g_j\right) \quad (21)$$

is the Laplace Stieltjes transform of σ_0^i at the point $g_i^{-1}\sum_{j=1}^{K} x^j g_j$ and is thus given by

$$\hat{E}^i \exp\left(-\frac{\sigma_0^i \sum_{j=1}^{K} x^j g_j}{g_i}\right) = \frac{\mu_i}{\mu_i + g_i^{-1}\sum_{j=1}^{K} x^j g_j}.$$

Hence, equation (6) reduces to

$$\begin{aligned} x^i &= \lambda_i\left[1 - \frac{\mu_i}{\mu_i + g_i^{-1}\sum_{j=1}^{K} x^j g_j}\right] \\ &= \frac{\lambda_i \sum_{j=1}^{K} x^j g_j}{\mu_i g_i + \sum_{j=1}^{K} x^j g_j} \\ &= \frac{\lambda_i y}{\mu_i g_i + y}, \ i = 1 \ldots K, \end{aligned} \quad (22)$$

where $y = \sum_{j=1}^{K} x^j g_j$. Multiplying (22) by g_i and taking the sum over $i = 1, ..., K$, we get the following single equation for the unknown y:

$$1 = \sum_{i=1}^{K} \frac{\lambda_i g_i}{\mu_i g_i + y}. \quad (23)$$

For the case of $K = 1$, this gives $y = g_1(\lambda_1 - \mu_1)$ and thus $x_1 = \lambda_1 - \mu_1$. This can be explained by the fact that for exponential case, the dynamics of the system has the same distribution as a FIFO system, for which the rate of growth of number of customers in the system is clearly $\lambda - \mu$.

For the case of $K = 2$, (23) yields a quadratic equation $y^2 + by + c = 0$ where

$$b = g_1(\mu_1 - \lambda_1) + g_2(\mu_2 - \lambda_2),$$

$$c = g_1 g_2(\mu_1\mu_2 - \lambda_1\mu_2 - \lambda_2\mu_1).$$

Note that the overload condition $\lambda_1/\mu_1 + \lambda_2/\mu_2 > 1$ implies that $c < 0$, so indeed the quadratic equation has a unique positive solution. The solution then determines x_1, x_2 through (22).

Whereas the exponential distribution allows us to obtain a simple solution for the overload, it does not describe well the distribution of the size (and hence service time) of a TCP session. The latter is known to be heavy tailed [1], [3] and in particular the Pareto and the Weibull distribution have been used to describe its distribution. Unfortunately, with these distributions we cannot get an explicit expression for (21) and therefore we cannot solve (6) explicitly. Nevertheless, one can use the hyper-exponential distribution which has been shown in [3] to approximate the heavy tailed distributions. We thus write the complementary probability distribution function of the service time σ_0^i of class i as

$$H_i^c(t) = \sum_{j=1}^{B(i)} p_j^i e^{-\mu_i^j t}.$$

It is thus a mixture of $B(i)$ exponentials weighted by the probabilities p_i^j, $j = 1, ..., B(i)$. The Laplace Stieltjes transform of σ_0^i is

$$\mathcal{L}_i(s) = \sum_{j=1}^{B(i)} \frac{p_j^i \mu_i^j}{s + \mu_i^j},$$

and thus

$$1 - \mathcal{L}_i(s) = 1 - \sum_{j=1}^{B(i)} \frac{p_j^i \mu_i^j}{s + \mu_i^j} = \sum_{j=1}^{B(i)} \frac{p_j^i s}{s + \mu_i^j}.$$

substituting in the above $s = y/g_i$, (6) then becomes:

$$\begin{aligned} x^i &= \lambda_i\left[1 - \hat{E}^i \exp\left(-g_i^{-1}\sigma_0^i \sum_{j=1}^{K} x^j g_j\right)\right] \\ &= \lambda_i \sum_{j=1}^{B(i)} \frac{p_j^i y}{y + g_i \mu_i^j}, \ i = 1 \ldots K, \end{aligned} \quad (24)$$

where again $y = \sum_{j=1}^{K} x^j g_j$. Multiplying by g_i and summing over $i = 1, ..., K$, we get the following single equation with the unknown y:

$$1 = \sum_{i=1}^{K} \lambda_i g_i \sum_{j=1}^{B(i)} \frac{p_j^i}{y + g_i \mu_i^j}. \quad (25)$$

It's solution gives us then the x^i of each class i by substituting in (24). Due to the nature of the super-exponential distribution, the above equation has in fact the *same form* of (23); this can be seen if we define $\lambda_i^j = \lambda_i p_j^i$, and $g_i^j = g_i$. Then (25) becomes

$$1 = \sum_{i,j} \frac{\lambda_i^j g_i^j}{y + g_i^j \mu_i^j}. \quad (26)$$

Finally, we note that even for more general distributions of the σ_0^i, our general results (e.g. Proposition 2) allow us to compute numerically the performance measures (such as increase rate of number of session).

VI. Application to TCP in overload conditions

In this section we apply the analysis for a single queue with DPS regime in overload to model several sources that send TCP sessions who share a common bottleneck link. (We note that the framework of DPS queues is not appropriate for studying TCP connections in a more complex topology with several bottlencek nodes since our model for a network assumes a store and forward mechanism: a connection starts to be treated in a DPS node only when it has ended to be treated in the previous uplink node; TCP however shares in practice all nodes along its path simultaneoulsy).

Consider the following network depicted in Figure 1. There are two groups of sessions all sharing the same bottleneck link (of 2Mbps). The sessions arrive at the bottleneck over high speed access links of 100Mbps. All TCP packets are of size 500bytes. We use the NewReno version of TCP without the delayed ack option.

Arrival rates The time between session inter-arrivals within each group is exponential and its expectation is 0.24sec. Thus $\lambda_i = 4.167, i = 1, 2$.

Service rates A file has an exponentially distributed size with average size of 40Kbytes. The system is indeed in overload: its load is given by twice λ times the average size of a flow in bits over the link's capacity. This gives $\rho = 2 \times 4.167 \times 40 \times 10^3 \times 8/2Mbps = 1.333$.

Delays and classes of sessions The propagation delay over the bottleneck link is 50ms. The two classes of sessions are distinguished by the delay of the access link (and thus by the RTT). The propagation delay at the access link of class 1 is 50ms. The propagation delay at the access link of class 2 is varied in our simulations from 50ms to 500ms. When computing the RTT, we will have to take into account the queueing delay. We note that the transmission time of a packet into the bottleneck link takes $500 \times 8/2Mbps = 2msec$. We wish the queueing delay to be small with respect to the propagation delay so that we can indeed change easily RTT over a large range. We thus chose a queue size of 30 packets. This implies a maximum queueing delay of 60msec. (As we shall see, the average queueing delay will be around 50msec.) The average RTT of the first group is 250msec, where RTT of the second one varies between 250msec to 1150msec in steps of 100msec.

Simulation were performed with ns simulator [9].

We begin by depicting in Fig. 2 the impact of the delay $D2$ (or equivalently of $RTT2$) of group 2 on the throughput (in terms of sessions per second) of each one of the classes. For each class, we depict the number of sessions that have been completed by the end of the simulations (50 sec). The isolated points correspond to simulations and the continuous curves to the numerical results discussed later on.

We see that as $D2$ increases, the general trend of the throughput of class 2 is to decrease and that of class 1 to increase. The throughput of class 1 is always larger than of class 2 when $D2 > D1$. For the largest $D2$, where it is 4.6 times $D1$, the throughput of class 1 is around the double of that of class 2.

Next, we look in more details on the simulations of the system where in average, $RTT2 = 550msec$ (and $RTT1 = 250msec$). Fig. 3 shows the cumulative number of departures as a function of time of each class. The upper curves corresponds to class 1. We see that the curves are very close to being linear, as predicted by our analytical results. As can be expected, longer RTTs imply lower throughputs of packets and thus also lower throughputs in terms of number of sessions terminating per second. In our formalism, we can say that the DPS weighting factor g_2 for class 2 is smaller than for class 1 (g_1). We shall show later how these parameters are selected.

Fig. 4 shows the number of sessions of each class present at the system. We see, as expected, a general linear trend in their growth. The upper curve corresponds to class 2. Since the arrival rates of both classes is the same, it is clear that the lower departure rates of class 2 that we saw in the previous figure will mean a larger number of class 2 sessions in the system.

Fig. 5 shows the average and instantaneous queue sizes at the entrance to the bottleneck link. We see that except for a short transient time, the average queue size stabilizes at 25 packets.

Finally, the cumulative number of lost packets at the bottleneck queue is depicted in Figure 6. It is seen to be quite close to linear.

Validation. We next computed numerically the throughputs. For equal delays we can take $g_1 = g_2 = 1$. We then obtain $y = \mu - \lambda_1 - \lambda_2 = 2.083$, so the rates of growth of the number of sessions in the system are $x_i = 1.0415, i = 1, 2$, and the throughputs are $\lambda_i - x_i = 3.125$. The simulated throughputs are $140/50 = 2.8$ so we get an error of around 10%.

For unequal delays the problem of choosing the g_i

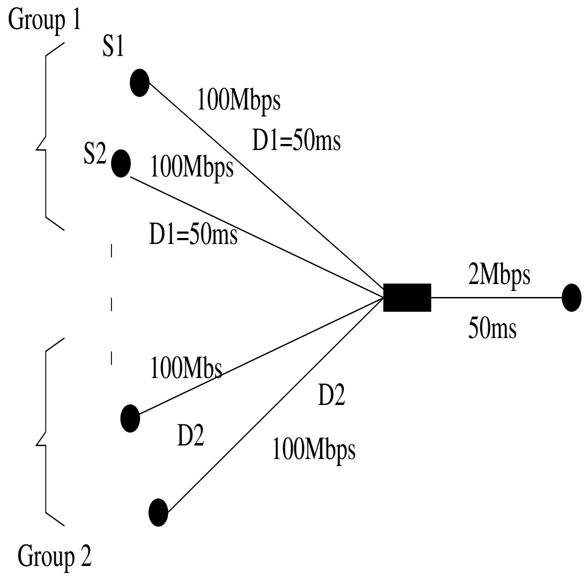

Fig. 1. Simulated Network

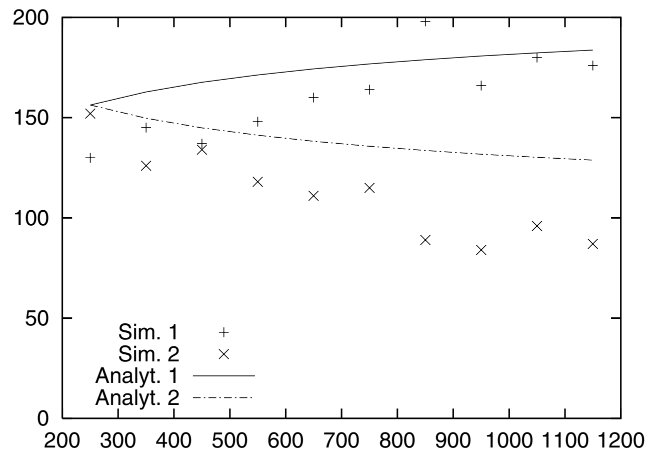

Fig. 2. Number of departures as function of RTT2 (in msec)

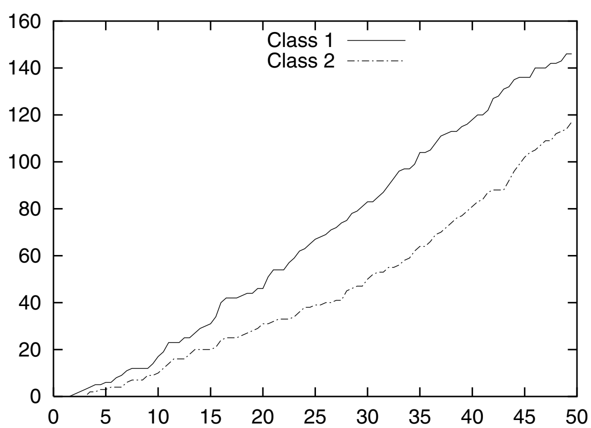

Fig. 3. Cumulative number of departures as a function of time (sec)

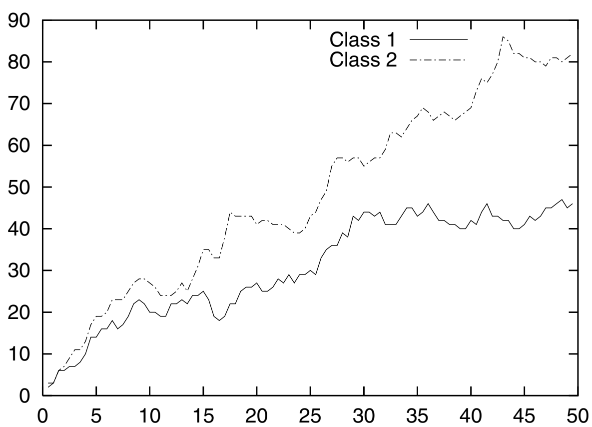

Fig. 4. Number of ongoing sessions as a function of time (sec)

is harder since we are not aware of simple explicit expressions for throughputs of short TCP sessions as a function of RTT. As an approximation, we choose $g_i = RTT_i^{-1}$, in line with the well known throughput formulae [10] that was derived for persistent sessions. We note that in our case, a session contains at the average 80 packets. Our simulations showed that the average duration of a session was much larger than the initial slow start phase, so we could indeed use results derived for persistent TCP. (This can be seen for example from the fact that the number of losses per session is much larger than 1. For example, in the simulation corresponding to $RTT2 = 550msec$ there were 3350 packets lost among around 400 sessions.)

With this choice of the g_i's, we see the comparison between simulations and our model prediction in Fig. 2. We see that our model tends to over estimates the sessions' throughput, with an average error of aound 10%.

We have obtained an even better prediction of the simulation results (not reported here) for $\rho = 1.333$ when using slightly higher load (than $\rho = 1.333$) in the analytical equations. A possible explanation for this phenomenon could be that since the queue has emptied during the simulation (especially during the first 10 seconds), some of the link capacity was wasted during

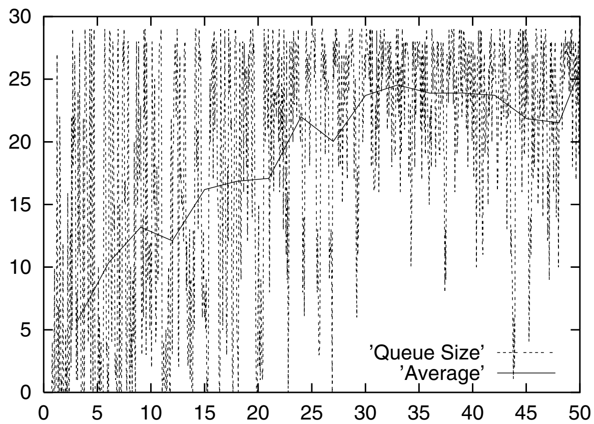

Fig. 5. Instantaneous and average queue size as a function of time (sec)

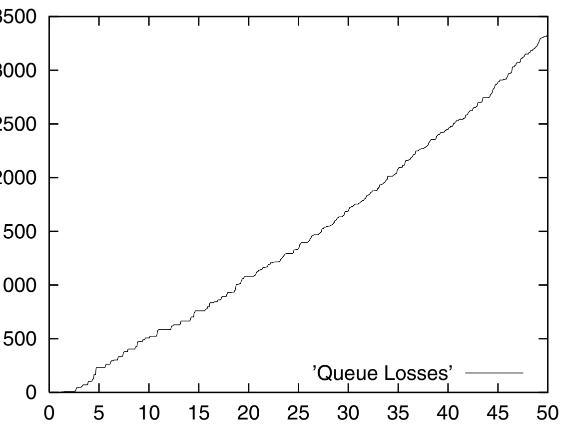

Fig. 6. Packets lost at bottleneck queue as a function of time (sec)

the simulation time so that the service time required by a session is larger (and hence a slightly larger load would better predict the simulation results).

VII. CONCLUDING REMARKS

The main goal of this paper has been to analyze the behavior of DPS queues in overload. Surprisingly simple expressions have been obtained for a general (not necessarily) Markovian setting for computing the growth rate and the throughput. These expressions were explicit for some special distributions of service times. We then tested the applicability of our results for modeling the behavior of TCP sessions sharing a common bottleneck link. When using our model with the DPS weights given by the reciprocal of the round trip times, our model did quite well, overestimated the throughputs by around 10%.

Acknowledgement. The authors wish to thank Dr. Arzad Alam Kherani for useful discussions on Section VI.

REFERENCES

[1] T. Bonald and J. Roberts, “Performance modeling of elastic traffic in overload”, *ACM Sigmetrics*, Cambridge, MA, USA, June, 2001.

[2] El Taha and Stidham, *Sample path analysis of queueing systems*, Kluwer, 2000.

[3] A. Feldmann and W. Whitt, ”Fitting Mixtures of Exponentials to Long-Tail Distributions to Analyze Network Performance Models”, *INFOCOM’97*, 1096-1104, 1997.

[4] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnié and J. W. Roberts, “Statistical bandwidth sharing: a study of congestion at flow level”, *SIGCOMM’*, 2001.

[5] M. Haddani, ”Etude de modeles probabilistes de reseaux de telecommunication”, Thesis, Univ. of Paris VI, 2001.

[6] A. Jean-Marie, *Numerical study of the processor sharing queue*, unpublished.

[7] A. Jean-Marie and P. Robert, *On the transient behavior of the processor sharing queue*, *QUESTA* **17**, No. 1-2, pp. 129-136, 1994.

[8] D. Mitra and A. Weiss, “A closed network with a discriminatory processor-sharing server”, *Performance Evaluation Review* **17**(1), pp. 200-208, 1989.

[9] Network Simulator v.2, University of California at Berkeley, Available via http://www-nrg.ee.lbl.gov/ns-2.

[10] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP Throughput: a Simple Model and its Empirical Validation”, in Proceedings of *ACM SIGCOMM’1998* conference.