Asymptotic Regimes and Approximations for Discriminatory Processor Sharing *

Gijs van Kessel[†] Rudesindo Núñez-Queija[†‡] Sem Borst[†‡§]

ABSTRACT

We study the joint queue length distribution of the Discriminatory Processor Sharing model, assuming all classes have phase-type service requirement distributions. We show that the moments of the joint queue length distribution can be obtained by solving linear equations. We use this to study the system in two asymptotic regimes. In the first regime, the different user classes operate on strictly separated time scales. Then we study the system in heavy traffic.

1. INTRODUCTION

The adoption of Processor Sharing (PS) as a modeling abstraction of TCP bandwidth sharing [8] – identifying a customer in the PS model with an active TCP flow – triggered a renewed interest in the analysis of PS models. The PS discipline assumes a perfectly egalitarian distribution of the bandwidth among all active flows. Because of TCP's distributed nature, however, the actual shares of flows sharing a common (bottleneck) link may show strong asymmetry, see for instance [1].

Discriminatory Processor Sharing (DPS) was introduced in [7] as a multi-class extension of (egalitarian) PS. The DPS discipline provides a natural approach for modeling the flow-level performance of heterogeneous TCP flows. The analysis of the DPS discipline is extremely difficult compared to that of ordinary PS. Most notably, the simple geometric queue length distribution for the standard PS discipline [10] does not have any counterpart for DPS. Mean sojourn times were studied in [4], in particular showing that, in the case of exponentially distributed service requirements, these can be obtained from linear equations. This result was then extended to higher moments of the queue length distribution [9], while also proving a heavy-traffic limit theorem. In this paper we further extend the results of [9] to *phase-type* distributions. In addition, we use the theory of nearly-completely decomposable Markov Chains to study DPS, assuming a strict separation of time scales among the different user classes.

2. MODEL DESCRIPTION

*This work is part of EQUANET, supported by the Dutch Ministry of Economic Affairs via its agency SenterNovem

[†]Dept. Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB, The Netherlands; r.nunez@tue.nl, sem@win.tue.nl

[‡]CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands; sindo@cwi.nl, sem@cwi.nl

[§]Bell Laboratories, Lucent Technologies, P.O. Box 636, Murray Hill, NJ 07974, USA; sem@research.bell-labs.com

Customers of K (≥ 1) different classes arrive at a single server (class k arrives at rate λ_k). We assume that the arrival processes are independent Poisson processes and denote the fraction of customers from class k by $p_k = \lambda_k/\Lambda$, where $\Lambda = \sum_{k=1}^{K} \lambda_k$ is the total arrival rate. A generic service time of class-k customers is denoted by B_k. The load of class k is $\rho_k = \lambda_k \mathbb{E} B_k$ and the stability condition reads $\rho < 1$, where $\rho = \sum_{k=1}^{K} \rho_k$ is the total load.

With N_k we denote the number of customers of class $k = 1, 2, \ldots, K$ (in steady state). All customers are served simultaneously. The weight of class k is denoted with $w_k (> 0)$. If there are $N_l = n_l$ customers of class $l = 1, 2, \ldots, K$, then each customer of class k receives a fraction $w_k / \sum_{l=1}^{K} n_l w_l$ of the server's capacity. Note that if all w_k are equal, the system reduces to that of egalitarian PS.

3. EXACT ANALYSIS FOR PHASE-TYPE SERVICE REQUIREMENTS

We assume that the service requirement distributions of all classes are of phase-type. (Similar results were independently obtained in [5].) Within a customer class we distinguish between customers that are in different service phases and refer to these as belonging to different customer *types*. Denoting the number of phases of the class-k phase-type distribution with m_k, the total number of types is $\sum_{k=1}^{K} m_k := J$. A class-i customer in its j^{th} service phase is of type $\sum_{k=1}^{i-1} m_k + j$. We use $k(j)$ to denote the customer class to which type-i customers belong.

Let p_{0j} be the probability that an arriving customer starts as a type-j customer, $j = 1, \ldots, J$. Thus, $\sum_{j:k(j)=l} p_{0j} = p_l$. The service phase corresponding to type j has mean duration $1/\mu_j$ and its service weight is g_j. It would be natural to take $g_i = g_j = w_{k(j)}$ if $k(i) = k(j)$, i.e., if types i and j belong to the same customer class, but this is not necessary for the analysis in this section. Furthermore, define p_{ij} $(i, j = 1, \ldots, J)$ as the probability that, after completing its current service phase a type-i customer becomes a type-j customer. In principle, no transitions are possible between types belonging to different customer classes but, again, we will not explicitly use this. Also, p_{i0} is the probability that a customer of type i will leave the system after the current service phase completion. We shall denote the number of type-j customers in the system by N'_j. Obviously, $\sum_{i=1}^{K} p_{0i} = 1$, $\sum_{j=0}^{J} p_{ij} = 1$, and $\sum_{j:k(j)=l} N'_j = N_l$.

Denoting by $\bar{N}'$ and $\bar{n}$ the vectors $(N'_1, N'_2, \ldots, N'_J)$ and $(n_1, n_2, \ldots, n_J) \geq \bar{0}$, respectively, the steady-state distribution $P(\bar{n}) := \mathbb{P}(\bar{N}' = \bar{n})$ satisfies

$$[\Lambda + \frac{\sum_{i=1}^J g_i n_i \mu_i}{\sum_{j=1}^J g_j n_j}] P(\bar{n})$$
$$= \sum_{i=1}^J \Bigg[\Lambda p_{0i} P(\bar{n} - \bar{e}_i) + \frac{g_i(n_i+1)\mu_i p_{i0}}{g_i + \sum_{k=1}^J g_k n_k} P(\bar{n} + \bar{e}_i)$$
$$+ \sum_{j=1}^J \frac{g_i(n_i+1)\mu_i p_{ij}}{g_i - g_j + \sum_{k=1}^J g_k n_k} P(\bar{n} + \bar{e}_i - \bar{e}_j) \Bigg]. \qquad (1)$$

(For notational convenience we define $P(\bar{n} - \bar{e}_i) = 0$ if $n_i = 0$.) It will prove convenient to use the following transformation for $\bar{n} \neq \bar{0}$:

$$R(\bar{n}) = \frac{P(\bar{n})}{\sum_{j=1}^J n_j g_j}.$$

We further define $R(\bar{0}) = 0$. Also, let $p(\bar{z})$ and $r(\bar{z})$ denote the generating functions of $P(\bar{n})$ and $R(\bar{n})$, respectively, where $\bar{z} = (z_1, \ldots, z_J)$ and $|z_i| < 1$ for $i = 1, \ldots, J$:

$$p(\bar{z}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_J=0}^{\infty} z_1^{n_1} \ldots z_J^{n_J} P(\bar{n}) = \mathbb{E}\left[z_1^{N_1'} \ldots z_J^{N_J'} \right],$$
$$r(\bar{z}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_J=0}^{\infty} z_1^{n_1} \ldots z_J^{n_J} R(\bar{n}).$$

It follows that

$$p(\bar{z}) = \sum_{i=1}^J g_i z_i \frac{\partial r}{\partial z_i} + 1 - \rho. \qquad (2)$$

From (1) we obtain a partial differential equation for $r(\bar{z})$:

$$\sum_{i=1}^K \left\{ \mu_i g_i (p_{i0} + \sum_{j=1}^K p_{ij} z_j - z_i) - \Lambda g_i z_i (1 - \sum_{j=1}^K p_{0j} z_j) \right\} \frac{\partial r}{\partial z_i}$$
$$= \Lambda(1-\rho)(1 - \sum_{j=1}^K p_{0j} z_j). \qquad (3)$$

This equation enables us to determine the moments of the queue length distribution by solving linear equations. Define the following partial derivatives of $p(\bar{z})$ and $r(\bar{z})$:

$$L^j_{i_1 \ldots i_j} = \lim_{\bar{z} \uparrow 1} \frac{\partial^j p(\bar{z})}{\partial z_{i_1} \ldots \partial z_{i_j}}, \quad R^j_{i_1 \ldots i_j} = \lim_{\bar{z} \uparrow 1} \frac{\partial^j r(\bar{z})}{\partial z_{i_1} \ldots \partial z_{i_j}}.$$

The next three theorems determine the mean numbers of customers of each type (L_i^1 for type i).

Theorem 1. *The j^{th} moment of the queue length at phase i can be expressed in terms of $R^j_{i_1 \ldots i_j}$ and $R^{j+1}_{i_1 \ldots i_{j+1}}$ as follows:*

$$L^j_{i_1 \ldots i_j} = \sum_{i=1}^J g_i R^{j+1}_{i_1 \ldots i_j i} + \sum_{l=1}^j g_{i_l} R^j_{i_1 \ldots i_j}.$$

Proof. The theorem can be directly obtained by differentiating (2) and letting $z_i \to 1$ for all i. □

Let a_{ij} be the accumulated amount of work (from arrival until departure) received at phase j assuming a start in phase i. Obviously, the a_{ij} are determined by $a_{ij} = \sum_{k=1}^K p_{ik} a_{kj}$, if $i \neq j$, and $a_{ii} = \frac{1}{\mu_i} + \sum_{k=1}^K p_{ik} a_{ki}$.

Theorem 2. *The R_i^1, $i = 1, 2, \ldots, J$, are given by*

$$R_i^1 = \frac{1}{g_i} \Lambda \sum_{i=1}^K p_{0i} a_{ij}.$$

Proof. From (3) it follows that the R_i^1 satisfy the following set of J equations with J unknowns where $i = 1, \ldots, J$,

$$\sum_{l=1}^K g_l \mu_l p_{li} R_l^1 - \mu_i g_i R_i^1 = -\Lambda p_{0i},$$

which admits a unique solution, given in the theorem.

An alternative informal argument is as follows. The capacity dedicated to type i is $g_i R_i^1 = \mathbb{E}\left[g_i N_i' / \sum_{k=1}^K g_k N_k' \right]$. This must be equal to the arriving amount of work at phase i, which is $\Lambda \sum_{i=1}^K p_{0i} a_{ij}$. □

Theorem 3. *The $R^2_{i_1 i_2}$ satisfy the following set of J^2 equations with J^2 unknowns where $i_1, i_2 = 1, \ldots, J$:*

$$(\mu_{i_1} g_{i_1} + \mu_{i_2} g_{i_2}) R^2_{i_1 i_2}$$
$$= \sum_{l=1}^K g_l (\Lambda p_{0 i_1} + \mu_l p_{l i_1}) R^2_{i_2 l} + \Lambda p_{0 i_1} g_{i_2} R^1_{i_2}$$
$$+ \sum_{l=1}^K g_l (\Lambda p_{0 i_2} + \mu_l p_{l i_2}) R^2_{i_1 l} + \Lambda p_{0 i_2} g_{i_1} R^1_{i_1}.$$

Proof. Follows from (3). □

From numerical experiments we found that this system of J^2 equations with J^2 unknowns is linearly independent (for a system with two customer types it is not hard to prove that this is true if $\Lambda < \mu$), implying that R^2_{ij} $(i, j = 1, \ldots, J)$ are uniquely determined by Theorem 3.

As a side remark, note that the set of equations in Theorem 3 can be reduced to one of $J(J+1)/2$ equations and equally many unknowns by using the fact that $R^2_{kl} = R^2_{lk}$.

4. TIME SCALE SEPARATION

Assuming that a strict *separation of time scales* is allowed, the DPS model can be analyzed exactly. Let the arrival rate and service requirements of class 1 be scaled with a positive parameter r: $\lambda_1^{(r)} = r \lambda_1^{(1)}$ and $\mathbb{P}^{(r)}(B_1 \leq x) = \mathbb{B}_1(rx)$, for some constant $\lambda_1^{(1)} > 0$ and probability distribution function $\mathbb{B}_1(\cdot)$. (To reflect the dependence on r we shall use a superscript $^{(r)}$.) Observe that the load of class 1 is not affected by r: $\rho_1^{(r)} = \lambda_1^{(r)} \mathbb{E}^{(r)} B_1 \equiv \lambda_1^{(1)} \mathbb{E}^{(1)} B_1 =: \rho_1$.

Theorem 4. *Let the service requirement distributions of both classes be of phase-type. Then, for $n_1, n_2 \in \mathbb{N}$, with $\Gamma(x) = \int_0^\infty e^{-u} u^{x-1} du$,*

$$\lim_{r \to \infty} \mathbb{P}^{(r)}(N_1 = n_1 | N_2 = n_2)$$
$$= \frac{\Gamma(n_1 + n_2 w_2/w_1 + 1)}{\Gamma(n_1+1)\Gamma(n_2 w_2/w_1 + 1)} \rho_1^{n_1} (1-\rho_1)^{\frac{n_2 w_2}{w_1} + 1}, \quad (4)$$

and

$$\lim_{r \to \infty} \mathbb{P}^{(r)}(N_2 = n_2) = \big(1 - \frac{\rho_2}{1-\rho_1}\big)\big(\frac{\rho_2}{1-\rho_1}\big)^{n_2}. \qquad (5)$$

In particular,

$$\lim_{r \to \infty} \mathbb{E}^{(r)}[N_1 | N_2 = n_2] = \Big(\frac{w_2}{w_1} n_2 + 1\Big) \frac{\rho_1}{1-\rho_1},$$

and

$$\lim_{r\to\infty} \mathbb{E}^{(r)} N_2 = \frac{\rho_2}{1-\rho},$$

so that

$$\lim_{r\to\infty} \mathbb{E}^{(r)} N_1 = \Big(\frac{w_2}{w_1}\frac{\rho_2}{1-\rho} + 1\Big)\frac{\rho_1}{1-\rho_1}.$$

Passing $r \to \infty$ corresponds to a perfect separation of time scales. N_2 evolves on a much slower time scale than N_1. For a fixed number $N_2 = n_2$ of class-2 users ("elephants"), class-1 users ("mice") equally share the service capacity $n_1/(n_1 + \frac{w_2}{w_1} n_2)$. (This heuristic argument is made rigorous in the proof of Theorem 4 for phase-type service requirement distributions, using the theories of *singular perturbation* and *nearly-completely decomposable Markov Chains*.) For fixed $N_2 = n_2$, with $n_2 w_2/w_1 \in \mathbb{N}$, the limiting dynamics of class-1 users correspond to that of the M/G/1 PS queue with $n_2 w_2/w_1$ permanent customers [2]. More generally, (4) agrees with [3, 6] when $n_2 w_2/w_1 \notin \mathbb{N}$.

In the limit $r \to \infty$, $1 - \rho_1$ is the capacity that is left for class-2 customers *for any fixed value of* $N_2 = n_2$, since class-1 customers will simply demand their average load ρ_1. Thus, in the limit, class 2 sees a processor sharing system with service capacity $1 - \rho_1$ and load ρ_2, which agrees with (5).

It is worth emphasizing that the limits in Theorem 4 are *insensitive* to characteristics of the service requirement distributions other than the means. In addition, we observe that, as $r \to \infty$, the limiting distribution of N_2 is independent of the weights, whereas that of N_1 does depend on the ratio of the weights.

$K \geq 2$ customer classes

The results of Theorem 4 can be extended to more than two customer classes. Let the arrival rates and service requirements of class k be scaled with r^{K-k}: $\lambda_k^{(r)} = r^{K-k}\lambda_k^{(1)}$ and $\mathbb{P}^{(r)}(B_k \leq x) = \mathbb{B}_k(r^{K-k}x)$. We then have, for class $i = 1, \ldots, K$,

$$\lim_{r\to\infty} \mathbb{E}^{(r)}[N_i | N_{i+1} = n_{i+1}, \ldots, N_K = n_K]$$
$$= \Big(\frac{\sum_{j=i+1}^K n_j w_j}{w_i} + 1\Big)\frac{\rho_i}{1-\sum_{j=1}^i \rho_j},$$

with

$$\lim_{r\to\infty} \mathbb{E}^{(r)} N_K = \rho_K/(1-\rho).$$

We can thus recursively compute $\lim_{r\to\infty} \mathbb{E}^{(r)} N_{K-k}$, $k = 1, \ldots K-1$.

5. HEAVY TRAFFIC

We now discuss a second asymptotic regime (no longer assuming a separation of time scales). In heavy traffic, with $\rho = \sum_{k=1}^K \rho_k \to 1$, the numbers of customers in the system of all classes will increase to infinity (with probability 1). However, when scaled with $1 - \rho$ we show that there is a proper limiting distribution.

Theorem 5. *If the service requirement distributions are of phase type, then for* $(\rho_1, \ldots, \rho_K) \to (\bar\rho_1, \ldots, \bar\rho_K)$, *with* $\sum_{k=1}^K \bar\rho_k = 1$,

$$(1-\rho)(N_1, N_2, \ldots, N_K) \xrightarrow{d} E \cdot (\frac{\bar\rho_1}{w_1}, \frac{\bar\rho_2}{w_2}, \ldots, \frac{\bar\rho_K}{w_K}),$$

where $\xrightarrow{d}$ *denotes convergence in distribution and* E *is an exponentially distributed random variable with mean*

$$\frac{\sum_k p_k \mathbb{E}[(B_k)^2]/\sum_k p_k \mathbb{E}B_k}{\sum_k \frac{1}{w_k}\bar\rho_k \mathbb{E}[(B_k)^2]/\mathbb{E}B_k},$$

which is equal to 1 in the case of standard PS, i.e., when $w_k = 1$ *for all* k.

This result reflects a so-called state-space collapse. In heavy traffic, the random vector $(1-\rho)(N_1, N_2, \ldots, N_K)$ converges in distribution to a constant vector multiplied with an exponentially distributed random scalar. This is proved by showing that the mean of $(N_1, N_2, \ldots, N_K)$ as well as the variance around this mean are of the order $O(\frac{1}{1-\rho})$. The exact form of the result can be understood by noting that $\rho_i = \mathbb{E}\left[\frac{w_i N_i}{\sum_k w_k N_k}\mathbf{1}_{\{\sum_k w_k N_k > 0\}}\right]$, where $\mathbf{1}_{\{A\}}$ equals 1 if A holds and it equals 0 otherwise. Then, if N_i/N_k tends to a constant, say n_i/n_k, it must be that $w_i n_i = \bar\rho_i \sum_k w_k n_k$, $i = 1, 2, \ldots, K$. Solving for n_i yields the result up to a multiplicative constant. The latter can be found by realizing that DPS is a work-conserving service discipline. Therefore, the mean workload is given by the Pollaczek-Khintchine formula

$$\mathbb{E}V = \frac{\rho}{2(1-\rho)}\sum_k p_k \mathbb{E}[(B_k)^2]/\sum_k p_k \mathbb{E}B_k,$$

which tends to $\frac{1}{2}\sum_k p_k \mathbb{E}[(B_k)^2]/\sum_k p_k \mathbb{E}B_k$, when scaling with $1-\rho$ and passing $\rho \to 1$.

6. REFERENCES

[1] Altman, E., Jimenez, T., Kofman, D. (2003). DPS queues with stationary ergodic service times and the performance of TCP in overload. In: *Proc. IEEE Infocom 2004*, Hong Kong.

[2] Van den Berg, J.L., Boxma, O.J. (1991). The $M/G/1$ queue with processor sharing and its relation to a feedback queue. *Queueing Systems* **9**, 365–401.

[3] Cohen, J.W. (1979). The multiple phase service network with generalized processor sharing. *Acta Informatica* **12**, 245–284.

[4] Fayolle, G., Mitrani, I., Iasnogorodski, R. (1980). Sharing a processor among many job classes. *J. ACM* **27**, 519–532.

[5] Haviv, M., Van der Wal, J. (2004). Waiting times in queues and in processor sharing systems with relative priorities. *Submitted for publication.*

[6] Kelly, F.P. (1979). *Reversibility and Stochastic Networks.* Wiley, Chichester.

[7] Kleinrock, L. (1967). Time-shared systems: a theoretical treatment. *Journal of the Association for Computing Machinery* **14**, 242–261.

[8] Massoulié, L., Roberts, J.W. (1999). Bandwidth sharing: objectives and algorithms. In: *Proc. IEEE Infocom '99*, New York NY, USA, 1395–1403.

[9] Rege, K.M., Sengupta, B. (1996). Queue length distribution for the discriminatory processor-sharing queue. *Oper. Res.* **44**, 653–657.

[10] Sakata, M., Noguchi, S., Oizumi, J. (1971). An analysis of the M/G/1 queue under round-robin scheduling. *Oper. Res.* **19**, 371–385.