Discriminatory Processor Sharing from Optimization Point of View *

J. Biro[2], T. Berczes[1], A. Korosi[3], Z. Heszberger[3], and J. Sztrik[1]

[1] University of Debrecen, Faculty of Informatics
[2] Inter-University Centre for Telecommunications and Informatics
Kassai u. 26., Debrecen, Hungary
[3] MTA-BME Information Systems Research Group
Budapest University of Technology and Economics
Magyar tudosok krt. 2., Budapest, Hungary
biro@tmit.bme.hu

Abstract. Discriminatory Processor Sharing models play important role in analysing bandwidth allocation schemes in packet based communication systems. Users in such systems usually have access rate limitations which also influence their bandwidth shares. This paper is concerned with DPS models which incorporate these access rate limitations in a bandwidth economical manner.

In this paper the interlock between access rate limited Discriminatory Processor Sharing (DPS) models and some constrained optimization problems is investigated. It is shown, that incorporating the access rate limit into the DPS model is equivalent to extending the underlying constrained optimization by constraints on the access rates. It also means that the available bandwidth share calculation methods for the access rate limited DPS are also non-conventional solution methods for the extended constrained optimization problem.

We also foreshadow that these results might be important steps towards obtaining efficient pricing and resource allocation mechanism when users are selfish and subject to gaming behavior when competing for communication resources.

1 Introduction

Processor Sharing (PS) models have been long studied in the queueing systems literature. One of the first reports on processor sharing have been performed in [10] motivated mainly by the modeling of time-shared computer systems. In this seminal work and also in [13] not only the egalitarian sharing (equal service rates are allocated for customers), but the (multi-class) discriminatory processor sharing has also been analyzed. In DPS the division of the processing capacity C can be controlled by a set of weights $(g_1, g_2, \ldots, g_K)$ in the following way. If there are $n_1, n_2, \ldots, n_K$ customers from the classes in the queuing system,

* This works was supported by FIRST TAMOP-4.2.2.C-11/1/KONV-2012-0001

they are served simultaneously, and the (instantaneous) service rate of a class$-i$ customer is

$$c_i = \frac{g_i}{\sum_{j=1}^{K} g_j n_j} C \quad .$$

Fayolle et al. in [5] completed the early efforts in solving discriminatory processor sharing models. They gave the correct characterizations for the M/G/1-PS system with respect to the steady-state average response times (by integro-differential equations), and also showed that in the special case of exponential distribution of the service time requirements, the steady-state average response times can be obtained by solving a system of linear equations. In [15] Rege and Sengupta showed how to obtain the moments of the queue length distributions as the solutions to linear equations in case of exponential service time requirements, and they also presented a heavy-traffic limit theorem for the joint queue length distribution. These results were extended to phase type distributions by van Kessel et al. [8]. A further remarkable milestone in DPS analysis is [1] in which the authors showed that the mean queue lengths of all classes are finite under reasonable stability conditions, regardless of the higher moments of the service requirements. It was also shown that the conditional response times of the different classes are stochastically ordered according to the DPS weights. In [3] Cheung, van den Berg and Boucherie for the egalitarian PS model obtained an exact and analytically tractable decomposition. This decomposition was applied to discriminatory PS to obtain an efficient and analytically tractable approximation of the queue length distribution and mean sojourn times.

In the discriminatory processor sharing models above, a common characteristics is that the users from the same class get equal share of the capacity allocated to that class, and the service capacity allocations of classes are proportional to their weights and to their number of users.

The other valuable direction of further developments of processor sharing model is to introduce capacity limits by which the customers can receive service from the server. This is mainly motivated by involving access rate limitations of users (e.g. in DSL-type access systems) in to the modeling framework. In [12] Lindberger introduced and analyzed the so-called M/G/R-PS model, which is a single-class processor sharing model with access rate limit b on the users ($R := C/b$ is the "number of servers" in this system). Several improvements of this model were studied for dimensioning purposes of IP access networks, e.g. in [16] and [4] still remaining at the single-class models.

Introducing access rate limitations in multi-class discriminatory processor sharing inherently raises the question of bandwidth re-distribution. This means that if customers in a class can not fully utilize their prospective service capacity share (bandwidth share) due to their access rate limit, how this unused (left) bandwidth is re-distributed among the other classes. In one of the extreme cases, there is no re-distribution at all meaning that the possible remaining unused bandwidth due to rate limits is wasted. One can also interpret this as the server capacity may not be fully utilized, even in those cases when there is "enough" customers in the system. This approach is followed for example in the papers [11],

[2]. In the other extreme case of multi-class discriminatory processor sharing, all the unused bandwidth by access rate limited customers are fully utilized by the other (non-limited) customers. In this scenario even the bandwidth shares can not be easily determined for given $\mathbf{n}$ and $\mathbf{g}$. Some of the classes will share certain amount of capacity proportional to their weights and number of users, whilst other classes' users are saturated at their access rate limits. Efficient algorithms have been presented for computing the capacity shares in [14] and [9].

In the original DPS model, the capacity shares of users for given $\mathbf{n}$ and $\mathbf{g}$ are solutions of a constrained optimization problem (see next chapter for details). Furthermore, because this solution is a proportional bandwidth allocation (among classes), this has a bridging role between the utilitarian social welfare maximization and individual payoff maximization, when classes of users are competing for the bandwidth shares.

2 Discriminatory Processor Sharing with Access Rate Limitations

Let n_i denote the number of class-i ($i = 1 \dots K$) flows (users, jobs) in the C capacity processor sharing system. In the original discriminatory processor sharing model [5] the server share its capacity among the classes in a discriminatory fashion by weights g_i, that is, a class-i flow bandwidth share from the total capacity C is

$$c_i = \frac{g_i}{\sum_{j=1}^{K} g_j n_j} C \ . \tag{1}$$

This weight-proportional share of bandwidth implies two important properties: For every pair of classes i, j the ratio of the service rates allocated to class-i and class-j users is equal to the ratio of the class weights, that is,

$$\frac{c_i}{c_j} = \frac{g_i}{g_j} \ , \ \forall i, j \in 1, \dots, K \ . \tag{2}$$

The total amount of capacity in a non-empty system used by the users of classes is C, that is,

$$\sum_{i=1}^{K} n_i c_i = C \ . \tag{3}$$

These properties can also be considered as requirements for a type of bandwidth share, which are uniquely fulfilled by (1).

It can also be shown (e.g. using Lagrangian multiplier method) that the solution of the following constrained optimization problem

$$\max_{\mathbf{x}} \sum_{i=1}^{K} n_i g_i \log x_i \ , \quad \text{s.t.} \quad \sum_{i=1}^{K} n_i x_i = C \quad , \quad x_i \geq 0 \tag{Opt}$$

is exactly that one in (1). Here it is worth noting that solving this optimization problem might be tedious at first glance, in fact, its solution is a closed-form proportional allocation-like formula in (1).

There can be several ways to introduce access rate limitations in the DPS model. One reasonable and simple enough approach is the following. Compute the bandwidth shares of class-i users according to (1) and cut at the access rate limits b_i, i.e.

$$c_i = \min\left(\frac{g_i}{\sum_{j=1}^{K} g_j n_j} C, b_i\right) . \tag{4}$$

which is in full accordance with the first extreme case (there is no bandwidth re-distribution at all) mentioned in the Introduction. Such kind of models were analyzed in [11] and [2]. From computational point of view the benefit of this case lies in the very simple calculation of service rates of classes, and as a consequence, the set of uncompressed (limited by their access rates) and the compressed (non-limited by their access rates, i.e. they can not reach their upper limit of their access rates b_i, hence the name "compressed") traffic classes directly follows from the service rate computation. The drawback of this simple access rate involvement is the possible waste of resources, that is it can occur that $\sum_{i=1}^{K} n_i c_i < C$ even if $\sum_{i=1}^{K} n_i b_i > C$. One can also interpret, that (3) as a requirement can not always be fulfilled by the bandwidth share governed by (4). [4]

Intuitively, a very straightforward way of introducing service rate limits would be to include further constraints $c_i \in [0, b_i]$ into the underlying optimization task, that is

$$\max_{\mathbf{x}} \sum_{i=1}^{K} n_i g_i \log x_i \ , \quad \text{s.t.} \ \sum_{i=1}^{K} n_i x_i = C \ \text{ and } \ x_i \in [0, b_i] \ , \quad i = (1, \ldots, K) \ . \tag{OptBounds}$$

Involving the access rate constraints into the optimization one may expect better utilization of the server capacity, when some users can not obtain its original proportional bandwidth share due to their access rate limits. In other words, the constraints of the optimization problem ensures that such an allocation represented by the optimal solution never waste server capacity if $\sum_{i=1}^{K} n_i b_i > C$, i.e. $\sum_{i=1}^{K} n_i c_i = C$ will hold.
Unfortunately, directly solving this extended optimization task can not be completed by Lagrange-multipliers, in this way it is not helpful in obtaining nice bandwidth share formula of the bandwidth-economical access-rate limited DPS. However, standard numerical methods could be applicable.

3 Determining the Bandwidth Shares

Instead of solving directly this optimization problem, first we recall two methods (presented in [14] and [9]) which can be used to determine the bandwidth shares of users in the bandwidth-economical access rate-limited DPS model.

[4] Henceforward we assume that $\sum_{i=1}^{K} n_i b_i > C$, that is, there is always at least one class, say class $-j$, for which $c_j < b_j$.

The first method is quite intuitive and provides an easily computable numerical method to determine the bandwidth shares. This is based on modifying the equation (2) into the following one:

$$c_j = \min\left(b_i, \frac{g_j}{g_i} c_i\right) \ , \ \forall i \in \mathcal{Z} \tag{5}$$

and keeping the capacity constraint as

$$\sum_{i=1}^{K} n_i c_i = C \ . \tag{6}$$

(Where $\mathcal{Z}$ denotes the set of compressed class indexes, that is, $\mathcal{Z} = \{i : c_i < b_i, i = 1, \ldots, K\}$). Then we show that this is also the solution of the optimization problem (OptBounds) above. An important observation is that if $\sum_{i=1}^{n} n_i b_i > C$ then there is always at least one compressed class (e.g. the class which has the smallest g_i/b_i ratio), let its index be K. Summing up both sides by j with $i = K$ we get (using (6))

$$C = \sum_{j=1}^{K} \min\left(b_K, \frac{g_j}{g_K} c_K\right) \tag{7}$$

from which c_K can be numerically determined (note that the right hand side is monotone increasing by c_K) and using (5) all the other c_i's can be determined. For more detailed analysis, see [14]. For the time being we don't know whether this is a solution of (OptBounds), we can only state that the equations (5) and (6) are fulfilled.

In the second, more deductive method (presented in [9]) , let us assume first without loss of generality that

$$\frac{g_1}{b_1} \geq \frac{g_2}{b_2} \geq \ldots \geq \frac{g_K}{b_K} \ . \tag{8}$$

Further, assume that $\left\{\sum_{i=1}^{K} n_i b_i > C\right\}$ holds, i.e. there is at least one class whose flows are compressed. In what follows, a method is given to determine the set of compressed flows $\mathcal{Z}$. For this, first consider the following inequality with respect to class-1:

$$\frac{g_1}{\sum_{j=1}^{K} n_j g_j} C \geq b_1$$

which could be rearranged as

$$\frac{g_1}{b_1} \geq \frac{\sum_{j=1}^{K} n_j g_j}{C} \ .$$

If this inequality holds, it means that class-1 is surely uncompressed, in other words, class-1 can not utilize its bandwidth share completely, there is excess bandwidth to be re-distributed among other classes.[5] It means that $C - n_1 b_1$ capacity is shared among the classes $\{2, \ldots, K\}$. Since

$$\frac{\sum_{j=1}^{K} n_j g_j - n_1 g_1}{C - n_1 b_1}$$

is constant, by following the order given in (8), class 2 has the highest chance to be uncompressed in the set $\{2, \ldots, K\}$. Thus, the following inequality should be checked in sequence:

$$\frac{g_2}{b_2} \geq \frac{\sum_{j=1}^{K} n_j g_j - n_1 g_1}{C - n_1 b_1} \ .$$

If it holds, class-2 is also uncompressed, and the similar inequalities should be checked for class-3.

It can be seen that for determining the compressed classes the following inequalities are to be checked, in increasing order of indexes

$$\frac{g_i}{b_i} \geq \frac{\sum_{j=i}^{K} n_j g_j}{C - \sum_{k=1}^{i-1} n_k b_k} \quad , \quad i = 1, 2, \ldots$$

Suppose that i^* is the last index for which the inequality above holds. On one hand, it follows that every class-i, $i \leq i^*$ is uncompressed, on the other hand, for every class-i, $i > i^*$ the inequalities

$$\frac{g_i}{b_i} < \frac{\sum_{j=i^*+1}^{K} n_j g_j}{C - \sum_{k=1}^{i^*} n_k b_k} \quad , \quad i = i^* + 1, \ldots, K$$

hold, that is, the set of compressed flows is $\mathcal{Z} = \{i^* + 1, \ldots, K\}$ and their bandwidth shares are

$$c_i = \frac{g_i}{\sum_{j=i^*+1}^{K} n_j g_j} (C - \sum_{k=1}^{i^*} n_k b_k) \quad , \quad i \in \mathcal{Z} \ . \tag{9}$$

It can also be shown in a straightforward manner that the c_i in (9) and $c_j = b_j$, $\forall j \in \mathcal{U}$ uniquely satisfy the requirements in (5) and (6).

4 The Main Result

Now, we are ready to state and prove the main result of the paper.

Theorem 1 *The solution of the optimization problem (OptBounds)* $\underline{c} = (c_1, \ldots, c_K)$ *is the following:*

[5] Also observe that if the inequality above would not hold, class-1 could exceptionally be attributed as surely compressed, because due to (8) all the other classes would also be compressed.

1. *If* $\sum_{k=1}^{K} n_k b_k \le C$*, then* $c_k = b_k$ *for* $1 \le k \le K$*.*
2. *If* $b_k \ge \frac{g_k}{\sum_{i=1}^{K} g_i n_i} C$ *for any* $1 \le k \le K$*, then* $c_k = \frac{g_k}{\sum_{i=1}^{K} g_i n_i} C$ *for* $1 \le k \le K$*.*
3. *If there exists an* i^**,* $1 \le i^* \le K-1$ *such that*

$$\sum_{k=1}^{i^*} n_k b_k + \sum_{k=i^*+1}^{K} n_k g_k \frac{b_{i^*}}{g_{i^*}} \le C \text{ and } \sum_{k=1}^{i^*+1} n_k b_k + \sum_{k=i^*+2}^{K} n_k g_k \frac{b_{i^*+1}}{g_{i^*+1}} > C,$$

then

$$c_k = b_k, \text{ if } k \le i^* \text{ and } c_k = \frac{g_k}{\sum_{i=i^*+1}^{K} g_i n_i} \left(C - \sum_{j=1}^{i^*} n_j b_j \right), \text{ if } \quad i^* < k.$$

Proof. We apply induction for the number of classes. If $K = 1$ the statement of the theorem is true. Now, assume that for $K-1$ classes the statement holds for arbitrary parameter setup. We prove the statement for K classes. By scaling g_i for each $1 \le i \le K$ with the same positive multiplicative factor does not affect the system. Hence without loss of generality we assume that the weights are chosen such that

$$\sum_{k=1}^{K} g_k n_k = 1. \tag{10}$$

We separate three cases.
In Case 1 we assume $b_1 \ge g_1 C$. This corresponds to the setup when all classes are compressed.
In Case 2 we assume that $b_1 < g_1 C$ and there exists $x_1^* \in [0, b_1]$ for which

$$c_2^{\mathcal{S}\setminus\{1\}, C-n_1 x_1} < b_2,$$

where $c_i^{\mathcal{S}\setminus\{1\}, C-n_1 x_1}$, $i \in \mathcal{S} \setminus \{1\}$ denotes the optimal solution of (OptBounds) with classes $\mathcal{S} \setminus \{1\}$ and capacity $C - n_1 x_1$. This corresponds to the setup when except the first class all classes are compressed.
In Case 3 we assume that $b_1 < g_1 C$ and for any $x_1 \in [0, b_1]$ we have

$$c_2^{\mathcal{S}\setminus\{1\},\ C-n_1 x_1} = b_2.$$

This corresponds to the setup when the first and the second class are compressed. Here, we have two subcases

a) $b_2 \le g_2 C$.
b) $b_2 > g_2 C$.

Case 1 Assume $b_1 \ge g_1 C$ holds. In this case the optimal solution of (Opt) is equal to the optimal solution of (OptBounds). Indeed, the optimal solution of (Opt) is $c_i = g_i C \ \forall i$. Since $g_1 C \le b_1$ and (8) it follows that $\forall i \ g_i C \le b_i$. Thus, $\forall i \ c_i = g_i C$ is optimal solution of (OptBounds) as well. We remark that we have not used induction in this case.

Case 2 Assume that $b_1 < g_1 C$ and there exists $x_1^* \in [0, b_1]$ for which

$$c_2^{\mathcal{S}\setminus\{1\}, C-n_1x_1^*} < b_2.$$

We will show that under these conditions the optimal solution of the optimization problem is

$$c_1 = b_1, \text{ and } c_i = \frac{g_i}{\sum_{j=2}^K n_j g_j}(C - n_1 b_1), \; i \geq 2, \tag{11}$$

where $c_i < b_i$ for $i \geq 2$.

First, we prove the following.

Lemma 1 *For any $1 \leq i \leq K$ we have $c_i \in [0, b_i]$.*

Proof. We clearly have $c_1 = b_1 \in [0, b_1]$. In the system with classes $S = \{2, 3, \dots, K\}$ and with capacity $C - n_1 x_1^*$ the optimal bandwidth for class 2 is smaller than b_2 since the assumption of Case 2. So class 2 is compressed. Using the induction one gets that each class is compressed, so

$$c_i^{\mathcal{S}\setminus\{1\}, C-n_1x_1^*} = \frac{g_i}{\sum_{j=2}^K n_j g_j}(C - n_1 x_1^*), \; 2 \leq i \leq K.$$

By the assumption $\frac{g_2}{\sum_{j=2}^K n_j g_j}(C - n_1 x_1^*) < b_2$ and by (8) we have

$$\frac{C - n_1 b_1}{\sum_{j=2}^K n_i g_i} \leq \frac{C - n_1 x_1 *}{\sum_{j=2}^K n_j g_j} < \frac{b_2}{g_2} \leq \frac{b_i}{g_i}, \; i \geq 2$$

$$\frac{g_i}{\sum_{j=2}^K n_j g_j}(C - n_1 b_1) < b_i, \; i \geq 2.$$

In the followings, we show that the optimal solution of (OptBounds) is given in (11). One can easily check that the following inequality holds

$$\max_{\substack{x_1 \in [0, b_1] \\ x_i \in [0, b_i], i \geq 2}} \sum_{i=1}^K n_i g_i \log x_i \leq \max_{\substack{x_1 \in [0, b_1] \\ x_i \geq 0, i \geq 2}} \sum_{i=1}^K n_i g_i \log x_i. \tag{12}$$

For $0 < x < C/n_1$ let

$$M(x) = \max_{\substack{x_1 = x \\ x_i \geq 0, i \geq 2}} \sum_{i=1}^K n_i g_i \log x_i.$$

Since the solution of problem (Opt) is known we have

$$M(x) = n_1 g_1 \log x + \sum_{i=2}^K n_i g_i \log\left(\frac{g_i}{\sum_{j=2}^K n_j g_j}(C - n_1 x)\right).$$

We will prove the following

Lemma 2 *$M(x)$ is increasing if and only if* $0 < x < g_1 C$.

Lemma 2 implies that on $[0, b_1]$ takes its maximum at $x = b_1$ since $b_1 < g_1 C$. Hence,

$$\max_{\substack{x_1 \in [0, b_1] \\ x_i \geq 0, i \geq 2}} \sum_{i=1}^{K} n_i g_i \log x_i \leq \max_{\substack{x_1 = b_1 \\ x_i \geq 0, i \geq 2}} \sum_{i=1}^{K} n_i g_i \log x_i .$$

From (12) we have

$$\max_{\substack{x_1 \in [0, b_1] \\ x_i \in [0, b_i], i \geq 2}} \sum_{i=1}^{K} n_i g_i \log x_i \leq \max_{\substack{x_1 = b_1 \\ x_i \geq 0, i \geq 2}} \sum_{i=1}^{K} n_i g_i \log x_i . \tag{13}$$

The optimal solution of the right hand side of (13) is c_i, $1 \leq i \leq K$ given in (11). c_i, $1 \leq i \leq K$ is a possible solution of the left side of (13) by Lemma 1, and because of the inequality (13) it is the optimal solution.

Proof (Proof of Lemma 2.). First, we show that $M(x)$ is increasing if $0 < x < g_1 C$. Take the derivative of $M(x)$:

$$M'(x) = n_1 g_1 \frac{1}{x} + \sum_{i=2}^{K} n_i g_i \frac{-n_1}{C - n_1 x} = n_1 g_1 \frac{1}{x} - n_1 \frac{\sum_{i=2}^{K} n_i g_i}{C - n_1 x}.$$

Find the interval on which $M'(x)$ is positive:

$$\begin{aligned}
n_1 g_1 \frac{1}{x} - n_1 \frac{\sum_{i=2}^{K} n_i g_i}{C - n_1 x} &> 0 \\
g_1 \frac{1}{x} &> \frac{\sum_{i=2}^{K} n_i g_i}{C - n_1 x} \\
g_1 C - n_1 g_1 x &> x \sum_{i=2}^{K} n_i g_i \\
g_1 C &> x \sum_{i=1}^{K} n_i g_i \\
g_1 C &> x.
\end{aligned}$$

Case 3 Assume that $b_1 < g_1 C$ and for any $x_1 \in [0, b_1]$ we have $c_2^{S \setminus \{1\},\, C - n_1 x_1} = b_2$ holds. In this case for any $x_1 \in [0, b_1]$ class 2 is uncompressed. Take away class 2, that is, consider the reduced system with classes $\{1, 3, \ldots, K\}$ and capacity $C - n_2 b_2$. Now, by induction the theorem holds for the reduced system. *We have to prove that in the reduced system class 1 is uncompressed.* If this holds, then Theorem 1 holds for the original system with classes $\{1, \ldots, K\}$ and capacity C. Indeed, if class 1 and 2 are uncompressed, then the badwidth shares of class 1 and class 2 are b_1 and b_2, respectively. We can apply the induction for the reduced system with classes $\{2, 3, \ldots, K\}$ and capacity $C - n_1 b_1$, that is, either case 1 or case 3 of Theorem 1 holds for the reduced system in the form:

1) If $\sum_{k=2}^{K} n_k b_k \leq C - n_1 b_1$, then $c_k = b_k$ for $2 \leq k \leq K$. For the original system we can rewrite the statement using simple rearrangement and using the fact that class 1 is uncompressed with bandwidth share b_1:
If $\sum_{k=1}^{K} n_k b_k \leq C$, then $c_k = b_k$ for $1 \leq k \leq K$.

3) For the reduced system: If there exists an i^*, $2 \leq i^* \leq K-1$ such that

$$\sum_{k=2}^{i^*} n_k b_k + \sum_{k=i^*+1}^{K} n_k g_k \frac{b_{i^*}}{g_{i^*}} \leq C - n_1 b_1$$

and

$$\sum_{k=2}^{i^*+1} n_k b_k + \sum_{k=i^*+2}^{K} n_k g_k \frac{b_{i^*+1}}{g_{i^*+1}} > C - n_1 b_1,$$

then

$$c_k = b_k, \text{ if } 2 \leq k \leq i^* \text{ and } c_k = \frac{g_k}{\sum_{i=i^*+1}^{K} g_i n_i} \left(C - \sum_{j=1}^{i^*} n_j b_j \right), \text{ if } i^* < k.$$

Using simple rearrangement and using the fact that class 1 is uncompressed with bandwidth share b_1, we can rewrite the statement and get the form that is written in the statement of Theorem 1.

In th rest of the proof we show that class 1 is uncompressed.

Case 3a Assume further that $b_2 \leq g_2 C$ holds. Assume that class 1 is compressed, that is,

$$\frac{g_1}{1 - g_2 n_2}(C - n_2 b_2) < b_1.$$

Taking into consideration condition that we have assumed $b_1 < g_1 C$, we have

$$\frac{g_1}{1 - g_2 n_2}(C - n_2 b_2) < g_1 C.$$

Rearranging the terms, we get

$$g_1(C - n_2 b_2) < g_1 C(1 - g_2 n_2)$$

$$g_1 C - g_1 n_2 b_2 < g_1 C - g_1 C g_2 n_2$$

$$b_2 > C g_2$$

which contradicts to assumption $b_2 \leq g_2 C$.

Case 3b Now, assume that $b_2 > g_2 C$ holds. The assumption, for any $x_1 \in [0, b_1]$ we have $c_2^{\mathcal{S} \setminus \{1\},\, C - n_1 x_1} = b_2$ in the reduced system with classes $\{2, \ldots, K\}$, means that for any $x_1 \in [0, b_1]$ we have

$$\frac{g_2}{\sum_{k=2}^{K} g_k n_k}(C - n_1 x_1) = \frac{g_2}{1 - n_1 g_1}(C - n_1 x_1) \geq b_2. \tag{14}$$

Especially for $x_1 = b_1$ we have

$$
\begin{aligned}
\frac{g_2}{1-n_1g_1}(C-n_1b_1) &\geq b_2 \\
\frac{g_2}{1-n_1g_1}(C-n_1b_1) - g_2C &\geq b_2 - g_2C \\
g_2C\left(\frac{1}{1-n_1g_1}-1\right) - \frac{g_2}{1-n_1g_1}n_1b_1 &\geq b_2 - g_2C \\
\frac{n_1g_2}{1-n_1g_1}(g_1C-b_1) &\geq b_2 - g_2C. \qquad (15)
\end{aligned}
$$

Now, assume class 1 is compressed in the system with classes $\{1,3,\ldots,K\}$ and capacity $C - n_2b_2$, that is,

$$
\begin{aligned}
\frac{g_1}{1-g_2n_2}(C-n_2b_2) &< b_1 \\
\frac{g_1}{1-g_2n_2}(n_2b_2-C) &> -b_1 \\
\frac{g_1}{1-g_2n_2}(n_2b_2-C) + g_1C &> g_1C - b_1 \\
\frac{n_2g_1}{1-g_2n_2}(b_2-g_2C) &> g_1C - b_1. \qquad (16)
\end{aligned}
$$

From (15) and (16):

$$
\begin{aligned}
g_1C - b_1 < \frac{n_2g_1}{1-n_2g_2}(b_2-g_2C) &\leq \frac{n_2g_1}{1-n_2g_2}\frac{n_1g_2}{1-n_1g_1}(g_1C-b_1) \\
1 &< \frac{n_2g_1}{1-n_2g_2}\frac{n_1g_2}{1-n_1g_1} \\
(1-n_1g_1)(1-n_2g_2) &< n_1n_2g_1g_2 \\
1 &< n_1g_1 + n_2g_2 \leq \sum_{i=1}^{K} n_ig_i = 1
\end{aligned}
$$

which is a contradiction.

5 Discussion

For simplifying the following presentation let us introduce $w_i = n_ig_i$, the bandwidth share of the class $C_i = n_ic_i$.

The bandwidth allocation in the original DPS can also be considered as a mechanism, in which players (the classes of users in our setup) are competing to the resources C, they give payments (also called bids) $w_i(> 0)$ to a central entity (often referred to as resource manager), and this manager chooses an allocation d_i, $i = 1, \ldots, K$ by charging each class (within the class the charge always distributed evenly among the users) by the same price μ (assuming there is no price discrimination between the players). Because we already assumed that the manager tries to allocate the whole capacity of the resources, we have

$$\sum_{j=1}^{K} \frac{w_j}{\mu} = C \tag{17}$$

from which we get for the price

$$\mu = \frac{\sum_{j=1}^{K} w_j}{C} \ . \tag{18}$$

Applying this price for the payment w_i the allocation is

$$C_i = \frac{w_i}{\sum_{j=1}^{N} w_j} C. \tag{19}$$

Note that we also assume here that the players are *price takers*, that is they accept the price and the resulted allocations.

This proportional allocation above is also the solution of the following optimization problem.

$$\max_{\mathbf{x}} \sum_{i=1}^{N} w_i \log x_i \ , \quad \text{s.t.} \ \sum_{i=1}^{K} x_i = C \quad , \quad x_i \geq 0 \tag{20}$$

which is a simple rewrite of (Opt) according to our new quantities w_i and C_i.

The proportional allocation and the price taker characteristic have a much more fundamental property, which can be explained through the notion of *competitive equilibrium* between users and the resource manager [7].

In competitive equilibrium players (classes in our case) $(i = 1, \ldots, K)$ act to maximize their payoff functions $P(w_i, \mu)$ for a given price μ over w_i, which payoff consists of the utility $U_i(\frac{w_i}{\mu})$ by obtaining $\frac{w_i}{\mu}$ resources minus the payment w_i for that amount of resource to the manager, that is $P(w_i, \mu) := U_i(\frac{w_i}{\mu}) - w_i$. A pair $(\mathbf{w}, \mu)$, $\mathbf{w} \geq 0, \mu > 0$ is said to be a competitive equilibrium between the users and the network, if users maximize their payoff functions and the network determines the price as

$$\mu = \frac{\sum_{i=1}^{N} w_i}{C} \ . \tag{21}$$

It can also be shown that if $((\mathbf{w}, \mu))$ forms a competitive equilibrium, then the proportional allocation $C_i = \frac{w_i}{\sum_i w_i} C$ is a solution of the following *utilitarian social welfare*[6] maximization

$$\max_{\mathbf{x}} \sum_{i=1}^{K} U_i(x_i) \ , \quad \text{s.t.} \ \sum_{i=1}^{K} x_i = C \quad , \quad x_i \geq 0 \tag{OptSocial}$$

[6] With a wide class of possible utility functions, for details see e.g. [7]

An important observation here is that the proportional allocation (19) spans a bridge between the individual user payoff maximization and the social welfare maximization, that is the two maximization coincide when proportional allocation applied. This remarkable property of proportional allocation also makes possible to investigate the case when players are *price anticipating*, that is the players take into account that the price is set according to (21) and they maximize their payoff

$$U_i\left(\frac{w_i}{\sum_{k=1}^N w_k}C\right) - w_i \tag{22}$$

with respect to their payments w_i [6].

As we can have seen through Theorem 1, the allocation in case of access rate limited shares (which is partially proportional) is the solution of (OptBounds), which is a quite natural extension of (Opt) by the constraints $x_i \le b_i$. Based on this, we conjecture that the (appropriately modified) competitive equilibrium and the social welfare can be coupled with this allocation in case of limited access to the resource, that is

Conjecture:

The competitive equilibrium $(\mathbf{w}^*, \mu^*)$ exists, in which w_i^* maximizes the payoff function

$$P_i(w_i, \mu^*) = U_i\left(\max(\frac{w_i}{\mu^*}, n_i b_i)\right) - w_i \tag{23}$$

where

$$\mu^* = \frac{\sum_{j=i^*+1}^K w_j}{C - \sum_{k=1}^{i^*} n_k b_k}\ , \tag{24}$$

and the previously calculated allocation (9)

$$C_k = n_k b_k,\ \text{if } k \le i^*\ (k \in \mathcal{U}) \tag{25}$$

and

$$C_k = \frac{w_k}{\sum_{i=i^*+1}^K w_i}\left(C - \sum_{j=1}^{i^*} n_j b_j\right) = \frac{w_k}{\mu^*},\ \text{if} \quad i^* < k\ (k \in \mathcal{Z})\ . \tag{26}$$

is a solution of the following utilitarian social welfare optimization extended by the access limits $x_i \le n_i b_i$.

$$\max_{\mathbf{x}} \sum_{i=1}^K U_i(x_i)\ , \quad \text{s.t.} \quad \sum_{i=1}^K x_i = C \quad , \quad x_i \in [0, n_i b_i] \qquad \text{(OptSocialBounds)}$$

6 Conclusion

In this paper we have proven that the bandwidth shares of the bandwidth-economical access rate limited discriminatory processor sharing is the solution of an extended constrained optimization task. The significance of this result lies in the fact that the DPS-related proportional bandwidth allocation play central role in analysing users in strategic environment when they are selfish and competing for the bandwidth. Our result could be used in analysing allocation games in which the players also have limits on their access to the network resources.

Acknowledgment

This research of the second author was also realized in the frames of TAMOP 4.2.4.A/2-11-1-2012-0001 National Excellence Program Elaborating and operating an inland student and researcher personal support system The project was subsidized by the European Union and co-financed by the European Social Fund.

References

1. Avrachenkov, K., Ayesta, U., Brown, P., Nez-Queija, R.: Discriminatory processor sharing revisited. In: In: Proc. IEEE Infocom 2005, Miami FL. pp. 784–795 (2005)
2. Ayesta, U., Mandjes, M.: Bandwidth-sharing networks under a diffusion scaling. Annals Operation Research 170(1), 41–58 (2009)
3. Cheung, S., van den Berg, H., Boucherie, R.: Decomposing the queue length distribution of processor-sharing models into queue lengths of permanent customer queues. Performance Evaluation 62(1-4), 100–116 (2005)
4. Fan, Z.: Dimensioning Bandwidth for Elastic Traffic. NETWORKING 2002: Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications pp. 826–837 (2006)
5. Fayolle, G., Mitrani, I., Iasnogorodski, R.: Sharing a processor among many job classes. J. ACM 27(3), 519–532 (1980)
6. Johari, R.: Algorithmic Game Theory, chap. The Price of Anarchy and the Design of Scalable Resource Allocation Mechanisms, pp. 543–568. Cambridge Univ. Press (2007)
7. Kelly, F.: Charging and rate control for elastic traffic. European Transactions on Telecommunications 8, 33–37 (1997)
8. van Kessel, G., Nunez-Queija, R., Borst, S.: Asymptotic regimes and approximations for discriminatory processor sharing. ACM SIGMETRICS Performance Evaluation Review 32(2), 44–46 (2004)
9. Kőrösi, A., Székely, B., Vmos, P., Bíró, J.: Characterization of peak-rate limited DPS with Pareto-efficient bandwidth sharing. Annales Mathematicae et Informaticae 39 (2012)
10. Kleinrock, L.: Time-shared systems: A theoretical treatment. J. of ACM 14(2), 242–261 (1967)
11. Lakshmikantha, A., Srikant, R., Beck, C.: Differential equation models of flow-size based priorities in internet routers. International Journal of Systems, Control and Communications 2(1), 170–196 (2010)
12. Lindberger, K.: Balancing quality of service, pricing and utilisation in multiservice networks with stream and elastic traffic. In: ITC-16: International Teletraffic Congress. pp. 1127–1136 (1999)

13. O'Donovan, T.M.: Direct solutions of M/G/1 processor-sharing models. Operations Research 22, 1232–1235 (1974)
14. Pályi, P.L., Kőrösi, A., Székely, B., Bíró, J., Rácz, S.: Characterization of peak-rate limited bandwidth efficient DPS. Acta Polytechnica Hungarica 9 (2012)
15. Rege, K., Sengupta, B.: Queue length distribution for the discriminatory processor-sharing queue. Operation Research 44, 653–657 (1996)
16. Riedl, A., Bauschert, T., Frings, J.: A framework for multi-service IP network planning. In: International Telecommunication Network Strategy and Planning Symposium (Networks). pp. 183–190 (2002)