Pricing for Heterogeneous Services at a Discriminatory Processor Sharing Queue

Yezekael Hayel[1] and Bruno Tuffin[1]

IRISA-INRIA Rennes, Campus universitaire de Beaulieu
35042 Rennes Cedex - FRANCE
{yhayel,btuffin}@irisa.fr
tel: (+33)299847134
fax: (+33)299842529

Abstract. In order to deal with applications with different quality of service requirements, service differentiation has to be implemented, especially in case of congestion. Different scheduling policies can be applied at a queue, such as strict priorities, generalized processor sharing, or discriminatory processor sharing. While prices optimizing the network revenue have been determined in the first two above cases, and the optimal revenue compared, nothing had been done yet on discriminatory processor sharing (DPS). Though, at the session level, processor sharing is known to properly model TCP flows behavior. DPS then models multiple TCP flows at a router providing differentiated services. We study here what pricing induces on a DPS router when two types of application compete for service, what is the resulting equilibrium, and explain how optimal prices can be found.

Key words: Economics, Queueing theory.

1 Introduction

Telecommunication networks such as the Internet are facing congestion due to increasing demand in terms of number of subscribers as well as in terms of more and more demanding applications. For this reason, quality of service (QoS) requirements may not be satisfied by the current best-effort service. To cope with this problem, architectures have been designed based on the idea of treating flows differently. IntServ [1] architecture applies resource reservation but is not scalable. DiffServ architecture [2] has then been designed to tackle this scalability problem. It consists in defining different classes of service (by marking packets with a DiffServ codepoint (DSCP)) and applying a given scheduling policy for serving the packets of the different classes at routers (according to the DSCP).

Surprisingly, no pricing procedure has been associated to the defined architectures, at least in the process of their design. However, considering free services or flat-rate pricing, selfish users would send all their packets to the class providing the "best" quality of service, so that congestion would persist. Pricing for network usage and/or congestion has recently received a lot of attention (see [3] and the references therein), but some of those works are not devoted to multi-class pricing like, among the most noteworthy schemes, auctions for bandwidth (see for instance [4]), or charging for elastic traffic based on transfer rate [5,6]. In the context of multi-class pricing, excepted pricing for guaranteed services [7], most papers deal with pricing for strict priority [8,9] due to its applicability. Nevertheless, priority queueing is not the only scheduling policy likely to be implemented in DiffServ architecture, and the good choice for a provider may be related to the financial benefits that it provides.

Based on this idea, we have compared in a previous paper [10] the respective impact of generalized processor sharing (GPS) and priority queueing (PQ) on the provider's revenue, when the QoS parameter studied is the delay. The model was taken from [9] where priority queueing was considered. Since no closed-form expression is available for the delay at a GPS queue [11], an approximation by separate queues has been used, for which the delay is known. This approximation is valid in the context of congested queues, which is actually the situation where service differentiation is relevant. We have especially proved that, at least for the studied class of utility functions and when modelling the router by an M/M/1 queue, PQ is the policy that always yields the highest revenue, and should then be preferred.

We wish to study in the present paper the impact of another scheduling policy, namely discriminatory processor sharing (DPS), on users' behavior and on the provider's revenue. DPS provides a more flexible way of giving priority than the PQ discipline. It has been introduced by Kleinrock [12] for a single server. It consists in serving classes in proportions controlled by weights, while packets within a given class are served according to a standard processor sharing strategy (which constitutes the main difference with respect to GPS, where packets within a class are served according to a FIFO scheme) [13]. A theoretical charm of DPS is that, unlike GPS, a closed-form solution of steady-state delay exists [14]. DPS has been justified in practice, at the flow level, as a fluid approximation of some weighted round-robin scheme [15]. It has also been shown in [16,17] that a queue with DPS discipline is well adapted, as an approximation, to the way TCP connections can share bandwidth. It uses the fact that, at the session level, TCP can be analyzed using a PS approach [18]. We thus use DPS at the flow level here in order to fit the TCP modelling.

The contribution of the paper is as follows. We consider two types of applications applying for service at two classes of service. In our analysis, we focus on dedicated classes, that is the case where a given application is directed to a given class of service. We consider that the two types of application both use TCP traffic, so that a DPS queue can be used to approach the system behavior. The modelling of TCP constitutes one of the main enhancements with respect to [9,10]. We are going to investigate the impact of DPS on the pricing model. Since, with respect to PQ and the approximation of GPS, delay over classes are *mutually* dependent, a deeper analysis of the model at the infinitesimal level is required. We show that, for fixed prices, there always exists a unique equilibrium for the average number of sessions of each type competing for service. This equilibrium is then a so-called Nash equilibrium often used in micro-economics studies [3], where no user/application class has any incentive to unilaterally deviate from its current policy. We show that, depending on the price values, we may have undesirable situations with only one class of traffic asking for service; a necessary and sufficient condition for avoiding that is provided. We then describe how optimal prices can be found and illustrate that PQ is (still) the border case maximizing the provider's revenue among all discriminatory processor sharing parameter choices.

The rest of the paper is organized as follows. In Section 2, we present the model, mainly taken from [9], and adapt it to TCP with DPS scheduling policy. Section 3 then discusses in details the equilibrium that can obtained depending on the prices selection. Section 4 discusses how optimal prices can be found and Section 5 illustrates numerically the previous results. Finally, Section 6 concludes the paper and gives some perspectives of research.

2 Model

Our general model closely follows [9]. We assume that service is differentiated by two classes of traffic. We also assume that there are two types of users/applications, delay-tolerant users that we will abusively (but for sake of simplicity) call data users, and delay-sensitive users, that we will call voice-users. We keep the notations voice and data to be consistent with [9, 10], but, especially as we consider TCP traffic, voice can/should be replaced by any application with low delay preferences. In this paper, we will force data users to one class and voice users to the other class. This is called the case of *dedicated* classes in [9]. The case of *open* classes where the users of each application can choose between classes is left for future work.

We consider a system with infinite population of potential customers applying for service as soon as the *cost* of sending packets is less than the benefits that they get from it. The benefits of sending packets are expressed by *utility* functions, $u_d(\cdot)$ for data and $u_v(\cdot)$ for voice, which depend on the mean delay in each class respectively denoted by $\mathbb{E}D_d$ and $\mathbb{E}D_v$. We consider the utility as a function of the average delay rather than the average utility of instantaneous delay since average delay seems to us the measure of interest, especially when dealing with data for instance (even for voice, average delay seems more relevant provided that an upper-threshold is not reached). Like in [9, 10], we consider utility functions $u_d(y) = y^{-\alpha_d}$ and $u_v(y) = y^{-\alpha_v}$ with $0 < \alpha_d < \alpha_v$, so that voice users value low delay more than data users, but conversely value high delay less. This choice of utility function is somewhat arbitrary, but is based on the idea that utility curves do intersect [9]: low delay is more valuable to voice than data, and conversely large delay is more valuable to data. Note also that this form of utility function can be related to the Cobb-Douglas function widely used in micro-economic analysis [19].

Let p_d and p_v be the per-packet price for access at respectively data and voice classes. The *residual utility* of a data user (resp. voice user) is given by

$$u_d(\mathbb{E}D_d) - p_d \quad (\text{resp.} \quad u_v(\mathbb{E}D_v) - p_v). \tag{1}$$

Sessions open as Poisson arrival processes. Let N_d and N_v be the mean number of data and voice users sending packets in steady-state. Let also λ_d and λ_v be the average size of data and voice flows (drawn independently between flows). In [9, 10], N_d and N_v were the *fixed* numbers of sources while λ_d and λ_v were the rates at which packets were sent. The model is then here slightly different in order to fit the usual assumptions for modelling TCP traffic [16–18] by using processor sharing. Even if using λ for the average size is unusual, we keep the notation here to be consistent with the notations in [9, 10]. As we consider Poisson flow arrivals, the number of flows in progress behaves like the number of customers in an M/M/1 processor sharing queue [12]. Moreover, as we assume two different class of traffic, the model behaves like an M/M/1 discriminatory processor sharing queue. Let μ be the service rate of the server/router.

We consider a discriminatory processor sharing (DPS) scheduling policy for differentiating services. DPS basically works as follows. There exists a nonnegative parameter γ representing *relative priority* of data customers and $1-\gamma$ for voice customers. Still, when packets of one class are not present in the queue, the server is fully allocated to the other class, but flows within a class are served according to a processor sharing (PS) scheme. A closed-form formula for the average delays in such M/M/1 queues are given in [20, page 86] by

$$\mathbb{E}D_v = \frac{\left(1 + \frac{\lambda_d N_d (2\gamma - 1)}{\mu - (1-\gamma)\lambda_v N_v - \gamma \lambda_d N_d}\right)}{\mu - \lambda_v N_v - \lambda_d N_d} \quad \text{and} \quad \mathbb{E}D_d = \frac{\left(1 - \frac{\lambda_v N_v (2\gamma - 1)}{\mu - (1-\gamma)\lambda_v N_v - \gamma \lambda_d N_d}\right)}{\mu - \lambda_v N_v - \lambda_d N_d}. \tag{2}$$

Remark that PQ (but with PS discipline within a class) is a special case of DPS, since $\gamma = 0$ gives strict priority to voice users, leading to expression of delay:

$$\mathbb{E}D_v = \frac{1}{\mu - \lambda_v N_v} \quad \text{and} \quad \mathbb{E}D_d = \frac{\mu}{(\mu - \lambda_v N_v)(\mu - \lambda_v N_v - \lambda_d N_d)},$$

while $\gamma = 1$ would give strict priority to data (note that the mean delay is the same for M/M/1/FIFO and M/M/1/PS queues).

3 Equilibrium analysis

Whereas determining the (mean) number of users of each type in equilibrium is quite easy for PQ and an approximation of GPS where the queues are logically separated, the analysis is much more intricate for DPS due to the influence of the number of users of one type on the delay of the other type of users, and vice versa.

The steady-state average numbers of sessions of each type has to ensure that mean residual utilities (1), are positive or null, otherwise the number of sessions naturally decreases. Following the same line, this average number increases until the residual utility approaches 0. It means that each type of application naturally adapts its steady-state number of sources in the sense the maximum (mean) number of sources of a given type cannot make negative its residual utility. We hence have a game between the different types of applications, for the maximum mean number of sessions, potentially leading to a Nash equilibrium. Let us now investigate the existence and uniqueness of this equilibrium.

Let

$$U_d(N_d, N_v) = u_d(\mathbb{E}D_d) - p_d \quad \text{and} \quad U_v(N_d, N_v) = u_v(\mathbb{E}D_v) - p_v$$

be the mean residual utilities, where we emphasize the role of the number of connections of each type. We obtain that N_v and N_d must verify

- If $N_d, N_v > 0$, then $U_d(N_d, N_v) = U_v(N_d, N_v) = 0$,
- If $N_d > 0,\ N_v = 0$, then $U_d(N_d, N_v) = 0,\ U_v(N_d, N_v) \leq 0$,
- If $N_d = 0,\ N_v > 0$, then $U_d(N_d, N_v) \leq 0,\ U_v(N_d, N_v) = 0$.

The last two relations mean that the type of application such that $N = 0$ has no incentive to open sessions since its mean residual utility is negative.

Figure 1 illustrates these equations, where the maximum number of customers of each type in the network is necessarily on the "minimum" of curves $U_d(N_d, N_v) = 0$ and $U_v(N_d, N_v) = 0$. Indeed, the mean number of sources of each type increases, decreasing then the utilities, until a residual utility reaches zero. Figure 1 depicts the three situations that will be used later on: either the

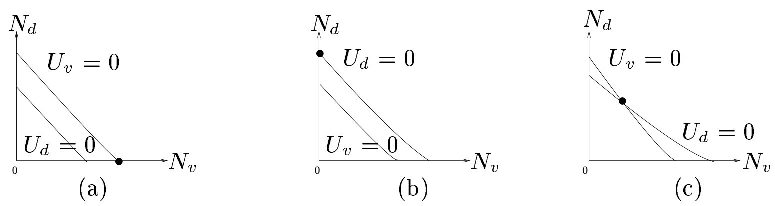


Fig. 1. Curves for the (maximum) number of customers of each type, and the associated Nash equilibrium. The three possible cases are displayed. The resulting Nash equilibrium is displayed by the thick point.

curves cross each other in the domain $\{(N_d, N_v) : N_d, N_v \geq 0\}$, or one curve is always under the other in this domain.

Lemma 1. *Due to the form of utility functions, the curves* $U_d(N_d, N_v) = 0$ *and* $U_v(N_d, N_v) = 0$ *either are one above the other or cross each other only once on* $N_d, N_v \geq 0$.

This lemma is proved in Appendix 1.

Wherever this (minimal) curve such that one type of application has reached his null (residual) utility, the system may continue to evolve since if the utility of the other class is positive, it will continue to increase its number of source, increasing then the delay of the first class (see Equations (2)), so that its utility will become negative, meaning that the number of sources will have to be reduced.

To simplify the notations, let $q_d = (p_d)^{1/\alpha_d}$ and $q_v = (p_v)^{1/\alpha_v}$. The following proposition establishes the existence of a unique Nash equilibrium, and gives an explicit solution in terms of q_d and q_v.

Proposition 1. *Consider fixed prices p_d and p_v. There exists a unique (Nash) equilibrium (N_d^*, N_v^*) that stabilizes the mean number of emitting sources.*

- *If the curve $U_d(N_d, N_v) = 0$ is always under $U_v(N_d, N_v) = 0$ on $N_v, N_d \geq 0$ (case (a) of Figure 1), the Nash equilibrium is given by*

$$(N_d^*, N_v^*) = \left(0, \frac{\mu - q_v}{\lambda_v}\right). \tag{3}$$

- *If the curve $U_v(N_d, N_v) = 0$ is always under $U_d(N_d, N_v) = 0$ on $N_v, N_d \geq 0$ (case (b) of Figure 1), the Nash equilibrium is given by*

$$(N_d^*, N_v^*) = \left(\frac{\mu - q_d}{\lambda_d}, 0\right). \tag{4}$$

- *If the curves have a crossing point (case (c) of Figure 1), it is unique, and its value is a Nash equilibrium $(N_d^*, N_v^*) > 0$, with*

$$\lambda_d N_d^* = \frac{\mu q_d}{\gamma q_d - (1-\gamma) q_v} - \frac{q_v q_d}{\gamma q_v - (1-\gamma) q_d}$$
$$\lambda_v N_v^* = -\frac{\mu q_v}{\gamma q_d - (1-\gamma) q_v} + \frac{q_v q_d}{\gamma q_v - (1-\gamma) q_d}.$$

We will call this last case the non-trivial equilibrium.

The proof of this proposition is given in Appendix 2.

An important remark is that the two border equilibriums (3) and (4) correspond to the cases where only one class of traffic eventually requests service. It then means that the model is a queue with one class of service, and no real service differentiation is applied (except that the other class is rejected). As a consequence, in order to obtain a gain by differentiating service among users, we have to look at prices under which the above border equilibriums are not reached.

The following proposition gives a necessary and sufficient condition over prices for being in a non-trivial equilibrium.

Proposition 2. *The Nash equilibrium is non-trivial if and only if q_v and q_d satisfy*

$$0 < q_v < \mu, \tag{5}$$

$$\sqrt{\mu q_v + \left(\frac{\mu - q_v}{2}\frac{1-\gamma}{\gamma}\right)^2} - \frac{\mu - q_v}{2}\frac{1-\gamma}{\gamma} < q_d, \tag{6}$$

$$\frac{\mu\gamma q_v + (1-\gamma) q_v^2}{\gamma q_v + \mu(1-\gamma)} > q_d. \tag{7}$$

The proof of this proposition is given in Appendix 3.

Remark: It can be verified with straightforward calculus that the traffic load ρ is less than one (which is the stability condition for the queue).

4 On optimal prices and revenues

The idea, from a network provider's perspective, is to find the p_d^*, p_v^* and γ^* giving the maximum network revenue

$$R^* = \max_{p_d,p_v,\gamma} \lambda_d N_d^*(p_d,p_v,\gamma)p_d + \lambda_v N_v^*(p_d,p_v,\gamma)p_v$$

subject to $p_v, p_d \geq 0$ and $\gamma \in [0,1]$.

It is possible to analytically determine the optimal γ for fixed values of p_d and p_v, but optimizing then in terms of p_d and p_v looks analytically intractable, so that the determination of γ is not helpful (particularly since it requires a decomposition in several cases).

In the next section we present numerical results illustrating the domains for border or non-trivial equilibrium, as well as the impact of prices and bandwidth partition on the network provider's revenue. The prices optimizing the revenue are also numerically determined.

5 Numerical illustrations

In this section we will present some numerical results obtained with the above model[1]. Unless stated otherwise, the following parameters will be used throughout this section: $\mu = 1$, $\alpha_v = 1.8$ and $\alpha_d = 1.5$, $\lambda_v = 0.1$ and $\lambda_d = 0.3$.

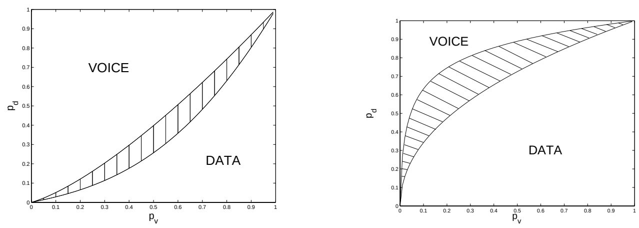
5.1 Price domains and associated equilibrium

Figure 2 illustrates the equilibrium domains in terms of prices for two values of γ, say $\gamma = 0.9$ and $\gamma = 0.2$. We have the three possible areas: VOICE corresponding to the (border) equilibrium where there is only voice traffic, DATA corresponding to the case where there is only data traffic, and the hatched area corresponding to a non-trivial equilibrium. Note that when γ is less than 0.5, the zone for non-trivial equilibrium is such that $p_v^{1/\alpha_v} > p_d^{1/\alpha_d}$, and that the converse is also true.

5.2 Optimal γ for fixed prices

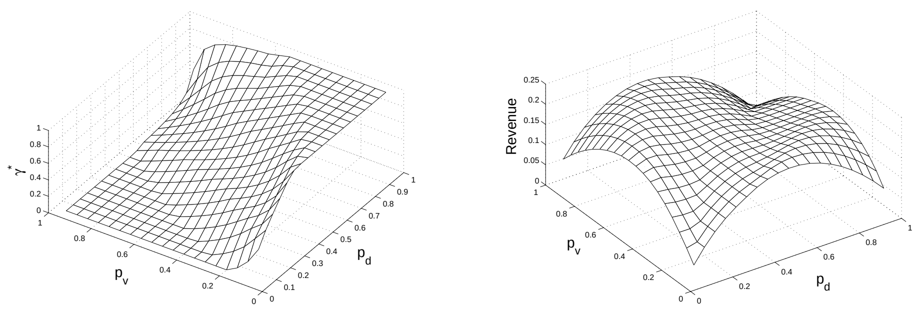
For given prices p_v and p_d, it might be interesting to look at the bandwidth-sharing parameter γ optimizing the revenue. This value is shown on Figure 3(a). It can be observed that when p_v is large and p_d is small, the optimal revenue is obtained when $\gamma^* = 0$, i.e. a strict priority is assigned to voice traffic, and conversely when p_v is low and p_d is large, $\gamma^* = 1$, meaning that a strict priority should be applied to data traffic.

[1] The standard numerical and optimization packages of any mathematical software can be used to solve this problem.



(a) $\gamma = 0.2$ (b) $\gamma = 0.9$

Fig. 2. Equilibrium domains with different values of parameter γ and varying prices.



(a) Optimal bandwidth-sharing (b) Provider's revenue

Fig. 3. Optimal bandwidth-sharing and provider's revenue in terms of prices.

5.3 Revenue

Figure 3(b) shows the revenue in terms of prices, using the optimal value of γ used for each couple (p_d, p_v). Since the curve of revenue is smooth, obtaining the optimal prices is numerically easy. For this example, we obtain the optimal prices $p_d^* = 0.24$ and $p_v^* = 0.68$, giving an optimal revenue $R^* = R(p_d^*, p_v^*, 0) = 0.21$.

It is important to note that, at optimal prices, the optimal γ is then 0, giving then strict priority to voice traffic. Actually, whatever the choices of parameters we have, this result has been verified by extensive simulations that cannot be reported here for lack of room. This is somehow in accordance with [10] where we have shown for a simple M/M/1 queue that PQ has to be preferred to GPS to optimize the revenue. In the present case, a formal proof of this conjecture has still to be found.

6 Conclusions and perspectives

We have studied in this paper the impact of pricing on a multi-class queue with DPS scheduling policy (known to properly model service differentiation between TCP flows). We have shown, in the case of two types of applications and of dedicated classes that the game between both types of traffic for setting their mean number of emitting sources always results in a unique Nash equilibrium. The location of this equilibrium depends on prices. We have determined three domains of prices for which either only voice requests service, either only data requests service or for which traffic is mixed (the so-called non-trivial equilibrium). We have numerically investigated the price selection in order to optimize the network revenue and illustrated that PQ does actually seem to always provide a higher revenue.

As directions for future work, we wish to prove theoretically that PQ is the limit case of DPS always providing the highest revenue, and should then be preferred, as done in [10] for GPS. Studying the behavior of other scheduling policies is also a topic of interest, in order to define the best strategy for optimizing the network's revenue. We also wish to compare/verify our results with ns-2 simulations.

References

1. Braden, R., Clark, D., Shenker, S.: Integrated Services in the Internet Architecture: an Overview. Technical report, RFC 1663 IETF (1994)
2. Wang, Z.: Internet QoS: Architectures and Mechanisms for Quality of Service. Morgan Kaufmann Publishers (2001)
3. Courcoubetis, C., Weber, R.: Pricing Communication Networks—Economics, Technology and Modelling. Wiley (2003)
4. Maillé, P., Tuffin, B.: Multi-bid auctions for bandwidth allocation in communication networks. In: Proceedings of IEEE INFOCOM. (2004)
5. Kelly, F., Mauloo, A., Tan, D.: Rate control in communication networks: shadow prices, proportional fairness and stability. Journal of the Operational Research Society **49** (1998) 237–252
6. Low, S., Lapsley, D.: Optimization Flow Control, I: Basic Algorithm and Convergence. IEEE/ACM Transactions on Networking **7** (1999)
7. Paschalidis, I., Liu, Y.: Pricing in Multiservices Loss Networks: Static Pricing, Asymptotic Optimality, and Demand Substitution Effects. IEEE/ACM TON **10** (2002)
8. Cocchi, R., Estrin, D., Shenker, S., Zhang, L.: Pricing in Computer Networks: Motivation, Formulation and Example. IEEE/ACM TON **1** (1993) 614–627
9. Mandjes, M.: Pricing Strategies under Heterogeneous Service Requirements. In: Proceedings of IEEE INFOCOM. (2003)
10. Hayel, Y., Ros, D., Tuffin, B.: Less-than-Best-Effort Services: Pricing and Scheduling. In: Proceedings of IEEE INFOCOM. (2004)
11. Borst, S., Boxma, O., Núnez Queija, R.: Heavy Tails: The Effect of the Service Discipline. In: Proceedings of Performance TOOLS. (2002)
12. Kleinrock, L.: Time-shared Systems: A Theoretical Treatment. Journal of the Association for Computing Machinery **14** (1967)

13. Rege, K., Sengupta, B.: Queue-Length Distribution for the Discriminatory Processor-Sharing Queue. Operations Research **44** (1996)
14. Fayolle, G., Mitrani, I., Iasnogorodski, R.: Sharing a Processor Among Many Job Classes. Journal of the Association for Computing Machinery **27** (1980)
15. Mitra, D., Weiss, A.: A Closed Network with a Discriminatory Processor-Sharing Server. In: Proceedings of ACM Sigmetrics. (1989)
16. Altman, E., Jimenez, T., Kofman, D.: DPS Queues with Stationary Ergodic Service Times and the Performance of TCP in Overload. In: Proceedings of IEEE INFOCOM. (2004)
17. Bu, T., Towsley, D.: Fixed point approximations for TCP behavior in an AQM network. In: Proceedings of ACM Sigmetrics. (2001)
18. Ben Fredj, S., Bonald, T., Proutiere, A., Regnie, G. Roberts, J.: Statistical Bandwidth Sharing: a Study of Congestion at Flow Level. In: Proceedings of IEEE SIGCOMM'. (2001)
19. Beckert, W.: Stochastic Demand Analysis. PhD thesis, UC Berkeley (2000)
20. Hassin, R., Haviv, M.: To Queue or Not To Queue. Kluwer's (2003)

Appendix

1 Proof of Lemma 1

Look at the non-linear system defined by $U_v(x,y) = U_d(x,y) = 0$ and solve it over $\mathbb{R}^2$. It can be rewritten as

$$\begin{cases} \frac{\mu - x\lambda_v - y\lambda_d}{q_v} = 1 + \frac{y\lambda_d(2\gamma-1)}{\mu-(1-\gamma)x\lambda_v - y\gamma\lambda_d} \\ \frac{\mu - x\lambda_v - y\lambda_d}{q_d} = 1 - \frac{x\lambda_v(2\gamma-1)}{\mu-(1-\gamma)x\lambda_v - y\gamma\lambda_d} \end{cases}$$

The unique solution (x^*, y^*) is:

$$x^* = \left(\frac{\mu q_d}{\gamma q_d - (1-\gamma)q_v} - \frac{q_v q_d}{\gamma q_v - (1-\gamma)q_d}\right)\frac{1}{\lambda_d},$$

$$y^* = \left(-\frac{\mu q_v}{\gamma q_d - (1-\gamma)q_v} + \frac{q_v q_d}{\gamma q_v - (1-\gamma)q_d}\right)\frac{1}{\lambda_v}.$$

Equations $U_d(x,y) = 0$ and $U_v(x,y) = 0$ both actually give y as a function of x since to each x corresponds only one y. These functions are continuous so that either one curve is always above the other on the possible domain of (N_d, N_v), either the curves cross each other only once from the unique solution of the system. □

2 Proof of Proposition 1

Consider the dynamics of the queue for the three possible curves configurations.

- If the curve $U_d(N_d, N_v) = 0$ is always under $U_v(N_d, N_v) = 0$ on $N_v, N_d \geq 0$ (with the condition that the queue is ergodic), (case (a) of Figure 1), the application types will increase their mean number of emitting sessions until a residual utility function approaches 0, that is U_d in the present case. The system continues then to evolve and only voice traffic increases its number of sources because its utility function is still strictly positive. This implies that the number of data sessions will naturally decrease as its utility becomes negative. The system evolves like that until $U_v(N_d, N_v)$ also reaches 0 or $N_d = 0$. Since we have assumed that the curves do not cross, $N_d = 0$ first. Then N_v increases until we reach the equilibrium $(N_d^*, N_v^*) = (0, \frac{\mu - q_v}{\lambda_v})$ such that $U_d < 0$ and $U_v = 0$ so that no type of application has an incentive to unilaterally change its mean number of emitting sources.
- If the curve $U_d(N_d, N_v) = 0$ is always above $U_v(N_d, N_v) = 0$ (case (b) of Figure 1), we obtain the symmetric case of the previous one, leading to the equilibrium $(N_d^*, N_v^*) = (\frac{\mu - q_d}{\lambda_d}, 0)$.
- If there is a crossing point (N_d^*, N_v^*) for the curves (case (c) of Figure 1), this crossing point is unique from Lemma 1. Note that $U_v(0, N_v) = 0$ for $N_v = \frac{\mu - q_v}{\lambda_v}$, $U_d(N_d, 0) = 0$ for $N_d = \frac{\mu - q_d}{\lambda_d}$ and that $U_d(0, \frac{\mu - q_v}{\lambda_v}) > 0$ and $U_v(\frac{\mu - q_d}{\lambda_d}, 0) > 0$.
  Again, the mean number of emitting sessions will (continuously) increase until one curve $U_d = 0$ or $U_v = 0$ is reached. Denote by $(\tilde{N}_d, \tilde{N}_v)$ the mean numbers of sources when such a point is reached. We can have two different cases:
  - If the first utility function to become null is U_d, then we have $U_d(\tilde{N}_d, \tilde{N}_v) = 0$ and $U_v(\tilde{N}_d, \tilde{N}_v) > 0$ so that N_v will increase, implying that N_d will decrease. Since $\tilde{N}_d > N_d^*$ and $\tilde{N}_v < N_v^*$, it will evolve this way until we eventually reach the point (N_d^*, N_v^*) where $U_d(N_d^*, N_v^*) = 0$ and $U_v(N_d^*, N_v^*) = 0$.
  - If the first utility function to become null is U_v, we can exchange the roles of data and voice, and we arrive to the same conclusion.

□

3 Proof of Proposition 2

In order to prove this proposition, we first prove that the condition is sufficient and then that it is necessary.

- To prove that the condition expressed by (5), (6) and (7) is sufficient, assume that q_d and q_v satisfy those relations.
  After some calculations, (6) can be rewritten as

$$\mu\gamma \frac{q_v}{q_d} - (1-\gamma)\mu < \gamma q_d - (1-\gamma) q_v, \tag{8}$$

  and (7) can be rewritten as

$$\mu\gamma - \mu(1-\gamma)\frac{q_d}{q_v} > \gamma q_d - (1-\gamma) q_v. \tag{9}$$

Indeed, from inequality (7),

$$\begin{aligned} q_d &< \frac{\mu\gamma q_v + (1-\gamma)q_v^2}{\gamma q_v + \mu(1-\gamma)}, \\ \frac{q_d}{q_v} &< \frac{\mu\gamma + (1-\gamma)q_v}{\gamma q_v + \mu(1-\gamma)}, \\ \gamma q_d - (1-\gamma)q_v &< \mu\gamma - \mu(1-\gamma)\frac{q_d}{q_v}. \end{aligned} \tag{10}$$

Also, from inequality (6),

$$\begin{aligned} \sqrt{\mu q_v + (\frac{\mu - q_v}{2}\frac{1-\gamma}{\gamma})^2} - \frac{\mu - q_v}{2}\frac{1-\gamma}{\gamma} &< q_d, \\ \gamma q_d^2 - q_d q_v(1-\gamma) &> \gamma\mu q_v - q_d\mu(1-\gamma), \\ \gamma q_d - q_v(1-\gamma) &> \gamma\mu\frac{q_v}{q_d} - \mu(1-\gamma). \end{aligned} \tag{11}$$

From (9), we obtain $q_v(\gamma q_d - (1-\gamma)q_v) < \mu\gamma q_v - (1-\gamma)\mu q_d$, and from (8) $\mu\gamma q_v - (1-\gamma)\mu q_d < q_d(\gamma q_d - (1-\gamma)q_v)$.
Our goal is to show that we are in a non-trivial equilibrium, meaning that the (unique) solution of the system $U_d(N_d, N_v) = 0$ and $U_v(N_d, N_v) = 0$ is such that $N_d > 0$ and $N_v > 0$. ¿From the proof of Lemma 1, the solution of the system verifies $\lambda_d N_d = \frac{\mu q_d}{\gamma q_d - (1-\gamma)q_v} - \frac{q_v q_d}{\gamma q_v - (1-\gamma)q_d}$, and $\lambda_v N_v = -\frac{\mu q_v}{\gamma q_d - (1-\gamma)q_v} + \frac{q_v q_d}{\gamma q_v - (1-\gamma)q_d}$. To prove that those values are strictly positive, consider the two complementary cases.

- If $\gamma q_d - (1-\gamma)q_v > 0$, then we have $\gamma q_v - (1-\gamma)q_d > 0$ by a simple manipulation. We deduce from Equation (9) that $\mu > q_v\frac{\gamma q_d - (1-\gamma)q_v}{\gamma q_v - (1-\gamma)q_d}$, which gives that $\lambda_d N_d > 0$. Moreover, we have from Equation (8) that $\mu < q_d\frac{\gamma q_d - (1-\gamma)q_v}{\gamma q_v - (1-\gamma)q_d}$, which gives that $\lambda_v N_v > 0$.
- Similarly, assume $\gamma q_d - (1-\gamma)q_v < 0$. We obtain that $\gamma q_v - (1-\gamma)q_d < 0$ by a simple manipulation. Then $\mu > q_v\frac{\gamma q_d - (1-\gamma)q_v}{\gamma q_v - (1-\gamma)q_d}$, from (9), which gives that $\lambda_d N_d > 0$. Moreover, (8) also gives $\mu < q_d\frac{\gamma q_d - (1-\gamma)q_v}{\gamma q_v - (1-\gamma)q_d}$, from which $\lambda_v N_v > 0$.

The sufficient condition is then proved.

– Second, we prove that the condition is necessary, i.e. that assuming the Nash equilibrium non-trivial, Relations (5), (6) and (7) are verified. N_d^* and N_v^* satisfy $\lambda_d N_d^* = \frac{\mu q_d}{\gamma q_d - (1-\gamma)q_v} - \frac{q_v q_d}{\gamma q_v - (1-\gamma)q_d}$, and $\lambda_v N_v^* = -\frac{\mu q_v}{\gamma q_d - (1-\gamma)q_v} + \frac{q_v q_d}{\gamma q_v - (1-\gamma)q_d}$. As both are non-negative, we obtain: $\frac{q_v}{\gamma q_v - (1-\gamma)q_d} < \frac{\mu}{\gamma q_v - (1-\gamma)q_d}$, and $\frac{\mu}{\gamma q_v - (1-\gamma)q_d} < \frac{q_d}{\gamma q_v - (1-\gamma)q_d}$, which is equivalent to: $\mu\gamma\frac{q_v}{q_d} - (1-\gamma)\mu < \gamma q_d - (1-\gamma)q_v$, and $< \mu\gamma - \mu(1-\gamma)\frac{q_d}{q_v}$. From (10), condition (7) can be rewritten as: $\frac{\mu\gamma + (1-\gamma)q_d}{\gamma q_d + \mu(1-\gamma)} < \frac{q_d}{q_v}$, and from (11), condition (6) becomes $\frac{q_d}{q_v} < \frac{\mu\gamma + (1-\gamma)q_v}{\gamma q_v + \mu(1-\gamma)}$. These inequalities give the two relations between q_d and q_v: $q_d < \frac{\mu\gamma q_v + (1-\gamma)q_v^2}{\gamma q_v + \mu(1-\gamma)}$, and $q_d > \sqrt{\mu q_v + (\frac{\mu - q_v}{2}\frac{1-\gamma}{\gamma})^2} - \frac{\mu - q_v}{2}\frac{1-\gamma}{\gamma}$ by solving the second order polynomial.
Moreover, $q_v > \mu$ is not possible since $N_v^* < \frac{\mu - q_v}{\lambda_v} < 0$. □