Numerical Key Results for Studying the Performance of TCP Traffic under Class-Based Weighted Fair Queuing System

Mohamed El Hedi Boussada[1], Jean Marie Garcia[2] and Mounir Frikha[1]

1. Mobile Network and Multimedia, SUP'COM, University of Carthage, Tunis, Tunisia

2. Services and Architectures for Advanced Networks, LAAS-CNRS, Toulouse, France

Abstract: This paper presents a fluid model to estimate the performances of TCP traffics under a two CBWFQ (class-based weighted fair queuing) with variable rate. The system is then equivalent to two virtual links with variable capacities. At first, we give some results to evaluate the performance of TCP traffics under "Best Effort" architecture, then we try to exploit these results to provide key results to qualify performances characteristics of TCP traffic under CBWFQ system. The core of our analysis is based on some numerical observations and relies on the conservation of the average total number of flows carried by the system. Detailed packet level simulations of TCP flows show the accuracy of our analysis. The results presented in this paper allow a rapid performance evaluation of TCP traffic circulating in the actual IP networks.

Key words: CBWFQ (class-based weighted fair queues), TCP traffic, fluid model, performance evaluation.

1. Introduction

Internet traffic is currently dominated by TCP controlled data transfers [1]. This traffic is elastic in the sense that the duration of each transfer depends on the state of the network [2]. Each document is transported by a flow of packets, whose rate is adapted, under the control of TCP protocol, according to the level of congestion in the network [3].

The quality of the transfer depends then on the time necessary to successfully transfer all packets of a flow. In this sense, the performance of elastic traffic is principally studied in flow level and can be translated by the average throughput of each flow [4].

Although this traffic does not have strict constraints in terms of time and bandwidth [2], operators and service providers generally offer all services at different service levels according to the QoS (quality of service) levels required by the client. For example

the SLA (service level agreement) created between the operator and the client is presented in three classes, bronze, silver, and gold levels with increasingly better response times. A service level is used to define the expected performance behavior of a service, where the performance metrics are, for example, average response time, supported throughput, service availability, etc. [5]. Then, operators have to use some congestion management techniques to provide the differentiation between these different service levels during high congestion periods. Congestion management entails the creation of queues, the assignment of packets to those queues, and the scheduling of packets in these queues for transmission [5].

The CBWFQ (class-based weighted fair queuing) is a scheduling discipline usually applied to QoS enabled routers [6]. The Class Based Weighted Fair Queuing is an advanced form of WFQ (weighted fair queuing) that supports user defined traffic classes. A queue is allocated for each class, and the traffic belonging to

Corresponding author: Mohamed El Hedi Boussada, Ph. D., research fields: large-scale modeling of telecommunications traffic in next generation networks.

that class is directed to the queue for that class [7]. The CBWFQ dynamically allocates the available bandwidth to each traffic class based on the class's weight [8].

This paper presents an analytical key result to evaluate and qualify performance characteristics of elastic traffics under a CBWFQ system. In the next section, we present our model. In section 3 we give useful results applying to a network whose resources are dedicated for elastic traffic only. Section 4 is devoted to presenting our approximations to evaluate the performance of elastic traffic carried by two CBWFQ queues. Some results are presented and validated with NS2 simulations in section 5.

2. Model

We consider a single link with capacity C (Mbits/Seconde) shared by a random number of elastic flow classes. Let E be the set of these elastic flow classes. Class-i flows, $i \in E$, arrive as an independent Poisson process with rate λ_i (Flows/Second) and have independent, exponentially distributed volumes with means σ_i (Mbits/flow). We refer to the product $\rho_i = \lambda_i \sigma_i$ (Mbits/Seconde) as the traffic intensity of class i. Let $\theta = \sum_{i \in E} \rho_i$ be the total traffic volume generated by all elastic flows. We assume that $\theta < C$ to ensure the stability of our system.

Each flow of a class i has a maximum bit rate $d_i \le C$. This is the actual rate of each flow in the absence of congestion, it means when $\sum_{i \in E} x_i d_i \le C$, with x_i is the number of class-i flows. Congestion forces TCP to reduce flow rates and thus to increase their duration. We refer to the vector $x = (x_i)_{i \in E}$ as the network state.

The evolution of the system state defines a multidimensional Markov process with transition rates λ_i from state x to state $x + e_i$ and $d_i(x)/\sigma_i$ from state x to state $x - e_i$ (provided $x_i > 0$), where $d_i(x)$ is the bit rate of class-i flows in a state x: $d_i(x) \le d_i$. In this paper, we will suppose that all

classes have the same maximum bit rate $d_i = d$ for all $i \in E$.

Users perceive performance essentially through the mean time necessary to transfer a document [3]. In the following, we evaluate performance in terms of throughput, defined as the ratio of the mean flow size to the mean flow duration in steady state. Assuming network stability and applying Little's formula, the throughput of a flow for any class $i \in E$ is related to the expected number of class-i flows in steady state, $(E[x_i])$, through the relationship [2]:

$$\gamma_i = \frac{\rho_i}{E[x_i]} \qquad (1)$$

3. Performance Evaluation of TCP Traffic under "Best Effort" Architecture

Let $N = \lfloor C/d \rfloor$ be the maximum number of flows that can be allocated with exactly d units on the link. Above this limit, congestion likely occurs and flows equally share the link capacity C. Our system will be identical to a "Processor sharing" queue. As we showed in Ref. [2], for a single class case (with traffic intensity ρ), the arrivals and departure of flows can be modeled by a birth-death process. The stationary distribution of this Markov process is given by:

$$\pi(x) = \begin{cases} \frac{\left(\frac{\rho}{d}\right)^x}{x!} \pi(0) \text{ if } x \le N \\ \frac{\rho^x}{d^N C^{x-N} N!} \pi(0) \text{ else} \end{cases} \qquad (2)$$

where x is the number of ongoing flows in steady state, and $\pi(0)$ is the probability that the link is empty:

$$\pi(0) = \left(\sum_{x=0}^{N-1} \frac{\left(\frac{\rho}{d}\right)^x}{x!} + \frac{\left(\frac{\rho}{d}\right)^N}{N!} \frac{C}{C-\rho} \right)^{-1} \qquad (3)$$

Therefore:

$$E[x] = \frac{\rho}{d} + \frac{\left(\frac{\rho}{d}\right)^N}{N!} \frac{C}{C-\rho} \frac{\rho}{C-\rho} \pi(0) \qquad (4)$$

Let B be the congestion probability on the link. It is written as follows:

$$B = \Pr[x \ge N] = \frac{\left(\frac{\rho}{d}\right)^N}{N!} \frac{C}{C-\rho} \pi(0) \qquad (5)$$

Numerical Key Results for Studying the Performance of TCP Traffic
under Class-Based Weighted Fair Queuing System

197

Thus:

$$E[x] = \frac{\rho}{d} + B\frac{\rho}{C-\rho} \qquad (6)$$

In the case of many classes with identical maximum bit rate, an aggregation of all flows can be done to have a single class with total load θ. The average number of flows of this class is obtained using Eq. (6) by replacing ρ with θ. $E[x_i]$ is then given as follows [1, 9]:

$$E[x_i] = \frac{\rho_i}{\theta} E[x] = \frac{\rho_i}{d} + B\frac{\rho_i}{C-\theta} \qquad (7)$$

4. Studying the CBWFQ Case

In a similar way to the configuration of Internet routers, we assume in this section that packets are affected to two CBWFQ queues before being sent into the link. Let ϑ_m, $m = 1, 2$, the weight of the CBWFQ queue number m. We assume that $\vartheta_1 + \vartheta_2 = 1$. The link capacity C can be viewed as a concatenation between two virtual links. We note by E_m the set of elastic flow classes traversing the virtual link m and by $\theta_m = \sum_{i \in E_m} \rho_i$ the load offered to this virtual link.

The capacity $C_m^* = \vartheta_m C$ is mainly dedicated to the flow classes of E_m, but if a part of this capacity becomes available, it will be shared between the other virtual links which need more resources to transmit its traffic in a better condition. Therefore, C_m^* can be viewed as the minimum capacity allocated to the virtual link m. This approach of coupling between the queues is adopted by many authors [6, 10, 11]. Bonald implicitly expressed this idea in Ref. [12] saying that the QoS offered to the flows is always better than that obtained if the link were divided into M virtual links (M is the number of CBWFQ queues) of capacity C_m^*, $1 \le m \le M$.

Such CBWFQ system is mostly treated under a decoupling approach, where the system is approximated by two servers [10, 13, 14]. In Refs. [2-4], we gave a simple approximation to evaluate the average number of ongoing connections for each queue based on this approach for any value of M. In

this section, we will differently treat a system of two CBWFQ queues. In our analysis, we will work with the entire system as a one entity and we aim to exploit the results proven in the previous section to analyze the performance of TCP traffic under two CBWFQ queues.

4.1 Performance Evaluation of TCP Traffic in a CBWFQ Queue with Very High Weight

When the weight of a WFQ queue tends toward 1, it can be considered as a priority queue [15]. So that, if a CBWFQ queue m, $m = 1, 2$, has a very high weight ($\vartheta_m \to 1$), the average number of flows traversing this queue (or this virtual link) can be approximated by :

$$E_m^{WFQ} = \frac{\theta_m}{d} + B_m \frac{\theta_m}{C-\theta_m} \qquad (8)$$

With:

$$B_m = \frac{\left(\frac{\theta_m}{d}\right)^N}{N!} \frac{C}{C-\theta_m} \pi_m(0) \qquad (9)$$

$$\pi_m(0) = \left(\sum_{x=0}^{N-1} \frac{\left(\frac{\theta_m}{d}\right)^x}{x!} + \frac{\left(\frac{\theta_m}{d}\right)^N}{N!} \frac{C}{C-\theta_m} \right)^{-1} \qquad (10)$$

4.2 Conservation of the Average Total Number of Flows

Although it is very difficult to prove it mathematically, all our numerical results prove that, when the system is stable, the average total number of flows traversing the system is approximately remains the same with or without the use of the CBWFQ mechanism. We consider, for example, a link of capacity $C = 100 (Mbits/Seconde)$ shared by a number of elastic flow classes having the same maximum bit rated $= 10 (Mbits/Seconde)$. In Fig. 1, we compare the numerical result of the average total number of flows traversing a two CBWFQ system, such as $\vartheta_1 = 0.7$ and $\vartheta_2 = 0.3$, and the numerical results of the average total number of flows traversing a "Best Effort" system for different values of θ. We assume that $\theta_1 = 0.6 \theta$. These results are given by a resolution of a two-dimensional Markov chain modelling
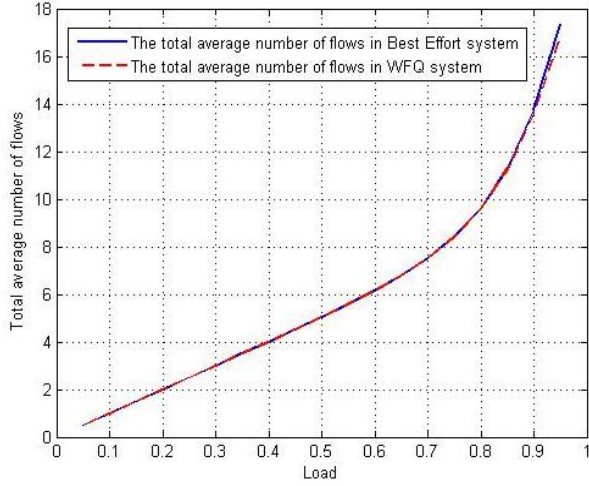
Fig. 1 Comparison between a "Best Effort" system and WFQ system in term of the average total number of flows.

the arrival and departure of flows in the system. We can note that for values of θ/C inferior or equal to 0.85, the error rate is practically negligible. This error rate increases in areas close to the instability zone, but we can always assume that there is a conservation of the average total number of flows in the system.

We define by E^{BE} and E^{WFQ} the average total number of flows for a "Best Effort" system and WFQ system respectively. It is clear that E^{BE} and E^{WFQ} are the sum of the two values of the average number of flows for the two virtual links.

Using Eq. (7), E^{BE} can be written as follows:

$$E^{WFQ} = E^{BE} = \frac{\theta}{d} + B\frac{\theta}{C-\theta} \qquad (11)$$

With B is the congestion probability in the system. It is given using Eq. (5) by replacing ρ by θ.

4.3 Numerical Based Approximation

The average number of flows traversing a queue increases when the weight assigned to this queue decreases. Our numerical observations show that this decrease is not abrupt. This can be explained by the coupling between the two queues: Every capacity unused by one queue can be used by the other.

Also, the numerical observations show that the average number of flows for each virtual link m, m = 1,2, can be expressed in function of its weight by:

$$E_m^{WFQ} = \frac{a}{\vartheta_m{}^\alpha+b} + c \qquad (12)$$

α is an integer superior to 1. α can not be equal to 1 because, in this case, the average number of flows for a class will likely decrease in brutal manner with the increase of the weight of the queue assigned to it, which is not the case in reality.

We tried several values of the order α, and finally we ended to the fact that α can be numerically adjusted to 3: With $\alpha = 3$, we have the least error rate, and then the best results.

When $\vartheta_m = 1/M = 0.5$, m = 1,2 , the system equally shares its resources. The system can be seen as a "Best-Effort" system. So that, we have:

$$\frac{a}{(1/M)^3+b} + c = \frac{\theta_m}{\theta}E^{BE} \qquad (13)$$

When $\vartheta_m \to 1$, E_m^{WFQ} is given by Eq. (8). So that, we have:

$$\frac{a}{1+b} + c = \frac{\theta_m}{d} + B_m\frac{\theta_m}{C-\theta_m} \qquad (14)$$

When $\vartheta_m \to 0$, we have:

$$E_{m^*}^{WFQ} = \frac{\theta_{m^*}}{d} + B_{m1}\frac{\theta_{m^*}}{C-\theta_{m^*}} \qquad (15)$$

With $m^* = (m \bmod 2) + 1$

Therefore:

$$\frac{a}{b} + c = E^{BE} - E_{m^*}^{WFQ} \qquad (16)$$

For Eqs. (13), (14) and (16) we deduce a, b and c:

$$a = \left(\frac{\theta_m}{\theta}E^{BE} - c\right)\left(\left(\frac{1}{M}\right)^3 + b\right) \qquad (17)$$

$$b = \frac{\theta_m\left(\frac{B_m}{C-\theta_m} - \frac{B}{C-\theta}\right)}{(M^3-1)\theta_{m^*}\left(\frac{B_{m^*}}{C-\theta_{m^*}} - \frac{B}{C-\theta}\right) - \theta_m\left(\frac{B_m}{C-\theta_m} - \frac{B}{C-\theta}\right)} \qquad (18)$$

$$c = \frac{\theta_m}{\theta}E^{BE} + \frac{1+b}{1-(1/M)^3}\theta_m\left(\frac{B_m}{C-\theta_m} - \frac{B}{C-\theta}\right) \qquad (19)$$

4.4 Testing the Accuracy of Our Analysis

To examine the accuracy of the proposed analysis, we consider a comparison with the results given by a numerical resolution of the two-dimensional Markov chain for a scenario of a link of capacity $C =$

100(Mbits/Seconde) shared by a two TCP flow classes having the same maximum bit rate $d = 10$(Mbits/Seconde). We assume then that each virtual link, $m = 1, 2$, is traversed by a class flows.

The value of $\theta/C = 0.7$ was chosen to be close to the saturation zone because the QoS differentiation is not significant for low values of θ [12].

Fig. 2 and Fig. 3 compare our approximation Eq. (12) with the exact result given by the numerical resolution of the Markov chain in term of E_m^{WFQ}, $m = 1, 2$, for various weight configurations and for two different values of θ_1/C.

The error rate does not exceed 6% for the two cases, which confirms our good behavior of the approximation. The approximation seems to be more accurate when the distribution of the load is almost the same (the error rate for the first case does not exceed 3.5% for both the two classes). It is important to note, also, that the results given by our approximation are more accurate than those given by a decoupling approach: if we treat the system as two independent virtual links and each virtual link is characterized by a mean capacity.

5. Simulations and Validity of Analytical Results

In this section, we aim to compare our analytical analysis with the real behavior of TCP traffic. We simulate then with NS 2 the case of a link of capacity $C = 100$(Mbits/Seconde) shared by two TCP flow classes having the same maximum bit rate $d = 10$(Mbits/Seconde). We assume then that each virtual link, $m = 1,2$, is traversed by a class flows. We assume that $\theta_1 = 0.6\,\theta$. The weight of the first queue varies from 0.7 to 0.9. Each simulation point is the average of 10 simulation runs and each simulation run lasts an hour and a half. In Figs. 4-6, we compare the simulation results with the analytical results in term of average throughput per queue: all flows traversing the same queue have the same average flow throughput.
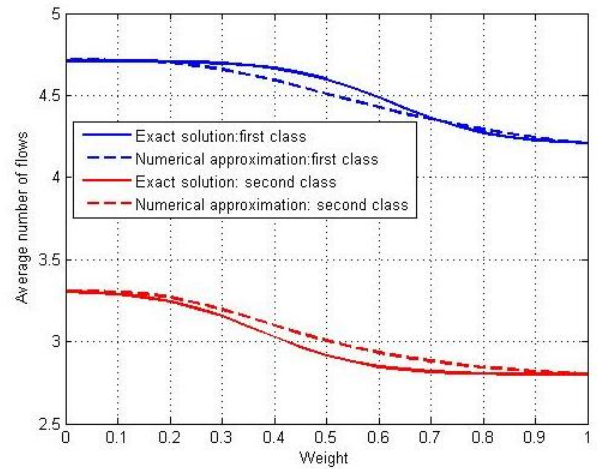

Fig. 2 Comparison between the exact solution and the numerical results in term of average number of flows for the two virtual links: case $\frac{\theta_1}{C} = 0.6$.
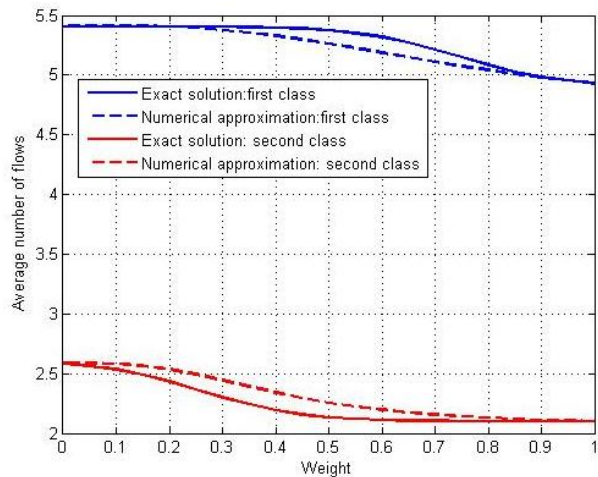

Fig. 3 Comparison between the exact solution and the numerical results in term of average number of flows for the two virtual links: case $\frac{\theta_1}{C} = 0.7$.

When the load does not exceed 80% of the capacity, the two results seem to be very close: the error rate is inferior to 5% for all cases. This error increases when the system is close to the instability zone. In practice, link bandwidth is not shared as precisely as assumed in the fluid models. TCP uses some complex algorithms (Slow Start, Congestion Avoidance...) to control congestion inside the network and restrict the throughput of flows. We maintain however that fluid models provide "very valuable insight into the impact on performance of traffic characteristics" [12]. The insensitivity of average performance to the detailed
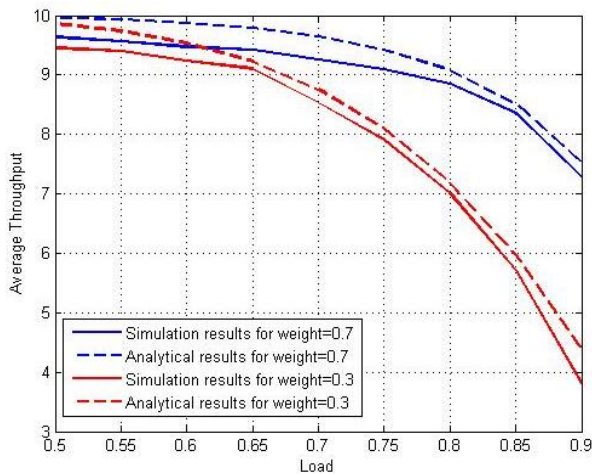
Fig. 4   Comparison between the analytical approximation and simulation in term of average flow throughput for the two virtual links: case $\vartheta_1 = 0.7$.
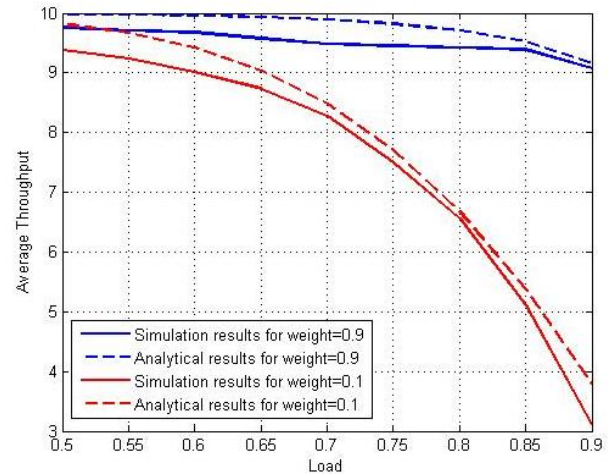

Fig. 6   Comparison between the analytical approximation and simulation in term of average flow throughput for the two virtual links: case $\vartheta_1 = 0.9$.
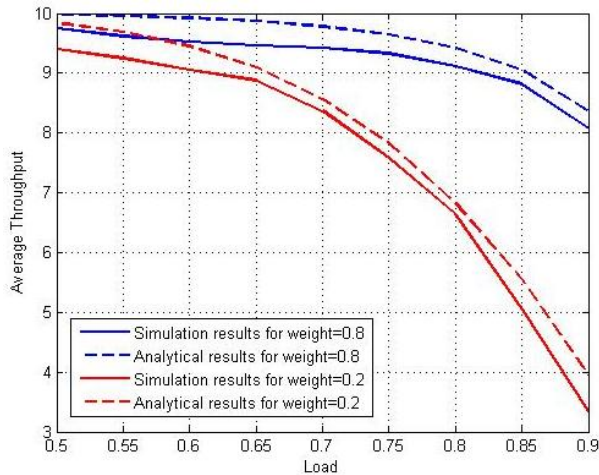
impact on the queueing performance.

The results show that our analyses provide a reasonable estimation of the average number of TCP flows for each class under a system of two CBWFQ queues. More work is needed to integrate these results to evaluate the performance of TCP traffic under a system of higher number of WFQ queues.

6. Conclusion

This paper provides an analytical technique to estimate the performance of TCP traffic under a system of two CBWFQ queues with variable service rates. The core of the analysis relies on some numerical observations for the evolution of the average number of flows for each queue in function of the weight assigned to it.

The important numerical result is that the average total number of flows traversing the system remains the same with or without the use of the WFQ policy. In this context, the results proven for a "Best Effort" system are used to estimate the average number of flows, for each virtual link. Comparison between analytical and simulation results shows the accuracy of our analytical approach.

The key result of the paper is to propose an analytical solution for studying the performance of TCP traffic under CBWFQ system. More work is

Fig. 5   Comparison between the analytical approximation and simulation in term of average flow throughput for the two virtual links: case $\vartheta_1 = 0.8$.

statistical properties of connections is of great importance for network engineering. This property is likely to be maintained approximately even when accounting for disparities due to packet level behavior [12].

The effect of the coupling is especially clear on Fig. 5. The fact that the average throughput of flows traversing the second queue does not tend quickly toward zero (although that the traffic intensity of this queue is superior to the capacity given to it) means that this queue exploited the capacity unused by the other queue. The proposed approximation can capture this effect and provide a reasonable estimation of its

Numerical Key Results for Studying the Performance of TCP Traffic
under Class-Based Weighted Fair Queuing System

201

needed to improve the accuracy of the proposed model and to provide new approximations for calculating the performance of TCP traffic with more than two CBWFQ queues.

References

[1] Garcia, J. M., and Boussada, M. E. H. 2016. "Evaluation des Performances Bout en bout du Trafic TCP Sous le régime. 'Equite Equilibrée'." In *Proceedings of the 11th Performance Evaluation Workshop*, March, LAAS-CNRS, Toulouse/France. (sciencesconf.org:aep11: 92416).

[2] Boussada, M. E. H., Garcia, J. M., and Frikha, M. 2015. "Flow Level Modelling of Internet Traffic in Diffserv Queuing." In *Proceedings of the 5th International Conference on Communications and Networking*, November, Hammamet/Tunisia. IEEE.

[3] Boussada, M. E. H., Garcia, J. M., and Frikha, M. 2016. "Evaluation des Performances Bout en Bout du Trafic TCP Dans une Architecture de Reseau Multi Files D'attente." Presented at the 17th Congress of the French Society of Operations Research and Decision Support (ROADEF), February, Compiegne/France.

[4] Garcia, J. M., and Boussada, M. E. H. 2016. "End-to-End Performance Evaluation of TCP Traffic under Multi-queuing Networks." *Int. J. Communications, Network and System Sciences* 9: 219-33.

[5] https://www.ibm.com/developerworks/library/ws-slafram/.

[6] Al-Sawaai, A., Awan, I., and Fretwell, R. 2009. *Performance Evaluation of Weighted Fair Queuing System Using Matrix Geometric Method.* In NETWORKING 2009 (pp. 66-78). Berlin Heidelberg: Springer.

[7] http://www.cisco.com/en/US/docs/ios/12_0t/12_0t5/feature/guide/cbwfq.html.

[8] Fischer, M. J., Masi, D. M. B., and Shortle, J. F. 2008. "Simulating the Performance of a Class-Based Weighted Fair Queueing System." In *Proceedings of the 40th Conference on Winter Simulation Winter Simulation Conference*, 2901-8.

[9] Brun, O., Ahmad, A. S., and Garcia, J. M. 2009. "Flow-Level Modelling of TCP Traffic Using GPS Queueing Networks." Presented at the 21st International. IEEE Teletraffic Congress, 2009. ITC 21 2009.

[10] Ashour, M., and Le-Ngoc, T. 2006. "Performance Analysis of Weighted Fair Queues with Variable Service Rates." Presented at the IEEE International Conference on Digital Telecommunications.

[11] Lapiotis, G., and Panwar, S. 1999. "Quality of Service Analysis of Shared Buffer Management Policies Combined with Generalized Processor Sharing." Presented at the Global Telecommunications Conference, 1999. GLOBECOM'99, 37-43.

[12] Bonald, T., and James, W. R. 2003. "Congestion at Flow Level and the Impact of User Behaviour." *Computer Networks* 42 (4): 521-36.

[13] Horvath, G., and Telek, M. 2003. "Approximate Analysis of Two Class WFQ Systems." Presented at the Workshop on Preformability Modeling of Computer and Communication Systems - PMCCS 2003, Arlington, IL, USA, September 2003, 43-6.

[14] Guillemin, F., and Pinchon, D. 2004. "Analysis of the Weighted Fair Queuing System with Two Classes of Customers with Exponential Service Times." *Journal of Applied Probability*.

[15] Bockstal, C., Garcia, J., and Brun, O. 2004. Approximation Du Régime Stationnaire D'un Systeme wfq. 6eme Rencontres Francophones Sur les Aspects algorithmiques des telecommunications (ALGOTEL'2004), Batz-sur-Mer (France), 75, 76.