Mohamed El Hedi Boussada, Jean-Marie Garcia, Mounir Frikha

To cite this version:

HAL Id: hal-01888313
https://hal.archives-ouvertes.fr/hal-01888313
Submitted on 10 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientiﬁc research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientiﬁques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A Flow-Level Performance Evaluation of Elastic Traffic under Low Latency Queuing System

Mohamed El Hedi Boussada Mounir Frikha
Mobile Network and Multimedia
SUPCOM, University of Carthage
Ariana-Tunisia
{med.elhadi.boussada, m.frikha}@supcom.tn

Jean Marie Garcia
Services and Architectures for Advanced Networks
LAAS-CNRS
Toulouse, France
jmg@laas.fr

Abstract—Internet tends progressively to be more interactive by supporting real-time communications into this packet-based environment. It is anticipated that online applications will rapidly develop and contribute by a significant amount of traffic in the near future. These applications have strict requirements in terms of delay and throughput which can’t be met only if they are prior. However, at high demand of real-time traffic, elastic traffic (which is transported generally by TCP) may not have sufficient resources to be transported with reasonable quality of service (QoS). Therefore, it is primordial to manage properly the resource-network in order to provide the QoS required by both each type of services. In this paper, we present a new fluid model to evaluate the performance of elastic traffic under Low Latency Queuing (LLQ) system combining a priority queue for delay-sensitive applications with a number of Class Based Weighted Fair Queues (CBWFQ) for elastic traffic. The originality of our contribution consists on focusing on the average total number of flows passing through the whole system by approximating it as a lossless best effort system with total load equal to the total carried load of the original system (no blocked load). The core of our analysis is based on some approximations proven for balanced fairness allocation, which provides a reasonable framework for estimating bandwidth sharing among elastic traffic for best effort allocations. Results issued from this allocation are then exploited to deduce the performance of CBWFQ system. Detailed packet level simulations are used to verify the effectiveness and the accuracy of our analysis. The approach presented in this paper allows a rapid performance evaluation of elastic or rate-adaptive traffic circulating in the actual networks.

Keywords—Elastic traffic, Real time traffic, Low Latency Queuing (LLQ), Class Based Weighted Fair Queuing (CBWFQ), Quality of Service (QoS), Balanced Fairness.

I. INTRODUCTION

Upgrades in Internet connection speeds, with the emergence of 4G LTE cellular and fiber optic communications, has led to a multiplication of services offered by modern telecommunication networks and to unprecedented growth in the number of users and traffic volumes that they generate [1], [2], [3], [4]. Internet traffic is forecasted to increase nearly threefold over the next 5 years. Major portion of this traffic will be dedicated mainly to various forms of video. Globally, IP video traffic will be 82 percent of all consumer Internet traffic by 2021. Live Internet video will contribute by 13 percent in this video traffic and online gaming will grow by 62 percent from 2016 to 2021 [5].

Internet tends, then, to be more interactive by supporting real-time communications in this packet-based environment. Therefore, we shall distinguish two broad categories of Internet traffic: real-time traffic and elastic traffic. Real-time traffic is generated by applications such as video-conference applications, Voice over Internet protocol applications (VoIP applications), online videos, online gaming, etc. These applications have strict bandwidth, end-to-end packet delay and jitter requirements for reliable operations [6]. Elastic traffic is generally transported by TCP (Transport Control Protocol) and it can adapt its rate to the available capacity [7]. Although that TCP is typically designed to carry data (transfer of documents such as files or web pages) , it is also suitable, today, for transporting modern video streaming services such YouTube and Netflix traffic for example [8], [9], [10]. It is reported that most commercial streaming traffic is carried over TCP [9], [11], [12].

Facing this diversity of applications supported by the actual IP networks, transporting all this traffic over one converged packet-oriented network is not built in the same manner. Operators and service providers generally offer all services at different service levels according to the Quality of Service (QoS) level required by both the service and the client [1]. The Diffserv architecture presents an architecture where packets are marked and divided into service classes as they enter the network [13]. Network nodes are then able to provide different services using some scheduling algorithms. In this context, different scheduling policy were proposed. In priority queueing (PQ), higher priority traffic is transmitted before lower priority traffic, with separate buffers for each class of traffic. Weighted Fair Queueing (WFQ) allocates an equal share of the bandwidth for each flow according to the weight attributed to it [14]. The Class Based Weighted Fair Queueing (CBWFQ) is an advanced form of WFQ that supports user defined traffic classes. A queue is allocated for each class, and the traffic belonging to that class is directed to the queue assigned to it [15]. The CBWFQ dynamically allocates the available bandwidth to each traffic class based on the queue’s weight. When one class does not use all its bandwidth, the other classes will share this residual capacity, which creates a coupling aspect between queues in order to optimize the resource sharing and ensure a more efficient traffic transport. Although that this policy may be able to meet the QoS measures needed by the different elastic traffic types, it seems to be not adopted to guarantee the strict QoS requirements of delay-sensitive applications [16]. The composition between these two policies of traffic management techniques (fixed
priority policy and bandwidth sharing-based policy) is consid-
ered by many telecommunication equipment constructors like
Cisco and Huawei [17], [18]. The low-latency queuing (LLQ),
which is being used frequently on the Internet, is a feature
developed by Cisco to bring strict priority queuing to class-
based weighted fair queuing [18]. Priority queuing is used to
satisfy the strict delay constraints for real-time traffic, whereas
CBWFQ is used to assure acceptable throughput for elastic
traffic classes.

During high demand of real-time traffic, the blocking
probability of coming streaming calls becomes important and
the number of ongoing elastic flows indefinitely increases.
Continuous monitoring of the network is then essential to
detect as quickly as possible any degradation of performance
and be able to reconfigure the network to cope with it [19].
At this level, end-to-end measurements are extremely complex to
be implemented [19], [20]. For operators, the solution is to use
traffic-engineering techniques to anticipate the degradation of
quality of service resulting from the phenomena of congestion.
However, the use of these techniques assumes to have models,
theoretical methods, analytical approximations and appropriate
software tools to properly manage the network [1], [20]. Until
now, operators do not have a real reliable model to modelize
and predict the performance of the traffic circulating in their
infrastructures. It is still difficult to predict the performance of
large-scale operator networks with millions of users, millions
of flows, various access terminal, various types of multimedia
applications, complex routing protocols and unexpected user
behaviours.

Flow level models represent very well the great dynam-
icity of the traffic as it makes it possible to represent the
connections and disconnections of users on an application
(or service). However, studying the flow level performance of
elastic traffic in a LLQ system (or an equivalent system)
has not received much attention in literature. Most works
evaluating the performance of elastic traffic in the presence
of real-time traffic assumed that TCP traffic should not be
transported in a differentiated manner [6], [21], [22], [23], [24],
[25], [26]. In general, this viewpoint remains insufficiently
appreciated in actual networks where TCP is used to carry
several types of traffic with different QoS requirements. In this
paper, we will overcome this limitation by proposing a general
flow-level model of a LLQ system where elastic flows are
distributed into a number of CBWFQ queues to be differently
managed. The main idea behind our analysis is to focus on
the average total number of flows crossing the system by
approximating the original system as a lossless Best Effort
system with a total load equal to the total carried load (no
blocked load) of the original system. The conservation of the
total average number of flows is not largely considered at the
previous works in this field which generally tends to analyse
the problem assuming that the dynamics of streaming requests
are much slower or much faster than those of elastic requests.
Our approach is based on adopting some approximations for
balanced fairness allocation, which provides a reasonable
framework for estimating bandwidth sharing among elastic
traffic for best effort allocations, to predict the performance
of elastic traffic under a LLQ system.
The rest of this paper is organised as follows: An overview
of related work is shown in the next section. In the third
section, we describe our model of the system to be studied. The
fourth section is devoted to give useful approximations proven
for balanced fairness allocation in best effort system. These
approximations will be exploited in the fifth section to evaluate
the performance of elastic traffic under a LLQ system. The
accuracy of the proposed analytical results is discussed in the
last section basing on some detailed packet level simulations
with Network Simulator (NS 2).

II. RELATED WORK

The coexistence between elastic and real time traffic is a vexed
problem [27]. Some authors have proposed that real
time traffic should be TCP-friendly, so that it can fairly share
network resources with TCP traffic [25], [27]. In practice,
real time applications often need some form of priority to
function adequately, which make this approach not applicable
in a real context. In [6], [26] and [21], the authors was
interested in studying the performance of elastic flows in a
network where real time traffic is prior and non-adaptive. In
addition, they justified the need for an appropriate admission
control mechanism for real-time traffic to not jeopardize the
performance of these sensitive flows. In [22] and [23], elastic
flows were supposed to have a guarantee bit rate (nominal
bit rate) and each arriving flow (elastic or non-elastic) will be
blocked if it may negatively affect the performances of ongoing
flows.

In the presence of variable service rate for elastic traffic, it
is so difficult to have a regular representation of the generating
matrix of the Markov chain modelling arrivals and departures
of flows in such system. Approximation techniques are pro-
posed then as an alternative to Markov chain analysis. The
integration of streaming and elastic traffic was mainly treated
under a quasi-stationary assumption. Under this assumption,
we assume that the number of ongoing elastic flows rapidly
evolves and attain the stationary regime while the number of
active streaming calls remains unchanged [22], [23], [21],
[26]. This assumption is reasonable when we consider the
combination of voice calls and web browsing or email applica-
tions [22]. In each state of real-time flows, the remaining
capacity (after satisfying the streaming flows requirements)
will be reserved for elastic flows and then we construct a new
elastic system with a reduced capacity. When the probability
that this remaining capacity is inferior to the total elastic
traffic intensity is not negligible, the elastic system becomes
unstable and the performances of elastic flows can not be
easily predicted [6]. Another drawback of this approach is
that it is difficult to guarantee the quasi-stationary assumption
in a system with more than one class of streaming flows
characterising by different arrival rates, different mean service
times and different bit rates.

Assuming now that the dynamics of elastic flows are
much slower than those of streaming flows (considering the
combination of voice calls and large file transfer applications),
another classical approach is proposed in the literature and it
is based on estimating of the mean available capacity for elastic
traffic. In general, this approach does not yield meaningful
results, especially when the average capacity used by real time
flows is close to the system capacity [6].

Other approaches are also proposed. By adopting an admis-
sion control for both elastic and real-time traffic, the authors
of [22] aim to find a weighted version of the two regime-approximations described above for neutral traffic regime (the relative dynamics of both request types are similar) in order to evaluate approximately the blocking probability of each type of flows. Malhotra et al. proposed in [6] a model where the streaming flows are prior and have the same rate as elastic one. The approximations proposed in [6] focus on the conservation of the total workload of the global system, which modeled as a General Processor Sharing (GPS) system without priorities having a total load equal to the elastic load plus the effective load of real time traffic (the carried real time load). The assumption of equal peak rate for both stream and elastic traffic is very restrictive and not always true in practice. Moreover, the Malhotra’s approach is completely sensitive to the detailed characteristics of traffics.

All these papers treated a limited system with two priority queues admitting that there is no differentiation on transporting TCP traffic. In [22], [23] and [26], elastic flows have not a rate limit and they equally share the remaining capacity. In a real context, user flows always have a rate limit which depends on the used access technique. The real behaviour of a general LLQ system is hard to be exactly modelled, especially because of the coupling aspect existing between the priority queue and the CBWFQ system on the one hand and between the CBWFQ queues itself on the other hand.

Our paper aims to overcome the limitations detailed above by :
- Presenting a general model for LLQ system that support different elastic and real time flow-classes.
- Providing a large-scale modelization of the LLQ system taking into account the strong coupling between the queues.

### III. Model

We consider a large number of ADSL/UMTS/LTE subscribers distributed over a single geographic area and connected to the backbone network through an access line of capacity $C$ (Mbits/Second). This access line is assumed to be the bottleneck link for all the traffic flows generated by users. This link is a full-duplex link, which means that the transmission capacity is $C$ (Mbits/Second) in both directions (See figure 1).

Users generate a random number of real-time/streaming and elastic flow-classes. Let $E$ be the set of elastic flow classes and $S$ the set of real-time flow classes.

Real-time flows are mainly defined by their rate and their mean holding-time. For each class $j \in S$, we define:

- $\tau_j(\text{Second})$ : The mean holding-time of flows.
- $d_j(\text{Mbits/Second})$ : The rate of each flow.

For each elastic class-$i$ flows ($i \in E$), we define:

- $\lambda_i(\text{Flows/Second})$ : The mean volume transferred by flows.
- $\theta_i(\text{Mbits/Second})$ : The maximum bit rate of each flow.

Flows arrive as an independent Poisson process with rate $\lambda_i(\text{Flows/Second})$ for streaming class $j$ and $\lambda_i(\text{Flows/Second})$ for elastic class $i$. We refer to the product $\rho_i = \lambda_i \theta_i(\text{Mbits/Second})$ as the load of the elastic class $i$. In the same way, we give by $\rho_j(\text{Flows/Second})$ the load of the streaming class $j$, where $\rho_j = \lambda_j \tau_j(\text{Flows/Second})$.

Let $x_i(\text{Flows/Second})$, $j \in S$, $x_i(\text{Flows/Second})$ be the number of class-$j$ flows in progress (respectively the number of class-$i$ flows in progress) . We refer to the vector $x_i^{(s)} = (x_i(\text{Flows/Second}))_{j \in S}$ (respectively $x_i^{(e)} = (x_i(\text{Flows/Second}))_{j \in E}$) as the state of real-time flow-classes (respectively the state of elastic classes). Let $d_i = (d_j(\text{Mbits/Second}))_{j \in S}$ and $d_i = (d_j(\text{Mbits/Second}))_{j \in E}$ be the vectors of rate limits for streaming flow-classes and elastic flow-classes, respectively.

We note by $\theta^e = \sum_{i \in E} \rho_i^e$ (respectively $\theta^s = \sum_{i \in S} \rho_i^s$) as the elastic load (respectively the real-time load) offered to the capacity $C$ in Mbits/Second. In order to maintain the stability of the system, we assume that the total load is strictly inferior to the link capacity:

$$\theta = \theta^e + \theta^s < C \quad (1)$$

In a similar way to the configuration of Internet routers, we assume that, for both two directions, packets are affected to a LLQ queue before being sent into the link. This LLQ queue combines a priority queue with a number of $M$ CBWFQ queues. Let $w_m$, $1 \leq m \leq M$, be the weight of the queue number $m$. We assume that:

$$\sum_{m=1}^{M} w_m = 1 \quad (2)$$

The priority queue is devoted to real-time flows, which have strict bandwidth and delay requirements that can be met only if the link capacity is completely allocated to them. Real-time flows whose requirements cannot be guaranteed will be blocked rather than allow them into the system and jeopardize the performance of ongoing real time traffic. This admission control coupled to the strict priority is generally considered sufficient to meet the quality of service requirements of the audio and video applications. Elastic traffic is distributed throughout the CBWFQ queues.

In a large scale modelization, each CBWFQ queue can be represented by a virtual link. We note by $E_m$ the set of elastic flow classes passing through the virtual link number $m$ and by $\theta^e_m = \sum_{i \in E_m} \theta_i^e$ the total elastic load offered to this virtual link. In the following, the virtual link $m$ refer to the CBWFQ queue $m$ and vice versa.

Since the above resource sharing is applied for both upload and download directions, our study will be limited to one
IV. APPROXIMATE RESULTS FOR BALANCED FAIRNESS ALLOCATION

In this section, we assume that the link capacity is only exploited by elastic traffic and there is no CBWFQ policy for this traffic.

The dynamic flow level modeling of Internet traffic in Best Effort architecture was practically introduced with Massoulié and Roberts with a model of fair bandwidth sharing (Processor-sharing Model) [32]. Theoretical allocations optimising an utility function of the instantaneous flow rates have also been proposed to estimate the TCP bandwidth sharing. Examples of such allocations are the classical max-min fairness and the Kelly’s proportional fairness [29]. In general, the analysis of a network operating under these allocations scheme is quite difficult. One reason is that they do not lead to an explicit expression for the stationary distribution of a vector $x$. We note also by $\pi(B)$ the stationary probability of a set $B$ and by $\pi(x)$ the stationary probability of a vector $x$.

A. Identical rate limits

We suppose here that all elastic classes have the same maximum bit rate $r$. Let $N = \lceil C/r \rceil$ be the maximum number of flows that can have its maximum bit rate. In [20], Henda proved that, using balanced fairness allocation, the average number of flows for each class $i$ is exactly given by:

$$E[x_i^{(e)}] = \frac{\rho_i^{(e)}}{r} + \pi(B) \frac{\rho_i^{(e)}}{C - \theta(e)}$$

(6)

where $B$ is the set of congestion stats:

$$\pi(B) = \pi(|x_i^{(e)}| \geq N) = \frac{(\frac{\theta}{r})^N C}{N! (C - \theta(e))} \pi(0)$$

(7)

$\pi(0)$ is the probability that the link is empty, and it is given by:

$$\pi(0) = \left( \sum_{k=0}^{N-1} \frac{\theta^{(c)} k}{r^k} + \frac{\theta^{(c)} N}{r^N} \frac{C}{C - \theta(e)} \right)^{-1}$$

(8)

The average total number of flows can be then deduced from (6) by:

$$E[x^{(e)}] = \frac{\theta^{(c)}}{r} + \pi(B) \frac{\theta^{(c)}}{C - \theta(e)}$$

(9)

And then, we have the following relation:

$$E[x_i^{(e)}] = \frac{\rho_i^{(e)}}{\theta^{(c)}} E[x^{(e)}]$$

(10)

B. General rate limits

For different transmission rates, an approximation was given in [20] and [21]:

$$E[x_i^{(e)}] \approx \frac{\rho_i^{(e)}}{d_i^{(e)}} + \pi(B_i) \frac{\rho_i^{(e)}}{C - \theta^{(c)}}$$

(11)

where $B_i$ is the set of congestion stats for the class $i$: $B_i = \{ x_i^{(e)} , x_i^{(e)} d_i^{(e)} \geq C - d_i^{(e)} \}$. The probability $\pi(B_i)$ is written as follows:

$$\pi(B_i) = \frac{1}{C - \theta^{(c)}} \sum_{k \in E} \rho_k^{(e)} \pi(W_k) + \pi(W_i)$$

(12)
\( W_k \) is defined as \( W_k = \{ x^{(e)}, C^{(e)} - d_k \leq x^{(e)}d^{(e)} < C \} \) and its probability is given by:

\[
\pi(W_k) = \pi(0) \sum_{x^{(e)} \in W_k} P(x^{(e)}d^{(e)})
\]  

(13)

Where \( P(x^{(e)}d^{(e)}) \) is given by the Kaufman-Roberts formula. This formula is written as follows [35]:

\[
P(n) = \sum_{k \in E} \frac{\rho_k}{n} P(n - d_k^{(e)}) \quad \forall 1 \leq n \leq C
\]  

(14)

with \( P(0) = 1 \) and \( P(n) = 0 \) for all \( n < 0 \).

The probability that the link is empty is then given by:

\[
\pi(0) = \left( \sum_{0 \leq x^{(e)}d^{(e)} < C} P(x^{(e)}d^{(e)}) \right) + \frac{1}{C - \theta^{(e)}} \sum_{k \in E} \rho_k^{(e)} \sum_{x \in W_k} P(x^{(e)}d^{(e)})^{-1}
\]  

(15)

V. ANALYSIS OF THE LLQ SYSTEM

A. High Priority traffic

If the capacity \( C \) is already occupied by high priority traffic then further incoming real-time flows are lost. For each class \( j \in S \), we note by \( B_j = \{ x^{(s)}, C - d_j^{(s)} < x^{(s)}d^{(s)} \leq C \} \) the set of blocking states. The probability of this set is given by:

\[
\pi(B_j) = \sum_{C - d_j^{(s)} < x^{(s)}d^{(s)} \leq C} \pi(x^{(s)})
\]  

(16)

For each state \( x^{(s)} \), the probability \( \pi(x^{(s)}) \) is given by:

\[
\pi(x^{(s)}) = \prod_{j \in S} \frac{\rho_j^{(s)}x_j^{(s)}}{x_j^{(s)!}} \pi(0)^{(s)}
\]  

(17)

Where:

\[
\pi(0)^{(s)} = \left( \sum_{0 \leq x^{(s)}d^{(s)} < C} \frac{\rho_j^{(s)}x_j^{(s)}}{x_j^{(s)!}} \right)^{-1}
\]  

(18)

Kaufman and Roberts proposed a more efficient solution than the direct calculation of the product form of the blocking probability [36], [37]. After splitting the link capacity into virtual circuits, their solution based on a recursive algorithm to determine the probability that \( m \) circuits are occupied. Let \( \delta d \) the greatest common divisor between \( d_j^{(s)} \) \( \forall j \in S \). We have then:

- \( K = C/\delta d \): The maximum number of circuits.
- \( m_j = d_j^{(s)}/\delta d \): The number of circuits required for one flow of class \( j \).

According to Kaufman and Roberts, the state probabilities are calculated by the following recursive algorithm:

\[
\pi^*(m) = \begin{cases} 
1 & \text{if } m = 0 \\
0 & \text{if } m < 0 \\
\frac{1}{m} \sum_{j \in S} \pi^*(m - m_j)m_j \rho_j^{(s)} & \text{else}
\end{cases}
\]  

(19)

The blocking probability for each class \( j \) is obtained by:

\[
\pi(B_j) = \sum_{K - m_j + 1 \leq m \leq K} \pi^*(m)
\]  

(20)

The average number of flows for each class \( j \) is written as follows:

\[
E[x_j^{(s)}] = \rho_j^{(s)}(1 - \pi(B_j))
\]  

(21)

Thus, the average total number of flows for real-time traffic is given by:

\[
E(s) = \sum_{j \in S} E[x_j^{(s)}]
\]  

(22)

B. Evaluation of the average total number of flows

Since each lost TCP packet is sent again until it reaches the destination, looking at flow-level, it is assumed that the total elastic traffic load is conserved in the system (there is no loss). Let \( \theta^{(mix)}(Mbits/Second) \) the average total load carried by the system. \( \theta^{(mix)} \) can be written as follows:

\[
\theta^{(mix)} = \sum_{j \in S} E[x_j^{(s)}]d_j^{(s)} + \theta^{(e)}
\]  

(23)

For a Best Effort system, the link capacity will be equally shared to carry the total traffic load \( \theta^{(mix)} \). For a LLQ system, the same capacity will be shared in a differentiated manner to carry the same total traffic load. This leads to the fact that the mean total number of flows is approximately conserved between the two systems. Let \( E_{\theta^{(mix)}} \) the average total number of flows carried by the system. We define a new set of elastic flow classes noted \( E' \) with the same cardinality as \( S' \) such that for each \( k \in E' \), \( \rho_k^{(e)} = E[x_k^{(s)}]d_k^{(s)} \) and \( d_k^{(e)} = d_k^{(s)} \cdot E_{\theta^{(mix)}} \) is given by:

\[
E_{\theta^{(mix)}}[x_i^{(e)}] = \sum_{i \in E \cup E'} E_{\theta^{(mix)}}[x_i^{(e)}]
\]  

(24)

Where \( E_{\theta^{(mix)}}[x_i^{(e)}] \) is deduced from 11 by replacing \( \theta^{(e)} \) with \( \theta^{(mix)} \).

The average total number of elastic flows can thus be approximated by:

\[
E_{\theta^{(mix)}}[x_i^{(e)}] \approx E_{\theta^{(mix)}} - \sum_{j \in S} E[x_j^{(s)}]
\]  

(25)

C. Evaluating the average total number of flows for each CBWFQ queue: General case

In this section we assume that all elastic flow-classes in \( E \) have the same maximum bit rate \( r \).

As we showed in [38] for a system of two CBWFQ queues carrying traffic with identical maximum bit rate, the evolution of the average total number of flows for each queue \( m \) in function of its weight can be estimated by the expression 26, where \( \beta_1, \beta_2 \) and \( \beta_3 \) are calculated according the system parameters and \( \alpha \) is numerically adjusted to 3. Our numerical observations show that this assumption remains true even for a greater number of queues:

\[
E_m^{CBWFQ} \approx \frac{\beta_1}{\beta_2 + \alpha} + \beta_3
\]  

(26)
The parameters $\beta_1$, $\beta_2$ and $\beta_3$ will be fixed basing on known results for three specific cases for $w_m$: $w_m$ tends towards 1, $w_m$ is equal to $1/M$ and $w_m$ tends towards 0.

When the weight of a queue $m$, tends towards 1, it can be considered as a priority queue [39]. Let $E_{CBWFQ}^{m/w_m\to1}$ the average number of flows passing through this queue in this case. $E_{CBWFQ}^{m/w_m\to1}$ can be approximated by:

$$E_{CBWFQ}^{m/w_m\to1} \approx E_{\theta^{(mz)}} - \sum_{k \leq m} \sum_{k \neq m} E [x^{(r)}]$$ (27)

Where $E_{\theta^{(mz)}} - \sum_{k \leq m} \sum_{k \neq m} E [x^{(r)}]$ refers to the average total number of flows for a system with two priority queues. The first priority queue is dedicated for real-time traffic while the second is devoted to elastic flow-classes in $E_m$. $E_{\theta^{(mz)}} - \sum_{k \leq m} \sum_{k \neq m} E [x^{(r)}]$ is given by:

$$E_{\theta^{(mz)}} - \sum_{k \leq m} \sum_{k \neq m} E [x^{(r)}] = \sum_{i \in E_m \cup E^r} E_{\theta^{(mz)}} - \sum_{k \leq m} \sum_{k \neq m} E [x^{(r)}]$$ (28)

Exploiting 26, $E_{CBWFQ}^{m/w_m\to1}$ can be written as follows:

$$E_{CBWFQ}^{m/w_m\to1} \approx \frac{\beta_1}{\beta_2 + 1} + \beta_3$$ (29)

When the weight of each queue $m$ is equal to $1/M$, the remaining capacity for elastic traffic is supposed to be equally shared among the queues, and we are practically in the best effort case. Given that all elastic flows have identical rate limits and using 10, we have:

$$E_{CBWFQ}^{m/w_m=1/M} \approx \frac{\theta^{(r)}}{\theta^{(r)}} E_{\theta^{(mz)}} [x^{(r)}]$$ (30)

Where $E_{\theta^{(mz)}} [x^{(r)}]$ is given by 25. Using 26 we obtain:

$$E_{CBWFQ}^{m/w_m=1/M} \approx \frac{\beta_1}{\beta_2 + (1/M)^\alpha} + \beta_3$$ (31)

When the weight of a queue $m$ tends towards zero, the average number of flows passing through this queue is obtained by:

$$E_{CBWFQ}^{m/w_m\to0} = E_{\theta^{(mz)}} [x^{(r)}] - E_{\theta^{(mz)} - \theta^{(m)}} [x^{(r)}]$$ (32)

Where $E_{\theta^{(mz)} - \theta^{(m)}} [x^{(r)}]$ refers to the average total number of elastic flows passing through the system regardless of the traffic of the CBWFQ queue number $m$. It is deduced from 25 as follows:

$$E_{\theta^{(mz)} - \theta^{(m)}} [x^{(r)}] \approx E_{\theta^{(mz)} - \theta^{(m)}} - \sum_{j \in S} E [x^{(s)}]$$ (33)

Where $E_{\theta^{(mz)} - \theta^{(m)}}$ is given by:

$$E_{\theta^{(mz)} - \theta^{(m)}} = \sum_{i \in (E/E_m)+E^r} E_{\theta^{(mz)} - \theta^{(m)}} [x^{(r)}]$$ (34)

$E_{\theta^{(mz)} - \theta^{(m)}} [x^{(r)}]$ is deduced from 11 by replacing $\theta^{(r)}$ with $\theta^{(mz)} - \theta^{(m)}$.

According to 26 , $E_{CBWFQ}^{m/w_m\to0}$ can be written as follows:

$$E_{CBWFQ}^{m/w_m\to0} \approx \frac{\beta_1}{\beta_2 + \beta_3}$$ (35)

The average total number of active flows in each queue can be then approximated by 26 where $\beta_1$, $\beta_2$ and $\beta_3$ are given from 29, 31 and 35 as follows:

$$\beta_1 \approx \left( E_{CBWFQ}^{m/w_m=1/M} - \beta_3 \right) \left( \frac{1}{M} \right) \alpha + \beta_2$$ (36)

$$\beta_2 \approx \frac{\delta_1}{(M^\alpha - 1)^2 - \delta_1 - \delta_1}$$ (37)

$$\beta_3 \approx E_{CBWFQ}^{m/w_m=1/M} - \frac{1 + \beta_2}{1 - (1/M)^\alpha} \delta_1$$ (38)

With:

$$\delta_1 = E_{CBWFQ}^{m/w_m=1/M} - E_{CBWFQ}^{m/w_m=1/M}$$ (39)

And:

$$\delta_2 = E_{CBWFQ}^{m/w_m=0} - E_{CBWFQ}^{m/w_m=1/M}$$ (40)

Since all flows passing through a queue $m$ are supposed to have the same maximum bit rate, the average number of flows for each class $i \in E_m$ can be deduced from 10 as follows:

$$E [x^{(r)}] = \frac{\beta^{(r)}_i}{\theta^{(r)}_i} E_{CBWFQ}$$ (41)

VI. SIMULATIONS AND VALIDITY OF ANALYTICAL RESULTS

In this section, we aim to compare our analytical analysis with the real behaviour of TCP traffic. We simulate then with NS 2 the case of a capacity link $C = 100(Mbits/Second)$ shared by two streaming flow classes (transporting by User Datagram Protocol (UDP)) and three TCP flow classes. We assume then that at the entry of the link there is a LLQ queue that combines a priority queue with three CBWFQ queues and each CBWFQ queue is dedicated to a single class of TCP flows. Let $w_1 = 0.7$, $w_2 = 0.25$ and $w_3 = 0.05$. Let $d_1^{(s)} = 10(\text{Mbits/Second})$, $d_2^{(s)} = 5(\text{Mbits/Second})$ and $r = 3(\text{Mbits/Second})$. We assume that UDP traffic constitutes 60% of the total system load which varies from 50% to 80% of the link capacity. The two streaming classes have the same load. In the same manner, we assume that all elastic flow classes have the same contribution in the total elastic load. For our simulation, we took $\lambda_j^{(s)} = \lambda_i^{(r)} = 1$ ($\forall i \in E, \forall j \in S$).

Each TCP flow is used to transfer a series of packets of 1000 bytes representing a document of a certain size. We chose FTP (File Transfer Protocol) and CBR (Constant Bit Rate) to be implemented respectively for TCP and UDP traffic. To ensure compatibility with our theoretical analysis, we assumed that there is no additional time added by the link.

Figure 2 plots a comparison between analytical and simulations results in term of the average throughput per queue for different ratios of $\theta/C$. At first, it is necessary to note that
In this paper, we propose a new fluid approach to evaluate the performance of elastic approach under a LLQ system. Our analysis focus on the total average flows carried by the system as a lossless Best Effort system with total load equal to the total carried load of the original system (no blocked load). Approximations issued from balanced fairness, which is considered by many authors as an efficient tool for estimating bandwidth sharing among elastic traffic for best effort allocations, are used to approximate the average total number of flows in this equivalent system and deduce the performance of CBWFQ system. Detailed packet level simulations proved the accuracy and the good behaviour of our analysis.

Our approach is original in the sense that it overcomes the limitations of previous works studying the integration of elastic and real-time traffic under classical approaches (quasi-stationary approach, calculating the average remaining capacity for elastic traffic) and assuming that all elastic traffic have the same requirement in quality of service. The problem that we studied reflects the reality (and the complexity) of the Internet multimedia processes with heterogeneous flows, differentiated classes of services and different transport protocols. The expression given to evaluate the average end-to-end throughput of elastic traffic under a multi-queue system is precise and can capture the strong coupling between the queues with a reasonable computation time.

Another key result is that the approximation proposed is insensitive to detailed traffic characteristics since it is based on approximations issued from balanced fairness allocations. This is particularly important for data network engineering since performance can be predicted from an estimate of overall traffic volume alone and is independent of changes in the mix of user applications. We expect results such as those presented in this paper to eventually lead to simple and robust traffic engineering rules and performance evaluation methods that are lacking for actual data networks.

VII. Conclusion

In this paper, we propose a new fluid approach to evaluate the performance of elastic approach under a LLQ system giving the head of priority to delay-sensitive applications. Our analysis focus on the total average flows carried by the whole system by approximating the system as a lossless Best

---

**References**


