End-to-End Quality of Service-based Admission Control Using the Fictitious Network Analysis

Pablo Belzarena[1], Paola Bermolen[1], Pedro Casas[1], Maria Simon[1]

[a]*Julio Herrera y Reissig 565, CP 11300, Montevideo, Uruguay*

Abstract

The performance analysis of a network link is a well studied problem. However, the most interesting issue for a service provider is to evaluate the end-to-end quality of service (QoS). The evaluation of the end-to-end QoS (e.g. loss probability or delay) depends on the traffic statistic which is constantly modified as the traffic traverse the network, making its analysis a very difficult problem. In this work we use a simplified framework known as fictitious network analysis that allows us to estimate on-line the end-to-end loss ratio from input traffic traces statistics. We prove that the defined estimator is consistent and that a Central Limit Theorem is verified. Based on these estimations an admission control mechanism can be implemented. More precisely, we propose a simply method to estimate the control admission region, i.e. which are the flows that can be accepted in the network that verifies that its end-to-end loss ratio is smaller than a given threshold.

While decisions based on the fictitious network analysis are safe, it may lead to network resources under-utilization (it generally overestimates the QoS parameters). In this work we establish sufficient conditions to assure that results obtained by means of the fictitious network coincide with real ones (there is no overestimation). We present first the conditions in the one-link case and extend them to the multilink case, necessary to evaluate the end-to-end loss ratio. When different results are obtained we define a method to find a bound for the overestimation. We also present numerical examples to compare the performance obtained in the real and the fictitious network, validating our main results.

Key words: Admission Control, End-to-End Quality of Service, Fictitious Network

1. Introduction

The admission control mechanisms proposed in the literature are mainly based on one link analysis [3]. However, for a service provider the most interesting issue is about

Email addresses: `belza@fing.edu.uy` (Pablo Belzarena), `paola@fing.edu.uy` (Paola Bermolen), `pcasas@fing.edu.uy` (Pedro Casas), `msimon@fing.edu.uy` (Maria Simon)

[1]ARTES Group: Joint Research Group of the Electrical Engineering and Mathematics and Statistics Departments. Facultad de Ingeniería, Universidad de la República. Contact: http://iie.fing.edu.uy/investigacion/grupos/artes or artes@fing.edu.uy

admission control mechanisms based on the end-to-end quality of service (QoS) evaluation. Recently some authors propose an end-to-end admission control mechanism based on active measurements [12]. In this paper we propose a different approach based on analytical models.

In order to design an admission control mechanism many issues must be addressed. The focus of this work is on the estimation of the admission control region. We look for a simply and efficient procedure for such estimation which can be applied on line.

To evaluate end-to-end QoS guarantees in a network, a performance model is required. In this work we will analyze the model called the "fictitious network model" in the context of many sources and small buffer asymptotic, introduced by Ozturk et al.[13]. We will show that this model allows to simple and on-line estimations of end to end QoS parameters, which will be in turn used to decide which flows can access the network.

Ozturk et al. find a useful way to analyze the overflow probability in a network interior link and show that when the fictitious network model is applied, an overestimation is obtained. The fictitious network analysis gives then a simple and efficient yet conservative way to implement on-line admission control mechanisms. However, the overestimation can translate into wasted network resources. If a flow is admitted, its QoS is guaranteed but the link capacities can be under-used.

In this work we analyzw in detail the fictitious network model and we find conditions to assure that the fictitious network analysis in an interior link gives the same overflow probability than the real network analysis, being much simpler. We also find a method to bound the overestimation when these conditions are not met. Preliminary results were analyzed in a previous work [2].

Next we analyze how to apply previous results to the evaluation of the end-to-end loss ratio. Again, we found conditions to assure that the calculations made over the fictitious network coincide with the real ones. Error bounds are also available in case these conditions are not verified. In addition, since no model is assumed for the input traffic, we define an estimator of the end-to-end Loss ratio based on traffic measurements. We show that this estimator is a good one, i.e. is consistent and verifies a Central Limit Theorem (CLT). These results allow us to define an admission control mechanism based on the expected end-to-end Loss Ratio a flow traversing the network will obtain.

In section 2 we summarize some related works in order to give a brief description of the problem´s context. The model introduced by Ozturk et al. is explained in section 3 where we summarize their main results. Our main results for the one link case are presented in section 4. In section 5 we analyze how these results can be extended to the multilink case. In section 6 we analyze how to estimate the end-to-end loss ratio from traces of the input traffic in the fictitious network. In section 7 we show some numerical examples to compare the real and the fictitious network results and validate our main results. Finally, we summarize and conclude our work in section 8.

2. Related works

Network designers and operators need a model to evaluate the end to end network performance in an Internet core backbone. Since losses are "rare" events, some researchers have proposed the use of Large Deviations Theory for network performance evaluation.

In this context, the effective bandwidth notion was introduced some years ago. The notion of equivalent bandwidth was formerly used to study the access control problem for some networks, as ATM. Many contributions following this approach were done during the 90's to analyze the access control in some networks based on the IntServ model or others. In that situation, the access node receives a connection request and has to estimate the resources it requires, in order to allow or deny the new connection. Kelly's Effective Bandwidth (EB) [10] may be used in such situations as the "equivalent capacity" needed by the new connection. In this context, the flow to be statistically characterized is an individual flow, and may be directly related with the data source (for instance, voice or video codecs). This situation was studied using the so called large buffer asymptotic, in which the link buffer grows to infinity, and its filling above some threshold is analyzed. This approach cannot be used in backbone links, where buffers are not devised to store bursts but to resolve simultaneous packet arrival, being consequently small. The application of Large Deviations Theory to the analysis of a network backbone must be performed on the basis of the many sources asymptotic. In this regime we take buffer size $B = Nb$ (with N the number of sources), output capacity $C = Nc$ and make N go to infinity. Results about loss probability in this regime can be found in [7, 8, 16, 17].

Using Large Deviations, Wischik [17] proves the following formula (called *inf sup* formula) for the overflow probability:

$$\log \mathbf{P}(Q_N > B) \approx -\inf_{t\geq 0} \sup_{s\geq 0}((B + Ct)s - Nst\alpha(s,t))$$

where Q_N represents the stationary amount of work in the queue, C is the link capacity, B is the buffer size and N is the number of incoming multiplexed sources of effective bandwidth $\alpha(s,t)$.

Wischik also shows in [18] that in the many sources asymptotic regime the aggregation of independent copies of a traffic source at the link output and the aggregation of similar characteristics at the link input, have the same effective bandwidth in the limit when the number of sources goes to infinity. This result allows to evaluate the end to end performance of some kind of networks like "in-tree" ones. Unfortunately this analysis can not be extended to networks with a general topology.

Eun y Shroff [9], have shown that in the many sources asymptotic regime, the probability of the buffer size to be grater than zero goes to zero when the number of sources goes to infinity. This result is valid for a discrete time queue and for a continuous time queue if the source is bounded or can be expressed as an integral of a stationary stochastic process.

A slightly different asymptotic with many sources and small buffer characteristics was proposed by Ozturk, Mazumdar and Likhanov in [13]. They consider an asymptotic regime defined by N traffic sources, link capacity increasing proportionally with N but buffer size such that $\lim \frac{B(N)}{N} \to 0$. In their work they calculate the rate function for the buffer overflow probability and also for the end to end loss ratio. This last result can be used to evaluate the end to end QoS performance in a network backbone in contrast with the Wischick result explained before, where it is necessary to aggregate at each link N i.i.d. copies of the previous output link.

Ozturk et al. also introduce the “fictitious network” model. The fictitious network is a network with the same topology than the real one, but where each flow aggregate goes to a link on its path without being affected by the upstream links until that link. The fictitious network analysis is simpler and so, more adequate to on-line performance evaluation and traffic engineering. Ozturk et al. show that the fictitious network analysis overestimates the overflow probability. In this work we analyze when, for an interior network link, the overflow probability calculated using the fictitious network is equal to the overflow probability of the real network. Ramon Casellas [5] has also studied the overestimation problem in the fictitious network. He found a condition to assure that there is no overestimation. This condition is a particular case of the sufficient condition proven in this work. In the next section we summarize Ozturk et al. main results.

3. Many sources and small buffer asymptotic performance model

3.1. Ozturk, Mazumdar and Likhanov work

Consider a network of L links which is accessed by M types of independent traffic. Consider a discrete time fluid FIFO model where traffic arrives at time $t \in Z$ and is served immediately if buffer is empty and is buffered otherwise. Each link k has capacity NC_k and buffer size $B_k(N)$ where $B_k(N)/N \to 0$ with $N \to \infty$. Input traffic of type m=1,...,M, denoted $X^{m,N}$ is stationary and ergodic and has rate $X_t^{m,N}$ at time t (workload at time t of N sources of type m).

Let $\mu_m^N = \mathbf{E}(X_0^{m,N})/N$ and $X^{m,N}(t_1,t_2) = \sum_{t=t_1}^{t_2} X_t^{m,N}$. We assume that $\mu_m^N \underset{N\to\infty}{\to} \mu_m$ and $X^{m,N}(0,t)/N$ satisfies the following Large Deviation Principle (LDP) with *good rate function* $I_t^{X^m}(x)$:

$$-\inf_{x\in\Gamma^o} I_t^{X^m}(x) \le \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\left(\frac{X^{m,N}(0,t)}{N} \in \Gamma\right) \tag{1}$$

$$\le \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\left(\frac{X^{m,N}(0,t)}{N} \in \Gamma\right) \le -\inf_{x\in\overline{\Gamma}} I_t^{X^m}(x) \tag{2}$$

where $\Gamma \subset \mathbb{R}$ is a Borel set with interior Γ^o and closure $\overline{\Gamma}$ and $I_t^{X^m}(x) : \mathbb{R} \to [0,\infty)$ is a continuous mapping with compact level sets. We also assume the following technical condition: $\forall$ m and $a > \mu_m$,

$$\liminf_{t\to\infty} \frac{I_t^{X^m}(at)}{\log t} > 0$$

Type m traffic has a fixed route without loops and its path is represented by the vector $\mathbf{k}^m = (k_1^m, ..., k_{l_m}^m)$, where $k_i^m \in (1,..,L)$. The set $\mathcal{M}_k = \{m : k_i^m = k, 1 \le i \le l_m\}$ denotes the types of traffic that goes through link k. To guarantee system stability it is assumed that

$$\sum_{m\in\mathcal{M}_k} \mu_m < C_k \tag{3}$$

The main result of Ozturk et al. work is the following theorem.

Theorem 3.1. *Let $X_{k,t}^{m,N}$ be the rate of type m traffic at link k at time t. There exist a continuous function $g_k^m : \mathbb{R}^M \to \mathbb{R}$ relating the instantaneous input rate at link k for traffic type m to all of the instantaneous external input traffic rates such that:*

$$\frac{X_{k,0}^{m,N}}{N} = g_k^m \left(\frac{X_0^{1,N}}{N}, ..., \frac{X_0^{M,N}}{N} \right) + o(1) \tag{4}$$

The buffer overflow probabilities are given by:

$$\lim_{N\to\infty} \frac{1}{N} \log P(\textit{overflow in link } k) = -\mathbf{I}_k =$$
$$-\inf \left\{ \sum_{m=1}^{M} I_1^{X^m}(x_m) : x = (x_m) \in \mathbb{R}^M, \sum_{m=1}^{M} g_k^m(x) \geq C_k \right\} \tag{5}$$

In (4), $o(1)$ verifies that $\lim\limits_{N\to\infty} o(1) = 0$ since $\frac{B_k(N)}{N} \underset{N\to\infty}{\to} 0$. The function $g_k^m(x)$ is constructed in the proof of the theorem. Ozturk et al. prove that the continuous function relating the instantaneous input rate at link i for traffic m to all of the instantaneous external input traffic rates is the same function relating these variables in a no buffers network. The function relating the instantaneous output rate at link i for traffic m to all of the instantaneous input traffic rates at this link is:

$$f_i^m(x, C_i) = \frac{x_m C_i}{\max(\sum\limits_{j\in\mathcal{M}_i} x_j, C_i)} \tag{6}$$

In a feed-forward network the function $g_k^m(x)$ can be written as composition of the functions of type (6) in a recursive way. Using equation (6) the buffer overflow probability can be calculated for any network link, by solving the optimization problem of equation (5). We need to know the network topology, the link capacities and, for each arrival traffic type m, the rate functions $I_1^{X^m}$.

Ozturk et al. define also the total (end to end) loss ratio as the ratio between the expected value of lost bits at all links along a route and the mean of input traffic in bits, for stream m identified by X^m. With the previous definition they find the following asymptotic for the loss ratio $\mathbf{L}^{m,N}$:

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbf{L}^{m,N} = -\min_{k\in k^m} \mathbf{I}_k \tag{7}$$

The main problem of this approach is that the optimization problem of equation (5) could be very hard to solve for real-size networks. The calculation of the function $g_k^m(x)$ is recursive and so, when there are many links it becomes complex. In addition, the virtual paths can change during the network operation. Therefore, it is necessary to recalculate on-line the function $g_k^m(x)$. To solve equation (5), it is also necessary to optimize a nonlinear function under nonlinear constraints. In order to simplify this problem, Ozturk et al. introduce the "fictitious network" concept, that is simpler and gives conservative results. In the next section we find conditions to assure that there is no overestimation in the calculus of the link overflow probability in the fictitious network analysis. We also find a bound for the error (difference between the rate function calculated for the real network and the fictitious one) in those cases where the previous condition is not satisfied.

The aim of our work is to define an admission control mechanism. Such a mechanism is simple a set of rules to accept or reject a flow that intend to access the network. This can be done by defining an acceptance region, i.e. which is the set of flows that can access the network. In [13] an acceptance region based on end-to-end QoS guarantees, is defined. This acceptance region is the traffic mix that can flow through the network without QoS violation. Assume that $X^{m,N}$ is the sum of Nn_m i.i.d. process. More formally, the acceptance region noted by $\mathcal{D}$ is the mix or collection $\{n_m\}_{m=1}^{M}$ of sources which can be flowing through the network while the QoS (loss ratio) for each class is met, that is:

$$\mathcal{D} = \{(n_m), m = 1, ..., M : \lim_{N\to\infty} \frac{1}{N} \log \mathbf{L}^{m.N} < -\gamma_m\} \quad \text{with} \quad \gamma_m > 0 \tag{8}$$

We will concentrate then in the estimation of this acceptance region. We aim not only to do it in a efficient way but also in a simple one in order to apply it on-line.

4. Fictitious network analysis

We analyze an interior network link k under the same assumptions that in Ozturk et al. work. $\mathcal{M}$ is the set of traffic types that access the network and $\mathcal{M}_i$ is the set of traffic types that go through link i. We suppose that the network is feed-forward, this means that each traffic type has a fixed route without loops. In the real network, the link k overflow probability large deviation function (or rate function) is given by:

$$I_k^R = \inf\left\{\sum_{i\in\mathcal{M}} I_1^{X^i}(x_i) : x = (x_i)_{i\in\mathcal{M}}, \sum_{i\in\mathcal{M}} g_k^i(x) \geq C_k\right\} \tag{9}$$

In the fictitious network this function is given by

$$I_k^F = \inf\left\{\sum_{i\in\mathcal{M}_k} I_1^{X^i}(x_i) : x = (x_i)_{i\in\mathcal{M}_k}, \sum_{i\in\mathcal{M}_k} x_i \geq C_k\right\} \tag{10}$$

In the following it is assumed that each traffic type is an aggregate of N $i.i.d$ sources. This implies that each rate function $I_1^{X^i}$ is convex and $I_1^{X^i}(\mu_i) = 0$ for all i. Then, (9) and (10) are convex optimization problems under constraints. The second one has the advantage that the constraints are linear and there are well known fast methods to solve it. The functions $I_1^{X^i}$ are continuous, so we solve the following problems corresponding to the real and fictitious network respectively.

$$P_R \begin{cases} \min \sum_{i \in \mathcal{M}} I_1^{X^i}(x_i) \\ \\ \sum_{i \in \mathcal{M}} g_k^i(x) \geq C_k \end{cases} \qquad P_F \begin{cases} \min \sum_{i \in \mathcal{M}_k} I_1^{X^i}(x_i) \\ \\ \sum_{i \in \mathcal{M}_k} x_i \geq C_k \end{cases}$$

Definition 4.1. *Consider two optimization problems*

$$P_1 \begin{cases} \min f_1(x) \\ x \in D_1 \end{cases} \quad \textit{and} \quad P_2 \begin{cases} \min f_2(x) \\ x \in D_2 \end{cases}$$

P_2 is called a relaxation of P_1 if $D_1 \subseteq D_2$ and $f_2(x) \leq f_1(x)$, $\forall x \in D_1$.

Proposition 4.2. *If P_2 is a relaxation of P_1 and x_2 is optimum for P_2 such $x_2 \in D_1$ and $f_2(x_2) = f_1(x_2)$, then x_2 is optimum for P_1.*

PROOF. $f_1(x_2) = f_2(x_2) \leq f_2(x) \leq f_1(x)\ \forall x \in D_1 \subseteq D_2$, so x_2 is optimum for P_1 because it minimizes f_1 and belongs to D_1.

Proposition 4.3. *P_F is a relaxation of P_R.*

PROOF. Since the functions $I_1^{X^i}$ are non negatives, it is clear that $\sum_{i \in \mathcal{M}_k} I_1^{X^i}(x_i) \leq \sum_{i \in \mathcal{M}} I_1^{X^i}(x_i)\ \forall x = (x_i)_{i \in \mathcal{M}}$. Then, we have to prove that

$$\left\{ x : \sum_{i \in \mathcal{M}} g_k^i(x) \geq C_k \right\} \subseteq \left\{ x : \sum_{i \in \mathcal{M}_k} x_i \geq C_k \right\}$$

By definition, $g_k^i(x) = 0\ \forall\ i \notin \mathcal{M}_k$ and $g_k^i(x) \leq x_i\ \forall\ i \in \mathcal{M}_k$ (since g_k^i can be written as composition of functions of type (6)) then

$$\sum_{i \in \mathcal{M}} g_k^i(x) = \sum_{i \in \mathcal{M}_k} g_k^i(x) \leq \sum_{i \in \mathcal{M}_k} x_i$$

and therefore $\sum_{i \in \mathcal{M}_k} g_k^i(x) \geq C_k$, implies $\sum_{i \in \mathcal{M}_k} x_i \geq C_k$.

Remark 4.1. If an optimum of the fictitious problem P_F verifies the real problem P_R constraints and the objective functions take the same value at this point, then it is an optimum of the real problem too.

Remark 4.2. The optimality conditions (KKT [2]) for the fictitious problem P_F are the following:

1. $\nabla \left(\sum_{i \in \mathcal{M}_k} I_1^{X^i}(x_i) + \lambda(C_k - \sum_{i \in \mathcal{M}_k} x_i) \right) = 0$, with λ Lagrange multiplier.

[2] Karush-Khum-Tucker [11]

2. $\lambda \geq 0$.

3. $\sum_{i \in \mathcal{M}_k} x_i \geq C_k$.

4. $\lambda \left(C_k - \sum_{i \in \mathcal{M}_k} x_i \right) = 0$.

First and second condition imply that:

$$\frac{\partial I_1^{X^i}}{\partial x_i}(x_i) = \lambda \geq 0 \quad \forall\, i \in \mathcal{M}_k$$

If $\lambda = 0$, $x_i = \mu_i\ \forall i$. In this case $\sum_{i \in \mathcal{M}_k} I_1^{X^i}(\mu_i) = 0$ and it is not considered. Then we supposed that $\frac{\partial I_1^{X^i}}{\partial x_i}(x_i) > 0$, which implies $x_i > \mu_i$. Finally, since $\lambda \neq 0$, the last condition implies that:

$$C_k - \sum_{i \in \mathcal{M}_k} x_i = 0$$

Then, $\widetilde{x} = (\widetilde{x}_i)_{i \in \mathcal{M}_k}$ optimum for P_F verifies:

$$\begin{cases} \widetilde{x}_i > \mu_i \quad \forall\, i \in \mathcal{M}_k \\ \sum_{i \in \mathcal{M}_k} \widetilde{x}_i = C_k \end{cases} \tag{11}$$

The following theorem gives conditions over the network to assure that link k overflow probability rate function for the real and for the fictitious network are equal ($E = I_k^R - I_k^F = 0$). Since the network is feed forward, it is possible to establish an order between the links. We say that link i is "previous to" or "less than" link j if for one path, link i is found before than link j in the flow direction.

Theorem 4.4 (Sufficient Condition). *If $\widetilde{x} = (\widetilde{x}_i)_{i \in \mathcal{M}_k}$ is optimum for P_F, and the following condition is verified for all links i less than k:*

$$C_k - \sum_{j \in \mathcal{M}_k \setminus \mathcal{M}_i} \mu_j \leq C_i - \sum_{j \in \mathcal{M}_i \setminus \mathcal{M}_k} \mu_j \quad \forall\, i < k \tag{12}$$

then x^* defined by:

$$(x^*)_i = \begin{cases} \widetilde{x}_i & \textit{if } i \in \mathcal{M}_k \\ \mu_i & \textit{if } i \notin \mathcal{M}_k \end{cases}$$

is optimum for P_R.

PROOF. The objective functions of the optimization problems (9) and (10) take the same values at x^* because $I_1^{X^i}(\mu_i) = 0 \quad \forall i$:

$$\sum_{i\in\mathcal{M}} I_1^{X^i}(x_i^*) = \sum_{i\in\mathcal{M}_k} I_1^{X^i}(\widetilde{x}_i) + \sum_{i\in\mathcal{M}\setminus\mathcal{M}_k} I_1^{X^i}(\mu_i) = \sum_{i\in\mathcal{M}_k} I_1^{X^i}(\widetilde{x}_i)$$

Considering proposition (4.3), it is enough to prove that x^* satisfy the real problem constraints:

$$\sum_{i\in\mathcal{M}} g_k^i(x_i^*) \geq C_k$$

By definition $g_k^i(x^*) = 0 \ \forall\ i \notin \mathcal{M}_k$. Moreover the function g_k^i can be written as composition of function of type (6), so if $\sum_{j\in\mathcal{M}_i}(x_j^*) \leq C_i \ \forall\ i$, then $g_k^i(x^*) = (x^*)_i$ $\forall\ i \in \mathcal{M}_k$ and

$$\sum_{i\in\mathcal{M}} g_k^i(x^*) = \sum_{i\in\mathcal{M}_k}(x^*)_i = \sum_{i\in\mathcal{M}_k}\widetilde{x}_i = C_k$$

proving the theorem. In the last equality we use that $\widetilde{x}$ verifies (11), since it is optimum for P_F. Then, it is sufficient to prove that $\sum_{j\in\mathcal{M}_i}(x^*)_j \leq C_i \quad \forall\ i < k$. Separating the sum,

$$\sum_{j\in\mathcal{M}_i}(x^*)_j = \sum_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j + \sum_{j\in\mathcal{M}_i\setminus\mathcal{M}_k}\mu_j \leq C_i \quad \forall\ i < k \tag{13}$$

and then we have to guarantee that

$$\sum_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j \leq C_i - \sum_{j\in\mathcal{M}_i\setminus\mathcal{M}_k}\mu_j \quad \forall\ i < k$$

Since $\widetilde{x}$ is optimum for P_F, it satisfy $C_k = \sum_{j\in\mathcal{M}_k}\widetilde{x}_j$, and therefore

$$\sum_{j\in\mathcal{M}_k\cap\mathcal{M}_i}\widetilde{x}_j = C_k - \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}\widetilde{x}_j$$

Also, $\widetilde{x}_j > \mu_j \ \forall\ j \in \mathcal{M}_k$

$$\sum_{j\in\mathcal{M}_k\cap\mathcal{M}_i}\widetilde{x}_j \leq C_k - \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}\mu_j$$

Using the hypothesis, we have that:

$$\sum_{j\in\mathcal{M}_k\cap\mathcal{M}_i}\widetilde{x}_j = C_k - \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}\mu_j \leq C_i - \sum_{j\in\mathcal{M}_i\setminus\mathcal{M}_k}\mu_j \quad \forall\ i < k$$

which proves (13) and the theorem.

Example 4.5. *Consider a network like in figure 1. We analyze the overflow probability at link k.*
If condition (12) is attained for link k, then $E = I_k^R - I_k^F = 0$. This condition is:

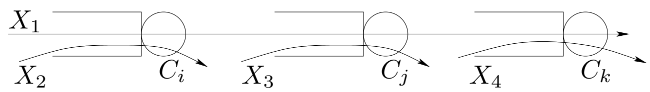


Figure 1:

$$\begin{cases} C_k - \mu_4 \le C_i - \mu_2 \\ \\ C_k - \mu_4 \le C_j - \mu_3 \end{cases}$$

4.1. Sufficient condition in terms of available bandwidth

Definition 4.6. *For a traffic type m in a link j, it is defined the available bandwidth ABW_j^m as the difference between the link j capacity and the mean value of the transmission rate of the other traffic types in j.*

In terms of the previous definition, the theorem condition (12) assures that the overflow probability rate function at link k on real and fictitious network are the same if for all link $j < k$, and for all m traffic type in $\mathcal{M}_j \cap \mathcal{M}_k$, $ABW_j^m > ABW_k^m$. This condition is represented in figure 2 for a simple network with two links.

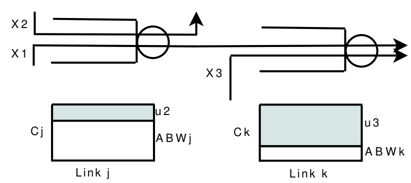


Figure 2:

4.2. Sufficient but not necessary condition

The theorem condition (12) is sufficient to assure that the overflow probability rate function at link k on real and fictitious networks are the same, but it is not a necessary condition. In fact, if $\widetilde{x}$ is optimum for the fictitious problem, and if x^* defined as:

$$(x^*)_i = \begin{cases} \widetilde{x}_i & \text{si } i \in \mathcal{M}_k \\ \mu_i & \text{si } i \notin \mathcal{M}_k \end{cases} \tag{14}$$

satisfies the real problem constraints, then x^* is optimum for the real problem. If x^* verifies the following condition

$$\sum_{j \in \mathcal{M}_i} (x^*)_j \le C_i \quad \forall\, i < k \tag{15}$$

it also verifies the real problem constraints and therefore is optimum for the real problem.

Therefore, in the case that the theorem condition is not fulfilled, if we found $\widetilde{x}$ optimum for the fictitious problem, then is easy to check if the rate functions are equal or no. It is enough to check (15), where x^* is defined in (14).

4.3. Error bound

Since the functions $I_1^{X^i}$ are non negatives, it is clear that the rate function for the real problem is always greater than the fictitious one. Then the error $E = I_k^R - I_k^F$ is always non negative. This implies that the fictitious network overestimates the overflow probability. We are interested in finding an error bound for the overestimation of the fictitious analysis when conditions (12) and (15) are not satisfied. A simple way to get this bound is to find a point x which verifies the real problem constraints. In this case, we have that:

$$E = I_k^R - I_k^F \leq \sum_{i \in \mathcal{M}} I_1^{X^i}(x_i) - \sum_{i \in \mathcal{M}_k} I_1^{X^i}(\widetilde{x}_i)$$

To assure that x verifies the real problem constraints, we have already seen that it is enough to show that $\sum_{j \in \mathcal{M}_i} x_j \leq C_i \; \forall \; i < k$ and $\sum_{j \in \mathcal{M}_k} x_j \geq C_k$. Therefore, we have to solve this inequalities system. From remark (4.2), it can be seen that the optimum of the fictitious problem is in the boundary of the feasible region ($\sum_{i \in \mathcal{M}_k} \widetilde{x}_i = C_k$). Since we are looking for a point near the optimum of the fictitious problem in the sense that the error bound be as small as possible, we solve the following system:

$$\left\{ \begin{array}{l} \sum_{j \in \mathcal{M}_i} x_j \leq C_i \quad \forall \; i < k \\ \\ \sum_{j \in \mathcal{M}_k} x_j = C_k \end{array} \right. \tag{16}$$

For the interesting cases, where there are losses at link k, this system always has a solution. In the following an algorithm to find a solution of this system is defined. We define the following point:

$$(x^*)_j = \left\{ \begin{array}{cl} \widetilde{x}_j & \text{if } \; j \in \mathcal{M}_k \\ 0 & \text{if } \; j \notin \mathcal{M}_k \end{array} \right.$$

If x^* verifies the conditions (16), we find a point that verifies the real problem constraints. In some cases this is not useful because $I_1^{X^j}(0) = \infty$ and we have that the error bound is infinite. If $P(X_1^{j,N} \leq 0) \neq 0$, the function $I_1^{X^j}(0) < \infty$ and a finite error bound is obtained. If x^* is not solution for system (16), then we redefine (by some small value) the coordinates where $\sum_{j \in \mathcal{M}_i} x_j > C_i$ in such a way that $\sum_{j \in \mathcal{M}_i} x_j = C_i$. The second equation must be verified too and, since some coordinates were reduced, others coordinates have to increase to get the total sum equal to C_k. Since the system is compatible, following this method, a solution is always found. There is no guarantee that the solution given by this method minimizes the error bound. However, this method has a very simple implementation and gives reasonable error bounds as we can see in the numerical examples of the last section.

4.4. Error bound in a particular case

We analyze a particular case in which the following conditions are verified:

1. $\mathcal{M}_i \backslash \mathcal{M}_k \neq \emptyset \; \forall \; i < k$ this means that for all link i less than k, there exists at least one traffic type going through link i and not arriving at link k.

2. $C_i - \left(C_k - \sum_{j \in \mathcal{M}_k \backslash \mathcal{M}_i} \mu_j \right) \geq 0 \quad \forall \; i.$

Consider $\widetilde{x} = (\widetilde{x}_i)_{i \in \mathcal{M}_k}$ optimum for (P_F) and x^* defined by:

$$(x^*)_j = \begin{cases} \widetilde{x}_j & \text{if } \; j \in \mathcal{M}_k \\ x_j^* & \text{if } \; j \notin \mathcal{M}_k \end{cases} \tag{17}$$

If x^* verify the real problem constraints, the following error bound is obtained:

$$\begin{aligned} E &\leq \sum_{i \in \mathcal{M}} I_1^{X^i}((x^*)_i) - \sum_{i \in \mathcal{M}_k} I_1^{X^i}(\widetilde{x}_i) \\ &= \sum_{i \in \mathcal{M}_k} I_1^{X^i}(\widetilde{x}_i) + \sum_{i \mathcal{M} \backslash \mathcal{M}_k} I_1^{X^i}(x_i^*) - \sum_{i \in \mathcal{M}_k} I_1^{X^i}(\widetilde{x}_i) \\ &= \sum_{i \in \mathcal{M} \backslash \mathcal{M}_k} I_1^{X^i}(x_i^*) \end{aligned} \tag{18}$$

By definition of x^* and the optimality conditions for the fictitious problem, it follows that:

$$\sum_{j \in \mathcal{M}_k} (x^*)_j = \sum_{j \in \mathcal{M}_k} (\widetilde{x})_j = C_k$$

Therefore, to prove that x^* verify the real problem constraints (16), it is enough to show that x^*, verify:

$$\sum_{j \in \mathcal{M}_i} (x^*)_j = \sum_{j \in \mathcal{M}_i \cap \mathcal{M}_k} \widetilde{x}_j + \sum_{j \in \mathcal{M}_i \backslash \mathcal{M}_k} x_j^* \leq C_i \quad \forall \; i \tag{19}$$

In this particular case, by the second condition, it is possible to define x_j^* for $j \in \mathcal{M}_i \backslash \mathcal{M}_k$ such that

$$\sum_{j \in \mathcal{M}_i \backslash \mathcal{M}_k} x_j^* \leq C_i - \left(C_k - \sum_{j \in \mathcal{M}_k \backslash \mathcal{M}_i} \mu_j \right) \quad \forall \; i$$

and therefore

$$\sum_{j \in \mathcal{M}_i} (x^*)_j \leq \sum_{j \in \mathcal{M}_i \cap \mathcal{M}_k} \widetilde{x}_j + C_i - C_k + \sum_{j \in \mathcal{M}_k \backslash \mathcal{M}_i} \mu_j$$

On the other hand, since $\widetilde{x}$ is optimum for P_F, $\sum_{j \in \mathcal{M}_k} \widetilde{x}_j = C_k$ and

$$\sum_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j = C_k - \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}\widetilde{x}_j$$

Replacing in the previous equation results

$$\begin{aligned}\sum_{j\in\mathcal{M}_i}(x^*)_j &\leq C_k - \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}\widetilde{x}_j + C_i - C_k + \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}\mu_j \\ &= C_i + \sum_{j\in\mathcal{M}_k\setminus\mathcal{M}_i}(-\widetilde{x}_j+\mu_j) < C_i\end{aligned}$$

since from (11), $\widetilde{x}_j > \mu_j \quad \forall\, j\in\mathcal{M}_k$.

Then x^* verifies (16) and therefore is optimum for the real problem. The error bound obtained is (18). We can found $(x^*)_{j\in\mathcal{M}\setminus\mathcal{M}_k}$ such that, the error bound (18) be minimum in the set of $(x^*)_{j\in\mathcal{M}}$ defined in (17) that verifies the real problem constraints. It is necessary to solve the following convex optimization problem:

$$\begin{cases}\min \sum\limits_{i\in\mathcal{M}\setminus\mathcal{M}_k} I_1^{X^i}(x_i) \\ \\ \sum\limits_{j\in\mathcal{M}_i\setminus\mathcal{M}_k} x_j \geq C_i - \sum\limits_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j \quad \forall\, i=1,\cdots,L\end{cases}$$

Once again it is sufficient to find $(x^*)_{j\in\mathcal{M}}$ that verifies the KKT optimality conditions:

1. $\frac{\partial}{\partial x_j}I_1^{X^j}(x_j) + \sum\limits_{i\in k^j}\lambda_i = 0\;\forall\, j\in\mathcal{M}\setminus\mathcal{M}_k$.
2. $\lambda_i \geq 0\;\forall\, i$.
3. $\sum\limits_{j\in\mathcal{M}_i\setminus\mathcal{M}_k} x_j \geq C_i - \sum\limits_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j \quad \forall\, i$.
4. $\lambda_i\left(\sum\limits_{j\in\mathcal{M}_i\setminus\mathcal{M}_k} x_j - \left(C_i - \sum\limits_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j\right)\right) = 0\;\forall\, i$.

We will define an algorithm to find such point. If for $j\in\mathcal{M}\setminus\mathcal{M}_k$, there is a link $i\in k^j$ that verifies

$$C_i - \sum_{h\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_h \leq \sum_{h\in\mathcal{M}_i\setminus\mathcal{M}_k}\mu_h$$

we define x^* as follows:

$$\sum_{h\in\mathcal{M}_i\setminus\mathcal{M}_k} x_h = C_i - \sum_{j\in\mathcal{M}_i\cap\mathcal{M}_k}\widetilde{x}_j$$

This determines a linear equations system that always has a solution, but it can be undetermined. The choice in that case it is not important because the optimum obtained is global. For the coordinates $j\in\mathcal{M}\setminus\mathcal{M}_k$ that are not determined with the previous equations, we define $x_j = \mu_j$. It is easy to check that $(x^*)_{j\in\mathcal{M}\setminus\mathcal{M}_k}$ defined by this algorithm verifies the KKT optimality conditions. Then we have defined an algorithm that gives the minimum error bound for this particular case.

5. End-to-End Loss Ratio Evaluation

In the previous section we found sufficient conditions to assure that results on the fictitious and on the real network analysis coincide for an interior link. However, to define an admission control mechanism based on the end-to-end quality of service, we must find a condition that guarantees that the end-to-end loss ratio coincides for both networks. A natural answer is that the sufficient condition found in theorem 4.4 must be verified for all links in the considered path. However, as equation 7 suggest, we will show that this is not necessary since it is enough that the sufficient condition is verified for the link with minimum overflow probability rate function. This link must be then identified, and clearly we aim to do it within the fictitious network context. We must then be sure that the link with minimum rate function is the same for the real and the fictitious network. In the sequel we address this two issues.

Proposition 5.1. *Let k_f be the link with minimum overflow probability rate function in the fictitious network for traffic type m:*

$$\overline{I}_{k_f} = \min_{k_i \in \mathbf{k}^m} \overline{\mathbf{I}}_{k_i}$$

If the conditions of theorem 4.4 are verified for link k_f, the minimum overflow probability rate function for traffic type m in the real network is also attained at link k_f.

PROOF. By definition of k_f and the already known relation between rate functions in the real and the fictitious network, results that:

$$\overline{I}_{k_f} \leq \overline{I}_{k_i} \leq I_{k_i} \quad \forall k_i \in \mathbf{k}^m$$

Since we assume that conditions of theorem 4.4 are verified for link k_f, the rate functions of both network coincide i.e. $I_{k_f} = \overline{I}_{k_f}$. By replacing $\overline{I}_{k_f}$ in the previous equation, we obtain that:

$$I_{k_f} \leq I_{k_i} \; \forall k_i \in \mathbf{k}^m$$

which completes the proof.

Proposition 5.2. *Let k be the link where $I_k = -\min_{k \in k^m} \mathbf{I}_k^m$ for the real network , i.e. the link where the minimum rate function of traffic type m is attained. Let $\overline{I}_k$ be the rate function of the same link k for the fictitious network. If the sufficient conditions of theorem 4.4 are verified for link k then $\mathbf{L}^m = \overline{\mathbf{L}}^m$, i.e. the end-to-end loss ratio for real and fictitious network coincide.*

PROOF. By equation 7 and proposition 5.1, we obtain that:

$$\begin{aligned}
\mathbf{L}^m &= \lim_{N\to\infty} \frac{1}{N} \log \mathbf{L}^{m,N} = -\min_{k_i \in k^m} \mathbf{I}_{k_i} = -\mathbf{I}_k \\
&= -\overline{\mathbf{I}}_k = -\min_{k_i \in k^m} \overline{\mathbf{I}}_{k_i} = \lim_{N\to\infty} \frac{1}{N} \log \overline{\mathbf{L}}^{m,N} = \overline{\mathbf{L}}^m
\end{aligned}$$

Remark 5.1. Previous propositions show that to evaluate the end-to-end loss ratio $\mathbf{L}^m$, it is enough that sufficient conditions of theorem 4.4 are verified by the link k where the minimum rate function of traffic type m path is attained. In this case, it results that $\mathbf{L}^m = \overline{\mathbf{L}}^m = \overline{I}_k$. If sufficient conditions are not verified, then the error bound obtained for the one link case can be applied.

6. Rate function estimation

In previous sections we show how to evaluate the end-to-end loos ratio in terms of the rate function for the fictitious network. In order to implement an on-line admission control based on this information, we must be able to accurately estimate the corresponding rate function. In this section we analyze how this estimation can be done using traffic traces of the input traffic.

Let $X_k^{m,N}(0,t)$ be the traffic type m workload at link k during the time interval $(0,t)$. We suppose that $X_k^{m,N}$ is the sum of $N\rho_m$ independent sources of type m:

$$X_k^{m,N}(0,t) = \sum_{i=1}^{N\rho_m} \widetilde{X}_k^{m,i}(0,t)$$

In this case, the instantaneous rate of traffic type m at time t is given by:

$$X_{k,t}^{m,N} = \sum_{i=1}^{N\rho_m} \widetilde{X}_{k,t}^{m,i}$$

Given the stationarity of the traffic, we can replace the t-index by 0 and for simplicity we omit the link index k. Then the instantaneous rate of total input traffic at link k is:

$$Z_0^N = \sum_{m\in\mathcal{M}_k} X_0^{m,N} = \sum_{m\in\mathcal{M}_k}\sum_{i=1}^{N\rho_m} \widetilde{X}_0^{m,i} = \sum_{j=1}^{N} \widetilde{Z}^j$$

where the variables $\widetilde{Z}^j$ are independent and identically distributed (*iid*) random variables. Each variable $\widetilde{Z}^j$ has the distribution of a mix of the variables $\widetilde{X}_0^{m,i}$ (given by the proportions ρ_m of each traffic type m present at link k). This means that instantaneous rate of input traffic at link k is the sum of N *iid* random variables and Cramer theorem (see for example [17]) can be applied. The variable $\frac{Z_0^N}{N}$ verifies then a large deviation principle with rate function:

$$I_t^Z(x) = \sup_{\theta\geq 0}\{\theta x - \Lambda(\theta)\} = \sup_{\theta\geq 0}\{\theta x - \log \mathbf{E}\left(e^{\theta \widetilde{Z}^1}\right)\} \tag{20}$$

Given the rate function of the LDP, $I_t^Z(x)$, we can calculate I_k^F:

$$\begin{aligned} I_k^F &= \inf\left\{I^Z(z) : z \geq C_k\right\} = \inf_{z\geq C_k}\sup_{\theta\geq 0}\left\{\theta z - \Lambda(\theta)\right\} \\ &= \sup_{\theta\geq 0}\left\{\theta C_k - \Lambda(\theta)\right\} \end{aligned} \tag{21}$$

Before solving the optimization problem 21, we must calculate or estimate $\Lambda(\theta)$. If some model is assumed for the traffic, $\Lambda(\theta)$ can be calculated analytically. In case no model is assumed as in our case, it must be estimated from measurements i.e. from traffic traces. A possible and widely used approach [6, 15] is to estimate the expectation as a temporal average of a given traffic trace $\{\widetilde{Z}^N(t)\}_{t=1:n}$:

$$\mathbf{E}\left(e^{\theta \widetilde{Z^1}}\right) = \mathbf{E}\left(e^{\theta \frac{Z_0^N}{N}}\right) \approx \frac{1}{n}\sum_{t=1}^{n} e^{\theta Z^N(t)/N}$$

Then $\Lambda(\theta)$ can be estimated by

$$\Lambda_n(\theta) = \log\left(\frac{1}{n}\sum_{t=1}^{n} e^{\theta Z^N(t)/N}\right)$$

Now, the rate function I_k^F can be estimated as:

$$I_{k,n}^F = \sup_{\theta \geq 0}\{\theta C_k - \Lambda_n(\theta)\}$$

However it remains unclear how good is this estimation. We will show that if $\Lambda_n(\theta)$ is a *good* estimator of $\Lambda(\theta)$, then $I_{k,n}^F$ is also a *good* estimator for the rate function I_k^F.

Theorem 6.1. *If $\Lambda_n(\theta)$ is an estimator of $\Lambda(\theta)$ such that both are C^1 functions and:*

$$\Lambda_n(\theta) \underset{n}{\longrightarrow} \Lambda(\theta)$$

$$\frac{\partial}{\partial\theta}\Lambda_n(\theta) \underset{n}{\longrightarrow} \frac{\partial}{\partial\theta}\Lambda(\theta)$$

where the convergence is almost surely and uniformly over bounded intervals, then $I_{k,n}^F$ is a consistent estimator of I_k^F. Moreover, if a functional Central Limit Theorem (CLT) applies to $\Lambda_n - \Lambda$, i.e,

$$\sqrt{n}\left(\Lambda_n(\theta) - \Lambda(\theta)\right) \underset{n}{\overset{w}{\Longrightarrow}} G(\theta)$$

where $G(\theta)$ is a continuous gaussian process, then:

$$\sqrt{n}\left(I_{k,n}^F - I_k^F\right) \underset{n}{\overset{w}{\Longrightarrow}} N(0,\sigma)$$

where $N(0,\sigma)$ is a centered normal distribution with variance σ.

PROOF. This theorem can be proved following the same ideas of the proof of Theorem 1. of [1].

Theorem 6.1 assures that under certain conditions we can obtain a consistent estimator of the rate function and then of the overflow probability and the end-to-end loss ratio. This conditions are verified for a large family of estimators like for example the time average estimator proposed here or Markov fluid sources estimators [14]. In addition to the consistence, a CLT is proved to be valid which allows the calculus of confidence intervals for the rate functions. From the previous analysis we conclude that the rate function and then the admission control region can be accurately estimated form traffic traces in a simple way. As we claimed before, this can be used in the definition of an admission control mechanism based in the end-to-end quality of service expected by the traffic.

7. Numerical examples

In this section we present some numerical examples to validate our results. There are many issues that could be evaluated to analyze the performance of an admission control mechanism. However, since the overall performance of our proposition depends on how accurate are the results obtained when the fictitious network model is considered, we will concentrate here only in this aspect.

Example 7.1. *Consider a network like in figure 3. We analyze the overflow probability at link k, assuming that $C_i > C_k$.*

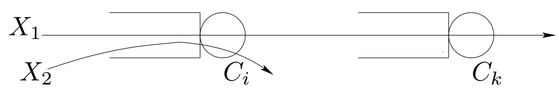


Figure 3: Example 7.1-Network topology

If condition (12) is attained for link k, then $E = I_k^R - I_k^F = 0$. This condition is:

$$C_k \leq C_i - \mu_2$$

If this condition is not satisfied, since $\widetilde{x} = C_k$ is optimum for P_F, we first verify if $x^* = (C_k, \mu_2)$ is optimum for (P_R). It is sufficient to show that x^* verifies the real problem constraints, i.e:

$$\begin{cases} C_k + \mu_2 \leq C_i \\ C_k = C_k \end{cases}$$

If $C_k + \mu_2 > C_i$, we look for $x^* = (x_1^*, x_2^*)$ that verifies

$$\begin{cases} x_1^* + x_2^* \leq C_i \\ x_1^* = C_k \end{cases}$$

It is possible to choose $x_1^* = C_k$ and $x_2^* = C_i - C_k > 0$ resulting in the following error bound:

$$E \leq I_1(C_k) + I_2(C_i - C_k) - I_1(\widetilde{x}_1) = I_2(C_i - C_k) \tag{22}$$

In the following numerical example, we calculate the overflow probability rate function for the real and fictitious network. Let $C_i = 16kb/s$ per source and C_k growing from 4 to $15.5kb/s$ per source. All traffic sources are on-off Markov processes. For X_1, the bit rate in the on state is $16kb/s$, and average times are $0.5s$ in the on state and $1.5s$

in the off state. For X_2, the bit rate in the on state is $16kb/s$, and average times are $1s$ in the on state and $1s$ in the off state. Since $\mu_1 = 4kb/s$ the stability condition is $C_k > \mu_1 = 4kb/s$. Using these values, the sufficient condition (12) is, $C_k \leq 8kb/s$. Figure 4 shows that while this condition is satisfied both functions match, but after $C_k \geq 8kb/s$ they separate. Figure 4 also shows the overestimation error ($E = I_k^R - I_k^F$) and the error bound (22) described before. In this case, the error bound is exactly the error.

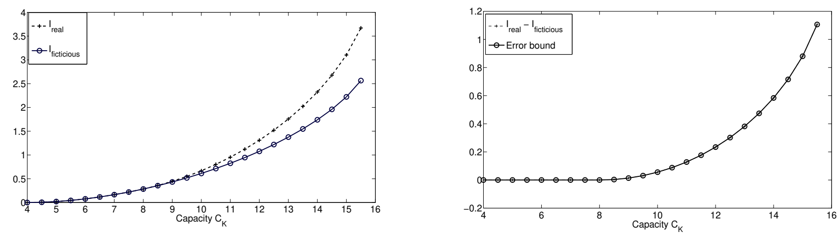


Figure 4: Example 7.1-Rate functions and error bound

Example 7.2. *Consider a network like in figure 5. We analyze the overflow probability at link k.*

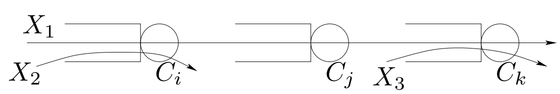


Figure 5: Example 7.2-Network topology

If condition (12) is attained for link k, then $E = I_k^R - I_k^F = 0$. This condition is:

$$\begin{cases} C_k - \mu_3 \leq C_i - \mu_2 \\ C_k - \mu_3 \leq C_j \end{cases}$$

If this condition is not satisfied, and $\widetilde{x} = (\widetilde{x}_1, \widetilde{x}_3)$ is optimum for P_F, we first verify if $x^* = (\widetilde{x}_1, \mu_2, \widetilde{x}_3)$ is optimum for P_R. It is sufficient to show that x^* verifies the real problem, i.e:

$$\begin{cases} \widetilde{x}_1 + \mu_2 \leq C_i \\ \widetilde{x}_1 \leq C_j \\ \widetilde{x}_1 + \widetilde{x}_3 = C_k \end{cases}$$

If these conditions are not satisfied, we look for $x^* = (x_1^*, x_2^*, x_3^*)$ that satisfies:

$$\begin{cases} x_1^* + x_2^* \leq C_i \\ x_1^* \leq C_j \\ x_1^* + x_3^* = C_k \end{cases}$$

We choose $x_1^* = \min(\widetilde{x}_1, C_i, C_j)$. Three different cases are identified. For the first case $x_1^* = \widetilde{x}_1$, we choose:

$$\begin{cases} x_1^* = \widetilde{x}_1 \\ x_2^* = C_i - \widetilde{x}_1 \\ x_3^* = C_k - \widetilde{x}_1 \end{cases} \tag{23}$$

In this case the error bound is:

$$\begin{aligned} E &\leq I_1(\widetilde{x}_1) + I_2(C_i - \widetilde{x}_1) + I_3(C_k - \widetilde{x}_1) - I_1(\widetilde{x}_1) - I_3(\widetilde{x}_3) \\ &= I_2(C_i - \widetilde{x}_1) + I_3(C_k - \widetilde{x}_1) - I_3(\widetilde{x}_3) \end{aligned} \tag{24}$$

Using that $\widetilde{x}_1 + \widetilde{x}_3 = C_k$, we have another possibility for determining an error bound. We can rewrite the first equation as $C_k - \widetilde{x}_3 + x_2^* \leq C_i$. And, since $\widetilde{x}_3 > \mu_3$ we can choose $x_2^* = C_i - (C_k - \mu_3)$ (or any lower value). In this case the error bound is:

$$\begin{aligned} E &\leq I_1(\widetilde{x}_1) + I_2(C_i - (C_k - \mu_3)) + I_3(C_k - \widetilde{x}_1) - I_1(\widetilde{x}_1) - I_3(\widetilde{x}_3) \\ &= I_2(C_i - (C_k - \mu_3)) + I_3(C_k - \widetilde{x}_1) - I_3(\widetilde{x}_3) \end{aligned} \tag{25}$$

The best error bound depends on the relative position of the points $C_i - (C_k - \mu_3)$ and $C_i - \widetilde{x}_1$. Since $C_i - \widetilde{x}_1 \leq C_i - (C_k - \mu_3)$, if both are less than μ_2 then the best error bound is (24).

For the second case $x_1^* = C_i$ ($C_i \leq C_j$), we choose:

$$\begin{cases} x_1^* = C_i \\ x_2^* = 0 \\ x_3^* = C_k - C_i \end{cases} \tag{26}$$

In this case the error bound is:

$$E \leq I_1(C_i) + I_2(0) + I_3(C_k - C_i) - I_1(\widetilde{x}_1) - I_3(\widetilde{x}_3) \tag{27}$$

For the last case $x_1^* = C_j$ ($C_j \leq C_i$), we choose:

$$\begin{cases} x_1^* = C_j \\ x_2^* = C_i - C_j \\ x_3^* = C_k - C_j \end{cases} \tag{28}$$

In this case the error bound is:

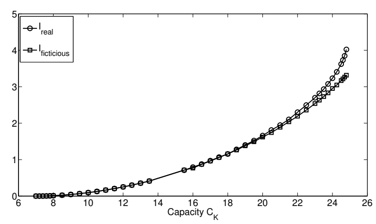
$$E \leq I_1(C_j) + I_2(C_i - C_j) + I_3(C_k - C_j) - I_1(\widetilde{x}_1) - I_3(\widetilde{x}_3) \tag{29}$$

For the following numerical example, we calculate the overflow probability rate function for the real and the fictitious network. Let $C_i = 5.5kb/s$, $C_j = 7kb/s$ per source and C_k ranging from 7 to $25kb/s$ per source. All traffic sources are on-off Markov processes. For X_1, the bit rate in the on state is $8kb/s$. For X_2, the bit rate in the on state is $10kb/s$. For X_3 the bit rate in the on state is $20kb/s$. The average time for all traffic types are $0.5s$ on the on state and $1.5s$ in the off state. Since $\mu_1 = 2kb/s$ and

$\mu_3 = 5kb/s$ the stability condition is $C_k > \mu_1 + \mu_3 = 7kb/s$. Using these values, the sufficient conditions (12) are:

$$\begin{cases} C_k \leq C_i - \mu_2 + \mu_3 = 8kb/s \\ C_k \leq C_j + \mu_3 = 12kb/s \end{cases}$$

The conditions are satisfied for values of C_k less than $8kb/s$. Figure 6 shows that both functions match even after the condition is not satisfied and up to $C_k \simeq 15kb/s$. The reason is that $x^* = (\widetilde{x}_1, \mu_2, \widetilde{x}_3)$ is optimum for P_R. From this point the functions begin to separate. Figure 6 also shows the functions I_k^R, I_k^F and $I_k^F + E'$, where E' is the error bound. Until $C_k = 24kb/s$, E' is calculated using (23), and then using (26). It is important to note that when $C_k > 14kb/s$ per source, the link utilization falls to less than 50% and therefore, as it can be seen in figure 6 the rate function $I_1^{X^k}$ takes values bigger than 0.5. If for example the number of sources feeding the network is $N = 100$, the losses are near 10^{-22}. Finally, we have seen that the estimated error bound is tight and when the error is big, the link overflow probability is small and, therefore, these links are not relevant for the QoS evaluation.



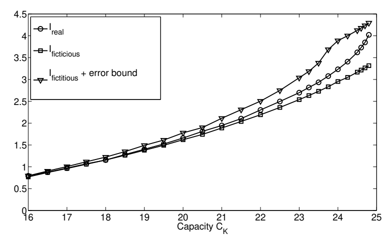


Figure 6: Example 7.2-Rate functions and error bound

Example 7.3. *Consider a network like in figure 7. We analyse the overflow probability at link k, assuming that $C_i + C_j > C_k$.*

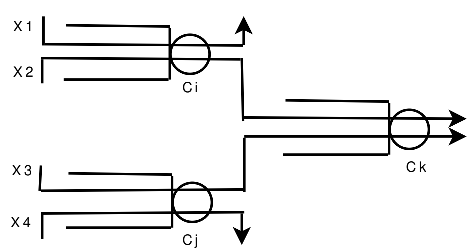


Figure 7: Example 7.3-Network topology

If condition (12) is attained for link k, then $E = I_k^R - I_k^F = 0$. This condition is:

$$\begin{cases} C_k - \mu_3 \leq C_i - \mu_1 \\ C_k - \mu_2 \leq C_j - \mu_4 \end{cases}$$

If this condition is not satisfied, since $\widetilde{x} = (\widetilde{x}_2, \widetilde{x}_3)$ is optimum for P_F, we first verify if $x^* = (\mu_1, \widetilde{x}_2, \widetilde{x}_3, \mu_4)$ is optimum for P_R. It sufficient to show that x^* verifies the real problem constraints, i.e:

$$\begin{cases} \widetilde{x}_2 + \mu_1 \leq C_i \\ \widetilde{x}_3 + \mu_4 \leq C_j \\ \widetilde{x}_2 + \widetilde{x}_3 = C_k \end{cases} \tag{30}$$

If these conditions are not satisfied, we look at first for $x^* = (x_1^*, \widetilde{x}_2, \widetilde{x}_3, x_4^*)$ that satisfies

$$\begin{cases} \widetilde{x}_2 + x_1^* \leq C_i \\ \widetilde{x}_3 + x_4^* \leq C_j \\ \widetilde{x}_2 + \widetilde{x}_3 = C_k \end{cases}$$

If $\widetilde{x}_2 > C_i$ or $\widetilde{x}_3 > C_j$ then it is not possible to choose such point. So, we look for $x^* = (x_1^*, x_2^*, x_3^*, x_4^*)$ that verifies

$$\begin{cases} x_1^* + x_2^* = C_i \\ x_3^* + x_4^* = C_j \\ x_1^* + x_3^* = C_k \end{cases}$$

One possible choice is:

$$\begin{cases} x_1^* = C_i \\ x_2^* = 0 \\ x_3^* = C_k - C_i \\ x_4^* = C_j - (C_k - C_i) \end{cases} \tag{31}$$

For the following numerical example, we calculate the overflow probability rate function for the real and fictitious network. Let $C_i = 12kb/s$, $C_j = 14kb/s$ per source and C_k growing from 8 to $25.5kb/s$ per source. All traffic sources are on-off Markov processes. For X_1, the bit rate in the on state is $20kb/s$. For X_2, the bit rate in the on state is $16kb/s$. For X_3, the bit rate in the on state is $16kb/s$. For X_4, the bit rate in the on state is $12kb/s$. The average times for all traffic types are $0.5s$ in the on state and $1.5s$ in the off state. Since $\mu_2 = 4kb/s$ and $\mu_3 = 4kb/s$ the stability condition is $C_k > \mu_1 + \mu_3 = 8kb/s$. Using these values, the sufficient condition (12) are:

$$\begin{cases} C_k < C_i - \mu_1 + \mu_3 = 11kb/s \\ C_k < C_j - \mu 4 + \mu_2 = 15kb/s \end{cases}$$

Figure 8 shows that both functions match even after the condition is not satisfied and up to $C_k \simeq 15kb/s$. The reason is that $x^* = (\mu_1, \widetilde{x}_2, \widetilde{x}_3, \mu_4)$ is optimum for the real problem. From this point the functions begin to separate.

Figure 8 also shows the functions I_k^R, I_k^F and $I_k^F + E'$, where E' is the error bound. Until $C_k = 24kb/s$, E' is calculated using (30), and then using (31). As in the previous example, when $C_k > 16kb/s$, the link utilization is less than 50% and that the error bound is tight for the relevant cases in the QoS evaluation.

All calculations of overflow probability rate functions were done with a software package developed by our group available in the web [4].

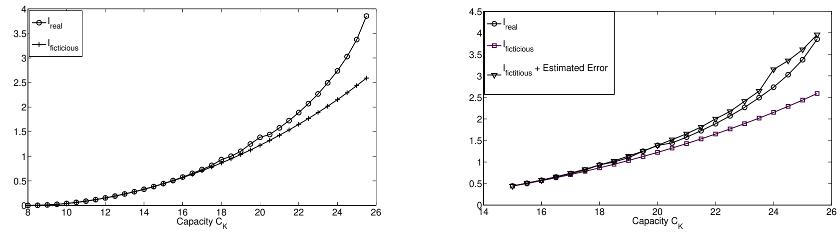


Figure 8: Example 7.3-Rate functions and error bound

8. Conclusions

In this paper we analyze how the fictitious network model can be applied to the definition of an admission control mechanism based on the end-to-end quality of service that a flow traversing the network will obtain. We have seen that the calculus of the overflow probability rate function of an interior network link is simpler and faster than the equivalent task in the real network. For this reason, on-line admission control systems will be easier to implement (or even feasible instead of impossible) using the fictitious network analysis.

While this approach is safe, network resources can be under-utilized. To solve this problem we find a condition, depending only on link capacities and mean traffic rates, to assure that the overflow probability calculated using the fictitious network has the same value that the one calculated in the real one. When this condition is not satisfied, the rate function of the link overflow probability calculated in the fictitious network can be smaller or equal than the same rate function calculated in the real network. We have shown that once the fictitious rate function is calculated, it is very simple to verify if both rate functions are equal or not. If they are not equal, a simple algorithm to find an error bound is described.

Previous results were extended to evaluate the end-to-end loss ratio. We find that this parameter can be calculated in terms of the moment generating function of the input traffic and we show how to estimate it from input traffic traces. We probe also that consistency and CLT properties of the moment generating function of the input traffic can be translated to the end-to-end loss ratio estimator through a natural procedure under very general hypothesis.

Finally, by means of numerical examples, we find that the error bound is tight. In these examples, it can be seen that when the error is big, the link overflow probability is very small and, therefore, these links are not relevant for the QoS evaluation. In spite of this, we can affirm that when the overflow probability at link k in the fictitious network is very small, even if the error is big, this link will not be considered for the QoS evaluation.

References

[1] L. Aspirot, P. Belzarena, P. Bermolen, A. Ferragut, G. Perera, and M. Simon, *Quality of service parameters and link operating point estimation based on effective bandwidths*, Performance Evaluation **59** (2005), no. 2-3, 103–120.

[2] P. Belzarena, P. Bermolen, M. Simon, and P. Casas, *Virtual path networks fast performance analysis*, ch. 17, pp. 359–386, River Publishers, 2009.

[3] L. Breslau, S. Jamin, and S. Shenker, *Comments on the performance of measurement-based admision control algorithm*, IEEE INFOCOM 2000 (Tel Aviv, Israel), 2000, pp. 1233–1242.

[4] D. Buschiazzo, A. Ferragut, A. Vazquez, and P. Belzarena, *ARCA*, http://iie.fing.edu.uy/investigacion/grupos/artes/arca, 2004.

[5] R. Casellas, *MPLS traffic engineering*, Ph.D. thesis, ENST, Paris, 2002.

[6] C. Courcoubetis and V. A. Siris, *Procedures and tools for analysis of network traffic measurements*, Performance Evaluation **48** (2002), no. 1/4, 5–23.

[7] C. Courcoubetis and R. Weber, *Buffer overflow asymptotics for a switch handling many traffic sources*, Journal of Applied Probability **33** (1996), 886–903.

[8] N.G. Duffield and N. O'Connell, *Large deviations and overflow probabilities for the general single server queue, with applications*, Math. Proc. Cam. Phil. Soc. **118** (1995).

[9] D. Y. Eun and N. Shroff, *Network decomposition: theory and practice*, IEEE/ACM Transactions on Networking **13** (2005), no. 3, 526–539.

[10] F.P. Kelly, *Notes on effective bandwidth*, Stochastic Networks: Theory and Applications, Royal Statistical Society Lecture Notes Series **4** (1996), 141–168.

[11] S. Mokhtar, D. Bazaraa Hanif, C. Sherali, and M. Shetty, *Nonlinear programming theory and algorithms*, John Wiley & Sons Inc, 1993.

[12] I. Más and G. Karlsson, *Probe-based admission control for a differentiated-services internet*, Computer networks **51** (2007), 3902–3918.

[13] O. Ozturk, R. Mazumdar, and N. Likhanov, *Many sources asymptotics in networks with small buffers*, Queueing Systems (QUESTA) **46** (2004), no. 1-2, 129–147.

[14] G. Perera, M. Simon, and J. Pechiar, *Effective bandwidth estimation and testing for markov sources*, Performance Evaluation **48** (2002), no. 1/4, 157–175.

[15] P. Rabinovitch, *Statistical estimation of effective bandwidth*, Master's thesis, University of Cambridge., 2000.

[16] A. Simonian and J. Guibert, *Large deviations approximations for fluid queues fed by a large number of on/off sources*, IEEE Journal on Selected Areas in Communications **13** (1995).

[17] D. Wischik, *Sample path large deviations for queues with many inputs*, Annals of Applied Probability **11** (2001), 379–404.

[18] Damon Wischik, *The output of a switch or effective bandwidths for networks*, Queueing Systems **32** (1999), no. 4, 383–396.