Computer Communications 24 (2001) 860–867

computer communications

www.elsevier.com/locate/comcom

On a congestion management scheme for high speed networks using aggregated large deviations principle

C. Ben Ahmed[a], N. Boudriga[a], M.S. Obaidat[b,*]

[a]Department of Networks, School of Communications, University of Carthage, Tunisia
[b]Department of Computer Science, Monmouth University, W. Long Branch, NJ 07764, USA

Received 20 July 2000; accepted 6 November 2000

Abstract

We consider a stochastic server fed by a set of parallel buffers offering dynamics evolving in discrete-time and following a service called "on demand processor sharing" (ODPS). This environment can be encountered in several situations in telecommunications including multiclass communication switches. In this paper, an analytic study of the multiclass model is proposed by using the concept of virtual queue that we have presented [Comput. Commun., 23 (2000) 912] and a multi level aggregation technique. The arrival process in each buffer is assumed to be arbitrary, and possibly auto-correlated, stochastic process. By using the large deviation principle to study the buffers' congestion, we provide a lower upper bound on the buffers overflow probabilities. We also consider the problem of finding a most likely sample path that leads to an overflow control problem, and give some applications of these results in the traffic engineering of high-speed networks.

Keywords: Agent-based rate coordination; Transmission control protocol; Asynchronous transfer mode; Available bit rate; Performance evaluation

1. Introduction

During the next few years, high speed networks and wireless networks will provide support for an even greater than today's varieties of traffic types including multimedia, real-time traffic, digitized voice, encoded video, and data. Some of these services are very sensitive to congestion phenomena, such as cell losses due to buffer overflows, and may require constrained quality of service (QoS) requirements that are difficult to be satisfied particularly when mobility is considered.

An essential step for preventing congestion through a variety of control mechanisms (e.g. buffer dimensioning, admission control, resource allocation, and re-routing tools) is to determine when it can occur by estimating its probability and to act to prevent congestion by choosing the appropriate technique.

In this paper, we consider network entities (such as switches and routers) that can support multiple service classes and provide different types of QoS for different traffic flows. We consider that a service class is characterized by the statistical properties of the incoming traffic and by its QoS requirements. The network entity has a dedicated buffer for each service class, each handling some number of media; and employing a service policy called the "on demand processor sharing" (ODPS) policy. This policy, also known as fair queuing, allocates at time n, a fraction $\theta_n^{j,i}$ of the available capacity to the jth media of class i. It is clear that, if θ_n^{ji} is the proportion of capacity allocated to the ith class, at instant n, then we have:

$$\theta_n^i = \sum_{j=1}^{N_n^i} \theta_n^{j,i}, \quad \sum_{j=1}^{L} \theta_n^i = 1, \ \forall n,$$

where N_n^i is the number of media of class i present at time n, and L is the number of classes.

In this paper, we address the problem of determining the buffer overflow probabilities for each class since these values can help the management of the QoS required by each class. Typical traffic in telecommunications networks is bursty. Thus, one can agree that stochastic processes with auto-correlation can be useful in modeling this traffic. The problem is particularly difficult to solve, since it essentially requires finding the distributions of waiting times and queue lengths in a multiclass $G/G/1$ setting with auto-correlated arrival processes and arbitrary service times.

In order to reduce such complexity, we will focus on the large deviations theory, determine the conditions for an efficient application of the large deviations principle, and

* Corresponding author. Tel.: +1-732-571-4482; fax: +1-732-263-5202.
E-mail address: obaidat@monmouth.edu (M.S. Obaidat).

0140-3664/01/$ - see front matter
PII: S0140-3664(00)00358-3

characterize computable asymptotic expressions for the overflow probabilities.

To this end, we will provide a lower upper bound on the buffer overflow probabilities (matching up to the first degree in the exponent). We will assume that we have L classes, and consider the exponent of the overflow probabilities as the optimal value of an appropriate optimal control problem, which we will explicitly solve and show that the optimal state trajectories of the control problem correspond to the most likely modes of overflow. Using the solution, we can obtain a detailed characterization of these modes [1].

Our results have important implications in the traffic management of high-speed networks, see Ref. [2] for QoS applications. Our results extend the deterministic worst-case analysis made in [3] by addressing the case where statistical measures of QoS are used to achieve more efficient utilization of the available resources. However, they can be used as a basis for an admission control mechanism, which provides class-dependent statistical QoS guarantees. The optimal control formulation was introduced in a somewhat more general setting by the authors of Ref. [4], where they made the emphasis on the analysis of a different scheduling policy for sharing resources (bandwidth) among media classes, which starts with the generalized longest queue first.

There is a growing interest in the literature for the applications of large deviations techniques in telecommunications systems, see Ref. [5], for a survey. The single class queue case has received extensive attention [6,7]. The extension of these ideas to single class networks, although much harder, has been treated in various versions and degrees of rigor in Refs. [5,8]. In Ref. [2], the authors obtain the asymptotic tails of the overflow probabilities for the Generalized Processor Sharing (GPS) policy with deterministic service capacity. The analysis they used was based on a large deviations result for the departure process from a $G/D/1$ queue [9]. More recently, Dupuis et al. [10] have developed a problem formulation for the large deviations analysis of the GPS policy in a different limiting regime.

In the sequel, we consider the ODPS policy in the presence of multiple media classes and a stochastic service capacity. This contributes to treat more complicated service disciplines. Consider for example the case where we have a deterministic server and three classes with dedicated buffers. We give priority to the first stream and use the GPS policy for the remaining two. The latter two streams face a server with stochastic capacity, a model of which can be obtained using the model for the arrival process of the first stream. Note that stochastic capacity significantly alters the way overflows occur. The reason is that the large deviation behavior of the departure process from a single class queue is different with deterministic and stochastic service capacity [5,8], and this affects the overflow probabilities in our model.

Among the main contributions of this work we can mention: (a) the formulation of the problem of the aggregated large deviation principle (LDP) and its application to ODPS model with multiple classes; and (b) the treatment of stochastic service capacities.

The remaining part of this paper is organized as follows. In Section 2, we present a brief review of the large deviations principle and present the mathematical basis. In Section 3, we present a description of our performance model. Then, we define formally the ODPS policy, and state the main result of the paper. In Section 4, we present the problem of optimal control, and show that for every overflow scheme, we can associate a sample path where both arrival and service process can take it. In Section 5, an approximation of the overflow probability is given. In Section 6, we present a congestion management policy scheme. Section 7 presents the conclusion of this paper.

2. Mathematical basics

In this section, we review some basic results on the theory of large deviations theory [7,11], that will be used in the sequel. Consider a sequence $\{V_1,V_2,\dots\}$ of random variables with values in $/R$ and define:

$$\Lambda_n(\lambda) = \frac{1}{n}\log E[e^{\lambda V_n}].$$

For the application under consideration, we assume that V_n is a partial sum process. Namely

$$V_n = \sum_{i=1}^{n} X_i,$$

where X_i, $i \geq 1$, are identically distributed and possibly dependent random variables. We will consider the following assumption:

Assumption A. $\Lambda(\cdot)$ is said to satisfy Assumption A, iff:

1. The limit $\Lambda(\lambda) = \lim_{n\to\infty}\Lambda_n(\lambda)$ exists $\forall\ \lambda \in /R$.
2. The origin is in the interior of the convergence domain $D_\Lambda = \{\lambda|\Lambda(\lambda) < \infty\}$ of $\Lambda(\lambda)$.
3. $\Lambda(\lambda)$ is differentiable in the interior of D_Λ and the derivate tends to infinity as λ approaches the boundary of D_Λ.
4. $\Lambda(\lambda)$ is a lower semi continuous, i.e. $\lim\inf_{\lambda_n\to\lambda}\Lambda(\lambda_n) \geq \Lambda(\lambda),\ \forall\lambda$.

Let us now define the Legendre transform $\Lambda^*(\cdot)$ of $\Lambda(\cdot)$, as follows:

$$\Lambda^*(a) = \sup_\lambda(a\lambda - \Lambda(\lambda)),$$

It is important to note that $\Lambda(\cdot)$ and $\Lambda^*(\cdot)$ are also called convex duals. Under Assumption A, the Gartner–Ellis theorem, [6] establishes that $\{V_n\}$ satisfies a LDP with the rate function $\Lambda^*(\cdot)$. In particular, this theorem intuitively asserts that for large n and small $\epsilon > 0$, the probability that V_n

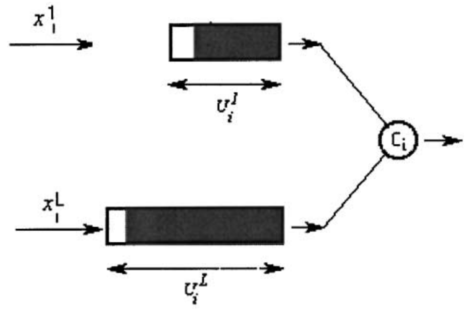

Fig. 1. A multiclass model.

belongs to the interval $[na - \epsilon, na + \epsilon]$ is given by:

$$P\{V_n \in [na - \epsilon, na + \epsilon]\} \approx e^{-n\Lambda^*(a)}.$$

The Gartner–Ellis theorem generalizes Cramer's theorem [12], which is applied to independent and identically distributed (*iid*) random variables. Similarly, we will consider the following assumption regarding the sample path LDP described in Ref. [13].

Assumption B. $\Lambda(\cdot)$ is said to satisfy Assumption B iff:

For all integer m and all real non-negative ϵ_1, ϵ_2 that are sufficiently small, and for every scalar $a_0, \ldots, a_{m-1}$, there exists $M > 0$ such that for all $n \geq M$ and all $k_0, \ldots, k_m$ satisfying $1 = k_0 \leq k_1 \leq \ldots \leq k_m = n$, we have:

$$\exp -\left(n\epsilon_2 + \sum_{i=0}^{m-1} (k_{i+1} - k_i)\Lambda^*(a_i)\right)$$

$$\leq P\Big[|V_{k_{i+1}} - V_{k_i} - (k_{i+1} - k_i)a_i| \leq \epsilon_1 n\Big]$$

$$\leq \exp\left(n\epsilon_2 - \sum_{i=0}^{m-1} (k_{i+1} - k_i)\Lambda^*(a_i)\right), \quad \forall 0 \leq i \leq m - 1.$$

Assumption B is a consequence of Mogulskii's theorem [11]. Intuitively speaking, Assumption B deals with the probability of sample paths that are constrained to be within a tube around a "polygonal" path made up with linear segments of slopes $a_0, \ldots, a_{m-1}$. A uniform bounding condition was given in Ref. [7] under which Assumption B can be satisfied.

It is widely accepted that the set of processes satisfying Assumptions A and B is large enough to include renewal, Markov-modulated, and stationary processes with mild mixing conditions. Such processes can model "burstiness" and are commonly used in modeling the input traffic to communication networks.

On a notational remark, in the rest of the paper will be denoted by

$$V_{i,j}^X = \sum_{k=i}^{j} X_k, \qquad i \leq j,$$

the partial sum of the random sequence $\{X_i,\ i \in /N\}$. We will be also denoting by $\Lambda_X(\cdot)$ and $\Lambda_X^*(\cdot)$ the limiting log-moment generating function and the large deviations rate function of the process X, respectively.

3. The multiple class model

The ODPS policy differs from the other policies in several aspects. Contrary to some other scheduling policies, such as the *generalized processor sharing* policy, the *longest queue first* policy, and the *generalized longest queue first* policy, which operate at higher levels by merging multiple media that share the same queue class, the ODPS scheduling policy, operates at lower levels by considering each media separately. This is essential in a system that has to guarantee a specific QoS to each media. According to ODPS policy, the scheduling procedure computes at each slot of time n the fraction $\theta_n^{j,i}$ of the available capacity to the jth media of class i, which depends on the delivered QoS until the slot of time n is over. Several policies can be implemented to compute $\theta_n^{j,i}$ easily that is why we are not concerned with this problem.

The scheduler can characterize, at each slot of time n, the set of media S_n of all tuples $(\theta_n^{k_1,\ldots,m_1}, \ldots \theta_n^{k_s,\ldots,m_s})$ defined by the following condition:

$$S_n = \left\{\theta_n^{k_s,m_s} | \theta_n^{k_1,m_{s1}} \leq \theta_n^{k_2,m_2} \leq \ldots \leq \theta_n^{k_r,m_r}, \text{ and } \sum_{s=1}^{r} \theta_n^{k_s,m_s} = 1\right\}.$$

S_n corresponds to the set of served media during the interval of time $[n, n + 1]$. Given S_n, the scheduler computes the fractions θ_n^i, $1 \leq i \leq L$, allowed to queue class i at slot time n, as shown below

$$\theta_n^i = \sum_{j \in S_i} \theta_n^{j,i}.$$

Now, we apply our model to the multiclass media operating under the ODPS policy that is to be analyzed. We then state the main result and provide a brief outline of the approach to be taken.

Let us assume that a slotted time model (i.e. discrete time) and consider the system depicted in Fig. 1. Let $\{Q_j,\ 1 \leq j \leq L\}$ be the set of queues in the system, and A_i^j denotes the number of class j cells, issued from the N_i^j media that enter queue Q_j at time i, for $1 \leq j \leq L$. Each queue Q_j has a finite dynamic buffer equal to U_i^j, allowed at time i such as

$$\sum_{j=1}^{L} U_i^j = U, \qquad \forall i \geq 0,$$

where U is the total amount buffer of the resource.

The set of queues $\{Q_j,\ 1 \leq j \leq L\}$ share the same server, which can process C_i cells during the time interval $[i, i + 1]$. We assume that the processes $\{A_i^j,\ 1 \leq j \leq L,\ i \in /N\}$ are stationary and mutually independent. However, we allow dependencies between streams A_i^j's issued from the same traffic media, for fixed j and different values of i.

Let O_i^j denote the queue length at time i (without counting arrivals at time i) in queue Q_j, for $1 \le j \le L$. We assume that the server allocates its capacity between queues Q_j dynamically according to a work-conserving policy (i.e. the server never stays idle when there is work in the system). We also assume that the queue length processes $\{O_i^j,\ 1 \le j \le L,\ i \in /N\}$ are stationary.

To simplify the analysis we consider a discrete-time "fluid" model, meaning that we will be treating x_i^j, O_i^j for $1 \le j \le L$ and C_i as non-negative real numbers (the amount of fluid entering in queue, or served). We assume the following stability condition:

$$\mathrm{E}[C_i] > \sum_{j=1}^{L} \mathrm{E}[x_i^j], \qquad \forall i \ge 0.$$

By assuming an ODPS policy, we have the following definition for sample paths.

Definition 1. If each queue Q_l has an allocated amount of buffer equals to U_n^l at time slot n, then we can define a sample path to be a tuple $\mathrm{SP}_n^l = (U_n^l, U, \vec{\theta}_n^l, \vec{x}_n^l)$, where $\vec{\theta}_n^l = (\theta_i^l)_{0\le i\le n}$, $\vec{\Theta}_n^{l-1} = (\Theta_i^l)_{0\le i\le n}$, $\vec{x}_n^l = (x_i^l)_{0\le i\le n}$ and $\vec{X}_n^{l-1} = (X_i^l)_{0\le i\le n}$. This leads to overflow of queue Q_l at time n, i.e.. $O_0^l = 0$, $O_n^l > U_n^l$.

4. The optimal control problem

We relate the control problem heuristically to the problem of obtaining an asymptotically tight estimate of the overflow probability of some queue assumed to be Q_L. For every overflow sample path leading to $O_i^L > U_i^L$ there exists some time $i > 0$ such that the queue Q_L is empty.

In order to reduce the computation complexity of the event $O_i^L > U_i^L$, we propose to use an aggregation technique consisting of $L-1$ levels of aggregation, see Ref. [1]. At the first level of aggregation we are concerned with the performance evaluation of $\{Q_1, Q_2\}$, which we represent by Ξ_1, named aggregate of the first level and for which we allocate at time n $\{(U_n^1, \theta_n^1), (U_n^2, \theta_n^2)\}$ where:

$$\begin{cases} U_n^1 + U_n^2 = B_n^1, \\ \theta_n^1 + \theta_n^2 = \Theta_n^1. \end{cases}$$

At the lth level of aggregation $1 < l \le L-1$, we are concerned with the performance evaluation of $\Xi_l = \{Q_l,\ \Xi_{l-1}\}$, named aggregate of the lth level of aggregation and for which we allocate at time n $\{(U_n^l, \theta_n^l), (B_n^{l-1}, \Theta_n^{l-1})$ where:

$$\begin{cases} U_n^l + B_n^{l-1} = B_n^l, & \text{where } B_n^{l-1} \le B_n^l \le U, \\ \theta_n^{\,l} + \Theta_n^{l-1} = \Theta_n^{\,l}, & \text{where } \Theta_n^{l-1} \le \Theta_n^{\,l} \le 1. \end{cases}$$

At the lth level of aggregation $1 < l \le L-1$, B_n^{l-1} has an intuitive interpretation, called *Virtual Buffer Size* (VBS). In fact, B_n^{l-1} represents the total amount of buffer space allocated to Ξ_{l-1} at time n, which means that, if its buffer contains B_n^{l-1} cells, the $(B_n^{l-1}+1)$th cell will be discarded. Setting B_n^{l-1} to $\mathrm{Min}_{j=1,\ldots,l-1}(U_n^j)$ will down size the VBS model. In the other hand, setting B_n^{l-1} to $\sum_{j=1}^{l-1} U_n^j$ will over size the VBS model since a cell rejection can occur within Ξ_{l-1} with less then

$$\sum_{j=1}^{l-1} U_n^j$$

cells.

For each value of B_n^{l-1}, we associate a probability distribution $P_n^{l-1}(B_n^{l-1})$ describing the probability that the aggregate queue Ξ_{l-1} behaves as a single queue with buffer space equal to B_n^{l-1}. In the following, we will define the expression of $P_n^{l-1}(B_n^{l-1})$.

The performance characterization of each level l of aggregation is based on the study and performance evaluation of the queue Q_l and the aggregated queue $\Xi_{l-1} = \{Q_{l-1},\ \Xi_{l-2}\}$ of the $(l-1)$th level of aggregation. Each level of aggregation l is conditioned by the probability $P_n^{l-1}(B_n^{l-1},\ \Theta_n^{l-1}) = P_n^{l-1}(B_n^{l-1})P_n^{l-1}(\Theta_n^{l-1})$.

This denotes the probability that the aggregated queue $\Xi_{l-1} = \{Q_{l-1},\ \Xi_{l-2}\}$ of the $(l-1)$th level of aggregation is attributed at time n with an amount of resources given by $(B_n^{l-1},\ \Theta_n^{l-1})$.

4.1. Characterization of the lth aggregate level

Let x_i^j denote the number of class j's cells issued from the N_i^j media that enter queue Q_j at time I (as mentioned in Section 3), and $\{X_i^{l-1},\ x_i^l,\ i \ge 0\}$ and x_i^{L+1} denote the empirical rates of the $(l-1)$th aggregated arrival process processes given by $X_i^{l-1} = \sum_{m=1}^{l-1} x_i^m$ of the lth queue and C, respectively. The probability of sustaining $\{X_i^{l-1},\ x_i^l,\ i \ge 0\}$ and x_i^{L+1}, in the interval $[0,n]$ for large values of $\{B_i^{l-1}, U_i^l\}$, where $B_i^{l-1} + U_i^l = B_i^l$, $B_i^{l-1} \le B_i^l \le U$, is given (up to first degree exponent) by:

$$\exp\left\{-\sum_{i=0}^{n} F_i(B_i^{l-1}, U_i^l, \Theta_i^{l-1}, \theta_i^l, X_i^{l-1}, x_i, x_i^{L+1})\right\} \tag{1}$$

where

$$\begin{aligned} &F_i(B_i^{l-1}, U_i^l; \Theta_i^{l-1}, \theta_i^l, X_i^{l-1}, x_i^l, x_i^{L+1}) \\ &\quad = U_i^l.\,[\Lambda_{A^l}^*(x_i^l) + \Lambda_C^*(\theta_i^l\, x_i^{L+1})] \\ &\qquad + P_i^{l-1}(B_i^{l-1}, \Theta_i^{l-1})B_i^{l-1}[\Lambda_{A^{l-1}}^*(X_i^{l-1}) + \Lambda_C^*(\Theta_i^{l-1} x_i^{L+1})]. \end{aligned}$$

and

$$P_i^{l-1}(B_i^{l-1}, \Theta_i^{l-1}) = \sum_{\substack{B_i^{l-2} + U_i^{l-1} = B_i^{l-1}, \\ \Theta_i^{l-2} + \theta_i^{l-1} = \Theta_i^{l-1}}} \pi_i^{l-1}(U_i^{l-1}, \theta_i^{l-1}) P_i^{l-2}(B_i^{l-2}, \Theta_i^{l-2}).$$

Also, let $\pi_i^{l-1}(U_i^{l-1}, \theta_i^{l-1}) = \pi_i^{l-1}(U_i^{l-1}).\pi_i^{l-1}(\theta_i^{l-1})$ denote the probability that Q_{l-1} is allocated $(U_i^{l-1}, \theta_i^{l-1})$ at time i.

This cost function is a consequence of Assumption B.

With the scaling introduced here as $B_i^{l-1}, U_i^l \to \infty$, the sequence of slopes $\{a_0, \ldots, a_{m-1}\}$ converges to the empirical rate $\{X_i^{l-1}, x_i^l, x_i^{L+1}\}$, and the sum of rate functions appearing in the exponent converges to the sum mentioned in expression (1).

At each level of aggregation l, we seek a path with maximum cost (here the cost function is given by the summation in the above expression). The optimization problem corresponding to the queue Q_l, is subject to the constraints $O_0^l = 0$, and $O_n^l = U_n^l$. The fluid levels in the queue Q_l, O_i^l, $1 \le i \le n$, are the state variables, the fractions θ_i^l, Θ_i^l and the allowed amount of resources at the class lth.The levels U_n^l, and B_n^{l-1} are the control variables. The dynamics of the system depend on the state and the employed scheduling policy. Later, we will discuss the optimization issue.

By applying the ODPS policy at the lth level of aggregation, while assuming that queue Q_l buffer is equal to U_n^l at time slot n, and conditioned by a total amount allowed to the lth aggregate Ξ_l equal to $\{(B_n^l, \Theta_n^l)\}$ at time slot n, we ca determine a finite number of sample paths of control trajectory defined by:

$$\mathrm{SP}_n^l = (U_n^l, B_n^{l-1}, \vec{\theta}_n^l, \vec{\Theta}_n^{l-1}, \vec{x}_n^l, \vec{X}_n^{l-1})$$

where $\vec{\theta}_n^l = (\theta_i^l)_{0 \le i \le n}$, $\vec{\Theta}_n^{l-1} = (\Theta_i^l)_{0 \le i \le n}$, $\vec{x}_n^l = (x_i^l)_{0 \le i \le n}$, and $\vec{X}_n^{l-1} = (X_i^l)_{0 \le i \le n}$, leading to an overflow of queue Q_l at time n, i.e. $O_0^l = 0$, $O_n^l = U_n^l$.

5. An outline of our approach

We assumed earlier that the arrival and service processes satisfy Assumptions A and B. As we have noted in Section 2, these assumptions are satisfied by processes that are commonly used to model bursty traffic in communication networks. Since the number of class j's cells A_i^j, $1 \le j \le L$, represents the number of arrivals, they can be assumed to be non-negative, which implies that their rate function $\Lambda_X^*(x)$, for $X \in \{x_i^j, 1 \le j \le L\}$ is infinity for all $x < 0$.

According to ODPS policy, the server allocates a fraction θ_i^j of its capacity to queue Q_j, at time i. The policy can be defined to be a work-conserving policy. More formally, this means that it satisfies:

$$O_{i+1}^j = [O_i^j + x_i{}^j - \theta_i{}^j.C_i]^+,$$

where $[x]^+ = \max(0, x)$.

At the lth level of aggregation, we are interested in estimating the overflow probability $P[O_n^l \ge U_n^l | \mathrm{SP}_n^l]$, for some sample path defined by SP_n^l $(U_n^l, B_n^{l-1}, \vec{\theta}_n^l, \vec{\Theta}_n^{l-1}, \vec{x}_n^l, \vec{X}_n^{l-1})$ that occurs at slot time n, for large values of U_n^l, andB_n^{l-1}. Finally, we will prove how the overflow probability can be minimized.

Let $n > 0$, and the variables $\{(U_n^l, \theta_n^l)$, and $(B_n^{l-1}, \Theta_n^{l-1})\}$ be the resources allocated to the queues $\{Q_l, \Xi_{l-1}\}$ of the lth level of aggregation at time n.

Let the constants $\{(\bar{x}_l, \bar{\theta}_l), (\bar{X}_{l-1}, \bar{\Theta}_{l-1}), \bar{x}_{L+1}\}$, and $\{\epsilon_l, \xi_{l-1}, \epsilon_{L+1}, \xi_{L+1}\}$ be strictly positive and consider the events:

$$e_l = \{|V_{1,n}^{A^l} - n\bar{x}_l| \le \epsilon_l n\},$$

$$E_l = \{|V_{1,n}^{\mathrm{A}^{l-1}} - n\bar{X}_{l-1}| \le \xi_{l-1} n\},$$

$$F_l = \{|V_{1,n}^{C^l} - n\bar{\theta}_l \bar{x}_{L+1}| \le \epsilon_{L+1} n\},$$

$$\Sigma_l = \{|V_{1,n}^{\bar{C}^l} - n\Theta_{l-1}\bar{x}_{L+1}| \le \xi_{L+1} n\},$$

Notice that:

(i) $\{\bar{x}_l, \bar{X}_{l-1}\}$ is the mean constant arrival rates of Q_l and Ξ_{l-1} during the interval $[0, n]$,
(ii) $\{\bar{\theta}_l, \bar{\Theta}_{l-1}\}$: is the mean constant proportions of services offered to Q_l and Ξ_{l-1} during the interval $[0, n]$,
(iii) $\bar{x}_{L+1}$: is the mean constant service rate during the interval $[0, n]$,

Under the ODPS scheduling policy, the queue Q_l builds up in the interval $[0, n]$. O_i^l is given by the following expression:

$$O_i^l = \sum_{k=1}^{i} x_k{}^l - \sum_{k=1}^{i} \theta_k^l x_k^{L+1}$$

Using e_1, E_l, F_l, Σ_1 and the above relation, we have $O_n^l \ge U_n^l - \epsilon_l'$, where $\epsilon_l' \to 0$ as $\epsilon_l \to 0$, $\xi_{l-1} \to 0$, $\epsilon_{L+1} \to 0$ and $\xi_{L+1} \to 0$.By calculating the probability $P[O_n^l \ge U_n^l | \mathrm{SP}_n^l]$ of the overflow of queue Q_l, over a sample path SP_n^l at the lth level of aggregation, we obtain:

$$P[O_n^l \ge U_n^l | \mathrm{SP}_n^l] \approx$$

$$P(e_l)P(E_l)P(F_l)P(\Sigma_l)$$

The expressions of $P(e_l)$, $P(E_l)$, $P(F_l)$ and $P(\Sigma_l)$ can be rewritten as:

$$P[O_n^l \ge U_n^l | \mathrm{SP}_n^l] \approx P\{|V_{1,n}^{A^l} - n\bar{x}_l| \le \epsilon_l n\} P\{|V_{1,n}^{\mathrm{A}^{l-1}} - n\bar{X}_{l-1}| \le \xi_{l-1} n\} P\{|V_{1,n}^{C^l} n\bar{\theta}_l \bar{x}_{L+1}| \le \epsilon_{L+1} n\} P\{|V_{1,n}^{\bar{C}^l} - n\Theta_{l-1}\bar{x}_{L+1}| \le \xi_{L+1} n\},$$

where n is large enough, and $\epsilon_l' \to 0$ as $\epsilon_l \to 0$, $\xi_{l-1} \to 0$, $\epsilon_{L+1} \to 0$ and $\xi_{L+1} \to 0$.

Then, using Assumption B, and for all $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$, there exists N such that for all $n > N$, we obtain the following framing of the probability of an overflow of queue Q_l, over a sample path SP_n^l, which is achieved at time n if the buffer space allowed to queue Q_l, is equal to U_n^j :

$$\exp(-\epsilon_1 n - \Lambda^*(\mathrm{SP}_n^l)) \le P[O_n^l \ge U_n^l | \mathrm{SP}_n^l] \le \exp(\epsilon_2 n - \Lambda^*(\mathrm{SP}_n^l))$$

where:

$$\Lambda^*(\mathrm{SP}_n^l) = \sum_{i=0}^{n} F_i(B_i^{l-1}, U_i^l, \Theta_i^{l-1}, \theta_i^l, X_i^{l-1}, x_i^l, x_i^{L+1}).$$

As a final step to this proof, if we assume that U_n^l, is sufficiently large, so we have the probability of an overflow of queue Q_l, over a sample path SP_n^l which is achieved at time n if the buffer space allowed to queue Q_l, is equal to SP_n^l :

$$P[O_n^l \geq U_n^l | \mathrm{SP}_n^l] \approx \exp(-\Lambda^*(\mathrm{SP}_n^l))$$

Hence, we present the following theorem.

Theorem 1. *Under the ODPS policy, and for each level of aggregation l $1 \leq l \leq L-1$, if we assume that the arrival and service processes satisfy Assumptions A and B, and for sample paths of control trajectory defined by:*

$$\mathrm{SP}_n^l = (U_n^l, B_n^{l-1}, \vec{\theta}_n^l, \vec{\Theta}_n^{l-1}, \vec{x}_n^l, \vec{X}_n^{l-1}),$$

the queue length O_n^l of queue Q_l at time n, is given by:

$$P[O_n^l \geq U_n^l | \mathrm{SP}_n^l] \approx \exp(-\Lambda^*(\mathrm{SP}_n^l)).$$

This is basically an optimal control problem trajectory that aims at minimizing the overflow probability.

5.1. Computation of $P_i^{l-1}(B_i^{l-1}, \Theta_i^{l-1})$

We defined $P_n^{l-1}(B_n^{l-1}, \Theta_n^{l-1}) = P_n^{l-1}(B_n^{l-1})P_n^{l-1}(\Theta_n^{l-1})$ as the probability that the aggregated queue $\Xi_{l-1} = \{Q_{l-1}, \Xi_{l-2}\}$ of the $(l-1)$th level aggregate, which attributes at time n an amount of resources given by $(B_n^{l-1}, \Theta_n^{l-1})$. Here, $P_n^{l-1}(B_n^{l-1})$ represents the probability that Ξ_{l-1} behaves as a single queue with VBS equal to B_n^{l-1}, and $P_n^{l-1}(\Theta_n^{l-1})$ is the probability that Ξ_{l-1} receives a fraction of Θ_n^{l-1} at time n.

Assuming that the allocation process $\{\theta_n^k,\ n \geq 0\}_{1 \leq k \leq L}$ is obeying a general probability distribution, then we can write:

$$P_n^{l-1}(\Theta_n^{l-1}) = \sum_{\sum \theta_n^k = \Theta_n^{l-1}} \prod_{k=1}^{l-1} \pi_n^k(\theta_n^k),$$

where $\pi_n^k(\theta_n^k)$ denotes the probability that class queue Q_k receives a proportion of resource equal to θ_n^k at time n. Similarly, since $\forall 2 \leq l \leq L$:

$$P_i^{l-1}(B_i^{l-1}, \Theta_i^{l-1}) = \sum_{\substack{B_i^{l-2} + U_i^{l-1} = B_i^{l-1},\\ \Theta_i^{l-2} + \theta_i^{l-1} = \Theta_i^{l-1}}} \pi_i^{l-1}(U_i^{l-1}, \theta_i^{l-1}) P_i^{l-2}(B_i^{l-2}, \Theta_i^{l-2})$$

and $\forall 1 \leq j \leq L-1$, $P_i^j(B_i^j, \Theta_i^j) = P_i^j(B_i^j)P_i^j(\Theta_i^j)$.

We have:

$$P_i^{l-1}(B_i^{l-1}) = \sum_{B_i^{l-2} + U_i^{l-1} = B_i^{l-1}} \pi_i^{l-1}(U_i^{l-1}).P_i^{l-2}(B_i^{l-2}),$$

where:

$$\pi_i^{l-1}(U_i^{l-1}) = P[O_i^{l-1} \geq U_i^{l-1} | \mathrm{SP}_i^{l-1}] - P[O_i^{l-1} \geq U_i^{l-1} + 1 | \mathrm{SP}_i^{l-1}]$$

and $P_i^{l-2}(B_i^{l-2})$ is obtained recursively by using the formula given above.

6. QoS management policy

Multimedia applications in high-speed networks share several important features, which represent a major departure from conventional computer data applications. First, they usually require real-time transmission of continuous media information such as digital video and audio. To support real-time traffic, the network must also support and guarantee the continuity of QoS while maintaining the integrity and the QoS of each traffic class in the system.

As the use of multimedia applications increases, so are the demand for resources required to support them. Network resources such as bandwidth, buffer space and processing time at each switch of the high speed network, should be allocated in a cost-effective manner. Each application expects the network to support a desired QoS. The service provider is interested to provide the desired QoS as efficiently as possible.

For the sake of simplicity, let us consider the case of ATM networks. In order to prevent gross misuse of ATM network resources, it is reasonable to include mechanisms for congestion management at all access points of a public ATM network. Let R be an ATM route defined by a set of ATM switches $\{r, r \in R\}$. In the remaining part of the paper we will use the symbol $X(r)$ to be the value of the parameter X at switch r.

Assume that at time slot i, the paths of control trajectories at some ATM node r, corresponding to queue $\{Q_j(r), 1 \leq j \leq L\}$, satisfy the following conditions:

$$\forall\ 1 \leq j \leq L : O_i^j(r) = \sum_{k=1}^{i} x_k^j(r) - \sum_{k=1}^{i} \theta_k^j(r) x_k^{L+1}(r) \leq U_i^j(r)$$

and

$$\sum_{k=1}^{L} U_i^k(r) = U(r),$$

where $U_i^k(r)$, $1 \leq k \leq L$, $r \in R$, denotes the buffer space allocated to queue class $Q_k(r)$ at time i. By using matrix form, we can rewrite the latter expression as:

$$\overrightarrow{O}_i(r) = \overrightarrow{X}_i(r) - \Theta_i(r).\overrightarrow{C}_i(r) \leq \overrightarrow{U}_i(r),$$

where

$$\vec{O}_i(r) = (O_i^j(r))_{1\le j\le L},\ \ \vec{X}_i(r) = \left(\sum_{k=1}^{i} x_k^j(r)\right)_{1\le j\le L},$$

and $\vec{U}_i(r) = (U_k^i(r))_{1\le k\le L}$, are L-column vectors. Also, $\vec{C}_i(r) = (x_k^{L+1}(r))_{1\le k\le i}$ is an i-column vector, and $\Theta_i(r) = (\theta_j^k(r))_{1\le k\le i\,1\le j\le L}$ is an Lxi-matrix.

Our objective is to ensure a strong predictive resource allocation scheme, such that the overflow at each class queue remains less than the threshold in the interval of time $[i+1, n]$. This can be done by:

- Predicting both arrival and service rates, i.e. $\{\bar{\bar{x}}_k^l(r), \bar{\bar{x}}_k^{L+1}(r), i+1 \le k \le n,\ 1 \le l \le L\}$, which can be done using Kalman filters.
- Choosing optimal values of resources $\{U_k^l(r), \theta_k^{L+1}(r), i+1 \le k \le n, 1 \le l \le L, r \in R\}$, which will be allocated to each queue class such that the following conditions are respected:

  $\forall\ i+1 \le k \le n, \forall\ 1 \le j \le L$

$$:\begin{cases} \mathrm{CLR}^j_{\mathrm{Overflow},k}(R) \le \mathrm{CLR}^j_{\mathrm{Threshold}}(R), \\ \mathrm{CDV}^j_{\mathrm{Overflow},k}(R) \le \mathrm{CDV}^j_{\mathrm{Threshold}}(R) \\ \mathrm{MDV}^j_{\mathrm{Overflow},k}(R) \le \mathrm{MDV}^j_{\mathrm{Threshold}}(R) \end{cases} \tag{2}$$

  where:
  - $\mathrm{CLR}^j_{\mathrm{Overflow},k}(R)$ is the overflow probability (Cell Loss Rate) of queue class j at time k over the route R, and $\mathrm{CLR}^j_{\mathrm{Threshold}}(R)$ is the threshold probability of the jth queue class over the route R.
  - $\mathrm{CDV}^j_{\mathrm{Overflow},k}(R)$ is the Cell Delay Variation of queue class j at time k over the route R, and $\mathrm{CDV}^j_{\mathrm{Threshold}}(R)$ is the threshold cell delay variation of the jth queue class over the route R.
  - $\mathrm{MDV}^j_{\mathrm{Overflow},k}(R)$ is the Mean Delay Variation of queue class j at time k over the route R, and $\mathrm{MDV}^j_{\mathrm{Threshold}}(R)$ is the threshold mean delay variation of the jth queue class over the route R.

Note that since the overflow probability is assumed to be small, an approximation of $\mathrm{CLR}^j_{\mathrm{Overflow},k}(R)$ is given as follows:

$$\mathrm{CLR}^j_{\mathrm{Overflow},k}(R) \approx \sum_{r\in R} CLR^j_{\mathrm{Overflow},k}(r)$$

where $\mathrm{CLR}^j_{\mathrm{Overflow},k}(r)$ is the overflow probability of queue class j at time k over the ATM switch r. Using Theorem 1, we have:

$$\mathrm{CLR}^j_{\mathrm{Overflow},k}(\mathrm{SP}^j_n(r)) \approx \exp[-\Lambda^*(\mathrm{SP}^j_n(r))].$$

Now, let us consider the evaluation of the expressions of the mean cell delay, and cell delay variation in our ATM network. We can prove easily how all of them are related to the probability distribution of the queue length $\mathrm{CLR}^j_{\mathrm{Overflow},k}(r) \approx \exp[-\Lambda^*(\mathrm{SP}^j_n(r))]$ for some sample path $\mathrm{SP}^j_n(r)$, at time slot n. The instantaneous expressions of these parameters that correspond to the jth queue class at time k of the rth ATM switch and for some sample path $\mathrm{SP}^j_n(r)$ are given by:

$$\mathrm{MCD}^j_k(\mathrm{SP}^j_n(r)) \approx \frac{\sum_{q=0}^{U^j_k(r)-1} P[O^j_n(r) \ge q(r) | \mathrm{SP}^j_n(r)]}{C^j_k(r)},$$

$$\mathrm{CDV}^j_n(\mathrm{SP}^j_k(r)) \approx \frac{\sum_{q=0}^{U^j_n(r)-1} (2q+1) P[O^j_k(r) \ge q(r) | \mathrm{SP}^j_k(r)]}{(C^j_k(r))^2} - \frac{\left\{\sum_{q=0}^{U^j_k(r)-1} P[O^j_k(r) \ge q(r) | \mathrm{SP}^j_k(r)]\right\}^2}{(C^j_k(r))^2}$$

Hence, the approximation of $\mathrm{MCD}^j_{\mathrm{Overflow},k}(R)$ can be given as:

$$\mathrm{MCD}^j_{\mathrm{Overflow},k}(R) \approx \sum_{r\in R} \mathrm{MCD}^j_{\mathrm{Overflow},k}(\mathrm{SP}^j_k(r)).$$

Similarly, an approximation of $\mathrm{CDV}^j_{\mathrm{Overflow},k}(R)$ can be given by:

$$\mathrm{CDV}^j_{\mathrm{Overflow},k}(R) \approx \sum_{r\in R} \mathrm{CDV}^j_{\mathrm{Overflow},k}(\mathrm{SP}^j_k(r)).$$

To do this, the resource manager and the scheduling policy must force the dynamics of the sample paths, i.e. $\mathrm{SP}^j_k(r)$, $i+1 \le k \le n$, $1 \le j \le L$, $r \in R$, such that the conditions given in Eq. (2) are respected. The predictive algorithm that we propose operates as following:

Predictive algorithm
For: $\forall\ i+1 \le k \le n,\ 1 \le j \le L,$
Find $\mathrm{SP}^j_k(r)$, $\forall r \in R$, such that: $\forall i+1 \le k \le n$, $\forall 1 \le j \le L:$

$$\sum_{r\in R} \mathrm{CLR}^j_{\mathrm{Overflow},k}(\mathrm{SP}^j_n(r)) \le \mathrm{CLR}^j_{\mathrm{Threshold}}(R)$$

$$\sum_{r\in R} \mathrm{MCD}^j_{\mathrm{Overflow},k}(\mathrm{SP}^j_k(r)) \le \mathrm{MCD}^j_{\mathrm{Threshold}}(R)$$

$$\sum_{r\in R} \mathrm{CDV}^j_{\mathrm{Overflow},k}(\mathrm{SP}^j_k(r)) \le \mathrm{CDV}^j_{\mathrm{Threshold}}(R)$$

and $\vec{O}_k(r) = \vec{X}_k(r) - \Theta_k(r)\vec{C}_k(r) \le \vec{U}_k(r), \forall r \in R.$

7. Conclusions

To conclude, we considered a multiclass model that allows the performance evaluation and analysis of several telecommunications entities including switches and routers. The entities are assumed to offer dedicated buffers for each service class.

Under the ODPS policy, we have obtained the asymptotic tail of the overflow probability for each buffer using the concept of VBS model and a technical aggregation. We have addressed the case of multiplexing L streams by using the VBS model and $(L-1)$ hierarchical aggregation scheme.

Using the standard large deviations methodology, we provided a lower upper bound on the buffer overflow probabilities (matching up to first degree of the exponent). We formulated the problem of congestion management by using the Kalman filters. Finally, this formulation provides a particular insight into the problem, as it offers an explicit characterization of the most likely traffic characteristics.

References

[1] C. Ben Ahmed, N. Boudriga, M.S. Obaidat, A methodology to optimize the design of telecommunication systems, Computer Communications 23 (2000) 912–921.

[2] I.Ch. Paschalidis, Class-specific quality of service guarantees in multimedia communication networks, Technical Report, Department of Manufacturing Engineering, Boston University, June 1998.

[3] A.K. Parekh, R.G. Gallager, A generalized processor sharing approach to overflow control in integrated services networks: the single node case, IEEE/ACM Transactions on Networking 1 (3) (1993) 344–357.

[4] D. Bertsimas, I.Ch. Paschalidis, J.N. Tsitsiklis, Asymptotic buffer overflow probabilities in multiclass multiplexers: an optimal control approach, IEEE Transactions on Automatic Control 43 (3) (1998) 315–335.

[5] A. Weiss, An introduction to large deviations for communication networks, IEEE Journal on Selected Areas in Communications 13 (6) (1995) 938–952.

[6] J.Y. Hui, Resource allocation for broadband networks, IEEE Journal on Selected Areas in Communications 6 (9) (1988) 1598–1608.

[7] R.J. Gibbens, P.J. Hunt, Effective bandwidths for the multi-type UAS channel, Queueing Systems 9 (1991) 7–28.

[8] A. Ganesh, V. Anantharam, Stationary tail probabilities in exponential server tandems with renewal arrivals, Queueing Systems 22 (1996) 203–248.

[9] G. de Veciana, C. Courcoubetis, J. Walrand, Decoupling bandwidths for networks: a decomposition approach to resource management, Memorandum, Electronics Research Laboratory, University of California Berkeley, 1993.

[10] P. Dupuis, K. Ramanan, A Skorokhod problem formulation and large deviation analysis of a processor sharing model, Technical Report, Division of Applied Mathematics, Brown University, 1997.

[11] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Jones and Bartlett, 1993.

[12] H. Cramer, Sur un nouveau théorème limite de la théorie des probabilités, Actualitées Scientifiques et Industrielles, Colloque Consacrée à la Théorie des Probabilités, vol. 736, Hermann, Paris, 1938, pp. 5–23.

[13] A. Dembo, T. Zajic, Large deviations: from empirical mean and measure to partial sums processes, Stochastic Processes and Applications 57 (1995) 191–224.