

Differentiated Services with Statistical Real-Time Guarantees in Static-Priority Scheduling Networks

Shengquan Wang[†], Dong Xuan[‡], Riccardo Bettati[†], and Wei Zhao[†]

[†] Department of Computer Science
Texas A&M University
College Station, TX 77843
{swang, bettati, zhao}@cs.tamu.edu

[‡] Department of Computer Science
The Ohio-state University
Columbus, OH 43210
xuan@cis.ohio-state.edu

Abstract

We propose and analyze a methodology for providing absolute differentiated services with statistical performance guarantees for real-time applications in networks that use class-based (as opposed to flow-aware) static-priority schedulers. We develop a method that can be used to derive statistical delay guarantees in a flow-unaware fashion. Traditionally, both deterministic and statistical delay analysis methods either depend on schedulers that keep per-flow state information, or require detailed information about flow population at delay analysis time. The fact that no such information is needed for delay analysis allows us to perform deadline tests during system (re-)configuration time. We are so able to reduce the runtime admission control to a simple utilization test. No explicit delay computation is necessary at admission time, making this approach scalable to large systems.

1 Introduction

Architectures for providing delay-guaranteed communication over internetworks either follow a connection-oriented model or at least rely on detailed per-flow information. In the IETF Integrated Services (`intserv`) architecture [2], for example, each connection is controlled both by admission control at connection admission time and by packet scheduling during the lifetime of the connection. At establishment time, the necessary resources must be allocated to the new connection, and during the lifetime, the connection is policed to ensure that the abnormal behavior

of a connection does not affect other connections. This necessitates that information about each connection is kept by each node along the path for admission control and packet forwarding.

It is agreed upon that Integrated Services do not scale. High-speed routers are required to maintain state and perform scheduling decisions for large numbers of connections. In addition, as the number of connections increases, the run-time overhead incurred in connection establishment and tear-down increases as well. The Integrated Services architecture therefore cannot provide scalable QoS guaranteed services, and the lack of scalability is due to overhead both at connection establishment and during the connection lifetime.

The Differentiated Services (`diffserv`) architecture is aimed at supporting service scalability through *aggregation of flows* into service classes [7]. Network nodes in `diffserv` display per-hop, per-class behavior, and need not maintain per-flow information. Since each router only guarantees that service agreements are *locally* maintained on a per-class basis, `intserv`-style end-to-end guarantees are difficult to provide. The end-to-end service might vary with networking conditions. A number of approaches have been proposed to provide `intserv`-like end-to-end performance without per-flow information in the network core. For example, assuming that no dynamic routing occurs, the *Premium Service*, which has been defined within the `diffserv` architecture, can offer the user a performance level that is similar to that of a leased-line, as long as the user's traffic is limited to a given bandwidth [18]. We call this type of service *absolute differentiated services*.

In order to provide service guarantees, an admission

control mechanism has to be in place, which makes sure that enough resources are available to satisfy the requirements of both the new and the existing connections after the new connection has been admitted. In order to keep steps with the scalability requirements for differentiated services networks, any admission control mechanism must be extremely light-weight so that can be realized in a scalable fashion. While the benefits of flow aggregation in terms of reduced overhead for packet scheduling are clear, the question of how to take advantage of the flow aggregation in differentiated-services networks for scaling of connection admission remains to be answered. In particular, methods must be found that work in the absence of per-flow information and when policing in the network is done on a per-class basis only. A number of approaches are being studied (e.g., [4, 24]) for admission control that relies on class-based information. In [24], we use appropriate system (re-)configuration steps to derive safe resource utilization bounds in each network node. Admission control is then reduced to a simple utilization-based test along the path of the flow: as long as the utilization of links along the path of a flow is not beyond a given bound, the probabilistic performance guarantee can be met. Thus, this approach (called Utilization-Based Admission Control – UBAC – in the following) eliminates explicit delay analysis at admission time, and helps the system scale up.

Utilization-based admission control (UBAC) is not new to networks. The fluid-flow model in the `intserv` framework, for example, allows various forms of utilization based admission control [20]. Such approaches cannot be used in a `diffserv` framework, however, because the fluid-flow model relies on guaranteed-rate schedulers, which need to maintain per-flow information and provide strong isolation between flows. The challenge of using the UBAC method is how to verify the correctness of a safe utilization bound at configuration time, when (i) schedulers in the nodes provide no policing and generally no isolation between flows belonging to the same class (we will be using class-based static-priority schedulers in this paper) and when (ii) no information is available yet about the flow population. For the case of *deterministic* performance guarantees for absolute differentiated services, we have derived methods to determine and/or assign safe utilization bounds (e.g., [24]).

While deterministic services provide a very simple model to the application, they tend to heavily overcommit resources because they account for the worst-case scenario. In real systems, this mostly results in significant portions of network resources (bandwidth etc.) being wasted [25]. *Statistical services*, on the other hand, significantly increase the efficiency of network usage by allowing increased statistical multiplexing of the underlying network resources. This comes at the expense of packets occasionally being dropped or excessively delayed. In this paper we will show how

such probabilistic guarantees can be provided in a differentiated services framework.

We will describe below how significant progress has been made to provide statistical performance guarantees within the `intserv` framework. All these results depend either on flow-aware schedulers or on information about flow population during admission control, and would therefore not be suited for use in a `diffserv` setting.

Our approach is based on *rate-variance envelopes* [11, 12], a simple and general traffic characterization. Such envelopes describe the variances of the flow rates as a function of the interval length. The results on delay violation probability derived in [12] depend on detailed information about the flow population. In this paper, we will develop flow-unaware versions of these results, and so develop a method that allows us to analyze delays and determine safe utilization bounds without depending on the dynamic status of the flow population. This will provide the basis for a scalable admission control for statistical performance guarantees in absolute differentiated services.

We will consider networks that use class-based static priority schedulers. Given that static priority scheduling is supported in many current routers (e.g., Packet Engines' RowerRail Routing Switch), our approach can be easily implemented in existing networks. As expected, our experimental data show that statistical services can achieve much higher utilization than deterministic services.

The rest of the paper is organized as follows. In Section 2, we describe previous work on absolute differentiated services and statistical service in general. The underlying network and traffic models for this study are introduced in Section 3. In Section 4, we develop a delay analysis methodology that is insensitive to flow information. In Section 5, we provide extensive experimental data to illustrate that the utilization achieved within a statistical model is much higher than within a deterministic model. A summary of the paper is given in Section 6.

2 Previous Work

A good survey with recent work on absolute differentiated services has been done in [21]. Nicols et. al. [18] propose the *premium service model*, which provides the equivalent of a dedicated link between two access routers. This model provides absolute differentiated services in priority-driven scheduling networks with two priorities, in which the high priority is reserved for premium service. Cruz [5] describes SCED+, which provides both guaranteed and statistical rate and delay bounds, and addresses scalability through traffic aggregation and statistical multiplexing. Stolica and Zhang [23] describe an architecture to provide guaranteed service without per-flow state management by using a technique called dynamic packet state (DPS). Our work is

based on static-priority scheduling, which is relatively simple and widely supported.

Statistical service has been studied via envelopes of bounding moment generating functions [3], exponentially bounded envelopes [22, 27], and envelopes consisting of families of bounding distributions [13, 28]. Statistical envelopes have also been applied to resource allocation for inter-class resource sharing [19] and video-on-demand services [9]. Much work has been done to generalize schedulability conditions for a deterministic service to a probabilistic framework. Several researchers made probabilistic extensions to deterministic service models. In [12], a *rate-variance envelope* is introduced, which describes the variance of the arrivals of a flow as a function of a time period of length. In [10], arrivals on a flow are assumed to be characterized by the rate-variance envelope and a long-term arrival rate. Then, applying the central limit theorem argument, a bound for the probability of a delay bound violation is derived for a static priority scheduler. In [12], the same framework is used to address bounds on the rate-variance envelope for regulated, adversarial traffic sources. In [11], the authors use a rate-variance envelope as a simple way to capture the second-moment properties of temporally correlated traffic flows, and to describe how quickly the rate-distribution becomes concentrated at the mean rate with increasing interval-length, a key factor for computing delay-bound-violation probabilities. They show empirical rate-variance envelopes for several long traces of compressed video to show how it captures the burstiness properties of realistic network traffic sources. In this paper, we will use this rate-variance envelope to do statistical delay analysis.

Admission control has been investigated widely [6, 8, 16]. The various approaches differ from each other in that they may require different scheduling schemes and so can be of vastly different complexity. For example, traditional admission control in networks with static priority scheduling is very complicated. Due to absence of flow separation, for any new flow request, admission control needs to explicitly compute and verify delays for the new and existing flows. This procedure is very expensive with increasing numbers of flows. The Utilization-Based Admission Control (UBAC) adopted dramatically reduces this complexity.

In its basic form, UBAC was first proposed in [15] for preemptive scheduling of periodic tasks on a simple processor. A number of utilization-based tests are known for centralized systems (e.g., 69% and 100% for periodic tasks on a single server using rate-monotonic and earliest-deadline-first scheduling, respectively [15]), or distributed systems (such as 33% for synchronous traffic over FDDI networks [1]). In this paper, we adopt utilization-based tests in providing differentiated services in static priority scheduling networks.

3 Network and Traffic Models

Our assumptions about network nodes and network traffic follow the Differentiated Services architecture: In the network, we differentiate flow-aware *edge routers* from *core routers*, which are only aware of aggregations of flows in form of *flow classes*. The network traffic consists of *flows*, which each belong to one of several flow classes. Examples of such flow classes are best-effort, voice-over-IP, video, or others. The flow class specifies both the traffic descriptor of a flow and its QoS requirements. We proceed below to describe our models in more detail.

3.1 Network Model

The IETF *diffserv* architecture distinguishes two types of routers: *Edge routers* are located at the boundary of the network, and provide support for traffic policing. *Core routers* are within the network, and maintain no flow information. Scheduling decisions within the core are therefore made based on class information only. Routers are connected through *links*, which we assume – for sake of simplicity of argument – to all be of capacity C , in bits per second.

When we compute delays, we follow standard practice and model a router as a set of *servers*, one for each router component where packets can experience delays. Packets are typically queued at some of these servers, typically at the output buffers, where they compete for the output link. Hence, we model a router as a set of output *link servers*. The other servers (input buffers, non-blocking switch fabric, wires, etc.) are typically not congested, and therefore can be eliminated by appropriately subtracting their constant delays from the deadline requirements of the traffic.

We assume that the schedulers in the nodes are not flow aware. Policing and scheduling therefore happens at class level, with class-based static priority schedulers assumed in this paper. Packets belonging to classes with higher priorities will be served prior to those with lower priorities.

Using class-based static priority schedulers alone makes it difficult to determine statistical end-to-end delays. The delay analysis methods described below rely on statistical independence among flows, and without additional mechanisms, the flows are no longer independent after being multiplexed at a server. Statistical independence can be re-established with the use of class-based jitter controller [14] at each core router. Class-based jitter control works by delaying each packet in a core router by its *ahead time* before passing the packet to the schedulers. The ahead time is the amount of time by which the packet left the previous server ahead of the deadline for its class. In this way, the statistical properties of the traffic arriving at the network edge is re-generated at each core router. It follows that the end-

to-end delays can be easily determined from local delays at each server along the path of a flow. For example, we will describe in Section 4.2 how the end-to-end delay probability is bounded by a product formula of the local delay probability of each server along the path of a flow. In the following discussion, we will therefore focus on the single-server analysis, and return to the end-to-end delay case in Section 4.2.

3.2 Traffic Model

We call a stream of packets between a sender and receiver a *flow*. Packets of a flow are transmitted along a single path, which we model as a list of link servers. Following the `diffserv` architecture, flows are partitioned into a number, say M , *classes*. QoS requirements and traffic specifications of flows are defined on a class-by-class basis¹. We assume that at each link server, a certain percentage of bandwidth is reserved for each particular traffic class. Let α_i denote the percentage of bandwidth reserved for Class i . It is the responsibility of the admission control module to ensure that the bandwidth usage of individual classes does not exceed the reserved portion. This is necessary to provide isolation among classes and hence to guarantee end-to-end delays to the flows in each class.

We model the traffic arrival for a flow as a stochastic arrival process $\mathcal{A} = \{\mathcal{A}(\tau), \tau \geq 0\}$, where random variable $\mathcal{A}(\tau)$ denotes the incoming traffic amount of the flow for a link server during time interval $[0, \tau]$. The arrival process \mathcal{A} is stationary and ergodic. The traffic arrivals for any two different flows are stochastically independent at the edge of the network and jitter controllers in the core routers preserve independence throughout the network core. The traffic arrival can be bounded either deterministically or stochastically by the arrival traffic envelopes as follows:

Definition 1 (Deterministic Arrival Traffic Envelope)

The function $b(t)$ is called the *deterministic arrival traffic envelope* of the traffic arrival \mathcal{A} if

$$\mathcal{A}(\tau + t) - \mathcal{A}(\tau) \leq b(t), \quad (1)$$

for any $\tau > 0$ and $t > 0$.

Definition 2 (Statistical Arrival Traffic Envelope) The distribution $B(t)$ forms the *stochastic arrival traffic envelope* of the traffic arrival \mathcal{A} if

$$\mathcal{A}(\tau + t) - \mathcal{A}(\tau) \leq_{st} B(t), \quad (2)$$

for any $\tau > 0$ and $t > 0$ ².

¹While at network level all flows within a class have identical traffic specifications and QoS requirements, user-level connections with varying bandwidth and burstiness requirements can be established by using more than one flow.

² $X \leq_{st} Y$ means $P\{X > Z\} \leq P\{Y > Z\}$ [13].

Since $\mathcal{A}(\tau)$ is stationary, $\mathcal{A}(\tau + t) - \mathcal{A}(\tau)$ possesses the same probability distributions for all τ . Therefore, we can define a random variable $R(t)$ as the stochastic arrival traffic rate as follows:

$$R(t) = \frac{\mathcal{A}(\tau + t) - \mathcal{A}(\tau)}{t}. \quad (3)$$

In this paper, we will be using the rate-variance envelope, which describes the variance of the arrival traffic rate during a time interval. The rate-variance envelope $RV(t)$ describes the variance of the arrival rate for the incoming flow over an interval of length t [11], i.e.

$$RV(t) = var(R(t)). \quad (4)$$

We assume that a class- i traffic flow is controlled by a leaky bucket with burst size σ_i and average rate ρ_i .

In the following, we will use the notation $G_{i,j}$ to denote a group of flows of Class i from Input Link j of a server and use $b_{i,j}$, $B_{i,j}$, and $RV_{i,j}$ to specify the deterministic arrival traffic envelope, the stochastic arrival traffic envelope, and the rate-variance envelope applied to the group of flows $G_{i,j}$ respectively.

4 Providing Absolute Differentiate Services with Statistical Guarantees

In the following, we will use a utilization-based test to realize admission control in a scalable fashion. During system (re-)configuration, a *safe utilization bound* is determined, which is then used for the tests during run time. As long as the utilization values of links along the path of a new flow do not exceed the safe utilization bound, the performance guarantees of all the flows will be met. The value for this utilization bound depends on network topology, traffic characteristics, and performance requirements of flows.

The challenge of using any utilization-based admission control method is to verify the correctness of a safe utilization bound at configuration time. Two critical issues have to be addressed:

- **Statistical Delay Analysis:** The main challenge for statistical delay analysis is how to clearly describe the traffic arrival process. The strong assumptions on the stochastic properties of traffic streams are inherently difficult for the network to enforce or police. Consequently, if a particular application does not conform to the chosen stochastic model, no guarantees can be made. Moreover, if admitted to the network, such a stream could adversely affect the performance of other applications if it is statistically multiplexed with them. In this paper, we will use an approach previously developed in [12] to do the statistical delay analysis.

- **Flow-Population-Insensitive Delay Analysis:** The delay violation probability derived in [12] depends on the information about flow population, i.e., the number of flows at input links and the traffic characteristics (e.g., the average rate and burst size) of flows. In our case the delay analysis is done at the system configuration time, when information about flow population is not available. Hence, it is necessary to derive a flow-population-insensitive formula. We will apply an approach similar to that used in [24] for deterministic traffic to solve this.

4.1 Statistical Delay Analysis at Configuration Time

First, we will address the issue of statistical delay analysis, and then derive the flow-population-insensitive statistical delay formula.

4.1.1 Statistical Delay Analysis

A probabilistic performance guarantee can be defined as a bound on the probability of exceeding a deadline as follows:

$$P\{D > d\} \leq \epsilon, \quad (5)$$

where the delay D suffered by a packet is a random variable, d is the given deadline, and ϵ is the given violation probability, which is generally small.

In statistical service, all input traffic conforms to a set of random processes. Suppose these processes are independent. If we know the mean value and the variance of each individual traffic random variable, and the number of flows is large enough, then by the *Central Limit Theorem*, we can approximate the random process of the combined flows. The *Central Limit Theorem* states that the summation of a set of independent random variables converges in distribution to a random variable that has a *Normal Distribution*. Actually, using rate-variance envelopes, the traffic arrival rate of each individual flow is a random variable, and the mean rate and the rate-variance of each individual flow can be determined using deterministic traffic models. The following theorem can be found in [12]:

Theorem 1 Consider a static-priority scheduler with L input links and link capacity C such that the traffic with Class i has an associated deadline d_i . Suppose that the group of flows $G_{i,j}$ has mean rate $\phi_{i,j}$ and rate-variance envelope $RV_{i,j}(t)$. With application of a Gaussian approximation over intervals, the delay violation probability for a random packet with Class i is approximately bounded by

$$P\{D_i > d_i\}$$

$$\begin{aligned} &\leq \max_{t < \beta_i} P\left\{\frac{1}{C} \left(\sum_{q=1}^{i-1} \sum_{j=1}^L B_{q,j}(t + d_i) + \sum_{j=1}^L B_{i,j}(t) \right) \geq td_i\right\} \\ &\leq \max_{t < \beta_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(C(t + d_i) - \mu_i(t))^2}{2\sigma_i^2(t)}\right), \end{aligned} \quad (6)$$

where

$$\mu_i(t) = (t + d_i) \sum_{q=1}^{i-1} \sum_{j=1}^L \phi_{q,j} + t \sum_{j=1}^L \phi_{i,j}, \quad (7)$$

$$\sigma_i^2(t) = (t + d_i)^2 \sum_{q=1}^{i-1} \sum_{j=1}^L RV_{q,j}(t + d_i) + t^2 \sum_{j=1}^L RV_{i,j}(t), \quad (8)$$

and

$$\beta_i = \min\left\{t : \sum_{q=1}^i \sum_{j=1}^L b_{q,j}(t) \leq Ct, t > 0\right\}. \quad (9)$$

By this theorem, the delay violation probability for any random packet can be computed approximately. In the above formula, the question left is how to get the values of *mean rate* and *rate-variance envelope*. In [12], two methods are presented for obtaining the rate-variance envelope:

- **Adversarial Mode:** the traffic arrival process conforms to a binomial distribution, where the rate-variance envelope is upper bounded.
- **Non-adversarial Mode:** the traffic arrival process conforms to a weighted uniform distribution, where the rate-variance envelope is approximated but non-worst-case.

Therefore, given the aggregated arrival traffic constraint function $b_{i,j}(t)$ of all $n_{i,j}$ flows, we can specify the *mean rate* and the *rate-variance envelope* as a function of $n_{i,j}$ etc.. By Theorem 2 in [24], the aggregated arrival traffic constraint function $b_{i,j}(t)$ is given as follows

$$b_{i,j}(t) = \begin{cases} Ct, & t \leq \tau_{i,j} \\ n_{i,j}(\sigma_i + \rho_i t), & t > \tau_{i,j} \end{cases} \quad (10)$$

where

$$\tau_{i,j} = \frac{n_{i,j}\sigma_i}{C - n_{i,j}\rho_i}, \quad (11)$$

Therefore, we have the following theorem:

Theorem 2 Given the same condition as Theorem 1, the mean rate $\phi_{i,j}$ is

$$\phi_{i,j} = n_{i,j}\rho_i, \quad (12)$$

and the rate-variance envelope is upper bounded by:

- *Adversarial Mode*

$$RV_{i,j}(t) \leq \frac{1}{t} (n_{i,j})^2 \rho_i \sigma_i \quad (13)$$

- *Non-adversarial Mode*

$$RV_{i,j}(t) \approx \frac{1}{12t} (n_{i,j})^2 \rho_i \sigma_i \quad (14)$$

The proof of Theorem 2 is given in Appendix.

At this point, the only undetermined is the number of flows on each link. In the following subsections, we describe how we eliminate the dependency on the number of flows on each link. The result is a delay formula that can be applied without knowledge of the flow population.

4.1.2 Flow-Population-Insensitive Statistical Delay Analysis

As we described earlier, admission control at run-time makes sure that the link utilization allocated to each class of flows is not exceeded. The total number N_i of flows of Class i from all input links is therefore subject to the following constraint:

$$N_i = \sum_{j=1}^L n_{i,j} \leq \frac{\alpha_i}{\rho_i} C, \quad (15)$$

where α_i is the ratio of the link bandwidth allocated to traffic of Class i . With this constraint, by (12), (13), and (14), the mean rate and the rate-variance can be upper-bounded. Therefore, the delay violation probability can be upper-bounded without relying on information about the flow-population. This is shown by the following theorem:

Theorem 3 Consider a static priority scheduler with L input links and link capacity C such that the traffic with Class i has an associated deadline d_i . Suppose that the group of flows $G_{i,j}$ has a stochastic envelope $B_{i,j}(t)$. Then the delay violation probability for a random packet of Class i is bounded by

- *Adversarial Mode*

$$P\{D_i > d_i\} \leq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \min_{t < \beta_i} \xi_i(t)\right), \quad (16)$$

- *Non-adversarial Mode*

$$P\{D_i > d_i\} \leq \frac{1}{\sqrt{2\pi}} \exp\left(-6 \min_{t < \beta_i} \xi_i(t)\right), \quad (17)$$

where

$$\xi_i(t) = \frac{(\eta_i t + \eta_{i-1} d_i)^2}{\zeta_i t + \zeta_{i-1} d_i}, \quad (18)$$

$$\beta_i = \frac{1}{\eta_i} \sum_{q=1}^i \alpha_q \frac{\sigma_q}{\rho_q}, \quad (19)$$

and

$$\eta_p = 1 - \sum_{q=1}^p \alpha_q, \quad (20)$$

$$\zeta_p = \sum_{q=1}^p (\alpha_q)^2 \frac{\sigma_q}{\rho_q}, \quad (21)$$

in (20) and (21), the value for p is either $i-1$ or i .

This is the main result of this paper. We observe that the formula does not depend on the flow population. Hence it can be used for utilization bound verification at configuration time. In the following, we will give the proof. The basic idea is that using inequality (15), flow-population will be removed.

Proof: Substituting (12), (13) and (14) into (7) and (8), we have

$$\begin{aligned} C(t + d_i) - \mu_i(t) &= t(C - \sum_{q=1}^i \sum_{j=1}^L n_{q,j} \rho_q) \\ &\quad + d_i(C - \sum_{q=1}^{i-1} \sum_{j=1}^L n_{q,j} \rho_q) \end{aligned} \quad (22)$$

and

- *Adversarial Mode*

$$\begin{aligned} \sigma_i^2(t) &= t \sum_{q=1}^i \sum_{j=1}^L (n_{q,j})^2 \rho_q \sigma_q \\ &\quad + d_i \sum_{q=1}^{i-1} \sum_{j=1}^L (n_{q,j})^2 \rho_q \sigma_q. \end{aligned} \quad (23)$$

- *Non-adversarial Mode*

$$\begin{aligned} \sigma_i^2(t) &= \frac{1}{12} \left(t \sum_{q=1}^i \sum_{j=1}^L (n_{q,j})^2 \rho_q \sigma_q \right. \\ &\quad \left. + d_i \sum_{q=1}^{i-1} \sum_{j=1}^L (n_{q,j})^2 \rho_q \sigma_q \right). \end{aligned} \quad (24)$$

By (15), we have

$$\sum_{j=1}^L n_{q,j} \rho_q \leq \alpha_q C, \quad (25)$$

and

$$\sum_{j=1}^L (n_{q,j})^2 \rho_q \sigma_q \leq (\alpha_q C)^2 \frac{\sigma_q}{\rho_q}, \quad (26)$$

therefore,

- for the adversarial mode

$$\frac{(C(t+d_i) - \mu_i(t))^2}{2\sigma_i^2(t)} \geq \frac{1}{2}\xi_i(t), \quad (27)$$

- and for the non-adversarial mode

$$\frac{(C(t+d_i) - \mu_i(t))^2}{2\sigma_i^2(t)} \geq 6\xi_i(t), \quad (28)$$

where $\xi_i(t)$ is defined in (18). ■

As an illustrative example, we apply the above theorem in the case of two classes, only one real-time class traffic. Suppose that $\alpha = \alpha_1$, $\sigma = \sigma_1$, $\rho = \rho_1$, $d = d_1$, $D = D_1$, $\xi(t) = \xi_1(t)$, $\beta = \beta_1$, for real-time class traffic. Simplifying the formula, we can get the following corollary:

Corollary 1 *In the case of a single real-time class, the delay violation probability for a random real-time class traffic packet is bounded by*

- adversarial mode

$$P\{D > d\} \leq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \min_{t < \beta} \xi(t)\right), \quad (29)$$

- and for the non-adversarial mode

$$P\{D > d\} \leq \frac{1}{\sqrt{2\pi}} \exp\left(-6 \min_{t < \beta} \xi(t)\right), \quad (30)$$

where

$$\xi(t) = \frac{((1-\alpha)t + d)^2}{\alpha^2 \frac{\sigma}{\rho} t}, \quad (31)$$

and

$$\beta = \frac{\alpha}{1-\alpha} \frac{\sigma}{\rho}. \quad (32)$$

Define the right hand side of (29) and (30) as ϵ_1 and ϵ_2 . We can find that $\xi(t)$ reaches its minimum at $t = t_0 (= \frac{d}{1-\alpha})$. By the property of function $\xi(t)$, we find

- if $\beta \geq t_0$, i.e., $\alpha \geq \frac{d}{\rho}$, then

$$\epsilon_1 = \frac{1}{\sqrt{2\pi}} \exp\left(-2 \frac{1-\alpha}{\alpha^2} \frac{d}{\rho}\right) \quad (33)$$

$$\epsilon_2 = \frac{1}{\sqrt{2\pi}} \exp\left(-24 \frac{1-\alpha}{\alpha^2} \frac{d}{\rho}\right) \quad (34)$$

- if $\beta < t_0$, i.e., $\alpha < \frac{d}{\rho}$, then

$$\epsilon_1 = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{1-\alpha}{\alpha^3} \left(\alpha + \frac{d}{\rho}\right)^2\right) \quad (35)$$

$$\epsilon_2 = \frac{1}{\sqrt{2\pi}} \exp\left(-6 \frac{1-\alpha}{\alpha^3} \left(\alpha + \frac{d}{\rho}\right)^2\right) \quad (36)$$

By the above formulas, we know that as α_k decrease to below $\frac{d}{\rho}$, ϵ_1 and ϵ_2 will quickly approach zero.

For the deterministic model, by Theorem 4 in [26], a bandwidth ratio formula can be derived as follows:

$$\alpha = \frac{d}{\frac{\sigma}{\rho}}. \quad (37)$$

Therefore, the bandwidth ratio value in the deterministic model is a critical point to the one in statistical model. Below this value, the delay violation probability is quite small and quickly approaches zero.

4.2 Verification of Utilization Bound

Having derived the flow-population-insensitive statistical delay formula, we can verify the utilization bound, and obtain the *worst-case achievable utilization* (WCAU). By the *worst-case achievable utilization*, we refer to the maximum of total utilization bound for all classes. Under the condition that the probabilistic delay guarantee can be met, we can compute the WCAU.

For a given delay violation probability ϵ_i and deadline d_i along route \mathcal{R} , we can split d_i into $\{d_i^k : k \in \mathcal{R}\}$, and the delay guarantee is met when

$$\begin{aligned} P\{D_i^{e^{2e}} > \sum_{k \in \mathcal{R}} d_i^k\} \\ &\leq 1 - \prod_{k \in \mathcal{R}} (1 - P\{D_i^k > d_i^k\}) \\ &\leq \epsilon_i. \end{aligned} \quad (38)$$

By Theorem 3, for the adversarial mode and the non-adversarial mode, substituting (16) and (17) into (39) respectively, we solve the inequalities and get the maximum value of α_i , for $i = 1, 2, \dots, M$. The WCAU is $\sum_{i=1}^M \alpha_i$.

5 Experimental Evaluation

In this section, we evaluate the performance of the approaches discussed in the previous sections. We will first define the performance metrics, and then describe the system configuration and present the performance results.

5.1 Experimental Model

- *Performance Metrics:* We are interested in two metrics:

- WCAU: The *worst-case achievable utilization* (WCAU) is the maximum of total utilization bound for all classes. We use this metric to measure the performance of the systems.
- Admission Probability: This is the probability that a flow is admitted in a stable system (all packets have bounded delays). The higher the admission probability, the better the network resources are being used.

- **Traffic Model:** We assume that the traffic belongs to a single real-time class. We simulate voice traffic, with bursts of 640 bits, an average rate of 32kbit/sec. We assume that the deadline is 5msec. We assume that requests for flow establishment form a Poisson process with rate λ , and that flow lifetimes are exponentially distributed with an average of 180 seconds. The real system would support best-effort traffic as well, which would not affect the results of this evaluation, and is therefore omitted in these experiments.
- **Evaluation Methods:** The WCAU can be computed by (16) and (17) using simple binary search. The admission probability of systems we consider can be analyzed by queuing theory and fixed point method [17].

5.2 Numerical Results and Observations

In this sub-section, we report performance results and make observations. Due to the limited space, we only present a limited number of cases here. However, we find that the conclusions we draw here generally hold for many other cases we have evaluated.

ϵ	WCAU		
	deterministic	adversarial	non-adversarial
0	0.250	–	–
10^{-6}	–	0.250	0.488
10^{-4}	–	0.250	0.563
10^{-2}	–	0.307	0.699

Table 1. Sensitivity of WCAU to Delay Violation Probability

5.2.1 Sensitivity of WCAU to Delay Violation Probability

Data on sensitivity of utilization to delay violation probability are given in Figures 1 and Table 1. Figure 1 and Table 1 show the data on the cases of the delay violation probability ϵ . From Figures 1 and Table 1, we have the following observations:

- As expected, the utilization for both adversarial and non-adversarial model increases as the delay violation probability increases. This means that the higher the delay violation probability, the higher utilization we can get for both adversarial and non-adversarial models. A higher delay violation probability allows for larger bandwidth allocations and, therefore, higher values for the WCAU. Both statistical models can achieve higher or equal value for WCAU than the deterministic

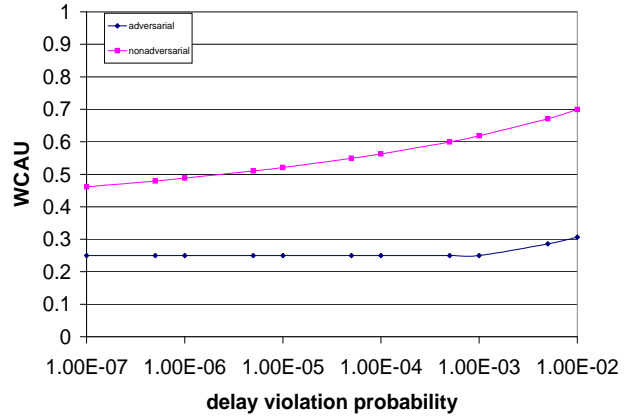


Figure 1. Sensitivity of WCAU to Delay Violation Probability

model. Since the deterministic model does not allow delay violations, more resources need to be reserved, which decreases WCAU.

- The non-adversarial models achieve much higher WCAU than deterministic ones for any delay violation probability. Non-adversarial models are much closer to real traffic models, and much better exploit available statistical multiplexing gain [12].

5.2.2 Admission Probability Comparison of Deterministic and Statistical Model

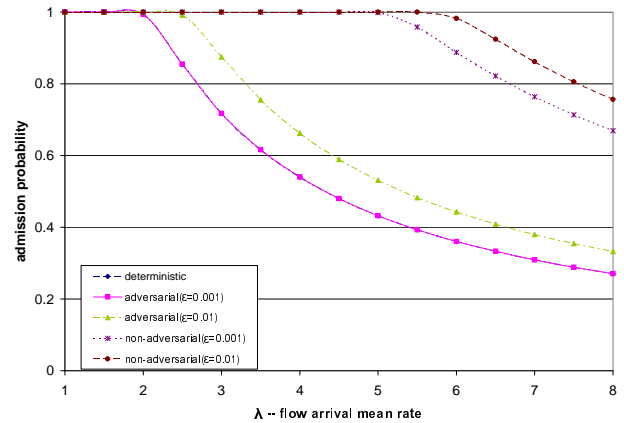


Figure 2. Admission Probability Comparison of Deterministic and Statistical Model

Data on sensitivity of admission probability is given in Figure 2 (Note that the curve for deterministic model and the curve for statistical model (adversarial, $\epsilon = 0.001$) over-

laps). From this figure, we make the following observations:

- Admission probability is sensitive to the flow arrival rate λ for all models. Admission probability decreases as λ increases in all models. The reason is obvious: A large λ value implies a large number of flows in the system. Since the bandwidth is limited, some flows are not allowed to enter the network, therefore, the admission probability decreases.
- Different models have different sensitivities to λ in terms of admission probability. The statistical models always achieve higher admission probabilities than the deterministic model. Non-adversarial models always achieve higher admission probabilities than adversarial models. This is because of the achievable utilization. The higher this utilization, the higher the admission probability.

6 Conclusions

In this paper, we have proposed a methodology for providing absolute differentiated services with *statistical* performance guarantees in networks that use static priority schedulers. Given that static priority schedulers are widely supported by current routers, we believe that our approach is practical and effective to support real-time applications in existing networks.

We have used Utilization-based Admission Control (UBAC) approach which uses a configuration-time verification to determine a safe utilization level of servers. Admission control at runtime then is reduced to simple utilization tests on the servers along the path of the new flow. Hence, the approach is scalable.

Obviously, the verification will have to be involved with delay analysis. The general delay derivation depends on flow population information, which is not available at the system configuration time. We have extended the general approach and developed a method that allows us to analyze the delays without depending on the dynamic status of flow population.

Acknowledgements

This work was partially sponsored by NSF under contract number EIA-0081761, by DARPA under contract number F30602-99-1-0531, and by Texas Higher Education Coordinating Board under its Advanced Technology Program.

References

- [1] G. Agrawal, B. Chen, W. Zhao, and S. Davari, *Guaranteeing Synchronous Message Deadlines with the Timed Token Protocol*, Proceedings of IEEE ICDCS, June 1992.
- [2] R. Braden and D. Clark and S. Shenker, *Integrated Services in the Internet Architecture: an Overview*, Internet RFC 1633, June 1994.
- [3] C. Chang, *Stability, queue length, and delay of deterministic and stochastic queueing networks*, IEEE Transactions on Automatic Control, 39(5):913-931, May 1994.
- [4] Anna Charny and J. Y. Le Boudec, *Delay Bounds in a Network with Aggregate Scheduling*, Proceedings of QOFIS, October 2000.
- [5] R. L. Cruz, *SCED+: efficient management of quality of service guarantees*, Proceedings of IEEE INFOCOM, March 1998.
- [6] A. Dailianas and A. Bovopoulos, *Real-time admission control algorithms with delay and loss guarantees in ATM networks*, Proceedings of IEEE INFOCOM, June 1994.
- [7] C. Dovrolis, D. Stiliadis and P. Ramanathan, *Proportional Differentiated Services: Delay Differentiation and Packet Scheduling*, Proceedings of ACM SIGCOMM, August 1999.
- [8] V. Firoiu, J. Kurose, and D. Towsley, *Efficient admission control for EDF schedulers*, Proceedings of INFOCOM, April 1997.
- [9] S. Kweon and K. Shin, *Video-on-demand service using a statistical traffic envelope*, Technical report, University of Michigan, Ann Arbor, MI, 1998.
- [10] E. Knightly, *H-BIND: A new approach to providing statistical performance guarantees to VBR traffic*, Proceedings of IEEE INFOCOM, March 1996.
- [11] E. Knightly, *Second Moment Resource Allocation in Multi-Service Networks*, Proceedings of ACM SIGMETRICS, June 1997.
- [12] E. Knightly, *Enforceable quality of service guarantees for bursty traffic streams*, Proceedings of IEEE INFOCOM, March 1998.
- [13] J. Kurose, *On computing per-session performance bounds in high-speed multi-hop computer networks*, ACM Sigmetrics, June 1992.

- [14] J. Liebeherr, S. D. Patek and E. Yilmaz, *Tradeoffs in Designing Networks with End-to-End Statistical QoS Guarantees*, Proceedings of IEEE/IFIP IWQoS, June 2000.
- [15] C. L. Liu and J. W. Layland, *Scheduling algorithms for multiprogramming in a hard real time environment*, Journal ACM, Vol. 20, No. 1, 1973, pp.46-61.
- [16] J. Liebeherr, D.E. Wrege, and D. Ferrari, *Exact admission control in networks with bounded delay services*, IEEE/ACM Transactions on Networking, 1996.
- [17] D. Mitra and J. A. Morrison, *Erlang capacity and uniform approximations for shared unbuffered resources* IEEE/ACM Transactions on Networking, 1994.
- [18] K. Nicols, V. Jacobson, L. Zhang, *A Two-bit Differentiated Services Architecture for the Internet*, Internet-Draft, Nov. 1997.
- [19] J. Qiu and E. Knightly, *Inter-class resource sharing using statistical service envelopes*, Proceedings of IEEE INFOCOM, March 1999.
- [20] S. Shenker, C. Partridge and R. Guerin, *Specification of Guaranteed Quality of Service*, RFC2212, September 1997.
- [21] R. Sivakumar, T. Kim, N. Venkitaraman and V. Bharghavan, *Achieving Per-Flow Weighted Rate Fairness in a Core Stateless Network*, Proceedings of IEEE ICDCS, March 2000.
- [22] D. Starobinski and M. Sidi, *Stochastically bounded burstiness for communication networks*, Proceedings of IEEE INFOCOM, March 1999.
- [23] I. Stoica, H. Zhang, *Providing Guaranteed Services Without Per Flow Management*, Proceedings of ACM SIGCOMM, August 1999.
- [24] S. Wang, D. Xuan, R. Bettati and W. Zhao, *Providing Absolute Differentiated Services for Real-Time Applications in Static Priority Scheduling Networks*, Proceedings of IEEE INFOCOM, April 2001.
- [25] D. Wrege, E. Knightly, H. Zhang, and J. Liebeherr, *Deterministic delay bounds for VBR video in packet-switching networks: Fundamental limits and practical tradeoffs*, IEEE/ACM Transactions on Networking, 4(3):352-362, June 1996.
- [26] D. Xuan, C. Li, R. Bettati, J. Chen and W. Zhao, *Utilization-Based Admission Control for Real-Time Application*, Proceedings of IEEE ICPP, 2000.
- [27] O. Yaron and M. Sidi, *Performance and stability of communication networks via robust exponential bounds*, IEEE/ACM Transactions on Networking, 1(3):372-385, June 1993.
- [28] H. Zhang and E. Knightly, *Providing end-to-end statistical performance guarantees with bounding interval dependent stochastic models*, Proceedings of ACM SIGMETRICS, May 1994.

Appendix: Proof of Theorem 2

The following lemmas (1, 2, and 3) [12] define the mean rate for a group of flows and show how an upper bound on the stochastic rate variance envelope can be derived from the deterministic parameters.

Lemma 1 (Mean Rate) *The mean rate of the group of flows $G_{i,j}$ can be defined as:*

$$\phi_{i,j} = \lim_{t \rightarrow \infty} \frac{b_{i,j}(t)}{t}. \quad (40)$$

Lemma 2 (Adversarial Mode) *The rate-variance envelope of the group of flows $G_{i,j}$ is upper bounded by:*

$$RV_{i,j}(t) \leq RV_{i,j}^*(t) = \phi_{i,j} \left(\frac{b_{i,j}(t)}{t} - \phi_{i,j} \right) \quad (41)$$

where $\phi_{i,j}$ is defined in (40).

Lemma 3 (Non-adversarial Mode) *The rate-variance envelope of the group of flows $G_{i,j}$ is approximately:*

$$RV_{i,j}(t) \approx \widehat{RV}_{i,j}(t) = \frac{\phi_{i,j}}{12} \left(\frac{b_{i,j}(t)}{t} - \phi_{i,j} \right) \quad (42)$$

where $\phi_{i,j}$ is defined in (40).

We know that the aggregated arrival traffic constraint function $b_{i,k,j}(t)$ is given as follows

$$b_{i,j}(t) = \begin{cases} Ct, & t \leq \tau_{i,j} \\ n_{i,j}(\sigma_i + \rho_i t), & t > \tau_{i,j} \end{cases} \quad (43)$$

Then applying (43) to (40), (41), and (42), Theorem 2 can be proved as claimed.