The output of a switch, or, effective bandwidths for networks

Damon Wischik, University of Cambridge

Statistical Laboratory, Mill Lane, Cambridge CB2 1SB, UK.
D.J.Wischik@statslab.cam.ac.uk, http://www.statslab.cam.ac.uk/~djw1005

February 1999

Abstract

Consider a switch which queues traffic from many independent input flows. We show that in the large deviations limiting regime in which the number of inputs increases and the service rate and buffer size are increased in proportion, the statistical characteristics of a flow are essentially unchanged by passage through the switch. This significantly simplifies the analysis of networks of switches. It means that each traffic flow in a network can be assigned an effective bandwidth, independent of the other flows, and the behaviour of any switch in the network depends only on the effective bandwidths of the flows using it.

Keywords. Effective bandwidths, feedforward networks, large deviations, decoupling bandwidths, output of a switch, many sources.

1 Introduction

A *switch* is a device that routes traffic. A switch has several input flows of traffic, each of which is routed to a specified destination; and inside the switch, work from all of the inputs is queued together. Switches are the building blocks of modern telecommunications networks.

The behaviour of isolated switches has been much studied. The theory of Large Deviations can be used to estimate the probability that the queue overflows, to study different queueing regimes, and to characterize the input flows.

In this paper, we study networks of switches. The fundamental result is that, under the many sources limiting regime, the large deviations characteristics of a flow of traffic are not changed by passing through a switch. This means that the techniques for analysing isolated switches can be applied inductively to networks. It also means that it is useful to talk about the characteristics of a type of traffic, without bothering about how many switches the flow has passed through or what other flows it has interacted with.

This work was carried out under a Research Studentship from the UK Engineering and Physical Sciences Research Council.

The theory of Large Deviations is concerned with limiting regimes, and it is the choice of limiting regime which makes possible such clean results for networks. In the *many sources* limiting regime studied here, the number of independent inputs to each switch increases, and the buffer size per input and service rate per input stay fixed. Such a regime is well-suited to studying modern high-speed telecommunications networks, in which switches typically have many inputs but only a small buffer, and the cell loss probability is required to be very small.

These results may be interpreted in terms of *effective bandwidths*, an intuitive and appealing way of viewing switches as resources. A flow entering a switch has effective bandwidth α if it has the same impact on the switch as would a flow of constant rate α; the effective bandwidth of the flow is a function which depends on the operating point of the switch (see Courcoubetis, Siris and Stamoulis [2] for an account of this dependence). The theory of effective bandwidths at isolated switches started a decade ago with a paper by Hui [7]. The results proved here show that the effective bandwidth function for the output flow is the same as for the input flow: so a flow has the same effective bandwidth function through the entire network.

The rest of this paper is in four sections. Section 2 describes the large deviations behaviour of an isolated switch. Section 3 proves the fundamental result, that the statistical characteristics of a flow are not changed by passing through a switch. Section 4 presents this result in the language of effective bandwidths. Section 5 considers limitations and extensions of these results, and compares them to results for another common limiting regime.

2 An Isolated Switch

This section introduces the large deviations theory used to describe the behaviour of an isolated switch. It summarises relevant results from Wischik [17]. For an introduction to large deviations, and definitions of the terms used here, see Dembo and Zeitouni [5]. We will content ourselves with explaining what is meant by a large deviations principle.

A sequence of random variables X^L in a Hausdorff space $\mathcal{X}$ with Borel σ-algebra $\mathcal{B}$ is said to satisfy a large deviations principle (LDP) with good rate function I if for any $B \in \mathcal{B}$,

$$\begin{aligned} -\inf_{x\in B^\circ} I(X) &\leq \liminf_{L\to\infty} \frac{1}{L} \log \mathbb{P}(X^L \in B) \\ &\leq \limsup_{L\to\infty} \frac{1}{L} \log \mathbb{P}(X^L \in B) \leq -\inf_{x\in \bar{B}} I(X), \end{aligned}$$

where $I : \mathcal{X} \to \mathbb{R}^+ \cup \{\infty\}$ has compact level sets. If $\mathcal{X}$ is a space of processes, this is called a sample path LDP.

The starting point is to give conditions under which a sample path large deviations principle for the average of independent processes holds. From the sample path LDP, it is easy to deduce LDPs for the amount of work queued in the switch and for many other quantities of interest in queueing theory. Essentially, it gives a complete characterization of a process, as far as the large deviations queueing theorist is concerned. It will be shown in Section 3 that if the input flows satisfy the conditions for a sample path LDP to hold, then the output flows also satisfy a sample path LDP.

The Sample Path LDP

First, the notation. We will be concerned with the space $\mathcal{X}$ of real-valued processes indexed by the natural numbers. Throughout this paper, t will denote a natural number. Denote a process in $\mathcal{X}$ by $\mathbf{X}(0,\infty)$ or $\mathbf{X}$, and its truncation to the set $\{1\dots t\}$ by $\mathbf{X}(0,t]$. Denote by X_t the value of the process at time t, and by $X(0,t]$ the cumulative process $X(0,t]=\sum_{i=1}^{t}X_i$, with $X(0,0]=0$. Consider a sequence of processes $\mathbf{X}^L$. Think of $\mathbf{X}^L$ as the average of L independent processes each distributed like $\mathbf{X}^{(L)}$.

Under the following conditions, $\mathbf{X}^L$ satisfies the sample path LDP stated in Theorem 1 below. Details of the proof may be found in Wischik [17], as well as several examples including long-range dependent processes. The proof works by finding an LDP for finite truncations $\mathbf{X}^L(0,t]$ using the Gärtner-Ellis theorem, extending this to $\mathcal{X}$ equipped with the projective limit topology using the Dawson-Gärtner theorem, and then strengthening the topology using the Inverse Contraction theorem. These steps are similar to those used by O'Connell [13] who finds a sample path LDP under a different limiting regime.

CONDITION **1 (Finite-time characteristics)** *For $\boldsymbol{\theta}\in\mathbb{R}^t$, define*

$$\mathbf{\Lambda}_t^L(\boldsymbol{\theta})=\frac{1}{L}\log\mathbb{E}\exp(L\boldsymbol{\theta}\cdot\mathbf{X}^L(0,t]).$$

Assume that for each t and $\boldsymbol{\theta}$, the limiting moment generating function

$$\mathbf{\Lambda}_t(\boldsymbol{\theta})=\lim_{L\to\infty}\mathbf{\Lambda}_t^L(\boldsymbol{\theta})$$

exists as an extended real number, and that the origin belongs to the interior of the effective domain of $\mathbf{\Lambda}_t$, and that $\mathbf{\Lambda}_t$ is an essentially smooth, lower-semicontinuous function.

CONDITION **2 (Large timescale characteristics)** *A* scaling function *is a function $v:\mathbb{N}\to\mathbb{R}$ for which $v(t)/\log t\to\infty$. For some scaling function v, define the scaled cumulant moment generating function*

$$\Lambda_t^L(\theta)=\frac{1}{v(t)}\mathbf{\Lambda}_t^L(\mathbf{1}\theta v(t)/t)$$

for $\theta\in\mathbb{R}$, where $\mathbf{1}$ is the constant vector of ones. From Condition 1, we know that for each t there is an open neighbourhood of the origin in which the limit

$$\Lambda_t(\theta)=\lim_{L\to\infty}\Lambda_t^L(\theta)$$

exists. Assume that there is an open neighbourhood of the origin in which those limits and the limit

$$\Lambda(\theta)=\lim_{t\to\infty}\Lambda_t(\theta)$$

exist uniformly in θ.

We also know from Condition 1 that for θ sufficiently small, the limit $\Lambda_t^L(\theta) - \Lambda_t(\theta) \to 0$ is uniform as $L \to \infty$. Assume that for θ sufficiently small the limit

$$\sqrt{\frac{v(t)}{\log t}}\Big(\Lambda_t^L(\theta) - \Lambda_t(\theta)\Big) \to 0$$

is uniform in θ as $t, L \to \infty$.

DEFINITION **3 (Stability)** *Define the* mean rate λ *of $\mathbf{X}^L$ to be the derivative $\Lambda'(0)$. Say that $\mathbf{X}^L$ is* stationary *if the limiting moment generating functions $\mathbf{\Lambda}_t$ correspond to a stationary process. We will also use these terms to describe $\mathbf{\Lambda}_t^L$. If $\mathbf{X}^L$ is stationary, λ is given by $\lambda = \frac{1}{t}\mathbf{\Lambda}_t'(\theta\mathbf{1})$ at $\theta = 0$ for all t.*

THEOREM **1 (Sample Path LDP)**
Suppose $\mathbf{X}^L$ satisfies Conditions 1 and 2. Then for any μ greater than its mean rate, $\mathbf{X}^L$ satisfy a LDP on the space

$$\mathcal{X}_\mu = \left\{\mathbf{x} \in \mathcal{X} : \frac{x(0,t]}{t} \le \mu \quad \textit{eventually}\right\}$$

equipped with the uniform topology

$$\|\mathbf{x}\| = \sup_{t>0}\left|\frac{x(0,t]}{t}\right|,$$

with good rate function

$$\mathbf{I}(\mathbf{x}) = \sup_{t>0}\sup_{\boldsymbol{\theta}\in\mathbb{R}^t} \boldsymbol{\theta}\cdot\mathbf{x}(0,t] - \mathbf{\Lambda}_t(\boldsymbol{\theta}).$$

The sample path LDP can be used to generate LDPs for a wide range of queueing problems using the Contraction Principle. This says that if f is a continuous function on $\mathcal{X}_\mu$, then $f(\mathbf{X}^L)$ satisfies an LDP with good rate function $I(y) = \inf\{\mathbf{I}(\mathbf{x}) : \mathbf{x} \in \mathcal{X}_\mu, f(\mathbf{x}) = y\}$. Wischik [17] uses this to study standard queues, likely sample paths to overflow, and priority queues. For the purposes of this paper, we will mainly be interested in standard queues with finite buffers, described in the next section.

The Queueing Model

Consider a standard first-in-first-out finite-buffer queue. Define $Q(\mathbf{x})$ to be the amount of work in the queue at time 0 when fed with input x_t at time $-t$ and served at a constant service rate C. Let B be the buffer size. Any work arriving when the queue is full is lost. We will be concerned with the behaviour of a queue fed with input process $\mathbf{X}^L$, the average of L independent processes.

To be concrete, let us assume that we are interested in the probability that the queue overflows. The following theorem gives a large deviations estimate for this event. Wischik [17] shows how it may be derived by contracting the sample path LDP. It is also possible to derive it directly from the above assumptions, as in Courcoubetis and Weber [3]. Duffield and Botvich [6] have proved a similar result for queues with infinite buffers, as have Simonian and Guibert [16] for the special case of Markov modulated fluid traffic. If we can show that the output process from a queue satisfies the above assumptions, we can immediately estimate the probability of overflow in a downstream queue.

THEOREM **2 (Buffer Overflow)** *If $\mathbf{X}^L$ satisfies Conditions 1 and 2, is stationary, and has mean rate strictly less than C, then the event that the queue overflows has rate*

$$\inf_t \sup_\theta \theta(B + Ct) - \mathbf{\Lambda}_t(\theta \mathbf{1}).$$

In a sense, though, it does not matter what the downstream queueing regime is: because for a wide range of queueing regimes there are similar results which can be derived from the sample path LDP. That is why it is so useful.

To analyse the output of a switch, we will need another large deviations result: an estimate for the probability that the queue is empty. If the queue is empty at the beginning and end of an interval (and does not overflow during the interval) then the amount of work leaving the queue in that interval is exactly the amount of work entering the queue. As usual, the following lemma can be proved by contracting the sample path LDP. It has also been proved directly by Courcoubetis and Weber, and Duffield and Botvich have proved a similar result for queues with infinite buffers.

LEMMA **3 (Buffer Empties)** *If $\mathbf{X}^L$ satisfies Conditions 1 and 2, is stationary, and has mean rate strictly less than C, then the event that the queue is non-empty has large deviations upper bound*

$$I = \sup_\theta \theta C - \mathbf{\Lambda}_1(\theta \mathbf{1}).$$

This section has stressed the usefulness of the sample path LDP in analysing the behaviour of isolated queues. It has also proposed a step in understanding networks of switches: show that if the inputs to a switch satisfy the conditions for the sample path LDP, then the outputs also satisfy a sample path LDP. This will be proved in the following section.

3 Switch outputs, decoupling, and networks

In this section, we show that in feedforward networks of queues with a mixture of traffic flows, under the many sources limiting regime, the large deviations characteristics of a flow are the same at all points along its route. This makes it easy to analyse such networks. First, we prove the fundamental result: that for a single switch with many independent inputs of the same type, the large deviations characteristics of an output are the same as those of an input.

3.1 The Output of a Switch

Consider the queueing model from the previous section, in which the total input $\mathbf{X}^L$ to queue L is the average of L independent identically distributed input process. Let $\mathbf{X}^{(L)}$ be a typical input process, and $\tilde{\mathbf{X}}^{(L)}$ the corresponding output. The generating function $\mathbf{\Lambda}_t^L$ for the aggregate input is therefore

$$\mathbf{\Lambda}_t^L(\boldsymbol{\theta}) = \log \mathbb{E} \exp(\boldsymbol{\theta} \cdot \mathbf{X}^{(L)}).$$

Similarly, the moment generating function for the aggregate of *independent copies* of a typical output is

$$\tilde{\mathbf{\Lambda}}_t^L(\boldsymbol{\theta}) = \log \mathbb{E} \exp(\boldsymbol{\theta} \cdot \tilde{\mathbf{X}}^{(L)}).$$

We consider this aggregate because, in our large deviations analysis, it is the natural way to describe the behaviour of a *single* output.

Let $\mathbf{\Lambda}_t^L$ satisfy Conditions 1 and 2, and be stationary with mean rate strictly less than the service rate. Let the limiting moment generating functions be $\mathbf{\Lambda}_t$. We will show that the output moment generating function $\tilde{\mathbf{\Lambda}}_t^L$ also satisfies a sample path LDP, with the same limiting moment generating functions $\mathbf{\Lambda}_t$.

The key observation is that the probability that the queue is empty over a fixed interval tends to 1, and so the probability that the input and output processes are identical over that interval tends to 1 also. The key result is Theorem 4, which says that over a fixed interval there is not only convergence in probability, but also convergence of the moment generating functions. This shows that $\tilde{\mathbf{\Lambda}}_t^L$ satisfies Condition 1, which is all that is needed to establish a sample path LDP for the output over a fixed interval.

To obtain the full LDP of Theorem 1, we would like to show that $\tilde{\mathbf{\Lambda}}_t^L$ satisfies Condition 2, which is a technical condition on the uniformity of convergence. In fact that condition is not satisfied, and we have not been able to establish Theorem 1 for the output. This is not actually a problem. A sample path LDP still holds, under the weak queue topology defined below. This topology is weaker than the uniform topology, but as noted in Wischik [17] it is strong enough to obtain all the results in that paper for queues with finite buffers, including Theorem 2 and Lemma 3. The sample path LDP is shown in Theorem 5.

The output process is stable in the same sense as the input process: that is, it does not exceed the mean rate in the long run. This is shown in Lemma 6.

THEOREM **4 (Finite-time characteristics of the output)**
If the input $\mathbf{X}^{(L)}$ satisfies Conditions 1 and 2, and is stationary with mean rate strictly less than C, then the output $\tilde{\mathbf{X}}^{(L)}$ satisfies Condition 1, with the same limiting moment generating function as $\mathbf{X}^{(L)}$. In other words,

$$\lim_{L\to\infty} \log \mathbb{E}\exp(\boldsymbol{\theta}\cdot\tilde{\mathbf{X}}^{(L)}(0,t]) = \mathbf{\Lambda}_t(\boldsymbol{\theta}).$$

Proof. First note that $\tilde{X}^{(L)}(0,t] \le X^{(L)}(0,t+\lfloor B/C\rfloor]$, since any work arriving before $-\lfloor B/C\rfloor$, even if it finds the queue full, must have left by time 0. In what follows, we drop the $\lfloor\cdot\rfloor$ notation.

For fixed t, the collection $\{\exp(\boldsymbol{\theta}\cdot\tilde{\mathbf{X}}^{(L)}(0,t])\}$ is uniformly integrable, since $0 \le \boldsymbol{\theta}\cdot\tilde{\mathbf{X}}^{(L)}(0,t] \le \max|\theta_i| X^{(L)}(0,t+B/C]$, and $X^{(L)}(0,t+B/C]$ is L^p-bounded for some $p>1$ (because the limiting moment generating function exists, by Condition 1).

For any $1\le s\le t$, $\mathbb{P}(\tilde{X}_s^{(L)} \ne X_s^{(L)})$ is bounded by the probability that the queue is non-empty at either $s-1$ or s. By Theorem 2, this tends to 0. So $\exp(\boldsymbol{\theta}\cdot\tilde{\mathbf{X}}^{(L)}(0,t]) - \exp(\boldsymbol{\theta}\cdot\mathbf{X}^{(L)}(0,t])$ converges to 0 in probability.

Thus $\mathbb{E}\exp(\boldsymbol{\theta}\cdot\tilde{\mathbf{X}}^{(L)}(0,t]) - \mathbb{E}\exp(\boldsymbol{\theta}\cdot\mathbf{X}^{(L)}(0,t]) \to 0$, and taking logarithms gives the result. □

DEFINITION **4 (Weak queue topology)** *Define the* weak queue topology *wq on $\mathcal{X}$ by the metric*

$$d(\mathbf{x},\mathbf{y}) = |Q(\mathbf{x})-Q(\mathbf{y})| + \sum_{t=1}^{\infty} \frac{1\wedge|x_t-y_t|}{2^t} \qquad (1)$$

and $d(\mathbf{x},\mathbf{y})=\infty$ *if* $Q(\mathbf{x})=\infty$ *or* $Q(\mathbf{y})=\infty$.

THEOREM **5 (Large timescale characteristics of the output)**
If the input $\mathbf{X}^{(L)}$ *satisfies Conditions 1 and 2, and is stationary with mean rate strictly less than* C*, then the output* $\tilde{\mathbf{X}}^{(L)}$ *satisfies satisfies an LDP in* $(\mathcal{X},wq)$ *with good rate function* $\mathbf{I}$.

Proof. First, by the Dawson-Gärtner theorem for projective limits (see [5] Theorem 4.6.1), the finite time LDPs of Theorem 4 can be extended to the full space $\mathcal{X}$ equipped with the projective limit topology, with good rate function $\mathbf{I}$. The projective limit topology corresponds to pointwise convergence of sequences, and can be made into a metric space with the metric given by the second term in (1). Denote this topology by p.

We want to strengthen this LDP from $(\mathcal{X},p)$ to $(\mathcal{X},wq)$. To do this we will use the Inverse Contraction Principle ([5] Theorem 4.2.4). Since wq is stronger than p, the identity map from $(\mathcal{X},wq)$ to $(\mathcal{X},p)$ is continuous. And $\tilde{\mathbf{X}}^{(L)}$ satisfies an LDP in $(\mathcal{X},p)$ with rate function $\mathbf{I}$. So if $\tilde{\mathbf{X}}^{(L)}$ is exponentially tight in $(\mathcal{X},wq)$ then it satisfies an LDP in $(\mathcal{X},wq)$ with the same rate function, and that rate function is good.

It remains to show that $\tilde{\mathbf{X}}^{(L)}$ is exponentially tight in $(\mathcal{X},wq)$: in other words that there exist compact sets K_α in $(\mathcal{X},wq)$ such that

$$\lim_{\alpha\to\infty}\limsup_{L\to\infty}\log\mathbb{P}(\tilde{\mathbf{X}}^{(L)}\notin K_\alpha)=-\infty. \tag{2}$$

Let μ be the mean rate of the $\mathbf{X}^L$, let $d_t=\sqrt{\log t/v(t)}$, and choose the sets

$$K_\alpha=\left\{\mathbf{x}:0\le\frac{x(0,t]}{t+B/C}\le\mu+\alpha d_{t+B/C}\right\}.$$

First, to show that K_α is compact. Since $\mathcal{X}$ is a metric space, it suffices to show that it is sequentially compact. So let $\mathbf{x}^k$ be a sequence of processes. Since the T-dimensional truncation of K_α is compact in $\mathbb{R}^t$, the intersection K_α is compact under the projective topology. That is, there is a subsequence $\mathbf{x}^j(k)$ which converges pointwise, say to $\mathbf{x}$. It remains to show that $\mathbf{x}^j\to\mathbf{x}$ under the weak queue topology. But if $\mathbf{x}\in K_\alpha$, there exists a t_0 such that for $t>t_0$, $x(0,t]/t<C$, and this t_0 can be chosen independently of $\mathbf{x}$. Therefore the queue size is just $Q(\mathbf{x}^j)=\sup_{t\le t_0}x^j(0,t]-Ct$, which converges because the $\mathbf{x}^j$ converge pointwise. Thus K_α is compact.

Next, to show the equation (2). Since $\tilde{\mathbf{X}}^{(L)}(0,t]\le\mathbf{X}^{(L)}(0,t+B/C]$, the left hand side is bounded above by the expression in the statement of [17] Lemma 5, which is there shown to equal $-\infty$. □

LEMMA **6 (Output stability)** *If the input* $\mathbf{X}^{(L)}$ *satisfies Conditions 1 and 2, and is stationary with mean rate strictly less than* C*, then for any* μ *greater than the mean rate, the output process* $\tilde{\mathbf{X}}^{(L)}$ *satisfies a sample path LDP in* $\mathcal{X}_\mu$ *equipped with the weak queue topology, with good rate function* $\mathbf{I}$.

Proof. We want to restrict the LDP of Theorem 5 to $\mathcal{X}_\mu$. By [5] Lemma 4.1.5. it suffices to show that $\mathbf{I}(\mathbf{x})=\infty$ if $\mathbf{x}\notin\mathcal{X}_\mu$, and that $\mathbb{P}(\tilde{\mathbf{X}}^L\in\mathcal{X}_\mu)=1$. The proof of the first is identical to [17] Theorem 6. For the second, that theorem also shows that for ε sufficiently small, $\mathbb{P}(\mathbf{X}^L(0,t]/t \le \mu - \varepsilon \text{ eventually}) = 1$, and since $\tilde{\mathbf{X}}^L(0,t] \le \mathbf{X}^L(0,t+B/C]$, we obtain the result. □

3.2 Traffic Mixes

In the last section we assumed that the aggregate input $\mathbf{X}^L$ to the switch was the average of L independent identically distributed input processes. This was used in two ways. First, it gave a large deviations estimate for the probability that the queue is non-empty. Second, it let us describe a typical input using the moment generating function for the aggregate, $\mathbf{\Lambda}_t^L$.

We can still estimate the probability that the queue is non-empty and describe a typical input, even when the aggregate input is not made up of independent identically distributed flows. Let $\mathbf{Y}^L$ be the aggregate input, and let $\mathbf{X}^{(L)}$ be the single input we are interested in. Define the moment generating functions $\mathbf{M}_t^L(\boldsymbol{\theta}) = \frac{1}{L}\log\mathbb{E}\exp(\boldsymbol{\theta}\cdot\mathbf{Y}^L)$ and $\mathbf{\Lambda}_t^L(\boldsymbol{\theta}) = \log\mathbb{E}\exp(\boldsymbol{\theta}\cdot\mathbf{X}^{(L)})$. Suppose that $\mathbf{M}$ and $\mathbf{\Lambda}$ satisfy Conditions 1 and 2, and are stationary, and that the mean rate of the aggregate input is less than the service rate. Then $\mathbf{M}$ gives a large deviations estimate for the event that the queue is non-empty, and $\mathbf{\Lambda}$ describes the input we are interested in. Theorems 4 and 5 go through unchanged, except that the rate I will depend on $\mathbf{M}$ rather than on $\mathbf{\Lambda}$.

There are many different ways of scaling the system to meet these conditions, with different numbers of inputs of different types. For example, let the aggregate input be made up of a mix of traffic types: $L\rho(j)$ copies of $\mathbf{X}^{(L)}(j)$ for $j = 1\dots J$, each traffic type satisfying Conditions 1 and 2. Then $\mathbf{M}$ is just a linear combination of the moment generating functions for the different traffic types.

Another example is when the aggregate input is made up of L flows that were independent and identical when they entered the network, but which have passed through several queues before reaching the queue Q we are considering. Allow each flow to follow a different route, possibly involving feedback and interaction with other flows. This is interesting because it makes the flows neither independent nor identical. Let the maximum delay that each flow can incur before reaching Q be less than $D < \infty$. Let the aggregate input to Q be $Y^L(0,t]$; this is less than the original aggregate input $X^L(0,t+D]$ over a longer time interval. It can be shown that if the mean rate of the $\mathbf{X}^L$ is less than C/D, a queue with service rate C fed with $X^L(0,t+D]$ still empties with high probability, and so Q empties with high probability, and the results of the last section apply. (Unfortunately, since the inputs to a queue are not independent, we cannot use this to find $\mathbf{Y}^L$ and estimate the probability of overflow.)

3.3 Decoupling of Flows

Consider two independent inputs $\mathbf{X}$ and $\mathbf{Y}$ to a switch whose aggregate input satisfies Conditions 1 and 2, and is stationary with mean rate less than the service rate. (The (L) notation has been dropped here.) We know from the previous sections that in the limit, $\tilde{\mathbf{X}}$ has the same distribution as $\mathbf{X}$, and that $\tilde{\mathbf{Y}}$ has the same distribution as $\mathbf{Y}$. We can also view $\mathbf{X}+\mathbf{Y}$ as a single input to the queue, note that $\tilde{\mathbf{X}}+\tilde{\mathbf{Y}}$ has the same distribution as $\mathbf{X}+\mathbf{Y}$, and deduce that in the limit $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ are independent.

It might be expected that traffic flows would influence each other. For example, if $\mathbf{X}$ is very bursty and $\mathbf{Y}$ is smooth, one might expect $\tilde{\mathbf{X}}$ to be less bursty and $\tilde{\mathbf{Y}}$ to be less smooth. But we see that this is not the case. In other words, $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ do *not* depend on the traffic mix at the switch (so long as the total mean input rate is less than the service rate). This is known as *decoupling.*

3.4 Networks of Switches

A feedforward network of switches is one in which the switches may be ordered in such a way that for every flow the sequence of switches through which it passes is strictly increasing. In the last section, it was shown that a flow passing through a switch is essentially unchanged, even if several different types of flows use the switch, in the limiting regime where the number of flows increases. This can be applied to a feedforward network of switches, as long as the network is scaled also.

Consider, for example, a simple network of two switches in tandem. Let the first switch have L independent inputs, each distributed like $\mathbf{X}^{(L)}$. Let one of the outputs $\tilde{\mathbf{X}}^{(L)}$ be fed into the downstream switch, along with a further $L-1$ independent copies of $\tilde{\mathbf{X}}^{(L)}$ from other upstream switches. Then the aggregate input to the downstream switch satisfies a sample path LDP with the same rate function as that appearing in the LDP for the aggregate input to the upstream switch, so we can estimate the overflow probability of the downstream queue with standard techniques.

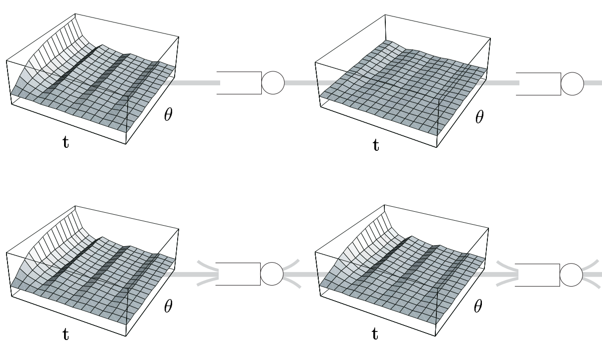

Figure 1: Switches in tandem. With just one input to each switch, the traffic is smoothed, and the behaviour of the downstream switch will depend on the degree of smoothing. With three inputs to each switch (independent and identically distributed), there is little smoothing, and the behaviour of the downstream switch is easy to predict.

Figure 1 gives simulation results to illustrate this, and to compare it to the case of switches handling a single flow of traffic. (The graphs show the effective bandwidth of a flow at different points in the network. The effective bandwidth $\alpha(\theta,t)$ is a convenient representation of the moment generating function: $(\theta t)^{-1}\mathbf{\Lambda}_t(\theta\mathbf{1})$.) The source illustrated is a periodic process of random phase, emitting one unit of work every fourth time step. The switches have service rate 0.4 per input and buffer size 1.5 per input. When there is only one flow of traffic using the switch, it is smoothed. When there are three flows of traffic using the switch, each flow is hardly smoothed at all.

For switches which are further downstream in the network, the proofs of Section 3 still work, if the maximum delay incurred by a flow at a switch B/C is replaced by the maximum delay incurred by a flow in reaching the switch under consideration.

4 Effective Bandwidth

In this section, we recast the results of Sections 2 and 3 into the language of effective bandwidths. For a full discussion of effective bandwidths, see Kelly [9]. We will see that the idea of effective bandwidth extends from describing flow-at-a-point to describing flow-through-the-network.

Effective Bandwidth for an isolated switch

As in Section 3.3, consider a switch with service rate C and buffer size B, and an input with moment generating functions $\mathbf{\Lambda}_t$. Recall from Theorem 2 that the rate for overflowing is

$$I = \inf_t \sup_\theta \theta(B + Ct) - \theta t \alpha(\theta, t), \tag{3}$$

where the effective bandwidth α is defined to be $\alpha(\theta, t) = (\theta t)^{-1} \mathbf{\Lambda}_t(\theta \mathbf{1})$.

(This is shorthand for the following. Consider a sequence of switches indexed by L, with switch L having service rate LC and buffer size LB. Let switch L have L independent inputs distributed like $\mathbf{X}^{(L)}$, where $\mathbf{\Lambda}_t$ is the limiting moment generating function of $\mathbf{X}^{(L)}(0, t]$. Then, $\lim_{L\to\infty} L^{-1} \log \mathbb{P}(\text{switch } L \text{ overflows})$ is equal to $-I$. But this is a cumbersome description, so we will stick with the shorthand.)

Consider replacing a small proportion δ of the inputs by flows of which produce work at a constant rate a; these have effective bandwidth a. The rate function for overflowing is now

$$I(\delta) = \inf_t \sup_\theta \theta(B + Ct) - \theta t((1 - \delta)\alpha(\theta, t) + \delta a).$$

If the optimizing parameters are $\hat{\theta}$ and $\hat{t}$, and under appropriate differentiability conditions, the value of a that makes $I'(0) = 0$ is $a = \alpha(\hat{\theta}, \hat{t})$. In other words, an input flow has the same effect on the system as would a constant flow of rate $\alpha(\hat{\theta}, \hat{t})$. This is why α is called the effective bandwidth function.

If the the switch has multiple input flows of different types, then the effective bandwidth function measures the tradeoff between the different flows. For example, if at the operating point $(\hat{\theta}, \hat{t})$ of the switch the effective bandwidth of flow 1 is twice that of flow 2, then replacing a small number of flows of type 1 by twice that number of flows of type 2 will not affect the cell loss probability.

Effective Bandwidths for Networks

It is simple to extend effective bandwidths from isolated switches to networks, using the results of Sections 3.3 and 3.4, which say that the statistical characteristics of a flow are not altered by passing through a switch, even when that switch is shared with different types of flow. Therefore the effective bandwidth function of a flow is the same at all points in the network (though the different switches will typically have different operating points, so the values of the function will be different). This simplifies the theory and engineering of networks in many ways.

It means, for example, that the effective bandwidth of a traffic flow in packet-switched networks plays a similar role to the bandwidth of a call in loss networks. This encourages the belief that well-understood techniques and insights from loss networks can be applied to packet-switched networks. For a review of those techniques and insights, see Kelly [8].

It also makes it easier to understand feedback and rate control for adaptive traffic, that is, traffic which can alter its rate in response to congestion-indicating signals from the network. It is natural to believe that feedback from a switch to a user should depend on the characteristics of the traffic from that user, *as seen by the switch*. If the effective bandwidth function changed along the route, depending on interactions with other flows at other switches, then the user might have difficulty in making effective use of the feedback signals, because she would not know how her traffic had been shaped by the intervening switches. But it does not change, and so she can better interpret feedback.

The key idea is that it is meaningful to talk about the characteristics of, say, video traffic, because the flow retains these characteristics regardless of its interactions with other flows in various switches throughout the network.

5 Discussion

The core of the argument is Theorem 4, which proves that the limiting moment generating function of the output process is the same as that of the input. It relies on the fact that when there are many independent sources, the queue empties regularly, with high probability. This is a reasonable engineering constraint for high-performance networks, in which delay and cell loss probabilities should be small. This constraint is satisfied by any work-conserving queue (i.e., any queue which does not idle when there is work waiting).

The theorem is proved for the case of a queue with a finite buffer. It seems likely that the result still holds for queues with infinite buffers and for other regimes like priority queues. The finiteness of the buffer is used to bound the amount of work that can leave the queue over a period of time, to give uniform integrability; for those other cases some other way of proving uniform integrability would be needed.

We have not dwelt on the question of how many input processes are needed for this limiting result to be accurate. Numerical simulation, illustrated in Figure 1, suggests that in some cases only a small number of independent inputs are needed to make the input and output look nearly identical. The real question, though, is: how many input processes are needed for reasonable convergence *over the scale of interest*? If we are interested in the probability of overflow at a downstream switch, we want reasonable convergence of the moment generating function at the critical timescale and spacescale for that switch. (The critical timescale and spacescale are the optimizing parameters θ and t appearing in Theorem 2.) For fixed θ and t, by keeping track of the bounds in Theorem 4, the difference between the moment generating functions for the input and output can be approximated by an expression which decays exponentially in LI, where L is the number of inputs and I is the rate function for the event that the upstream queue is nonempty. The accuracy of the large deviations estimate of Theorem 2 must also be taken into account; this has been studied by Likhanov and Mazumdar [10].

When feedforward networks are so simple, it is tempting to conjecture that similar results might hold in general networks. There are numerous examples of pathalogical behaviour in finite networks. But in large networks, under this many sources regime, we expect that switches will still empty sufficiently often, and the main result will still hold.

Other limiting regimes

The limiting regime studied here is called the many sources asymptotic, because it looks at the limit where the number of independent sources increases (and buffer size and service rate increase also). It is this choice of limiting regime that permits such clean results for networks. Another limiting regime which has been widely studied is the *large buffer asymptotic*, which looks at a single process over increasing timescales (with the buffer size increasing also).

The output of a queue under this regime has been widely studied: see, for example, O'Connell [14, 12], Paschalidis [15], Bertsimas, Paschalidis and Tsitsiklis [1], Majewski [11] and de Veciana, Courcoubetis and Walrand [4]. Under this regime it can be shown that the statistical characteristics of a flow *are* changed by passing through a queue: the output is less bursty than the input. If the input is made up of several flows, they *do not* decouple. So although it is possible to define an effective bandwidth for a flow at a single switch, this effective bandwidth does not extend naturally to networks, and a more complicated network calculus is needed.

Conclusion

We have shown that the relevant statistical characteristics of a flow of traffic are preserved by passage through a switch, in the limit where the number of inputs to that switch increases.

This is a limiting result. But simulation suggests that it can still be reasonably accurate even for a handful of independent sources. And the theory is useful at least as much for the insights it gives as for numerical estimates.

It dramatically simplifies the analysis of networks of switches with different classes of traffic.

Acknowledgements

I am grateful to Frank Kelly for many valuable discussions, and to the referees for their helpful comments.

References

[1] D. Bertsimas, I. C. Paschalidis and J. N. Tsitsiklis. On the large deviations behaviour of acyclic networks of G/G/1 queues. Tech. Report LIDS-P-

2278, MIT Laboratory for Information and Decision Systems (December 1994).

[2] C. Courcoubetis, V. A. Siris and G. D. Stamoulis. Application of the many sources asymptotic and effective bandwidth to traffic engineering. To appear in Telecommunication Systems. Preprint available from http:// www.ics.forth.gr/ netgroup/publications/.

[3] C. Courcoubetis and R. Weber. Buffer overflow asymptotics for a buffer handling many traffic sources. J. Appl. Prob. (1996). 33 886–903.

[4] G. de Veciana, C. Courcoubetis and J. Walrand. Decoupling bandwidths for networks: a decomposition approach to resource management for networks. In: *Proc. IEEE Infocom.* volume 2, pages 466–474, 1994 pages 466–474.

[5] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett (1993).

[6] N. Duffield and D. Botvich. Large deviations, the shape of the loss curve, and economies of scale in large multiplexers. Queueing Systems (1995). 20 293–320.

[7] J. Y. Hui. Resource allocation for broadband networks. IEEE J. Sel. Areas Commun. (1988). 6 1598–1608.

[8] F. Kelly. Loss networks. The Annals of Applied Probability (1991). 1(3) 319–378. Special invited paper.

[9] F. Kelly. Notes on effective bandwidths. In: *Stochastic Networks: Theory and Applications*, eds. F. P. Kelly, S. Zachary and I. Ziedins, Oxford, Royal Statistical Society Lecture Note Series. 1996 pages 141—168.

[10] N. Likhanov and R. R. Mazumdar. Cell loss asymptotics for buffers fed with a large number of independent stationary sources. Journal of Applied Probability (1999). To appear.

[11] K. Majewski. *Large Deviations of Feedforward Queueing Networks.* Ph.D. thesis, Ludwig-Maximilians-Universität München (May 1996).

[12] N. O'Connell. Queue lengths and departures at single-server resources. In: *Stochastic Networks: Theory and Applications*, eds. F. P. Kelly, S. Zachary and I. Ziedins, Oxford. 1996 pages 91–104.

[13] N. O'Connell. A large deviation principle with queueing applications. Technical Report HPL-BRIMS-97-05, BRIMS, Hewlett Packard Labs, Bristol (March 1997).

[14] N. O'Connell. Large deviations for departures from a shared buffer. J. Appl. Prob. (1997). 34 753–766.

[15] I. C. Paschalidis. *Large Deviations in High Speed Communications Networks.* Ph.D. thesis, MIT Laboratory for Information and Decision Systems, Cambridge, MA, USA (May 1996).

[16] A. Simonian and J. Guibert. Large deviations approximation for fluid sources fed by a large number of on/off sources. IEEE J. Selected Areas Commun. (1995). 13 1017–1027.

[17] D. J. Wischik. Sample path large deviations for queues with many inputs. To appear.