RC 20649 (91385) 12/5/96

Computer Science/Mathematics 41 pages

IBM Research Report

Performance Bounds for Guaranteed and Adaptive Services

Rajeev Agrawal
Department of Electrical and Computer Engineering
University of Wisconsin–Madison
Madison, WI 53706, USA.
agrawal@engr.wisc.edu.

Rajendran Rajan
T.J. Watson Research Center
IBM Research Division
Yorktown Heights, NY 10598, USA.
raju@watson.ibm.com.

IBM Research Division
Almaden · T.J. Watson · Tokyo · Zurich

Performance Bounds for Guaranteed and Adaptive Services

Rajeev Agrawal*
Department of Electrical and Computer Engineering
University of Wisconsin – Madison
Madison, WI 53706-1691
agrawal@engr.wisc.edu

Rajendran Rajan
IBM T. J. Watson Research Center
Hawthorne, NY 10532
raju@watson.ibm.com

Abstract

In this paper, we investigate issues related to the efficient support of the guaranteed and adaptive service categories in integrated services networks. The guaranteed service category is targeted at real-time applications that require hard end-to-end delay bounds for burstiness constrained traffic, while the adaptive service category is intended for sessions that have a minimum bandwidth requirement, and use a window based flow control mechanism to avail themselves of unused bandwidth in the network. The analytic framework presented in the paper uses two basic network elements – regulators and schedulers. Regulators enforce *burstiness constraints* on the traffic flow while schedulers ensure that a certain level of service (quantified by a *service curve*) is provided to the session by each network element it traverses. We model guaranteed service sessions as feed-forward networks of regulators and schedulers, and adaptive service sessions as similar networks with feedback. Our formulation is more general than prior work on schedulers and regulators, and allows for easy aggregation rules for networks of these elements. We present these rules and discuss several corollaries of this network calculus. In particular, we obtain bounds on delay, queue length and burstiness for unicast guaranteed sessions, and also discuss the relationship between throughput and buffering requirements for unicast and multicast adaptive sessions with end-to-end and hop-by-hop window flow control.

Keywords: Guaranteed service, adaptive service, service curve, regulator, scheduler, delay, burstiness, queueing.

*Research supported in part by NSF Grant No. NCR-9305018.

1. Introduction

Integrated services networks are meant to support a range of applications with diverse traffic characteristics and quality of service (QoS) requirements. Applications could have hard requirements on certain QoS parameters, while being tolerant of moderate lapses with regards to others. For instance, voice calls require bandwidth and delay guarantees while being tolerant of occasional cell loss, while e-mail and file transfers have the opposite requirements. Support for such application specific requirements is being introduced through a variety of networking standards; work by the RSVP and Integrated Services working groups in the IETF being the most relevant to this paper [2, 14, 16].

Networks that support diverse QoS requirements are provisioned to support fixed *service categories*, rather than allow sessions to negotiate each QoS parameter individually. Each service category assumes a constellation of application requirements, allowing sessions to precisely specify certain QoS parameters, mandating pre-specified network behavior with regards to certain others, and allowing individual network elements freedom in their treatment of the remaining. Moreover, service categories are specified at a high level of generality, so as to allow a variety of implementations through different scheduling, policing and buffer management mechanisms. In this paper, we present a general framework that may be used to describe and analyze two such service categories, which we term *guaranteed service* and *adaptive service*.

The guaranteed service is designed for applications that have stringent delay and loss requirements on a per-packet basis, while being able to specify reasonable upper limits on their traffic. A specific instance of a guaranteed service has been adopted by the IETF Int-Serv working group [14], where sessions specify an upper bound on their traffic by means of a peak rate, a mean rate and a burst size. Network elements characterize their level of service through a latency and a rate, which depend on the session's requirements. Typical applications that are targeted by such a service include audio and video teleconferencing, as well as playback.

In the first part of this paper we consider a general framework for guaranteed service, where a session specifies a *burstiness constraint* on the traffic generation process, while the network elements employ *regulators* and *schedulers* to protect individual sessions from one another. Regulators reshape traffic to conform to a specified burstiness constraint, and may be used to model devices such as leaky buckets that enforce traffic contracts at access points to the network or subnetworks. Further, regulators may also be used within the network, for instance, when link scheduling is performed using a *rate controlled service discipline* [17, 18] which uses regulators to isolate incoming traffic streams before arbitrating between competing streams through a scheduling mechanism. This use of regulators has the additional advantage of reducing jitter and buffering requirements at downstream network elements [4, 6]. Service curve schedulers model the operation of link scheduling mechansims that ensure that a certain level of service (quantified by a *service curve*) is provided to each of the sessions sharing the link. Such service allocation may reflect active scheduler involvement in isolating flows (as in virtual clock, generalized processor sharing (GPS), packet GPS, self-clocked fair queueing, weighted round-robin, etc.) or may be obtained through a careful accounting of the service capacity and arrival constraints on other connections (as when several burstiness constrained flows share a FIFO or priority scheduler). In Section 2, we introduce regulators and service curve schedulers as characterized by certain mathematical relationships between their arrival and departure processes. We then obtain bounds on queue lengths, delay and output burstiness for a session traversing a single network element (regulator or scheduler). In Section 3 we model a unicast guaranteed service session as a burstiness constrained flow traversing an *acyclic tandem of network elements*. We obtain a composition rule that reduces this analysis to the earlier single element analysis. We explore a number of consequences of the composition rule, especially in relation to the use of regulators to reduce buffering requirements at intermediate network elements.

The (σ, ρ) regulator was introduced by Cruz [3, 4] and the notion of service curve for GPS schedulers was introduced by Parekh and Gallager [11, 12, 13]. Of greatest relevance to the above results on guaranteed services is work by Parekh and Gallager [11, 12, 13] on GPS schedulers, by Cruz [5] on service curve schedulers and (σ, ρ) regulators in discrete time, and by Stiliadis and Varma [15] on *latency-rate* (a special class of service curve) schedulers. Other related work may be found in [7],[19], [13], [8], [1] and [6]. In this paper, we present a more general formulation of regulators and service curve schedulers than considered in previous work. This formulation holds for *general fluid* arrival and departure processes which could be mixtures of packetized or purely fluid flows. We present a number of new results for guaranteed sessions, apart from providing simpler derivations of some well-known performance bounds.

A number of applications on today's networks do not have stringent delay requirements, but would benefit from guaranteed bandwidth that they could always use. Furthermore, these applications, such as web browsers or bulk e-mail, use flow control mechanisms to estimate and share excess bandwidth in the network. An adaptive service, designed for use by such applications, would ensure that sessions that operated in compliance with the flow control mechanism receive lossless delivery of packets with a guarantee on the minimum bandwidth. In order to deliver such a service, network elements would have to provision buffers to store outstanding packets, and also employ scheduling mechanisms designed to provide some isolation between sessions.

In the second part of our paper we consider a framework for the adaptive service category, where besides regulators and service curve schedulers, the network also uses window based flow control mechanisms (end-to-end or hop-by-hop). In this case, the session topologies may be modelled as *networks with feedback*, where the flow of traffic is constrained by acknowledgments. Our general formulation of schedulers and regulators allows us to derive novel pathwise performance bounds for such networks. These results are presented in Sections 4 and 5 where we consider unicast as well as multicast sessions, with end-to-end and hop-by-hop window flow control, and lower bound their departure processes by that of an appropriately chosen service curve scheduler (without flow control). This allows us to obtain performance results by applying results obtained in the single element case. In particular, we obtain the relationship between the buffering requirements at network elements and the throughput achievable by the session for the above mentioned session topologies.

Finally, in Section 6 we show how it may be possible to obtain a tighter service curve characterization by incorporating packetization.

2. Network Elements in Isolation

In this section we formalize the operation of regulators and schedulers. We then discuss some simple performance measures for these network elements in isolation.

2.1 Regulators

Regulators are devices used to shape or smooth traffic to an *upper envelope* or *burstiness constraint*. Typically, such devices are used at network (or internetwork) boundaries to ensure that the stream of packets entering the network conforms to the traffic contract. Regulators may also be employed within a network, for instance, when link scheduling is performed using a *rate controlled service discipline* which uses regulators to isolate incoming traffic streams before arbitrating between competing streams through a scheduling mechanism. This has the advantage of reducing buffering requirements at downstream network elements.

A simple example of a regulator is a (σ, ρ) regulator which has the burstiness constraint $B(t) := \sigma + \rho t, t > 0$. This regulator is a device that holds up arrivals just long enough to ensure that the departing stream is (σ, ρ) constrained, i.e., the departure process D satisfies $D(t) - D(s) \leq B(t-s) = \sigma + \rho(t-s)$ for all $0 \leq s < t$. A popular implementation of a (σ, ρ) regulator is called a *leaky bucket*, where tokens arrive into a bucket of size σ at constant rate ρ. Departures occur only when tokens are available, and when they occur, cause the contents of the token bucket to be decremented by an equal volume. The (σ, ρ) regulator was introduced by Cruz [3, 5] and further generalized by Anantharam and Konstantopoulos [10].

In general, we allow a *burstiness constraint* $B : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ to be a non-decreasing and sub-additive function (i.e., $B(t) + B(s) \geq B(t+s)$ for all $t, s \geq 0$) with $B(0) = 0$. We specify a regulator in terms of a burstiness constraint B and treat it as a map that takes the general fluid arrival process A to a general fluid departure process D. Throughout this paper we assume all arrival and departure processes $A, D : \mathbb{R} \to \mathbb{R}_+ \cup \{\infty\}$ to be cumulative and therefore non-decreasing, and to be zero for negative times. We first characterize this regulator map through a set of properties that the departure process should satisfy, and then obtain an explicit solution to this characterization. This explicit solution allows us to treat regulators as a special case of service curve schedulers in Section 2.2.

The departure process D of a regulator with burstiness constraint B and arrival process A satisfies the following conditions:

R1. $D(t) - D(s) \leq B(t-s)$ for all $s \leq t$.

R2. $D(t) \leq A(t)$ for all $t \in \mathbb{R}$.

R3. D is the (pointwise) maximal function satisfying **R1** and **R2**.

The departure process satisfying the above conditions **R1**–**R3** is explicitly obtained in the following theorem.

Theorem 2.1. *There exits a unique D satisfying* **R1**–**R3** *and is given by*

$$D(t) = \inf_{s \leq t} [B(t-s) + A(s)], \qquad t \in \mathbb{R}. \tag{2.1}$$

Proof. From the sub-additivity of B it follows that for all $u \leq t$,

$$\begin{aligned} D(t) &= \inf_{s \leq t} [B(t-s) + A(s)] \\ &\leq \inf_{s \leq u} [B(t-s) + A(s)] \\ &\leq \inf_{s \leq u} [B(t-u) + B(u-s) + A(s)] \\ &= B(t-u) + \inf_{s \leq u} [B(u-s) + A(s)] \\ &= B(t-u) + D(u). \end{aligned}$$

From the condition $B(0) = 0$, it follows that

$$D(t) \leq B(0) + A(t) = A(t), \quad t \in \mathbb{R}.$$

Thus D satisfies **R1** and **R2**. Let D' be any departure process that also satisfies **R1** and **R2**. Then, for all $s \leq t$,

$$\begin{aligned} D'(t) &\leq B(t-s) + D'(s) \leq B(t-s) + A(s) \\ &\leq \inf_{s \leq t} [B(t-s) + A(s)] = D(t). \end{aligned}$$

Thus, D satisfies **R3**. □

2.2 Service Curve Schedulers

Service curves were introduced by Parekh and Gallagher [12, 13] to study generalized processor sharing (GPS), and extended by Cruz [5] and Stiliadis and Varma [15] who use service curves more generally, to characterize the level of service obtained by a session from the scheduler. The service curve characterization introduced in this section is more general than those found in the works cited before, and includes both discrete as well as continuous traffic models within the same analytic framework.

In Section 2.3 we show that pathwise, per-session performance bounds at a network element may be obtained from this service curve characterization of the scheduler, in conjunction with burstiness bounds on arriving traffic. In later sections, we show how service curves may be easily composed to obtain tight end-to end bounds across complex session topologies including unicast and multicast with or without end-to-end and hop-by-hop window flow control.

Example 2.2 (Latency Rate Service Curves). The simplest example of a service curve is a *latency-rate*. This service curve is defined by two parameters, a latency θ and a rate r. A scheduler allocates a latency-rate (θ, r) service curve to a session if for any cumulative arrival process A from that session, the volume of cumulative departures $D(t)$ at any time t is no smaller than $\inf_{s \leq t}\{r(t-s-\theta)^+ + A(s)\}$. One way to interpret this lower bound is to imagine the arrival process A passing through a constant rate server of rate r, followed by a delay element with latency θ. In this case, the cumulative departure process from the constant rate server of rate r is $D^r(t) = \inf_{s \leq t}\{r(t-s) + A(s)\}$ (see [9], for instance). Consequently, the cumulative volume of departures from the delay element with latency θ, up to time t is given by the expression $D^{\mathrm{l}}(t) = D^r(t-\theta) = \inf_{s \leq t-\theta}\{r(t-s-\theta)^+ + A(s)\}$. In other words, a latency rate service is one that ensures that departures occur no later than from the above-described tandem of constant rate server followed by a delay element.

In order to see how a latency rate service curve could arise from the operation of a scheduler, consider a simple example of two sessions – the first served as higher priority traffic with arrival process A' and the second with arrival process A served as lower priority traffic – both sharing a link of capacity C. Assume that the arrivals of the first session satisfy a (σ, ρ) burstiness constraint $B'(t) = \sigma + \rho t, t > 0$, with $\rho < C$. Now, the cumulative departure process D' of session 1 is given by

$$D'(t) = \inf_{s \le t} \{C(t-s) + A'(s)\}, \quad t \in \mathbb{R}.$$

It is well known that this departure process satisfies the burstiness constraint (see also Corollary 3.8)

$$D'(t) - D'(s) \le \min\{\sigma + \rho(t-s), C(t-s)\}, \quad 0 \le s < t.$$

The cumulative service $D^*(t)$ available for session 2 upto time t is the residual capacity of the link after serving the higher priority traffic, i.e., $D^*(t) := Ct - D'(t), t \ge 0$, $D^*(t) := 0, t < 0$. Then, by the burstiness constraint on D', it follows that D^* has the following lower envelope:

$$D^*(t) - D^*(s) \ge [(C-\rho)(t-s) - \sigma]^+ = (C-\rho)(t - s - \sigma/(C-\rho))^+, \quad 0 \le s < t.$$

The departure process D of the lower priority traffic is given in terms of its arrival process A and the service D^* available to it as

$$D(t) = \inf_{s \le t} \{D^*(t) - D^*(s) + A(s)\}, \quad t \in \mathbb{R}.$$

Combining this with the lower envelope on D^*, we get

$$D(t) \ge \inf_{s \le t} \{(C-\rho)(t - s - \sigma/(C-\rho))^+ + A(s)\}, \quad t \in \mathbb{R}.$$

and thus, the priority scheduler appears to the second connection as a latency rate scheduler, with latency $\theta = \sigma/(C-\rho)$ and rate $r = C - \rho$.

In general, we assume that the level of service allocated to a session by a scheduler is quantified by a *service curve* $S : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ which is non-decreasing with $S(0) = 0$. Let $A, D : \mathbb{R} \to \mathbb{R}_+ \cup \{\infty\}$ denote the session's arrival and departure processes into and from the scheduler, respectively. Then we say that the scheduler satisfies a service curve constraint S (or is a *service curve S scheduler*) if its departure process satisfies the following conditions:

$$A(t) \ge D(t) \ge \inf_{s \le t} [S(t-s) + A(s)], \qquad t \in \mathbb{R}. \tag{2.2}$$

Remark 2.3. *Note that a regulator with burstiness constraint B is a special case of a scheduler with service curve S, where the right inequality in* (2.2) *is replaced by an equality and $S = B$ is sub-additive.*

2.3 Analysis of Regulators and Schedulers in Isolation

In this section we study the performance of a session passing through a single network element – regulator or scheduler – whose operation is characterized through a service curve. In particular, we obtain tight bounds on the delay queue length and departure process burstiness given a burstiness constraint on the arrival process. Several of these results are well known for the service curve frameworks of Cruz [5], and Stiliadis and Varma [15]. Our principal motivation in re-working them is not merely to extend them to our more general service curve model, but to illustrate the simplicity of proofs obtained by the regulator/scheduler framework presented in the previous subsections.

In all results below, the cumulative arrival process of the session is denoted by A, its cumulative departure process by D, and the service curve offered by the network element by S. The burstiness constraint of A is denoted by B, i.e.,

$$A(t) - A(s) \leq B(t-s) \quad \forall s \leq t. \tag{2.3}$$

First, we bound the worst case delay associated with a network element, when the burstiness of the arrival process is known. Denote by $\Delta(t)$ the the delay suffered by the last arrival at time t, i.e.,

$$\Delta(t) := \inf\{\delta \geq 0 : D(t+\delta) \geq A(t)\}. \tag{2.4}$$

Proposition 2.4 (Upper Bound on Delay).

$$\Delta(t) \leq \Theta(B,S) := \inf\{\delta \geq 0 : S(u+\delta) - B(u) \geq 0 \quad \forall u \geq 0\}, \quad t \in \mathbb{R}. \tag{2.5}$$

Proof.

$$\begin{aligned} \Delta(t) &= \inf\{\delta \geq 0 : D(t+\delta) \geq A(t)\} \\ &\leq \inf\{\delta \geq 0 : \inf_{s \leq t+\delta} [S(t+\delta-s) + A(s)] \geq A(t)\} \end{aligned}$$

$$= \inf\{\delta \geq 0 : \inf_{s \leq t+\delta}[S(t+\delta-s) + A(s) - A(t)] \geq 0\}$$

$$= \inf\{\delta \geq 0 : \inf_{s \leq t}[S(t+\delta-s) + A(s) - A(t)] \geq 0 \quad \text{and} \inf_{t < s \leq t+\delta}[S(t+\delta-s) + A(s) - A(t)] \geq 0\}$$

Note that $A(s) - A(t) \geq 0$ if $t < s \leq t+\delta$. So the second term in the above expression may be ignored. Thus,

$$\begin{aligned}\Delta(t) &\leq \inf\{\delta \geq 0 : \inf_{s \leq t}[S(t+\delta-s) - (A(t) - A(s))] \geq 0\} \\ &\leq \inf\{\delta \geq 0 : \inf_{s \leq t}[S(t+\delta-s) - B(t-s)] \geq 0\} \\ &= \inf\{\delta \geq 0 : \inf_{u \geq 0}[S(u+\delta) - B(u)] \geq 0\}.\end{aligned}$$

□

The above delay bound may be seen to be the largest horizontal distance between the burstiness constraint and the service curve in Figure 1.

The next proposition obtains upper bounds on the queue size at the network element, and hence an upper bound on the buffers to be allocated at the network element for this session.

Proposition 2.5 (Upper Bound on Queue Size). *For each arrival process A, the queue length Q is bounded above by*

$$Q(t) := (A(t) - D(t)) \leq \sup_{s \leq t}((A(t) - A(s)) - S(t-s)), \quad t \in \mathbb{R}. \tag{2.6}$$

If the arrival process has a burstiness constraint B, then the queue length is bounded above by

$$Q(t) \leq \sup_{s \geq 0}(B(s) - S(s)) =: \Sigma(B, S), \quad t \geq 0.$$

Proof. The first part (pathwise bound) follows by simply substituting the lower bound for D in (2.2). If A satisfies the burstiness constraint B then substituting (2.3) in the pathwise bound concludes the proof. □

The above queue length bound may be seen to be the largest vertical distance between the burstiness constraint and the service curve in Figure 1.

As a session traverses multiple nodes, its traffic characteristics are altered by the operation of the network element. In general, the network element may reduce the burstiness by spacing out packets, or increase the burstiness by bunching up packets and transmitting them back-to-back. It is important to obtain tight bounds on the burstiness of the departure process from a network element, as this has clear implications for the buffering required at downstream nodes. In the next result we obtain such a bound.

Proposition 2.6 (Burstiness Constraint on Departure Process). *If A satisfies the burstiness constraint B, then the departure process D from a service curve S scheduler satisfies the burstiness constraint $B^D(B,S)$ given by*

$$(D(t+u) - D(t)) \leq \sup_{s \geq 0}\{B(u+s) - S(s)\} =: B^D(B,S)(u), \quad \forall t \geq 0, u > 0.$$

Proof. Replacing $D(t+u)$ and $D(t)$ with their upper and lower bounds, respectively, we get for all $t \geq 0, u > 0$,

$$\begin{aligned}
(D(t+u) - D(t)) &\leq A(t+u) - \inf_{r \leq t}\{A(r) + S(t-r)\} \\
&= \sup_{r \leq t}\{A(t+u) - A(r) - S(t-r)\} \\
&\leq \sup_{r \leq t}\{B(u+t-r) - S(t-r)\} \\
&= \sup_{s \geq 0}\{B(u+s) - S(s)\} =: B^D(B,S)(u).
\end{aligned}$$

□

Note that $B \leq B^D(B,S) \leq B + \Sigma(B,S)$. Thus, the departure process burstiness constraint is at least as large as the arrival process burstiness constraint, but no larger than the arrival process burstiness constraint plus the maximum queue length.

Remark 2.7. *The bounds obtained in Propositions 2.4 – 2.6 are tight.*

3. Guaranteed Service Sessions

Having obtained tight bounds on worst case delay, queue length and departure process burstiness for a single network element, we now wish to extend those results to a guaranteed service session traversing multiple network elements, each of which is characterized through a service curve. The principal observation that simplies this analysis is that the operation of a tandem of network elements, each described by a service curve, is analytically equivalent to the operation of a single network element whose service curve is obtained by composing the elements of the tandem. Thus, one may first compose different service curves and obtain tight end-to-end bounds from the results for the single element scheduler. In this section, we first obtain the above composition rule and then examine how buffering requirements of a guaranteed service session may be reduced by regulating its burstiness within the network.

Consider a unicast guaranteed service session traversing an open tandem of m service curve schedulers and/or regulators as shown in Figure 2. Let the arrival process to the tandem be denoted by A, the departure process by D, the arrival process into scheduler k by A_k and the departure process by D_k, $1 \le k \le m$. Then

$$A_1 = A, \quad A_k = D_{k-1}, k = 2, \ldots, m, \quad D := D_m,$$

and the departure process from scheduler k $(1 \le k \le m)$ is related to the arrival process into that scheduler by

$$A_k(t) \ge D_k(t) \ge \inf_{s \le t} [S_k(t-s) + A_k(s)], \qquad t \in \mathbb{R}. \tag{3.1}$$

Then, the tandem of schedulers behaves like a single scheduler with service curve S defined by the following composition rule.

Composition Rule for Tandems:

$$S(t) := \mathsf{T}(S_1, \cdots, S_m)(t) := \inf_{\substack{t_1, \cdots, t_m \ge 0 \\ t_1 + \cdots + t_m = t}} \{S_1(t_1) + \cdots + S_m(t_m)\}, \quad t \ge 0. \tag{3.2}$$

THEOREM 3.1 (Tandem of Schedulers as a Single Scheduler). *The composition S of service curves $S_1, \ldots, S_m$ is itself a service curve. Further, the departure process D from the tandem is related to the arrival process as*

$$A(t) \ge D(t) \ge \inf_{s \le t} [S(t-s) + A(s)], \qquad t \in \mathbb{R}, \tag{3.3}$$

and hence, the tandem of service curve schedulers is a scheduler with the service curve S.

PROOF. It is straightforward to show that the tandem of service curve schedulers behaves like a single scheduler by substituting in (3.1) the inequalities for $A_k = D_{k-1}$ in the corresponding inequalities for D_k. □

Remark 3.2. *The lower bound in* (3.3) *is attained and is hence tight.*

Corollary 3.3 (Tandem of Regulators as a Single Regulator). *Consider a tandem of regulators with burstiness constraints* $B_1, \ldots B_m$. *Their composition* $B := \mathsf{T}(B_1, \ldots, B_m)$ *is itself a burstiness constraint. Further, the departure process* D *from the tandem is related to the arrival process* A *as*

$$D(t) = \inf_{s \le t}[B(t-s) + A(s)], \qquad t \in \mathbb{R}, \tag{3.4}$$

and hence, the tandem of regulators is a regulator with burstiness constraint B.

Proof. We may show that B is sub-additive by noting that

$$\sum_{i=1}^{m} B_i(s_i) + \sum_{i=1}^{m} B_i(t_i) \ge \sum_{i=1}^{m} B_i(s_i + t_i),$$

for any $s_i \ge 0$ and $t_i \ge 0$, $1 \le i \le m$. Hence, with $s = \sum_i s_i$ and $t = \sum_i t_i$,

$$\begin{aligned} B(s) + B(t) &= \inf_{\substack{\sum_i s_i = s \\ s_1, \cdots, s_m \ge 0}} \sum_i B_i(s_i) + \inf_{\substack{\sum_i t_i = t \\ t_1, \cdots, t_m \ge 0}} \sum_i B_i(t_i) \\ &\ge \inf_{\substack{\sum_i s_i = s \\ s_1, \cdots, s_m \ge 0}} \inf_{\substack{\sum_i t_i = t \\ t_1, \cdots, t_m \ge 0}} \sum_i B_i(s_i + t_i) \\ &= \inf_{\substack{\sum_i u_i = s+t \\ u_1, \cdots, u_m \ge 0}} \sum_i B_i(u_i) = B(s+t). \end{aligned}$$

The rest of the corollary follows on the same lines as the derivation of (3.3). □

Corollary 3.4. *The end-to-end delay, queue length and burstiness bounds may be obtained from Propositions* 2.4, 2.5 *and* 2.6 *with* $S = \mathsf{T}(S_1, \cdots, S_m)$.

Proof. The result follows by combining Theorem 3.1 with the corresponding bounds in Propositions 2.4, 2.5 and 2.6. □

Example 3.5 (Tandem of latency-rate schedulers fed by a sigma-rho process). Consider a (σ, ρ) constrained arrival process feeding a tandem of m latency-rate schedulers $(\theta_k, r_k), 1 \le k \le m$. Then, it is easy to verify that $S = \mathsf{T}(S_1, \cdots, S_m)$ is a latency-rate service curve with latency $\theta = \sum_{k=1}^{m} \theta_k$ and rate $r = \min_{1 \le k \le m} r_k$. In this case the end-to-end delay bound is $\theta + \sigma/r$, the end-to-end queue length bound is $\sigma + \rho\theta$, and the burstiness constraint on the departure process is given by $B^D(t) = \sigma + \rho\theta + \rho t, t > 0$.

3.1 Applications of the Composition Rule for Tandems

In the rest of the section, we investigate a number of applications of the composition rule (3.2) that allows us to study a tandems of schedulers and regulators.

Proposition 3.6. *The order of network elements in the tandem may be changed with no effect on the composite service curve S and hence on the end-to-end bounds obtained in Corollary 3.4.*

Proof. Follows immediately from (3.2) as $S = \mathsf{T}(S_1, \cdots, S_m)$ does not depend on the order of the network elements. □

Proposition 3.7. *For any tandem of service curves $S_1, \ldots, S_m$, $S = \mathsf{T}(S_1, \cdots, S_m) \leq S_{\min} := \min\{S_1, \ldots, S_m\}$ with equality if $S_{\min}$ is sub-additive.*

Proof. It is easy to verify that $\mathsf{T}(S_1, \cdots, S_m) \leq S_k, 1 \leq k \leq m$, by picking $t_k = t$ in (3.2). Hence the first part follows. Now, in case $S_{\min}$ is sub-additive, then

$$\begin{aligned} S(t) = \mathsf{T}(S_1, \cdots, S_m)(t) &= \inf_{\substack{t_1, \ldots, t_m \geq 0 \\ t_1 + \cdots + t_m = t}} \{S_1(t_1) + \cdots + S_m(t_m)\} \\ &\geq \inf_{\substack{t_1, \ldots, t_m \geq 0 \\ t_1 + \cdots + t_m = t}} \{S_{\min}(t_1) + \cdots + S_{\min}(t_m)\} \\ &\geq \inf_{\substack{t_1, \ldots, t_m \geq 0 \\ t_1 + \cdots + t_m = t}} S_{\min}(t) = S_{\min}(t), \quad t \geq 0. \end{aligned}$$

□

Corollary 3.8. *For a tandem of of burstiness constraints $B_1, \ldots, B_m$ $B = \mathsf{T}(B_1, \cdots, B_m) = \min\{B_1, \ldots, B_m\}$ if either $B_1 = \cdots = B_m$ or if $B_1, \ldots, B_m$ are all concave.*

Proof. The first case is obvious. In the second case $\min\{B_1, \ldots, B_m\}$ is also concave and hence sub-additive. □

The above result may be used to show that a tandem of σ, ρ regulators gives rise to a concave piecewise linear regulator and conversely any such regulator may be obtained by a tandem of appropriately chosen σ, ρ regulators.

In the following proposition we establish that the end-to-end delay and queue length bounds obtained for the tandem by composing service curves is no greater than the quantity obtained by adding the element-wise bounds, when each element is fed by the original arrival process.

Proposition 3.9.

$$\Theta(B, S) \leq \sum_{k=1}^{m} \Theta(B, S_k).$$
$$\Sigma(B, S) \leq \sum_{k=1}^{m} \Sigma(B, S_k).$$

Proof. From (2.5), and that $B(0) = 0$, we see that

$$B((t - \Theta(B, S_k))^+) \leq S_k(t), \quad t \geq 0, 1 \leq k \leq m. \tag{3.5}$$

Now, by definition of the composition rule (3.2) we obtain

$$S(t + \sum_{k=1}^{m} \Theta(B, S_k)) = \lim_{n \to \infty} \sum_{k=1}^{m} S_k(t_k^n + \Theta(B, S_k))$$

for some real valued sequences $\{t_k^n\}_{n=1}^{\infty}$ converging to t_k, $1 \leq k \leq m$, with $\sum_{k=1}^{m}(t_k^n + \Theta(B, S_k)) = t$ for each n. From (3.5) and the sub-additivity of B we obtain

$$S(t + \sum_{k=1}^{m} \Theta(B, S_k)) \geq \lim_{n \to \infty} \sum_{k=1}^{m} B((t_k^n)^+) \geq B(\sum_{k=1}^{m} (t_k^n)^+) \geq B(t), \forall t \geq 0.$$

The corresponding inequality for the queue length bounds also follows similarly from the definition of the composition rule. □

Next, we consider the volume of buffering required at a network element to support a guaranteed service session. From Corollary 3.4 it follows that volume of packets in the tandem is bounded by $\Sigma(B, S)$. Clearly, this will be a bound on the queue length at network element m, and hence a buffer of that size would suffice at network element m. It may appear that another sufficient buffer allocation may be obtained by bounding the burstiness of the departures from the upstream element using Corollary 3.4 ($B^{D_{m-1}} := B^D(B, \mathsf{T}(S_1, \cdots, S_{m-1}))$) and then computing the queue length bound for element m ($\Sigma(B^{D_{m-1}}, S_m)$) fed by an arrival process with this burstiness constraint. This allocation may also seen to be necessary as we may construct a sample path where this queue length bound is attained. However, it turns out that these two bounds are equal. In a similar vein, the burstiness constraint on the depature process from element m, obtained from the composition rule ($B^D(B, S)$) turns out to be the same as that obtained by doing an element by element analysis ($B^D(B^{D_{m-1}}, S_m)$). We establish these two results in the following proposition.

Proposition 3.10.

$$\begin{aligned} \Sigma(B^D(B, \mathsf{T}(S_1, \cdots, S_{m-1})), S_m) &= \Sigma(B, S) \\ B^D(B^D(B, \mathsf{T}(S_1, \cdots, S_{m-1})), S_m) &= B^D(B, S) \end{aligned}$$

Proof. For any $u \geq 0$,

$$\begin{aligned} \sup_{s \geq 0}\{B(s+u) - S(s)\} &= \sup_{s \geq 0}\{B(s+u) - \inf_{0 \leq r \leq s}(\mathsf{T}(S_1, \cdots, S_{m-1})(r) + S_m(s-r))\} \\ &= \sup_{s \geq 0}\{\sup_{r \geq 0}\{B(r+s+u) - \mathsf{T}(S_1, \cdots, S_{m-1})(r) - S_m(s)\}\} \\ &= \sup_{s \geq 0}\{B^D(B, \mathsf{T}(S_1, \cdots, S_{m-1}))(s+u) - S_m(s)\}\}. \end{aligned}$$

This proves both the above equalities. □

It is clear that the above queue size bound gets larger as more network elements are traversed, requiring ever larger buffer allocations along the path. For instance, in the case of a (σ, ρ) traffic feeding a tandem of m identical latency-rate (θ, r) schedulers (see Example 3.5), the buffering requirement at the last element is $\sigma + m\rho\theta$ which grows linearly in m. In general, we cannot do with any smaller buffers at downstream elements. In certain cases however, more careful accounting combined with the introduction of regulators in the network, lead to smaller buffering requirements.

A simple instance of such reduction occurs when a network element factors in the transmission rate of the input link from the upstream element. In this case, the link may be treated as a regulator with constant slope, and the burstiness constraint of departures from the upstream node (obtained in Corollary 3.4) may be further composed with the burstiness constraint of the regulator to obtain a smoother arrival process. This inclusion of the peak rate of arrivals decreases the buffer allocation required by the connection. There are a number of intances where regulators are usefully deployed within a network, either as part of the scheduling mechanism such as rate controlled service disciplines or to police the flow at intermediate points. The introduction of regulators within the network reduces buffering requirements at downstream nodes. While this could increase the delay of individual packets, and perhaps even the mean delay of the stream, it is important to note that properly chosen regulators do not increase the worst case perfomance bounds guaranteed by the network. In the following proposition, we obtain an *envelope* E, such that regulators with the burstiness constraint $B \geq E$ may be inserted at multiple points in the tandem without affecting the worst case end-to-end delay suffered by the session.

PROPOSITION 3.11. *Let S be some service curve. The envelope E of S, denoted by $\mathcal{E}(S)$ is defined as*

$$E(t) := \sup_{s \geq 0}\{S(t+s) - S(s)\}, \quad t \geq 0. \tag{3.6}$$

Then, E is a non-decreasing sub-additive function with $E(0) = 0$, i.e., a burstiness constraint function, and $\mathsf{T}(E, S) = S$. Moreover, if $S = B$ is itself a burstiness constraint then, $\mathcal{E}(B) = B$.

PROOF. First note that $E(t+u) = \sup_{s\geq 0}\{S(t+u+s) - S(s)\} \leq \sup_{s\geq 0}\{S(t+u+s) - S(t+u)\} + \sup_{s\geq 0}\{S(t+u) - S(s)\} \leq E(t) + E(u)$, which proves the sub-additivity of E. As E is clearly non-decreasing with $E(0) = 0$, E is a burstiness constraint function. Next, observe that $S(t-s) + E(s) \geq S(t-s) + S(t) - S(t-s) = S(t)$. From this it follows that $\mathsf{T}(E, S) = S$. The last part is obvious. □

PROPOSITION 3.12. *Define the envelope $E := \mathcal{E}(\mathsf{T}(B, S))$. We claim that*

$$\Theta(B, \mathsf{T}(E, S_1, E, S_2, \ldots, E, S_m)) = \Theta(B, \mathsf{T}(E, S)) \tag{3.7}$$

$$= \Theta(B, S) \tag{3.8}$$

i.e., introducing network elements with service curve E between adjacent elements in the tandem does not alter the end-to-end delay.

PROOF. We may use Corollary 3.6 to rearrange the order of network elements to obtain

$$\mathsf{T}(E, S_1, E, S_2, \ldots, E, S_m) = \mathsf{T}(\mathsf{T}(E, \ldots, E), \mathsf{T}(S_1, \ldots, S_m)).$$

Then, (3.7) follows from Propositon 3.11.

Now to show (3.8), first we show that $\Theta(B, \mathsf{T}(E, S)) = \Theta(B, \mathsf{T}(B, E, S))$. As $\mathsf{T}(B, E, S) \leq \mathsf{T}(E, S)$, it immediately follows that $\Theta(B, \mathsf{T}(B, E, S)) \geq \Theta(B, \mathsf{T}(E, S))$. To see the reverse inequality, note that $\Theta(B, B) = 0$. Hence,

$$\Theta(B, \mathsf{T}(E, S)) = \Theta(B, B) + \Theta(B, \mathsf{T}(E, S)) \geq \Theta(B, \mathsf{T}(B, E, S)). \tag{3.9}$$

Now,

$$\Theta(B, \mathsf{T}(B, E, S)) = \Theta(B, \mathsf{T}(E, \mathsf{T}(B, S))) = \Theta(B, \mathsf{T}(B, S)).$$

the last equality following form Proposition 3.11. Finally, we again apply Proposition 3.9 to show that $\Theta(B, S) = \Theta(B, \mathsf{T}(B, S))$, which proves the result. □

We may insert a regulator with any burstiness constraint larger than $\mathcal{E}(\mathsf{T}(B, S))$ and still not increase the worst-case end-to-end delays. In particular, it may be shown that $\mathcal{E}(\mathsf{T}(B, S))$ is smaller than B. Hence, reshaping the session to the burstiness constraint of the arrival process at intermediate nodes leaves worst case end-to-end delays unchanged.

Example. For a tandem of m identical latency-rate (θ, r) schedulers fed by a sigma-rho (σ, ρ) constrained arrival process considered in Example 3.5 (with $\rho \leq r$), the envelope $E := \mathcal{E}(\mathsf{T}(B, S))$ is given by $E(t) = \min\{rt, \sigma + \rho t\}, t \geq 0$. The use of regulators with this burstiness constraint give a buffering requirement of $E(2\theta) = \min\{2r\theta, \sigma + 2\rho\theta\}$ at element $m \geq 2$. If we used regulators with burstiness constraint $B(t) = \sigma + \rho t$ of the arrival process, then the the buffering requirement would be $\sigma + 2\rho\theta$. By the previous proposition, the use of any of either of these regulators leaves the end-to-end delay unaltered. Both these buffering requirements are smaller than the original buffering requirement of $\sigma + m\rho\theta$. More importantly they do not increase with the number of elements traversed by the session.

4. Adaptive Service Sessions with End-to-End Flow Control

4.1 Single Hop Sessions with End-to-End Flow Control

Consider an adaptive service session traversing a single network element (scheduler or regulator), with end-to-end window flow control. The network element guarantees a minimum level of service to the session, quantified through a service curve, and the flow control mechanism allows the session to use excess bandwidth whenever available. Moreover, the network provisions buffers to provide a lossless service to any session that is compliant with the flow control mechanism. To study the performance of this session, we have to incorporate the flow control mechanism into our earlier network model involving service curve schedulers and regulators. This we do by allowing the source to transmit only when it has the requisite *tokens* in its token buffer (See Figure 3). There are W tokens present in the token buffer initially, where W is the fixed window size of the adaptive session. Tokens are consumed by packets that gain access to the network and are replenished by packets departing from the network.

The reason for studying the above session topology is two-fold. First it represents the simplest model of an adaptive service session, and is thus a natural starting point for studying more complex adaptive service session topologies and flow control mechanisms. More interestingly, as we show in the subsequent subsections, it turns out that both unicast and multicast sessions with multiple hops and window flow control can be reduced to a single element cycle.

Let the arrival process to the closed cycle be denoted by A and the departure process by D. Then, the windowing and service curve constraints are captured by

$$A(t) \geq D(t) \geq \inf_{s \leq t}[S(t-s) + (A(s) \wedge (D(s) + W))], \qquad t \in \mathbb{R}. \tag{4.1}$$

The main problem with the above characterization of the departure process from the closed cycle is that it involves an implicit inequality. We would like to show that there exists a solution D to the above and obtain a tight lower bound on all such solutions D. To this end we first consider the second inequality in (4.1) with an equality instead, i.e.,

$$\check{D}(t) = \inf_{s \leq t}[S(t-s) + (A(s) \wedge (\check{D}(s) + W))], \qquad t \in \mathbb{R}. \tag{4.2}$$

Clearly, any solution to (4.2) also satisfies (4.1).

Theorem 4.1. *There exists a unique departure process $\check{D}$ satisfying* (4.2) *and for any departure process D satisfying* (4.1), *$D \geq \check{D}$. Further $\check{D}$ may be obtained by the method of successive approximations, i.e., $\check{D} = \lim_{d \to \infty} \check{D}^d = \inf_{d \geq 0} \check{D}^d$, where $\check{D}^0(t) = \infty, t \geq 0$, and*

$$\check{D}^{d+1}(t) = \inf_{s \leq t}[S(t-s) + (A(s) \wedge (\check{D}^d(s) + W))], \qquad t \in \mathbb{R}, d \geq 0. \tag{4.3}$$

Proof. See Appendix A. □

Next, we present the rule for composing the service curve scheduler and the end-to-end windowing mechanism into another service curve scheduler. We will then show that the single element with feedback may be replaced by the composite service curve scheduler without feedback.

Composition Rule for Single Hop Adaptive Service Sessions : Define the service curves $S_l, l \geq 1$ to be $S_1 = S$ and $S_l = S + W, l \geq 2$. Let $\mathsf{T}(S_1, \cdots, S_d)$ represent the service curve of a tandem of d schedulers with service curves $S_l, l = 1, \ldots, d$ as defined by (3.2). Define the composite service curve

$$\check{S} := \wedge_{d=1}^{\infty} \mathsf{T}(S_1, \cdots, S_d), \tag{4.4}$$

where $\wedge$ denotes a pointwise minimum.

Remark 4.2. We have used the term service curve to denote functions S_l which do not neccessarily satisfy $S_l(0) = 0$.

The expression $\mathsf{T}(S_1, \cdots, S_d)$ is obtained by composing service curves along the path from the source to destination which loops through the cycle $d-1$ times, thus traversing S once and $(S+W)$ $(d-1)$ times. $\check{S}$ is the pointwise minimum taken over all paths from the origin to destination of the composition of schedulers along that path. This form of the composite service curve also holds for other session topologies considered later in the section.

Corollary 4.3. *Any departure process D satisfying* (4.1) *also satisfies a service curve constraint* (2.2) *with service curve $\check{S}$. If the window size W is such that $\mathsf{T}(S, S+W) \geq S$, then $\check{S} = S$, i.e., the closed cycle behaves like an open session.*

Proof. Define the service curves

$$\check{S}^d = \wedge_{m=1}^{d} \mathsf{T}(S_1, \cdots, S_m), \quad d \geq 1.$$

The first part of the result would follow from the above theorem, if we can show that $\check{D}^d, d \geq 1$ defined in the successive approximations procedure in the above theorem is given by

$$\check{D}^d(t) = \inf_{s \leq t} [\check{S}^d(t-s) + A(s)], \qquad t \in \mathbb{R}. \tag{4.5}$$

This is easily verified for $d = 1$. Now assume it is true for some $d \geq 1$. Then for $t \in \mathbb{R}$,

$$
\begin{aligned}
\check{D}^{d+1}(t) &= \inf_{s \leq t}[S(t-s) + (A(s) \wedge (\check{D}^d(s) + W))] \\
&= \inf_{s \leq t}[S(t-s) + A(s)] \wedge \inf_{s \leq t}[S(t-s) + (\check{D}^d(s) + W)] \\
&= \inf_{s \leq t}[S(t-s) + A(s)] \wedge \inf_{s \leq t}[S(t-s) + (\inf_{u \leq s}[\check{S}^d(s-u) + A(u)] + W)] \\
&= \inf_{s \leq t}[S(t-s) + A(s)] \wedge \inf_{u \leq t}[\inf_{u \leq s \leq t}[S(t-s) + \check{S}^d(s-u) + W] + A(u)] \\
&= \inf_{s \leq t}[S(t-s) + A(s)] \wedge \inf_{u \leq t}[\mathsf{T}(\check{S}^d, S_{d+1})(t-u) + A(u)] \\
&= \inf_{s \leq t}[(S(t-s) \wedge \mathsf{T}(\check{S}^d, S_{d+1})(t-s)) + A(s)] \\
&= \inf_{s \leq t}[\check{S}^{d+1}(t-s) + A(s)].
\end{aligned}
$$

The second part follows by observing that $T(S, S+W) \geq S$ implies that $\mathsf{T}(S_1, \cdots, S_d) \geq S$ for all $d \geq 1$ and hence $\check{D}^d = S$ for all $d \geq 1$. □

4.2 Unicast Sessions with End-to-End Flow Control

Now consider a unicast session traversing a tandem of m network elements (schedulers and/or regulators) with end-to-end window flow control as shown in Figure 4. This generalizes the system studied in Section 4.1 where $m = 1$. As before let the window size be W. Note that the unicast session without feedback may be thought of as a special case with the window size $W = \infty$.

Let the arrival process to the closed tandem be denoted by A, the departure process by D, the arrival process into scheduler k by A_k and the departure process by D_k, $1 \leq k \leq m$. Then, the equations for a unicast session without feedback also hold for this case with the important difference that now $A_1 = A \wedge (D + W)$ instead of $A_1 = A$. Let $S = \mathsf{T}(S_1, \cdots, S_m)$ be obtained by the aggregation rule (3.2). As in the case of the unicast session without feedback, we can reduce the unicast session traversing a tandem of m network elements with window flow control into a cycle of one scheduler with the service curve S.

Proposition 4.4. *The departure process $D := D_m$ from the unicast session with window flow control satisfies the single element cycle constraint* (4.1) *with* $S =$

$\mathsf{T}(S_1, \cdots, S_m)$. *Moreover, if* $\check{D}$ *is an achievable lower bound on all solutions* D *of* (4.1), *then it is also an achievable lower bound on all departure processes* D_m *from the unicast session with window flow control.*

Example. Consider a unicast session traversing a tandem of m identical latency-rate (θ, r) schedulers with end-to-end window flow control. For this example, the service curve that collapses the tandem to a single element is as for the unicast session without feedback, i.e., it is a latency rate service curve with latency $m\theta$ and rate r (see Example 3.5). The service curve $\mathsf{T}(S_1, \cdots, S_d)$ defined in (4.4) are given by

$$\mathsf{T}(S_1, \cdots, S_d)(t) = (d-1)W + r(t - dm\theta)^+, \quad t \geq 0, d \geq 1.$$

Thus,

$$\check{S}(t) = \begin{cases} 0, & 0 \leq t < m\theta, \\ (d-1)((rm\theta) \wedge W) + (r(t - dm\theta)) \wedge W, & dm\theta \leq t < (d+1)m\theta, d \geq 1. \end{cases}$$

If $W \geq rm\theta$ then $T(S, S+W) \geq S$ and hence, by Corollary 4.3, $\check{S} = S$. Note that $rm\theta$ is the bandwidth-delay (or rate-latency) product of the tandem. Thus a window size that equals the bandwidth-delay product recovers the same (service curve) performance as that of a session without feedback. A larger window size is unnecessary. Note that without any burstiness constraint on the arrival process into the system, we require a buffer equal in size to the window at each element (as the entire window may arrive in one burst at any network element). Thus we need a buffer of size $rm\theta$ at each element to realize the open session service curve.

4.3 Multicast Sessions with End-to-End Flow Control

Just as the earlier unicast session with end-to-end window flow control could be reduced to a single element cycle, in this section we show that a similar reduction is possible in the case of a multi-cast session where the source transmission is modulated by acknowledgments from each of the destinations. As shown in Figure 5, data packets flow from the source to a number of destinations. Multiple copies of packets are forwarded at branch nodes, reducing the burden of sending multiple identical packets from the source to each of the destinations. Further, we assume that network element k assures a service curve S_k. Destinations acknowledge each packet received, and for simplicity of exposition we assume that the volume of acknowledgments is the same as the volume of data received. Acknowledgments from different destinations may (or may not) be merged on the return path. Data transmission at the source is constrained by the arrival of acknowledgments from all destinations, in that the source not transmit more than a certain window of unacknowledged packets to each destination.

The multicast flow with m network elements (schedulers and/or regulators) may be modeled as a fork-join queueing network with $m + 1$ nodes, where traffic flows from a source node, through a feed-forward forking network of intermediate nodes to a number of destinations, and thence through another feed-forward join network back to the source, where it is joined with an exogenous arrival process. Assume that the source node is numbered 0, and other nodes from 1 to m in some order. We ignore the destinations in numbering the nodes as their arrivals equal departures. Let $\mathcal{F}$ denote the set of final nodes in the reverse path encountered by acknowledgments just before they reach the data source. Thus, we have a feed-forward fork-join network connecting the source node 0 to the set $\mathcal{F}$ of final nodes. For $1 \leq k \leq m$, let $\mathcal{P}(k)$ denote the set of nodes immediately preceding node k. Now, we may write the dynamics of each node as follows.

The departure process D_0 from the source is constrained by the arrival of acknowledgments from the final nodes $\{D_k, k \in \mathcal{F}\}$, the corresponding window sizes $\{W_k, k \in \mathcal{F}\}$, and the exogenous data generation process A. More precisely,

$$\begin{aligned} D_0(t) &= A(t) \wedge \min_{k \in \mathcal{F}} [D_k(t) + W_k] \\ &= A(t) \wedge [D(t) + W], \quad t \in \mathbb{R}, \end{aligned} \tag{4.6}$$

where $W = \min_{k \in \mathcal{F}} W_k$ denotes the smallest of the window sizes, $\tilde{W}_k := W_k - W$ for $k \in \mathcal{F}$, and

$$D(t) := \min_{k \in \mathcal{F}} [D_k(t) + \tilde{W}_k], \quad t \in \mathbb{R}. \tag{4.7}$$

At each of the nodes in the feed-forward fork-join network, the departure process satisfies the service curve constraint

$$\min_{l\in\mathcal{P}(k)} D_l(t) \geq D_k(t) \geq \inf_{s\leq t}\{S_k(t-s) + (\min_{l\in\mathcal{P}(k)} D_l(s))\}, \quad t\in\mathbb{R}, k=1,\ldots,m. \quad (4.8)$$

Next, we show how the analysis of the multicast session may be reduced to that of a single element cycle.

Composition Rule for Multicast Session with Window Flow Control: Define a path from the source to any node $k=1,\ldots,m$, as a string $\pi=(k_0,k_1,\ldots,k_{d(\pi)})$, where $k_0=0$, $k_{d(\pi)}=k$, and for $l=1,\ldots,d(\pi)$, $k_{l-1}\in\mathcal{P}(k_l)$. Note that the length of this path is denoted by $d(\pi)$. Let $\Pi(k)$ denote the set of all paths from the source to k, and let $\Pi := \cup_{k\in\mathcal{F}}\Pi(k)$ denote the set of all paths from the source node to any of the final nodes. Let $k(\pi)$ be the final node for path $\pi\in\Pi$. Then the composite service curve S for the multicast session is defined by

$$\begin{aligned} S(t) &:= \min_{\pi\in\Pi}(\tilde{W}_{k(\pi)} + \mathsf{T}(S_{k_0},S_{k_1},\ldots,S_{k_{d(\pi)}})(t)) \\ &= \min_{k\in\mathcal{F}}\min_{\pi\in\Pi(k)}(\tilde{W}_k + \mathsf{T}(S_{k_0},S_{k_1},\ldots,S_{k_{d(\pi)}})(t)), \quad t\geq 0, \end{aligned}$$

where T is the composition rule for a tandem of service curves as defined in (3.2), and $S_0(t)=\infty$ for $t>0$ and $S_0(t)=0$ for $t=0$. See Figures 6 and 7 which illustrate this composition rule.

Proposition 4.5. *The departure process D satisfies the single element cycle constraint* (4.1) *with S given by the composition rule above. Moreover, if $\check{D}$ is an achievable lower bound on D satisfying* (4.1)*, then $\check{D}_0$ defined by* (4.6) *and $\check{D}_k, 1\leq k\leq m$, defined iteratively by the right hand side of* (4.8) *(with D_l replaced by $\check{D}_l$) is also an achievable lower bound on the departure processes $D_k, 0\leq k\leq m$, from the multicast session with window flow control.*

Proof. In order to prove the proposition we first show that $\{D_k, k=1,\ldots,m\}$ satisfy the following service curve constraint in terms of D_0.

$$D_0(t)\geq D_k(t)\geq \inf_{s\leq t}\{S^k(t-s)+D_0(s)\}, \quad t\in\mathbb{R}, 1\leq k\leq m. \quad (4.9)$$

where

$$S^k := \min_{\pi\in\Pi(k)} \mathsf{T}(S_{k_0},S_{k_1},\ldots,S_{k_{d(\pi)}})$$

Let $d(k)$ be the length of the longest path in $\Pi(k)$. Then, node k is said to be at depth $d(k)$ from the source. The proof of (4.9) proceeds by an induction on the depth d of nodes. Let V_d be all the nodes at depth $d, d = 0, 1, \ldots$. Note $V_0 = \{0\}$ and thus the result is trivially true for $k \in V_0$. Assume it is true for all $k \in \cup_{c=0}^{d} V_c$ for some $d \geq 0$. We shall show that the result continues to hold for $k \in V_{d+1}$. Note that $\mathcal{P}(k) \subset \cup_{c=0}^{d} V_c$ (otherwise k would have a depth greater than $d+1$). Thus, by (4.8) and the induction hypothesis,

$$\begin{aligned} D_0(t) \geq D_k(t) &\geq \inf_{r \leq t}\{S_k(t-r) + \min_{l \in \mathcal{P}(k)} \inf_{s \leq r}\{S^l(r-s) + D_0(s)\}\} \\ &= \inf_{s \leq t}\{\min_{l \in \mathcal{P}(k)} \inf_{s \leq r \leq t}\{S_k(t-r) + S^l(r-s)\} + D_0(s)\} \\ &= \inf_{s \leq t}\{S^k(t-s) + D_0(s)\}, \quad t \in \mathbb{R}. \end{aligned}$$

Substituting by the above inequalities in (4.7) and using (4.6), we get

$$\begin{aligned} A(t) \geq D_0(t) \geq D(t) &\geq \min_{k \in \mathcal{F}}[\inf_{s \leq t}\{S^k(t-s) + D_0(s)\} + \tilde{W}_k] \\ &= \inf_{s \leq t}\{\min_{k \in \mathcal{F}}[S^k(t-s) + \tilde{W}_k] + D_0(s)\} \\ &= \inf_{s \leq t}\{S(t-s) + D_0(s)\} \\ &= \inf_{s \leq t}\{S(t-s) + A(s) \wedge (D(s) + W)\}, \quad t \in \mathbb{R}. \end{aligned}$$

Thus, D satisfies (4.1). The fact that an achievable lower bound on the solutions to (4.1) yield an achievable lower bound on the departure processes from the multicast session, as described in the proposition, may be verified easily. □

5. Sessions with Hop-by-Hop Flow Control

In this section, we consider unicast and multicast sessions with hop-by-hop flow control. We first present the unicast case to elucidate the problem and notation, and then proceed to obtain results for the general multicast session topology.

Consider the unicast connection studied in Section 4.2 with the difference that instead of end-to-end window flow control, there are hop-by-hop windows that constrain transmission across adjacent network elements. Under this flow control mechanism each network element k, $1 \leq k \leq m-1$, requires a token from the downstream network element $k+1$ for a packet to be eligible for transmission. Each of these elements, k, has a certain number $u_{k+1,k}$ of tokens initally; tokens are consumed by packets arriving into that element and are replenished by packets departing from the immediately downstream element $k+1$. This ensures that there are never more than $u_{k+1,k}$ packets in the immediately downstream element $k+1$. The final element m delivers packets to the destination and thus does not require any tokens. If we define the *effective arrival process* $A_k(t)$ into node k to be the total number of eligible packets received thus far, i.e., packets for which tokens have been received, then

$$\begin{aligned} A_1(t) &:= \min\{A(t), D_2(t) + u_{2,1}\}, \\ A_k(t) &:= \min\{D_{k-1}(t), D_{k+1}(t) + u_{k+1,k}\}, \quad 1 \leq k \leq m-1, \\ A_m(t) &:= D_{m-1}(t), \quad t \in \mathbb{R}. \end{aligned}$$

The departure process from element k satisfies the service curve constraint with service curve S_k and arrival process A_k, i.e.,

$$A_k(t) \geq D_k(t) \geq \inf_{s \leq t}\{S_k(t-s) + A_k(s)\}, \quad t \in \mathbb{R}, 1 \leq k \leq m.$$

Now, in a multicast adaptive session with hop-by-hop windows departures from a particular network element are constrained by the departures of immediately upstream elements, as well as by departures from immediately downstream ones. Every time a departure from an immediately downstream element occurs, it sends a token to this element. The upstream element is constrained by tokens it has outstanding from each of its immediately downstream elements, and the number of outstanding tokens is diminished on the transmission of data from the upstream element. Consequently, the upstream element can transmit as many packets as the minimum of all the outstanding tokens it has. Clearly, a network element is also constrained by the number of packets it has received. Hence, we may erase the distinction between tokens and packets, defining the *effective arrival process* into a network element as the minimum of all packets and tokens received by the element thus far.

Both the unicast and multicast sessions considered above, as well as all the session topologies in the previous sections, are special cases of a more general class of networks called *fork-join* networks, described below. Let D_k denote the departure process from element k of a network comprising m elements. The departure process D_k from network element k is "forked" into m copies, the copy for element j entering buffer U_{kj}. Let these buffers have an initial content of u_{kj} which may be infinite. Note that the buffer may contain either packets or tokens. Departures from all buffers $U_{jk}, 1 \leq j \leq m$ are "joined" together with an *exogenous arrival process* X_k to form the *effective arrival process* A_k into element k. Thus,

$$A_k(t) := X_k(t) \wedge \min_{1 \leq j \leq m} \{D_j(t) + u_{jk}\} =: M_k(\mathbf{X}(t), \mathbf{D}(t)), \quad t \in \mathbb{R}, \tag{5.1}$$

where $\mathbf{D} := (D_1, \ldots, D_m)$ and $\mathbf{X} := (X_1, \ldots, X_m)$, and $\mathbf{X}(t) = \mathbf{0}$, $t < 0$. Note that if any of the u_{jk} are infinite, then the corresponding departure process D_j does not directly constrain the effective arrival process A_k. The departure process from element k satisfies the service curve constraint with service curve S_k and arrival process A_k, i.e.,

$$M_k(\mathbf{X}(t), \mathbf{D}(t)) \geq D_k(t) \geq \inf_{s \leq t} \{S_k(t-s) + M_k(\mathbf{X}(s), \mathbf{D}(s))\}, \quad t \in \mathbb{R}. \tag{5.2}$$

We need to show that there exists at least one solution $\mathbf{D}$ to the above constraints and obtain a service curve lower bound on any such $\mathbf{D}$. To do this, we need to assume that the fork-join network is *deadlock free*, i.e., given any cycle of network elements $k_1, \ldots, k_d, k_1$, for some $d \geq 1$, we have $\sum_{j=1}^{d-1} u_{k_j k_{j+1}} + u_{k_d k_1} > 0$. In other words, at one of the elements $\{u_{k_1 k_2}, \ldots, u_{k_{d-1} k_d}, u_{k_d k_1}\}$ should be positive.

Now, consider replacing (5.2) by

$$D_k(t) = \inf_{s \leq t} \{S_k(t-s) + M_k(\mathbf{X}(s), \mathbf{D}(s))\}, \quad t \in \mathbb{R}. \tag{5.3}$$

Clearly, any solution to (5.3) is also a solution to (5.2). Below, we shall show that there is a unique solution to (5.3), that this may be obtained by the method of successive approximations (this gives an explicit solution), and that it is in fact a lower bound on any $\mathbf{D}$ satisfying (5.2).

Theorem 5.1. *There exists a unique departure process $\check{\mathbf{D}}$ satisfying* (5.3), *and for any departure process $\mathbf{D}$ satisfying* (5.2), $\mathbf{D} \geq \check{\mathbf{D}}$. *Further $\check{\mathbf{D}}$ may be obtained by* *the method of successive approximations, i.e.,* $\check{\mathbf{D}} = \lim_{d\to\infty} \check{\mathbf{D}}^d = \inf_{d\geq 0} \check{\mathbf{D}}^d$, *where* $\check{\mathbf{D}}^0(t) = \infty, t \geq 0$, *and*

$$\check{D}_k^{d+1}(t) = \inf_{s\leq t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}^d(s))], \qquad t \in \mathbb{R}, 1 \leq k \leq m, d \geq 0.$$

PROOF. See Appendix B. □

Composition Rule for Fork-Join Networks: Based on the successive approximations sequence defined in the above theorem, we obtain a service curve lower bound on any $\mathbf{D}$ satisfying (5.2). Let $\Pi(j,k;d)$ be the set of all paths from the exogenous arrival X_j to the departure D_k of length less than or equal to d. Thus $\pi \in \Pi(j,k;d)$ implies that $\pi = (k_1, \ldots k_{d'})$ for some $1 \leq d' \leq d$ and $k_1 = j$, $k_{d'} = k$, $1 \leq k_2, \ldots, k_{d'-1} \leq m$. Similarly, define $\Pi(j,k)$ to be the set of all paths from the exogenous arrival X_j to the departure D_k of any finite but arbitrary length. Define

$$\begin{aligned}
R_{jk} &:= u_{jk} + S_k, \quad 1 \leq j,k \leq m \\
S_{jk}^d &:= \min_{\pi=(k_1,\ldots,k_{d'})\in\Pi(j,k;d)} \mathsf{T}(S_{k_1}, R_{k_1k_2}, \cdots, R_{k_{d'-1}k}), \quad 1 \leq j,k \leq m,\ d \geq 1, \\
S_{jk}^\infty &:= \min_{\pi=(k_1,\ldots,k_{d'})\in\Pi(j,k)} \mathsf{T}(S_{k_1}, R_{k_1k_2}, \cdots, R_{k_{d'-1}k}), \quad 1 \leq j,k \leq m,
\end{aligned}$$

where the min over an empty set is understood to be ∞. S_{jk}^∞ will be seen to be the service curve between the exogenous arrival X_j and the departure D_k.

COROLLARY 5.2. $\check{\mathbf{D}}^d, d \geq 1$ *defined as the successive approximations sequence in the above theorem is given by*

$$\check{D}_k^d(t) = \min_{1\leq j\leq m} \inf_{s\leq t}\{S_{jk}^d(t-s) + X_j(s)\}, \quad t \in \mathbb{R},$$

and consequently, the fixed point is given by

$$\check{D}_k(t) = \min_{1\leq j\leq m} \inf_{s\leq t}\{S_{jk}^\infty(t-s) + X_j(s)\}, \quad t \in \mathbb{R}.$$

PROOF. The proof proceeds by an induction on $d \geq 1$. It is easily verified for $d = 1$. Now assume it is true for some $d \geq 1$. Then for $t \in \mathbb{R}$,

$$\check{D}_k^{d+1}(t) \quad = \quad \inf_{s\leq t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}^d(s))]$$

$$
\begin{aligned}
&= \inf_{s\le t}[S_k(t-s)+X_k(s)] \wedge \inf_{s\le t}[S_k(t-s)+\min_{1\le j\le m}(\check{D}_j^d(s)+u_{jk})] \\
&= \inf_{s\le t}[S_k(t-s)+X_k(s)] \\
&\quad \wedge \inf_{s\le t}[S_k(t-s)+\min_{1\le j\le m}((\min_{1\le i\le m}\inf_{r\le s}\{S_{ij}^d(s-r)+X_i(r)\})+u_{jk})] \\
&= \inf_{s\le t}[S_k(t-s)+X_k(s)] \\
&\quad \wedge \min_{1\le i\le m}\inf_{r\le t}[\min_{1\le j\le m}\inf_{r\le s\le t}(S_{ij}^d(s-r)+(u_{jk}+S_k(t-s))+X_i(r)] \\
&= \inf_{s\le t}[S_k(t-s)+X_k(s)] \\
&\quad \wedge \min_{1\le i\le m}\inf_{s\le t}[\min_{1\le j\le m}\mathsf{T}(S_{ij}^d,R_{jk})(t-s)+X_i(s)] \\
&= \min_{1\le i\le m}\inf_{s\le t}[S_{ik}^{d+1}(t-s)+X_i(s)].
\end{aligned}
$$

□

Corollary 5.3 (Composition Rule for Unicast Session with Hop-by-hop Windows).

Let $\Pi=\Pi(1,m)$ be the set of all paths from the source to the destination. Without loss of generality we may restrict attention to paths $\pi=(k_1,\ldots,k_d)$ with $k_1=1, k_d=m, 1\le k_2,\ldots,k_{d-1}\le m$ with $|k_{j+1}-k_j|=1$, $1\le j\le d-1$. Let

$$\check{S}=S_{1m}^{\infty}=\min_{\pi\in\Pi}\mathsf{T}(S_{k_1},S_{k_2}+u_{k_1,k_2},\ldots,S_{k_d}+u_{k_{d-1},k_d}),$$

where we set $u_{k-1,k}=0$, $1\le k\le m$. Then, the departure process D of the unicast session with hop-by-hop flow control satisfies the service curve constraint (2.2) *with arrival process A, and service curve $\check{S}$.*

Example. Consider a unicast connection traversing a tandem of m identical latency-rate (θ,r) schedulers with hop-by-hop (instead of the end-to-end in Section 4.2) window flow control. Assume all the windows are of the same size W. For this example the paths $\pi\in\Pi$ traverse exactly $m+2l$ elements and l windows for some $l\ge 0$. Thus, for any such path

$$\mathsf{T}(S_{k_1},S_{k_2}+u_{k_1,k_2},\ldots,S_{k_d}+u_{k_{d-1},k_d})(t)=lW+r(t-(m+2l)\theta)^+,\quad t\ge 0, l\ge 0.$$

Thus, the composite service curve for this session is given by

$$
\begin{aligned}
\check{S}(t) &= \min_{l\ge 0}(lW+r(t-(m+2l)\theta)^+),\quad t\ge 0 \\
&= \begin{cases} 0, & 0\le t<m\theta, \\ l((r2\theta)\wedge W)+(r(t-(m+2l)\theta))\wedge W, & (m+2l)\theta\le t<(m+2(l+1))\theta, l\ge 0. \end{cases}
\end{aligned}
$$

In this case if $W \geq r2\theta$ then $\check{S} = S$, where S is the service curve of the open session. Since the buffering requirement at each element is the window size W, we are now able to realize the open session service curve with a buffer of size 2θ compared with $m\theta$ for the end-to-end window flow control scheme. This reduction in the buffereing requirement is precisely the justification for using the more complex hop-by-hop flow control mechanism.

6. Packetization

In this section we present a generalization of the service curve constraint (2.2) introduced in Section 2.2 which allows us to explicitly take into account the effect of packetization. This modelling feature makes it possible to obtain tighter worst case delay bounds.

Consider a packetized arrival stream with successive packet sizes denoted by $L_k, k \geq 1$. For this arrival process define the packetization function P as follows:

$$P(x) = \sum_{j=1}^{k} L_j, \quad \text{if } \sum_{j=1}^{k} L_j \leq x < \sum_{j=1}^{k+1} L_j, k \geq 0.$$

If $X(t)$ represents the number of bits that have arrived till time t then $P(X(t))$ is precisely the number of bits that correspond to packets that have arrived completely till this time. When considering a store-and-forward mechanism at each network element, we must consider arrivals to count only complete packets. Thus, if A denotes such a packetized arrival process, then we have $P(A(t)) = A(t), \ \forall t \in \mathbb{R}$, i.e., P preserve the arrival process in that it does not pack A into larger packets. (For this reason the packetization function may be different for different paths of the arrival process.) Also note that $P : \mathbb{R}_+ \to \mathbb{R}_+$ defined above is non-decreasing with $P(x) \leq x, \ \forall x \in \mathbb{R}_+$. More generally, we consider P to be any function satisfying these two properties. Note that $P = I$, the identity function, also satisfies these conditions. We now generalize the previous definition of a service curve scheduler to:

$$A(t) \geq D(t) \geq P(\inf_{s \leq t}[S(t-s) + A(s)]), \qquad t \in \mathbb{R}, . \tag{6.1}$$

As before, we shall assume that $S(0) = 0$ and S in non-decreasing. By choosing $P = I$, we obtain the basic formulation introduced earlier. For other choices of P

we can still obtain the basic formulation, albeit with a smaller service curve. In particular, define the maximum length of a packet to be $L_{\max} := \sup_{x \in \mathbb{R}_+}(x - P(x))$. Then

$$\begin{aligned} P(\inf_{s \le t}[S(t-s) + A(s)]) &= \inf_{s \le t} P(S(t-s) + A(s)) \\ &\ge \inf_{s \le t}[(S(t-s) - L_{\max})^+ + A(s)] \\ &= \inf_{s \le t}[(S'(t-s) + A(s)] \end{aligned}$$

where $S'(u) := (S(u) - L_{\max})^+, u \in \mathbb{R}_+$. Note that the inequality above follows from the observation that for $x \ge 0$, $P(x + A) \ge P(A) = A$ and $P(x + A) \ge x + A - L_{\max}$. Thus, (6.1) implies (2.2) with a smaller service curve. When obtaining service curve characterizations for specific schedulers, it may be possible to get a larger service curve in the above formulation (6.1) than with the original formulation (2.2). This difference can be exploited to obtain tighter delay bounds.

Let $\Delta(t)$ be the delay experience by any bit arriving at time t as defined in (2.4). Note that this delay is the largest over all D satisfying (6.1), when D equals the right hand side of (6.1). Let $\check{D}$ and $\hat{\Delta}$ denote this worst case departure and delay.

Proposition 6.1 (Upper Bound on Delay from a Packetized Scheduler). *The worst case delay $\hat{\Delta}(t)$ does not depend on the packetization function P. In particular, the bound $\Theta(B, S)$ obtained in (2.5) with $P = I$ is valid for any other packetization function.*

Proof.

$$\begin{aligned} \hat{\Delta}(t) &= \inf\{\delta \ge 0 : \check{D}(t + \delta) \ge A(t)\} \\ &= \inf\{\delta \ge 0 : P(\inf_{s \le t+\delta}[S(t + \delta - s) + A(s)]) \ge A(t)\} \\ &= \inf\{\delta \ge 0 : \inf_{s \le t+\delta}[S(t + \delta - s) + A(s)] \ge A(t)\} \end{aligned}$$

where the last equality follows from the observation that $x \ge A$ iff $P(x) \ge P(A) = A$, which in turn follows from the condition that P is non-decreasing, and $x < A$ implies that $P(x) \le x < A = P(A)$. □

Example. Consider the example of a constant rate server with rate r. Then with the use of the packetization function defined in terms of the packet lengths, we may easily see that the departing packetized bit stream D to satisfy the constraint (6.1) with $S(t) = rt$. However, if we wish D to satisfy (2.2), then the largest possible choice of a service curve that works is $S'(t) := (S(t) - L_{\max})^+ = r(t - L_{\max}/r)^+, t \in \mathbb{R}_+$. This latter characterization introduces an additional latency of $\theta = L_{\max}/r$. This basic idea can be used to get a tighter characterization of a number of schedulers.

Appendix A – Proof of Theorem 4.1

Let $\{D^\alpha, \alpha \in \mathcal{A}\}$ be the set of all departure processes (i.e., D^α non-decreasing, $D^\alpha(t) = 0, t < 0$) satisfying the inequality

$$D^\alpha(t) \geq \inf_{s \leq t}[S(t-s) + (A(s) \wedge (D^\alpha(s) + W))], \qquad t \in \mathbb{R}. \tag{6.2}$$

Clearly, the set is non-empty as $D(t) = \infty, t \geq 0$ satisfies (6.2). Next, it is easy to verify that $\check{D} := \inf_{\alpha \in \mathcal{A}} D^\alpha$ also satisfies (6.2). We claim that $\check{D}$ satisfies (4.2). Assume not. Define $\check{D}'$ to be

$$\check{D}'(t) := \inf_{s \leq t}[S(t-s) + (A(s) \wedge (\check{D}(s) + W))], \qquad t \in \mathbb{R}. \tag{6.3}$$

First, note that $\check{D}'$ is non-decreasing as $A' := A \wedge (\check{D} + W)$ is non-decreasing and hence for any $u > t$,

$$\begin{aligned}
\check{D}'(u) &= \inf_{s \leq u}[S(u-s) + A'(s)] \\
&= \inf_{s \leq t}[S(u-s) + A'(s)] \wedge \inf_{t < s \leq u}[S(u-s) + A'(s)] \\
&\geq \inf_{s \leq t}[S(t-s) + A'(s)] \wedge A'(t) \\
&= \inf_{s \leq t}[S(t-s) + A'(s)] \\
&= \check{D}'(t).
\end{aligned}$$

Moreover, since $\check{D}'(t) = 0$ for any $t < 0$, $\check{D}'$ is a departure process. Also, by construction $\check{D}' \leq \check{D}$ and for some $t \in \mathbb{R}$, $\check{D}'(t) < \check{D}(t)$. Thus,

$$\check{D}'(t) \geq \inf_{s \leq t}[S(t-s) + (A(s) \wedge (\check{D}'(s) + W))], \qquad t \in \mathbb{R}.$$

Thus, $\check{D}'$ also satisfies (6.2), and hence contradicts the minimality of $\check{D}$ as a solution to (6.2). This contradiction establishes that $\check{D}$ satisfies (4.2).

Next, we establish that there is a unique non-decreasing solution to (4.2). Assume not. Let $\check{D}$ and $\check{D}'$ be two different departure processes satisfying (4.2). Let

$$t_0 := \inf\{t \in \mathbb{R} : \check{D}(t) \neq \check{D}'(t)\}.$$

Note that $t_0 \geq 0$ as $\check{D}(t) = \check{D}'(t)$ for all $t < 0$. We now show that $\check{D}(t_0) = \check{D}'(t_0)$ as

$$\begin{aligned}
\check{D}(t_0) &= \inf_{s \leq t_0} [S(t_0 - s) + (A(s) \wedge (\check{D}(s) + W))] \\
&= \inf_{s < t_0} [S(t_0 - s) + (A(s) \wedge (\check{D}(s) + W))] \wedge [S(0) + A(t_0)] \wedge [S(0) + \check{D}(t_0) + W] \\
&= \inf_{s < t_0} [S(t_0 - s) + (A(s) \wedge (\check{D}(s) + W))] \wedge [S(0) + A(t_0)] \\
&= \inf_{s < t_0} [S(t_0 - s) + (A(s) \wedge (\check{D}'(s) + W))] \wedge [S(0) + A(t_0)] \\
&= \check{D}'(t_0).
\end{aligned}$$

The third equality above holds because $[S(0)+\check{D}(t_0)+W] > \check{D}(t_0)$, the fourth because $\check{D}(s) = \check{D}'(s)$ for $s < t_0$, and the last by interchanging $\check{D}$ and $\check{D}'$ in the preceeding steps. Now let $\check{D}(t_0^+) := \lim_{t \searrow t_0} \check{D}(t)$ and $\check{D}'(t_0^+) := \lim_{t \searrow t_0} \check{D}'(t)$. Then, there exist a $\delta > 0$ such that for all $t \in (t_0, t_0 + \delta)$, $\check{D}(t) < \check{D}(t_0^+) + W$ and $\check{D}'(t) < \check{D}'(t_0^+) + W$. Thus, for $t \in (t_0, t_0 + \delta)$,

$$\begin{aligned}
\check{D}(t) &= \inf_{s \leq t} [S(t - s) + (A(s) \wedge (\check{D}(s) + W))] \\
&= \inf_{s \leq t} [S(t - s) + A(s)] \wedge \inf_{s \leq t_0} [S(t - s) + \check{D}(s) + W] \wedge \inf_{t_0 < s \leq t} [S(t - s) + \check{D}(s) + W] \\
&= \inf_{s \leq t} [S(t - s) + A(s)] \wedge \inf_{s \leq t_0} [S(t - s) + \check{D}(s) + W]
\end{aligned}$$

as $\inf_{t_0 < s \leq t} [S(t - s) + \check{D}(s) + W] \geq [S(0) + \check{D}(t_0^+) + W] > \check{D}(t)$. Since $\check{D}(s) = \check{D}'(s)$ for $s \leq t_0$, we get that $\check{D}(t) = \check{D}'(t)$ for $t \in (t_0, t_0 + \delta)$. This contradicts the definition of t_0 and the uniqueness is established as a consequence.

Finally, we show that $\check{D}^d, d \geq 0$ defined in the theorem converge to this unique solution. It is easy to establish by induction that $\check{D}^d, d \geq 0$ is a non-increasing sequence and hence has a limit $\check{D}^\infty := \lim_{d \to \infty} \check{D}^d = \inf_{d \geq 0} \check{D}^d$. Thus,

$$\begin{aligned}
\check{D}^\infty(t) &= \inf_{d \geq 0} \check{D}^d(t) = \inf_{d \geq 0} \check{D}^{d+1}(t) \\
&= \inf_{d \geq 0} \inf_{s \leq t} [S(t - s) + (A(s) \wedge (\check{D}^d(s) + W))] \\
&= \inf_{s \leq t} [S(t - s) + (A(s) \wedge (\inf_{d \geq 0} \check{D}^d(s) + W))] \\
&= \inf_{s \leq t} [S(t - s) + (A(s) \wedge (\check{D}^\infty(s) + W))].
\end{aligned}$$

□

Appendix B – Proof of Theorem 5.1

Let $\{\mathbf{D}^\alpha, \alpha \in \mathcal{A}\}$ be the set of all departure processes (i.e., $\mathbf{D}^\alpha$ non-decreasing, $\mathbf{D}^\alpha(t) = \mathbf{0}, t < 0$) satisfying the right inequality in (5.2). Clearly, the set is non-empty as $\mathbf{D}(t) = \infty, t \geq 0$, satisfies this inequality. It is easy to verify that $\check{\mathbf{D}} := \inf_{\alpha\in\mathcal{A}} \mathbf{D}^\alpha$ also satisfies this inequality. We claim that $\check{\mathbf{D}}$ satisfies (5.3). Assume not. Define $\check{\mathbf{D}}'$ to be

$$\check{D}'_k(t) := \inf_{s\leq t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}(s))], \qquad t \in \mathbb{R}.$$

First, note that $\check{\mathbf{D}}'(t) = \mathbf{0}$ for $t < 0$ and that $\check{\mathbf{D}}'$ is non-decreasing as shown in the proof of Theorem 4.1 and hence is a departure process. Also, by construction $\check{\mathbf{D}}' \leq \check{\mathbf{D}}$ and for some $t \in \mathbb{R}$, $\check{\mathbf{D}}'(t) < \check{\mathbf{D}}(t)$. Thus,

$$\check{D}'_k(t) \geq \inf_{s\leq t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}'(s))], \qquad t \in \mathbb{R}.$$

Hence $\check{\mathbf{D}}'$ also satisfies the right inequality in (5.2), and therefore contradicts the minimality of $\check{\mathbf{D}}$ as a solution to that inequality. This contradiction establishes that $\check{\mathbf{D}}$ satisfies (5.3).

Next, we establish that there is a unique non-decreasing solution to (5.3). Assume not. Let $\check{\mathbf{D}}$ and $\check{\mathbf{D}}'$ be two different departure processes satisfying (5.3). Since $\check{\mathbf{D}} \wedge \check{\mathbf{D}}'$ is also a departure process satisfying (5.3), we may assume $\check{\mathbf{D}} \leq \check{\mathbf{D}}'$ without loss of generality. Let

$$t_0 := \inf\{t \in \mathbb{R} : \check{\mathbf{D}}(t) \neq \check{\mathbf{D}}'(t)\}.$$

It is easy to see that $t_0 \geq 0$. Next, note that $\check{\mathbf{D}}(t_0) = \check{\mathbf{D}}'(t_0)$. We prove this by contradiction. Let $\mathcal{B}_0 := \{1 \leq j \leq m : \check{D}_j(t_0) \neq \check{D}'_j(t_0)\}$ be the set of departure processes which differ in the two at time t_0. If $\check{\mathbf{D}}(t_0) \neq \check{\mathbf{D}}'(t_0)$, then $\mathcal{B}_0$ is non-empty. Then, by the deadlock-free assumption, there exists a $k \in \mathcal{B}_0$ which, under $\check{\mathbf{D}}(t_0)$, is not constrained by any $j \in \mathcal{B}_0$, i.e., $\check{D}_k(t_0) < \min_{j\in\mathcal{B}_0}(\check{D}_j(t_0) + u_{jk})$. (In order to see why this follows from the deadlock-free condition, assume this conclusion is false. Then it is not hard to see that there exists a cycle of network elements

$k_1, \ldots, k_d, k_1 \in \mathcal{B}_0$, for some $d \geq 1$, such that k_j is constrained by k_{j+1}, $j = 1, \ldots, d-1$ and k_d is constrained by k_1, i.e.,

$$\begin{aligned} D_{k_j} &= D_{k_{j+1}} + u_{k_j k_{j+1}}, \quad j = 1, \ldots, d-1, \\ D_{k_d} &= D_{k_1} + u_{k_d k_1}. \end{aligned}$$

Thus,

$$D_{k_1} = D_{k_1} + \sum_{j=1}^{d-1} u_{k_j k_{j+1}} + u_{k_d k_1},$$

which implies that $\sum_{j=1}^{d-1} u_{k_j k_{j+1}} + u_{k_d k_1} = 0$, thereby contradicting the deadlock free condition.) Thus,

$$\begin{aligned} \check{D}_k(t_0) &= \inf_{s \leq t_0} [S_k(t_0 - s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}(s))] \\ &= \inf_{s < t_0} [S_k(t_0 - s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}(s))] \wedge [S_k(0) + X_k(t_0)] \\ &\quad \wedge [S_k(0) + \min_{j \notin \mathcal{B}_0} (\check{D}_j(t_0) + u_{jk})] \wedge [S_k(0) + \min_{j \in \mathcal{B}_0} (\check{D}_j(t_0) + u_{jk})] \\ &= \inf_{s < t_0} [S_k(t_0 - s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}(s))] \wedge [S_k(0) + X_k(t_0)] \\ &\quad \wedge [S_k(0) + \min_{j \notin \mathcal{B}_0} (\check{D}_j(t_0) + u_{jk})] \\ &= \inf_{s < t_0} [S_k(t_0 - s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}'(s))] \wedge [S_k(0) + X_k(t_0)] \\ &\quad \wedge [S_k(0) + \min_{j \notin \mathcal{B}_0} (\check{D}'_j(t_0) + u_{jk})] \\ &\geq \inf_{s < t_0} [S_k(t_0 - s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}'(s))] \wedge [S_k(0) + X_k(t_0)] \\ &\quad \wedge [S_k(0) + \min_{j \notin \mathcal{B}_0} (\check{D}'_j(t_0) + u_{jk})] \wedge [S_k(0) + \min_{j \in \mathcal{B}_0} (\check{D}'_j(t_0) + u_{jk})] \\ &= \inf_{s \leq t_0} [S_k(t_0 - s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}'(s))] \\ &= \check{D}'_k(t_0) \geq \check{D}_k(t_0). \end{aligned}$$

Thus, $\check{D}_k(t_0) = \check{D}'_k(t_0)$, which is a contradiction.

Next, let $\mathcal{B}$ be the set of elements j such that $\inf\{t \in \mathbb{R} : \check{D}_j(t) \neq \check{D}'_j(t)\} = t_0$. By the definition of t_0, there exists a $\delta' > 0$, such that for all $t \in (t_0, t_0 + \delta')$ and $j \notin \mathcal{B}$, $\check{D}_j(t) = \check{D}'_j(t)$. Now let $\check{\mathbf{D}}(t_0^+) := \lim_{t \searrow t_0} \check{\mathbf{D}}(t)$ and $\check{\mathbf{D}}'(t_0^+) := \lim_{t \searrow t_0} \check{\mathbf{D}}'(t)$. As before, by the deadlock-free assumption, there exists a $k \in \mathcal{B}$ which under $\check{\mathbf{D}}(t_0^+)$ is not constrained by any $j \in \mathcal{B}$, i.e., $\check{D}_k(t_0^+) < \min_{j \in \mathcal{B}} (\check{D}_j(t_0^+) + u_{jk})$. Let $\epsilon := \min_{j \in \mathcal{B}} (\check{D}_j(t_0^+) + u_{jk}) - \check{D}_k(t_0^+) > 0$. Let $\delta > 0$ be such that for all $t \in (t_0, t_0 + \delta)$,

$\check{D}_k(t) < \check{D}_k(t_0^+) + \epsilon$. Thus, for $t \in (t_0, t_0 + (\delta' \wedge \delta))$,

$$
\begin{aligned}
\check{D}_k(t) &= \inf_{s \le t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}(s))] \\
&= \inf_{s \le t}[S_k(t-s) + X_k(s)] \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \notin \mathcal{B}}(\check{D}_j(s) + u_{jk})] \\
&\quad \wedge \inf_{s \le t_0}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}_j(s) + u_{jk})] \wedge \inf_{t_0 < s \le t}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}_j(s) + u_{jk})] \\
&\ge \inf_{s \le t}[S_k(t-s) + X_k(s)] \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \notin \mathcal{B}}(\check{D}_j(s) + u_{jk})] \\
&\quad \wedge \inf_{s \le t_0}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}_j(s) + u_{jk})] \wedge [S_k(0) + \min_{j \in \mathcal{B}}(\check{D}_j(t_0^+) + u_{jk})] \\
&\ge \inf_{s \le t}[S_k(t-s) + X_k(s)] \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \notin \mathcal{B}}(\check{D}_j(s) + u_{jk})] \\
&\quad \wedge \inf_{s \le t_0}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}_j(s) + u_{jk})] \wedge (\check{D}_k(t_0^+) + \epsilon) \\
&= \inf_{s \le t}[S_k(t-s) + X_k(s)] \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \notin \mathcal{B}}(\check{D}_j(s) + u_{jk})] \\
&\quad \wedge \inf_{s \le t_0}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}_j(s) + u_{jk})] \\
&= \inf_{s \le t}[S_k(t-s) + X_k(s)] \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \notin \mathcal{B}}(\check{D}'_j(s) + u_{jk})] \\
&\quad \wedge \inf_{s \le t_0}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}'_j(s) + u_{jk})] \\
&\ge \inf_{s \le t}[S_k(t-s) + X_k(s)] \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \notin \mathcal{B}}(\check{D}'_j(s) + u_{jk})] \\
&\quad \wedge \inf_{s \le t}[S_k(t-s) + \min_{j \in \mathcal{B}}(\check{D}'_j(s) + u_{jk})] \\
&= \check{D}'_k(t) \ge \check{D}_k(t).
\end{aligned}
$$

Thus, $\check{D}_k(t) = \check{D}'_k(t)$ for $t \in (t_0, t_0 + (\delta' \wedge \delta))$. This contradicts the fact that $k \in \mathcal{B}$, and the uniqueness is established as a consequence.

Finally, we show that $\check{\mathbf{D}}^d, d \ge 0$ defined in the theorem converge to this unique solution. It is easy to establish by induction that $\check{\mathbf{D}}^d, d \ge 0$ is a non-increasing sequence and hence has a limit $\check{\mathbf{D}}^\infty := \lim_{d \to \infty} \check{\mathbf{D}}^d = \inf_{d \ge 0} \check{\mathbf{D}}^d$. Thus,

$$
\begin{aligned}
\check{D}_k^\infty(t) &= \inf_{d \ge 0} \check{D}_k^d(t) = \inf_{d \ge 0} \check{D}_k^{d+1}(t) \\
&= \inf_{d \ge 0} \inf_{s \le t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}^d(s))] \\
&= \inf_{s \le t}[S_k(t-s) + M_k(\mathbf{X}(s), (\inf_{d \ge 0} \check{\mathbf{D}}^d(s)))] \\
&= \inf_{s \le t}[S_k(t-s) + M_k(\mathbf{X}(s), \check{\mathbf{D}}^\infty(s))].
\end{aligned}
$$

□

References

[1] J. C. R. Bennett and H. Zhang. Hierarchical packet fair queueing algorithms. In *Proc. ACM SIGCOMM'96 Conference*, pages 143–156, Stanford University, CA, August 1996.

[2] R. Braden, et. al. Resource ReSerVation Protocol (RSVP) - version 1, functional specification. Internet Draft, draft-ietf-rsvp-spec-08.ps, Nov 1995.

[3] R. L. Cruz. A calculus for network delay, part I: Network elements in isolation. *IEEE Trans. Inform. Theory*, 37:114–131, 1991.

[4] R. L. Cruz. A calculus for network delay, part II: Network analysis. *IEEE Trans. Inform. Theory*, 37:132–141, 1991.

[5] R. L. Cruz. Quality of service guarantees in virtual circuit switched networks. *IEEE Journal on Selected Areas in Communications*, 13:1048–1056, August 1995.

[6] L. Georgiadis, R. Guerin, V. Peris, and R. Rajan. Efficient support of delay and rate guarantees in an internet. In *Proc. ACM SIGCOMM'96 Conference*, pages 106–116, Stanford University, CA, August 1996.

[7] S. J. Golestani. A self-clocked fair queueing scheme for broadband applications. In *Proceedings of IEEE INFOCOM'94*, pages 636–646, Toronto, Canada, June 1994.

[8] P. Goyal, S. S. Lam, and H. M. Vin. Determining end-to-end delay bounds in heterogeneous networks. *ACM/Springer-Verlag Multimedia Systems Journal*, 1996. To appear.

[9] J. M. Harrison. *Brownian motion and stochastic flow systems.* Wiley, New York, 1985.

[10] T. Konstantopoulos and V. Anantharam. Optimal flow control schemes that regulate the burstiness of traffic. *IEEE/ACM Transactions on Networking*, 3(4):450–458, August 1995.

[11] A. K. Parekh. *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks.* PhD thesis, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, February 1992. No. LIDS-TH-2089.

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

[13] A. K. Parekh and R. G. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The multiple node case. *IEEE/ACM Transactions on Networking*, 2(2):137–150, April 1994.

[14] S. Shenker and C. Partridge. Specification of guaranteed quality of service. Internet Draft draft-ietf-intserv-guaranteed-svc-03.txt, Nov 1995.

[15] D. Stiliadis and A. Varma. Latency-rate servers: A general model for analysis of traffic scheduling algorithms. In *Proc. IEEE INFOCOM'96*, pages 111–119, San Francisco, March 1996.

[16] J. Wroclawski. Specification of the controlled-load network element service. Internet Draft draft-ietf-intserv-ctrl-load-svc-01.txt, Nov 1995.

[17] H. Zhang and D. Ferrari. Rate-controlled static priority queueing. In *Proceedings of IEEE INFOCOM'93*, pages 227–236, San Francisco, April 1993.

[18] H. Zhang and D. Ferrari. Rate-controlled service disciplines. *Journal of High Speed Networks*, 3(4):389–412, 1994.

[19] L. Zhang. Virtual clock: A new traffic control algorithm for packet switching networks. In *Proceedings of the ACM SIGCOMM'90*, pages 19–29, Philadelphia, PA, Sept. 1990.

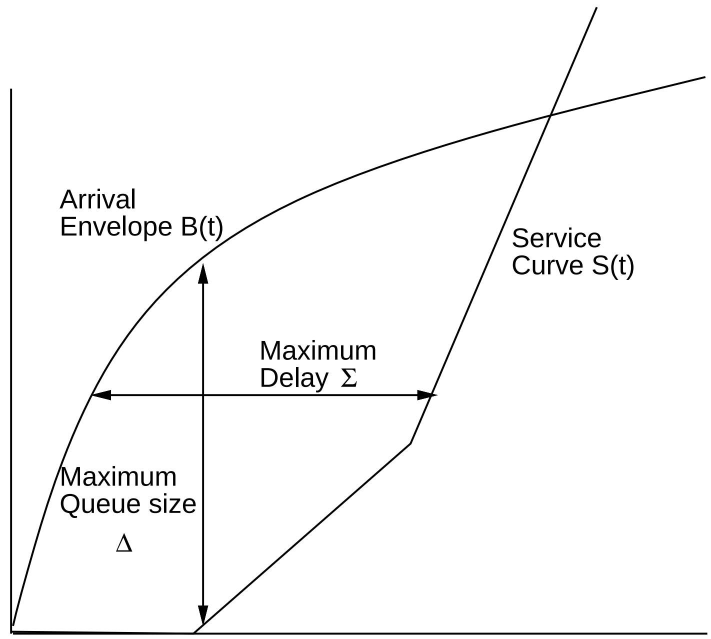

Figure 1. The worst case delay and queue length bounds.

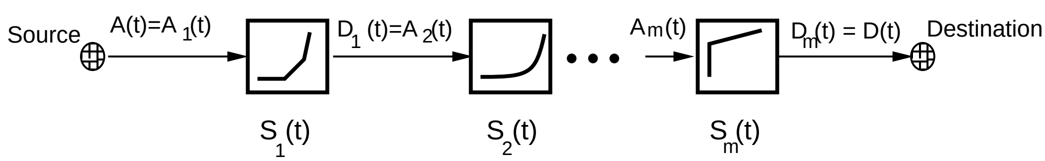

Figure 2. A unicast session without feedback.

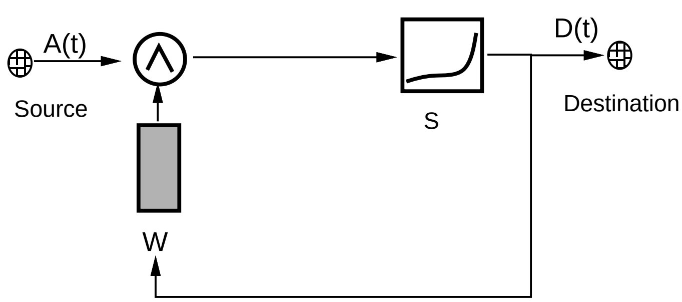

Figure 3. A single hop session with window flow control.

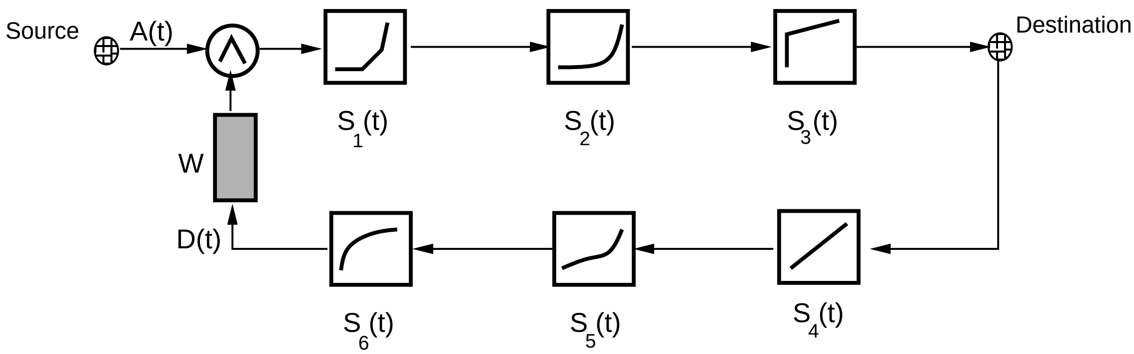

Figure 4. A unicast session with window flow control.

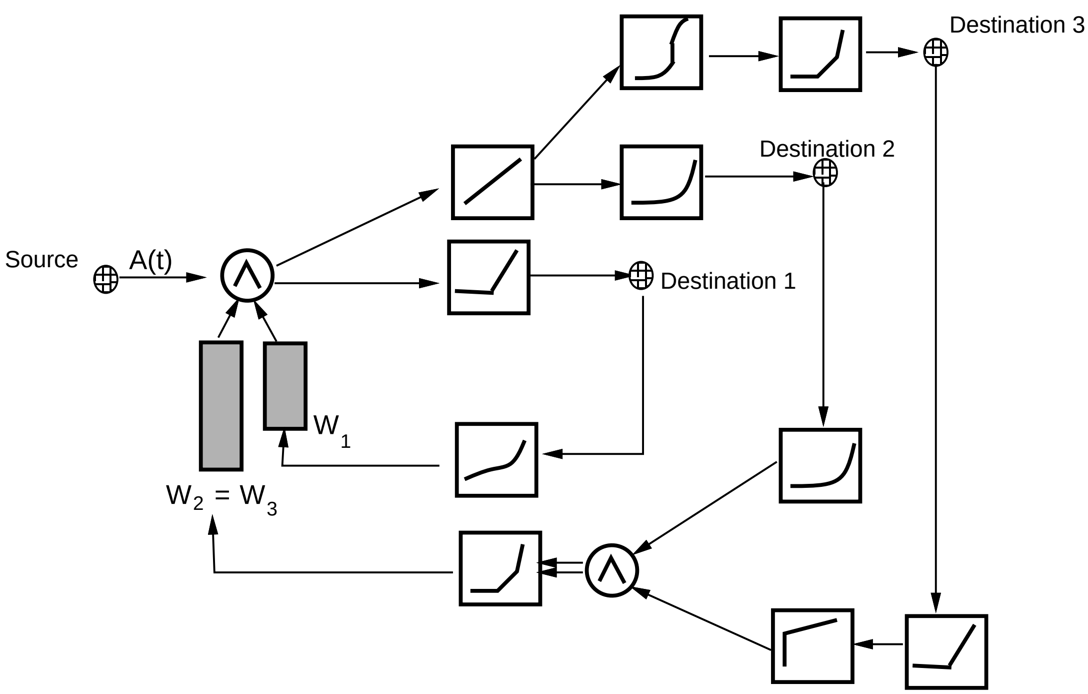

Figure 5. A multicast session with window flow control.

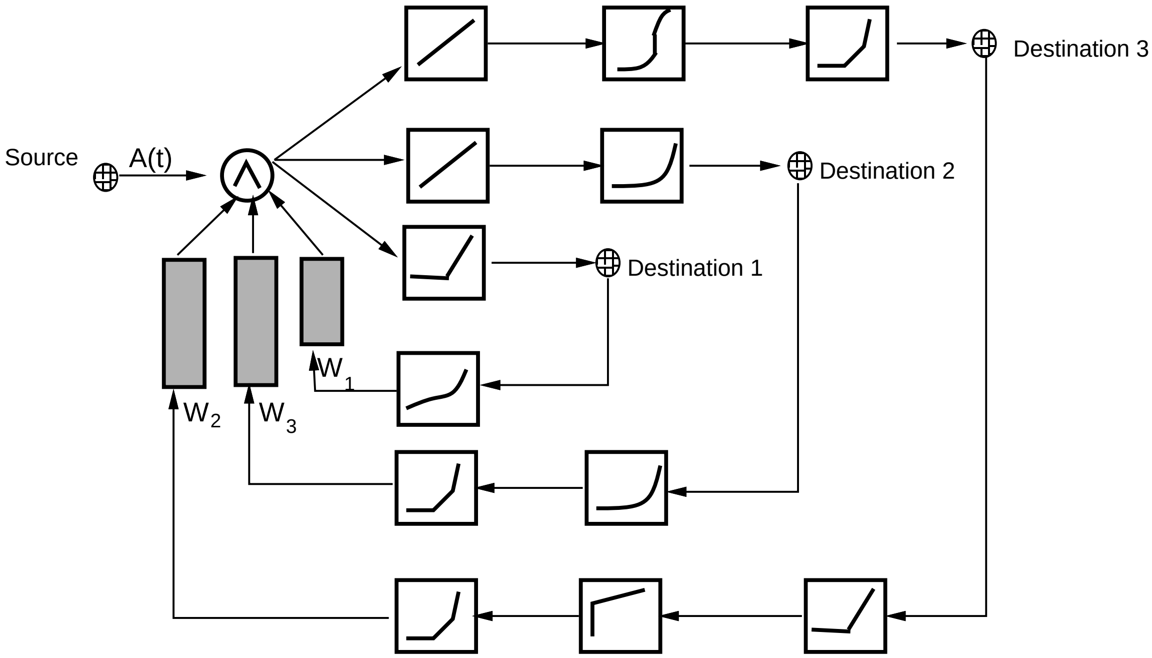

Figure 6. Illustration of the composition rule for multicast sessions.

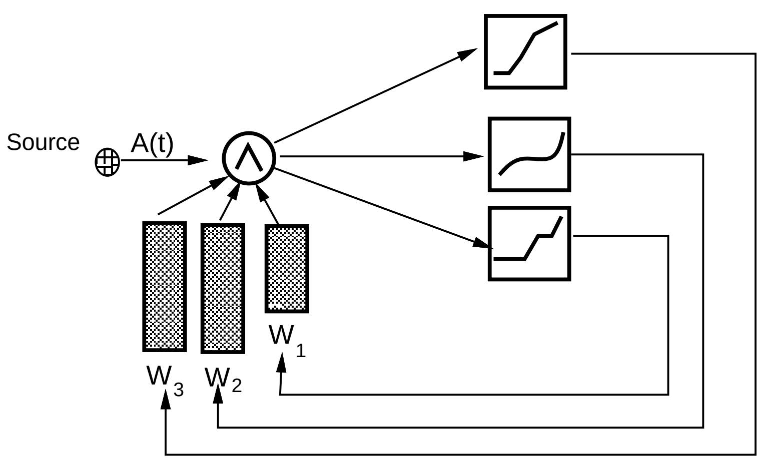

Figure 7. Illustration of the composition rule for multicast sessions.

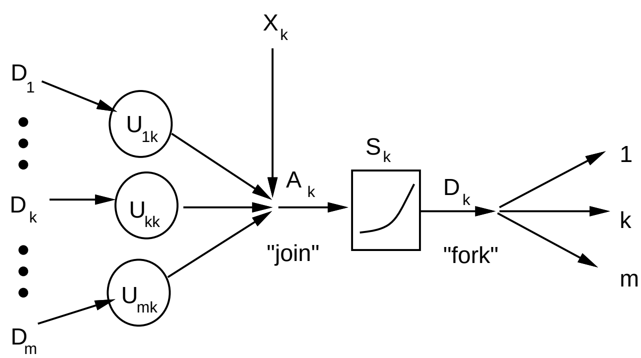

Figure 8. Schematic illustration of fork-join network element.