Packet Scale Rate Guarantee for non-FIFO Nodes

Jean-Yves Le Boudec and Anna Charny

Abstract—**Packet Scale Rate Guarantee (PSRG) is a generic node model which underlies the definition of Expedited Forwarding (EF) proposed in the context of Internet Differentiated Services. For the case of FIFO nodes, PSRG is equivalent to the well-understood concept of adaptive service curve. However, in practice, many devices do not necessarily preserve the FIFO property, and therefore known FIFO results do not hold. This paper analyzes the properties of PSRG in the absence of FIFO assumption. Our analysis is based on a novel characterization of PSRG which avoids the use of virtual finish times; it is obtained by min-max algebra. We use it to show that delay bounds previously obtained for the FIFO case are still valid; in contrast, we find that this is not true for the characterization of the concatenation of two nodes.**

Keywords: Expedited Forwarding, Differentiated Services, Min-Max, Network Calculus

I. INTRODUCTION

Packet Scale Rate Guarantee (PSRG) [1] was first introduced in the context of Differentiated Services (Diff-Serv) to define the Expedited Forwarding per-hop behaviour (EF PHB) [2]. It describes the service provided by a node to an aggregate of traffic destined to a particular output from one or more inputs. PSRG can be viewed as a characterization of how far a node differs from an ideal node that would implement Generalized Processor Sharing (GPS)[3]. Previous abstractions such as Guaranteed Rate Clock (GRC)[4] or service curves [5], [6], [7] capture how much a node can be late with respect to GPS. PSRG goes one step further and captures how much a node is either late or early with respect to GPS, thus providing a better way of defining a node as being "approximately GPS". For a discussion of PSRG versus GRC, see [1] and comment 2 after Theorem II.1 in this paper.

Note that PSRG does *not* imply that the node is work-conserving; indeed, it is intended to model complex nodes, such as Internet routers, that consist of many components; viewed as blackboxes, such nodes are generally not work conserving (see also the examples that can be built from Section V). It is a merit of PSRG to be able to derive the bounds mentioned in the next paragraph in the absence of any work conservation assumption.

PSRG has two parameters: a rate and a latency; the Internet Engineering Task Force (IETF) has published a document [8] that gives the values of these parameters for

Author Affiliations: Jean-Yves Le Boudec <jean-yves.leboudec@epfl.ch> EPFL-INN, CH-1015 Lausanne, Switzerland; Anna Charny <acharny@cisco.com>, Cisco, Chelmsford, MA 01776

nodes of various types that are found in practice.

PSRG is related to the concept of adaptive service curve, another abstract node model introduced in [9], [10]. An adaptive service curve is a wide-sense increasing function of time $\beta(t)$ that, roughly speaking, expresses the amount of service guaranteed over any time interval of length t. It is shown in [11], Theorem 7.3.1 on page 235, that an adaptive service curve of the form $\beta(t) = \max[r(t - e), 0]$ implies PSRG with rate r and latency e; the converse is true only for a FIFO node. This is used in [1], [11] to establish the following properties of PRSG for FIFO nodes: a delay-from-backlog bound (same as in Section III); a delay bound with arrival curve constraint on the input (same as in Section IV); a concatenation rule (see proof of Theorem V.1). The first of these three results is typical of PSRG, and cannot be obtained with the concepts of GRC or service curve. Being able to bound delay from backlog is useful in networks that use statistical multiplexing; typically there, the available buffer B at a node is less than the worst case buffer, and bounding the backlog by B can be used to deduce a bound on delay, using the delay-from-backlog bound [12]. Applications that need a strict delay bound and small loss are interactive audio and circuit emulation, which employ some form of FEC to correct for losses. For such applications, small network buffers are required. The alternative of having potentially large buffers in the network and having the destination drop late packets at destination does not work [13].

The goal of this paper is to examine these properties for *non-FIFO* nodes (work-conserving or not). Indeed, in practice, many devices cannot be accurately described by a FIFO model. Even when all packets of the aggregate share a single FIFO queue at the output (a standard Diff-Serv assumption), packets of the aggregate arriving from different input ports may experience variable amount of delay before they can be delivered at the output. These variable delays may cause reordering of packets as they arrive to the output queue compared to their order of arrivals to the input ports. A popular high-speed switch and router architecture is based on input-output buffered crossbars [14], [15]. In this architecture, packets arriving to different inputs are typically stored in queues at the input ports while they await for an opportunity to be transferred to the output queue. The crossbar architecture imposes a constraint that at any scheduling opportunity at most one packet can leave a given input, and at most one

packet can be transferred to any output. Since packets from any input can be destined to any output, a crossbar arbiter typically needs to solve a bipartite matching problem at any scheduling opportunity, and frequently does not preserve the order of packet arrivals at the input when choosing which packets to transfer from which input. For this architecture the out-of-order packet delivery happens as a rule and is by no means a rare exception. Note that, in this case, this is not contradictory with the fact that end-user flows may not be reordered; this is because an EF flow is an aggregate of many individual end-user flows.

Packet reordering within an aggregate also occurs as a rule in multistage fabrics, where different packets may follow different paths through the network and get reordered due to the resulting variability of cross-fabric delays. Note that in this case even packets arriving to the same input may be easily reordered. Finally, if the scheduler providing PRSG to an aggregate is more complicated than a FIFO, the scheduler itself may cause packet reordering. An example of such a scheduler is described in [16].

Hence, understanding the properties of PSRG in the non-FIFO case is of significant practical importance. Yet, the techniques used to demonstrate the properties of PSRG in the FIFO case do not appear to extend to the non-FIFO case, and the properties of the non-FIFO node providing PSRG to an aggregate have not been well understood. In this paper we introduce some new techniques that enable us to explore the properties of PSRG in the absence of the FIFO assumption. They are based on min-max algebra.

PSRG is defined as follows. The service curve of a given queue conforms to PSRG if the departure time d_n of the nth packet to arrive to the queue satisfies

$$d_n \leq f_n + e \qquad (1)$$

where e is the so-called latency (or error) term and f_n is given by the following recursion:

$$\begin{cases} f_0 = 0 \\ f_n = \max\left[a_n, \min\left(d_{n-1}, f_{n-1}\right)\right] + \frac{l_n}{r} \text{ for } n \geq 1 \end{cases} \qquad (2)$$

Here a_n is the time of the nth arrival to the queue, r is the guaranteed rate of the scheduler, and l_n is the length of the nth packet. The choice of indices assumes that there are no packets in the node at time zero. In this paper, we consider that packet n is the nth packet to arrive at the node, with some unspecified rule for breaking ties. This is called the "packet identity aware" definition in the proposed definition of EF [2].

Because PSRG gives an upper bound on d_n, but no lower bound (other than $d_n \geq a_n$), it is possible for a node to satisfy a PSRG property and not be FIFO. In fact, Section V provides a way to build such examples.

It is instructive to compare the definition in Equation (2) to a well-known Guaranteed Rate Curve (GRC) [4]. GRC is also defined by Equation (1), except f_n is given by a different recursion:

$$\begin{cases} f_0 = 0 \\ f_n = \max\left[a_n, f_{n-1}\right] + \frac{l_n}{r} \text{ for } n \geq 1 \end{cases} \qquad (3)$$

It can be shown that both in FIFO and non-FIFO cases, PSRG implies GRC [1]. Therefore, all properties of GRC also hold for PSRG. However, most of the results known for GRC, such as backlog bound as a function of an arrival curve and the concatenation theorem [4], have been studied for the FIFO case [11], and the techniques used for the FIFO case do not appear to extend to the non-FIFO case. Although the motivation of this paper is to study PSRG in a non-FIFO case, the techniques developed here also find application to non-FIFO GRC nodes. In particular, we show that the delay bound as a function of arrival constraints for the non-FIFO case is the same as for the FIFO case.

Since PSRG is stronger than GRC, some properties of PSRG do not hold for GRC. In particular, one of the most useful properties of PSRG is that unlike GRC, it permits expressing the bound on delay as a function of backlog ("delay-from-backlog bound"). It was shown in [1] that, for the FIFO case, a packet finding a queue Q upon its arrival to a PSRG server will be delayed at most by $Q/r + e$, where r and e are the guaranteed rate and the latency of PSRG respectively. In this paper, we demonstrate that the same result holds for the non-FIFO case as well. Note that this result is entirely not obvious – in the FIFO case a packet arriving to queue is delayed only by packets that arrived prior to it to this queue, while in the non-FIFO case packets that arrived later may nevertheless be transmitted ahead of the packet of interest. It turns out this does not change the delay bound previously known for the FIFO case.

In contrast, the behavior of a concatenation of two PSRG nodes is different in the FIFO and non-FIFO cases. It is known ([11], Proposition 7.3.2 on page 236) that the concatenation of two FIFO PSRG nodes with latencies e_1 and e_2 and rates r_1 and r_2 is a PSRG system with rate $r = \min(r_1, r_2)$ and latency $e = e_1 + e_2 + \frac{l_{\max}}{r_1}$, where $l_{\max}$ is the maximum packet size for the flow under consideration. Our analysis in Section V shows that this result no longer holds for the non-FIFO case.

The paper is organized as follows. In Section II we describe an alternative characterization of PSRG which is crucial in developing the results for the non-FIFO case. It avoids the use of the virtual finish times f_n, and is our main theoretical tool. It is based on min-max algebra. In Section III we apply this characterization to show that the delay-from-backlog bound also holds for non-FIFO nodes. In Section IV we show a similar conclusion for the backlog bound as a function of an arrival curve; this

latter result is in reality a property of GRC nodes. In Section V we study a specific concatenation scenario, where the first node is a variable delay node, and the second node is a FIFO PSRG node. This models many of the nodes mentioned earlier in this section (the general concatenation scenario is left for further study). If the first node would be FIFO, there would be simple results, derived from the theory developed before this paper. However, this does not hold if the first node is not FIFO; we find another concatenation result in that case. In Section VI we give parallel results for non-FIFO GRC nodes, that are obtained with the same technique. Some of the proofs are long and are given in appendix.

II. AN ALTERNATE CHARACTERIZATION OF PACKET SCALE RATE GUARANTEE

In this section we obtain a characterization of packet scale rate guarantee which, unlike the original definition ([1] or Equation (2)), does not contain the virtual finish times f_n. It is the basis for most results in this paper. We start with an expansion of the recursive definition of packet scale rate guarantee,

Lemma II.1 (Min-max expansion of PSRG) Consider three arbitrary sequences of non-negative numbers $(a_n)_{n \geq 1}$, $(d_n)_{n \geq 0}$, and $(m_n)_{n \geq 1}$, with $d_0 = 0$. Define the sequence $(f_n)_{n \geq 0}$, by

$$\begin{cases} f_0 = 0 \\ f_n = \max\left[a_n, \min\left(d_{n-1}, f_{n-1}\right)\right] + m_n \text{ for } n \geq 1 \end{cases}$$

Also define

$$\begin{cases} A_j^n = a_j + m_j + ... + m_n \text{ for } 1 \leq j \leq n \\ D_j^n = d_j + m_{j+1} + ... + m_n \text{ for } 0 \leq j \leq n - 1 \end{cases}$$

For all $n \geq 1$, we have

$$\begin{aligned} f_n = \min \quad [\quad &\max(A_n^n, A_{n-1}^n, ..., A_1^n), \\ &\max(A_n^n, A_{n-1}^n ..., A_2^n, D_1^n), \\ &... \\ &\max(A_n^n, A_{n-1}^n ..., A_{j+1}^n, D_j^n), \\ &... \\ &\max(A_n^n, A_{n-1}^n, D_{n-2}^n), \\ &\max(A_n^n, D_{n-1}^n) \\ &] \end{aligned}$$

The proof is technical and is given in appendix; it is based on min-max algebra.

Comment: The expansion in Lemma II.1 can be interpreted as follows. The first term $\max(A_n^n, A_{n-1}^n, ..., A_1^n)$ corresponds to the guaranteed rate clock recursion (see Lemma VI.1). The following terms have the effect of reducing f_n, depending on the values of d_j.

We now apply the previous lemma to packet scale rate guarantee and obtain the required characterization without the virtual finish times f_n:

Theorem II.1: Consider a system where packets are numbered $1, 2, ...$ in order of arrival. Call a_n, d_n the arrival and departure times for packet n, and l_n the size of packet n. Define by convention $d_0 = 0$. The packet scale rate guarantee with rate r and latency e is equivalent to: For all n and all $0 \leq j \leq n - 1$, one of the following holds

$$d_n \leq e + d_j + \frac{l_{j+1} + ... + l_n}{r} \tag{4}$$

or there is some $k \in \{j + 1, ..., n\}$ such that

$$d_n \leq e + a_k + \frac{l_k + ... + l_n}{r} \tag{5}$$

The proof is also given in appendix. It is a straightforward application of Lemma II.1.

Comment 1: The original definition of EF in [17] was based on the informal intuition that a node guarantees to the EF aggregate a rate equal to r, at all time scales (this informal definition was replaced by PSRG). Theorem II.1 makes the link to the original intuition: a rate guarantee at all time scales means that either Equation (4) or Equation (5) must hold. For a simple scheduler, the former means that d_j, d_n are in the same backlogged period; the latter is for the opposite case, and here a_k is the beginning of the backlogged period. But note that we do not assume that the PSRG node is a simple scheduler; as mentioned earlier, it may be any complex, non work conserving node. It is a merit of the abstract PSRG definition to avoid using the concept of backlogged period, which is not meaningful for a composite node [8], [1].

Comment 2: In Section VI we give a similar result for GRC. It is instructive to compare both in the case of a simple scheduler, where the interpretation in terms of backlogged period can be made. Let us assume the latency term is 0, to make the comparison simple. For such a simple scheduler, PSRG means that during *any* backlogged period, the scheduler guarantees a rate at least equal to r. In contrast, and again for such simple schedulers, GRC means that during the backlogged period starting at the first packet arrival that finds the system empty (this is called "busy period" in queuing theory), the average rate of service is at least r. GRC allows the scheduler to serve some packets more quickly than at rate r, and take advantage of this to serve other packets at a rate smaller than r, as long as the overall average rate is at least r. PSRG does not allow such a behaviour.

III. DELAY FROM BACKLOG FOR NON-FIFO NODE

A main feature of the packet scale rate guarantee definition is that it allows to bound delay from backlog. A delay from backlog bound was obtained in [1] for a FIFO node. The proof does not extend to non-FIFO nodes. However, using Theorem II.1, we can now show that the result does also hold for non FIFO nodes.

Theorem III.1: Consider a node offering the Packet Scale Rate Guarantee with rate r and latency e, not nec-

essarily FIFO. Call Q the backlog at time t. All packets that are in the system at time t will leave the system no later than at time $t + Q/r + e$,

Proof: Consider a fixed packet n which is present at time t. Call a_j [resp. d_j] the arrival [resp. departure] time of packet j. Thus $a_n \leq t \leq d_n$. Let $\mathcal{B}$ be the set of packet numbers that are present in the system at time t, in other words:

$$\mathcal{B} = \{k \geq 1 | a_k \leq t \leq d_k\}$$

The backlog at time t is $Q = \sum_{i \in \mathcal{B}} l_i$. The absence of FIFO assumption means that $\mathcal{B}$ is not necessarily a set of consecutive integers. However, define j as the minimum packet number such that the interval $[j, n]$ is included in $\mathcal{B}$. There is such a j because $n \in \mathcal{B}$. If $j \geq 2$ then $j - 1$ is not in $\mathcal{B}$ and $a_{j-1} \leq a_n \leq t$ thus necessarily

$$d_{j-1} < t \tag{6}$$

If $j = 1$, Equation (6) also holds with our convention $d_0 = 0$. Now we apply the alternate characterization of packet scale rate guarantee (Theorem II.1) to n and $j - 1$. One of the two following equations must hold:

$$d_n \leq e + d_{j-1} + \frac{l_j + ... + l_n}{r} \tag{7}$$

or there exists a $k \geq j$, $k \leq n$ with

$$d_n \leq e + a_k + \frac{l_k + ... + l_n}{r} \tag{8}$$

Assume that Equation (7) holds. Since $[j, n] \subset \mathcal{B}$, we have $Q_n \geq l_j + ... + l_n$. By Equation (6) and Equation (7) it follows that

$$d_n \leq e + t + \frac{Q}{r}$$

which shows the result in this case. Otherwise, use Equation (8); we have $Q \geq l_k + ... + l_n$ and $a_k \leq t$ thus

$$d_n \leq e + t + \frac{Q}{r}$$

$\diamond$

IV. Delay Bound for Arrival Curve Constraints

In this section we give a bound on delay which does not depend on the queue size, but on an arrival curve constraint. It is well known that such a constraint can be used to derive delay bounds, assuming that the node offers a service curve guarantee [9], [6], [7]. If a node offering a packet scale rate guarantee is FIFO, it follows from [11], Theorem 7.3.1, that it also offers a rate-latency service curve, from which a delay bound can be derived. In this section, we show that the same delay bound holds in the absence of FIFO assumption.

Contrary to the bound in Section III, the delay bound for arrival curve constraints is not typical of packet scale rate guarantee, but can be derived under the weaker assumption that the node satisfies a guaranteed rate clock property.

Theorem IV.1: Consider a guaranteed rate clock node with rate r and latency e. Assume that the input is constrained by an arrival curve $\alpha(\cdot)$. The node need not be FIFO. The delay for any packet is bounded by

$$\sup_{t>0}[\frac{\alpha(t)}{r} - t] + e \tag{9}$$

Note that Equation (9) is the horizontal deviation between the arrival curve α and the rate-latency service curve with rate r and latency e, which is a known result for FIFO nodes. The new information in the theorem is that it also holds for non FIFO nodes.

Proof: Call $a_n \geq 0, d_n \geq 0$ the arrival and departure times, with $n \geq 1$. Packets are numbered in order of arrival. We use Lemma VI.1, which is proved in Section VI. Thus, for any fixed n, we can find a $1 \leq j \leq n$ such that

$$f_n = a_j + \frac{l_j + ... + l_n}{r}$$

The delay for packet n is

$$d_n - a_n \leq f_n + e - a_n$$

Define $t = a_n - a_j$. By hypothesis

$$l_j + ... + l_n \leq \alpha(t+)$$

where $\alpha(t+)$ is the limit to the right of α at t. Thus

$$d_n - a_n \leq -t + \frac{\alpha(t+)}{r} + e \leq \sup_{t \geq 0}[\frac{\alpha(t+)}{r} - t] + e$$

Now $\sup_{t>0}[\frac{\alpha(t)}{r} - t] = \sup_{t \geq 0}[\frac{\alpha(t+)}{r} - t]$. $\diamond$

Since the packet scale rate guarantee implies the GRC property (Section I), we have:

Corollary IV.1: Consider a node satisfying the packet scale rate guarantee, with rate r and latency e. Assume that the input is constrained by an arrival curve $\alpha(\cdot)$. The node need not be FIFO. The delay for any packet is bounded by Equation (9).

Comment: If $\alpha(t) = \rho t + \sigma$, with $\rho \leq r$, then $D = \sigma/r + e$, which shows the statement on delay in [8]. If we know more constraints (typically the peak rate), then we get a better bound.

V. Composite Node With Variable Delay Component

In this section we consider a composite node, made of two components. The former ("variable delay component") imposes to packets a delay in the range $[\delta_{\max} - \delta, \delta_{\max}]$. The latter is FIFO and offers to its input the packet scale rate guarantee, with rate r and latency e. As

mentioned in Section I, this node is frequently found in practice. We show that, if the variable delay component is known to be FIFO, then we have a simple result. In contrast, if we cannot make this assumption, the analysis is more involved and requires Theorem II.1. We first give the following lemma, which has some interest of its own.

Lemma V.1 (Variable Delay as PSRG) Consider a node which is known to guarantee a delay $\leq \delta_{\max}$. The node need not be FIFO. Call $l_{\min}$ the minimum packet size. For any $r > 0$, the node offers the packet scale rate guarantee with latency $e = [\delta_{\max} - \frac{l_{\min}}{r}]^+$ and rate r.

Proof: With the standard notation in this paper, the hypothesis implies that $d_n \leq a_n + \delta_{\max}$ for all $n \geq 1$. Define f_n by Equation (2). We have $f_n \geq a_n + \frac{l_n}{r} \geq a_n + \frac{l_{\min}}{r}$, thus $d_n - f_n \leq \delta_{\max} - \frac{l_{\min}}{r} \leq [\delta_{\max} - \frac{l_{\min}}{r}]^+$. ◇

We will now apply known results on the concatenation of FIFO elements and solve the case where the variable delay component is FIFO.

Theorem V.1: (Composite Node with FIFO Variable Delay Component) Consider the concatenation of two nodes. The former imposes to packets a delay $\leq \delta_{\max}$. The latter offers the packet scale rate guarantee to its input, with rate r and latency e. Both nodes are FIFO. The concatenation of the two nodes, in any order, offers the packet scale rate guarantee with rate r and latency $e' = e + \delta_{\max}$.

Proof: We use [11], Proposition 7.3.2, which says that the concatenation of two *FIFO* nodes, each offering the packet scale rate guarantee, with rates r_1, r_2 and latencies e_1, e_2, offers the packet scale rate guarantee, with rate $\min(r_1, r_2)$ and latency $e_1 + e_2 + l_{\max}/r_1$ ($l_{\max}$ was defined earlier as the maximum packet size). Combine this with Lemma V.1: for any $r' \geq r$, the combined node offers the packet scale guarantee with rate r and latency $e(r') = e + \delta_{\max} + \frac{l_{\max} - l_{\min}}{r'}$. Define f_n for all n by Equation (2). Consider some fixed but arbitrary n. We have $d_n - f_n \leq e(r')$, and this is true for any $r' \geq r$. Let $r' \to +\infty$ and obtain $d_n - f_n \leq \inf_{r' \geq r} e(r') = e + \delta_{\max}$ as required. ◇

If we relax the FIFO assumption for the variable delay component, then Theorem V.1 does not hold (we give an example later in this section). However, if we can express an arrival curve constraint for the input traffic, then we still have a result, even for the non-FIFO case:

Theorem V.2: (Composite Node with non-FIFO Variable Delay Component) Consider the concatenation of two nodes. The first imposes to packets a delay in the range $[\delta_{\max} - \delta, \delta_{\max}]$. The second is FIFO and offers the packet scale rate guarantee to its input, with rate r and latency e. The first node is not assumed to be FIFO, so the order of packet arrivals at the second node is not the order of packet arrivals at the first one. Assume that the fresh input is constrained by a continuous arrival curve $\alpha(\cdot)$. The concatenation of the two nodes, in this order,
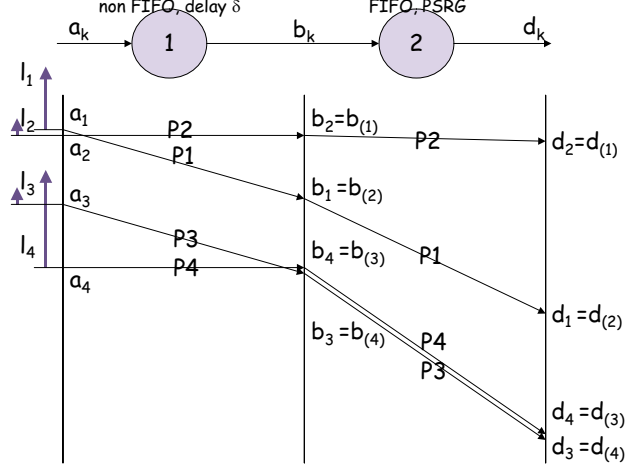

Fig. 1. Composite Node with non-FIFO Variable Delay Component. Packet n arrives at times a_n at the first component, at time b_n at the second component, and leaves the system at time d_n. Since the first component is not FIFO, overtaking may occur; (k) is the packet number of the kth packet arriving at the second component.

satisfies the packet scale rate guarantee with rate r and latency

$$
\begin{aligned}
e' = \ & e + \delta_{\max} + \\
& \min\{\sup_{t \geq 0}[\tfrac{\alpha(t+\delta) - l_{\min}}{r} - t], \\
& \sup_{0 \leq t \leq \delta}[\tfrac{\alpha(t) + \alpha(\delta) - 2l_{\min}}{r} - t]\}
\end{aligned}
\tag{10}
$$

The proof is long, and is given in appendix.

Figures 2 to 4 show numerical applications when the arrival curve includes both peak rate and mean rate constraints.

Special Case : For $\alpha(t) = \rho t + \sigma$, a direct computation of the suprema in Theorem V.2 gives:

$$
\begin{aligned}
\text{if } \rho \leq r \text{ then } \quad & e' = e + \delta_{\max} + \tfrac{\rho\delta + \sigma - l_{\min}}{r} \\
\text{else} \quad & e' = e + \delta_{\max} - \delta + 2\tfrac{\rho\delta + \sigma - l_{\min}}{r}
\end{aligned}
$$

The latency of the composite node has a discontinuity equal to σ/r at $\rho = r$. For the relevance of the case $\rho > r$, see the discussion at the end of Section VI.

Comment 1 : We now justify why Theorem V.2 is needed, in other words: if we relax the FIFO assumption for the variable delay component, then Theorem V.1 does not hold any more. Intuitively, this is because a tagged packet (say $P3$ on Figure 1) may be delayed at the second stage by packets ($P4$ on the figure) that arrived later, but took over our tagged packet. Also, the service rate may appear to be reduced by packets ($P1$ on the figure) that had a long delay in the variable delay component. Formally, we have:

Proposition V.1 (Tightness) The bound in Theorem V.2 is tight in the case of an arrival curve of the form $\alpha(t) = \rho t + \sigma$ and if $l_{\max} \geq 2l_{min}$.

Proof: Case $\rho \leq r$: Assume that the source is greedy from time 0, with packet $n = 1$, of size $l_1 = l_{\min}$, $a_1 = 0$,
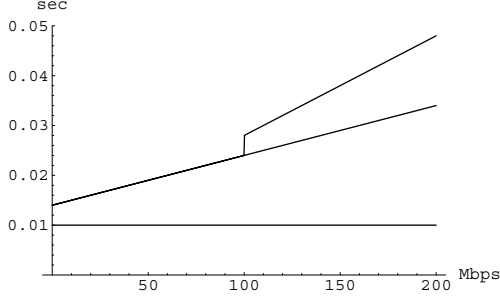
Fig. 2. Numerical Application of Theorem V.2 and Theorem VI.3, showing the additional latency $e' - e$ for a composite node, made of a variable delay element ($\delta = \delta_{\max} = 0.01s$) followed by a PSRG or GRC component with rate $r = 100$Mb/s and latency e. The fresh traffic has arrival curve $\rho t + \sigma$, with $\sigma = 50$KBytes. The figure shows $e' - e$ as a function of ρ, for $l_{\min} = 0$. Top graph: delay element is non-FIFO, second component is PSRG (Theorem V.2); middle graph: delay element is non-FIFO, second component is GRC (Theorem VI.3); bottom line: delay element is FIFO, both cases (Theorem V.1 and Theorem V.2). Top and middle graph coincide for $\rho \leq r$.
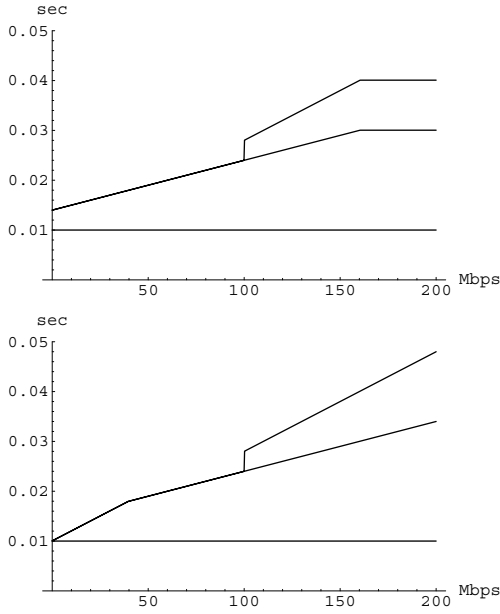

Fig. 4. Latency increase as a function of peak rate and mean rate. The parameters are the same as for Figure 3.

Fig. 3. Same as Figure 2, but the fresh traffis has a peak rate limit. The arrival curve for the fresh traffic is $\min(pt + MTU, \rho t + \sigma)$, with $MTU = 500$B, $p = 200Mb/s$ (top picture) or $p = 2\rho$ (bottom picture).

$b_1 = \delta_{\max}$. Assume all subsequent packets have a delay in the first component equal to $\delta_{\max} - \delta$. We can build an example where packet 1 is overtaken by packets $n = 2, ..., n_1$ that arrive in the interval $(0, \delta]$, with $l_2 + ... + l_{n_1} = \rho\delta + \sigma - l_1$. Assume that packet 1 undergoes the maximum delay allowed by PSRG at the second component. It follows after some algebra that $d_1 = e + \delta_{\max} + \frac{\rho\delta + \sigma}{r}$. Now $f_1 = \frac{l_{\min}}{r}$ thus $d_1 - f_1 = e'$ and the characterization is tight.

Case $\rho > r$: We build a worst case scenario as follows. We let $e = 0$, without loss of generality (add a delay el-
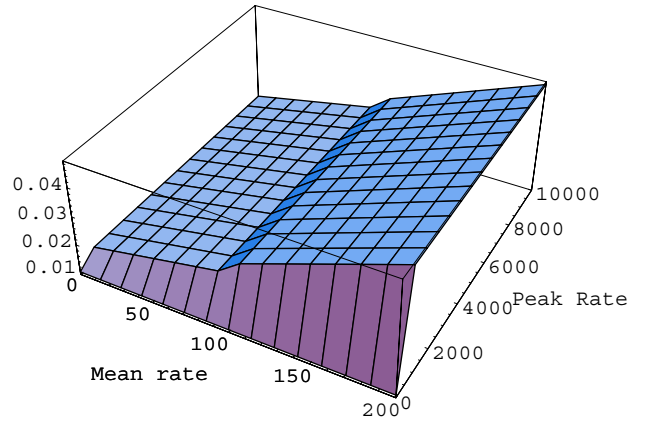
ement to this example and obtain the general case). The principle is to first build a maximum size burst which is overtaken by a tagged packet j. Later, a tagged packet n is overtaken by a second maximum size burst. Between the two, packets arrive at a data rate r; the second burst is possible because $r < \rho$ and $a_n - a_j$ is long enough. Details are in Figure 5 and Table I. We find finally $d_n - f_n = 2(\rho\delta + \sigma - l_{\min})/r$ which shows that the bound is achieved.
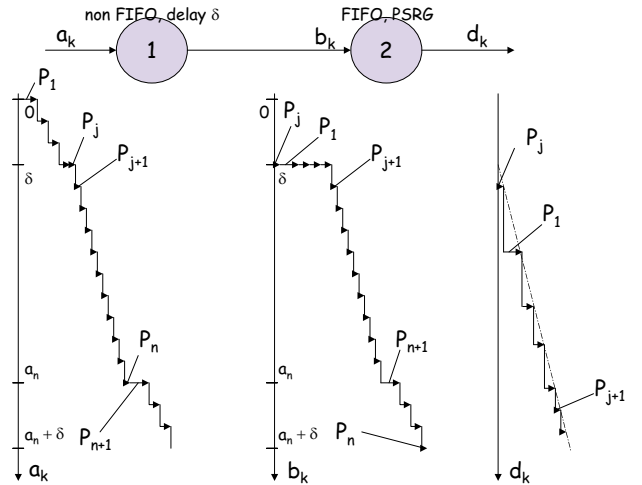

Fig. 5. Worst-case example for Theorem V.2. All packets have 0 delay through the first component except packets $1...j - 1$ and n.

$\diamond$

The proposition shows that the concatenation of non-FIFO PSRG nodes does not follow the rule as for FIFO nodes, which is recalled in the proof of Theorem V.1. Note that if the condition $l_{\max} \geq 2l_{\min}$ is not satisfied then the bound in Theorem V.2 is tight up to a tolerance

k	a_k	l_k	b_k	f_k	d_k
1	0	$\sigma - l_{\min}$	δ^+	not relevant	$d_j + l_1/r$
2	l_2/ρ	l_2	δ^+	not relevant	$d_j + (l_1 + l_2)/r$
...	...	...	...	...	...
$j-1$	$(l_2 + ... + l_{j-1})/\rho = \delta$	l_{j-1}	δ^+	not relevant	$d_j + A$
j	δ	$l_{\min}$	δ	$\geq \delta + l_{\min}/r$	$\delta + l_{\min}/r$
$j+1$	$\delta + l_{\min}/r$	$l_{\min}$	a_{j+1}	$\delta + 2l_{\min}/r$	$f_{j+1} + A$
...	...	...	...	...	...
$n-1$	$\delta + (n-j-1)l_{\min}/r$	$l_{\min}$	a_{n-1}	$\delta + (n-j)l_{\min}/r$	$f_{n-1} + A$
n	$\delta + (n-j)l_{\min}/r$	$l_{\min}$	$a_n + \delta$	$\delta + (n-j+1)l_{\min}/r$	$f_n + 2A$
$n+1$	a_n^+	$\sigma - l_{\min}$	a_{n+1}	not relevant	$f_{n-1} + A + (\sigma - l_{\min})/r$
$n+2$	$a_n + a_2$	l_2	a_{n+2}	not relevant	$f_{n-1} + A + (\sigma - l_{\min} + l_2)/r$
...	...	...	...	...	...
$n+j-1$	$(a_n + a_{j-1})^- = (a_n + \delta)^-$	l_{j-1}	$(a_n + \delta)^-$	not relevant	$f_{n-1} + 2A$

Notes: $A = (\rho\delta + \sigma - l_{\min})/r$

$(j, l_2, ..., l_{j-1})$ is a solution to $l_2 + ... + l_{j-1} = \rho\delta$, sc $l_2, ..., l_{j-1} \in [l_{\min}, l_{\max}]$. For example, let $j = 2 + \lfloor \frac{\rho\delta}{l_{\min}} \rfloor$,

$l_2 = \rho\delta - (j-3)l_{\min}, l_3 = ... = l_{j-1} = l_{\min}$. We have $l_2 \leq l_{\max}$ because $l_{\max} \geq 2l_{\min}$

TABLE I

DETAILS FOR FIGURE 5. ASSUME FOR THIS TABLE THAT $\sigma - l_{\min} \leq l_{\max}$, OTHERWISE REPLACE PACKETS 1 AND $n+1$ BY A NUMBER OF SMALLER PACKETS ARRIVING IN BATCH.

of $2l_{\min}/r$.

Comment 2 : Equation (10) for the latency is the minimum of two terms. In the case $\alpha(t) = \rho t + \sigma$, for $\rho \leq r$, the bound is equal to its former term, otherwise to its second term. For a general α however, such a simplification does not occur.

Comment 3 : If α is not continuous (thus has jumps at some values), then it can be shown that Theorem V.2 still holds, with Equation (10) replaced by

$$e' = e + \delta_{\max} + \min\{\sup_{t\geq 0}[\frac{\alpha(t+\delta)}{r} - t], \sup_{0\leq t\leq \delta}[\frac{\alpha_0(t)+\alpha_0(\delta)}{r} - t]\}$$

with $\alpha_0(u) = \min[\alpha(u+) - l_{\min}, \alpha(u)]$.

VI. ARE THERE COMPARABLE RESULTS FOR GUARANTEED RATE CLOCK SCHEDULERS ?

Our initial motivation in this paper is to understand packet scale rate guarantee in the non-FIFO case; in this section, as a basis for comparison, we analyze whether similar results hold for GRC.

First, we have already seen (Section IV) that the delay bound with arrival constraints is in reality obtained for the guaranteed rate clock guarantee, and we do have the same bound for packet scale rate guarantee. Second, in contrast, it is known [1] that for GRC there cannot exist a delay-from-backlog bound as in Section III. Third, it remains to analyze the concatenation of nodes as in Section V; we find a similar (but simpler) result, which we give now. The proofs for the next 4 results mimic those of their PSRG counterparts, and are not given here. The proof of the last theorem is slightly different and is given in appendix.

Lemma VI.1: Consider two arbitrary sequences of non-negative numbers $(a_n)_{n\geq 1}$, and $(m_n)_{n\geq 1}$. Define the sequence $(f_n)_{n\geq 0}$, by

$$\begin{cases} f_0 = 0 \\ f_n = \max[a_n, f_{n-1}] + m_n \text{ for } n \geq 1 \end{cases}$$

Also define

$$A_j^n = a_j + m_j + ... + m_n \text{ for } 1 \leq j \leq n$$

For all $n \geq 1$, we have

$$f_n = \max(A_n^n, A_{n-1}^n, ..., A_1^n)$$

Theorem VI.1: Consider a system where packets are numbered $1, 2, ...$ in order of arrival. Call a_n, d_n the arrival and departure times for packet n, and l_n the size of packet n. Define by convention $d_0 = 0$. The guaranteed rate clock definition with rate r and latency e is equivalent to saying that for all n there is some $k \in \{1, ..., n\}$ such that

$$d_n \leq e + a_k + \frac{l_k + ... + l_n}{r} \qquad (11)$$

Lemma VI.2 (Variable Delay as GRC) Consider a node which is known to guarantee a delay $\leq \delta_{\max}$. The node need not be FIFO. Call $l_{\min}$ the minimum packet size. For any $r > 0$, the node is a guaranteed rate clock scheduler with latency $e = [\delta_{\max} - \frac{l_{\min}}{r}]^+$ and rate r.

Theorem VI.2: (*Composite GRC Node with FIFO Variable Delay Component*) Consider the concatenation of two nodes. The former imposes to packets a delay $\leq \delta_{\max}$. The latter is a guaranteed rate clock scheduler with rate r and latency e. Both nodes are FIFO. The concatenation of the two nodes, in any order, is a guaranteed rate clock scheduler with latency rate r and latency $e'' = e + \delta_{\max}$.

Theorem VI.3: Consider the concatenation of two nodes. The first imposes to packets a delay in the range $[\delta_{\max} - \delta, \delta_{\max}]$. The second is FIFO and offers the guaranteed rate clock service to its input, with rate r and latency e. The first node is not assumed to be FIFO, so the order of packet arrivals at the second node is not the order of packet arrivals at the first one. Assume that the fresh input is constrained by a continuous arrival curve $\alpha(\cdot)$. The concatenation of the two nodes, in this order, offers to the fresh input the guaranteed rate clock service with rate r and latency

$$e'' = e + \delta_{\max} + \frac{\alpha(\delta) - l_{\min}}{r}$$

The proof is given in appendix. It uses a similar method as the proof of Theorem V.2.

Application: For $\alpha(t) = \rho t + \sigma$, we find

$$e'' = e + \delta_{\max} + \frac{\rho\delta + \sigma - l_{\min}}{r}$$

and this is true even for $\rho > r$.

Comment: Compare with the value of e', obtained for packet scale rate guarantee. For $\rho \leq r$, the latencies e' and e'' are identical. However, the guarantee expressed by PSRG is stronger than that of GRC. Thus the stronger guarantee of PSRG comes at no cost, in that case.

It may seem irrelevant to consider the case $\rho > r$. Indeed, the GRC node model is not very useful here, since bounds for this model (such as Corollary IV.1) require arrival curve constraints and are infinite for $\rho > r$. However, the same cannot be said for PSRG, since it gives a delay from backlog bound; there may be cases where the only information available on the aggregate input is a bound on sustainable rate ρ, with $\rho > r$. In such cases, there are probably other mechanisms (such as window flow control [18]) to prevent buffer overflow; here, it is useful to be able to bound e' as in Theorem V.2.

VII. Conclusion

We have introduced a new technique, based on min-max algebra, for the analysis of non-FIFO PSRG nodes. We have shown that both delay-from-backlog bounds and delay bound from arrival curve constraints continue to hold for non-FIFO cases. In contrast, this is not true for the concatenation of non-FIFO PSRG nodes. We have analyzed a specific concatenation scenario, where the first node is a variable delay node, and the second node is a FIFO PSRG node. We have found a PSRG characterization of the concatenated node, in both cases where the first node is FIFO or not. The latency term in the latter case is larger than in the former. We have shown that similar results hold for GRC nodes. Further work focuses on more general concatenation scenarios.

References

[1] J. C. R. Bennett, Benson K., Charny A., Courtney W. F., and J.-Y. Le Boudec, "Delay jitter bounds and packet scale rate guarantee for expedited forwarding," *ACM/IEEE Transactions on Networking*, vol. to appear, 2002.

[2] B. Davie, A. Charny, J.C.R. Bennett, K. Benson, J.Y. Le Boudec, W. Courtney, S. Davari, V. Firoiu, and D. Stiliadis, "An expedited forwarding phb (per-hop behavior)," RFC 3246, March 2002.

[3] 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 Trans. Networking, vol 1-3*, pp. 344–357, June 1993.

[4] P. Goyal and H. Vin, "Generalized guaranteed rate scheduling algorithms: a framework," *IEEE/ACM Trans. Networking, vol 5-4*, pp. 561–572, August 1997.

[5] R. Agrawal, R. L. Cruz, C. Okino, and R. Rajan, "Performance bounds for flow control protocols," *IEEE/ACM Transactions on Networking (7) 3*, pp. 310–323, June 1999.

[6] C.S. Chang, "On deterministic traffic regulation and service guarantee: A systematic approach by filtering," *IEEE Transactions on Information Theory*, vol. 44, pp. 1096–1107, August 1998.

[7] J.-Y. Le Boudec, "Application of network calculus to guaranteed service networks," *IEEE Transactions on Information Theory*, vol. 44, pp. 1087–1096, May 1998.

[8] A. Charny, J.C.R. Bennett, K. Benson, J. Y. Le Boudec, A. Chiu, W. Courtney, S. Davari, V. Firoiu, C. Kalmanek, and K.K. Ramakrishnan, "Supplemental information for the new definition of the ef phb (expedited forwarding per-hop behavior)," RFC 3247, March 2002.

[9] R. Agrawal, R. L. Cruz, C. Okino, and R. Rajan, "A framework for adapative service guarantees," in *Proc. Allerton Conf on Comm, Control and Comp, Monticello, IL*, Sept 1998.

[10] Clayton M. Okino, "A framework for performance guarantees in communication networks," 1998, Ph.D. Dissertation, UCSD.

[11] J.-Y. Le Boudec and P. Thiran, *Network Calculus*, Springer Verlag Lecture Notes in Computer Science volume 2050 (available online at http://lcawww.epfl.ch), July 2001.

[12] M. Vojnovic and Le Boudec J.-Y., "Stochastic analysis of some expedited forwarding networks," in *Proceedings of Infocom 2002*, New-York, June 2002.

[13] P. Hurley, *The Provision of a Low Delay Service within the Best-effort Internet*, Ph.D. thesis, Swiss Federal Institute of Technology (EPFL), December 2001.

[14] A. Charny, P. Krishna, N. Patel, and R. Simcoe, "Algorithms for providing bandwidth and delay guarantees in input-buffered crossbars with speedup," in *Proc. IWQOS'98*, 1998.

[15] P. Krishna, N. Patel, A. Charny, and R. Simcoe, "On the speedup required for work-conserving crossbar switches," *IEEE J.Sel. Areas in Communications*, vol. 17 no 6, pp. 1057–1066, June 1999.

[16] J. M. Blanquer and B. Ozden, "Fair queuing for aggregated multiple links," in *Proc. Sigcomm 2001*, September 2001.

[17] V. Jacobson, K. Nichols, and K. Poduri, "An expedited forwarding phb," June 1999, RFC 2598, IETF.

[18] Cruz R. L. and Okino C. M, "Service guarantees for window flow control," in *34th Allerton Conf of Comm., Cont., and Comp. Monticello, IL*, Oct 1996.

[19] Baccelli F., Cohen G., Olsder G. J., , and Quadrat J.-P., *Synchronization and Linearity, An Algebra for Discrete Event Systems*, John Wiley and Sons, 1992.

VIII. Appendix: Proofs

We use the operator notation $A \wedge B := \min(A, B)$ and $A \vee B := \max(A, B)$. Thus, the expression

$$\min\{\max(A, B), \max(C, D)\}$$

can be re-written

$$(A \vee B) \wedge (C \vee D)$$

A. Proof of Lemma II.1

In order to further simplify the notation, we use, locally to this proof, the following convention: first, $\vee$ has precedence over $\wedge$; second, we denote $A \vee B$ with AB. Thus, in this proof only, the expression

$$AB \wedge CD$$

means

$$(A \vee B) \wedge (C \vee D)$$

The reason for this convention is to simplify the use of the distributivity of $\vee$ with respect to $\wedge$ [19], which is here written as

$$A(B \wedge C) = AB \wedge AC$$

Our convention is typical of "min-max" algebra, where min takes the role of addition and max the role of multiplication. Armed with this facilitating notation, the proof becomes simple, but lengthy, calculus. In the rest of the proof we consider some fixed n and drop superscript ^{n}.

For $0 \leq j \leq n-1$, define

$$F_j = f_j + m_{j+1} + ... + m_n$$

and let $F_n = f_n$. Also let $D_0 = d_0 + m_1 + ... + m_n = m_1 + ... + m_n$

First note that for all $j \geq 1$:

$$f_j = (a_j + m_j) \vee [(f_{j-1} + m_j) \wedge (d_{j-1} + m_j)]$$

then, by adding $m_{j+1} + ... + m_n$ to all terms of the right hand side of this equation, we find

$$F_j = A_j \vee (F_{j-1} \wedge D_{j-1})$$

or, with our notation:

$$F_j = A_j (F_{j-1} \wedge D_{j-1})$$

and by distributivity:

$$F_j = A_j F_{j-1} \wedge A_j D_{j-1} \qquad (12)$$

Now we show by downwards induction on $j = n-1, ..., 0$ that

$$
\begin{aligned}
f_n = \quad & A_n A_{n-1}...A_{j+1} F_j \\
\wedge \quad & A_n A_{n-1}...A_{j+1} D_j \\
\wedge \quad & ... \\
\wedge \quad & A_n A_{n-1}...A_{k+1} D_k \\
\wedge \quad & ... \\
\wedge \quad & A_n A_{n-1} D_{n-2} \\
\wedge \quad & A_n D_{n-1} \qquad (13)
\end{aligned}
$$

where k ranges from j to $n-1$. For $j = n-1$, the property follows from Equation (12) applied for $j = n$. Assume now that Equation (13) holds for some $j \in \{1, ..., n-1\}$. By Equation (12), we have

$$
\begin{aligned}
A_n A_{n-1}...A_{j+1} F_j = \\
A_n A_{n-1}...A_{j+1}(A_j F_{j-1} \wedge A_j D_{j-1})
\end{aligned}
$$

thus

$$
\begin{aligned}
A_n A_{n-1}...A_{j+1} F_j = \\
A_n A_{n-1}...A_{j+1} A_j F_{j-1} \wedge A_n A_{n-1}...A_{j+1} A_j D_{j-1}
\end{aligned}
$$

which, combined with Equation (13) for j shows the property for $j - 1$.

Now we apply Equation (13) for $j = 0$ and find

$$
\begin{aligned}
f_n = A_n A_{n-1}...A_1 F_0 \wedge A_n A_{n-1}...A_1 D_0 \wedge ... \\
\wedge A_n A_{n-1} D_{n-2} \wedge A_n D_{n-1}
\end{aligned}
$$

First note that $F_0 = D_0$ so we can remove the first term in the right hand side of the previous equation. Second, it follows from $a_1 \geq 0$ that $D_0 \leq A_1$ thus

$$A_n A_{n-1}...A_1 D_0 = A_n A_{n-1}...A_1$$

thus finally

$$
\begin{aligned}
f_n = A_n A_{n-1}...A_1 \wedge A_n A_{n-1}...A_2 D_1 \wedge ... \\
\wedge A_n A_{n-1} D_{n-2} \wedge A_n D_{n-1}
\end{aligned}
$$

which is precisely the required formula.

B. Proof of Theorem II.1

First, assume that the packet scale rate guarantee holds. Apply Lemma II.1 with $m_n = \frac{l_n}{r}$. It follows that, for $1 \leq j \leq n-1$.

$$f_n \leq \max \left[A_n^n, A_{n-1}^n, ..., A_{j+1}^n, D_j^n \right]$$

thus f_n is bounded by one of the terms in the right hand side of the previous equation. If it is the last term, we have

$$f_n \leq D_j^n = d_j + \frac{l_{j+1} + ... + l_n}{r}$$

now $d_n \leq f_n + e$, which shows Equation (4). Otherwise, there is some $k \in \{j + 1, ..., n\}$ such that

$$f_n \leq A_k^n = a_k + \frac{l_k + ... + l_n}{r}$$

which shows Equation (5). For $j = 0$, Lemma II.1 implies that

$$f_n \leq \max \left[A_n^n, A_{n-1}^n, ..., A_1^n \right]$$

and the rest follows similarly.

Second, assume conversely that Equation (4) or Equation (5) holds. Consider some fixed n, and define

A_j^n, D_j^n, F_j^n as in Lemma II.1, with $m_n = \frac{l_n}{r}$. For $1 \leq j \leq n-1$, we have

$$d_n - e \leq \max \left[A_n^n, A_{n-1}^n, ..., A_{j+1}^n, D_j^n \right]$$

and for $j = 0$:

$$d_n - e \leq \max \left[A_n^n, A_{n-1}^n, ..., A_1^n \right]$$

thus $d_n - e$ is bounded by the minimum of all right-handsides in the two equations above, which, by Lemma II.1, is precisely f_n.

C. Proof of Theorem V.2

We first introduce some notation (see Figure 1). Call $a_n \geq 0$ the arrival times for the fresh input. Packets are numbered in order of arrival, so $0 \leq a_1 \leq a_2 \leq ....$ Let l_n be the size of packet n. Call b_n the arrival time for packet n at the second component; b_n is not assumed to be monotonic with n, but for all n:

$$a_n \leq b_n \leq a_n + \delta \qquad (14)$$

Also call d_n the departure time of packet n from the second component. By convention, $a_0 = d_0 = 0$.

Then, define

$$e_1 = e + \delta_{\max} + \sup_{t \geq 0} [\frac{\alpha(t+\delta) - l_{\min}}{r} - t]$$

and

$$e_2 = e + \delta_{\max} + \sup_{0 \leq t \leq \delta} [\frac{\alpha(t) + \alpha(\delta) - l_{\min}}{r} - t]$$

so that $e' = \min[e_1, e_2]$. It is sufficient to show that the combined node separately satisfies the packet scale rate guarantee with rate r and with latencies e_1 and e_2. To see why, define f_n by Equation (2). If $d_n - f_n \leq e_1$ and $d_n - f_n \leq e_2$ for all n, then $d_n - f_n \leq e'$.

Part 1: We show that the combined node satisfies the packet scale rate guarantee with rate r and latency e_1.

An arrival curve for the input traffic to the second component is $\alpha_2(t) = \alpha(t+\delta)$. Thus, by Corollary IV.1, $d_n \leq b_n + D_2$, with

$$d_n \leq b_n + e + \sup_{t \geq 0} [\frac{\alpha(t+\delta)}{r} - t]$$

By Equation (14):

$$d_n - a_n \leq e + \delta_{\max} + \sup_{t \geq 0} [\frac{\alpha(t+\delta)}{r} - t]$$

Now we apply Lemma V.1 which ends the proof for this part.

Part 2: We show that the combined node satisfies the packet scale rate guarantee with rate r and latency e_2.

Let $\delta_{\min} = \delta_{\max} - \delta$ the constant part of the delay. We do the proof for $\delta_{\min} = 0$ since we can eliminate the constant delay by observing packets $\delta_{\min}$ time units after their arrival, and adding $\delta_{\min}$ to the overall delay.

Part 2A:

We assume in this part that there cannot be two arrivals at the same instant; in part 2B, we will show how to relax this assumption.

For a time interval $(s, t]$ (resp. $[s, t]$), define $A(s, t]$ as the total number of bits at the fresh input during the interval $(s, t]$ (resp. $[s, t]$); similarly, define $B(s, t]$ and $B[s, t]$ at the input of the second node. We have the following relations:

$$A(s, t] = \sum_{n \geq 1} 1_{\{s < a_n \leq t\}} l_n \ , \ A[s, t] = \sum_{n \geq 1} 1_{\{s \leq a_n \leq t\}} l_n$$

$$B(s, t] = \sum_{n \geq 1} 1_{\{s < b_n \leq t\}} l_n \ , \ B[s, t] = \sum_{n \geq 1} 1_{\{s \leq b_n \leq t\}} l_n$$

Note that

$$A(a_j, a_n] = \sum_{i=j+1}^{n} l_i$$

but, by lack of FIFO assumption, there is no such relation for B.

By definition of an arrival curve, we have $A(s, t] \leq \alpha(t - s)$.

Lemma VIII.1: For $0 \leq t, u$ and $0 \leq v \leq t$, if there is an arrival at t, then $A(t, t + u] \leq \alpha(u) - l_{\min}$ and $A[t - v, t) \leq \alpha(v) - l_{\min}$

Proof: First note that $A[t, t+u] \leq \inf_{\epsilon > 0} A(t-\epsilon, t+u] \leq \inf_{\epsilon > 0} \alpha(u + \epsilon) = \alpha(u)$ (the last equality is because α is continuous).

Second, let l be the packet length for one packet arriving at time t. Then $A(t, t + u] + l \leq A[t, t + u] \leq \alpha(u)$ thus $A(t, t + u] \leq \alpha(u) - l \leq \alpha(u) - l_{\min}$. The same reasoning shows the second inequality in the lemma. $\diamond$

Now we apply Theorem II.1. Consider some fixed packets numbers $0 \leq j < n$. We have to show that one of the following holds:

$$d_n \leq e_2 + d_j + \frac{A(a_j, a_n]}{r} \qquad (15)$$

or there is some $k \in \{j + 1, ..., n\}$ such that

$$d_n \leq e_2 + a_k + \frac{A[a_k, a_n]}{r} \qquad (16)$$

(Case 1:) Assume that $b_j \geq b_n$. Since the second node is FIFO, we have

$$d_n \leq d_j$$

and thus Equation (15) trivially holds.

(Case 2:) Assume that $b_j < b_n$. By Theorem II.1 applied to the second node, we have

$$d_n \le e + d_j + \frac{1}{r}B(b_j, b_n]$$ (17)

or there exists some k such that $b_j \le b_k \le b_n$ and

$$d_n \le e + b_k + \frac{1}{r}B[b_k, b_n]$$ (18)

(Case 2a:) Assume that Equation (17) holds. By Equation (14), any packet that arrives at node 2 in the interval $(b_j, b_n]$ must have arrived at node 1 in the interval $(a_j - \delta, b_n] \subset (a_j - \delta, a_n + \delta]$. Thus

$$\begin{aligned} B(b_j, b_n] &\le A(a_j - \delta, a_n + \delta] \\ &\le A(a_j, a_n] + A[a_j - \delta, a_j) + A(a_n, a_n + \delta] \\ &\le A(a_j, a_n] + 2\alpha(\delta) - 2l_{\min} \end{aligned}$$

the last part being due to Lemma VIII.1. Thus

$$\begin{aligned} d_n &\le e + \delta + \frac{\alpha(\delta)}{r} - \delta + \frac{\alpha(\delta)}{r} + d_j \\ &\quad + \frac{1}{r}A(a_j, a_n] - 2l_{\min} \\ &\le e_2 + d_j + \frac{1}{r}A(a_j, a_n] \end{aligned}$$

which shows Equation (15).

(Case 2b:) Assume that Equation (18) holds. Note that we do not know the order of k with respect to j and n. However, in all cases, by Equation (14):

$$B[b_k, b_n] \le A[b_k - \delta, a_n + \delta]$$ (19)

We further distinguish three cases.

(Case 2b1:) $k \le j$:

Define

$$u = a_j - b_k + \delta$$ (20)

By hypothesis, $a_k \le a_j$ and $b_k - \delta \le a_k$ so that $u \ge 0$. Note also that $a_j \le b_j \le b_k$ and thus $u \le \delta$.

By Equation (19):

$$B[b_k, b_n] \le A[b_k - \delta, a_j) + A[a_j, a_n] + A(a_n, a_n + \delta]$$

Now by Lemma VIII.1 $A(a_n, a_n + \delta] \le \alpha(\delta)$ and $A[b_k - \delta, a_j) \le \alpha(u) - l_{\min}$. Thus

$$B[b_k, b_n] \le A[a_j, a_n] + \alpha(u) + \alpha(\delta) - 2l_{\min}$$

Combine with Equation (18), Equation (20) and obtain

$$d_n \le a_j + \frac{A[a_j, a_n]}{r} + e_2$$

which shows that Equation (16) holds.

(Case 2b2:) $j < k \le n$:

Define $u = \delta - b_k + a_k$. By Equation (19)

$$B[b_k, b_n] \le A[a_k, a_n] + \alpha(u) + \alpha(\delta) - 2l_{\min}$$

which shows that

$$d_n \le e_2 + a_k + \frac{1}{r}A[a_k, a_n]$$

(Case 2b3:) $k > n$:

Define $u = \delta - b_k + a_n$. By $b_k \le b_n$ and $b_n \le a_n + \delta$ we have $u \ge 0$. By $b_k \ge a_k$ and $a_k \ge a_n$ we have $u \le \delta$.

Now by Equation (18):

$$d_n \le e + b_k + \frac{1}{r}B[b_k, b_n] = e + \delta - u + a_n + \frac{1}{r}B[b_k, b_n]$$

By Equation (19)

$$\begin{aligned} B[b_k, b_n] &\le A[a_n - u, a_n + \delta] \\ &= A[a_n - u, a_n) + l_n + A(a_n, a_n + \delta] \\ &\le \alpha(u) + l_n + \alpha(\delta) - 2l_{\min} \end{aligned}$$

which shows that

$$d_n \le e_2 + a_n + \frac{l_n}{r}$$

Part 2B: Now it remains to handle the case where packet arrivals at either component may be simultaneous. We assume that packets are ordered at component 2 in order of arrival, with some unspecified mechanism for breaking ties. Packets also have a label which is their order of arrival at the first component; we call (k) the label of the kth packet in this order (see Figure 1 for an illustration).

Call $\mathcal{S}$ the original system. Fix some arbitrary integer N. Consider the truncated system $\mathcal{S}^N$ that is derived from the original system by ignoring all packets that arrive at the first component after time $a_N + \delta$. Call $a_n^N, b_n^N, d_n^N, f_n^N$ the values of arrival, departure, and virtual finish times in the truncated system (virtual finish times are defined by Equation (2)). Packets with numbers $\le N$ are not affected by our truncation, thus $a_n^N = a_n, b_n^N = b_n, d_n^N = d_n, f_n^N = f_n$ for $n \le N$. Now the number of arrival events at either component 1 or 2 in the truncated system is finite; thus we can find a positive number η which separates arrival events. Formally: for any $m, n \le N$:

$$a_m = a_n \text{ or } |a_m - a_n| > \eta$$

and

$$b_m = b_n \text{ or } |b_m - b_n| > \eta$$

Let $\epsilon < \frac{\eta}{2}$. We define a new system, called $\mathcal{S}^{N,\epsilon}$, which is derived from $\mathcal{S}^N$ as follows.

• We can find some sequence of numbers $x_n \in (0, \epsilon)$, $n \le N$ such that: (1) they are all distinct; (2) if the packet labeled m is ordered before the packet labeled n in the order of arrival at the second component, then $x_m < x_n$. Building such a sequence is easy, and any sequence satisfying (1) and (2) will do. For example, take $x_n = \frac{k}{N+1}\epsilon$

where k is the order of arrival of packet n (in other words, $(k) = n$).

- Define the new arrival and departure times by

$$a_n^\epsilon = a_n + x_n \ , \ b_n^\epsilon = b_n + x_n \ , \ d_n^\epsilon = d_n + x_n$$

It follows from our construction that all a_n^ϵ are distinct for $n \le N$, and the same holds for b_n^ϵ. Also, the arrival order of packets at the second component is the same as in the original system.

Thus we have built a new system $\mathcal{S}^{N,\epsilon}$ where all arrivals times are distinct, the order of packets at the second component is the same as in $\mathcal{S}^N$, arrival and departure times are no earlier than in $\mathcal{S}^N$, and differ by at most ϵ.

For $k \le N$, call $F_{(k)}^\epsilon$ the virtual finish times at the second component. By definition:

$$\begin{cases} F_{(0)}^\epsilon = 0 \\ F_{(k)}^\epsilon = \max \left[b_{(k)}^\epsilon, \min \left(d_{(k-1)}^\epsilon, F_{(k-1)}^\epsilon \right) \right] \\ \quad + \frac{l_{(k)}}{r} \text{ for } k \ge 1 \end{cases}$$

and a similar definition holds for $F_{(k)}$ by dropping ϵ. It follows by induction that

$$F_{(k)}^\epsilon \ge F_{(k)}$$

thus

$$d_{(k)}^\epsilon \le d_k + \epsilon \le e + F_{(k)} \le e + F_{(k)}^\epsilon + \epsilon$$

Similarly, $b_k^\epsilon \le a_k^\epsilon + \delta$. This shows that $\mathcal{S}^{N,\epsilon}$ satisfies the assumptions of the theorem, with e replaced by $e + \epsilon$

Thus the conclusion of Part 2A holds for $\mathcal{S}^{N,\epsilon}$. Define now f_n^ϵ by Equation (2) applied to a_n^ϵ and d_n^ϵ. We have:

$$d_n^\epsilon \le f_n^\epsilon + e_2 + \epsilon \qquad (21)$$

It also follows by induction that

$$f_n^\epsilon \le f_n + \epsilon$$

Now $d_n \le d_n^\epsilon$ thus

$$d_n - f_n \le d_n^\epsilon - f_n^\epsilon + \epsilon$$

Combining with Equation (21) gives:

$$d_n - f_n \le e_2 + 2\epsilon$$

Now ϵ can be arbitrarily small, thus we have shown that for all $n \le N$:

$$d_n - f_n \le e_2$$

Since N is arbitrary, the above is true for all n.

D. Proof of Theorem VI.3

We use the same notation and convention as in the proof of Theorem V.2. We can also assume that all packet arrivals are distinct, using the same type of reduction.

Fix some $n \ge 1$; due to Theorem VI.1, it is sufficient to show that there is some $k \in \{1, ..., n\}$ such that

$$d_n \le e_2 + a_k + \frac{l_k + ... + l_n}{r} \qquad (22)$$

By hypothesis, there exists some j such that $b_j \le b_n$ and

$$d_n \le b_j + e + \frac{B[b_j, b_n]}{r} \qquad (23)$$

We cannot assume that $j \le n$; thus, define k as the oldest packet arrived in the interval $[b_j, b_n]$, in other words: $k = \inf\{i \ge 1 : b_j \le b_i \le b_n\}$. Necessarily, we have now $k \le n$.

Any packet that arrives at the second node in $[b_j, b_n]$ must have arrived at node 1 after or with packet k, and before b_n. Thus $B[b_j, b_n] \le A[a_k, b_n]$. Now $b_n \le a_n + \delta$. Thus by Lemma VIII.1 in this appendix:

$$B[b_j, b_n] \le A[a_k, a_n] + A(a_n, b_n]$$
$$\le A[a_k, a_n] + \alpha(\delta) - l_{\min}$$

Also, $b_j \le b_k \le a_k + \delta$ and by Equation (23):

$$d_n \le a_k + e + \delta + \alpha(\delta) + A[a_k, a_n] - l_{\min}$$

which shows Equation (22).