Generalised Processor Sharing networks fed by heavy-tailed traffic flows

Miranda van Uitert*,[1], Sem Borst*,†,‡,

*CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

†Department of Mathematics & Computing Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

‡Bell Laboratories, Lucent Technologies
P.O. Box 636, Murray Hill, NJ 07974, USA

Abstract—**We consider networks where traffic is served according to the Generalised Processor Sharing (GPS) principle. GPS-based scheduling algorithms are considered important for providing differentiated quality of service in integrated-services networks.**

We are interested in the workload of a particular flow i at the bottleneck node on its path. Flow i is assumed to have long-tailed traffic characteristics. We distinguish between two traffic scenarios, (i) flow i generates instantaneous traffic bursts and (ii) flow i generates traffic according to an on/off process. In addition, we consider two configurations of feed-forward networks. First we focus on the situation where other flows join the path of flow i. Then we extend the model by adding flows which may branch off at any node, with cross traffic as a special case. We prove that under certain conditions the tail behaviour of the workload distribution of flow i is equivalent to that in a *two*-node tandem network where flow i is served in isolation at *constant* rates. These rates only depend on the traffic characteristics of the other flows through their average rates. This means that the results do not rely on any specific assumptions regarding the traffic processes of the other flows. In particular, flow i is not affected by excessive activity of flows with 'heavier-tailed' traffic characteristics. This confirms that GPS has the potential to protect individual flows against extreme behaviour of other flows, while obtaining substantial multiplexing gains.

Keywords— **Generalised Processor Sharing (GPS), heavy-tailed traffic, regular variation, Weighted Fair Queueing (WFQ)**

I. Introduction

Integrated-services networks carry a large amount of different services. Each of these services has its own traffic characteristics and requires its own quality of service (QoS) guarantees. This heterogeneity in traffic characteristics and QoS guarantees creates the need for traffic control mechanisms to regulate the usage of network resources. In particular, scheduling mechanisms play an important role in achieving differentiated QoS. One of the most important scheduling algorithms is the Generalised Processor Sharing (GPS) mechanism, which was first studied by Parekh and Gallager [11], [12]. GPS is characterised by two attractive properties, (i) each backlogged flow is guaranteed a minimum service rate and (ii) the excess service rate is redistributed among the backlogged flows in proportion to their minimum service rates. Because of the second property GPS is work-conserving. Commonly-used scheduling mechanisms in packet-switched networks, such as Weighted Fair Queueing (WFQ) and other algorithms [14], are based on GPS.

Achieving differentiated QoS is a challenging task due to the highly bursty traffic characteristics in high-speed communication networks. In contrast to traditional assumptions, the burstiness extends over a wide range of time scales. Statistical data analysis [13], [16] has in fact shown that traffic patterns may look similar when observed on various time scales. This behaviour is usually referred to as self-similarity. Several studies, e.g. [9], further offered evidence of a closely related property called long-range dependence, which means that correlations in the traffic activity decay slowly over time. These findings caused a fundamental shift in modelling traffic behaviour. Classical models mostly assume traffic processes with a Markovian structure, which are inherently short-range dependent. Recently though, the focus has shifted to traffic processes with long-tailed characteristics, which provide a useful paradigm for modelling long-range dependence and self-similarity. An example of such a model is an on/off process where the on periods are regularly varying with index $-\nu, \nu \in (1, 2)$.

It is not clear to what extent long-tailed traffic may impact the potential for scheduling mechanisms to help achieve differentiated QoS. To be able to guarantee end-to-end QoS, it is particularly relevant to understand to what degree traffic flows may be negatively affected as they traverse the network. Anantharam [1] was one of the first to study the influence of scheduling strategies on the extent to which long-tailed traffic affects network performance. He showed the influence can be significant, depending on whether or not preemption is admissible.

In this study we investigate the impact of long-tailed traffic on performance in GPS networks. Existing work on GPS networks is largely restricted to a deterministic setting. In [12] Parekh and Gallager show that the first GPS property, minimum guaranteed rates, translates into worst-case bounds on delay and workload for leaky bucket controlled traffic flows. It is clear that the second GPS property, work conservation, yields statistical multi-

[1] Also with KPN Research, P.O. Box 421, 2260 AK Leidschendam, The Netherlands.

plexing gains. In order to quantify these gains however, and to examine how they are possibly influenced by the occurrence of long-tailed traffic, a stochastic analysis of GPS networks is required.

Networks of fluid flows seem to defy exact analysis for all but a few specific cases, and in particular we are not familiar with any stochastic analysis of GPS networks. In the present paper we specifically focus on GPS networks fed by several traffic flows, of which at least one has long-tailed traffic characteristics. Under certain conditions we show that the tail distribution of the workload of the long-tailed flow at the bottleneck node on its path is equivalent to that in a *two*-node tandem network where it is served in isolation at *constant* rates. These rates are the service rates of the two bottleneck nodes for the long-tailed flow in the original network, reduced by the average traffic intensities of the other flows. Hence, the long-tailed flow is only affected by the traffic characteristics of the other flows through their average rates and is not influenced by excessive behaviour of any of the other flows. This result extends the results in Borst, Boxma and Jelenković [**?**], [4] for a single GPS node fed by traffic with long-tailed characteristics. Agrawal, Makowski and Nain [2] establish a similar reduced-load equivalence result for a fluid queue fed by a flow with subexponentially distributed on periods and a general light-tailed flow.

The remainder of this paper is organised as follows. In the next section we consider a simple two-node tandem network, which is fed by a single flow. As alluded to above, this model will play a key role in analysing more complex network configurations. We relate the tail behaviour of the busy-period distribution at node 1 to the arrival process. Then we determine the tail behaviour of the workload distribution at node 2 in terms of the residual busy-period distribution at node 1. Two traffic processes are considered, (i) a traffic flow generating instantaneous bursts and (ii) a traffic flow behaving according to an on/off process. We describe the GPS mechanism in more detail in Section III. In Sections IV and VI we extend the model of Section II to a GPS tandem network that is fed by multiple flows. We consider two network configurations: in Section IV we assume that all flows which are served at node 1 proceed to node 2, while in Section VI we allow for flows which are only served at node 1. In both sections we determine an upper and a lower bound for the workload distribution of the long-tailed flow at node 2. In Section V we prove a general lemma which shows that the lower and upper bounds for the workload distribution asymptotically coincide. We use this lemma to derive the asymptotics for the other models in this paper as well. In the subsequent sections we extend the analysis to more general GPS networks with the long-tailed flow traversing more than two nodes.

II. Two-node network fed by a single flow

In this section we consider a simple tandem network, which is fed by a single flow. We analyse the tail behaviour of the workload distribution at the first and second node. Admittedly, this model represents the simplest possible network scenario, but it plays a central role in the further analysis. We need the results concerning the tail behaviour of the workload distribution in this tandem network to analyse more general networks, where multiple flows share the capacity according to the GPS principle. Surprisingly, it turns out that in the GPS networks that we consider, the tail behaviour of the workload distribution of an individual flow is equivalent to that in a tandem network where the flow is served in isolation at constant rates.

First we introduce some notation. Denote by d_1 and d_2 the constant service rates at node 1 and node 2, respectively. We assume $d_1 > d_2$ to exclude the trivial case where the workload at node 2 is always zero. We define ρ to be the traffic intensity, i.e., the mean amount of traffic offered to the network per unit of time. For stability we assume $\rho < d_2$. Denote by $A(s,t)$ the amount of traffic generated during the time interval $(s,t]$. We define $W^c(t)$ to be the workload at time t if the flow were fed into a queue of rate c,

$$W^c(t) := \sup_{0\le s\le t}\{A(s,t) - c(t-s)\},$$

assuming $W^c(0) = 0$. For $c > \rho$, W^c is a stochastic variable with the limiting distribution of $W^c(t)$ for $t \to \infty$. We define P to be the busy period in this queue. Observe that the total workload in the tandem network at time t is $W^{d_2}(t)$, while the workload at node 1 is $W^{d_1}(t)$. Thus the workload at node 2 at time t is

$$W^{d_1,d_2}(t) := W^{d_2}(t) - W^{d_1}(t), \tag{1}$$

assuming the system is empty at time 0. For $d_2 > \rho$, let W^{d_1,d_2} be a stochastic variable with the limiting distribution of $W^{d_1,d_2}(t)$ for $t \to \infty$.

For any two real functions $f(\cdot)$ and $g(\cdot)$, we use the notational convention $f(x) \sim g(x)$ to denote $\lim_{x\to\infty} f(x)/g(x) = 1$, or equivalently, $f(x) = g(x)(1 + o(1))$ as $x \to \infty$. For any stochastic variable X with distribution function $F(\cdot)$ and $\mathbb{E}X < \infty$, denote by $F^r(\cdot)$ the distribution function of the residual lifetime of X, i.e., $F^r(x) = \frac{1}{\mathbb{E}X}\int_0^x (1 - F(y))\mathrm{d}y$, and by X^r a stochastic variable with that distribution.

The classes of *long-tailed, subexponential, regularly varying*, and *intermediately regularly varying* distributions are denoted with the symbols $\mathcal{L}, \mathcal{S}, \mathcal{R}$ and $\mathcal{IR}$, respectively. See [15] for the detailed definitions of these classes and [6] for general background on heavy-tailed distributions.

We now state some results for the distribution of the workload and the busy period at a single node. We need these results to determine the asymptotic behaviour of W^{d_1,d_2}, and later that of the workload in more general networks.

A. Instantaneous arrivals

Suppose the flow generates instantaneous traffic bursts according to a Poisson process with rate λ. Let K be the stochastic variable representing the burst size. We assume that the burst size distribution $K(\cdot)$ is intermediately regularly varying with mean κ. The traffic intensity is $\rho = \lambda\kappa$. The following three results play a crucial role in the analysis in subsequent sections.

Theorem II.1 (Pakes [10]) If $K^r(\cdot) \in \mathcal{S}$ and $\rho < c$, then

$$\mathbb{P}(W^c > x) \sim \frac{\rho}{c-\rho}\mathbb{P}(K^r > x).$$

Theorem II.2 (Zwart [17]) If $K(\cdot) \in \mathcal{IR}$ and $\rho < c$, then

$$\mathbb{P}(P > x) \sim \frac{c}{c-\rho}\mathbb{P}(K > x(c-\rho)).$$

The above theorem immediately gives the tail distribution of the residual busy period.

Theorem II.3 (residual busy period) If $K(\cdot) \in \mathcal{IR}$ and $\rho < c$, then

$$\mathbb{P}(P^r > x) \sim \frac{c}{c-\rho}\mathbb{P}(K^r > x(c-\rho)).$$

B. On/off processes

Suppose the flow generates traffic according to an on/off process. We assume the off periods to be exponentially distributed with mean $1/\lambda$. While on, the flow produces traffic at a constant rate r. Assume the stochastic variable representing the on period K to have an intermediately regularly varying distribution with mean κ. Because the fraction of off time is equal to $p = \frac{1}{1+\lambda\kappa}$, the traffic intensity is equal to $\rho = \frac{\lambda\kappa r}{1+\lambda\kappa}$. The following three results are the analogues of Theorems II.1, II.2 and II.3, respectively.

Theorem II.4 (Jelenković and Lazar [7]) If $K^r(\cdot) \in \mathcal{S}$ and $\rho < c < r$, then

$$\mathbb{P}(W^c > x) \sim p\frac{\rho}{c-\rho}\mathbb{P}(K^r > \frac{x}{r-c}).$$

Theorem II.5 (Boxma and Dumas [5], [17]) If $K(\cdot) \in \mathcal{IR}$ and $\rho < c < r$, then

$$\mathbb{P}(P > x) \sim p\frac{c}{c-\rho}\mathbb{P}(K > \frac{x(c-\rho)}{r-c}).$$

The following theorem immediately follows from Theorem II.5.

Theorem II.6 (residual busy period) If $K(\cdot) \in \mathcal{IR}$ and $\rho < c < r$, then

$$\mathbb{P}(P^r > x) \sim p\frac{c}{c-\rho}\mathbb{P}(K^r > \frac{x(c-\rho)}{r-c}).$$

C. Workload distribution

The above results completely specify the tail behaviour of the workload distribution at node 1. Moreover, we can use them to analyse the workload distribution at node 2. Observe that the input process at node 2 is an on/off process with as on periods the busy periods at node 1. The on rate is equal to the service rate at node 1, d_1. The off periods correspond to the idle periods at node 1, which are exponentially distributed. In addition, the on and off periods at node 2 are independent.

For both traffic scenarios the tail distribution of the residual busy period at node 1 is intermediately regularly varying. Hence, we can apply Theorem II.4 to determine the tail behaviour of the workload distribution at node 2, which is given in the following lemma.

Lemma II.1 (workload second node) If $K(\cdot) \in \mathcal{IR}$, then

$$\mathbb{P}(W^{d_1,d_2} > x) \sim p'\frac{\rho}{d_2-\rho}\mathbb{P}(P^r > \frac{x}{d_1-d_2}),$$

with the fraction of off time $p' = \frac{d_1-\rho}{d_1}$.

In order to derive the workload distribution of flow i in more general networks, we will need three properties of $\mathbb{P}(W^{d_1,d_2} > x)$. These properties are given in the following lemma.

Lemma II.2: For the traffic scenarios described in Subsections II-A and II-B the following three properties hold: (i) for α, β sufficiently small,

$$\lim_{x\to\infty} \frac{\mathbb{P}(W^{d_1+\alpha,d_2+\beta} > x)}{\mathbb{P}(W^{d_1,d_2} > x)} = G(\alpha,\beta), \tag{2}$$

with $\lim_{\alpha,\beta\to 0} G(\alpha,\beta) = 1$;
(ii) for any real y,

$$\lim_{x\to\infty} \frac{\mathbb{P}(W^{d_1,d_2} > x-y)}{\mathbb{P}(W^{d_1,d_2} > x)} = 1; \tag{3}$$

(iii) for any $c > \rho$ there exists a finite constant C such that,

$$\limsup_{x\to\infty} \frac{\mathbb{P}(W^c > x)}{\mathbb{P}(W^{d_1,d_2} > x)} = C < \infty. \tag{4}$$

Proof: Theorems II.3, II.6 and Lemma II.1 have to be used for all properties. In addition, we use for (ii) that $P^r(\cdot) \in \mathcal{IR} \subset \mathcal{L}$ for both traffic scenarios. Finally, for (iii) we use Theorems II.1, II.4 and Lemma II.1, and the fact that $K^r(\cdot) \in \mathcal{IR}$. ∎

III. PRELIMINARIES

In the next sections we extend the model which we described in the previous section. We consider again a two-node tandem network, but now fed by multiple flows, where traffic is scheduled according to the GPS mechanism. We focus on the workload distribution of a particular flow i which passes through both nodes. In this section we introduce the notation which we use throughout the paper and we explain how the GPS mechanism operates. Although the network that we consider in Sections IV and VI has only two nodes, we introduce notation for networks where flow i traverses N nodes.

At each node of the network, traffic is served according to the GPS mechanism which operates as follows. Define c_n to be the service rate of node n and $S^{(n)}$ to be the set of all flows that receive service at node n, $n = 1, \ldots, N$. Each flow $q \in S^{(n)}$ is assigned a weight $\hat{\phi}_{q,n}$. If every flow at node n is backlogged at time t, then flow $q \in S^{(n)}$ is served at node n at rate

$$\phi_{q,n} := \frac{\hat{\phi}_{q,n}}{\sum_{q\in S^{(n)}} \hat{\phi}_{q,n}} c_n.$$

If some of the flows that are served at node n are not backlogged at time t, then the excess service rate is redistributed among the backlogged flows at node n in proportion to their respective weights. This means that the server always operates at the full service rate when there is work, and thus GPS is work-conserving.

Denote by $A_Q(s,t) := \sum_{q\in Q} A_q(s,t)$ the amount of traffic generated by flows $q \in Q$ in the time interval $(s,t]$, and denote by $A_{q,n}(s,t)$ the amount of traffic that arrives at node n originating from flow q during $(s,t]$. In particular, $A_{q,n}(s,t) = A_q(s,t)$ if node n is the first node that flow q feeds into and we define $A_{Q,n}(s,t) := \sum_{q\in Q} A_{q,n}(s,t)$. Let $B_{q,n}(s,t)$ be the amount of service that flow q receives at node n during the time interval $(s,t]$. Define $V_{q,n}(t)$ as the workload of flow q at node n at time t, and $V_{q,n}$ as a stochastic variable with the limiting distribution of $V_{q,n}(t)$ for $t \to \infty$ (assuming it exists). Similarly, we define $V_{Q,n}(t) := \sum_{q \in Q} V_{q,n}(t)$ and we denote by $V_n(t) := \sum_{q \in S^{(n)}} V_{q,n}(t)$ the total workload at node n at time t.

Using the above definitions, the following identity relation holds for $0 \le s \le t$,

$$V_{q,n}(t) = A_{q,n}(s,t) + V_{q,n}(s) - B_{q,n}(s,t). \quad (5)$$

Using (5), the following relation exists between the arrival processes at two successive nodes,

$$A_{q,n+1}(s,t) = B_{q,n}(s,t) = A_{q,n}(s,t) + V_{q,n}(s) - V_{q,n}(t). \quad (6)$$

The total workload at node n at time t is given by,

$$V_n(t) = \sup_{0 \le s \le t} \{A_{S^{(n)},n}(s,t) - c_n(t-s)\}, \quad (7)$$

assuming that $V_n(0) = 0$.

We define ρ_q to be the average rate of flow q and $\rho_Q := \sum_{q \in Q} \rho_q$ to be the aggregate average rate of all flows $q \in Q$. Let $W_Q^c(t)$ be the workload at time t in a queue with service rate $c \ge 0$ which is fed by the flows $q \in Q$. Then, for $c > \rho_Q$, W_Q^c is a stochastic variable with the limiting distribution of $W_Q^c(t)$ for $t \to \infty$. Analogously we denote by $W_Q^{d_1,d_2}(t)$ the workload at time t at node 2 of a tandem network fed by the flows $q \in Q$. For $d_2 > \rho_Q$, $W_Q^{d_1,d_2}$ is a stochastic variable with the limiting distribution of $W_Q^{d_1,d_2}(t)$ for $t \to \infty$.

We make the following two crucial assumptions throughout the remainder of this paper.

Assumption 3.1: We assume for each flow q, $\phi_{q,n} > \rho_q$ for all $n = 1, \ldots, N$.

This assumption implies that each flow is guaranteed a higher rate than its average rate, which ensures stability. Define $\tilde{c}_n := c_n - \rho_{S^{(n)} \setminus \{i\}}$ as the average service rate available at node n for flow i, i.e., the service rate at node n minus the aggregate average rate of all flows $\in S^{(n)}$ other than i.

Assumption 3.2: We assume $\tilde{c}_N < \tilde{c}_n$ for all $n = 1, \ldots, N-1$.

The above assumption implies that node N can be viewed as the bottleneck node for flow i. The next lemma gives an upper bound for $V_{q,n}(t)$ which follows immediately from the GPS discipline. Since this lemma is a special case of Lemma VII.3, we omit the proof.

Lemma III.1 (GPS upper bound workload) For $n \in \{1,2\}$,

$$V_{q,n}(t) \le W_q^{\tilde{\phi}_q}(t),$$

with $\tilde{\phi}_q = \phi_{q,1}$ if $n = 1$ and $\tilde{\phi}_q = \min\{\phi_{q,1}, \phi_{q,2}\}$ if $n = 2$.

IV. Merging flows

We distinguish between the following two scenarios. In this section we assume the other flows which feed into the network to join the path of flow i, i.e., they are not allowed to leave this path (see Fig. 1). In Section VI flows *are* allowed to leave the path of flow i. The latter model includes cross traffic as a special case.

In particular, we consider the following scenario in this section. We assume the GPS network to be fed by flow i and by two

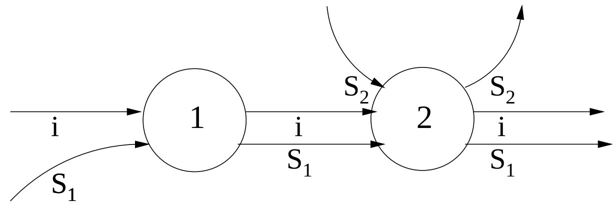

Fig. 1. Two-node network with merging.

additional sets of flows. The set S_1 and flow i feed into node 1 and are served both at nodes 1 and 2, while the set of flows S_2 feed into node 2 and receive only service at this node. We are interested in the distribution of the workload of flow i at node 2, $V_{i,2}$.

In this section we derive both a lower and an upper bound for $\mathbb{P}(V_{i,2} > x)$. The idea can be described as follows. If the flows other than i always showed exactly average behaviour, then $V_{i,2}$ would in distribution be equal to $W_i^{\tilde{c}_1,\tilde{c}_2}$. In reality, stochastic fluctuations in the activity of the other flows will cause $V_{i,2}$ to deviate somewhat from $W_i^{\tilde{c}_1,\tilde{c}_2}$. Accordingly, the bounds will relate $V_{i,2}$ to $W_i^{\tilde{c}_1,\tilde{c}_2}$ with some additional correction terms. In the subsequent section, we show that these terms can be neglected asymptotically, resulting in the exact workload asymptotics.

In both the upper and lower bound for $V_{i,2}(t)$ we need a manageable expression for the total workload at node 2. The following lemma, which follows immediately from (6) and (7), provides such an expression.

Lemma IV.1 (alternative expression $V_2(t)$)

$$\begin{aligned} V_2(t) = & \sup_{0 \le s \le t} \{A_i(s,t) + A_{S_1}(s,t) + A_{S_2}(s,t) \\ & + V_1(s) - c_2(t-s)\} \\ & - \sup_{0 \le s \le t} \{A_i(s,t) + A_{S_1}(s,t) - c_1(t-s)\}. \end{aligned}$$

Before presenting the lower and upper bound, we introduce an additional variable. For $c < \rho_Q$, U_Q^c is defined to be a stochastic variable with the limiting distribution of $U_Q^c(t)$ for $t \to \infty$, with

$$U_Q^c(t) = \sup_{0 \le s \le t} \{c(t-s) - A_Q(s,t)\}. \quad (8)$$

Throughout the analysis, we use the following properties of the sup operator,

$$\sup_t \{f(t) + g(t)\} \le \sup_t \{f(t)\} + \sup_t \{g(t)\}, \quad (9)$$

which also implies

$$\sup_t \{f(t) + g(t)\} \ge \sup_t \{f(t)\} - \sup_t \{-g(t)\}. \quad (10)$$

Lemma IV.2 (lower bound $\mathbb{P}(V_{i,2} > x)$) For any $\delta > 0$, $\epsilon > 0$ sufficiently small and any y,

$$\mathbb{P}(V_{i,2} > x) \ge \mathbb{P}(W_i^{\tilde{c}_1 - \epsilon, \tilde{c}_2 + 2\delta} > x + y) \mathbb{P}(Y^{\delta,\epsilon} \le y),$$

with $Y^{\delta,\epsilon}$ a stochastic variable with the limiting distribution of $Y^{\delta,\epsilon}(t)$ for $t \to \infty$, where

$$Y^{\delta,\epsilon}(t) := U_{S_1}^{\rho_{S_1} - \delta}(t) + U_{S_2}^{\rho_{S_2} - \delta}(t) + W_{S_1}^{\rho_{S_1} + \epsilon}(t)$$

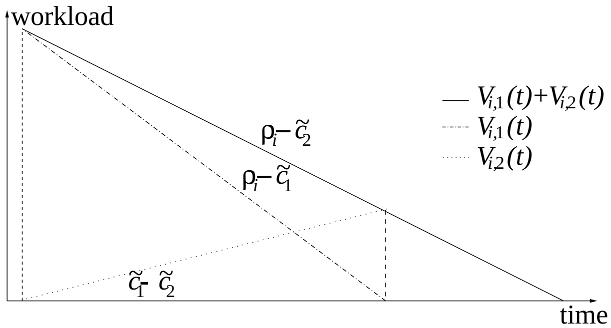

Fig. 2. Overflow scenario instantaneous traffic bursts.

$$+ \sum_{q \in S_1} W_q^{\tilde{\phi}_q}(t) + \sum_{q \in S_2} W_q^{\tilde{\phi}_q}(t).$$

The stochastic variable $Y^{\delta,\epsilon}$ can be seen as the 'correction term' mentioned earlier, accounting for scenarios where $V_{i,2}(t)$ is smaller than $W_i^{\tilde{c}_1-\epsilon,\tilde{c}_2+2\delta}(t)$.

Proof: The proof follows by writing $V_{i,2}(t) = V_2(t) - V_{S_1,2}(t) - V_{S_2,2}(t)$ and then using Lemmas III.1, IV.1 and (9), (10) to obtain $V_{i,2}(t) \geq W_i^{\tilde{c}_1-\epsilon,\tilde{c}_2+2\delta}(t) - Y^{\delta,\epsilon}(t)$. Then use independence of $Y^{\delta,\epsilon}$ and the traffic process of flow i. ∎

Lemma IV.3 (upper bound $\mathbb{P}(V_{i,2} > x)$) For any $\eta > 0$, $\nu > 0$ sufficiently small and any y,

$$\begin{aligned} \mathbb{P}(V_{i,2} > x) \quad &\leq \quad \mathbb{P}(W_i^{\tilde{c}_1+\eta,\tilde{c}_2-2\nu} > x - y) \\ &+ \quad \mathbb{P}(W_i^{\tilde{\phi}_i} > x)\mathbb{P}(Z^{\eta,\nu} > y), \end{aligned}$$

with $Z^{\eta,\nu}$ a stochastic variable with the limiting distribution of $Z^{\eta,\nu}(t)$ for $t \to \infty$, where

$$Z^{\eta,\nu}(t) := U_{S_1}^{\rho_{S_1}-\eta}(t) + W_{S_1}^{\rho_{S_1}+\nu}(t) + W_{S_2}^{\rho_{S_2}+\nu}(t).$$

Analogously to $Y^{\delta,\epsilon}$ in the lower bound, the stochastic variable $Z^{\eta,\nu}$ can be seen as the correction term, corresponding to situations where $V_{i,2}(t)$ is larger than $W_i^{\tilde{c}_1+\eta,\tilde{c}_2-2\nu}(t)$.

Proof: The proof follows by observing $V_{i,2}(t) \leq V_2(t)$ and then using Assumption 3.2, Lemmas III.1, IV.1 and (9), (10) to find

$$V_{i,2}(t) \leq \min\{W_i^{\tilde{\phi}_i}(t), W_i^{\tilde{c}_1+\eta,\tilde{c}_2-2\nu}(t) + Z^{\eta,\nu}(t)\}.$$

Then use independence of $Z^{\eta,\nu}$ and the traffic process of flow i. ∎

V. Tail behaviour workload distribution

We now state our key theorem concerning the tail behaviour of the workload distribution.

Theorem V.1 (asymptotic equivalence) For the traffic scenarios described in Subsections II-A and II-B, under Assumptions 3.1 and 3.2,

$$\mathbb{P}(V_{i,2} > x) \sim \mathbb{P}(W_i^{\tilde{c}_1,\tilde{c}_2} > x),$$

where $\tilde{c}_1$ and $\tilde{c}_2$ represent the total service rate minus the aggregate average rate of all flows other than flow i at nodes 1 and 2 respectively, as defined in Section III.

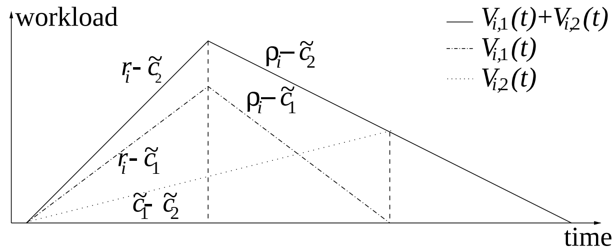

Fig. 3. Overflow scenario on/off process.

According to this theorem, the workload distribution of flow i at node 2 is asymptotically equivalent to that in a tandem network where flow i is served in isolation at rates $\tilde{c}_1$ and $\tilde{c}_2$. Hence, the workload of flow i at node 2 is only affected by the characteristics of the other flows through their average rates, even when the other flows are 'heavier tailed'. This suggests that an extremely large workload of flow i is most likely due to either a long on period or a large burst size of flow i itself. During the subsequent congestion period, the other flows continue to receive service at approximately their average rates. In the theorem this is represented by the constant rates $\tilde{c}_1$ and $\tilde{c}_2$. This result extends the result of [**?**] for the single-node case and shows that GPS is capable of isolating flows in networks as well.

The typical overflow scenario is schematically depicted in Fig. 2. At some point, flow i generates a large burst, causing $V_{i,1}(t)$ to reach some large value. After that, flow i returns to its average behaviour, producing traffic at rate ρ_i. Consequently, $V_{i,1}(t)$ will start to decrease at roughly rate $\rho_i - \tilde{c}_1$, and $V_{i,2}(t)$ will start to increase approximately at rate $\tilde{c}_1 - \tilde{c}_2$, until $V_{i,1}(t)$ reduces to zero at some point. From then on, $V_{i,1}(t)$ will remain relatively small, and $V_{i,2}(t)$ will also start to decrease, roughly at rate $\rho_i - \tilde{c}_2$, until $V_{i,2}(t)$ becomes zero as well. The corresponding behaviour for an on/off process is illustrated in Fig. 3.

A similar reduced-load equivalence result is obtained in [2] for a flow with subexponential on periods and a general light-tailed flow. In our situation, the other flows need *not* be light-tailed because of the GPS properties. Note however that Assumption 3.1 is crucial. If $\rho_q > \phi_{q,n}$ for some flows q other than flow i, then these flows may not receive service at a stable rate when other flow i generates a large amount of traffic. These flows can take away less capacity than ρ_q. Alternatively, if $\rho_i > \phi_i$, then flow i may not receive service at a stable rate when other flows generate a large amount of traffic. In the latter case, flows with an on period distribution or a burst size distribution which is heavier tailed than that of flow i will potentially affect the workload of flow i, see [3].

The above theorem follows from a general lemma which shows that the bounds of Lemmas IV.2 and IV.3 asymptotically coincide. Before giving this lemma, we first introduce some additional notation. For $a > 0$, $a_1 > a_2 > 0$, let R_i be some stochastic variable. For δ, ϵ, η and $\nu > 0$ let $C_{-i}^{\delta,\epsilon}$ and $D_{-i}^{\eta,\nu}$ also be stochastic variables.

Lemma V.1: If for δ, ϵ, η and $\nu > 0$ sufficiently small and any y,

$$\mathbb{P}(R_i > x) \geq \mathbb{P}(W_i^{a_1-\epsilon,a_2+\delta} > x + y)\mathbb{P}(C_{-i}^{\delta,\epsilon} \leq y), \quad (11)$$

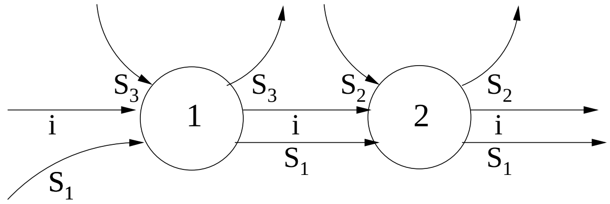

Fig. 4. Two-node network with splitting.

$$\begin{aligned}\mathbb{P}(R_i > x) &\leq \mathbb{P}(W_i^{a_1+\eta, a_2-\nu} > x - y) \\ &+ \mathbb{P}(W_i^a > x)\mathbb{P}(D_{-i}^{\eta,\nu} > y), \end{aligned} \quad (12)$$

and $\mathbb{P}(W_i^a > x)$ and $\mathbb{P}(W_i^{a_1,a_2} > x)$ satisfy Properties (2), (3) and (4), then

$$\mathbb{P}(R_i > x) \sim \mathbb{P}(W_i^{a_1,a_2} > x).$$

Proof: Using the lower bound (11) and Properties (2) and (3), we obtain for any $\delta, \epsilon > 0$ sufficiently small and y,

$$\liminf_{x\to\infty} \frac{\mathbb{P}(R_i > x)}{\mathbb{P}(W_i^{a_1,a_2} > x)} \geq G_i(-\epsilon, \delta)\mathbb{P}(C_{-i}^{\delta,\epsilon} \leq y),$$

which tends to 1 when we let $y \to \infty$ and then $\delta, \epsilon \downarrow 0$. Analogously, using the upper bound (12) and Properties (2), (3) and (4), we have for any $\eta, \nu > 0$ sufficiently small and y,

$$\limsup_{x\to\infty} \frac{\mathbb{P}(R_i > x)}{\mathbb{P}(W_i^{a_1,a_2} > x)} \leq G_i(\eta, -\nu) + C\mathbb{P}(D_{-i}^{\eta,\nu} > y),$$

with $C < \infty$. The first term tends to 1 and the second term vanishes when we let $y \to \infty$ and $\eta, \nu \downarrow 0$. ∎

VI. SPLITTING FLOWS

Consider again a tandem network in which the following flows are served according to the GPS principle (see Fig. 4). As in Section IV, flow i and the set of flows S_1 feed into node 1 and are served both at nodes 1 and 2, and the set of flows S_2 feed into node 2. In addition, we consider in this section the set of flows S_3 which feed into node 1 but do not move on to node 2 after receiving service at node 1.

In the following lemma we give an alternative expression for $V_2(t)$ which follows directly from (6) and (7).

Lemma VI.1 (alternative expression $V_2(t)$)

$$\begin{aligned} V_2(t) &= \sup_{0\leq s\leq t}\{A_i(s,t) + A_{S_1}(s,t) + A_{S_2}(s,t) + \\ &\qquad V_{i,1}(s) + V_{S_1,1}(s) - c_2(t-s)\} + V_{S_3,1}(t) \\ &- \sup_{0\leq s\leq t}\{A_i(s,t) + A_{S_1}(s,t) + A_{S_3}(s,t) - c_1(t-s)\}. \end{aligned}$$

Now we derive both an upper and a lower bound for $\mathbb{P}(V_{i,2} > x)$. These bounds are similar to the bounds in Lemmas IV.2 and IV.3, except that the structure of the correction terms $Y^{\delta,\epsilon}$ and $Z^{\eta,\nu}$ is more complicated due to the presence of the additional set of flows S_3.

For $\delta, \epsilon > 0$, redefine $Y^{\delta,\epsilon}$ to be a stochastic variable with the limiting distribution of $Y^{\delta,\epsilon}(t)$ for $t \to \infty$, with

$$\begin{aligned} Y^{\delta,\epsilon}(t) &:= U_{S_1}^{\rho_{S_1}-\delta}(t) + U_{S_2}^{\rho_{S_2}-\delta}(t) + W_{S_1}^{\rho_{S_1}+\epsilon}(t) \\ &+ W_{S_3}^{\rho_{S_3}+\epsilon}(t) + \sum_{q\in S_1} W_q^{\tilde\phi_q}(t) + \sum_{q\in S_2} W_q^{\tilde\phi_q}(t). \end{aligned}$$

For $\eta, \nu > 0$, redefine $Z^{\eta,\nu}$ to be a stochastic variable with the limiting distribution of $Z^{\eta,\nu}(t)$ for $t \to \infty$, with

$$\begin{aligned} Z^{\eta,\nu}(t) &:= U_{S_1}^{\rho_{S_1}-\eta}(t) + U_{S_3}^{\rho_{S_3}-\nu}(t) + U_{S_3}^{\rho_{S_3}-\eta}(t) \\ &+ \sum_{j=1}^{3} W_{S_j}^{\rho_{S_j}+\nu}(t) + \sum_{q\in S_3} W_q^{\tilde\phi_q}(t). \end{aligned}$$

Lemma VI.2 (lower bound $\mathbb{P}(V_{i,2} > x)$) For any $\delta > 0$, $\epsilon > 0$ sufficiently small and any y,

$$\mathbb{P}(V_{i,2} > x) \geq \mathbb{P}(W_i^{\tilde c_1 - 2\epsilon, \tilde c_2 + 2\delta} > x + y)\mathbb{P}(Y^{\delta,\epsilon} \leq y).$$

Proof: The proof follows by observing that $V_{i,2}(t) = V_2(t) - V_{S_1,2}(t) - V_{S_2,2}(t)$ and then using Assumption 3.2, Lemmas III.1, VI.1 and (9), (10) to obtain $V_{i,2}(t) \geq W_i^{\tilde c_1-2\epsilon,\tilde c_2+2\delta}(t) - Y^{\delta,\epsilon}(t)$. Then use independence. ∎

Lemma VI.3 (upper bound $\mathbb{P}(V_{i,2} > x)$) For any $\eta > 0$, $\nu > 0$ sufficiently small and any y,

$$\begin{aligned}\mathbb{P}(V_{i,2} > x) &\leq \mathbb{P}(W_i^{\tilde c_1+2\eta, \tilde c_2-4\nu} > x - y) \\ &+ \mathbb{P}(W_i^{\tilde\phi_i} > x)\mathbb{P}(Z^{\eta,\nu} > y).\end{aligned}$$

Proof: The proof follows by observing that $V_{i,2}(t) \leq V_2(t)$ and then using Lemmas III.1, VI.1 and (9), (10). Writing $A_{S_3}(r,s) = A_{S_3}(r,t) - A_{S_3}(s,t)$ and splitting some corresponding terms, we then find

$$V_{i,2}(t) \leq \min\{W_i^{\tilde\phi_i}(t), W_i^{\tilde c_1+2\eta,\tilde c_2-4\nu}(t) + Z^{\eta,\nu}(t)\}.$$

Then use independence. ∎

Now we have all the ingredients to use Lemma V.1, which gives the main result of this section.

Theorem VI.1 (asymptotic equivalence) For the traffic scenarios described in Subsections II-A and II-B, under Assumptions 3.1 and 3.2,

$$\mathbb{P}(V_{i,2} > x) \sim \mathbb{P}(W_i^{\tilde c_1,\tilde c_2} > x),$$

where $\tilde c_1$ and $\tilde c_2$ represent the total service rate minus the aggregate average rate of all flows other than flow i at nodes 1 and 2 respectively, as defined in Section III.

VII. PRELIMINARIES GENERAL NETWORKS

In the next two sections we extend the model of Section VI and focus on the Nth node on the path of flow i. We assume this node to be the bottleneck node for flow i. Again we assume the flows to be served at each node according to the GPS mechanism. First we introduce some additional notation and present a number of lemmas which we use in the next sections. Then we analyse the behaviour of the workload of flow i at the bottleneck node on its path, if no other flows feed into any of the nodes on this path. Although this model is quite simple, it provides some useful intuition for the results in Sections VIII and IX.

We define S_j to be the set of flows that feed into node j and S_m^p to be the set of flows that feed into node m and leave the path of flow i at node p (so flows in S_m^p receive service at node p). For $q \in S_m^p$ we define $\tilde\phi_q := \min\{\phi_{q,m}, \ldots, \phi_{q,p}\}$, which is the minimum rate guaranteed to flow q on its path along node m up to and until p.

We now present some lemmas which we use in the next sections. For lack of space, we omit most of the proofs. For details we refer to [15]. The following lemma, which can be proven using induction on $n-m$, gives a lower bound for the amount of service that flow q receives at node n during the time interval $(s,t]$.

Lemma VII.1 (lower bound $B_{q,n}(s,t)$) For $q \in S_m^p$, $1 \le m \le n \le p$ and $\gamma_q \le \tilde{\phi}_q$,

$$B_{q,n}(s,t) \ge \gamma_q(t-s) - \sup_{s \le s_m \le t}\{\gamma_q(s_m - s) - A_q(s,s_m)\}.$$

Using this lemma, we can derive an upper bound for the total workload of flow $q \in S_m^p$ at nodes $m,\ldots,n$. This upper bound is presented in the next lemma.

Lemma VII.2 (upper bound total workload flow q) For $q \in S_m^p$, $1 \le m \le n \le p$ and $\gamma_q \le \tilde{\phi}_q$,

$$\sum_{j=m}^{n} V_{q,j}(t) \le W_q^{\gamma_q}(t).$$

The above lemma immediately implies the following lemma, which includes Lemma III.1 as a special case.

Lemma VII.3 (GPS upper bound workload) For $q \in S_m^p$, $1 \le m \le n \le p$,

$$V_{q,n}(t) \le W_q^{\tilde{\phi}_q}(t).$$

From Lemma VII.2 we can derive an upper bound for the amount of service that flow q receives during the interval $(s,t]$ as well. This upper bound is given in the following lemma.

Lemma VII.4 (upper bound $B_{q,n}(s,t)$) For $q \in S_m^p$, $1 \le m \le n \le p$ and $\gamma_q \le \tilde{\phi}_q$,

$$B_{q,n}(s,t) \le \gamma_q(t-s) + \sup_{0 \le s_m \le t}\{A_q(s_m,t) - \gamma_q(t-s_m)\}.$$

We now briefly discuss the workload behaviour at the Nth node of a network which is fed only by flow i. Take $m^* \in \arg\min_{n=1,\ldots,N-1}\{\tilde{c}_n\}$. In Section III we assumed that $\tilde{c}_n > \tilde{c}_N$ (Assumption 3.2) for all $n = 1,\ldots,N-1$, so that $\tilde{c}_{m^*} > \tilde{c}_N$.

Theorem VII.1 (workload node N)

$$\mathbb{P}(V_{i,N} > x) = \mathbb{P}(W_i^{c_{m^*},c_N} > x).$$

Proof: Observe that, because of the definition of m^*, the total workload at nodes $1,\ldots,m^*$ is equal to that at a node with service rate c_{m^*} which is fed by the original traffic process of flow i. Hence, $\sum_{j=1}^{m^*} V_{i,j}(t) = W_i^{c_{m^*}}(t)$. Since $c_N < c_{m^*}$ (Assumption 3.2) we can apply the same reasoning to the total workload at nodes $1,\ldots,N$ and we have $\sum_{j=1}^{N} V_{i,j}(t) = W_i^{c_N}(t)$. In [8] the following observation is made. If $c_k > c_j$ for $k > j$ then the backlog at node k will always be zero in stationarity and this node can be removed from the tandem network. Because the nodes succeeding node m^* (except N) have a service rate which is larger than c_{m^*},

$$\sum_{j=m^*+1}^{N-1} V_{i,j}(t) = 0$$

and we have, using (1),

$$V_{i,N}(t) = \sum_{j=1}^{N} V_{i,j}(t) - \sum_{j=1}^{m^*} V_{i,j}(t) = W_i^{c_{m^*},c_N}(t).$$

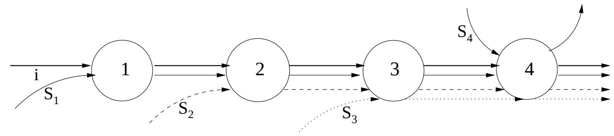

Fig. 5. General network with merging.

■

The workload at node N in this network is equal to that at node 2 in a *two*-node tandem network serving flow i at rates c_{m^*} and c_N. Thus the distribution of the workload is entirely determined by the bottleneck nodes. Intuitively, this can be explained as follows. The workload at a particular node depends on two rates, the rate at which traffic is sent into the node and the rate at which traffic is served by the node. The first rate is the rate at which traffic is served by the bottleneck node on the path to the relevant node, i.e., c_{m^*}. The other rate is the service rate for flow i, which is c_N. Asymptotically, this is still true for the more general networks which we discuss in the next sections.

VIII. General network with merging

Analogously to Sections IV and VI we distinguish between two network scenarios. In this section we consider an extension of the network described in Section IV and assume that each node on the path of flow i in the GPS network is fed by an additional set of flows (see Fig. 5 for the case where flow i traverses 4 nodes). These sets follow the path of flow i and do not leave before node N, the bottleneck node. In Section IX we consider an extension of this network and the network described in Section VI and allow the flows feeding into a node on the path of flow i to leave this path before the bottleneck node.

We first derive bounds for $\mathbb{P}(V_{i,N} > x)$. The idea is similar to that in Section IV. If the flows other than i always showed exactly average behaviour, then $V_{i,n}$ would in distribution be equal to $W_{i,N}^{\tilde{c}_1,\ldots,\tilde{c}_N}$. In Section VII we showed that $W_{i,N}^{\tilde{c}_1,\ldots,\tilde{c}_N}$ has the same distribution as $W_i^{\tilde{c}_{m^*},\tilde{c}_N}$. In addition to $W_i^{\tilde{c}_{m^*},\tilde{c}_N}$, the bounds contain some correction terms accounting for the stochastic fluctuations of the flows other than flow i, which we later show can be asymptotically neglected.

Recall that in the two-node model the upper and lower bounds for $V_{i,2}(t)$ were derived from bounds for $V_1(t)$ and $V_2(t)$. Similarly, in the N-node case, the lower and upper bounds for $V_{i,N}(t)$ rely on bounds for the total workload at each node $n \in \{1,\ldots,N\}$. Define $X_n(t) :=$

$$\sup_{0 \le s_1 \le \ldots \le s_{n+1}=t}\{A_i(s_1,t) + \sum_{j=1}^{n}[A_{S_j}(s_j,t) - c_j(s_{j+1}-s_j)]\}.$$

In the next lemma, which can be proven using induction, we give an expression for $V_n(t)$ in terms of $X_n(t)$. This expression will be used in deriving the upper and lower bounds for $V_{i,N}(t)$.

Lemma VIII.1 (workload node n) For $n \ge 2$,

$$V_n(t) = X_n(t) - X_{n-1}(t).$$

In order to determine a lower and an upper bound for $V_n(t)$ we have to find a lower and an upper bound for $X_n(t)$.

Lemma VIII.2 (lower bound) For any $\theta_1, \ldots, \theta_n$,

$$X_n(t) \geq W_i^e(t) - \sum_{j=1}^{n} U_{S_j}^{\theta_j}(t),$$

with $e := \min_{m=1,\ldots,n}\{c_m - \sum_{j=1}^{m} \theta_j\}$.

Proof: Exploit the fact that $\sum_{j=1}^{n} \theta_j(t - s_j) = \sum_{j=1}^{n} \theta^j(s_{j+1} - s_j)$ with $\theta^j := \sum_{m=1}^{j} \theta_m$, to rewrite $X_n(t)$ and then use (8) and (10). ∎

Lemma VIII.3 (upper bound) For any $\xi_1, \ldots, \xi_n$,

$$X_n(t) \leq W_i^d(t) + \sum_{j=1}^{n} W_{S_j}^{\xi_j}(t),$$

with $d := \min_{m=1,\ldots,n}\{c_m - \sum_{j=1}^{m} \xi_j\}$.

Proof: Similar to that of the lower bound. ∎

We now use the bounds for $X_n(t)$ to construct a lower and an upper bound for $\mathbb{P}(V_{i,N} > x)$. We first introduce some additional notation similar to Section IV. For $\delta, \epsilon > 0$, define $Y^{\delta,\epsilon}$ as a stochastic variable with the limiting distribution of $Y^{\delta,\epsilon}(t)$ for $t \to \infty$, with $Y^{\delta,\epsilon}(t) :=$

$$\sum_{j=1}^{N} U_{S_j}^{\rho_{S_j}-\delta}(t) + \sum_{j=1}^{N-1} W_{S_j}^{\rho_{S_j}+\epsilon}(t) + \sum_{j=1}^{N} \sum_{q \in S_j} W_q^{\tilde{\phi}_q}(t).$$

For $\eta, \nu > 0$, define $Z^{\eta,\nu}$ as a stochastic variable with the limiting distribution of $Z^{\eta,\nu}(t)$ for $t \to \infty$, with

$$Z^{\eta,\nu}(t) := \sum_{j=1}^{N} W_{S_j}^{\rho_{S_j}+\nu}(t) + \sum_{j=1}^{N-1} U_{S_j}^{\rho_{S_j}-\eta}(t).$$

Lemma VIII.4 (lower bound $\mathbb{P}(V_{i,N} > x)$) For any $\delta > 0$, $\epsilon > 0$ sufficiently small and any y,

$$\mathbb{P}(V_{i,N} > x) \geq \mathbb{P}(W_i^{\tilde{c}_{m^*} - m^*\epsilon, \tilde{c}_N + N\delta} > x + y)\mathbb{P}(Y^{\delta,\epsilon} \leq y).$$

Proof: By definition, $V_{i,N}(t) = V_N(t) - \sum_{j=1}^{N} \sum_{q \in S_j} V_{q,N}(t)$. This is lower bounded by $X_N(t) - X_{N-1}(t) - \sum_{j=1}^{N} \sum_{q \in S_j} W_q^{\tilde{\phi}_q}(t)$ using Lemmas VII.3 and VIII.1. Now use the lower bound in Lemma VIII.2 for $X_N(t)$ with $\theta_j = \rho_{S_j} - \delta$ and the upper bound in Lemma VIII.3 for $X_{N-1}(t)$ with $\xi_j = \rho_{S_j} + \epsilon$. Then use independence. ∎

Note that the lower bound for $V_{i,2}(t)$ in Lemma IV.2 is indeed a special case of the lower bound for $V_{i,N}(t)$.

Lemma VIII.5 (upper bound $\mathbb{P}(V_{i,N} > x)$) For any $\eta > 0$, $\nu > 0$ sufficiently small and any y,

$$\begin{aligned} \mathbb{P}(V_{i,N} > x) &\leq \mathbb{P}(W_i^{\tilde{c}_{m^*} + m^*\eta, \tilde{c}_N - N\nu} > x - y) \\ &+ \mathbb{P}(W_i^{\tilde{\phi}_i} > x)\mathbb{P}(Z^{\eta,\nu} > y). \end{aligned} \quad (13)$$

Proof: By definition, $V_{i,N}(t) \leq V_N(t)$. Thus, because of Lemma VIII.1, $V_{i,N}(t) \leq X_N(t) - X_{N-1}(t)$. Analogously to the proof of the lower bound use the upper bound in Lemma VIII.3 for $X_N(t)$ with $\xi_j = \rho_{S_j} + \nu$ and the lower bound in Lemma VIII.2 for $X_{N-1}(t)$ with $\theta_j = \rho_{S_j} - \eta$. Then use independence. ∎

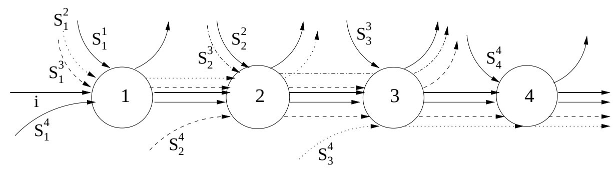

Fig. 6. General network with splitting.

Note that the upper bound for $V_{i,2}(t)$ in Lemma IV.3 is a special case of the upper bound for $V_{i,N}(t)$.

We are now able to characterise the tail behaviour of $\mathbb{P}(V_{i,N} > x)$. It follows immediately from Lemma V.1 and the lower and upper bound given in Lemmas VIII.4 and VIII.5.

Theorem VIII.1 (asymptotic equivalence) For the traffic scenarios described in Subsections II-A and II-B, under Assumptions 3.1 and 3.2,

$$\mathbb{P}(V_{i,N} > x) \sim \mathbb{P}(W_i^{\tilde{c}_{m^*}, \tilde{c}_N} > x),$$

where $\tilde{c}_{m^*}$ and $\tilde{c}_N$ represent the total service rate minus the aggregate average rate of all flows other than flow i at nodes m^* and N, respectively, as defined in Section III.

Remarkably, the workload distribution of flow i at the bottleneck node is asymptotically equivalent to that in a *two*-node tandem network where flow i is served in isolation at constant rates. In Sections V and VI these rates are simply $\tilde{c}_1$ and $\tilde{c}_2$. For the N-node network we have to take the two smallest service rates for flow i when reduced by the aggregate average rates of the other flows, $\tilde{c}_{m^*}$ and $\tilde{c}_N$. Hence, for the network described in this section as well, the workload of flow i at the bottleneck node is only affected by the characteristics of the other flows through their average rates. This suggests that an extremely large workload of flow i at its bottleneck node is most likely due to either a long on period or a large burst of the flow itself and the other flows showing roughly their average behaviour. Consequently, we can consider flow i to be served in isolation at constant rates $\tilde{c}_1, \ldots, \tilde{c}_N$. Following the reasoning of [8] as in the proof of Theorem VII.1 we can remove all nodes with capacity $\tilde{c}_n > \tilde{c}_{m^*}$ after which we are left with a two-node tandem network.

IX. General network with splitting

In this section we extend the model of the previous section and assume that each node on the path of flow i is fed by an additional set of flows, which may leave this path before node N (see Fig. 6 for the case where flow i traverses 4 nodes). We omit most of the proofs and refer to [15] for details.

We first introduce some additional notation. Define $\hat{A}_k^p(s,t)$ to be the amount of work arriving at node k during the interval $(s,t]$ associated with flows entering the path of flow i at node k and passing through node $p \geq k$, i.e.,

$$\hat{A}_k^p(s,t) := \sum_{m=p}^{N} A_{S_k^m}(s,t).$$

Similarly we define $V_k^p(t)$ to be the workload at node k at time t associated with flows passing through node $p \geq k$ (including flow i), i.e.,

$$V_k^p(t) := \sum_{j=1}^{k} \sum_{m=p}^{N} V_{S_j^m,k}(t) + V_{i,k}(t).$$

Finally we define $c_k^p(s,t)$ to be the amount of service available in node k during the interval $(s,t]$ for flows passing through node $p \geq k$, i.e.,

$$c_k^p(s,t) := c_k(t-s) - \sum_{j=1}^{k} \sum_{m=k}^{p-1} B_{S_j^m,k}(s,t).$$

The following lemma expresses the workload at node n at time t associated with the flows passing through node p, in terms of $X_n^p(t)$, with $X_n^p(t) :=$

$$\sup_{0 \leq s_1 \leq \ldots \leq s_{n+1}=t} \{A_i(s_1,t) + \sum_{k=1}^{n} [\hat{A}_k^p(s_k,t) - c_k^p(s_k,s_{k+1})]\}.$$

Lemma IX.1 (workload node n) For $2 \leq n \leq p$,

$$V_n^p(t) = X_n^p(t) - X_{n-1}^p(t). \tag{14}$$

Proof: Similar to that of Lemma VIII.1. ■

If $\sum_{j=1}^{k} \sum_{m=k}^{N-1} B_{S_j^m,k}(s,t) = 0$ and we take p equal to N in (14) so that $c_k^p(s_k,s_{k+1}) = c_k(s_{k+1}-s_k)$ for $k = 1, \ldots, N-1$, then we see that it reduces to the result in Lemma VIII.1 where we assumed that flows cannot leave the path of flow i before node N.

As before, we can derive lower and upper bounds for $X_n^p(t)$, using Lemmas VII.1 and VII.4 to obtain bounds for the terms $B_{S_j^m,k}(s_k,s_{k+1})$ occurring in $c_k^p(s_k,s_{k+1})$ in $X_n^p(t)$. These bounds are similar to those in Lemmas VIII.2 and VIII.3. We can then use these bounds to obtain lower and upper bounds for $\mathbb{P}(V_{i,N} > x)$. Define

$$\sigma_k := \sum_{j=1}^{k} \sum_{m=k}^{N} |S_j^m| + 2 \sum_{f=1}^{k-1} \sum_{j=1}^{f} \sum_{m=f}^{N-1} |S_j^m|.$$

Lemma IX.2 (lower bound $\mathbb{P}(V_{i,N} > x)$) For any $\delta, \epsilon > 0$ sufficiently small and any y,

$$\mathbb{P}(V_{i,N} > x) \geq \mathbb{P}(W_i^{\tilde{c}_{m^*} - \epsilon\sigma_{m^*}, \tilde{c}_N + \delta\sigma_N} > x + y) \mathbb{P}(Y^{\delta,\epsilon} \leq y),$$

with $Y^{\delta,\epsilon}$ some random variable independent of the traffic process of flow i.

Proof: $V_{i,N}(t) = V_N(t) - \sum_{j=1}^{N} \sum_{q \in S_j^N} V_{q,N}(t)$, by definition. This is lower bounded by $X_N^N(t) - X_{N-1}^N(t) - \sum_{j=1}^{N} \sum_{q \in S_j^N} W_q^{\tilde{\phi}_q}(t)$ using Lemmas VII.3 and IX.1. Now we can use a lower bound for $X_N^N(t)$ and an upper bound for $X_{N-1}^N(t)$. ■

Lemma IX.3 (upper bound $\mathbb{P}(V_{i,N} > x)$) For any $\eta, \nu > 0$ sufficiently small and any y,

$$\begin{aligned} \mathbb{P}(V_{i,N} > x) &\leq \mathbb{P}(W_i^{\tilde{c}_{m^*} + \eta\sigma_{m^*}, \tilde{c}_N - \nu\sigma_N} > x - y) \\ &+ \mathbb{P}(W_i^{\tilde{\phi}_i} > x) \mathbb{P}(Z^{\eta,\nu} > y), \end{aligned}$$

with $Z^{\eta,\nu}$ some random variable independent of the traffic process of flow i.

Proof: By definition, $V_{i,N}(t) \leq V_N(t) = V_N^N(t)$. Using Lemma IX.1, $V_{i,N}(t) \leq X_N^N(t) - X_{N-1}^N(t)$. Analogously to the proof of Lemma IX.2 we then use an upper bound for $X_N^N(t)$ and a lower bound for $X_{N-1}^N(t)$. ■

The lower and upper bound for $V_{i,N}(t)$ in Lemmas IX.2 and IX.3 reduce to the lower and upper bound in Lemmas VIII.4 and VIII.5, in case we assume that no flows leave the path of flow i, i.e., $S_j^m = \emptyset$ for $m < N$.

We now have gathered all the elements to characterise the tail behaviour of the workload distribution in the most general class of networks that we consider.

Theorem IX.1 (asymptotic equivalence) For the traffic scenarios described in Subsections II-A and II-B, under Assumptions 3.1 and 3.2,

$$\mathbb{P}(V_{i,N} > x) \sim \mathbb{P}(W_i^{\tilde{c}_{m^*}, \tilde{c}_N} > x),$$

where $\tilde{c}_{m^*}$ and $\tilde{c}_N$ represent the total service rate minus the aggregate average rate of all flows other than i at nodes m^* and N, respectively, as defined in Section III.

Again the workload distribution of flow i at the bottleneck node is asymptotically equivalent to that in a *two-node* tandem network where flow i is served in isolation at constant rates.

X. Concluding remarks

In this paper we analysed the workload behaviour under the GPS mechanism in networks fed by multiple flows. Specifically, we considered a particular flow i traversing the network and assumed it to have heavy-tailed traffic characteristics. We showed that the tail behaviour of the workload distribution of flow i at its bottleneck node is equivalent to that in a *two*-node tandem network where flow i is served in isolation at *constant* rates. In case flow i traverses only two nodes and the second node is the bottleneck node, these rates are the service rates in the original network reduced by the average rates of the other flows. However, when flow i traverses more than two nodes, we have to take the rates from the nodes which are bottleneck when the service rate is reduced by the average rates of the other flows. Hence, flow i is only affected by the characteristics of the other flows through their average rates. This suggests that the GPS mechanism is capable of isolating individual flows in networks, even when they have heavy-tailed traffic characteristics, while achieving significant multiplexing gains.

The results in this paper may be extended in several directions. We assumed for each flow the minimal rate guaranteed by the GPS mechanism to be larger than the average input rate. It may be possible to relax this assumption for a certain class of flows as in [4]. In this paper we only considered the workload distribution at nodes with the minimum average service rate for flow i on its path. The tail behaviour of the workload distribution of flow i at a node following the node with the minimal average service rate is an interesting topic for further research.

References

[1] V. Anantharam. Scheduling strategies and long-range dependence. *Queueing Systems* 33 (1-3), 73 – 89, 1999.

[2] R. AGRAWAL, A.M. MAKOWSKI AND P. NAIN. On a reduced load equivalence for fluid queues under subexponentiality. *Queueing Systems* 33 (1-3), 5 – 41, 1999.

[3] S.C. BORST, O.J. BOXMA AND P.R. JELENKOVIĆ. Induced burstiness in Generalized Processor Sharing queues with long-tailed traffic sources. *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, Illinois, 1999.

[4] S.C. BORST, O.J. BOXMA AND P.R. JELENKOVIĆ. Asymptotic behavior of Generalized Processor Sharing with long-tailed traffic flows. *Proceedings of Infocom 2000*, Tel Aviv, 912 – 922, 2000.

[5] O.J. BOXMA AND V. DUMAS. The busy period in the fluid queue. *Performance Evaluation Review* 26, 100 – 110, 1998.

[6] P. EMBRECHTS, C. KLÜPPELBERG AND T. MIKOSCH. *Modelling Extremal Events,* Springer Verlag, Berlin, 1997.

[7] P.R. JELENKOVIĆ AND A.A. LAZAR. Asymptotic results for multiplexing subexponential on-off processes. *Advances in Applied Probability* 31, 394 – 421, 1999.

[8] O. KELLA AND W. WHITT. A tandem fluid network with Lévy input. *Queues and Related Models*, I. Basawa and U. Bhat (eds.), Oxford, Oxford University Press, 112 – 128, 1992.

[9] W.E. LELAND, M.S. TAQQU, W. WILLINGER AND D.V. WILSON. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking* 2, 1 – 15, 1994.

[10] A.G. PAKES. On the tails of waiting-time distributions. *Journal of Applied Probability* 12, 555 – 564, 1975.

[11] 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, 1993.

[12] 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, 1994.

[13] V. PAXSON AND S. FLOYD. Wide area traffic: the failure of Poisson modelling. *IEEE/ACM Transactions on Networking* 3 (3), 226 – 244, 1995.

[14] D. STILIADIS AND A. VARMA. Efficient fair queueing algorithms for packet-switched networks. *IEEE/ACM Transactions on Networking* 6 (2), 175 – 185, 1998.

[15] M.J.G. VAN UITERT AND S.C. BORST. A reduced-load equivalence for Generalised Processor Sharing networks with heavy-tailed input flows. CWI report PNA-R0007, 2000.

[16] W. WILLINGER, M.S. TAQQU, W.E. LELAND AND D.V. WILSON. Self-similarity in high-speed packet traffic: analysis and modeling of Ethernet traffic measurements. *Statistical Science* 10, 67 – 85, 1995.

[17] A.P. ZWART. Tail asymptotics for the busy period in the GI/G/1 queue. Technical report COSOR 99-12 Eindhoven University of Technology, 1999, to appear in *Mathematics of Operations Research* 2001.