Centrum voor Wiskunde en Informatica

REPORT *RAPPORT*

Generalized Processor Sharing Queues with Heterogeneous Traffic Classes

S.C. Borst, M.R.H. Mandjes, M.J.G. van Uitert

Probability, Networks and Algorithms (PNA)

PNA-R0106 May 31, 2001

Report PNA-R0106
ISSN 1386-3711

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

Generalized Processor Sharing Queues with Heterogeneous Traffic Classes

Sem Borst*,†,‡, Michel Mandjes*,†,*, Miranda van Uitert*
email: sem@cwi.nl, michel@cwi.nl, miranda@cwi.nl

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

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

‡*Department of Mathematics & Computer Science*
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

**Faculty of Mathematical Sciences*
University of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands

ABSTRACT

We consider a queue fed by a mixture of light-tailed and heavy-tailed traffic. The two traffic flows are served in accordance with the Generalized Processor Sharing (GPS) discipline. GPS-based scheduling algorithms, such as Weighted Fair Queueing (WFQ), have emerged as an important mechanism for achieving service differentiation in integrated networks.

We derive the asymptotic workload behavior of the light-tailed traffic flow under the assumption that its GPS weight is larger than its traffic intensity. The GPS mechanism ensures that the workload is bounded above by that in an isolated system with the light-tailed flow served in isolation at a constant rate equal to its GPS weight. We show that the workload distribution is in fact asymptotically equivalent to that in the isolated system, multiplied with a certain pre-factor, which accounts for the interaction with the heavy-tailed flow. Specifically, the pre-factor represents the probability that the heavy-tailed flow is backlogged long enough for the light-tailed flow to reach overflow. The results provide crucial qualitative insight in the typical overflow scenario.

2000 Mathematics Subject Classification: 60K25 (primary), 68M20, 90B18, 90B22 (secondary).

Keywords and Phrases: fluid queues, Generalised Processor Sharing (GPS), heavy-tailed traffic, large deviations, queue-length asymptotics, regular variation, Weighted Fair Queueing (WFQ).

Note: Work carried out under the project PNA2.1 "Communication and Computer Networks".

1 Introduction

Integrated networks have important advantages over dedicated networks. In the first place, their return on investment is less sensitive to the popularity of the individual applications. Moreover, new applications can be introduced rapidly in integrated networks. And also, the heterogeneity of the individual traffic flows allows for a higher level of statistical multiplexing, such that less capacity is needed to support all applications.

However, there are intrinsic difficulties in supporting heterogeneous applications on a single network. An important complication is that different applications need different quality of service (QoS), usually expressed in terms of performance metrics as (packet) delay and loss. One could decide to *not* differentiate and treat all traffic flows in the same way, such that all flows receive the QoS needed by the flows with the most stringent requirements. If the peak rates of the individual flows are small compared to the link rate, then this could still lead to a fairly high utilization. However, in the access network, where less flows are multiplexed, this approach will inevitably lead to inefficient use of resources.

An obvious alternative is to pursue differentiation, by defining a number of QoS classes and treating these classes differently. One of the instruments which can be used to accomplish QoS differentiation is the scheduling mechanism. The simplest is strict priority, which has the disadvantage that traffic of the QoS class with the lowest priority can be completely overwhelmed by traffic of the other QoS classes. A popular alternative is Weighted Fair Queueing (see [18], [19]), which is the packet-based version of a fluid mechanism called Generalized Processor Sharing (GPS). With GPS, each class (or individual traffic flow) is assigned a positive weight. Because traffic is served according to these weights, each class is guaranteed a minimum service rate. In addition, the excess capacity is redistributed in proportion to the weights, meaning that GPS is work-conserving.

Besides achieving service differentiation, scheduling mechanisms also play a potential role in controlling the performance impact of bursty traffic. Extensive traffic measurements have shown that burstiness may extend over a wide range of time scales (see [20], [22]). This typically manifests itself in long-range dependence and self-similarity, which can be modeled using fluid models with heavy-tailed arrival processes. Since long-range dependence and self-similarity seem to be an intrinsic feature of certain types of traffic, the issue is not so much trying to eliminate these phenomena, but rather to minimize the impact on the performance of other traffic classes. The FIFO discipline clearly does not accomplish this goal, as smooth traffic would extremely suffer from bursty traffic. It is shown in [15, 23] that if the input exclusively consists of heavy-tailed flows, then the queue distribution 'inherits' the heavy-tailed characteristics. For the situation of heavy-tailed input mixed with light-tailed input, more detailed traffic characteristics determine whether the queue will have a heavy tail or not [6, 8, 23]. In contrast to FIFO, GPS does seem to have the capability to reduce the performance impact of heavy-tailed traffic.

Borst, Boxma and Jelenković [2, 3, 4] analyze GPS queues with heavy-tailed traffic flows. They show a sharp dichotomy in qualitative behavior, depending on the traffic intensities and the relative values of the weight parameters. For certain weight combinations, an individual flow with heavy-tailed traffic characteristics is effectively served at a *constant* rate, which is only influenced by the average rates of the other flows. In particular, the flow is essentially immune from excessive activity of flows with 'heavier'-tailed traffic characteristics. For other weight combinations however, a flow may be strongly affected by the activity of 'heavier'-tailed flows, and may inherit their traffic characteristics. The latter result also holds for light-tailed flows when their traffic intensity exceeds their GPS weight. Unfortunately, the result does not indicate to what extent light-tailed flows are affected by heavy-tailed flows in the more plausible situation when their GPS weight is larger than their traffic intensity.

In the present paper we consider a GPS queue fed by a mixture of light-tailed and heavy-tailed traffic. We derive the asymptotic workload behavior of the light-tailed flow under the assumption that its GPS weight is larger than its traffic intensity. In the analysis, we reduce the space of all possible sample paths to overflow to a single 'most-likely' path which occurs with overwhelming probability, yielding valuable insight in the typical overflow scenario. We examine how the performance experienced by the light-tailed flow is affected by possibly badly behaving heavy-tailed input. In particular, we identify conditions under which the performance of the light-tailed flow does not degrade under the influence of heavy-tailed input.

In some ways, a two-queue system may provide a useful model for integrated-services networks with two traffic classes, which deserves special attention for the following reason. Because of scalability issues, it is practically infeasible to manipulate packets at the granularity level of individual traffic flows in the backbone of any large-scale network. To avoid these complexity problems, traffic flows may instead be aggregated into a small number of classes with roughly similar features, with scheduling acting at the coarser level of aggregate flows.

The remainder of the paper is organized as follows. In Section 2, we present a detailed model description and state some important preliminary results. In Section 3, we provide an overview of the main results of the paper, which characterize the exact asymptotic behavior of the workload distribution of the light-tailed traffic flow.

The subsequent sections are devoted to the detailed proofs. We start in Section 4 with deriving lower and upper bounds for the workload distribution of the light-tailed flow. In Section 5, we proceed to prove some auxiliary results for the light-tailed flow in isolation. Although the bounds seem quite crude by themselves, we show in Section 6 that they asymptotically coincide, yielding the exact asymptotic behavior.

One of the asymptotic terms involves the probability that the heavy-tailed flow is backlogged long enough for overflow to occur. In order to determine the distribution of the backlog period, we first establish in Section 7 some preliminary results for the heavy-tailed flow in isolation. We then compute the specific form of the distribution for various traffic processes in Sections 8, 9 and 10.

2 Model description

We now present a detailed model description. We consider two traffic flows sharing a link of unit rate. Traffic from the flows is served in accordance with the Generalized Processor Sharing (GPS) discipline, which operates as follows. Flow i is assigned a weight ϕ_i, $i = 1, 2$, with $\phi_1 + \phi_2 = 1$. As long as both flows are backlogged, flow i is served at rate ϕ_i, $i = 1, 2$. If one of the flows is not backlogged, however, then the capacity is reallocated to the other flow, which is then served at the full link rate (if backlogged). (It may occur that one of the flows is not backlogged, while generating traffic at some rate $r_i < \phi_i$. In that case, only the *excess* capacity, i.e., $\phi_i - r_i$, is reallocated to the other flow.) Denote by $A_i(s,t)$ the amount of traffic generated by flow i during the time interval $(s,t]$. We assume that the process $A_i(s,t)$ is reversible and has stationary increments. Denote by $V_i(t)$ the backlog (workload) of flow i at time t. Let $\mathbf{V}_i$ be a stochastic variable with as distribution the limiting distribution of $V_i(t)$ for $t \to \infty$ (assuming it exists). Define $B_i(s,t)$ as the amount of service received by flow i during $(s,t]$. Then the following identity relation holds, for all $s \leq t$,

$$V_i(t) = V_i(s) + A_i(s,t) - B_i(s,t). \tag{1}$$

For any $c \geq 0$, denote by $V_i^c(t) := \sup_{s \leq t}\{A_i(s,t) - c(t-s)\}$ the workload at time t in a queue of capacity c fed by flow i. Denote by ρ_i the traffic intensity of flow i (as will be defined in detail below). For $c > \rho_i$, let $\mathbf{V}_i^c$ be a stochastic variable with as distribution the limiting distribution of $V_i^c(t)$ for $t \to \infty$. Then a similar identity relation as above holds, for all $s \leq t$,

$$V_i^c(t) = V_i^c(s) + A_i(s,t) - B_i^c(s,t). \tag{2}$$

In the next two subsections we describe the traffic model that we consider. We first introduce some additional notation. For any two real functions $g(\cdot)$ and $h(\cdot)$, we use the notational convention $g(x) \sim h(x)$ to denote $\lim_{x\to\infty} g(x)/h(x) = 1$, or equivalently, $g(x) = h(x)(1+o(1))$ as $x \to \infty$. We use $f(x) \stackrel{<}{\sim} g(x)$ to denote $\limsup_{x\to\infty} f(x)/g(x) \leq 1$. Also, $f(x) \stackrel{>}{\sim} g(x)$ denotes $\liminf_{x\to\infty} f(x)/g(x) \geq 1$. For any two stochastic variables $\mathbf{X}$ and $\mathbf{Y}$, we write $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$ to denote that they have the same distribution function. For any stochastic variable $\mathbf{X}$ with distribution function $F(\cdot)$, $\mathrm{E}\{\mathbf{X}\} < \infty$, denote by $F^r(\cdot)$ the distribution function of the residual lifetime of $\mathbf{X}$, i.e., $F^r(x) = \frac{1}{\mathrm{E}\{\mathbf{X}\}}\int_0^x (1-F(y))\mathrm{d}y$, and by $\mathbf{X}^r$ 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. The definitions of these classes can be found in Appendix A.

2.1 Traffic model flow 1

We assume that flow 1 is light-tailed. Specifically, we make the assumption that the input process $A_1(s,t)$ is a *Markov-modulated fluid*. Such a process can be described as follows. There is an irreducible Markov chain with a finite state space $\{1,2,\ldots,d\}$. The corresponding transition rate matrix is denoted by $\Lambda := (\lambda_{ij})_{i,j=1,\ldots,d}$, where we follow the convention that $\lambda_{ii} := -\sum_{j\neq i}\lambda_{ij}$. Since the Markov chain is irreducible, there is a unique stationary distribution, which we denote by the vector π. When the source is in state i, traffic is generated (as fluid) at constant rate $R_i < \infty$. Let R be the diagonal matrix with the coefficients R_i on the diagonal. Denote the mean rate by $\rho_1 := \sum_{i=1}^{d}\pi_i R_i$. Denote the peak rate by $R_P := \max_{i=1,\ldots,d} R_i$. It is important to observe that the class of Markov fluid input is closed under superposition, i.e., the superposition of Markov fluid sources can again be modeled as a Markov fluid source.
Results from Kosten [14], Kesidis *et al.* [13], and Elwalid & Mitra [10] yield the following standard properties.

Property 2.1 *Take $\rho_1 < c_1 < R_P$. Then*

- *The moment generating function of traffic generated in an interval of length t is given by, in matrix notation,*

$$\mathrm{E}\{\exp(sA_1(0,t))\} = \pi \exp((\Lambda + sR)t)\mathbf{1},$$

  with $\mathbf{1}$ the all one vector of dimension d.

- *There exists a limiting moment generating function:*

$$\frac{1}{t}\log \mathrm{E}\{\exp(s(A_1(0,t) - c_1 t))\} \to M_{c_1}(s).$$

  This function is continuous and differentiable. It also holds that there is a finite C such that

$$\mathrm{E}\{\exp(s(A_1(0,t) - c_1 t))\} \leq C\mathrm{e}^{M_{c_1}(s)t}.$$

- *The large-buffer asymptotics of a queue with Markov fluid input are given by*

$$\liminf_{x\to\infty} \frac{1}{x}\log \Pr\{\mathbf{V}_1^{c_1} > x\} = -s^*(c_1).$$

  Here $s^*(c_1)$ is the unique positive root of $M_{c_1}(s) = 0$. Moreover, $M'_{c_1}(s^*(c_1)) > 0$.

Although we restrict ourselves to Markov fluid input, we believe that our results are valid for a more general class of light-tailed input. We will comment on this issue in Remark 5.1.

2.2 Traffic model flow 2

We assume that flow 2 is heavy-tailed. We make the assumption that the input process $A_2(s,t)$ is either instantaneous or On-Off, with heavy-tailed burst sizes or On-periods, respectively.

Instantaneous input

Here, flow 2 generates instantaneous traffic bursts according to a renewal process. The interarrival times between bursts have distribution function $U_2(\cdot)$ with mean $1/\lambda_2$. The burst sizes $\mathbf{B}_2$ have distribution function $B_2(\cdot)$ with mean $\beta_2 < \infty$. Thus, the traffic intensity is $\rho_2 := \lambda_2\beta_2$. We assume that $B_2(\cdot)$ is regularly varying of index $-\nu_2$, i.e., $B_2(\cdot) \in \mathcal{R}_{-\nu_2}$ for some $\nu_2 > 1$. The next result which is due to Pakes [17] then yields the tail behavior of the workload distribution of flow 2 in isolation.

Theorem 2.1 *If $B_2^r(\cdot) \in \mathcal{S}$, and $\rho_2 < c$, then*

$$\Pr\{\mathbf{V}_2^c > x\} \sim \frac{\rho_2}{c - \rho_2}\Pr\{\mathbf{B}_2^r > x\}.$$

Fluid input

Here, flow 2 generates traffic according to an On-Off process, alternating between On- and Off-periods. The Off-periods $\mathbf{U}_2$ have distribution function $U_2(\cdot)$ with mean $1/\lambda_2$. The On-periods $\mathbf{A}_2$ have distribution function $A_2(\cdot)$ with mean $\alpha_2 < \infty$. While On, flow i produces traffic at constant rate r_2, so the mean burst size is $\alpha_2 r_2$. The fraction of time that flow 2 is Off is

$$p_2 = \frac{1/\lambda_2}{1/\lambda_2 + \alpha_2} = \frac{1}{1 + \lambda_2 \alpha_2}.$$

The traffic intensity is

$$\rho_2 = (1 - p_2) r_2 = \frac{\lambda_2 \alpha_2 r_2}{1 + \lambda_2 \alpha_2}.$$

We assume that $A_2(\cdot)$ is regularly varying of index $-\nu_2$, i.e., $A_2(\cdot) \in \mathcal{R}_{-\nu_2}$ for some $\nu_2 > 1$. The next result which is due to Jelenković & Lazar [12] then yields the tail behavior of the workload distribution of flow 2 in isolation.

Theorem 2.2 *If $A_2^r(\cdot) \in \mathcal{S}$, and $\rho_2 < c < r_2$, then*

$$\Pr\{\mathbf{V}_2^c > x\} \sim p_2 \frac{\rho_2}{c - \rho_2} \Pr\{\mathbf{A}_2^r > \frac{x}{r_2 - c}\}.$$

3 Overview of the results

In this section we provide an overview of the main results of the paper which characterize the exact asymptotic behavior of $\Pr\{\mathbf{V}_1 > x\}$ as $x \to \infty$. At the end of this section, we present an example. Throughout, we assume $\rho_i < \phi_i$, $i = 1, 2$, which ensures stability of both flows. In addition, we make the assumption that $r_2 > \phi_2$ in case of fluid input of flow 2. Otherwise, the workload of flow 2 would be zero, so the workload of flow 1 would be equal to the total workload $\mathbf{V}$. The tail distribution of the latter quantity has been obtained in [6].

To put things in perspective, we first briefly review the case that $\rho_1 > \phi_1$, while $\rho_1 + \rho_2 < 1$. If either (i) $B_2^r(\cdot) \in \mathcal{IR}$ (instantaneous input of flow 2), or (ii) $A_2^r(\cdot) \in \mathcal{IR}$, $r_2 > \phi_2$ (fluid input), then from [2],

$$\Pr\{\mathbf{V}_1 > x\} \sim \frac{\phi_2 - \rho_2}{\phi_2} \frac{\rho_2}{1 - \rho_1 - \rho_2} \Pr\{\mathbf{P}_2^r > \frac{x}{\rho_1 - \phi_1}\},$$

with $\mathbf{P}_2$ a random variable with as distribution the busy-period distribution in a queue of constant capacity ϕ_2 fed by flow 2.

The above result suggests that the most likely way for flow 1 to build a large queue is that flow 2 generates a large burst, or experiences a long On-period, while flow 1 itself shows roughly average behavior. Note that when flow 2 produces a large amount of traffic, so it becomes backlogged for a long period of time, it receives service at rate ϕ_2. Thus it will experience a busy period as if it were served at constant rate ϕ_2. During that period, flow 1 receives service at rate ϕ_1, while it generates traffic roughly at rate ρ_1, so its queue

will grow approximately at rate $\rho_1 - \phi_1$. When flow 2 is not backlogged, its queue will drain approximately at rate $1 - \rho_1$.
Thus, the backlog of flow 2 behaves as that in a queue of constant capacity $1 - \rho_1$ fed by an On-Off source with as On- and Off-periods the busy and idle periods of flow 2 when served at constant rate ϕ_2, respectively. That is reflected in the above result if we use Theorem 2.2 to interpret the right-hand side.

We now focus on the case $\rho_1 < \phi_1$. Before presenting the main result, we first provide a heuristic derivation of the asymptotic behavior of $\Pr\{\mathbf{V}_1 > x\}$ based on large-deviations arguments, see for instance Anantharam [1]. The overflow scenario described above for the case $\rho_1 > \phi_1$ cannot occur, and now flow 1 too must deviate from its 'normal' behavior in order for the queue to grow. Specifically, large-deviations results suggest that flow 1 must behave as if its traffic intensity is temporarily increased from ρ_1 to some larger value $\hat{\rho}_1$ (as will be specified below). During that time period, flow 2 is continuously backlogged, consuming capacity ϕ_2, thus leaving capacity ϕ_1 for flow 1. (Notice that if flow 2 were not permanently backlogged, then flow 1 would have to show even greater anomalous activity in order for a given backlog level to occur.)
To summarize, the intuitive argument is as follows: a large backlog of level x of flow 1 occurs as a consequence of two rare events: (i). Flow 1 shows similar 'abnormal' behavior as is the typical cause of overflow when served in isolation, thus behaving as if its traffic intensity is increased from ρ_1 to $\hat{\rho}_1$ for a period of time $x/(\hat{\rho}_1 - \phi_1)$. (ii). During that time period, flow 2 is constantly backlogged, demanding capacity ϕ_2, with capacity ϕ_1 remaining for flow 1.
These considerations lead to the following characterization of the asymptotic behavior of $\Pr\{\mathbf{V}_1 > x\}$:

$$\Pr\{\mathbf{V}_1 > x\} \sim \Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1 - \phi_1}\}. \tag{3}$$

The second term represents the probability that flow 2 is continuously backlogged during a period of time $x/(\hat{\rho}_1 - \phi_1)$. Here $\mathbf{T}_2^{1-\rho_1}$ is a stochastic variable with as distribution the limiting distribution of $T_2^{1-\rho_1}(t)$ for $t \to \infty$, with

$$T_2^c(t) := \inf\{u \geq 0 : V_2^c(t) + A_2(t, t+u) - \phi_2 u = 0\}$$

representing the drain time in a queue of capacity ϕ_2 fed by flow 2 with initial workload $V_2^c(t)$.
Thus, the workload distribution is asymptotically equivalent to that in an isolated system, multiplied with a certain pre-factor. The isolated system consists of flow 1 served in isolation at constant rate ϕ_1. The pre-factor represents the probability that flow 2 is backlogged long enough for flow 1 to reach oveflow. The combination of light-tailed and heavy-tailed large deviations is similar to that in the 'reduced-peak equivalence' result derived in Borst & Zwart [6] as well as that for an M/G/2 queue with heterogeneous servers studied in Boxma *et al.* [7].
Note that the general decompositional form of (3) holds irrespective of the detailed traffic characteristics of the two flows. However, the specific form of the two individual terms in (3) *does* depend on the detailed properties of the traffic processes. In particular, we need to distinguish whether flow 2 generates instantaneous or fluid input. In the latter

case, it also depends on whether the peak rate r_2 exceeds $1-\rho_1$ or not.

We now state the main theorem of the paper.

Theorem 3.1 *Suppose that the input process $A_1(s,t)$ satisfies Property 2.1 and that the input process $A_2(s,t)$ is either instantaneous or On-Off, with regularly varying burst sizes or On-periods, respectively. Assume that $\rho_i < \phi_i$, $i = 1, 2$, and $r_2 > \phi_2$ in case of fluid input of flow 2. Then*

$$\Pr\{\mathbf{V}_1 > x\} \sim \Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1 - \phi_1}\},$$

where $\hat{\rho}_1 := M'_{\phi_1}(s^*(\phi_1)) + \phi_1$.
Case I: If $B_2^r(\cdot) \in \mathcal{IR}$ (instantaneous input), then

$$\Pr\{\mathbf{T}_2^{1-\rho_1} > x\} \sim \frac{\rho_2}{1-\rho_1-\rho_2}\Pr\{\mathbf{B}_2^r > x(\phi_2 - \rho_2)\}. \tag{4}$$

Case II-A: If $A_2^r(\cdot) \in \mathcal{IR}$ with $r_2 < 1-\rho_1$ (fluid input), then

$$\Pr\{\mathbf{T}_2^{1-\rho_1} > x\} \sim (1-p_2)\Pr\{\mathbf{A}_2^r > \frac{x(\phi_2-\rho_2)}{r_2-\rho_2}\}. \tag{5}$$

Case II-B: If $A_2^r(\cdot) \in \mathcal{IR}$ with $r_2 > 1-\rho_1$ (fluid input), then

$$\Pr\{\mathbf{T}_2^{1-\rho_1} > x\} \sim p_2\frac{\rho_2}{1-\rho_1-\rho_2}\Pr\{\mathbf{A}_2^r > \frac{x(\phi_2-\rho_2)}{r_2-\rho_2}\}. \tag{6}$$

Noting that $p_2\rho_2 = (1-p_2)(r_2-\rho_2)$, we can observe that in the limiting regime $r_2 \to 1-\rho_1$ cases II-A and II-B coincide. Also, case I can be seen as the limiting case of II-B if we use $r_2\mathbf{A}_2 = \mathbf{B}_2$ and let $r_2 \to \infty$ so that $p_2 \downarrow 1$. In [5] a qualitatively similar result as in case I is derived for a system with two coupled queues, one having heavy-tailed input, the other one exhibiting light-tailed properties.

Before proceeding to the formal proof of Theorem 3.1, we first give an example. Assume flow 1 to behave according to an On-Off process with exponentially distributed On- and Off-periods with means $1/\mu_1$ and $1/\mu_2$, respectively. When the flow is in the On-state, it generates traffic at rate R_1. We assume flow 2 to generate instantaneous input with regularly varying burst sizes of index $-\nu_2$, i.e.,

$$\Pr\{\mathbf{B}_2 > x\} \sim C_2 x^{-\nu_2} l_2(x),$$

with $l_2(\cdot)$ some slowly varying function.
First we determine the deviant traffic intensity $\hat{\rho}_1$ using [16],

$$\hat{\rho}_1 = \frac{\frac{R_1\phi_1^2}{\mu_2}}{\frac{\phi_1^2}{\mu_2} + \frac{(R_1-\phi_1)^2}{\mu_1}}.$$

Using [9], we obtain for flow 1,

$$\Pr\{\mathbf{V}_1^{\phi_1} > x\} \sim \frac{R_1}{\phi_1}\frac{\mu_2}{\mu_1+\mu_2}\exp(-(\frac{\mu_1}{R_1-\phi_1} + \frac{\mu_2}{\phi_1})x).$$

For flow 2, from (4),

$$\Pr\{\mathbf{T}_2^{1-\rho_1} > x\} \sim \frac{\rho_2}{1-\rho_1-\rho_2}\frac{C_2}{\beta_2(\nu_2-1)}(x(\phi_2-\rho_2))^{1-\nu_2} l_2(x(\phi_2-\rho_2)).$$

Now we have all the ingredients for $\Pr\{\mathbf{V}_1 > x\}$.

The next sections are devoted to the formal proof of Theorem 3.1. We start in Section 4 by deriving lower and upper bounds for the workload distribution of flow 1. We then proceed in Section 5 to prove some auxiliary results for flow 1 in isolation. Although the bounds derived in Section 4 seem quite crude by themselves, we show in Section 6 that they asymptotically coincide, yielding the exact asymptotic behavior of $\Pr\{\mathbf{V}_1 > x\}$.
In order to determine the drain time distribution of flow 2 as specified in Theorem 3.1, we first establish in Section 7 some preliminary results for flow 2 in isolation. Note that the specific form of the drain time distribution depends on whether flow 2 generates instantaneous or fluid input. In the latter case, we also need to distinguish whether the peak rate r_2 exceeds $1-\rho_1$ or not. We calculate the drain time distribution for the various cases in Sections 8, 9 and 10.

4 Bounds

In this section we derive lower and upper bounds for the workload distribution of flow 1. The bounds will be instrumental in obtaining the asymptotic behavior of $\Pr\{\mathbf{V}_1 > x\}$ as given in Theorem 3.1.

4.1 Lower bound

We start with a lower bound for the workload distribution of flow 1.

Lemma 4.1 *Suppose there exist $r^* \leq s^* \leq t$ and y such that*

$$A_1(s^*, t) - \phi_1(t - s^*) > x,$$

$$A_1(r^*, s^*) - (\rho_1 - \epsilon)(s^* - r^*) \geq -y,$$

$$\inf_{s^* \leq u \leq t}\{A_2(r^*, u) - (1-\rho_1+\epsilon)(s^* - r^*) - \phi_2(u - s^*)\} \geq y.$$

Then $V_1(t) > x$.

Proof
From (1), for all $s \leq t$,

$$V_1(t) = V_1(s) + A_1(s, t) - B_1(s, t).$$

By definition,

$$B_1(s, t) + B_2(s, t) \leq t - s.$$

Because of the GPS discipline,

$$B_2(s,t) \geq \min\{\phi_2(t-s), V_2(s) + \inf_{s\leq u\leq t}\{A_2(s,u) + \phi_2(t-u)\}\}.$$

Substituting,

$$\begin{aligned} V_1(t) &\geq A_1(s,t) - (t-s) + \min\{\phi_2(t-s), V_1(s) + V_2(s) + \inf_{s\leq u\leq t}\{A_2(s,u) + \phi_2(t-u)\}\} \\ &= A_1(s,t) - \phi_1(t-s) + \min\{0, V_1(s) + V_2(s) + \inf_{s\leq u\leq t}\{A_2(s,u) - \phi_2(u-s)\}\}. \end{aligned}$$

From (1), for all $r \leq s$,

$$V_1(s) + V_2(s) = V_1(r) + V_2(r) + A_1(r,s) + A_2(r,s) - B_1(r,s) - B_2(r,s).$$

By definition,

$$B_1(r,s) + B_2(r,s) \leq s - r.$$

Thus,

$$V_1(s) + V_2(s) \geq A_1(r,s) + A_2(r,s) - (s-r).$$

Substituting,

$$\begin{aligned} V_1(t) &\geq A_1(s,t) - \phi_1(t-s) + \min\{0, A_1(r,s) + A_2(r,s) - (s-r) + \\ &\quad \inf_{s\leq u\leq t}\{A_2(s,u) - \phi_2(u-s)\}\} \\ &= A_1(s,t) - \phi_1(t-s) + \min\{0, A_1(r,s) - (\rho_1 - \epsilon)(s-r) + \\ &\quad \inf_{s\leq u\leq t}\{A_2(r,u) - (1-\rho_1+\epsilon)(s-r) - \phi_2(u-s)\}\} \end{aligned}$$

for all $r \leq s \leq t$.

$\square$

We now translate the above sample-path result into a probabilistic lower bound. We first introduce some additional notation. For any c and $w \geq 0$, define

$$\mathbf{V}_i^c(w) := \sup_{0\leq s\leq w}\{A_i(-s,0) - cs\}.$$

Note that, for $c > \rho_i$, $\mathbf{V}_i^c(\infty) \stackrel{d}{=} \mathbf{V}_i^c$ as defined earlier.
For any c, $v \geq 0$, and y, define

$$\mathbf{T}_2^c(v,y) := \inf\{u \geq 0 : \sup_{0\leq r\leq v}\{A_2(-r,0) - cr\} + A_2(0,u) - \phi_2 u \leq y\}.$$

Thus, $\mathbf{T}_2^c(v,y)$ represents the drain time in a queue of capacity ϕ_2 fed by flow 2 with initial workload $\sup_{0\leq r\leq v}\{A_2(-r,0) - cr\} - y$.
Note that, for $c > \rho_2$,

$$\mathbf{T}_2^c(y) := \mathbf{T}_2^c(\infty,y) = \inf\{u \geq 0 : V_2^c(0) + A_2(0,u) - \phi_2 u \leq y\},$$

and that $\mathbf{T}_2^c(0) \stackrel{d}{=} \mathbf{T}_2^c$ as defined earlier.
Also, define

$$\mathbf{T}_2(y) := \mathbf{T}_2^c(0,y) = \inf\{u \geq 0 : A_2(0,u) - \phi_2 u \leq y\}.$$

(note that the latter quantity does not depend on the value of c), and denote $\mathbf{T}_2 := \mathbf{T}_2(0)$.
Denote

$$P^{\rho_1-\epsilon}(s^*,v,x,y) := \Pr\{\sup_{s^*-v\leq r\leq s^*}\{(\rho_1-\epsilon)(s^*-r) - A_1(r,s^*)\} \leq y \mid A_1(s^*,0)+\phi_1 s^* > x\}.$$

Corollary 4.1 *For any $v \geq 0$ and y,*

$$\Pr\{\mathbf{V}_1 > x\} \geq \Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1}) > x\}\Pr\{\mathbf{T}_2^{1-\rho_1+\epsilon}(v,y) > \frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1}\}P^{\rho_1-\epsilon}(s^*,v,x,y).$$

Proof
Using Lemma 4.1, the independence of $A_1(s,t)$ and $A_2(s,t)$, and the fact that $A_1(s,t)$ and $A_2(s,t)$ have stationary increments, for all $v \geq 0$, $w \geq 0$, and y,

$$\begin{aligned}
& \Pr\{V_1(t) > x\} \\
\geq\ & \Pr\{\exists s^* \in [t-w,t], r^* \in [s^*-v,s^*] : A_1(s^*,t) - \phi_1(t-s^*) > x, \\
& \quad A_1(r^*,s^*) - (\rho_1-\epsilon)(s^*-r^*) \geq -y, \\
& \quad \inf_{s^*\leq u\leq t}\{A_2(r^*,u) - (1-\rho_1+\epsilon)(s^*-r^*) - \phi_2(u-s^*)\} \geq y\} \\
=\ & \Pr\{\exists s^* \in [-w,0], r^* \in [s^*-v,s^*] : A_1(s^*,0) + \phi_1 s^* > x, \\
& \quad A_1(r^*,s^*) - (\rho_1-\epsilon)(s^*-r^*) \geq -y, \\
& \quad \inf_{s^*\leq u\leq 0}\{A_2(r^*,u) - (1-\rho_1+\epsilon)(s^*-r^*) - \phi_2(u-s^*)\} \geq y\} \\
\geq\ & \Pr\{\exists s^* \in [-w,0], r^* \in [s^*-v,s^*] : A_1(s^*,0) + \phi_1 s^* > x, \\
& \quad \inf_{s^*-v\leq r\leq s^*}\{A_1(r,s^*) - (\rho_1-\epsilon)(s^*-r)\} \geq -y, \\
& \quad \inf_{s^*\leq u\leq s^*+w}\{A_2(r^*,u) - (1-\rho_1+\epsilon)(s^*-r^*) - \phi_2(u-s^*)\} \geq y\} \\
\geq\ & \Pr\{\exists s^* \in [-w,0] : A_1(s^*,0) + \phi_1 s^* > x, \inf_{s^*-v\leq r\leq s^*}\{A_1(r,s^*) - (\rho_1-\epsilon)(s^*-r)\} \geq -y, \\
& \quad \exists r^* \in [s^*-v,s^*] : \inf_{s^*\leq u\leq s^*+w}\{A_2(r^*,u) - (1-\rho_1+\epsilon)(s^*-r^*) - \phi_2(u-s^*)\} \geq y\} \\
\geq\ & \Pr\{\exists s^* \in [-w,0] : A_1(s^*,0) + \phi_1 s^* > x\} \\
& \quad \Pr\{\inf_{s^*-v\leq r\leq s^*} A_1(r,s^*) - (\rho_1-\epsilon)(s^*-r) \geq -y \mid A_1(s^*,0) + \phi_1 s^* > x\} \\
& \quad \Pr\{\exists r^* \in [s^*-v,s^*] : \inf_{s^*\leq u\leq s^*+w}\{A_2(r^*,u) - (1-\rho_1+\epsilon)(s^*-r^*) - \phi_2(u-s^*)\} \geq y\} \\
\geq\ & \Pr\{\exists s \in [0,w] : A_1(-s,0) - \phi_1 s > x\} \\
& \quad \Pr\{\sup_{s^*-v\leq r\leq s^*}\{(\rho_1-\epsilon)(s^*-r) - A_1(r,s^*)\} \leq y \mid A_1(s^*,0) + \phi_1 s^* > x\} \\
& \quad \Pr\{\exists r \in [0,v] : \inf_{0\leq u\leq w}\{A_2(-r,u) - (1-\rho_1+\epsilon)r - \phi_2 u\} \geq y\} \\
\geq\ & \Pr\{\sup_{0\leq s\leq w}\{A_1(-s,0) - \phi_1 s\} > x\}P^{\rho_1-\epsilon}(s^*,v,x,y) \\
& \quad \Pr\{\sup_{0\leq r\leq v}\inf_{0\leq u\leq w}\{A_2(-r,u) - (1-\rho_1+\epsilon)r - \phi_2 u\} \geq y\}
\end{aligned}$$

$$
\begin{aligned}
&\geq \Pr\{\mathbf{V}_1^{\phi_1}(w) > x\} P^{\rho_1-\epsilon}(s^*, v, x, y) \\
&\quad \Pr\{\inf_{0\leq u\leq w}\{\sup_{0\leq r\leq v}\{A_2(-r,0) - (1-\rho_1+\epsilon)r\} + A_2(0,u) - \phi_2 u\} \geq y\} \\
&= \Pr\{\mathbf{V}_1^{\phi_1}(w) > x\}\Pr\{\mathbf{T}_2^{1-\rho_1+\epsilon}(v,y) > w\} P^{\rho_1-\epsilon}(s^*, v, x, y).
\end{aligned}
$$

Taking $w = \frac{(1+\alpha)x}{\hat{\rho}_1 - \phi_1}$ completes the proof.

□

4.2 Upper bound

We proceed to derive an upper bound for the workload distribution of flow 1.

Lemma 4.2 *Suppose* $V_1(t) > x$.
Then for all y *there exist* $r^* \leq s^* \leq t$ *such that*

$$
A_1(s^*, t) - \phi_1(t - s^*) > x, \tag{7}
$$

and at least one of the three following events occurs

$$
A_1(r^*, s^*) - (\rho_1 + \epsilon)(s^* - r^*) > y, \tag{8}
$$

or

$$
V_1^{\phi_1}(t) > x + y, \tag{9}
$$

or

$$
\inf_{s^*\leq u\leq t}\{A_2(r^*, u) - (1 - \rho_1 - \epsilon)(s^* - r^*) - \phi_2(u - s^*)\} > -2y. \tag{10}
$$

Proof
First we show that (7) is implied by $V_1(t) > x$. Because of the GPS discipline,

$$
V_1(t) \leq V_1^{\phi_1}(t) = \sup_{s\leq t}\{A_1(s,t) - \phi_1(t-s)\}.
$$

Hence, there exists an $s \leq t$ such that $A_1(s,t) - \phi_1(t-s) > x$. Define

$$
s^* := \inf\{s : A_1(u,t) - \phi_1(t-u) \leq x \; \forall u > s\} = \sup\{s : A_1(s,t) - \phi_1(t-s) > x\}.
$$

We now show that $V_1(t) > x$ implies that at least one of the events (8), (9) or (10) must occur. We distinguish between the following two cases.
i. Flow 1 is continuously backlogged during the interval $[s^*, t]$.
We first show that (a) $V_1(t) > x$ implies that either (9) holds or

$$
\forall u \in [s^*, t] : B_2(s^*, u) - \phi_2(u - s^*) > -y.
$$

Next we show that (b) the latter event implies that either (8) or (10) holds.
(a) We prove that the events

$$
\exists u^* \in [s^*, t] : B_2(s^*, u^*) - \phi_2(u^* - s^*) \leq -y \tag{11}
$$

and

$$\forall q \le s^* \le t : A_1(q,t) - \phi_1(t-q) \le x + y \qquad (12)$$

imply $V_1(t) \le x$.
Since flow 1 is continuously backlogged during $[s^*, t]$,

$$V_1(t) = V_1(s^*) + A_1(s^*, t) - (t - s^*) + B_2(s^*, u^*) + B_2(u^*, t)$$

and

$$B_2(u^*, t) \le \phi_2(t - u^*).$$

Because of the GPS discipline,

$$V_1(s^*) \le \sup_{r \le s^*}\{A_1(r, s^*) - \phi_1(s^* - r)\}.$$

Hence, using (12),

$$\begin{aligned} V_1(t) &\le \sup_{r \le s^*}\{A_1(r, s^*) - \phi_1(s^* - r)\} + A_1(s^*, t) - (t - s^*) + \phi_2(t - s^*) - y \\ &= \sup_{r \le s^*}\{A_1(r, t) - \phi_1(t - r)\} - y \le x + y - y = x, \end{aligned}$$

which is in contradiction with $V_1(t) > x$. Since (12) is the complement of (9), it remains to be shown that (11) implies (8) or (10).
(b) By definition,

$$\begin{aligned} B_2(s^*, u) &\le V_2(s^*) + A_2(s^*, u) \le V(s^*) + A_2(s^*, u) \\ &= \sup_{r \le s^*}\{A_1(r, s^*) + A_2(r, s^*) - (s^* - r)\} + A_2(s^*, u). \end{aligned}$$

Hence,

$$\begin{aligned} & \inf_{s^* \le u \le t}\{B_2(s^*, u) - \phi_2(u - s^*)\} \\ \le\ & \inf_{s^* \le u \le t}\{\sup_{r \le s^*}\{A_1(r, s^*) + A_2(r, s^*) - (s^* - r)\} + A_2(s^*, u) - \phi_2(u - s^*)\} \\ =\ & \sup_{r \le s^*}\inf_{s^* \le u \le t}\{A_1(r, s^*) + A_2(r, u) - (s^* - r) - \phi_2(u - s^*)\} \\ \le\ & \sup_{r \le s^*}\inf_{s^* \le u \le t}\{A_2(r, u) - (1 - \rho_1 - \epsilon)(s^* - r) - \phi_2(u - s^*)\} \\ +\ & \sup_{r \le s^*}\{A_1(r, s^*) - (\rho_1 + \epsilon)(s^* - r)\}. \end{aligned}$$

ii. Flow 1 is not continuously backlogged during the interval $[s^*, t]$.
Thus, there exists a $u \in [s^*, t]$ such that $V_1(u) = 0$. Define $u^* := \sup\{u \in [s^*, t] : V_1(u) = 0\}$.
Then,

$$V_1(t) = A_1(u^*, t) - B_1(u^*, t),$$

combined with

$$B_1(u^*, t) \ge \phi_1(t - u^*)$$

yields

$$V_1(t) \le A_1(u^*, t) - \phi_1(t - u^*).$$

In view of $V_1(t) > x$, we have $A_1(u^*, t) - \phi_1(t - u^*) > x$, which contradicts the definition of s^*.

□

We now use the above sample-path relation to obtain a probabilistic upper bound. Denote

$$Q^{\rho_1+\epsilon}(s^*, x, y) := \Pr\{\sup_{r \le s^*}\{A_1(r, s^*) - (\rho_1 + \epsilon)(s^* - r)\} > y \mid A_1(s^*, 0) + \phi_1 s^* > x\}.$$

Corollary 4.2 *For any y,*

$$\begin{aligned}
\Pr\{\mathbf{V}_1 > x\} \quad &\le \quad \Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1-\epsilon}(-2y) > \frac{(1-\alpha)x}{\hat{\rho}_1 - \phi_1}\} \\
&+ \quad \Pr\{\mathbf{V}_1^{\phi_1} > x + y\} + \Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1-\alpha)x}{\hat{\rho}_1 - \phi_1}) > x\} + \Pr\{\mathbf{V}_1^{\phi_1} > x\}Q^{\rho_1+\epsilon}(s^*, x, y).
\end{aligned}$$

Proof
Using Lemma 4.2, the independence of $A_1(s,t)$ and $A_2(s,t)$, and the fact that $A_1(s,t)$ and $A_2(s,t)$ have stationary increments, for all $w \ge 0$ and y (the numbers indicate the events in the corresponding equations in Lemma 4.2),

$$\begin{aligned}
\Pr\{V_1(t) > x\} \quad &\le \quad \Pr\{(7) \wedge \{(8) \vee (9) \vee (10)\}\} \\
&= \quad \Pr\{(7), (8)\} + \Pr\{(7), (9)\} + \Pr\{(7), (10)\} \\
&\le \quad \Pr\{(7), (8)\} + \Pr\{(9)\} + \Pr\{(7), (10)\} \\
&= \quad \Pr\{\exists r^* \le s^* \le t : A_1(s^*, t) - \phi_1(t - s^*) > x, A_1(r^*, s^*) - (\rho_1 + \epsilon)(s^* - r^*) > y\} \\
&+ \quad \Pr\{V_1^{\phi_1}(t) > x + y\} \\
&+ \quad \Pr\{\exists r^* \le s^* \le t : A_1(s^*, t) - \phi_1(t - s^*) > x, \\
&\qquad \inf_{s^* \le u \le t}\{A_2(r^*, u) - (1 - \rho_1 - \epsilon)(s^* - r^*) - \phi_2(u - s^*)\} > -2y\} \\
&= \quad \Pr\{\exists r^* \le s^* \le 0 : A_1(s^*, 0) + \phi_1 s^* > x, A_1(r^*, s^*) - (\rho_1 + \epsilon)(s^* - r^*) > y\} \\
&+ \quad \Pr\{V_1^{\phi_1}(0) > x + y\} \\
&+ \quad \Pr\{\exists r^* \le s^* \le 0 : A_1(s^*, 0) + \phi_1 s^* > x, \\
&\qquad \inf_{s^* \le u \le 0}\{A_2(r^*, u) - (1 - \rho_1 - \epsilon)(s^* - r^*) - \phi_2(u - s^*)\} > -2y\} \\
&\le \quad \Pr\{\exists s^* \le 0 : A_1(s^*, 0) + \phi_1 s^* > x, \sup_{r \le s^*}\{A_1(r, s^*) - (\rho_1 + \epsilon)(s^* - r)\} > y\} \\
&+ \quad \Pr\{V_1^{\phi_1}(0) > x + y\} \\
&+ \quad \Pr\{\exists s^* \le 0 : A_1(s^*, 0) + \phi_1 s^* > x, \\
&\qquad \sup_{r \le s^*} \inf_{s^* \le u \le 0}\{A_2(r, u) - (1 - \rho_1 - \epsilon)(s^* - r) - \phi_2(u - s^*)\} > -2y\} \\
&\le \quad \Pr\{\exists s^* \le 0 : A_1(s^*, 0) + \phi_1 s^* > x\}
\end{aligned}$$

$$
\begin{aligned}
& \Pr\{\sup_{r\le s^*}\{A_1(r,s^*)-(\rho_1+\epsilon)(s^*-r)\}>y \mid A_1(s^*,0)+\phi_1 s^*>x\} \\
+ \quad & \Pr\{V_1^{\phi_1}(0)>x+y\}+\Pr\{\exists s^*\in[-w,0]: A_1(s^*,0)+\phi_1 s^*>x\} \\
+ \quad & \Pr\{\exists s^*\le -w: A_1(s^*,0)+\phi_1 s^*>x, \\
& \sup_{r\le s^*}\inf_{s^*\le u\le s^*+w}\{A_2(r,u)-(1-\rho_1-\epsilon)(s^*-r)-\phi_2(u-s^*)\}>-2y\} \\
\le \quad & \Pr\{\exists s\ge 0: A_1(-s,0)-\phi_1 s>x\}Q^{\rho_1+\epsilon}(s^*,x,y) \\
+ \quad & \Pr\{V_1^{\phi_1}(0)>x+y\}+\Pr\{\exists s\in[0,w]: A_1(-s,0)-\phi_1 s>x\} \\
+ \quad & \Pr\{\exists s^*\le -w: A_1(s^*,0)+\phi_1 s^*>x\} \\
& \Pr\{\sup_{r\le s^*}\inf_{s^*\le u\le s^*+w}\{A_2(r,u)-(1-\rho_1-\epsilon)(s^*-r)-\phi_2(u-s^*)\}>-2y\} \\
\le \quad & \Pr\{\sup_{s\ge 0}\{A_1(-s,0)-\phi_1 s\}>x\}Q^{\rho_1+\epsilon}(s^*,x,y) \\
+ \quad & \Pr\{V_1^{\phi_1}(0)>x+y\}+\Pr\{\sup_{0\le s\le w}\{A_1(-s,0)-\phi_1 s\}>x\} \\
+ \quad & \Pr\{\exists s\ge 0: A_1(-s,0)-\phi_1 s>x\} \\
& \Pr\{\sup_{r\ge 0}\inf_{0\le u\le w}\{A_2(-r,u)-(1-\rho_1-\epsilon)r-\phi_2 u\}>-2y\} \\
\le \quad & \Pr\{\mathbf{V}_1^{\phi_1}>x\}Q^{\rho_1+\epsilon}(s^*,x,y)+\Pr\{V_1^{\phi_1}(0)>x+y\}+\Pr\{\mathbf{V}_1^{\phi_1}(w)>x\} \\
+ \quad & \Pr\{\sup_{s\ge 0}\{A_1(-s,0)-\phi_1 s\}>x\} \\
& \Pr\{\inf_{0\le u\le w}\{\sup_{r\ge 0}\{A_2(-r,0)-(1-\rho_1-\epsilon)r\}+A_2(0,u)-\phi_2 u\}>-2y\} \\
\le \quad & \Pr\{\mathbf{V}_1^{\phi_1}>x\}\Pr\{\mathbf{T}_2^{1-\rho_1-\epsilon}(-2y)>w\} \\
+ \quad & \Pr\{V_1^{\phi_1}(0)>x+y\}+\Pr\{\mathbf{V}_1^{\phi_1}(w)>x\}+\Pr\{\mathbf{V}_1^{\phi_1}>x\}Q^{\rho_1+\epsilon}(s^*,x,y).
\end{aligned}
$$

Taking $w=\frac{(1-\alpha)x}{\hat{\rho}_1-\phi_1}$ completes the proof.

□

5 Preliminary results for the light-tailed flow

In this section we prove some auxiliary results for flow 1 in isolation. The results will be crucial in obtaining the asymptotic behavior of $\Pr\{\mathbf{V}_1>x\}$ in the GPS model as given in Theorem 3.1.

The following result is proven in [6] (for a more general class of input processes than just Markov fluid sources).

Proposition 5.1 *If Property 2.1 holds with $c_1=\phi_1$, then, for any $\alpha>0$,*

$$
\liminf_{x\to\infty}\frac{\Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1})>x\}}{\Pr\{\mathbf{V}_1^{\phi_1}>x\}}=1, \tag{13}
$$

where $\hat{\rho}_1:=M'_{\phi_1}(s^*(\phi_1))+\phi_1$.

Lemma 5.1 *For any* $\gamma > 0$, $\epsilon > 0$, $t^* < 0$,

$$\lim_{x\to\infty} \Pr\{\sup_{r\le t^*}\{(\rho_1-\epsilon)(t^*-r) - A_1(r,t^*)\} \le \gamma x \mid A_1(t^*,0)+\phi_1 t^* > x\} = 1.$$

Proof

Recall that flow 1 is a Markov fluid source. We condition on the state of the underlying Markov chain at time t^*. Let $E_j(t^*)$ be the event that the state at time t^* is j, $j = 1,\dots,d$, and $\pi_j(t^*) := \Pr\{E_j(t^*) \mid A_1(t^*,0)+\phi_1 t^* > x\}$. Then,

$$\begin{aligned} & \Pr\{\sup_{r\le t^*}\{(\rho_1-\epsilon)(t^*-r) - A_1(r,t^*)\} \le \gamma x \mid A_1(t^*,0)+\phi_1 t^* > x\} \\ = & \sum_{j=1}^{d} \Pr\{\sup_{r\le t^*}\{(\rho_1-\epsilon)(t^*-r) - A_1(r,t^*)\} \le \gamma x \mid A_1(t^*,0)+\phi_1 t^* > x, E_j(t^*)\}\pi_j(t^*) \\ = & \sum_{j=1}^{d} \Pr\{\sup_{r\le t^*}\{(\rho_1-\epsilon)(t^*-r) - A_1(r,t^*)\} \le \gamma x \mid E_j(t^*)\}\pi_j(t^*). \end{aligned}$$

The statement of the lemma then follows by observing that

$$\lim_{x\to\infty} \Pr\{\sup_{r\le t^*}\{(\rho_1-\epsilon)(t^*-r) - A_1(r,t^*)\} \le \gamma x \mid E_j(t^*)\} = 1$$

for all $j = 1,\dots,d$, since $\mathrm{E}\{A_1(-t,0)\} = \rho_1 t$.

□

Lemma 5.2 *For any* $\gamma > 0$, $\epsilon > 0$, $\mu > 0$, $t^* < 0$,

$$\lim_{x\to\infty} x^{\mu}\Pr\{\sup_{r\le t^*}\{A_1(r,t^*) - (\rho_1+\epsilon)(t^*-r)\} > \gamma x \mid A_1(t^*,0)+\phi_1 t^* > x\} = 0.$$

Proof

As in the proof of Lemma 5.1, let $E_j(t^*)$ be the event that the state at time t^* is j, $j = 1,\dots,d$, and $\pi_j(t^*) := \Pr\{E_j(t^*) \mid A_1(t^*,0)+\phi_1 t^* > x\}$. Then,

$$\begin{aligned} & \Pr\{\sup_{r\le t^*}\{A_1(r,t^*) - (\rho_1+\epsilon)(t^*-r)\} > \gamma x \mid A_1(t^*,0)+\phi_1 t^* > x\} \\ = & \sum_{j=1}^{d} \Pr\{\sup_{r\le t^*}\{A_1(r,t^*) - (\rho_1+\epsilon)(t^*-r)\} > \gamma x \mid A_1(t^*,0)+\phi_1 t^* > x, E_j(t^*)\}\pi_j(t^*) \\ = & \sum_{j=1}^{d} \Pr\{\sup_{r\le t^*}\{A_1(r,t^*) - (\rho_1+\epsilon)(t^*-r)\} > \gamma x \mid E_j(t^*)\}\pi_j(t^*). \end{aligned}$$

The statement of the lemma then follows by observing that there exist constants C, s^{**} (independent of j) such that

$$\Pr\{\sup_{r\le t^*}\{A_1(r,t^*) - (\rho_1+\epsilon)(t^*-r)\} > \gamma x \mid E_j(t^*)\} \le C\mathrm{e}^{-s^{**}x},$$

where $s^{**} > 0$ is the solution of $M_{\rho_1+\epsilon}(s) = 0$. In [16, Section 4] it is shown that C can be expressed in terms of the dominant eigenvalue of the matrix $\Lambda + s^{**}R$ and the corresponding (component-wise positive) eigenvalue.

□

Lemma 5.3 *For any $\gamma > 0$, $\mu > 0$,*

$$\limsup_{x\to\infty} \frac{x^\mu \Pr\{\mathbf{V}_1^{\phi_1} > (1+\gamma)x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}} = 0.$$

Proof
The proof follows immediately from the fact that $\Pr\{\mathbf{V}_1^{\phi_1} > x\}$ decays exponentially at rate s^*, where $s^* > 0$ is the solution of $M_{\phi_1}(s) = 0$ [14].

□

Lemma 5.4 *For any $\alpha > 0$, $\mu > 0$,*

$$\limsup_{x\to\infty} \frac{x^\mu \Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1-\alpha)x}{\hat{\rho}_1-\phi_1}) > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}} = 0.$$

Proof
The proof consists of three steps. First we give a sufficient condition for the lemma to hold, explicitly using the fact that the Markov fluid source has a bounded peak rate R_P. Then we estimate the decay rate of the event that a queue of capacity ϕ_1 fed by a Markov fluid source reaches overflow at time t. Finally we identify the most likely epoch of overflow, and show that this implies the required property.

- Obviously,

$$\begin{aligned}\Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1-\alpha)x}{\hat{\rho}_1-\phi_1}) > x\} &\leq \Pr\{\exists t \leq T_x(\alpha) : A_1(0,t) - \phi_1 t > x\} \\ &\leq \sum_{t=0}^{T_x(\alpha)} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\},\end{aligned}$$

  with

$$T_x(\alpha) := \left\lceil \frac{(1-\alpha)x}{\hat{\rho}_1-\phi_1} \right\rceil.$$

  From

$$\begin{aligned}&\max_{t=0,\dots,T_x(\alpha)} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\} \\ \leq\ &\sum_{t=0}^{T_x(\alpha)} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\} \\ \leq\ &(T_x(\alpha)+1) \max_{t=0,\dots,T_x(\alpha)} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\}\end{aligned}$$

  and

$$\lim_{x\to\infty} \frac{1}{x} \log(T_x(\alpha)+1) = 0,$$

we find that

$$
\begin{aligned}
& \limsup_{x\to\infty} \frac{1}{x} \log \sum_{t=0}^{T_x(\alpha)} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\} \\
= & \limsup_{x\to\infty} \frac{1}{x} \log \max_{t=0,\ldots,T_x(\alpha)} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\} \\
\leq & \limsup_{x\to\infty} \frac{1}{x} \log \sup_{t\in[0,T_x(\alpha)]} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\} \\
\leq & \limsup_{x\to\infty} \frac{1}{x} \log \sup_{t\in[S_x,T_x(\alpha)]} \Pr\{A_1(0,t) - \phi_1 t > x - (R_P - \phi_1)\}. \qquad (14)
\end{aligned}
$$

with $S_x := (x - R_P)/(R_P - \phi_1)$. Notice that we can indeed exclude all t smaller than S_x from the optimization, because in that range no overflow is possible. Clearly, we have proven the stated if we show that the latter decay rate is strictly smaller than s^* for all $\alpha > 0$.

- For x large enough, and all t between S_x and $T_x(\alpha)$, due to Chebychev's inequality, and Property 2.1,

$$\Pr\{A_1(0,t) - \phi_1 t > x - (r_P - \phi_1)\} \leq \inf_{s>0} \frac{\mathrm{E}\{\mathrm{e}^{s(A_1(0,t)-\phi_1 t)}\}}{\mathrm{e}^{s(x-(r_P-\phi_1))}} \leq C \inf_{s>0} \frac{\mathrm{e}^{M_{\phi_1}(s)t}}{\mathrm{e}^{s(x-(r_P-\phi_1))}}.$$

Now replace t in (14) by

$$t_x(\beta) = \frac{(1-\beta)x}{\hat\rho_1 - \phi_1},$$

then the supremum is over $\beta \in [\alpha, 1]$. The infimum over $s > 0$ is calculated by differentiation. We get the first-order condition

$$M'_{\phi_1}(s) = \frac{(x - (R_P - \phi_1))(\hat\rho_1 - \phi_1)}{(1-\beta)x}.$$

It is easily verified that the right-hand side of the previous display equals $(\hat\rho_1 - \phi_1)(1+\beta)$ for x large and β small. Call the solution $s^*(\beta)$.

Now recall that s^* solves $M_{\phi_1}(s) = 0$, and that $M'_{\phi_1}(s^*) = \hat\rho_1 - \phi_1 > 0$, see Property 2.1. Using

$$M'_{\phi_1}(s) = M'_{\phi_1}(s^*) + M''_{\phi_1}(s^*)(s - s^*) + \mathrm{O}((s-s^*)^2) =$$

$$\hat\rho_1 - \phi_1 + M''_{\phi_1}(s^*)(s - s^*) + \mathrm{O}((s-s^*)^2),$$

it is elementary to show that

$$s^*(\beta) = s^* + \delta\beta + \mathrm{O}(\beta^2), \text{ where } \delta := \frac{\hat\rho_1 - \phi_1}{M''_{\phi_1}(s^*)};$$

the convexity of $M_{\phi_1}(\cdot)$ implies that δ is positive. We also get that

$$M_{\phi_1}(s^*(\beta)) = M_{\phi_1}(s^*) + M'_{\phi_1}(s^*)\delta\beta + \mathrm{O}(\beta^2) = M'_{\phi_1}(s^*)\delta\beta + \mathrm{O}(\beta^2)$$

and

$$\lim_{x\to\infty}\frac{1}{x}\log\inf_{s>0}\frac{\mathrm{e}^{t_x(\beta)M_{\phi_1}(s)}}{\mathrm{e}^{s(x-(r_P-\phi_1))}} = \lim_{x\to\infty}\frac{1}{x}\left(t_x(\beta)M_{\phi_1}(s^*(\beta)) - s^*(\beta)x\right) =$$

$$\left(\frac{1-\beta}{\hat{\rho}_1-\phi_1}M'_{\phi_1}(s^*)-1\right)\delta\beta - s^* = -\delta\left(\frac{M'_{\phi_1}(s^*)}{\hat{\rho}_1-\phi_1}\right)\beta^2 - s^* = -\delta\beta^2 - s^*.$$

- Recall that we have to perform the optimization over $\beta \in [\alpha, 1]$. The supremum over β is clearly attained at $\beta = \alpha > 0$. Now the stated follows from the fact that $\Pr\{\mathbf{V}_1^{\phi_1} > x\}$ decays at rate s^*, as explained in the first step of the proof.

□

Remark 5.1 The results of Glynn & Whitt [11] suggest that the derived properties hold for a more general class of arrival processes than just Markov fluid. Upon inspection of the proofs in the present section, we see that two properties were explicitly exploited. In the first place it was repeatedly used that the source has a bounded peak rate. Secondly, it is required that the dependence between $A_1(r, t^*)$ and $A_1(t^*, 0)$ is rather mild. This leads us to believe that the lemmas still hold if the exponential sojourn times of the Markov fluid source are replaced by other light-tailed random variables.

6 Asymptotic analysis

We now use the results from the previous section to show that the lower and upper bounds for $\Pr\{\mathbf{V}_1 > x\}$ of Section 4 asymptotically coincide, resulting in the decompositional form of (3). For the proof, we need to make certain assumptions on the behavior of the drain time distribution $\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}$. In later sections, we will determine the specific form of the drain time distribution, and find that flow 2 indeed satisfies these assumptions. For notational convenience, we frequently switch to a variable $\hat{x}$, which should be thought of as playing the role of $\frac{x}{\hat{\rho}_1-\phi_1}$.

Lemma 6.1 *Suppose that the input process $A_1(s,t)$ satisfies Property 2.1 with $c_1 = \phi_1$ and that flow 2 satisfies Assumptions 6.1-6.3 listed below with $c = 1-\rho_1$. Assume that $\rho_i < \phi_i$, $i = 1, 2$, and $r_2 > \phi_2$ in case of fluid input of flow 2. Then*

$$\Pr\{\mathbf{V}_1 > x\} \sim \Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}.$$

Assumption 6.1 *For any $\alpha > 0$, $\gamma > 0$, $\epsilon > 0$, either (a)*

$$\liminf_{\hat{x}\to\infty}\frac{\Pr\{\mathbf{T}_2^{c+\epsilon}(\gamma\hat{x}) > (1+\alpha)\hat{x}\}}{\Pr\{\mathbf{T}_2^{c} > \hat{x}\}} = F^c(\alpha,\gamma,\epsilon),$$

with $\lim_{\alpha,\gamma,\epsilon\downarrow 0} F^c(\alpha,\gamma,\epsilon) = 1$, *or (b)*

$$\liminf_{\hat{x}\to\infty} \frac{\Pr\{\mathbf{T}_2 > (1+\alpha)\hat{x}\}}{\Pr\{\mathbf{T}_2^c > \hat{x}\}} = F(\alpha),$$

with $\lim_{\alpha\downarrow 0} F(\alpha) = 1$.

Assumption 6.2 *For any* $\alpha > 0$, $\gamma > 0$, $\epsilon > 0$,

$$\limsup_{\hat{x}\to\infty} \frac{\Pr\{\mathbf{T}_2^{c-\epsilon}(-\gamma\hat{x}) > (1-\alpha)\hat{x}\}}{\Pr\{\mathbf{T}_2^c > \hat{x}\}} = G^c(\alpha,\gamma,\epsilon),$$

with $\lim_{\alpha,\gamma,\epsilon\downarrow 0} G^c(\alpha,\gamma,\epsilon) = 1$.

Assumption 6.3 *For some* $\mu > 0$,

$$\liminf_{x\to\infty} \hat{x}^{\mu}\Pr\{\mathbf{T}_2^c > \hat{x}\} \geq 1.$$

Proof of Lemma 6.1

The proof consists of a lower bound and an upper bound which asymptotically coincide. We start with the lower bound. We distinguish between two cases: Assumption 6.1 (a); Assumption 6.1 (b).

(a) Using Corollary 4.1 with $v = \infty$, $y = \frac{\gamma x}{\hat{\rho}_1 - \phi_1}$, Proposition 5.1, and Lemma 5.1,

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_1 > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1 - \phi_1}\}} \geq$$

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1}) > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}} \liminf_{x\to\infty} \frac{\Pr\{\mathbf{T}_2^{1-\rho_1+\epsilon}(\frac{\gamma x}{\hat{\rho}_1-\phi_1}) > \frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1}\}}{\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}}$$

$$\liminf_{x\to\infty} \Pr\{\sup_{r\leq s^*}\{(\rho_1-\epsilon)(s^*-r) - A_1(r,s^*)\} \leq \frac{\gamma x}{\hat{\rho}_1-\phi_1} \mid A_1(s^*,0) + \phi_1 s^* > x\} =$$

$$F^{1-\rho_1}(\alpha,\gamma,\epsilon).$$

Letting $\alpha,\gamma,\epsilon \downarrow 0$ completes the proof.

(b) Using Corollary 4.1 with $v = 0$, $y = 0$, and Proposition 5.1, noting that $P^{\rho_1-\epsilon}(s^*,0,x,0) = 1$,

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_1 > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1 - \phi_1}\}} \geq$$

$$\liminf_{x\to\infty} \frac{\Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1}) > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}} \liminf_{x\to\infty} \frac{\Pr\{\mathbf{T}_2 > \frac{(1+\alpha)x}{\hat{\rho}_1-\phi_1}\}}{\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}} = F(\alpha).$$

Then let $\alpha \downarrow 0$.

We now turn to the upper bound. Using Corollary 4.2 with $v = \infty$, $y = \frac{\gamma x}{2(\hat{\rho}_1 - \phi_1)}$, Lemmas 5.2-5.4, and Assumptions 6.2, 6.3, for some $\mu > 0$,

$$\begin{aligned}
&\limsup_{x\to\infty} \frac{\Pr\{\mathbf{V}_1 > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}} \\
\leq\ &\limsup_{x\to\infty} \frac{\Pr\{\mathbf{T}_2^{1-\rho_1-\epsilon}(\frac{-\gamma x}{\hat{\rho}_1-\phi_1}) > \frac{(1-\alpha)x}{\hat{\rho}_1-\phi_1}\}}{\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}} + \limsup_{x\to\infty} \frac{x^\mu \Pr\{\mathbf{V}_1^{\phi_1} > (1+\frac{\gamma}{2(\hat{\rho}_1-\phi_1)})x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}} \\
+\ &\limsup_{x\to\infty} \frac{x^\mu \Pr\{\mathbf{V}_1^{\phi_1}(\frac{(1-\alpha)x}{\hat{\rho}_1-\phi_1}) > x\}}{\Pr\{\mathbf{V}_1^{\phi_1} > x\}} \\
+\ &\limsup_{x\to\infty} x^\mu \Pr\{\sup_{r\leq s^*}\{A_1(r,s^*) - (\rho_1+\epsilon)(s^*-r)\} > \frac{\gamma x}{2(\hat{\rho}_1-\phi_1)} \mid A_1(s^*,0) + \phi_1 s^* > x\} \\
=\ &G^{1-\rho_1}(\alpha,\gamma,\epsilon).
\end{aligned}$$

Letting $\alpha, \gamma, \epsilon \downarrow 0$ completes the proof.

□

In order to complete the proof of Theorem 3.1, it remains to be shown that flow 2 satisfies Assumptions 6.1-6.3 above, with $\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}$ as in (4)-(6). This is done in the next four sections.

7 Preliminary results for the heavy-tailed flow

To determine the behavior of $\Pr\{\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}\}$ as $x \to \infty$, we will reduce the space of all relevant sample paths to a single most-likely scenario, which occurs with overwhelming probability. In this section, we establish some preliminary results which we will use to neglect the contribution of all non-dominant scenarios.

Large-deviations arguments for heavy-tailed distributions suggest that a persistent backlog as associated with the event $\mathbf{T}_2^{1-\rho_1} > \frac{x}{\hat{\rho}_1-\phi_1}$, for large x, is most likely due to just a single large burst or long On-period. To formalize this idea, we first introduce some additional notation. A burst is called large if the size exceeds $\kappa\hat{x}$, with $\kappa > 0$ some small constant, independent of $\hat{x}$. Also, an On-period is called long if the length exceeds $\kappa\hat{x}$. In case of instantaneous input, we denote by $\mathcal{N}_{\kappa\hat{x}}[l,r]$ the number of large bursts of flow 2 arriving in the time interval $[l,r]$. In case of an On-Off process, we define $\mathcal{N}_{\kappa\hat{x}}[l,r]$ as the number of long On-periods overlapping with the time interval $[l,r]$, including the On-period which may be in progress at time l.

Depending on the traffic scenario, we denote by $N(t)$ either the number of bursts or the number of On-periods of flow 2 in the time interval $[0,t]$. An upper bound for this process is given by

$$N(t) \leq N_U(t) := \{n : \sum_{i=1}^{n} \mathbf{U}_{2i} \leq t\} + 1,$$

with $\mathbf{U}_{2i}$ i.i.d. random variables representing either interarrival times or Off-periods of flow 2, depending on the traffic scenario.

We now state a crucial lemma which will allow us to limit the attention to large bursts and long On-periods, and replace all remaining traffic activity by its average rate. The lemma is a minor modification of Lemma 3 in Resnick & Samorodnitsky [21].

Lemma 7.1 *Let $\mathbf{S}_n = \mathbf{X}_1 + \ldots + \mathbf{X}_n$ be a random walk with i.i.d. step sizes such that* $\mathrm{E}\{\mathbf{X}_1\} < 0$ *and* $\mathrm{E}\{\mathbf{X}_1^p\} < \infty$ *for some $p > 1$. Then, for any $\mu < \infty$, there exists a $\kappa^* > 0$ and a function $\phi(\cdot) \in \mathcal{R}_{-\mu}$ such that for all $\kappa \in (0, \kappa^*]$,*

$$\Pr\{\mathbf{S}_n > \hat{x} | \mathbf{X}_i \leq \kappa\hat{x}, i = 1, \ldots, n\} \leq \phi(\hat{x})$$

for all n and $\hat{x}$.

Note that if $\mathbf{X}_i$ can be represented as the difference of two non-negative independent random variables $\mathbf{X}_i^1$ and $\mathbf{X}_i^2$, then the lemma remains valid if the $\mathbf{X}_i$'s are replaced by the $\mathbf{X}_i^1$'s.

We now use the above lemma to show that the workload of flow 2 cannot significantly deviate from the normal drift over intervals of the order $\hat{x}$ when there are no large bursts.

Lemma 7.2 *If $B_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $\eta > 0$, $\theta > 0$, there exists a $\kappa^* > 0$ such that for all $\kappa \in (0, \kappa^*]$,*

$$\Pr\{\mathbf{T}_2((\theta - (\phi_2 - \rho_2)\eta)\hat{x}) > \eta\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0, \eta\hat{x}] = 0\} = \mathrm{o}(\Pr\{\mathbf{B}_2^r > \hat{x}(\phi_2 - \rho_2)\})$$

as $\hat{x} \to \infty$.

Proof
The event $\mathbf{T}_2((\theta - (\phi_2 - \rho_2)\eta)\hat{x}) > \eta\hat{x}$ means that

$$\inf_{0 \leq u \leq \eta\hat{x}} \{A_2(0, u) - \phi_2 u\} > (\theta - (\phi_2 - \rho_2)\eta)\hat{x},$$

which in particular implies that

$$A_2(0, \eta\hat{x}) - \phi_2\eta\hat{x} > (\theta - (\phi_2 - \rho_2)\eta)\hat{x},$$

or equivalently,

$$A_2(0, \eta\hat{x}) - (\rho_2 + \theta/2\eta)\eta\hat{x} > \theta\hat{x}/2,$$

so that also

$$\sup_{0 \leq u \leq \eta\hat{x}} \{A_2(0, u) - (\rho_2 + \theta/2\eta)u\} > \theta\hat{x}/2.$$

Now let $\mathbf{S}_n := \mathbf{X}_1 + \ldots + \mathbf{X}_n$ be a random walk with step sizes $\mathbf{X}_i := \mathbf{B}_{2i} - (\rho_2 + \theta/2\eta)\mathbf{U}_{2i}$, with $\mathbf{U}_{2i}$ and $\mathbf{B}_{2i}$ i.i.d. random variables representing the interarrival times and burst sizes of flow 2, respectively. Note that $\mathbf{X}_i$ represents the net increase in the workload in a queue of capacity $\rho_2 + \theta/2\eta$ between two consecutive bursts, and that $\mathrm{E}\{\mathbf{X}_i\} < 0$.
Because of the saw-tooth nature of the process $\{A_2(0, u) - (\rho_2 + \theta/2\eta)u\}$, we have

$$\sup_{0 \leq u \leq t} \{A_2(0, u) - (\rho_2 + \theta/2\eta)u\} \leq \mathbf{B}_{20} + \sup_{1 \leq n \leq N(t)} \mathbf{S}_n.$$

Thus,

$$
\begin{aligned}
& \Pr\{\mathbf{T}_2((\theta-(\phi_2-\rho_2)\eta)\hat{x}) > \eta\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,\eta\hat{x}]=0\} \\
\leq\ & \Pr\{\mathbf{B}_{20} + \sup_{1\leq n\leq N(\eta\hat{x})} \mathbf{S}_n \geq \theta\hat{x}/2, \mathcal{N}_{\kappa\hat{x}}[0,\eta\hat{x}]=0\} \\
\leq\ & \Pr\{\mathbf{B}_{20} + \sup_{1\leq n\leq N(\eta\hat{x})} \mathbf{S}_n \geq \theta\hat{x}/2 \mid \mathcal{N}_{\kappa\hat{x}}[0,\eta\hat{x}]=0\} \\
\leq\ & \Pr\{\mathbf{B}_{20} + \sup_{1\leq n\leq N(\eta\hat{x})} \mathbf{S}_n \geq \theta\hat{x}/2 \mid \mathbf{B}_{2i} \leq \kappa\hat{x}, i\geq 0\} \\
\leq\ & \Pr\{\sup_{1\leq n\leq N(\eta\hat{x})} \mathbf{S}_n \geq (\theta/2-\kappa)\hat{x} \mid \mathbf{B}_{2i} \leq \kappa\hat{x}, i\geq 1\} \\
\leq\ & \Pr\{\sup_{1\leq n\leq (\lambda_2+\epsilon)\eta\hat{x}} \mathbf{S}_n \geq (\theta/2-\kappa)\hat{x} \mid \mathbf{B}_{2i} \leq \kappa\hat{x}, i\geq 1\} + \Pr\{N(\eta\hat{x}) > (\lambda_2+\epsilon)\eta\hat{x}\} \\
\leq\ & \sum_{i=1}^{(\lambda_2+\epsilon)\eta\hat{x}} \Pr\{\mathbf{S}_n \geq (\theta/2-\kappa)\hat{x} \mid \mathbf{B}_{2i} \leq \kappa\hat{x}, i=1,\ldots,n\} + \Pr\{N(\eta\hat{x}) > (\lambda_2+\epsilon)\eta\hat{x}\}.
\end{aligned}
$$

The second term decays exponentially fast as $\hat{x}\to\infty$. According to Lemma 7.1, there exists a $\kappa^* > 0$ and a function $\phi(\cdot) \in \mathcal{R}_{-\mu}$, $\mu > \nu_2$, such that for all $\kappa \in (0,\kappa^*]$, each of the probabilities in the first term is upper bounded by $\phi(\hat{x})$. The statement then follows.

□

We now formulate the counterpart of the above lemma for On-Off processes, meaning that the workload of flow 2 closely follows the drift over intervals of the order $\hat{x}$ when there are no long On-periods.

Lemma 7.3 *If $A_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $\eta > 0$, $\theta > 0$, there exists a $\kappa^* > 0$ such that for all $\kappa \in (0,\kappa^*]$,*

$$
\Pr\{\mathbf{T}_2((\theta-(\phi_2-\rho_2)\eta)\hat{x}) > \eta\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,\eta\hat{x}]=0\} = \mathrm{o}(\Pr\{\mathbf{A}_2^r > \frac{\hat{x}(\phi_2-\rho_2)}{r_2-\rho_2}\})
$$

as $\hat{x}\to\infty$.

Proof
Let $\mathbf{S}_n := \mathbf{X}_1 + \ldots + \mathbf{X}_n$ be a random walk with step sizes $\mathbf{X}_i := (r_2-\rho_2-\theta/2\eta)\mathbf{A}_{2i} - (\rho_2 + \theta/2\eta)\mathbf{U}_{2i}$, with $\mathbf{A}_{2i}$ and $\mathbf{U}_{2i}$ i.i.d. random variables representing the On-periods and Off-periods of flow 2, respectively. Note that $\mathbf{X}_i$ represents the net increase in the workload in a queue of capacity $\rho_2+\theta/2\eta$ during an Off-period and consecutive On-period, and that $\mathrm{E}\{\mathbf{X}_i\} < 0$.
Because of the saw-tooth nature of the process $\{A_2(0,u) - (\rho_2+\theta/2\eta)u\}$, we have

$$
\sup_{0\leq u\leq t}\{A_2(0,u) - (\rho_2+\theta/2\eta)u\} \leq (r_2-\rho_2)\mathbf{A}_{20} + \sup_{1\leq n\leq N(t)} \mathbf{S}_n \leq
$$

$$
(r_2-\rho_2)\mathbf{A}_{20} + \sup_{1\leq n\leq N_U(t)} \mathbf{S}_n.
$$

The remainder of the proof is similar to that of Lemma 7.2.

□

We now prove that it is relatively unlikely for flow 2 to generate two large bursts in an interval of order $\hat{x}$.

Lemma 7.4 *If $B_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $\alpha < 1$, $\kappa > 0$,*

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} = \mathrm{o}(\Pr\{\mathbf{B}_2^r > \hat{x}(\phi_2 - \rho_2)\})$$

as $\hat{x} \to \infty$.

Proof
By definition,

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} = \Pr\{\#\{j \in \{1,\ldots,N((1-\alpha)\hat{x})\} : \mathbf{B}_{2j} \geq \kappa\hat{x}\} \geq 2\}$$

$$\leq \Pr\{\#\{j \in \{1,\ldots,N_U((1-\alpha)\hat{x})\} : \mathbf{B}_{2j} \geq \kappa\hat{x}\} \geq 2\}.$$

Now condition on $N_U((1-\alpha)\hat{x})$. This yields the following upper bound

$$\mathrm{E}\{N_U((1-\alpha)\hat{x})^2\}\Pr\{\mathbf{B}_2 \geq \kappa\hat{x}\}^2.$$

Finally, observe that $\mathrm{E}\{N_U((1-\alpha)\hat{x})^2\}$ is quadratic in $\hat{x}$ for $\hat{x} \to \infty$.

□

We now state the counterpart of the above lemma for On-Off processes, meaning that the probability that flow 2 experiences two long On-periods during an interval of order $\hat{x}$ is negligibly small.

Lemma 7.5 *If $A_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $\alpha < 1$, $\kappa > 0$,*

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} = \mathrm{o}(\Pr\{\mathbf{A}_2^r > \frac{\hat{x}(\phi_2 - \rho_2)}{r_2 - \rho_2}\})$$

as $\hat{x} \to \infty$.

Proof
This lemma is a variant of Proposition 6.3 of [23]. Note that

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} \leq$$

$$(1-p_2)\Pr\{\mathbf{A}_2^r \geq \kappa\hat{x}\}\Pr\{\#\{j \in \{1,\ldots,N_U((1-\alpha)\hat{x})\} : \mathbf{A}_{2j} \geq \kappa\hat{x}\} \geq 1\} +$$

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2, \text{flow 2 is Off at time 0}\}.$$

By conditioning upon $N_U((1-\alpha)\hat{x})$, one can bound the second probability in the first term by $\mathrm{E}\{N_U((1-\alpha)\hat{x})\}\Pr\{\mathbf{A}_2 \geq \kappa\hat{x}\}$. The first factor is linear in $\hat{x}$ for $\hat{x} \to \infty$, whereas the second is in $\mathcal{R}_{-\nu_2}$. Hence, the first term is in $\mathcal{R}_{2(1-\nu_2)}$. To bound the second term, condition (again) on $N_U((1-\alpha)\hat{x})$. This yields

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2, \text{flow 2 is Off at time } 0\} \leq \mathrm{E}\{N^U((1-\alpha)\hat{x})^2\}\Pr\{\mathbf{A}_2 \geq \kappa\hat{x}\}^2.$$

Finally, note that, as in Lemma 7.4, $\mathrm{E}\{N_U((1-\alpha)\hat{x})^2\}$ is quadratic in $\hat{x}$ for $\hat{x} \to \infty$.

□

We now prove that the amount of traffic generated by flow 2 after turning Off is not below average by any significant margin.

Lemma 7.6 *Suppose that flow 2 turns Off at time v. Then for any $\delta > 0$, $\theta > 0$,*

$$\lim_{\hat{x}\to\infty} \Pr\{\sup_{u \geq v}\{(\rho_2 - \delta)(u-v) - A_2(v,u)\} \leq \theta\hat{x}\} = 1.$$

Proof
Let $\mathbf{S}_n := \mathbf{X}_1 + \ldots + \mathbf{X}_n$ be a random walk with step sizes $\mathbf{X}_i := (\rho_2 - \delta - r_2)\mathbf{A}_{2i} + (\rho_2 - \delta)\mathbf{U}_{2i}$, with $\mathbf{A}_{2i}$ and $\mathbf{U}_{2i}$ i.i.d. random variables representing the On-periods and Off-periods of flow 2, respectively. Note that $\mathbf{X}_i$ represents the net decrease in the workload in a queue of capacity $\rho_2 - \delta$ fed by flow 2 during an On-period and consecutive Off-period, and that $\mathrm{E}\{\mathbf{X}_i\} < 0$.
Now observe that

$$\sup_{u \geq v}\{(\rho_2 - \delta)(u-v) - A_2(v,u)\} \leq (\rho_2 - \delta)\mathbf{U}_{20} + \sup_{n \geq 1}\mathbf{S}_n,$$

so that

$$\begin{aligned} & \Pr\{\sup_{u \geq v}\{(\rho_2 - \delta)(u-v) - A_2(v,u)\} \leq \theta\hat{x}\} \\ = \; & 1 - \Pr\{\sup_{u \geq v}\{(\rho_2 - \delta)(u-v) - A_2(v,u)\} > \theta\hat{x}\} \\ \geq \; & 1 - \Pr\{(\rho_2 - \delta)\mathbf{U}_{20} + \sup_{n \geq 1}\mathbf{S}_n > \theta\hat{x}\}. \end{aligned}$$

The probability in the last term goes to 0 as $\hat{x} \to \infty$ for any $\theta > 0$, since the maximum of a random walk with negative drift is finite with probability 1.

□

Lemma 7.7 *If $B_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $0 < \xi < 1 - \alpha$, $\zeta > 0$, $\kappa > 0$,*

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[\xi\hat{x},(1-\alpha)\hat{x}] \geq 1, V_2^c(0) > \zeta\hat{x}\} = \mathrm{o}(\Pr\{\mathbf{B}_2^r > \hat{x}(\phi_2 - \rho_2)\})$$

as $\hat{x} \to \infty$.

Proof
Because of independence, the probability equals

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[\xi\hat{x},(1-\alpha)\hat{x}] \geq 1\}\Pr\{V_2^c(0) > \zeta\hat{x}\}.$$

By conditioning upon $N_U((1-\alpha-\xi)\hat{x})$, we have

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[\xi\hat{x},(1-\alpha)\hat{x}] \geq 1\} \leq \mathrm{E}\{N_U((1-\alpha-\xi)\hat{x})\}\Pr\{\mathbf{B}_2 > \kappa\hat{x}\}.$$

As before the first term is linear in $\hat{x}$ for $\hat{x} \to \infty$. The statement then follows from the fact that $B_2(\cdot) \in \mathcal{R}_{-\nu_2}$ in combination with Theorem 2.1.

□

Lemma 7.8 *If $A_2^r(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $c \in (\rho_2, r_2)$, $0 < \xi < 1-\alpha$, $\zeta > 0$, $\kappa > 0$,*

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] = 1, \mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}] = 0, V_2^c(0) \geq \zeta\hat{x}\} = \mathrm{o}(\Pr\{\mathbf{A}_2^r > \frac{\hat{x}(\phi_2-\rho_2)}{r_2-\rho_2}\})$$

as $\hat{x} \to \infty$.

Proof
The event $\mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}] = 0$ in conjunction with $\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] = 1$ implies that flow 2 has switched On at some time t in the interval $[\xi\hat{x},(1-\alpha)\hat{x}]$. Therefore, an upper bound is given by

$$\Pr\{\mathcal{N}_{\kappa\hat{x}}[\xi\hat{x},(1-\alpha)\hat{x}] = 1, \text{long On-period started after time } \xi\hat{x}, V_2^c(0) \geq \zeta\hat{x}\} =$$

$$\Pr\{\#\{j \in \{1,\ldots,N_U((1-\alpha-\xi)\hat{x})\} : \mathbf{A}_{2j} > \kappa\hat{x}\} = 1\}\Pr\{V_2^c(0) \geq \zeta\hat{x}\}.$$

By conditioning upon $N_U((1-\alpha-\xi)\hat{x})$, the first term can be bounded by

$$\mathrm{E}\{N_U((1-\alpha-\xi)\hat{x})\}\Pr\{\mathbf{A}_2 > \kappa\hat{x}\}.$$

Combining the fact that $A_2(\cdot) \in \mathcal{R}_{-\nu_2}$ with Theorem 2.2 then completes the proof.

□

8 Case I: instantaneous input

In this section we consider the case where flow 2 generates instantaneous traffic bursts of regularly varying size. The next theorem shows that flow 2 then satisfies Assumptions 6.1-6.3 and that (4) holds.

Theorem 8.1 *If $B_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $c > \rho_2$, $\alpha > 0$, $\gamma > 0$,*

$$\Pr\{\mathbf{T}_2^c(\gamma\hat{x}) > (1+\alpha)\hat{x}\} \gtrsim \frac{\rho_2}{c-\rho_2}\Pr\{\mathbf{B}_2^r > ((\phi_2-\rho_2)(1+\alpha)+\gamma)\hat{x}\}, \tag{15}$$

$$\Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \lesssim \frac{\rho_2}{c-\rho_2}\Pr\{\mathbf{B}_2^r > ((\phi_2-\rho_2)(1-\alpha)-\gamma\frac{c+\phi_2-2\rho_2}{\phi_2-\rho_2})\hat{x}\}, \tag{16}$$

and

$$\Pr\{\mathbf{T}_2^c > \hat{x}\} \sim \frac{\rho_2}{c-\rho_2}\Pr\{\mathbf{B}_2^r > \hat{x}(\phi_2-\rho_2)\}. \tag{17}$$

Before giving the formal proof of the above theorem, we first provide an intuitive argument. Consider a queue of capacity ϕ_2 fed by the arrival process of flow 2. In order for the event $\mathbf{T}_2^c > \hat{x}$ to occur, the workload must remain positive throughout the interval $[0, \hat{x}]$, given that the initial workload is $V_2^c(0)$. Note that the normal drift in the workload is $\rho_2 - \phi_2 < 0$. Thus, there is a 'deficit' $(\phi_2 - \rho_2)\hat{x}$, which must be compensated for by the initial workload $V_2^c(0)$ plus possibly flow 2 showing above-average activity during the interval $[0, \hat{x}]$.
We claim that the most likely way for the gap to be filled is by a large initial workload only, i.e., $V_2^c(0) > (\phi_2 - \rho_2)\hat{x}$. This in turn is most probably due to an extremely large burst of flow 2 somewhere before time 0, which is consistent with the usual situation for heavy-tailed distributions that a large deviation is caused by just a single exceptional event. Using Theorem 2.1, we see that the probability of this event is indeed exactly the right-hand side of (17).
Note that it is unlikely for the gap to be filled by flow 2 producing extra traffic during the interval $[0, \hat{x}]$, because this would require a large burst arriving almost immediately after time 0. The probability of this event is negligibly small compared to that of $V_2^c(0) > (\phi_2 - \rho_2)\hat{x}$. A combination of both is even less likely, since this would amount to two rare events occurring simultaneously.
The above arguments will be formalized in the proof below. We first prove that the event $V_2^c(0) > (\phi_2 - \rho_2)\hat{x}$ indeed implies that $\mathbf{T}_2^c > \hat{x}$ for large $\hat{x}$, thus obtaining a lower bound for the probability of the latter event. Next we show that for large $\hat{x}$ the event $V_2^c(0) > (\phi_2 - \rho_2)\hat{x}$ is also necessary for $\mathbf{T}_2^c > \hat{x}$ to occur, by proving that the probability of all other possible scenarios is negligibly small.

Proof of Theorem 8.1
We start with the proof of (15). We first prove that for any $\alpha > 0$, $\gamma > 0$, $\delta > 0$, $\theta > 0$, the event

$$\mathbf{T}_2^c(\gamma\hat{x}) > (1+\alpha)\hat{x} \tag{18}$$

is implied by the events

$$V_2^c(0) > ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \gamma + \theta)\hat{x},$$

and

$$\sup_{0 \le u \le (1+\alpha)\hat{x}} \{(\rho_2 - \delta)u - A_2(0,u)\} \le \theta\hat{x}.$$

The second event means that for all $u \in [0, (1+\alpha)\hat{x}]$,

$$A_2(0,u) \ge (\rho_2 - \delta)u - \theta\hat{x}.$$

Thus, for all $u \in [0, (1+\alpha)\hat{x}]$,

$$\begin{aligned} V_2^c(0) + A_2(0,u) - \phi_2 u &> ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \gamma + \theta)\hat{x} + (\rho_2 - \delta)u - \theta\hat{x} - \phi_2 u \\ &= (\phi_2 - \rho_2 + \delta)((1+\alpha)\hat{x} - u) + \gamma\hat{x} \\ &\ge \gamma\hat{x}, \end{aligned}$$

so that

$$\inf\{u \ge 0 : V_2^c(0) + A_2(0,u) - \phi_2 u \le \gamma\hat{x}\} > (1+\alpha)\hat{x},$$

which gives (18).
Hence, using independence of $V_2^c(0)$ and $A_2(0,u)$,

$$\Pr\{\mathbf{T}_2^c(\gamma\hat{x}) > (1+\alpha)\hat{x}\} \geq$$

$$\Pr\{V_2^c(0) > ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \gamma + \theta)\hat{x}\}\Pr\{\sup_{0\leq u\leq(1+\alpha)\hat{x}}\{(\rho_2 - \delta)u - A_2(0,u)\} \leq \theta\hat{x}\}.$$

Using Theorem 2.1,

$$\Pr\{V_2^c(0) > ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \gamma + \theta)\hat{x}\} \sim$$

$$\frac{\rho_2}{c-\rho_2}\Pr\{\mathbf{B}_2^r > ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \gamma + \theta)\hat{x})\}.$$

Also, for all $\alpha > 0$, $\delta > 0$, $\theta > 0$,

$$\Pr\{\sup_{0\leq u\leq(1+\alpha)\hat{x}}\{(\rho_2 - \delta)u - A_2(0,u)\} \leq \theta\hat{x}\} \geq \Pr\{\sup_{u\geq 0}\{(\rho_2 - \delta)u - A_2(0,u)\} \leq \theta\hat{x}\} \to 1,$$

as $\hat{x} \to \infty$, since $\mathrm{E}\{A_2(0,u)\} = \rho_2 u$.
Thus, for all $\alpha > 0$, $\gamma > 0$, $\delta > 0$, $\theta > 0$,

$$\Pr\{\mathbf{T}_2^c(\gamma\hat{x}) > (1+\alpha)\hat{x}\} \gtrsim \frac{\rho_2}{c-\rho_2}\Pr\{\mathbf{B}_2^r > ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \gamma + \theta)\hat{x}\}.$$

Letting $\delta \downarrow 0$, $\theta \downarrow 0$, using the fact that $B_2^r(\cdot) \in \mathcal{IR}$, (15) follows.
We now turn to the proof of (16).
By partitioning, we obtain for any $\alpha > 0$, $\gamma > 0$, $\zeta > 0$, $\theta > 0$, $\kappa > 0$, $w \geq 0$,

$$\begin{aligned}
& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \\
= \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(w) > ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw\} \\
+ \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(w) \leq ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw, \\
& \quad \mathcal{N}_{\kappa\hat{x}}[0,w] \leq 1, \mathcal{N}_{\kappa\hat{x}}[w,(1-\alpha)\hat{x}] = 0\} \\
+ \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(w) \leq ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw, \\
& \quad \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, \mathcal{N}_{\kappa\hat{x}}[w,(1-\alpha)\hat{x}] = 1, V_2^c(0) \leq \zeta x\} \\
+ \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(w) \leq ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw, \\
& \quad \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, \mathcal{N}_{\kappa\hat{x}}[w,(1-\alpha)\hat{x}] = 1, V_2^c(0) > \zeta x\} \\
+ \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(w) \leq ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw, \\
& \quad \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\},
\end{aligned}$$

which is obviously upper bounded by

$$\begin{aligned}
& \Pr\{V_2^c(w) > ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw\} \\
+ \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(w) \leq ((\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta)\hat{x} - cw, \\
& \quad \mathcal{N}_{\kappa\hat{x}}[w,(1-\alpha)\hat{x}] = 0\} \\
+ \;& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, V_2^c(0) \leq \zeta x\} \\
+ \;& \Pr\{\mathcal{N}_{\kappa\hat{x}}[w,(1-\alpha)\hat{x}] \leq 1, V_2^c(0) > \zeta x\} \\
+ \;& \Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} \\
= \;& (A) + (B) + (C) + (D) + (E).
\end{aligned}$$

Take $w = \xi x$, with

$$\xi := \frac{\gamma + \zeta + \theta}{\phi_2 - \rho_2} < 1 - \alpha.$$

Now consider term (A). Using Theorem 2.1,

$$\begin{aligned}
(A) &= \Pr\{\mathbf{V}_2^c > ((\phi_2 - \rho_2)(1 - \alpha) - \gamma - \theta - (c - \rho_2)\xi)\hat{x}\} \\
&\sim \frac{\rho_2}{c - \rho_2}\Pr\{\mathbf{B}_2^r > ((\phi_2 - \rho_2)(1 - \alpha) - \gamma - \theta - (c - \rho_2)\xi)\hat{x}\} \\
&= \frac{\rho_2}{c - \rho_2}\Pr\{\mathbf{B}_2^r > ((\phi_2 - \rho_2)(1 - \alpha) - \gamma - \theta - \frac{(c - \rho_2)(\gamma + \zeta + \theta)}{\phi_2 - \rho_2})\hat{x}\}.
\end{aligned}$$

Next, consider term (B). The event $\mathbf{T}_2^c(-\gamma\hat{x}) > (1 - \alpha)\hat{x}$ means that

$$\inf_{0 \le u \le (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0, u) - \phi_2 u\} > -\gamma\hat{x}.$$

Now observe that

$$\begin{aligned}
&\inf_{0 \le u \le (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0, u) - \phi_2 u\} \\
\le\ & V_2^c(0) + \inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(0, u) - \phi_2 u\} \\
\le\ & V_2^c(0) + \inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(0, w) - \phi_2 w + A_2(w, u) - \phi_2(u - w)\} \\
\le\ & V_2^c(0) + A_2(0, w) - \phi_2 w + \inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(w, u) - \phi_2(u - w)\} \\
\le\ & V_2^c(w) + (c - \phi_2)w + \inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(w, u) - \phi_2(u - w)\}.
\end{aligned}$$

Thus, the event $\mathbf{T}_2^c(-\gamma\hat{x}) > (1 - \alpha)\hat{x}$ implies

$$\inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(0, u) - \phi_2 u\} > -V_2^c(w) - (c - \phi_2)w - \gamma\hat{x},$$

so that

$$\begin{aligned}
(B) &\le \Pr\{\inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(w, u) - \phi_2(u - w)\} > -V_2^c(w) - (c - \phi_2)w - \gamma\hat{x}, \\
&\qquad V_2^c(w) \le ((\phi_2 - \rho_2)(1 - \alpha) - \gamma - \theta)\hat{x} - (c - \rho_2)w, \mathcal{N}_{\kappa\hat{x}}[w, (1 - \alpha)\hat{x}] = 0\} \\
&\le \Pr\{\inf_{w \le u \le (1-\alpha)\hat{x}} \{A_2(w, u) - \phi_2(u - w)\} > \theta\hat{x} - (\phi_2 - \rho_2)((1 - \alpha)\hat{x} - w), \\
&\qquad \mathcal{N}_{\kappa\hat{x}}[w, (1 - \alpha)\hat{x}] = 0\} \\
&= \Pr\{\inf_{0 \le u \le (1-\alpha)\hat{x} - w} \{A_2(0, u) - \phi_2 u\} > \theta\hat{x} - (\phi_2 - \rho_2)((1 - \alpha)\hat{x} - w), \\
&\qquad \mathcal{N}_{\kappa\hat{x}}[0, (1 - \alpha)\hat{x} - w] = 0\} \\
&= \Pr\{\inf_{0 \le u \le (1-\alpha-\xi)\hat{x}} \{A_2(0, u) - \phi_2 u\} > (\theta - (\phi_2 - \rho_2)(1 - \alpha - \xi))\hat{x}, \\
&\qquad \mathcal{N}_{\kappa\hat{x}}[0, (1 - \alpha - \xi)\hat{x}] = 0\} \\
&= \Pr\{\mathbf{T}_2(\theta - (\phi_2 - \rho_2)(1 - \alpha - \xi))\hat{x}) > (1 - \alpha - \xi)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0, (1 - \alpha - \xi)\hat{x}] = 0\}.
\end{aligned}$$

Finally, consider term (C). The event $\mathbf{T}_2^c(-\gamma\hat{x}) > (1 - \alpha)\hat{x}$ means that

$$\inf_{0 \le u \le (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0, u) - \phi_2 u\} > -\gamma\hat{x}.$$

Now observe that

$$\inf_{0\le u\le(1-\alpha)\hat{x}}\{V_2^c(0)+A_2(0,u)-\phi_2 u\}\le V_2^c(0)+\inf_{0\le u\le w}\{A_2(0,u)-\phi_2 u\}.$$

Thus, the event $\mathbf{T}_2^c(-\gamma\hat{x})>(1-\alpha)\hat{x}$ implies

$$\inf_{0\le u\le w}\{A_2(0,u)-\phi_2 u\}>-V_2^c(0)-\gamma\hat{x},$$

so that

$$\begin{aligned}(C)\ &\le\ \Pr\{\inf_{0\le u\le w}\{A_2(0,u)-\phi_2 u\}>-V_2^c(0)-\gamma\hat{x},\mathcal{N}_{\kappa\hat{x}}[0,w]=0,V_2^c(0)\le\zeta\hat{x}\}\\ &\le\ \Pr\{\inf_{0\le u\le w}\{A_2(0,u)-\phi_2 u\}>-(\gamma+\zeta)\hat{x},\mathcal{N}_{\kappa\hat{x}}[0,w]=0\}\\ &=\ \Pr\{\inf_{0\le u\le\xi\hat{x}}\{A_2(0,u)-\phi_2 u\}>-(\theta+(\phi_2-\rho_2)\xi)\hat{x},\mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}]=0\}\\ &=\ \Pr\{\mathbf{T}_2((\theta-(\phi_2-\rho_2)\xi)\hat{x})>\xi\hat{x},\mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}]=0\}.\end{aligned}$$

Thus, taking $\eta=\xi$ and $\eta=1-\alpha-\xi$ in Lemma 7.2, and using Lemma 7.4, we obtain

$$\Pr\{\mathbf{T}_2^c(-\gamma\hat{x})>(1-\alpha)\hat{x}\}\lesssim\frac{\rho_2}{c-\rho_2}\Pr\{\mathbf{B}_2^r>((\phi_2-\rho_2)(1-\alpha)-\gamma-\theta-\frac{(c-\rho_2)(\gamma+\zeta+\theta)}{\phi_2-\rho_2})\hat{x}\}.$$

Letting $\zeta\downarrow 0$, $\theta\downarrow 0$, using the fact that $B_2^r(\cdot)\in\mathcal{IR}$, (16) follows.
Finally, note that (17) follows from (15) and (16) by letting $\alpha\downarrow 0$, $\gamma\downarrow 0$, and using the fact that $B_2^r(\cdot)\in\mathcal{IR}$.

□

9 Case II-A: fluid heavy-tailed input with $r_2<1-\rho_1$

We now consider the case where flow 2 generates traffic according to an On-Off process with peak rate $r_2<1-\rho_1$. The next theorem shows that flow 2 satisfies Assumptions 6.1-6.3 and that (5) holds.

Theorem 9.1 *If $A_2(\cdot)\in\mathcal{R}_{-\nu_2}$, then for any $\alpha>0$, $\gamma>0$,*

$$\Pr\{\mathbf{T}_2>(1+\alpha)\hat{x}\}\gtrsim(1-p_2)\Pr\{\mathbf{A}_2^r>\frac{\phi_2-\rho_2}{r_2-\rho_2}(1+\alpha)\hat{x}\},\tag{19}$$

$$\Pr\{\mathbf{T}_2(-\gamma\hat{x})>(1-\alpha)\hat{x}\}\lesssim(1-p_2)\Pr\{\mathbf{A}_2^r>(\frac{(\phi_2-\rho_2)(1-\alpha)-\gamma}{r_2-\rho_2}-\frac{\gamma}{\phi_2-\rho_2})\hat{x}\},\tag{20}$$

and

$$\Pr\{\mathbf{T}_2>\hat{x}\}\sim(1-p_2)\Pr\{\mathbf{A}_2^r>\frac{\phi_2-\rho_2}{r_2-\rho_2}\hat{x}\}.\tag{21}$$

Before giving the formal proof of the above theorem, we first provide an intuitive argument. Consider a queue of capacity ϕ_2 fed by the arrival process of flow 2. In order for the event $\mathbf{T}_2^c > \hat{x}$ to occur, the workload must remain positive throughout the interval $[0, \hat{x}]$, given that the initial workload is 0. Note that the normal drift in the workload is $\rho_2 - \phi_2 < 0$. Thus, there is a 'deficit' $(\phi_2 - \rho_2)\hat{x}$, which must be made up for by flow 2 showing above-average activity during the interval $[0, \hat{x}]$.
We claim that the most likely way for the gap to be filled is by a single long On-period of flow 2 covering the entire interval $[0, v]$, with $v := \frac{(\phi_2 - \rho_2)\hat{x}}{r_2 - \rho_2}$. (When On, flow 2 generates above-average traffic at rate $r_2 - \rho_2 > 0$, so this event (call it $E(\hat{x})$) makes up for the entire deficit.) This is consistent with the usual situation for heavy-tailed distributions that a large deviation is caused by just a single exceptional event. Observe that the probability of this event is indeed exactly the right-hand side of (21). Note that it is unlikely for the gap to be filled by several long On-periods, since the probability of this happening is an order of magnitude smaller.
The above arguments will be formalized in the proof below. We first prove that the event $E(\hat{x})$ indeed implies that $\mathbf{T}_2^c > \hat{x}$ for large $\hat{x}$, thus obtaining a lower bound for the probability of the latter event. Next we show that for large $\hat{x}$ the event $E(\hat{x})$ is also necessary for $\mathbf{T}_2^c > \hat{x}$ to occur, by proving that the probability of all other possible scenarios is negligibly small.

Proof
We first prove that for any $\alpha > 0$, $\delta > 0$, $\theta > 0$, the event

$$\mathbf{T}_2 > (1 + \alpha)\hat{x} \tag{22}$$

is implied by the event $E(\hat{x})$ that flow 2 is On at time 0 and turns Off again at time $v > \tau\hat{x}$, with

$$\tau := \frac{(\phi_2 - \rho_2 + \delta)(1 + \alpha) + \theta}{r_2 - \rho_2 + \delta},$$

combined with

$$\sup_{v \leq u \leq (1+\alpha)\hat{x}} \{(\rho_2 - \delta)(u - v) - A_2(v, u)\} \leq \theta\hat{x}.$$

The second event means that for all $u \in [v, (1 + \alpha)\hat{x}]$,

$$A_2(v, u) \geq (\rho_2 - \delta)(u - v) - \theta\hat{x}.$$

We distinguish between two cases.
i. $0 \leq u \leq v$.
Then

$$A_2(0, u) - \phi_2 u = r_2 u - \phi_2 u \geq 0.$$

ii. $v \leq u \leq (1 + \alpha)\hat{x}$.

Then

$$\begin{aligned}
A_2(0,u) - \phi_2 u &= A_2(0,v) + A_2(v,u) - \phi_2 u \\
&\geq r_2 v + (\rho_2 - \delta)(u-v) - \theta\hat{x} - \phi_2 u \\
&= (r_2 - \rho_2 + \delta)v - (\phi_2 - \rho_2 + \delta)u - \theta\hat{x} \\
&> ((\phi_2 - \rho_2 + \delta)(1+\alpha) + \theta)\hat{x} - (\phi_2 - \rho_2 + \delta)u - \theta\hat{x} \\
&\geq (\phi_2 - \rho_2 + \delta)(1+\alpha)\hat{x} - (\phi_2 - \rho_2 + \delta)(1+\alpha)\hat{x} \\
&= 0.
\end{aligned}$$

So,

$$\inf\{u \geq 0 : A_2(0,u) - \phi_2 u \leq 0\} > (1+\alpha)\hat{x},$$

which gives (22).
Hence, because of independence, using Lemma 7.6, for any $\alpha > 0$, $\delta > 0$, $\theta > 0$,

$$\begin{aligned}
\Pr\{\mathbf{T}_2 > (1+\alpha)\hat{x} \geq 0\} &\geq \Pr\{E(\hat{x})\}\Pr\{\sup_{v \leq u \leq (1+\alpha)\hat{x}} \{(\rho_2 - \delta)(u-v) - A_2(v,u)\} \leq \theta\hat{x}\} \\
&\gtrsim \Pr\{E(\hat{x})\} \\
&= (1-p_2)\Pr\{\mathbf{A}_2^r > \frac{(\phi_2 - \rho_2 + \delta)(1+\alpha) + \theta}{r_2 - \rho_2 + \delta}\hat{x}\}.
\end{aligned}$$

Letting $\delta \downarrow 0$, $\theta \downarrow 0$, using the fact that $A_2^r(\cdot) \in \mathcal{IR}$, (19) follows.
We now turn to the proof of (20).
By partitioning, we obtain for any $\alpha > 0$, $\gamma > 0$, $\theta > 0$, $\kappa > 0$, $v \geq w \geq 0$,

$$\begin{aligned}
&\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \\
= \; &\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \mathcal{N}_{\kappa\hat{x}}[0,w] \geq 1, \\
&\quad \mathcal{N}_{\kappa\hat{x}} v, (1-\alpha)\hat{x}] \geq 1\} \\
+ \; &\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \mathcal{N}_{\kappa\hat{x}}[0,w] \geq 1, \\
&\quad \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\} \\
+ \; &\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0\} \\
+ \; &\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\},
\end{aligned}$$

which is clearly upper bounded by

$$\begin{aligned}
&\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \mathcal{N}_{\kappa\hat{x}}[0,w] \geq 1, \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] \geq 1\} \\
+ \; &\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\} \\
+ \; &\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0\} \\
+ \; &\Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} \\
= \; &(A) + (B) + (C) + (D).
\end{aligned}$$

Take $v = \tau\hat{x}$ and $w = \xi\hat{x}$, with

$$\tau := \frac{(\phi_2 - \rho_2)(1-\alpha) - \gamma - \theta}{r_2 - \rho_2} < 1 - \alpha,$$

and

$$\xi := \frac{\gamma+\theta}{\phi_2-\rho_2} < \tau.$$

Now consider term (A). For the relevant events to occur, flow 2 must be On during the entire interval $[w, v]$, so that

$$\begin{aligned}
(A) &\leq (1-p_2)\Pr\{\mathbf{A}_2^r > v-w\} \\
&= (1-p_2)\Pr\{\mathbf{A}_2^r > (\frac{(\phi_2-\rho_2)(1-\alpha)-\gamma-\theta}{r_2-\rho_2} - \frac{\gamma+\theta}{\phi_2-\rho_2})\hat{x}\}.
\end{aligned}$$

Next, consider term (B). The event $\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}$ means that

$$\inf_{0\leq u\leq(1-\alpha)\hat{x}}\{A_2(0,u)-\phi_2 u\} > -\gamma\hat{x}.$$

Now observe that

$$\begin{aligned}
&\inf_{0\leq u\leq(1-\alpha)\hat{x}}\{A_2(0,u)-\phi_2 u\} \\
\leq\ &\inf_{v\leq u\leq(1-\alpha)\hat{x}}\{A_2(0,u)-\phi_2 u\} \\
\leq\ &\inf_{v\leq u\leq(1-\alpha)\hat{x}}\{A_2(0,v)-\phi_2 v + A_2(v,u)-\phi_2(u-v)\} \\
=\ &A_2(0,v)-\phi_2 v + \inf_{v\leq u\leq(1-\alpha)\hat{x}}\{A_2(v,u)-\phi_2(u-v)\} \\
\leq\ &(r_2-\phi_2)v + \inf_{v\leq u\leq(1-\alpha)\hat{x}}\{A_2(v,u)-\phi_2(u-v)\}.
\end{aligned}$$

Thus, the event $\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}$ implies

$$\inf_{v\leq u\leq(1-\alpha)\hat{x}}\{A_2(v,u)-\phi_2(u-v)\} > -(r_2-\phi_2)v-\gamma\hat{x},$$

so that

$$\begin{aligned}
(B) &\leq \Pr\{\inf_{v\leq u\leq(1-\alpha)\hat{x}}\{A_2(v,u)-\phi_2(u-v)\} > -(r_2-\phi_2)v-\gamma\hat{x}, \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}]=0\} \\
&= \Pr\{\inf_{0\leq u\leq(1-\alpha)\hat{x}-v}\{A_2(0,u)-\phi_2 u\} > -(r_2-\phi_2)v-\gamma\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}-v]=0\} \\
&= \Pr\{\inf_{0\leq u\leq(1-\alpha-\tau)\hat{x}}\{A_2(0,u)-\phi_2 u\} > (\theta-(\phi_2-\rho_2)(1-\alpha-\tau))\hat{x}, \\
&\qquad \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha-\tau)\hat{x}]=0\} \\
&= \Pr\{\mathbf{T}_2((\theta-(\phi_2-\rho_2)(1-\alpha-\tau))\hat{x}) > (1-\alpha-\tau)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha-\tau)\hat{x}]=0\}.
\end{aligned}$$

Finally, consider term (C).

$$\begin{aligned}
(C) &\leq \Pr\{\mathbf{T}_2(-\gamma\hat{x}) > w, \mathcal{N}_{\kappa\hat{x}}[0,w]=0\} \\
&= \Pr\{\mathbf{T}_2((\theta+(\phi_2-\rho_2)\xi)\hat{x}) > \xi\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}]=0\}.
\end{aligned}$$

Thus, taking $\eta=\xi$ and $\eta = 1-\alpha-\tau$ in Lemma 7.3, and using Lemma 7.5, we obtain

$$\Pr\{\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \lesssim (1-p_2)\Pr\{\mathbf{A}_2^r > (\frac{(\phi_2-\rho_2)(1-\alpha)-\gamma-\theta}{r_2-\rho_2} - \frac{\gamma+\theta}{\phi_2-\rho_2})\hat{x}\}.$$

Letting $\theta \downarrow 0$, using the fact that $A_2^r(\cdot) \in \mathcal{IR}$, (20) follows.

□

10 Case II-B: fluid heavy-tailed input with $r_2 > 1 - \rho_1$

We now consider the case where flow 2 generates traffic according to an On-Off process with peak rate $r_2 > 1-\rho_1$. The next theorem shows that flow 2 satisfies Assumptions 6.1-6.3 and that (6) holds.

Theorem 10.1 *If $A_2(\cdot) \in \mathcal{R}_{-\nu_2}$, then for any $c \in (\rho_2, r_2)$, $\alpha > 0$, $\gamma > 0$,*

$$\Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \gtrsim p_2 \frac{\rho_2}{c-\rho_2} \Pr\{\mathbf{A}_2^r > (\frac{(\phi_2-\rho_2)(1+\alpha)}{r_2-\rho_2} + \frac{\gamma}{r_2-c})\hat{x}\}, \quad (23)$$

$$\Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \lesssim p_2 \frac{\rho_2}{c-\rho_2} \Pr\{\mathbf{A}_2^r > \frac{(\phi_2-\rho_2)(1-\alpha)}{r_2-\rho_2}\hat{x}\}, \quad (24)$$

and

$$\Pr\{\mathbf{T}_2^c > \hat{x}\} \sim p_2 \frac{\rho_2}{c-\rho_2} \Pr\{\mathbf{A}_2^r > \frac{\phi_2-\rho_2}{r_2-\rho_2}\hat{x}\}. \quad (25)$$

Before giving the formal proof of the above theorem, we first provide an intuitive argument. Consider a queue of capactity ϕ_2 fed by the arrival process of flow 2. In order for the event $\mathbf{T}_2^c > \hat{x}$ to occur, the workload must remain positive throughout the interval $[0, \hat{x}]$, given that the initial workload is $V_2^c(0)$. Note that the normal drift in the workload is $\rho_2 - \phi_2 < 0$. Thus, there is a 'deficit' $(\phi_2 - \rho_2)\hat{x}$, which must be compensated for by the initial workload $V_2^c(0)$ plus possibly flow 2 showing above-average activity during the interval $[0, \hat{x}]$.
As before, we claim that the most likely way for the gap to be filled is by an extremely long On-period of flow 2 which started somewhere before time 0. Unfortunately, it is harder to pin down exactly how long that On-period must last, since it depends on when it started. No matter when the On-period started however, it turns out that we must always have $V_2^c(v) > (r_2 - c)v$, with $v := \frac{(\phi_2-\rho_2)\hat{x}}{r_2-\rho_2}$. Using Theorem 2.2, we see that the probability of this event is indeed exactly the right-hand side of (25).
The above arguments will be formalized in the proof below. We first prove that the event $V_2^c(v) > (r_2 - c)v$ indeed implies that $\mathbf{T}_2^c > \hat{x}$ for large $\hat{x}$, thus obtaining a lower bound for the probability of the latter event. Next we show that for large $\hat{x}$ the event $V_2^c(v) > (r_2 - c)v$ is also necessary for $\mathbf{T}_2^c > \hat{x}$ to occur, by proving that the probability of all other possible scenarios is negligibly small.

Proof of Theorem 10.1
We start with the proof of (23). For compactness, denote $v = \tau\hat{x}$, with

$$\tau := \frac{(\phi_2 - \rho_2 + \delta)(1+\alpha) + \theta}{r_2 - \rho_2 + \delta}.$$

We first prove that for any $\alpha > 0$, $\gamma > 0$, $\delta > 0$, $\theta > 0$, the event

$$\mathbf{T}_2^c(\gamma\hat{x}) > (1+\alpha)\hat{x} \quad (26)$$

is implied by the events

$$V_2^c(v) > (r_2 - c)v + \gamma\hat{x},$$

and

$$\sup_{v \le u \le (1+\alpha)\hat{x}} \{(\rho_2 - \delta)(u - v) - A_2(v, u)\} \le \theta\hat{x}.$$

The first event implies that there is an $-r \le v$ such that

$$A_2(-r, v) - c(r + v) > (r_2 - c)v + \gamma\hat{x},$$

so

$$A_2(-r, v) > cr + r_2 v + \gamma\hat{x}.$$

Because $A_2(-r, v) \le r_2(r + v)$, we then find $cr < r_2 r$, which gives $r > 0$ since $c < r_2$.
Thus, $V_2^c(0) \ge A_2(-r, 0) - cr$.
As $A_2(u, v) \le r_2(v - u)$ for all $u \in [0, v]$, we have

$$\begin{aligned} V_2^c(0) + A_2(0, u) &\ge A_2(-r, 0) - cr + A_2(0, u) \\ &= A_2(-r, u) - cr \\ &= A_2(-r, v) - A_2(u, v) - cr \\ &> cr + r_2 v + \gamma\hat{x} - r_2(v - u) - cr \\ &= r_2 u + \gamma\hat{x}. \end{aligned}$$

The second event means that for all $u \in [v, (1 + \alpha)\hat{x}]$,

$$A_2(v, u) > (\rho_2 - \delta)(u - v) - \theta\hat{x}.$$

We distinguish between two cases.
i. $0 \le u \le v$.
Then

$$\begin{aligned} V_2^c(0) + A_2(0, u) - \phi_2 u &\ge r_2 u + \gamma\hat{x} - \phi_2 u \\ &= (r_2 - \phi_2)u + \gamma\hat{x} \\ &\ge \gamma\hat{x}. \end{aligned}$$

ii. $v \le u \le (1 + \alpha)\hat{x})$.
Then

$$\begin{aligned} V_2^c(0) + A_2(0, u) - \phi_2 u &= V_2^c(0) + A_2(0, v) + A_2(v, u) - \phi_2 u \\ &> r_2 v + \gamma\hat{x} + (\rho_2 - \delta)(u - v) - \theta\hat{x} - \phi_2 u \\ &= (r_2 - \rho_2 + \delta)v - (\phi_2 - \rho_2 + \delta)u + (\gamma - \theta)\hat{x} \\ &= ((\phi_2 - \rho_2 + \delta)(1 + \alpha) + \theta)\hat{x} - (\phi_2 - \rho_2 + \delta)u + (\gamma - \theta)\hat{x} \\ &\ge (\phi_2 - \rho_2 + \delta)(1 + \alpha)\hat{x} - (\phi_2 - \rho_2 + \delta)(1 + \alpha)\hat{x} + \gamma\hat{x} \\ &= \gamma\hat{x}. \end{aligned}$$

So,

$$\inf\{u \ge 0 : V_2^c(0) + A_2(0, u) - \phi_2 u\} > (1 + \alpha)\hat{x},$$

which gives (26).

Hence, because of independence, using Lemma 7.6, for any $\alpha > 0$, $\gamma > 0$, $\delta > 0$, $\theta > 0$,

$$
\begin{aligned}
\Pr\{\mathbf{T}_2^c(\gamma\hat{x}) \geq (1+\alpha)\hat{x}\} &\geq \Pr\{V_2^c(v) \geq (r_2 - c)v + \gamma\hat{x}\} \\
&\quad \Pr\{\sup_{v \leq u \leq (1+\alpha)\hat{x}} \{(\rho_2 - \delta)(u - v) - A_2(v,u)\} \leq \theta\hat{x}\} \\
&\gtrsim \Pr\{V_2^c(v) \geq (r_2 - c)v + \gamma\hat{x}\} \\
&= \Pr\{\mathbf{V}_2^c \geq ((r_2 - c)\tau + \gamma)\hat{x}\} \\
&\sim p_2 \frac{\rho_2}{c - \rho_2} \Pr\{\mathbf{A}_2^r > (\tau + \frac{\gamma}{r_2 - c})\hat{x}\} \\
&= p_2 \frac{\rho_2}{c - \rho_2} \Pr\{\mathbf{A}_2^r > (\frac{(\phi_2 - \rho_2 + \delta)(1+\alpha) + \theta}{r_2 - \rho_2 + \delta} + \frac{\gamma}{r_2 - c})\hat{x}\}.
\end{aligned}
$$

Letting $\delta \downarrow 0$, $\theta \downarrow 0$, using the fact that $A_2^r(\cdot) \in \mathcal{IR}$, (24) follows.
We now turn to the proof of (24).
By partitioning, we obtain for all $\alpha > 0$, $\gamma > 0$, $\theta > 0$, $\kappa > 0$, $v \geq w \geq 0$,

$$
\begin{aligned}
& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) \geq (1-\alpha)\hat{x}\} \\
= & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) \geq (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, V_2^c(v) > (r_2 - c)(v - w)\} \\
+ & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) \geq (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\} \\
+ & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) \geq (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 1\} \\
+ & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) \geq (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\}.
\end{aligned}
$$

Now consider the third term. Suppose that $\mathcal{N}_{\kappa\hat{x}}[0,w] \geq 1$, i.e., there is a long On-period in the interval $[0,w]$. Since $\mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 1$, $\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1$, this long On-period must then last till at least time v. However, this contradicts the fact that $V_2^c(v) \leq (r_2 - c)(v - w)$.
Hence, the third term may be rewritten as

$$
\begin{aligned}
& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \leq 1, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 1\} \\
= & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] = 1, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 1, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0\} \\
= & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] = 1, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 1, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, V_2^c(0) \leq \zeta\hat{x}\} \\
+ & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] = 1, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 1, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, V_2^c(0) > \zeta\hat{x}\}.
\end{aligned}
$$

We thus arrive at the upper bound, for all $\zeta > 0$,

$$
\begin{aligned}
& \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}\} \\
\leq & \Pr\{V_2^c(v) > (r_2 - c)(v - w)\} \\
+ & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\} \\
+ & \Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, V_2^c(0) \leq \zeta\hat{x}\}
\end{aligned}
$$

$$\begin{aligned}
&+ \quad \Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] = 1, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0, V_2^c(0) > \zeta\hat{x}\} \\
&+ \quad \Pr\{\mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x}] \geq 2\} \\
&= \quad (A) + (B) + (C) + (D) + (E).
\end{aligned}$$

Take $v = \tau\hat{x}$ and $w = \xi\hat{x}$, with

$$\tau := \frac{(\phi_2 - \rho_2)(1-\alpha)}{r_2 - \rho_2} + \frac{\gamma + \theta}{\phi_2 - \rho_2} + \frac{(r_2 - \phi_2)\zeta}{(\phi_2 - \rho_2)(r_2 - \rho_2)} < 1 - \alpha,$$

and

$$\xi := \frac{\gamma + \zeta + \theta}{\phi_2 - \rho_2} < \tau.$$

Now consider term (A). Using Theorem 2.2,

$$\begin{aligned}
(A) &= \Pr\{\mathbf{V}_2^c > (r_2 - c)(\tau - \xi)\hat{x}\} \\
&\sim p_2 \frac{\rho_2}{c - \rho_2} \Pr\{\mathbf{A}_2^r > (\tau - \xi)\hat{x}\} \\
&= p_2 \frac{\rho_2}{c - \rho_2} \Pr\{\mathbf{A}_2^r > \frac{(\phi_2 - \rho_2)(1-\alpha) - \zeta}{r_2 - \rho_2}\hat{x}\}.
\end{aligned}$$

Next, consider term (B). The event $\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}$ means that

$$\inf_{0 \leq u \leq (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0,u) - \phi_2 u\} > -\gamma\hat{x}.$$

Now observe that

$$\begin{aligned}
& \inf_{0 \leq u \leq (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0,u) - \phi_2 u\} \\
\leq & \inf_{v \leq u \leq (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0,u) - \phi_2 u\} \\
\leq & \inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(0,v) - \phi_2 v + A_2(v,u) - \phi_2(u - v)\} \\
= \; & A_2(0,v) - \phi_2 v + \inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\} \\
= \; & A_2(0,v) - cv + (c - \phi_2)v + \inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\} \\
\leq & V_2^c(v) + (c - \phi_2)v + \inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\}.
\end{aligned}$$

Thus, the event $\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}$ implies

$$\inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\} > -V_2^c(v) - (c - \phi_2)v - \gamma\hat{x},$$

so that

$$\begin{aligned}
(B) \leq \; & \Pr\{\inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\} > -V_2^c(v) - (c - \phi_2)v - \gamma\hat{x}, \\
& \quad V_2^c(v) \leq (r_2 - c)(v - w), \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\} \\
\leq \; & \Pr\{\inf_{v \leq u \leq (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\} > -(r_2 - c)(v - w) - (c - \phi_2)v - \gamma\hat{x}, \\
& \quad \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\}
\end{aligned}$$

$$
\begin{aligned}
&= \Pr\{\inf_{v \le u \le (1-\alpha)\hat{x}} \{A_2(v,u) - \phi_2(u-v)\} > -(r_2 - \phi_2)v + (r_2 - c)w - \gamma\hat{x}, \\
&\qquad \mathcal{N}_{\kappa\hat{x}}[v,(1-\alpha)\hat{x}] = 0\} \\
&= \Pr\{\inf_{0 \le u \le (1-\alpha)\hat{x}-v} \{A_2(0,u) - \phi_2 u\} > -(r_2 - \phi_2)v - \gamma\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha)\hat{x} - v] = 0\} \\
&= \Pr\{\inf_{0 \le u \le (1-\alpha-\tau)\hat{x}} \{A_2(0,u) - \phi_2 u\} > (\theta - (\phi_2 - \rho_2)(1-\alpha-\tau))\hat{x}, \\
&\qquad \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha-\tau)\hat{x}] = 0\} \\
&= \Pr\{\mathbf{T}_2^c((\theta - (\phi_2 - \rho_2)(1-\alpha-\tau))\hat{x}) > (1-\alpha-\tau)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,(1-\alpha-\tau)\hat{x}] = 0\}.
\end{aligned}
$$

Finally, consider term (C). The event $\mathbf{T}_2^c(-\gamma\hat{x}) > (1-\alpha)\hat{x}$ means that

$$
\inf_{0 \le u \le (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0,u) - \phi_2 u\} > -\gamma\hat{x}.
$$

Now observe that

$$
\inf_{0 \le u \le (1-\alpha)\hat{x}} \{V_2^c(0) + A_2(0,u) - \phi_2 u\} \le V_2^c(0) + \inf_{0 \le u \le w} \{A_2(0,u) - \phi_2 u\}.
$$

Thus, the event $\mathbf{T}_2(-\gamma\hat{x}) > (1-\alpha)\hat{x}$ implies

$$
\inf_{0 \le u \le w} \{A_2(0,u) - \phi_2 u\} > -V_2^c(0) - \gamma\hat{x},
$$

so that

$$
\begin{aligned}
(C) &\le \Pr\{\inf_{0 \le u \le w} \{A_2(0,u) - \phi_2 u\} > -V_2^c(0) - \gamma\hat{x}, V_2^c(0) \le \zeta\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0\} \\
&\le \Pr\{\inf_{0 \le u \le w} \{A_2(0,u) - \phi_2 u\} > -(\gamma + \zeta)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,w] = 0\} \\
&= \Pr\{\inf_{0 \le u \le \xi\hat{x}} \{A_2(0,u) - \phi_2 u\} > (\theta - (\phi_2 - \rho_2)\xi)\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}] = 0\} \\
&= \Pr\{\mathbf{T}_2((\theta - (\phi_2 - \rho_2)\xi)\hat{x}) > \xi\hat{x}, \mathcal{N}_{\kappa\hat{x}}[0,\xi\hat{x}] = 0\}.
\end{aligned}
$$

Thus, taking $\eta = \xi$ and $\eta = 1 - \alpha - \tau$ in Lemma 7.3, and using Lemma 7.5, we obtain

$$
\Pr\{\mathbf{T}_2^c(-\gamma\hat{x}) \ge (1-\alpha)\hat{x}\} \lesssim p_2 \frac{\rho_2}{c - \rho_2} \Pr\{\mathbf{A}_2^r > \frac{(\phi_2 - \rho_2)(1-\alpha) - \zeta}{r_2 - \rho_2}\hat{x}\}.
$$

Letting $\zeta \downarrow 0$, using the fact that $A_2^r(\cdot) \in \mathcal{IR}$, (24) follows.

□

11 Conclusion

We analyzed a GPS queue with two flows, one having light-tailed characteristics, the other one exhibiting heavy-tailed properties. We showed that the workload distribution of the light-tailed flow is asymptotically equivalent to that when served in isolation at its minimum guaranteed rate, multiplied with a certain pre-factor. The pre-factor may be interpreted as the probability that the heavy-tailed flow is backlogged long enough for the light-tailed flow to reach overflow. We did not consider the case where the traffic intensity of the heavy-tailed flow exceeds its minimum guaranteed rate. In this case, the pre-factor – representing again the probability that the heavy-tailed flow is continuously backlogged during the period to overflow of the light-tailed flow – is likely to be some constant. Determining the exact value of the constant seems however a rather challenging task.

In the present paper we have focused on a scenario with two flows. Observe however that the light-tailed flow may be thought of as an aggregate flow, given that the class of Markov-modulated fluid input is closed under superposition of independent processes. In case of instantaneous input, the heavy-tailed flow too may actually represent an aggregate flow, since the superposition of independent Poisson streams with regularly varying bursts produces again a Poisson stream with regularly varying bursts. Unfortunately, the class of On-Off sources is clearly not closed under superposition. In fact, the superposition exhibits a fundamentally more complex structure than a single On-Off-source, which drastically complicates the analysis of the queueing behavior, see [23].

Despite the above and earlier observations, it would still be interesting to extend the analysis to general scenarios with several light-tailed flows, let's say $N_1 \geq 1$, and $N_2 \geq 1$ heavy-tailed flows.

In case $N_1 = 1$, $N_2 > 1$, we expect that the workload distribution of the light-tailed flow is equivalent to that when served in isolation at its minimum guaranteed rate, multiplied with a certain pre-factor, exactly as before. In this case however, the pre-factor represents the probability that each of the heavy-tailed flows is constantly backlogged during the period to overflow of the light-tailed flow. Calculating this probability seems a demanding task, since the most likely scenario cannot be easily pinned down due to the complicated interaction of the heavy-tailed flows prior to the overflow period.

In case $N_1 > 1$, $N_2 = 1$, we conjecture that the workload distribution of the light-tailed flows is equivalent to that in an isolated GPS queue consisting of the light-tailed flows only, multiplied again with a pre-factor. The pre-factor reflects the probability that the heavy-tailed flow is constantly backlogged during the time to overflow of the light-tailed flows. Unfortunately however, there are only logarithmic asymptotics known for a GPS queue with several light-tailed flows.

Not surprisingly, the two above-described complicating circumstances conspire in scenarios with $N_1 > 1$, $N_2 > 1$.

References

[1] Anantharam, V. (1988). How large delays build up in a GI/G/1 queue. *Queueing Systems* **5**, 345–368.

[2] Borst, S.C., Boxma, O.J., Jelenković, P.R. (1999). Induced burstiness in Generalized Processor Sharing queues with long-tailed traffic flows. In: *Proc. 37th Annual Allerton Conference on Communications, Control, and Computing*, Urbana-Champaign, Illinois, USA.

[3] Borst, S.C., Boxma, O.J., Jelenković, P.R. (2000). Asymptotic behavior of Generalized Processor Sharing with long-tailed traffic sources. In: *Proc. Infocom 2000 Conference*, Tel-Aviv, Israel, 912–921.

[4] Borst, S.C., Boxma, O.J., Jelenković, P.R. (2000). Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows. CWI Report PNA-R0016. Submitted for publication.

[5] Borst, S.C., Boxma, O.J., Van Uitert, M.J.G. (2001). Two coupled queues with heterogeneous traffic. Accepted for publication in: *Proc. ITC-17*.

[6] Borst, S.C., Zwart, A.P. (2000). A reduced-peak equivalence for queues with a mixture of light-tailed and heavy-tailed input flows. Technical Report SPOR 2000-04, Eindhoven University of Technology. Submitted for publication.

[7] Boxma, O.J., Deng, Q., Zwart, A.P. (1999). Waiting-time asymptotics for the M/G/2 queue with heterogeneous servers. Technical Memorandum COSOR 99-20, Eindhoven University of Technology. Submitted for publication.

[8] Dumas, V., Simonian, A. (2000). Asymptotic bounds for the fluid queue fed by subexponential on/off sources. *Adv. Appl. Prob.* **32**, 244–255.

[9] Elwalid, A.I., Mitra, D. (1991). Analysis and design of rate-based congestion control of high speed networks, I: stochastic fluid models, access regulation. *Queueing Systems* **9**, 29–64.

[10] Elwalid, A.I., Mitra, D. (1993). Effective bandwidth of general Markovian traffic sources and admission control of high speed networks. IEEE/ACM *Trans. Netw.* **1**, 329–343.

[11] Glynn, P.W., Whitt, W. (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. *J. Appl. Prob.* **31A**, 131–156.

[12] Jelenković, P.R., Lazar, A.A. (1999). Asymptotic results for multiplexing subexponential on-off processes. *Adv. Appl. Prob.* **31**, 394–421.

[13] Kesidis, G., Walrand, J., Chang, C.-S. (1993). Effective bandwidths for multi-class Markov fluids and other ATM sources. IEEE/ACM *Trans. Netw.* **1**, 424–428.

[14] Kosten, L. (1986). Liquid models for a type of information buffer problem. *Delft Progress Report* **11**, 71–86.

[15] Mandjes, M., Borst, S.C. (2000). Overflow behavior in queues with many long-tailed inputs. *Adv. Appl. Prob.* **32**, 1150–1167.

[16] Mandjes, M., Ridder, A. (1995). Finding the conjugate of Markov fluid processes. *Prob. Eng. Inf. Sci.* **9**, 297–315.

[17] Pakes, A.G. (1975). On the tails of waiting-time distributions. *J. Appl. Prob.* **12**, 555–564.

[18] Parekh, A., Gallager, R. (1993). A generalized processor sharing approach to flow control in integrated services networks: The single node case. IEEE/ACM *Trans. Netw.*, **1**, 344–357.

[19] Parekh, A., Gallager, R. (1994). A generalized processor sharing approach to flow control in integrated services networks: The multiple node case. IEEE/ACM *Trans. Netw.*, **2**, 137–150.

[20] Paxson, A., Floyd, S. (1995). Wide area traffic: the failure of Poisson modeling. IEEE/ACM *Trans. Netw.* **3**, 226–244.

[21] Resnick, S., Samorodnitsky, G. (1999). Activity periods of an infinite server queue and performance of certain heavy tailed fluid queues. *Queueing Systems* **33**, 43–71.

[22] Willinger, W., Taqqu, M.S., Sherman, R., Wilson, D.V. (1997). Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. IEEE/ACM *Trans. Netw.* **5**, 71–86.

[23] Zwart, A.P., Borst, S.C., Mandjes, M. (2000). Exact asymptotics for fluid queues fed by multiple heavy-tailed On-Off flows. Technical Report SPOR 2000-14, Eindhoven University of Technology. Submitted for publication. Shortened version in: *Proc. Infocom 2001 Conference*, Anchorage, Alaska, USA, 279–288.

A Definitions

Definition A.1 A distribution function $F(\cdot)$ on $[0,\infty)$ is called *long-tailed* ($F(\cdot) \in \mathcal{L}$) if

$$\lim_{x\to\infty} \frac{1-F(x-y)}{1-F(x)} = 1, \quad \text{for all real } y.$$

Definition A.2 A distribution function $F(\cdot)$ on $[0,\infty)$ is called *subexponential* ($F(\cdot) \in \mathcal{S}$) if

$$\lim_{x\to\infty} \frac{1-F^{2*}(x)}{1-F(x)} = 2,$$

where $F^{2*}(\cdot)$ is the 2-fold convolution of $F(\cdot)$ with itself, i.e., $F^{2*}(x) = \int_0^x F(x-y)F(\mathrm{d}y)$.

A useful subclass of $\mathcal{S}$ is the class $\mathcal{R}$ of *regularly-varying* distributions (which contains the Pareto distribution):

Definition A.3 A distribution function $F(\cdot)$ on $[0,\infty)$ is called *regularly varying of index* $-\nu$ ($F(\cdot) \in \mathcal{R}_{-\nu}$) if

$$F(x) = 1 - \frac{l(x)}{x^{\nu}}, \quad \nu \geq 0,$$

where $l : \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly-varying function, i.e., $\lim_{x\to\infty} l(\eta x)/l(x) = 1$, $\eta > 1$.

Examples of subexponential distributions which do not belong to $\mathcal{R}$ include the Weibull, lognormal, and Benktander distributions. A technical extension of $\mathcal{R}$ is the class $\mathcal{IRV}$ of *intermediately regularly-varying* distributions:

Definition A.4 A distribution function $F(\cdot)$ on $[0, \infty)$ is called *intermediately regularly varying* ($F(\cdot) \in \mathcal{IRV}$) if

$$\lim_{\eta \uparrow 1} \limsup_{x \to \infty} \frac{1 - F(\eta x)}{1 - F(x)} = 1.$$