TECHNICAL RESEARCH REPORT

Characterization of General TCP Traffic under a Large Number of Flows Regime

by Peerapol Tinnakornsrisuphap, Richard J. La, Armand M. Makowski

CSHCN TR 2002-21
(ISR TR 2002-43)

Characterization of General TCP Traffic Under a Large Number of Flows Regime

Peerapol Tinnakornsrisuphap, Richard J. La, and Armand M. Makowski *
Department of ECE & ISR, University of Maryland, College Park
{peerapol, hyongla, armand}@isr.umd.edu

September 23, 2002

Abstract

Short-lived TCP traffic (*e.g.,* web mice) composes the majority of the current Internet traffic. Accurate traffic modeling of a large number of short-lived TCP flows is extremely difficult due to (i) the interaction between session, transport, and network layers; and (ii) the explosion of the size of state space when the number of flows is large. Typically, ad-hoc assumptions are required for the analysis to be tractable under a certain regime.

We introduce a stochastic model of a bottleneck ECN/RED gateway under a large number of competing TCP flows. Our main result shows that as the number of flows becomes large, the queue dynamics and the aggregate traffic are simplified and can be accurately described by simple statistical recursions. These recursions can be evaluated independently of the number of flows, and hence the resulting traffic model is scalable. Furthermore, the limiting model is also consistent with other previously proposed models in their respective regime. Simulation results are also presented to confirm the results.

1 Introduction

Due to the growing size and popularity of the Internet, Internet traffic modeling has become an important research area. Internet traffic consists of many heterogeneous traffic sources, the majority of which utilize TCP congestion control mechanism [4]. One type of applications, such as FTP and Telnet, are relatively long-lived, while another type of applications are typically short-lived, *e.g.,* web browsing.

Characterization and modeling of TCP traffic yields an understanding of the interaction between the transport layer (TCP) and the network layer. Such interaction is well-understood in the context of a single long-lived TCP flow. However, when the number of

*This material is based upon work supported by the Space and Naval Warfare Systems Center – San Diego under Contract No. N66001-00-C-8063. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Space and Naval Warfare Systems Center – San Diego.

TCP flows is large, straightforward modeling usually results in a model that is not scalable because of the explosion of the size of the state space required to model all flows. Furthermore, for short-lived TCP flows, one has to further take into account an extra layer of dynamics, i.e., the session layer. These present two major obstacles to modeling short-lived TCP flows accurately.

Existing literature on short-lived TCP traffic modeling usually relies on ad-hoc assumptions, which render the model to be accurate only in certain regimes. Hollot *et al.* model short-lived TCP flows as exponential pulses, whose interarrival times are exponentially distributed, i.e., Poisson process [3]. They characterize the statistics of these exponential pulses using their time reversal, which is described by the *shot noise* process. This model assumes that the short-lived flows last only a few round-trip times and do not experience packet drops or marks, and hence implicitly assumes that congestion level is relatively low. Furthermore, flows are always in either congestion avoidance (long-lived connections) or slow start (short-lived connections), and they do not transition from one to the other. In other words, they do not explicitly model the session dynamics, where connections arrive and leave the network after transfers are completed. A similar approach to modeling short-lived flows is also taken in [7].

On the other end of the spectrum, Kherani and Kumar suggest that as the bottleneck capacity becomes very small, the queueing model for the bottleneck queue can be accurately described as a processor sharing queue [6]. When the capacity is large, however, the processor sharing model becomes less accurate because newly arrived TCP flows cannot fully utilize their allocated bandwidth. In fact, in the large capacity regime these short-lived flows may terminate even before they can increase their transmission rates to fully utilize their allocated bandwidth due to slow start.

The shortcomings of these models suggest a need for a unified model that is accurate in all regimes, instead of being restricted to one regime or another. Since the number of connections that share a bottleneck link is likely to be large, we follow the approach in [9] and [10], which consider "macroscale" modeling of aggregate TCP flows competing for the capacity of a bottleneck link [2]. Macroscale TCP models can be developed by systematically applying limit theorems to derive a limiting traffic model when the number of TCP flows is large. The potential benefits of doing so are three-fold. First, model simplification (with the promise of scalability) typically occurs when applying limit theorems, with irrelevant details filtered out without relying on ad-hoc assumptions. Second, limit theorems are central to the modern theory of probability, and as such have been the focus of a huge literature that contains a large number of results and techniques. Hence, given this large body of knowledge, it is reasonable to expect the existence of suitable limit theorems (under very weak assumptions) which can be applied to the situation of interest. Finally, in the networking context, resource allocation problems are interesting in networks operating at high utilization, *e.g.,* when the number of users is large. In such a scenario, the limit behavior will become increasingly more accurate as the number of users increases.

In this paper we extend the model in [10] and incorporate an additional layer of dynamics, namely the session layer. We show that the queue size per session and the workload per session brought in during a round-trip time converge to deterministic processes asymptotically as the number of sessions increases. Furthermore, we demonstrate that the sessions becomes asymptotically independent, indicating that the RED mechanism does alleviate the synchronization problem among the connections. The rest of the paper is organized as follows. The model and dynamics of network, transport, and session layers are described in Section 2, followed by our asymptotic results in Section 3. Section 4 gives a brief discussion on the results and a comparison with the other previously proposed models mentioned earlier. A numerical example and simulation results are presented in Section 5 and 6, respectively. We conclude the paper with a suggestion for future work in Section 7.

2 The Model

In our model, we have three layers of dynamics, namely network, transport, and session layers, which interact with each other through mechanisms that will be specified shortly. At the lowest level, the network is simplified to be a single bottleneck router with an ECN/RED marking mechanism controlling the congestion level. The traffic injected into the network is controlled by TCP congestion control mechanism in the transport layer, which reacts to the marks from the network. Each TCP connection is initiated by a session. A session can be either active or idle. If a session is busy, a file or an object is transferred through a TCP connection. A busy period of a session lasts until it no longer has any more data to transfer, at which time it goes idle. The duration of an idle period is random and represents the idle time between consecutive file transmissions. When a new file/object to be transferred arrives, the session becomes active again and sets up a new TCP connection. We now give detailed descriptions of the model for each layer and the interaction of these three layers.

Let $\mathcal{N} = \{1, \cdots, N\}$ be the set of sessions that share a bottleneck RED gateway. Time is assumed to be discrete and slotted in contiguous time slots of duration equal to the round-trip delay of TCP connections. We write $X^{(N)}$ to indicate the explicit dependence of the quantity X on the number N of sessions. Equivalence in law or in distribution between random variables (rvs) is denoted by $=_{st}$. The indicator function of an event A is given by $\mathbf{1}[A]$, and we use $\stackrel{P}{\to}_n$ (resp. $\Longrightarrow_n$) to denote convergence in probability (resp. weak convergence or convergence in distribution) with n going to infinity.

2.1 Session Dynamics

Each session $i \in \mathcal{N}$ is either active or idle. An idle session at the beginning of time slot $[t, t+1)$ does not have any packets to transmit in the time slot. An idle session in time slot $[t, t+1)$ becomes active at the beginning of time slot $[t+1, t+2)$ with probability $P_{ar}, 0 < P_{ar} < 1$, independently of the past. In other words, the duration of an idle period is geometrically distributed with parameter P_{ar} and has a mean of $1/P_{ar}$. This attempts to capture the dynamics of connection arrivals, where the interarrival times are reported to be exponentially distributed [8]. [1] Let $\{U_i(t), i \in \mathcal{N}; t = 0, 1, \cdots\}$ be a collection of i.i.d. rvs uniformly distributed on [0, 1]. Let $\mathbf{1}_{\{U_i(t+1) < P_{ar}\}}$ be the indicator function of the event that a new file/object arrives in the time slot $[t+1, t+2)$ for an idle session i.

Let $\{F_i(t), i \in \mathcal{N}; t = 0, 1, \cdots\}$ be a collection of i.i.d. non-negative integer-valued rvs with a general distribution function F. The workload of a connection of session i that

[1] Recall that one can approximate an exponential rv X with parameter α with $\lceil X \rceil$, which is a geometric rv with parameter $p = 1 - e^{-\alpha}$.

becomes active at the beginning of time slot $[t, t+1)$ is given by $F_i(t)$. This workload represents the *total* amount of workload a TCP connection brings in before it is torn down rather than workload brought in by an object or a file. In other words, if a same TCP connection is used to transfer more than one object while is alive, $F_i(t)$ represents the total amount of workload brought in by all objects during the active period. We denote the remaining workload of connection i at the beginning of time slot $[t, t+1)$ by $X_i(t)$. Clearly, $X_i(t) = 0$ if session i is idle during $[t, t+1)$. The evolution of $X_i(t)$ is given by the following:

$$X_i^{(N)}(t+1) = \mathbf{1}_{\{X_i^{(N)}(t)>0\}} \left(X_i^{(N)}(t) - A_i^{(N)}(t)\right) + \mathbf{1}_{\{X_i^{(N)}(t)=0\}} \mathbf{1}_{\{U_i(t+1)<P_{ar}\}} F_i(t+1) \ , \quad (1)$$

where $A_i^{(N)}(t)$ denotes the number of packets injected into the network by connection i at the beginning of time slot $[t, t+1)$. This will be explained in the following subsection.

2.2 TCP Dynamics

For each $i \in \mathcal{N}$, let $W_i^{(N)}(t)$ be an integer-valued rv that encodes the congestion window size (in packets) at the beginning of time slot $[t, t+1)$. We assume that the range of rv $W_i^{(N)}(t)$ is $\{0, 1, \cdots, W_{\max}\}$, where $W_{\max}$ is a finite integer representing the receiver advertised window size of the TCP connection. We assume that the congestion window size of an idle session is zero. When an idle session becomes active at the beginning of time slot $[t, t+1)$, the congestion window size of TCP connection is set to one at the beginning of time slot $[t+1, t+2)$. This models one round-trip delay for three-way handshake. Here we describe how the congestion window sizes of active connections evolve. Each TCP source transmits as many of the remaining data packets as possible that are allowed by its congestion window in that time slot. In other words, suppose that connection i has $X_i^{(N)}(t)$ remaining packets (or workload) waiting to be transmitted at the beginning of time slot $[t, t+1)$,[2] the number of packets connection i transmits at the beginning of time slot $[t, t+1)$, denoted by $A_i^{(N)}(t)$, is given by

$$A_i^{(N)}(t) = \min\left(W_i^{(N)}(t), X_i^{(N)}(t)\right) . \quad (2)$$

The congestion control mechanism of TCP operates in two different modes: slow start (SS) and congestion avoidance (CA). A new TCP connection starts in SS. In SS, the congestion window size is doubled every round-trip time until one or more packets are marked. If a mark is received, then the congestion window size is halved and TCP switches to CA. The congestion window size is limited by the receiver advertised window size $W_{\max}$. Hence, the evolution of the congestion window of connection i in SS can be written as

$$\begin{aligned} W_{i,SS}^{(N)}(t+1) &= \min\left(2W_i^{(N)}(t) \vee 1, W_{\max}\right) M_i^{(N)}(t+1) \\ &+ \min\left(\lceil \frac{W_i^{(N)}(t)}{2} \rceil, W_{\max}\right) (1 - M_i^{(N)}(t+1)) \ , \end{aligned} \quad (3)$$

where $a \vee b = \max(a, b)$ and $M_i^{(N)}(t+1)$ is an indicator function of the event that no marks have been received in time slot $[t, t+1)$, i.e., $M_i^{(N)}(t+1) = 1$ when no packet from Session

[2]We refer to a TCP connection of an active Session i by connection i when there is no confusion.

i is marked in the time slot and $M_i^{(N)}(t+1) = 0$ when at least one packet is marked. The marking mechanism will be explained in more detail in Subsection 2.3.

In CA, the congestion window size in the next time slot is increased by 1 if no marks are received in time slot $[t, t+1)$, and if one or more packets are marked in time slot $[t, t+1)$ the congestion window in the next time slot is reduced by half. The congestion window size in CA can be described by the following:

$$\begin{aligned} W_{i,CA}^{(N)}(t+1) &= \min\left(W_i^{(N)}(t) + 1, W_{\max}\right) M_i^{(N)}(t+1) \\ &+ \min\left(\lceil \frac{W_i^{(N)}(t)}{2} \rceil, W_{\max}\right)(1 - M_i^{(N)}(t+1)). \end{aligned} \tag{4}$$

We use $\{0,1\}$-valued rvs $\{S_i^{(N)}(t), i \in \mathcal{N}\}$ to encode the state of TCP connections. We interpret $S_i^{(N)}(t) = 0$ (resp. $S_i^{(N)}(t) = 1$) as connection i being in CA (resp. in SS) at the beginning of the time slot $[t, t+1)$. Therefore, the complete recursion of the congestion window size can be written as

$$\begin{aligned} W_i^{(N)}(t+1) &= \mathbf{1}_{\{X_i^{(N)}(t) - W_i^{(N)}(t) > 0\}} \\ &\times [S_i^{(N)}(t) W_{i,SS}(t+1) + (1 - S_i^{(N)}(t)) W_{i,CA}(t+1)], \end{aligned} \tag{5}$$

where the first indicator function is used to reset the congestion window size to zero when Session i runs out of data to transmit and returns to its idle state.

Finally, the evolution of $S_i^{(N)}(t)$ is given by

$$S_i^{(N)}(t+1) = \mathbf{1}_{\{X_i^{(N)}(t) - W_i^{(N)}(t) \le 0\}} + \mathbf{1}_{\{X_i^{(N)}(t) - W_i^{(N)}(t) > 0\}} S_i^{(N)}(t) M_i^{(N)}(t+1) \tag{6}$$

This equation can be interpreted as follows. Connection i is in SS in time slot $[t+1, t+2)$ if either (1) there is no packet left to transmit (so the connection resets) at the beginning of the time slot or (2) the connection was active and in SS in time slot $[t, t+1)$ and received no mark in the time slot. From (6) we assume that a new TCP connection in SS is ready to be set up after the previous connection is torn down after finishing its workload, and the new TCP connection becomes active when a new file/object arrives initiating three-way handshake.

2.3 Network Dynamics

In this subsection we explain how packets are marked to provide the congestion notification to the active TCP connections. The capacity of the bottleneck link is NC packets/slot for some positive constant C. The buffer size is assumed to be infinite so that no packets are dropped due to buffer overflow. Thus, congestion control is achieved solely through the random marking algorithm of the RED gateway.

Let $Q^{(N)}(t)$ denote the number of packets queued in the buffer at the beginning of time slot $[t, t+1)$. Connection i injects $A_i^{(N)}(t)$ packets into the network, and they are put in the buffer at the beginning of time slot $[t, t+1)$. Let the rv

$$A^{(N)}(t) := \sum_{i=1}^{N} A_i^{(N)}(t) \tag{7}$$

denote the aggregate number of packets offered to the network by the N sessions at the beginning of time slot $[t, t+1)$. Hence, $Q^{(N)}(t)+A^{(N)}(t)$ packets are available for transmission during that time slot. Since the bottleneck link has a capacity of NC packets/time slot, $\left[Q^{(N)}(t) + A^{(N)}(t) - NC\right]^+$ packets will not be served during time slot $[t, t+1)$, and will remain in the buffer. Hence, their transmission is deferred to subsequent time slots. The number of packets in the buffer at the beginning of time slot $[t+1, t+2)$, $Q^{(N)}(t+1)$, is therefore given by

$$Q^{(N)}(t+1) = \left[Q^{(N)}(t) - NC + A^{(N)}(t)\right]^+ . \tag{8}$$

Each incoming packet into the router in time slot $[t, t+1)$ is marked with a probability $f^{(N)}\left(Q^{(N)}(t)\right)$, depending on the queue length at the beginning of the time slot $[t, t+1)$. This model approximates the case where the memory of the queue averaging mechanism is long, which is the case for the recommended parameter settings of RED [1]. We represent this possibility by the $\{0,1\}$-valued rvs $M_{i,j}^{(N)}(t+1)$ $(j = 1, ..., A_i^{(N)}(t))$ with the interpretation that $M_{i,j}^{(N)}(t+1) = 0$ (resp. $M_{i,j}^{(N)}(t+1) = 1$) if the jth packet from source i is marked (resp. not marked) in the RED buffer. To do so we introduce the collection of i.i.d. $[0,1]$-*uniform* rvs $\{V_{i,j}(t+1),\ i,j = 1, \cdots;\ t = 0, 1, \cdots\}$ that are assumed to be independent of other rvs. The process by which packets are marked is as follows. For each $i \in \mathcal{N}$ and $j = 1, 2, \ldots$, we define the marking rvs

$$M_{i,j}^{(N)}(t+1) \quad = \quad \mathbf{1}_{\{V_{i,j}(t+1) > f^{(N)}(Q^{(N)}(t))\}},$$

so that the rv $M_{i,j}^{(N)}(t+1)$ is the indicator function of the event that the jth packet from source i is *not* marked in time slot $[t, t+1)$. The indicator function of the event that no packets from connection i in time slot $[t, t+1)$ are marked can now be written as

$$M_i^{(N)}(t+1) = \prod_{j=1}^{A_i^{(N)}(t)} M_{i,j}^{(N)}(t+1) \ . \tag{9}$$

3 The Asymptotics

The main result of the paper consists of the asymptotics for the normalized buffer content as the number of sessions becomes large. This result is discussed under the following assumptions (A1)-(A2):

(A1) There exists a continuous function $f : \mathbb{R}_+ \to [0,1]$ such that for each $N = 1, 2, \ldots$,

$$f^{(N)}(x) = f(N^{-1}x), \quad x \geq 0;$$

(A2) For each $N = 1, 2, \ldots$, the dynamics (1), (5), (6) and (8) start with the initial conditions

$$Q^{(N)}(0) = 0,\ W_i^{(N)}(0) = 0,\ S_i^{(N)}(0) = 1, \text{and } X_i^{(N)}(0) = 0; \quad i = 1, \ldots, N. \tag{10}$$

Assumption (A1) is a structural condition while (A2) is made essentially for technical convenience as it implies that for each N and all $t = 0, 1, \cdots$, the rvs $(W_1^{(N)}(t), S_1^{(N)}(t), X_1^{(N)}(t))$, $\cdots$, $(W_N^{(N)}(t), S_N^{(N)}(t), X_N^{(N)}(t))$ are *exchangeable*. Assumption (A2) can be omitted but at the expense of a more cumbersome discussion.

Theorem 1 *Assume that (A1)-(A2) hold. Then, for each $N = 1, 2, \ldots$ and $t = 0, 1, \ldots$, there exists a (non-random) constant $q(t)$ and rvs $(W(t), X(t), S(t))$ such that the following holds: (i) The following convergences take places:*

$$\frac{Q^{(N)}(t)}{N} \xrightarrow{P}_N q(t) \quad \text{and} \tag{11}$$

$$\left(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)\right) \Rightarrow_N (W(t), X(t), S(t)) \tag{12}$$

(ii) For any bounded function $g : \mathbb{N}^3 \to \mathbb{R}$

$$\tfrac{1}{N}\textstyle\sum_{i=1}^{N} g\left(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)\right) \xrightarrow{P}_N \mathbf{E}\left[g\left(W(t), X(t), S(t)\right)\right], \tag{13}$$

$$\text{Also,}\ \tfrac{1}{N}\textstyle\sum_{i=1}^{N} A_i^{(N)}(t) \xrightarrow{P}_N \mathbf{E}\left[\min\left(W(t), X(t)\right)\right]. \tag{14}$$

Moreover, if the workload distribution F has a finite second moment, then

$$\tfrac{1}{N}\textstyle\sum_{i=1}^{N} X_i^{(N)}(t) \xrightarrow{P}_N \mathbf{E}\left[X(t)\right] \tag{15}$$

(iii) For any integer $I = 1, 2, \ldots$, the triplets $\{(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)), i = 1, \ldots, I\}$ becomes asymptotically independent as N becomes large, with

$$\lim_{N\to\infty} \mathbf{P}[(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)) = (w_i, x_i, s_i),\ i = 1, \cdots, I]$$
$$= \prod_{i=1}^{I} \mathbf{P}\left[(W(t), X(t), S(t)) = (w_i, x_i, s_i)\right] \tag{16}$$

for any $w_1, \ldots, w_I, x_1, \ldots, x_I$ in $\mathbb{N}$ and $s_1, \ldots, s_I$ in $\{0, 1\}$.

In addition, with initial conditions $q(0) = 0,\ W(0) = 0,\ S(0) = 1,\ X(0) = 0$, it holds that

$$q(t+1) = \left(q(t) - C + \mathbf{E}\left[A(t)\right]\right)^+ \tag{17}$$

where $A(t) = \min\left(W(t), X(t)\right)$. Further, the recurrence

$$X(t+1) =_{st} \mathbf{1}_{\{X(t)=0\}} \mathbf{1}_{\{U(t+1)<P_{ar}\}} F(t+1) + \mathbf{1}_{\{X(t)>0\}} \left(X(t) - A(t)\right), \tag{18}$$

$$A(t) =_{st} \min\left(W(t), X(t)\right), \tag{19}$$

$$\begin{aligned} W_{SS}(t+1) =_{st}\ & \min\left(2W(t) \vee 1, W_{\max}\right) M(t+1) \\ & + \min\left(\lceil \frac{W(t)}{2} \rceil, W_{\max}\right)\left(1 - M(t+1)\right), \end{aligned} \tag{20}$$

$$\begin{aligned} W_{CA}(t+1) =_{st}\ & \min\left(W(t) + 1, W_{\max}\right) M(t+1) \\ & + \min\left(\lceil \frac{W(t)}{2} \rceil, W_{\max}\right)\left(1 - M(t+1)\right), \end{aligned} \tag{21}$$

$$W(t+1) =_{st} \mathbf{1}_{\{X(t)-W(t)>0\}} \cdot \left(S(t) W_{SS}(t+1) + (1 - S(t)) W_{CA}(t+1)\right), \tag{22}$$

$$S(t+1) =_{st} \mathbf{1}_{\{X(t)-W(t)>0\}} S(t) M(t+1) + \mathbf{1}_{\{X(t)-W(t)\le 0\}} \tag{23}$$

holds in law, where

$$M(t+1) = \mathbf{1}_{\{V(t+1) \le (1-f(q(t)))^{A(t)}\}} \tag{24}$$

for i.i.d. $[0, 1]$-uniform rvs $\{U(t+1), V(t+1);\ t = 0, 1, \ldots\}$.

Proof: The proof is given in A. ∎

4 Discussion

Theorem 1 shows that the dynamics of the queue at time t, denoted by $Q^{(N)}(t)$, can be approximated by $Nq(t)$ with $q(t)$ determined via a simple deterministic recursion, which is independent of the number of sessions. The offered traffic into the network during the time slot, $A^{(N)}(t)$, can also be approximated by $N \cdot \mathbf{E}[A(t)]$. These approximations become more accurate as the number of sessions becomes large, and the computational complexity does not depend on N. The limiting model is therefore "scalable" as it does not suffer from the explosion of state space, nor does it require any ad-hoc assumptions.

Theorem 1 also shows that the dependency between each session becomes negligible under a large number of sessions, i.e., "RED breaks the global synchronization when the number of sessions is large."

Although the sequence $\{(q(t), W(t), X(t), S(t)),\ t = 0, 1, \ldots\}$ is a time-homogeneous Markov chain with values in $\mathbb{R}_+ \times \{1, \ldots, W_{\max}\} \times \mathbb{N}_+ \times \{0, 1\}$, we shall not address here the existence of the steady-state when $t \to \infty$ as complications arise due the fact that the first component is degenerate (i.e., deterministic). However, we note that the numerical calculations for the limiting model are very simple. The number of steps required for the calculation for each time step is independent of N. We can determine $q(t)$ through the following steps:

(i) Let $t = 0$. Assume $q(0) = 0, W(0) = 0, X(0) = 0$, and $S(0) = 1$. Use $\mathbf{E}[A(0)] = 0$ to calculate $q(1) = 0$;

(ii) Use (18) through (23) with the corresponding $q(t)$ to calculate the transition probabilities and $\mathbf{P}[(W(t+1) = w, X(t+1) = x, S(t+1) = s)]$ for $w \in \{0, \ldots, W_{\max}\}$, $x \in \mathbb{N}_+$, and $s \in \{0, 1\}$. Then calculate $\mathbf{E}[A(t+1)]$;

(iii) Use $\mathbf{E}[A(t+1)]$ from (ii) to update $q(t+2)$ using (17);

(iv) Increase t by one then repeat steps (ii)-(iv).

We now consider the resulting model from Theorem 1 in the regime when C is either very large or very small with the following assumption

(A3) The marking function $f : \mathbb{R} \to [0, 1]$ is monotonically increasing with $f(0) = 0$ and $\lim_{x \to \infty} f(x) = 1$;

(A4) The average session layer input traffic $P_{ar} \int x dF(x)$ is strictly non-zero.

As $C \to \infty$, it is easy to see that $q(t) \to 0$. Hence the marking probability per flow also converges to zero from (A3) for all t. Therefore, each incoming flow will always operate in the slow start (exponential growth) mode and hence the resulting input traffic into the network will become a multiplexing of (discrete-time) Poisson arrival of streams of random number of packets, each of which doubles its window size every round-trip. The aggregate input traffic is therefore similar to the shot-noise processes, which agrees with [3].

On the other hand, as $C \to 0$ the queue will start building up, and hence as t goes to ∞, $q(t)$ approaches ∞. Thus, for large t, all TCP flows (including incoming TCP flows) will experience marking probability close to one from Assumption (A3). Therefore, all active connections will be able to inject only one packet per round-trip into the network because every packet transmitted will be marked, and TCP congestion window cannot grow larger than one. Since the bottleneck router will transmit packets non-selectively, any active flow will receive roughly equal throughput and hence the queue behavior approaches that of processor-sharing, which is in agreement with [6].

5 Numerical Example

This section presents a numerical example to study the behavior of the queue size per flow. The system and control parameters are set as follows:

$$C = 3 \text{ (packets/RTT)},\ q^N_{min} = 2 \cdot N,\ q^N_{max} = 10 \cdot N, p_{max} = 0.1,$$

The initial values are set to $W_i(0) = 0, X_i(0) = 0$, and $S_i(0) = 1$, and the queue size is initialized to $Q(0) = 0$ at the beginning. The variables $W_i(t), X_i(t)$, and $S_i(t)$ evolve according to (5), (1), and (6), respectively, and the queue size $Q(t)$ is updated according to (8). We assume the gentle mode of the RED gateway in order to have a continuous marking function. The workload $F_i(t) \sim Geometric(p), i = 1, 2, \cdots$ and $t = 0, 1, \cdots$, where $p = 0.01$, i.e., $E[F_i(t)] = 100$ packets. The idle periods of sessions are geometrically distributed with a mean of 20 time slots. The receiver advertised window size W_{max} is set to 64.

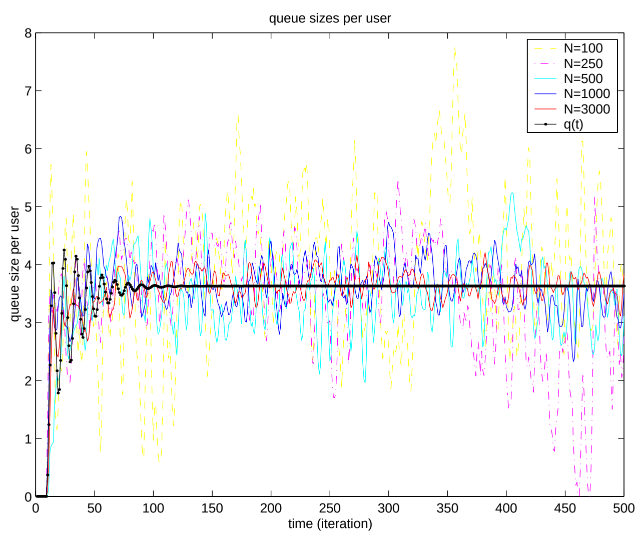

Figure 1: Evolution of queue size per flow.

Fig. 1 plots the evolution of the queue size per flow with $N = 100$, 250, 500, 1,000, and 3,000 as well as the deterministic process $q(t)$ computed through steps (i)-(iv) in Section 4. As expected, the oscillation in the queue size per flow decreases and the queue process converges to $q(t)$ with increasing N. Although we have not formally stated the convergence of $q(t)$ to a steady state, numerical examples indicate that for all reasonable set of parameters, the queue size per flow exhibits stationary behavior after a short transient period for all sufficiently large N.

6 Simulation

In this section we verify our analysis through *ns*-2 simulation results. In the simulation we gradually vary the number of sessions from 25 to 1,000, and study the queue behavior. Other system parameters are scaled with the number of sessions N as follows:

$$C^N = 0.24 \cdot N \text{ Mbps}, \; B^N = 25 \cdot N,$$
$$q_{min}^N = 2 \cdot N, \; q_{max}^N = 10 \cdot N, \; p_{max} = 0.1 \; ,$$

where C^N is the bottleneck link capacity, q_{min}^N and q_{max}^N are the threshold values used by the RED gateway [2], and B^N is the buffer size. The receiver advertised window size is set to 64 packets, and the packet size is set to 1,000 bytes. The exponential averaging weight of the RED gateway is set to $0.02/N$ in order to have a similar time constant in all cases. A session generates a workload that is exponentially distributed with a mean of 100 packets, and the interarrival times of the new workloads for each connection are exponentially distributed with a mean of 3.3 seconds. When a session runs out of data to transfer, it terminates the TCP connection. A new TCP connection is initiated by the session when the next workload arrives for the session. The round-trip propagation delays of the sessions are randomly selected from [52, 121.5] ms, with a mean of approximately 87 ms. The gateway implements the RED mechanism with ECN option, and the gentle mode of the RED mechanism is turned on so that the marking function is continuous in the average queue size. Also, the *drop_front_* option of the RED gateway is enabled, i.e., the RED gateway marks the packet at the front of the queue rather than the packet that has just arrived, in order to reduce the feedback delay. The TCP connections are Reno connections that are ECN capable, i.e., connections react to ECN marks set by the gateway.

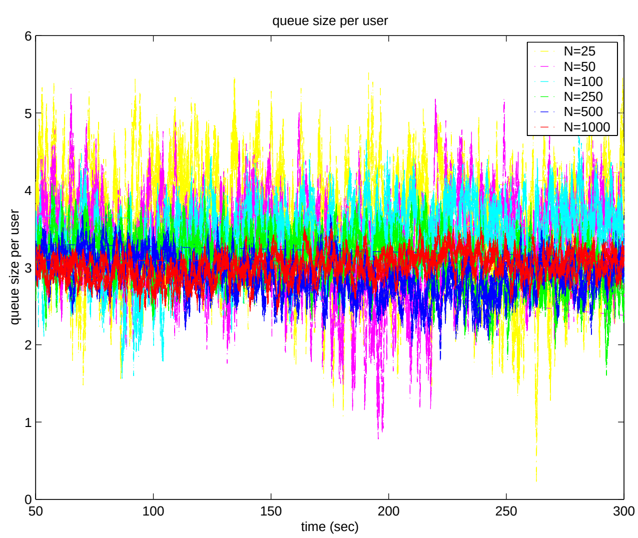

Figure 2: Evolution of queue size.

Fig. 2 shows the evolution of the queue size per flow with different number of sessions. As one can see the magnitude of oscillation in the queue size per flow decreases with increasing N, and the queue size per flow converges to a deterministic process as predicted by our results.

7 Conclusions

In this paper, we have developed a stochastic model for general TCP flows, which has taken into account the interaction between three layers, namely the network, transport, and session layers. The resulting model is scalable and is more accurate as the number of sessions grows large. A sharper approximation can also be developed with a central limit theorem-type complement similar to [10].

While the limiting result applies only to TCP flows with identical round-trip, it will be of use in a number situations, *e.g.,* the buffer dimensioning problem in an intercontinental Internet link, where the intercontinental link is typically a bottleneck, its large propagation delay dominates the round-trip delays of the connections, and the number of sessions is extremely large. Also, our limited simulation results indicate a similar limiting behavior for connections with heterogeneous round-trip delays.

Although we have yet to prove the existence of a steady state regime for the limiting recursion identified here, the limited simulation results suggest the existence of such a steady state under some conditions. Future work on this class of models also includes the incorporation of random round-trip delays and non-responsive UDP flows.

References

[1] Sally Floyd, "RED: Discussions of setting parameters," http://www.icir.org/floyd/REDparameters.txt.

[2] Sally Floyd and Van Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, Aug. 1995.

[3] C.V. Hollot, Vishal Misra, Y. Chait, and Don Towsley, "Effect of short-lived TCP flows on AQM performance," Preprint.

[4] Van Jacobson, "Congestion avoidance and control," in *Proceedings of SIGCOMM'88 Symposium*, Aug. 1988, pp. 314–332.

[5] Alan F. Karr, *Probability*, Springer-Verlag, 1993.

[6] Arzad Kherani and Anurag Kumar, "Stochastic models for throughput analysis of randomly arriving elastic flows in the Internet," in *Proceedings of IEEE INFOCOM*, New York City, NY, July 2002.

[7] M. Mellia, I. Stoica, and H. Zhang, "TCP model for short lived flows", *IEEE Communications Letters*, vol. 6(2), pp. 85-7, 2002.

[8] Vern Paxson and Sally Floyd, "Wide area traffic: The failure of Poisson modeling," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 226–244, 1995.

[9] Peerapol Tinnakornsrisuphap and Armand Makowski, "Queue dynamics of RED gateways under large number of TCP flows," in *Proceedings of IEEE GLOBECOM*, 2001.

[10] Peerapol Tinnakornsrisuphap and Armand Makowski, "On the behavior of ECN/RED gateways under a large number of TCP flows: Limit theorems," Tech. Rep., The Institute for Systems Research, University of Maryland, 2002.

A A. Proof of Theorem 1

A.1 Some simple and useful facts

To facilitate the presentation of the proof of Theorem 1, we begin with a few simple and useful facts. Fix $i = 1, \ldots, N$ and consider an arbitrary *bounded* mapping $g : \mathbb{N}^3 \to \mathbb{R}$: Through careful case analysis, it follows from (1), (5) and (6) that

$$\begin{aligned}
& g(W_i^{(N)}(t+1), X_i^{(N)}(t+1), S_i^{(N)}(t+1)) \\
& = \mathbf{1}_{\{X_i^{(N)}(t)=0\}} g\left(0, \mathbf{1}_{\{U_i(t+1)<P_{ar}\}} F_i(t+1), 1\right) \\
& + \mathbf{1}_{\{X_i^{(N)}(t)>W_i^{(N)}(t), S_i^{(N)}(t)=0\}} \\
& \quad \times \left(M_i^{(N)}(t+1) F_g^1\left(W_i^{(N)}(t), X_i^{(N)}(t)\right) + \left(1 - M_i^{(N)}(t+1)\right) F_g^2\left(W_i^{(N)}(t), X_i^{(N)}(t)\right)\right) \\
& + \mathbf{1}_{\{X_i^{(N)}(t)>W_i^{(N)}(t), S_i^{(N)}(t)=1\}} \\
& \quad \times \left(M_i^{(N)}(t+1) F_g^3\left(W_i^{(N)}(t), X_i^{(N)}(t)\right) + \left(1 - M_i^{(N)}(t+1)\right) F_g^2\left(W_i^{(N)}(t), X_i^{(N)}(t)\right)\right) \\
& + \mathbf{1}_{\{0<X_i^{(N)}(t)\le W_i^{(N)}(t)\}} g(0,0,1),
\end{aligned} \tag{25}$$

where the $\mathbb{N} \times \mathbb{N}$ mappings F_g^1, F_g^2 and F_g^3 are associated with g and defined as follows:

$$\begin{aligned}
F_g^1\left(W_i^{(N)}(t), X_i^{(N)}(t)\right) &= g\left(\min\left(W_i^{(N)}(t)+1, W_{\max}\right), X_i^{(N)}(t) - W_i^{(N)}(t), 0\right) \\
F_g^2\left(W_i^{(N)}(t), X_i^{(N)}(t)\right) &= g\left(\min\left(\lceil \frac{W_i^{(N)}(t)}{2} \rceil, W_{\max}\right), X_i^{(N)}(t) - W_i^{(N)}(t), 0\right) \\
F_g^3\left(W_i^{(N)}(t), X_i^{(N)}(t)\right) &= g\left(\min\left(2W_i^{(N)}(t) \vee 1, W_{\max}\right), X_i^{(N)}(t) - W_i^{(N)}(t), 1\right).
\end{aligned} \tag{26}$$

Despite the complicate form of (25), it can be constructed intuitively. First, note that the state of the triplet $(X_i^{(N)}(t+1), W_i^{(N)}(t+1), S_i^{(N)}(t+1))$ depends on the state of the triplet at time t. For example, if $X_i^{(N)}(t) = 0$, then $W_i^{(N)}(t+1)$ has to be zero and $S_i^{(N)}(t+1)$ is always one while the value of $X_i^{(N)}(t+1)$ depends on whether a new request arrives or not, hence we have the first term on the RHS of (25). The rest of (25) can also be understood in a similar fashion and the detail will be omitted here. This simple fact that any arbitrary mapping of the triplet at time $t+1$ can be expanded as a function of the triplet at time t with additional random elements that are independent of any events up to and including time t is crucial later in the proof of Theorem 1.

Let $\mathcal{F}_t$ denote the σ-field generated by the rvs $\{Q^{(N)}(0), W_i^{(N)}(0), X_i^{(N)}(0), U_i(s), F_i(s), V_i(s), V_{i,j}(s)\ , \ i,j = 1,2,\ldots\ ;\ s = 1,\ldots,t\ \}$. with the rvs $Q^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)$ and $W_i^{(N)}(t)$ $(i = 1,\ldots,N)$ being all $\mathcal{F}_t$-measurable, it holds under the enforced independence assumptions that

$$\mathbf{E}\left[M_{i,j}^{(N)}(t+1)|\mathcal{F}_t\right] = 1 - f^{(N)}(Q^{(N)}(t)), \quad j = 1,2,\ldots$$

so that

$$\mathbf{E}\left[M_i^{(N)}(t+1)|\mathcal{F}_t\right] = Z_i^{(N)}(t) \tag{27}$$

by conditional independence, where we have set

$$Z_i^{(N)}(t) = \left(1 - f^{(N)}(Q^{(N)}(t))\right)^{A_i^{(N)}(t)}. \tag{28}$$

It is now clear that

$$M_i^{(N)}(t+1) =_{st} \mathbf{1}_{\{V_i(t+1)\le Z_i^{(N)}(t)\}}. \tag{29}$$

It readily follows from (25) that

$$\begin{aligned}
&\mathbf{E}\left[g(W_i^{(N)}(t+1), X_i^{(N)}(t+1), S_i^{(N)}(t+1))|\mathcal{F}_t\right] \\
&= \mathbf{1}_{\{X_i^{(N)}(t)=0\}}\mathbf{E}\left[g\left(0, \mathbf{1}_{\{U_i(t+1)<P_{ar}\}}F_i(t+1), 1\right)\right] \\
&+ \mathbf{1}_{\{X_i^{(N)}(t)>W_i^{(N)}(t), S_i^{(N)}(t)=0\}}\left(Z_i^{(N)}(t)F_g^1(W_i^{(N)}(t), X_i^{(N)}(t)) + \left(1 - Z_i^{(N)}(t)\right)F_g^2(W_i^{(N)}(t), X_i^{(N)}(t))\right) \\
&+ \mathbf{1}_{\{X_i^{(N)}(t)>W_i^{(N)}(t), S_i^{(N)}(t)=1\}}\left(Z_i^{(N)}(t)F_g^3(W_i^{(N)}(t), X_i^{(N)}(t)) + \left(1 - Z_i^{(N)}(t)\right)F_g^2(W_i^{(N)}(t), X_i^{(N)}(t))\right) \\
&+ \mathbf{1}_{\{0<X_i^{(N)}(t)\le W_i^{(N)}(t)\}}g(0,0,1) \\
&= F_g(Z_i^{(N)}(t), W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))
\end{aligned} \tag{30}$$

where the mapping $F_g : [0,1]\times\mathbb{N}\times\mathbb{N}\times\{0,1\}\to\mathbb{R}$ is associated with g through

$$\begin{aligned}
F_g(z,w,x,s) &= \mathbf{1}_{\{x=0\}}\mathbf{E}\left[g\left(0, \mathbf{1}_{\{U_i(t+1)<P_{ar}\}}F_i(t+1), 1\right)\right] \\
&+ \mathbf{1}_{\{x>w\}}\mathbf{1}_{\{s=0\}}\left(zF_g^1(w,x) + (1-z)F_g^2(w,x)\right) \\
&+ \mathbf{1}_{\{x>w\}}\mathbf{1}_{\{s=1\}}\left(zF_g^3(w,x) + (1-z)F_g^2(w,x)\right) \\
&+ \mathbf{1}_{\{0<x\le w\}}g(0,0,1).
\end{aligned} \tag{31}$$

We note that $\mathbf{E}\left[g\left(0, \mathbf{1}_{\{U_i(t+1)<P_{ar}\}}F_i(t+1), 1\right)\right]$ always exists and is finite because the mapping g is bounded. Further, the mapping F_g is continuous with respect to the product topology on $[0,1]\times\mathbb{N}\times\mathbb{N}\times\{0,1\}$.

Upon taking expectations on both sides of (30) we see that

$$\mathbf{E}\left[g\left(W_i^{(N)}(t+1), X_i^{(N)}(t+1), S_i^{(N)}(t+1)\right)\right] = \mathbf{E}\left[F_g\left(Z_i^{(N)}(t), W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)\right)\right]. \tag{32}$$

A.2 A Weak Law of Large Numbers

We introduce the following terminology to facilitate the discussion: For each $t = 0, 1, \ldots$, the statements **[A:t]**, **[B:t]**, **[C:t]** and **[D:t]** below refer to the following convergence statements:

[A:t] For some non-random $q(t)$, it holds that

$$\frac{Q^{(N)}(t)}{N} \xrightarrow{P}{}_N q(t); \tag{33}$$

[B:t] For some $\{1,\ldots,W_{\max}\}$-valued rv $W(t)$, non-negative integer-valued rv $X(t)$, and $\{0,1\}$-valued rv $S(t)$, it holds that

$$\left(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)\right) \Rightarrow_N (W(t), X(t), S(t))\,; \tag{34}$$

[C:t] For any integer $I = 1, 2, \ldots$, the rvs $\{\left(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)\right),\ i = 1,\ldots,I\}$ become asymptotically independent with large N as described by (16) where $(W(t), X(t), S(t))$ are the rvs occurring in **[B:t]**;

[D:t] For any bounded mapping $g : \mathbb{N}^3 \to \mathbb{R}$, the convergence (13), (14) hold with $(W(t)$, $X(t)$, $S(t))$ the rvs occurring in **[B:t]**. Moreover, if the file arrival distribution has a finite second moment, then the convergence (15) also holds.

With the help of a series of lemmas, we shall prove the validity of the statements **[A:t]**–**[D:t]** for all $t = 0, 1, \ldots$. We do so by induction on t and in the process we establish Theorem 1.

Lemma 1 *Under (A1), if* **[A:t]** *and* **[B:t]** *hold for some* $t = 0, 1, \ldots$, *then* **[B:t+1]** *holds with* $W(t+1), X(t+1), S(t+1)$ *related in distribution to* $W(t), X(t), S(t)$ *by (22), (18), and (23).*

Proof: Together the convergence **[A:t]** and **[B:t]** imply [5, Thm. 5.28, p. 150] the joint convergence $(N^{-1}Q^{(N)}(t), W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)) \Rightarrow_N (q(t), W(t), X(t), S(t))$. Next the continuity of the mapping f implies that of $(y, w, x) \to (1 - f(y))^{\min(w,x)}$ on $\mathbb{R}_+ \times [0,\infty) \times [0,\infty)$, so that

$$(Z_1^{(N)}(t), W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)) \Rightarrow_N (Z(t), W(t), X(t), S(t)) \tag{35}$$

by the Continuous Mapping Theorem [5, Thm. 5.29, p. 150] with

$$Z(t) = (1 - f(q(t)))^{\min(W(t),X(t))}\ .$$

Consider (32) for any *bounded* arbitrary mapping $g : \mathbb{N}^3 \to \mathbb{R}$, and recall that the mapping F_g defined by (31) is continuous on $[0,1] \times \mathbb{N} \times \mathbb{N} \times \{0,1\}$. Consequently, the Continuous Mapping Theorem can again be invoked to yield

$$F_g(Z_1^{(N)}(t), W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)) \Rightarrow_N F_g(Z(t), W(t), X(t), S(t)), \tag{36}$$

whence

$$\lim_{N\to\infty} \mathbf{E}\left[F_g(Z_1^{(N)}(t), W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t))\right] = \mathbf{E}\left[F_g(Z(t), W(t), X(t), S(t))\right] \tag{37}$$

by the Bounded Convergence Theorem [5, Thm. 4.16, p. 108]. Combining (32) and (37) we get

$$\lim_{N\to\infty} \mathbf{E}\left[g(W_1^{(N)}(t+1), X_1^{(N)}(t+1), S_1^{(N)}(t))\right] = \mathbf{E}\left[F_g(Z(t), W(t), X(t), S(t))\right] \tag{38}$$

and since the bounded mapping g is arbitrary, it follows immediately that

$$\left(W_1^{(N)}(t+1), X_1^{(N)}(t+1), S_1^{(N)}(t+1)\right) \Rightarrow_N (W(t+1), X(t+1), S(t+1))$$

for some $\{1,\ldots,W_{\max}\}$–valued rv $W(t+1)$, non-negative integer-valued $X(t+1)$ and {0,1}-valued rv $S(t+1)$ with

$$\mathbf{E}\left[g\left(W(t+1), X(t+1), S(t+1)\right)\right] = \mathbf{E}\left[F_g\left(Z(t), W(t), X(t), S(t)\right)\right]. \tag{39}$$

A moment of reflection and a comparison to the analysis in (30)-(32) will convince the reader that (39) is equivalent to (18)-(23). ■

Lemma 2 *Under (A1), if* **[A:t]** *and* **[D:t]** *hold for some* $t = 0, 1, \ldots$, *then* **[A:t+1]** *also holds.*

Proof: From **[A:t]** and **[D:t]** (specifically, (14)), we conclude that

$$\frac{Q^{(N)}(t)}{N} - C + \tfrac{1}{N}\textstyle\sum_{i=1}^N A_i^{(N)}(t) \quad \xrightarrow{P}_N \quad q(t) - C + \mathbf{E}\left[A(t)\right] \tag{40}$$

and the desired result is a simple consequence of the continuity of the function $x \to x^+$ as we note that since

$$\frac{Q^{(N)}(t+1)}{N} = \left[\frac{Q^{(N)}(t)}{N} - C + \tfrac{1}{N}\textstyle\sum_{i=1}^N A_i^{(N)}(t)\right]^+$$

for all $N = 1, 2, \ldots$. ■

The proof of Lemma 2 also shows that

$$\frac{Q^{(N)}(t+1)}{N} \xrightarrow{P}_N q(t+1)$$

with non-random $q(t+1)$ determined by (17).

Lemma 3 *Under (A1)–(A2), if* **[A:t]**, **[B:t]** *and* **[C:t]** *hold for some* $t = 0, 1, \ldots$, *then* **[C:t+1]** *also holds.*

Proof: We first observe that for a fixed N, the triplets $(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)$ are coupled only through the marking probability which depends only on $Q^{(N)}(t)$. Fix a positive integer I. The rvs $V_1(t+1), \ldots, V_I(t+1)$ are i.i.d. $[0,1]$-uniform rvs which are independent of $\mathcal{F}_t$. Thus, upon making use of the representation (3)-(4) with (29), we see that the rvs $(W_1^{(N)}(t+1), X_1^{(N)}(t+1), S_1^{(N)}(t+1)), \ldots, (W_I^{(N)}(t+1), X_I^{(N)}(t+1), S_I^{(N)}(t+1))$ are mutually independent given $\mathcal{F}_t$. Consequently, for arbitrary bounded mappings $g_1, \ldots, g_I : \mathbb{N} \to \mathbb{R}$, we get

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=1}^I g_i(W_i^{(N)}(t+1), X_i^{(N)}(t+1), S_i^{(N)}(t+1)|\mathcal{F}_t\right] \\
=& \prod_{i=1}^I \mathbf{E}\left[g_i(W_i^{(N)}(t+1), X_i^{(N)}(t+1), S_i^{(N)}(t+1)|\mathcal{F}_t\right] \\
=& \prod_{i=1}^I F_{g_i}(Z_i^{(N)}(t), W_i^{(N)}(t), X_i^{(N)}(t+1), S_i^{(N)}(t+1))
\end{aligned}$$

with the help of (30) and (31).

Now it follows from (16) in **[C:t]** that the joint convergence

$$
\begin{aligned}
&(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t) \ldots, W_I^{(N)}(t), X_I^{(N)}(t), S_I^{(N)}(t)) \\
&\quad \Rightarrow_N (W_1(t), X_1(t), S_1(t) \ldots, W_I(t), X_I(t), S_I(t))
\end{aligned}
$$

holds with limiting rvs $(W_1(t), X_1(t), S_1(t)) \ldots, (W_I(t), X_I(t), S_I(t))$ which are i.i.d. rvs each distributed according to $(W(t), X(t), S(t))$. As in the proof of Lemma 1, the arguments leading to the convergence (36) also lead to

$$
\begin{aligned}
&(F_{g_1}(Z_1^{(N)}(t), W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)), \ldots, F_{g_I}(Z_I^{(N)}(t), W_I^{(N)}(t), X_I^{(N)}(t), S_I^{(N)}(t))) \\
&\Rightarrow_N (F_{g_1}(Z_1(t), W_1(t), X_1(t), S_1(t)), \ldots, F_{g_I}(Z_I(t), W_I(t)), X_I(t), S_I(t))
\end{aligned}
$$

where the limiting rvs $(Z_1(t), W_1(t), X_1(t), S_1(t)), \ldots, (Z_I(t), W_I(t), X_I(t), S_I(t))$ are i.i.d. rvs each distributed according to the pair $(Z(t), W(t), X(t), S(t))$. Therefore, by the Bounded Convergence Theorem,

$$
\begin{aligned}
& \lim_{N\to\infty} \mathbf{E}\left[\prod_{i=1}^{I} g_i(W_i^{(N)}(t+1), X_i^{(N)}(t+1), S_i^{(N)}(t+1))\right] \\
= \ & \lim_{N\to\infty} \mathbf{E}\left[\prod_{i=1}^{I} F_{g_i}(Z_i^{(N)}(t), W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))\right] \\
= \ & \mathbf{E}\left[\prod_{i=1}^{I} F_{g_i}(Z_i(t), W_i(t), X_i(t), S_i(t))\right] \\
= \ & \prod_{i=1}^{I} \mathbf{E}\left[F_{g_i}(Z_i(t), W_i(t), X_i(t), S_i(t))\right] \\
= \ & \prod_{i=1}^{I} \mathbf{E}\left[g_i(W_i(t+1), X_i(t+1), S_i(t+1))\right] \qquad (41)
\end{aligned}
$$

where the last equality made use of the relation (39). The desired result **[C:t+1]** now follows from (41) given that the mappings $g_1, \ldots, g_I$ are arbitrary. ∎

Lemma 4 *Under (A1)–(A2), if* **[A:t]**, **[B:t]** *and* **[C:t]** *hold for some $t = 0, 1, \ldots$, then* **[D:t]** *holds.*

Proof: Pick a mapping $g : \mathbb{N}^3 \to \mathbb{R}$. We begin by observing that under (A2) the rvs $(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))$; $i = 1, \ldots, N$ are exchangeable. As a result, we get

$$
\begin{aligned}
& \mathrm{var}\left[\tfrac{1}{N}\textstyle\sum_{i=1}^{N} g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))\right] \\
= \ & N^{-2}\sum_{i=1}^{N} \mathrm{var}[g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))] \\
+ \ & N^{-2}\sum_{i,j=1, i\neq j}^{N} \mathrm{cov}[g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)), g(W_j^{(N)}(t), X_j^{(N)}(t), S_j^{(N)}(t))] \\
= \ & N^{-1}\mathrm{var}[g(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t))] \\
+ \ & \frac{N-1}{N}\mathrm{cov}[g(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)), g(W_2^{(N)}(t), X_2^{(N)}(t), S_2^{(N)}(t))]. \qquad (42)
\end{aligned}
$$

Now let N go to infinity in (42): The validity of **[C:t]** and the Bounded Convergence Theorem already imply

$$\begin{aligned} & \lim_{N\to\infty} \mathrm{cov}[g(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t)), g(W_2^{(N)}(t), X_2^{(N)}(t), S_2^{(N)}(t))] \\ = \;& \mathrm{cov}[g(W_1(t), X_1(t), S_1(t)), g(W_2(t), X_2(t), S_2(t))] = 0 \end{aligned} \tag{43}$$

by asymptotic independence. On the other hand,

$$\limsup_{N\to\infty} \mathrm{var}[g(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t))] < \infty$$

since g is bounded.

Combining these observations we readily see that

$$\lim_{N\to\infty} \mathrm{var}\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))\right] = 0,$$

whence, by Chebyshev's Inequality,

$$\tfrac{1}{N}\textstyle\sum_{i=1}^N g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)) - \mathbf{E}\left[\tfrac{1}{N}\textstyle\sum_{i=1}^N g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t))\right] \stackrel{P}{\rightarrow}{}_N 0.$$

This last convergence is equivalent to

$$\tfrac{1}{N}\textstyle\sum_{i=1}^N g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)) - \mathbf{E}\left[g(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t))\right] \stackrel{P}{\rightarrow}{}_N 0$$

by exchangeability, and the desired convergence result (13) is now immediate once we remark under **[B:t]** that $\lim_{N\to\infty} \mathbf{E}\left[g(W_1^{(N)}(t), X_1^{(N)}(t), S_1^{(N)}(t))\right] = \mathbf{E}\left[g(W(t), X(t), S(t))\right]$. It is then straightforward to get (14) with $g(W_i^{(N)}(t), X_i^{(N)}(t)) = A_i^{(N)}(t) = \min(W_i^{(N)}(t), X_i^{(N)}(t))$ which is bounded by $W_{\max}$.

Finally, if the file arrival distribution F has a finite second moment, the convergence (15) follows from the dominated convergence theorem on (43) with $g(W_i^{(N)}(t), X_i^{(N)}(t), S_i^{(N)}(t)) = X_i^{(N)}(t)$ and note that $X_i^{(N)}(t)$ is dominated by a random variable with the distribution F which has a finite second moment. ■

We now conclude with a proof of Theorem 1: We first note that under (A1)-(A2) the statements **[A:t]**–**[D:t]** trivially hold for $t = 0$. Moreover, if **[A:t]**–**[C:t]** hold for some $t = 0, 1, \ldots$, then so do the statements **[D:t]** **[B:t+1]**, **[A:t+1]** and **[C:t+1]** by Lemma 4, Lemma 1, Lemma 2 and Lemma 3, respectively. Consequently, the statements **[A:t]**–**[D:t]** do hold for all $t = 0, 1, \ldots$ by induction and the validity of Claims (i)-(iii) of Theorem 1 is now established. The dynamics (17) is a byproduct of the proof of Lemma 2, while (18)-(24) are already contained in Lemma 1.