Letter

Performance Analysis of the User Part Congestion Control Scheme in SS7

Branimir M. Trenkić

Abstract **This paper focuses on the transient performance analysis of the User Part (UP) congestion and flow control mechanisms in ITU-T Signaling System No.7 (SS7). In particular, we developed an analytic model that we use to study the steady-state performance of the UP congestion control scheme. Numerical results of the investigation are presented and discussed.**

Keywords **Digital communication systems, Signaling system No. 7, Congestion control**

1. Introduction

The Blue Book congestion controls are complex and their description is spread widely and piecemeal across the recommendations, [1]. Here we focus on the UP control (used in conjunction with the international Message Transfer Part (MTP) control). The international MTP option is based on a single MTP message priority, a single congestion onset threshold (O) and single congestion abatement threshold (A), with $A < O < L$ (L is the buffer capacity). The link congestion status is therefore either congested or uncongested. For every message that arrives at a congested link, a TFC (TransFer Controlled) message is sent to the source SP (Signaling Point), indicating the destination affected by the congestion. On receipt of a TFC message the MTP informs all local UP's about the congestion situation by means of MTP-STATUS primitives with indication CONGESTION (CI's). Both TUP (Telephony User Part) and ISUP (ISDN User Part) have recommendations for UP congestion control which are very similar, [1], and we consider them together. The UP reduces traffic to the affected destination in steps, under the control of two timers designated in ISUP as T_{29} and T_{30} (in TUP as $Tue1$ and $Tue2$) which are started on receipt of the first CI. If a congestion indication is received after the expiry of T_{29} but before T_{30} expires, the traffic load is reduced one more step and both T_{29} and T_{30} are restarted. If T_{30} expires, then the traffic load is increased by one step and T_{30} is restarted. This is repeated until full load has been resumed.

In this Letter, we propose an accurate method for steady-state probabilities calculation of the overall UP congestion controlled process. Next, we use that method to study the steady-state performance of the UP congestion control scheme.

Received October 18, 2000.

Branimir M. Trenkić, IRITEL Telecommunications and Electronics Institute, 11080 Zemun, Batajnički put 23, Yugoslavia.
E-mail: trenkic@iritel.com

2. System description

In order to develop tractable model we consider the queueing behavior of SS7 messages in an outgoing link of an STP (Signaling Transfer Point) (like as shown in Fig. 4, [2]). Let S denote the number of source-destination UP using this link. In case of full traffic load, messages arrive to the link according to a Poisson process with rate $S\lambda_0$ (we assume that the message generation process for each UP is identical). For the sake of simplicity, we ignore the $(TFC + CI)$ notification delay. For this reason, it is unreasonable to study the effect of the T_{29} timer and we restrict our attention to the case where $T_{29} = 0$. Further, even though the lengths of T_{30} is deterministic, we assume that the length of T_{30} has exponential distribution with mean $\theta = 1/E[T_{30}]$.

We assume that during the period of the congestion status n (uncongested $(n = 0)$ or congested $(n = 1)$), multiplexed messages arrive according to a Markovian arrival process (MAP) with representation (C_n, D_n), [3]. Let K be the maximum reduction step of traffic load in an UP. Define the state of an UP as k $(1 \le k \le K)$ if the UP has reduced its traffic load k times since beginning with full load. We assume that each UP whose state is k sends messages according to a Poisson process with rate λ_k $(\lambda_1 > \lambda_2 > \ldots > \lambda_K)$. Let $X_k(t)$ be the number of UPs in state k at time t. Hence $J_n(t) = (X_1(t), \ldots, X_K(t))$ $(n = 0, 1)$ can be defined as the underlying process of the arrival to with state space consisting of $(i_1, \ldots, i_K)$. Conditions $0 \le i_j \le S$ and $0 \le \sum_{j=1}^{K} i_j \le S$ determined the total number of states in the underlying processes. The state such that $\sum_{j=1}^{K} i_j = 0$ is the impossible in the $J_1(t)$ because we ignored the $(TFC + CI)$ notification delay. According to the UP congestion control scheme, non-zero entries of parameter matrices are given by

C_0: $(i_1, i_2, \ldots, i_K) \longrightarrow (i_1, \ldots, i_{k-1}+1, i_k-1, \ldots, i_K)$ with rates $i_k\theta$, $2 \le k \le K$.

D_0: $(i_1, i_2, \ldots, i_K) \longrightarrow (i_1, i_2, \ldots, i_K)$ with rates $\left(S - \sum_{j=1}^{K} i_j\right)\lambda_0 + \sum_{j=1}^{K} i_j\lambda_j$.

C_1: $(i_1, i_2, \ldots, i_K) \longrightarrow (i_1, \ldots, i_{k-1}+1, i_k-1, \ldots, i_K)$ with rates $i_k\theta$, $1 \le k \le K-1$.

D_1: $(i_1, i_2, \ldots, i_K) \longrightarrow (i_1, i_2, \ldots, i_K)$ with rates $\left(S - \sum_{j=1}^{K} i_j\right)\lambda_0$, and

D_1: $(i_1, i_2, \ldots, i_K) \longrightarrow (i_1, \ldots, i_{k-1}+1, i_k-1, \ldots, i_K)$ with rates $i_k\lambda_k$, $2 \le k \le K$.

1434-8411/02/56/3-208 $15.00/0

The diagonal elements of the matrix C_0 and C_1 are negative values to make $C_0e + D_0e = 0$ and $C_1e + D_1e = 0$ respectively (where e is a unit column vector with the appropriate dimension).

3. Transient analysis of the UP Congestion Control

We assume that the message lengths are exponentially distributed with mean C/μ bits where C is the transmission speed of the STP link in bits/second. Then the steady-state model of the UP congestion control scheme can be described by an continuous time Markov chain $Z(t) = (B(t), \psi(t), J(t), t \geq 0)$. $B(t)$ and $\psi(t)$ denote the queue length and the congestion status of the STP buffer at time t. All the transitions can be bidirectional except for those entering the congestion and uncongestion regions (indicated by $\psi(t) = 0 \to \psi(t) = 1$ and $\psi(t) = 1 \to \psi(t) = 0$) which mast be unidirectional. Dividing the overall chain $Z(t)$ at those unidirectional transitions we obtained two alternating transient subprocesses $Z_0(t) = (0 \leq B(t) \leq O, 0, J_0(t), t \geq 0)$ and $Z_1(t) = (A \leq B(t) \leq L, 1, J_1(t), t \geq 0)$.

The steady-state probabilities of the overall controlled process, $Z(t)$, can then be represented by means of the solution of the mean sojourn time in both subprocesses. Denote $x^0 = \left(x_0^0, x_1^0, \ldots, x_O^0\right)$ and $x^1 = \left(x_A^1, x_{A+1}^1, \ldots, x_L^1\right)$ as the mean sojourn time matrices of both transient subprocesses. Since that $Z_0(t)$ and $Z_1(t)$ always starts on level $A-1$ and level $O+1$, respectively, x^0 and x^1 we can obtain, for instance, by the Generalized Folding-algorithm, [4]. Finally, we can obtain the queue length distribution as follows

$$\pi_i = \begin{cases} cp^0x_i^0, & 0 \leq i \leq A-1 \\ cp^0x_i^0 + cp^1x_i^1, & A \leq i \leq O \\ cp^1x_i^1, & O+1 \leq i \leq L \end{cases} \tag{1}$$

where p^0 and p^1 are probability vectors of the initial phases of the level where each subprocess starts. To obtain this, we introduce an embedded Markov chain at those instances. c is determined by the normalization condition.

Once the queue length distribution is obtained, the performance measures can be easily obtained. We define $\beta_E = \mu(1 - \pi_0 e)$ as the effective throughput (information-bearing messages only). In the UP congestion control scheme the effective throughput is equal to the total throughput. Let b and w be the random variables that denote the queue length and the message delay in the link respectively. Mean values of these quantities are obtained as follows.

$$E[b] = \sum_{i=0}^{L} i\pi_i e \qquad E[w] = \frac{E[b]}{\beta_E} \tag{2}$$

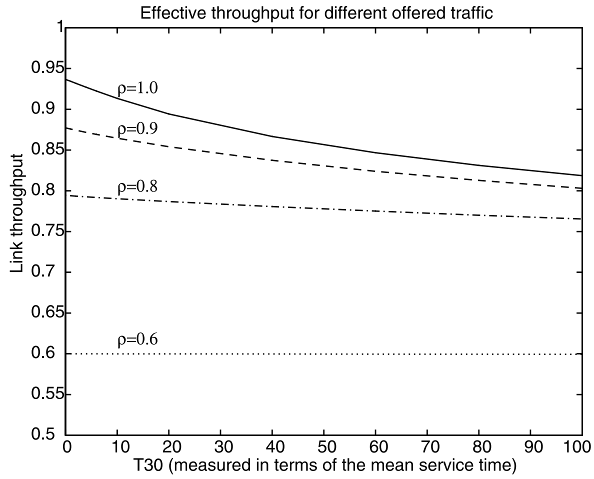

Fig. 1. The effective throughput against T_{30} for different offered traffic.

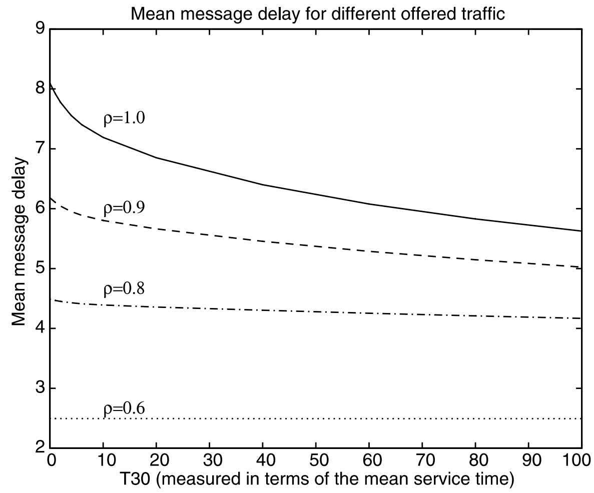

Fig. 2. The mean message delay against T_{30} for different offered traffic.

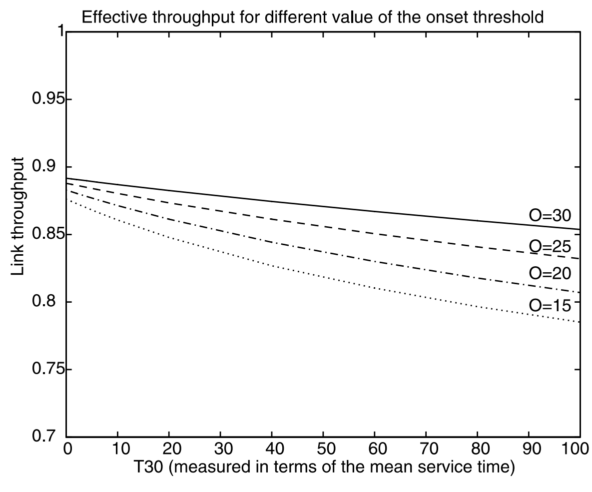

Fig. 3. Perturbation of onset threshold at $\rho = 0.9$.