Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT

Journal of Algorithms 50 (2004) 232–245

Journal of Algorithms

www.elsevier.com/locate/jalgor

Instability of FIFO in session-oriented networks

Matthew Andrews

Bell Laboratories, 600–700 Mountain Avenue, Murray Hill, NJ 07974, USA

Received 30 June 2000

Abstract

We show that the First-In-First-Out (FIFO) scheduling discipline can be unstable in the (σ, ρ)-regulated session model for packet-switched networks. In this model packets are injected into the network in fixed sessions. The total size of the session-i packets injected during the time interval $[x, y)$ is at most $\sigma_i + \rho_i(y - x)$ for some burst parameter σ_i and rate ρ_i. The sum of the rates of sessions passing through a server is at most the server speed.

Previous work on FIFO stability either allowed for dynamically changing session paths or else assumed that session-i packets are injected at a constant rate. Our result shows that FIFO can be unstable for static paths as long as the injections into a session can be temporarily suspended.

1. Introduction

First-In-First-Out (FIFO) is the simplest of all queueing disciplines. Indeed, its simplicity means that it is widely used to provide best-effort service in packet-switched networks. Hence, in order to characterize the performance of many networks it is important to understand the performance of FIFO. One crucial issue is that of *stability*. Under what conditions is there a bound on the total size of packets in the network? Equivalently, under what conditions can we provide a bound on the delay suffered by packets in the network?

In this paper we show that FIFO can be unstable in the (σ, ρ)-regulated session model. This model was proposed by Cruz [8,9] and is one of the most commonly studied models in the networking literature (e.g., [2,7–9,12,15,16,19,21,22,25,27,28]). We now describe the model and contrast it with other network models that have been defined. In particular, we show how it differs from the adversarial queueing model [3] for which FIFO stability was analyzed in [1].

E-mail address: andrews@research.bell-labs.com.

0196-6774/$ – see front matter
doi:10.1016/S0196-6774(03)00096-8

1.1. The model

We model the network as a collection of servers. (Equivalently, one can think of the servers as the edges or links of the network.) In the (σ, ρ)-regulated session-model packets are injected by an *adversary* along fixed paths in the network. Each path consists of a sequence of servers. We refer to these paths together with the packets injected into them as *sessions*. Each session i is associated with a *burst parameter*, σ_i and a *rate* ρ_i. Let $A_i(x, y)$ be the total size of the session-i packets injected into the network during the time interval $[x, y)$. We assume that session i is (σ_i, ρ_i)*-regulated* which means that,

$$A_i(x, y) \leqslant \sigma_i + \rho_i(y - x), \quad \forall x, y. \tag{1}$$

In other words, the long-term rate at which traffic is injected into the network along session i is bounded by ρ_i. The term σ_i allows us to inject short *bursts* of packets into the network. A rate condition of this form can be enforced using a "leaky bucket" (e.g., [8,21,26]).

If two or more packets are queued at a server then they are processed in FIFO order. We normalize the speed of the servers to 1, i.e., the total size of the packets that can be processed in one time unit is 1. Although this is with loss of generality, it suffices for our purposes since our aim is to give an example of FIFO instability.

In order for the stability question to be interesting, no server in the network should be overloaded. Let $S(e)$ be the set of sessions that pass through server e. We define the load, ρ_G, on the network by

$$\rho_G = \max_e \sum_{i \in S(e)} \rho_i. \tag{2}$$

We require the following *load condition*

$$\rho_G \leqslant 1. \tag{3}$$

Clearly, if the load condition does not hold then we cannot hope for stability. Packets that require some server e can be injected faster than they can be processed. If $\rho_G = 1$ then we say that the network is *critically* loaded. If $\rho_G < 1$ then we say that the network is *subcritically* loaded.

1.2. Main result

We state the main result of this paper.

Theorem 1. *There exists a network, a set of sessions, and an injection pattern on each session such that*:

(1) *the injections conform to the (σ, ρ)-regulated session model defined above*;
(2) $\rho_G = 1 - 3 \times 10^{-4}$, *i.e., the network is subcritically loaded*;
(3) *the network is* unstable, *i.e., the total size of packets in the network is unbounded. This immediately implies that the delays experienced by packets are unbounded.*