Queue length based pacing of internet traffic

Yan Cai, *University of Massachusetts - Amherst*

Abstract

As the Internet evolves, there is a continued demand for high Internet bandwidth. This demand is driven partly by the widely spreading real-time video applications, such as on-line gaming, teleconference, high-definition video streaming. All-optical switches and routers have long been studied as a promising solution to the rapidly growing demand. Nevertheless, buffer sizes in all-optical switches and routers are very limited due to the challenges in manufacturing larger optical buffers. On the other hand, Internet traffic is bursty. The existence of burstiness in network traffic has been shown at all time scales, from tens of milliseconds to thousands of seconds. The widely existing burstiness has a very significant impact on the performance of small buffer networks, resulting in high packet drop probabilities and low link utilization.^ There have been many solutions proposed in the literature to solve the burstiness issue of network traffic. Traffic engineering techniques, such as traffic shaping and polishing, have been available in commercial routers/switches since the era of Asynchronous Transfer Mode (ATM) networks. Moreover, TCP pacing, as a natural solution to the TCP burstiness, has long been studied. Furthermore, several traffic conditioning and scheduling techniques are proposed to smooth core network traffics in a coordinated manner. However, all the existing solutions are inadequate to efficiently solve the burstiness issue of high-speed traffic.^ In this dissertation we aim to tackle the burstiness issue in small buffer networks, which refer to the future Internet core network consisting of all-optical routers and switches with small buffers.^ This dissertation is composed of two parts. In the first part, we analyze the impact of a general pacing scheme on the performance of a tandem queue network. This part serves as a theoretical foundation, based on which we demonstrate the benefits of pacing in a tandem queue model. Specifically, we use the Infinitesimal Perturbation Analysis (IPA) technique to study the impact of pacing on the instantaneous and average queue lengths of a series of nodes. Through theoretical analyses and extensive simulations, we show that under certain conditions there exists a linear relationship between system parameters and instantaneous/average queue lengths of nodes and that pacing improves the performance of the underlying tandem queue system by reducing the burstiness of the packet arrival process. ^ In the second part, we propose a practical on-line packet pacing scheme, named Queue Length Based Pacing (QLBP). We analyze the impact of QLBP on the underlying network traffic in both time and frequency domains. We also present two implementation algorithms that allow us to evaluate the performance of QLBP in real experimental and virtual simulation environments. Through extensive simulations, we show that QLBP can effectively reduce the burstiness of network traffic and hence significantly improve the performance of a small buffer network. More important, the network traffic paced with QLBP does not exhibit a weakened competition capability when competing with non-paced traffic, which makes the QLBP scheme more attractive for ISPs.^

Subject Area

Engineering, Computer

Recommended Citation

Yan Cai, "Queue length based pacing of internet traffic" (January 1, 2011). *Doctoral Dissertations Available from Proquest.* Paper AAI3482591. http://scholarworks.umass.edu/dissertations/AAI3482591