IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Hershey PA 17033-1117, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB7312

Chapter XI

RAC: A Soft-QoS Framework for Supporting Continuous Media Applications

Wonjun Lee
Ewha Womans University, Korea

Jaideep Srivastava
University of Minnesota, USA

In this chapter, we present a novel disk admission control algorithm that exploits the degradability property of continuous media applications to improve the performance of the system. The algorithm is based on setting aside a portion of the resources, i.e., disk I/O bandwidth, as reserves and managing it intelligently so that the total utility of the system is maximized. This reserve-based admission control strategy (RAC) is a compromise between purely greedy and non-greedy strategy, and it leads to an efficient protocol that improves the performance of the system. While the protocol is simple for admission decision, it also results in better performance for the system by reserving some resources for the important future applications. The framework consists of two parts. One is the reserve-based admission control mechanism in which new streams, arriving during periods of congestion, are offered lower QoS, instead of being blocked. The other one is a resource scheduler for continuous media with dynamic resource allocation to achieve higher utilization than non-dynamic schedulers by effectively sharing available resources among contending streams and by reclamation, which is a scheduler-initiated negotiation to reallocate resources among streams to improve overall QoS.

This chapter appears in the book, Multimedia Networking: Technology, Management and Applications by Syed Mahbubur Rahman.

INTRODUCTION

The emergence of varied high-speed networked multimedia systems opens up the possibility that a much more varied collection of continuous media applications could be handled in real-time. However, due to the limitation of available resources in such systems, we still need to develop more intelligent mechanisms for efficient admission control, negotiation, resource allocation and resource scheduling, so that we could optimize the total system utilization. In particular, there has been increased interest in I/O issues for multimedia or continuous media (CM). The goal of conventional disk scheduling policies is to reduce the cost of seek operations and to achieve a high throughput, while providing fair access to every process that seeks its services. The goal of disk scheduling for continuous media is to meet the deadlines of the periodic I/O requests generated by the stream manager to meet rate requirements. An additional goal is to minimize buffer requirements. In order to ensure continuous and stringent real-time constraints of video delivery in CM servers (Harinath et al., 2000; Haskin & Schmuck, 1996; and Lelend et al., 1994), several factors such as disk bandwidth, buffer capacity, network bandwidth, etc., should be considered carefully and should be handled efficiently. The reservations of these resource factors are required for supporting an acceptable level of display quality and for providing on-time delivery constraints. In particular, disk bandwidth constraint may be the most important factor, given that the I/O bandwidth reserved for each stream on disk depends on latency overhead time, transfer time, defined cycle length and contention of multiple streams. Hence, we should be able to guarantee that the request of each stream can be fairly supported with good disk utilization and server cost-performance.

The goal of disk/server scheduling for CM is to satisfy QoS requirements (Vin, Goyal & Goyal, 1994) by meeting deadlines of periodic I/O requests generated by some server resident CM stream manager with minimum buffers and with a fair scheduling algorithm (Kenchammana-Hosekote & Srivastava, 1997). A video stream being viewed requires timely delivery of data, but it is able to tolerate some loss of the data for small amounts of time. The problem here is dual (i.e., admission control and disk I/O bandwidth management in CM server systems) to that of network call admission control and dynamic bandwidth management (Lau & Lui, 1995). Admission control in CM servers or video-on-demand systems restricts the number of applications supported on the resources. For example, the applications may be video streams, and the resources used could be connections on the network or on continuous media servers (Lee & Sabata, 1999, 2001; Lee et al., 1998). To handle these issues in CM server systems, we propose an adaptive admission control strategy which achieves better performance than the conventional greedy admission control strategies generally used for CM servers. It recognizes that CM (e.g., video) applications can tolerate certain variations on QoS parameters. It develops an algorithm for sharing processing resources at the server to share available resources effectively among contending streams (Lee & Srivastava, 1998). The proposed algorithm provisions are for *reclamation* (i.e., scheduler-initiated negotiation) to reallocate resources among streams to improve the QoS overall.

In this chapter, we present a dynamic and adaptive admission control strategy for providing a fair disk bandwidth scheduling and better performance for video streaming. It efficiently services multiple clients simultaneously and satisfies the real-time requirement for continuous delivery of video streaming at specified bandwidths in distributed environments. This new stream scheduler provides a proficient admission control functionality which can optimize disk utilization with a good response ratio for the requests of clients--