Adaptive Video on Demand

Sudhansu Aggarwal* Juan A. Garay† Amir Herzberg†

We formulate the problem of Video on Demand (VOD) from a resource allocation perspective. In particular, we introduce the *decision element* into a movie-vending environment, which complements the current approaches. In contrast with the more traditional resource allocation problems (e.g., machine scheduling, call control, etc.), the problem possesses the distinctive *bunching* property, which stands for the feasibility of several requests being served by one resource (channel).

Most published works and on-going projects seem to share a common view on the basic architecture for VOD services: The movies are stored in a central *video server*, which may be connected to other servers over a high bandwidth WAN. The *video server* is connected by a high-capacity fiber link to local distribution centers (hubs), from which coax cables are used to broadcast to the households. In this architecture, there appears to be two major bottlenecks: a) The limited number of broadcast channels available on the coax cable (shared by many homes); and b) the number of movies which the server is able to transmit concurrently. Previous works and current implementations attempt to achieve, under these constraints:

- **Full Video on Demand:** Whenever a user's request for a movie arrives, it is immediately served, provided there is capacity (a channel) available, rejected otherwise.
- **Near Video on Demand:** A fixed set of movies is played regularly, at fixed time intervals.

It is immediate that either policy may poorly utilize the available resources. The purpose of this work is to investigate the intermediate terrain that lies between the two extreme policies. Indeed, we introduce the *decision element* in the video service architecture, which leads to the concept of:

- **Adaptive Video on Demand:** Upon arrival, a movie request is accepted and served (possibly with some delay), or rejected. The decision as to whether accept or reject (and the amount of delay) is made by a scheduling algorithm.

Our motivation is to complement approaches based on predicting demand for movies (i.e., Near VOD), are best suited for popular movies, but inappropriate for more unusual selections. This leads us to investigate the problem in an *on-line* fashion, namely, having to accept or reject a movie request without the knowledge of future requests.

COMP. RATIO	*No ref.*	*Ref. by choice*	
		Imm. Not.	*Delayed Not.*
Lower bound	μ	$\frac{u}{c+1} \approx \mu$	$\ln \mu/2$
Algorithm	μ	μ	$O(\ln \mu)$

Table 1: Our results.

We show upper and lower bounds on the *competitive ratio* of on-line movie scheduling algorithms for the most natural scenarios. Let c denote the number of available channels, m the number of movies and u the number of users. Also let $\mu \stackrel{\text{def}}{=} u/c$. Our results are summarized in Table 1. In particular, for the more natural case of *refusal by choice* with *delayed notification*, we present Harmonic, a channel pre-partition algorithm that exhibits an asymptotically optimal behavior. The algorithm stipulates that any movie shown on a channel in class C_i must have been requested by *at least i* users.

Furthermore, we compare the performances of the different algorithms under more realistic distributions of requests. Our studies show a strong performance advantage of Harmonic compared to the more straightforward Exponential algorithm, Near, and variants of Full VOD.

*LCS, MIT, 545 Technology Square, Cambridge, MA 02139. E-mail: sonu@theory.lcs.mit.edu. Work done while the author was visiting the IBM T.J. Watson Research Center.

†IBM T. J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598. {amir,garay}@watson.ibm.com.

PODC 94 - 8/94 Los Angeles CA USA
© 1994 ACM 0-89791-654-9/94/0008..$3.50