CS384M: Multimedia Systems

Fall 2001

<u>Project 2</u>

Buffer Management Techniques for Video Servers

Due: Tuesday, October 25, 2001
Time: In Class

Objective of this project

The objective of this project is to design, implement and evaluate a buffer management policy for video servers.

Project Description

Consider a video server that services video retrieval requests by proceeding in terms of periodic rounds. As we have discussed in class, a video server can utilize the available buffer space for various purposes.

- Video servers utilize the available buffer space to buffer data retrieved during a round prior to transmission.
- Video servers utilize the available buffer space to prefetch blocks from future rounds to reduce the load imbalance across disks. Prefetching data blocks also allow a server to eliminate deadline violations in future rounds.

These uses of buffer space affect the total number of clients a video server can support simultaneously. For instance, increasing the round duration, and thereby increasing the amount of buffer space allocated to hold data retrieved during a round, increases the number of clients that the server can service from the disks. However, this reduces the amount of buffer space available for prefetching to balance load across disks within a round. This can reduce the overall number of clients the video server can support. Similarly, static partitioning of the available buffer space between the round-buffering and prefetching tasks can lead to underutilization of buffers in the presence of variable bit rate (VBR) encoded video streams.

Your task is to explore and expose these tradeoffs. <u>You need to develop a buffer management scheme that allocates dynamically the available buffer space to these competing tasks such that the total number of clients that can be supported by the video server is maximized.</u>

You are free to choose any disk scheduling algorithm that you seem fit for your video server. Assume that a data block that cannot be accessed prior to the completion of a round (i.e., prior to its deadline) is discarded (i.e., not accessed from disk at all). Similarly, a data block for which no buffer space is available is discarded. In this setting, you should determine the number of clients that a particular video server configuration supports such that the total number of frames that will be discarded is smaller than 0.1% (i.e., 99.9% of the frames are accessed from the server).

You should implement your buffer management scheme in a simulator and evaluate its effectiveness. For the purposes of the simulation, you can make the following assumptions.

- When you simulate a video server servicing n clients, assume that all the n clients arrive at time t=0 and remain in the system until the end of the simulation.
- The video file that each client accesses is randomly selected (from the 4 video file traces provided to you). Further, each client requests the retrieval to begin from a random point in the file. You can use a uniform random number generator to determine the file to access and the starting point within the file.

- Assume that a client that reaches the end of the file (i.e., the end of the video trace) continues playback from the beginning of the file until the end of your simulation.

You should carry out your simulations for different video server configurations (i.e., number of disks available at the server, the amount of buffer space available at the server, etc.).

Your simulations should answer the following questions.

1. For a given a server configuration (i.e., number of disks, amount of buffer space, and characteristics of client requested files), how will you select a round duration?
2. How does the total buffer size required increases with the round duration? How does the number of clients supported by the server increase with the round duration? What is the average utilization of the buffer space (i.e., what is the average of the percentage of buffers utilized during each round) with your buffer management scheme? What is the average buffer space utilization if you had partitioned your available buffer space statically between the buffering and prefetching functions?
3. How does increasing the number of disks in the system affect your answer to Question 2?
4. What is the minimal server configuration that you would need to design a video server that can service up to n clients? Provide your answer for n=32, 64, 128, 256, and 512.

Extra Credit (25%)

In practice, video servers can also use the available buffer space as a cache to maintain video streams (or part thereof) to service future requests. The larger the number of requests that can be served from such a cache, the smaller is the load imposed on the disks (and hence, the larger is the total number of clients that the server can service simultaneously).

How will you modify your buffer management algorithm to accommodate the need for caching video streams? How much buffer space would you allocate for caching? What is the impact of your scheme on the number of clients the server can support?

Simulator and Input Traffic

For your simulations, you can build upon the disk array simulator available from here. If you have any problems in downloading or installing the simulator, please contact the TA.

As for inputs to your simulator, you can use the following video file traces. These traces provide frame sizes (in bytes). Three traces are for Variable Bit Rate (VBR) encoded video streams and one sequence is for a Constant Bit Rate (CBR) encoded video stream.

Here are the traces: VBR Video Trace 1, VBR Video Trace 2, VBR Video Trace 3, CBR Video Trace

Submission Requirements

You will need to submit a report describing your buffer management scheme and results.

Your report should contain the following.

- A clear description of the buffer management scheme that you have developed.
- A description of the experiments that you have conducted to evaluate your scheme.
- Explanation/justification for your results.

You must typeset your report and include any graphs that you think are necessary to support your results. You should clearly answer each of the above questions and any others that you think are important.

Grading

Your report will be graded for clarity, quality of the solution, and the quality of the experimental evaluations. You should explain clearly your buffer management scheme, its impact on the admission control algorithm and placement scheme, the experimental setup and the results. As always, it is not only important to “have the right answer” but to also explain your answer crisply. Include the graphs that you think are important; explain their importance clearly and provide a crisp explanation of what they show.

Good luck!!