A DELIVERY METHOD OF VIDEOS WITH REQUIRED MINIMUM BANDWIDTHS

Hideaki ITO Teruo FUKUMURA

School of Computer and Cognitive Sciences, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan
E-mail:{itoh, fukumura}@sccs.chukyo-u.ac.jp

ABSTRACT

Networked video delivering systems have been developing. When videos are delivered from servers to client sites over a network, they are played without skew time at client sites. Moreover, link overflow does not occur, since capability of links are restricted. The systems allocate resources for meeting these requirements. For effective delivery, it is necessary that a policy of scheduling is established. A proposed policy is that the video is transmitted with the minimum bandwidth. The minimum bandwidth is corresponding to a consuming ratio of the video. The scheduling method consists of three steps. They are to find paths for delivering the video clips, to calculate required time to deliver them without link overflow, and to set their start times from servers without skew time. Moreover, simulation of this method is achieved by making schedules to deliver video. This policy is simple and flexible to deliver videos.

1. INTRODUCTION

Networked video delivering systems and distributed video-on-demand systems have been developing, which deal with video, as application systems in a network environment[7]. To improve the efficiency of such systems, several kinds of scheduling algorithms for delivering video are also proposed[4, 5]. Videos are continuous data so that clients are able to watch them without skew time. The skew time is the time difference between the time that the latest arrival video clip is consumed, and the time that the next video clip starts to be played, at a client site. Schedules are required to deliver videos more effectively. In order to deliver the videos without the skew time, several kinds of resources in a network are allocated before the video clips leave video servers storing each video clip. Since the resources are restricted, the scheduling for utilizing them is required for effective streaming of videos.

In order to improve the efficiency of the usage of link bandwidths, many scheduling policies have been proposed. The policy shown in this paper is that the link bandwidth used for delivering a video is its consuming ratio. This policy is called the minimum bandwidth delivery method. If the consuming ratio of a certain video is a constant ratio, the bandwidth used for delivering it is the same ratio. Therefore, any buffer space is not required in intermediate servers. Moreover, there are the cases that an entire video is divided into some video clips, and that these video clips are stored in different video servers. Schedules to deliver these video clips are constructed for each video clip. In order to select a path to be used for delivering each clip, the scheduler selects one path according to costs of paths from a server to a client site, or to initial latency time when a certain path is used for delivering the clip. The scheduler which is built based on the minimum bandwidth delivery method provides two kinds of strategies for selecting the path. The cost is computed in terms of parameters on quality of services. A path to deliver a video clip is defined as the sequence of a video server, some intermediate servers, and a client site. Another type of the strategies is concerning with initial latency times that are waiting times at a client. Furthermore, link overflow and skew time do not occur also in the schedules. They are working as constraints for scheduling.

Some scheduling algorithms for delivering video streams are proposed, already. These algorithms treat scheduling of resources which are disk bandwidths and network bandwidths. Moreover, features of video data, disks, and network facilities are reflected. These methods pay attention to analysis of their characters from the viewpoint of the network analysis. On the other hand, some routing algorithms are proposed by [1, 2, 9], which satisfy quality of service constraints. It is necessary to control quality of service constraints for delivering video streams. In order to decide the paths, it is considered that such algorithms are useful. The method shown in [8] deals with the method for transmitting videos and buffering in servers by assuming a hierarchical network structure as a network topology.

This paper is organized as follows. Section 2 describes some assumptions on a network environment and assumptions to construct schedules. The policy and the procedures of the minimum bandwidth delivery method are presented in Section 3. Simulating results of the scheduling are shown in Section 4. Finally, Section 5 presents some concluding remarks.
2. A MODEL OF A NETWORKED VIDEO DELIVERING SYSTEM

An overview of a model for delivering videos is shown in Figure 1, which is a networked video delivering system. Clients request to watch a certain video through a network. The network consists of video servers, intermediate servers, client sites and links which connect them. A requested video is delivered from the video server to the client site through several intermediate servers. The delivered video ought to be consumed without the skew time at the client site.

On the other hand, it is required that parameters on quality of services are satisfied. These parameters are specified in terms of failure ratios, packet loss ratios and delay times of intermediate servers and links. They seem as the costs of paths to achieve delivery of videos.

In order to decide the path, the costs of paths are computed as the following cost function:

\[
\text{cost(path)} = W_{FR} \times \text{FailureRate} + W_{PL} \times \text{PacketLoss} + W_{DT} \times \text{DelayTime}.
\]

\[
W_{FR} + W_{PL} + W_{DT} = 1.
\]

In these functions, \text{FailureRate}, \text{PacketLoss} and \text{DelayTime} are parameters for specifying the quality of services on the path. These parameters are failure rate, packet loss and delay time of the path, and \(W_{FR}\), \(W_{PL}\) and \(W_{DT}\) are weights of these parameters, respectively. By adjusting these weights one suitable path is selected.

There are cases that an entire video is divided into some video clips, and that these clips are stored in different video servers. Because the entire video is a huge amount of data, e.g., movies, it is difficult to deliver the video from the same video servers. Moreover, by dividing the videos, an initial latency time is able to be reduced and flexibility to construct schedules is improved since the size of a video becomes small.

3. AN OVERVIEW OF A SCHEDULING METHOD

3.1. Policies of the scheduling

The policies of the minimum bandwidth delivery method are as follows:

- The minimum required bandwidths in links are used for delivering videos. The minimum bandwidth of the video is corresponding to its consuming ratio by assuming constant bit ratio. Required link bandwidth of the link is allocated to achieve the delivery. The rest of the link bandwidth of the link is able to be used for delivering other video clips, also. Moreover, buffer spaces are not required in intermediate servers, since the video is transmitted with its consuming ratio.

- An entire video is divided into some video clips. In addition, small sized clips are treated easily to construct a schedule. Furthermore, an initial latency time is able to be reduced, since resource allocation becomes flexible because usage time of a link is reduced. However, the scheduler has to maintain the order of the video clips to be played, and the video clips are delivered continuously without skew time at a client site, even if different intermediate servers are used for delivering each video clip of an entire video.

Schedules are constructed for replying to a set of requests. These requests are processed in the order of their occurrences. However, there is the case that the initial latency time for an early accepted request is longer than the initial latency time of a latter accepted request for satisfying constraints. The skew time occurs when a certain link on the delivery path becomes a bottleneck, and when the link between connecting a video server and an intermediate server does not have sufficient bandwidth. It is assumed that link bandwidth of all links are greater than consuming ratios of all video clips, also.

3.2. Scheduling phases

Scheduling procedure consists of the following three steps:

1. Finding paths from a video server to a client site. These paths are called candidate paths, and constructed for each video clip of the requested video.

2. Deciding the delivery path, and computing the usage time of links on the path. One of the candidate paths is selected as a delivery path. The delivery path is used for delivering video clips. Since links are shared, not all clients use the link at the same time. Moreover, it is difficult that a client waits until transmissions of all videos required by other clients are finished.
There are two types of strategies to decide the delivery path. They are (a) cost-effective and (b) latency-time-effective. In the first strategy, the path whose cost is the lowest among candidate paths is selected preferentially than others if possible. When this selected path is finished to deliver other videos, the requested video clip starts the video server. On the other hand, in the second strategy, a path is selected, so that the video starts from the video server as soon as possible. As a result, the initial latency time is reduced.

In this step, the start times of video clips from servers are computed, then to prevent the link overflow is thought over. The start times of the clips are set to the times that the clips are able to be left by using its delivery path as soon as possible. They are used as initial start times in the next step.

3. Computing exact start times of clips, which satisfies the constraints. The start times of the clips are computed exactly, so that the skew time does not occur at client sites. Between two consequent video clips, the time that the previous video clip is consumed is equal to the time that the next clip starts to be played. If it is impossible to satisfy this relationship, step 2. is applied again to find another delivery path.

A scheduler constructs two kinds of schedules. One specifies the transmission schedules with respect to each video clip. Another specifies usage times and used bandwidth of links with respect to resources. After scheduling, usage time and used bandwidth are noted in a representation of a network. This representation is used for seeking available time to deliver other videos, also.

4. SOME SIMULATING RESULTS

Some simulating results are shown by making some assumptions. An assumed network structure is shown in Figure 2, which consists of three video servers, VS0, ..., VS2, nine intermediate servers which are S3, ..., S12, and links connecting servers. The labels of the links indicate their bandwidths. The client sites are connected to intermediate servers. Its bandwidth is equal to the consuming ratio of the video. Moreover, the size of each video is 40 [MByte]. All videos consist of two video clips whose size is 20 [MByte], and their consuming ratios are 1 [MBps]. These two video clips are stored in the same server. Furthermore, the requests occur during 100 seconds. The request times are generated as uniform random variables. The number of clients site is changed to 1000.

For making schedules, occurrences of delay and packet loss in a network environment are not reflected, i.e., the scheduler does not handle these phenomena. These parameters are only used for selecting delivery paths in current.

Figure 3 shows latency times when the number of clients
is varied. The scheduler provides the two strategies to select a delivery path, which are the cost-effective strategy and the latency-time-effective strategy. The initial latency times using the latency-time-effective strategy are smaller than ones using the cost-effective strategy.

Figure 4 shows costs of the delivery paths in the constructed schedules. A path whose cost is lower than others is preferably selected as a delivery path. Even if the number of clients increases, the costs of two kinds of schedules constructed according to two strategies do not much differ from each other. It seems that their difference is fewer than the difference between latency times when each strategy is used, in this simulating result.

Figure 5 shows the link usage ratios. This ratio is the proportion between the mean values of usage times of all links connecting servers, and the total working time of the network. The usage time of each link is a sum of actually used time to deliver videos. The total working time is time difference between the time that the first request occurs, and the time that the last video clip is consumed in the last client site.

Figure 6 shows a network usage ratio. The network usage ratio is a proportion of the amount of delivered videos using the scheduler to the amount of videos that are able to be delivered from all of the video servers during the total working time. The amount of delivered video using scheduler is corresponding to the amount of consumed video at all of clients. On the other hand, the amount of videos which are able to be delivered is \((\text{the sum of bandwidths of the links between video servers and intermediate servers}) \times (\text{the total working time})\). Each video server is connected to only one intermediate server, as shown in Figure 2.

5. CONCLUDING REMARKS

An overview of the scheduling method, the minimum bandwidth delivery method which achieves to deliver videos with the minimum bandwidth is described. Moreover, some simulating results are shown.

As a future work, it is required that utilization of bandwidths of links makes be more effective. Moreover, if the requested video clip is transmitted to a suitable intermediate server by storing the video into the disk of the servers as temporally storage, initial latency times will be reduced.

6. ACKNOWLEDGEMENTS

We thank Y. Ikoma for his cooperation to implement simulators and for discussions. We acknowledge Professor S. Iida for his many suggestions on this paper.

7. REFERENCES