The Strategy of Batch Using Dynamic Cache for Streaming Media

Zhiwen Xu, Xiaoxin Guo, Yunjie Pang, Zhengxuan Wang

Faculty of Computer Science and Technology, Jilin University,
Changchun City, 130012, Jilin Province, China
xuzhiwen@public.cc.jl.cn

Abstract. The batch is an important technique for delivering video over Internet or VoD. It is a key method to improve effect for video multicast. In this paper, we research the batch strategy of proxy cache for streaming media using dynamic cache, proposed the three kind of cache algorithm for proxy cache: window batch, size batch, efficient batch. These methods increased the length of batch, solved the problem for latency time of batch in video muticast, improved the byte-hit ratio of proxy cache for streaming media, and economized the resources of network backbone. Event-driven simulations are conducted to evaluate these kinds of strategy are better than prefix cache and segment cache.

1 Introduction

The batch is one important technique for transmitting video multicast and VoD. It makes it possible to transmit video and the media required by all the clients within b of the duration time. Therefore through only one transmission, multi-application within duration time b can be satisfied. Although this method saves the transmission resource, users need to wait during the time interval b. What's more, in the process to transmit video, the length of b is not allowed to be too long, which limits the use of batch technique. In the paper, we make use of the technique based on the proxy cache for streaming media and dynamic cache to solve the following problems: to lengthen time of batch, enhance the byte-hit ratio of proxy cache, ensure that we can transmit the video to the users without waiting time.

2 The Batch Using Dynamic Cache for Streaming Media

In the proxy cache based on segmented streaming media[2], we cache the prefix first to ensure there's no startup delay. The segmented cache, employing prefetch technique caches only a part of video, according to the popularity for the users' requests. When two or more users apply for the same video, we can save network resource as long as we can transmit the entire requirement at a time. In proxy cache, we employ dynamic movable time windows to cache video, and make replacement by using an algorithm called FIFO in order to make sure that applicants can directly receive video from the proxy cache, needless to wait. In this paper, we put forward three batch strategies: window batch, size patch, efficient batch.

2.1 Window Batch

If the time length of the video is L_h minutes, then the simplest strategy of batch is used to transmit the video in batch with a set length in the window. We set W to be the width of batch for each video clip, and L_h/W will determine the times of multicast required. When two users (or more) apply to video, than that video required by the very first applicant is transmitted through the server by the proxy. The latter users required the video that is transmitted by the proxy cache. If the average request ratio reaches λ, then the average of the users who require video service can be simply believed to be λW. If two or less applicants apply for a certain video when we are window batch of the video, then the data stream in batch can be omitted. There exists one problem about how to select. The number of applications from the users is uncertain, which is determined by the random changes in the requesting process. When there are no more than two applicants within the window batch time, we don't employ the batch technique. In the proxy cache of segmented streaming video, the length of the video cached in the proxy is 2^ib. Set the length of the window based on window batch to be W, and W=2^ib. when batch takes place in time W, then we should make a judgment whether the batch saves the value. Let us suppose the requests used of batch are respectively $A_1, A_2, ..., A_j$, then the value of j/W should be smaller than the threshold value of proxy cache, otherwise we can not transmit the video used of batch. In this way, we can surely enhance the efficiency of proxy cache and simplify the replacement algorithm of cache.

2.2 Size Batch

Generally, for the service providers, to cover as many users as possible in the batch is their great concern, because the income of the system is directly related to the average length of batch. Set C_s represents the value of the video stream of batch when seeing video and N' is the average of the users within batch. If P is the worth for each delivering video in the batch, then the worth of transmitted video in the batch is N'P. In order to maintain the income for network resource, N'P $\geq C_s$ is necessary. Let us define K $\doteq C_s$/P, and N'$\geq$K, then the longer the length of batch, the larger income of the system. In size batch, there involves one question, automatic selection. Higher number of the users in batch results from higher arrival ratio. Low byte-hit ratio results in low yield of the system. Before multicast, the yield is ensured by the requests from M$\geq$K users in batch. The time length needed by dynamic cache in the proxy cache is the time brought by collecting M users. When M=2, the practicability is useful.

2.3 Efficient Batch

In the batch process, the smaller capacity the proxy cache is taken up, the larger the quantity within batch. The higher the byte-hit radio of proxy cache, the higher the efficiency. The length of video with popularity λ is L_i, and that length in the proxy cache is V_i. The users make applications at time $A_0, A_1, ..., A_n$ in turn and A_0=0. When k=1,2,...,n, $(A_k - A_{k-1}) \leq L_i/\lambda V_i$, we can ensure it is beneficial for every video request within batch. That is to say, we transmitted those requested video in batch which help to save network resources.

3 Analysis for Batch Strategy

We analyze the performance of the batch schemes mentioned above. We focus on a particular video in the system, with the request's arrival process being Poisson with rate λ. In the various schemes for evaluation, we are interested in the following closely related performance measure: (a) Its distribution is represented by f_s(s), and the mean value is S'. The average of the concurrent stream indicates the average required in the system. Probably it is also related to the saved bandwidth, compared with pure VoD, because in the pure VoD, the average number of concurrent stream's requested is λL_h, the saved bandwidth can be defined as $\eta = 1 - S'/(\lambda L_h)$. (b) The batch size N, both its distribution f_N(n) and its mean N'. Note that N' is related to S' by the λT_h= S' N'. Indeed, by the above formula, within the display time of a certain video, the arrival average of the users' requests can be simply represented by λL_h, which is equal S'N'. (c) For a random user, the delay is represented by D, whose distribution f_D(t) is expressed by D'. The proxy cache requires the smallest dynamically cache batch size for D.

3.1 Analysis for Window Batch

In the fixed window batch strategy, the distribution of the batch size is a Poisson process, expressed by N'= λ W. The video is accurate batch once in every W second, therefore S'=L_h/W and $\eta = 1 - 1/(\lambda W)$. D is a $\sim$ U(0,W) uniformly distribution and D'=W/2. in fixed gating, the possibility of the request i(i $\geq$ 1)in batch is given by $\lambda W^i e^{-\lambda W}/(i!(1 - e^- \lambda W))$, therefore N'=$\lambda W/(1 - e^{-\lambda w})$ and

$$S' = \frac{L_h}{W}(1 - e^{-\lambda w}) \tag{1}$$

As expect ,when $\lambda \longrightarrow 0, S' \longrightarrow \lambda L_h$, and when $lim_{\lambda \longrightarrow \infty} S' = L_h/W$, then S'$\leq$ L_h/W. The bandwidth saved is $\eta = 1 - (1 - (1 - e^{-\lambda w})(\lambda W))$. Given one arrival in the batch window, and this arrival time in the window is distributed evenly, therefore the distribution function related to the users' delay for a certain video is f_D(t)=1/W, for $0 \leq t \leq$ W and D'=W/2. T=0, when the proxy cache can't transmit the video in the batch. In automatic selection, the distribution of the batch size is $P(N=i)=(\lambda W)^{i-1}e^{-\lambda w}/(i-1)!$, for $i\geq 1$, therefore N'=1+λ W and

$$S'=\frac{\lambda T_h}{1+\lambda w} \tag{2}$$

and $\eta=\lambda W/(1+\lambda W)$, the request distribution is given by the following formula:

$$f_D(t)=\frac{1}{1+\lambda W}\delta(t-W)+\frac{\lambda W}{1+\lambda W}\frac{1}{W} \tag{3}$$

for $0\leq t\leq W$, where $\delta(t)$ is the usual impulse function with $\delta(t)=0$ for t≠ 0 and $\int_{-\infty}^{\infty}\delta(t)dt=1$. It accounts for the first user in each batch having a delay of W while the remaining users in the batch having a delay uniformly distributed between o and W, then the function works out that for the first user of each batch, there is a time interval W. Therefore D' can be given:

$$D'=\frac{\lambda W+2}{2(1+\lambda W)}W \tag{4}$$

The cache size of batch for proxy cache is D', and the request number in batch is N'.

3.2 Analysis for Size Batch

The video films collect M requests at a time in batch, therefore N is decided, which is equal to M. As a result N'=M, S'=$\lambda L_h/M$,$\eta=1-1/M$. Let W represent the batch period, which is variable at random and equal (M-1), the summation of exponential variables. Its distribution can be given:

$$g_w(W)=\frac{\lambda(\lambda W)^{M-2}}{M-2}e^{-\lambda W} \tag{5}$$

The users' request duration can be given according to the movement status on W:

$$f_D(x)=\frac{1}{M}g_w(x)+\frac{M+2}{M}\int_x^{\infty}\frac{g_w(W)}{W}dW+\frac{\delta(W)}{M} \tag{6}$$

D'=(M-1)/(2λ), which is equal to half of average batch period. The cache size in batch is T=(M-1)/λ.

3.3 Analysis for Efficient Batch

In the efficient batch scheme, N=j, its distribution can be given:

$$P(N=j)=\frac{(\lambda A_j)^{j-1}}{(j-1)!} \tag{7}$$

The users' request distribution is

$$f_D(x)=\frac{L_i}{\lambda V_i}e^{-\lambda x} \tag{8}$$

All the requests is transmitted in this batch scheme. The time of duration is smaller than $L_i/\lambda V_i$, when M=2. The proxy cache size T= A_n, and the request number of batch is S(t)=n.

4 Performance and Conclusion

We compared these three batch policies with the full video approach, the variable-sized segment approach, and the prefix schemes in terms of the impact they imposed on byte hit ratio and startup delay, in the following aspects: the cache size, the popularity of the video.

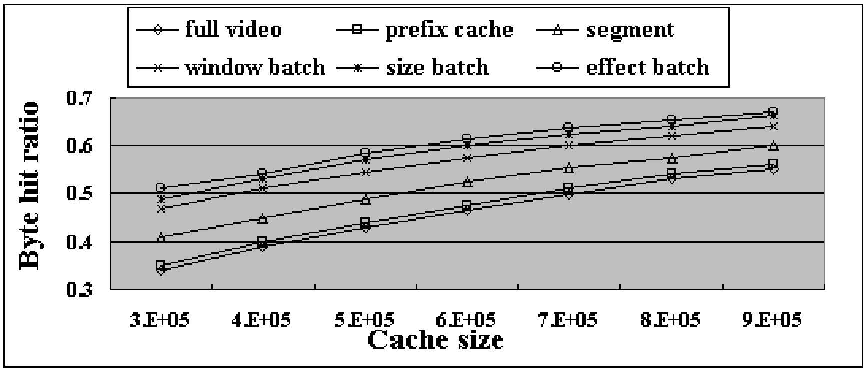

Fig.1 The relation of byte hit radio and cache size

The window batch of the proxy cache is the simplest and most practical strategy. Efficient batch of the proxy cache has the highest byte hit ratio(Fig.1).Within the whole range of cache size, the effect of the dynamic cache approach, variable-size segment-based strategy and the prefix strategy all effectively solve the problem of startup delay.

Efficient batch, size batch, window batch further improve the byte hit ratio of the proxy cache and solve the wait delay existed in batch for VoD. Efficient batch has the highest byte hit ratio. Window batch has the following characteristics: convenient usage, simply and practical replacement algorithm of cache, high byte hit ratio. These three kinds of batch strategies can enhance the efficiency of the proxy cache together with the segment strategy.

References

1. Z.Miao, A.Ortega: Proxy caching for efficient video services over the Internet.In Proc. Of Int, Web Caching Workshop, Apr. 1999.
2. K.L.Wu, P.S.Yu: Segment-Based Proxy Caching of Multimedia Streams. In: Proc. of IEEE INFOCOM, May, 2001.
3. S.Sen, J.Reforrd, D.Towsley: Proxy Prefix Caching for Multimedia Streaming. In: Proc. Of IEEE INFOCOM, Mar.1999.
4. S. Ramesh, I. Rhee, and K. Guo: Multicast with Cache (MCache): Anadaptive Zero-Delay Video-on-Demand Service. in: Proc. IEEE INFOCOM,April 2001.