

An Adaptive Data Retrieval Scheme for Reducing Energy Consumption in Mirrored Video Servers

Minseok Song

School of Computer Science and Engineering,
Inha University, Korea
mssong@inha.ac.kr

Abstract. We propose a new energy-aware data retrieval scheme (EDR) for mirrored video servers. We start by analytically determining the retrieval period that balances bandwidth and buffer size. We then propose a data retrieval scheme in which the period can be dynamically changed to reflect server utilization, with the aim of giving disks more chance to enter low-power mode. Simulation results show that it saves up to 36% energy compared with conventional video server operation.

1 Introduction

Today's data centers typically consume a lot of power. Large power consumption is a serious economic problem to service providers [5]. For example, a medium-sized 30,000 ft² data centers requires 15 MW which costs \$ 13 million per year [5]. Cooling systems for higher heat densities are prohibitively expensive and running them makes up a significant proportion of the power cost of running a data center [2, 3]. In addition, it has been shown that 15 °C above ambient can double the failure rate of a disk drive [2].

To reduce power consumption in servers, modern disks have multiple power modes [2, 5]: in active mode the platters are spinning and the head is seeking, reading or writing data, in idle mode the disk spins at full speed but there is no disk request, and in low-power mode the disks stops spinning completely and consumes much less energy than active or idle mode. Most of the schemes for energy conservation in data centers involve switching disks to low-power modes whenever that is possible without adversely affecting performance [2, 5]. These schemes primarily aim to extend the period during which a disk is in low-power mode. Because returning from low-power to active mode requires spinning up the disk, the energy saved by putting the disk into low-power mode needs to be greater than the energy needed to spin it up again. But this concept is hardly relevant to video servers due to the long duration of video streams.

To address this, we propose an energy-aware data retrieval (EDR) scheme for mirrored video servers. Based on an analytical model for the optimal data retrieval size, we propose a data placement scheme in which an appropriate block size can be dynamically selected and a data retrieval scheme which adaptively retrieves data from the primary and backup copies to permit disks to go to low-power mode as often as possible.

The rest of this paper is organized as follows. We explain video server model in Section 2. We then propose a new data retrieval scheme in Section 3. We validate the proposed scheme through simulations in Section 4 and conclude the paper in Section 5.

2 Video Server Model

We use round-based scheduling for video data retrieval: time is divided into equal-sized periods, called rounds, and each client is served once in each round. We partition a video object into blocks and distribute them over multiple disks. The size of the striping unit denoted as the maximum amount of contiguous data stored on a single disk is the amount of data retrieved during a round so that only one seek overhead is incurred in each data retrieval [1]. Data retrieved during the period R of a round are grouped together and constitute a segment. Then each segment is stored in a round-robin fashion across disks.

For fault-tolerance, we use a replication technique where the original data is duplicated on separate disks. We refer to the original data as the primary copy (PMC) and call the duplicated data the backup copy (BCC). The server reads data from the PMC when all disks are operational; but when a disk fails, it uses the BCC for data retrieval. Among various replication schemes, chained declustering (CD) shows the best performance [4]. Let us assume that a disk array consists of D homogeneous disks. In the CD scheme, the primary copy on disk i has a backup copy on disk $(i + 1) \bmod D$. We place the backup copy on one disk as in the CD scheme.

3 Energy-Aware Data Retrieval

The round length plays an important role in determining the requirements for server resources (i.e. disk bandwidth and buffer) [1, 4]. We now describe the optimal round length that balances the requirements for disk bandwidth and buffer. Let us assume that each video V_i has data rate of dr_i (in bits/sec), ($i = 1, \dots, N$). Let p_i be the access probability of video V_i , where $\sum_{i=1}^N p_i = 1$. Let B be the total buffer size. We use SCAN scheduling. Since double buffering is used for SCAN scheduling [1], the buffer utilization BU for c clients is expected as follows:

$$BU = \sum_{i=1}^N \frac{2R \times p_i \times dr_i}{B} c. \quad (1)$$

The bandwidth utilization for a disk is usually defined as the ratio of total service time to R [1]. We use a typical seek time model in which a constant seeking overhead (seek time + rotational delay) of T_s is required for one read of contiguous data [1]. Let tr be the data transfer rate of the disk. Since there are D disks in the server, the disk bandwidth utilization DU for c clients can be estimated as follows:

$$DU = \frac{1}{D} \sum_{i=1}^N p_i \left(\frac{T_s + R \times \frac{dr_i}{tr}}{R} \right) c. \quad (2)$$

If $DU = BU$, the server is able to balance the use of buffer and disk bandwidth. From Equations (1) and (2), we obtain the optimal round length OR as follows:

$$OR = \frac{\frac{B}{tr} + B \sqrt{\frac{1}{tr^2} + \frac{8 \times D \times T_s}{B \times \sum_{i=1}^N p_i dr_i}}}{4 \times D}. \quad (3)$$

To allow disks to go to low-power mode, EDR permits dynamic data retrieval from either the PMC or the BCC. Let CDU_i be the bandwidth utilization of disk i . We define two states: if $\forall i \text{ } CDU_i \leq 0.5$, then the server is in *energy reduction (ER)* state; otherwise, the server is in *normal* state. In the normal state, data is retrieved from the PMC on every disk; while in the ER state, data is retrieved only from odd-numbered disks. Since the data on disk i is replicated as a backup copy on disk $(i + 1) \bmod D$, the disk loads incurred in reading the PMC of disks $2i$ are shifted to the BCC of disks $(2i + 1) \bmod D$ ($i = 1, \dots, \lfloor \frac{1}{2}D \rfloor$). The disks $2i$ are able to go to low-power mode because they are not accessed. Even though the disk loads of disks $2i$ are shifted to disks $(2i + 1) \bmod D$, the utilization of disk $(2i + 1) \bmod D$ does not exceed 1 because $\forall i, CDU_i \leq 0.5$.

If $\forall i, CDU_i \leq 0.5$, even-numbered disks go to low-power mode, and so reducing the disk bandwidth utilization below 0.5 is important for energy reduction. We observe from Equation (2) that extending the round length reduces disk bandwidth utilization; but this increases the buffer overhead, by Equation (1). EDR has *common, major* and *minor* cycles, with lengths of $3OR$, OR and $\frac{3}{2}OR$, respectively. From Equation (3), we observe that a cycle of length OR leads to the balanced use of disk bandwidth and buffers. Thus, choosing the major cycle results in the balanced use of disk bandwidth and buffers. But by selecting the minor cycle for the round length we reduce the energy consumption because a long round is advantageous in terms of disk throughput.

To extend the time spent in low-power mode, the server adaptively adjusts the round length. Assuming that all data is retrieved from the PMC, EDR calculates ADU , the maximum disk bandwidth utilization among all disks when the major cycle is selected, and IDU , the maximum bandwidth utilization when the minor cycle is chosen. Additionally, let ABU and IBU be the buffer utilizations when major and minor cycles are selected, respectively. Use of the major cycle may produce the normal state, in circumstances under which the minor cycle would permit the system to go to ER state. As a consequence, if possible, EDR tries to select the minor cycle as the round length for energy reduction. But the server needs to check the buffer condition (i.e. that $IBU \leq 1$) because if that constraint is not met, using the minor cycle may lead to buffer overflow. If $IBU \leq 1$ and $IDU \leq 0.5$, EDR selects the minor cycle as the round length; otherwise, the major cycle is chosen. Because the minor cycle is only chosen if $IDU \leq 0.5$, we can easily see that *the server is in ER state if the minor cycle is chosen*.

The change of round length should not cause additional seek overhead. To meet this constraint, we split a data segment into 6 sub-segments, where the size of each sub-segment corresponds to the data retrieved during a round of length $\frac{1}{2}OR$. The 6 sub-segments are stored contiguously and constitute a segment. In normal state, two contiguous subsegments are read during OR , but in ER state, the server retrieves three contiguous subsegments during $\frac{3}{2}OR$.

4 Experimental Results

To evaluate the effectiveness of our scheme, we performed simulations. Our server has 16 disks, each of which employs an IBM Ultrastar36Z15 disk whose parameters are shown in [5]. We choose 12 ms for T_s and 1 GB for B . The arrival of client requests is assumed to follow a Poisson distribution. We also assume that all videos are 90 minutes

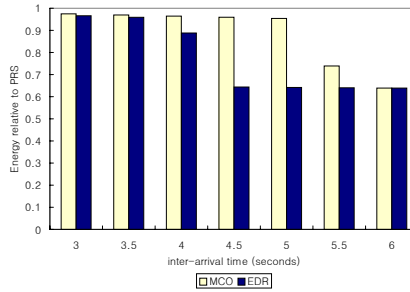


Fig. 1. Energy consumption relative to PRS for various inter-arrival times

long and have the same bit-rate of 1MB/sec. We compare EDR with two other methods; the PRS scheme that does not allow data retrieval from the BCC, and the MCO scheme that permits data retrieval from the BCC but does not allow adaptive cycle adjustment. We assess how the energy consumption depends on the inter-arrival time of requests over 24 hours. Fig. 1 shows the energy consumption of MCO and EDR schemes relative to PRS. The EDR scheme exhibits the best performance under all workloads, saving between 4% to 36% over the conventional PRS scheme. Compared with MCO, EDR reduces energy consumption by up to 33%. This is because by adaptively increasing the round length from the major to the minor cycle, EDR reduces the disk bandwidth utilization, which results in more opportunities to stay in low-power mode.

5 Conclusions

We have proposed a new energy-aware data retrieval scheme for mirrored video servers. An analytical model shows how to balance the use of disk bandwidth and buffer and based on this, we have proposed a data retrieval scheme which adaptively retrieves data from the primary and backup copies to give the server more chances to operate in low-power mode. Experimental results show that our scheme enables the server to achieve appreciable energy savings under a range of workloads.

References

1. E. Chang. *Storage and Retrieval of Compressed Video*. PhD thesis, University of California at Berkeley, 1996.
2. S. Gurumurthi. *Power Management of Enterprise Storage Systems*. PhD thesis, Pennsylvania State University, 2005.
3. E. Pinheiro and R. Bianchini. Energy conservation techniques for disk-array-based servers. In *Proceedings of ACM/IEEE Conference on Supercomputing*, pages 88–95, June 2004.
4. M. Song and H. Shin. Replica striping for multi-resolution video servers. In *Proceedings of IDMS/PROMS (LNCS 2515)*, pages 300–312, November 2002.
5. Q. Zhu and Y. Zhou. Power aware storage cache management. *IEEE Transactions on Computers*, 54(5):587–602, 2005.