Zoned-RAID for Multimedia Database Servers

Ali E. Dashti[1], Seon Ho Kim[2], and Roger Zimmermann[3]

[1] Electrical and Computer Engineering Department, Kuwait University,
Safat, 13060, Kuwait
dashti@eng.kuniv.edu.kw
[2] Department of Computer Science, University of Denver,
Denver, CO 80208, U.S.A.
seonkim@cs.du.edu
[3] Department of Computer Science, University of Southern California,
Los Angeles, CA 90089, U.S.A.
rzimmerm@usc.edu

Abstract. This paper proposes a novel fault-tolerant disk subsystem named *Zoned-RAID* (Z-RAID). Z-RAID improves the performance of traditional RAID system by utilizing the *zoning* property of modern disks which provides multiple zones with different data transfer rates in a disk. This study proposes to optimize data transfer rate of RAID system by constraining placement of data blocks in multi-zone disks. We apply Z-RAID for multimedia database servers such as video servers that require a high data transfer rate as well as fault tolerance. Our analytical and experimental results demonstrate the superiority of Z-RAID to conventional RAID. Z-RAID provides a higher effective data transfer rate in normal mode with no disadvantage. In the presence of a disk failure, Z-RAID still performs as well as RAID.

1 Introduction

Recent years have witnessed the proliferation of multimedia databases, especially handling streaming media types such as digital audio and video, with the wide acceptance of the public and the industry. These media have become a part of everyday life including not only electronic consumer products but also online streaming media services on the Internet. Due to 1) successful standards for compression and file formats, such as MPEG (Motion Picture Expert Group), 2) increased network capacity for local area networks (LAN) and the Internet, and 3) advanced streaming protocols (e.g., Real Time Streaming Protocol, RTSP), more and more multimedia database applications, combined with the Internet, are providing streaming media services such as remote viewing of video clips.

Streaming media (SM) have two main characteristics. First, SM data must be displayed at a pre-specified rate. Any deviation from this real-time requirement may result in undesirable artifacts, disruptions, and jitters, collectively termed *hiccups*. Second, SM objects are large in size. For example, the size of a two-hour MPEG-2 encoded digital movie requiring 4 Mb/s for its display is

3.6 GBytes. Due to these characteristics, the design of SM servers has been different from that of conventional databases, file servers, and associated storage systems [5, 3] to provide a *hiccup-free display*, a higher throughput, a shorter startup latency, and a more cost-effective solution.

Magnetic disk drives have been the choice of storage devices for SM servers due to their high data transfer rate, large storage capacity, random access capability, and low price. Therefore, many studies have investigated the design of SM servers using magnetic disk drives [5, 3]. Due to the essential role of disk storage systems in SM servers, understanding recent trends in disk technologies can be helpful. First, the capacity and speed of magnetic disk drives have improved steadily over the last decade. According to [9] on the recent trends in data engineering, the storage capacity of magnetic disks has increased at the rate of about 60% per year. At the same time, the data transfer rate of magnetic disks has increased at the rate of about 40% per year. Thus, the imbalance between disk space and data transfer rate has widened. Because data transfer rate (bandwidth) is the scarce resource in the applications that intensively access disks, one wants to optimize for bandwidth rather than for space [9].

Another important physical characteristic of modern disks is *Zoned recording* (or *zoning*). This is an approach utilized by disk manufactures to increase the storage capacity of magnetic disks [12]. This technique groups adjacent disk cylinders into zones. Tracks are longer towards the outer portions of a disk platter as compared to the inner portions, hence, more data can be recorded in the outer tracks when the maximum linear density, i.e., bits per inch, is applied to all tracks. A zone is a contiguous collection of disk cylinders whose tracks have the same storage capacity, i.e., the number of sectors per track is constant in the same zone. Hence, outer tracks have more sectors per track than inner zones. Different disk models have different number of zones. Different zones in a disk provide different transfer rates because: 1) the storage capacity of the tracks for each zone is different, and 2) the disk platters rotate at a fixed number of revolutions per second. We can observe a significant difference in data transfer rates between the minimum and maximum (around 50% difference) [3, 7, 12].

Last, since disk prices are approaching tape prices and tape backup takes a far longer time, disks are replacing tapes for backup and fault tolerant systems. Thus, many applications have been using RAID (Redundant Array of Independent Disks) [14]. Out of multiple levels of RAID, especially, both RAID level 1 and level 5 have been commonly used for a fault tolerant disk system [9].

In large scale multimedia database servers in support of streaming media, it is obviously critical both to optimize disk bandwidth and to provide disk-based fault tolerance. Many studies [16, 1, 6, 17, 7] discussed data placement on multi-zone disks to maximize the effective data transfer rate. [10] provided MRS (Multi-Rate Smoothing) data placement on multi-zone disks for a smooth transmission of variable-bit-rate data over network. However, none of above studies includes reliability issue. RAID has been widely used for faut-tolerant streaming servers as well as conventional file servers. Various reliability strategies in video servers, including RAID, were surveyed and compared in [4]. However, no study considered one of the most important characteristics of disk drives, variable data transfer rates from multiple zones in a disk. Therefore, conventional techniques place data blocks without any constraints inside a disk. This may result in less optimized disk performance because the data transfer rate significantly varies depending on the location of data block in multi-zone disks.

This study proposes a novel data placement scheme to optimize the data transfer rate of RAID systems using multi-zone disks by constraining data placement, especially for streaming media server that require a high data transfer rate as well as fault tolerance. To our knowledge, combining data placement on RAID with multi-zone disks is a new approach. The main ideas of the proposed constrained data placement are 1) to store primary data blocks (for normal access) in faster zones and secondary blocks or parity blocks (for standby in case of a disk failure) in slower zones, and 2) to store frequently accessed data blocks (such as popular video clips) in faster zones and infrequently accessed blocks in slower zones. Our experimental results demonstrate a significant increase in the effective data transfer rate of RAID in normal mode with no disk failure.

2 Z-RAID

Since RAID [14] was proposed in 1988, it has been widely implemented in many systems requiring fault tolerance. Originally, RAID levels 1-5 were proposed but many variants such as level 0 and 6 have been studied and commercialized. However, level 1 (mirroring) and 5 (block-based parity encoding) received most attention in many applications due to their cost-effectiveness and implementation efficiency [9]. Thus, this study focuses on extending RAID level 1 and 5 to our proposed Zoned-RAID (Z-RAID) approach.

A multi-zone disk can be modelled as follows: A disk with total space, S, has n zones, where zone 0 is the innermost (slowest) and zone $n-1$ is the outermost (fastest). The number of cylinders in each zone is $Cyl(i)$, $0 \le i < n$, and the total number of cylinders is Cyl. Cylinders are numbered from the innermost to the outermost. The size of a cylinder is $S(i)$ bytes, $0 \le i < Cyl-1$. The data transfer rate of each cylinder is $R_c(j)$, $0 \le j < Cyl$, $(R_c(0) \le R_c(1) \le ... \le R_c(Cyl-1))$. Note that all cylinders in the same zone have the same data transfer rate. A rotational latency, l_{rot}, is one disk revolution time of a disk. A seek time between two locations in a disk, say x cylinders apart, can be calculated using a practical non-linear approximation, $seek(x)$ [15]. Then, an actual block retrieval time consists of a seek time, a rotational latency, and block reading time.

2.1 Z-RAID Level 1

RAID level 1 utilizes a replication of disks, called *mirroring*. When we have two disks, d_0 and d_1, then a primary copy of a block, B_i, is placed on d_0 and the secondary copy, say B'_i, is placed on d_1. Blocks are arbitrarily distributed across cylinders inside a disk. This implies the system uses the average data transfer rate of a multi-zone disk and the average seek time (one half of the worst seek which is

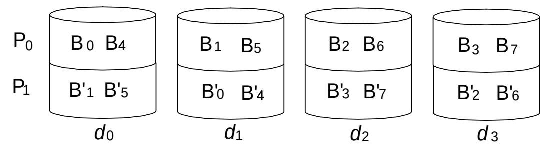


Fig. 1. Z-RAID level 1 with four disks.

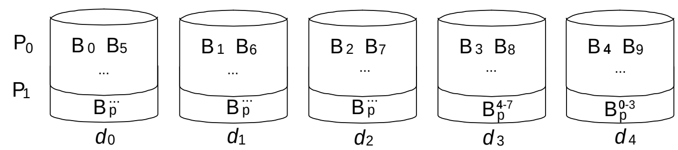


Fig. 2. Z-RAID level 5 with five disks.

from the outermost cylinder to the innermost cylinder). Then, the effective data transfer rate of a disk with no overhead (no seek time, no rotational latency) is:

$$R_R = \sum_{i=0}^{Cyl-1} (R_c(i) \times \frac{S(i)}{S}) \tag{1}$$

In a streaming media server whose access unit is a block (B), each block access includes the worst seek time and rotational latency to support realtime block retrieval even in the worst case [5, 8]. Thus, the effective data transfer rate of RAID level 1 in a streaming media server is:

$$R_{RB} = \frac{B}{seek(Cyl) + l_{rot} + B/R_R} \tag{2}$$

Z-RAID level 1 also uses mirroring like RAID level 1. However, it utilizes only faster zones of disks for primary copies of blocks. All secondary copies are placed on slower zones. With Z-RAID 1, each disk is divided into two logical partitions of equal size ($P_0 = P_1 = S/2$), P_0 which occupies the faster zones ($S/2$ from the outermost cylinders) and P_1 which occupies the slower zones (remaining $S/2$). All primary blocks, B_i, are assigned to P_0 while all secondary blocks, B'_i, are stored in P_1, see Figure 1. Let us say that P_0 consists of cylinders from m to $Cyl - 1$, where m is the cylinder number that divides the disk space in half (i.e., $\sum_{i=0}^{m-1} S(i) = S/2$). Note that the value of m and Cyl should be determined using real disk characteristics because different disk models have different zone characteristics. A more general allocation of blocks is as follows: when Z-RAID consists of k disks, if B_i resides on P_0 of disk j, B'_i is stored in P_1 of disk $(j + 1)\ mod\ k$.

In normal mode without disk failure, blocks are retrieved from P_0s of disks. Because P_0s are located in faster zones of a disk, Z-RAID will increase the effective data transfer rate of the disk. Moreover, because the maximum cylindrical distance inside P_0 is far shorter than Cyl, Z-RAID will decrease the required seek time between two adjacent block retrievals. Both will result in a significantly enhanced effective data transfer rate:

$$R_{ZR} = \sum_{i=m}^{Cyl-1} (R_c(i) \times \frac{S(i)}{S/2}) \quad (3)$$

$$R_{ZRB} = \frac{B}{seek(Cyl - m - 1) + lrot + B/R_{ZR}} \quad (4)$$

2.2 Z-RAID Level 5

RAID level 5 uses a block-based parity encoding. It distributes parity blocks across disks in a parity group so that both normal blocks and parity blocks can be placed on a disk. Blocks are arbitrarily distributed in a disk. Thus, in normal mode, the effective data transfer rate of RAID level 5 is identical to RAID level 1, i.e., Equations 1 and 2.

Z-RAID level 5 follows the same way as RAID level 5 to distribute parity blocks across disks. However, the location of parity blocks inside a disk is constrained to the slower zone areas. For example, when we form a parity group with 5 disks, 4 data blocks and a parity block will be distributed across 5 disks. Thus, 20% of each disk space consisting of corresponding innermost tracks will store all parity blocks while 80% of the disk space with outer tracks stores data blocks. For example, each disk has two logical partitions, P_0 (outer 80% of disk space) and P_1 (inner 20% space). Normal data blocks are stored in P_0 and all parity blocks are in P_1, see Figure 2. The same advantages of Z-RAID level 1 in Section 2.1 are expected: higher effective data transfer rate and shorter average seek time in normal mode.

When d disks are in a parity group, 1/d of each disk space will be used to store parity blocks. Then, P_0 consists of cylinders from m (where $\sum_{i=0}^{m-1} S(i) = S/d$) to $Cyl - 1$. Equation 3 and 4 for Z-RAID level 1 can be used for Z-RAID level 5 with a different value of m that is a function of d.

2.3 Z-RAID for Multimedia Databases

Because Z-RAID can provide a higher effective data transfer rate with the same fault tolerant disk system compared to a conventional RAID, it can be used where ever a RAID can be used. However, some applications such as streaming applications that require a large page (block) size mostly benefit from Z-RAID because a block retrieval time depends more on data transfer time than other near constant factors such as seek time and rotational latency. Note that B/R_{ZR} becomes a dominant factor (see Equation 4) as B grows larger.

Another important observation in real streaming applications is that objects may have different popularity or access frequency. For example, in a movie-on-demand system, more than half of the total user requests might reference only a handful of recently released hot movies. It is widely understood that the popularity distribution among objects in video-on-demand systems can be well represented by the Zipf distribution [13], which is a very skewed distribution.

Z-RAID can well take advantage of this skewed popularity distribution because the distribution of data transfer rates across zones is also skewed. With n objects in the system, one can sort objects in descending order based on their popularity. Then one assigns blocks of objects from the outermost tracks in a disk which has the fastest data transfer rate towards the inner tracks, track by track. When the blocks of the first, most popular, object are all assigned, then the next object is assigned in the same way from the next track. This process is repeated until all objects are assigned.

3 Comparisons

In our experiments, we used two Seagate disk models, the Cheetah X15 and the Barracuda 7200.7 plus. The Cheetah X15 provides one of the fastest rotation speeds at 15,000 revolutions per minute (RPM), with a very short average seek time of 3.6 milliseconds. This model exemplifies a typical high performance disk and was introduced in 2000. The Barracuda 7200.7 is a typical cost-effective high capacity disks with 7,200 RPM and 8.5 milliseconds of average seek time (introduced in 2004). Table 1 and Figure 3 show the zone characteristics of Cheetah X15 and Barracuda 7200.7.

3.1 Analytical Comparison

First, we calculated and compared the effective data transfer rates of RAID and Z-RAID with the two disk drives detailed in Table 1 using equations from Sections 2.1 and 2.2. We compared our design with two conventional approaches widely used for streaming media servers. With the guaranteed approach that supports 100% hiccup-free displays, one must assume the worst case seek time and the maximum rotational latency for each data block retrieval. Many round robin data placement and retrieval schemes [5, 8] follow this guaranteed approach, hence they fall into the category of worst case analysis. To quantify the effective data transfer rates of this approach, we performed a worst case analysis assuming the maximum seek time (7.2 ms for Cheetah X15 and 17 ms for Barracuda 7200.7) and the worst rotational latency (4 ms for Cheetah X15 and 8.3 ms for Barracuda 7200.7). Second, with the statistical approach that tolerates a non-zero hiccup probability, one can take advantage of the average seek time and average rotational latency per data block retrieval. Many random data placement and retrieval schemes [11] follow this statistical approach to enhance the performance of the system at the expense of a minor degradation of display quality, i.e., occasional hiccups. For this approach, we performed an average case analysis assuming the average seek time (3.6 ms for Cheetah X15 and 8.5 ms for Barracuda 7200.7) and average rotational latency (2 ms for Cheetah X15 and 4.16 ms for Barracuda 7200.7).

It is well established that the performance of streaming media servers – especially their disk subsystems – significantly varies depending on the data block size that is the unit of access to the disks. Thus, we calculated the effective

Table 1. Parameters for two Seagate disks.

Model	ST336752LC	ST3200822A
Series	Cheetah X15	Barracuda 7200.7 plus
Manufacturer	Seagate Technology	Seagate Technology
Capacity S	37 GB	200 GB
Transfer rate R_c	See Table 3.a	See Table 3.b
Spindle speed	15,000 rpm	7,200 rpm
Avg. rotational latency	2 msec	4.16 msec
Worst case seek time	7.2 msec	17 msec

Zone #	Size (GB)	Read Transfer Rate (MB/s)
0	12	57.5
1	3.5	55.4
2	3.0	54.7
3	4.0	52.7
4	3.0	50.6
5	2.5	48.1
6	3.0	45.6
7	2.5	43.6
8	2.5	41.9

a. Cheetah X15

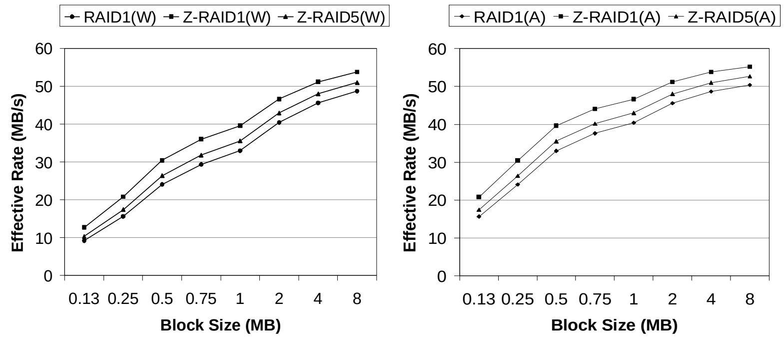
Zone #	Size (GB)	Read Transfer Rate (MB/s)
0	48	65.2
1	17	63.8
2	14	61.5
3	21	58.2
4	9	56.0
5	12	54.1
6	14	52.4
7	9	50.6
8	6	49.5
9	13	46.8
10	9	44.1
11	6	42.2
12	8	39.7
13	8	37.6
14	6	35.3

b. Barracuda 7200.7

Fig. 3. Zoning information of two Seagate disks.

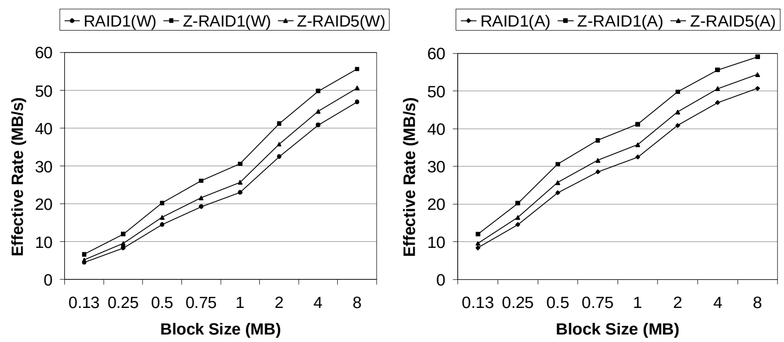
data transfer rate as a function of the data block size varying from 128 Kbytes to 8 Mbytes (a reasonable range for streaming media servers).

Figure 4 shows the effective data transfer rates of RAID and Z-RAID with the Cheetah X15. RAID1 denotes the traditional RAID level 1, Z-RAID1 means the proposed Z-RAID level 1, and Z-RAID5 refers to the proposed Z-RAID level 5. Note that the effective rate of RAID5 in normal mode is identical to that of RAID1 because all data blocks are arbitrarily distributed across all zones without any constraints. In our calculation, the size of the parity group of Z-RAID5 was 5 disks so that 20% of disk space (from the slowest zone) in each disk is dedicated to store parity blocks. Figures 4.a and 4.b show the results from the worst case and the average case analysis, respectively. Compared to RAID1, Z-RAID1 demonstrates enhanced rates from 10.5% to 38.6% in the worst case analysis, and from 9.5% to 33.1% in the average case analysis. Compared to RAID5, the percentage enhancement of Z-RAID5 ranges from 4.8% to 12.7% in the worst case analysis, and from 4.5% to 11.4% in the average case analysis. Figure 5 shows the analytical results with the Barracuda 7200.7. The results and trends are similar to those of the Cheetah X15. Z-RAID1 improves over RAID1 from 18.5% to 46.8% in the worst case analysis, and from 16.5% to 43.6% in the average case analysis. Compared to RAID5, the percentage enhancement of Z-RAID5 ranges from 7.9% to 14.7% in the worst case analysis, and from 7.3% to 14.1% in the average case analysis.



a. Worst analysis b. Average analysis

Fig. 4. Effective data rate of a Seagate X15 disk.



a. Worst analysis b. Average analysis

Fig. 5. Effective data rate of a Seagate 7200.7 disk.

As shown, for all comparisons, Z-RAID outperforms RAID. The percentage improvement of the effective data transfer rate is greater for small block sizes where the reduced seek time is the dominant factor in determining the rate. The dominant factor shifts from the seek time to the actual block reading time as the block size increases, see the divisors in Equations 2 and 4. The reduced seek time is also the reason why Z-RAID1 gains a higher percentage increase than Z-RAID5. With Z-RAID5, the performance enhancement decreases as the size of the parity group increases. With a smaller group such as three disks, a higher effective rate is achieved than with larger groups.

3.2 Simulation Results

The analytical models of the previous section provide some compelling evidence that Z-RAID provides increased performance. However, they cannot encompass the full complexity of a storage system and hence are based on some arguable simplifying assumptions. Hence, to further evaluate the performance of the Z-RAID technique we implemented a simulator. It includes a detailed disk model

Table 2. Experimental parameters for the Z-RAID Level 1 simulator.

Z-RAID Level 1	18 Disks (Seagate Cheetah X15)
Block size $\mathcal{B}$	0.25, 0.5, 1, 2, 4, 8 MB
Time period T_p	$(\frac{\mathcal{B}}{1.5\text{Mb/s}})$ sec
Throughput $\mathcal{N}_{Tot}$	< 4800
No. of stored clips	47
Object type	MPEG-1 (1.5 Mb/s)
Object size (length)	675 MB (1 hour)
Access distribution	Zipf

that was calibrated with parameters extracted from commercially available disk drives. To model user behavior, the simulator included a module to generate synthetic workloads based on various Poisson and Zipf distributions [18].

The simulator was implemented using the C programming language on a Sun server running Solaris and it consists of the following components. The *disk emulation* module imitates the response and behavior of a magnetic disk drive. The level of detail of such a model depends largely upon the desired accuracy of the results. Our model includes mechanical positioning delays (seeks and rotational latency) as well as variable transfer rates due to the common zone-bit-recording technique. The *file system* module provides the abstraction of files on top of the disk models and is responsible for the allocation of blocks and the maintenance of the free space. Either random or constrained block allocation were selectable with our file system. The *loader* module generates a synthetic set of continuous media objects that are stored in the file system as part of the initialization phase of the simulator. The *scheduler* module translates a user request into a sequence of real-time block retrievals. It implements the concept of a time period and enables the round-robin movement of consecutive block reads on behalf of each stream. Furthermore, it ensures that all real-time deadlines are met. Finally, the *workload* generator models user behavior and produces a synthetic trace of access requests to be executed against the stored objects. Both, the distribution of the request arrivals as well as the distribution of the object access frequency can be individually specified. For the purpose of our simulations, the request inter-arrival times were Poisson distributed while the object access frequency was modeled according to Zipf's law [18].

For the evaluation of RAID1 and Z-RAID1, the simulator was configured with a total of 18 disks of the Cheetah X15, each with 37 GB of space. Table 3.2 summarizes the rest of the simulation parameters.

For regular RAID1 mirroring, the data blocks were randomly distributed across all the zones of a disk. For Z-RAID1 mirroring, the primary copies of the data were constrained to the faster half of the disk drives. We tested retrieval block sizes of 0.25, 0.5, 1, 2, 4, and 8 MB and we executed the simulation with a nominal workload of $\lambda = 2,000$ requests per hour. The simulated database consisted of video clips whose display time was one hour long and which required a constant retrieval rate of 1.5 Mb/s (e.g., MPEG-1). This resulted in a uniform storage requirement of 675 Mbytes per clip. We also performed simulations of RAID5 and Z-RAID5 with the parity group size 5.

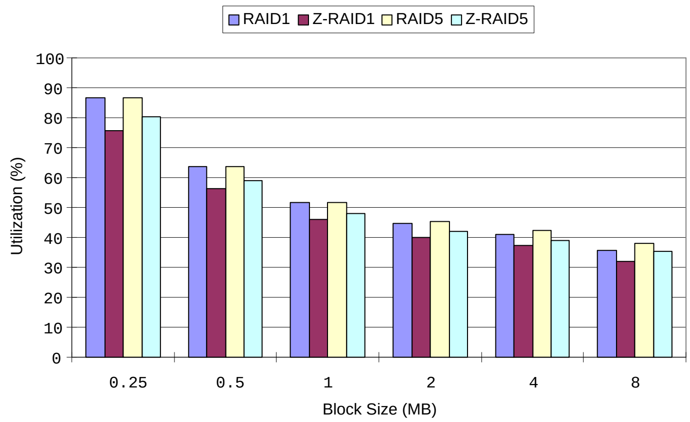


Fig. 6. Simulation results using Seagate Cheetah X15 disks.

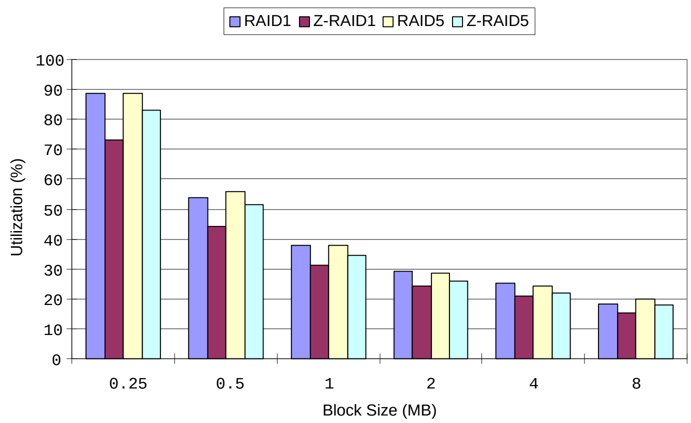


Fig. 7. Simulation results using Seagate Barracuda 7200.7 disks.

The frequency of access to different media clips is usually quite skewed for a video-on-demand system, i.e., a few newly released movies are very popular while most of the rest are accessed infrequently. The distribution pattern can be modeled using Zipf's law, which defines the access frequency of movie i to be $F(i) = \frac{c}{i^{1-d}}$, where c is a normalization constant and d controls how quickly the access frequency drops off. In our simulations, d was set to equal 0.271, which was chosen to approximate empirical data for rental movies [2]. For each experiment, the server had to service requests that arrived based on a Poisson distribution to simulate human behavior.

We focused on the disk utilization to compare the two techniques. A lower disk utilization – given a fixed workload – indicates a higher effective data transfer rate and a higher maximum throughput for the overall system. Because the effective bandwidth of a disk drive increases with larger block sizes, we expected to see a drop in disk utilization with increased block sizes. Figure 6 shows the results of the simulations using 18 Cheetah X15 disks, which depicts the reduction of the overall disk utilization of Z-RAID1 and 5 with a constant workload as compared with standard RAID 1 and 5. Z-RAID1 and 5 outperformed RAID1 and 5, respectively. For example, when the block size is 0.5 megabytes, the disk utilization of RAID1 was 64% while that of Z-RAID1 was 56% to service the same number of request. The percentage reduction of disk utilization between Z-RAID1 and RAID1 ranges from 11.1% (8 Mbytes of block size) to 13.6% (0.25 Mbytes of block size). Similar to the analytical comparisons, Z-RAID5 was performing lower than Z-RAID1 but still performing higher than RAID5.

We performed more simulations with different configuration using the Barracuda 7200.7. We used 33 disks and the workload was the same, $\lambda = 2,000$ requests per hour. Figure 7 shows similar results as the previous simulations with the Cheetah X15. The percentage reduction of disk utilization between Z-RAID1 and RAID1 ranges from 16.8% (8 Mbytes of block size) to 17.9% (0.25 Mbytes of block size).

Finally, we compared the performance of two disk models using RAID1 and Z-RAID1. The configuration used 18 disks and the workload was $\lambda = 1,500$ requests per hour. With a small block size the X15 provided a lower utilization than the 7200.7, because of its exceptionally small retrieval overhead (seek time plus rotational latency). However, as the block size increases the higher transfer rate of the 7200.7 becomes the dominant factor and allows it to achieve a lower utilization than the X15.

4 Conclusion

Our proposed Z-RAID system constrains the data block placement in a RAID system utilizing the zone characteristics of multi-zone disk drives. The constrained data placement and retrieval incur a shorter seek time between two adjacent block retrievals, which results in a reduced overhead for each block retrieval. Moreover, because the blocks are retrieved from the faster zones of a disk, the effective data transfer rate is increased further. Our analytical and simulation results for a streaming media server application demonstrate that both Z-RAID level 1 and 5 outperform the traditional RAID level 1 and 5, respectively.

The practical aspect of Z-RAID can be a more cost-effective and affordable system. Typically, RAID systems have been constructed from high performance disk drives such as SCSI disks. In general, those disks provide a higher transfer rate than other inexpensive disks such as IDE models. The drawback is a higher price. For cost-effectiveness, more economical RAIDs with IDE disks (IDE-RAID) have been recently introduced. We conclude that a Z-RAID system with IDE disks can provide the same high performance as a RAID system with high-end SCSI disks, but at the lower cost of IDE-RAID. Considering the recent trend showing that the performance gap between SCSI disks and IDE disks is narrowing (while the price gap still remains very significant), Z-RAID can provide an even better solution for a disk subsystem with inexpensive disks.

References

1. Y. Birk. Track-pairing: a novel data layout for vod servers with multi-zone-recording disks. In *IEEE International Conference on Multimedia Computing and System*, June 1995.
2. A. Dan, D. Sitaram, and P. Shahabuddin. Scheduling Policies for an On-Demand Video Server with Batching. In *Proceedings of the ACM Multimedia*, pages 391–398, 1994.
3. Ali Dashti, Seon Ho Kim, Cyrus Shahabi, and Roger Zimmermann. *Streaming Media Server Design*. Prentice Hall PTR, 2004.
4. J. Gafsi and E. W. Biersack. Modeling and Performance Comparison of Reliability Strategies for Distributed Video Servers. *IEEE Transactions on Parallel and Distributed Systems*, 11(4):412–430, 2000.
5. D. J. Gemmell, H. M. Vin, D. D. Kandlur, P. V. Rangan, and L. A. Rowe. Multimedia Storage Servers: A Tutorial. *IEEE Computer*, May 1995.
6. S. Ghandeharizadeh, S. H. Kim, C. Shahabi, and R. Zimmermann. Placement of Continuous Media in Multi-Zone Disks. In Soon M. Chung, editor, *Multimedia Information Storage and Management*, chapter 2. Kluwer Academic Publishers, Boston, August 1996. ISBN: 0-7923-9764-9.
7. S. Ghandeharizadeh and S.H. Kim. A Comparison of Alternative Continuous Display Techniques with Heterogeneous Disks. In *Proceedings of the International Conference on Information and Knowledge Management*, pages 442–449, 1999.
8. S. Ghandeharizadeh, S.H. Kim, W. Shi, and R. Zimmermann. On Minimizing Startup Latency in Scalable Continuous Media Servers. In *Proceedings of Multimedia Computing and Networking*, pages 144–155. Proc. SPIE 3020, Feb. 1997.
9. Jim Gray and Prashant Shenoy. Rules of thumb in data engineering. In *Proceedings of IEEE International Conference on Database Engineering*, Feb. 2000.
10. S. Kang and H.Y. Yeom. Storing Continuous Media Objects to Multi-Zone Recording Disks Using Multi-Rate Smoothing Technique. *IEEE Transactions on Multimedia*, 5(3):473–482, 2003.
11. R. Muntz, J. Santos, and S. Berson. RIO: A Real-time Multimedia Object Server. *ACM Sigmetrics Performance Evaluation Review*, 25(2), Sep. 1997.
12. S. W. Ng. Advances in Disk Technology: Performance Issues. *IEEE Computer Magazine*, pages 75–81, May 1998.
13. J. Nussbaumer, B. Patel, F. Schaffa, and J. Sternbenz. Network Requirement for Interactive Video-on-Demand. *IEEE Transactions on Selected Areas in Communications*, 13(5):779–787, 1995.
14. D. Patterson, G. Gibson, and R. Katz. A case for Redundant Arrays of Inexpensive Disks (RAID). In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, May 1988.
15. C. Ruemmler and J. Wilkes. An Introduction to Disk Drive Modeling. *IEEE Computer*, March 1994.
16. M.F. Mitoma S.R. Heltzer, J.M. Menon. Logical data tracks extending among a plurality of zones of physical tracks of one or more disk devices. In *U.S. Patent No. 5,202,799*, April 1993.
17. W.-C. Wang and et al. Fast data placement scheme for video servers with zoned-disks. In *Proceedings of Multimedia Storage and Archiving Systems II*, pages 92–102. Proc. SPIE 3229, 1997.
18. G. K. Zipf. *Human Behavior and the Principle of Least Effort*. Addison-Wesley, Reading MA, 1949.