Optimal Data Placement on Disks: A Comprehensive Solution for Different Technologies *

P. Triantafillou S. Christodoulakis C.A. Georgiadis

Laboratory of Distributed Multimedia Information Systems and Applications, and
Department of Electronics and Computer Engineering
Technical University of Crete
Crete, Greece †

Abstract

The problem of optimally placing data on disks (**ODP**) to maximize disk-access performance has long been recognized as important. Solutions to this problem have been reported for some widely available disk technologies, such as magnetic CAV and optical CLV disks. However, important new technologies such as **multi-zoned** magnetic disks, have been recently introduced. For such technologies no formal solution to the ODP problem has been reported. In this paper, we first identify the fundamental characteristics of disk-device technologies which influence the solution to the ODP problem. We develop a comprehensive solution to the problem that covers all currently available disk technologies. We show how our comprehensive solution can be reduced to the solutions for existing disk technologies, contributing thus a solution to the ODP problem for multi-zoned disks. Our analytical solution has been validated through simulations and through its reduction to the known solutions for particular disks. Finally, we study how the solution for multi-zoned disks is affected by the disk and data characteristics.

Index Terms: performance modeling and prediction, disk technologies, zoned disks, data placement

1 Introduction

In this paper, we consider a set of objects stored on a disk device. A disk drive partitions a platter's surface into a set of concentric tracks, containing a number of sectors. Sectors are the minimum unit of disk data that can be accessed. Consecutive tracks from all the platters form cylinders. Information is read and written onto the platter surfaces using a per-platter-surface read/write head. Each head is attached to a head-arm mechanical assembly which positions all the heads onto the desired cylinder. Finally, the disk pack is constantly revolving with the help of a spindle. When a request for a sector is issued, the head-arm assembly must *seek* to the cylinder containing the target sector. After the *rotational delay* needed for the target sector to be brought under the head, the sector can be transferred.

*This work was done in the context of and supported by the ESPRIT Long Term Research Project HERMES (project number 9141).

†Authors address: P.O. Box 133 Chania, 731 10 Crete, Greece. Phone: +30-821/64803, 69737, 69738. FAX: +30-821/64846. E-mail address: {peter,stavros,georgiad}@ced.tuc.gr

Disk Technologies

CAV disks consist of concentric tracks and have a constant angular velocity. The sectors at the outer tracks are longer but with a smaller recording density (bits/inch). This results in a waste of storage space. Therefore, the track capacity and the time to read/write a sector remain constant.

Recently, a new technology of magnetic disks emerged, namely, the **multi-zoned CAV** disk technology (**Zoned-CAV**). In such a disk, successive groups of cylinders are formed, called *zones*. Within each zone, the track capacity, in terms of number of sectors, and the transfer rate are the same. However, a track in a cylinder of a given zone contains more sectors than the track in a cylinder of the neighboring zone closer to the platter center. As a result, since the angular velocity remains constant, the transfer rate of the outer zones is higher than the transfer rate of the inner ones. This happens because in the outer cylinders the linear velocities are higher, so that more sectors per unit of time pass beneath the disk head. Zoned-CAV disks have become very popular due to their higher storage capacities and average transfer rates. These two characteristics make Zoned-CAV disks very attractive for many applications such as multimedia which demand large storage capacities and transfer rates [5].

Another disk type, often found in the optical disks, is the **C**onstant **L**inear **V**elocity (**CLV**) disk. In this technology, the recording density (bits/inch) remains constant throughout the disk platter. Thus, outer tracks have higher storage capacities than inner ones. The constant linear velocity of optical disks results in a constant transfer rate. In order for this to happen, the CLV format disks adjust their angular velocity, decreasing it as we move to outer tracks. Furthermore, instead of concentric tracks, these disks usually have a single spiral track. We can still define *tracks* by considering a radial line of the disk. A "track" is lying between two successive intersections of the spiral track with the radial line. The set of the successive tracks which all have the same capacity (in number of sectors) can be viewed as forming a zone, similarly to Zoned-CAV disks.

The following table summarizes the technologies presented earlier.

	CAV	**CLV**	**Zoned-CAV**
Type	magnetic	optical	magnetic
Constant	angular velocity storage capacity transfer rate	linear velocity transfer rate sector length	angular velocity
Variable	linear velocity sector length	angular velocity storage capacity	linear velocity storage capacity sector length transfer rate

The Problem

A measure of the performance of disks is the expected latency for randomly accessing a set of objects. Such a set of objects could be those that qualify in a query. The cost for accessing a set of objects depends on the placement of these objects. Among all possible permutations of the positions of the objects, there is one that minimizes the total expected cost. We refer to that placement as the *optimal data placement*.

The problem is formulated as follows: given the access probability for each object, we wish to determine their optimal placement on the disk so that accessing randomly any number of these objects incurs the minimum cost. In general, the fundamental characteristics of disk device technologies are their magnetic/optical nature, the angular and linear velocities, their storage capacities, and their transfer rates. Current disk technologies differ in terms of these characteristics. Angular and linear velocities, storage capacities, and transfer rates are made constant or variable in different existing disk products.

As previously stated, the optimal placement depends on the disk technology. The rotational delay and the transfer cost components of the CAV disks cannot influence the optimal data placement. Since the angular velocity and the storage capacity per track are constant, it follows that the rotational delay to access an object is constant. For the same reasons, the transfer rate is also independent of the object's position on the disk. Therefore, for CAV disks, the optimal data placement is only affected by the seek cost component. The solution for the optimal data placement is known as *the organ pipe permutation* ([10]).

For CLV disks, in addition to the seek cost, the rotational latency increases as we move towards the outer disk edge, since the angular velocity decreases as we move to the outer tracks. Thus, to obtain the optimal data placement the average sum of these costs should be minimized. In [4], an expression was developed for the optimal placement for CLV disks. In that work, the problem was solved analytically for a special case of the probability access distribution. The data were divided into a high access probability group and a low access probability group. Then, the optimal position for the data group with the higher access probability was determined.

In Zoned-CAV disks, again more than one cost components must be considered when solving the ODP problem. Zoned-CAV disks have both variable storage capacities and variable transfer rates. So, to obtain an optimal placement the transfer cost must be considered in the optimization process in addition to the seek cost.

Goals and Motivations

Neither CAV nor CLV optimal data placement results can be applied to Zoned-CAV disks, since none of them captures the variable transfer rates of tracks in Zoned-CAV disks. On the other hand, although the skewed storage capacity distribution seems to imply variations in the expected rotational delay, this is not true. The reason is that the angular velocity remains constant, which results in a constant average rotational delay equal to half of the disk rotation time. Therefore, the variable cost parameters involved are the seek and the transfer cost. Finally, because of the constant angular velocity and the almost constant length (inches) of the sectors of the Zoned-CAV disks, the rate of increase of the transfer rate is (almost) equal to the rate of the increase of the track storage capacity.

Thus, there are available disk technologies for which there is no known analytical solution to the ODP problem. In addition, even if a solution for Zoned-CAV disks was developed, new solutions might be needed for future disk technologies. In this paper, we address these concerns by developing a unifying solution for optimal data placement. This comprehensive solution is reducible to the known solutions for the CAV and CLV disks. In addition, it is reducible to the solution for the Zoned-CAV disks. Subsequently, we will study in detail the ODP problem for Zoned-CAV disks as it is affected by the disk and data characteristics, such as the seek cost and storage capacity slopes, the transfer times, the block size, the access distributions, etc.

2 A Comprehensive Solution

There are three costs that constitute a random access cost on a disk: *seek*, *rotational delay* and the *transfer* costs. It is generally agreed that a general access probability distribution would make very hard every attempt to develop a closed formula solution for the problem. Therefore, in our analysis, we will restrict the object access probabilities to two values ([4]), a high one and a low one. Next, we present the disk and the access models and a brief description of an analysis that results in the optimal data placement determination. Finally, we will show how one can derive the optimal data placement solution for some of the currently available disk technologies (i.e., CAV, CLV, Zoned-CAV).

2.1 The Disk and the Access Cost Models

Solving for the optimal data placement using a discrete expression, is difficult. These difficulties can be overcome by mapping the discrete model into a continuous one. The large capacities of today's disks make such a mapping reasonable. We will thus discuss probability masses, instead of qualifying sectors.

2.1.1 Storage Capacity and Transfer Rate Distribution

In the continuous model, a disk position ranges from 0, for the innermost track, to 1, for the outermost track. We will determine the storage capacity distribution in terms relative to the middle position. The storage capacity of the middle position is by definition one unit. The slope of the storage distribution is denoted by k and the intercept, i.e., the capacity of the innermost position, is denoted by l (figure 1). Therefore, the capacity of position x is:

$$C(x) = k \cdot x + l, \quad 0 \leq x \leq 1,\ 0 \leq l \leq 1,\ 0 \leq k \leq 2 \tag{1}$$

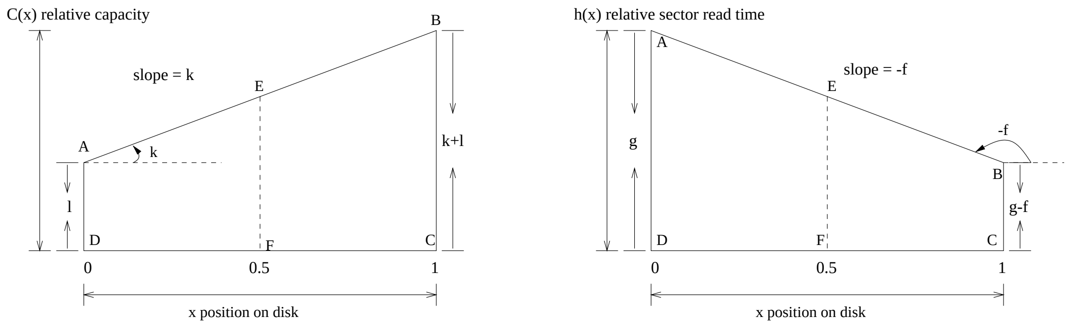

Figure 1: Model of storage capacity and transfer rate distribution

Obviously, l cannot be less than zero. In addition, the middle track capacity is always greater than or equal to the innermost track capacity, resulting in a maximum l value of 1. The minimum value for the slope k is zero (i.e., $C(x)$ is parallel to the x-axis) and occurs when l equals 1. The maximum value for k occurs when l equals to 0. When this happens, the vertices A and D coincide and from the similarity of the triangles ABC and AEF we gain the analogy: $\frac{BC}{EF} = \frac{AC}{AF} \Rightarrow \frac{k}{1} = \frac{1}{0.5} \Rightarrow k = 2$. Note that, because the area of the trapezoid is equal to one, we can express the intercept l as a function of the slope k: $(AD + BC)\frac{CD}{2} = (l + (k + l))\frac{1}{2} = (EF) \cdot (CD) = 1 \cdot 1 \Rightarrow \frac{k+2l}{2} = 1 \Rightarrow l = \frac{2-k}{2}$. Thus, given the storage capacity of the innermost track and the outermost track of any real disk, we can compute the slope k of our model. For example, consider the disk HP 2240 with innermost track capacity of 56 sectors and outermost track capacity of 96 sectors. Then, the middle track has $(96 + 56)/2 = 76$ sectors and therefore the capacity of the innermost track relatively to the middle track is $l = 56/76 = 0.7368$. Finally, from equation $l = \frac{2-k}{2}$, we derive that $k = 0.5263$.

We will express the read sector time in terms also relative to the middle position of the disk. Given the relative slope $-f$ of the change in the time to read a sector across the disk, and the relative time g to read a sector from the innermost disk position, we have:

$$h(x) = -f \cdot x + g, \quad 0 \leq x \leq 1,\ 1 \leq g \leq 2,\ 0 \leq f \leq 2 \tag{2}$$

By definition, the time to read a sector from the middle disk position is 1 time unit. The sector read time of the innermost position can never be less than the sector read time of the middle disk position, resulting in a minimum value for g equal to 1. We can see from figure 1 that when g increases, the slope $-f$ decreases and, therefore, f increases, and vice versa. Because $h(1) = 0 \Leftrightarrow g = f$ holds, g is maximized meaning that f is also maximized. When this happens, the vertices B and C coincide and f is equal to the tangent $\frac{EF}{FC} = 1/0.5 = 2$. The minimum value for f is observed in uniform sector read time distribution ($-fx + g = g$, $\forall x \Rightarrow f = 0$). As in the case with the storage capacity distribution, the intercept g can be expressed as a function of the slope $-f$. The area of the trapezoid again equals to 1 and therefore $(AD + BC)\frac{CD}{2} = (g + (g - f))\frac{1}{2} \Rightarrow \frac{2g-f}{2} = 1 \Rightarrow g = \frac{2+f}{2}$.

2.1.2 Random Access Probability Distribution

Our model assumes that sector requests are independent of each other. In addition, we restrict the access probabilities to two values, P_1 and P_2 ($P_1 > P_2$), relative to each other (e.g., $P_1 = 16$ and $P_2 = 1$ means that P_1, relatively to P_2, is 16 times greater). We denote the proportion of point masses (e.g., the proportion of sectors in the discrete model) with relative access probability P_1 by r and the proportion accessed with probability P_2 by $1 - r$. Letting μ be the number of probability points (its discrete equivalent is the total number of the sectors) and $N = P_1 r + P_2(1-r)$ be a normalizing factor, then the absolute access probability of a probability point having relative probability density value P_1 is $\frac{P_1}{N}\frac{1}{\mu}$, and for a point with relative probability P_2 is $\frac{P_2}{N}\frac{1}{\mu}$.

$$P(\text{a point is accessed with relative prob. } P_1) = \frac{P_1}{N}\frac{1}{\mu} \tag{3}$$

$$P(\text{a point is accessed with relative prob. } P_2) = \frac{P_2}{N}\frac{1}{\mu} \tag{4}$$

2.1.3 Seek, Rotational and Transfer Cost Functions

The formulas that we develop here refer to the costs incurred at the middle track (position) of the disk. The seek cost function has two branches, one for the acceleration phase (square root function of distance) and one for the linear phase (linear function of distance) ([6]). However, in order to handle the mathematics that follow, we consider a simplified seek model. The seek cost function is given by the cost function of the linear phase and for a distance d is given by

$$Sk(d) = a \cdot d + b, \qquad 0 \leq d < 1 \tag{5}$$

Note that since $0 \leq d < 1$ our seek cost slope values will be significantly different than those reported for real disk products. For example, the reported slope for the (long) seek cost slope of the HP C2200A is 0.012. In our model, this corresponds to a seek cost slope value of 17.38[1]. In other words, setting $a = 17.38$ our model can be used to match the seek cost model for the above disk product.

A second simplification is that, we will not model proximal window accesses[2] of optical disks. The reason is that the number of window accesses is negligible compared to the number of seek accesses.

[1] Let us illustrate the derivation for the value of a. Denote the seek cost slope a_{real}, the total number of cylinders T_{real}, and a given seek distance (in number of cylinders) d_{real} for the HP C2200A disk. We must have $a_{real}\, d_{real} = a\, d$. We have that $d = d_{real}/T_{real}$. Thus $a = a_{real}\, T_{real}$.

[2] In optical disks, there are two ways of accessing a track: a *proximal window* mode, for tracks very close to the source track, and a *seek* mode, for the rest of the track accesses ([2, 3]). In the proximal window access mode, there is no seek operation. Instead, a nearby targeted track can be accessed by simply diverting the laser beam onto it.

As mentioned, the average rotational delay is usually modeled as half revolution. Therefore, the rotational delay, at disk position x, is:

$$Rd(x) = \frac{1}{2}h(x) \cdot F_1 \cdot C(x), \qquad 0 \leq x \leq 1 \tag{6}$$

where $h(x)$ is the read sector time at position x, F_1 is the time to read the middle disk position, and $C(x)$ is the storage capacity of disk position x.

Letting B be the block size in number of sectors and F_2 the time to read a sector from the middle disk position, the transfer cost function becomes:

$$Tr(x) = h(x) \cdot F_2 \cdot B, \qquad 0 \leq x \leq 1 \tag{7}$$

2.1.4 The Expected Random Access Cost

The expected access cost is computed by summing the costs for successive accesses to each possible pair of initial and destination positions on the disk. Each term of the sum must be weighted by the product of the access probability of the initial and the destination position. The cost function[3], is:

$$Cost = \int_0^1 \int_0^1 Pm(x)Pm(y)(Sk(|x-y|) + Rd(y) + Tr(y))dxdy \tag{8}$$

where x and y are the initial and the target positions on the disk respectively, $Pm(x)$ is the probability of accessing position x, $Sk(|x-y|)$ is the seek cost between positions x and y, $Rd(y)$ is the rotational delay at position y, and $Tr(y)$ is the transfer cost at position y. Incorporating these equations into (8) we get:

$$Cost \;=\; b + 2a\int_0^1 \int_x^1 Pm(x)Pm(y)(y-x)dydx + \int_0^1 Pm(x)Rd(x)dx + \int_0^1 Pm(x)Tr(x)dx \tag{9}$$

2.2 Consecutivity and Unimodality

Below we give two results about the form which an optimal object placement will take[4]. Specifically, a position on the continuous disk must contain only probability masses of one of the two subsets, P_1 and P_2. Additionally, the form of any solution must be *unimodal*, meaning that in an optimal placement, the probability masses from the P_1 group must be placed as close together as possible.

Consecutivity : *In an optimal arrangement, there cannot exist two disk portions x and y ($x \neq y$), such that there are two probability points p_1 and p_3 at position x and one point p_2 at position y, such that $p_1 < p_2 < p_3$.*

Unimodality : *There cannot exist three disk portions x, y and z with access probabilities P_1, P_2 and P_1 ($P_1 > P_2$) respectively such that the inequality $x < y < z$ holds.*

The above theorems imply that the hot data must be placed together in a single group. This means that the cold data form at most two (cold) groups (if the hot group is placed between them)[5].

[3] The details of mapping the discrete cost function into that for the continuous model can be found in [8].

[4] A proof for these theorems for CLV disks can be found in [4]. The proof for the comprehensive disk model can be found in ([8]).

[5] With the help of these theorems, a general access probability distribution can be handled by a heuristic similar to the one developed in [4]. After determining the placement of the three groups of data on the disk with the use of the placement algorithm, we restrict on each group and re-apply the placement algorithm on this group and on the portion of the disk it occupies. The above procedure continues recursively until the placement problem becomes trivial, i.e., until one or two objects (masses) have to be placed.

2.3 Analysis

The idea behind the following analysis is to break the limits of the integrals already developed (see equation (9)) in order to incorporate the middle (m) of the P_1 area (for which the optimal position is sought) in the expected cost formula. Then, we will be able to find out for which value of m this cost is minimized.

Because of the unimodality theorem, the solution of the optimal data placement problem will be given by determining the position of the P_1 probability mass on the disk. More precisely, the solution will be the estimation of the middle point of the P_1 area. We denote this point as m:

$$m = \frac{X_l + X_r}{2}$$

where X_l and X_r is the left and right boundary, respectively, of the P_1 area (see figure 2).

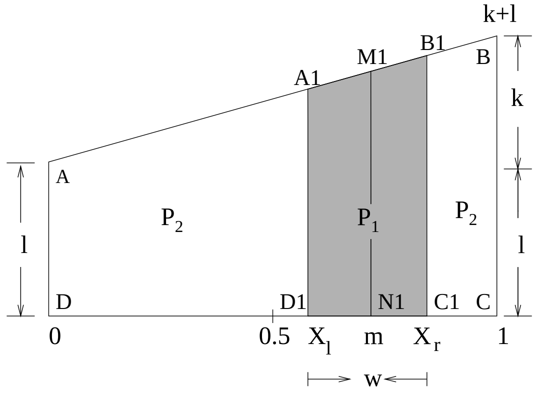

Figure 2: Optimal placement parameters

Let w be the width of the P_1 group. Then,

$$\begin{aligned} r &= \frac{\text{area of } P_1}{\text{total area}} = \frac{\frac{(A_1D_1+B_1C_1)}{2}(C_1D_1)}{\frac{(AD+BC)}{2}\cdot(CD)} = \frac{(M_1N_1)(C_1D_1)}{\frac{(AD+BC)}{2}(CD)} \\ &= \frac{C(m)w}{C(0.5)\cdot 1} = \frac{C(m)w}{1\cdot 1} = C(m)w \end{aligned} \tag{10}$$

and therefore:

$$w = \frac{r}{C(m)} \tag{11}$$

Note, that the total area equals to 1, i.e., $\mu = 1$. Additionally, the middle position of the P_1 area must be such that the following inequalities hold:

$$m - \frac{w}{2} \geq 0 \quad \text{and} \quad m + \frac{w}{2} \leq 1$$

From the above we can derive the extreme values of the m variable:

$$m_{min} = \frac{w}{2} \quad \text{and} \quad m_{max} = 1 - \frac{w}{2}$$

Using the last equalities and equations (11) and (5) we get:

$$m_{min} = \begin{cases} \frac{-2+k+\sqrt{4-4\,k+k^2+8\,k\,r}}{4\,k} & \text{, if } k \neq 0 \\ \frac{r}{2} & \text{, otherwise} \end{cases} \text{, } m_{max} = \begin{cases} \frac{-2+3\,k+\sqrt{4+4\,k+k^2-8\,k\,r}}{4\,k} & \text{, if } k \neq 0 \\ 1-\frac{r}{2} & \text{, otherwise} \end{cases}$$

Simple expressions for X_l and X_r can now be derived:

$$X_l(r,m) = m - \frac{w}{2} = m - \frac{1}{2}\frac{r}{C(m)} \qquad X_r(r,m) = m + \frac{w}{2} = m + \frac{1}{2}\frac{r}{C(m)}$$

Given these bounds and that $N = P_1 r + P_2(1-r)$, and using the expressions (3) and (4) we can specify our probability assignment function:

$$Pm(x) \quad = \quad \begin{cases} \frac{P_1}{N}\frac{1}{\mu}C(x) & \text{, if } X_l \le x \le X_r \\ \frac{P_2}{N}\frac{1}{\mu}C(x) & \text{, otherwise} \end{cases} = \begin{cases} \frac{P_1}{N}C(x) & \text{, if } X_l \le x \le X_r \\ \frac{P_2}{N}C(x) & \text{, otherwise} \end{cases} \tag{12}$$

We then substitute the three cost components in the cost expression and a random access cost expression is derived, as a function of the middle of the P_1 area. In order to simplify the expression, we set $P_2 = 1$. Using the Maple symbolic mathematics package ([1]) we obtain the expression for the cost metric.

$$Cost(k,f,a,b,P_1,r,B,F_1,F_2,m) = -\frac{C_1(k,f,a,b,P_1,r,B,F_1,F_2,m)}{120\ (P_1\, r + 1 - r)^2\,(2\,km + 2 - k)^5} \tag{13}$$

$C_1(k,f,a,b,P_1,r,B,F_1,F_2,m)$ is an expression computed by Maple which is ten pages long and therefore is omitted. The optimal value for m will be derived from

$$\frac{\partial Cost(k,f,a,b,P_1,r,B,F_1,F_2,m)}{\partial m} = 0 \tag{14}$$

Only one root is lying in the interval $[0,1]$. This is our optimal position ($m_{optimal}$).

3 Reductions of the Comprehensive Solution

3.1 Optimal Position for CAV Disks

To find $m_{optimal}$ for the CAV disks, we need to set $k = 0$ since the track capacity of the CAV is constant. Additionally, the constant transfer rate in a CAV disk implies that $f = 0$. Substituting $k = f = 0$ in equation (14) and solving for m, we get: $m_{optimal} = 0.5$. This result agrees with the known optimal organ pipe placement for CAV disks.

3.2 Optimal Position for CLV Disks

For the CLV disks, the transfer rate is constant while the track capacity increases when we travel away from the disk center. This means that the slope of the transfer rate should be set to 0 ($f = 0$), while the storage capacity slope is greater that zero. Setting $f = 0$ in equation (14) and solving for m, we get the optimal position which matches the one developed in [4].

3.3 Optimal Position for Zoned-CAV Disks

For the Zoned-CAV disks, both the storage capacity and the transfer rate vary. The constraint, however, is that, because the angular velocity and the physical size of a sector are constant, the rate with which the track capacity increases is equal to the rate with which the read sector time decreases. In terms of mathematics, we set $f = -k$. Additionally, $F_1 = 0$, since the rotational delay is constant.

Using the Maple package and the constraint $f = -k$, we solve for the optimal position:

$$m_{optimal} = \frac{1}{2} - \frac{1}{k} + \frac{1}{12\,k}\sqrt{\frac{6\,C_2}{a}}$$

$$\begin{aligned} C_2 &= 3\,ak^2 + 6\,k^2 BF_2\,P_1\,r + 6\,k^2 BF_2 - 6\,k^2 BF_2\,r + 12\,a + \sqrt{3\,C_3} \\ C_3 &= 24\,k^4 B^2 F_2^2 P_1\,r - 12\,ak^4 BF_2\,r + 12\,k^4 B^2 F_2^2 r^2 + 12\,ak^4 BF_2\,P_1\,r + 3\,a^2 k^4 + 12\,k^4 B^2 F_2^2 \\ &\quad + 12\,ak^4 BF_2 + 12\,k^4 B^2 F_2^2 P_1^2 r^2 - 24\,k^4 B^2 F_2^2 r - 24\,k^4 B^2 F_2^2 r^2 P_1 + 48\,ak^2 BF_2 \\ &\quad - 48\,ak^2 BF_2\,r - 48\,a^2 k^2 r^2 + 48\,ak^2 BF_2\,P_1\,r + 24\,a^2 k^2 + 32\,a^2 k^2 P_1\,r^2 + 48\,a^2 \end{aligned}$$

3.3.1 Model Validation

Here, we report on our efforts to validate the analytical solution and the Maple computations for the ODP solution for the Zoned-CAV disk. Our simulator used the model of Section 2 and run many times, for various disk configurations and user behaviours. In each run, we let only one of the parameters k, a, P_1, B, F_2, and r, be a free variable. For each value of a free variable, we placed our hot data in consecutive positions on the disk surface, and we measured the average access cost per request. Then we plotted the disk position for which the average access cost is minimized, as a function of the free variable. In figures 3, we can see that there is a very close match[6].

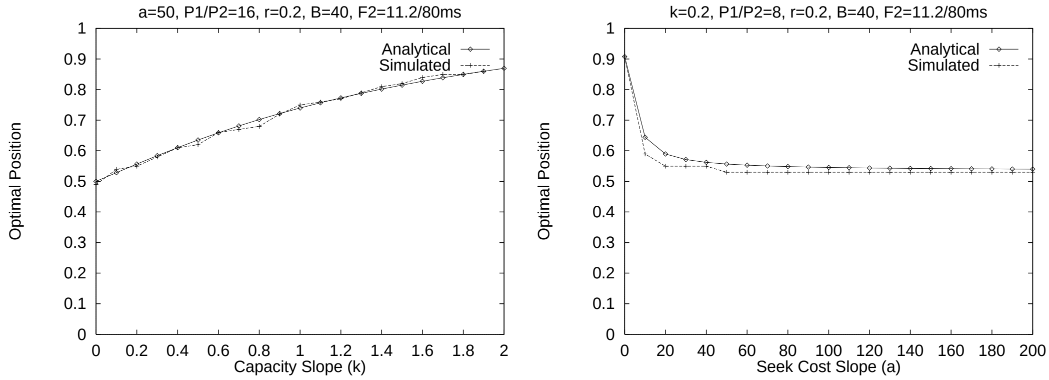

Figure 3: Validation

4 Performance and Data Placement in Zoned-CAV Disks

In figures 4-10, we show how the optimal position is affected by the disk and data parameters.

Impact of Storage Capacity Slope

In figure 4, we can see that an increase of the storage capacity slope k 'pushes' ($m_{optimal}$) to the outer edge. This happens for two reasons. First, because the outer tracks have larger capacities and thus higher access probabilities. If our hot data are placed in the outer tracks then the head movements will be fewer and shorter. Second, the transfer cost in those tracks is reduced. Increasing the seek cost slope a, 'pushes' $m_{optimal}$ to the outer track slower. Note that when $k = 0$, $m_{optimal} = 0.5$, as expected, since then our Zoned-CAV reduces to a CAV disk.

[6] For space reasons, we show only the validation for the cases where the capacity slopes and the seek cost slopes are a free variable. The validation of other cases can be found in [8].

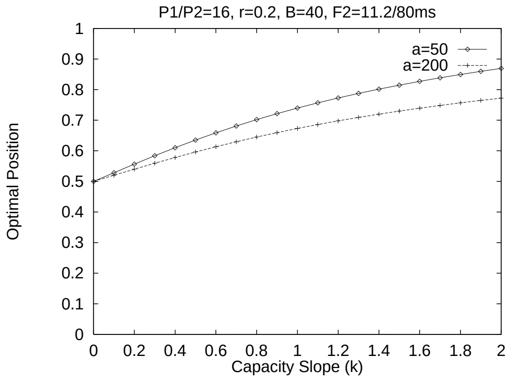

Figure 4: Impact of storage capacity slope.

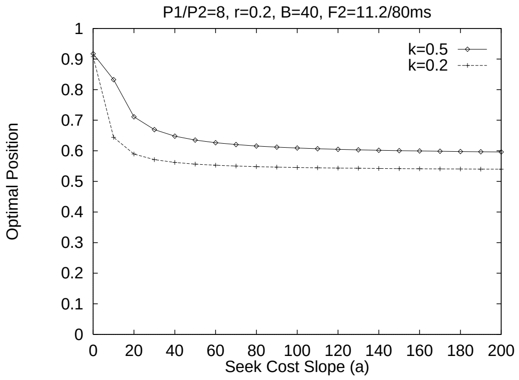

Figure 5: Impact of seek cost slope.

Impact of Seek Cost Slope

In figure 5, $m_{optimal}$ is shifted towards 0.5 when a increases. When a is high, seeks become more expensive if the hot area is placed at the outer tracks, because seeks to the cold data are longer. Thus, when the seek dominates the access cost, we should place hot data closer to the middle tracks. For greater k values, $m_{optimal}$ tends to remain closer to the outer edge.

Impact of High Access Probability

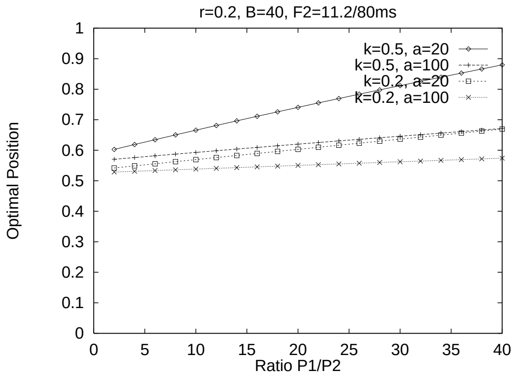

Figure 6: Impact of high access probability.

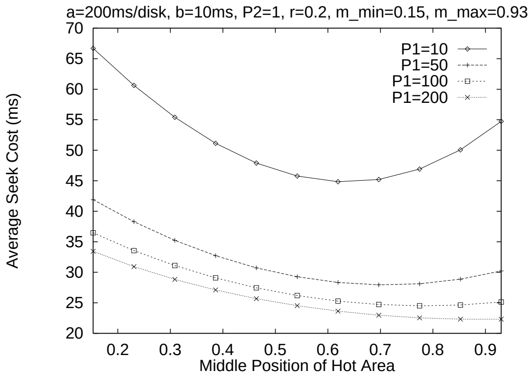

Figure 7: Impact of placement of the hot area on the seek cost.

The hotter the most frequently accessed data, the closer to the outer tracks is the optimal position (figure 6). For very hot data the head will almost always be lying around these data, thus performing very few long seeks. If we seldomly leave the hot area, then the expected random access cost is minimized close to the outer edge, since both the transfer cost and seek cost are minimized there. The reduction of the transfer cost is obvious since transfer rates are higher in the outer tracks. The seek cost is minimum near the edge because there the data occupy fewer tracks. Thus, the seeks within the hot area will now be shorter. The minimization of the seek cost can be seen in figure 7.

Impact of Proportion of the Hot Data

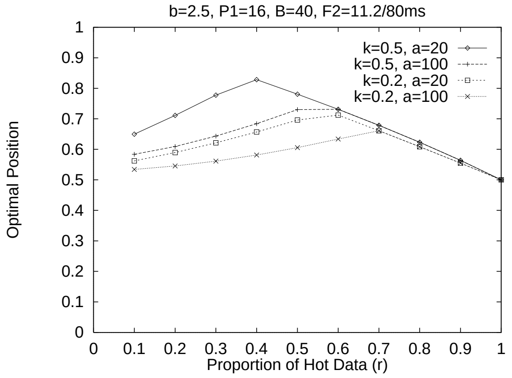

Figure 8: Impact of proportion of the most frequently accessed data.

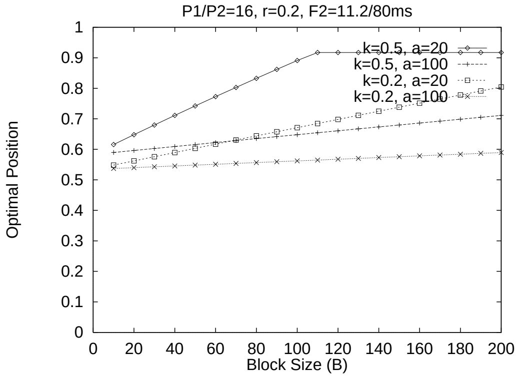

Figure 9: Impact of block size.

When the proportion of the hot data increases, the optimal position moves away from the center in order to occupy as many "faster" and "larger" tracks as possible. When r becomes the whole of the disk, then the middle position m of the P_1 is shifted to the middle of the disk.

Impact of Block Size

In figure 9, we can see that larger block sizes imply a shift of the optimal position towards the outer disk edge. This follows since for large blocks the transfer cost is more significant and the transfer rate at the outer tracks is greater. For large values of the storage capacity slope (k), the optimal position shifts faster away from the center. For the same values of the storage capacity slope, larger seek cost slopes result in the optimal position staying closer to the middle of the disk.

Impact of Read Sector Time

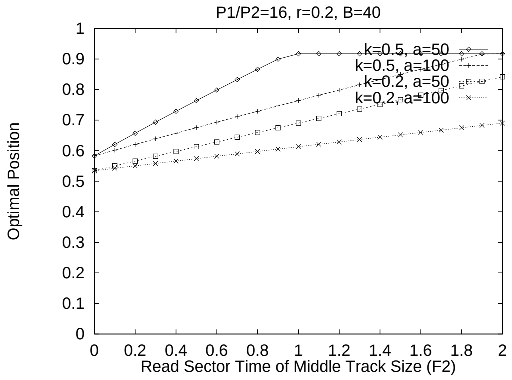

Figure 10: Impact of read sector time of the middle track.

Slower disk rotation implies a higher transfer cost. The minimization of the transfer cost occurs at the outer tracks (figure 10). The ratio of the slopes k and a is a measure of the speed with which $m_{optimal}$ is shifted to the outer tracks. Note that if the read sector time is very small, the seek cost slope does not have any impact on $m_{optimal}$ since only the seek cost needs minimization.

Generally, the optimal position depends strongly on the slope of the storage capacity, the relative access probability between the hot and the cold data, and the slope of the seek cost. Large values of the first two parameters push the optimal position towards the outer edge of the disk while large values of the latter tend to keep the optimal position near the center of the disk.

5 Conclusions

Solutions to the problem of optimal data placement (ODP) on disks are important since they can help significantly improve the performance of the I/O system, which is critical for many applications. Solutions to the ODP problem for 'traditional' disk devices have been developed and reported in the literature. However, newer technologies have emerged (such as the multi-zone CAV disks) for which no formal solutions have been developed. These technologies offer significant advantages in terms of higher transfer rates and larger storage capacities. (Furthermore, others may emerge in the future, since it is recognized by all that the I/O subsystems have become the performance bottlenecks). Thus, there is a need to view the ODP problem comprehensively, instead of developing ad hoc solutions for particular technologies.

This paper addresses this issue. We identified the fundamental characteristics of disks, how these affect the access costs, and how do the solutions to the ODP problem depend on these characteristics. Subsequently, we developed a formal and comprehensive solution to the ODP problem. Our formal analysis has been validated against simulations of disk accesses. The comprehensive solution is reducible to the known solutions for magnetic CAV and optical CLV disks (thus, providing another form of validation).

Our solution is also reducible to the solution for the multi-zone CAV disks (and because it is based on the fundamental characteristics of disks can also cover possible future technologies which are based on a combination of these characteristics). In this way, we developed a solution which can be used for existing disk technologies (for which no solution exists) and for possible/future technologies. Our analysis also gave us the opportunity to study in detail the performance of multi-zone disks as that is affected by the data placement and the particular characteristics of the disks and of the data.

References

[1] B.C. Char, K.O.Geddes, M.B. Monagan, S.M. Watt, "MAPLE Reference Manual", 5^{th} Edition, University of Waterloo, Waterloo, Ontario, March 1988.

[2] S. Christodoulakis, "Analysis of Retrieval Performance for Records and Objects Using Optical Disk Technology", *ACM Transaction on Database Systems*, Vol. 12, No. 2, pp. 137-169, June 1987.

[3] S. Christodoulakis and D.A. Ford, "Performance Analysis and Fundamental Trade Offs for Optical Disks", *Proceedings of SIGMOD'88*, Chicago, pp. 286-294, June 1988.

[4] S. Christodoulakis and D.A. Ford, "Optimal Data Placement on CLV Optical Discs", *ACM Transactions on Information Systems (ACM TOIS)*, 1991.

[5] G. Nerjes, P. Muth and G. Weikum, "Stochastic Service Guarantees for Continuous Data on Multi-Zone Disks", *ACM Symposium on Principles of Database Systems*, May, 1997.

[6] C. Ruemmler and J. Wilkes, "An Introduction to Disk Drive Modeling", *IEEE Computer*, pp. 17-28, March 1994.

[7] Ralf Steinmetz and Klara Nahrstedt, "Multimedia: Computing, Communications and Applications", *Prentice Hall PTR* , 1995.

[8] P. Triantafillou, S. Christodoulakis, and C. Georgiadis, "Optimal Data Placement on Disks: A Comprehensive Solution for Different Technologies", *MU.S.I.C. Technical Report*, 1996.

[9] C. Yu, et al, "Efficient Placement of Audio Data on Optical Disks for Real-Time Applications", *CACM*, July 1989, Vol. 32, No. 7.

[10] C.K. Wong, "Algorithmic Studies in Mass Storage Systems", *Springer-Verlag*, Berlin, 1983.