

# Conditions of Devices for Effective Tiered Storage

Kazuho Fujii

## 1 Introduction

Storage tiering is a technique to create a high performance and cheap storage by combining several kinds of storage devices. Tiered storage system allocates frequently accessed data to high performance but expensive devices, and it allocates rarely accessed data to cheaper but slower device. The system collects access statistics and sometimes optimizes data allocation. For example, SSD, SAS HDD and SATA HDD are usually used for each tiers.

Caching is a similar technique with tiering. Caching duplicates data from poor performance devices to high performance device when the data is accessed. Compared to tiering, caching reallocates data more frequently. Additionally we can use volatile memory for caching tiers such as RAM.

Only kinds of devices do not make storage tiers. Tiers can be made with data formats (e.g., compressed data or raw data) and data positions (e.g., over the network or not). All schemes of saving data can be regarded as tiers. That is, storage tiering is not technique only for storage systems; all computer system is related to it.

Most of tiering technologies concern about data replacement. It is a popular problem which data should be on high performance storage. Many schemes are proposed that predict the next accessed data and enhance probability of "cache hit". LRU (least recently used) method and LFU (least frequently used) method are well known.

In contrast constitution of tiers has not been concerned. This is because configuration of tiers was static; it could not be changed flexibly. However, vitalization is general technology now. We can create virtual tiered storage consisted as we want.

Although there are many kinds of storage devices (including data format), it is unclear what kinds of devices we should combine or how we should decide the ratio of devices. If we can know them, we can design effective storage.

In this report, I theoretically derive the optimum configuration of tiered storage. The rest of this report is organized as follows: Section 2 setups the analyzed system; Section 3 introduces the assumption of the Pareto distribution access; Section 4 calculates the mean performance of storage under the Pareto distribution access; Section 5 derives theoretically optimized configuration of tiered storage; Section 6 discusses about the condition of devices for effective tiering; Section 7 describes the conclusion.

## 2 Setup

Now I think about creating tiered storage with  $N$  kinds of recording devices respectively named Device-1, Device-2,  $\dots$ , Device- $N$ . Device-1 has the lowest performance but it is the most cost-effective. Device- $N$  has the highest performance but it is the most expensive.

I introduce performance values of devices. The performance value of Device- $i$  is expressed with  $q_i$ . Then,  $0 < q_1 < q_2 < \dots < q_N$ . The performance value of a device represents any performance characteristic of the device (e.g., random read, sequential read) with a single value. Of course measuring this value is difficult, but defining it is convenient for a theoretical discussion.

Similarly I express the cost per a unit capacity of Device- $i$  with  $c_i$ . Then,  $0 < c_1 < c_2 < \dots < c_N$ . Caching copy data from a device to another device. In this case  $c_i$  includes costs of both of the devices.

I express the capacity ratio of Device- $i$  to the total capacity of the volume with  $r_i$ . By definition,

$$\sum_{i=1}^N r_i = 1. \quad (1)$$

The mean cost per a unit capacity  $\bar{c}$  is

$$\bar{c} = \sum_{i=1}^N r_i c_i. \quad (2)$$

The goal is to get the optimum values of  $r_i$ . Values  $q_i, c_i$  are given as catalog specifications of storage devices. And  $\bar{c}$  is given as a budget. Hence, the required values are  $r_i$  where the tiered storage get the highest performance.

## 3 Pareto Distribution

The optimum values of  $r_i$  are different with access pattern to the storage. We have to assume a probability distribution of access to blocks of the storage.

Here, I assume that the access obeys the Pareto distribution. It is sometimes called the law of 80:20 or the Zipf's law. Under the Pareto distribution just few data are accessed mostly and most of data are rarely accessed. This distribution can express "hot spots" and "long tails". Tiering is effective in such cases.

In the Pareto distribution, the probability  $F$ , where the number of access to a block is not more than  $x$ , is

$$F = 1 - \left(\frac{x_0}{x}\right)^a, \quad (a > 1, x_0 > 0). \quad (3)$$

$a$  and  $x_0$  are parameters of the probability distribution.  $a$  means bias of access. It is called Pareto coefficient. As  $a \approx 1$  access is concentrated to a single block; as  $a \rightarrow \infty$  access is uniformly distributed, that is the number of access at each block is  $x_0$ .

The probability density function  $f$  is

$$f := \frac{\partial F}{\partial x} = \frac{ax_0^a}{x^{a+1}}. \quad (4)$$

## 4 Mean Performance

For the next, I calculate the mean performance of the storage under the Pareto distribution access pattern.

I assume that the tiering is perfect. That is the higher accessed block is always allocated to the higher performance device. Let Device- $i$  to have blocks which access count  $x$  is  $x_{i-1} \leq x < x_i$ . Note that  $x_N = \infty$ . Then, the relationship between  $r_i$  and  $x_i$  is

$$r_1 + r_2 + \dots + r_i = 1 - \left(\frac{x_0}{x_i}\right)^a. \quad (5)$$

The number of access to the tier of Device- $i$  equals  $\int_{x_{i-1}}^{x_i} x f dx$ , and the total number of access to the storage equals  $\int_{x_0}^{\infty} x f dx$ . Therefore, the probability where an access is to the tier of Device- $i$  is

$$\begin{aligned} p_i &= \frac{\int_{x_{i-1}}^{x_i} x f dx}{\int_{x_0}^{\infty} x f dx} \\ &= (1 - r_1 - \dots - r_{i-1})^{\frac{a-1}{a}} - (1 - r_1 - \dots - r_i)^{\frac{a-1}{a}}. \end{aligned} \quad (6)$$

Consequently, I can get the mean performance of tiered storage  $\bar{q}$ . That is

$$\begin{aligned} \bar{q} &= \sum_{i=1}^N p_i q_i \\ &= \sum_{i=1}^N \left\{ (1 - r_1 - \dots - r_{i-1})^{\frac{a-1}{a}} - (1 - r_1 - \dots - r_i)^{\frac{a-1}{a}} \right\} q_i. \end{aligned} \quad (7)$$

## 5 The Optimum Ratio of Devices

In this section I derive the optimum values of  $r_i$ . Now I want to derive the ratio of each drives  $r_1, r_2, \dots, r_N$  which maximizes the mean performance value  $\bar{q}$  under the constraints

(1) (2). I can use the method of Lagrange multiplier for this problem. With multiplier  $\mu$  and  $\lambda$ , I define a function  $g$  as

$$g := \bar{q} - \lambda \left( \bar{c} - \sum_{i=1}^N r_i c_i \right) - \mu \left( 1 - \sum_{i=1}^N r_i \right). \quad (8)$$

I can obtain  $r_1, r_2, \dots, r_N$  where  $\bar{q}$  get an extreme value by solving the simultaneous equations of (1) (2) and following  $N$  equations:

$$\frac{\partial g}{\partial r_i} = 0, \quad (i = 1, 2, \dots, N). \quad (9)$$

By solving the simultaneous equations (1) (2) (7) (9), I can get the optimum ratio of each device  $r_1, r_2, \dots, r_N$  as following:

$$r_1 = (\bar{c} - c_1) \left\{ \sum_{i=1}^{N-1} \Delta c_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a} \right\}^{-1} \left\{ - \left( \frac{\Delta c_1}{\Delta q_1} \right)^{-a} \right\} + 1, \quad (10)$$

$$r_i = (\bar{c} - c_1) \left\{ \sum_{i=1}^{N-1} \Delta c_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a} \right\}^{-1} \left\{ \left( \frac{\Delta c_{i-1}}{\Delta q_{i-1}} \right)^{-a} - \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a} \right\}, \quad (11)$$

$$r_N = (\bar{c} - c_1) \left\{ \sum_{i=1}^{N-1} \Delta c_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a} \right\}^{-1} \left\{ \left( \frac{\Delta c_{N-1}}{\Delta q_{N-1}} \right)^{-a} \right\}, \quad (12)$$

where

$$\Delta c_i := c_{i+1} - c_i, \quad (13)$$

$$\Delta q_i := q_{i+1} - q_i. \quad (14)$$

Note that  $r_i$  in the equation (11) are in cases of  $i = 2, 3, \dots, N - 1$ . In this case,  $\bar{q}$  gets the maximum value

$$\bar{q} = q_1 + (\bar{c} - c_1)^{\frac{a-1}{a}} \left\{ \sum_{i=1}^{N-1} \Delta c_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a} \right\}^{-\frac{a-1}{a}} \left\{ \sum_{j=1}^{N-1} \Delta q_j \left( \frac{\Delta c_j}{\Delta q_j} \right)^{-a+1} \right\}. \quad (15)$$

## 6 Discussion

Seeing equations (10) (11) (12),  $x_0$  doesn't appear but only the Pareto coefficient  $a$  appears. Therefore, we have to decide only  $a$  when assuming the probability distribution. It differs in situations where storage used.

Discussing the condition where  $r_i > 0$ , we can know which devices we should select. The condition where  $r_1 > 0$  in the equation (10) is

$$\bar{c} < c_1 + \left(\frac{\Delta c_1}{\Delta q_1}\right)^a \sum_{i=1}^{N-1} \Delta c_i \left(\frac{\Delta c_i}{\Delta q_i}\right)^{-a}. \quad (16)$$

This means that the lowest performance device has no effect when having enough budget. In this case we should remove the lowest performance device and apply equations (10)(11)(12) again.

The condition where  $r_i > 0$  ( $i = 2, 3, \dots, N - 1$ ) in the equation (11) is

$$\frac{\Delta c_{i-1}}{\Delta q_{i-1}} < \frac{\Delta c_i}{\Delta q_i}. \quad (17)$$

This condition means that points  $(c_i, q_i)$  must be on a convex curve in a graph where  $x$  and  $y$  axes are respectively cost and performance. We can not get effective tiered storage if we combine drives that do not meet this condition. The combination of SSD, SAS HDD and SATA HDD might not meet this condition because SSD becomes cheaper.

The equation (12) always satisfies  $r_N > 0$ . Hence, the highest performance device always has effect. When a higher performance device newly appears, we should use it.

## 7 Conclusion

In this report the optimum ratio of devices are derived that maximize a tiered storage performance. It can be applied to designing tiered storage or developing a software that automatically adjusts tiering system configurations.

Additionally the fact is cleared that kinds of devices in tiers require some conditions for effective performance. Some tiered storage systems might take a penalty by unsatisfying the conditions.

## A The Case Provided a Performance Requirement

In the case where a required performance value  $\bar{q}$  ( $q_1 \leq \bar{q} \leq q_N$ ) is provided, we can get the optimum ratio with the same method. It is the ratio  $r_1, r_2, \dots, r_N$  which minimize mean costs  $\bar{c}$  under the constraints (1) (7).

The result is as follows:

$$r_1 = (\bar{q} - q_1)^{\frac{a}{a-1}} \left\{ \sum_{i=1}^{N-1} \Delta q_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a+1} \right\}^{-\frac{a}{a-1}} \left\{ - \left( \frac{\Delta c_1}{\Delta q_1} \right)^{-a} \right\} + 1, \quad (18)$$

$$r_i = (\bar{q} - q_1)^{\frac{a}{a-1}} \left\{ \sum_{i=1}^{N-1} \Delta q_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a+1} \right\}^{-\frac{a}{a-1}} \left\{ \left( \frac{\Delta c_{i-1}}{\Delta q_{i-1}} \right)^{-a} - \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a} \right\}, \quad (19)$$

$$r_N = (\bar{q} - q_1)^{\frac{a}{a-1}} \left\{ \sum_{i=1}^{N-1} \Delta q_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a+1} \right\}^{-\frac{a}{a-1}} \left\{ \left( \frac{\Delta c_{N-1}}{\Delta q_{N-1}} \right)^{-a} \right\}. \quad (20)$$

In this case, the mean cost gets the minimum value

$$\bar{c} = c_1 + (\bar{q} - q_1)^{\frac{a}{a-1}} \left\{ \sum_{i=1}^{N-1} \Delta q_i \left( \frac{\Delta c_i}{\Delta q_i} \right)^{-a+1} \right\}^{-\frac{a}{a-1}} \left\{ \sum_{j=1}^{N-1} \Delta c_j \left( \frac{\Delta c_j}{\Delta q_j} \right)^{-a} \right\}. \quad (21)$$