

Energy-Aware Deduplication in Backup Storage Systems

Yizhou Yan^{1,a}, Wenjun Wu^{2,b}

^{1,2}Beihang University, Beijing, 100191, China

email: {yanyizhou, wwj}@nlsde.buaa.edu.cn

Keywords: Deduplication, Energy Model, Data center, Backup

Abstract. Deduplication technology is widely used to eliminate duplicate data and save resources in storage systems of modern data centers. However, from the energy point of view, deduplication cannot always achieve good energy efficiency for storage systems. In this paper, we propose an energy-aware deduplication strategy for backup tasks to achieve the balance of the process in terms of energy consumption, backup throughput and deduplication ratio. Experiments have been made on different datasets to verify the effectiveness of the new strategy.

Introduction

Many studies demonstrate that there is up to 80% to 90% redundant data in modern backup and archival storage systems [1]. Especially in the data backup tasks, the storage space required can be reduced by 10 to 20 times, or even more through deduplication [2]. Energy consumption for data centers has always been an important issue. According to statistics, the energy consumption of data centers accounted for 1.5% of world energy consumption [3], among which 40% of the energy is from storage system [4]. Whether in a storage system or backup system, deduplication will consume a lot of computing and storage resources, which significantly affects the energy consumption [5], for generating chunk fingerprints and exchanging them between memory and disk drives.

However, most research efforts on data deduplication focus on deduplication efficiency and backup throughput. They don't explore the trade-off in the data deduplication process between the energy consumed for additional computation and the energy saved by lower storage. Therefore, in this paper, we develop an energy model for the data deduplication process and propose novel data backup strategies based on criteria of the throughput and energy consumption.

This paper is organized as follows: Section 2 summarizes the related work in deduplication technique; Section 3 presents the energy model for storage system; Section 4 describes the backup strategy; Section 5 provides the experiments and analysis of results; The last section concludes this paper.

Related Work

Data deduplication technique is widely used in backup and storage systems. There are many previous work focused on enhancing the effect of deduplication, reducing the cost of time and increasing system throughput. However, little work has been done to optimize the energy efficiency of deduplication process.

Many previous researchers have focused on index structure to improve the efficiency of exchange and search. Tianming Yang, et al., developed DEBAR system with distributed fingerprint index structure arranged by binary sequence [6]. In order to overcome the bottleneck of large index look-up, Bhagwat D, et al., proposed Extreme Binning splitting the index by file and select one fingerprint to represent the file in memory [7]. MAD2, described in [8] by Jiansheng Wei, et al., uses HBM (Hash Bucket Matrix), BFA (Bloom Filter Array) and Dual Cache to preserve and take advantage of the data locality. Other researchers are concerned about the impact on deduplication ratio of the data set classification [9][10]. Nohhyun Park, et al., characterized the backup data set by new metrics in a deduplication process and proposed an algorithm to predict deduplication ratio [10]. The algorithm divides the data set into three parts based on the deduplication ratio. From the

statistics upon the size changes of each part in multiple deduplication processes of the data set, it can compute the average rate of change to predict the deduplication ratio of next round.

However, deduplication researchers pay less attention on utilization control and energy consumption. Costal, et al [11] presents a simple energy model on data deduplication to enable administrators to reason about the energy and performance break-even points when configuring a storage system. Compared to the previous work, the major contribution of our work is to provide a more realistic energy model and develop a run-time backup strategy for dynamically assessing the trade-off between energy-saving and throughput of data deduplication and decide whether to invoke deduplication in backup tasks. Our method enables system administrators to configure their backup systems based on characteristics of the dataset in backup tasks.

Energy Model for Storage System. Storage systems typically consist of multiple parts, including control nodes, disk expansion cabinet (hard drives) and backup power. An EMC storage system (VNX5500), for example, includes the following parts, storage processor, control station, NAS controller and DAEs (disk array enclosures).

As reported by D.Floyer in [13], the main part of energy consumption of the storage system pertaining to its usage comes from its Control Station and DAEs (disks). For server (Control Station) energy part, our previous study [5] proposed an energy model based on system usage in Eq(1). We got the parameters (a, b, c) through Lasso regression with more than 2000 sets of energy consumption data in the form of (Utilization_{CPU}, Utilization_{MEM}, Power). The power consumption of an active disk can be defined as three states, idle (spin-down) power, standby power and the power on max workload. They can be represented as (P_{idle} , $P_{standby}$, P_{maxIO}).

$$P_{server} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \begin{Bmatrix} Utilization_{cpu} Utilization_{mem} \\ Utilization_{cpu}^2 Utilization_{mem}^2 \\ Utilization_{cpu}^3 Utilization_{mem}^3 \end{Bmatrix} + c(1)$$

Energy consumption of the storage system is mainly from the several components mentioned above and can be divided into two parts: workload independent part and workload relevant part that contributed by control nodes and disks.

$$P_{Storage\ System} = P_{System\ Base} + \sum_{nodes} P_{server} + \sum_{disks} P_{disk}(2)$$

In Eq(2), $P_{System\ Base}$ represents the workload independent part that can be measured directly. P_{server} can be computed by corresponding energy model in Eq(1). P_{disk} can be measured in corresponding state among (P_{idle} , $P_{standby}$, P_{maxIO}).

Intelligent Deduplication based Backup Strategies

Deduplication efficacy in a backup process heavily depends upon its workload. It is necessary to dynamically estimate the cost of deduplication operation and make balance between cost and benefit incurred by deduplication. Based on the two major metrics including throughput and energy efficiency, the paper develops two algorithms to determine whether to run the deduplication process: Throughput oriented backup Strategy (TS) and Energy aware backup Strategy (ES).

Each backup task often includes specific data set, adaptive strategy and independent metadata. A typical backup scene described in [12] is that full backups are done on weekends while incremental backups are performed from Monday to Friday. Most backup tasks are cyclical, which means that the same data set is supposed to be backed up for multiple times.

As backup tasks are periodic, statistical regularity can be discovered in deduplication result. The research in [10] shows that the composition of data set in next round of deduplication can be predicted based on the previous result. Since deduplication operation is very time consuming and resource-intensive, to obtain significant effect, we need to filter out the effective data set from being deduplicated.

A. Throughput oriented Backup Strategy

We set up a “deduplication eliminate throughput”(DET) to measure the efficiency of deduplication in the backup system. It can be calculated by duplicate data eliminated and time cost as shown in Eq(3). DET is adopted to evaluate deduplication effect in backup task and help to figure out the data sets that are unsuitable for deduplication.

$$DET = \frac{\text{eliminated duplicate data (ED)}}{\text{time cost (T)}}(3)$$

$$ED = \text{Logical size} - \text{Physical size}(4)$$

To assess a dataset with DET, we need to take backup throughput and time window into account. We define a threshold that is equal to the bandwidth of the storage system. To get a stable and accurate prediction of DET for a backup task, we need to run deduplication operation for N times. After that, we can compare the DET with the threshold and make the decision whether it should be deduplicated or not.

Based on this idea, we can define TS as follows:

- (1) Every task is initialized with the deduplication process. Every time when the task is accomplished, the algorithm counts the deduplication ratio and processing duration.
- (2) After the task runs deduplication for at least N times, the algorithm can start predicting the deduplication efficiency of the next round based on the accumulated historical data. Empirically, we set N to 3 in our experiments.
- (3) DET value can be calculated from Eq(3). When the predicted DET value of the next deduplication in the task becomes more than the threshold R, the algorithm decides to suspend the deduplication.

Basically, the value of R is defined as the same as the IO speed of the storage system to ensure the benefit of deduplication goes beyond the IO capacity of the storage. Therefore, when $DET > R$, the deduplication can definitely result in a better performance in throughput than a simple backup process without deduplication.

B. Energy-aware Backup Strategy

The TS takes throughput optimization into account. But it doesn't consider energy saving for deduplication. The influence of the deduplication process on energy consumption of the backup storage system is actually double-fold: On one hand, the computationally intensive operations in the deduplication process including file chunking, hash computing and index lookup lead to the increase in energy consumption. On the other hand, the deduplication can reduce IO access and storage utilization, which contributes to the decrease in energy consumption. Therefore, to achieve the high power efficiency of backup storage systems, it is necessary to filter out backup tasks or data sets that generate higher energy consumption.

The increase in energy consumption caused by deduplication is defined as $P_{\text{incremental}}$ in Eq (5). It occurs on the control servers that run bit-wise deduplication operations including chunking, hashing and recovery during deduplication.

$$P_{\text{incremental}} = \sum_{\text{nodes}} ((P_{\text{dedup}} - P_{\text{standby}}) * T_{\text{dedup}})(5)$$

P_{dedup} and P_{standby} can be computed by Eq(3.1). The only thing we need to measure is the resource utilization during deduplication process. T_{dedup} is the time cost of deduplication process. It can be got in the first three backup rounds.

The decrease in energy consumption caused by deduplication is defined as P_{reduced} in Eq(6). Due to the reduction in storage space after deduplication, more disks are freed and can be put in a power saving state of standby, spin-down or off state, thus reducing the overall energy consumption of the backup system.

$$P_{\text{reduced}} = (P_{\text{maxIO}} - P_{\text{standby}}) * T_{\text{iosave}} * N_{\text{active}} + (P_{\text{standby}} - P_{\text{idle}}) * T_{\text{backup}} * \frac{S_{\text{dup}}}{S_{\text{energy_unit}}}(6)$$

The first half of Eq(6) represents the energy consumption of disks occupied by duplicate data during the regular time for IO access. T_{iosave} represents the time overhead for transferring the data and N_{active} denotes the number of working disks. The second half represents the power consumption caused by storing repetition of data within the backup cycle. The idle hard drives can be turned from standby state to spin-down or off state for the purpose of power saving. T_{backup} represents the time

period of backup tasks. S_{dup} represents the volume of duplicate data that have been deleted and S_{energy_unit} is for the entire disk volume.

Based on the energy prediction model defined in Eq(5, 6), we can propose the Energy-aware backup Strategy (ES) for data set filtering. The first two steps of the algorithm are defined the same as in TS to infer the inherent data set redundancy. At the third step, the algorithm can calculate $P_{incremental}$ and $P_{reduced}$ using Eq(5) and Eq(6). If $P_{incremental} > P_{reduced}$, the algorithm determines that this data set is energy inefficient and cancel the deduplication during the following backup rounds.

Evaluation

Our main test environment is a cluster with eight Xeon E5620 servers, each of which has a 16-core CPU at 2.4 GHz and 24GB RAM, running CentOS 6.3. Four of them run as the backup server. The cluster is connected to an EMC VNX5500 storage system consisting of 60 SAS disks with 300GB volume and 1.5k RPM speed. We implemented the de-duplication module and integrate it into an open-source backup software “UrBackup”[15] to set up an experimental backup environment. The module runs as a middleware on the server side.

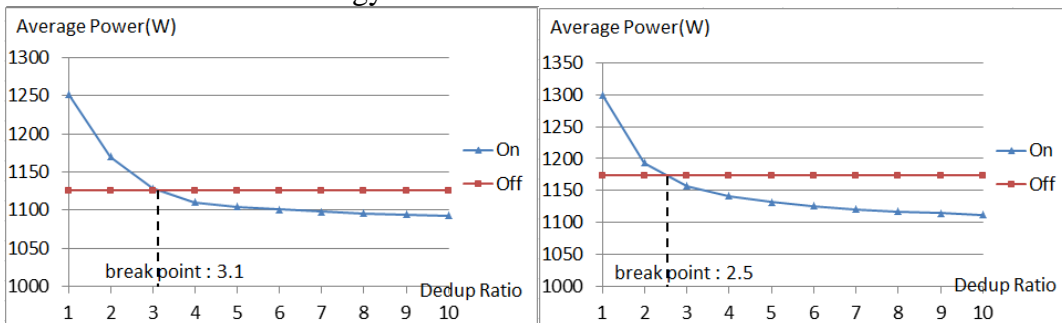
Table 1. Backup Task Dataset Description

No.	Type	Logical Size(MB)	backup rule	data source
1	txt/doc/pdf, etc	2,238,184	Daily differential & Weekly full backup	From students, working sets on the course of Hadoop.
2	iso, mkv/avi, jpg/png, etc	2,229,510	Weekly incremental backup	From the video conference system used by our lab.
3	zip/tar, exe/bin, rpm, etc	1,829,383	Weekly incremental backup	From archived data of FTP in our lab.
4	iso, txt/doc/pdf, zip/tar, etc	1,840,271	Daily incremental & Weekly full backup	From escience center services

A. Model Evaluation

The first group of experiment is designed to demonstrate how deduplication operation affects the energy consumption of the storage system when it processes datasets with different data similarities. To this end, we executed the backup process on 1TB and 2TB data with the level of data similarity ranging from 1 to 10.

Both Fig. 1(a) and Fig.1(b) plots two curves of energy consumption with deduplication being turned either on or off. On each figure, there is a break-even point where the energy consumption of the backup process with deduplication equals to that of the backup process without deduplication. This break-even point suggests the exact moment when the benefit of deduplication outweighs the energy cost caused by deduplication. It is noted that the break-even point in Fig. 1(a) occurs at the deduplication ratio of 3.1 whereas the point in Fig. 1(b) occurs at the ratio of 2.5. The disparity can be explained by the effect of deduplication. Because if more redundant data are eliminated, more idle disks can be shut down to save energy.



(a) Average Power with 1TB data (b) Average Power with 2TB data
Fig. 1. The Average Power of VNX5500 under Different Deduplication Ratio

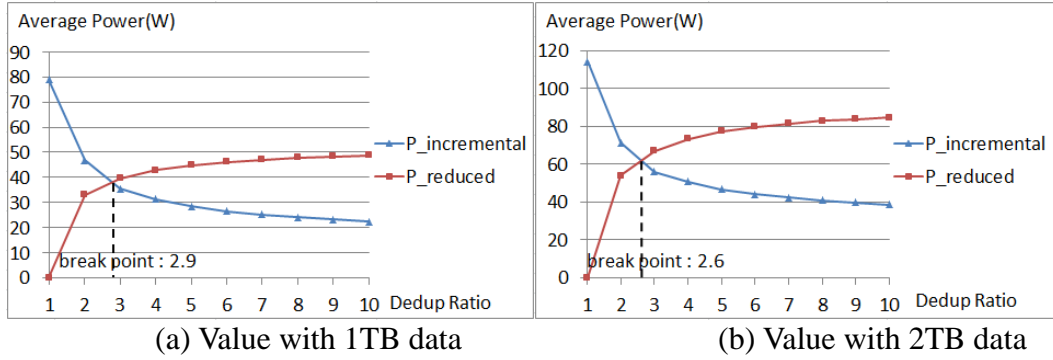


Fig. 2. The Value of $P_{incremental}$ and $P_{reduced}$ under Different Deduplication Ratio

The break-even point is very important for the energy-aware backup strategy to appropriately decide whether to run deduplication operation. Fig. 2(a) and 2(b) present the curves of $P_{incremental}$ and $P_{reduced}$ that are calculated through Eq(5) and Eq(6) with the varying level of deduplication ratio. The predicated break-even point also occurs at the similar spot as the experimental measurement in Fig. 1. The result indicates that the model proposed above is accurate for the energy-aware backup storage to make the decision whether the data set need to be deduplicated or not.

B. Throughput Evaluation

We set up four backup scenarios whose data sets are described in the Table 1. We carried out backup tests for three months according to the rules in Table 1. During these backup tests, three different strategies including all deduplication strategy (all-dedup), TS and ES were separately enforced to the data sets.

Fig.3 shows the comparison between the logical size of the data set and physical size. We calculate the cumulative logical and physical size of the data sets every week. Four data sets have similar size but different deduplication performance.

The throughput and deduplication ratio of different strategies for the backup process is shown in Fig.4(a). Under TS, the second, third and fourth data set didn't go through deduplication after three weeks because their DET values are below the threshold. While the fourth data set will be deduplicated using ES. From Fig.4(a), we can find out that both strategies achieved a higher throughput than all-dedup strategy. In addition, the backup system achieved the highest throughput under TS. The results turns out that for the data sets that have a DET value lower than IO throughput, cancelling the deduplication can get a throughput higher than simple backup.

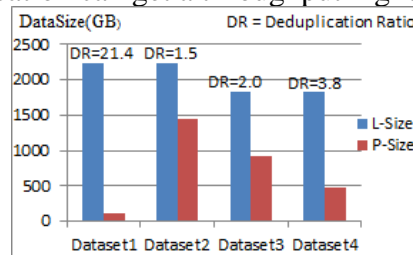
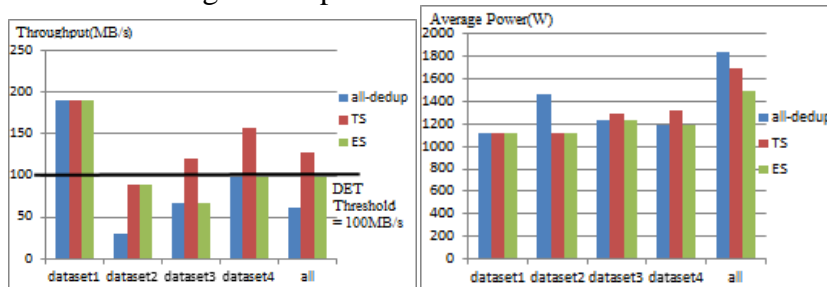


Fig.3. Deduplication ratio of four data sets



(a) Throughput under different strategies (b) Power under different strategies
Fig.4. Throughput and Average Power of data sets under different strategies

C. Energy Evaluation

In order to measure the effect of the number of disks used, we observed the energy consumption of storage system and servers with different strategies. Fig.4(b) presents the average power for the backup process in our experiment. ES gains the result of lowest energy consumption by avoiding potential energy consumption from data-intensive deduplication.

TS reduced the power of the whole system by about 150W, 8% of all. While ES saved about 350W due to more disks shut down. The data scale of our experiment is not large. And the space saved by deduplication cannot free bigger devices like DAEs. If so, the energy saving will be more significant.

When backup activities are limited in a narrow time window, TS can be used to satisfy the requirement of time. And when energy saving is the first priority, ES can make a better balance between throughput and energy consumption.

Conclusion

This paper proposes an energy model for describing energy consumption of storage systems and evaluating the power cost incurred by deduplication of backup tasks. Based on the energy model, we introduce two novel backup strategies that can dynamically estimate the throughput and the energy saving benefit of deduplication in the backup process to make effective decision on whether to launch deduplication operations. Extensive experiments have been done to verify the performance of these strategies. System administrators can utilize the strategies to optimize the usage of deduplication in terms of both throughput and energy saving.

Acknowledgement

This work was supported by National Important Research Plan of Cloud Storage System Development Grant No.2013AA01A210 and the State Key Laboratory of Software Development Environment.

References

- [1] McKnight J, Asaro T, Babineau B. Digital Archiving: End-User Survey and Market Forecast 2006–2010[J]. The Enterprise Strategy Group, 2006.
- [2] Biggar H. Experiencing data de-duplication: Improving efficiency and reducing capacity requirements[J]. The Enterprise Strategy Group, 2007.
- [3] Koomey J. Growth in data center electricity use 2005 to 2010[J]. Oakland, CA: Analytics Press. August, 2011, 1: 2010.
- [4] Schulz G. Storage industry trends and it infrastructure resource management (irm), 2007[J].
- [5] Luo L, Wu W, Tsai W T, et al.. Simulation of power consumption of cloud data centers[J]. Simulation Modelling Practice and Theory, 2013, 39: 152-171.
- [6] Yang T, Jiang H, Feng D, et al.. DEBAR: A scalable high-performance de-duplication storage system for backup and archiving[C]//Parallel & Distributed Processing (IPDPS), 2010 IEEE International Symposium on. IEEE, 2010: 1-12.
- [7] Bhagwat D, Eshghi K, Long D D E, et al.. Extreme binning: Scalable, parallel deduplication for chunk-based file backup[C]//Modeling, Analysis & Simulation of Computer and Telecommunication Systems, 2009. MASCOTS'09. IEEE International Symposium on. IEEE, 2009: 1-9.
- [8] Wei J, Jiang H, Zhou K, et al.. MAD2: A scalable high-throughput exact deduplication approach for network backup services[C]//Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on. IEEE, 2010: 1-14.
- [9] Zhu G, Zhang X, Wang L, et al. An intelligent data de-duplication based backup system[C]//Network-Based Information Systems (NBIS), 2012 15th International Conference on. IEEE, 2012: 771-776.
- [10] Park N, Lilja D J. Characterizing datasets for data deduplication in backup

- applications[C]//Workload Characterization (IISWC), 2010 IEEE International Symposium on. IEEE, 2010: 1-10.
- [11]Costa L B, Al-Kiswany S, Lopes R V, et al. Assessing data deduplication trade-offs from an energy and performance perspective[C]//Green Computing Conference and Workshops (IGCC), 2011 International. IEEE, 2011: 1-6.
- [12]Wallace G, Douglis F, Qian H, et al. Characteristics of backup workloads in production systems[C]//FAST. 2012: 4.
- [13]David Floyer. Green Validation Report: 3PAR InServ Storage Arrays.[EB/OL]http://wikibon.org/wiki/v/Green_Validation_Report:_3PAR_InServ_Storage_Arrays
- [14]Murugan M, Kant K, Raghavan A, et al. Group based energy adaptation for storage systems[C]//Proceedings of the 9th International Workshop on Feedback Computing (USENIX). 2014.
- [15]UrBackup.[EB/OL] <http://www.urbackup.org/>.