A New Approach to Storage Management Restrictions Using the "Data Quality" Concept

Research and Implementation

Koby Biller

DISKLACE Ltd., Rehovot, Israel

koby@disklace.com

Abstract- **Storage management is the control of the capacity, policy and events in storage devices. Its purpose is to achieve maximum efficiency when allocating or deleting files, while always retaining an accurate view of the repository. Storage management was developed along with the development of storage devices to solve new problems that might be encountered with new storage technologies and software changes. One technique developed to increase storage availability even when storage utilization has increased is "Disk Fragmentation". Fragmentation was invented to split file allocation within the storage device in such a way that there would always be free space for new files if the device capacity was not fully used. The fragmentation concept works perfectly with a new device, but over time the fragments of free space become very small and every file is split into smaller fragments. The effort to access an existing file is added to the effort of finding available fragments, and the result is a delay in the service. Excessive fragmentation of data in storage devices carries with it a restriction which represents a major problem underlying computer slowdown and other unpredictable storage-related symptoms.**

This paper presents the "Data Quality" approach, which measures the above restriction on the basis of research performed during the period 2005 to 2008, in which we surveyed the effectiveness of ourpatented method for measuring the quality of data affected by disk fragmentation. The approach incorporates a novel way to manage fragmented data using file system statistics. We show clear advantages over the current alternatives such as replacing the device or "hard-wiping" it. Furthermore, we present a unified scale that allows comparison of data quality on an organization-wide level.

Keywords- Disk Fragmentation; Preventive Maintenance; Data Disorder Measurement; Defragmentation; Storage Management

I. INTRODUCTION

Storage management [1] incorporates a set of programs and management tools used to manage the capacity, policy and events in storage devices. Its purpose is to achieve maximum flexibility during normal use of any file system. The introduction of any new storage device is accompanied by improvements in storage management procedures to dynamically reflect the new functions. Disk fragmentation was invented to address the need for on-going availability of free space for new files until the storage device is exhausted. The fragmentation concept works perfectly when the device is empty. After several file deletions and reallocations the fragmentation level increases, and over time the fragments of free space become progressively smaller as every new file splits into smaller fragments. Software efforts to access an existing file are added to the effort of finding its fragments, and the result is a delay in the service. Excessive fragmentation of data in storage devices thus becomes a restriction that represents a major problem underlying computer slowdown and other unpredictable storage-related symptoms.

Computer disk fragmentation [2] is a phenomenon that has attracted considerable research attention [3-5]. Disk fragmentation occurs in hard drives as a user saves, changes, or deletes files. Over time, both the files and the storage device become fragmented. As a consequence, the computer slows down as it has to look in many different places to open or create a file. This fragmentation takes place "under the hood", and is an inevitable feature of the normal lifecycle of data on disks.

Currently, users benefit from almost infinite usage of disk resources. It is commonly known, however, that fragmentation leads to unpredictable and degraded application performance [6] without any indication or warning from the operating system and regardless of the amount of storage space. In this work, we show the futility of attempting to avoid this problem by manual maintenance *after* these symptoms appear.

We have investigated the phenomenon of disk fragmentation and introduced a comprehensive solution that improves data lifecycles in both current and future storage devices. This is achieved by employing a novel technique that measures data quality and implements a concept of optimized preventive maintenance. The result is a prolonged computer lifespan, easier maintenance, and improved user satisfaction. After undergoing a thorough test by the industry (300,000 users), this Data Quality model has already been adopted by some customers including a large consumer products company and is completing its 4th year in a laboratory at a well-known research institute.

II. MEASUREMENT METHOD

The formula (Eqn. 1), used for measurement of data quality, is part of a granted patent [7] and is based on a statistical second-level moment equation combining fragmentation of the files with free space on the storage device. The result is presented on a logarithmic scale, in which a lower number reflects a better overall data quality. Units on the scale, termed "Lace Level", are normalized and are independent of the size of the storage device.

$$\text{Lace Level} = \log\left(\frac{\left(\sum_{k=1}^{M} f_k\right)^2}{\sum_{k=1}^{M} f_k^2}\prod_{j=1}^{N}\frac{\left(\sum_{i=1}^{n_j} c_{ij}\right)^2}{\sum_{i=1}^{n_j} c_{ij}^2}\right)$$

Equation 1 Calculation of fragmentation levelM is the number of free space extents (size is variable), f_k is the size of the k^{th} extent of free space, N is the number of files on the storage device. n_j and c_{ij} are, respectively,the numbers of extents taken by file j and the size of the i^{th} extent belonging to it

The value of the measured Lace Level may vary from 1 (for a perfect disk having all the files in order) to 15 (the highest value measured until now).

III. DISTRIBUTION OF RESULTS OF THE RESEARCH

To study the global picture of the fragmentation phenomena, and to assess the viability of our measurement scale, we released a free software solution to measure the Lace Level. This software was downloaded more than 300,000 times, and about 100,000 users voluntarily and anonymously reported back with their results. The distribution of results presented here (Fig. 1) has a distinct hat shape with a large peak. Given the large sample size, we consider that these data constitute a reliable sample, and that it hints at the underlying structure of the global behaviour of file systems. The *inflection point* of the graph is further discussed below.

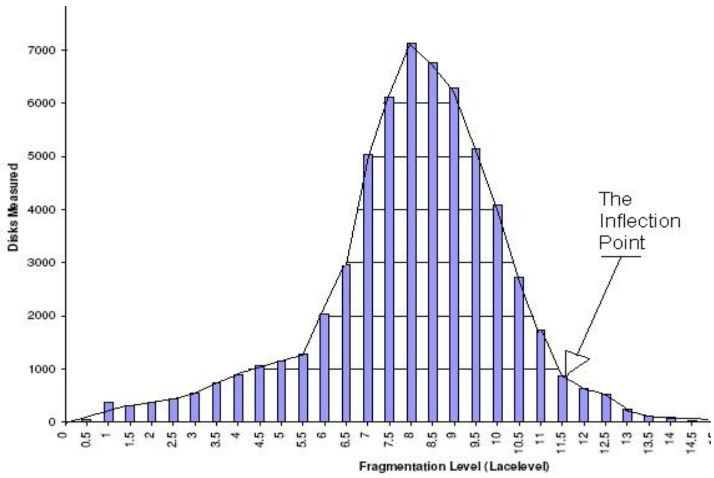

Fig. 1 Distribution of Lace Level measurement in100,000 disks

IV. EFFECTIVENESS OF DEFRAGMENTATION

To analyze the relative success of maintenance operations on disks (i.e., defragmentation), we chose to compare Lace Level measurements taken before and after maintenance. For this, we isolated 30,000 hard drives from our samples for which the before/after measurements for the same disk could be coupled (by their unique network address and disk size). Our results, presented in Fig. 3, revealed that maintenance of a storage device with a high level of fragmentation yields poor results. The average decline in defragmentation effectiveness measured in our research appears to be beyond Lace Levels of 12. Based on this observation, we deduce that the inflection point in the global fragmentation distribution (Fig. 1) can be understood as a "point of no return" in disk maintenance. Since the global distribution of fragmentation could very well be multimodal the actual inflection can shift, depending on the usage pattern of storage in the device (e.g. server, workstation, cloud device, mobile phone, etc.).

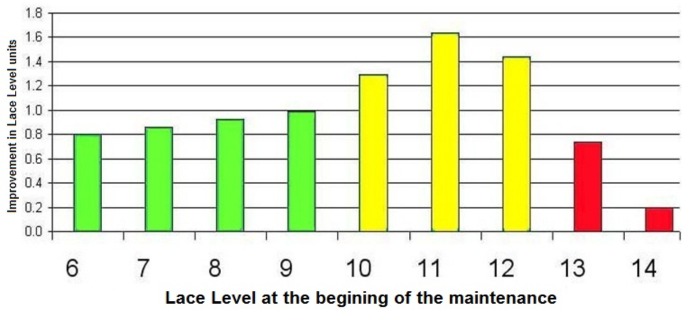

Fig. 2 Maintenance results according to level at the starting points

V. IMPLEMENTATION

The ability to uniformly quantify the data quality of storage devices, combined with our indicator for performance of preventive measures, offers a new route of optimized preventive maintenance. Essentially, each storage device needs to be maintained before its data quality degrades beyond a certain threshold, which should be calibrated on the basis of its purpose. Thus, the responsibility for computer maintenance should no longer be in the hands of the end users, but should rather become a part of the central management.

To test this new philosophy, we conducted a long-term on-site test (1 year) on two hard drives on the same server using a mirror update (which means that the data on those volumes are identical).One disk was automatically maintained by the product, while the other disk reported only the fragmentation level. The results of this test are presented in Fig. 3. The green curve shows the Lace Level (and therefore the data quality) of a drive maintained with the optimized preventive scheme. When the fragmentation level reached the global inflection point described earlier, a simple defragmentation was automatically performed. As the fragmentation did not stop increasing, repeated maintenance was performed, resulting in a saw-tooth-like behavior. The red curve represents a drive that was not managed according to our method. Over time its data quality became severely degraded and even defragmentation did not help, as predicted by our before/after study (compare with Fig. 2). The only available means of remedying this disk was by reformatting it and starting anew.

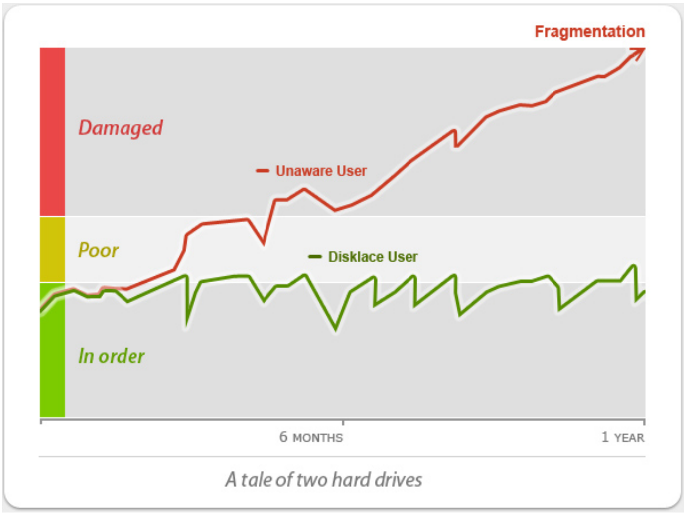

Fig. 3 Maintenance results according to level at the starting points

With the aim of offering such dynamic maintenance regimes in large-scale operations, we implemented the approach described in this paper in our *Lace Watcher PRO* product [5]. The implementation relies on the collection of Lace Level data from all of the computers in a network (automatic topology study), and not only from volunteers as described above.

The initial threshold value is set to the inflection point found in this research, but it can be overridden by the administrator depending on the specific finding in the particular network. Once a Lace Level exceeds the threshold, an alert is sent or an automatic defragmentation is performed.

The difference between the theoretical finding of this research and its practical implementation in the field is displayed upon request (Fig. 4), and users can always compare the global results to the nature of the fragmentation in their systems.

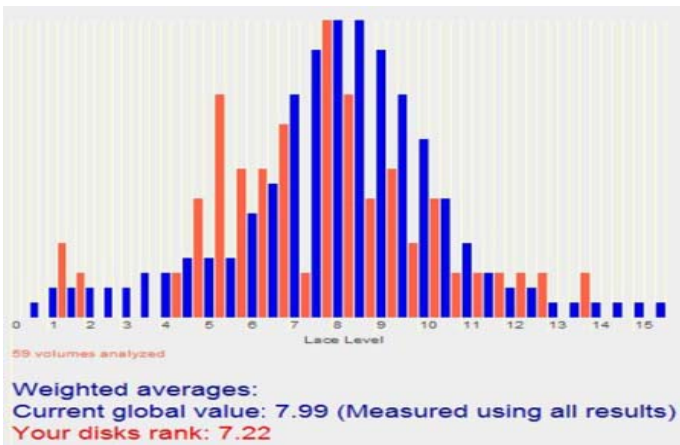

Fig. 4 Example showing a comparison of a specific customer (RED) with the global fragmentation distribution (BLUE)

VI. FUTURE PERSPECTIVES

Since the survey for this research ended, new technologies have been developed in the storage world. These new technologies suffer from the same problem as that identified in this research. One example is seen on storage devices with virtual disks that replace the directly attached mechanical disks. The technology provides a better utilization of storage devices, but the disks used in this method are of the same type and therefore suffer from the same fragmentation problem. The increasing usage of SSD (solid state devices) in portable computers and tablets represents a more interesting problem, since the disk management software used for these devices remains the same, but the storage technology is much more sensitive to excessive writes. Therefore manufacturers ask users to refrain from utilizing periodical defragmentation [9]. Thus, optimized preventive maintenance, as proposed here, appears to be the most beneficial course of action.

Mobile devices have large storage capacities. These devices hold sensitive data but their fragmentation level is not maintained. Every storage device that allows deletion of data and uses the released space to write new data will eventually suffer from a fragmentation problem that will have to be managed. The overall capacity of storage capabilities on cellular devices and tablets will shortly reach the point at which performance will decrease. When this happens, a solution can be provided by the Data Quality approach described in this paper.

VII. CONCLUSIONS

We have presented a novel Data Quality approach to the problem of file system fragmentation. A new measurement scale was used to analyse data from ca. 100,000 samples. Global data properties were identified and a regime of optimized preventive maintenance was developed. Based on our results we released a lightweight management and control software solution, which was shown to provide effective defragmentation schedules that take into account the unique properties of the storage device.

ACKNOWLEDGMENT

Our grateful thanks go to all of the 300,000 users who downloaded the test program, and in particular to the 100,000 among them who shared their results with us. There was no way that we could have completed this research without their active support.

REFERENCES

[1] ROWSTRON, A.I.T. Storage management and caching in PAST, a large-scale, persistent peer-to-peer storage utility. SOSP Proceedings of the 18th ACM Symposium on Operating Systems Principles (2001), pp. 188–201.

[2] RANDALL, B. A Note on Storage Fragmentation and Program Segmentation*IBM Thomas J. Watson Research Center Yorktown Heights, New York* (1959) 12(7): 365-ff.

[3] BHADKAMKAR, M., FARFAN, F., HRISTIDIS. V., AND RANGASWAMI R. Storing semi-structured data on disk drives. ACM Transactions on Storage (2009) 5(2).

[4] TABERO, J., SEPTIEN, J., MECHA,H., AND MOZOS, D. Allocation heuristics and defragmentation measures for reconfigurable systems management. Integration, VLSI Journal (2008) 41(2): 281–296.

[5] SHORE, J.E. On the external storage fragmentation produced by first-fit and best-fit allocation strategies. Communications of the ACM (1975) 18(8): 433–440.

[6] SEARS R. AND VAN INGEN, C. Fragmentation in large object repositories. Cornell University Library arXiv:cs/0612111 [cs.DB].

[7] US PATENT 8.051, 115, B2: Measuring fragmentation on direct access storage devices and defragmentation thereof.

[8] A full functioning demo of Lace watcher-PRO for windows supporting 1TB storage and partitions up to 250GB is available at http://www.disklace.com.

[9] BETTS, B. Solid state of the art. Journals & Magazines Engineering & Technology, (2012) 7(9): 84–87.

Koby Biller is the founder of DISKLACE Ltd. He was born and raised in Jerusalem, ISR, and studied at the Hebrew University of Jerusalem. In 1973 he joined IBM ISRAEL, starting as a program analyst, and spent 27 years working there as a systems analyst and systems engineer. During that time he developed novel working methods and productivity tools for database and data communication systems. In 1982 he played a key role in a special infrastructure project at IBM headquarters in Paris, in which the company's distributed offices were transferred to a central location. He was among the first to study the IBM 370 operating systems. As a systems programmer he implemented new technologies and developed supporting packages in the fields of databases, data communications, data protection and automated operations and provided technical support services for the IT development team. He has won the IBM Professional Excellence award several times and the IBM 100% Club Membership for marketing achievement twice.

The method and implementation described in this paper were submitted to international patent offices, and a final patent for the concept was granted by the US patent office.