Title

Single-Snapshot File System Analysis

ISBN

9780769551029

Authors

Wildani, Avani
Adams, Ian F
Miller, Ethan L

Publication Date

2013-08-01

DOI

10.1109/mascots.2013.47

Peer reviewed

Single-Snapshot File System Analysis

Avani Wildani
Storage Systems Research Center
University of California, Santa Cruz
Santa Cruz, CA, USA
avani@cs.ucsc.edu

Ian F. Adams
Storage Systems Research Center
University of California, Santa Cruz
Santa Cruz, CA, USA
iadams@cs.ucsc.edu

Ethan L. Miller
Storage Systems Research Center
University of California, Santa Cruz
Santa Cruz, CA, USA
elm@cs.ucsc.edu

***Abstract*—Metadata snapshots are a common method for gaining insight into filesystems due to their small size and relative ease of acquisition. Since they are static, most researchers have used them for relatively simple analyses such as file size distributions and age of files.**

We hypothesize that it is possible to gain much richer insights into file system and user behavior by clustering features in metadata snapshots and comparing the entropy within clusters to the entropy within natural partitions such as directory hierarchies. We discuss several different methods for gaining deeper insights into metadata snapshots, and show a small proof of concept using data from Los Alamos National Laboratories. In our initial work, we see evidence that it is possible to identify user locality information, traditionally the purview of dynamic traces, using a single static snapshot.

I. Introduction

Metadata snapshots are a simple way to gain understanding of the structure and contents of a filesystem. Since these snapshots are typically much smaller than dynamic traces, static metadata is relatively easy to collect and store. This size advantage, along with the lower performance overhead for collecting static snapshots, makes them relatively easy to obtain for analysis. While there have been numerous studies [2, 6] that attempt to reconstruct dynamic trace information from a series of snapshots by interpolating the inter-snapshot accesses, we focus instead on what can be learned about a system by looking at metadata correlations within a single snapshot. Our initial work indicates that clustering metadata within a single snapshot may provide valuable insight into a storage workload.

With the basic POSIX-like metadata produced from a `stat` (or similar command) of the files in the system, one can glean a variety of useful statistics of interest to researchers and administrators, such as file size distributions and namespace layout. With multiple snapshots taken over time, it is even possible to see how file systems evolve [2] or to calculate the inter-reference intervals between files [6].

While there have been many useful studies that analyzed metadata snapshots, most have focused on simple statistics, such as file size, age, or extension. Even with basic metadata, we hypothesize that there is much more insight that can be gained. For example, timestamps can give insight into the dynamic activity of the system from a purely static viewpoint. UIDs can be used in conjunction with file paths to figure out if there is a "typical" namespace structure users create. Entropy between members of a namespace can help us relate different segments of a trace.

In this work, we apply techniques from unsupervised learning and information theory to learn correlations within static metadata traces. As an example, consider trying to identify the access locality within a file system. If files track their last modification time with reasonable accuracy, we can take a simple two step process to start learning about their modification locality. First, we group files by similar modification times, using a density based technique such as DBSCAN [5]. We can then analyze each of these clusters along a variety of dimensions. It can be as simple as comparing the number of unique user IDs within each cluster to see if UID is a good predictor of modification locality, to more in depth techniques such as agglomerative clustering to examine the namespace locality within files modified at a similar time. In the remainder of this paper, we discuss the motivations behind single-snapshot analysis and present a series of views, which are projections of datasets into three dimensions, to analyze what clustering algorithms are most likely to predict user locality in both archival and HPC data sets.

II. Background

We seek to characterize workloads based on single snapshots. A number of recent studies have used dynamic traces of storage systems to identify working sets [4, 8]. Identifying working sets accurately and reliably can greatly improve both the performance and the efficiency of storage systems [3]. Additionally, understanding workload characteristics is essential for optimal storage management and provisioning [1].

Keeping complete logs of accesses is prohibitive in many systems, however, because of the computational

overhead to collect the logs and the storage overhead to keep them. For a modern storage system with hundreds of thousands of I/Os per second, storing even minimal representations of the I/O without any metadata is very costly. For example, an enterprise storage system creates over 16 GB of block-level I/O logs per day [8].

Storing metadata is even harder than storing raw accesses because there is more overhead both in terms of size and performance. As a result, metadata is almost exclusively stored as snapshots – static read-outs of stored data elements along with metadata such as `atime`, `ctime`, and `path` – and most analyses attempt to interpolate dynamic traces from these snapshots.

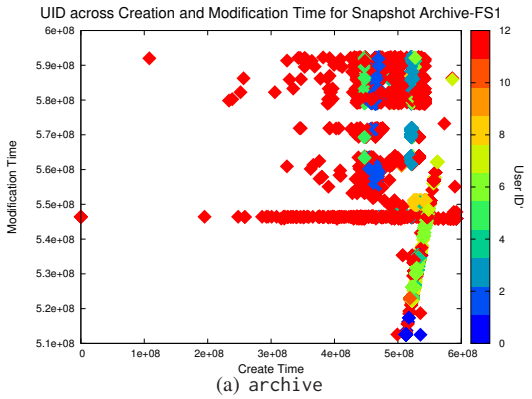
For example, Gibson and Miller [6] calculate interreference intervals from daily snapshots to obtain long term trends. While their work contained valuable insights, such as file usage over time, they require dynamic data and remark that they met resistance from system administrators when requesting even a daily trace. Similarly, Agrawal *et al.* [2] performed a long-term study of file system metadata using annual metadata snapshots from thousands of enterprise desktops. They were able to watch the changes both of their users and the file system, utilizing a very large dataset. Our work to identify trends within snapshots augments this type of work; for example, clusters could be tracked across multiple snapshots to track how the apparent usage patterns change over time.

One area that has been analyzing single snapshots of systems is computer forensics. Often, the goal in a forensics environment is to identify particular files or users that are anomalous compared to the rest of the snapshot. Many of the techniques these researchers use apply to our problem. For instance, Rowe and Garfinkel [7] point out that files that have close proximity in creation or modification times can have causal relationships, and they also look for co-occurrence of files that are duplicated in a snapshot.

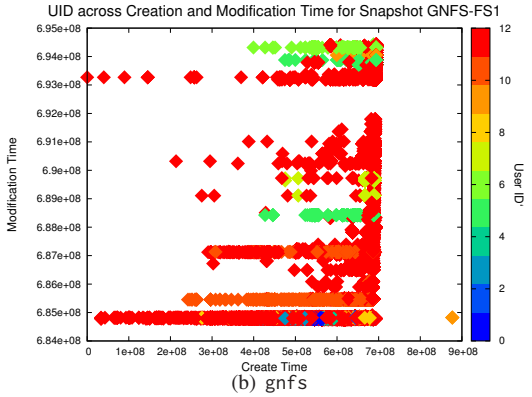
III. Experiment Design

Metadata snapshots present us with a highly nonlinear, multidimensional data set. To meaningfully compare snapshots and discuss correlations within them, we introduce the concept of a *view*. A view is simply a projection of the snapshot space onto three dimensions. Views allow us to visually isolate different aspects of user locality, such as when users modify files they created, and help define what clustering algorithms are likely to perform well on each snapshot.

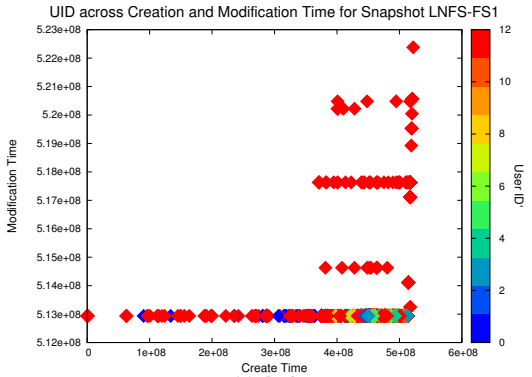
We examined a series of views using three anonymized snapshots from Los Alamos National Laboratory (LANL); Table I contains details. From these



(a) `archive`



(b) `gnfs`



(c) `lnfs`

Fig. 1. Create time vs Modification Time views of LANL snapshots with the top 10 UIDs identified by color. The concentrations of accesses by the same UID represent the type of patterns unsupervised learning should find.

TABLE I. LANL Snapshot Statistics

	Type	# Records	# Machines
`archive`	Archive	112020366	13
`gnfs`	Global NFS	6437081	13
`lnfs`	Local NFS	306340	2

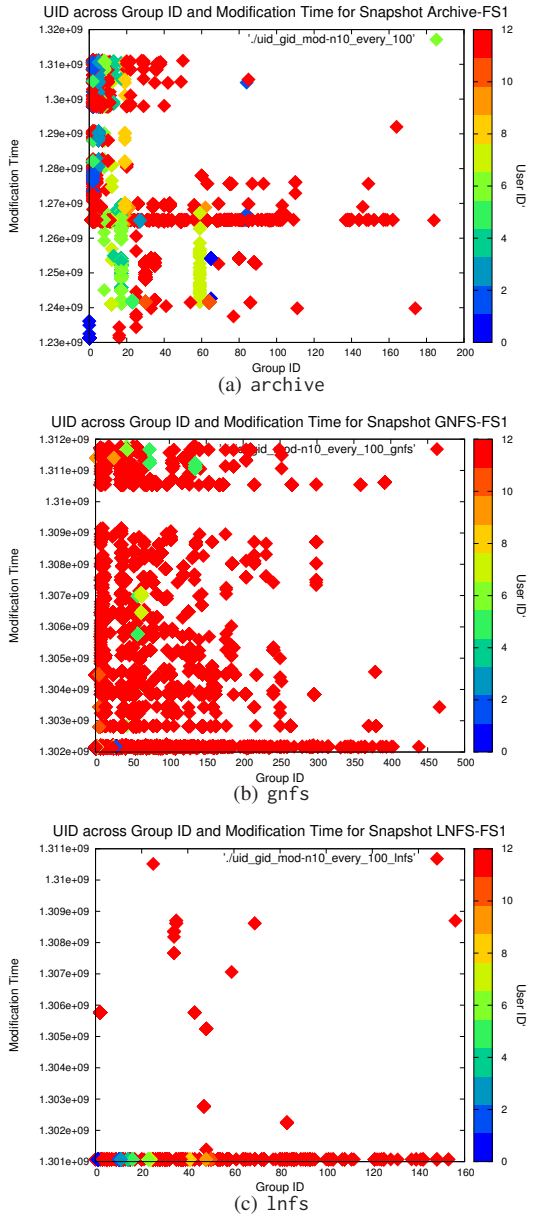


Fig. 2. Group ID vs Modification Time views of LANL snapshots with the top 10 UIDs identified by color

snapshots, we focus on the fields "create_time," "modification_time," "UID," "group_id," and "file_id" to explore user locality.

A. Results

Figure 1 shows the ten most popular UIDs (relabeled as 1-10), represented by color, graphed against file creation and modification times within our three snapshots. Remaining UIDs were placed into a separate group, "Other", represented by UID 11. Since we are at an early stage in our work, we have focused on views of our data sets that provide the most evidence for interesting user locality behaviors. Each graph represents an entire snapshot, though for clarity we subsample by only plotting every hundredth snapshot entry. The shapes and locations of the apparent clusters are driving our search for a high validity clustering method that is generalizable to other snapshots with minimal parameter selection. As such, at this step we are focusing on high level patterns that are representative of the trace.

We restrict our preliminary analysis to high activity UIDs to obtain a sense of how much noise our clustering algorithms will need to tolerate. In the `archive` snapshot, we see distinct clusters of file creation by the same user – for example the vertical line of blue accesses at approximately 1.2×10^9 seconds. We also see a range of users creating files late in the time covered by the snapshot. The `gnfs` case shows a few outlying creates that happen a long time after most do. We also see a very different modification pattern compared to the archival case; instead of many modifications by a user mapping to the same create time, we see the opposite, horizontal bands of user creation activity across a single modification time. Moreover, these bands overlap. This indicates that any clustering that accurately classifies usage requires more than three dimensions to separate distinct user groups. In terms of real usage patterns, the vertical clusters of the `archive` case could be a log file, whereas the horizontal clusters of `gnfs` could correspond to batch processes on the shared filesystem.

Finally, the `lnfs` case shows a local system with the high activity users doing almost all of their file modifications within a very small window of time. This indicates a very active system, where files are touched frequently, in contrast to the archival system where there is a large spread of modification times between users. Other views we have taken of these snapshots, including looking at UID and modification time plotted against File_ID or Group_ID (Fig. 2), follow very similar patterns. In the Group_ID views, for instance, we see that popular users are generally only modifying files that belong to one or two groups, and, particularly in the `lnfs` case, most of these modifications are within a small span of time.

Based on these results, we are investigating clustering algorithms that can support:

- *n*-dimensional, non-linear heterogeneous data
- Encoding hierarchical relationships between both data and labels
- Data with low inter-cluster separation, as in the `lnfs` case, without overfitting

Additionally, our clustering should handle arbitrary numbers of sparse, binary dimensions to encode "yes/no" questions about the files in the snapshot derived from the permissions and path fields.

IV. Current Directions

We have defined the concept of a view, and calculated views of archival and NFS for HPC datasets in order to direct our clustering methods, with the goal of showing that there is analytical value within a single system snapshot. Clustering a single snapshot is a novel approach that is useful given the size and performance requirements of modern systems. We believe that our clustering will be able to hint at working sets based on modifications grouped by labels such as user or path. These working sets could then reveal hidden characteristics of a workload, such as project interrelationships, from a single snapshot. Given enough data, our ultimate goal is to reliably classify snapshots by the type of workload they represent.

Moving further, we would like to combine the clusterings from a single-snapshot analysis with the methodologies in place, such as learning inter-reference intervals, to learn trends between snapshots. For instance, we could compare a series of clusterings to refine the clustering parameters. Finally, we are interested in studying the effect of changes in snapshots on the clustering to obtain a rigorous validity metric.

Acknowledgments

This research was supported in part by the NSF under awards CNS-0917396 and IIP-0934401, by the DOE under Award Number DE-FC02-10ER26017/DE-SC0005417, and by CRSS sponsors. We also thank all SSRC members for their valuable input and support.

References

[1] I. Adams, B. Madden, J. Frank, M. W. Storer, and E. L. Miller, "Usage behavior of a large-scale scientific archive," in *Proceedings of 2012 International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov. 2012.

[2] N. Agrawal, W. J. Bolosky, J. R. Douceur, and J. R. Lorch, "A five-year study of file-system metadata," in *Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST)*, pp. 31–45, Feb. 2007.

[3] M. Bhadkamkar, J. Guerra, L. Useche, S. Burnett, J. Liptak, R. Rangaswami, and V. Hristidis, "Borg: block-reorganization for self-optimizing storage systems," pp. 183–196, 2009.

[4] S. Doraimani and A. Iamnitchi, "File grouping for scientific data management: lessons from experimenting with real traces," pp. 153–164, 2008.

[5] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining*, 1996.

[6] T. Gibson, E. L. Miller, and D. D. E. Long, "Long-term file activity and inter-reference patterns," in *Proceedings of the 24th International Conference for the Resource Management and Performance and Performance Evaluation of Enterprise Computing Systems (CMG98)*, (Anaheim, CA), pp. 976–987, CMG, Dec. 1998.

[7] N. Rowe and S. Garfinkel, "Finding anomalous and suspicious files from directory metadata on a large corpus," in *3rd Intl. ICST conference on digital forensics and cyber crime*, 2011.

[8] A. Wildani, E. Miller, and O. Rodeh, "Hands: A heuristically arranged non-backup in-line deduplication system," *International Conference on Data Engineering (ICDE'13)*.