Describing HPC Filesystem Trees with the Grand Unified File-Index

Author: Jenna M. Kline | The Ohio State University, Department of Computer Science and Engineering
Advisors: Jason Lee & Rusty Davis | Los Alamos National Lab, HPC-DES

THE OHIO STATE UNIVERSITY

GUFI GitHub

Introduction

PROBLEM STATEMENT: Analyze, characterize, & describe modern HPC filesystem trees

GOALS

1. Exploratory analysis to characterize HPC filesystems
2. Propose first step in determining how HPC filesystems evolve

MOTIVATION

- HPC filesystems are extremely complex & difficult to manage with existing tools
- Previous studies of filesystem characteristics pre-date modern HPC filesystems of the last decade [1-3]

With GUFI, it is now possible to quickly query HPC filesystem metadata for large-scale system analysis.

GRAND UNIFIED FILE-INDEX (GUFI) [4]

- Enables file system querying that would be impossible to accomplish with POSIX tools
- Provides a novel methodology to collect & analyze filesystem metadata
- Single index for all filesystems

SCORECARD COMMAND LINE TOOL

- **Purpose:** Quickly generate high-level view of filesystem directories & subtrees with GUFI
- **Target Audience:** HPC system administrators & filesystem researchers
- **Recent use-case:** Identified a subtree using a disproportionate amount of metadata

Methodology

HPC FILESYSTEMS ANALYZED

LANL maintains massive filesystems to store scientific simulation data.

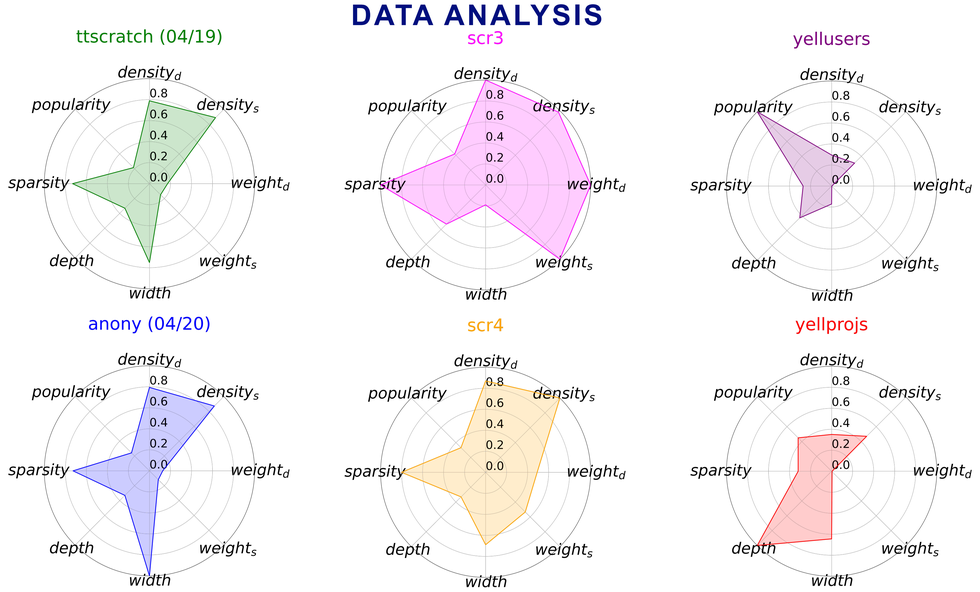
Filesystem	Total Size	Directories Count	Files Count	Depth	Type
anony	411 TB	7.4 M	157.6 M	37	Lustre
scr3	975 TB	2.2 M	59.2 M	59	Lustre
scr4	1 PB	5.1 M	118.9 M	37	Lustre
ttscratch	423 TB	5.5 M	117.5 M	37	Lustre
yellprojs	28 TB	14.1 M	133.3 M	112	NFS
yellusers	2 TB	1.6 M	12.9 M	48	NFS
Totals	**2.8 PB**	**36 M**	**600 M**	**-**	

METRICS TO DESCRIBE HPC FILESYSTEMS

- Used GUFI to query the metadata of 6 HPC filesystems listed above
- Identified 8 metrics that uniquely distinguish filesystem trees

Metric	Short Name	Definition
Directory Avg File Count	**Density$_d$**	Average number of files in a directory by level
Subtree Avg File Count	**Density$_s$**	Average number of files in a subtree by level
Directory Avg Size by Level	**Weight$_d$**	Size (bytes) of a directories averaged by level
Subtree Avg Size by Level	**Weight$_s$**	Size (bytes) of subtrees averaged by level
Directory Width	**Width**	Total number of directories at each level
Tree Depth	**Depth**	Maximum level of a tree
Directory Sparsity	**Sparsity**	Count of empty directories & files in a tree
User Popularity	**Popularity**	Count of unique user ids

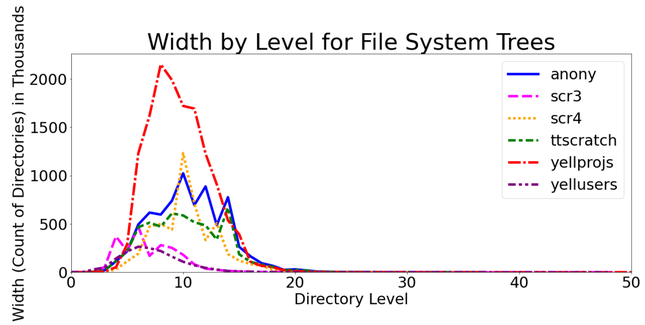
DATA ANALYSIS

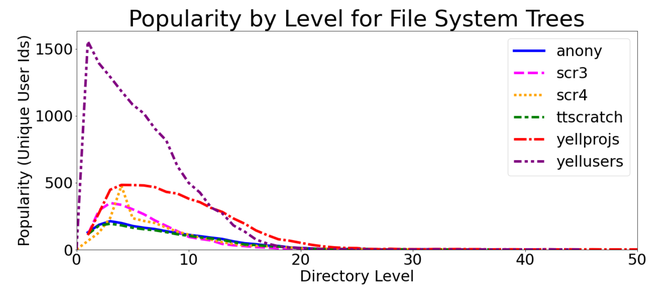


Results

KEY TAKEAWAYS

- **Top ~20 levels are most unique to each filesystem**
- **Branching factor not a meaningful distinguishing measure**
- **Directory shape stays consistent over time**
- **Directories may widen as they age**
- **Scratch directories tend to be larger & denser**





Charts. Width and Popularity by Level for File System Trees

Future Work

1. Collect additional HPC filesystem traces to analyze their evolution
2. Use GUFI to build more complex metrics to describe filesystems
3. Predict requirements of future HPC filesystems with GUFI

Los Alamos NATIONAL LABORATORY
NNSA National Nuclear Security Administration
Managed by Triad National Security, LLC, for the U.S. Department of Energy's NNSA.
LA-UR-23-28917

References
[1] A. Leung, S. Pasupathy, G. Goodson, and E. L. Miller, "Measurement and Analysis of Large-Scale Network File System Workloads," presented at the USENIX Annual Technical Conference, Jun. 2008.
[2] S. Dayal, "Characterizing HEC Storage Systems at Rest," 2008.
[3] N. Agrawal, W. J. Bolosky, J. R. Douceur, and J. R. Lorch, "A five-year study of file-system metadata," ACM Trans. Storage, vol. 3, no. 3, pp. 9-es, Oct. 2007, doi: 10.1145/1288783.1288788.
[4] D. Manno et al., "GUFI: Fast, Secure File System Metadata Search for Both Privileged and Unprivileged Users," in SC22: International Conference for High Performance Computing, Networking, Storage and Analysis, Nov. 2022, pp. 1–14. doi: 10.1109/SC41404.2022.00062.