A Distributed Digital Body Farm for Collecting Deleted File Decay Data

By:
Omoche Cheche Agada (George Mason University)

From the proceedings of

The Digital Forensic Research Conference

DFRWS USA 2022

July 11-14, 2022

DFRWS USA 2022

A DISTRIBUTED DIGITAL BODY FARM FOR COLLECTING DELETED FILE DECAY DATA.

AGADA Omoche Cheche

July12, 2022

GEORGE MASON UNIVERSITY
College of Engineering and Computing

Agenda

- ❑ Introduction
- ❑ Methodology / Implementation
- ❑ Data Analysis
- ❑ Conclusion and Future Work

Introduction

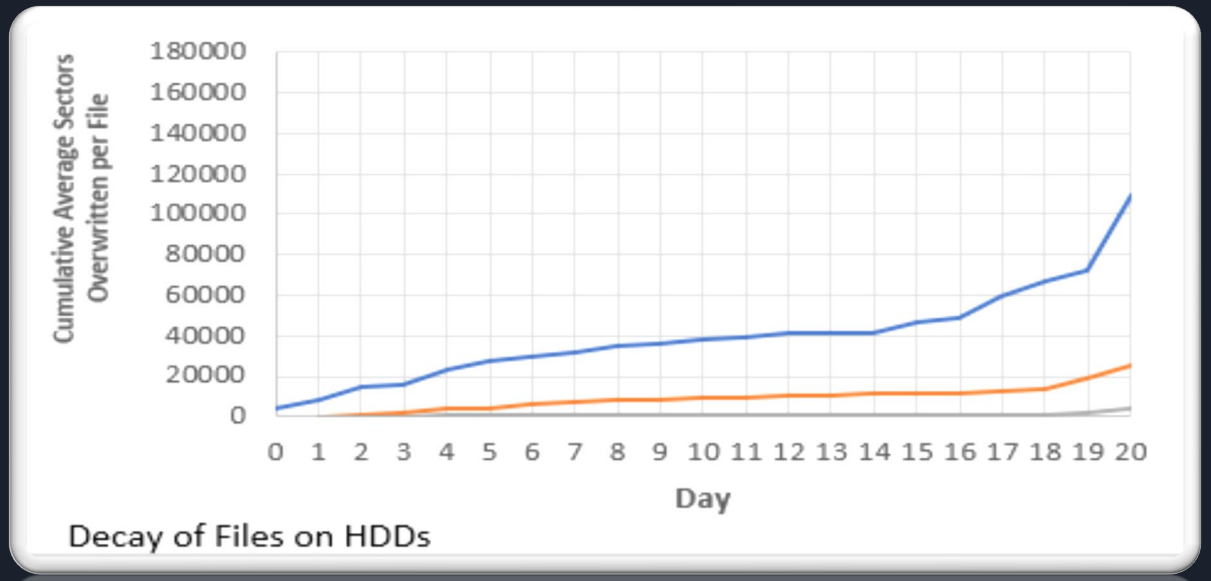
The Concept of a Body Farm

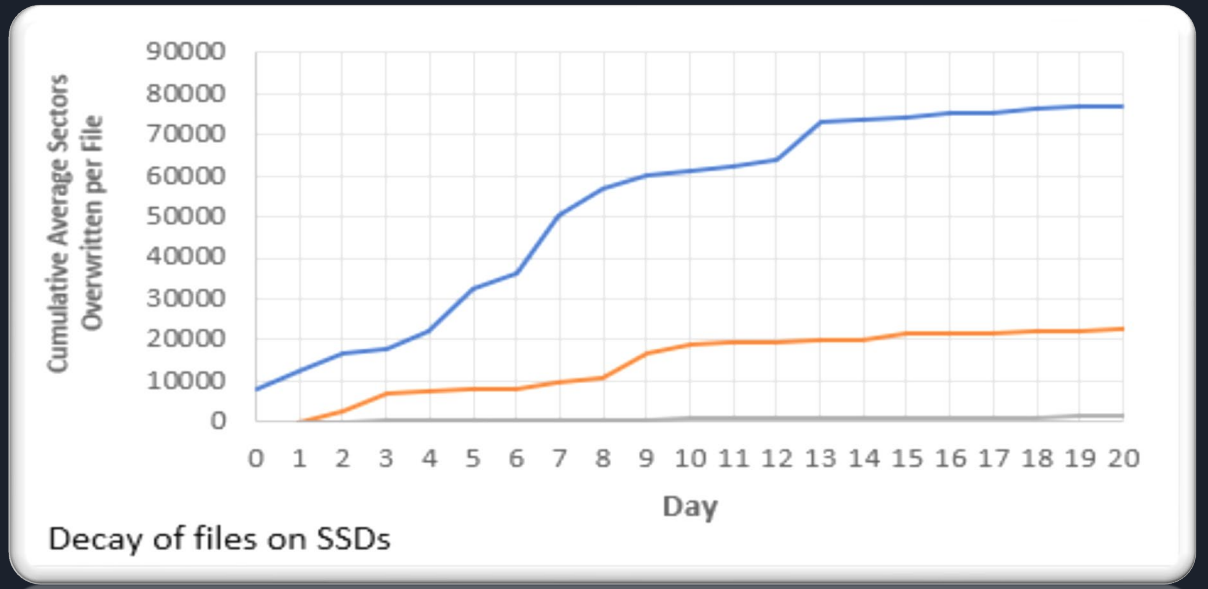
- Native to Forensic Anthropology.
- A facility where human remains is studied for decay patterns.
- Used in law enforcement for homicide investigations.

Introduction

The Concept of a Digital Body Farm

- A novel application for studying patterns of decay in deleted digital files.
- Application is deployed on systems belonging to volunteers.
- No PII or user data is collected.

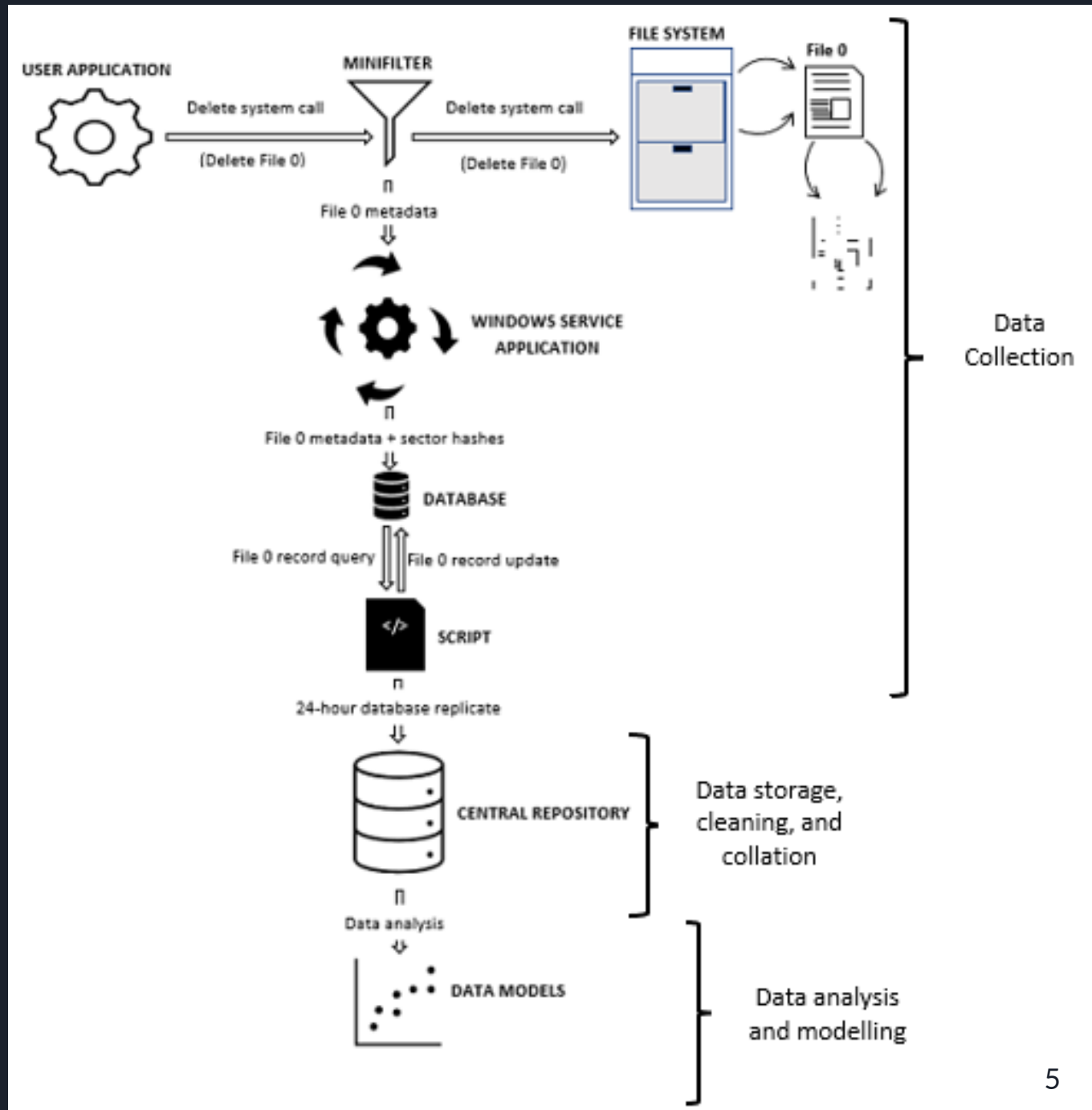




Methodology/ Implementation

The DDBF

1. Filesystem filter driver
2. Windows service application
3. Database
4. Script
5. Central Database or Repository



Data Analysis

Linear regression with percentile cutoffs

- File type (extension)
- File size
- File category
- Sector count
- Disk type

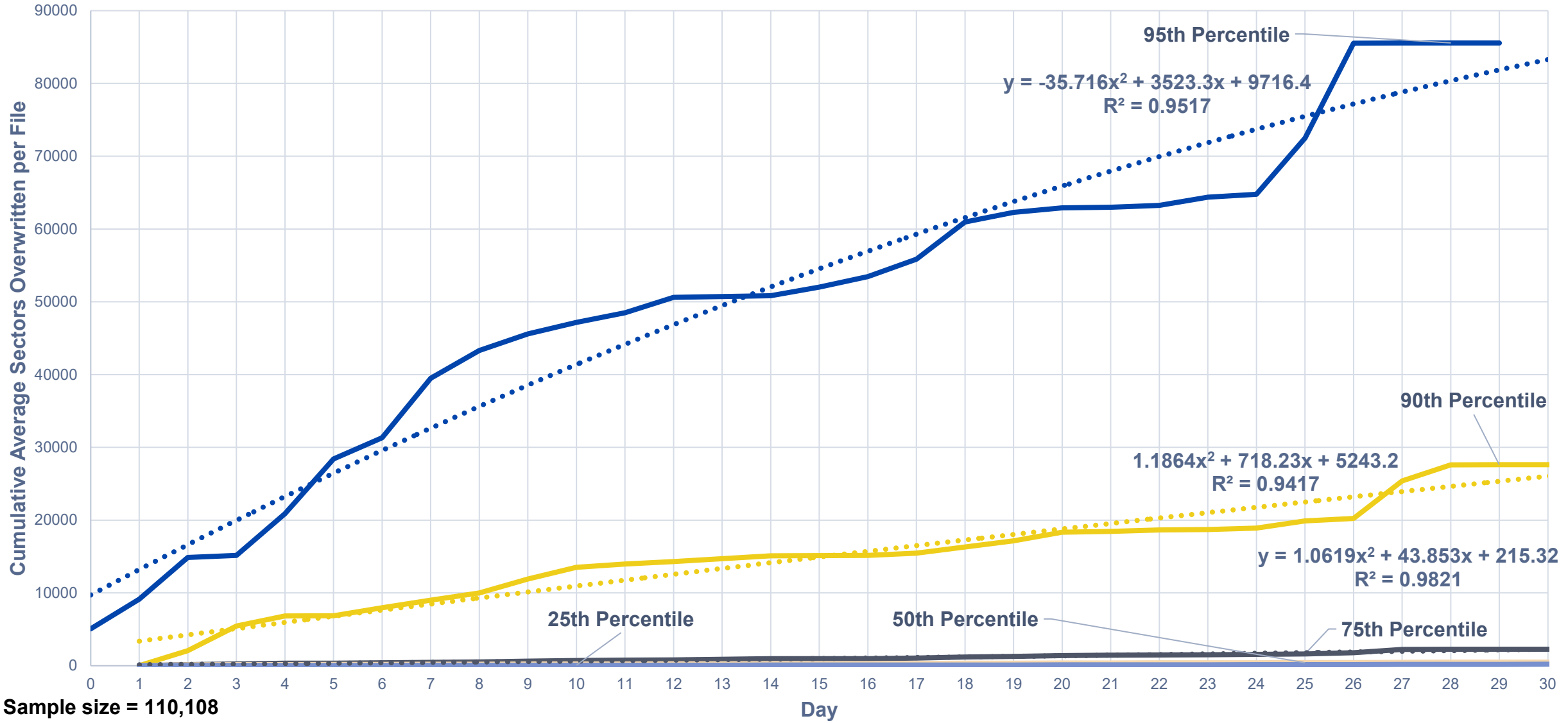
12 PCs – 3 Time Zones

Cumulative Aggregates of Daily Sector Decay for 110,108 files (monitored for 30 days)

- 25th Percentile
- 50th Percentile
- 75th Percentile
- 90th Percentile
- 95th Percentile

Data Analysis

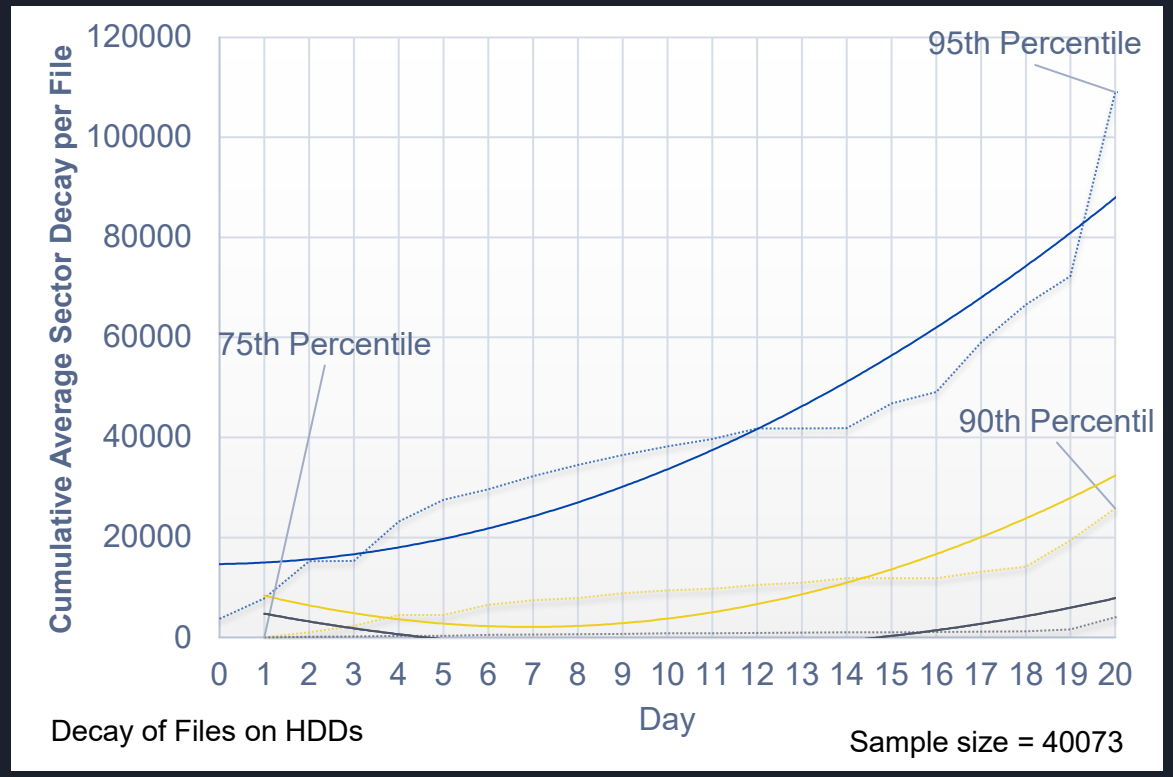
Regression Model



Sample size = 110,108

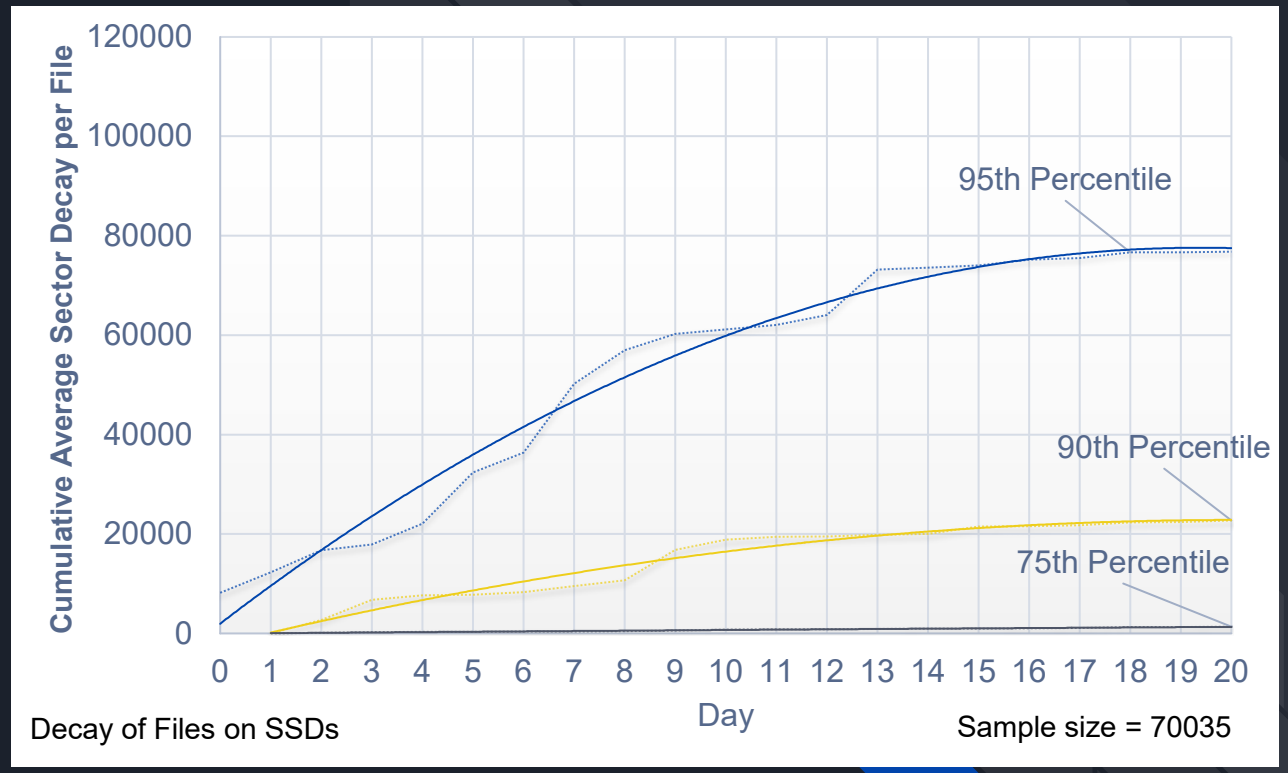
Data Analysis

Modelling Decay by Media Type



Decay of Files on HDDs

Sample size = 40073



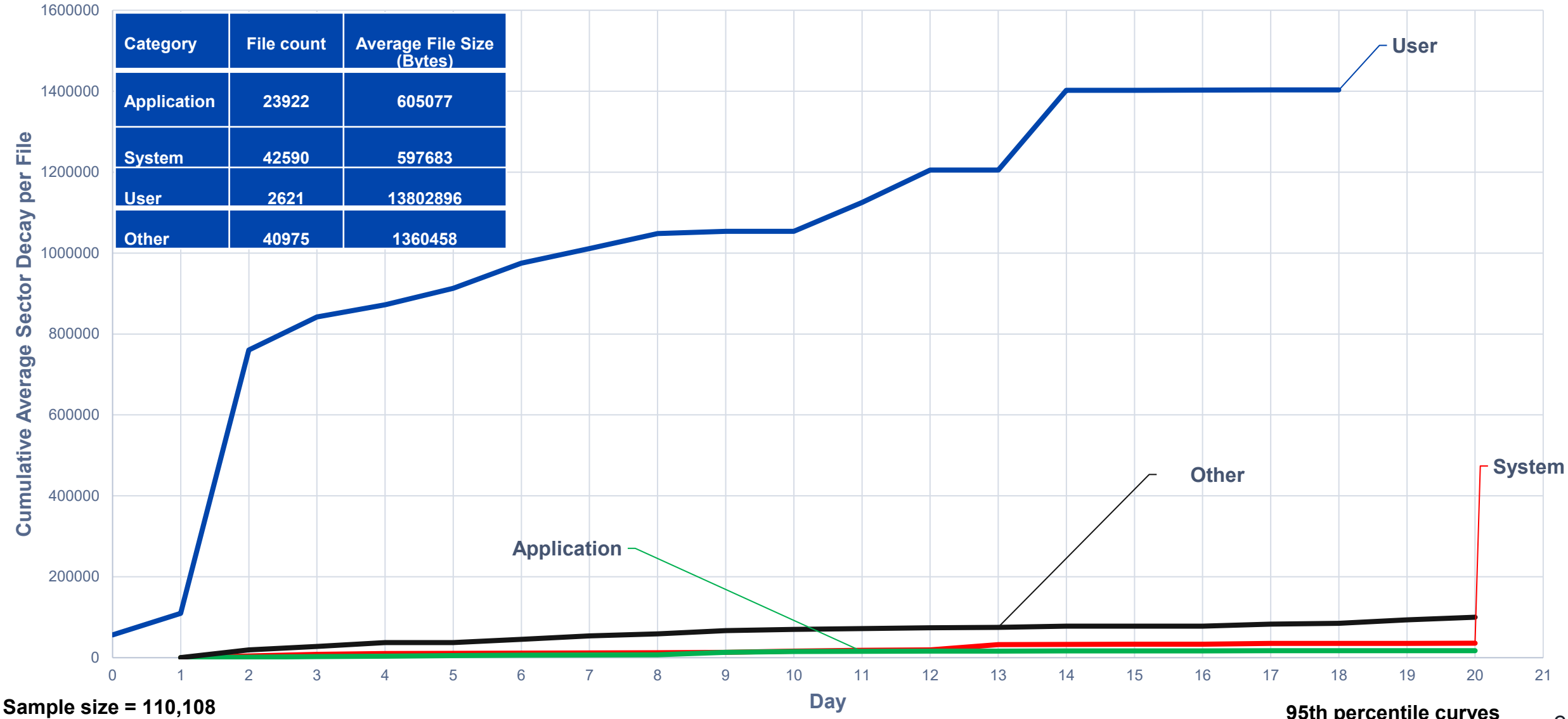
Decay of Files on SSDs

Sample size = 70035

Data Analysis

Modelling Decay by File Categories

Category	File count	Average File Size (Bytes)
Application	23922	605077
System	42590	597683
User	2621	13802896
Other	40975	1360458

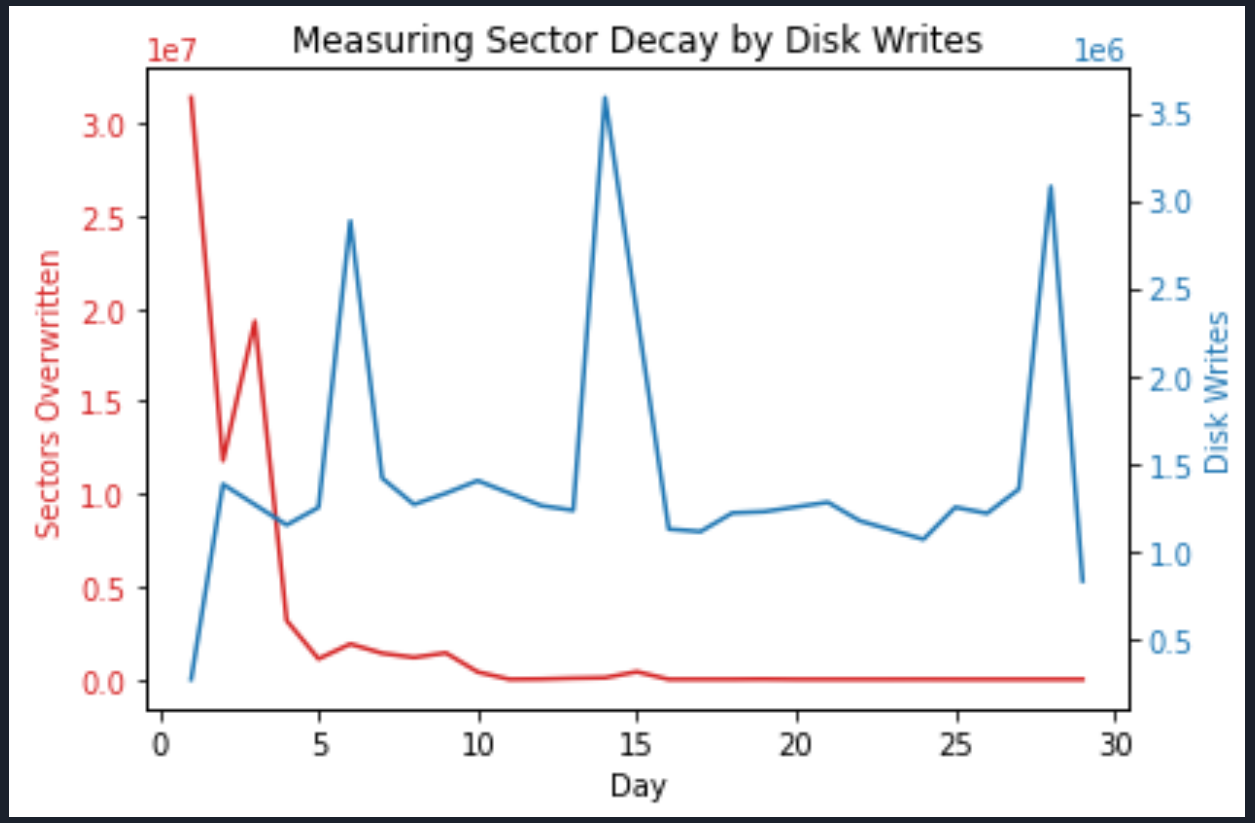


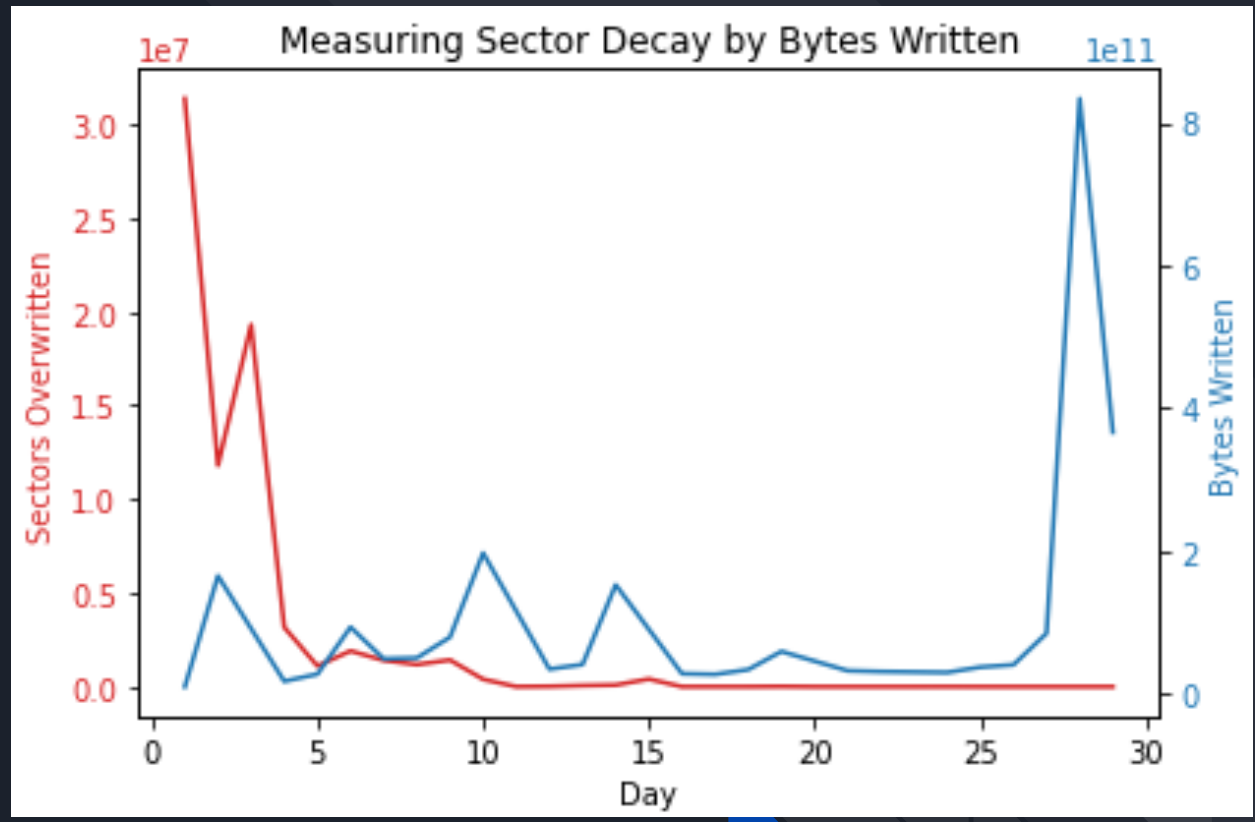
Sample size = 110,108

95th percentile curves

Data Analysis

System Load Analysis

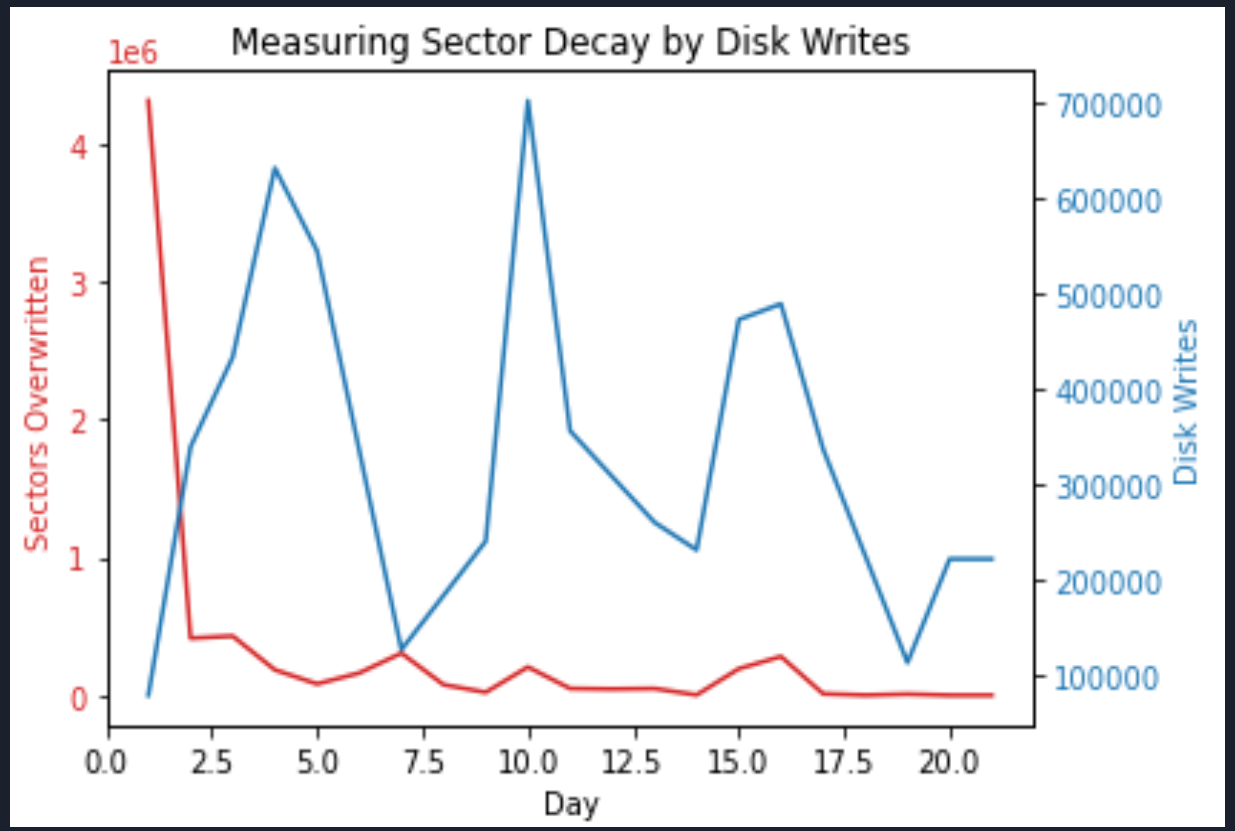


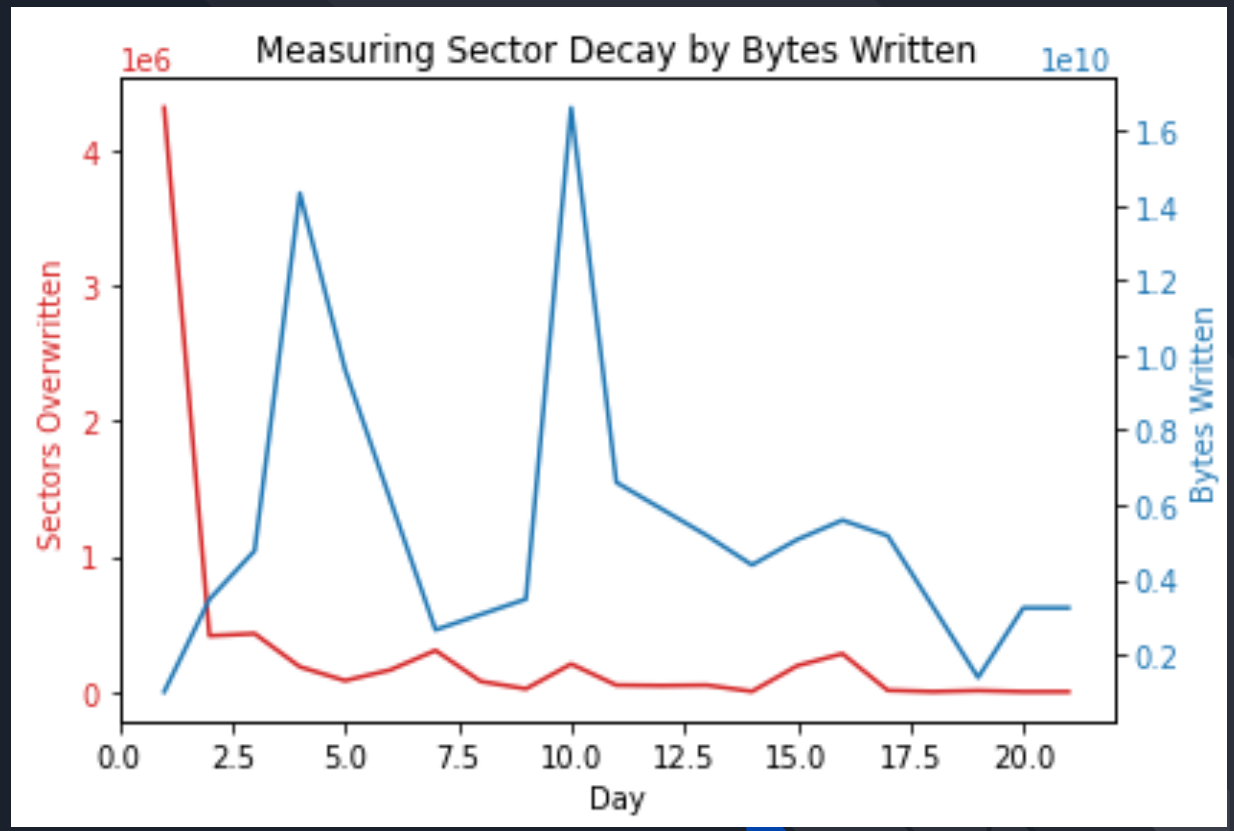


PC	Disk Type	Disk Brand	Disk Size	Disk Usage	No of Files
1	SSD	KIOXIA	512 GB	74 %	16506

Data Analysis

System Load Analysis

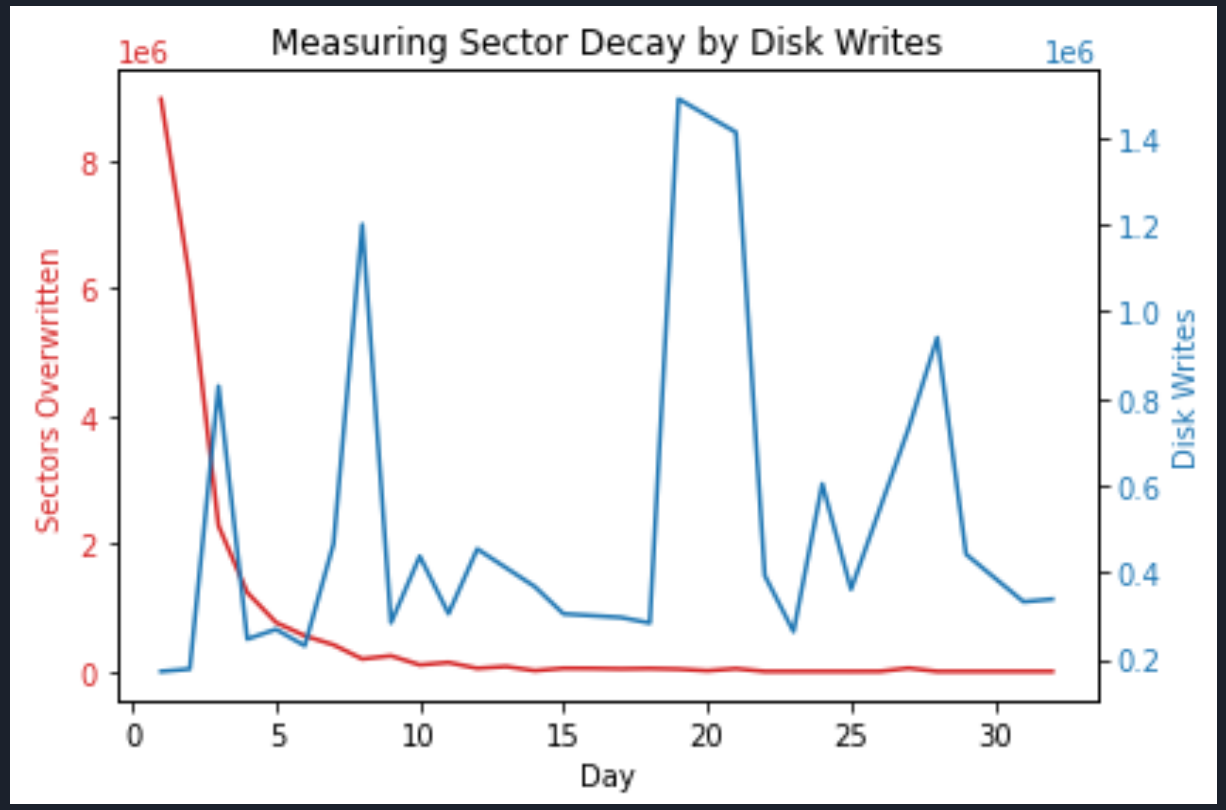


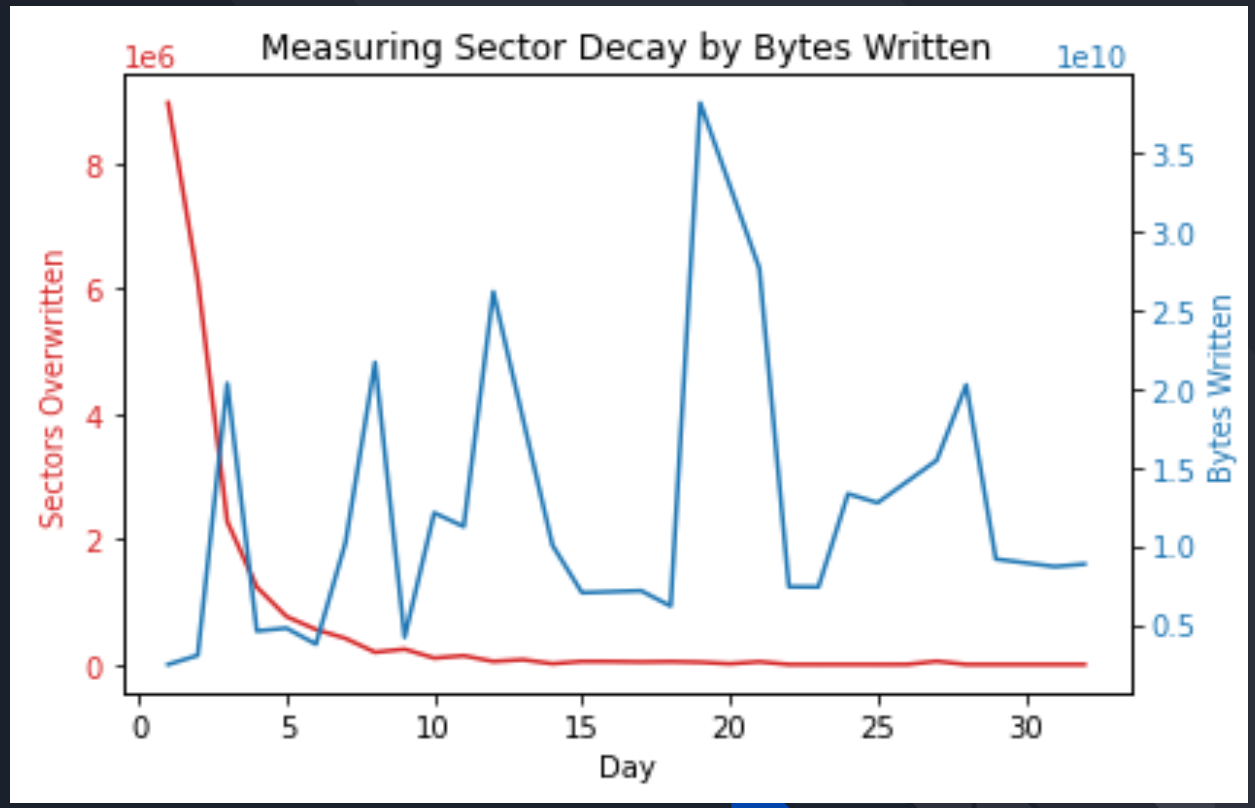


PC	Disk Type	Disk Brand	Disk Size	Disk Usage	No of Files
2	SSD	KINGSTON	240 GB	89.2%	4987

Data Analysis

System Load Analysis

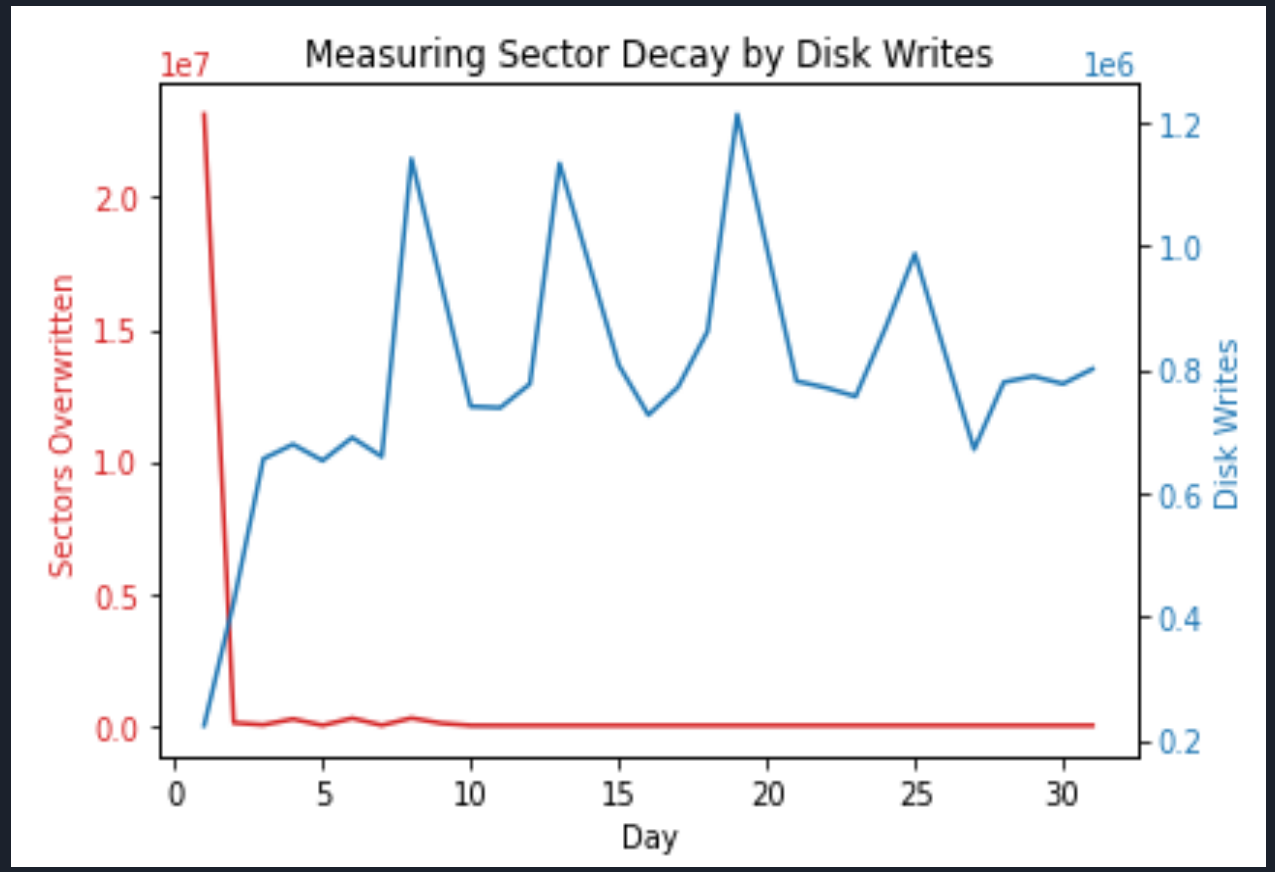


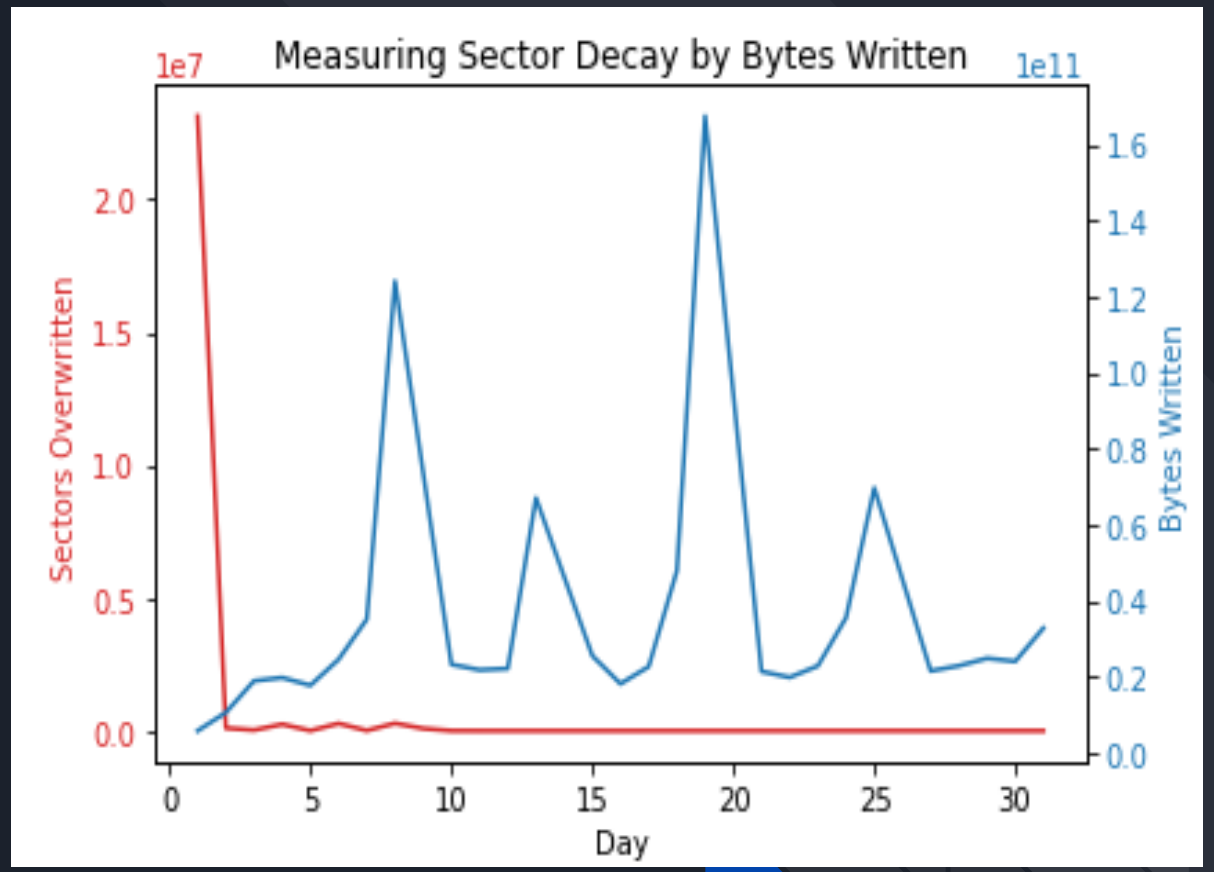


PC	Disk Type	Disk Brand	Disk Size	Disk Usage	No of Files
4	HDD	SEAGATE	1 TB	8.4%	6960

Data Analysis

System Load Analysis





PC	Disk Type	Disk Brand	Disk Size	Disk Usage	No of Files
5	SSD	SAMSUNG	256 GB	69.5 %	9735

Data Analysis

Predictions (Sector Decay on SSD)

Test Data
12483

No of Days	75th		90th		95th	
0	149	56.1%	2555	50.6%	1959	50.7%
5	500	90.3%	12162	88.3%	35951	88%
10	834	93.6%	18716	92.3%	59874	92.2%
15	1149	93.9%	22215	92.9%	73726	92.7%
20	1446	93.6%	22660	92.9%	77509	92.7%

Prediction | Accuracy

Conclusion and Future Work

- This phase of our work proposed a methodology and tool for studying deleted file content decay.
- The distributed nature of the tool allowed for remote data collection while preventing the exposure of sensitive user data.
- The models developed provide insight into the data decay process.
- The goal of future work is to generate a large dataset of file decay data for constructing a self-learning model to be employed in file decay prediction.

References

Digital Body Farm Project (Github repo)
https://github.com/chex2chex/Digital-Body-Farm

Omoche C. Agada, Ibifubara Iganibo, James Jones, and Kevin D. Fairbanks. 2022. A Digital Body Farm for collecting Deleted File Decay Data. In IFIP International Conference on Digital Forensics. Published in the "*Advances in Digital Forensics XVIII: 18th IFIP WG 11.9 International Conference, Virtual Event, January 3-4, 2022*". (Release date - August 15, 2022).

Omoche Cheche Agada, James H. Jr Jones, and Kevin D. Fairbanks. 2022. The Distributed Digital Body Farm: Enabling the Analysis of Deleted File Decay Patterns. In Cyber Security Experimentation and Test Workshop (CSET 2022), August 8, 2022, Virtual, CA, USA.

Contact
oagada@gmu.edu