

Summary of Video and Photo Storage Usage

Report by Crystal Hollis

This report provides a detailed analysis of our storage utilization, particularly focusing on fiscal years 2015 to 2025. The primary goal was to quantify the extent of our stored footage and understand the implications for future storage needs, especially as we continue to adopt higher-quality equipment.

Data Analysis Process

The analysis was conducted using a comprehensive dataset that included Job Numbers, Neg Numbers, File Paths, and Folder Sizes across various directories. Notably, the dataset does not include data from a few select hard drives and Box. It is important to note that an unknown quantity of video folder sizes may only be accessible on Box, as we have primarily backed up video to Box and then deleted it locally to save server and hard drive space. While this approach was operationally beneficial, it might not be ideal to rely solely on one backup location. The analysis was performed using Python and Microsoft Excel, with the final dataset saved as DEAPhotoStorage_2024-08-18_v1_final.csv.

Key Steps:

1. **Data Collection:** Python was used to scan all relevant directories, except for a few inaccessible hard drives and Box. Photography Jobs data was gathered from FileMaker.
2. **Data Segmentation:** The data was segmented by fiscal year, job number, and neg number.
3. **Size Calculation:** The sizes of video and photo folders were calculated to determine average storage requirements over time. Columns include Folder Sizes in Bytes, Gigabytes (Decimal system), and "Readable" (Binary system).

Key Numbers for Comparison

1. Total Job Numbers in FY 2023 and FY 2024	1,341 in FileMaker; 766 identified in dataset
2. Total Job Numbers in FY 2025 as of 08/19/24	23 in FileMaker; 1 identified in dataset
3. Total Jobs and Neg Numbers Between FY 2015 and FY 2025	Total Job Numbers: 5,431 in FileMaker; 2,038 identified in dataset Total Neg Numbers: 8,555 in FileMaker; 3,331 identified in dataset
4. Total Neg Numbers Associated with "Donor" and "Campaign" in Folder Names (Between FY 2015-2025)	275 in FileMaker; 127 identified in dataset
5. Average Size (Gigabytes) of Donor and Campaign Related Jobs (Between FY 2020-2025)	5.97 GB
6. Average Size (Gigabytes) of Jobs in FY 2023 and FY 2024	9.42 GB
7. Average Size (Gigabytes) of all Jobs Between FY 2015-2025	7.67 GB
8. Average Size (Gigabytes) of Photos in FY 2023 and Between FY 2015-2025	FY 2023: 3.68 GB FY 2015-2025: 3.04 GB
9. Average Size (Gigabytes) of Videos in FY 2023 and Between FY 2015-2025	FY 2023: 50.13 GB FY 2015-2025: 17.04 GB
10. Total Number and Size of Folders Pending Processing Without Job Numbers Identified	File Count: 106 folders Folder Size: 32.37 TB
11. Total Number and Size of Folders Not Located in /Volumes/DEAPhoto\$	File Count: 3454 folders Folder Size: 57.96 TB
12. Total Number and Size of Folders in /Volumes/DEAPhoto\$	File Count: 302,933 folders Folder Size: 29.18 TB

Factors Contributing to Discrepancy

Several factors could contribute to the discrepancy between the numbers in FileMaker and the numbers in the dataset:

- **Incomplete / Inaccurate Data Entry:** Some job numbers or neg numbers may not have been entered correctly.
- **Pending Jobs:** There are jobs that are still pending and not yet fully processed or archived.
- **Duplicates/Temporary Files:** Duplicate entries could inflate the numbers.
- **Jobs stored in directories not included in dataset:** Some jobs might be stored elsewhere that was not part of the directory scan.

Given these discrepancies, the confidence level in the accuracy of the dataset is moderate. Further investigation and regular updates are recommended to maintain data accuracy.

Additional Insights – Factors Affecting Video Sizes

According to Senior Videographer Garrett Holton, his average project size ranges from 100 GB to 200 GB, with larger projects requiring up to 500 GB to 700 GB due to extensive b-roll and RAW 4K footage. This differs significantly from other video types, such as final exports, compressed copies, and older videos that have yet to be processed and archived. While our analysis shows smaller average project sizes, it's important to consider future expectations and upgraded equipment. If we anticipate producing more high-quality 4K RAW and LOG footage, we should prepare for increased storage demands.

Random Samples

File Path	Folder Size (GB)	File Count
/Volumes/DEAphoto\$/Video/1 Back And Forth/GH_EDITS/2-READY FOR UPLOAD/35670V-CoxDavidBMillerBusinessQuadrangleDedication(GH)05.03.24/35670V-RAW	106.5 GB	40 files
/Volumes/5TB/FOOTAGE NOTE: This is SMU DEA 007 (SMU Dallas and Beyond, Minus Red(MR))	2.78 TB	1,199 files
/Volumes/SMU DEA 010/Ryan Blitzer/21-291 Ware/WARE4	1.02 TB	3,343 files
/Volumes/2TB/22067-SMU Cox Event Video	758 GB	19,390 files
/Volumes/SMU Meadows/SMU Meadows/01_Footy/B Cam	704 GB	352 files
/Volumes/SMU DEA 031/smu_20daa_waldrep	654 GB	1,073 files

Feel free to review the attached dataset: DEAPhotoStorage_2024-08-18_v1_final.csv. Keep in mind that "Size (GB)" column uses the decimal system (which is what Mac OS uses) and "Folder Size (Readable)" column uses the binary system.

Conclusion

OIT's projection that 10-14 TB of storage space may be required annually is a reasonable estimate. The large amount of unprocessed data (32.37 TB) and non-centralized data (57.96 TB) represent a combination of many years' worth of video and photography, dating as far back as 2012. This is not representative of annual storage demands but rather the demand for over a decade's worth of footage and project files. The growing use of high-quality equipment and the increasing volume of data suggest that storage needs will continue to rise, supporting the idea of requiring an additional 10-14 TB of storage space annually.