SCENE SAVERS
Making History Modern

What Is LTFS and Should I Be Using It In My Archive?

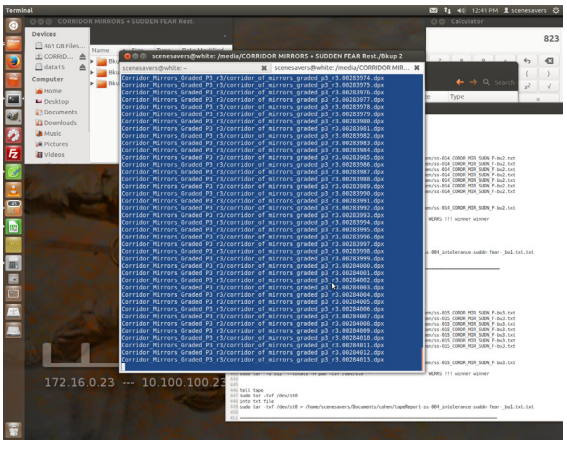
As you are evaluating different methods of storing and preserving your digital data, I'm sure you've come across LTFS. In this article, we will discuss what LTFS is and whether you should be considering it for your archive.

LTFS stands for Linear Tape File system. This is a file system that was developed in 2010 by IBM to address challenges with storing and recovering data from LTO data tapes. According to Wikipedia and other sources, LTFS refers to both the format of data recorded on magnetic tape media and the implementation of specific software that uses this data format to provide a file system interface to data stored on magnetic tape. Additional information can be found at LTFS on Wikipedia.

So, what does this mean in laymen's terms? In order to understand what this means, you have to look at the background of storing data on magnetic tapes. For purposes of this article, I'm going to stick to discussing LTO data tapes, as these are commonly used in archives for backing up and storing files for long-term preservation. If you want more information on LTO data tapes, please go to: LTO on Wikipedia

Image of LTO-5 data tape.

LTFS works with LTO-5 and LTO-6 data tapes. It does not work with LTO-4 or earlier versions of LTO. LTFS works by creating two partitions on the tape. Partition 0 contains a human readable XML index of the content on the tape. Partition 1 contains the digital files.



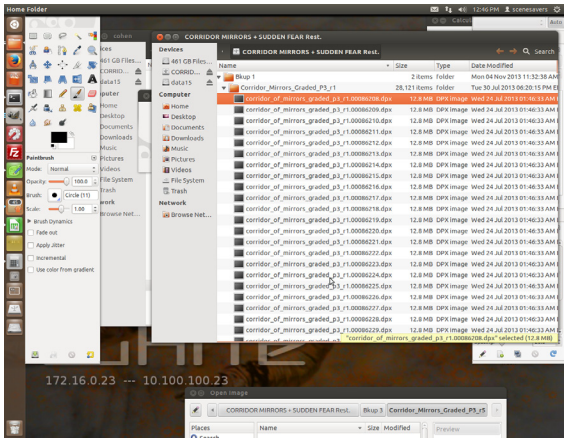
Example of USTar and how it appears in browser window.

Prior to the development of LTFS, there were a number of ways to write data to a magnetic tape, but all suffered from various challenges. TAR or USTar (Uniform Standard Tape Archive) is one open-source and widely used/widely available method of writing data to a magnetic tape. This is an older format, standardized in 1988, that has its roots in Unix. While it accomplished the task of writing data to magnetic tapes, and allowed that data to be accessed by multiple systems, it was slow. It also required users to have some technical knowledge when writing data to, or recovering data from magnetic tapes. Its main benefit is that it is open source and widely supported.

424 SCOTT ST. COVINGTON, KY 41011
859.291.5100
www.scenesavers.com

Other ways to write data exist, such as BackupExec. While these programs would allow incremental backups or daily updates, they are proprietary to the company(s) that developed them. We all know that relying on any one company's products for all of our backup, storage and retrieval is dangerous, as it could mean we would be unable to access that data in the future if the company goes out of business or changes how their systems/software work.

Enter LTFS. LTFS allows users to view a data tape as they would a hard drive. As long as the tape drive is accessible by the computer they are working on, the tape will appear as another drive. Users have the ability to drag and drop files to/from the tape, just as they would from a drive. LTFS contains the index of the contents of the data right on the tape. Using LTFS, no other indexing system is required, as was the case with other solutions. For example, if you need to share a tape where the data was written as USTar, the whole tape would need to be indexed prior to recovery of any files by the recipient. For a full LTO-5 tape, this would require several hours. Alternately, complex management systems could be used to maintain an index for all tapes, but these systems are prohibitively expensive for most archives.



Example of LTFS and how it appears in browser window.

With LTFS, since the index is already included, this search and retrieval is very quick, allowing user's almost instant access to any content on the tape. This makes it perfect as a stand-alone solution, without the need for a complex management system in the background. It also enables sharing of tapes between organizations, since the tape contains all the information needed for the recipient to view and recover files.

So, you might be asking "What are the downsides?" Two or three disadvantages come to mind when using LTFS. One is that, even though users can see a directory listing of what is on the tape, it is still a tape with its inherent delay in seeking files. So, even though it appears as a hard drive, there will be a delay in retrieving and playing back your content. Second, the format is an append only format, meaning that you can add or delete data, but you can only free up the space used by the deleted data with a complete tape erase. Since the format should be used as a long-term preservation strategy, this really shouldn't be an issue. The third challenge is a little more complicated. LTFS doesn't support unique identifiers that might be required by other systems in your organization, such as media asset management systems. The MAM might use unique identifiers to keep track of all the assets at your location as an example. Even if using LTFS, you may still be required to use your MAM and databases for efficient management of all assets in your organization.

Still, LTFS has a lot to offer an archive. From being a stand-alone, self-contained, self-describing solution to being open source and supported by many different manufacturers and operating systems, LTFS definitely has a place in your archive for long-term preservation management.

424 SCOTT ST. COVINGTON, KY 41011
859.291.5100
www.scenesavers.com