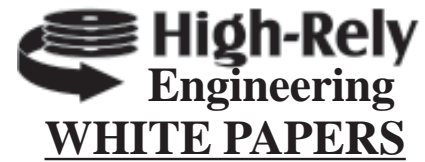


Juner 21, 2006

RE: The Best Media for Archival

by Darren McBride



Summary: What truly is the best media for Archival purposes and why? This paper is intended to help the reader understand the issues and criteria of relevance in making this decision.

When we talk to our colleagues about moving away from tape backups, often, a number of fallacious arguments are presented. Here, we will discuss two typical objections and refute them, point by point.

The two common objections are 1.) Tape is the best thing available for archival, and 2.) Everyone else uses it

Tape is the best thing available for archival.

When you get down to it, any media is poor for true archival purposes. Paper printed with archival ink on non-acid paper and hermetically sealed will last hundreds of years. Stone tablets are even better. Obviously, this creates a size problem, so let's look at other types of media. Here are some stats as claimed by vendors:

Quantum DLT Tapes 30 Years
Manufacturer Reference: <http://www.quantum.com/am/products/dlt/media/default.htm>

CD-R/DVD (in general) 10 Years
Independent Reference: http://www.athensmusician.net/archive/2001-10-17_cdr01.shtml

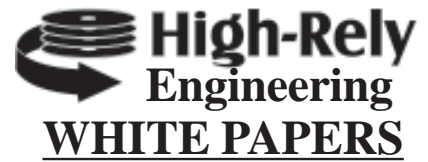
Western Digital Hard Drives 5 Years
Manufacturer Reference: <http://www.westerndigital.com/en/products/current/drives.asp?Model=WD2500JB>

At first, it looks as though tape wins, and this is just what the tape vendors want you to think. However, there are 2 major problems with these numbers.

The first problem is the "30 year" figure for tape, and the "5 year" figure for hard drives. They're referring to different things. In the case of the tape, it is how long the tape is supposed to last, sitting on a shelf somewhere, in a proper environment. The hard drive, on the other hand, is talking about continuous use in a server or desktop PC. How does that translate into years on a shelf? The truth is we can't really be sure. In my shop, there are several PC's from the mid 1980's that still boot and run just fine from some of the first hard drives that were made available to consumers.

Here is an article discussing [hard drive reliability factors](http://www.storagereview.com/guide2000/ref/hdd/perf/index.html), such as MTBF's and Service Lifetimes: <http://www.storagereview.com/guide2000/ref/hdd/perf/index.html>

1. Since we are talking about three different types of media here, let's discuss what goes bad with each.
2. With CD-R's and DVD's, it's fairly well understood that 10 years is what you can expect before oxidation of the metallic foil sets in and starts to kill the data.
3. Older hard drives had a problem requiring the need to be rewritten every few years to "freshen" the magnetic field that contains the data. Newer hard drives don't have this particular problem. The problem with the newer hard drives is "freezing." That is, after a number of years, it is possible for the platters to get stuck due to deterioration of the mechanical components that turn them.



With tape, there are two main problems. The first is the “shedding” of material from the surface of the tape. The second is the separating of the layers of the tape. Because tape formulations change all the time, their “30 year” number is purely theoretical.

Ultimately, in even 15 years, it is going to be very tough to find a working DLT drive. It may even be hard to find an IDE or SCSI controller. In the end, the media is not the problem. It is the hardware and software you use to read the media. That is why the only valid archival program consists of transferring the data every 5 years or so to a slightly newer medium.

With all of that said, let’s take a look at one of the institutions that has the highest standards for the preservation of data, The Library of Congress. They have examined this issue carefully and conclude the best way to preserve data, is to have a massive 185 Terabyte RAID array running all the time. In other words, they aren’t even going to take the stuff offline. They have all sorts of offsite mirrors of this array, but ultimately they realize that they will have to have a “living” system to reliably archive our nation’s history.

Realistically, Hard Drives, DVDs, CDs, and Tapes are all equally good for short term archival of 5 or so years. Beyond that, you aren’t “truly” archiving the data; you are simply following the letter, and not the spirit, of your archiving requirements.

“Everyone else uses it?”

It’s understandable, in a way. New people coming into the IT field have been told that tape is just something they are going to have to deal with in their career. I myself was resigned to such a fate for many years. I concocted several alternatives of my own over time. These alternatives would have certainly worked better than tape, but the trump card my boss or CFO would always throw down was “That’s just too expensive. I see tape drives for less on CDW.” I could never convince them that **IT’S ONLY CHEAPER IF IT ACTUALLY WORKS.**

Many colleagues in the IT profession have lost a lot of important data to tape malfunctions. I found myself in a position where I would certainly not trust my job to tape. I began to do things like copy all of the data to unused workstations after the normal tape job had run. I would cobble together anything I could to provide a backup for my backup. I know many IT people who do this very same thing because they are put in an awkward position by their bosses.

I see now, that the fundamental problem was that I involved tape at all. Now, I use removable hard drives and do a straight file copy to the drives. At this point in my career, I won’t accept responsibility for someone’s data if the only backup solution they will provide is a tape drive. I’ve been burned too many times, and I’m sure you have too.

