# Cost of Ownership of Disk-Based vs. Digital Data Tape-Based Video Archives

By Christopher A. Stone, Brillig Consulting

The cost of storage has steadily fallen as new developments in magnetic recording have been introduced for both disk drives and digital data tape. With the enticing low price offered for high capacity Serial Advanced Technology Attachment (SATA) hard disks there is a temptation to just store your video archive on rotating media. But is this a wise move? This white paper looks at some of the obvious and not-so-obvious issues of archiving video data on both disk-based and tape-based systems.

Let's first address some of the traditional comparisons between disk and tape, leaving out cost for the time being. These are:

- Capacity
- Performance
- Data Availability
- Volatility

In the past, data tape cartridges always had more capacity than a hard disk. With the advent of SATA drives this is no longer the case. 1 Terabyte (TB) drives are readily available, while the largest capacity tape cartridge in common use is 800 Gigabytes (GB) (with a road map to 1.6 TB and 3.2 TB already planned).

Traditionally, data could be transferred faster from hard disk than from tape, but the latest LTO 4 tape drives can deliver real life sustained data rates greater than 100 Megabytes (MB) per second. This is much faster than SATA RAID 5 disk systems which can only manage approximately half that transfer rate (with 7200 RPM drives and a 5-drive RAID set).

So two of our traditional values, capacity and speed, are turned "topsy turvy". What about the others, data availability and volatility?

The primary advantage that disk drives enjoy today is fast access to small amounts of data. Data stored on a disk drive is positioned randomly and the data transfer rate is a function of the rotational speed of the disk and the time to move the read-write head. Tape drives, on the other hand, record data sequentially along the whole length of the tape. Let's look at the time to retrieve a two-hour DV25 video clip from both devices:

To access the video from disk we have a seek time to get to the beginning of the data and then a transfer time to read all of the clip:

Average seek time: 5-10 milliseconds and is insignificant
DV25 bit rate : 25 Megabit (Mb) per second = 3.125 MB per second
2 hours of video : 3.125 x 3600 seconds per hour x 2 hours = 22,500 MB

At 50 MB per second it will take 450 seconds or 7.5 minutes to read the clip.

To access the same video from a tape housed in a robotic tape library incurs time for the robot to fetch the tape, time to load the tape into the drive, time to seek to the beginning of the data, and then the transfer time to read the clip:

| | |
|---|---|
| Robotic move: | 10 seconds |
| Tape load time: | 22 seconds |
| Average seek time: | 62 seconds |

At 110 MB per second it will take 205 seconds to transfer the data. Adding the access time yields a total of 299 seconds or almost 5 minutes to read the clip, which is 33% faster that getting the clip from a hard disk RAID. Similarly, a one hour clip takes 3.75 minutes to transfer from disk and 3.27 minutes from a tape library.

This is important to note because as more content is recorded in High Definition and the required storage capacity increases for a clip of the same duration, data transfer rates will become even more significant to overall archive performance. Digital video archive systems based on robotic tape libraries and modern high performance tape drives will significantly outperform disk-based storage systems in most applications.

Finally, data volatility, or more especially non-volatility, is certainly a key consideration. Losing archived content is simply not acceptable. The different RAID schemes provide varying levels of security. The compromise is, as always, speed versus cost. For example, RAID 10 improves performance but is very costly, while RAID 5 provides more cost effective storage but with lower performance.

Modern tape technology with error correction and Read after Write verification is very good and stable. Apart from physical damage to tapes the data can be considered to be non-volatile. Making a second copy of a tape is simple, inexpensive and can be an automatic process using video archive software such as the XenData package. Indeed, two copies can be retained on site and a third copy kept in a separate location for disaster protection, simply for the cost of the media.

The same cannot be said for disk based systems. The SATA disk drives used for large capacity RAIDs needed for video archives do fail, and given the number of individual drives required it is not unlikely that two drives may fail within a very short period of each other as they age or are subjected to the same environmental conditions. This effectively rules out using a single RAID 5 because the second failure in the same RAID set effectively destroys all the data.

Using a RAID 6 system will improve security, but the only way to maintain overall performance is to use replicated RAID sets, be it RAID level 5 or 6. Why do you need to go to this extreme? Availability is the key reason. If you have a RAID set in rebuild the performance is seriously degraded. The percentage of system time spent on rebuilding the RAID set can usually be adjusted. Set the parameter too high and the rebuild takes less time but the operating performance during the time it takes to do the rebuild can be as low as 40% of normal. Set it to a low percentage to maintain performance and the rebuild time becomes very long and you risk losing another second drive during the rebuild which means risking (RAID 6) or losing (RAID 5) your data. Rebuilding 1 TB disks can be a long, slow process often taking days to complete, while you struggle to operate with reduced system performance.

Moving beyond traditional comparisons of disk versus tape, let’s look at some of the less tangible components that add up to the total cost of ownership of a video archive:

- Initial purchase cost
- Operating costs to run the archive, including air conditioning
- Cost of replicated media and disaster recovery

- Transportability costs
- Post warranty replacement costs
- Upgrade costs

Figure 1 shows a typical small video archive containing 3000 hours of DV25 content. The LTO 4 tape-based archive uses XenData video archive software running on a Windows server with 2 TB of disk as a temporary cache. Forty-four LTO 4 digital data tapes are located inside the robotics tape library, together with a second replicated set of data stored on LTO 4 tapes on the shelf. This system is compared to a pure disk based solution. The disk solution uses replicated RAID sets to ensure data availability and non-volatility of the archived content. It still doesn't protect from physical disaster at the site unless the second replicated RAID system is located remotely. That would require a dedicated broadband connection and additional systems that are not covered in this study.

Only brand name systems were researched for this white paper. In all cases, the servers and the disk subsystems are IBM, Dell or HP models. As you can see there is a significant initial cost savings when purchasing the tape based system. The 3000 hour DV25 archive costs $82,000 for the tape-based system compared to $121,000 for the disk-based system. Both systems include the first 3 years of on-site support.

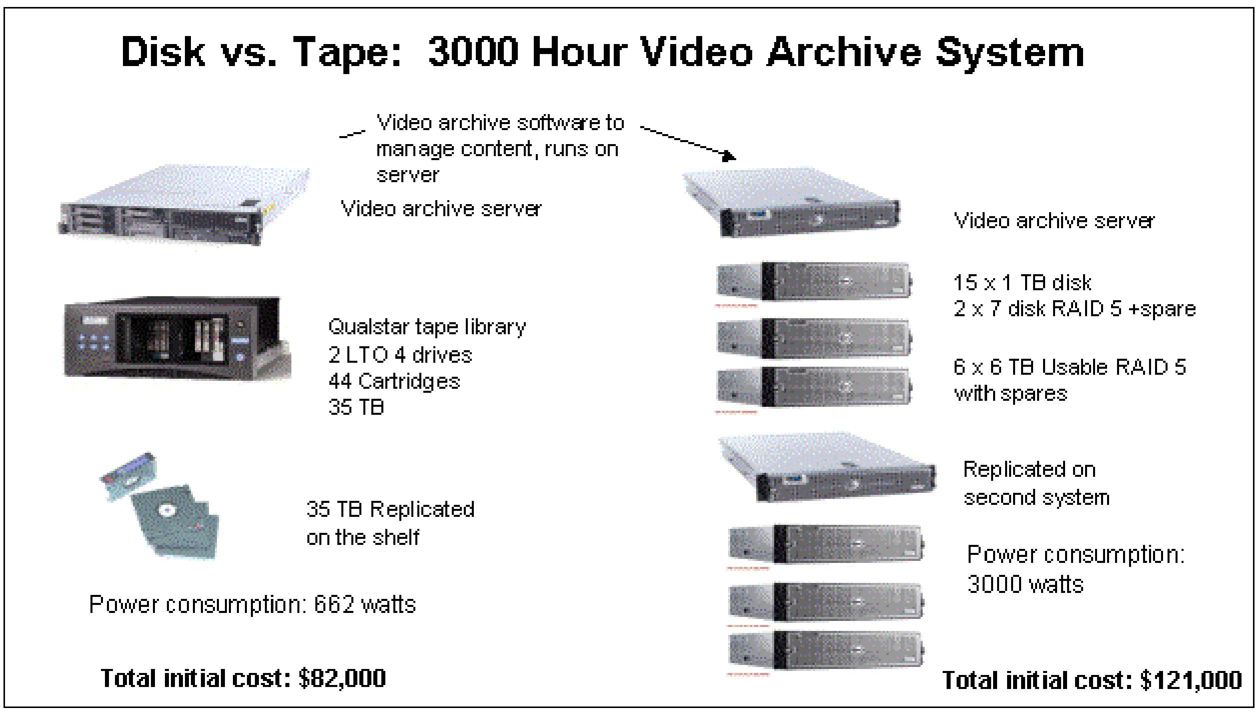


**Figure 1**

Significant operating costs are associated with the electricity needed to run the system and air conditioning to cool the location where the system resides. The tape-based archive consumes a maximum of 662 watts which is mostly from the archive server and the disk drives in the cache. The tape library's consumption is minimal. In contrast, the disk-based system requires 3KW of power to operate. What does this mean in terms of cost? Using California commercial electricity rates the tape-based system costs $811 a year to operate compared to $3,679 for the equivalent disk system..

The cost differences do not stop there. Cooling the systems is a necessity and current methodology assigns a 1:1 factor for dense disk-based rack systems and 0.5:1 for a low density system such as the tape library. So our cost for cooling the tape-based archive is $406 per year versus $3,679 for the disk-based archive. Over five years the cost to operate the tape-based archive is $6,085 and the disk-based system is $36,790, more than six times greater.

Figure 2 compares 9000 hour archive systems. Here the cost savings are even more dramatic: $172,000 for a tape-based archive versus $328k for the disk-based one. Power consumption is an even greater factor with these larger archives. The tape-based unit consumes 816 watts while the disk based system uses 6.5 kW. Adding in the cooling expense raises the annual cost to operate the tape-based archive to $1,500, compared to $15,942 for the disk-based system. Over five years this amounts to $7,500 and $79,710, respectively, making the disk-based system more than 10 times more expensive to operate.

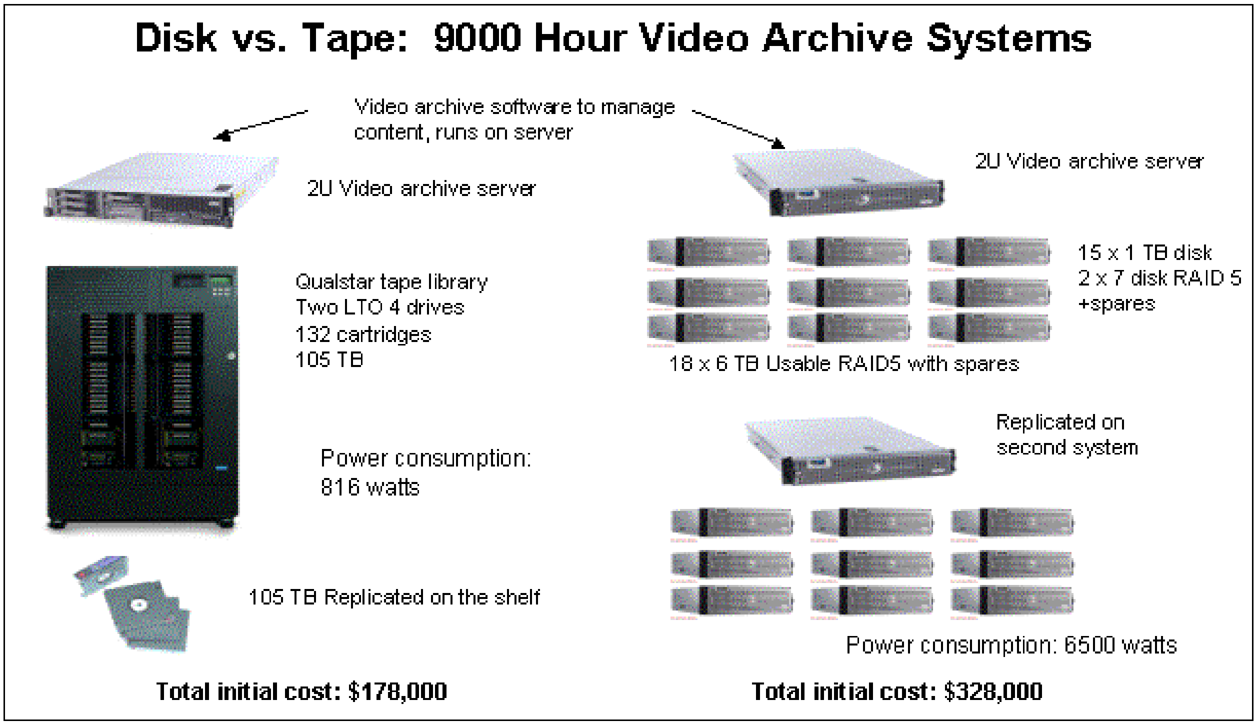


**Figure 2**

With a disk-only system it is extremely expensive to provide multiple copies of the archive's contents. While two copies do reside on the mirrored RAIDs to protect against local hardware failures, there is no provision for disaster protection unless one of the RAID systems is remotely located. The cost of adding a third copy is prohibitive.

Making multiple copies of the archive with a tape-based system is inexpensive, just the cost of the media. A third set of tapes for offsite disaster protection does not add much to the overall cost of the system while providing significant failsafe protection for the video archive investment.

There are additional benefits to being able to create multiple copies on tape. It is easy and inexpensive to transfer video between facilities, particularly group stations and broadcast facilities located in countries that have limited broadband infrastructures. Indeed, several facilities use LTO tapes as a cost effective means to transfer material between remote locations. LTO 4 tape costs less than $1.90 (MSRP) per hour of DV25 content.

Now let's take a look at the cost of post-warranty support. The systems used in the comparisons have the first three years of on-site support built into the initial purchase price, but what happens when we look at years 4 and 5 of our 5 year life cycle costs?

The cost to provide post-warranty support for the 3000 hour systems is $4,375 per year for the tape-based system and $7,500 for the disk-based system, for a total cost of $8,750 and $15,000 respectively. That's more than enough to pay for a third set of media.

The post-warranty support cost comparison for the 9000 hour systems are even more striking. The tape-based system costs $5,115 per year while the disk-based system cost is $19,500. This amounts to $10,230 for the tape-based archive and a whopping $39,000 for the disk-based system for years four and five. The reason that the disk-based systems are so expensive to maintain is that service and support costs are figured on a per-chassis basis and there are eighteen chassis in the hard disk solution.

Disk drives are powered and rotating 24 hours a day. Given that the Average Failure Rate is 0.73% for enterprise class SATA disk drives, and there are 270 disks in the arrays for the 9000 hour archive, two disk drives will likely fail each year during the warranty period. The probability of multiple disk drive failures increases significantly during the post warranty period. There is a much greater danger of complete loss of data on one of the RAID systems due to concurrent drive failures. Fortunately, using mirrored systems with replicated data mitigates this issue, but at a significant cost.

What happens if a tape drive fails? Well, the data is secure unless a tape is damaged in the failed drive, a very remote occurrence, but replicated media solves that problem. Since the video archive system uses multiple tape drives the system can continue to operate while a replacement tape drive is changed out. The replacement tape drive is approximately 10 times the cost of a single replacement disk drive, but remember it isn't running 24 hours a day seven data a week, either. Post-warranty replacement cost will be about equal since multiple disks will be replaced for every tape drive failure.

Upgrading the system is another factor to be considered when selecting a video archive system. Expanding disk-based systems adds a lot more hardware because the mirror must be maintained, incurring proportionally higher operating costs. Tape-based systems can be upgraded much less expensively. LTO tape drives are read/write compatible back one generation and read compatible two generations back. Each new generation typically doubles capacity and transfer rates, so a 3000 hour archive becomes a 6000 hour system by simply replacing the tape drives and media. Upgrade the system a second time and the capacity quadruples while still reading the original media. Because tapes can sit on the shelf, previous generation media can be exported out of the system. If any content is required to be kept near-line then it is a simple matter to transfer that content onto new, higher density media.

Let's look at an example of doubling the capacity of both the 9000 hour systems, perhaps because we are now recording in HD versus DV. The tape-based system will require two new tape drives and 132 pieces of media. Based on previous pricing models the next generation tape cartridges will cost approx 80% of two of the older tape cartridges. Using $210 for a 1.6 TB cartridge and $10,000 per tape drive, the cost to go from a 99TB to 198 TB system is: 2 x $10,000 + 264 x $210 = $75,440. And we still have our existing material stored on tape.

Assuming that disk drive prices fall by 25% while capacity doubles from 1 TB to 2 TB, then to double the disk-based system cost $231,000 if we want to retain our existing data. Similarly, the 3000 hour system will cost $75,750 to upgrade the disk-based system and $38,480 to upgrade the tape-based archive.

In summary, using digital data tape for a video archive is substantially less expensive to purchase, operate, support and upgrade compared to a similar sized disk-based system. Disks are clearly not the best medium for a long term archive. Tables 1 and 2 summarize the life cycle costs of both the 3000 and 9000 hour video archives discussed in this white paper.

| Description | Digital Tape-based Archive | Disk-based Archive |
|---|---|---|
| Initial purchase cost | $82,000 | $121,000 |
| | | |
| Operating power costs per year | $811 | $3,679 |
| Cooling power costs per year | $406 | $3,679 |
| 5 year operating costs | $7,095 | $36,790 |
| | | |
| Service and support costs<br>24 x 7 years 4 & 5 | $8,750 | $15,000 |
| | | |
| Total cost of archive | $96,835 | $172,790 |
| | | |
| Cost to replicate an hour of content | $32 | $58 |
| Cost to double the capacity of the archive | $38,480 | $75,550 |

**Table 1. Typical 3000 hour video archive system life cycle costs**

| Description | Digital Tape-based Archive | Disk-based Archive |
|---|---|---|
| Initial purchase cost | $178,000 | $328,000 |
| | | |
| Operating power costs per year | $1000 | $7,971 |
| Cooling power costs per year | $500 | $7,971 |
| 5 year operating costs | $7,500 | $79,710 |
| | | |
| Service and support costs<br>24 x 7 years 4 & 5 | $10,230 | $39,000 |
| | | |
| Total cost of archive | $195,730 | $446,710 |
| | | |
| Cost to replicate an hour of content | $21.70 | $49.60 |
| Cost to double the capacity of the archive | $75,440 | $231,000 |

**Table 2. Typical 9000 hour video archive system life cycle costs**