



Removable Disk in the Digital Archive: A Cost/Benefit Analysis

Research Note

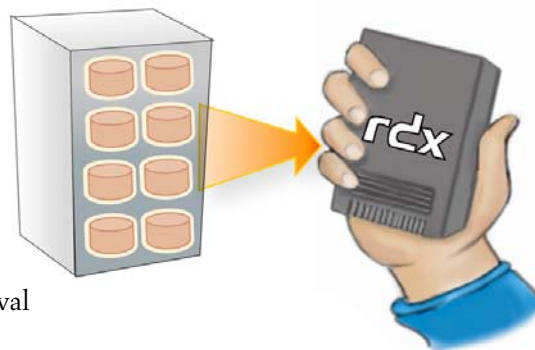
John Webster

October 9, 2007

Licensed to ProStor Systems, Inc. for web posting. Do not reproduce. All opinions and conclusions herein represent the independent perspective of Illuminata and its analysts.

Disk drives that used removable disk media (a.k.a. disk packs) were once popular as an alternative to drives that used a fixed amount of sealed media.¹ At that point in computing history, disk drives were incredibly expensive by today's standards, so IT administrators paid close attention to capacity usage and often used removable disk media to increase their total quantity of storage without adding more drives. Removable media were also used for archival storage as the media could be removed and stored for long time periods; in situations requiring high security, removable media were stored overnight in vaults.

Removable disk media declined for years and finally disappeared from data centers with the advent of the RAID array featuring "inexpensive" capacity and scalability built around fixed disk drives that far exceeded the limits of previous generations of disk. However, we can now see that RAID arrays have their own limitations, particularly when used in archival storage applications.



The need for archival storage has been made particularly acute by regulatory compliance and litigation protection requirements. Disk arrays, even those using inexpensive SATA disks, are a relatively expensive proposition for archival storage. They're also powered and cooled whether or not the data stored inside is being accessed. Furthermore, typical five-year drive lifetimes force at least one replacement cycle, including a complete data migration of archived data sets, to accommodate required retention periods of seven years or more.

Tape is the obvious alternative to disk for archival storage, but it has its own limitations. Tape media are normally refreshed in three-to-four year cycles—a task that forces users to copy data from old to new tapes and that becomes increasingly time-consuming as the archive grows. Tape can also slow a company's response to situations requiring timely and precise data retrieval (litigation support, for example).

So, is it time for removable disk to make a comeback? We think so.

¹ Typically called fixed disk drives or Winchester drives.

The Impact of Removable Disk on Archiving

It is well known that many regulatory agencies (FTC, SEC, DHHS, etc.) now require that certain types of records, including email, be retained for seven years. What is less well known is that a majority of IT administrators now report having to retain data for periods of up to 50 years² for various other reasons. Today, long-term archival storage really has to deliver long-term data availability. The problem this creates for IT administrators is that the costs to maintain and support a long-term digital archive can add up quickly. These costs include:

- Initial expenditure for hardware and software, including initial installation.
- People costs for planning and integration of the hardware/software into the existing IT environment.
- Hardware/software maintenance and support costs required beyond the initial warranty period.³
- Power and cooling for all digital archive hardware components (arrays, tape libraries, networking gear, etc.).
- Hardware/software upgrades while in service.
- Media maintenance and refresh (disk and tape replacement).
- Entire subsystem refresh—i.e. costs to remove the obsolete subsystem and replace it—a process that will include all of the same initial costs noted above that were required for the obsolete subsystem.

Recently, ProStor Systems introduced the first removable disk media designed to replace entry-to-mid-range-level tape cartridges in backup applications. ProStor currently delivers its 2.5" RDX ruggedized, removable disk cartridges in capacities ranging from 40 to 300 GB. To support media longevity requirements, ProStor offers

independent lab results that show that one can expect a 30-year usable archive life from RDX disks—between eight and ten times longer than the tape cartridges in current use. Imation and Tandberg Data have designed RDX removable disk technology into their storage devices. RDX cartridges are also available from multiple system OEMs and channel outlets. In addition, ProStor has now introduced its own InfiniVault archive system based on RDX removable disk, and aimed at the small-medium enterprise (SME) user, described later in this research note.

Using removable disk as an archival storage media can significantly lower the costs of supporting a long term digital archive as compared with the fixed disk and tape cartridge alternatives. Let's consider some of those cost benefits in detail.

Energy and Floorspace

Arrays based on conventional fixed disks consume power and generate heat regardless of whether or not their data is in active use. The data is online and consuming energy resources, the cost of which is escalating at an average rate of 10% per year, whether it needs to be or not. A modest EMC Centera (a SATA disk-based digital archival subsystem) configuration, for example, can consume 9.6 kVA and generate 32,800 Btu/hr. With removable disk, IT administrators can choose which data sets to keep online vs. offline so power and cooling are required only for the disks containing data that *needs* to be immediately available.

Fixed disk subsystems also require increasing amounts of data center floorspace and floor-loading support as they will inevitably scale upward in capacity. For example, a single EMC Centera's cabinet weighs in at three-quarters of a metric ton. With removable disk, offline capacity can be stored outside the data center such that floorspace and loading support are devoted only to online datasets.

Media Maintenance and Refresh Cycles

Using tape in a long-term archival application requires that IT administrators monitor and manage the useful life of the media itself. Tape can

² SNIA Data Management Forum survey of 200+ IT administrators published 7/07.

³ Normally only two to three years for disk and one year for tape hardware.

begin to degrade and lose data after three years. Tape can also become brittle and break if not stored properly. Specialized tape storage designed to mitigate these risks adds to the cost of media maintenance. Fermi Lab for example, whose tape storage requirements scale to the multiple petabyte range, refreshes its tape cartridges every three years by copying data from old tapes to new, and then physically destroying (as in “crushing”) the old media for security reasons.

With removable disk media that remains viable for up to thirty years, the costs associated with tape maintenance and disposal are either eliminated or significantly reduced. These costs include:

- The cost of replacement media required for every refresh cycle—a process that will be required at least once for datasets that must be retained for seven years. The actual cost will be highly variable depending on the storage density of the media and whether the data is compressed.
- The cost of staff and hardware/software resources required to migrate datasets from old to new media. The current average fully-burdened cost of a storage administrator is approximately \$90,000.00 per year. However, SME’s don’t have specialized storage administration staff, so their costs will be a higher fully-burdened cost of \$120,000 per year for a general IT administrator.
- The costs associated with destruction and specialized disposal of the old media. This has become a critical requirement, particularly for unencrypted tapes, because of the highly-publicized leaks of personal data related to tapes that have inadvertently fallen into the hands of data thieves. The financial risk (in terms of legal fees, civil penalties, and loss of good will) associated with leakage of personal data can dwarf all of the other costs by several orders of magnitude.
- The costs associated with maintaining a specialized tape storage environment to retard tape deterioration—likely off-site. These costs include physical storage fees at an average \$6.00 per tape per year, auditing of the tape inventory by an outside firm, and transportation to and from the off-site location

at the nominal rate of \$0.50 per cartridge. Some of these costs could also apply to removable disk stored off-site.

Tape vendors also have a habit of either changing or retiring tape formats from time to time, forcing users to migrate data from tapes created in the old format to new format tapes—and probably having to replace the underlying tape subsystem in the bargain. With removable disk, media format is not an issue, so all costs—hard and soft—associated with a forced migration due to format change go away. This includes administrative costs associated with planning for and integrating the new tape format into the existing archival environment.

Fixed Disk Subsystem Life Extension or Refresh

The useful life expectancy for most fixed disk subsystems is five years. Many users have gotten into the habit of replacing them after three years to avoid escalating maintenance charges that kick-in after an extended warranty period expires. Continuing that practice with fixed disk in archival storage applications will put a significant burden on the archival storage budget. Remember that archived data sets associated with compliance mandates alone will likely be retained for seven years—two years beyond the useful life expectancy of fixed disk subsystems. This means that if fixed disk is used to archive these data sets, IT administrators are forced to make one of two choices:

1. Try to extend the useful life of the fixed disk subsystem to seven years, paying escalating maintenance fees (average 17% of subsystem cost) from years four through seven (assuming that the vendor will still be offering maintenance during years six and/or seven). If the subsystem is leased, then the IT administrator is at the mercy of the lessor when negotiating the lease charges for extending the original lease term through seven years.
2. Replace the entire subsystem sometime during the seven-year dataset retention period, commonly known as a “forklift” upgrade. Subsystem replacement entails the following costs:

- Decommissioning and deinstallation of the old subsystem.
- Possible destruction of the old subsystem for security purposes.
- Cost of the replacement subsystem including associated software charged separately and that may only run the subsystem.
- Transportation, installation, and integration of the new subsystem.
- Possible training on the new subsystem.

Using removable disk media with a life expectancy of up to thirty years does not completely eliminate all of these charges, but the fact that removable disk subsystem life expectancy can be extended to seven years and possibly beyond reduces or eliminates many of them. The costs associated with a forklift upgrade are eliminated. Maintenance costs compared to the fixed disk alternative can be significantly reduced as they will only apply to the subsystem controller and rack, and not to the disk drives. To be fair though, storage of removable disk media will, like storage of tape, engender additional costs that are not applicable to fixed disk.

Justifying the Cost of a Digital Archive with ROI

For the reasons outlined above, we believe that a digital archive built around removable disk will significantly reduce an archive’s total cost of ownership (TCO) as measured over the relevant seven-year life span compared with using available fixed disk and tape alternatives. Indeed, given the fact that IT administrators report more elongated retention requirements, extending a TCO analysis beyond seven years is justifiable and can yield even greater benefit. IT administrators can, we think, rightfully conclude that the cost of a

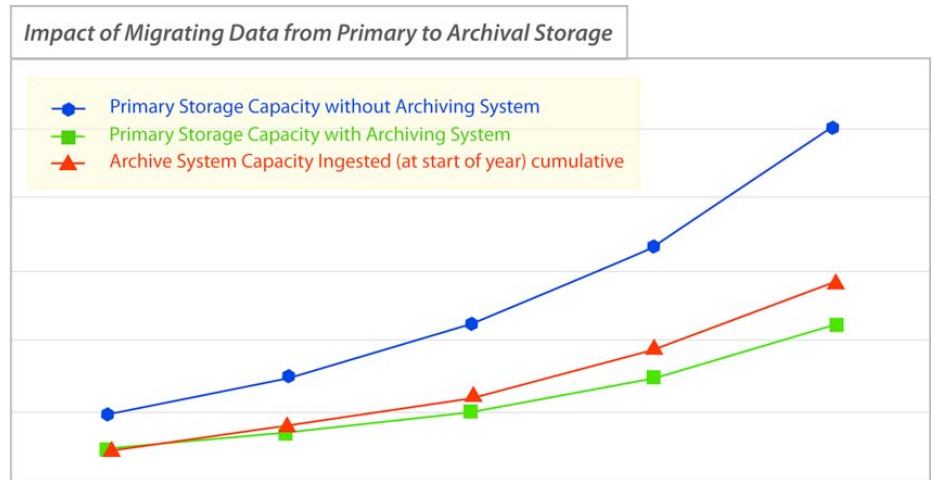
removable disk-based digital archive is within financial reach, even for SME budgets. However, we have yet to mention the ROI benefits of removable media because they apply mainly to instituting a digital archive in the first place, regardless of the underlying type of storage used.

It is well known that funding and managing the underlying infrastructure to support data growth remains the number one storage issue for all IT administrators. Capacity devoted to primary storage grows at an average rate of 70 percent per year. This growth rate continues to burden both capital and operational expense budgets—increasingly so with the escalating cost of power and cooling. Slowing the growth rate of primary storage capacity is therefore highly desirable. Enter the digital archive.

Using software to identify inactive or rarely accessed datasets, IT administrators can migrate these data sets to a digital archive.

Migrating datasets off primary storage results in a number of ROI-related benefits:

- Additional primary storage purchases can be deferred, thereby reducing yearly capital expense and operation expense associated with new storage acquisition and integration.
- Backups of primary storage run more efficiently as they are not backing up the same inactive data sets, over and over.



- Energy costs associated with primary storage are higher on a per GB basis than for archival storage, regardless of the media. Spending the same energy dollars on inactive datasets as on active ones is a waste that can be recaptured by moving slow and inactive datasets off primary storage. Energy savings realized along with the resulting improved ROI are simply amplified by using removable disk.

This is not to say that there are no specific ROI benefits to be realized from the use of removable disk—especially compared to tape. IT departments are routinely tasked to respond to litigation that requires discovery of electronic records and email. If data is stored on removable disk, this process can be reduced to hours as opposed to the days it might take to spin through tapes looking for and extracting the exact files needed. This reduction in labor, relative to using tape, is just one way that removable disk pays for itself.

ProStor Removable Disks

ProStor Systems has been marketing removable disk as an alternative to fixed disk and tape for both archive and data protection applications. ProStor has also recently introduced its InfiniVault Archive Appliance for SME users. The InfiniVault system combines fixed disks and RDX removable disk cartridges, all managed by ProStor's Archive and Compliance Engine software. The InfiniVault has a variety of features aimed at reducing administrative and other costs. These include:

- Rapid discovery and retrieval of archived files with file attribute and content searching when IT administrators are tasked to respond to compliance and legal discovery mandates. The software also supports the assignment of legal holds to specified files.
- File compression and file-level de-duplication software that further expands the ability of administrators to keep archive data on disk without incurring the environmental costs associated with fixed disk arrays.
- Creation of an audit trail of ingested files that relieves administrators from performing this

essential and time-consuming process when done manually.

- Built-in encryption capabilities that mitigate the legal risk of lost disk cartridges.

When both TCO and ROI benefits are considered, InfiniVault will likely put digital archival storage within the financial reach of the SME. A basic InfiniVault configuration that includes 5.2 TB of online storage (2.2TB of fixed active archive and 3TB of RDX cartridges in one cabinet) comes in at less than \$30K.

Conclusion

It's high time for removable disk to make a comeback. Fixed disk is an expensive answer to archival storage requirements. Yes, it's fast, but its speed in archival applications comes at a considerable cost when seven-year file preservation requirements are factored into a TCO analysis. Tape has traditionally been the archival media of choice, but as we have shown, it becomes increasingly costly to support and manage as time goes on and the archival store grows. And, in the increasingly likely event of a legal discovery exercise, spinning through piles of tapes to find the required documents is highly disruptive and puts the organization at risk of noncompliance. Removable disk can deliver the speed-to-information of fixed disk in archival applications without the management overhead of tape.

Like large enterprises, SME users are exposed to regulatory non-compliance and litigation risk and must cope with paying for and managing escalating amounts of primary storage. At an entry cost of less than \$30K for the mid-range system, InfiniVault puts a unified fixed and removable disk archive storage appliance within the reach of the SME IT budget. This lets the ROI benefits of implementing a digital archive be realized immediately—to say nothing of being cheaper to operate for years to come.