



Economical Long-Term Archive For Digital Content

Digital Content Archive: Ideal for MAM & DAM Workflows

Rorke Data's archive solutions work with all types of digital content workflows including:

- Video
- Film
- EMR
- Medical images including DICOM
- Web
- Print and Pre-press
- Document Imaging
- Endoscopic

This paper will overview how our digital content archive can be used with digital video, however our solution can be leveraged in many content workflows and as such, it is a perfect archive solution for MAM and DAM workflows for any vertical market.

DCA for Video

Archiving digital content continues to be a major expense item for organizations responsible for protection of those valuable assets. Changes in technology including high definition and 3-D are driving up the capacity demands which exacerbate an already increasing expense area. Storage technology transitions that occur for some technologies require migration of videos with additional operational expenses for the migration as well as the capital investments for new devices.

A digital content archive can be described simply as archiving video information to a storage system that can retain the video information with integrity for long periods of time. Today's videos exist as a collection of files that can be managed in standard file systems that may be utilized in servers with various types of operating systems. The origin of the video files still comes from cameras or computer generated images. There are multiple paths for how video files from cameras get to their final, protected archiving location. Videos typically are retained for varying periods of time depending on the specific usage situation. Ultimately, videos will be retained for very long periods of time – measured in decades. The computer storage technology in use today for most video archiving is measured in single digit years for viable retention before data migration is required which does not meet the needs for the decades of storage.

In addition to the long-term requirements for storing information on a technology, the video information in files must be protected from disaster, have secure, controlled access, and have verifiable integrity to insure that no data loss or degradation has occurred. While these elements are a given as requirements, the methods to meet them can vary widely, especially if they are not integrated, automatic capabilities of the digital content archiving system. The extra costs of adding these capabilities or even manually performing the required tasks to accomplish them may be prohibitive or not included in the budgetary processes for the organization responsible. The importance of protecting the asset which a digital content represents can be recognized with the selection of the archiving system that addresses the need. Managing digital content archives is accomplished by software applications that control the placement of the video files on the archiving system and the movement of files between editing and broadcast systems to and from the archive. Media asset management software is a special version of digital asset management that is focused on the digital content space. Included with media asset management may be digital rights management for validation of access or control.

- » Content user audit trail reporting
- » Automatic single-instance file feature
- » Removable RDX disk based media for offsite vaulting, preservation and disaster recovery
- » Content Addressable Storage (CAS) with sophisticated hash algorithms is used to track and verify all content – ensures the immutability of the data.
- » Can be configured to make copies of data automatically
- » Powerful AES-256 encryption for RDX removable cartridges – provides data security.
- » Encryption keys are managed internally simplifies management of very secure data.
- » Media is backward/forward compatible
- » Reduces backup costs by 50%, cooling and power saves 75%
- » Improved operational efficiencies
- » Ease of use set and forget GUI
- » Data compression available
- » On-the-fly file level data de-duplication
- » 1TB/cartridge

The archiving system for digital content must work seamlessly with the media asset management software. Prior storage technologies in use did not natively store file-based video data and required middleware and encapsulation software usually in the form of backup software to store data. New technology archiving systems can store data in the native file organization of videos without the need or costs of the middleware or backup software. While this may be resisted greatly by vendors of that software, the overall simplification of the workflow and complexity are a great value to the organizations responsible for the digital content archive.

Sources for Digital Content Archives

Video files to be archived can be from computer generation as in animation but the majority come from cameras through various routes. The ultimate archive location is a video library that is managed with asset management software for decades. The production side video data from cameras usually requires an offload from the digital camera to a technology that allows the videos to be transported to post production. This offload may be to a media that is durable enough to be handled by less-than-careful individuals. The durability and reliability are critically important as the camera-based storage is usually erased for continued usage while the captured data is being transported. In addition to the physical transport, network movement of the digital content files may be utilized as well.

In post-production, the sources and intermediary files may be archived at known points for recovery in the case of some disaster or problem. These archives may be of a transient nature with lifespan measured in months. The protection is just as critical as the long-term protection from the value of the asset the digital content data represents.

The finished videos in their many different forms including languages, editing variations, etc. are the files that are moved to the video library. These represent the complete investment in the production of the videos and are the asset to be used in continued revenue production. In addition to the value for commercial purposes, many of the videos in a video library may have historic significance and the preservation is an industry initiative. Maintaining the integrity in a readily accessible media asset managed copy of the digital content is a requirement on the archiving storage system being used.

Economics of Long-Term Archival

Capital is the money spent, typically at one time, to purchase the technology for storage of the digital contents. If only looking at the initial capital expenditure, the true cost of archiving digital contents is grossly misrepresented. The operational costs of management, power, cooling, etc. are relatively well understood and should be included in the economic considerations. The more important and bigger impact issue is the element of time.

“How long will the video data be retained?” is the first question to be asked to understand the economic implications. The next is “How long will the archiving storage technology be able to store data and be accessed with current technology systems?” A final significant question is “What will it cost in capital and migration expense to move digital videos to a new technology?” Other questions about the reliability of access of data for the transition to the new technology and the availability of the older technology systems or personnel to support them fall in the shadow of the three basic questions.

When looking at the economics for long-term archival, the formula is simplified to an equation that is the expense for a given capacity over a defined period of time which is simply written as:

\$ / TB / year.

In this form, the capital expense for the acquisition of the archiving system, plus the migration expense, plus the operational expenses can be divided by the capacity, usually in terabytes, and that divided by the total number of years of expected archive life. The capital expense must include every technology transition that will occur during the lifespan for the archive. As example, if a tape system is chosen, the technology lifespan before migration is required is no more than seven years. This means that even though the tape media may have a much longer lifespan, the tape technology represented by the manufacture of new tape drives capable of reading the tape media that was written is only available for up to seven years. If the lifespan for the archival life chosen is set at 20 years, this means there will be three technology migrations required (the initial plus two more) and the expenses must include the capital and migration costs for those transitions. The number of years can be any expected number that technologists are comfortable in establishing. Twenty to thirty years have been used as practical numbers with the understanding that underlying systems technology undergo major transitions in that timeframe and will provide opportunities for new methods and capabilities.

In addition to measuring the true cost of long-term archiving by using the time element, the ability to scale as more capacity is required is another factor in the economics of long-term archiving. Scaling can be done in a number of ways but the economics are the most advantageous when existing systems can be expanded as needed, when needed, in increments that minimize the impact. One means to do this is to capitalize on the Research and Development investments being made in storage technologies and be able to seamlessly incorporate the new, denser technologies with no disruption to the existing archiving system or already-archived data. This implies that as new developments are made which include larger capacity devices, they can be incorporated directly into the archiving system and not have to wait until another major technology transition is undertaken.

Addressing Digital Video Archiving Needs

Two technologies exist that can address the digital video archiving needs and provide the lowest-cost long-term archiving economics. One is the RDX removable disk cartridge. RDX has been in market since 2005 and has been in wide-spread usage in many markets. Developed by ProStor Systems, the RDX is available from multiple manufacturers and is sold by many different vendors including Dell, IBM, HP, Fujitsu, Imation, and Tandberg.

The other technology is the ProStor InfiniVault digital archiving system that delivers a complete archiving system with built in automation around self-protection, integrity verification, and data management. The ProStor InfiniVault uses the RDX disk cartridges for both online storage for long-term retention of data and protected, removable copies.

RDX

ProStor's RDX removable disk technology is designed to provide a unique level of protection for valuable information over an extended period of time. An RDX system consists of one or more rugged, removable disk cartridges and either an individual dock or multi-cartridge enclosure, the latter associated with the ProStor InfiniVault archival system. When the RDX cartridge is inserted into the dock or enclosure, it functions as a normal disk drive. Yet it can also be used and removed for off-line storage just like a traditional tape or optical drive with backups or for archival data. The media, which includes a 2.5-inch mobile disk drive housed in a shock-proof casing, can be safely archived for 30 years. The newest standalone devices have the USB-3 interface on the RDX dock and support 150MB/s transfer rates.



RDX's high-performance, enterprise-class data protection and low cost meet the reliability and business requirements for all data archive and backup applications. RDX technology capitalizes on the best characteristics of both tape and disk:

- Tape—removability, portability, and low cost
- Disk—reliability, simplicity, fast data transfer rates and random access

Replacement for Tape or Optical Media

The RDX media is randomly accessed so stored files can be immediately retrieved without the time-consuming sequential seeks of tape technology. Single files and selected volumes can be restored in seconds instead of hours.

RDX Unique Economic Value Proposition for Long-Term Data

RDX offers disk technology with an unmatched economic value proposition – a productive life-span that can be amortized and exploited across 10 to 30 years. Traditional disk storage systems become obsolete and are generally provisioned in advance and then replaced after 3-5 years, resulting in a regular cycle of large capital outlays. Similarly, highly complex tape mechanisms give tape media and storage solutions a short life expectancy and lead to frequent technology upgrades. In both cases, upgrading to the next generation requires the complete purchase of new drives and media. In contrast, the mechanical simplicity of a multi-generational RDX system results in a productive lifespan that is 10 times longer than the most reliable tape or optical drives. And because the RDX cartridge houses a completely self-contained disk drive, any advancement in disk capacity technology will be immediately usable by the system. This means that tomorrow's higher capacity RDX cartridges will operate in today's RDX systems. In fact, RDX users don't need to provision all their storage capacity upfront, but can rather provision as new capacity is needed taking advantage of the very latest pricing, capacity, density, and performance of the most current SATA disk technology. Follow-on RDX technology will seamlessly introduce Solid State Disks into the same cartridge case and will be forward and backwards compatible with all existing RDX systems.

Removable and Durable for a 30-Year Lifespan

The removable RDX media is uniquely designed for portability with its highly ruggedized cartridge. The ruggedized, protective, and shock-proof RDX cartridge meets industry-standard drop test requirements of one meter (39 inches) onto a tiled floor without damage.

The RDX removable disk drive is a perfect choice to be used as a camera offload device used for transport and as an interchange device for sending large amounts of data to other locations on a random access, high-speed medium.

Removable and Durable for a 30-Year Lifespan cont.

- **Offload Copies**
 - **Protections - multiple copies**
 - **Transfer to post**
 - **Free up camera resource**
- **Sources**
 - **P2 cards**
 - **CF cards**
 - **Direct USB**
- **Durable, reusable, long-life RDX disk cartridges**



RDX as Camera Offload / Interchange Device

ProStor InfiniVault

The ProStor InfiniVault is a self-managing, comprehensive and scalable archiving storage system that exploits the capabilities and economics of RDX for managing long-term data. InfiniVault is a particularly natural fit for the needs of the digital content archiving space because of its:

- Ease of use
- Economical packaging
- Automation of protection and integrity verification functionality
- Scalability to address growing capacity demands

InfiniVault is an ideal solution for replacing tape and optical-based systems to reduce costs and simplify operations. It also provides significantly greater functionality than most tape, optical, and fixed disk-based archival systems thanks to superior data management and security features and its ability to span multiple locations with both network data transfer and the use of removable media.

Customers can select from four InfiniVault models targeted at typical storage capacity requirements from a few TB to over 100TB in base systems and the ability to federate systems into petabyte capacity requirements. The InfiniVault system is attached to the network and folders are set up at deployment to match automated long-term data management rules to particular classes of data. As data is transferred into the system, it always appears available for access. Behind the scenes, data is authenticated using digital signatures, secured through access controls and optionally selectable encryption, and copied to alternate locations where

it can be stored either on- or offline. In the event of a failure or disaster, the system can automatically failover to an alternate replicated InfiniVault system or can be restored from RDX media. The system includes a comprehensive set of auditing functions to provide a complete audit trail for every digital content file.

Highlighted InfiniVault Automated Data Management Features

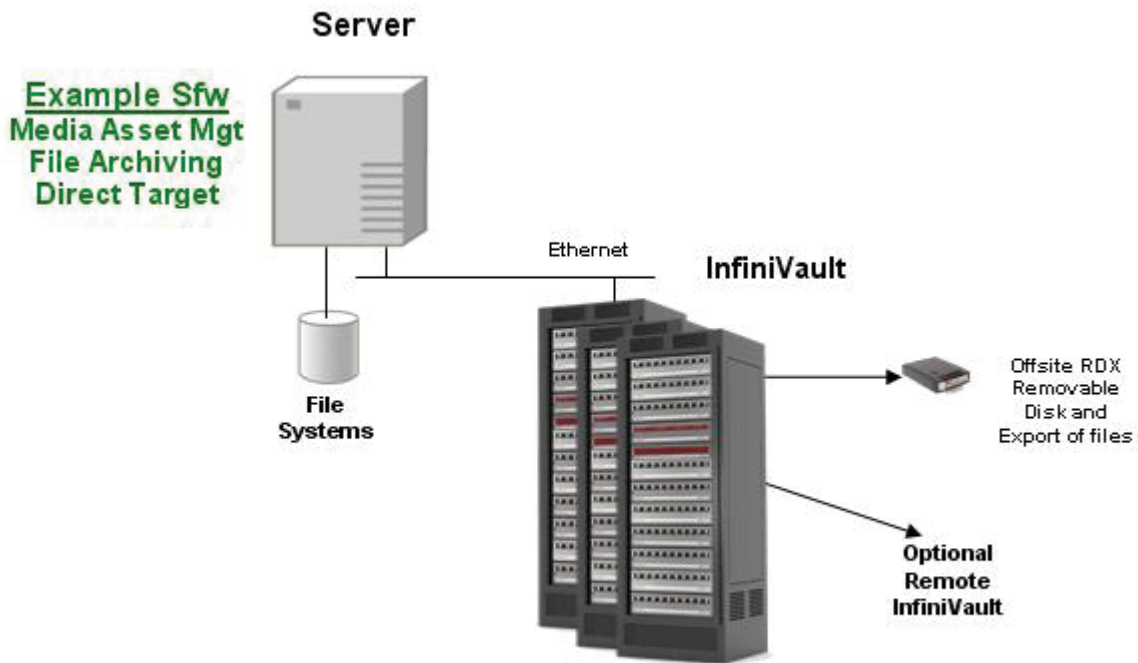
Multiple Data Copies	InfiniVault allows the user to make up to four copies of data within each system, with each copy being stored on a separate removable RDX cartridge. This is essentially an immediate backup copy – removable by virtue of being on RDX -- that eliminates the need to routinely backup the InfiniVault or copies of the data on primary storage.
Selectable Data Encryption	InfiniVault has a built-in AES 256-bit encryption engine that allows for secure data storage, in the event a cartridge is lost or stolen. The encryption can be selectively enabled on a per-vault basis
Built-in Replication engine	Each InfiniVault can be automatically replicated to secondary InfiniVault(s) located across the network, supporting both disaster recovery and continuous data access strategies.
Active Directory integration	InfiniVault integrates seamlessly into Active Directory environments to allow for easy security management. Administrators can use Access Control Lists (ACLs) to grant individual users, groups and guests to access the data storage system.
Seamless infrastructure integration	As a network storage device, InfiniVault presents CIFS / NFS interfaces to allow any application to write data to InfiniVault

Highlighted InfiniVault Automated Data Management Features cont.

Selectable WORM storage	The InfiniVault can use industry standard RDX cartridges in a Write Once Read Many fashion. This means that while the retention period for a given dataset is active, the data cannot be altered. Once the retention period expires, the RDX cartridge can be reused again to store more data. The re-usable nature of RDX cartridges makes InfiniVault more economically attractive for storage projects than tape or optical storage systems.
Low power consumption	On a per Gigabyte scale, InfiniVault uses significantly less power than any other storage array on the market. With the automated "spin-down" of spinning media when not accessed, the removable nature of the RDX cartridges and the robust media management application built-in to the InfiniVault, customers can easily use between 30% - 85% less power than traditional disk storage systems.
RDX Cartridge Compatibility	Since RDX cartridges were first introduced in 40GB capacities to the now current 1 TB capacity to the future 1.5 TB and 2 TB cartridge capacities, InfiniVault is read / write compatible with all of the RDX cartridge capacities.
Long term media shelf life	Independent lab results indicate that data stored on an RDX cartridge has a high probability of retrieval 30+ years into the future. Traditional disk systems average around 6 years of life.

Physical separation of data classes	With InfiniVault, customers can define "storage vaults" which are physically separated units of storage from other "storage vaults" This means that data will never be co-mingled unless a person desires such an end result. Physical separation of data provides customers the ultimate assurance their data is both safely protected and secure.
Standardized Administrator Interface	InfiniVault provides administrators the ability to simplify support and operational procedures across all locations.
Standardized media	Use of RDX helps users regain control over their media purchasing capabilities.

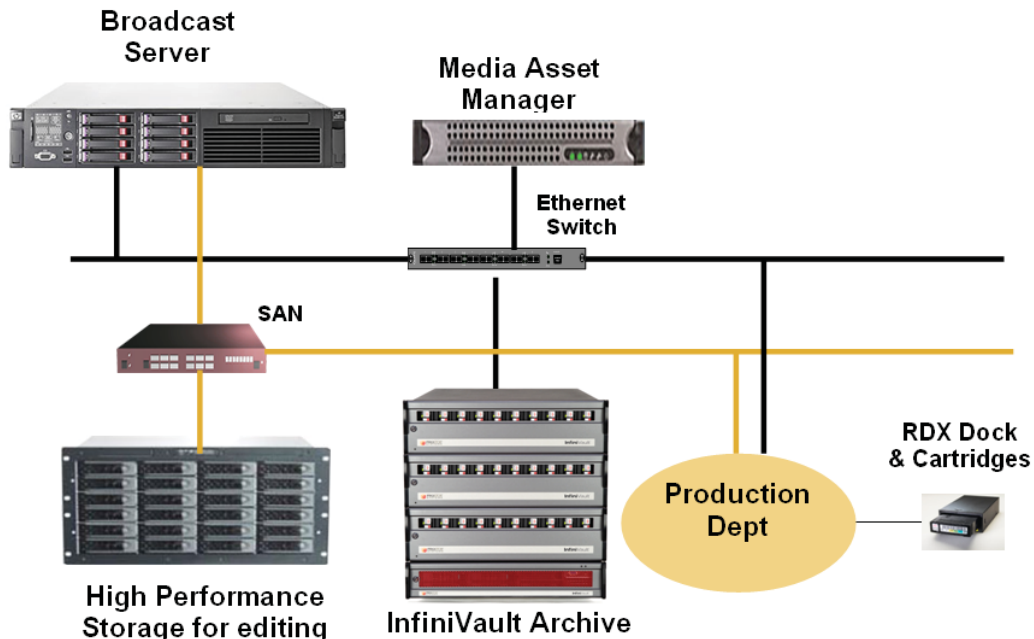
The ProStor InfiniVault seamlessly works directly with media asset management software and integrated system software such as Final Cut. There is no need for middleware or conversion software as InfiniVault directly stores the video file data sent. InfiniVault appears as an infinite capacity Network Attached Storage system with the integrated, automatic management of multiple copies and remote replication for disaster protection, security mechanisms, and scalability.



ProStor InfiniVault Archiving System

In addition to the ability to make multiple copies on RDX disk cartridges and remote replication, the InfiniVault has a combined three layer data integrity verification capability. For every file sent to InfiniVault, a hash algorithm is performed to produce a digital fingerprint. The digital fingerprint is stored in the InfiniVault's internal control information which contains metadata for every file. As the file is written to RDX disk cartridges, a Cyclical Redundancy Check algorithm is applied to every 32K segment of data and the CRC result is also written to the RDX disk cartridge. On the RDX disk, every disk drive sector has an Error Correcting Code generated so that any single bit errors can be corrected as they are found using a scan operation. When files are retrieved from InfiniVault, the ECC will be applied, the CRC will be calculated and compared against the stored value and the hash algorithm will be applied to the newly read data and compared against the original digital fingerprint. All these actions will verify that the data is exactly the same as written initially.

InfiniVault and RDX can be used in the automation of the workflows in the digital content industry. The following diagram is just one example of using the ProStor InfiniVault in a digital content environment.



InfiniVault and RDX in Example Workflow

Summary

The RDX disk cartridge and the ProStor InfiniVault archiving system represent new technologies that have been successfully utilized in areas such as healthcare and traditional information technology. Characteristics and ongoing developments with these technologies enhance their applicability in the digital content archive area. When the effects of long-term retention of digital contents are considered, no other technology comes close to the economics and usability of the InfiniVault with RDX.

From an economics standpoint:

- Long-term storage as \$/TB/year measure
- > 20 years without migration
- Capacity scaling with the next larger disk devices with same characteristics
- Seamless introduction of Solid State Disks

From the usability standpoint:

- Direct file archiving – no transformations, middleware, or backup software required
- Complete inventory of file data with audit trails of access
- Automated data protection
- Multi-level integrity verification

With the growth in capacity demand and the forced technology transitions that occur with existing systems, the use of RDX and the ProStor InfiniVault are the best answers to digital content archiving needs now and into the future.



/ An Avnet Company /

» www.rorke.com