







Contents

1.	The need for archiving
2.	Archival storage considerations
3.	The smartDAX On-Line Archive Approach
4.	Comparisons
5.	Conclusions13

1. The need for archiving

Business analysts' state that most companies continue to live in the Digital Dark Ages as they fail to understand why having a digital archive is essential for their business. As a result of this, they risk loosing corporate data every potentially causing both productivity issues and business continuity problems. Is the current situation really that worrying and why should each professional company consider putting a professional digital archive into place?

1.1 Data protection & availability

As the economy is evolving from being industrial-based to information-based, almost all information companies generate, has a specific value. Nearly 90% of this information is also produced in electronic format, where e-mail is occupying a significant part of this. Simply said, your core business depends on the content and organization of your digital data. Contracts and agreements do not longer exist in physical format, only in electronic format. What about your client and supplier database? Do you have a printed version of this? Are you or any of your employees able to reproduce this information in case of retrieval failure? Data protection entails all these incidental disruptions, but it also includes how to deal with intended violations on your mission-critical data whereby internal and external harmful influences are taken into account.

Let's talk about data availability. Productivity rates leap as organizations streamline their business processes in order to eliminate duplication and to automate the predictable, repeatable steps. Furthermore, ensuring data availability in specific formats or retrieving data within determined time spans is also required by law in certain industry sectors. Do you know how much time is spent weekly on looking for digital information and what the hidden, associated costs are? See table 1. Email, as most information workers will agree, is by far the most time-consuming activity, followed by creating documents and then finding and analyzing information.

Task	Average hours per worker	Cost per worker	Cost per worker
	per week	per week*	per year*
E-mail: read & answer	14.5	418.3	21,752.9
Create documents	13.3	333.7	19,952.7
Analyze information	9.6	277.0	14,401.9
Search	9.5	274.1	14.251.9
Edit/review	8.8	253.9	13,201.8
Gather information for documents	8.3	240.0	12,481.7
File and organize documents	6.8	196.2	10,201.4
Create presentations	6.7	193.3	10,051.3
Create images	5.6	162.7	8,461.1
Data entry to e-forms	5.6	162.4	8,446.1
Manage document approval	4.3	124.1	6,450.9
Publish to web	4.2	121.2	6,300.8
Manage document routing	4.0	115.4	6,000.8
Publish to other channels	3.9	112.5	5,850.8
Create rich media	2.8	80.8	4,200.6
Translate	1.0	29.7	1,545.2

 Table 1: The Cost of Information tasks to the enterprise

Source IDC: Proving the value of content technologies study.

*Based on average salary of \$60,000 per year plus benefits.



Task	Average weekly hours	Cost per worker per week (\$)	Cost per worker per year (\$)	Annual cost to enterprise with 1,000 information workers
Reformatting from multiple	3.8	110	5,701	5,700,760
formats into one document				
format.				
Search but not find	3.5	101	5,251	5,250,700
Recreating content	3	87	4,501	4,500,600
Multi-channel publishing with multiple applications	2.8	81	4,201	4,200,560
Moving documents from one	2.4	69	3,600	3,600,480
format to another				
Acquiring archived records with	2.3	66	3,450	3,450,460
little or no automation				
Version control issues	2.2	63	3,300	3,300,440

In table 2 you can see how many hours are wasted per week per task.

Source IDC: Proving the value of content technologies study.

* Based on average salary of \$60,000 per year plus benefits.

1.2 Increasing storage demands

As the volume of electronic information continues to grow rapidly every year, the amount of new storage capacity installed follows this trend. Email and email attachments are largely responsible for the explosion in data storage needs. IT departments do not develop in the same growth rate however and are ever more faced with limited resources and budget for managing all the (new) storage solutions.

Consequently, IT managers must become more efficient. A first step in the right direction would be, increasing the awareness of alternatives. As the hard disk prices continue to decline, many companies have resigned themselves to simply adding servers in stead of putting policies into place for a more effective utilization of existing storage.

Moreover, 80% of the data on most networks is inactive and as it all resides on hard disk it is a waste of valuable online storage capacity. This includes a high level of redundancy as (duplicate) data is often stored and backed-up several times.

Corporations have been backing up multiple copies of identical documents and can radically reduce their storage (management) costs with easy, cost-efficient solutions for the automatic backup and centralized management of their user data. This will allow organizations to reduce the amount of file and email servers by creating shared pools of common files and backing up only the changes to existing documents.

1.3 Compliance & legislation

Electronic records play as evidence in legal disputes and regulatory audits an increasingly important role. Several industry sectors are tied down by specific compliance and legislation standards which complicate the already challenging storage management a company faces even more.

Table 3 provides an overview of key laws. Some requirements have to be satisfied by all companies, although these requirements can be stricter for specific industries such as financial services, healthcare etc. Mostly they involve some or all of the following requisites when creating compliant digital archives:

- Storage media must be non-rewritable and non-erasable to demonstrate the authenticity of a document.
- Duplicate copies of (all) data must be made.
- An index of all data must be made
- Data and index must be downloadable
- Data needs to be retained for minimum periods.
- Data must be kept confidential
- Data must be kept safe and accessible
- Data must be readily accessible (upon request of regulator/auditor).

Table 3: Summary of key laws

Law	Sector	Territory
Basel II	Banking	Europe/global
Can Spam Act 2003	All doing US business	US
Civil Contingencies Act 2004	Public sector Infrastructure providers	UK
Data Protection Act 1998	All	Country specific
Environmental Information regulations	All	UK/Europe
European Privacy Directive	All	EU
Financial Services and Market Act 2000	Financial Services	
Freedom of Information 2000	Public	UK
IFRS - International Financial Reporting Standards	All	International
Regulation of Investigatory Powers Act (RIPA)	Mainly service providers	UK
Sarbanes-Oxley (SOX)	Listed companies	US/Global
SB1386	All doing business in US	US CA/global
SEC Rule 17a-4	Financial services	US

2. Archival storage considerations

2.1 Archive vs. Back-up

As mentioned earlier, alternatives to traditional storage have to be looked at when searching for answers to solve the storage dilemma. Consolidation and transformation are key factors in this discussion.

Many times, companies are aware of the need to back-up their data but not all are conscious of the benefits of having a digital archive. There is a also a lot of confusion between these entirely different concepts.

Back-up is the process where an additional copy of files is made for a short period of time. The main purpose is to cope with hardware failures or incidental loss of the original data.

Archiving is the process where a permanent copy of data is made. The copied data is intended to be used in the future reference and/or the primary source of that data. In many cases, the original data is being deleted from the primary storage.

So, digital archiving is intended for the long-term storage, preservation and access to information. By consolidating traditional storage tiers you can create cost-effective and user-friendly digital archives which serve as your long-term storage. This sounds complex, but actually involves the transformation of current and future storage repositories into active digital archives.

2.2 Digital archiving benefits and considerations

By creating a centralized digital archive you will obtain the following benefits:

- Improved data protection
- Continuous data availability (tracking, retrieval and audit)
- Legal compliancy
- Increased storage capacity
- Simplified storage management and administration
- Reduced backup & recovery time
- Improved network (server) performance
- Increased user and administrator productivity
- Decrease in subsequent storage purchase & costs of maintenance
- Increase in profits

There are several technologies and products available to enable the transformation from passive storage repositories into active digital archives. In chapter five more detailed comparisons will be performed.

However, there are several key considerations that have to be taken into account, when selecting an archival storage solution:

Reliability / complete solution.

Is your corporate data at all times protected and available? Do you acquire a total solution (software, hardware, support) from one supplier? What about maintenance & support possibilities?

User-friendly.

Do you need complicated implementation plans or is it easy to install and use by both network administrators and end-users? Do end-users need to do things differently or do they continue to work as always while a digital archive operates in the background?

Open standard.

Do you want to incrust your data in proprietary vendor-formats or would you rather utilise open standards to minimize risks?

Scalability & future proof.

How flexible is your solution? Can it be adjusted easily when the companies' requirements change? What is the technology and media life span and how often do you need to migrate data?

Costs.

What are expected the costs in the short and long run? Does it deliver immediate benefits by automating tasks? What are the associated costs of administration and maintenance? What are the media costs?

Naturally, there are many more considerations, but some others could be:

- All types of data can be archived
- Easy integration into your current network
- Possibility to automate manual tasks
- Compliant with current legislation
- Availability of remote management tools
- Etc...

smartdax

3. The smartDAX On-Line Archive Approach

3.1 Tiered storage concept

Tiered archival storage is the single solution for unstructured data growth as no particular archival storage technology can completely address all of the internal and external compliance requirements that organizations nowadays face.

80-20%: Tests prove that 80% of historical data is accessed infrequently but regularly. This means in most companies that 80% of primary storage, back-up and disaster recovery requirements could be relieved.

How? DAX Archiving Software simply adds a third tier to your current storage capacity. Complete (software and hardware) solutions ensure in tier 3 virtually unlimited terabytes of archival storage space while ensuring the optimum combination of RAID and Blu-ray/DVD Optical Technologies.

Tier 1: Recent data is retained on hard disk for rapid access in read-only format, while at the same time this archived data is copied to low-cost and removable DVD/BD media in tier 3.

Tier 2: Only recent data of tier 1 has to be backedup in tier 2.

Tier 3: In the permanent optical archive all your data will be stored safely and preserved for future usage. WORM (write once, read many) technology fulfills compliance requirements and protects against data losses caused by viruses, hack attacks and human error.

Immediate Gains:

- Instant access to read-only, on-line hard disk archive.
- Permanent access to optical archive.
- Optical archive can grow without limits.
- Plug-and-play network-attached solutions
- Optimum combination of open standard RAID and Blu-ray/DVD Optical Technologies.
- Seamless integration into any network.
- WORM technology makes data loss impossible
- Improved network performance.
- More cost-efficient storage management.
- Blu-ray optical technology offers massive storage capacity: 25-50GB per Blu-ray Disc.



smart dax

3.2 Components smartDAX On-Line Archive

The smartDAX On-Line Archive consists of:

1) DAX Archiving Software; Archive management software for creating a virtual digital archive and efficiently managing your archival storage applications.

2) Complete Windows XP System for building a read-only, on-line hard disk archive providing immediate access to the archive.

3) Professional Disc Publisher for creating a permanent optical archive by storing data on BD/DVD media, ensuring non-erasable, permanent and infinite archival storage.



3.3 How does it work?

The smartDAX Technology is strongly entrenched in our archive storage management software. DAX Archiving Software offers a user-friendly software interface and it allows network administrators to centrally manage their archival storage applications.

It works with the e**X**tended **F**ile **S**ystems (XFS) technology which groups any number of BD/DVD media together to visualize them as one virtual archive. Users see the created digital archive simply as another shared file system (F:\), from which they read and write files to as usual.

Files are always visible in the directory cache of the virtual archive (F:/) but can reside on hard disk, on Blu-ray Disc (BD)/DVD or on both. When a file is placed into the archive, our software will check it's presence in the (hard disk) file cache. If it's a new file, it will be written to the file cache and later burned to BD/DVD media.

Professional disc publishers of various suppliers (check daxarchiving.com) can be used to automate the archiving and (full-color) labeling of BD/DVD media.

Burning to disc depends on the configurable high and low water marks in the cache. When the cache reaches the high water mark, it will automatically purge the least used files first to make room for more new files to come into the system. The files then, depending on how often they are accessed, may be in the hard drive cache for immediate retrieval.

If the requested files are no longer in the hard drive cache, a message will be generated informing the user and the system manager with the label information of the BD/DVD containing the requested files. This requires manual intervention to reinsert the BD/DVD into any BD/DVD reader.







3.4 Unique features smartDAX On-Line Archive

The smartDAX On-Line Archive is an intelligent optical disc archiving system, providing best-in-class archival storage by utilising open standard technologies. Advanced cache management ensures quick on-line access. Low-cost and true WORM media guarantee compliance and portability. Flexibility and scalability are necessary to meet changing business requirements and maximize return on investment.

User-friendly, transparent solution;

The smartDAX On-Line Archive concept offers a network-attached digital archive as if it were a hard disk with a file system based interface. This makes porting applications to the archive extremely easy. In addition, the flexible setup features allow performance tuning based on standard available components.

Scalability and extendibility;

The on-line archive capacity is completely scalable (1TB for example) and extendible in the future. The off-line BD/DVD archive can grow infinitely and provides a permanent, WORM-compliant archive.

Future proof, cost-effective investment;

As the smartDAX On-Line Archive Concept is based on of-the-shelf RAID and BD/DVD Technology, all your files are stored according to market standards. This enhances rapid adjustments to innovation but also assures a productive approach to archiving. Furthermore you have the assurance that your data will always be able available.

End-to-end solution;

The smartDAX On-Line Archive solution is a complete solution including software, hardware and support. Even though product collaborations are made, it requires no extra integration efforts from these parties. DAX Archiving Solutions will provide personal assistance in setup, training and support issues.

smartdax

4. Comparisons

4.1 Archiving Technology

As mentioned earlier, the tiered storage concept is the basis for choosing RAID and Optical Blu-ray/DVD Technology as the optimal combined digital archiving solution. Below the advantages and disadvantages of the current technologies are explained.

Tape systems have been used traditionally for back-ups. Tape is however not suitable for long-term archiving. Tape is a magnetic media and can easily be damaged by electromagnetic radiation. In addition, random and fast access to data on tape is impossible as the drive needs to rewind the tape first. Tape is in most cases rewritable and as a consequence not applicable in most professional environments where regulatory WORM (write once, read many) archiving is required. Compatibility is also an issue for tape users as every tape format is proprietary. As the tape standards change approximately every 5 years, this reduces extremely its efficiency as data has to be migrated to new standards and exchange between different suppliers could be complicated.

RAID systems are extremely suitable for archiving on-line data as it provides immediate and random access. WORM compliant archiving is however not a standard option in every hard disk which makes extra investments in specific software necessary. Even though initial costs of hard disks continue to decrease, they still have an average life span of a couple of years. This demands frequent data migration which does not benefit the overall maintenance and management costs of storage. These systems are also more susceptible to the harmful influences of viruses, hackers and possible data loss caused by human error.

UDO looks promising at a first glance, it is however a proprietary, expensive medium which is not widely accepted yet and offered by one single supplier. This restricts extremely its utilization and efficiency as an optimal archive solution.

Optical Blu-ray Disc (BD)/DVD Technology offers random access to any file in the digital archive. It is a low cost, standardized open solution as the presence of DVD technology is widely used in the industrialized business environment. Additionally, optical media has industry standard formats; ISO 9660 and UDF (Universal Disk Format) which are supported on every major operating system today including Windows, Linux, UNIX and MAC OS X natively without any additional client software. Optical media is also less vulnerable than magnetic media as it is less susceptible to environmental conditions. It also provides a true WORM recording medium which is essential considering today's requirements for multi-year data retention periods and growing regulatory compliancy requirements for unalterable, non-erasable storage capabilities. Optical technology meets these requirements as part of its standard features.

Blu-ray is the next-generation optical disc format. With 25GB on a single-layer disc, 50GB on a dual-layer disc and even greater capacities in the future, the storage capacity has expanded enormously while maintaining all other benefits of optical technology. It is still a standardized media which will replace the DVD media within the coming years. Thanks to advanced coating technologies, an average media life of >100 years (under normal usage conditions) is predicted by most suppliers.

	Таре	UDO	RAID	BD/DVD
Regulatory Compliance (WORM)	X	\checkmark	X	\checkmark
Random Access	X	\checkmark	\checkmark	\checkmark
Durability Media	X	\checkmark	X	\checkmark
Removable Media	\checkmark	\checkmark	X	\checkmark
Open Standard	X	X	\checkmark	\checkmark
Cost-Effective	\checkmark	X	\checkmark	\checkmark

4.2 On-Line Archive vs. Jukebox Archive

The traditional solution to an optical archive is a jukebox. Jukeboxes mostly include integrated servers providing a relatively small cache to the system. Hence most of the data will be kept near-line in the media libraries or on the shelf for off-line retrieval.

The tables below summarize the key differences:

Technical comparison				
Aspect	On-Line Archive w/5TB storage	100 slot Jukebox		
Storage space	5 TB + unlimited off-line space	5TB		
Retrieval speed	On-Line. Immediate from hard disk	Near-line. Slower due to jukebox robotics and DVD drive speed.		
Automatic labeling of media	yes	no		
Automatic duplicates for disaster recovery	yes	no		

Commercial comparison				
Aspect	On-Line Archive w/5TB storage	100 slot BD jukebox		
Components	PC w/ 5 TB storage Duplicator DAXfs software	PC Jukebox Jukebox mgt sw.		
Cost	+ (less)	-		
Extendibility	Yes, simply increase hard disk and software license	no		
Cost of support	Retrieval of data from archive is not dependant on jukebox mechanics.			

The smartDAX On-Line Archive also provides:

- **Retrieval speed** of all files at the speed of a network drive. So all archived files are now kept on-line. Their retrieval doesn't require any robot to put the disk in a drive etc. All data simply remains in the cache. If the data exceeds the total space, it is automatically purged off-line.
- **Retrieval of off-line files** is easy and doesn't interrupt the operation of the system for other users. The procedure to bring back off-line data is done via the DVD drive in the server.
- **Easy identification** of off-line media. Data is burned to BD/DVD. The media is automatically labeled by the printer embedded in the disc publisher simplifying the identification of off-line media.

5. Conclusions

From several points of views, the smartDAX On-Live Archive represents the most cost-effective, scalable and user-friendly digital archiving solution.

It satisfies the main drivers for digital archiving. Data availability is no longer a problem as it offers immediate access to the on-line hard disk archive. Protection and compliance issues are met by the read-only hard disk archive and the true write-once nature of blu-ray and DVD technology.

Acquisition and management costs are greatly reduced by utilizing open standard technologies. This also makes it very user-friendly towards both end-users and network administrators.

Moreover, it is a very scalable and future-proof investment. You can scale your on-line hard disk archive to any size (at any moment) and it includes a permanent optical archive which can grow infinitely.