archive by DISC

WHITEPAPER
Sustainable Archival Storage

"The Benefits of Optical Archiving"

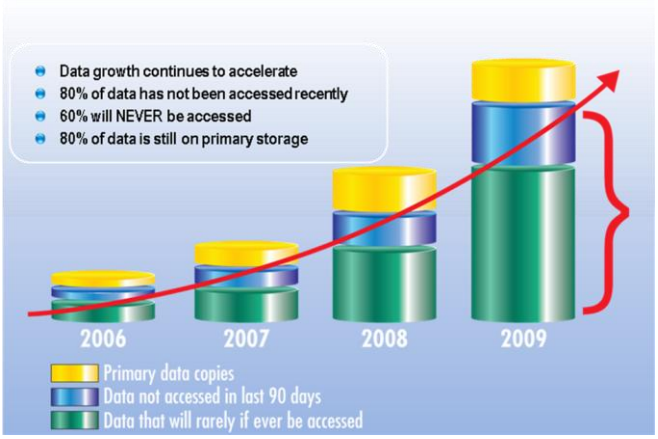
www.disc-group.com

Contents

1. Executive Summary

Archiving is simply the active management of information. As information gets older its value to the organisation reduces and at some point it becomes more of a burden than a benefit.

The timeframe for this depends on the information itself but numerous studies have shown that only 40% of digital information once created is ever viewed and only 20% is altered.



The difficulty is identifying the data that will never be accessed again, if we could do that we could erase it and save a lot of time and money. Unfortunately most organisations can't predict the information they won't need in the future and hence prefer to retain everything. As a result the amount of digital storage continues to grow exponentially which creates financial and logistically issues for organisations as their demand for digital data storage increases.

Archiving alleviates this problem by identifying the older or little used information and manages it separately from the organisations active information. This archive information can then be managed on long term storage. The main benefits of this are:

- Reduce the risk
- Reduce the cost
- Increase the availability of information.

There are three main reasons organisations archive data:

1) Records retention, for compliance and good corporate governance
2) Long term retention, for preservation of information assets
3) IT Optimisation

Based on the reasons why organisations archive information, the fundamental requirements of archiving are:

- Lower cost than existing primary storage
- Long term retention
- Compliance with regulations
- Transparent access to archival information
- Offline management of very old information

The currently available options for archival storage are:

- Nearline RAID arrays

Nearline RAID arrays are basically cheaper versions of the primary RAID array. They are typically created with lower cost magnetic disk drives with larger individual capacities. Like the primary RAID array these will need to be regularly backed up as part of the organisations disaster recovery procedures. Nearline RAID arrays are best suited to short term archives, less than 5 years, where the data is not vital and there is sufficient capacity in the existing data backup process to include this data.

- Tape nearline and offline libraries

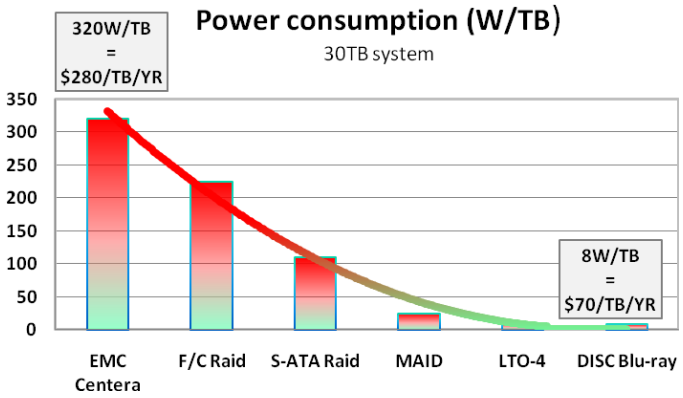
Tape systems have been used traditionally for making data back-ups and are best suited for transferring large amounts of data in a single go. They are not designed for writing or reading individual files, which are typical for an archive. This reduces both reliability and speed as they don´t provide random access. They are best suited for backing up large, contiguous data sets.

- Optical nearline and offline libraries

Optical archive configurations are similar to tape but the main difference is in the media itself. Blu-ray technology is designed for wide use within the consumer/video market and provides backward compatibility. Hence it is reliable, robust and like CD/DVD before it, will be widely available for coming years. Furthermore, an average media life of 100 years is predicted for current 50GB Blu-ray Disc, thanks to its advanced coating technologies. In combination with its low energy consumption and cooling costs, optical media is extremely suitable for long-term and green archiving.

Organisations archive information for a combination of regulatory compliance, long term retention and IT optimisation reasons. These reasons drive a set of requirements and the three main archiving options were compared against them.

With the current concern about global warming and the drive to reduce carbon omissions it is also important to compare the power consumption of the 3 options.



	Magnetic Hard Drive	Tape	Blu-ray Optical
Lower cost than existing primary storage	✓	✓	✓
Long term retention	⊘	✓	✓✓
Compliance with regulations	✓	✓	✓✓
Transparent access to archival information	✓	✓	✓
Offline management of very old information	⊘	✓	✓✓

History has shown that whilst near line RAID and tape archives fulfil some of these requirements they have significant drawbacks and only Blu-ray optical archives are uniquely suited to sustainable long term information archival.

The DISC Blu-ray Optical Library Series is also the archiving solution with the lowest energy consumption of all the options. Hence, with all the standard features of optical archiving (100+ year media life, low power consumption, random access etc), the reliable DISC Blu-ray Libraries provide the safest place for your information. And at the same time you have a sustainable archival storage solution that won't cost the earth.

2. The need for archiving

2.1 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.

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.

Table 1: The Cost of Information tasks to the enterprise

Task	Average hours per worker per week	Cost per worker per week*	Cost per worker per year*
E-mail: read & answer	14.5	$ 418.30	$21,752.90
Create documents	13.3	$ 333.70	$19,952.70
Analyze information	9.6	$ 277.00	$14,401.90
Search	9.5	$ 274.10	$14.251.90
Edit/review	8.8	$ 253.90	$13,201.80
Gather information for documents	8.3	$ 240.00	$12,481.70
File and organize documents	6.8	$ 196.20	$10,201.40
Create presentations	6.7	$ 193.30	$10,051.30
Create images	5.6	$ 162.70	$ 8,461.10
Data entry to e-forms	5.6	$ 162.40	$ 8,446.10
Manage document approval	4.3	$ 124.10	$ 6,450.90
Publish to web	4.2	$ 121.20	$ 6,300.80
Manage document routing	4.0	$ 115.40	$ 6,000.80
Publish to other channels	3.9	$ 112.50	$ 5,850.80
Create rich media	2.8	$ 80.80	$ 4,200.60
Translate	1.0	$ 29.70	$ 1,545.20

*Based on average salary of $60,000 per year plus benefits. *Source IDC: Proving the value of content technologies study.*

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

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 formats into one document format	3.8	110	5,701	5,700,760
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 format to another	2.4	69	3,600	3,600,480
Acquiring archived records with little or no automation	2.3	66	3,450	3,450,460
Version control issues	2.2	63	3,300	3,300,440

** Based on average salary of $60,000 per year plus benefits. Source IDC: Proving the value of content technologies study.*

2.2 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 instead of putting policies into place for a more effective utilization of existing storage.

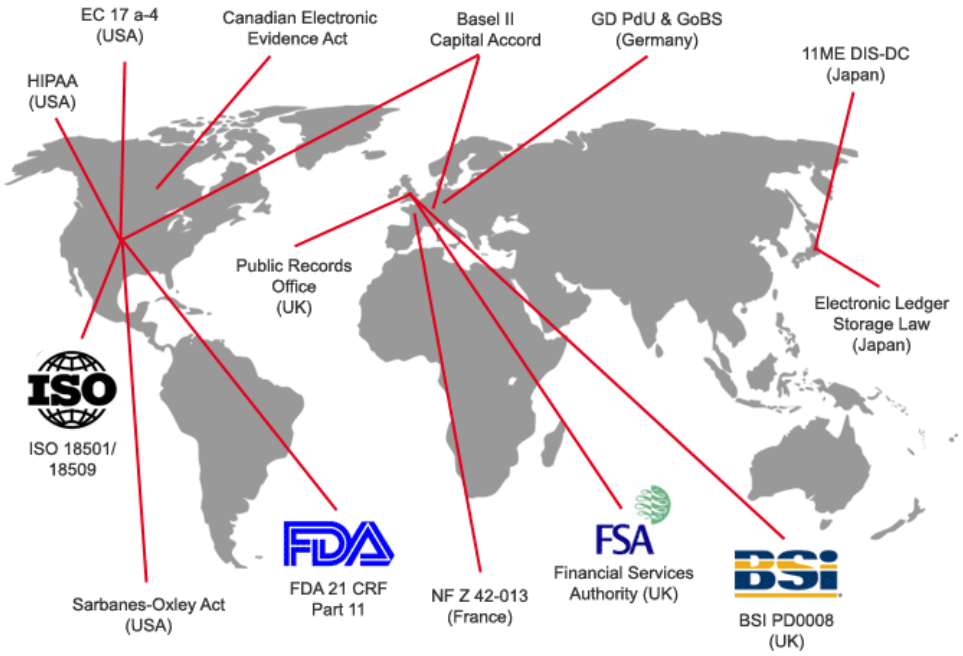
Moreover, 80% of the data on most networks is inactive and as it all resides on hard disk it is a waste of valuable online storage capacity. This includes a high level of redundancy as (duplicate) data is often stored and backed-up several times.

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.

2.3 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.



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).

3. Archival storage considerations

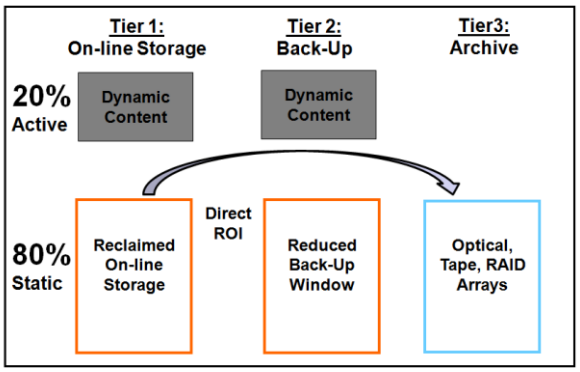
3.1 Archiving Benefits

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.

Archiving alleviates this problem by identifying the older or little used information and manages it separately from the organisations active information. This archive information can then be managed on long term storage. The main benefits of this are:

- Reduce the risk
- Reduce the cost
- Increase the availability of information.

By separating active and static information different protection strategies can be adopted. As there will be less active information to manage an enhanced disaster recovery strategy can be implemented. The backup window is reduced but more importantly the recovery time of a RAID array crashing is reduced. By separating active and static information it frees up space on the file servers and hence reduces the need to purchase additional storage capacity



Cost is reduced as the archival storage is a lower cost than the primary, RAID array, storage. As information gets older the likelihood of it being accessed reduces considerably hence it is not necessary to store this on primary RAID with instant access.

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. Sometimes, there is also some 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.

3.2 Archiving requirements

There are three main reasons organisations archive data:

1) Records retention, for compliance and good corporate governance
    - Retain information for legal or regulatory requirements
    - Original records must be retained in the original form – unaltered
    - Typical information includes:
        - Contracts, financial transactions
        - Also emails, invoices and purchase orders
        - Medical records and images

2) Long term retention, for preservation of information assets
    - Retain information for future use
    - Information is required for 10s of years or indefinitely
    - Typical information includes:
        - Media content, movies, music and TV recordings
        - Experimental data, pharmaceutical trails and oil exploration findings
        - Historical records, maps and legal documents

3) IT Optimisation
    - Separating inactive and active data, to improve efficiency
    - Increases available space on file servers, reduces backup windows and improves disaster recovery time
    - Typical information includes:
        - Unstructured data, user files

In reality, organisations archive for all these reasons but one will be the primary drive to initiate the archiving of information.

Based on the reasons why organisations archive information the fundamental requirements of archiving are:

- Lower cost than existing primary storage
  Information stored in an archive should cost less than leaving it on the primary system. This should not only consider the acquisition cost of additional primary storage but also the disaster recovery procedures and the operational costs.

- Long term retention
  Long term retention is generally agreed to mean longer than the reliably life time of the original system it is stored on. Hence, with a primary RAID system this is beyond 4-5 years. After this period to maintain access to the data the storage system would need to be replaced and the data migrated to this new system.

- Compliance with regulations
  A detailed analysis of the thousands of regulations is beyond the scope of this paper but the essence of these regulations is to ensure that the original record can be reproduced without being altered. The ultimate test is to produce the record in a court of law, prove it has not been altered and be able to use it as evidence.

- Transparent access to archival information
  However the archive is implemented, the end users demand constant access to their information and hence the archive must be transparent to the end user. That is they don't have to look elsewhere for their data and it appears to remain where they put it. In fact IT departments often implement these transparent archives without the end user being aware of any changes.

- Offline management of very old information
  As information gets older the likelihood of it being accessed reduces considerably hence it is not necessary to store this on primary RAID with instant access. Furthermore, as archives tend to be filled up sequentially with little modification, the oldest information was often written to the archive first. Hence, once the organisation no longer needs direct access to this information, it can be off-lined from the archive to further reduce costs. Off-lined data can be retained outside of the archival system and can be restored if the data is required. This allows organisations to avoid making decisions about deleting information.

4. Archival storage technologies

4.1 Introduction

The currently available options for archival storage are:

- Nearline RAID arrays
- Tape nearline and offline libraries
- Optical nearline and offline libraries

There have been some recent developments in the Cloud Storage, where organisations virtualised their storage allowing them to consolidate and manage the storage more effectively.
This approach makes implementing and managing an archive much easier, as the information is consolidated into a single unified storage structure. Identifying old or little used data, migrating it into the archive and managing the access becomes truly transparent within the cloud.
However, the basic archival options remain and are discussed below.

4.2 Nearline RAID arrays

Nearline RAID arrays are basically cheaper versions of the primary RAID array. They are usually created with lower cost magnetic disk drives with larger individual capacities. Like the primary RAID array these will need to be regularly backed up as part of the organisations disaster recovery procedures.

Nearline RAID arrays meet the archive requirements as follows:

1) Lower cost than existing primary storage
    - Lower acquisition cost, due to lower reliability drives
    - Operational costs will be similar to primary storage, in terms of
        - High energy consumption and cooling
            - Energy reduction by powering down drives (MAID=Massive Array of Idle Disks)
        - Backup procedures
        - Monitoring and replacement of failed drives

2) Long term retention
    - Not possible
    - Entire RAID array would need to be replaced and the data migrated to new system every 2-5 years.

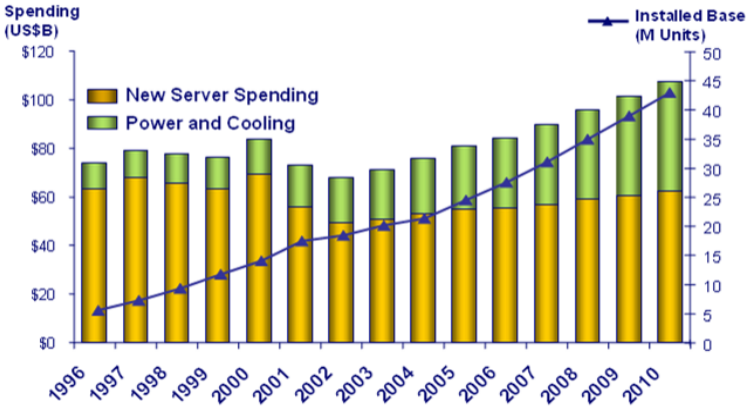
3) Compliance with regulations
    - Complicated software algorithms provide WORM type behaviour

4) Transparent access to archival information
    - Yes, provide a link between existing primary RAID and archive

5) Offline management of very old information
    - Not possible
        - Magnetic hard disks can't be safely removed and stored long term outside of the archive

Nearline RAID arrays are best suited to short term archives, less than 5 years, where the data is not vital and there is sufficient capacity in the existing data backup process to include this data.

Continuous power consumption and the lack of removability of the media make RAID technology less suited as long-term archiving technology.



- **Power & Cooling cost in 1996 is only 15% of total**
- **Power & Cooling cost in 2010 will be 40%**

(source IDC)

4.2 Tape nearline and offline libraries

Tape nearline archives consist typically of a server which is network attached (NAS) and is running software that manages the movement of data from the primary RAID to the tape archive.

The tape library is a large cabinet with shelves for the tapes and a smaller number of tape drives. A robotic mechanism moves the tape media from the shelves to the drives allowing data to be accessed.

They meet the archive requirements as follows:

1) Lower cost than existing primary storage
   - Yes, most companies have some form of tape backup already installed so only additional software required
   - Tape has a low power consumption

2) Long term retention
   - Yes, tape media life is estimated at 7-10 years
   - However, tape technology is not always backwards compatible
     - For example a tape written with LTO1 technology cannot be read with the current LTO5 drive

3) Compliance with regulations
   - No, tape can be altered as tape is fundamentally re-writable media
     - It requires additional software to prevent alternation of information

4) Transparent access to archival information
   - Yes, but not suited for random access, tape needs to spool which results in:
     - A lot of tape winding also reduces reliability
   - Access to files is slow due to long delays in finding the start of individual files

5) Offline management of very old information
    - Yes, removability of media allows for off-lining of information,
    - However, tape needs periodic re-tensing to prevent tape adhering together
    - Tape vulnerable to (electro) magnetic radiation

Tape systems have been used traditionally for making data back-ups and best suited for transferring large amounts of data in a single go. They are not designed for writing or reading individual files which is typical for an archive and this reduces their reliability. Tape is in most cases rewritable and as a consequence not applicable in many 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 7-10 years, this reduces extremely its efficiency as data has to be migrated to new standards and exchange between different suppliers could be complicated.

4.3 Blu-ray optical nearline and offline libraries

Optical archives configurations are similar to tape. The big difference is in the media itself. Blu- ray technology is designed for wide use within the consumer/video market.

Hence it is reliable, robust and like its predecessors CD/DVD, it will be widely available for coming years.

Optical/Blu-ray technology meets the archive requirements as follows:

1) Lower cost than existing primary storage
    a. Yes, total cost of ownership is lower than primary RAID
        i. Low energy consumption and cooling costs
        ii. No need for constant backup

2) Long term retention
    - Yes, Media designed for up to 100 years
    - Blu-ray is the successor to CD/DVD which has a long history

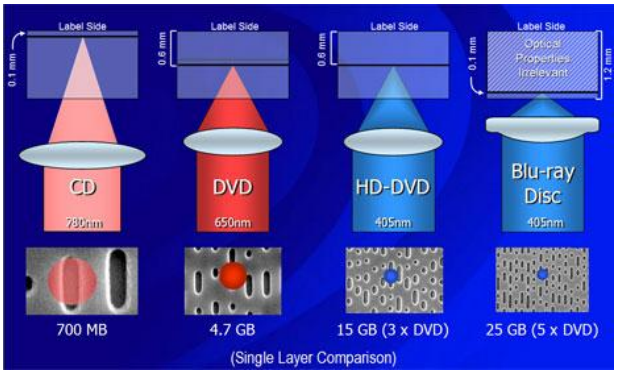
3) Compliance with regulations
    a. Yes, BD-R is unalterable, WORM (Write Once Read Many) media

4) Transparent access to archival information
    a. Yes, provide a link between existing primary RAID and archive
    b. Optical media is random access, giving fast access to individual files
    c. Access time to files is slower than nearline RAID but faster than tape

5) Offline management of very old information
    a. Yes, Removability of media allows for off-lining of information
    b. Media has a hard coat to protect it
    c. Magazine, **SmartPack**, used to handle media outside of library

The benefit of using a blue-violet laser (405nm) is that it has a shorter wavelength than a red laser (650nm), which makes it possible to focus the laser spot with even greater precision.

This allows data to be packed more tightly and stored in less space, so it's possible to fit more data on the disc even though it's the same size as CD/DVD. This together with the change of numerical aperture to 0.85 us is what enables Blu-ray Discs to hold 25GB/50GB.

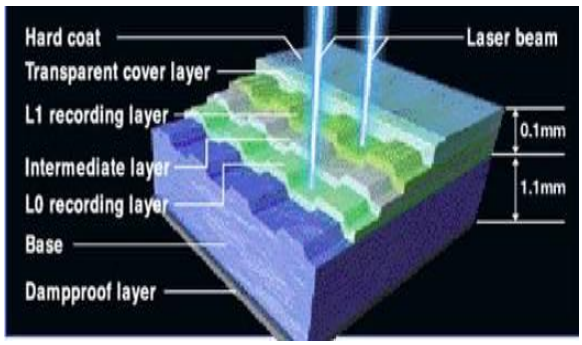


Blu-ray Discs are currently available in two versions: single layer and double layer. A double layer disc may hold up to twice the amount of data or video compared to a single layer disc, and uses two independent layers placed on one side of the disc to store its information. A single-layer disc holds up to 25 gigabytes, while a double-layer disc holds up to 50 gigabytes of data, without the need to flip the disc.

Blu-ray Disc format is easily extendable (future-proof) as it includes support for multi-layer discs. This allows the storage capacity to be increased to 100-200GB in the near future simply by adding more layers to the discs.

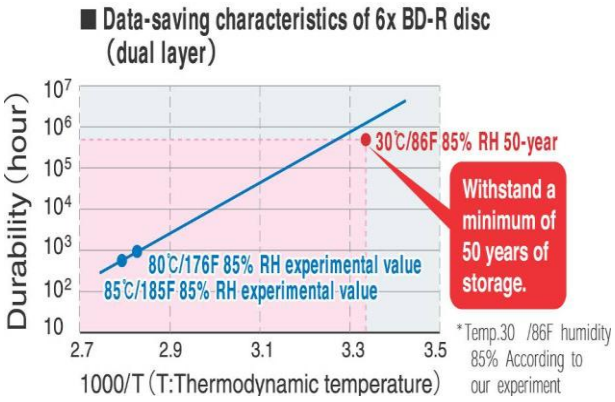
Blu-ray technology makes use of new low cost **hard-coating** for protection, which makes the discs even more resistant to scratches and fingerprints than today's DVDs, while still preserving the same look and feel.

It also adopts a new error correction system which is more robust and efficient than the one used for DVDs.



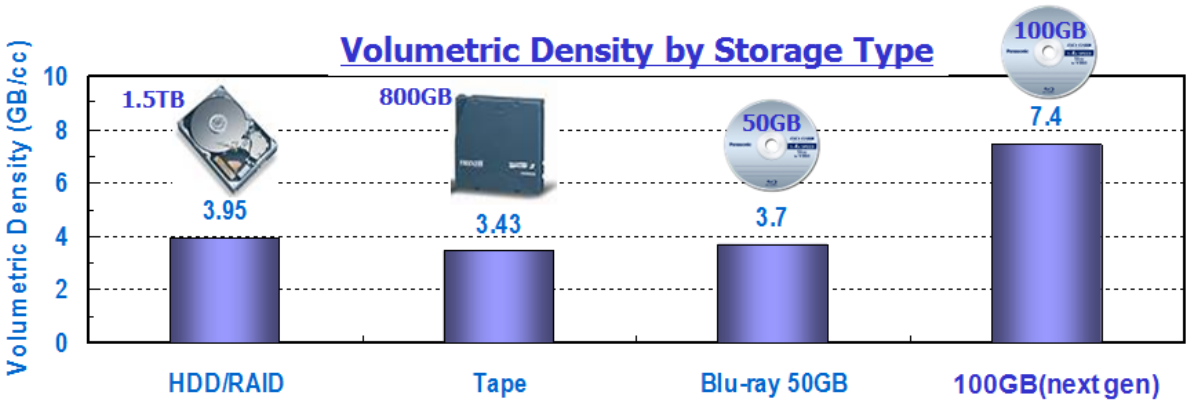
Thanks to its advanced coating technologies, an average media life of 100 years is predicted for current 50GB Blu-ray Discs by most suppliers.

Tests show that BD-R media are extremely suitable for long-term archival storage purposes. In a typical office environment (normal room conditions), the projected archival lifespan is at least 50 years.
(Source Panasonic).



Furthermore, 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. This avoids the direct need for any additional client software as data can still be read in a standard drive in a standard PC.

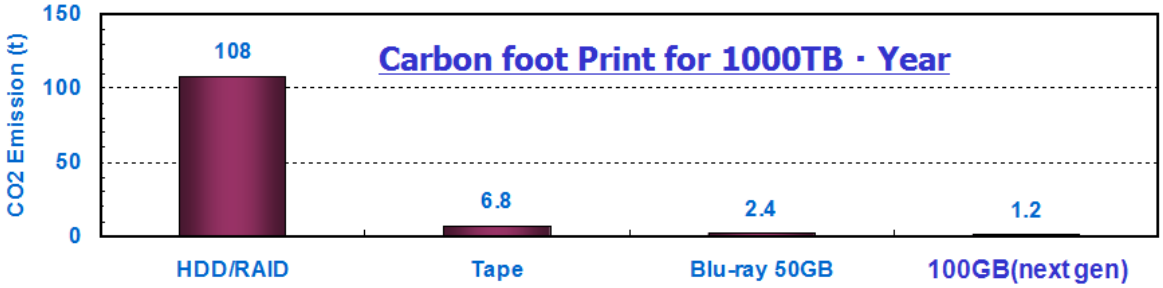
Optical media is also less vulnerable than magnetic media as it is less susceptible to environmental conditions. Blu-ray Technology offers random access to any file on BD Discs in the digital archive which does not require regular re-spooling. Media is normally not spinning and hence less subject to physical damage. Together with its advanced media life (100 yrs) and future roadmap of 100-200GB per Disc, costly data migrations are fewer hence there is also less chance of losing data.



Source: Panasonic

In addition, it provides a true WORM recording medium which can't be altered with software hacks or viruses. This 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.

Last but not least, Blu-ray Disc Technology offers a sustainable archiving technology. Blu-ray media consumes zero power when not being accessed and the energy consumption is low due to shared resources. This generates low heat and little or no cooling required.



Source: Panasonic

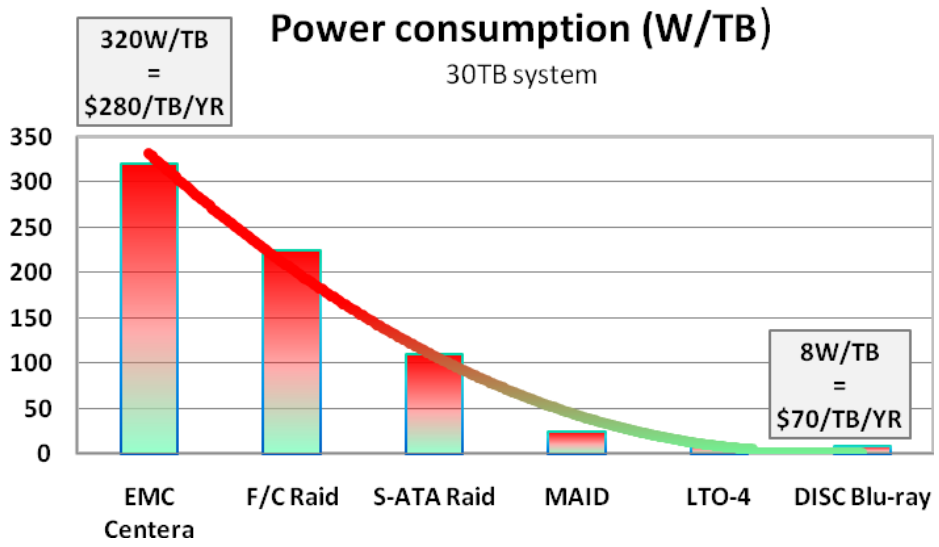
5. Conclusion

The continuous growth in digital data and industry-specific legislation compel many professional organisations to include the topic of data archiving and storage in today´s board meetings.

Organisations archive information for a combination of regulatory compliance, long term retention and IT optimisation reasons. These reasons drive a set of requirements and the three main archiving options were compared against them.

	Magnetic Hard Drive	Tape	Blu-ray Optical
Lower cost than existing primary storage	✓	✓	✓
Long term retention	⊘	✓	✓✓
Compliance with regulations	✓	✓	✓✓
Transparent access to archival information	✓	✓	✓
Offline management of very old information	⊘	✓	✓✓

The **DISC Blu-ray Optical Library Series is** also the archiving solution with the lowest energy consumption of all the options.



Taking into account the main requirements of an efficient and long-term archiving technology,

Blu-ray Technology fulfills most of these as part of its standard features.
Key benefits are:

- Longevity
  - 50+ years media lifetime guarantees long-term data retention
  - Promising technology roadmap of 100-200GB/disc
- Compliance
  - BD-R is write-once, offering highly reliable and tamper-proof archives
- Standard
  - Standardized technology (UDF) offers long-term data read.
  - Widespread acceptance by users
  - Multi-platform support
- Green
  - Less CO2 emission requiring minimal power and air-conditioning consumption
- Low TCO
  - Use of standard technology simplifies implementation and administration
  - Proven compatibility (CD, DVD, BD, +) reduces the need for costly and inefficient data migrations.
  - Economically priced media.
  - Longevity, of data life

The reliable DISC Blu-ray Libraries and meanwhile a sustainable archival

provide the safest place for your information storage solution that won't cost the earth.