

ARCHIVAL STORAGE TOP 10 QUESTIONS

INTRODUCTION

Organisations across virtually every industry are struggling with requirements to retain corporate records for ever greater periods of time. It is not uncommon for companies to need information access for decades and where public safety or cultural heritage is a factor, records may need to be retained indefinitely.

The rapid pace of technological evolution has forced companies to take a short-term, tactical view of IT, which can clash dramatically with the growing need for a long-term archival storage strategy. Products and processes used to solve short-term storage requirements are often inappropriate for long-term archival storage.

This questionnaire is designed to raise some of the issues most critical to building and maintaining a secure data archive. The following pages provide additional background information and summarise the role UDO plays relative to each question.

THE TOP 10 QUESTIONS

1. How much data is being stored in your network and what percentage of the total is archive data?
2. What is the growth of your archive data on a weekly or monthly basis?
3. Is your organisation subject to external industry regulations or internal corporate policies on record retention and authenticity?
4. What is the profile of your archive data (i.e. its format, size, retention period and access requirements)?
5. How quickly do archive records need to be accessed to meet your business requirements?
6. What is the maximum number of concurrent users likely to be accessing your data archive at any given point in time?
7. What is the disaster recovery strategy for your archive?
8. What is your plan for maintaining older media and migrating data to newer storage technologies over the life of the information?
9. Who within your organisation is responsible for ensuring that archival storage compliance is being properly addressed?
10. How do you calculate the Total Cost of Ownership (TCO) for creating and maintaining a long-term data archive?

THE TOP 10 QUESTIONS IN DETAIL

1. How much data is being stored in your network and what percentage of the total is archive data?

Quantifying the volume of data in a network is an important place to begin. It seems a simple question, but many system administrators do not have a good understanding of exactly how much data they are trying to manage. Capturing the total data volume requires determining the actual disk utilisation minus duplication, orphaned data and system objects.

The next step is to identify what percentage of the total data volume is active data and what percentage is inactive data. Active data is defined as data currently being created, modified or frequently accessed. Inactive data can be defined as data that is no longer being modified and is less frequently accessed. Some inactive data may have very little value to an organisation, but other inactive data is valuable and must be retained for long-term availability. Valuable inactive data can be classified as “archive” data and requires an archival storage strategy if it is to be reliably stored for long periods of time.

Archive data typically has very different storage priorities and requirements and by identifying this data and separating it from active data, a unique strategy can be defined to ensure that the archive data is retained for as long as required and its authenticity is protected.

The Role of UDO

Administrators are often surprised at how much of their online storage is consumed by archive data. Separating archive data from active data can have major benefits in administration and cost reduction. For example, since archive data is no longer being modified, it is not necessary for it to be part of a normal backup cycle. Archive data can be placed outside the standard backup schedule so that backup operations can focus on protecting the active data that is changing on a frequent basis. Moving archive data to non-volatile UDO storage dramatically reduces the average backup window, administrative overhead and the total cost of the backup process.

Plasmon G-Series libraries Powered by UDO provide archive capacities ranging from under 1TB to over 19TB in a single footprint. The combination of UDO technology in a G-Series optical library offers a reliable, scalable archive solution at a very competitive cost.

[Back to The Top 10 Questions](#)

2. What is the growth of your archive data on a weekly or monthly basis?

The long-term nature of an archive means that you have to design for current requirements while planning for future growth. Having said this, archive growth calculations can be difficult. It may not be adequate to base your growth estimate on current data types since history has demonstrated that the archive net is being cast ever wider to include new data types not previously considered. Estimates need to include the growth rate of currently identified archive data, as well as, the potential addition of new records that may need to be managed in the future.

The Role of UDO

Constantly growing volumes of archive records demand solutions with high initial capacity and the flexibility to scale over time. Both UDO technology and G-Series libraries Powered by UDO have been designed specifically to accommodate archive growth. With current 30GB UDO media, the Enterprise G-Series can expand from just under 5TB to over 19TB through upgradeable modules. Overall system capacity will grow even greater as the UDO technology roadmap increases media capacity to 60GB and 120GB in the future.

This strategy provides incremental archive capacity while protecting your initial hardware investment. The removable nature of UDO is another important advantage to growing archives. Rather than purchasing new or expanding an existing library, organisations have the option of taking older data sets offline to make room for newer data. This is simply not an option with magnetic disk based archives without purchasing additional hardware to transfer the data. Whether expanding an existing library or taking data offline, both techniques are non-disruptive and cost-effective strategies to incrementally expand archive capacity.

[Back to The Top 10 Questions](#)

3. Is your organisation subject to external industry regulations or internal corporate policies on record retention and authenticity?

Another way of clearly identifying archive data is by regulations and policies that define the processes and retention periods for specific data types. More and more organisations are required to adhere to government and industry regulations on data storage for legal or public safety purposes. In addition to external regulations, organisations often have their own data retention policies to ensure availability of corporate knowledge and to protect against potentially damaging litigation. Understanding the role that regulations and policies play within your environment is essential to establishing a defensible data archive.

The Role of UDO

Whether dictated by industry regulation or corporate policy, one overriding goal is to establish and maintain data authenticity. True authenticity is established through processes and practices that provide well documented audit trails for all data. In addition, the storage technology also plays a crucial role. In areas where authenticity is most critical, many regulations call for unalterable, Write Once storage. UDO media is available in a true Write Once format, providing a level of data authenticity that is unmatched by rewritable magnetic disk and tape technologies using Write Once emulation.

[Back to The Top 10 Questions](#)

4. What is the profile of your archive data (i.e. its format, size, retention period and access requirements)?

Once the volume and scope of your archive data has been determined, the next step is to sub-divide the archive by establishing profiles for each primary data type. A profile is a set of attributes that characterise a document format, size, retention period and access requirements.

Archives being used for a single task may only require two or three profiles since the document types may be very limited. By contrast, archives being used by many applications will require additional profiles to reflect the wider range of document attributes. The objective of profiling is to add structure to previously unstructured data. Well-defined profiles allow the administrator to properly manage both the physical and logical storage of the archive data.

For example, healthcare institutions handle both patient records, as well as, high volume digital images generated by modern medical equipment such as x-rays, MR or CT scans. Most of this data must be retained for the lifetime of the patient. Fast data retrieval is a high priority in order to ensure that medical staff have ready access to patient information. Data integrity is of foremost importance as it is an essential criterion in maintaining a secure and unadulterated Healthcare Enterprise Archive.

Engineering and manufacturing companies are also dealing with large and complex data, such as 3-D models, simulations, mock-ups and other technical documents, but the retention period and access requirements attributes will be different from those of medical records.

The Role of UDO

One advantage of data profiling is the ability to physically group similar profiles on the same piece of UDO media. By doing this, older or less frequently accessed records could be removed from a G-Series library making room for newer data while still retaining the media for potential access. For example, an engineering company can choose to archive all design, technical and administration records for one or more related projects on a single 30GB cartridge, allowing projects to be taken offline for future reference or distribution to other site locations.

Cartridged UDO media is extremely robust and the media has a higher tolerance to environmental storage conditions than magnetic tape and other removal media. UDO media is designed to accommodate offline storage, providing flexible options for companies to meet demanding data profile requirements.

[Back to The Top 10 Questions](#)

5. How quickly do archive records need to be accessed to meet your business requirements?

Since data written to most archives is part of a background process, the ingestion rate of records into an archive is often not critical path. By contrast, timely retrieval of archived records is business critical. While it is true that most archive data does not require the same access performance as active data, just how quickly do your archive records need to be available? Is a 5-minute retrieval time adequate or do you need to have access in 10 seconds? The answer to this question will be closely tied to your business requirements and the discovery period mandated by regulations. It is also a critical factor in choosing the proper hardware. Failure to meet access time service level agreements can have a significant impact on the overall success of products or services that are dependent on your archive.

The Role of UDO

UDO delivers random access performance far higher than that of magnetic tape archives. With a G-Series library you can have access to any file in less than 10 seconds where a tape archive typically takes several minutes, as random access to individual files is very slow on sequential media. If your business requirements demand quick access, tape may simply not be an option.

[Back to The Top 10 Questions](#)

6. What is the maximum number of concurrent users likely to be accessing your data archive at any given point in time?

Timely retrieval of archived records is business critical and an important factor that will impact on data access time is the number of concurrent users accessing the archive. The ability to predict access patterns and user levels is important to properly configure the archive hardware. An archive with insufficient "horse power" may cause request queuing, impacting the overall performance of the archive and consequently the business operations.

The Role of UDO

To accommodate a range of concurrent access requirements, G-Series libraries Powered by UDO offer highly adaptable configurations, supporting from 2 to 12 drives. The more UDO drives in the library, the more concurrent read /write operations can be accommodated. The quick exchange times and fast seek operations of UDO drives result in a very low drive to capacity ratio. By contrast, tape drives have very high exchange, seek and rewind times, requiring many more tape drives than UDO drives to satisfy the same level of concurrent access (if it can be satisfied at all). Understanding the access demands on the archive allows a UDO library to be properly configured to meet demanding business requirements from simultaneous users without expensive hardware overkill.

[Back to The Top 10 Questions](#)

7. What is the disaster recovery strategy for your archive?

By definition, archive data has intrinsic value to a company. Lost data could cost an organisation a huge amount of time and money to recreate from scratch and some records, such as legal or historic documents, can simply not be replaced. Even the most secure storage media cannot protect your data from fire, flood or sabotage. If your archive records are valuable to your business you cannot afford to have a single copy. For this reason, it is important to implement a disaster recovery strategy that provides for a secondary copy of your archive in a protected vault or at another geographic location. A disaster recovery strategy costs money, but what is the cost of loosing your archive?

The Role of UDO

The stability of UDO media cannot protect data from site disasters, so like all other media there should be at least one secure secondary copy of your valuable archive data. There are two typical disaster recovery strategies: mirrored systems and offline vaulted media and G-Series libraries with UDO support both. The removability and stability of UDO media makes offline vaulted media a particularly secure and cost effective option that magnetic disk archives cannot support without additional hardware and administrative overhead. Where fast access to disaster recovery data is required, mirrored G-Series libraries can be installed at two independent locations to provide virtually uninterrupted failover.

[Back to The Top 10 Questions](#)

8. What is your plan for maintaining older media and migrating data to newer storage technologies over the life of the information?

Applying short-term IT solutions to a long-term archive can prove to be inefficient, costly and dangerous to the longevity of the data. The time periods involved in an archive pose unique considerations that are not always part of the typical IT infrastructure. Since archive data may need to be available for many decades, you need to plan for the maintenance of your archive environment and for the potential migration of data from one storage architecture to another. The choice of storage technology, both software and hardware, will have a major impact on maintenance overhead and migration frequency. Choosing products that have not been designed with a long-term support strategy will result in higher costs, increased maintenance and may compromise the authenticity of the archive data. Choosing products in conformity with industry-standards is another wise consideration, since international standards provide an assurance for quality and interoperability.

The Role of UDO

The stable recording and media properties of UDO provide data life in excess of 50 years, which cannot be matched by magnetic disk or tape technologies. While nobody expects archive data to be stored on the same piece of UDO media for 50 years, the stability of the media reduces the frequency of data migration. With UDO it is practical to migrate data once every 8 to 10 years, far less frequently than is required by magnetic tape or disk. And unlike tape, which must be frequently re-tensioned and stored in carefully controlled environmental conditions, UDO requires virtually zero maintenance over the life of the media, reducing the Total Cost of Ownership of a UDO archive.

Reduced migration frequency is not only a function of the UDO media, but of Plasmon's overall business philosophy. Specialising in archival storage for almost 20 years means that Plasmon takes a very long-term view of product support. Unlike magnetic disks that can have an 18 to 24-month life, Plasmon's archival storage products have a proven track record for providing support in excess of a decade.

In addition, next generation UDO drives are read backward compatible. This means that enabling a G-Series library to support higher capacity UDO media requires only a drive upgrade, protecting the library and media investment while allowing the use of older media until data migration is appropriate (if at all). Organisations archiving valuable data for many years need to be confident that the company providing their technology has a long-term support philosophy.

Based on these values, Plasmon submitted the UDO media format to ISO, IEC and Ecma International for certification and the technology was accepted by these most trusted technical standards organisations as an industry recognised standard. Having a fully documented and open international standard is another benefit to UDO users. An ISO standard for UDO provides assurance for those companies investing in the technology and defines a framework for backward compatible future UDO generations offering 60GB and 120GB capacities.

Back to The Top 10 Questions

9. Who within your organisation is responsible for ensuring that archival storage compliance is being properly addressed?

For organisations that place a high priority on data compliance, there is a growing trend to define a role that is responsible for the compliance process. Often referred to as the "Compliance Officer", this person takes responsibility for researching compliance requirements and for the design and implementation of both procedures and technology to meet compliance obligations.

For smaller organisations this role may be a part time responsibility, but it still makes good sense to establish the position since it can serve more than one purpose. First and foremost, it will help to ensure that the organisation is meeting all necessary compliance regulations. Secondly, it demonstrates corporate commitment to legal and regulatory bodies. Thirdly, it better prepares an organisation to respond professionally in the event of litigation. Those organisations with a Compliance Officer are in a stronger position to defend the integrity of their audit trails and data, avoiding the potential of costly penalties and jail sentences.

The Role of UDO

The more tools a Compliance Officer has at his disposal, the better able an organisation is to demonstrate data integrity. Clearly documented audit trails are fundamental and while this is largely a question of procedure, UDO also plays an important role. UDO true Write Once media uses an irreversible recording technique so that data can never be changed or altered, which facilitates audit trail management. In addition, each piece of UDO media has a unique, unalterable, serial number. Software application can include this serial number as part of the audit trail data, recording when and on which piece of media records have been written. This extra level of authentication prevents someone from writing modified records to a new piece of media in an attempt to tamper with archived data.

Back to The Top 10 Questions

10. How do you calculate the Total Cost of Ownership (TCO) for creating and maintaining a long-term data archive?

Cost is always a key factor in choosing any IT solution, but here again the criteria by which you judge the cost of a data archive is different than that for active data storage. Since an archive needs to operate for many years, you cannot simply look at the initial acquisition cost and because an archive consists of many components you cannot compare it to the raw storage cost of magnetic disk, tape and optical. To fairly compare different technology options, you must look at all the pieces necessary to meet the longevity and authenticity requirements demanded by an archive environment.

The table below summarises the primary costs that are more easily quantifiable, as well as, the secondary costs that can be more difficult to quantify since they are often dependent on local resources and the selected technology.

Costs	Frequency
Primary Costs	
Hardware Acquisition	One time
Software Acquisition	One time
Media Acquisition	One time
Hardware Maintenance	Annual
Software Maintenance	Annual
Floor Space	Annual
Power and Cooling	Annual
Secondary Costs	
Installation and Training	On time
System Administration	Annual
Data Migration	Periodic
Hardware / Software Upgrades	Periodic
Media Maintenance	Annual
Data Authentication / Audit Trail Management	Annual
Disaster Recovery Strategy	Annual

When estimating the TCO of a particular configuration or when using TCO to compare different technology options it is important to factor in as many of these costs as possible and to do so in an even handed way.

The Role of UDO

Unlike magnetic disk and tape, UDO has been designed specifically for archival storage. As such, the fundamental attributes of UDO are a natural fit into archive environments. Scalability, longevity and authenticity are inherent traits of UDO so do not require special hardware or software to emulate these qualities. In addition, Plasmon's long-term support commitment to the technology is in keeping with the needs of organisations that must maintain their archive for many years in the future.

These key design benefits have a direct impact on the TCO of a UDO archive. When factoring in all the appropriate costs over the years of operation, UDO proves extremely cost competitive. Compared to magnetic disk archive solutions, UDO is typically much less expensive and offers the authenticity and removability not possible with rewritable magnetic disk. In the case of magnetic tape, UDO offers a very similar TCO with the added benefits of unmatched longevity, authenticity and much quicker access performance.

Visit the Plasmon website to download a copy of a detailed Archival Storage TCO Analysis based on the costs outlined above. www.plasmon.com/ref/tco

Back to The Top 10 Questions

EMEA Headquarters

UK / Scandinavia / Israel / South Africa
 Plasmon Data Limited
 Whiting Way, Melbourn, Herts, SG8 6EN
 United Kingdom
 Tel: +44 (0)1763 262963
 Fax: +44 (0)1763 264444
sales@plasmon.co.uk

Regional Sales Offices

Benelux / France / North Africa
 Plasmon Data sa/nv
 Draaiboomstraat 6, 2160 Wommelgem
 Belgium
 Tel: +32 (0)3 355 1980
 Fax: +32 (0)3 354 2155
plasmon.ne@plasmon.co.uk

Central and Eastern Europe

Plasmon Data Limited
 Ansberger Strasse 36B, 84048 Mainberg
 Germany
 Tel: +49 8751 875 100
 Fax: +49 8751 875111
plasmon.ce@plasmon.co.uk

Southern Europe and Middle East

Plasmon Data SRL
 Via Marsala 36 (Torre B), 21013 Gallarate (VA)
 Italy
 Tel: +39 0331 782904
 Fax: +39 0331 792905
plasmon.se@plasmon.co.uk

www.plasmon.co.uk
www.udo.com

Copyright Plasmon Data 2004. All trademarks listed are registered with their respective companies.