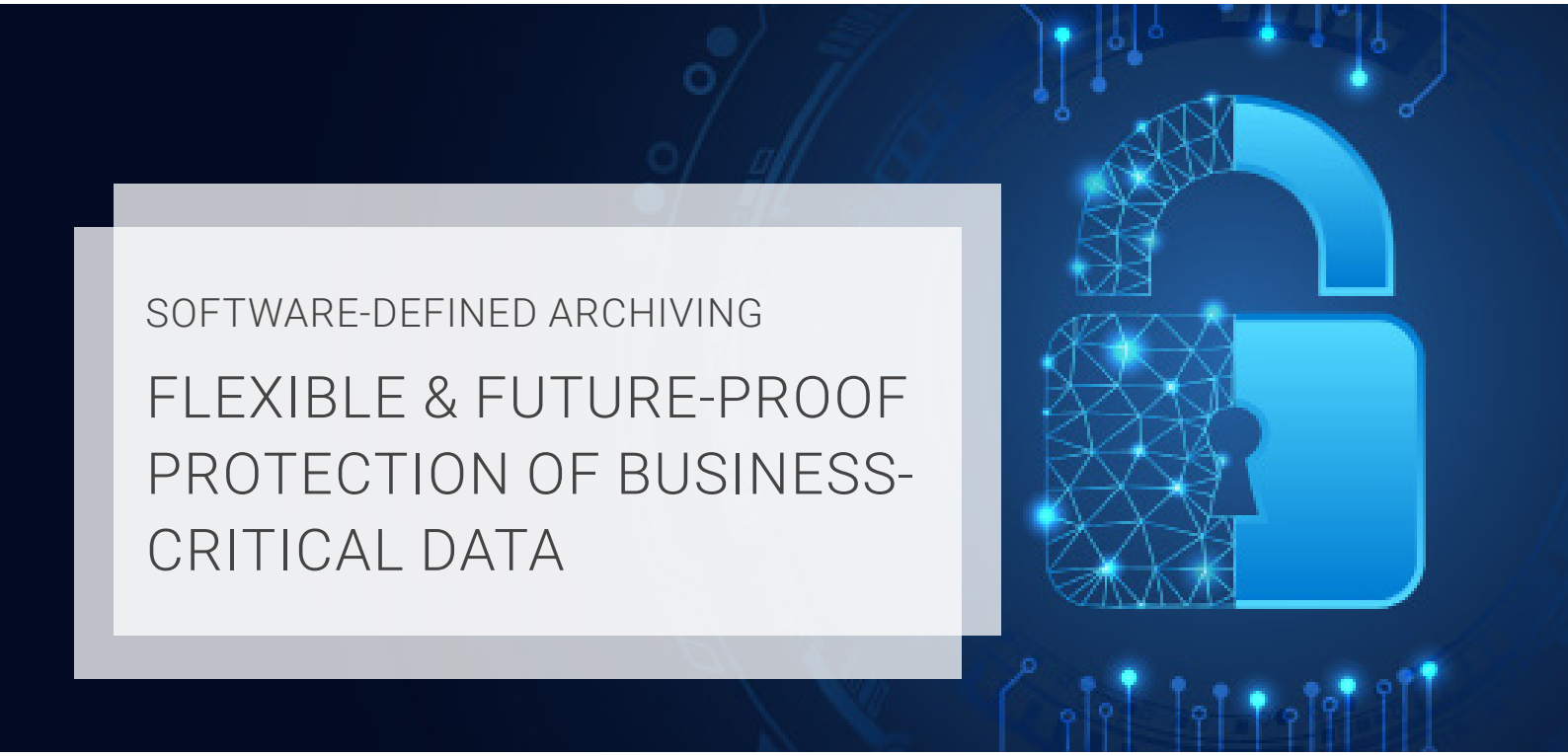
SOFTWARE-DEFINED ARCHIVING

FLEXIBLE & FUTURE-PROOF PROTECTION OF BUSINESS-CRITICAL DATA

As data volumes grow, so do the requirements for long-term data storage and archiving. Learn in this whitepaper how a software-defined approach can help you protect business-critical data in a future-proof, flexible, and cost-efficient way.

CONTENT



SOFTWARE-DEFINED ARCHIVING

FLEXIBLE & FUTURE-PROOF PROTECTION OF BUSINESS-CRITICAL DATA

1. THE CURRENT SITUATION

Today, information is largely available in a company as unstructured digital data. This mass of information in the form of emails, production data, patient data, business documents, research data, etc. is growing unabated. Organizations in all sectors are faced with the task of solving growing challenges to their information management and information security, while making do with fewer resources and lower budgets.

IT departments not only have to process, store, and protect these volumes of data, but also partly to archive them in an audit-proof manner. The focus here is on data security, integrity, and availability - which must be permanently guaranteed. However, the storage requirements usually grow faster than the available budgets.

To keep costs and efforts under control, companies are forced to increase the efficiency of their IT in all areas and to optimize their infrastructure. The ability to scale storage capacity at a reasonable cost plays a crucial role in this.

In this whitepaper, we therefore discuss the question of what challenges companies face today in long-term data storage and archiving and how these can be solved in a future-proof manner.

2. CHALLENGES IN LONG-TERM DATA STORAGE

FLEXIBILITY

Many documents and data are subject to retention periods of ten, thirty, or more years. At the same time, the storage industry is experiencing rapid technological progress. In this area of tension, the question of flexible handling of business-critical data arises. It is clear that archive data often has to be migrated several times during the retention period - but it remains unclear which storage technology will be the standard in the future. The only way out lies in increasing flexibility and increased adaptability in terms of infrastructure.

But it is not only the storage technology which requires the greatest possible flexibility; the application level above it must also be considered in order to avoid data silos and dependencies in the future. For example, all data from the business applications used - such as emails, documents, invoices, patient data, videos, images, SAP data - must be stored securely. As technologies are constantly evolving, companies and institutions should rely on open interfaces and standards to ensure that future applications can also be connected.

When looking into the future of data storage and archiving, the question of hardware and software procurement is also exciting – the keyword is „as-a-service". Models with usage-based payment for infrastructure and software solutions have now reached all areas of IT and offer companies further flexibility in their IT - always geared to actual needs.

COST & EFFORT

Money matters, and this is not only the case with long-term data storage. Rapidly increasing data volumes and ever more complex demands on IT are pushing the issue of cost efficiency high up the IT priority list.

It is crucial to evaluate the total cost of ownership (TCO). This perspective goes beyond the mere consideration of the acquisition costs of hardware and software and also integrates, for example, system administration, training, or support services. Thus, the focus is not only on cost reduction, but rather on increasing benefits and efficiency.

Since IT costs are distributed over the entire life cycle, procurement and administration play a central role, because they account for a significant part of the total costs. In the long term, however, the ongoing operating costs form the center of the TCO view. In the area of archiving and long-term storage, it is therefore not „only" about the license costs of the software or the procurement costs of the hardware. The costs for administration, maintenance, and operation must be reduced in the long term. This aspect becomes even more important when the lack of skilled workers and resources in IT is included in the consideration.

The topic of data migrations must not be neglected either. Due to the long storage periods, increasing data volumes, and the technological progress of storage systems, migrations of the data pool to new systems are unavoidable - an aspect which often receives too little attention and which, unfortunately, is often still a major time and cost eater.

COMPLIANCE & DATA INTEGRITY

To store data in an audit-proof manner, it is not enough to store it in a protected location on a robust storage medium. Compliance-conform archiving requires not only the storage of data according to the WORM principle - Write Once Read Many - but also reliable retention management. The keywords „data integrity" and „long-term availability" play the key role here.

Companies are bound by different guidelines and legal requirements depending on the industry. Often, specified retention periods must be adhered to and the integrity of the archived data must be guaranteed. Many financial institutions, for example, must comply with the strict requirements of the US Securities and Exchange Commission for data archiving, SEC 17a-4. In the pharmaceutical industry, GxPData is an important requirement; and in the healthcare sector there are the requirements of HIPAA, KRITIS, and X-ray regulations. Other laws and requirements such as the German Commercial Code (HGB), the German Tax Code (AO), the IT Security Act (IT-Sicherheitsgesetz), SOX, Basel III, GoBD, IDW PS 330, IDW PS 880, BDSG, the Swiss Code of Obligations (OR), and the Business Records Ordinance (GeBüV) are also important.

Data integrity describes certain requirements for the protection and quality of digital data. To maintain integrity, the consistency, completeness, accuracy, and validity of data must be ensured throughout the entire retention period. All changes must be documented in a traceable manner so that data cannot be changed or manipulated unnoticed or without authorization. Data integrity has the overarching goal of protecting data from internal and external breaches and changes. Loss of data integrity can lead to compromised data and have far-reaching consequences.

Regardless of the industry, the European General Data Protection Regulation (GDPR) is highly relevant for all companies based in the European Union. The regulation governs the collection and processing of personal data of EU citizens. Violations of data protection regulations are sanctioned under the GDPR with fines of up to €20 million or 4% of global group turnover, whichever is higher in each individual case. Among other things, the regulation requires that IT systems be set up from the ground up to effectively implement data protection principles (privacy-by-default and privacy-by-design).

Some regulations and laws also require the encryption of certain corporate data, which is why it is advantageous if the archive storage directly includes encryption. The goal of archive compliance efforts is to minimize business risks which can arise from a loss of data and to permanently ensure the provability of archived data. In addition, product liability requirements and the company's own guidelines must be fulfilled - all in all a monumental task.

COMPLEXITY

No large company can avoid the issue of IT complexity. But the question of a simple infrastructure and transparent processes also plays an important role in hospitals, public administration, or medium-sized businesses. But how can complexity be reduced?

First of all, it helps to sort things out. Data silos must be abolished, too many different configurations of hardware, system software, and middleware must be reduced. In the area of long-term storage, economies of scale can thus be achieved, e.g. by storing business-critical data, such as backups and archive data, in a scalable system.

It is fundamental to maintain openness and independence from certain technologies and providers, for example in the choice of storage, interfaces, or the connection of applications. A storage system should also be multi-client capable in order to accommodate different requirements in one system. The complexity of the IT landscape can hardly be avoided nowadays, but it can be reduced in many areas.

SILENT DATA CORRUPTION & RANSOMWARE

If bits on the storage medium change their state randomly and without external influence, this is called silent data corruption. This can distort stored information or even make it unreadable. This „tilting" of individual bits can be triggered by ageing of the media, chemical, or electro-magnetic processes, among other things. With the rapidly increasing amounts of data in companies, the danger of losing data or no longer having it as a correct and legally compliant 1:1 copy is also constantly growing.

Without a secure backup and/or the ability to detect silent data corruption, errors often only become apparent when the data is already corrupted or lost altogether. Therefore, corporate archives should always be able to check the integrity of the stored information and automatically restore faulty data if necessary. This ability to self-heal is essential for an audit-proof archive.

External threats also play a central role in data storage today. Ransomware is on everyone's lips, is becoming increasingly perfidious, and leaves no industry untouched. It is almost impossible to stop a targeted malware attack on a local infrastructure. Even with the use of anti-virus protection, MFA, detection algorithms, versioning, and snapshots, there remains a significant residual risk.

But especially in data storage and archiving, various approaches and mechanisms help to establish comprehensive protection against malware. Be it redundant storage on media outside the local infrastructure, the immutability of data by applying the WORM principle (Write Once Read Many), encryption, limitation of data access points, the choice of operating system, or a comprehensive backup and disaster recovery strategy - each component is helpful and together they form a robust foundation against the increasing danger of ransomware attacks.

3. ARCHIVE STORAGE EVOLUTION

In the past, mainly two storage concepts, jukeboxes and proprietary hardware systems, have established themselves for long-term storage and archiving. Today, due to increasing IT requirements and rapid technological progress, more sustainable and flexible storage models have emerged.

THE FIRST GENERATION - JUKEBOXES

Colloquially known as „jukeboxes", magneto-optical removable storage systems were the first answer to the question of unalterable storage in the early to mid-2000s. The storage media were in stable cartridges, which the jukebox changed automatically as needed, and were considered durable and tamper-proof.

However, this technology soon showed several limitations. Different types of media were not compatible, which is why lengthy recopying of data was necessary when replacing them. The management of the partly outsourced media was as cumbersome as it was error-prone. It also soon became apparent that: the limited scalability of the jukeboxes could not cope with the growing amounts of data; and their access speed could not satisfy user requirements. These limitations eventually led to the replacement of the various jukebox versions from around 2005 onwards.

THE SECOND GENERATION - PROPRIETARY HARDWARE SYSTEMS

After jukeboxes, many companies invested in hard disk-based, proprietary archive appliances. These systems were much more scalable than jukeboxes and are still widely used as archive platforms today, although many solutions have now been discontinued (end-of-life).

However, as data silos, they also have limitations. A major disadvantage is that the archive data is only protected within the dedicated (black) box. Porting or migrating to new hardware is complicated, lengthy, and expensive, because the data usually has to be completely rewritten by the applications when the system is changed.

In addition, customers are strongly tied to a specific technology and thus to a manufacturer, as proprietary storage systems and APIs are used. Furthermore, the archive must be re-licensed with each new hardware generation. Since archive storage which has not yet been used also requires a license, considerable and recurring follow-up costs arise over the storage period. The large amount of time and effort required to operate and maintain these proprietary stand-alone solutions should also not be underestimated.

THE NEXT GENERATION & PERFECT WORLD SCENARIO

Due to the limitations and disadvantages of the previous archive storage solutions, the question arises as to what future-proof systems could look like. There are enough technologies and solutions, whether on-premises, in the cloud, hybrid, or object storage, just to scratch the surface of the possibilities. If long-term data storage and archiving were a dream and not a nightmare, what would the solution look like in a perfect world?

- The solution does what it is supposed to do: Store and protect data from different sources for the long term and meet all regulatory requirements.
- Sources of error are detected and corrected independently by the system.
- Users and administrators have no work with the system.
- Capacity grows with requirements and can be expanded easily and indefinitely.
- The total costs are transparent and, in particular, low.
- The system is robust, secure, and highly available.
- The solution is designed for future changes and is independent of specific manufacturers and technologies.
- Investments in software, hardware, and services are covered.
- IT can rely on the system and concentrate on the core business.

This scenario is closely interwoven with the idea of the software-defined approach, which is why we want to take a closer look at this principle. When choosing long-term storage, it is crucial for companies to remain adaptable and agile. As in other areas of professional IT, a software-defined solution approach (Software-Defined Archiving - SDA) offers flexible and future-proof options.

4. SOFTWARE-DEFINED STORAGE & ARCHIVING

THE PRINCIPLE

Long retention periods, the technological progress of storage systems, unavoidable data migrations, and rapidly increasing data volumes – With such conflicting priorities, the question arises as to how to handle business-critical data securely and economically. Which storage technology will prevail in the future? How can storage costs be controlled despite data growth? How can future unknown requirements be flexibly implemented? The only way out of this conflict lies in making the IT infrastructure more flexible and more adaptable.

Companies are already gaining important flexibility through virtualization and can exploit cost advantages through the demand-driven use of IT resources. Both virtualized environments in the

company's own data center and offers from external cloud service providers are playing an increasingly important role. Companies can also easily take advantage of these benefits in the area of archiving and long-term data storage - by using Software-Defined Archiving (SDA). But how does SDA work?

Software-Defined Storage separates the storage software from the underlying hardware. The architecture is generally designed for industry-standard or x86 systems, which are intended to avoid dependence on proprietary hardware and APIs. The decoupling of software and hardware allows for flexible expansion of storage capacity. The great benefit of SDS lies in the independence from specific (expensive) hardware, the far-reaching flexibility, and the scale-out capacity.

Similar to the established concept of Software-Defined Storage, Software-Defined Archiving is based on the combination of specialized software and industry-standard storage systems. Users gain long-term independence from specialized hardware (which forms the basis of many archiving solutions) through an additional software layer between their business applications and the infrastructure.

Due to their relatively rigid design, increasing the capacity of proprietary systems, for example, is usually associated with high costs. Software archive solutions, on the other hand, run on physical or virtual servers and enable storage according to the WORM principle independently of the storage hardware used. This means that they are not only easier to scale, but also require less effort. Upgrading the archive with more storage is possible even during operation, which minimizes downtimes. Moreover, even storage systems from different manufacturers can be combined to form an overall archive, as there is no dependence on hardware or a manufacturer. The storage layer becomes interchangeable, as the intelligence of the archive lies in the software layer.

In a nutshell: Software-Defined Archiving offers companies the opportunity to react flexibly to their increased archiving requirements without being tied to a storage manufacturer or the hardware.

ADDED VALUE OF A SOFTWARE-DEFINED APPROACH

The software-defined approach to storing and archiving business-critical data brings several advantages:

- No hardware lock-in for specific platforms
- Free choice of storage provider, technology, and model (independence)
- Archiving functionalities and intelligence are decoupled from the hardware
- No technical restrictions in terms of specific protocols or APIs
- Unlimited scalability through flexible licensing depending on actual demand
- Non-disruptive memory update and data migration
- Optimization of unused capacities

Software-Defined Archiving solutions not only offer technological advantages, but are also worthwhile from an economic point of view. With WORM-compliant data storage, which is independent of the hardware used, companies can use existing storage capacities more efficiently and protect their

investments in hardware, software, and know-how. In addition, the implementation effort is considerably reduced if the archive solution can be easily integrated into the existing IT infrastructure as Windows software or a virtual appliance, for example.

Using the Software-Defined Archiving approach, providers can realize transparent licensing models which, in contrast to hardware data silos, are closely oriented to actual customer needs. If the software archive solution is also able to store and manage data redundantly, a company does not need to invest in additional replication solutions at storage level.

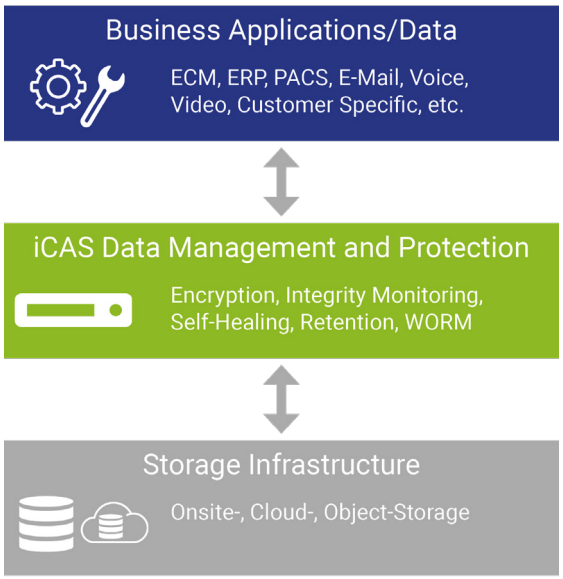
Doing away with complex, often oversized, and rigid hardware archive systems in favor of flexible, virtual environments with the best possible utilization has further advantages for companies: The space required in their own data center is reduced, as is power consumption. In addition, if access to archived information is fast and powerful, the productivity of employees and business processes increases.

ITERNITY SOLUTION OVERVIEW

iCAS Middleware

The software-defined solution iCAS offers a flexible and economical solution for the increasing demands on digital archives. The middleware protects and stores archive data from a wide range of business applications. All well-known ECM, ERP, DMS, PACS, and email systems have been certified for iCAS, so that even system or manufacturer changes do not pose a problem for the long-term existence of the archive platform. iCAS can be easily integrated into any IT infrastructures, supporting NAS, DAS, SAN, object, and cloud storage solutions. The middleware is also able to perform secure and seamless data migrations.

The integrated data replication of iCAS ensures high availability without additional mirror technologies. The archive can be replicated to different storage devices, whereby iCAS always checks the objects on the storage targets for their integrity. Optionally, the software encrypts all data to be archived according to the AES256 standard, so that the increasingly strict data protection requirements can be met.



iCAS is based on the latest Microsoft server technology and can be easily integrated into existing system landscapes. This means that the solution automatically benefits from the currently supported functions and improvements at operating system level. In addition, iCAS can be implemented as a virtualized, cost-efficient solution, for example on VMware.

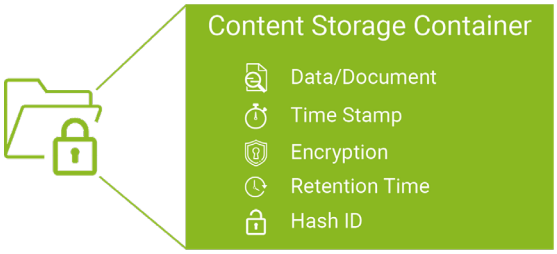
Applications are connected to the iCAS middleware via an application programming interface (API) or a file system interface. This makes iCAS compatible with CIFS/SMB and NFS. The software achieves reliable archive protection by providing functions such as WORM, retention management, encryption, self-healing mechanisms, and compression.

iCAS stores the data and documents to be archived together with associated meta information such as retention period, time stamp, and hash ID in self-sufficient archive containers. The patented and certified CSC technology (Content Storage Container) is used. The data and documents always carry all important additional information directly with them and are protected in the containers for their long journey. Migrations or storage changes cannot harm the data and its integrity. Data containers become unique objects by the software calculating a unique HMAC-SHA 512-bit hash value based on the content, which is also used for verification. Finally, iCAS stores the CSCs on a data carrier. This is where the main advantage of iCAS comes into play: the decoupling of the content-related objects from the storage medium.

It is important to note that iCAS is **not a document management system (DMS/ECM)**. Rather, the solution forms a functional software layer between the various business applications and the storage platforms used. For audits and the evidential value in court, data must be verifiably stored in an immutable manner. It is therefore not enough to simply *not* delete data from the DMS.
Read more about DMS and archive

Since the archive objects contain all important additional descriptive information, it is irrelevant which current or future storage technology is used for storage. Data verification is possible at any time via the content-related hash values.

The multi-client iCAS is able to distribute the archive storage per client to up to 128 different LUNs and archive paths, allowing users to build up even large archives. The flexible architecture of the system enables the archive storage to be expanded as needed during operation.



With their existing standard backup tools, companies can quickly back up and restore their archive – and so special WORM tapes and the necessary silo hardware are no longer required. The integrated backup optimization reduces the archive objects and further accelerates the backup. Thanks to CSC technology, the restored data is verifiably identical to the original versions.

iCAS FS Scale-Out Storage Platform

The iCAS FS platform is designed for maximum scalability and efficiency. The software-based architecture is built on standard hardware and a hardened Linux operating system. Thanks to WORM storage, retention management, encryption, and audit trail, iCAS FS also focuses on compliance and security.

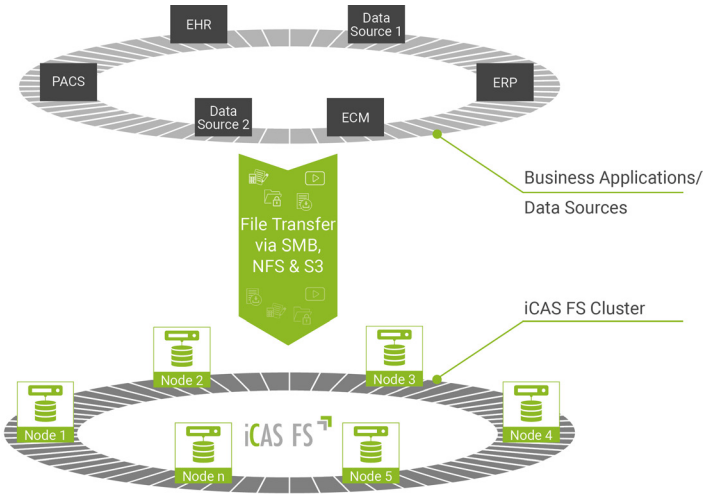
iCAS FS is a scale-out cluster for long-term data storage and archiving without any effort for IT. The complete set-up is monitored automatically - errors are detected and solved by the system itself. Thus, iCAS FS enables „Cloud Experience On-Premises“: high flexibility, user-friendly application, and low overall costs in your own data center.

iCAS FS is infinitely scalable from 20 TB and one node and expands the use case of audit-proof archiving. A wide range of requirements can be covered with the platform, such as:

- Secure backup target, e.g. as a verified „Veeam Ready - Repository“
- Relief of the primary storage
- Big Data storage, e.g. as a warm tier for Splunk applications

- Audit-proof archiving with WORM storage and retention management

The overall performance of iCAS FS is provided by a multitude of storage server nodes. This modular architecture enables capacity and performance to be expanded at will. Thanks to transparent read/write failover and active-active clusters, iCAS FS offers high availability for archiving. If a node is not available, the data is written to the other available nodes. As soon as the node is available again, the data is automatically synchronized and full redundancy is restored. iCAS FS switches between synchronous and asynchronous replication, depending on the latency of the data centers, and thus also enables stretched clusters over long distances.



All iTernity solutions are KPMG-certified. The requirements for audit-proof storage can be met in full, which is also confirmed by the certifications issued. This ensures the integrity of the data in compliance with the highest security and compliance standards.

iCAS and iCAS FS also have a self-healing function: The archive software is able to archive data synchronously on any two storage systems. The hash procedures used make it possible to continuously monitor replicated data, identify defective archive objects and then repair them independently. Creeping data corruption is thus a thing of the past and business-critical data remains available and valid in the long term.

The cost efficiency and benefits of iCAS FS were also scrutinized by the IT analysts of the Enterprise Strategy Group (ESG). The result: companies can reduce their overall costs by over 53% and IT expenditure by up to 76%.

You can read the complete ESG analyst report here

iTernity's archive and storage solutions are also available globally through the HPE Complete price list and HPE GreenLake. HPE GreenLake is a consumption-based as-a-service offering which combines the advantages of the cloud with the security and performance of on-premises solutions. The benefits are usage-based payment, reduction of total cost of ownership, and high agility - while maintaining data control and security.

5. CONCLUSION

The right strategy for long-term data storage and archiving is an important component in minimizing total cost of ownership (TCO) and business risks. Organizations in all industries are well advised not to rely on proprietary, hardware-based archiving systems for information which needs to be stored long-term. Solutions which are not directly technology-dependent and are based on open industry standards offer more flexibility and future-proofing. After all, no one can predict how storage technologies will evolve in the future, where rapid data growth will lead us, and what compliance requirements will be added.

With a software-defined approach, companies can streamline and simplify their archive and storage infrastructure while increasing performance and reliability. Business-critical data is protected in the long term, overall costs can be significantly reduced, and the archive and storage infrastructure is built on a future-proof and flexible foundation.

READ MORE

- Find out in our reference stories how companies benefit in practice from a software-defined approach
- Discover our software-defined archive and storage solutions iCAS Middleware and iCAS FS Scale-Out Storage Platform
- ESG Cost Comparison: Public cloud vs. On-premises with surprising findings and potential savings for long-term data storage
- Sign up to the iTernity newsletter to stay up to date with all news and events

We protect your business-critical data. The trust you place in us is our motivation and an investment in the future. The result: more security, less effort, no worries.

Our DNA is archiving, our mission the long-term availability and integrity of all types of corporate data. Our focus is on your challenges, whether data protection, cost pressure, data growth, cyber attacks, lack of time, or complexity – we take your data securely into the future.

CONTACT OUR EXPERTS

Heinrich-von-Stephan-Straße 21 | 79100 Freiburg | Germany
info@iTernity.com | +49 761 590 34 810 | www.iTernity.com