New Storage Strategies for Long-Term Data Retention

As data volumes dramatically increase, enterprises are in danger of being overrun by unstructured data. But the issue isn't having the capacity to store all that data, but rather the need to manage it properly and cost effectively for business value—especially in long-term retention use cases.

No one should be surprised to learn that digital data is growing at a breakneck pace, driven heavily by rapid proliferation of unstructured data such as files, video and social media. And unstructured data is growing much faster than other forms of data— upwards of 40-80% annually, depending upon the application and use case.[1]

While there are many factors contributing to these trends, one of the major catalysts is long-term retention (LTR) requirements in regulatory, legal and governance scenarios, including data protection

measures such as backups and archives. Just how "long term" is LTR? Consider the following data retention requirements in just these few instances:

- In the United Kingdom, internet service providers must keep communications records for a minimum of one year.

- U.S. federal guidelines on research records mandate that research data be kept at least three years after the research work has concluded.

- Under the Health Insurance Portability and Accountability Act, patient data must be retained a minimum of six years.

---

1   "Unstructured Data Growth Fueling Massive Migration to Object Storage," Enterprise Storage Forum, December 2017

VERITAS™ The truth in information.

TechTarget | **Custom Media**

Compliance and regulatory requirements aren't the only motivations for LTR of enterprise data. For instance, it's not unusual for aircraft manufacturers—or suppliers of other, long-life durable goods—to retain their records for 30 years or more, depending upon the lifetime of the product. And many organizations retain research and development records, including patent applications, for even longer periods of time.

So, it's easy to envision a scenario where organizations have to make tough choices about what data to keep and for how long—and more importantly, where and how to store it for easy and fast access, as affordably as possible.

This problem impacts many functions within the typical enterprise. For instance, backup administrators need to ensure all data is safely backed up in a cost-efficient manner, while risk and compliance managers want to avoid problems and penalties associated with not being able to retrieve relevant data for regulatory and legal requirements. But properly, efficiently and securely managing all that unstructured data for LTR must account for three major needs:

- **Reducing time** needed to restore archived data under ever-more-stringent service-level agreements, recovery-point objectives and recovery-time objectives.
- **Managing risk** associated with data that must be stored and produced for compliance, security, governance, legal and business continuity.
- **Improving the economic picture** by extracting insights from the data for analytics, customer experience and other customer-facing workloads, resulting in improved business opportunities.

As LTR requirements become more precise and more demanding, IT professionals find themselves under intense pressure to not only make sense of

their storage infrastructure, but also to determine how to mix and match technologies with different LTR use cases. Clearly, this piecemeal approach no longer can be justified from economic, management, manpower or data management perspectives. Maintaining the status quo for rapidly changing LTR workloads is no longer an option.

At the same time, storage administrators and IT executives agree that meeting LTR storage requirements can't be met using a one-size-fits-all model. Enterprises need to have storage solutions incorporating both on-premises environments and the cloud, in order to accommodate such issues as storage tiering, workload migration, multi-cloud architectures, tightening backup windows and storage management—all as cost efficiently as possible.

The need for flexible options

Taking a hybrid approach to LTR storage provides organizations with the greatest flexibility, cost efficiency and management simplicity. By combining on-premises storage with cloud, enterprises can save both CapEx and OpEx, reduce complexity and enable storage functionality that meets the needs of expanding workloads and more stringent use cases.

By employing a policy-based, automated tiering protocol to either public or on-premises private clouds—or to both—organizations have the greatest flexibility to handle LTR workloads. For instance, let's say you need to back up everything that's been stored for at least five years, but that you still need

access on an infrequent basis. An automated policy can move those files to the cloud if the context and cost requires.

Taking a hybrid approach—allowing organizations to select the right architectural, deployment and management models for their unique LTR workloads and use cases—reduces complexity by aligning storage strategies in backup, archiving, recovery and retention with the needs of each enterprise, often down to the business unit or workgroup level.

What to look for in a solution

If your storage architecture relies too extensively on primary storage to achieve high capacity, fast throughput and low latency, you'll likely overpay for those benefits for a significant part of your LTR workloads. But if you shift too much of your LTR requirements to tape libraries or off-premises public cloud services, performance bottlenecks will occur and producing essential data for anything from a compliance mandate to an e-discovery demand will become a huge headache. This could easily result in being out of compliance with service-level agreements or not meeting recovery-time objectives—or worse, such as not making data available at all if a tape can't be read.

That's why you need solutions that are flexible enough to accommodate different LTR requirements

at different times, while balancing demands for availability, performance and security with smart spending.

A particularly cost-efficient model for supporting LTR workloads is a purpose-built appliance.

By using an appliance, storage administrators can deploy a turnkey solution for LTR, as well as for use cases such as tape replacement and archiving.

Storage appliances are a flexible way to fill the cost/performance gap between increasingly expensive primary storage—especially for LTR workloads that do not always require instantaneous access around the clock—and lower-performance, offline storage options such as tape libraries and public cloud storage. Appliances also are easy to deploy and help avoid long-term technology lock-in to proprietary, vendor-specific architectures that rob organizations of agility and flexibility.

Ideally, an LTR-aimed appliance helps seamlessly migrate data to and from different infrastructure environments, such as between on-premises storage to the cloud, from the cloud back to on-premises or between multiple clouds. It's also important for the appliance to be tightly tied to core storage management functions and applications, such as backup, archiving, disaster recovery and automated storage tiering.

An appliance model allows organizations to easily migrate data using such protocols as object, block, file, NFS and CIFS, while dynamic storage tiering and policy management can be configured to move aged data to the most cost-efficient storage medium—without manual intervention.

Storage and IT administrators evaluating storage appliances should look for key features to support LTR workloads straining under the weight of dramatically increasing unstructured data volumes, including:

- Cost-optimized storage.
- Pre-tuned, tested and installed operating system.
- Embedded security.
- Simple hardware and software monitoring.

Another important feature to look for is support for software-defined storage, which de-couples software from the underlying hardware to ease management and to facilitate continuous hardware upgrades without impacting software licenses. Over time, this can even allow organizations to deploy "do it yourself" hardware upgrades for load rebalancing or routine technology refresh.

You'll also want a solution that supports the ability to work among multiple public cloud service providers and/or an on-premises private cloud. LTR workloads not only are always expanding, but they also are constantly in need of re-evaluation of the most cost-efficient venue for the data. LTR data may need to move from on-premises to a private cloud or to any number of different public clouds, so you want to make sure your solution's cloud tiering capabilities are flexible, automated and policy-driven.

Conclusion

Storage has become a critical element in an organization's digital transformation and infrastructure modernization strategies, but the sheer volume and complexity of storage, combined with mounting regulatory demands, has put enormous pressure on organizations to take a second look at their approach to long-term data retention.

To overcome the challenges of cost, complexity and risk, enterprises should look for a more flexible set of options to manage LTR storage. IT and storage administrators need to be able to select from among such options as software-defined storage architecture, purpose-built storage appliances and cloud tiering to architect their LTR storage solutions.

Veritas provides enterprise decision-makers with a number of different cost-effective storage options that fill the gap between expensive primary storage and lower-cost but potentially complex solutions such as tape libraries and cloud storage.

For more information, please visit www.veritas.com.

About Veritas

Veritas Technologies empowers businesses of all sizes to discover the truth in information—their most important digital asset. Using the Veritas platform, customers can accelerate their digital transformation and solve pressing IT and business challenges including multi-cloud data management, data protection, storage optimization, compliance readiness and workload portability—with no cloud vendor lock-in. Eighty-six percent of Fortune 500 companies rely on Veritas today to reveal data insights that drive competitive advantage.