

Solution Brief**RainStor + Caringo**

Big Data Retention

For Structured and Unstructured Data

RainStor + Caringo Benefits

- Unified storage and on demand granular query for multi-structure data with no restore or re-inflation
- Software based, use any type or combination of x86 hardware
- Massive scalability with commodity economics
- Automated, metadata driven data lifecycle management at the object and record level
- WORM capable with auditable access for compliance with regulatory or governance policies

Store and query *any* data type on the same storage system

Managing disparate storage solutions for structured data (databases, log files, text, value based data) and unstructured data (audio, documents, emails, images, video) has always been a challenge for IT Organizations. Historically, this issue was manageable by pushing information to low cost and inaccessible archives, however this approach doesn't provide granular query access and increases complexity by adding backup/restore processes that lengthen as data sets grow. Enterprises now require both types of information to remain accessible and online for years and sometimes decades with data growth rates rapidly outstripping budgets and resources. It is clear that just archiving on tape or other forms of inaccessible storage or tying up expensive SAN and NAS devices doesn't work anymore.

What if you can store unstructured, semi-structured, and structured data on the same storage system? And what if this data were to remain queryable on demand for as long as you need on any combination of standard x86 server hardware that can scale by the terabyte to petabytes, and billions of files at the lowest ongoing TCO of ownership aligned with your business needs.

A new way to retain data while allowing access at the lowest cost

Today large amounts of structured data from relational databases or data warehouses that are no longer being updated or modified by applications are typically removed from their production environments in an effort to improve the performance and reduce the cost of operating those applications and systems. Typically the data is either archived to tape or moved to a tier of storage that requires that the data be re-instated to its original production system before queries or reports can be executed. This makes the process of accessing the information cumbersome, prone to error and ultimately costly, since the expense of accessing data quickly outweighs the initial low cost of archiving it.

Additionally the era of Big Data and machine-generated data from logs, sensors and other devices means that traditional technologies are not able to keep up with rapidly growing data volumes and retention requirements. Furthermore, since the data is immediately historical and does not require any updates, using high-performance relational databases is not only overkill, but also a bottleneck to the main objective of efficiently storing the data while still allowing the necessary queries to analyze and report on the data.

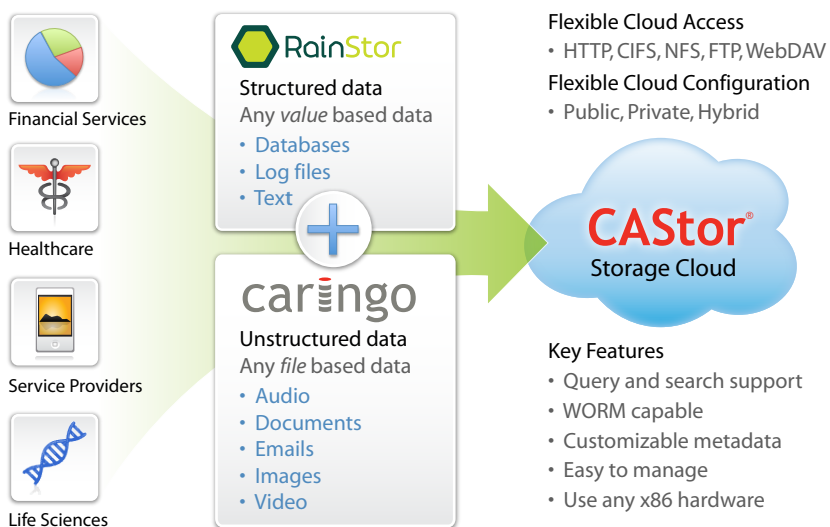
RainStor's Big Data software combined with Caringo CASTor Object Storage software provides a long-term data solution to store and query a mix of data types within a single, unified storage system.

RainStor + Caringo

Flexible and accessible long-term storage

Caringo® Inc., a global leader in object storage, and RainStor™, a provider of Big Data management software, have partnered to deliver a solution for retaining and accessing unstructured, semi-structured, and structured data in Caringo object storage software powered by CASTor. The combined solution provides a horizontal platform that enterprises can use to increase the efficiency and dramatically reduce the cost of their infrastructure required to retain and manage data. The solution is ideal for use cases where database data must be stored with file-based data for and remain accessible for long periods of time such as in:

- **Financial Services:** Store complete customer account information including stock tick, trade data, transactional information, and ATM/check images together for compliance purposes or dispute resolution.
- **Healthcare:** Store patient medical images as well as patient documents and structured database medical records for years or decades for continued patient care.
- **ISVs and Service Providers:** Store server logs, customer usage logs and customer files for a variety of cloud service implementations on infrastructure that scales by the terabyte with no-downtime expansion and maintenance.
- **Life Science:** Store value based research information with file based documents and images in a highly scalable solution with commodity economics.



Footprint reduction with rapid search

RainStor's Big Data Management database contains patented compression and de-duplication techniques, reducing the data footprint size by 95% or more compared to traditional approaches. For example, data stored in a typical relational database is bloated by system indexes that can take up an additional 60% of storage. This means that 5 terabytes of user data might actually utilize 8 terabyte of high performance expensive storage. RainStor requires no indexes and a compression rate of 95% results in a 20 times reduction of the 5 terabytes user data footprint to just 250 gigabytes on CASTor running on lower cost storage. With these capabilities, RainStor provides the perfect complementary database to production systems needing to "skinny down" while keeping data accessible. In addition the combined solution provides a highly scalable primary store for the capture and reporting of massive volumes of machine-generated big data. Unlike other forms of binary compression, data retained in RainStor can be stored and queried on-demand at a granular level via SQL or standard Business Intelligence tools without the need for re-inflation.

Big Data Retention for Structured and Unstructured Data

RainStor

45 Belden Place, 2nd Floor
San Francisco
CA 94104
www.rainstor.com

Caringo, Inc.

6801 North Capital of Texas Highway
Building 2, Suite 200
Austin, Texas 78731
www.caringo.com

Massively scalable, easy to manage storage with commodity economics

CAStor is a software-based storage solution that installs on any combination of x86 hardware with up to 98% disk utilization. Employing a symmetric architecture means that all storage nodes are the same so there are no bottlenecks or single points of failure. This also means that expansion can happen a single drive at a time. There are no limits in capacity, object count or object size. With integrated automated optimization, data protection and zero-provision expansion, one system administrator can easily manage over 10 PB of storage.

Automated data protection and compliance

Integrated data protection and compliance functionality enable adherence to even the strictest corporate policies or industry regulations. CAStor utilizes configurable metadata driven retention and expiration policies per object and is WORM capable supporting immutable objects with auditable access. The integrated Health Processor continuously checks the integrity of data and all nodes participate in a rapid, self-healing recovery process if an issue is detected. RainStor adds to CAStor's capabilities with granular configurable retention and expiry policies down to the record level for structured data, thereby allowing flexibility in data retention beyond an object level where appropriate.

Reduced complexity and costs

The combined RainStor and Caringo solution delivers unified storage that enables organizations to break free from the reliance on expensive databases and storage devices that are both overkill for multi-structured information that is no longer being updated. Decisions on data retention can now be driven by the value of the data itself and not by the affordability of storage.

For more information on RainStor visit www.rainstor.com and for more information on Caringo visit www.caringo.com.