

Datos IO RecoverX

Datos IO provides the industry's first cloud-scale, application-centric, data management platform enabling organizations to protect, mobilize, and monetize all their application data across private cloud, hybrid cloud and public cloud environments.

To learn more, visit www.datos.io



Key Features

- Elastic scale-out architecture
- Parallel streaming data movement
- App-consistent backups
- Table level backup & recovery
- Any point-in-time recovery
- Globally distributed metadata catalog
- Semantic Deduplication
- Enterprise policy management
- NFS/Object Storage support
- Broad application support
- Automated deployment



Data Sources Supported

- Non-relational databases
 - MongoDB
 - Apache Cassandra
 - DataStax Enterprise
- Big Data Filesystems
 - Apache HDFS
 - Commercial Distributions (Cloudera, Hortonworks)
- Relational databases
 - Microsoft SQL Server



Overview

Enterprises are increasingly adopting next-generation applications and migrating traditional applications to multi-cloud environments. As a result, enterprise IT infrastructure now consists of multiple clouds (public, private and hybrid) in distributed geographies, all connected over multiple networking links. However, every cloud uses unique technology for infrastructure services. For example, private clouds are predominantly powered by the VMware ESX virtualization suite, while AWS and Azure public clouds use Xen and Hyper-V for their virtualization technology. Accordingly, there are no ESX virtual machines and SCSI LUNs in the public cloud, and the only common denominator binding all the clouds together is the data itself, which does not change across VMWare, AWS and Azure.

Because of these shifts, customers now need a comprehensive Cloud Data Management solution.

Datos IO addresses this challenge with the industry's first cloud-scale, application-centric, data management platform, RecoverX, enabling organizations to protect, mobilize, and monetize their application data across private cloud, hybrid cloud and public cloud environments. RecoverX is built on top of our seminal data management architecture called Consistent Orchestrated Distributed Recovery (CODRTM) engine, which is not dependent on media servers and transfers data in parallel to and from file-based and object-based secondary storage. The architecture is fully

distributed in nature, which provides high availability in failure scenarios and uses elastic compute resources for scalable performance. Using this architecture, RecoverX delivers application-aware data management that allows massive storage efficiency and sub-table level recovery/mobility solutions at scale for traditional and next-generation applications. The key use cases that RecoverX addresses are: point-in-time backup & recovery, test/dev refresh, cross-cloud data mobility, and archival / long-term retention – either in the cloud, on-premise / private clouds, or for hybrid cloud environments.



Benefits

Unlike legacy solutions, RecoverX is built from the ground-up to address the requirements of cloud-first applications. This allows RecoverX to deliver massive benefits for enterprises.

- Software-only elastic architecture allows private cloud, public cloud and multi-cloud deployment
- Fully orchestrated and granular recovery reduces application downtime by ~4-6x
- Industry-first semantic deduplication results in 80-90% storage cost savings
- Native failure handling provides operational resiliency when there are node or database failures
- Application-awareness enables automated refresh of test/dev environments and results in ~8x operational efficiency
- Application-awareness enables universal data mobility across private and public cloud
- Globally distributed metadata catalog allows for backup anywhere, recover anywhere, and migrate anywhere



Product Features

Datos IO RecoverX provides data protection, single-click granular recovery, industry-first semantic deduplication and cross-cloud data mobility for next-generation non-relational databases (e.g. Cassandra, MongoDB), big data filesystems (e.g. Apache HDFS and commercial distributions from Cloudera) and relational databases (e.g. Microsoft SQL Server). RecoverX uses semantic deduplication to achieve massive storage efficiency, thereby, reducing capital costs at scale. The enterprise rich policy management capabilities give database administrators (DBAs) and storage / backup administrators operational leverage to create workflows in line with their application and workload requirements. Cross cloud data mobility allows universal visibility and portability of data and results in massive operational efficiency.

RecoverX can be deployed on physical servers, virtual machines or any compute instance in the cloud as a single node or clustered software (3 nodes or 5 nodes) configuration. The infrastructure requirements are minimal and a wide variety of backup storage – NFS or Object – are supported. Clustered configuration provides high-availability and high performance to handle massive scale environments (100TB+). Finally, in-line with cloud-first design methodology, RecoverX performance also scales elastically with the underlying infrastructure (compute and memory) that is provided. This allows for dynamically changing the infrastructure footprint to cater to the changing application workload to get requisite data protection performance from RecoverX.

Features	Cassandra Database	MongoDB Database	Hadoop Filesystem	Microsoft SQL Server
Backup Granularity	Table	Collection	Directory	Table
Backup Interval	15 mins and above			
Recovery Granularity	Sub-Table	Collection	File	Table
Recovery Options	Fixed point-in-time Any point-in time Time-range	Fixed point-in-time Any point-in-time	Fixed point-in-time	Fixed point-in-time Any point-in-time
Recovery to different topology	Yes; for test/dev use cases			
Cross Cloud Mobility	Yes; for test/dev, compliance, cloning, migration use cases			
Secondary Storage	NFS, AWS S3, Google Cloud Storage			
Advanced Features	Compaction handling for storage efficiency TTL data handling for restore	Continuous data protection	File-level deduplication	Advanced compression for 10x storage efficiency
RecoverX Deployment	Single node or clustered (3 or 5 nodes)			
User Interface	Graphical User Interface, CLI, Restful API			
Supported Versions	Apache Cassandra 2.0, 2.1, 3.X DSE 4.5 - 5.1	MongoDB 3.0, 3.2, 3.4	CDH 5.8+ HDP 2.5+	SQL Server 2012+

Multi-cloud is the new normal, and to keep pace with this cloud transformation, enterprises must adopt a cloud-first data management strategy. Datas IO's mission is to help organizations accelerate their adoption of multi-cloud by enabling them to protect, mobilize, and monetize their traditional and next-generation applications. Datas IO has leapfrogged the data management industry, creating the world's first hyper-scale, distributed data management platform – Datas IO RecoverX. Datas IO RecoverX is elastic, scale-out software for multi-cloud environments, delivering efficient data protection and data management services for traditional and next-generation applications.



About Datas IO

Datas IO is the application-centric data management company for the multi-cloud world. Our flagship Datas IO RecoverX delivers a radically novel approach to data management helping organizations embrace the cloud with confidence by delivering solutions that protect, mobilize, and monetize their data — at scale. Datas IO was recently awarded Product of the Year by Storage Magazine, and was recognized by Gartner in the 2016 Hype Cycle for Storage Technologies. Backed by Lightspeed Venture Partners and True Ventures, Datas IO is headquartered in San Jose, California.