

Highly Scalable Data Management for Hadoop

Hadoop has enabled companies to deploy large-scale, distributed applications that help enterprises identify a number of things, including customer shopping patterns, tackle potential fraud, and even process human genome data across hundreds of terabytes and even petabytes of data. Although Hadoop provides basic data management features (e.g. multiple replicas to protect against hardware failures), these features do not protect against data loss resulting from user errors, ransomware, or application corruption. Loss of data and increased downtime can severely impact your revenue and business reputation.

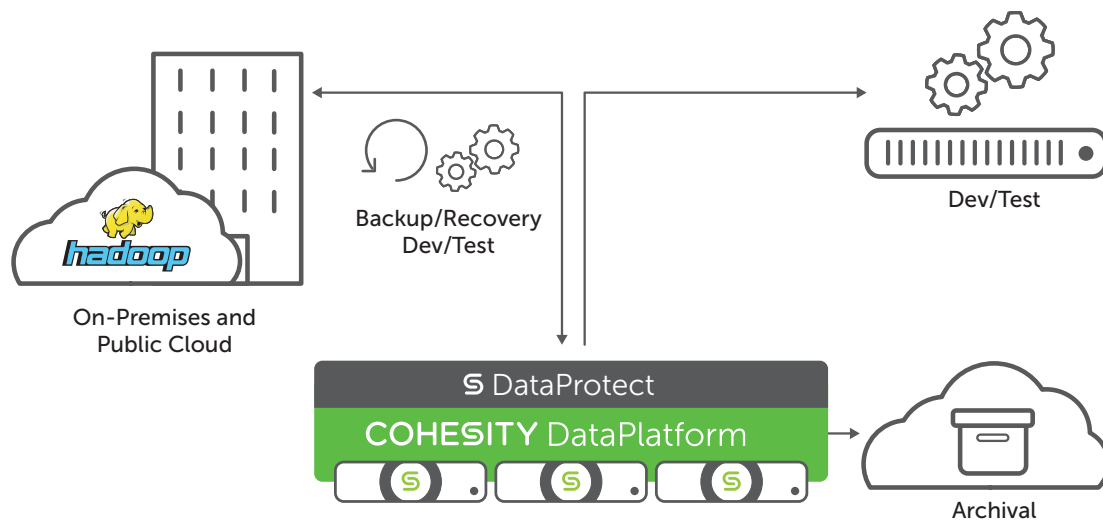
Organizations require an enterprise-class data management solution that spans across multiple use cases: self-service access to data, sophisticated backup and granular recovery, and a simplified and economical way to archive older data. Native and homegrown solutions cannot handle these requirements. Companies need a new approach to manage data on their Hadoop platforms.

Simplified Management and Protection

Cohesity provides the industry's fastest data management software with the ability to handle massive data sets – terabytes, petabytes and beyond – residing on a Hadoop distributed file system. Cohesity ensures data resiliency in the event of disasters or corruption, enabling companies to get back online faster. Cohesity backs up, clones, and recovers terabyte and petabyte-sized data sets and beyond faster than any other solution on the market, minimizing the impact of data loss associated with human and application errors and reducing downtime to minutes and hours, as opposed to days and weeks. Cohesity offers extreme scale, rapid recovery, and smart storage optimization, one of the many reasons why Cohesity is used by the leading Fortune 500 businesses in retail, financial services, and healthcare industries.

KEY BENEFITS

- *Enterprise-class data management and protection*
- *Rapid recovery*
- *Limitless scalability for petabytes and beyond*
- *Make backup data productive*
- *Reduce data footprint with Content-aware storage optimization*



Key Features

- **Extremely Scalable**, using a scale-out architecture that can support a few terabytes to petabytes of data
- **Flexible Backup and Recovery** allows backup and recovery of individual tables, databases, or the whole database to the original or alternate location of a different size or configuration
- **Rapid Recovery** leverages the scale-out architecture, fully materialized restore points, so enterprises can quickly locate the backed-up data and restore rapidly.
- **Cloud Native architecture** supports deployment on various Cloud platforms, including AWS, Azure, and Google Cloud.
- **Database Aware** means backing up & recovering the schema, metadata, and the data in a consistent manner for Hive & Hbase
- **Content-aware Storage Optimization** with compression, de-duplication, and erasure coding, saving significant resources and cost.
- **Enterprise-class solution** including key security capabilities (Kerberos, LDAP integration, RBAC, encryption, masking), auditing, alerting, high availability with erasure coding, and call-home technologies
- **Network Efficient** reduces network costs by employing techniques such as throttling, incremental-forever transfers, and compression.
- **Heterogeneous Backup** enables the backup of multiple applications and databases using a single user interface.
- **Consolidate Backup & Cloning to simplify data** management on a single platform.