

# LiveData Migrator for Azure

Automatically migrate from on-premises HDFS to Azure with zero downtime, zero data loss and best time-to-value

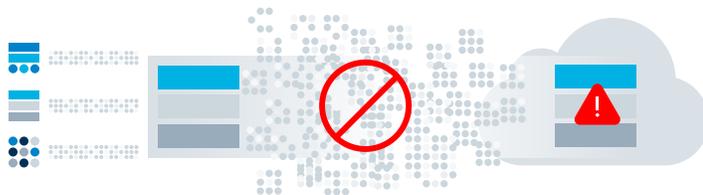


Business success now depends on how effectively organizations utilize their data. To do so companies are modernizing their data architecture, which means bringing petabytes of constantly changing data to work in the cloud. As Microsoft continues to add more robust analytics and data services to Azure, it has become one of the fastest growing cloud platforms.

While the move to cloud is making companies more competitive, lean and nimble, the transformation is fraught with risks of disruption to the existing business, data inconsistency issues, and risk of data loss. You need a solution that can let you maintain business operations while you migrate massive volumes of data efficiently. LiveData Migrator for Azure is a new core service on the Microsoft Azure platform enabling automated data migration at scale and with a customer experience similar to other native Azure services.

## Manual Cloud Migration

This common approach is based on custom programming to copy data to the cloud. Manual cloud migration includes incomplete strategies such as change data capture, which has no visibility of existing data, only of changes, tools to copy static data such as DistCp, which do not account for changes being made to the data under migration, or dual-estate architectures which require that your applications are modified significantly to perform ingest into both environments.



### Manual Data Migration Risk

- **Business Disruption** – Large data sets will take time to migrate to the cloud. This time lag makes it challenging to accurately bring continuously changing large scale data sets to the cloud without causing unacceptable amounts of downtime.
- **Data Inconsistency** – Data that is actively changing during migration is difficult to validate to ensure it's accuracy. Manual reconciliation at scale is costly and does not guarantee consistent data outcomes.
- **High IT Resources** – The overhead of activities to attempt a nondisruptive, no-downtime big data migration are significant. It requires resources to create, test, manage, schedule and maintain custom migration scripts, and validate ongoing results. Manual migrations often exceed budgets, are prone to delays, or fail altogether.

## Automated Cloud Migration

LiveData Migrator for Azure eliminates the risks of big data migration to the cloud by continuously migrating your data to Azure Data Lake Storage (ADLS) Gen2.



### Automates Data Migration to Azure Cloud

- The native integration of LiveData Migrator for Azure with other Azure resources simplifies deployment on Azure and provides a customer experience equivalent to other native Azure Services.
- LiveData Platform for Azure automates your Hadoop to Azure cloud migration at scale with no application downtime and no risk of data loss, even when your data sets are under active change.
- As changes occur anywhere in the source system, LiveData Migrator for Azure migrates the changes and applies them to the target environment.
- Additionally, LiveData Migrator for Azure minimizes IT resources with a non-intrusive migration solution that requires zero changes to applications, cluster, node configuration or operation.
- LiveData Migrator for Azure works in conjunction with LiveData Plane for Azure to establish an active-active data plane consistency between an enterprise's on-premises Hadoop and ADLS Gen2 environments. This enables enterprises to deploy a true hybrid data estate and LiveData environment.



## LiveData Migrator for Azure Business Benefits

### Business Continuity

- 100% data availability during migration; no business disruption
- Immediate availability of migrated data
- High scalability and performance for migration at any scale

### Complete and Continuous Migration

- Data migration with single pass of source storage
- Ongoing migration of any subsequent data changes
- Ensures zero data loss of source data and changes

### Cost Avoidance/IT efficiency

- Minimizes the need for IT resource involvement
- Automated data migration
- No custom code maintenance
- Faster time-to-value and adoption of AI & ML initiatives
- Works in conjunction with LiveData Plane for Azure for active-active replication enabling a hybrid data estate
- Support for phased migration, no "Big Bang" cutover for users or applications

## LiveData Migrator for Azure – Why It's Different

- **Core service within Microsoft Azure:** Running on top of the LiveData Platform for Azure, which is natively integrated with Azure resources enables automated deployment of the LiveData Platform for Azure services, and with an equivalent user experience and manageability via Azure Portal
- **Support for Native Azure Security and Manageability:** Leveraging Azure features such as Role Based Access Control, Active Directory, Azure Policy enforcement and Activity Log integration
- **Billing Integration:** Customers are billed through Azure, eliminating the need for them to add a new vendor contract or require additional vendor approvals
- **Hadoop Support:** Works across a variety of big data source environments, including all major Hadoop distribution vendors
- **Any Scale:** Migrates big data sets at any scale to Azure Data Lake Storage Gen2 without needing to halt changes made to the data sets during migration
- **Selective Migration:** Allows selection of which data sets should be migrated. Administrators define migration policies that control what data is migrated between clusters and selectively exclude data from migration to specific clusters in the Azure environment, or move it off to be archived
- **Complete and Continuous data migration:** Migrates existing data sets with a single pass through the source storage system, eliminating the overhead of repeated scans, while also supporting continuous migration of any ongoing changes from source to target
- **Immediate Availability:** Enables data to become immediately available for use in the target environment as soon as it has been migrated, without having to wait for all data set migrations to complete
- **Interfaces and API:** Migrations can be managed through a comprehensive and intuitive command-line interface, or from the self-documenting REST API, allowing users to track and monitor migration progress

## LiveData Migrator for Azure Use Cases

### Adopt Azure Data Lake Storage Gen2

ADLS is the only cloud storage service that is purpose-built for big data analytics. It is designed to integrate with a broad range of analytics frameworks enabling a true enterprise Data Lake, maximized performance via true filesystem semantics, and scalability to meet the needs of the most demanding analytics workloads.

### Cloud Migration Without Downtime

Non-disruptive automated migration reduces the risks involved with large data migration initiatives and allows your users and systems to continue operating while migration is underway.

### Cloud Analytics

Bring on-premises Hive workloads to Azure HDInsight, which provides a managed, full-spectrum, open-source analytics service in the Azure cloud.

### Hybrid-cloud Enablement

Maintain a hybrid architecture where on-premises and cloud environments are kept continuously in sync for as long as necessary.

### Disaster Recovery

Enterprises that wish to decommission certain environments that have to be maintained for regulatory purposes can utilize the cloud as the new secure repository.

### Backup

Taking advantage of cloud economics to maintain scheduled backups.

### Cloud Burst

Enabling Enterprises to access compute and storage in the cloud when required rather than building infrastructure for peak capacity.

WANDISCO is the LiveData company. WANDISCO solutions enable enterprises to create an environment where data is always available, accurate, and protected, creating a strong backbone for their IT infrastructure and a bedrock for running consistent, accurate machine learning applications. With zero downtime and zero data loss, WANDISCO LiveData Cloud Services keep geographically dispersed data at any scale consistent between on-premises and cloud environments allowing businesses to operate seamlessly in a hybrid or multi-cloud environment. WANDISCO has over a hundred customers and significant go-to-market partnerships with Microsoft Azure, Amazon Web Services, Google Cloud, Oracle, and others as well as OEM relationships with IBM and Alibaba. For more information on WANDISCO, visit [www.wandisco.com](http://www.wandisco.com).

5000 Executive Parkway, Suite 270  
San Ramon, CA 94583

[www.wandisco.com](http://www.wandisco.com)

### Talk to one of our specialists today

**US** +1 877 WANDISCO (926-3472)  
**EMEA** +44 (0) 114 3039985  
**APAC** +61 2 8211 0620  
**All other** +1 925 380 1728

Join us online to access our extensive [resource library](#) and view our webinars.

### Follow us to stay in touch

