

Data Migrator for Azure

Microsoft's preferred and native solution for Hadoop-to-Azure data migrations with zero business disruption and reduced business risk

Increasingly, organizations are migrating what they have implemented on-premises using Hadoop, Spark, and other technologies to more innovative solutions in the Azure cloud. Yet unlike building cloud-native applications from scratch, organizations that want to move from their existing Hadoop deployments to new solutions in Azure are faced with significant challenges. These challenges include the scale of the data that needs to be migrated, the amount of data changes occurring (ingests, updates, deletes), and the critical nature of these environments, which can't be disrupted.

We make data migration non-disruptive

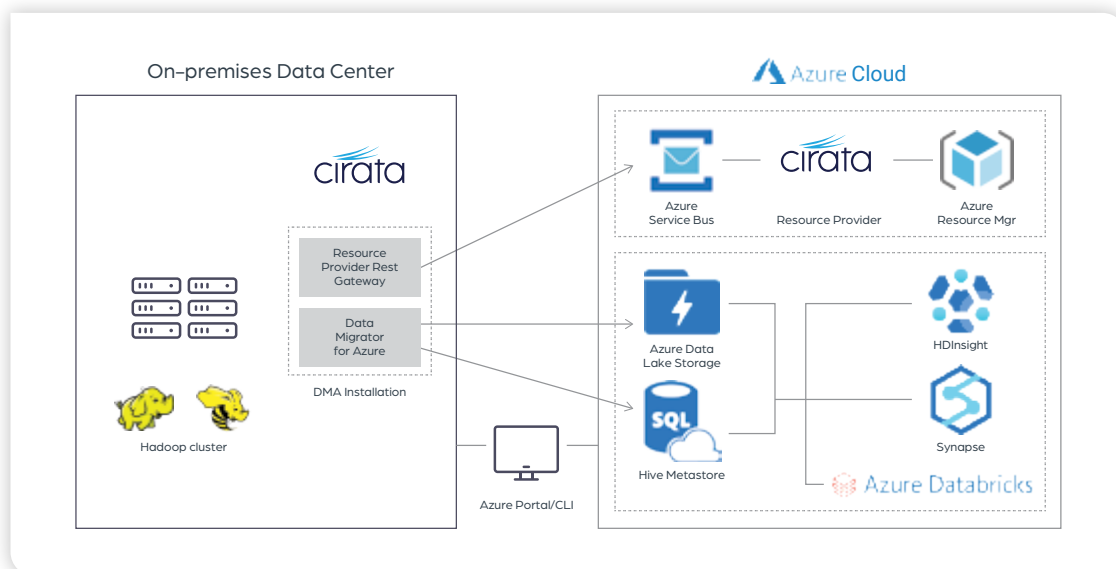
Cirata's unique approach to migrating data at scale without disrupting the use of those datasets while an organization adopts cloud infrastructure and services has been a critically important answer to these challenges.

Data Migrator for Azure is a native Azure service that enables users to migrate petabyte-scale Hadoop data and Hive metadata to the Azure cloud with zero application downtime

and zero risk of data loss, even while the source data is under active change. With Data Migrator for Azure, you can deploy and manage your data lake migrations using the same Azure management experience you enjoy today through the Azure Portal and Azure CLI.

We reduce business risks

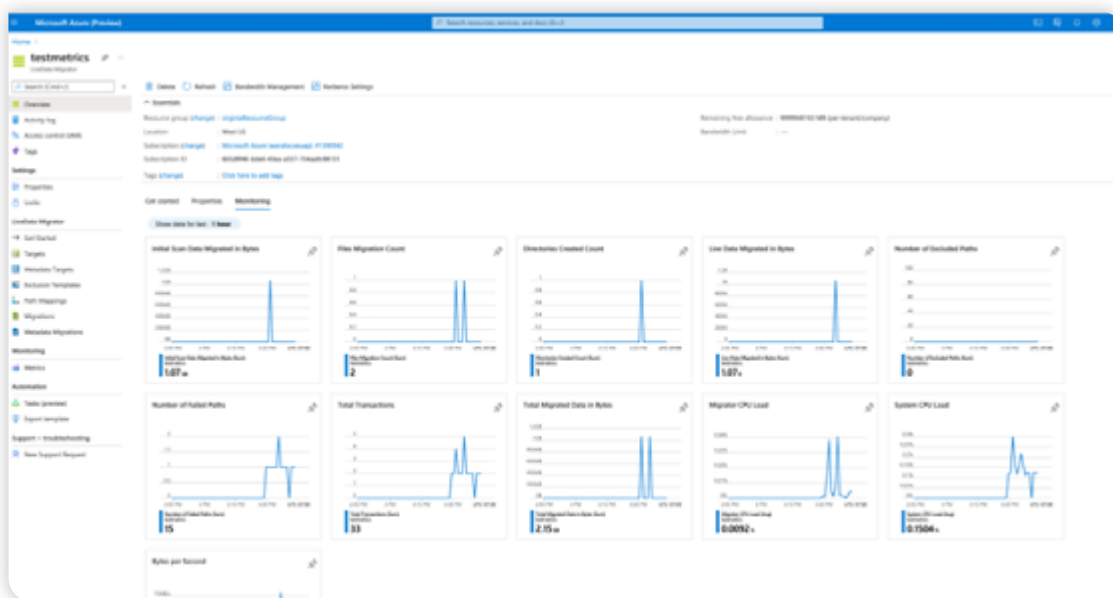
According to Bloor Research, more than 80% of all data migration projects run over time and budget. Migrating large data volumes with traditional approaches, such as transfer devices or DistCp (distributed copy), requires disrupting the operation of on-premises applications and doesn't cater to data that are modified or created during migration. Reconciliation at scale is costly and does not guarantee a completely consistent data outcome. The overhead required to achieve non-disruptive, no-downtime big data migration is significant due to repeated scans, systems out of sync, and manual intervention for anticipated failures and interruptions. Cirata Data Migrator for Azure minimizes these business risks and enables successful data migration with no business disruption and best time-to-value.



Data Migrator for Azure capabilities

- **Core Service within Microsoft Azure:** Deep integration with Azure resources enables Data Migrator for Azure to be deployed at the same time as other native Azure services and with an equivalent user experience.
- **Support for native Azure security and manageability:** Data Migrator for Azure leverages 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 you to add a new vendor contract or require additional vendor approvals.
- **Quick deployment and operation:** The Data Migrator for Azure resource can be created directly from the Azure portal, and the Data Migrator for Azure service is installed on your chosen source host(s). Deployment can be performed in minutes without impacting current operations, so users can begin moving data immediately.
- **Complete and continuous migration:** Migration of the selected data sets are performed 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, with zero disruption to current production systems.

- **Support for data and metadata transfer:** Data Migrator for Azure supports the transfer of unstructured datasets to Azure Data Lake Storage Gen2, as well as optional metadata transformation to various metadata targets such as Azure Databricks and Azure SQL Database.
- **Migration at any scale:** Data Migrator for Azure migrates data sets at any scale, from terabytes to multi-petabytes, without impact to current production environments. Begin risk-free for small migrations and scale up to multi-petabyte initiatives without needing any additional installation requirements.
- **Azure Portal and CLI Data Extension:** Users can manage the full data migration directly from the Azure portal. Additionally, Data Migrator for Azure can be configured and operated from the Azure CLI.
- **Configurability and control:** You will have the ability to configure the migrations to meet your specific needs. Data Migrator for Azure includes standard configuration – such as defining sources, targets, and data to be migrated or excluded – as well as advanced capabilities such as path mapping, scheduling and network bandwidth management controls.
- **Metrics and monitoring:** Data Migrator for Azure enables hands-off operations by providing information to keep you updated on the migration jobs, from health and status metrics that provide estimates for migration completion, files transferred over time, excluded paths, items that failed to transfer, as well as other real-time insights regarding usage.



Data Migrator for Azure business benefits

This turnkey solution enables automated data transfer at any scale with zero business disruption, minimized risk, and best time-to-value.

Business continuity

- Automated data transfer of changing data
- No downtime or business disruption
- Immediate data availability

Scalability

- Supports data transfer at any scale
- Horizontal scaling with multiple transfer agents
- Maximizes use of available network bandwidth

Cost and risk avoidance

- Fully automated data transfer
- No custom coding nor application changes
- Minimizes need for IT resource involvement

Use cases

Cloud data migration

Fully-automated data migration with zero disruption allows your users and systems to continue operating while migration is underway. Migrate to Azure Data Lake Storage, the only cloud storage service purpose-built for big data analytics.

Hybrid data estate

Maintain on-premises assets for as long as necessary. Organizations can maintain critical on-premises applications while expanding their investment and innovation in the cloud.

Cloud-burst

Enables enterprises to access compute and storage in the cloud when required, rather than building out on-premises infrastructure for peak capacity.

Cloud analytics

Once in ADLS, the data is available to Azure analytics services such as HDInsight, Synapse, and Azure Databricks. Optionally transform the data and metadata directly to the Delta Lake format on Databricks so that it becomes immediately available in Databricks' preferred form.

Disaster Recovery

Enterprises that wish to decommission costly on-premises environments that sit idle and are only maintained for regulatory purposes can utilize the cloud as the new secure repository.