

# Data Migration Using Azure Data Factory

## Team Members:

- Aravind Reddy Ramidi
- Yogi Siva Geethika Chekuri
- Bindu Yaddula
- Radha Gude
- Venkata Sai Charan Karasala
- Srikanth Piniseti

# Content

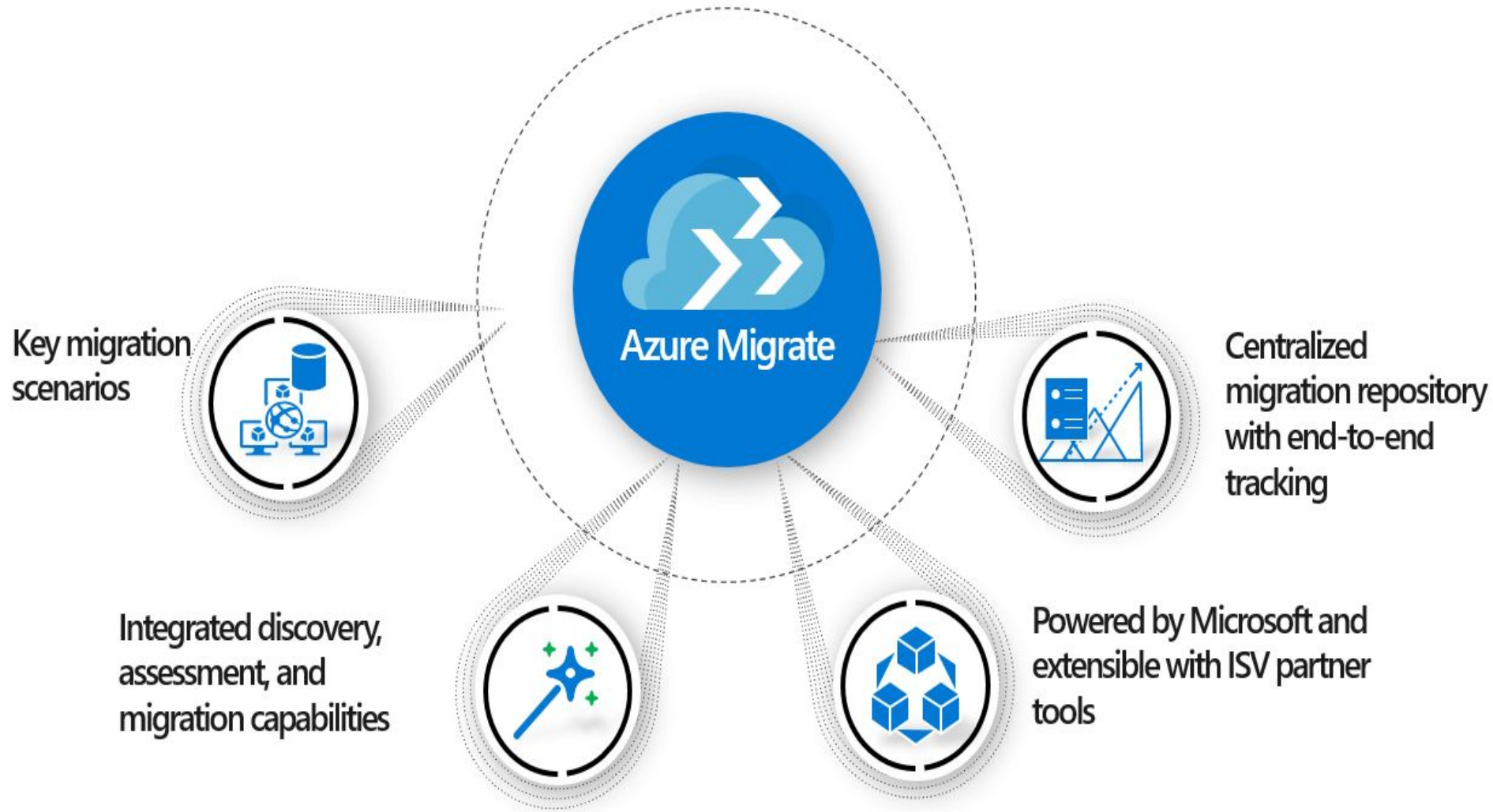
- ▶ Introduction
- ▶ Preparing for Data Migration
- ▶ Creating a Data Migration pipeline
- ▶ Executing a Data Migration Pipeline
- ▶ Best Practices For Data Migration using Azure Data Factory
- ▶ Conclusion

# Introduction

- ▶ Data migration is the process of moving data from one location to another, typically from an older system or technology to a newer one.
- ▶ Azure Data Factory is a cloud-based data integration service that enables you to create, schedule, and manage data pipelines.
- ▶ With Azure Data Factory, you can move and transform data from on-premises and cloud-based sources to various data stores such as Azure SQL Database, Azure Data Lake Storage, and Azure Blob Storage.
- ▶ Azure Data Factory provides a fully managed, serverless environment for building and deploying data pipelines quickly and easily.
- ▶ In this presentation, we will explore the features and capabilities of Azure Data Factory and how it can be used for data migration.



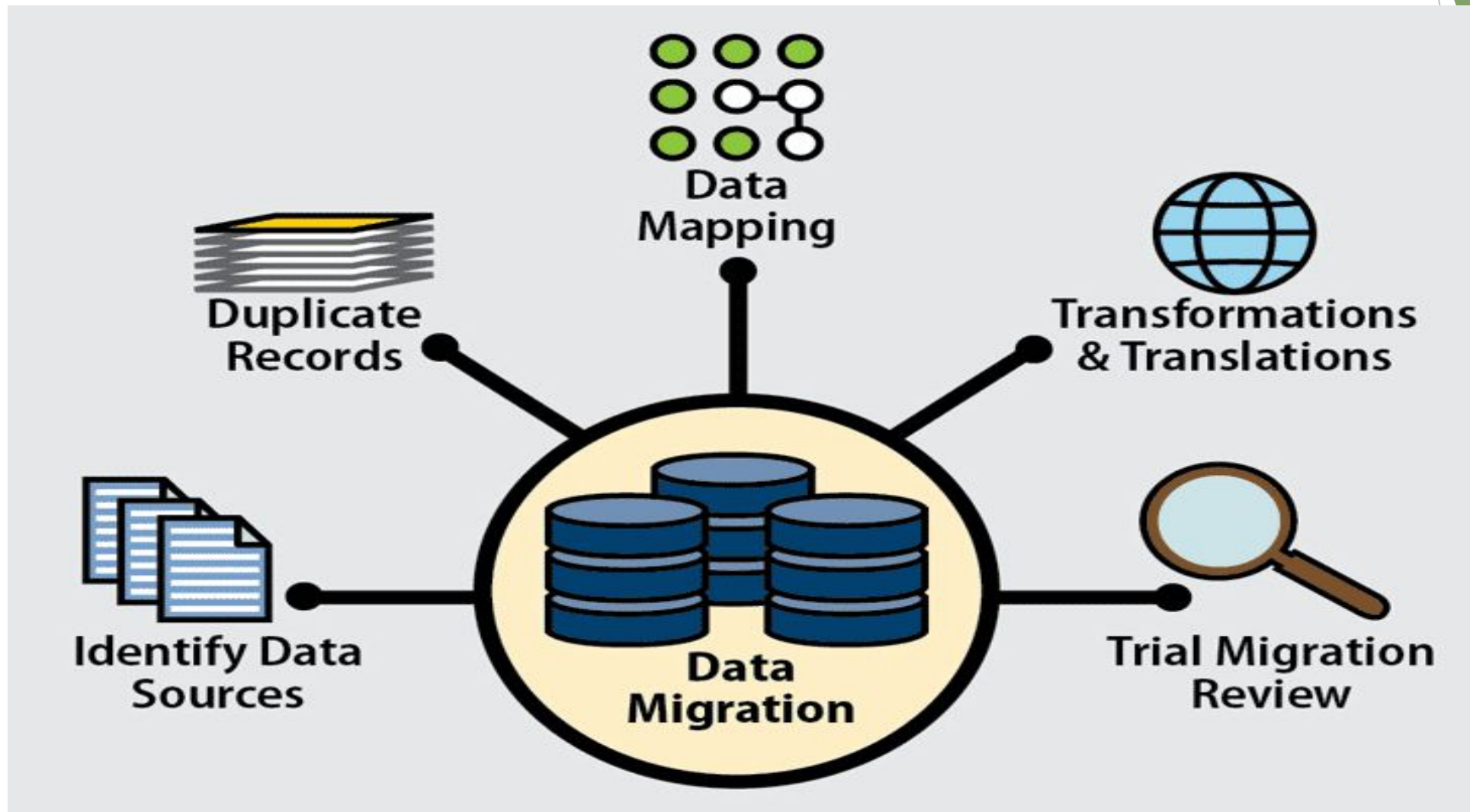
# Azure Migrate—Hub for Datacenter Migration



# Preparing for Data Migration

Some of the key features of Azure Data Factory include:

- ▶ Data migration is made easier by Azure Data Factory's serverless architecture, which eliminates the need for manual intervention at every stage of the process, as well as its support for a wide range of data sources and destinations, data transformation tools, integration with other Azure services for security and compliance, and automatic scaling.
- ▶ They also need to ensure that the necessary permissions and access credentials are available for the source and destination data stores.
- ▶ Finally, organizations need to consider data governance and security policies, such as compliance with data protection regulations and secure data transfer practices.



# Explanation of the data migration pipeline:

- ▶ A data migration pipeline is a series of activities that move data from a source to a destination. data migration pipeline is a series of steps or processes used to ensure that data is accurately and efficiently transferred from one location to another. It includes the following stages:
- ▶ Data Source: This is the system or application where the data currently resides. It can be a database, file system, cloud storage or any other type of data storage.
- ▶ Data extraction: Data is extracted from the source data store.
- ▶ Data transformation: Data is transformed as per the business rules and transformation logic defined.
- ▶ Data loading: Data is loaded into the destination data store.
- ▶ Data Verification: Once the data is loaded, it needs to be verified to ensure that it has been migrated correctly.
- ▶ Data Cleanup: After the data has been verified, any temporary files, tables or other objects created during the migration process need to be removed.
- ▶ Data Archiving: It's a good practice to keep a backup of the data being migrated. Data archiving helps in retaining the data in its original format, and it can be used as a reference in the future.



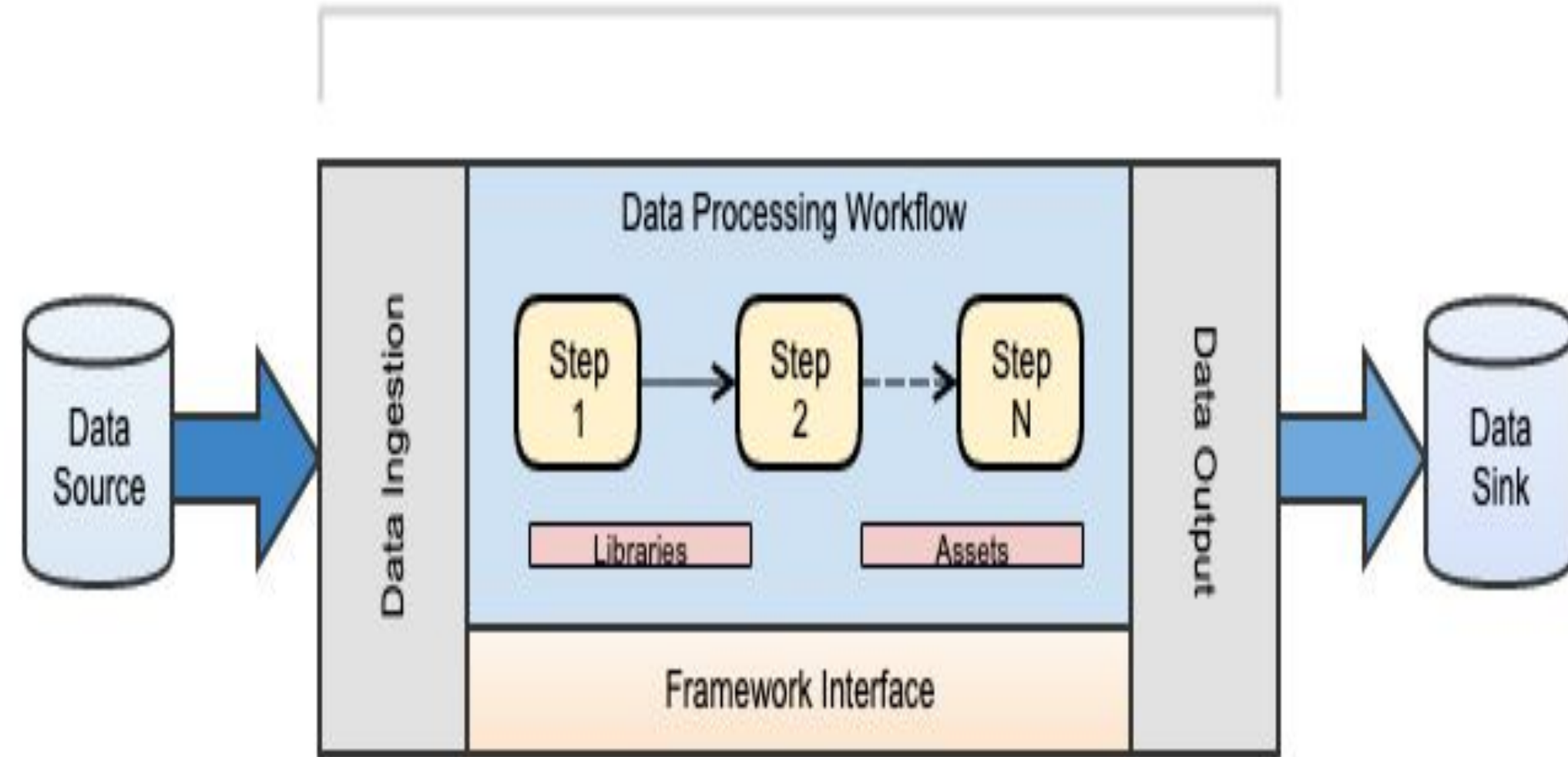
# Creating a Data Migration Pipeline

To create a data migration pipeline in Azure Data Factory, follow these steps:

- ▶ Create a new pipeline: Open Azure Data Factory and create a new pipeline.
- ▶ Add activities: Add activities to the pipeline, such as copy and transform activities, to move and transform data.
- ▶ Configure source and destination: Configure the source and destination data stores and provide access credentials.
- ▶ Configure transformation logic: Configure the transformation logic to transform data as per business rules.
- ▶ Run the pipeline: Run the pipeline to start the data migration process.
- ▶ Monitor and troubleshoot: Monitor the pipeline and troubleshoot any issues that may arise.



## PIPELINE



# Executing a Data Migration Pipeline

Here are the several steps for executing a Data Migration pipeline

- ▶ Define the scope and requirements of the data migration project: It is essential to define the scope and requirements of the project. This includes identifying the data sources, data volume, data types, and the target system.
- ▶ Prepare the source data: Once the scope and requirements are defined, the source data needs to be prepared for migration. This includes cleaning and transforming the data to ensure it is in the correct format for the target system.
- ▶ Choose a data migration tool: Choose a tool that meets the project requirements and supports the source and target systems.
- ▶ Set up the data migration pipeline: Configure the data migration tool to connect to the source and target systems. Define the mapping between the source and target data fields.
- ▶ Run the data migration pipeline: Execute the data migration pipeline to transfer the data from the source system to the target system. Monitor the pipeline for any errors or data discrepancies.

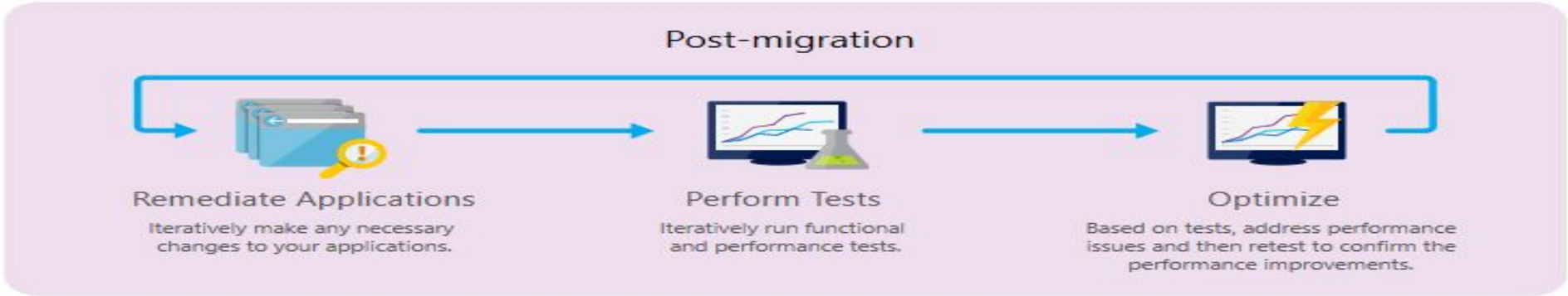
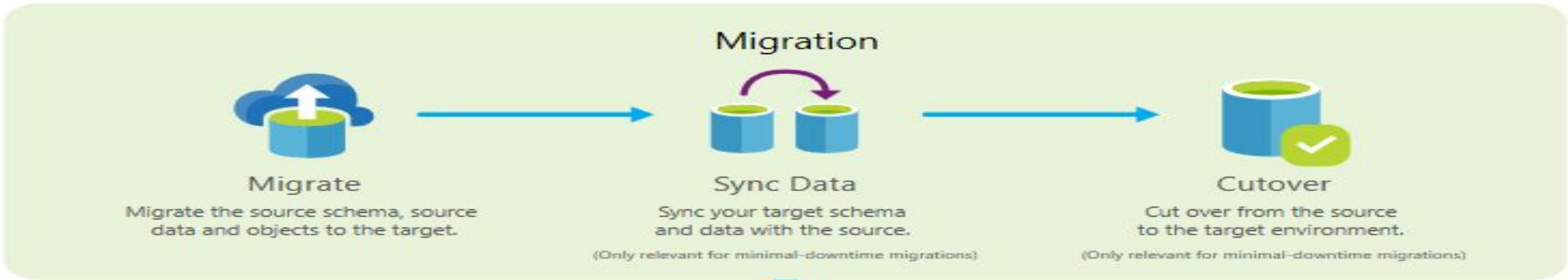
- ▶ **Verify the migrated data:** Once the data migration is complete, verify the migrated data in the target system. This includes checking the data accuracy, completeness, and consistency.
- ▶ **Perform data validation and testing:** Conduct data validation and testing to ensure the migrated data meets the business requirements and is functional.
- ▶ **Go live:** Once the migrated data is validated and tested, go live with the new system. Plan for any downtime and communicate the change to the stakeholders.
- ▶ **Monitor and maintain the new system:** After going live, monitor and maintain the new system to ensure it continues to function correctly and meets the business requirements.

# Best Practices For Data Migration using Azure Data Factory

Here are some best practices for data migration using Azure Data Factory:

- ▶ **Plan and prepare:** Before starting the migration process, it is important to plan and prepare for the migration. This includes identifying the data sources, data volumes, migration timelines, and defining the migration strategy.
- ▶ **Use incremental migration:** This involves migrating data in batches or increments, which can help to reduce downtime and minimize the risk of data loss.
- ▶ **Test the migration:** This can be done by setting up a test environment and performing a trial migration.
- ▶ **Use Azure Data Factory pipelines:** Use Azure Data Factory pipelines to automate the migration process and to ensure that the data is migrated accurately and efficiently.
- ▶ **Monitor the migration:** Monitor the migration process closely to identify any issues or errors that may arise.

- ▶ **Optimize performance:** This can help to improve the speed and efficiency of the migration process.
- ▶ **Ensure data integrity:** This can be done by performing data validation checks and using data quality tools to identify and correct any data issues.
- ▶ **Maintain data security:** This includes implementing appropriate access controls, encryption, and data protection measures to prevent unauthorized access or data breaches.
- ▶ **Test the data migration pipeline thoroughly:** Before deploying the data migration pipeline to a production environment, it's important to test it thoroughly. This includes running the pipeline in a test environment, verifying that the data is migrated correctly, and testing the pipeline's performance under different load conditions.



# Conclusion

Azure Data Factory is a powerful tool for data migration and integration projects. It offers a wide range of connectors and tools to help move data between various sources and targets seamlessly.

By leveraging Azure Data Factory, businesses can streamline their data migration process, reducing the time, cost, and effort required to move data from one system to another.

The platform provides features such as data mapping, scheduling, and monitoring, ensuring that data is moved securely and accurately.

Azure Data Factory is also scalable and flexible, allowing businesses to adjust their data migration strategy based on their evolving needs.



Thank you

The background features abstract, overlapping geometric shapes in various shades of green, primarily concentrated on the right side of the image, creating a modern and dynamic visual effect.