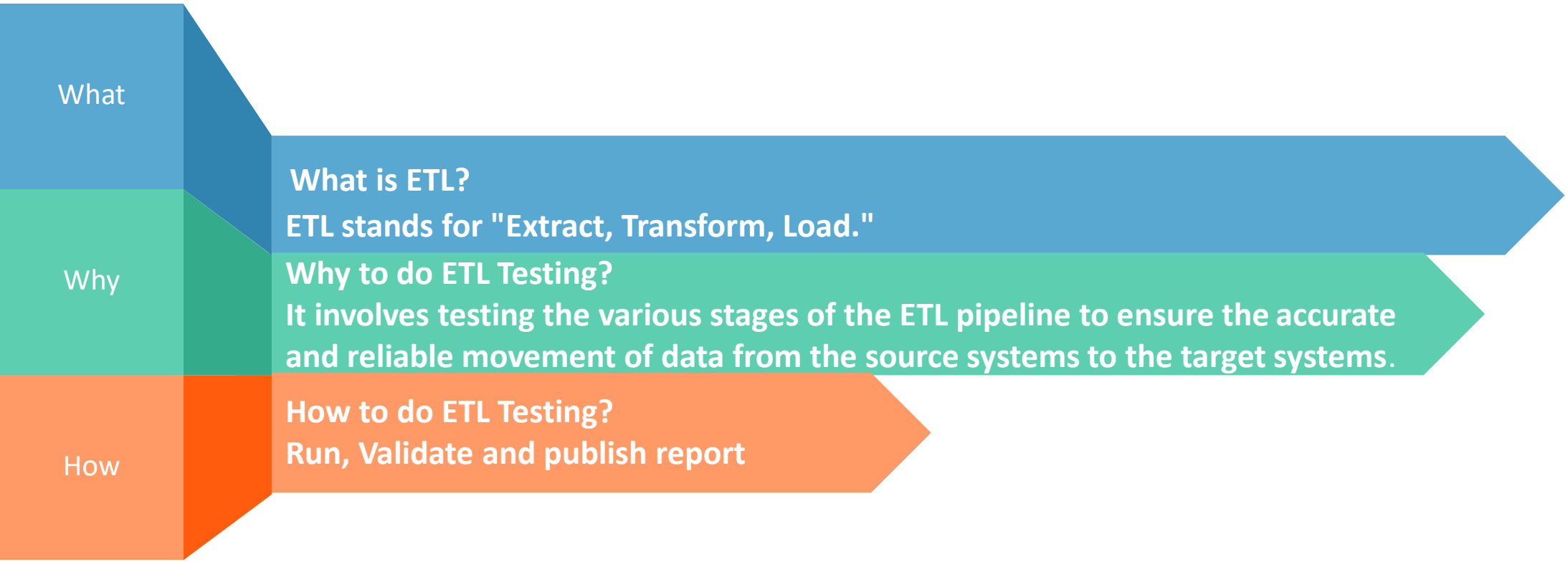
ETL Training

ETL

What ? Why? How?

Testleaf
Always Ahead

The Golden Circle

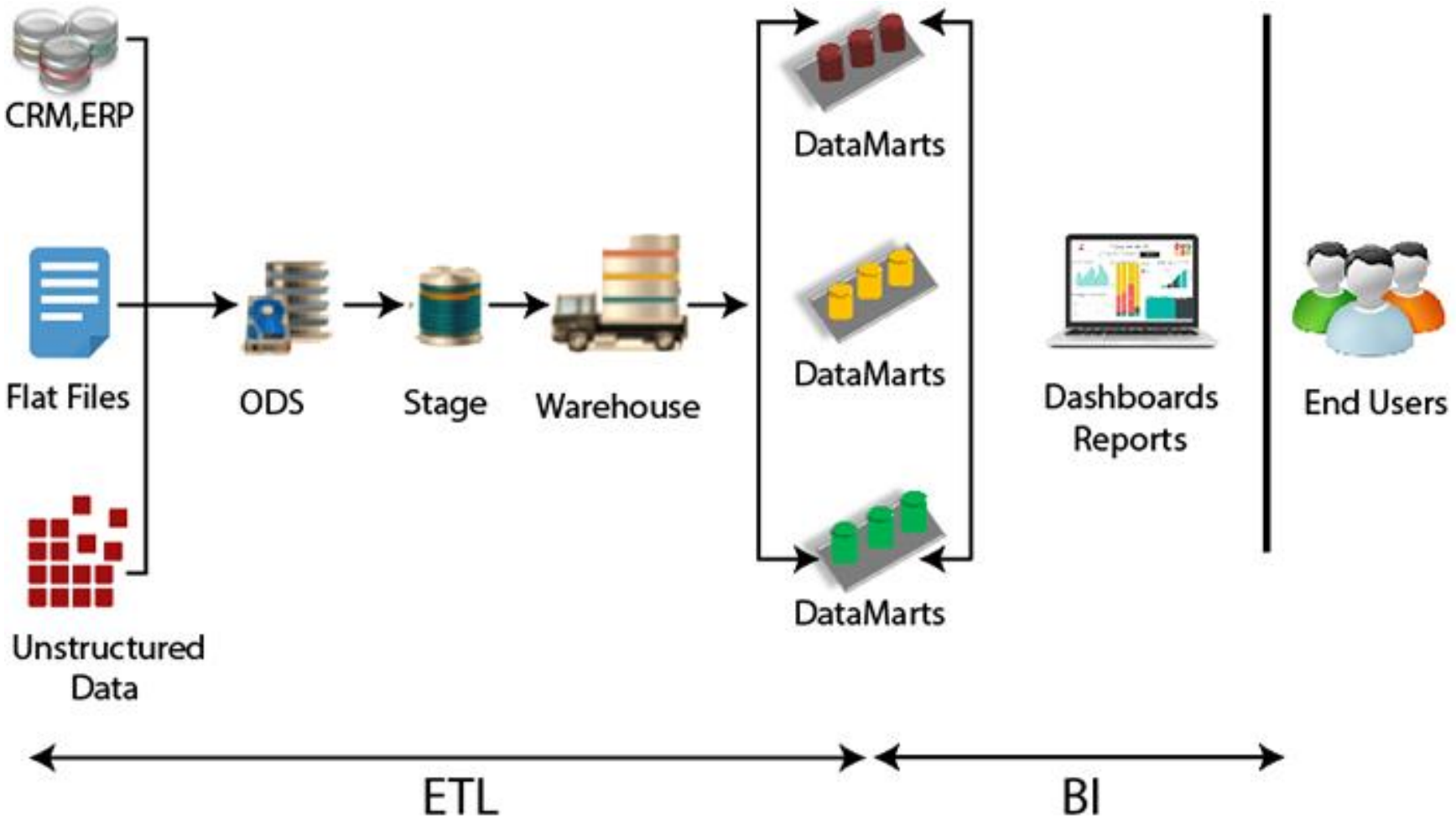


When to do ETL testing?

ETL testing is used to ensure the accuracy, completeness, and reliability of the data during the Extract, Transform, and Load process

- **New ETL Implementations**
- **Data Integration Projects-** to ensure that the integrated data is accurate and consistent.
- **Data Migration -** to verify that data is correctly transformed and loaded into the new system without any data loss or corruption.
- **Regular Data Updates -** the source data is regularly updated or modified, ETL testing is necessary to ensure that the data integration process continues to function as expected
- **Data Reconciliation-** ETL testing includes comparing the data between source and target systems to verify data accuracy and identify any discrepancies.
- **Performance Optimization-**helps identify performance bottlenecks and ensures that the data movement process is optimized for efficiency and speed.

DWH Architecture



Types of ETL Testing

Unit testing:

Tests individual components of the ETL process

Integration testing:

Tests the interaction between various components of the ETL process for identifying defects in the interaction between components and ensuring that the ETL process functions as a whole.

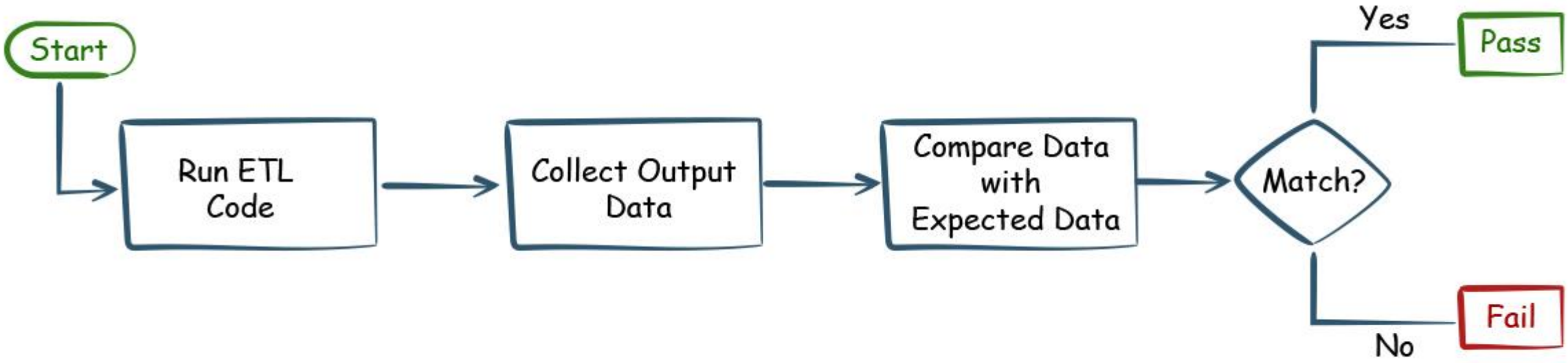
Performance testing:

Tests the ETL process under different load conditions. identifies performance bottlenecks and ensuring that the ETL process can handle large volumes of data.

Regression testing:

Tests the ETL process after any changes or upgrades

ETL testing process



Data Validations

- Metadata testing
- Data Completeness testing
- Data Accuracy testing
- Data Transformation testing
- Data Quality testing
- Duplicate Check
- Performance testing

ETL testing challenges

- Data volume and complexity is huge
- Unavailability of inclusive test bed at times
- Data completeness check for transformed data is tricky
- Test coverage for huge volume of data
- Automation for data testing

ETL testing best practices

Following these ETL best practices provides a better ETL testing experience.

- ✓ **Test early and often**
- ✓ **Use realistic test data**
- ✓ **Automate your testing**
- ✓ **Collaborate with stakeholders**
- ✓ **Follow a structured approach**
- ✓ **Test for performance**
- ✓ **Test after any changes or upgrades**

ETL tester's roles and responsibilities

As an ETL tester, you play a critical role in ensuring the accuracy and completeness of data during the ETL process. Here are some of the most important responsibilities of an ETL tester:

- **Test planning and preparation**
- **Data analysis**
- **Test execution**
- **Defect management**
- **Communication**
- **Continuous improvement**

ETL Testing Tools

- iCEDQ
- Informatica Data Validation
- DataQ
- BiG EVAL
- QuerySurge
- DbFit