[ENG] PM4 Proposal of phases, scope of migration tests and method of switching to the new tool

Contents

Purpose of the Document

The purpose of this document is to describe the methods for testing data migration, taking into account the environments that will be used to carry out individual tests and the migration itself.

Due to the assumption that credit processes initiated earlier will continue in the current systems, the document will outline the general principles and approach to multi-phase migration, including incremental migration if necessary.

Purpose of the Tests

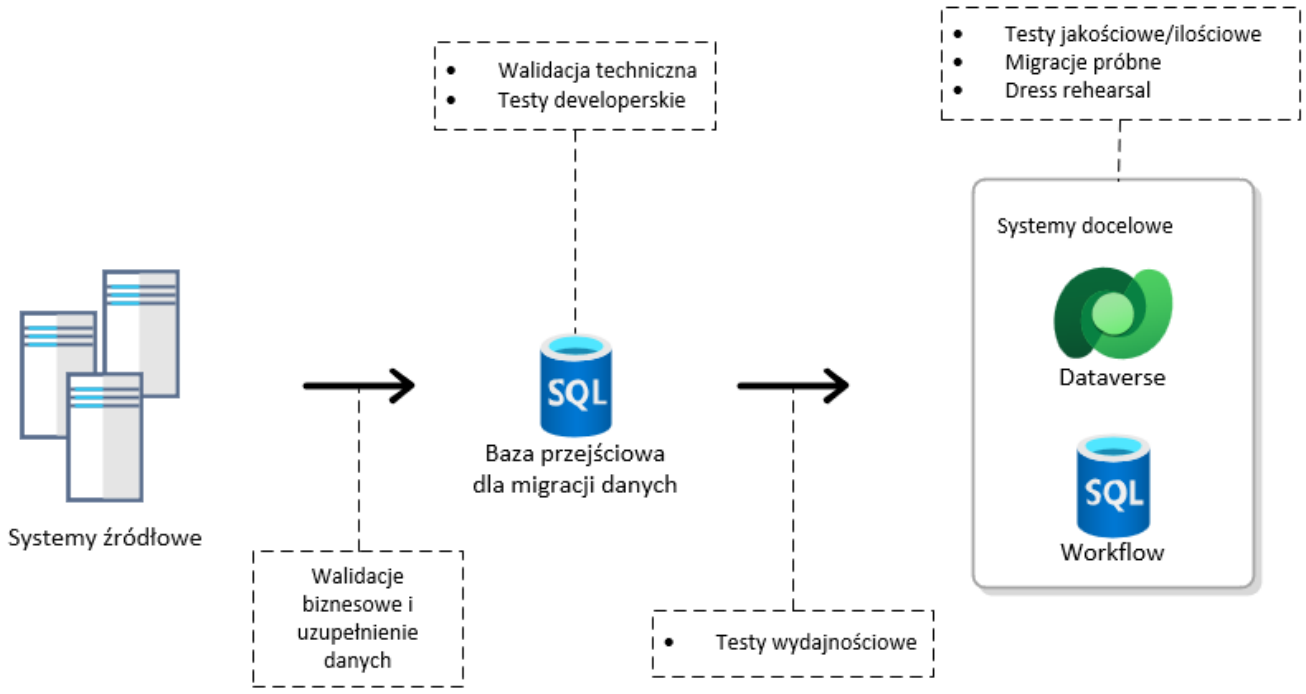
The aim of the tests is to verify whether the provided migration solution enables the execution of business processes in Workflow.

Detecting and eliminating data migration errors will ensure the highest possible quality of the migration process and the proper functioning of the Falcon solution.

Testing and Production Environments

A list of environments is available here: PT16 Proces CI/CD i sposobu przenoszenia zmian pomiędzy środowiskami testowymi oraz na produkcję - New CRM and Credit Worflow for Corporate - DISCOVERY - Confluence).

Data Migration Testing



Types of Data Migration Tests

Developer Tests

This type of testing is conducted by the development team and aims to verify the correctness of migration tools and processes in the development environment (DEV), including:

- Checking whether ETL scripts/tools function according to expectations
- Identifying errors in data mapping (e.g., whether source data has been correctly assigned to the appropriate fields in SQL database)
- Identifying errors in the logic and process of data transformation
- Early-stage verification of the planned migration process logic
- Validating data integrity after transferring it to the testing environment
- Simulating errors, such as missing required fields, corrupted records, or data inconsistencies to verify error handling mechanisms

Due to the nature of the tests and practical development considerations, errors identified in this environment will not be reported.

Responsibility: Provider

Migration Performance Tests

This type of testing is used to assess whether the data migration process is efficient, scalable, and capable of processing production volume data within an acceptable timeframe (deadline to be determined during the project implementation phase).

Performance tests will be carried out in the UAT or DMTEST environment (if a decision is made to build such an environment).

The selection of tools and specific metrics will depend on the chosen method of migration and will be precisely established during the implementation project analysis phase.

The main elements that will be checked during performance tests include:

- Migration time - measuring the data processing for various data volumes (e.g., 10%, 25%, 100%)
- Migration time - ensuring migration occurs within the appropriate, pre-established timeframe
- Stability of ETL processes - checking whether migration processes operate stably under high load
- Checking for potential scalability of the migration process
- Monitoring infrastructure loads (e.g., CPU / RAM / network)

Responsibility: Provider and Bank (on-premise infrastructure)

Quantitative and Qualitative Tests

Tests conducted in the UAT or DMTEST environment (if the decision to build such an environment is made) during trial migrations and dress rehearsals aim to verify:

- Data consistency - whether data is accurately represented in the target systems (e.g., hierarchy, relationships between entities)
- Data completeness - whether the number of records to be migrated matches the planned number of corresponding objects in the target systems
- Correctness of data mapping and transformation against business requirements for target systems
- Whether any duplicates were created as a result of migration
- Consistency of numerical data - e.g., the number of transactions per client

Responsibility: Provider and Bank

Functional Testing of the Falcon Solution on Migrated Data

Data migration supports functional tests of the system aimed at verifying the correctness of the solution using migrated data.

Trial Migrations

Trial data migration aims to conduct the actual data transfer process to target systems - at this stage, data resides in an intermediate database and is transferred to the target system.

This allows for the verification of all stages, procedures, and tools used throughout the data migration process. The goal is to identify potential issues and risks, as well as to verify compliance with established requirements before executing the actual migration in the production environment.

Data used in trial migrations must comply with all requirements for data intended for production migration (taking into account the progress of the project /implementation work).

For repeated data loading testing, it is assumed that migration tools/processes will have the capability to roll back (in particular, delete) previously loaded data in the target systems - the exact rollback mechanism will be described during the detailed analysis phase of implementation.

Within trial migrations, we can distinguish between: basic and general trial migrations.

Basic Trial Migrations

Conducting this type of test aims to check the entire migration pipeline, the correctness of the data reading and loading processes from source systems to target system. These migrations are performed across various environments (see Table 1 - comparison of trial migrations) and may include simulated full or incremental migrations. Naturally, trial migrations can be part of other types of migration tests and the overall process of evaluating the solution's quality.

The expected result of trial migrations is to identify issues regarding the migration process and subsequently resolve them before the next stages/iterations of the migration process.

General Trial Migrations (Dress Rehearsal)

This is a special type of trial data migration that uses a full set of data compliant with production data (data anonymization may depend on the environment). For these types of tests, it is recommended that testing occurs using non-anonymized production data, e.g., in UAT, DMTEST, or PROD environments (Phase 1 only) to ensure maximal compliance with production conditions.

A prerequisite for conducting Dress Rehearsal tests is the completion of the analysis and implementation phase of the Falcon system's functionalities corresponding to the specific phase of the project (see - PT6 Przypisanie modułów do fazy projektu w ramach, której musi być dostarczony - New CRM and Credit Worflow for Corporate - DISCOVERY - Confluence), which translates into the scope of data migrated in that phase.

The table below compares data migrations.

Feature	Trial migration	Dress rehearsal
Objective	Technical verification of data migration [1]	Simulation of full data migration
Environment	INT, UAT, DMTEST, PROD (only Phase 1)	UAT, QUA, DMTEST, PROD
Scope of data	Selected data sample[1, 2] (e.g. 5%, 25%... 100%)	Full range of data [2]
Type of data	Test data	Production data [2]
Responsibility	Supplier and Bank's technical team	Supplier, Bank

[1] - adapted to the requirements of a specific environment
[2] - for INT, UAT, QUA environments full anonymization, DMTEST environment (optional) and PROD - non-anonymized data

Test reporting

Each type of data migration testing performed will have its own reporting and error reporting system to be determined in the implant project.

The table below compares the test reporting.

Test	General purpose of reporting	Examples of error types	Recipients
Development tests	Technical information for the development team regarding issues that occurred during the building and launching of migration processes. This may include details such as: error occurrence location, description, uniqueness, resolution status, priority, and assignment to developers.	• errors loading the intermediate database • errors processing and mapping data in the intermediate database • connection errors • infrastructure problems • errors loading data to target systems • problems with prepared/used migration tools	Supplier (development team)
Migration performance tests	Technical information describing the performance of the migration process and any deviations from the established assumptions. It contains information such as: process identifier, target system, entities participating in the test, error description, standard description	• entity migration timeout to target system • attempt to exceed the limits of the target environment/system (e.g. Request rate limit exeeded, Too Many Requests, CPU Quoata exeeded)	Supplier Bank
Quantitative /qualitative testing	Assessment of completeness, correctness and compliance with business requirements of data after trial migrations (including Dress Rehearsal)	• missing x% records • checksum for field X does not match • no relationship between entity X and Y • invalid data format	Supplier Bank
Test migrations	Description of the test performed along with any errors noticed, may include sample lists of records with errors, the amount of data subject to migration, the amount of data migrated, quality metrics, etc.	• problems in mapping or transformations for entity X • here is a data inconsistency between the target system and the data source • connection problems (integration tests)	Supplier Bank
General trial migrations	Description of the test performed along with any errors noticed, may include sample lists of records with errors, the amount of data subject to migration, the amount of data migrated, quality metrics, etc.	• problems in mapping or transformations for entity X • there is a data inconsistency between the target system and the data source • connection problems (integration tests) • migration time exceeded for entity X	Supplier Bank

Production implementation plan

Implementation strategy

Data migration to target systems (CRM / Workflow) was planned in a phased manner, taking into account full and possibly incremental migrations.

Preparing data migration

As part of preparing the data migration, the necessary environments and tools selected by the Bank will be prepared

The detailed analysis phase of the implementation stage will provide:

- business requirements for data
- migration non-functional requirements for data migration, in particular performance and security
- detailed migration plan for each entity to be migrated
- source analysis with mappings and transformations
- characteristics and requirements for the data model in Falcon solution target system
- details regarding the structure of extracts and the destination for data extracts
- GO / NO-GO criteria
- definition of a detailed migration schedule
- definition of possible migration windows
- definition of technical break rules for target systems, including code freeze and business freeze
- definition of rules for restoring systems to their pre-migration state

Phase 1 of the project: Pilot / Friends & Family

Full data migration is assumed

Full data migration is assumed

- All data from the source systems that is necessary to run the features and processes defined for Phase 1 will be migrated to Falcon).

Assumptions:

- In Phase 1, the production environment (PROD) is initially clean and ready to receive data. Until production goes live, the PROD environment can be used to test the migration of production data.

Incremental migrations in Phase 1

During Phase 1, it may be necessary to perform **incremental migrations** for objects that have already been migrated during a full migration. The purpose of incremental migrations is to:

- Data alignment between source systems and target systems when changes are made to the source systems (e.g., new records are added, existing data is updated) between data extraction and Falcon production launch.

Assumptions:

- Incremental migrations will include objects that have already been migrated as part of a full migration.
- Data in the target system will be updated with the latest state of the source systems within established time windows that will be defined as part of the detailed analysis of the implementation phase.
- The migration strategy assumes **no reverse communication** between the target and source systems:
    - Data is migrated only in the direction **source system → target system.**

The current assumptions of the coexistence of CRM system do not provide for the need to perform data migration in subsequent project phases.

Therefore, there is no need to create and use the environment marked as DMTEST.

Due to the possible change of these assumptions, we present the approach to migration in the next Phases of the project as a potential variant.

Phase 2 and subsequent phases of project implementation

Full migration for Phase 2 and beyond

For each subsequent implementation phase, **a full migration** is planned, covering new data ranges, specified in ??PM2 Doprecyzowanie zakresu danych do migracji. During this stage:

- All data from the source systems that are necessary to run the features and processes defined for phase X will be migrated to the target systems (CRM – Dataverse, Workflow - SQL Server).

Incremental migrations Phase 2 and beyond

In situations where a legacy system remains active after a full migration, incremental migrations can be performed to align data between source and target systems.

The data scope definition is specified in ??PM2 Doprecyzowanie zakresu danych do migracji.

Assumption:

- There is no reverse communication – data is migrated unidirectionally, from the source system to the target system.

Migration tests

The entire migration process in the Falcon project is supported by detailed testing, which was described in the previous chapters.

- **Developer testing:** Verification of technical correctness of migration processes (mapping, transformations, validations).
- **Performance testing:** Checking migration time and system load with full data volume.
- **Quantitative and qualitative testing:** Validation of completeness and correctness of data after migration.
- **Test migrations:** Testing migration processes for a limited set of data.
- **General migration testing:** Simulation of a full migration in a production-like test environment.

Framework plan for data migration to production

Assumptions

1. It is assumed that the PROD environment is prepared for data migration
2. It is assumed that the data migration process has been prepared in accordance with business requirements, tested and approved for production use by the BANK
3. At the detailed analysis stage of the project implementation phase, the following issues will be specified
    a. time to perform production migration
    b. time to prepare data for migration (extraction from source systems)
    c. responsibility matrices for migration processes, in particular for production implementation
4. It is assumed that all migration tests have been successfully completed

Plan:

1. Checking and confirming the readiness of the production environment in accordance with the requirements of Phase 1, including:
    a. checking the environment configuration
    b. checking whether the target databases / target objects are free of test data
2. Technical break of source systems before data extraction for migration (code & business freeze)
3. The prepared data extract from the source systems is transferred to the intermediate migration database
4. Technical verification of the transferred extract
5. Performing planned data processing and transformation
6. Obtaining GO / NO-GO acceptance from all stakeholders responsible for the production migration process
7. Implementing the technical break on the target systems (code & business freeze)
8. (incremental migration, subsequent phases) Creating snapshots / backup copies for objects that will be affected by the migration
9. Saving data to the target systems, including:
    a. migrating data to the NewCRM production environment (Dataverse database)
    b. migrating data to the Workflow production environment (Microsoft SQL Server database)
    c. Monitoring the migration process
10. Validating data after migration
    a. Performing quality tests and quantitative on data in target systems
    b. Confirmation of correctness of the migration
11. Launching target systems with production data
12. Checking key system functionalities in the PROD environment
13. (optional) Carrying out incremental migrations
14. Switching to intensive monitoring and support mode
15. Closing the implementation of data migration

Approach to Troubleshooting Errors During Transition to Production

The migration process will be monitored in terms of the course of the process and the compliance of data placed in target systems with data transferred for migration from source systems.

Actions taken in the event of adverse events:

1. Problem identification
    a. Checking and analyzing ETL process logs / tools used in migration
    b. Identifying the source of the problem / Locating the responsible component
    c. Determining the scope of the error (entity, number of records, selection parameters)
2. Error classification - detailed classification rules will be described at the analysis stage of the implementation phase
3. Securing the target system - the rules for stopping and withdrawing migrated data will be described at the analysis stage of the implementation phase. Assuming that there is only one, full data migration to a clean PROD environment, these rules can be very simplified.
4. Detailed analysis of the cause of the error
    a. Verification of data transformation logic
    b. Verification of data mapping logic
    c. Checking the integrity of the source data
    d. Verification of target schemas and implemented constraints
5. Error repair / data improvement
6. Repair tests
7. Patch acceptance process
8. Migration process consistent with the migration plan to the production environment
9. Tests

Acceptances

(acceptance by inserting ✅ after the name of the accepting person)

1. DevOps Databases
2. Application teams
    a. Milledesk Corporate: Janusz Norbert Cichocki
    b. Workflow Corporate
3. Architecture team: Tomasz Kochanek ✅