Data Migration

Overview

Data migration is the process of transferring data between storage types, formats, or computer systems. In many situations this is complicated process, involving data in different formats, multiple databases and other data sources, and active applications that rely on the data.

This document summarizes DialogWorks process for data migration.

5 Step Data Migration Process:

1. Kickoff

   Covers:
   - Scope of work
   - People involved, their roles and responsibilities
   - Preliminary schedule and key dates
   - Important issues and concerns such as system downtime.

2. Discovery

   Investigate and document the current data situation:
   - Databases
   - Fields, data types, values, relationships
   - Size/volume of data
   - Data sources
   - Security
   - Validity, data cleanliness
   - Data related processes, manual and automated.
   - Applications and users
   - Business impacts of the migration.

3. Detailed Migration Plan

   Specification detailing exactly what the data migration will involve:
   - Systems involved
   - Key activities, transition steps, schedule
   - Dependencies
   - Identify equipment and tools: computers and storage devices, databases, software applications and utilities, custom software utilities and application extensions.
   - Testing plan, verification procedures
   - Completion criteria.

4. Develop

   - Target data schema and database configuration.
   - Create data mapping templates.
   - Create any necessary software tools/templates.
   - Test the various tools and processes that will be used.
   - If data migration will happen on an ongoing basis then create applications, and processes, to automate the data transfer. Consider history logging, data validation, error handling, rolling back to prior state, notifications.

5. Execute

   - Data exports
   - Populate data mapping templates,
   - Data processing/transformation (e.g., dedupping and data cleanup, calculated fields)
   - Data import into new database
   - Post import data processing
   - Provide history logs and error alerts and appropriately handle any issues.
   - Validation:
     - Compare general statistics for source and target data. record counts, range of values, min/max/average values.
     - Manually check all values in a random sample of records, plus check records with unusual origins or boundary conditions.
     - Test applications that rely on the data