**Data Migration**

**What is Data Migration?**

* Data migration is the process of moving data from one system to another. It involves a change in storage and database or application.
* In the context of the extract/transform/load (ETL) process, any data migration will involve at least the transform and load steps.
* This means that extracted data needs to go through a series of functions in preparation, after which it can be loaded in to a target location.

**Why a Data Migration?**

* Organizations undertake data migrations for a number of reasons.
* They might need to overhaul an entire system, upgrade databases, establish a new data warehouse, or merge new data from an acquisition or other source.
* Data migration is also necessary when deploying another system that sits alongside existing applications.

**Data Migration Strategy**

**Knowing the data**

Before migration, source data needs to undergo a complete audit.

Unexpected issues can surface if this step is ignored.

**Cleanup**

Once you identify any issues with your source data, they must be resolved.

This may require additional software tools and third-party resources because of the scale of the work.

**Maintenance and protection**

Data undergoes degradation after a period of time, making it unreliable.

This means there must be controls in place to maintain data quality.

**Governance**

Tracking and reporting on data quality is important because it enables a better understanding of data integrity.

The processes and tools used to produce this information should be highly usable and automate functions where possible.

**Data Migration Strategies**

**“Big Bang” Migration**

In a big bang data migration, the full transfer is completed within a limited window of time.

Live systems experience downtime while data goes through ETL processing and transitions to the new database.

**“Trickle” Migration**

Trickle migrations, complete the migration process in phases.

During implementation, the old system and the new are run in parallel, which eliminates downtime or operational interruptions.

Processes running in real-time can keep data continuously migrating.

Best Practices for Data Migration

**Back up the data before executing.**

In case something goes wrong during the implementation, the data may lose.

Make sure there are backup resources and that they’ve been tested before you proceed.

**Stick to the strategy.**

Too many data managers make a plan and then abandon it when the process goes “too” smoothly or when things get out of hand.

The migration process can be complicated and even frustrating at times, so prepare for that reality and then stick to the plan.

**Test.**

During the planning and design phases, and throughout implementation and maintenance, test the data migration to make sure to achieve the desired outcome.

6 Key Steps in a Data Migration Strategy

**1. Explore and Assess the Source**

Before migrating data, you must know (and understand) what you’re migrating, how it fits within the target system.

Understand how much data is pulling over and what that data looks like.

There may be data with lots of fields, some of which won’t need to be mapped to the target system.

There may also be missing data fields within a source that will need to pull from another location to fill a gap.

Must be clear on what needs to migrate over, what can be left behind and what might be missing.

Beyond meeting the requirements for data fields to be transferred, run an audit on the actual data contained within.

**2. Define and Design the Migration**

The design phase is where organizations define the type of migration to take on — big bang or trickle.

Considering the design, the data to be pulled over, and the target system, begin to define timelines and any project concerns.

By the end of this step, the whole project should be documented.

During planning, it’s important to consider security plans for the data.

Any data that needs to be protected should have protection threaded throughout the plan.

**3. Build the Migration Solution**

A common tactic is to break the data into subsets and build out one category at a time, followed by a test.

If an organization is working on a particularly large migration, it might make sense to build and test in parallel.

**4. Conduct a Live Test**

The testing process isn’t over after testing the code during the build phase.

It’s important to test the data migration design with real data to ensure the accuracy of the implementation and completeness of the application.

**5. Flipping the Switch**

After final testing, implementation can proceed, using the style defined in the plan.

**6. Audit**

Once the implementation has gone live, set up a system to audit the data in order to ensure the accuracy of the migration.

Data Mapping

Data mapping is the process of matching fields from one database to another.

It's the first step to facilitate data migration, data integration, and other data management tasks.

Data now comes from many sources, and each source can define similar data points in different ways.

For example, the state field in a source system may show Illinois as "Illinois," but the destination may store it as "IL."

Data mapping bridges the differences between two systems, or data models, so that when data is moved from a source, it is accurate and usable at the destination.

If not properly mapped, data may become corrupted as it moves to its destination.

Quality in data mapping is key in getting the most out of your data in data migrations, integrations, transformations, and in populating a data warehouse.

What are the steps of data mapping?

**Step 1: Define**

Define the data to be moved, including the tables, the fields within each table, and the format of the field after it's moved.

For data integrations, the frequency of data transfer is also defined.

**Step 2: Map the Data**

Match source fields to destination fields.

**Step 3: Transformation**

If a field requires transformation, the transformation formula or rule is coded.

**Step 4: Test**

Using a test system and sample data from the source, run the transfer to see how it works and make adjustments as necessary.

**Step 5: Deploy**

Once it's determined that the data transformation is working as planned, schedule a migration or integration go-live event.

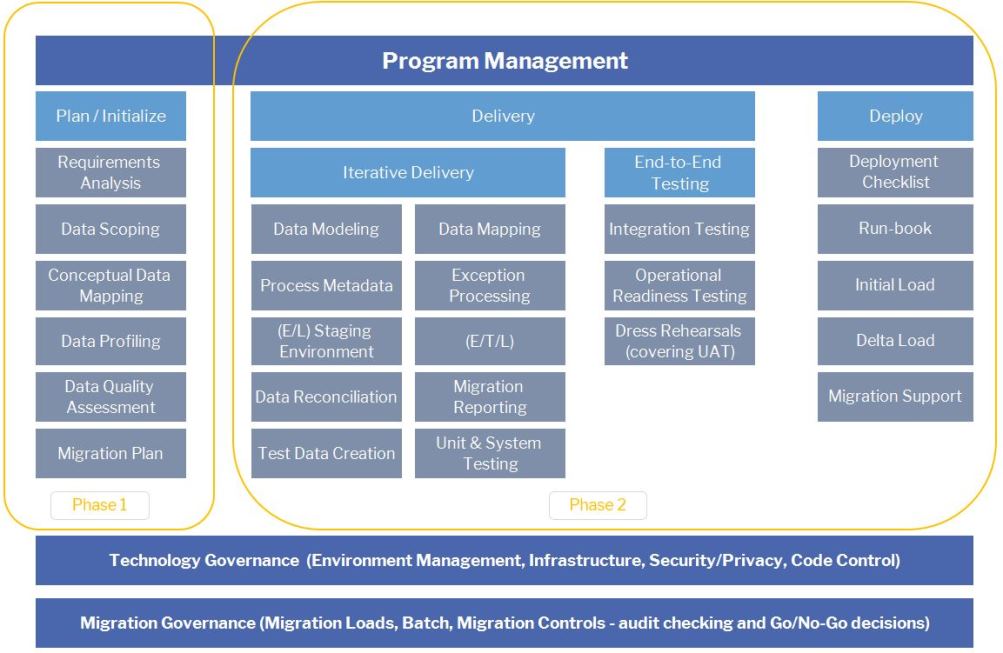
**Step 6: Maintain and Update**

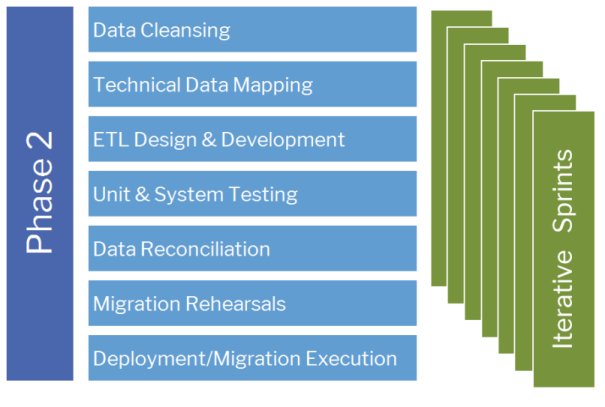
For ongoing data integration, the data map is a living entity that will require updates and changes as new data sources are added, as data sources change, or as requirements at the destination change.

**Client-side work of data preparation**

**Stage 1** involves review of the locations of existing data, the condition of that data and the approach that most suits the Client. During Stage 1 we would do high level data mapping, create a data migration inventory / plan; and perform some data profiling (i.e. existence of field, nullability, keys). At the end of Stage 1, we would supply the following Deliverables:

* High Level Data Mapping between the Source and Target
* Data Profiling results from source data (which will form part of the data mapping).





**Interview Questions**

**What is Agreement in Data Migration Process?**

This is also called as Statement of Work (SOW) or Contract with the customer.

It contains the task, summary of the work, environment factors where the migration will happen, technical and other requirements, Period of performance (SLA), point of contact, Cost/Pricing

**What are the types of Migration?**

Sample Migration

Actual Migration

Delta Migration

**What is Delta Migration?**

A delta migration only moves modified files since the last migration that was executed. This is done by comparing the modified timestamps between the source and destination platforms

**What is the flow after getting the ticket?**

1. After receiving the ticket, we have to response to the customer that we have received a ticket
2. Getting all the details/information about the ticket from the customer
3. Check/validate all the details/info that are given by the customer
4. We have to check the number of records that needs to migrate
5. Have to send an agreement/Contract to the customer with cost/pricing
6. We have to get the signoff for the agreement from the customer before start the migration
7. Give respond to the customer about the signoff agreement
8. We have to do the Sample migration to validate where the data are properly populated from source to destination. The Sample migration needs to be done if the count of records are in millions
9. Check with customer about the Sample Migration
10. If we don’t have the records in millions, then we can start actual migration, after the respond of the SOW.