Legacy Issues in Migrating Data for New Data Warehouse

Sushant S. Sule
MBA-IT Student
Department of Business Administration
SICSR, Symbiosis International University (SIU), Pune, Maharashtra, India

Pravin S. Metkewar
Associate Professor
Department of Business Administration
SICSR, Symbiosis International University (SIU), Pune, Maharashtra, India

Syed Khizer
Assistant Professor
Department of Computer Science
College of Computer, Qassim University, Al-Qassim, Kingdome of Saudi Arabia (KSA)

Abstract

This paper tries to propose a solution for various issues that uncover while migrating the data from old legacy systems to new systems. Business organizations implement new Software Application System to replace the functionalities of their major processes by their old legacy systems from time to time. Data Migration is the procedure of relocating data from one framework then onto the other while changing the capacity, database or application. Complexities arise when there is a try to take the information (data) right from the legacy framework (system) and change or modify it to fit into the new framework (system). Mostly the structure and the data types of the old legacy systems are different in relation to the new system being implemented; the difference is simply not constrained to the table names, field names, properties or sizes. The types of databases are distinct, as also the entity relationships in the new framework may not compatible with the earlier legacy systems. To get the legacy data into its new application format, a certain number of modifications and transformations must take place. These modifications and transformations are known as 'Data Conversion'. During the implementation of new framework, the current structure which is being used by the old legacy system is taken into account and is mapped to the new framework being designed and implemented.
Keywords- Data Migration, Data Warehouse, Legacy, Issues, Systems

I. LEGACY SYSTEM ISSUES AND CONSTRAINTS

A. Standardization

After many of years of implementing a data warehouse, the business organization generally feels the need migrate to a new warehouse as the old legacy system do not match the new standards introduced in the market.

B. Security

Security is one of the major issues where the information or data is leaked out of data warehouse where the currently implemented warehouse is outdated.

C. Maintenance

It sometimes becomes cumbersome to maintain the old warehouses because the numbers of experts available for the same are less as they have already moved to the new technology.

II. DATA MIGRATION CHALLENGES

A. Data Loss Risk

At a certain point when the information in legacy system is accessible but the same data is not accessible in the new target frame work, then it is termed as 'Data loss'. Data loss is dangerous in data relocation.

B. Data Corruption and Integrity

If the configuration and quality of data in the legacy systems and target systems is diverse because of the migration process then it is adulterated information. Because of data migration, irregularities or excess or copied information or vicinity of non-important information are identified with information honesty issues. Data corruption and data integrity influences business and operation effectiveness and it absolutely beats motivation behind movement.

C. Semantics Risk

In the course of migration of data, it generally happens at a certain point that the legacy framework and target segment hold same importance yet unit of estimation is different and the data types have completely changed. It is important to see in such a case that the data is loosed and the migration is successful but the value of data is not lost at the target.

III. MAJOR CHALLENGES

A. Unmatched Data Types (Number, Date, Sub-Records)

Different data types can be taken care of effectively approximating the nearest sort from the target database to keep up to the integrity of data. If the legacy systems support complex data types but the target systems do not, then it is necessary to look up to the compatibility of both the systems. Similarly, if the legacy system supports different type of encoding but the target system doesn't then again it is necessary to look up to the compatibility of the systems.

When a database or a data warehouse is used not only for storage purposes but also to enable decision-making for the top authorities, i.e. business logic, it is then not only necessary but also mandatory to perform a feasibility study of the data prior to migrating it in the new target system thoroughly.

B. Different Character Sets (encoding)

Majority of the frameworks are created on PC-based platform encoding or national extension based on ASCII. The most recent one is UTF-8 which keeps ASCII mapping for alpha and numerical characters however empowers stockpiling of characters for the greater part of the national letters in order including Chinese, Japanese and Russian. Centralized server frameworks are for the most part in view of EBCDIC encoding which is contradictory with ASCII and transformation is required to show the information. ETL instruments ought to bolster the transformations between character sets, including EBCDIC.

IV. CONVERSION AND MIGRATION PROCESS

The following steps are to be followed while migrating from the old legacy systems to new target systems:

1) Plan
2) Analyze
3) Convert
4) Migrate

A. Plan

To accomplish the objective of a successful relocation of data from legacy framework to target framework through the process of conversion, it is necessary that a considerable amount of planning is done before the actual movement of data from the legacy systems to the new systems which includes keeping in mind the risks as well.

The Migration Plan characterizes:

1) Prerequisites i.e. what information is moved, where it is moved, how it is moved, when it is moved, and around to what extent the move will take.
2) Team i.e. the Users, Business Analysts, System/Data Analysts, Testers, Target Application Users.
3) Environment under which the Legacy (Source) and the Target Application system operates, what will be the Data Conversion Stage Area.
4) Schedules - and task distribution, staffing levels.
5) Configuration Management Plan – to maintain control over changes needed due to various reasons.

B. Analyze

The primary phase of data modification and transition is Data Classification; by making Data Profile for the information components utilized as a part of Legacy Framework and Target Application.

1) Source and Target Data Profiling

In this stage, a detailed inside out study of the source system and target system is made with reference to the Data Structures model. This detailed investigation is carried out by the Data Analysts with a major assistance of Technical Analysts and recorded in the Data Profile vault. This storehouse, which incorporates source and target information in shared format, gives a profile of various elements which includes data utilization, limit and development designs.

2) Data Mapping

Data mapping is the procedure where every source information components are allotted to one or more target information component. A group of Business Analysts, Technical Analysts attempts to map the data in the old legacy systems to new target systems after the investigation (analysis) of data. Here, every data field that will be relocated from the source framework to the objective framework must be characterized and inspected to guarantee consistence with field lengths, information sorts, area values allowed, framework rules, uprightness checks and some other conceivable issues.

3) Data Cleansing

The Data Cleansing procedure results in important and precise data being changed over from the Legacy frameworks into the Target System. The Data Profile Repository arranged during Data profiling stage (by Data Analysis and System Analysis), to lead the Data Cleansing – in which Legacy application information to check that information are right, finished, steady and convertible.

4) Build Converted Data Test Plan

Once the framework is decided, execution model i.e. the apparatus course or improvement of the custom-made arrangement, Test Process and Plan taking into account the Data Map (Transformation Rules) are created. Testing process in Data Conversion and Migration venture could be arranged into two:

- Physical mistakes: These are the consequence of grammatical blunders of the scripts/programs, which can be effortlessly distinguished and determined.
- Logical mistakes: These are distinguished and determined amid Test stage. Such mistakes are consequence of the nature of the mapping exertion. Amid Implementation/Testing scripts/programs created in light of the Data Map i.e. Change detail, are executed.

C. Convert

The following steps are to be followed during the process of conversion:

1) Design and Develop Migration Tool

The data migration procedure can be proficient by the accompanying strategies:

- Using an information change instrument
- Scripts grew particularly with the end goal of transformation in the venture.
- Manual information transformation and relocation.

2) Pre-Test and Recalibration Tool

1) Mock Migration

Conduct dress practices for each arranged discharge. Mock relocations might be fractional or complete end-to-end cycles to check movement methodology and benchmark the process durations for every relocation undertaking.

2) Pilot Migration

Complete end-to-end relocation in the pilot environment. Coordinate with business clients in doing information acceptance, confirm and assess the control instrument and measurements.

3) Test Extract and Load

Once the Pilot Migration is finished, test the concentrate of the moved information and test the heap into the fake target environment.

3) Configure Staging Area

The following are stages to configure the staging area:

- Import Legacy System Data
- Data Conversion
- Converted

4) Execute Data Conversion

This is the most critical period of information change process. Once the information mapping is finished, a general information change procedure is intended to change over information into a structure required by the objective framework. Mapping determination serves as an information to this stage. Every one of the principles in mapping particular are consolidated in the change process outline.

5) Validate by Test Plans

Testing the changed data is an imperative movement of information transformation. The changed over information, a consequence of the transformation process, in view of mapping particulars ought to be tried.

D. Migrate

The following steps are to be followed during the process of migration:

1) Develop Migration Statistics

There are certain objectives connected with executing a effective data relocation system. Fundamentally, information must be relocated from the source stage to the objective stage totally and precisely, and as indicated by organization and administrative approaches on data controls and security. This implies no dropped or fragmented records, and no information handle that fall flat approval or other quality controls in the objective environment. . Another objective of information movement is that the

procedure be done rapidly, with as short a downtime window as could be expected under the circumstances. Finally, the expense of data migration must be sensible, as far as innovation and staff prerequisites.

2) Load the Data in Target System

Despite the fact that the data in the staging area is available in the objective information structure model, it is required to be stacked to target framework. This is by and large finished by information load scripts that pick information from the organizing range and load it to the objective framework straightforwardly. The information load scripts that heap information into target framework are tried after the stacking is done to the objective framework.

3) Validate the Target Application

Reconciliation reports are run.

Acknowledgment criteria are checked; if compromise mistakes or other criteria are not met, the framework is moved back to the first information source. Something else, interfaces and handling from the source is ceased and afterward initiated on the objective.

4) Implementing the Target System

The activities which have to carry out after successful migration of data from old legacy systems to new target systems include:

- Metrics and Control chart
- Scorecards or dashboards
- Design documents
- Technical approach
- Setup required
- The basic source and the ownership of source documents

V. CONCLUSION

Thus, we have shown the challenges of data migration, the major constraints in migrating the data from legacy systems to new systems and the entire procedure how the data from legacy systems can be migrated to construct a new data warehouse.

REFERENCES

Website References

[1] Legacy System Data Conversion and Migration
http://www.kmgin.com/Docs/White-Paper/data_migration.pdf
[2] Consolidating Data from Diverse Legacy Systems
http://www.cloveretl.com/sites/applicationcraft/files/files/solutions/data/Clover_Data_Migration_Case_Study.pdf
[2] Data migration from old to new application
http://www.gplivna.eu/papers/legacy_app_migration.htm
[3] Data Migration Challenges and solution for successful implementation
https://www.linkedin.com/pulse/20140918151302-65816706-data-migration-challenges-and-solution-for-successful-implementation