Data Integrity Strategies for Migrating Data from Legacy Systems

By **Ivan Soto Sep 26, 2017 6:29 am PDT**

Getty Images

Introduction

Due to the dynamic nature of technological changes and data integrity requirements companies are faced with the challenge of frequently migrating data from legacy systems.

Data migration can be very challenging and can create data integrity issues if not executed with an adequate strategy.

One of the biggest challenges is that companies fail to understand the magnitude and impact of migrating data from legacy systems while maintaining data integrity.

Prior to the execution of projects that require data migration companies need to understand the risk impact of the legacy data that requires migration.

Based on the risk impact of the legacy data companies need to define a data migration strategy that enables compliance with data integrity regulatory requirements.

This article will discuss data integrity strategies during the data migration of legacy systems.

Data Migration Challenges

The drivers for migrating data from legacy systems may be different from one company to the other but they typically include the following:

- Obsolete technology
- New regulatory requirements that legacy systems are unable to meet
- Vendor no longer supporting the technology
- Vendor goes out of business
- Vendor is taken over by another business
- Business decision to replace legacy system

There are several challenges that need to be well understood and managed to reduce or eliminate any data integrity compliance risk.

Some of these challenges include the following:

- Clear understanding of which data must be migrated
- Clear understanding of the risk impact of the data
- Ensuring data integrity during the migration process
- Defining an adequate data migration strategy
- Defining an adequate data migration verification strategy

Another challenge is developing a clear understanding of the type of data that will be migrated including the following attributes:

- Legacy data format
- New system data format
- Legacy data size
- Legacy data regulatory impact
- Legacy data retention period

One of the biggest issues with data migration is that companies normally don't understand the data that needs to be migrated. This creates the challenge that very often companies end up migrating data that does not need to be migrated to the new environment. This issue is driven by the lack of understanding of the record retention requirements of the legacy data. Data migration should only include data that is currently under record retention requirements that must be migrated to the new system. Migrating only relevant data enables efficiency and cost control during data migration projects.

The risk impact of the legacy data need to be very well understood prior to data migration . Data that has direct impact on Critical Process Parameters and Critical Quality Attributes is considered critical data that has a high-risk impact . The risk impact of the legacy data isn a critical input into the data migration strategy and planning .

Maintaining data integrity during data migration projects is critical from a compliance and business perspective. Data integrity is a critical regulatory requirement that needs to be part of the data migration strategy. The data migration strategy needs to be documented along with deliverables required to execute the data migration. The data migration strategy needs to ensure that the data will not be altered or lost during the data migration activities. The data migration strategy need to include alerts and notification of errors and failures during the data migration process. The data migration strategy needs to be documented along with deliverables required to execute the data migration.

To ensure that the goals of the data migration strategy are properly executed companies need to develop a solid data migration verification process. The data migration verification activities are intended to demonstrate and provide documented evidence that data integrity was maintained during the data migration process.

Data formats are a critical attribute that needs to be very well understood prior to data migration and included as an input in the strategy. Compatibility issues related to data format are a common problem often overlooked by companies prior to data migration.

The size of the legacy data it is another critical attribute that needs to be well understood prior to data migration. The size of the legacy data can be impacted by any data conversion that need to occur during the migration process. When migrating data to a cloud environment companies need to clearly understand the legacy data size requirements and future disk and memory space needed to manage the data during its lifecycle.

In summary, one of the biggest challenges with data migration projects is the lack of adequate understanding of this task, including planning, strategy and understanding of the critical data attributes that need to be well managed and understood.

The next section of this article will provide strategies and ideas about managing and controlling the data integrity risk associated with data migration projects.

Data Migration Strategies & Solutions

The migration strategy and planning can be initiated once companies develop an understanding of the data that needs to be

migrated, the risk and regulatory impact.

During the strategy and planning companies need to evaluate their options prior to data migration which includes the following:

- Archiving
- Keeping legacy data read-only
- Hybrid approach
- Full migration

Data archiving is the process of moving data that is no longer actively used to a separate storage location for long term retention. Before considering archiving, companies need to consider the following factors:

- How often the company need to access the data
- Criticality of the data
- High amount of data manipulation required to move to the new solution
- Project budget constraints
- Regulatory impact
- Data must be retrievable and accessible during retention period

Keeping the data read-only also requires consideration of the following factors:

- Cost associated with maintenance and licenses
- Limiting access to data
- No data changes
- Data record retention requirements

Archiving and making data read-only are options that should be consider as part of the strategy and planning phase during data migration projects.

Part of the data migration strategy and planning requires documented deliverables that are intended to define and document the data migration strategy and related verification.

For data migration projects companies need to create and execute the following deliverables:

- Data risk and impact assessment
- Data Migration Plan
- Data Migration Protocol
- Data Migration Summary Report

The data risk impact assessment is a critical activity that needs to be documented very early prior to any data migration activities. The data risk impact assessment is intended to provide the risk impact and level of the data that needs to be migrated. The data risk impact assessment is a key input to the overall migration strategy and plan.

Once the data risk and impact assessment is completed then the Data Migration Plan. The Data Migration Plan is intended to define and document the overall migration strategy and deliverables.

The Data Migration Plan should include the following information:

- Purpose
- Scope
- Roles and Responsibilities
- Migration Strategy
- Deliverables
- Data Migration Verification Strategy
- Data Risk Impact Assessment
- Identify Data to be Migrated
- Migration Tools
- Sampling Strategy
- Acceptance Criteria

- Deviation Handling

The Data Migration Plan is a key deliverable of the project that must be approved prior to initiating any migration activities. The Data Migration Plan requires Quality review and approval. The plan needs to define the sampling strategy which should be driven by the risk impact of the data. 100% sampling is not feasible and adequate therefore standards such as ANSI/AQL should be used to define the risk based sampling strategy. The data migration verification strategy need to provide evidence that the migration was successfully completed and data integrity was maintained. The data migration verification strategy should define any tools that will be used for the migration activities. Tool or techniques that can be used for data migration verification include source vs target data integrity checks that can be performed using the following tools:

- Cryptographic hash function
- CheckSum

Cryptographic hash function is a mathematical function that takes an input of any size and returns as an output an alpha numeric digest of a fixed size. Any alterations to the inputs will drastically change the digest. Cryptographic hash functions can be used to compute a digest of a data set at the source and then compute a digest at the target.

The Data Migration Protocol is a critical deliverable that is intended to provide documented evidence that migration activities were successfully executed and that they meet the established acceptance criteria. The Data Migration Protocol is an executable document that need to include how deviations and failures will be managed during the execution.

The Data Migration Protocol should include the following information:

- Purpose
- Scope
- Roles & Responsibilities
- Data Sampling Approach Strategy
- Migration Verification Tools
- Migration Verification Strategy
- Deviations
- Acceptance Criteria

Once the Data Migration Protocols is executed a Data Migration Summary Report need to be created. The Data Migration Summary Report is intended to summarize the results and deviation found during the execution of the protocol.

The Data Migration Summary Report should include the following information:

- Purpose
- Scope
- Roles & Responsibilities
- Migration Verification Results
- Migration Verification Summary
- Deviation Summary
- Acceptance Criteria Summary

These deliverables and related activities provide a structured and consistent approach to manage and control data migration projects. These deliverables are also intended to provide objective documented evidence that data integrity is maintained during data migration projects.

Summary

Data migrations requires adequate strategic planning to reduce the business and compliance risk. The Data Migration Plan is a critical deliverable that needs to be created to document the migration strategy.

The data risk impact must be assessed and documented and it is an input into the migration strategy and plan.

In summary, data migration is a critical activity that must be well managed and controlled to maintain the integrity of the data that need to be migrated.