



7/15/2016

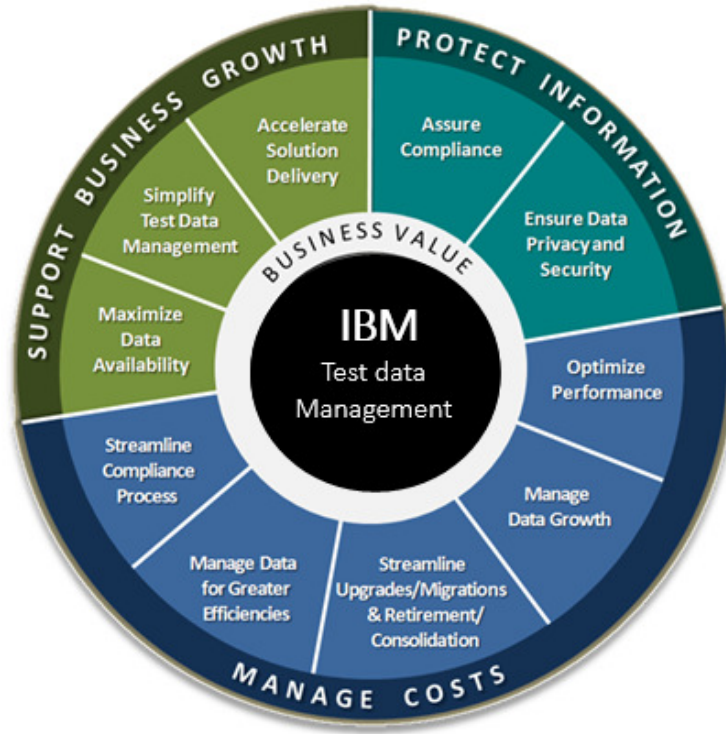
IBM Test Data Management View Point



Raghavendra N S & Rajasekhar M Atmuri
IBM CORPORATION

Background

With the changing trends in technology and software development approach in the modern IT segment, test data management has become an essential and integral part of software development life cycle. The adoption of a progressively more supple (e.g. Waterfall to Agile) and vibrant development process requires rapid access to the appropriate test data. There is a paying attention on test data due to its large dependency on quality of testing. However, quality of test data remains one of the big challenges for testers. Information technology or software industry statistics reveals that nearly 30% of test execution failures are due to improper test data. In order to prepare or generate the test data, one should have good strategy in each of the TDM practices around test data creation and maintenance, namely Data Subset, Data Masking, Test Data Ageing, Test Data Refresh, Data Archive, Master Copy and

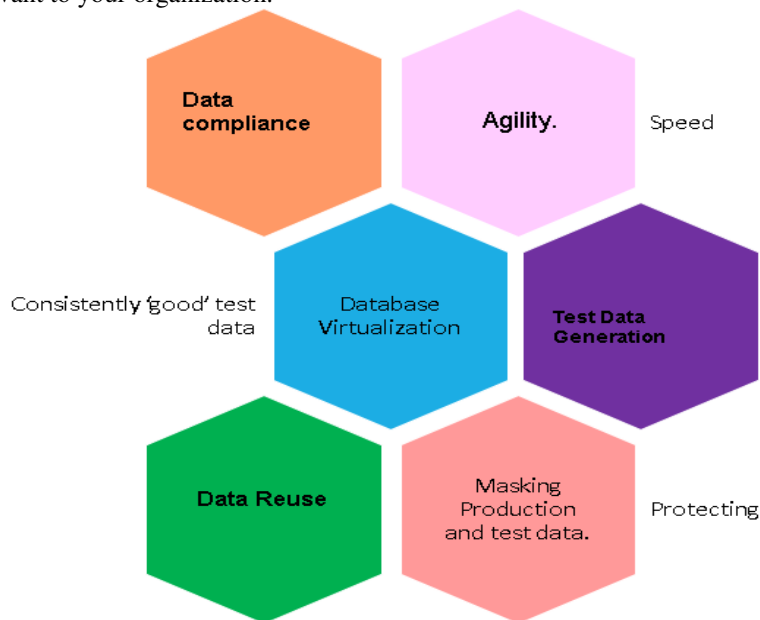


Synthetic Data Generation.

Test Data Management - Strategy

Identify Strategy that works for you

An organization can deploy one or more Test Data Management strategies depending on their needs. Any Test Data Management strategy deployed should target to address all or some of the following key industry challenges that are relevant to your organization.



Set Your Goals

The Test Data Management strategy like all strategies must have goals and ways to achieve those goals. Strategic components are paired with goals to achieve targeted business results. To meet our goals, the Strategy must yield results that are reusable, consumable, scalable and measurable.

Test Data management Goals, Strategies and Targeted Results

Consumable

Goal	Strategies	Out come
Quality Testing User Experience	<ul style="list-style-type: none"> • Self-service data • Data Catalog 	<ul style="list-style-type: none"> • Decrease post-release defects • Mapping across Business scenarios, requirement test cases and data request • Streamline data acquisition process • Increase coverage for testing • Provide supporting metrics

Scalable

Goal	Strategies	Out come
Compliance Traceability	<ul style="list-style-type: none"> • Service model • Governance 	<ul style="list-style-type: none"> • Meet and exceed industry standards • Build enterprise wide knowledge • Increase operational efficiency • Increase bandwidth • Data Quality analysis to identify typical Inconsistencies in test environment

Reusable

Goal	Strategies	Out come
Standardization and Modularity	<ul style="list-style-type: none"> • Data Modeling • Data Management - create Gold data provide de-identified test data for versioning test beds • Knowledge transfer - enablement on tools and best practices 	<ul style="list-style-type: none"> • Reliable, Reusable data models • Flexible reusable data extracts • Knowledge base artifacts • Component based solution • Reusable formats to ensure consistency in data request

The above table lists some of the more broad-based goals that the strategy will address. There are certainly other goals and other benefits that can be derived from the deployed strategies

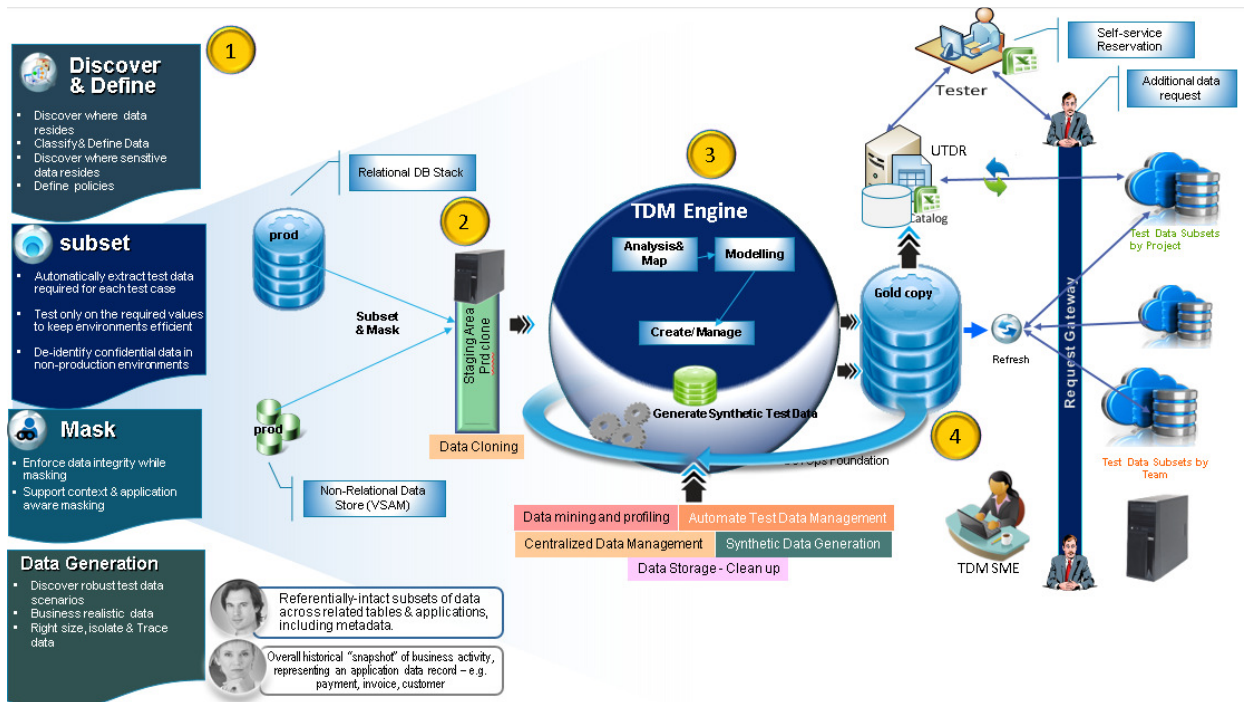
Test Data Management Strategy Prelude

There are essentially three significant elements to the TDM Strategy, each with detailed components that define broad-based strategies that are both foundational and tactical. The three elements of the TDM strategy are:

- Beneficiary Analysis – understanding past and present TDM demands, who the consumer is, What business requirements and priorities exist and how these factors will influence which consumers, from which key application systems, become beneficiaries. Also, consider the following factors while choosing a strategy as part of the Beneficiary analysis
 - Cost of implementing & maintaining chosen TDM approach vs. immediate and long term benefits projected
 - Phased approach in implementing TDM addressing business needs & challenges in priority vs. big bang implementation
 - Cost of in-house tools vs. commercial tools / packages

- Future proof your approach considering the fast pace of changes in industry / technology within the context of your overall organization vision
- The Test Data Management Approach – the process for decomposing complex systems, crafting data models, identifying relationships, extracting data based on accurate modeling, and ultimately the final product – data creation.
- Service Model and Supporting Governance – the governing process for the TDM operational model, including roles and responsibilities for executing on the strategy.

An Effective Test Data Management Approach



Following is a high level sequence of steps for an effective Test Data Management Approach

Discover Data

Identify right source of data (databases, flat files, mainframes, etc) and create data models understanding how relationships among data are structured across systems / data sources. These data models are going to provide the basis to identify the source for various business objects with data scattered across systems. Consider using tools with Automated Data-Relationship Discovery function.

Subset & Mask

Using entire production data for testing is not cost effective and can often become counterproductive during testing cycles. Identify a subset of data that provides coverage for maximum test scenarios. A good Test Data Management tool provide capability to extract data from different data sources for an entire business object maintaining the referential integrity.

Data breach is one of the biggest challenges in the industry, that limits the as-is usage of production data for testing. The Production data typically contains sensitive information (SI) and personally identifiable information that you

are forbidden to use in testing environments. Identify the sensitive data within the context of your organization's data security & compliance requirements and masking functions to transform all sensitive information into data acceptable for testing. Capabilities for de-identifying confidential data must provide a realistic look and feel, and should consistently mask complete business objects such as customer orders across test systems.

Create Gold Copy

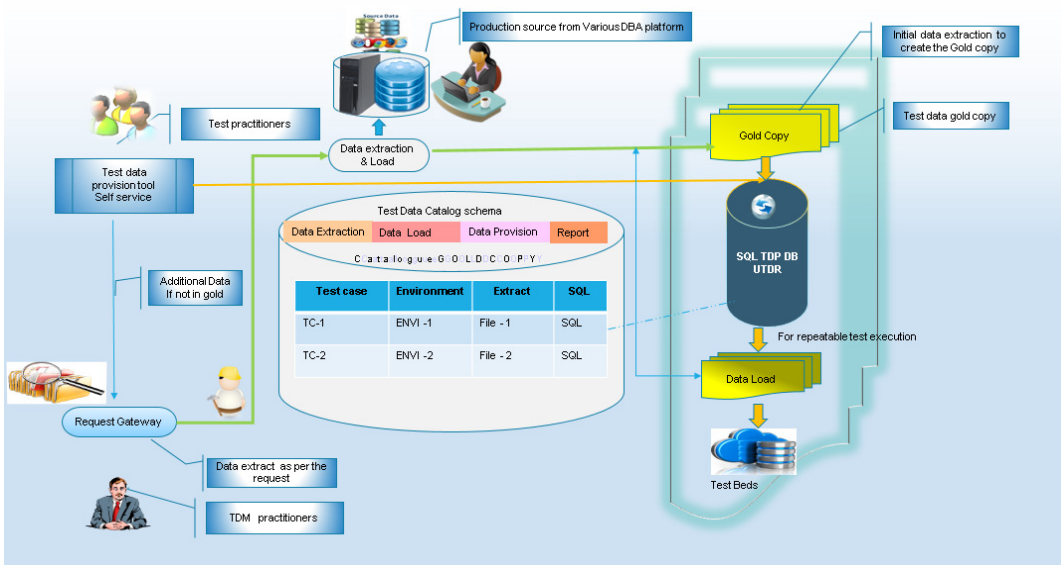
Use Gold copy as a repository of your test data that can be reused on-demand fulfillment of all future test data requests. Test Data on Demand play a vital role in minimizing the testing cycles as it cuts down the time spent on test data identification and test data generation / creation from each test cycle.

Test Environment refresh

Over a period of time, the test data diverges from the realistic data due to new releases in production. Enforce governance to ensure the test data in gold copy is refreshed with production data regularly to ensure the efficiency of testing. Also, governance should be in place to periodic / on demand refresh of test environments from the gold copy.

Workflow based request provisioning system

A workflow based request provisioning system provides a common interface for placing the test data requests and data management activities. This ensures request management, automated provisioning, mapping data requests to test cases and optional reservations of data. Internally the solutions follows a modular structure for organizing data management artifacts so that reuse of techniques can be maximized. Well defined data template ensure consistent format for data request. The solution maintain federated data realization and ensure they are adhered to while provisioning. Further, it is important for solution to provide adequate visibility into test data management activities, through a centralized test management tool.



Business Benefit



IBM Test Data Management Implementation Approach

Engage TDM consultant to choose the right TDM toolset

Identification of right tools is vital to the success of Test Data Management implementation in any organization. Recommend engaging a TDM consultant with broad experience of implementing Test Data Management solutions to drive the TDM tools identification process

- Determining their goals and tool selection criteria.
- Proposing a shortlist of TDM tools that meet the client's needs.
- Scoring the shortlisted tools against the selection criteria.
- Analyzing the scores and providing selection advice to the client's stakeholders.

Proof of Concept

Conduct proof of concept to validate the feasibility of the Test Data Management Strategy and the selected tool sets are ideal candidate for implementing Test Data Management solution, delivering the required test data. IBM provide this proof of concept, including:

- Determining the solution models for selecting, masking, generating and distributing test data, and drawing up the data design.
- Creating data scenarios for selecting, masking and generating test data, and creating and testing the scripts.
- Preparing the distribution of the test data.
- Executing the prepared scripts and distributing the test data to the designated test environment.
- Presenting the results and evaluating them with the client.

Pilot

Demonstrate that the proposed TDM process is workable and can deliver the expected benefits, by implementing the process for one system / application area. IBM offers the services to implement TDM Pilot that include:

- Defining the TDM Pilot – determining its goals, tool environment, scope, data coverage and control measures used.
- Preparing the TDM Pilot – installing and configuring the tooling, defining the TDM process outlines for the pilot
- Organization workshops for determining and analyzing the general test data requirements to be implemented as part of the pilot
- Performing the TDM Pilot – executing all stages of the TDM model within a pilot environment, including performing the most risk-full activities in all stages of the model.

Implementation

During this phase, IBM introduces the selected toolset and the TDM service acts as an instrument for test data. Services provided include Tools Implementation and full TDM Rollout.

IBM introduces the full TDM cycle into the customer organization. Once the TDM process has been set up, it can be applied to new software development or maintaining existing applications. IBM sets up the data for an application/new release via a four-step process:

Discovery. We understand the demand for test data in the particular client situation, then determine and detail the test-specific requirements – the number of test cases, scope of the data set, etc.

Design. To turn demand into a design of the test data, IBM determines the solutions for selecting, masking, generating and distributing the test data based on what is in scope.

Prepare. This phase consists of creating the data scenarios for selecting, masking and generating the test data scope, and creating and testing the scripts. Tools are set up to generate the exact dataset required – either by selecting the right data or creating the data, or a combination of both. The distribution of the test data is also prepared.

Provision. After the TDM tooling has been set up and tested, data is generated and provided into the right environment at the right time. Other available services during this stage include:

- Managing the provision of test environments with test data.
- Keeping test data sets up-to-date.
- Reseeding test environments with data from the central test data repository.
- Rolling back the environment to the virgin data state