

SELF-SERVICE DATA PROVISIONING AND MASKING

A DATA MASKING SERIES CASE STUDY

Developers and QA mask and copy the production data they need to test in non-production environments using a self-service model.

Introduction For a Financial Services Company, the goal was to self-service subsets of production data to test applications. For this client, Axis customized DMSuite™ 2.0. The application is metadata driven, integrates data obfuscation, and provides repeatable process of provisioning of production data to non-production environments. The Axis' client with DMSuite™ 2.0 reduced time to provision and refresh production quality data to non-production environments from 2 weeks to 24 hours.

1 Project Details

The goal of the 12 week engagement was to provide a self-service provisioning solution which fit exactly to the client's existing development and testing infrastructure. The application had to generate and import jobs to Datastage, schedule and submit those jobs with Maestro, and move the data between mixed Oracle environments of 9i and 10g databases. The application also had to support provisioning/refresh of data within 24 hours.

2 Value Statement:

- DMSuite™ is web based application and handles the entire process of provisioning data. No application installation is required.
- A user is validated using MS Active Directory. The time to get a user up and running is only a matter of adding the user to the application by an administrator.
- DMSuite™ automatically identifies the schema differences between source and target database objects such as tables, views, triggers, and sequences. Therefore, DBAs when they first create the target database can identify any outstanding issues. This radically reduces the go back and forth between the DBAs and application development & QA teams.
- DMSuite™ automatically schedules the jobs with a scheduling console, such as Maestro. The jobs run in parallel based on resource definition, and can be scheduled to run during off-peak hours.
- Data refresh is a one button execution. The application identifies the delta between source and target before provisioning and alerts users to changes before loading data.
- The application manages sequences, especially useful when loading data to the target database for the first time.
- The learning curve is minimal. All a user needs to provision data is to identify the tables they need. The application can identify what additional objects are necessary based on the source DDL.
- Users have ability to filter data for a given table by defining a where clause.
- All the logs are consolidated in DMSuite's status page. Within one application users can view ETL tool logs, as well as database specific logs to identify any issues with data.
- DMSuite is bundled with SecureDev module. This module can both mask and provision data. DMSuite automatically schedules jobs with SecureDev, when the ETL tool might not support a data type such as CLOBs, NUMBER without precision and scale, etc.
- For all environments teams can define a 'Collection' of reference tables. As a result, it is easy to reuse existing collections by all application users.

In Closing The project was a stunning success for Axis and our client partner, culminating in company-wide deployment of DMSuite™ 2.0 by the CIO of the company. The turn-around time to provision production quality data was reduced from 2 weeks to 24 hours. First two application teams provisioned 157 GB of data, ~450 million rows from 252 tables.

