

2015



PROLOGIS
Experts in Aviation Consultancy

Low-Cost Enterprise Data Warehousing

Authors:

Bernd Tieke
Senior Consultant
tieke@prologis.aero

Matthias Hansen
Management Consultant &
Business Development
hansen@prologis.aero

Copyright © 2015 PROLOGIS AG

All rights reserved. This study or any portion thereof may not be reproduced or used in any manner whatsoever without the express written permission of the publisher except for the use of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law.



Low-Cost Enterprise Data Warehousing

The enterprise data warehouse (EDW) as an integrated information pool of disparate data sources in a company has been highly appreciated by our customers over the last few years. However, the road to creating a working and stable data warehouse can be a rocky one.

Depending on the modeling technique and the number of connected source systems, the number of objects in the underlying database can easily exceed several hundreds.

In the past, the structure and data flows for all these objects had to be coded manually, which resulted in a huge amount of work in terms of development and testing. For the developers, this has been quite tedious and it also stands in contradiction to the “don't repeat yourself” paradigm. Also, with today's fast changing business, it is inevitable to constantly adjust the data warehouse to meet current needs. Given the manual approach, the effort for the adjustment processes is sometimes way too high, thus resulting in high costs or even worse, loss of interest by the business, because of simply having to wait too long.

To solve this problem, PROLOGIS changed its approach to data warehousing in the past few years. From an architectural point of view, we use the Data Vault Methodology invented by Daniel Linsted for taking a clear and consistent approach to the core data warehouse.

We also collect all relevant meta data information in one application. This information consists of

- ☺ statements to get the source data out of the staging area
- ☺ columns and their respective data types
- ☺ the data vault model
- ☺ comments and documentation.

From that information pool, the application can generate code for data flows and database structures automatically. This output can be easily deployed to development, test and production environments.

Most importantly, the underlying templates for code generation can be modified to suit the customer's needs. We can adapt to specific databases or ETL (Extract, Transform, Load) tools, depending on what has to be used due to the individual IT landscape/architecture. The files containing the information about the data warehouse can be source-controlled and thus provide a history of past changes. Finally, regarding changes of the data warehouse, the concentration of all relevant information in one application helps in identifying the objects and data flows to be changed.

By taking this approach, PROLOGIS was able to successfully cut development costs for the enterprise data warehouse significantly in real world projects. It allowed us to concentrate on business modeling instead of code debugging. If you are having difficulties in setting up an enterprise data warehouse, extremely high costs or you are experiencing long time gaps between describing your requirements and delivery of the changes, please get in touch with us at it-services@prologis.aero.