

## > Don't let your Warehouse go Stale!

>Raphael Klebanov, *Customer Experience at WhereScape USA*

### >Preface

In the 1960s, Thomas Kuhn – American physicist, historian, and philosopher of science – coined the term '*paradigm shift*', which became indispensable in the world of Information Technology. The high-tech advancements in the data-warehousing realm are not evolutionary – like many authors claim – but rather a '*series of peaceful interludes punctuated by intellectually violent revolutions*',<sup>1</sup> and in those revolutions '*one conceptual world view is replaced by another*'<sup>1</sup>, hence the '*paradigm shift*'.

<sup>1</sup> *The Structure of Scientific Revolutions* by Thomas S. Kuhn, 50th Anniversary Edition, 2012

This article describes the shifting paradigm of the modern data warehousing and how WhereScape is positioning itself to cope with it. Both WhereScape products - 3D and RED - are undergoing constant improvement and adaptation to the "intellectually violent revolutions" of the Data information world.

Many new features are added in every RED and 3D release; it is hard to catch up with fast-ripening WhereScape software. This article aims to shed light on those features so our customers and prospects have a comprehensible picture as to why they would want to use them.

### >Paradigm is Shifting...

The old-fashioned analytical data warehouse is not catching up with ever-accelerating business changes as a result of exploding data. Essentially, what business needs in paradigm-shifting times is to mitigate the risk in decision making—pretty much the same as always, but with new tools and a greater volume of analytic insight.

And in this quest for more effective decision-making, data warehouses become even *more* critical to running business operations. Whether analyzing sales performance or studying weather patterns, enterprises rely on warehouse information to support operational analytics and make sound business decisions even faster than they could before. Highly available and secure data warehouse and information systems solutions became crucial for the modern business.

### >Any Data is Good Data

Data warehouses can't limit themselves to relational/structured/"clean" data. Often, the data warehouse needs to deal with data in its raw form, including semi-structured, first-class formats such as XML and JSON. Data warehouses must be able to not only bulk-load large data sets as is, but also scale linearly in a cost-effective manner as volumes grow, without upfront planning. To this end, distributed environments such as Hadoop® enable businesses to run reliable, linearly-scalable computing clusters in tandem with WhereScape software. This approach allows businesses to preserve the wealth of the data while avoiding the cost of custom code in order to integrate data from various sources before analysis.

What  
Modern Business  
Needs?

The ability to utilize semi-structured, hierarchical, and "flat" formats in addition to traditional (relational) data stores, in a scalable and cohesive way.

What  
WhereScape  
Offers?

Big Data Technologies Support

1. Oracle Loads from Hadoop - UNIX/Linux script based via OSCH (Oracle SQL Connector for HDFS) and OLH (Oracle Loader for Hadoop).
2. Teradata TPT Loads from Hadoop - UNIX/Linux script based.
3. Creating Models from Hadoop and HDFS connections (in 3D)
4. More...

## >Accelerated Processing

Today's enterprise analysis cannot be accomplished with yesterday's data. Business needs insights based on real-time or near-real-time information. Increasingly, companies unlock significant value by making information usable at much higher frequency, hence supporting faster decision-making to adjust business "switches" immediately.

What  
Modern Business  
Needs?

High performance solutions for supporting new BI analytic (huge) workloads: real-time or near-real-time processing.

What  
WhereScape  
Offers?

High Performance Solutions

1. MAPS PDW as a target for a number of objects and associated stored procedures (Microsoft)
2. SQL Server Integration Services (SSIS) as a load method for Flat Files from Windows connections SQL Server Integration Services (SSIS), with loads enabled for Teradata (Microsoft/Teradata)
3. SQL Server Integration Services (SSIS) Exports
4. Greenplum Exports
5. SQL Server 2014 Clustered Columnstore indexes supported
6. SSIS Loads, now with fastload options (Microsoft)
7. View Objects that can now be replicated to another DB. Schema (Microsoft)
8. Alter View syntax when creating/recreating views on SQL Server
9. No longer forcing conversion of user/pwd to UPPER when defining SQL Server Linked Server Login detail for Linked Server definition
10. Teradata 15 as a DW DB
11. Support for column transformations in TPT Update loads (Teradata)
12. Passing in parameters to stored procedures in Netezza.
13. More ...

## >Data Discovery before Modelling/Implementation

Historically, data warehouses operate on a single data design model so as to be loaded and analyzed. In reality, data sources are varying, requirements change, and the "insight value" of various data are frequently unknown at the time they are collected. New data warehouses must allow for the discovery of heterogeneous data for analysis without "prettifying" the data; the analyst needs to be able to create the whole iteration of the data warehouse from data sources through EDW and the presentation layer.

What  
Modern Business  
Needs?

Improved data modeling and integration approaches: automation; integration between design and implementation stages.

What  
WhereScape  
Offers?

Automation and Integration

1. Models exported from WhereScape 3D as a "RED Builder Model" can now be imported into Oracle database via the RED Application Deployment tool
2. Models exported from WhereScape 3D as a "RED Builder Model" can now be imported into Teradata database via the RED Application Deployment tool
3. RED integration supported (Joins on stage tables; Unique short names; Change detection fields; RED projects; Syncing of object types; Business key flags; Surrogate key flags)
4. The query join type can now be specified on the source mappings of entities.
5. Search and replace functionality has been added to modify model conversion rules.
6. A model conversion rule was added to remove indexes.
7. A model conversion rule was added to remove attributes.
8. The model conversion rule 'Add relationships from source models' was added.
9. The model conversion rule 'Add source mappings from FKs' has been added.
10. Functionality was added to remove all subtype relationships.
11. More...

## >Data Analysis Cost should not Break the Bank

Traditionally, data warehouses were designed for small amounts of well-structured, relational data. With modern-days high volumes and speeds, the costs escalate rapidly. Newer MPP-style analytics platforms using columnar storage to predictably handle high volumes are becoming more common. Big Data is already taking us to uncharted territory. New data warehouses utilize a parallel and distributed approach to handle larger volumes of ever-growing and ever-diversifying data.

What  
Modern Business  
Needs?

Low cost solutions to motivate even more useful processing and analysis of Big Data.

What  
WhereScape  
Offers?

Data Governing, Security & Standardization

1. New Oracle multiple user login method for the RED repository
2. Maximum Groups have been increased to 250; and maximum Projects increased to 1000 (Common)
3. RED now handles tables with more than 512 columns (Common)
4. Models can now be exported in the Common Warehouse Metamodel format (common)
5. More...

## >Standardized Approach

Without a data model in place, you can't create a data warehouse. Without SQL language – and its multiple dialects – businesses are forced to write expensive proprietary code that is inevitably discarded when technology or tools change. Data warehouses should be able to run SQL queries against data of agnostic nature, residing in Oracle, Netezza, or on a distributed file system.

What  
Modern Business  
Needs?

Modify data governance tactics; enhanced security approaches; conformity with the information platforms and methodologies.

What  
WhereScape  
Offers?

WhereScape offers Usability and User-friendliness

1. Improved diagram printing issues (Common)
2. Usability/ease of use redesign (3D):
3. Improved handling of surrogate keys;
4. Adjustment of default values;
5. Allowing diagrams to open as read only;
6. Navigating from the tree to the diagram;
7. Introduction of a category manager;
8. Restoring the tree state at start-up
9. The background color of entities can now be manually set (3D)
10. Model categories can now be managed more easily (3D)
11. Table background colors can be set in the diagram (3D)
12. View Objects can now be replicated to another DB. Schema (Oracle)
13. "Display Data" from the Object Pane now handles display of Oracle Timestamp datatypes
14. More...

## >Conclusion

As our clients are faced with ever-growing volumes and varieties of data, WhereScape aims to organize and process this valuable information to provide critical decision-making insight. Every new RED and 3D release comes with a collection of invaluable features to make your data warehouse more flexible, responsive, transparent, and standardized—while at the same time lowering costs to promote scalability. With all these paradigm shifts in the data warehouse world, it is important to remain current and aware of the value of your data.

If you have any questions about any of these tools and how they can be helpful to your business, feel free to contact me, Raphael Klebanov, at [raphael.klebanov@wherescape.com](mailto:raphael.klebanov@wherescape.com) or 303.968.0703. And if you have not already done so, upgrade WhereScape to latest and greatest version to take advantage of these brand new features—and other improvements: it is free and non-intrusive. Use this link: <https://www.wherescape.com/support/download/>