

Data Integration

Automation and Acceleration

WhereScape®
and
TERADATA®

WhereScape®

Jack Howard – Principal Architect
jack.howard@wherescape.com
Roger Walton – Global Enterprise Accounts
roger.walton@wherescape.com

The trouble with data warehousing

“The business can't articulate what they want and/or they keep changing their requirements”

“It's too difficult and complex to get the data ready for analysis”

“Can't validate the source data”

“Can't iterate because everything takes too long”

“No time to manage governance”

“No time to create documentation”

“The business doesn't understand and appreciate IT's difficulties”

“IT doesn't understand and appreciate the business's problems”

“IT tries to prepare for every eventuality”

“Give me access to the data and we'll do the rest”



Founded in 1999 with
offices in USA, Europe
& Asia Pacific



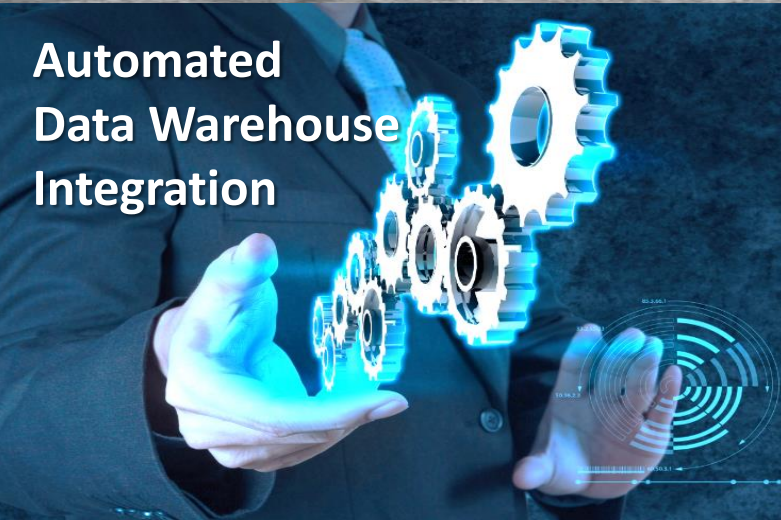
Deep roots in data
ecosystem expertise



Time to value
driven



Our mission is to **dramatically reduce** the risk & time to
deliver **high quality** data solutions.



Automated
Data Warehouse
Integration

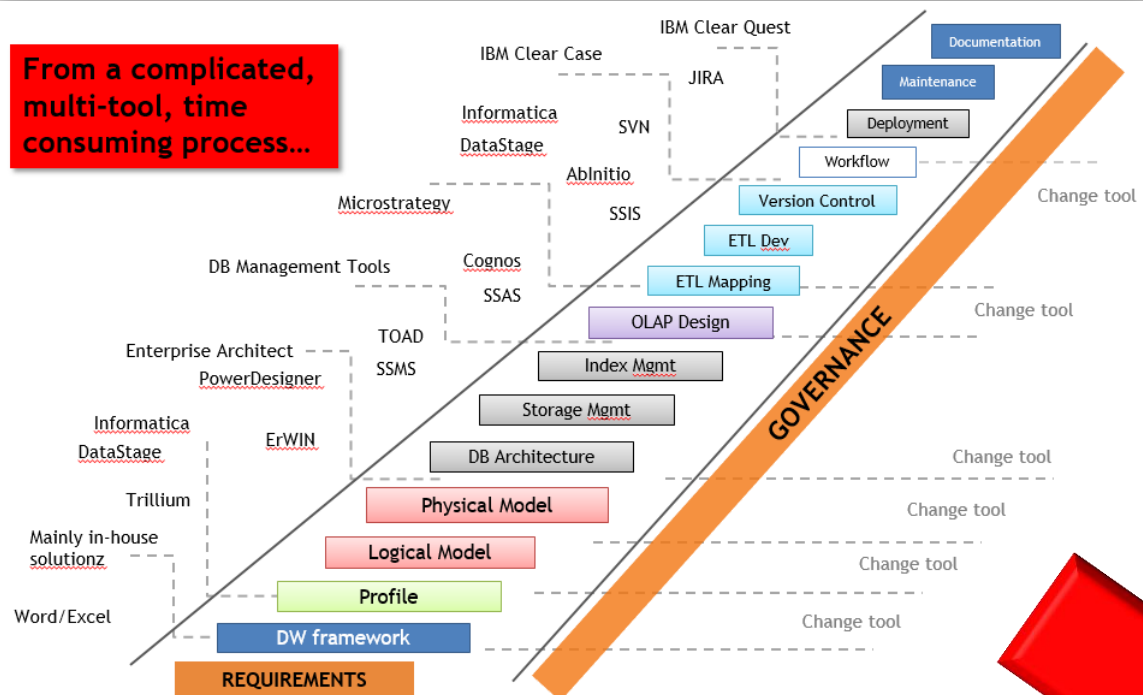


500+ Customers
Worldwide



Worldwide
Recognition

From a complicated, multi-tool, time consuming process...

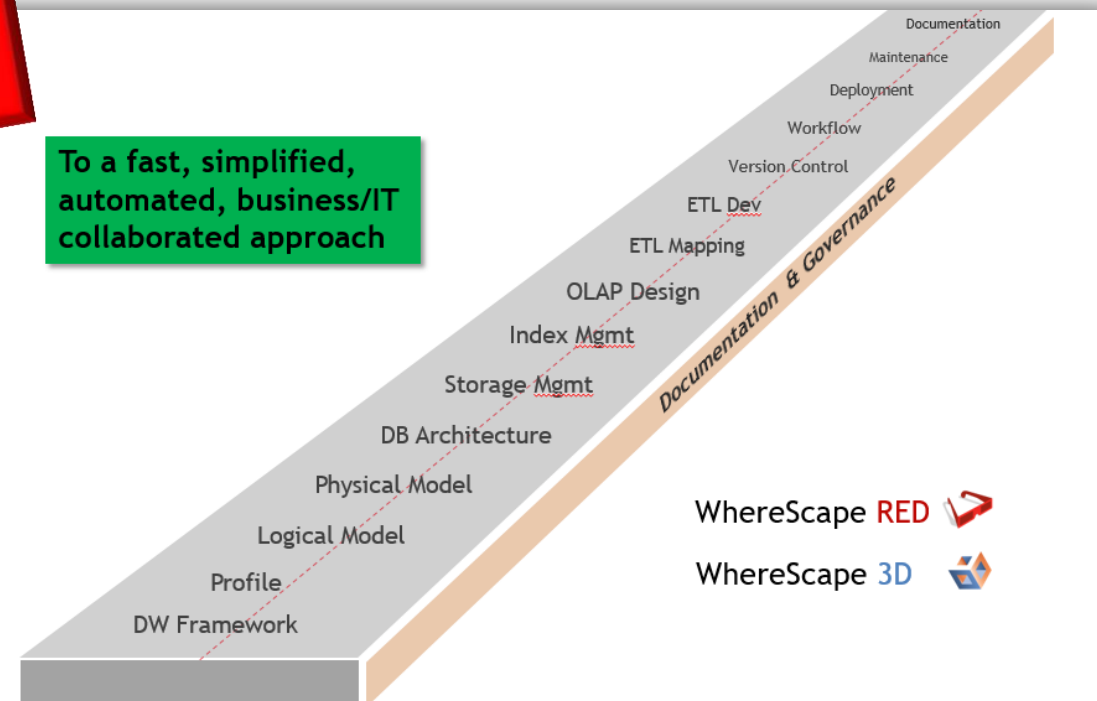


WhereScape's software is a proven and sophisticated application:

- that designs, builds and operates data warehouses
- automates repeatable best practice development standards
- with built in data governance
- using metadata to drive agile delivery
- with consistent quality and full documentation
- and delivers value, faster, to business & IT stakeholders
- that saves time and money for our customers

A fully integrated approach to Simplify, automate and accelerate

To a fast, simplified, automated, business/IT collaborated approach



Product approach



3D - Discovery, Design, Model

- Discover, profile, explore and document any potential source system
- Auto generate models
- Design, model and test any target data warehouse schema using source system data
- Investigate a data model in conjunction with source system data
- Perform a complete source-to-target mapping
- Test planned schema populated with real or obfuscated test data
- Create and manipulate conceptual and logical views
- Generate task-based views of the project while maintaining data-based and model-based views of the project and assign resources, time and cost
- Generate complete project documentation - automatically



RED - Build, Manage, Document

- Integrated rapid development environment driven by robust metadata with full code reusability
- Drives standardization of processes from source to target
- Easily handles extracting and moving data from almost any source system to target - supporting fastload, multiload and tpt.
- Supports customization so that existing procedures and code can be integrated into the metadata layer.
- Incorporates flexible scheduler and supports re-startability inherently. And can be called from other schedulers like Tivoli and Control-M
- Supports version control with checking in and out of code – including code promotion from one environment to another.
- Easy export of data warehouse metadata to another target system.

Typical customer experiences

- Global Retailer
 - “3-year project backlog being delivered in just 6 months
- Global Insurer
 - “I wish we had this 10 years ago”
- Global Automotive
 - “For every 6 weeks of work I can now deliver the same in 6 days”
- Teradata Professional Services
 - “We spent several man-months trying to do what was demonstrated in 90 minutes”

TESCO

Micron

MedAssets

BlueCross BlueShield of North Carolina

BlueCross BlueShield Minnesota

GE Aviation



vodafone

RBS The Royal Bank of Scotland

GroupHealth



IRVING

Wynn LAS VEGAS

COSTCO WHOLESALE



f5 NETWORKS

NORDSTROM

HSBC



ZURICH

UnionBank

EnergyAustralia



RESERVE BANK OF AUSTRALIA

UNIVERSITY of WASHINGTON

WELLS FARGO

Wake Forest Baptist Health



LOTTO

Fonterra Dairy for life

xerox

FIRST CAPITAL

OMNI HOTELS

JUNIPER NETWORKS

Sainsbury's

Bank of America

SingTel



CANADIAN WESTERN BANK

NetworkRail

United Rentals



COWBOYS



OPTUS COMMUNICATIONS

SKYCITY ENTERTAINMENT GROUP

SKYCITY ENTERTAINMENT GROUP

Telstra

CN

General Cable

HIGHMARK BLUE SHIELD



Asahi PREMIUM BEVERAGES

eHealth

AIR NEW ZEALAND

WhereScape

Data Warehouse Automation

WhereScape®

Technology Timeline

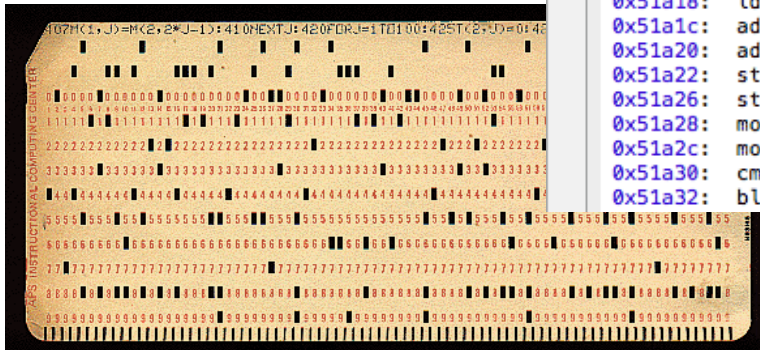
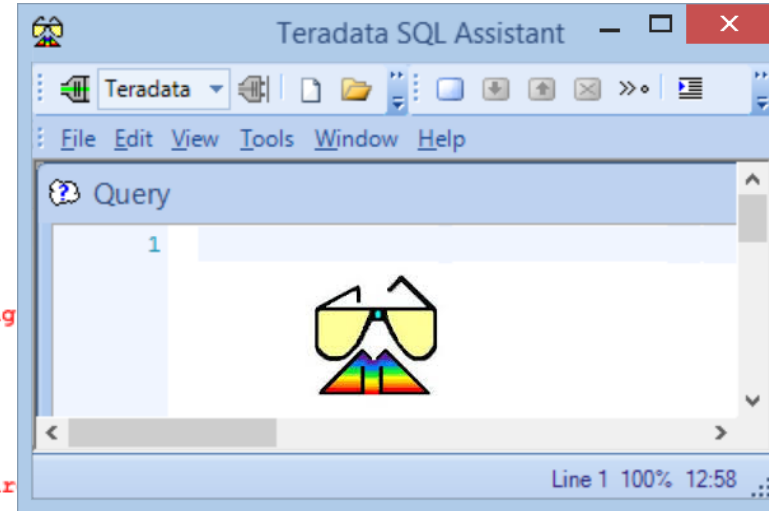
Assembler

```
0x519fe: mov    r0, r10
0x51a00: bl     0x51dd4
0x51a04: ldrex r0, [r6]
0x51a08: cmp   r0, r5
0x51a0a: bne   0x51a14
0x51a0c: strex r1, r4, [r6]
0x51a10: cmp   r1, #0
0x51a12: bne   0x51a04
0x51a14: cmp   r0, #0
0x51a16: bne   0x519fe
0x51a18: ldr.w r0, [r11]
0x51a1c: add.w r8, r8, #1
0x51a20: adds  r0, #1
0x51a22: str.w r0, [r11]
0x51a26: str   r5, [r6]
0x51a28: movw  r0, #38527
0x51a2c: movt  r0, #152
0x51a30: cmp   r8, r0
0x51a32: ble   0x519fe
```

flag

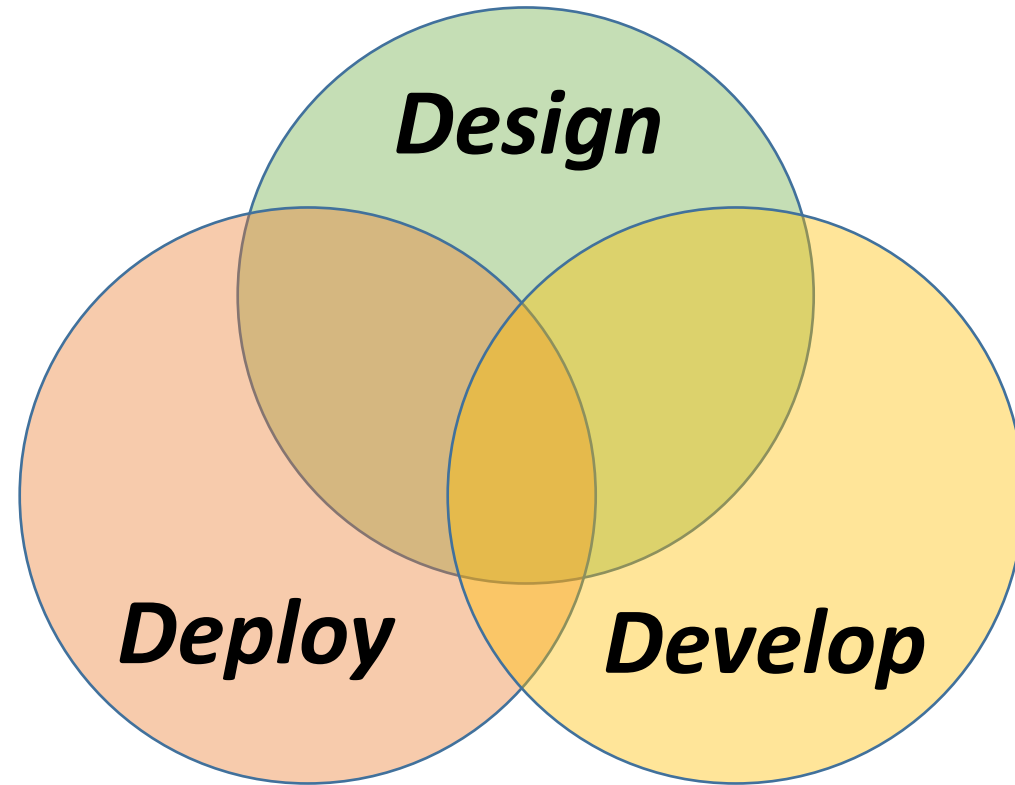
share

flag.store(...)



Running: faster, but tiring and eventually you get injured

Domains Automation must address



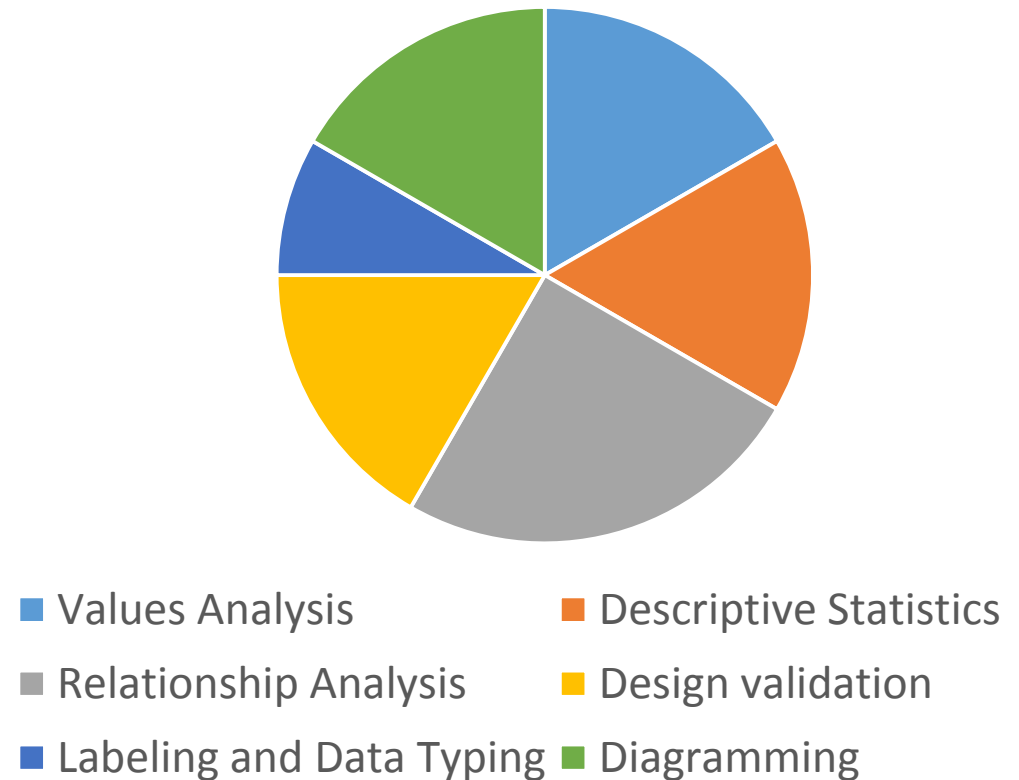
Design Automation

WhereScape®

Database Design is Data Intensive

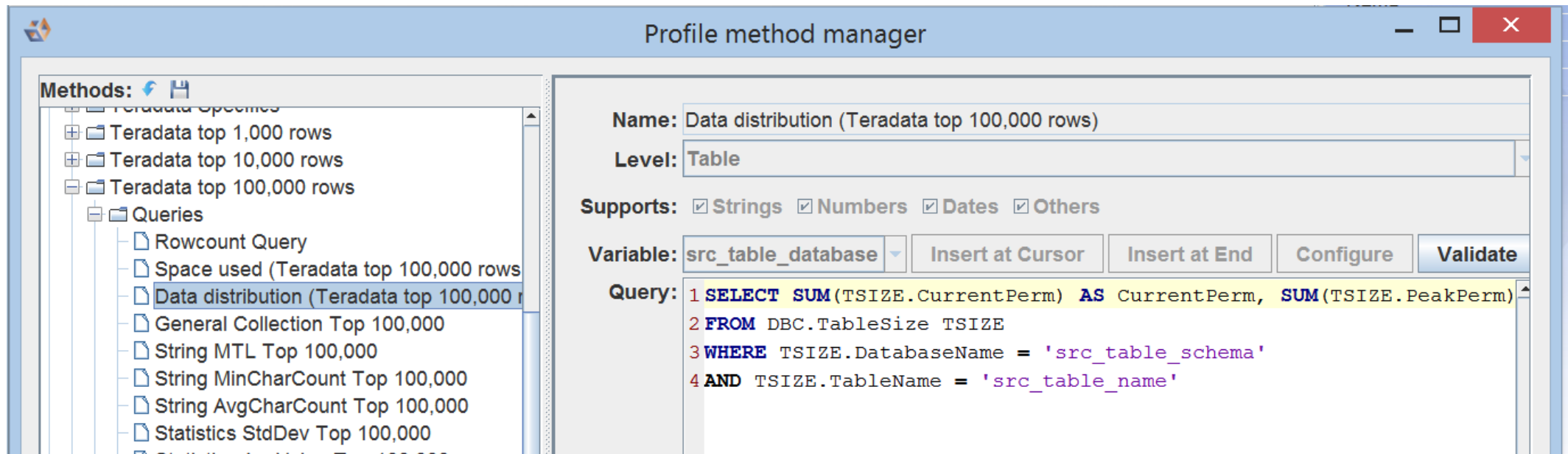
- *Complex queries are needed to understand data for design needs*
- *Diagramming is a small percentage of the work*

Database Design Effort



Profiling Automation

- *Extensible, scalable profiling*



The screenshot displays the 'Profile method manager' window. On the left, a tree view under 'Methods:' shows a hierarchy of profiling methods, with 'Data distribution (Teradata top 100,000 rows)' selected. The main area shows the configuration for this method:

- Name:** Data distribution (Teradata top 100,000 rows)
- Level:** Table
- Supports:** Strings Numbers Dates Others
- Variable:** src_table_database (with buttons for 'Insert at Cursor', 'Insert at End', 'Configure', and 'Validate')
- Query:**

```
1 SELECT SUM(TSIZE.CurrentPerm) AS CurrentPerm, SUM(TSIZE.PeakPerm)
2 FROM DBC.TableSize TSIZE
3 WHERE TSIZE.DatabaseName = 'src_table_schema'
4 AND TSIZE.TableName = 'src_table_name'
```


Profiling Automation

- *Relationship profiling!*

The screenshot displays a database diagram with two tables: **Shippers** and **Big_Orders**. The **Shippers** table has columns: ShipperID (PK), CompanyName, and Phone. The **Big_Orders** table has columns: OrderID (PK), CustomerId, EmployeeID, OrderDate, RequiredDate, ShippedDate, ShipVia, and Freight. A relationship line connects ShipperID in Shippers to ShipVia in Big_Orders.

Below the diagram is a "Metric profiling results filter" window. The "Metrics" dropdown is set to "Relationship - single value". The results table is as follows:

Profiling metric	PK Entity	PK Attribute	FK...	FK Entity	FK Attribute	Profiling res...
In use %	Shippers	ShipperID	dbo	Big_Orders	ShipVia	100.0
Referentially invalid values %	Shippers	ShipperID	dbo	Big_Orders	ShipVia	0.0

Below the results is a SQL query editor showing the following query:

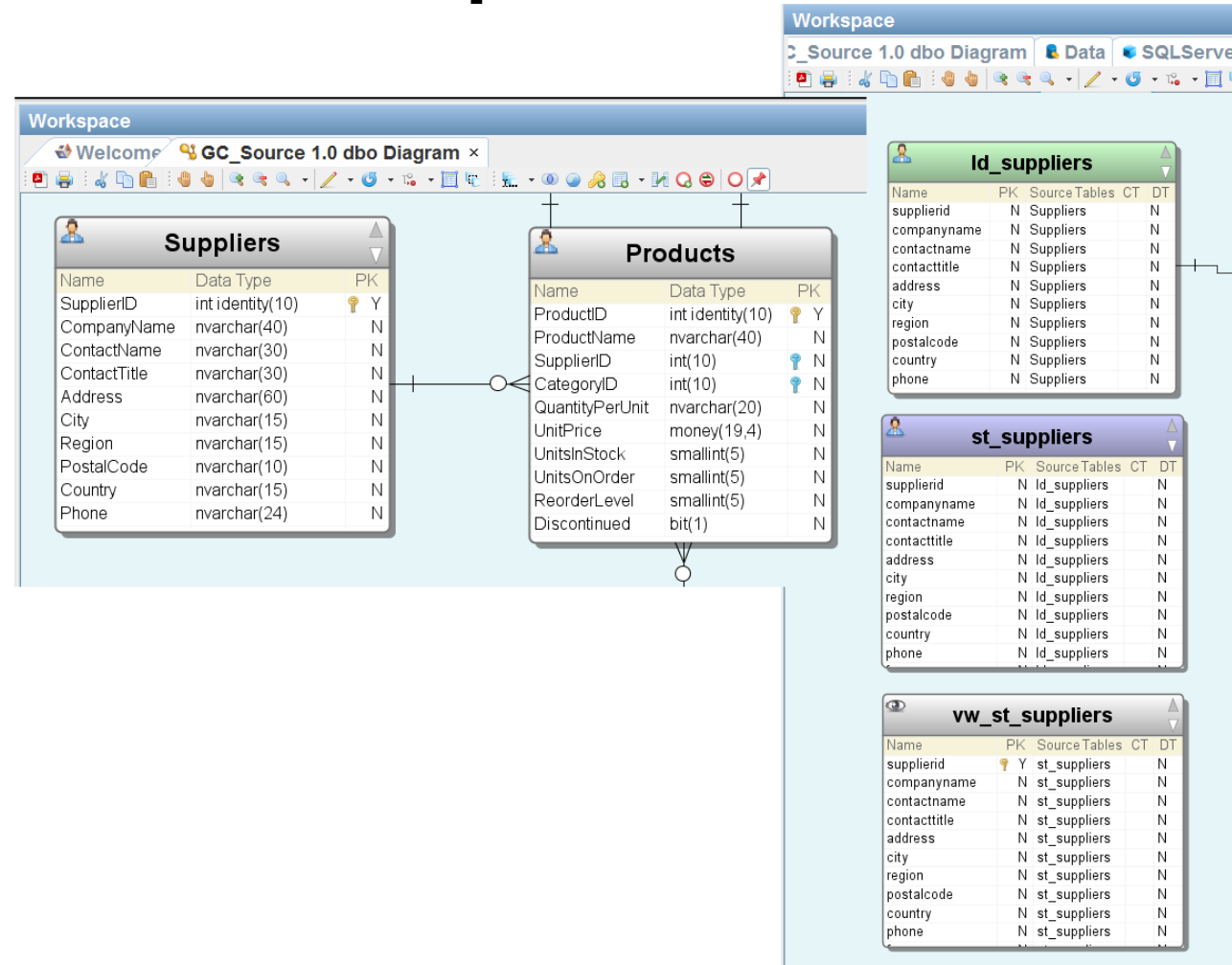
```
1 SELECT fk_count.fk_cnt, pk_count.pk_cnt FROM
2 (SELECT COUNT(*) fk_cnt FROM
3 (SELECT DISTINCT pk.[ShipperID] FROM [dbo].[Shippers] pk LEFT OUTER JOIN [dbo].[Big_Orders]
4 (SELECT COUNT(*) pk_cnt FROM [dbo].[Shippers] pk) pk_count
```

At the bottom, a small table shows the results of the query:

fk_cnt	pk_cnt
3	3

Rules Based Development

- *“Build me a Load, Stage and View object for each entity in my design”*



Build Automation

WhereScape®

No layer dead ends

LOADING

DATA TRANSFORM

DATA FOUNDATION

BUSINESS TRANSFORM

END USER LAYER

ACCESSING



Load Tables



Stage Tables

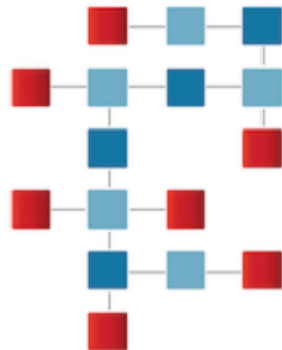


Stage Views

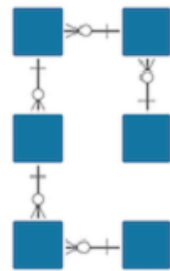
Data Store



Data Vault



3NF

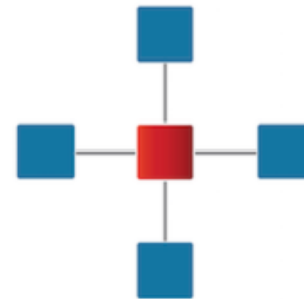


Stage Tables



Stage Views

Star Schema



Views



Performance Views/Indexes



Aggregates



OLAP Cubes

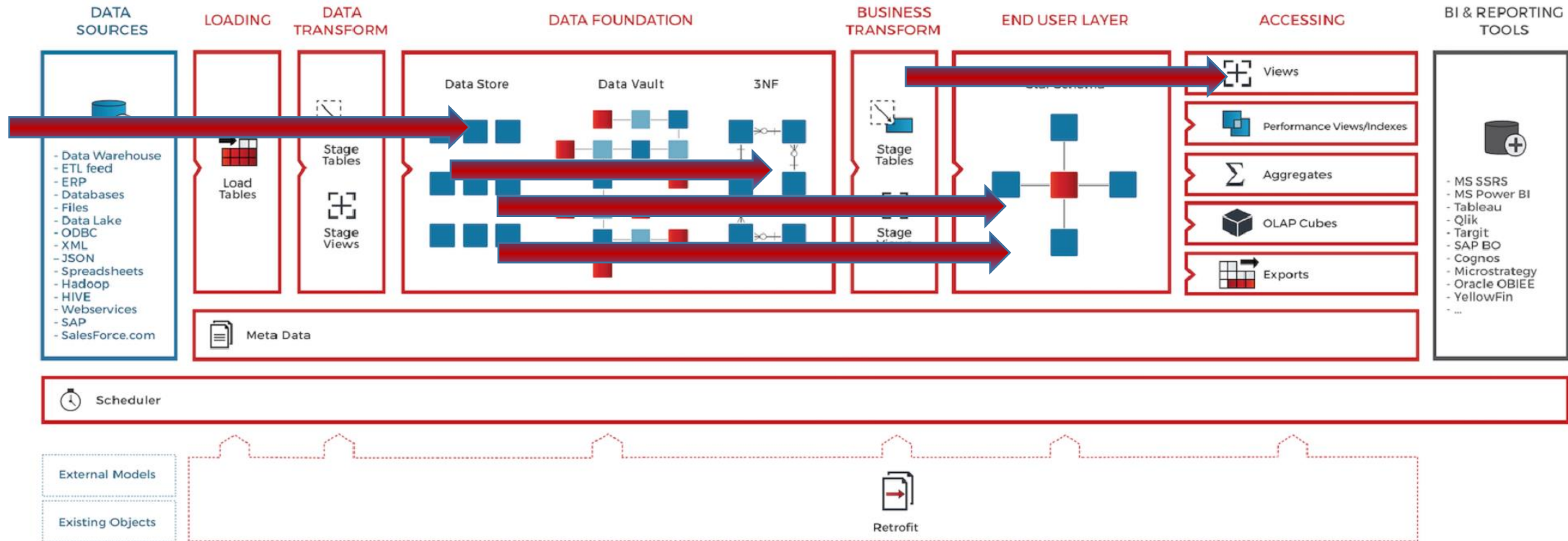


Exports



Meta Data

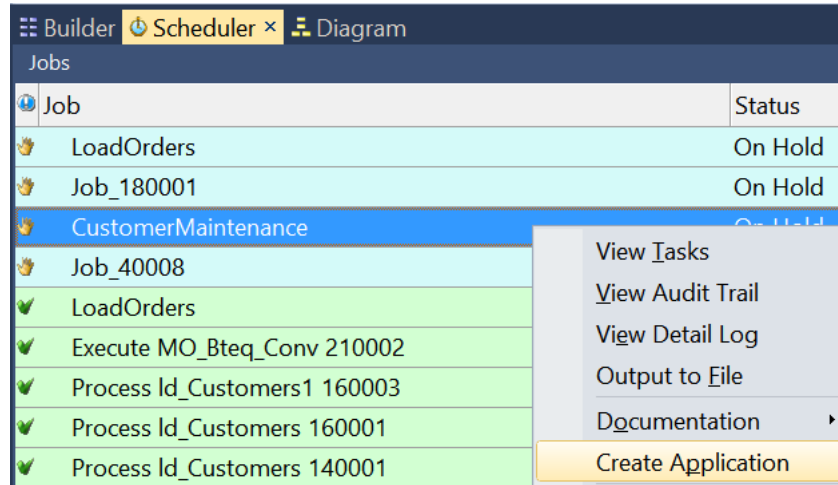
Seamless Design Integration



Deployment Automation

WhereScape®

Rock solid builds



The screenshot shows the 'Scheduler' window in the WhereScape Builder. It displays a table of jobs with columns for 'Job' and 'Status'. The 'CustomerMaintenance' job is selected, and a context menu is open over it, showing options like 'View Tasks', 'View Audit Trail', 'View Detail Log', 'Output to File', 'Documentation', and 'Create Application'.

Job	Status
LoadOrders	On Hold
Job_180001	On Hold
CustomerMaintenance	On Hold
Job_40008	On Hold
LoadOrders	Completed
Execute MO_Bteq_Conv 210002	Completed
Process Id_Customers1 160003	Completed
Process Id_Customers 160001	Completed
Process Id_Customers 140001	Completed

- **Build Application by Job**
 - All dependencies are included

Rock solid builds

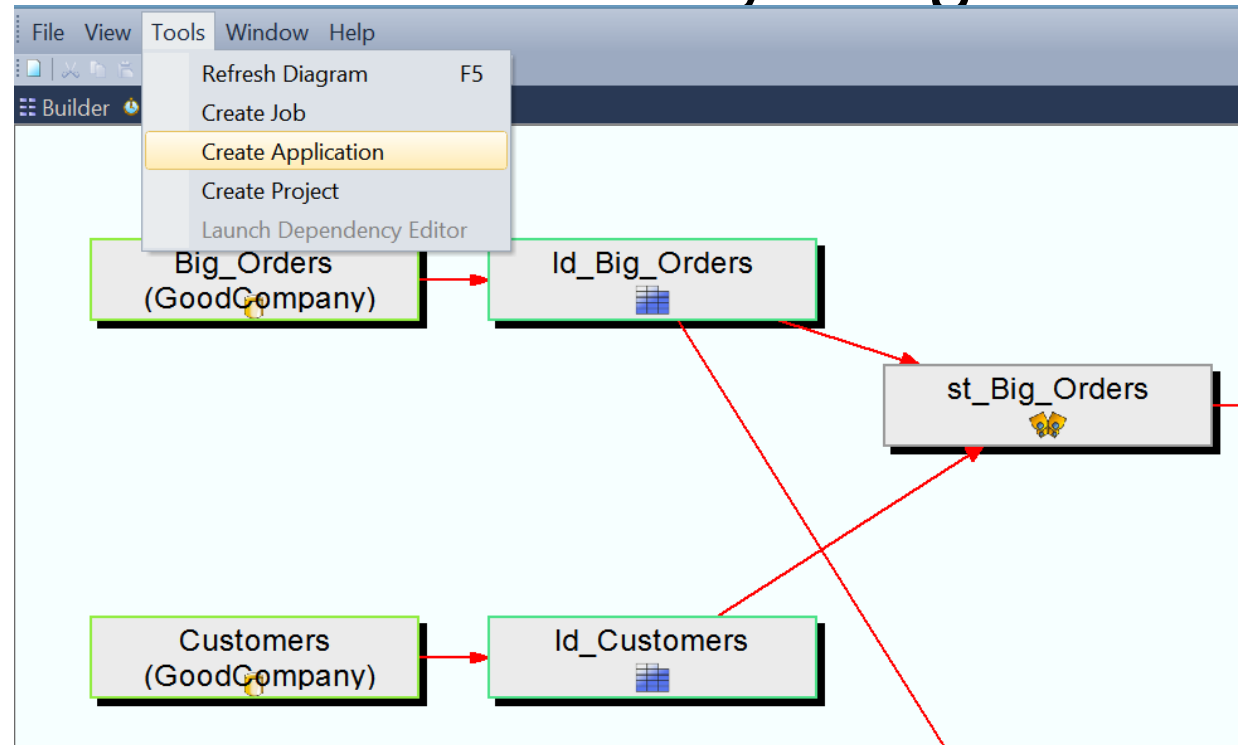
The screenshot shows the 'Jobs' list in the Builder interface. A context menu is open over the 'Process Id_Customers 140001' job, with 'Create Application' highlighted.

Job	Status
LoadOrders	On Hold
Job_180001	On Hold
CustomerMaintenance	On Hold
Job_40008	On Hold
LoadOrders	On Hold
Execute MO_Bteq_Conv 210002	On Hold
Process Id_Customers1 160003	On Hold
Process Id_Customers 160001	On Hold
Process Id_Customers 140001	On Hold

- View Tasks
- View Audit Trail
- View Detail Log
- Output to File
- Documentation
- Create Application**

Build Release by Job

Build Release by Diagram



Automation in the hands of the customer!