

Leverage Big Data

Part 2

DEMO

Migrating Optim Archive File Data
to Hadoop

Hari Kasina
Optim Consultant



Hari Kasina

Optim Consultant

Also, supporter of Green Planet
&
Fight against Global Warming

Worked at IBM for 7+ years on Optim Technology

Optim & Big Data Specialist

20+years in Information Technology Expertise

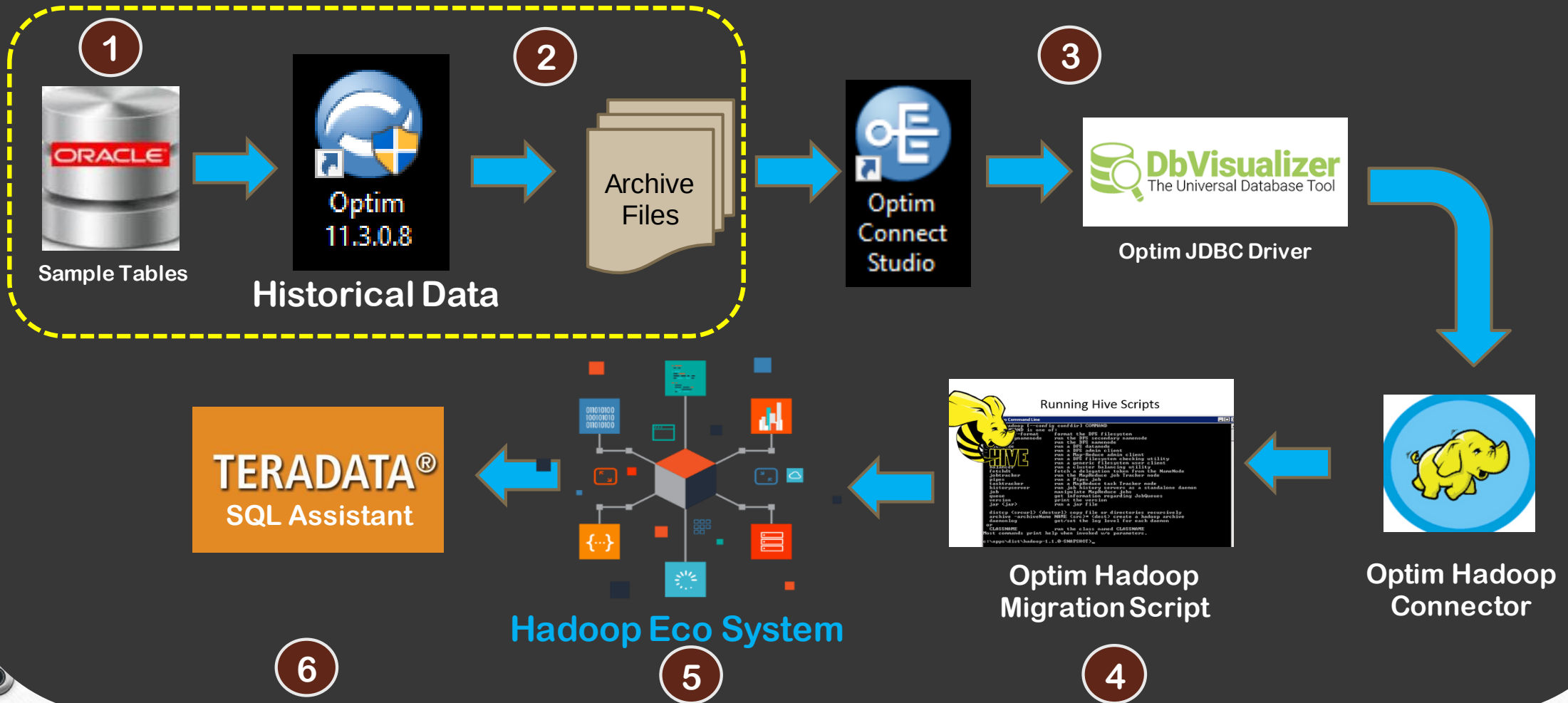
Speaker at: Oracle OAUG Conference, IBM IOD & Webinars

Working on US Design Patent for Archiving Data using Table Partitions

Recap: Summary of Part1

- History of Optim archive and ever expanding data under utilized
- Benefits of Utilizing Optim Archived Data in Hadoop
- Systems Architecture of Optim and Hadoop Eco System
- Case Study review of Hadoop Hive Reports running on migrated data

DEMO: Process and Steps



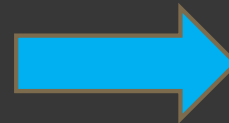
Software Tools used for DEMO

Tool	Version	Usage Notes
Oracle Database	12.2.x	Optim Sample Tables for Archive
IBM Optim	11.5.8	Archive Data out of Oracle Tables
Optim Connect Studio	11.5.8	Establish JDBC connection for Archive Files
Optim JDBC NAV Driver		JDBC Driver to connect to Archive Files
DB Visualizer	10.0.13	Query Optim Archive File Data using JDBC Connection
Apache HIVE	3.1.0.3	Hadoop HIVE tables to store migrated data
Horton Works HIVE ODBC Driver	2.06.01	Connect to HIVE Table using ODBC
Teradata SQL Assistant	16.20.0.4	Query HIVE Tables Data

DEMO Test Case

Optim to Hadoop

Optim Data	Data Count
OPTIM_SALES	110
OPTIM_ORDERS	9321
OPTIM_ITEMS	510



HIVE Table	Table Count
OPTIM_SALES_ORC	110
OPTIM_ORDERS_ORC	9321
OPTIM_ITEMS_ORC	510

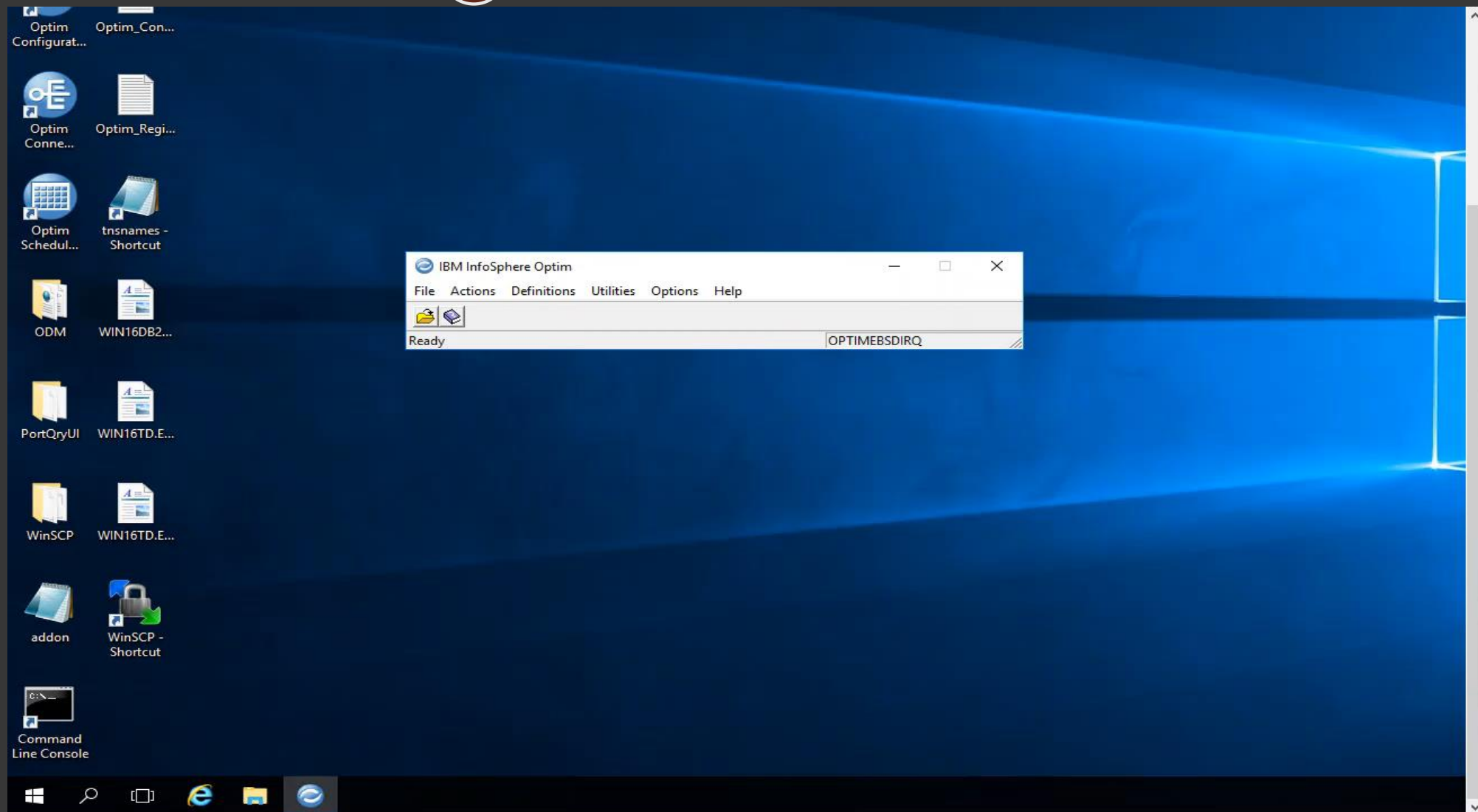
1

Data in Oracle

The screenshot displays the Oracle SQL Developer interface. On the left, the 'Connections' pane shows a tree of database objects under the 'PSTARCHDATA' connection. The 'Worksheet' pane on the right contains a SQL query that counts rows in three tables: 'OPTIM_SALES', 'OPTIM_ORDERS', and 'OPTIM_ITEMS'. Below the query, the 'Query Result' pane shows the output of the query, which is a table with two columns: 'TABLE_NAME' and 'TABLE_COUNT'. The results are as follows:

	TABLE_NAME	TABLE_COUNT
1	OPTIM_SALES	110
2	OPTIM_ORDERS	9321
3	OPTIM_ITEMS	510

2 Optim Archive



3

Optim Connect Studio

Recycle Bin
EBS_UAT4

Optim 11.3.0
notes

Optim Configurat...
Optim_Con...

Optim_Regi...

Optim Schedul...
tnsnames -
Shortcut

ODM
WIN16DB2...

PortQryUI
WIN16TD.E...

WinSCP
WIN16TD.E...

Optim Conne...

Design - CH3QW607819/bindings/OPTIM_SAMPLE/binding.bnd - Optim Connect Studio

File Edit Navigate Search Project Window Help

Configuration Metadata

- OPTIM_SAMPLE
 - Adapters
 - Data sources
 - OPTIM_SAMPLE
 - Events
- OPTIMDEMO
- OPTIMMON
- OPTIMMONS
- OPTTRACE

Properties

Property	Value
Name	OPTIM_SAMPLE
Path	\CH3QW607819\bindings\OPTIM_SA...
Read-only	true

OPTIM_SAMPLE (OPTIM_SAMPLE)

Data Source Configuration

Type: PST_GDB Binding: OPTIM_SAMPLE

Properties

Property	Value
ARCV_FILE	
ARCV_GUID	
ARCV_ID	
COLLECTION	OPTIM.SAMPLE
DB2SSN	
DIRDB	OPTIMEBSDIRQ
PLAN	
PST_AF_SUBSET	
PSTTRACE	
TRACEOUT	
TRACEPARAM	

Restore defa

Configuration Advanced

4

Load Data to Hadoop



Hadoop@EcoSystem> █

I

5 Optim Data now on Hadoop

```
Hadoop@EcoSystem> hive
```

I

ODBC Data Source Administrator (64-bit)

User DSN System DSN File DSN Drivers Tracing Connection Pooling About

User Data Sources:

Name	Platform	Driver
CData MongoDB Source	32/64-bit	CData ODBC Driver for MongoDB
dBASE Files	32-bit	Microsoft Access dBASE Driver (*.dbf, *.ndx)
Excel Files	32-bit	Microsoft Excel Driver (*.xls, *.xlsx, *.xlsm, *.xlt, *.xltm)
HiveTest	32/64-bit	Hortonworks Hive ODBC Driver
MongoLocalhost2	32/64-bit	CData ODBC Driver for MongoDB
MS Access Database	32-bit	Microsoft Access Driver (*.mdb, *.accdb)
Visio Database Samples	32-bit	Microsoft Access Driver (*.mdb, *.accdb)

An ODBC User data source stores information about how to connect to the indicated data source. User data source is only visible to you and can only be used on this computer.

OK Cancel

Hortonworks Hive ODBC Driver DSN Setup

Data Source Name: HiveTest

Description:

Hive Server Type: Hive Server 2

Service Discovery Mode: No Service Discovery

Host(s): Knox-Dev-01

Port: 8443

Database: optdb

ZooKeeper Namespace:

Authentication

Mechanism: User Name and Password

Realm:

Host FQDN: _HOST

Service Name: hive

☒ Canonicalize Principal FQDN

☐ Delegate Kerberos Credentials

User Name: et

Password:

☐ Save Password (Encrypted)

Delegation UID:

Thrift Transport: HTTP

HTTP Options SSL Options

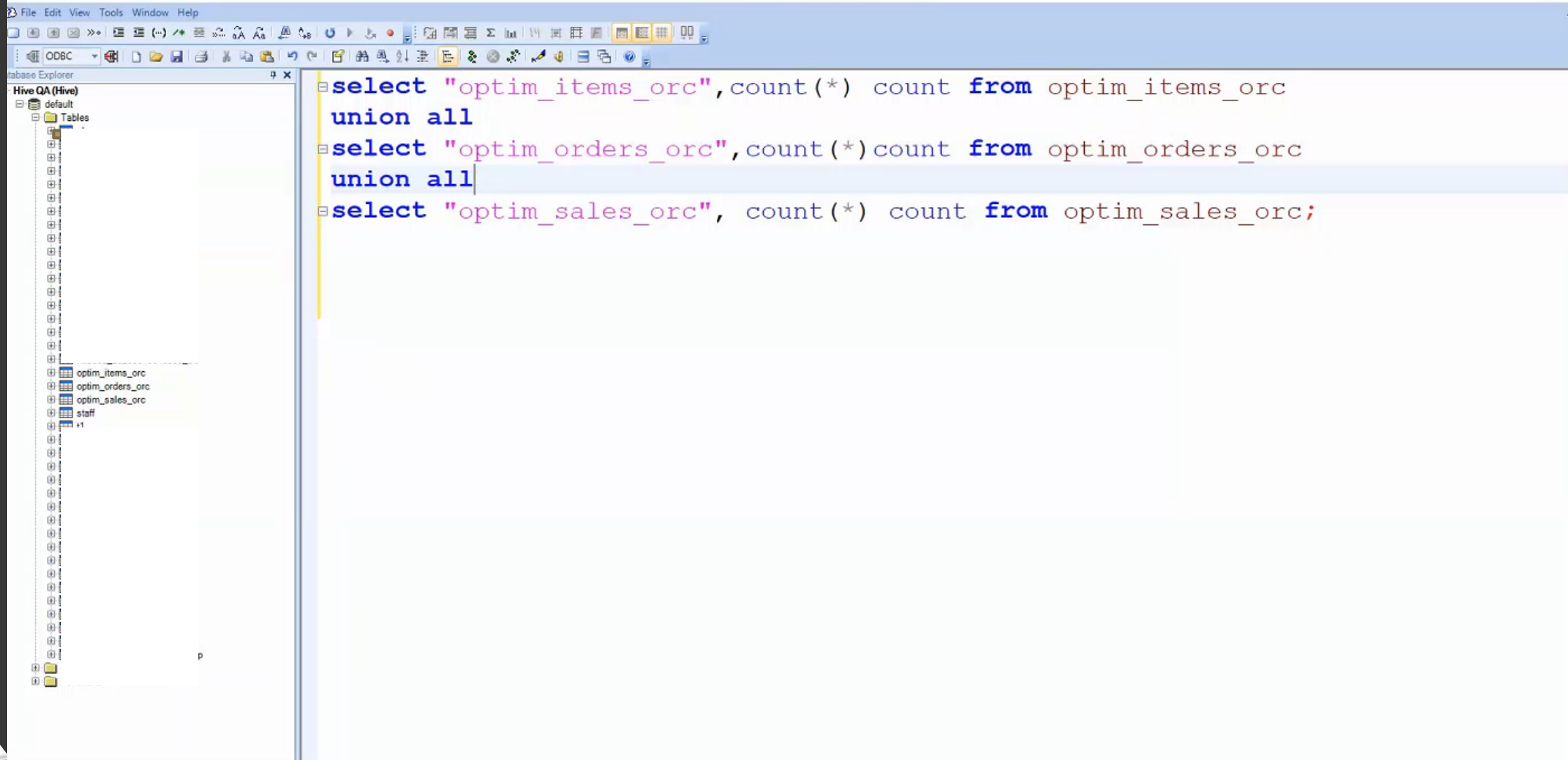
Advanced Options... Logging Options...

v2.6.1.1001 (64 bit)

Test OK Cancel

6

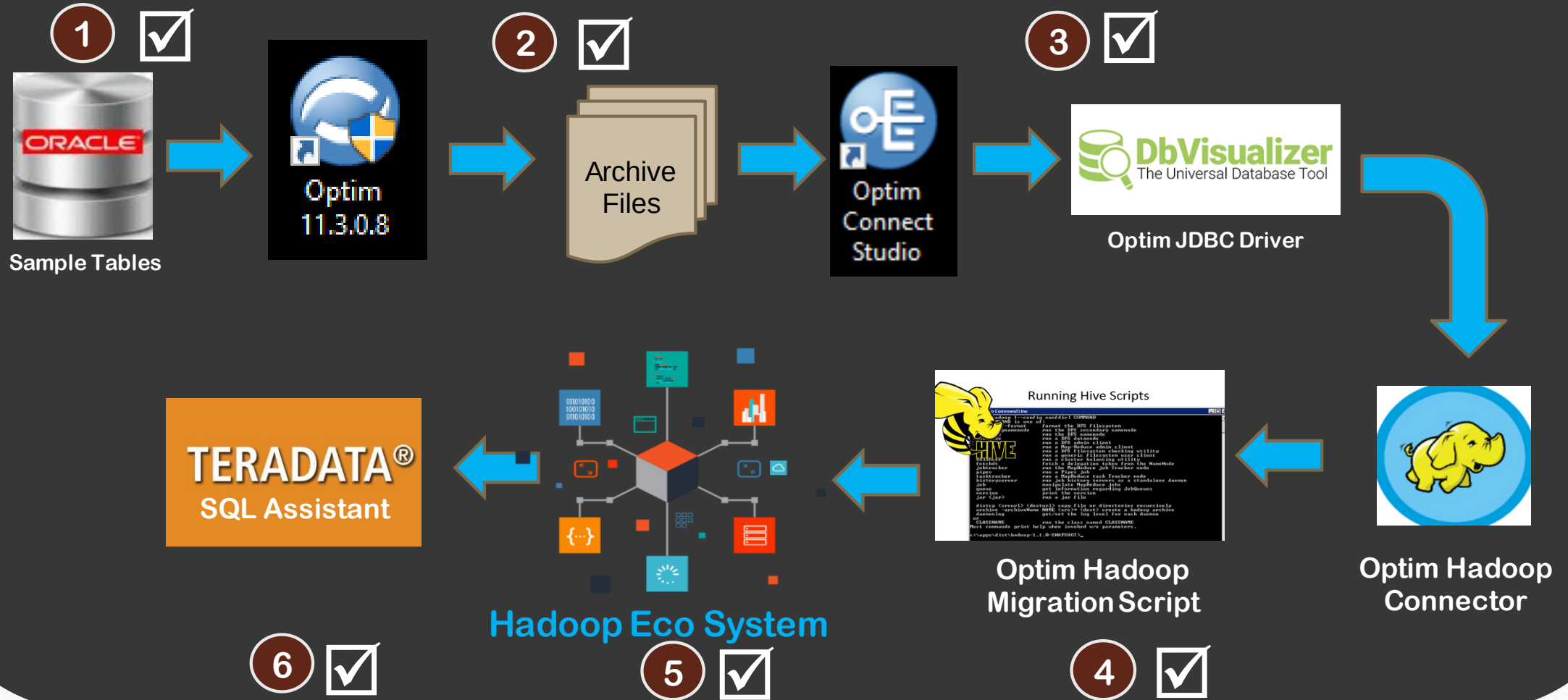
Query HIVE Data in SQL Assistant



The screenshot displays the SQL Assistant application interface. On the left, the 'Database Explorer' pane shows a connection to 'Hive QA (Hive)' with a 'default' schema. Under the 'Tables' folder, the following tables are listed: 'optim_items_orc', 'optim_orders_orc', 'optim_sales_orc', and 'staff'. The main area on the right is a SQL query editor containing the following query:

```
select "optim_items_orc", count(*) count from optim_items_orc
union all
select "optim_orders_orc", count(*) count from optim_orders_orc
union all
select "optim_sales_orc", count(*) count from optim_sales_orc;
```

Demo Summary





Q & A

Hari Kasina

Optim Consultant

hari_kasina@yahoo.com

Phone: +1-732-865-2960