

SQL Server 2012 BI, Big Data, BI Semantic Model, and SSAS

Rafal Lukawiecki
Strategic Consultant
Project Botticelli Ltd
rafal@projectbotticelli.com @rafaldotnet



Acknowledgments & Notes

This seminar is based on a number of sources including a few dozen of Microsoft-owned presentations, used with permission. Thank you: Chris Volckerick, Henrik Kraght, Phil Battersby, Peter Simpson, Arun Justus, TK Anand, Kuleen Bharadwaj, Michael Tejedor, and Diwakar Rajagopal.



The information herein is for informational purposes only and represents the opinions and views of Project Botticelli and/or Rafal Lukawiecki. The material presented is not certain and may vary based on several factors. Microsoft makes no warranties, express, implied or statutory, as to the information in this presentation.

Portions © 2012 Project Botticelli Ltd & entire material © 2012 Microsoft Corp unless noted otherwise. Some slides contain quotations from copyrighted materials by other authors, as individually attributed or as already covered by Microsoft Copyright ownerships. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Project Botticelli Ltd as of the date of this presentation. Because Project Botticelli & Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft and Project Botticelli cannot guarantee the accuracy of any information provided after the date of this presentation. Project Botticelli makes no warranties, express, implied or statutory, as to the information in this presentation. E&OE.

Resources on projectbotticelli.com

- New video training: [Introduction to SQL Server 2012 Business Intelligence](http://projectbotticelli.com/sqlbi) + offer (valid to 30 April 2012):
projectbotticelli.com/sqlbi
10% off 1 or 3 month Full Access Membership:
SWISS2012TEN
20% off 1 year Full Access Membership:
SWISS2012TWENTY
- PPTs (free registration required):
projectbotticelli.com/ppt
- Articles for This Seminar (free):
projectbotticelli.com/knowledge/microsoft-sql-server-2012-business-intelligence

projectbotticelli.com/knowledge/apache-hadoop-big-data-microsoft



Objectives

- Introduce Top 10 BI Features of SQL Server 2012
- Introduce Big Data
- Introduce BI Semantic Model
- Show demos of Tabular Models



Ten Key SQL 2012 BI Innovations

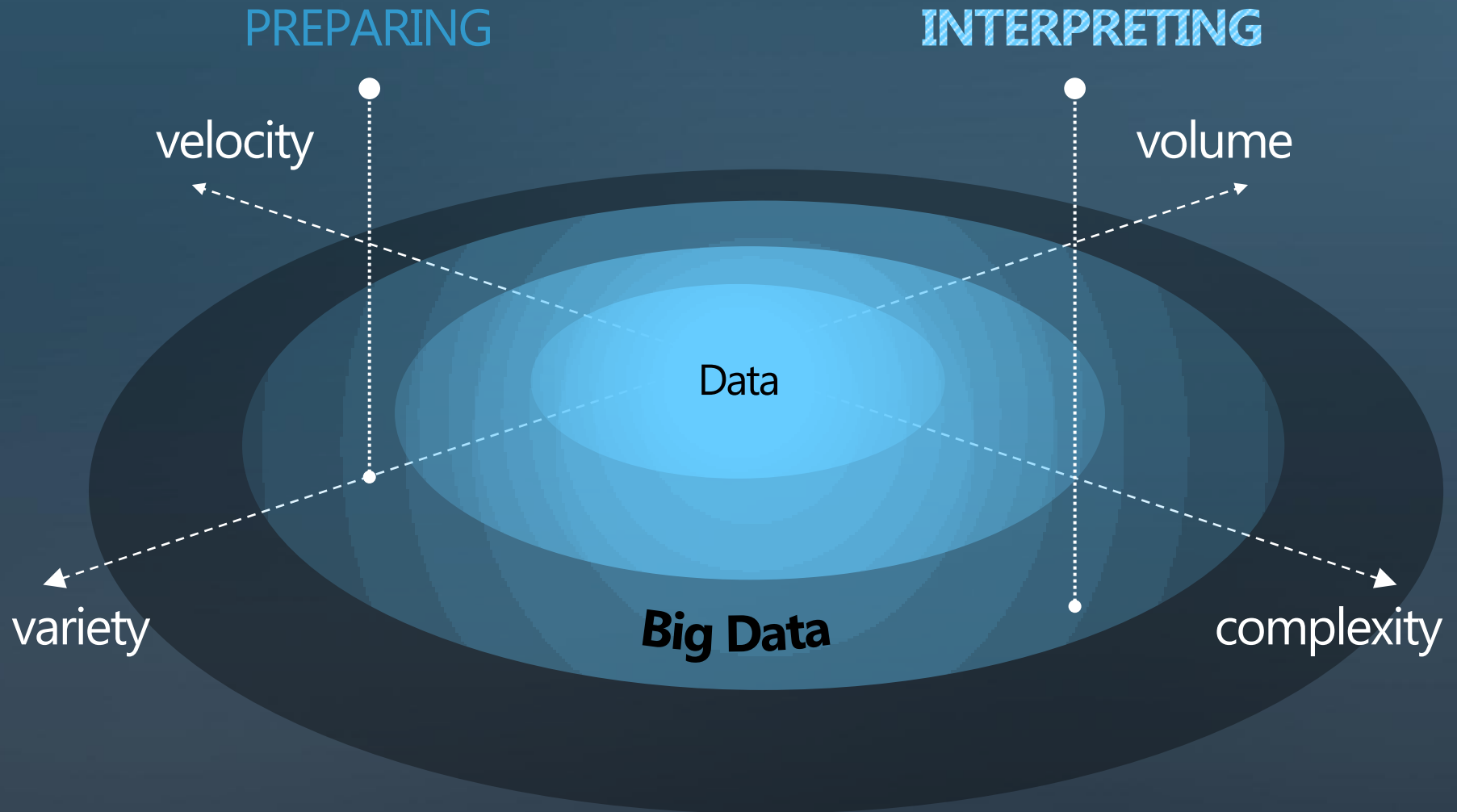
- BI Semantic Model
- Analysis Services
Tabular Mode
- PowerPivot "2"
- Power View (SSRS)
- Self-Service Data Alerts
(SSRS)
- Hadoop Big Data
Integration
- xVelocity–Columnstore
Indexes
- Geospatial Indexes
- Unstructured Data
Queries
- Data Quality Services
 - Updated Master Data
Services

xVelocity for DW: Columnstore Indexes

- Innovation that drives xVelocity for SSAS (aka VertiPaq in PowerPivot) but for SQL tables
- Data grouped by column, not row
 - Makes sense when data in a column is highly repetitive
 - E.g. some dimension tables
- Ultra-fast queries
- Read-only in SQL Server 2012
- CREATE COLUMNSTORE INDEX
- [msdn.microsoft.com/en-us/library/gg492088\(v=SQL.110\).aspx](http://msdn.microsoft.com/en-us/library/gg492088(v=SQL.110).aspx)

Interpreting Big Data

Requires New Approaches: Hadoop

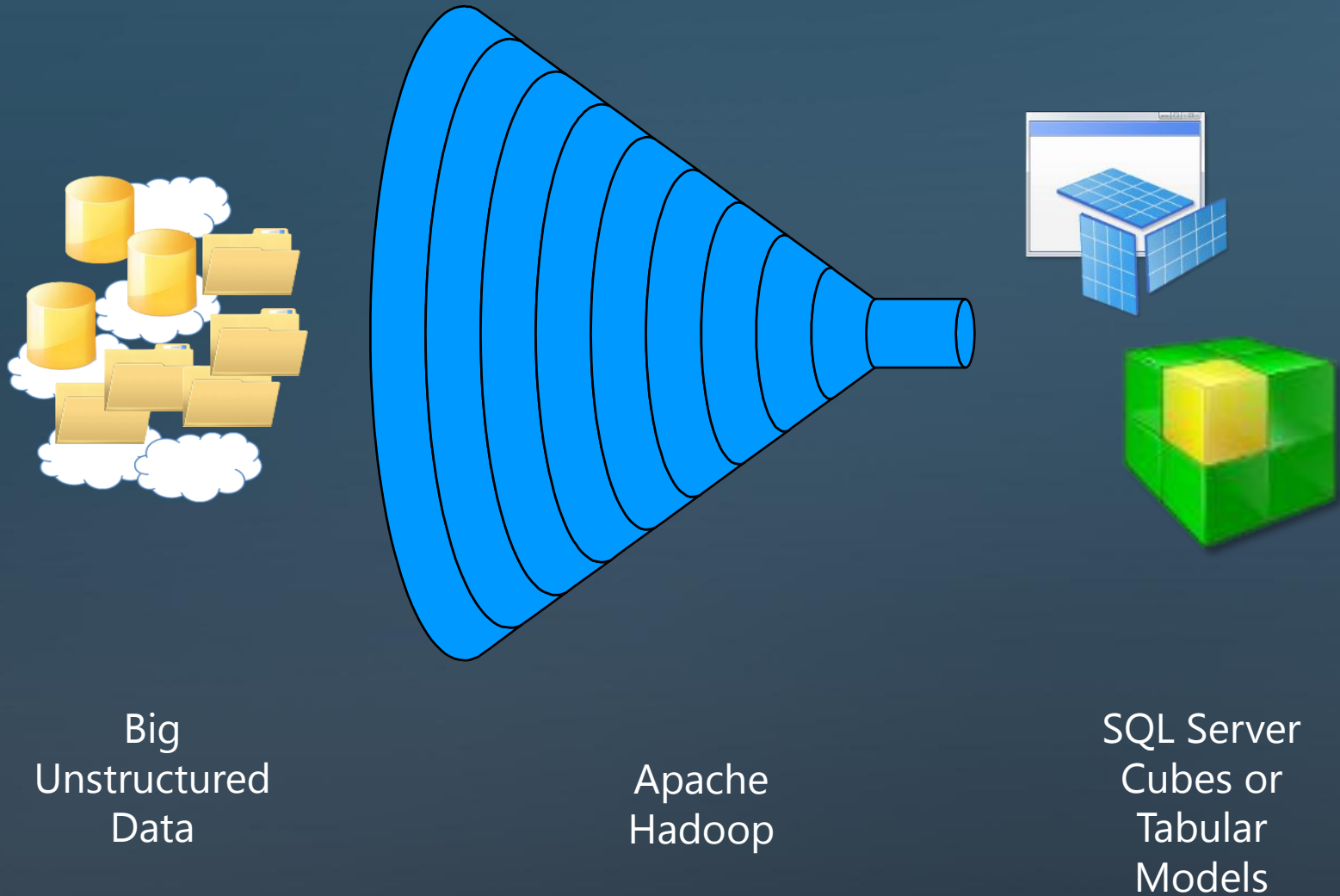


Hadoop Principles

Powers analytics for Microsoft, Yahoo, Google, Facebook

- Practical method for massive data processing
parallelisation
- Data stored **close** to processing power
 - Overcomes network limitations
- **MapReduce** is a plain-old **Divide-and-Conquer**
 - **Map** your problem into smaller (divide)
 - **Reduce** results into higher-level aggregates (conquer)
 - **Your** processing logic is split between the map and reduce functions

Key Scenario for Hadoop and Microsoft



Hadoop — Nodes in Cluster

- Distributed **data**
- Distributed **processing**



Yahoo! Hadoop Cluster, Source: developer.yahoo.com, 2007

Hadoop — Nodes in a Cluster

- Distributed data
- Distributed processing



“Buster Cluster”, an early research project by Miles Osborne, University of Edinburgh, School of Informatics. Picture used with permission. <http://homepages.inf.ed.ac.uk/miles/>

Hadoop — Main Components

- Data

- Stored in **HDFS** (Hadoop File System)
 - Which is even network rack location aware to minimise transfers
- Queried like a warehouse (Hadoop **Hive**)

- Processing

- **MapReduce**: code *your* data processing abstraction at:
 - *lower-level* in Java or anything that reads/writes stdin/stdout
 - *higher-level* in **Pig Latin**
 - Extended with your User Defined Functions (**UDF**) in Java, JavaScript, or Python

Hadoop and Microsoft

Broadening Access to Hadoop

- Loading data in-out of HDFS with **SQL Server Connector for Apache Hadoop**
 - Uses Hadoop **Sqoop**, works with SQL 2008 R2 and 2012
- **ODBC connector** for querying Hive
- **Excel Add-in** for querying Hive
- **JavaScript* support** for MapReduce and Pig, .NET soon
 - JS console to Hadoop
- Azure and Windows **Hadoop distributions**
 - Enables Windows apps in MapReduce jobs, e.g. StreamInsight event processing
- **Enterprise features**
 - Active Directory security, System Center management

Hadoop on Azure

<http://www.hadooponazure.com> (CTP)

The screenshot displays the Windows Azure Platform management console. The top navigation bar includes links for English, Billing, and Sign Out. Below this is a toolbar with various actions like New Hosted Service, New Production Deployment, New Staging Deployment, Upgrade, Configure, Delete, Start, Stop, Swap VIP, Configure OS, Reboot, Reimage, Enable, Configure, and Connect.

The main content area is divided into three sections: a left-hand navigation pane, a central table of hosted services, and a right-hand properties pane.

Left-hand navigation pane:

- Deployment Health
- Affinity Groups
- Management Certificates
- Hosted Services (5)**
- Storage Accounts (5)
- User Management
- VM Images
- CDN
- Home
- Hosted Services, Storage Accounts & CDN
- Database
- Reporting
- Service Bus, Access Control & Caching
- Virtual Network

Central table of hosted services:

Name	Type	Status	Environment
Project Ubik	Subscription	Active	
BlackOps	Hosted Service	Created	
Isotope Demo Cluster	Hosted Service	Created	
Isotope	Hosted Service	Created	
Certificates			
IsotopeAzureKit - 9/8/2011 12:37:10 AM	Deployment	Ready	Production
IsotopeWorkerNode	Role	Ready	Production
IsotopeWorkerNode_IN_0	Instance	Ready	Production
IsotopeWorkerNode_IN_1	Instance	Ready	Production
IsotopeWorkerNode_IN_2	Instance	Ready	Production
IsotopeWorkerNode_IN_3	Instance	Ready	Production
IsotopeWorkerNode_IN_4	Instance	Ready	Production
IsotopeWorkerNode_IN_5	Instance	Ready	Production
IsotopeWorkerNode_IN_6	Instance	Ready	Production
IsotopeWorkerNode_IN_7	Instance	Ready	Production
IsotopeHeadNode	Role	Ready	Production
IsotopeHeadNode_IN_0	Instance	Ready	Production
UbikDemo	Hosted Service	Created	
Windows Azure Sandbox 8-238	Subscription	Active	
isotope sandbox	Hosted Service	Created	
Win and SQL Azure -- Denny Lee	Subscription	Active	

Right-hand properties pane (selected instance: IsotopeHeadNode_IN_0):

- Environment:** Production
- Name:** IsotopeHeadNode_IN_0
- Start result:** Started
- Start time:** 9/8/2011 8:26:06 AM UTC
- Size:** Extra Large
- Status:** Ready
- Update domain:** 0
- Type:** Instance

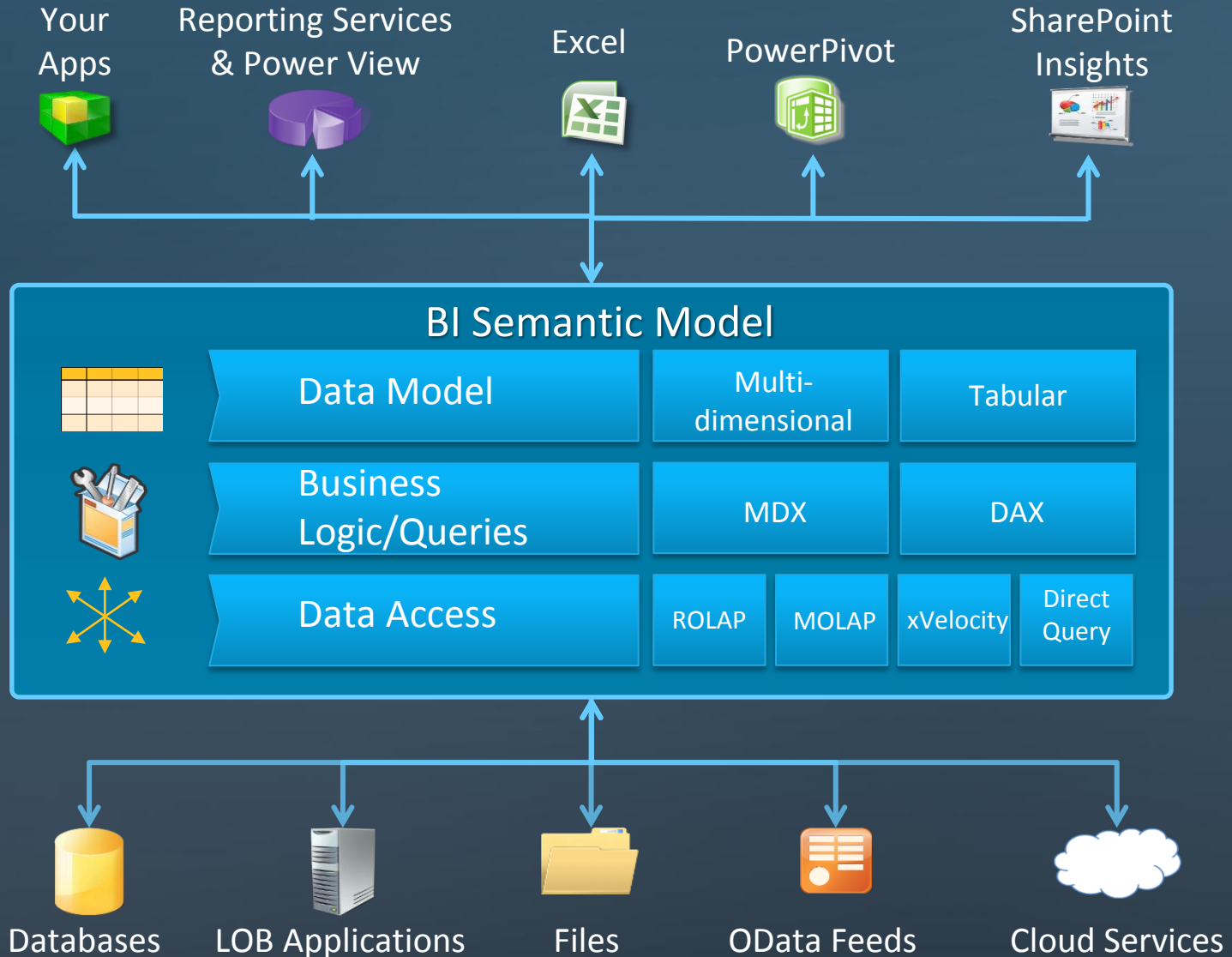
At the bottom of the console, there is a status bar that reads "Done. 0s to next refresh." and a footer with copyright information: "© 2011 Microsoft Corporation Privacy Statement Terms of Use Help and Support Feedback".

BI Semantic Model

- One Model for All User Experiences
- Visualise analysis using your favourite tools
- Model data the way you like
- Store analytical data however it is best done
- BISM is a *concept*, not a product
 - Can be hosted in PowerPivot or SSAS

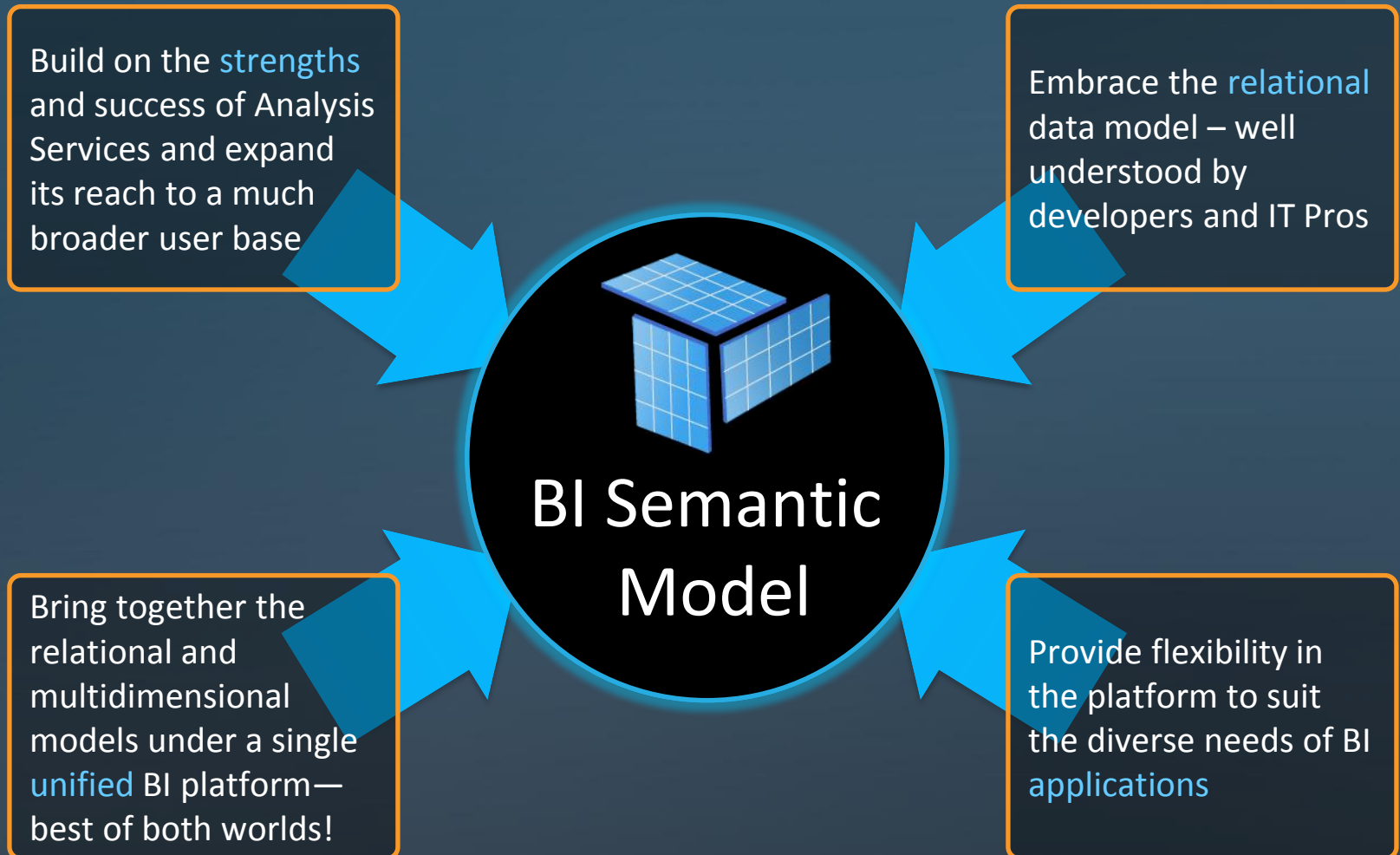
BI Semantic Model

One Model for All User Experiences



Evolving Analysis Services

What Were Microsoft's Goals?



Tabular Mode

Scaling PowerPivot to Enterprise Needs

- Model in PowerPivot or SSDT
 - PowerPivot as source of SSAS Tabular Models
 - Excel for browsing and testing in SSDT
- All new PowerPivot features:
 - Diagrams, Measure Grid, KPIs, Hierarchies, Perspectives, 30+ New DAX Functions
- ...and, unique to SSAS Tabular Mode:
 - Row-level Security, Partitions, Large Tables (>2 billion rows), Images, Memory Paging

BI Semantic Model

What about existing Analysis Services applications?



Existing
applications

Based on Unified
Dimensional Model

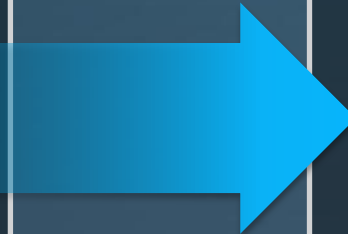


Microsoft®
SQL Server® 2008 R2



Microsoft®
SQL Server® 2008

Microsoft®
SQL Server™ 2005



Existing
applications

Every UDM becomes a BI
Semantic Model



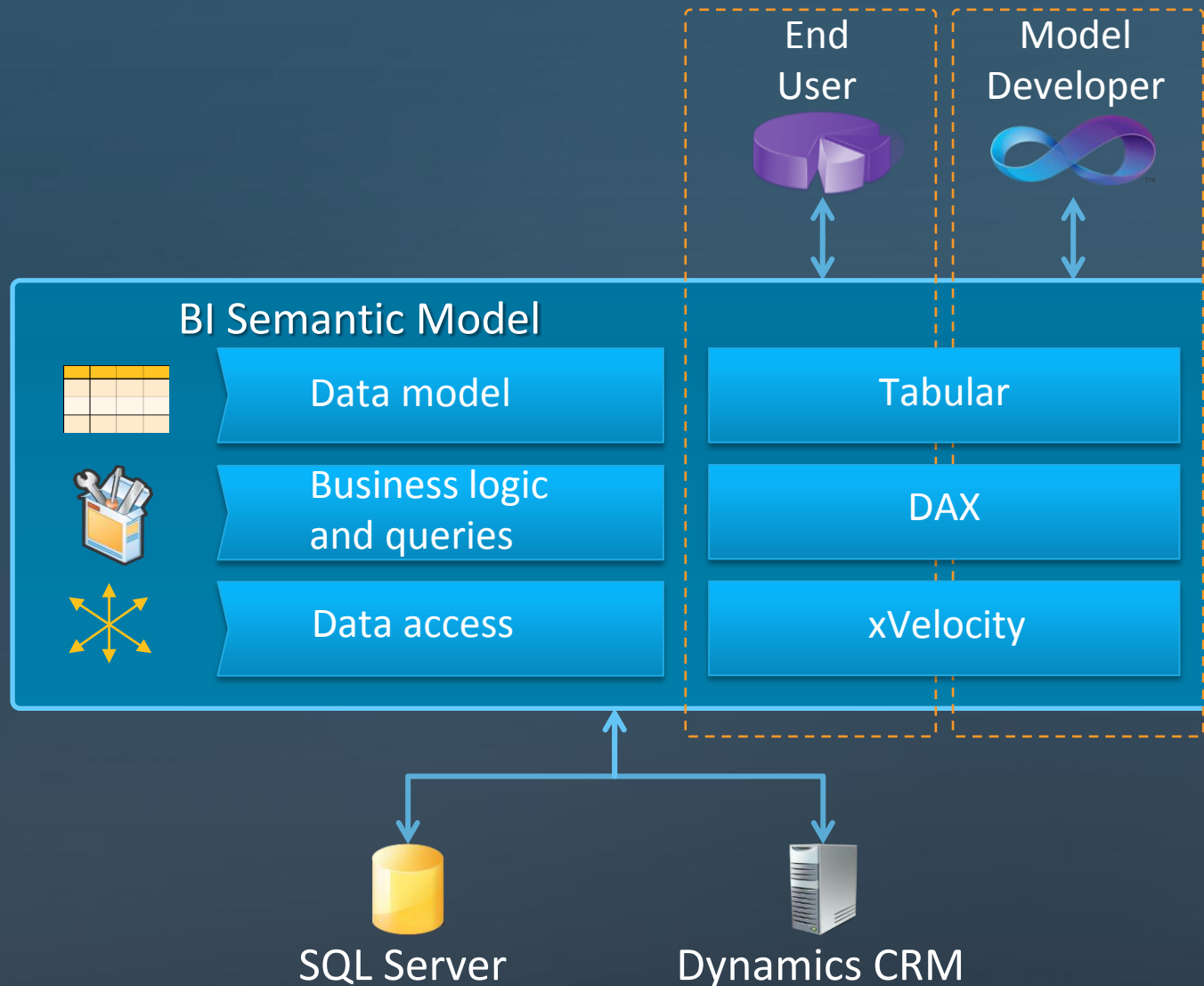
New
applications

New technology
options

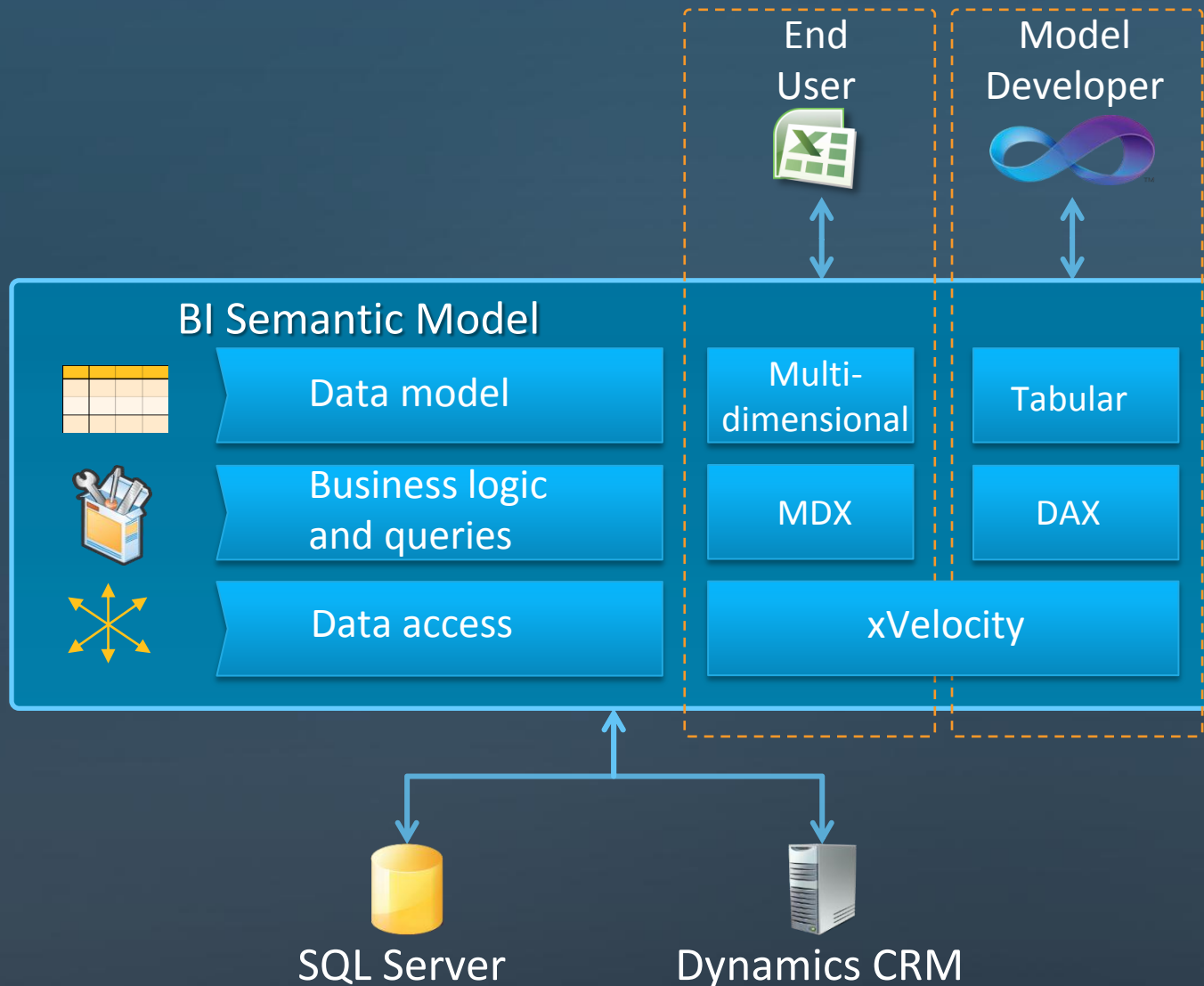


Microsoft®
SQL Server® 2012

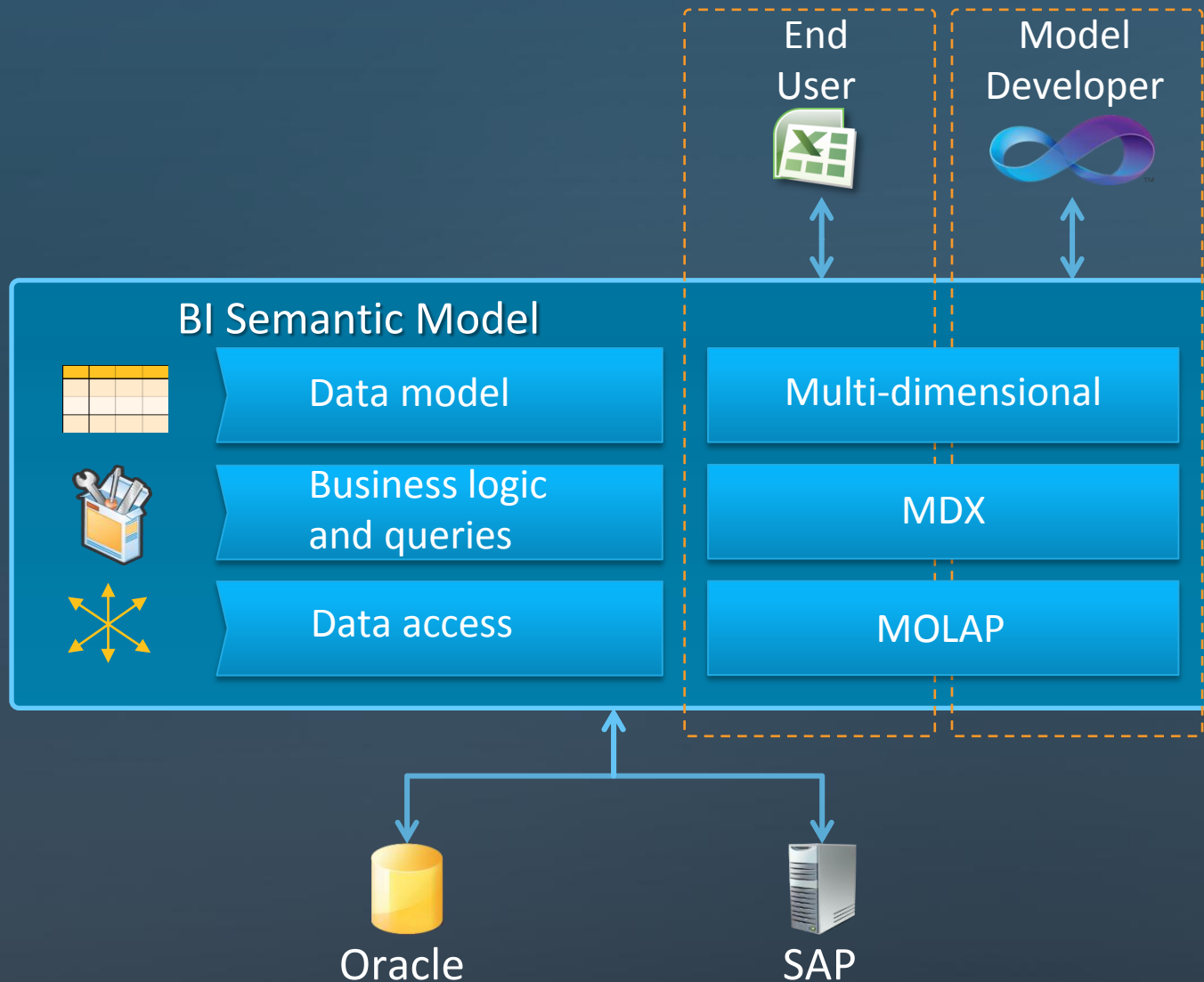
Example: Power View Over a Sales Model



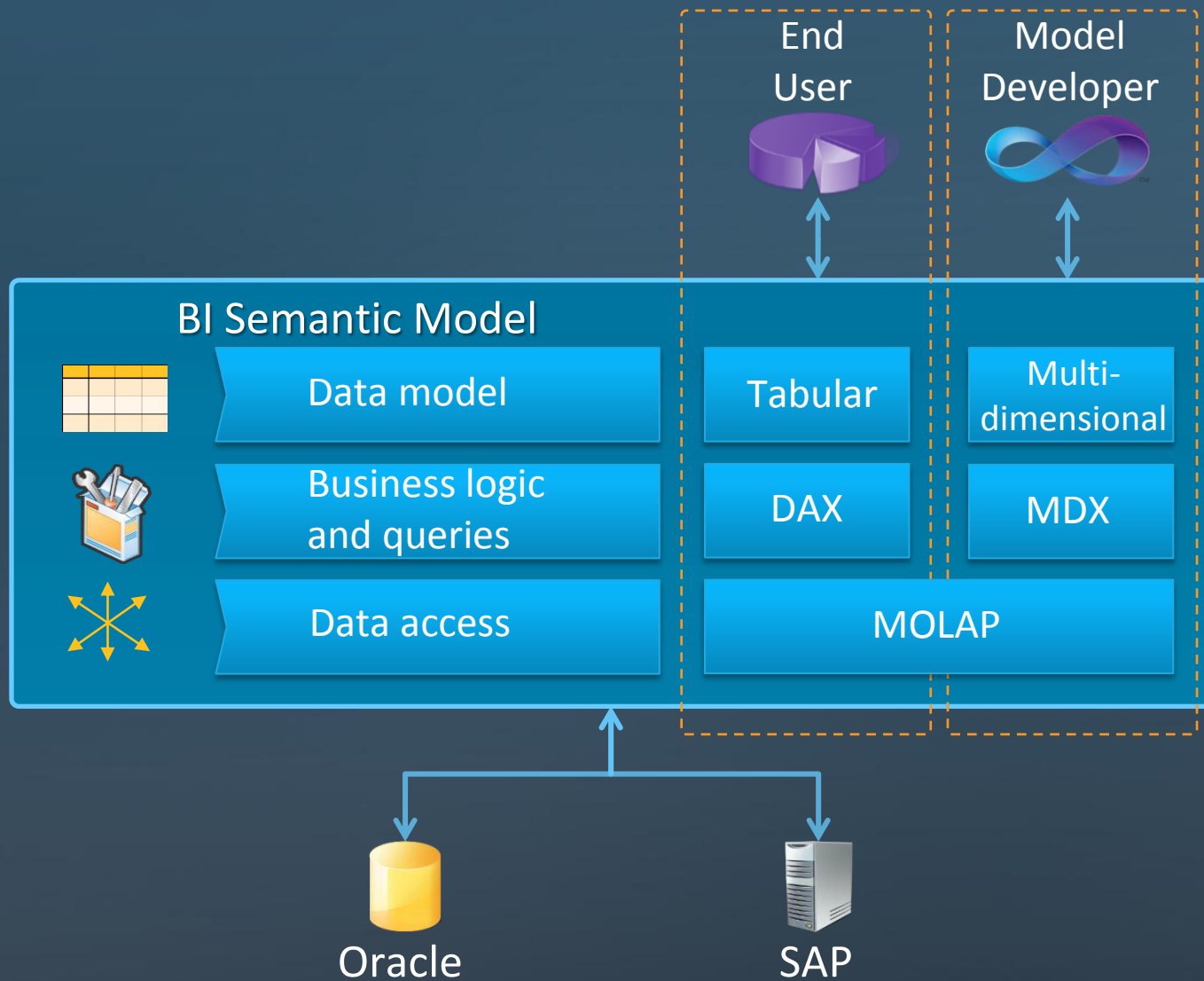
Example: Excel Over a Sales Model



Example: Excel Over a Finance Model



Example: Power View Over a Finance Model



Row-Level Security

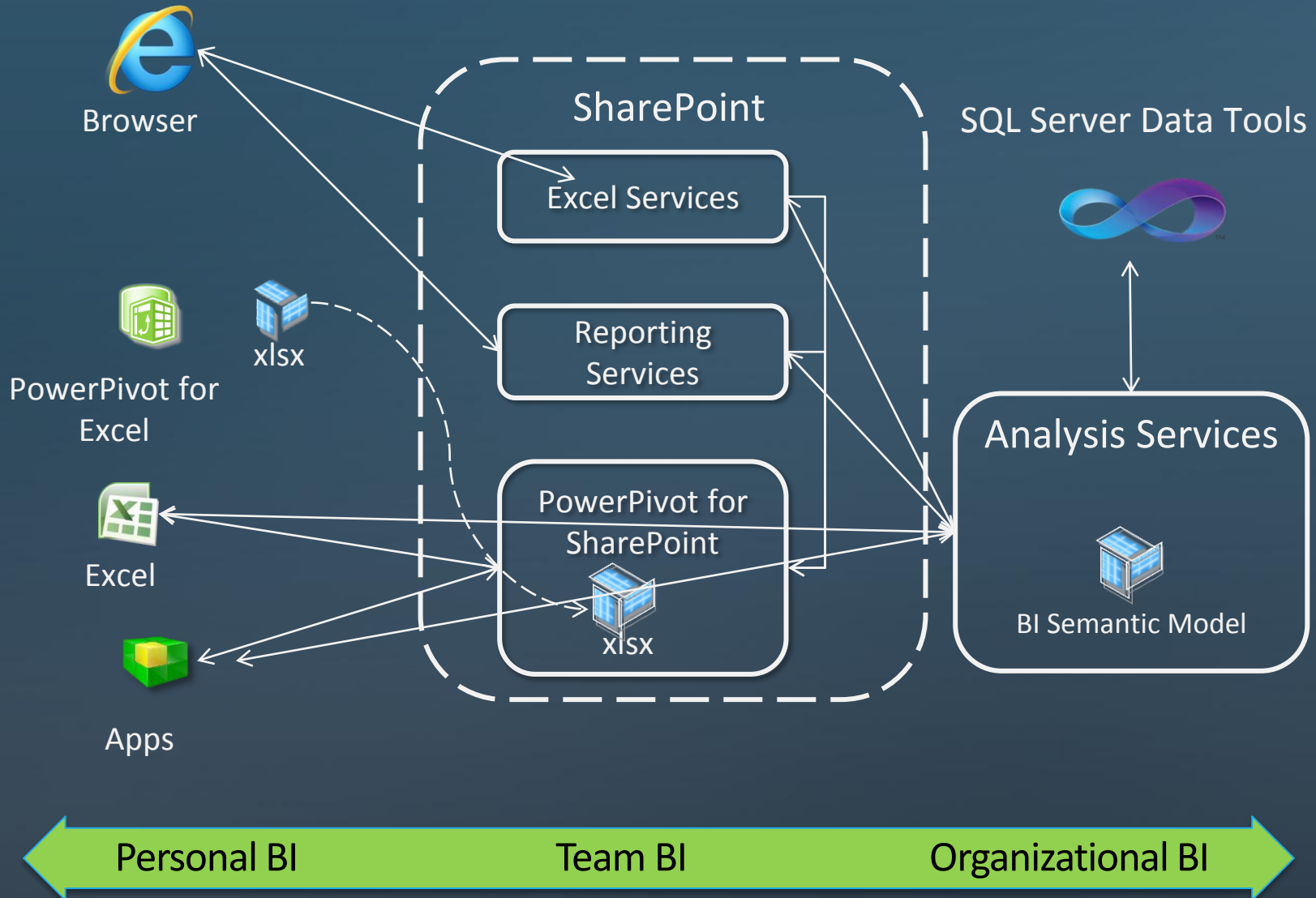
Much demanded enterprise feature

- Granular security:
 - Create named roles
 - Assign membership in terms of AD groups etc
 - Assign row filters using Boolean DAX formulas
- Or file-based security:
 - SharePoint and PowerPivot role-based security for all-or-nothing access

Tour of SQL Server 2012 BI Semantic Model and Analysis Services

demo

Analysis Services Architecture



Summary

- SQL Server 2012 is all about BI
- Big Data with SQL Analytics: powerful insight at a fraction of the traditional cost
- BI Semantic Model: one model for all analytical user experiences

Q&A

Contact: rafal@projectbotticelli.com, @RAFALDOTNET

Register on projectbotticelli.com for our newsletter and online training

PPTs: projectbotticelli.com/ppt

Watch: projectbotticelli.com/video

Study: projectbotticelli.com/sqlbi

Read: projectbotticelli.com/knowledge/microsoft-sql-server-2012-business-intelligence

Thank You



Microsoft®