



SENTURUS

IS OLAP DEAD?



Is OLAP Dead?

Quotes on the state of OLAP Cubes

- *”In Memory databases are killing OLAP”*
- *“The OLAP Cube is history”*
- *“Is OLAP still Relevant?”*
- *“Building Cubes for Tableau can be a waste of time!”*

Today's Agenda

- OLAP Defined
- Why OLAP Pros/Cons
- Current state of OLAP architectures
- New Visualization Tools
- Fast Columnar/In-Memory databases
- Big OLAP on Big Data

OLAP, OLTP, DWH Defined

- OLTP - On Line Transaction Processing
 - Modern ERP systems
 - Core data repository of data flowing into business systems
 - Optimized for quick data entry and relational integrity
 - Not Optimized for reporting and data analysis
 - Typically very complex schema design with many normalized tables that facilitate high volume throughput of transactions

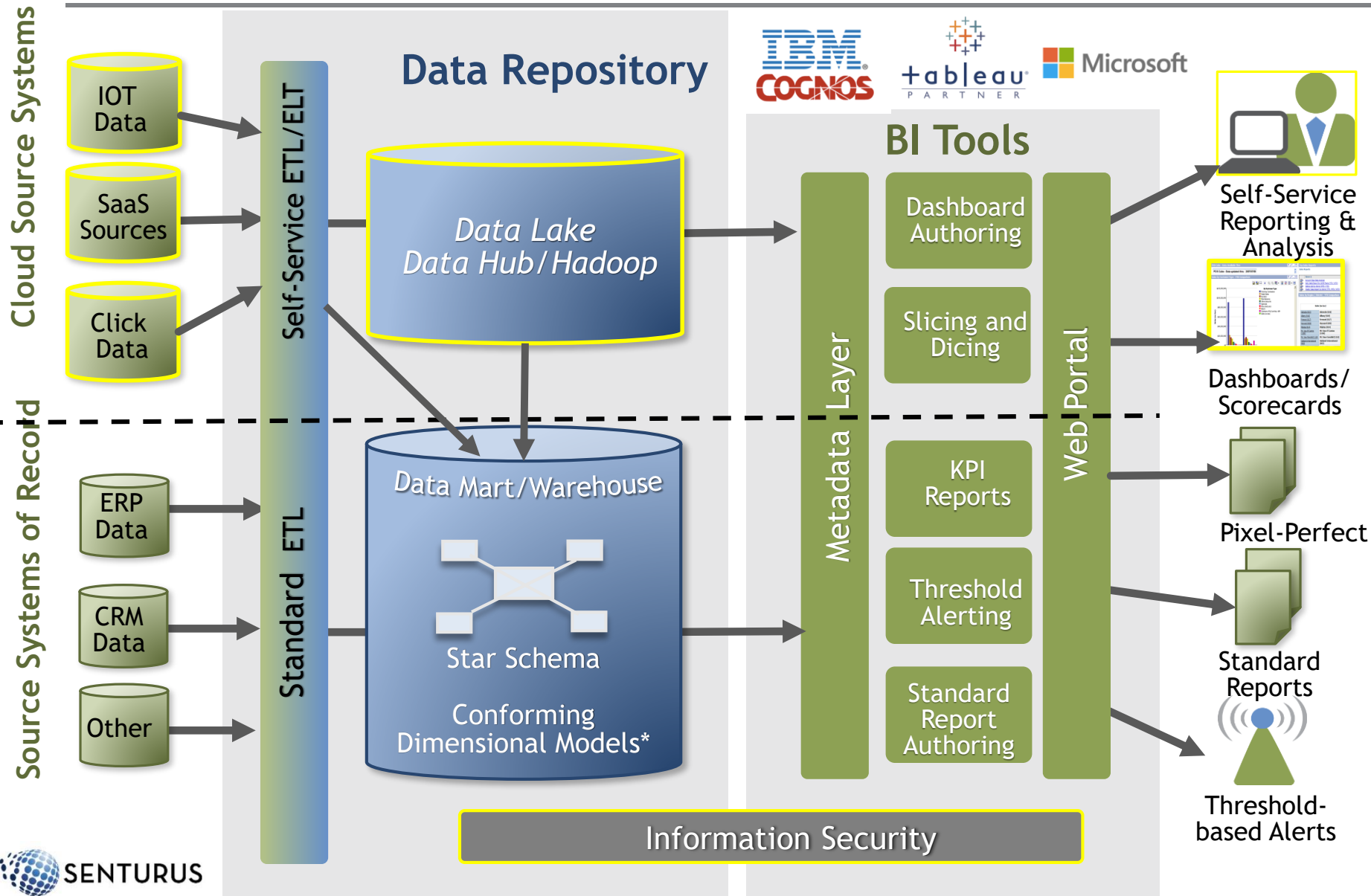
OLAP, OLTP, DWH Defined

- Data Warehouse
 - Central repository of multiple ERP systems or other data sources
 - The Data Warehouse (DWH) is a database system separate from the OLTP
 - Generally stored in a relational database
 - Architected in an optimized fashion for easy reporting and analysis
 - Generally dimensionally modeled
 - Organizes and hides the complexity of the OLTP for efficient, timely and accurate reporting

OLAP, OLTP, DWH Defined

- OLAP - Online Analytical Processing
 - Edgar F. Codd ‘Father of Relational Database’ coined the term OLAP. Arbor/Essbase went on to market the term
 - The data warehouse plus a **central repository** that defines the relationships between tables (facts/dimensions) and complex business rules/calculations (e.g. YTD, YTD LY, Margins, % etc.)
 - Allows for **high performing** interactive analysis
 - Generally referred to as ‘Cubes’

MODERN BI ARCHITECTURE



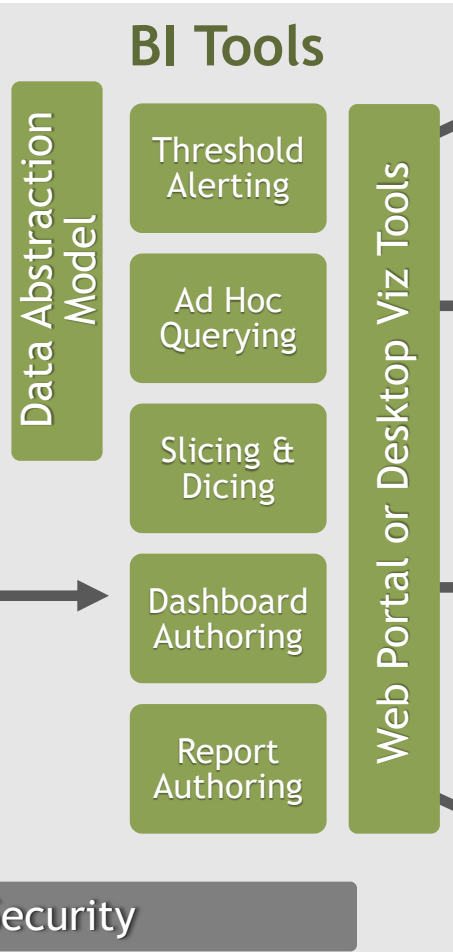
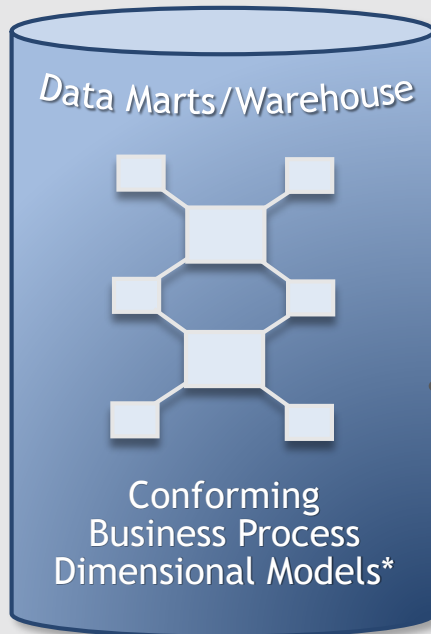
Typical Best Practices BI System with OLAP Layer

Architected Data Warehouse & BI Tools



Properly Staged Data

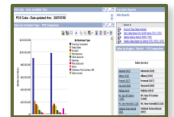
Single Version of the Truth



Standard Reports



Self-Service Reporting & Analysis



Dashboards/Scorecards



Threshold-based Alerts

Why OLAP

- Historically, the size and speed limitations of databases limited query performance
- **Central Repository** for relationships and complex business calculations
- Buffers the business user from complex native database structures and sensitive calculation logic
- Cubes generally have higher performance vs. relational queries
- Fast, simple, drag-and-drop ad-hoc analysis and reporting
- Visual Exploration
 - Multi-dimensional view of data
 - Drill-Down on hierarchies
- Complex Calculations are stored in the cube so that complex SQL statements are avoided
- **Many business users love the interface and are used to querying by governed data dimensions and measures that are prebuilt for them**

OLAP familiar Interfaces

Pre-Defined Calculations

- Selected Period
- Pr Yr
- Pr Yr Chg
- Pr Yr Chg %
- MTD
- MTD Pr Yr
- MTD Chg
- MTD Chg %
- YTD
- YTD Pr Yr
- YTD Chg
- YTD Chg %

Drill-Down

Visual display of members & hierarchies Drag & Drop Interfaces

Internet Sales Amount	2006	2007	2008	2009
Accessories			\$667,015.32	\$330,000.00
Bikes	\$7,060,199.57	\$5,762,134.30	\$15,483,926.11	\$11,000,000.00
Clothing			\$322,676.62	\$150,000.00
Components				\$100,000.00

Excel on SSAS Tabular

EOP Inventory Units	FY11	FY12	FY13	FY14	FY15	FY16	FY17	FY18	Grand Total
Wholesale	2,174,209	2,806,489	2,608,486	2,738,406	2,599,614	3,603,290	4,192,066	4,592,036	4,592,036
Retail	820,223	965,405	1,080,666	864,650	949,405	1,095,104	1,339,655	1,403,934	1,403,934
eCommerce	322,109	415,434	396,300	374,053	414,902	612,639	593,121	1,126,475	1,126,475
Intercompany	141,379	130,313	89,941	111,039	152,855	118,562	118,599	167,842	167,842
Missing					47			792,593	792,593
Grand Total	3,457,920	4,317,641	4,255,393	4,088,148	4,116,821	5,429,595	6,241,441	8,082,880	8,082,880

OLAP Limitations

- Massive increase in data volumes
 - Latency - Large Cubes increase cube build times thus impacting SLAs
 - Large cardinality dimensions and many dimensions
 - Real-time updates are difficult if not impossible
- Movement of data into another proprietary structure
- Upfront investment in cube modeling
 - Measures, dimensions, hierarchies all defined upfront
 - Not a flexible agile BI environment
 - New cube builds and designs required as business changes
- Continued developer maintenance and administration
- Today CPU power, memory and powerful servers are very affordable - do we still need the OLAP layer



SENTURUS

THE CURRENT STATE OF OLAP ARCHITECTURE

Traditional OLAP

TRADITIONAL OLAP ARCHITECTURES

MOLAP - Multi Dimension OLAP

- Most traditional OLAP design
- Data is stored in the multidimensional cube
- Data is moved from the relational database to the cube
- Data is pre-aggregated and allows for very fast analysis

ROLAP - Relational OLAP

- Modeled on top of the relational star schema database
- Data storage is kept in the relational database
- Utilizes SQL to query the DB in an OLAP manner
- May use proprietary in-memory caching techniques

HOLAP - Hybrid OLAP

- Combines the advantages of MOLAP and ROLAP
- Stores summary data in MOLAP structure
- Can "drill-through" to relational database for more detail

Top OLAP Products

- For Dimensional BI Uses
 - IBM Cognos Transformer Cubes (MOLAP)
 - Microsoft SQL Server Analysis Services (SSAS)
 - Dimensional and Tabular (MOLAP/HOLAP)
 - IBM Cognos Dynamic Cubes (ROLAP)
 - MicroStrategy (ROLAP)
- Typically For Finance Use (less “free form” BI)
 - IBM Cognos TM1 (writeback)
 - Hyperion Essbase (writeback)

COGNOS POWERPLAY (TRANSFORMER)

Advantages

- Performance (vs. relational)
- Easy to use and develop
- ETL-like capabilities (limited) - i.e. no star schema needed
- Can act as meta-data layer
- Great relative-time calc capabilities (YTD, Rolling 13 months...)
- Less intensive hardware requirement

Challenges

- Significant cube size limitations
- Limited categories per dimension level
- Cube builds take time & Cubes exist as separate files (.mdc)
- Lacks capabilities now available in other OLAP tools
- Row-level (dimensional) security is very challenging to maintain
- Unclear product support going forward
- Only works in the IBM Cognos stack

IBM COGNOS DYNAMIC CUBES

- IBM Cognos Dynamic Cubes was added to the Cognos 10.2 BI suite as an in-memory Relational OLAP product that could address the challenge of high-performance/low latency interactive analysis against terabytes of data
- The last significant update to Dynamic Cubes occurred in version 10.2.2. IBM has since focused most development efforts on the Cognos 11 release
- Currently, no current plans by IBM to enhance the Dynamic Cubes product

IBM COGNOS DYNAMIC CUBES

Advantages

- **Scalability** - limited only by database and RAM cache sizing
- Handles **large dimensions** with drill to detail - It also handles dimensions with hundreds of millions of records. This allows for user “drill to detail” which allows users to see the fact table level detail.
- Allows Dimension **attributes**
- Built-in **Relative Time** calcs on par with Transformer
- **MDX Scripting** Can set up just about any type of calculation that is necessary.
- **Dynamic Security** - you can set up dimensional filtering so that all security is derived from sql tables.
- **Aggregate advisor** helps tune database
- **Aggregate Aware** - can dynamically select database aggregate tables or in-memory aggregates for fast results.

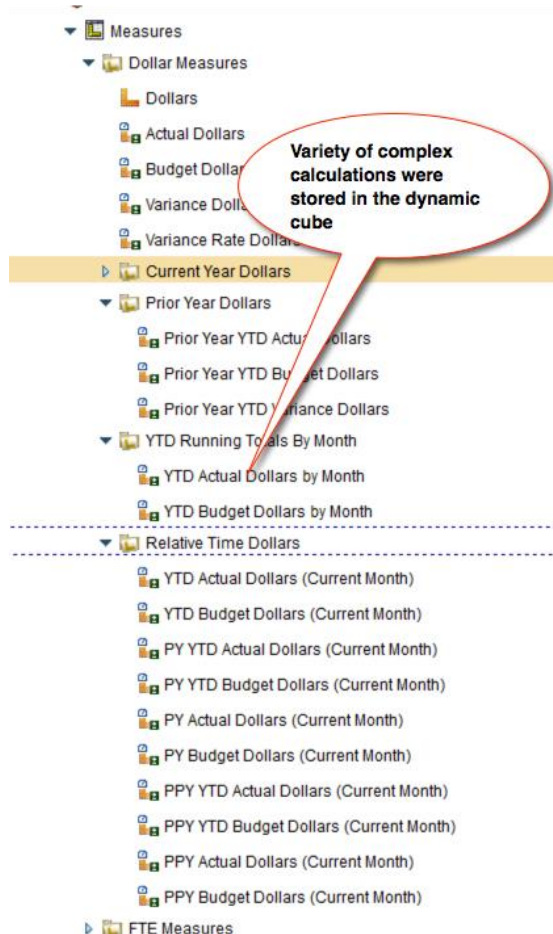
IBM COGNOS DYNAMIC CUBES

Challenges

- Requires **star or snowflake schema** as data source
- **Cache** needs to be “warmed” for decent performance
- Requires **64-bit** application server and may require significant memory footprint for large cubes (e.g. 64-128GB)
- Does not support Visual Totals when member security is used
- Report Authors require dimensional reporting experience
- **CAN ONLY BE USED BY COGNOS BI STACK**

Dynamic Cubes in play

Large Health Insurance Company Deployed Dynamic Cubes



- Finance Project used IBM Cognos Dynamic Cubes to replace legacy Cognos Transformer cubes. went into production Q1 2017
- Large number of reports were converted or created on top of the Dynamic Cube to provide a guided set of highly formatted reports that allowed drill-down
- Many complex business calculations were developed in the cube so that report writers can leverage a central set of calculations without having to write them in the report

MICROSOFT SQL SERVER ANALYSIS SERVICES (SSAS)



MultiDimensional	Tabular
Dimensions and Measure Groups	Tables and Relationships
Fast Design and in-Memory	Highly scalable and mature
Easy to get started	Feature Rich and Complex

MICROSOFT SQL SERVER ANALYSIS SERVICES (SSAS)

TABULAR MODEL



- Introduced in SQL Server 2012
- Model paradigm = Tables and Relationships
- Data stored in-memory
- Uses a different engine (xVelocity) and uses a columnar DB structure
- Combines the functionality of MOLAP Cubes and Relational DBs

MICROSOFT SQL SERVER ANALYSIS SERVICES (SSAS) TABULAR

Advantages

- Simpler data development model. Faster to develop
- Generally much faster than MOLAP
- DAX learning curve is easier than MDX
- Fast COUNT DISTINCT queries

Challenges

- Dependent on server memory footprint. (DirectQuery Mode now available in 2016)
- Some multidimensional features are not available (e.g. Many-to-Many)
- Complex calculations may be difficult to implement
- Large datasets



MICROSOFT SQL SERVER ANALYSIS SERVICES (SSAS) MULTI DIMENSIONAL

Advantages

- **Scalability** - Example: 1 billion fact table records on a server with only 8 GIG of memory, and performance was decent.
- Handles **large dimensions** with drill to detail - It also handles dimensions with hundreds of millions of records. This allows for user “drill to detail” which allows users to see the fact table level detail.
- Live “**count distinct**” capability at any level (e.g. unique invoice header count)
- **Dynamic Date Calculations** - via MDX extensions
- **Complex Calcs** Can develop complex calcs via MDX scripting
- **Dynamic Security** - you can set up dimensional filtering so that all security is derived from sql tables. This is handy when you have a large cube with security that needs to be table driven

Challenges

- Limited to **Microsoft** SQL Server platform
- Not all features get **exposed** thru other BI tool reporting layers
- **MDX coding** required for some common functions (e.g. relative time)

SSAS in play

Major American Clothing Company

- Re-architected an older Oracle based data warehouse to a SQL Server
- User community already very familiar with cube technologies
- Wanted to use SSAS OLAP cubes for their advanced relative time calcs
- Ability to create complicated advanced inventory calcs and on the fly currency conversions
- Ability to set defaults for certain dimensions such as currency type
- SSAS fits into their corporate strategy for multiple tools
- SSAS Tabular was chosen for performance and flexibility



SENTURUS

New Generation

Visualization Tools

New Generation Visualization Tools



- Over the last few years desktop visualization tools have sprouted on desktops throughout the enterprise
- IBM Cognos Analytics 11 now allows similar functionality over a Web interface
- Rich visualizations are now easily created by business users without the help of IT
- Decentralized model of data governance
- No waiting on developers to create next iteration of an OLAP cube
- Allows users to integrate data on the desktop/web
- Creation of desktop 'micromodels' (Tableau Data Extracts)
- Can use OLAP datasources but works best with non-OLAP sources
- Can begin to have performance issues when creating large data extracts or going against large datasources

Tableau Data Extracts

“Building Cubes for Tableau can be a waste of time!”

- DTEs is a compressed snapshot of data stored on disk and loaded into memory as required
- Data Engine can be described as its own “in-memory analytic database”
- Stores data in a Columnar Store Structure.
- Dramatically reduces the input/output time required to access an aggregate values

Reasons to use DTEs

- Better Performance vs. connected datasources
- Reduced load on connected datasources
- Portability - can be bundled in a packaged workbook for easy sharing
- Pre-Aggregation - option to aggregate data for visible dimensions “Aggregated Extract”

Tableau and Cubes

- OLAP Cubes want to do all the calculations
- Tableau will work if you stay within the structure of the cube
- Cubes are centrally developed
- Cubes can be the only primary source; No data blending
- No Cube Extracts
- Supports:
 - Oracle Essbase
 - Teradata OLAP
 - Microsoft Analysis Services (SSAS)
 - SAP NetWeaver Business Warehouse
 - Microsoft PowerPivot
 - Analytical Views in SAP Hana

IBM Cognos Analytics 11

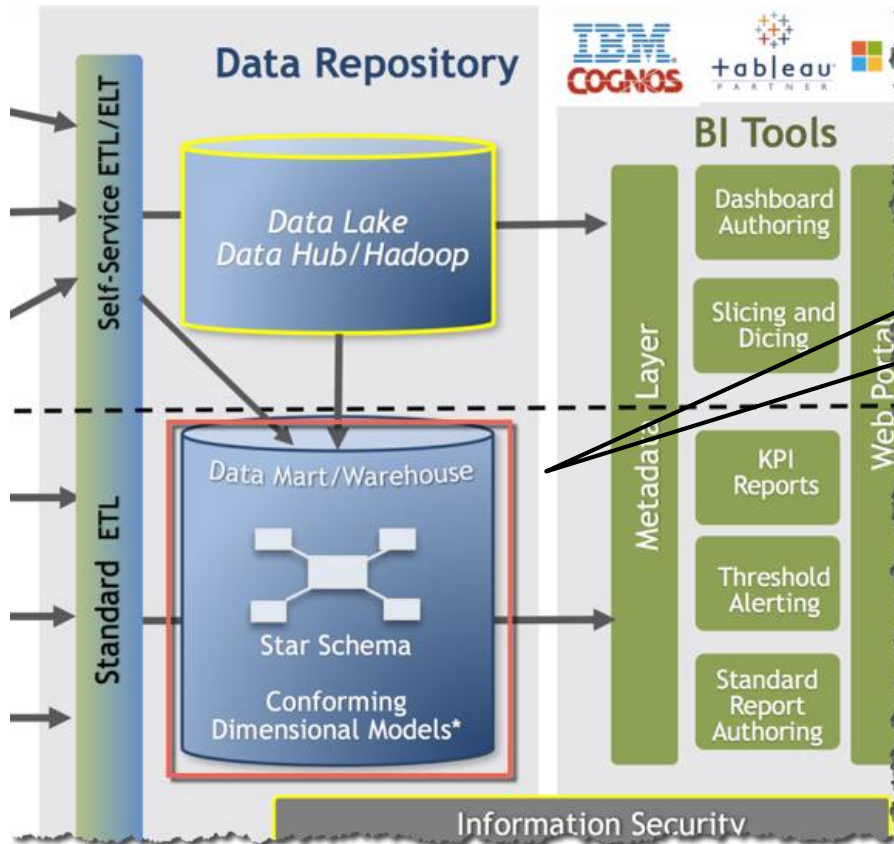
- New Cognos 11 architecture adds **Data Modules** which represent a major shift in the central metadata layer (framework) paradigm
- **Data Modules** now allows end users to quickly add new data sources and quickly model new data subjects without having to wait for DWH changes
- Uploaded files and data sources can be stored as ‘snapshots’ on the server’s file system using the **Apache Parquet** columnar file storage mechanism
- Allows for fast query response times



SENTURUS

Fast Columnar and In-Memory databases

“IN-MEMORY DATABASES ARE KILLING OLAP”



COLUMNAR & IN-MEMORY DATABASES

Columnar Databases

- Traditional databases store data by each row
- Columnar databases store data in columns rather than in rows
- This storage architecture can result in high-performing queries especially aggregation queries
- Example DBS:
 - Sybase IQ
 - IBM DB2 with BLU Acceleration
 - A capability built into DB2. Not a separate install component
 - Focus on Analytics
 - Dynamic In-Memory. Does not require all data to be in-memory
 - Columnar and Traditional Row-Based Tables
 - SQL Server 2014/16
 - Columnar Store indexes
 - In-Memory OLTP tables

BUT IF REMOVE THE OLAP LAYER

- Raw queries will be fast but what about the semantic layer?
- You could use relational models with some level of metadata and calculations
- But complex calcs, dimensions, drill downs would be missing



SENTURUS

**THE CURRENT STATE OF OLAP
ARCHITECTURE**

***THE NEW SCHOOL
BIG OLAP ON BIG DATA***

Big OLAP on Big Data

- Several new vendors and open source solutions are creating new scalable OLAP on Hadoop products

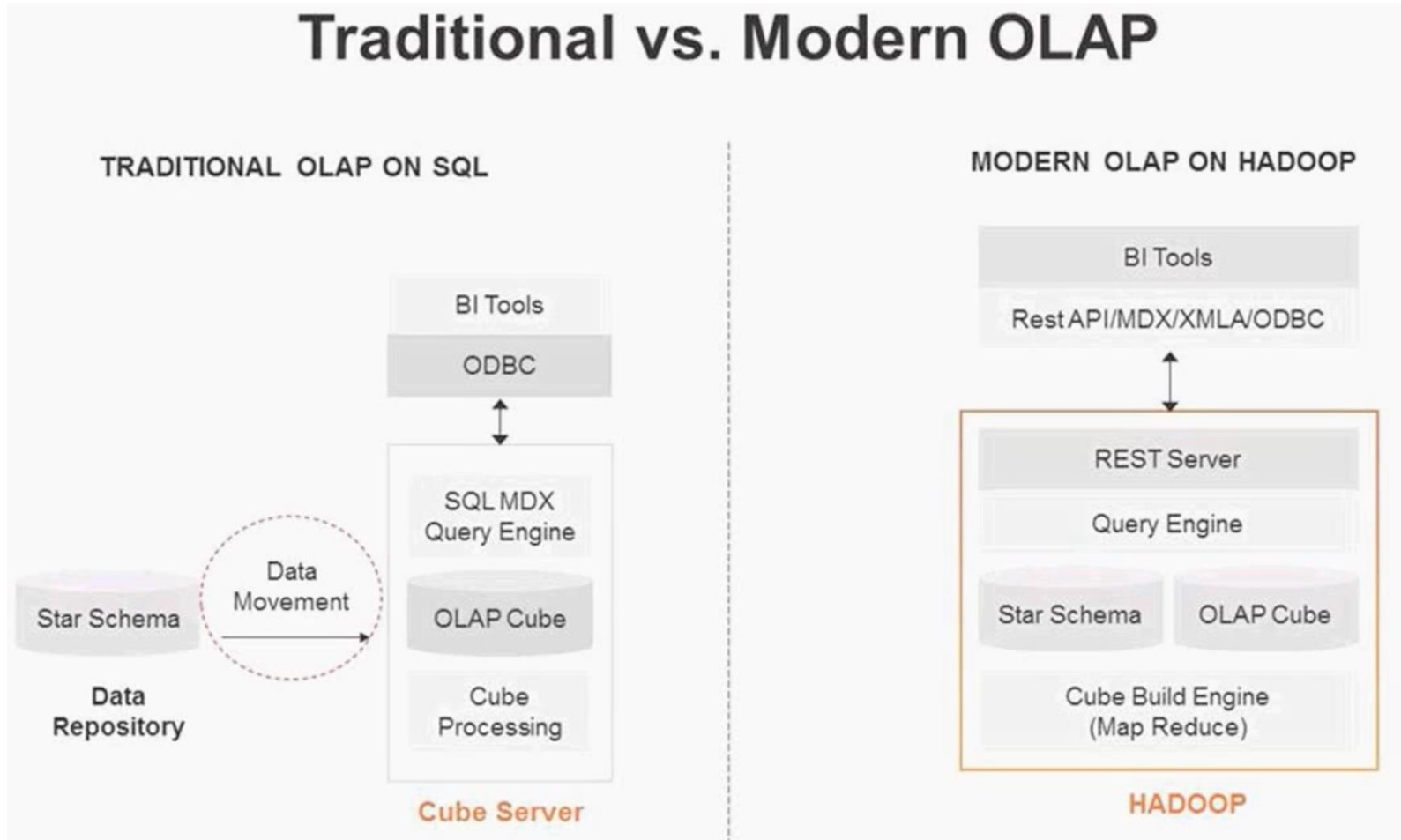


Big OLAP on Big Data

- Slow performing queries on Big Data implementations are driving new OLAP technologies
- Classic OLAP technologies on Big Data necessitated movement of Hadoop data into traditional Relational Data Warehouses further increasing latency.
- New OLAP technologies are architected to be part of the Hadoop stack and allow queries across Hadoop with not additional movement of data

Big OLAP on Big Data

Traditional vs. Modern OLAP





SENTURUS

Summary Wrap-up

So is OLAP Dead?

- In general, the key concepts of OLAP - dimensions, measures, hierarchies, drill-down etc. are still alive and well. But the technology that surfaces those concepts are changing
- Business users will always want a high performing BI layer that is easy to use and allows for interactive BI
- Some will want a central repository that contains all the relationships, hierarchies and complex business rules already developed
- Other users like data scientists, advanced business analysts will want a more agile free form solution yet still have high performance