

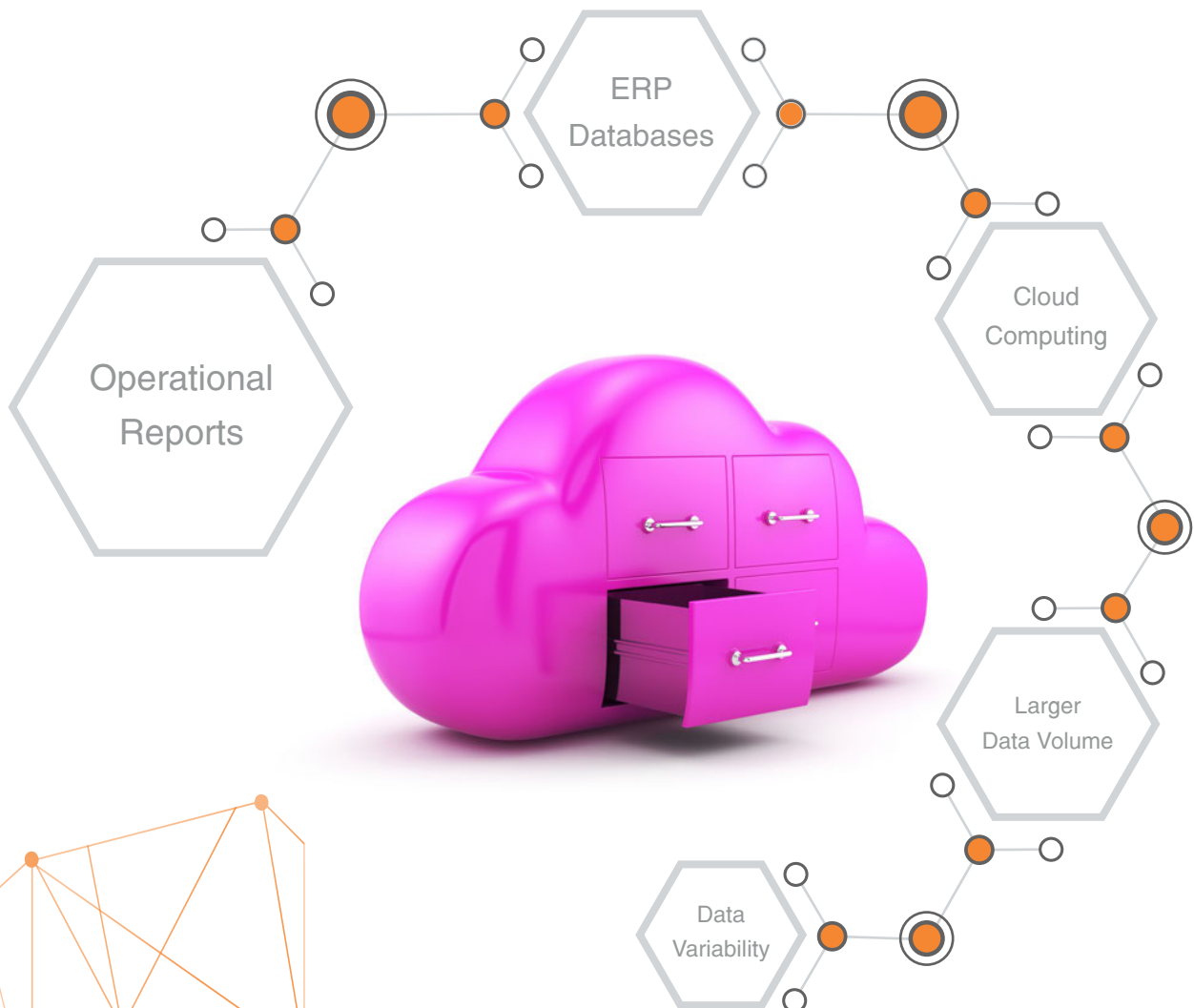


OPERATIONAL DATA STORE

Summary

Traditionally Operational Data Store (ODS) is a database that is designed to integrate data from multiple sources for additional operations on a time sensitive data. Because the data originates from various sources, consistency and cleaning is usually required. ODS delivers cross system operational reports and dashboards, and also acts as a key data warehouse source. ODS stores data at lower level of granularity with minimum possible history based on organization's specific needs.

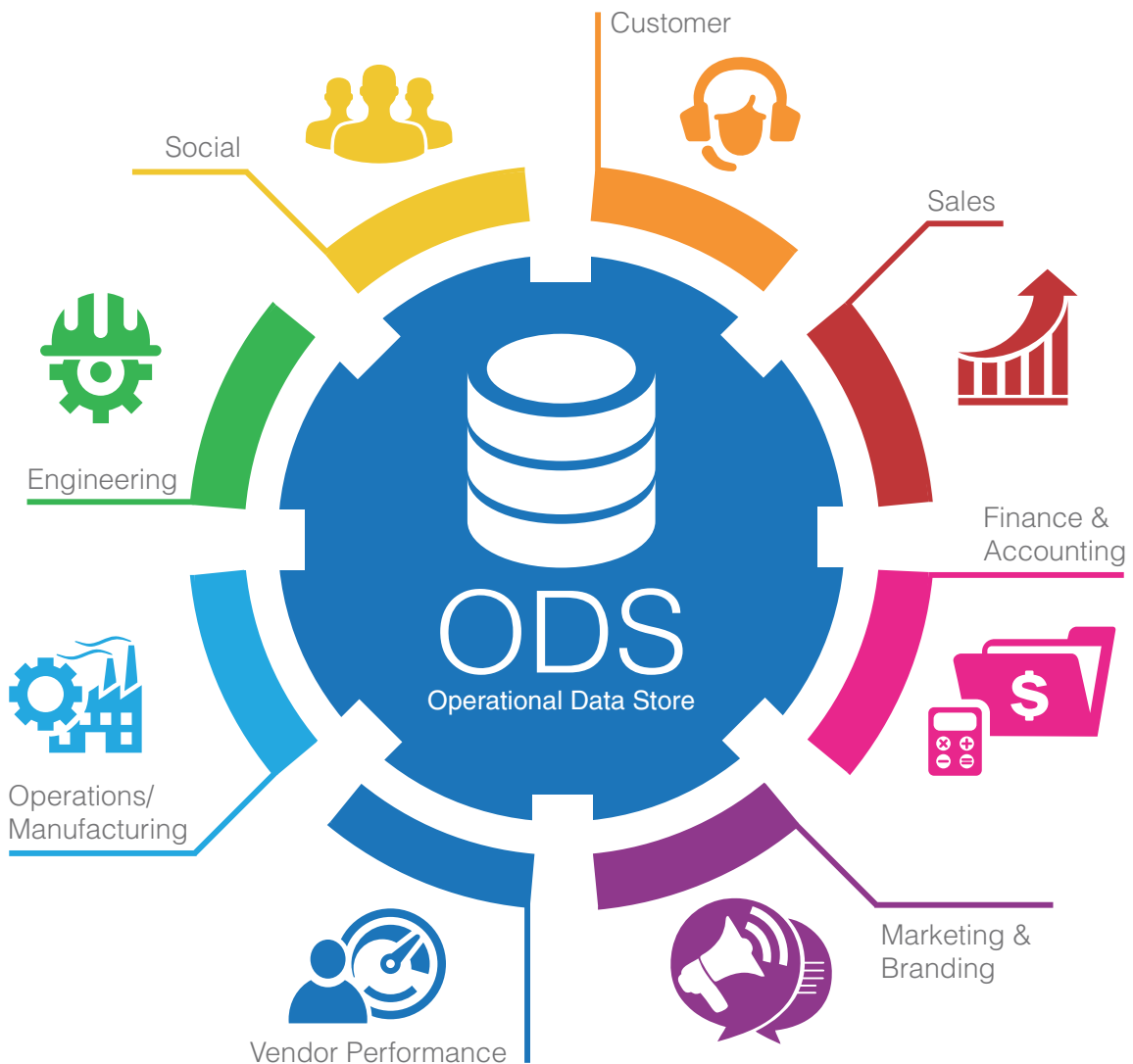
With the technology advancement in areas of server virtualization and parallel processing, ODS can be leveraged above and beyond the source for operational reports. However, on the flip side, the evolution of proprietary 3rd party ERP databases, cloud computing and much larger data volume, velocity and data variability is adding tremendous complexity and is compelling data architects and other IT pundits to perceive ODS as a strategic but a complex solution domain. This document discusses the ODS evolution roadmap and key considerations that can be referred in developing ODS strategy and plans.



ODS Evolution Roadmap

ODS can evolve from a simple integrated operational data store with few data sources to support operational reporting to a complex and the state of the art solution with bi-directional feeds, rules engine, triggers and provides a staging area for enterprise data warehouse.

Following section presents a reference roadmap, organizations can develop their own roadmap depending upon the ODS cost, complexity and the business benefits.



Reference Roadmap

The reference roadmap under following listed points is in the logical order of ODS maturity.

- ODS as an integrated operational reporting database for selected business functions
- Data visualization tools accessing the ODS for operational reporting and dashboards
- Iterative addition of new source systems
- ODS as with integrated metadata management capabilities
- ODS as 'System of Reference' for master data, enabling Bi-directional data exchange
- ODS with integrated rules engine and triggers
- ODS with integrated data mining capabilities
- ODS as a data warehouse staging area
- ODS as an enterprise OLTP (Online Transaction Processing) system, enabling enterprise OLAP (Online Analytical Processing) capabilities

Reference Development Plan

Depending upon the organizational skills, capabilities, ODS complexity and cost considerations, some of the listed activities can be executed in parallel.

- Develop ODS roadmap
- Develop ODS data model approach and the data architecture
- Identify source system integration sequence, priority and data refresh frequency
- Identify data atomicity and transactional events for each source system
- Develop data security model
- Develop data governance model
- Develop data integration architecture
- Develop ODS availability and disaster recovery options
- Develop conceptual data model
- Develop logical data model
- Identify technology stack for database, integration, metadata, master data, rules engine and visualization
- Identify infrastructural needs including, network, compute, storage
- Provision infrastructure and 3rd party software
- Install and configure 3rd party software
- Data model physical development
- Enable integration
- Other implementation activities per the defined roadmap

Key Considerations

In addition to the major ODS components like Database, ETL and integration, following are some of the key considerations that should not be overlooked while developing ODS strategy and the plan.

Metadata Management

As ODS gets populated with data from different sources, it becomes important to know and maintain the key properties of data like structural metadata, business metadata and the governance metadata e.g. definition, version, location, admin authorities, classification, codes and mappings for data linkage, data lineage, data security and for overall data governance purposes.

ODS Data model

Data model design needs to be based on the ultimate expectations from the ODS. Typical ODS design is a hybrid between OLAP and OLTP. While ODS is catering to reports (design for faster reads), it may not hold the history to enable data analytics and trending. The evolution of data model can start with simple copies of source system logical data model and physicalizing those for faster reads. The data model can evolve into an enterprise OLTP (EOLTP) system, a long term and an iterative process. The EOLTP system will be the ultimate system of reference encompassing the logical model for the entire organization's business. EOLTP can be assumed as an integrated and a unified data model for CRM, ERP, MRP and other key business functions.

Data Governance

While ODS may not be the system of record for compliance reasons, however, ODS as system of reference imposes a need for a formal data governance. Formal data governance is needed to manage the master data changes, any source system structural or data processing changes and to implement data security.

Data Security

Data visualization with self-service capabilities, integration of cloud applications, web services and the organization wide access to decision support systems are imposing major data security issues. Security needs to be implemented either at application / visualization layer, data layer or the combination of two and in collaboration with the organization's security policies and functions.

Data Linkage

Data / record linkage is an activity of finding records in a data set that refers to the same entity across different data sources. Data linkage is necessary when integrating data sets from different source that may or may not share common identifiers.

Data Lineage

With multiple sources, targets and reporting & consolidation system with embedded business rules, it is important to manage the data lifecycle to avoid data integrity issues. Data lineage is the methodology and an approach for finding more effective use of the business information and it brings data lifecycle transparency to all the stakeholders. The metadata abstraction layer enables the data lineage information. Data lineage will track the data from its origin to destination and it will also depict the different processes, business rules and functions involved in the data flow and their dependencies.

Change Data Capture

As ODS gets populated from different sources, it's required for ODS or integration layer to know and propagate the source system data changes to ODS. Change data capture (CDC) is an approach to data integration that is based on the identification, capture and delivery of the changes made to data sources. Multiple software design patterns exist and a thoughtful consideration is needed to select an appropriate option. Some of the examples are: change timestamp, row level versioning, event programming and data level triggers.



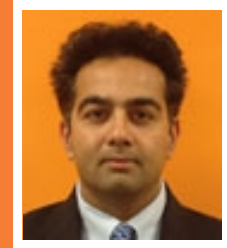
A well-executed ODS strategy will solve multiple underpinning data integration and quality problems, deliver business with a consistent, clean and integrated data and will also create a foundation for enterprise level OLAP and OLTP systems.

The key success factor is developing a comprehensive strategy that encompasses all relevant ODS building blocks and considerations, organizational capabilities like infrastructure technology, staff and skills, maturity and variability of the source systems.

About the Author

Amit Soni, VP and Chief Solutions Architect

Amit is leading Solutions Architecture and Business Intelligence practice as well as developing cloud based business, analytics and IT solutions for Manufacturing and Services industries, and delivering SaaS, PaaS and IaaS solutions to improve our client's business results.



About Netlink

Netlink is a global IT company that provides Information Technology and Business solutions using our proprietary cloud platforms to bring Immediate Business Results. Please visit www.Netlink.com for details.

Citations

Wikipedia Article on 'Data lineage': http://en.wikipedia.org/wiki/Data_lineage

Wikipedia Article on 'Change data capture':

http://en.wikipedia.org/wiki/Change_data_capture



USA:

Madison Heights, 999 Tech Row, Madison Heights, MI 48071
Tel: 1-800-485-4462

Bhopal:

Center Point, New Market, T.T. Nagar, Bhopal-462003 (MP) India
Tel: +91755-6671931

New Delhi:

7th Floor, Enkay Tower, Udyog Vihar - Phase V, Gurgaon-122016 Haryana, India
Tel: +91755-6671931