Oracle Data Integration Strategy and Roadmap

Oracle Fusion Middleware Product Management

ORACLE®

Agenda

- **Introduction to Oracle Data Integration**
  - Business Drivers for Data Integration
  - Benefits from a Modern Data Integration Platform
  - Key Oracle Data Integration Products
  - Oracle Data Integration Solution
- **Oracle GoldenGate Overview**
- **Data Integrator Overview**
- **ODI & GG Together**
  - Best of Breed Integration for Batch and Realtime Data Integration
  - Support any Type of Data Integration Use Case
  - Implementing Best-Practice Technical Pattern for Data Warehousing
- **Technical Details – How it Works**
- **Demonstration and Q&A (if available)**

ORACLE

Data Integration Portfolio

INTRODUCTION

Business Drivers for Data Integration
Essential Ingredient for <u>Information Agility</u>
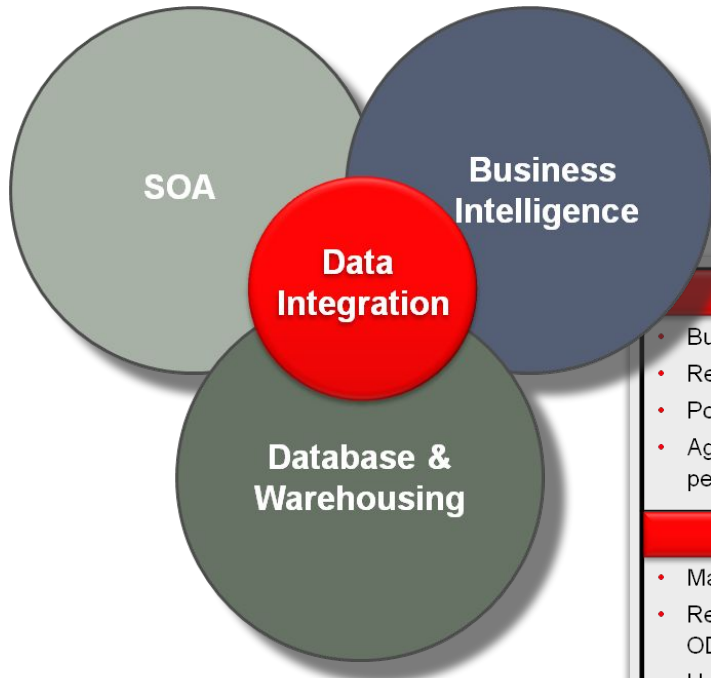
SOA

Business Intelligence

Data Integration

Database & Warehousing

Key Data Integration Use Cases

• SOA, Enterprise Integration & Modernization

BI & DW Projects	DB Availability & Replication
• Bulk-load historical data to new application • Real-time Data Warehouse for BI • Populate Warehouse with High Performance ODI • Aggregates and aligns data for operational analytics, performance management, etc	• Real-time DB High-Availability & Backup • Application data consistency (bi-directional) • Disaster recovery and data protection • Query offloading, DB performance optimization
Modernization Initiatives	**Migrations / Consolidation**
• Mainframe / iSeries sources to Oracle RAC • Re-architect legacy batch processing to SOA and ODI-EE • Unification of structured and unstructured data	• Upgrade Applications or Migrate to New Schema • Single-time Bulk load and/or keep in Sync with ongoing translation, delivery of data. • Applications merge support
SOA Projects	**Master Data Management**
• Establish CDC framework for pushing database events into SOA, BAM and CEP • Incorporate Efficient Bulk Data Processing with ODI • Providing data access, large payload transformation etc. as services within SOA	• Create Single View of the Truth in real time • Aggregating multiple operational sources • Synchronize Data with ODI-EE • Supports master data stores, publishes master data changes to all consumers

ORACLE

Benefits from a Modern DI Platform

Data Integration is Infrastructure that enables Business Value

Do More with Less

> Design metadata-driven integration
> Leverage skills & dictate patterns

Compete Globally 24X7

> Ensure continuous uptime
> Access data in real time

Use Data for Competitive Advantage

> Ensure the quality of your data
> Actively govern most valuable asset

Automate and Adapt Business Processes

> Expose data services for reuse
> Orchestrate processes using SOA

ORACLE®

Key Data Integration Products

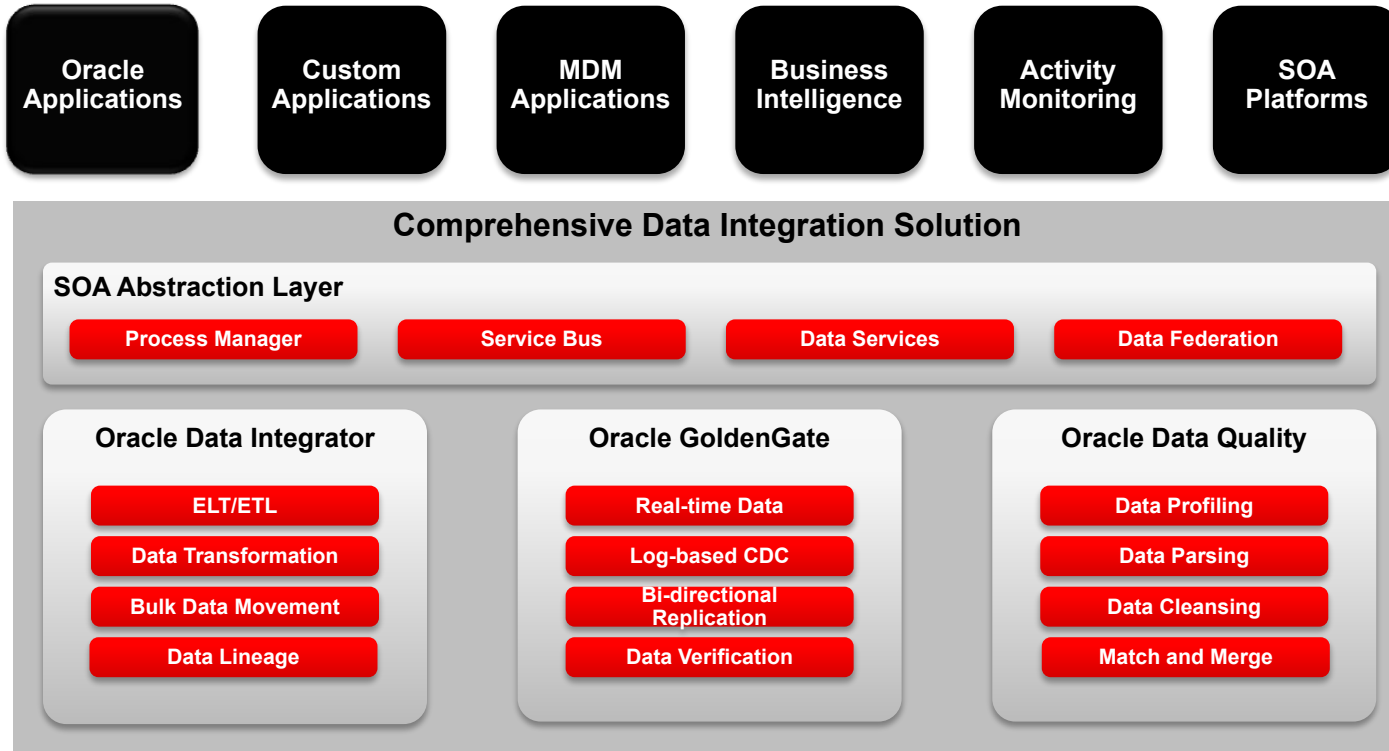| ORACLE® GOLDENGATE	• Real Time Data Replication • Changed Data Capture	• DBMS High Availability • Disaster Tolerance
ORACLE® DATA INTEGRATOR ENTERPRISE EDITION	• Heterogeneous E-LT & ETL • High-speed Transformations	• OLAP Data Loading • Data Warehouse Loading
ORACLE® DATA INTEGRATION SUITE	• Comprehensive Integration • ELT/ETL for Bulk Data • Service Bus	• Process Orchestration • Human Workflow • Data Grid
ORACLE® DATA SERVICE INTEGRATOR	• Data Service Modeling • XQuery Data Federation	• Data Security/Redaction • XA Compliance
ORACLE® DATA PROFILING	• Business Data / Metadata • Statistical Analysis	• Time Series Reporting • Integrated Data Quality
ORACLE® DATA QUALITY	• Cleansing & Parsing • De-duplication	• High Performance • Integrated w/ODI

Oracle Data Integration Solution
Best-in-class Heterogeneous Platform for Data Integration

Oracle Applications	Custom Applications	MDM Applications	Business Intelligence	Activity Monitoring	SOA Platforms

Comprehensive Data Integration Solution

SOA Abstraction Layer

Process Manager	Service Bus	Data Services	Data Federation

Oracle Data Integrator
- ELT/ETL
- Data Transformation
- Bulk Data Movement
- Data Lineage

Oracle GoldenGate
- Real-time Data
- Log-based CDC
- Bi-directional Replication
- Data Verification

Oracle Data Quality
- Data Profiling
- Data Parsing
- Data Cleansing
- Match and Merge

Storage | Data Warehouse/ Data Mart | OLTP System | OLAP Cube | Flat Files | Web 2.0 | Web and Event Services, SOA

Data Integration Portfolio
G O L D E N G A T E

Oracle GoldenGate Overview
Enterprise-wide Solution for Real Time Data Needs

Disaster Recovery, Data Protection — Standby (Open & Active)

Zero Downtime Migration and Upgrades

Log Based, Real-Time Change Data Capture

OGG

Operational Reporting — Reporting Database

ODS → ETL → EDW

ETL → EDW

Real-time BI

Query Offloading

Data Distribution

Heterogeneous Source Systems
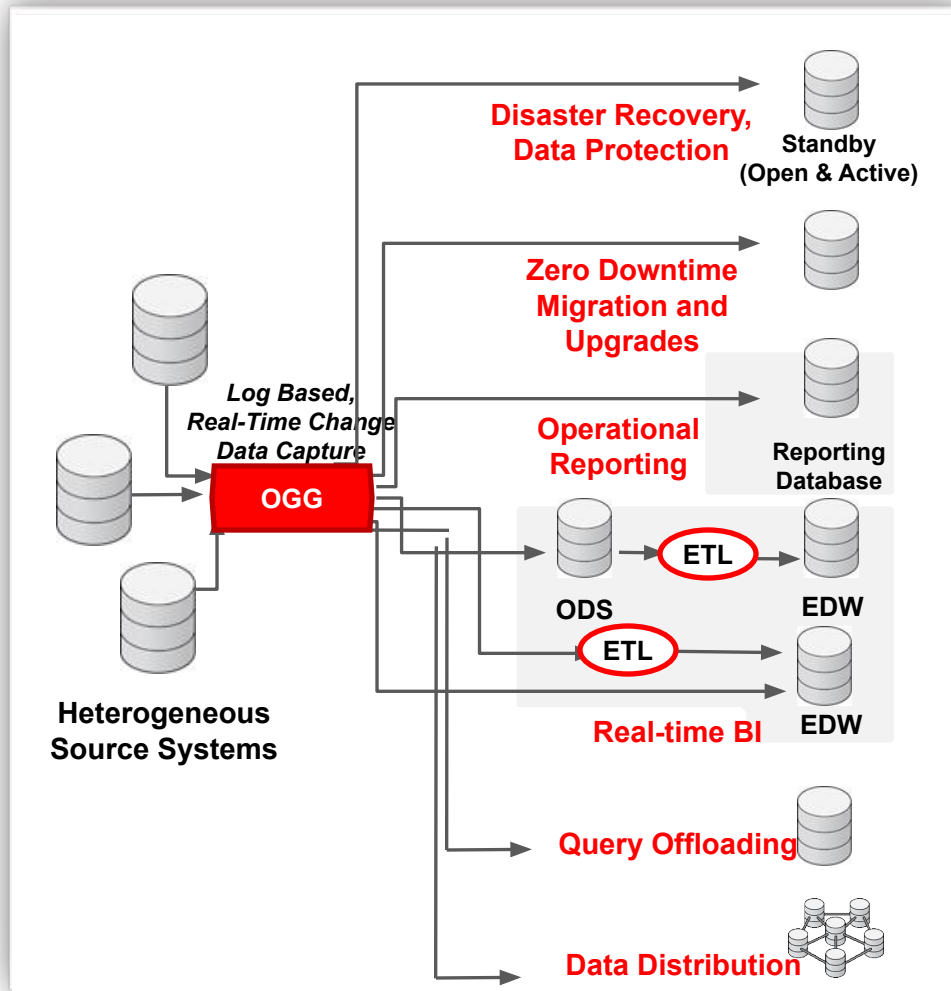
- Standardize on Single Technology for Multiple Needs
- Deploy for Continuous Availability and Real-time Data Access for Reporting / BI

- Highly Flexible
- Fast Deployments
- Lower TCO & Improved ROI

How Oracle GoldenGate Works
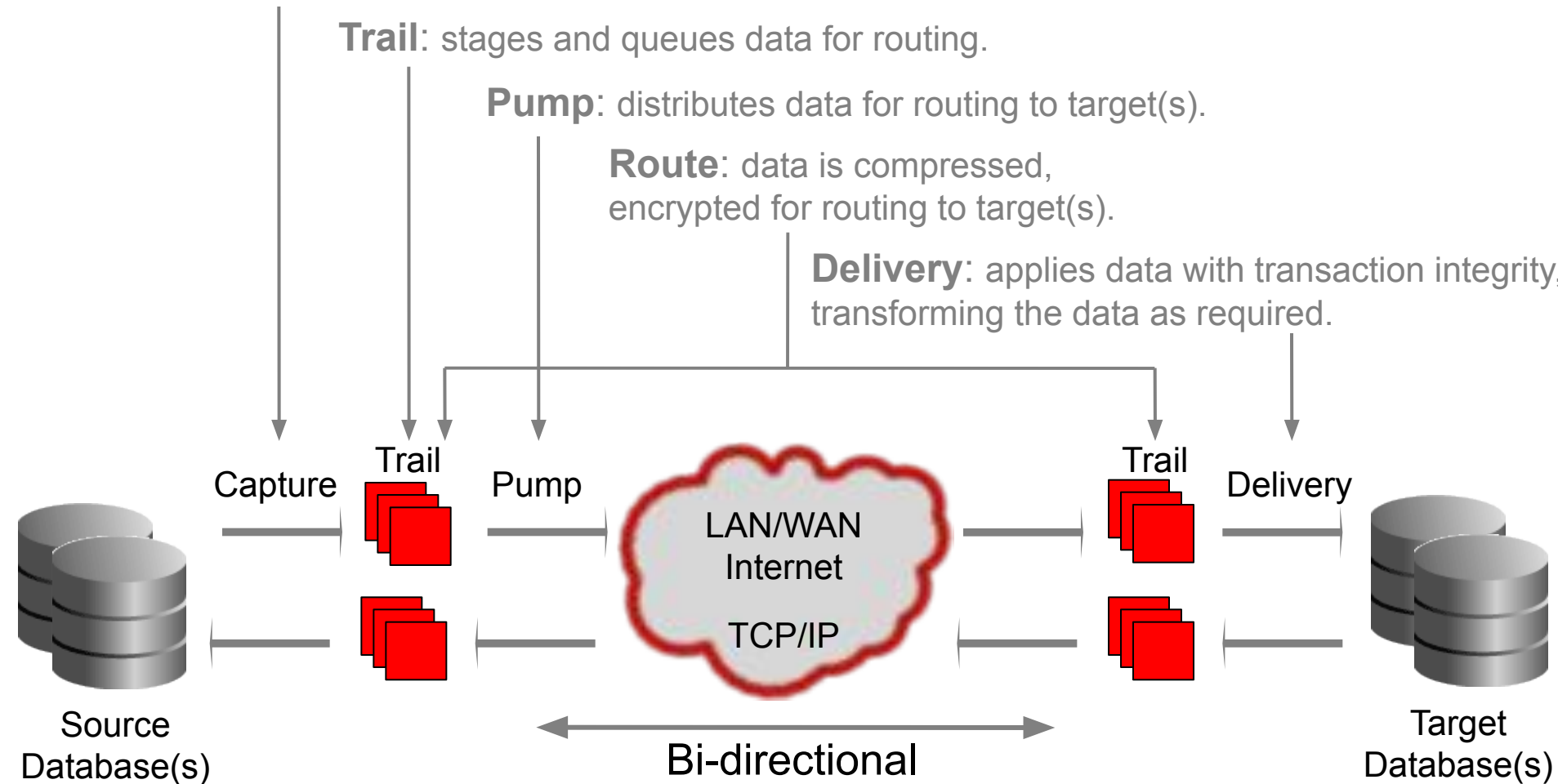
Modular De-Coupled Architecture

Capture: committed transactions are captured (and can be filtered) as they occur by reading the <u>transaction logs</u>.

Trail: stages and queues data for routing.

Pump: distributes data for routing to target(s).

Route: data is compressed, encrypted for routing to target(s).

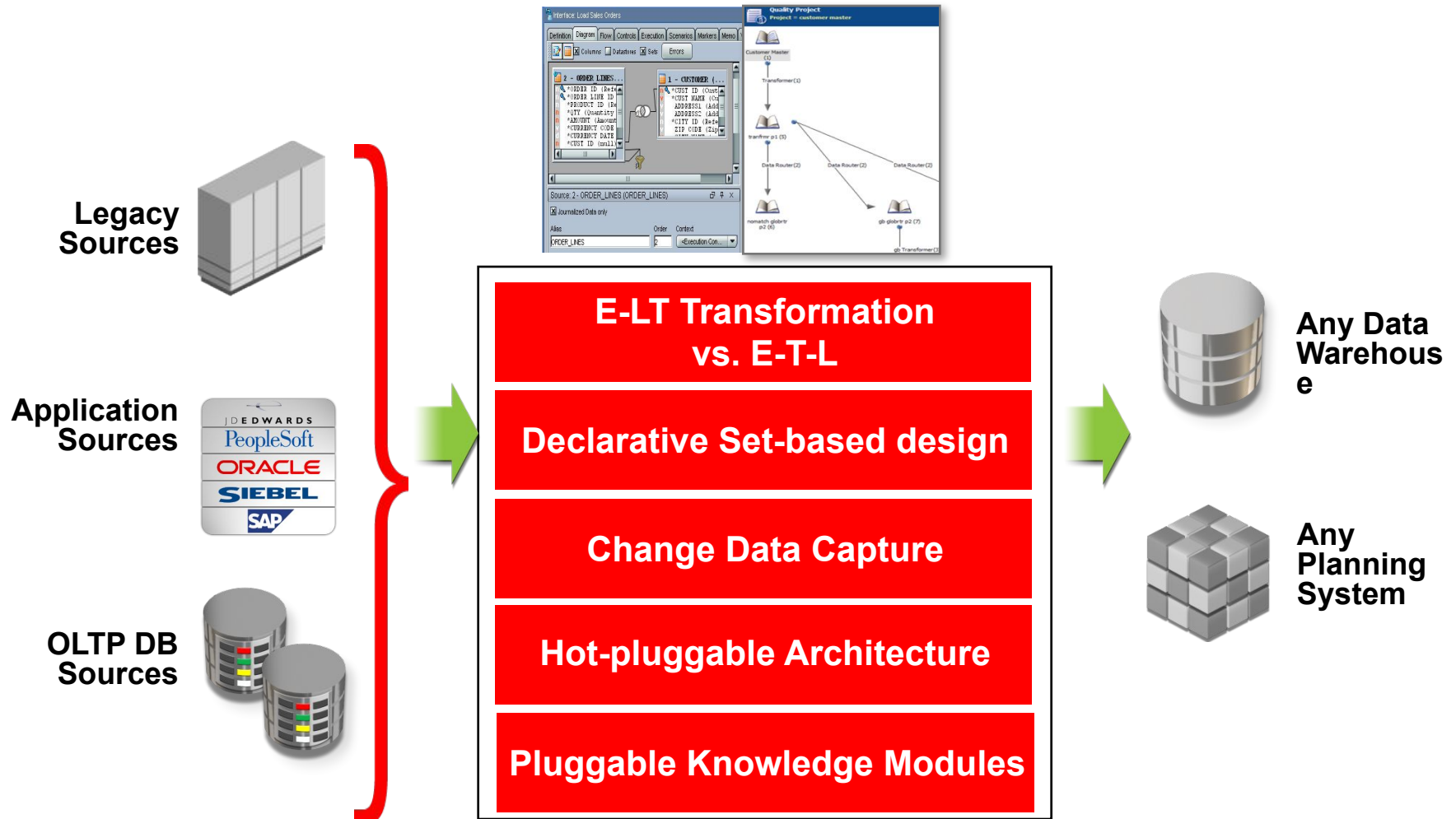**Delivery**: applies data with transaction integrity, transforming the data as required.

Capture   Trail   Pump   LAN/WAN Internet TCP/IP   Trail   Delivery

Source Database(s)   Bi-directional   Target Database(s)

Data Integration Portfolio
DATA INTEGRATOR

Oracle Data Integrator Enterprise Edition
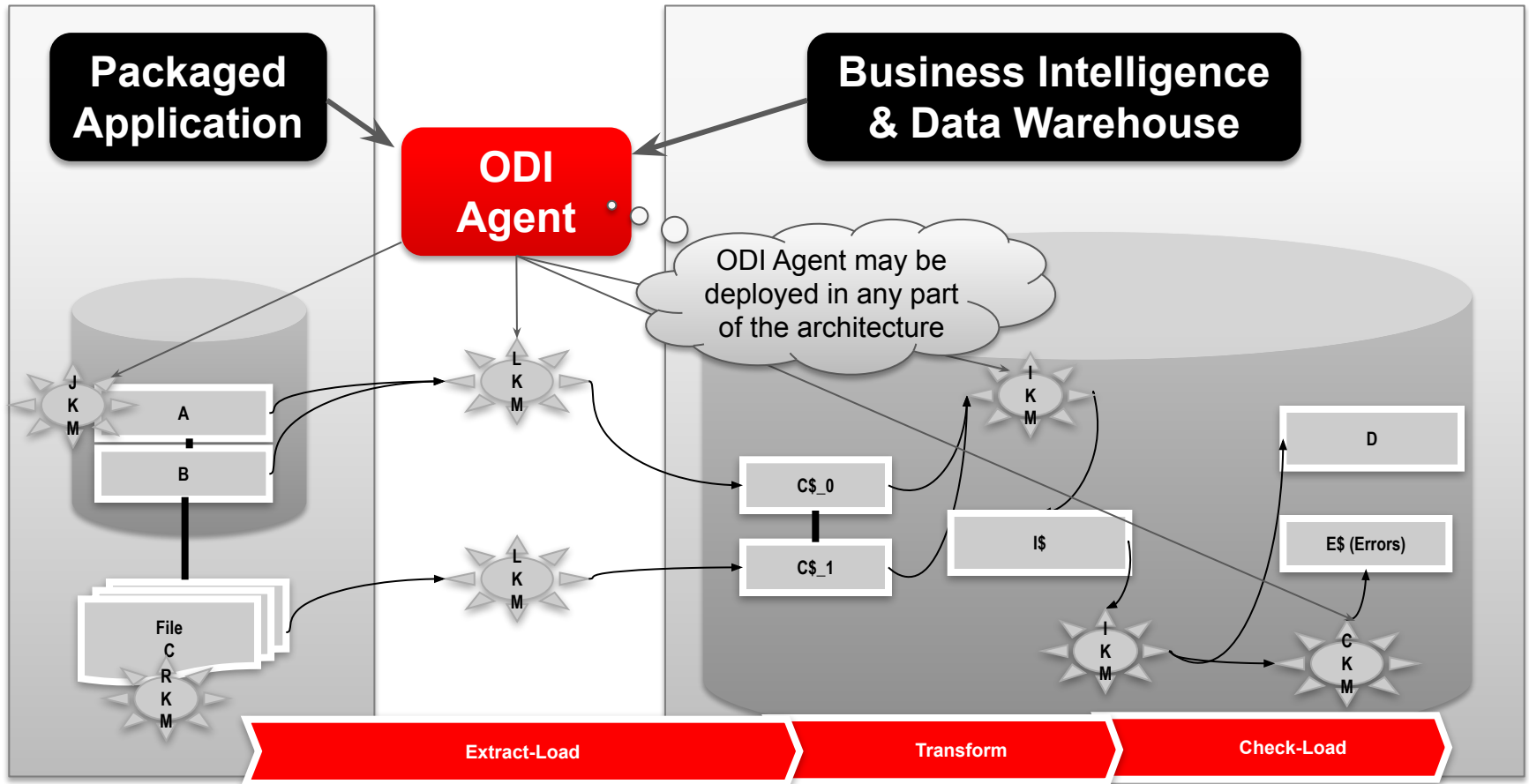## Optimized E-LT for High Performance, Productivity and Low TCO

Legacy Sources

Application Sources

OLTP DB Sources

E-LT Transformation vs. E-T-L

Declarative Set-based design

Change Data Capture

Hot-pluggable Architecture

Pluggable Knowledge Modules

Any Data Warehouse

Any Planning System

How ODI Works: E-LT Architecture
High Performance, Flexible, Lightweight Architecture

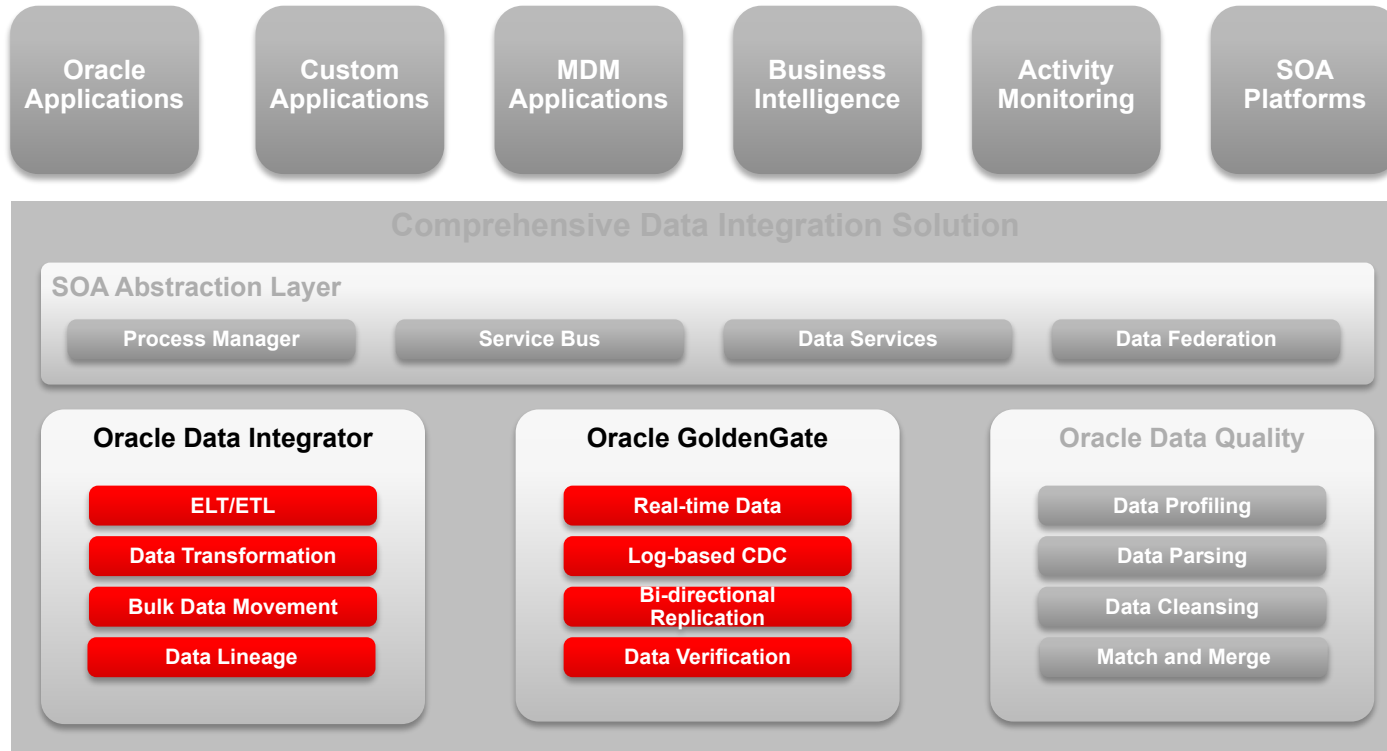- Key Architecture Benefits: 100% Java, Open APIs, fast E-LT

Data Integration Portfolio
O D I & G G T O G E T H E R

Oracle Data Integration Solution
Best-in-class Heterogeneous Platform for Data Integration

Oracle Applications	Custom Applications	MDM Applications	Business Intelligence	Activity Monitoring	SOA Platforms

Comprehensive Data Integration Solution

SOA Abstraction Layer

Process Manager	Service Bus	Data Services	Data Federation

Oracle Data Integrator
- ELT/ETL
- Data Transformation
- Bulk Data Movement
- Data Lineage

Oracle GoldenGate
- Real-time Data
- Log-based CDC
- Bi-directional Replication
- Data Verification

Oracle Data Quality
- Data Profiling
- Data Parsing
- Data Cleansing
- Match and Merge

Storage | Data Warehouse/ Data Mart | OLTP System | OLAP Cube | Flat Files | Web 2.0 | Web and Event Services, SOA
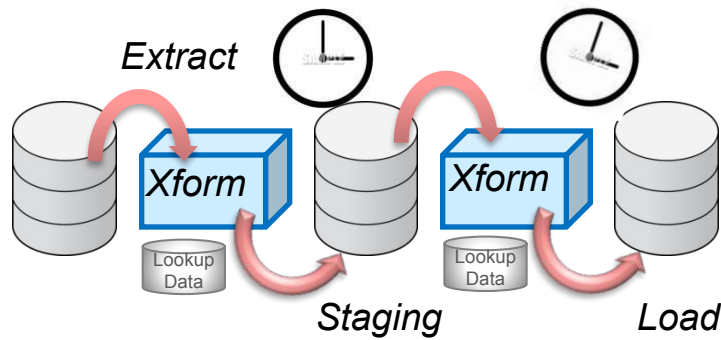
Best-of-Breed Data Integration

Heterogeneous, Real-time, Non-Invasive, High Performance E-LT, and Low Hardware Costs
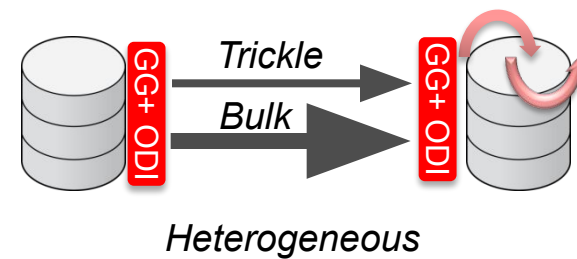
Traditional ETL + CDC

- Invasive Capture on OLTP systems using complex Adapters
- Transformations in ETL engine on expensive middle tier servers
- Bulk load to the data warehouse with large nightly/daily batch
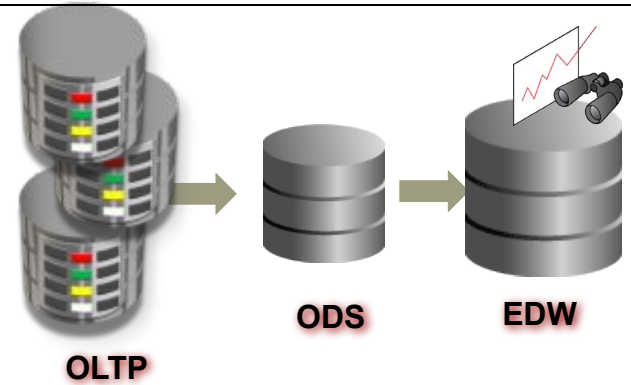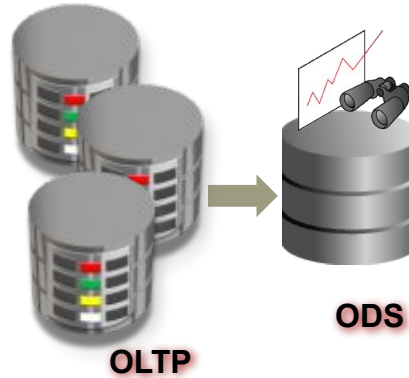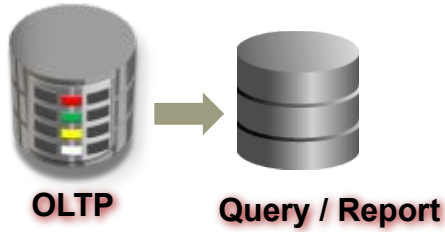
Oracle E-LT + Real-time

- Continuous feeds from operational systems
- Non-invasive data capture
- Thin middle tier with transformations on the database platform (target)
- Mini-batches throughout the day or bulk processing nightly

Extract

Xform

Lookup Data

Xform

Lookup Data

Staging   *Load*

Trickle
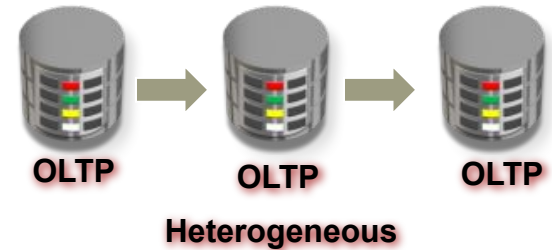
Bulk

GG+ ODI

GG+ ODI

Heterogeneous

ORACLE

Support Any Type of Data Integration
Best of Breed means using the Right Tools for the Job!

Analytical

OLTP → Query / Report

OLTP → ODS

OLTP → ODS → EDW

Operational

OLTP Old ⇄ OLTP New
Heterogeneous

OLTP → OLTP
Heterogeneous

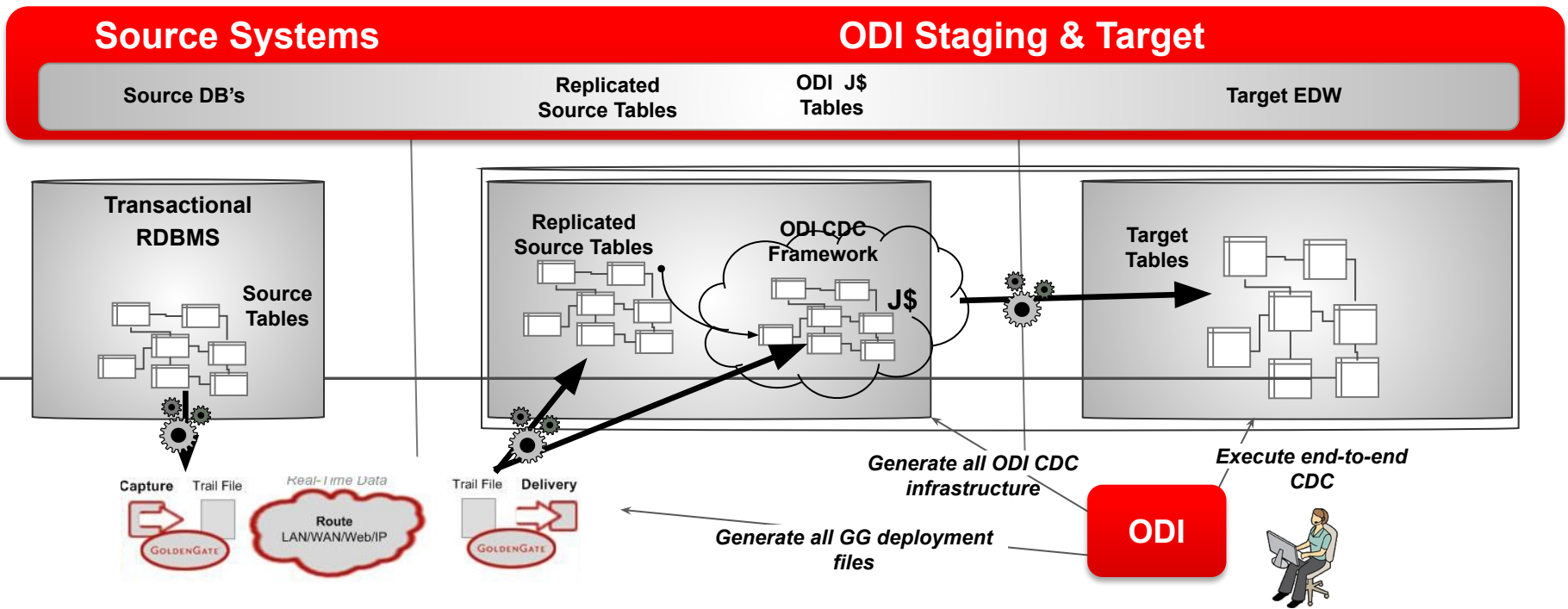OLTP → OLTP → OLTP
Heterogeneous

ORACLE
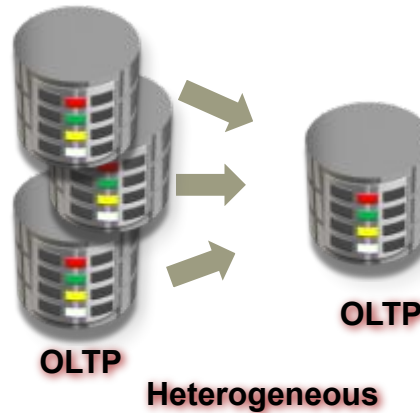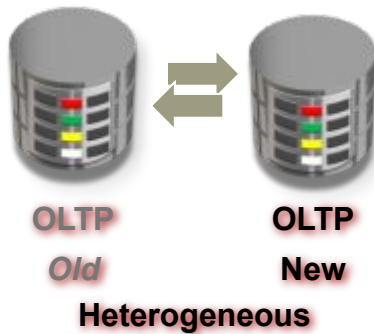
ODI-EE Integration with GoldenGate

Non-invasive Data Capture combined with ODI ELT strengths

Key Benefits:

1. **Eliminate Overhead** ⇒ no need for DB API overhead on the Source, or the invasiveness of the ODI J$ objects on the Source system,
2. **Automate GoldenGate** ⇒ automation of GG deployment *directly from* ODI GUI
3. **Provide Common DW Pattern** ⇒ supplies a common pattern for mini-batch style (non-real-time) DW aggregate loads

Source Systems

ODI Staging & Target

Source DB's

Replicated Source Tables

ODI J$ Tables

Target EDW

Transactional RDBMS

Source Tables

Replicated Source Tables

ODI CDC Framework

J$

Target Tables

Capture   Trail File   *Real-Time Data*   Trail File   Delivery

Route LAN/WAN/Web/IP

Generate all ODI CDC infrastructure

Execute end-to-end CDC

Generate all GG deployment files

ODI

Oracle's Data Integration Joint Solution

Best-of-Breed and *Proven*

Oracle GoldenGate

Oracle Data Integrator Enterprise Edition

Technology Differentiators:

Performance

Extensible & Flexible

Enterprise

- **Lowest latency** and highest throughput; non-invasive, low overhead

- **De-coupled** architecture; multiple deployment styles; open and extensible

- Maintain transactional **integrity**; resilient against interruptions and failures

- **E-LT architecture** for best performance of high data volume transformations

- **Knowledge Module** architecture for extensibility and flexible connectivity

- **SOA**-native, integrated with Fusion MW to fit future enterprise architectures

ORACLE

ODI & GG Together

TECHNICAL DETAILS

Overview of the ODI KM Framework
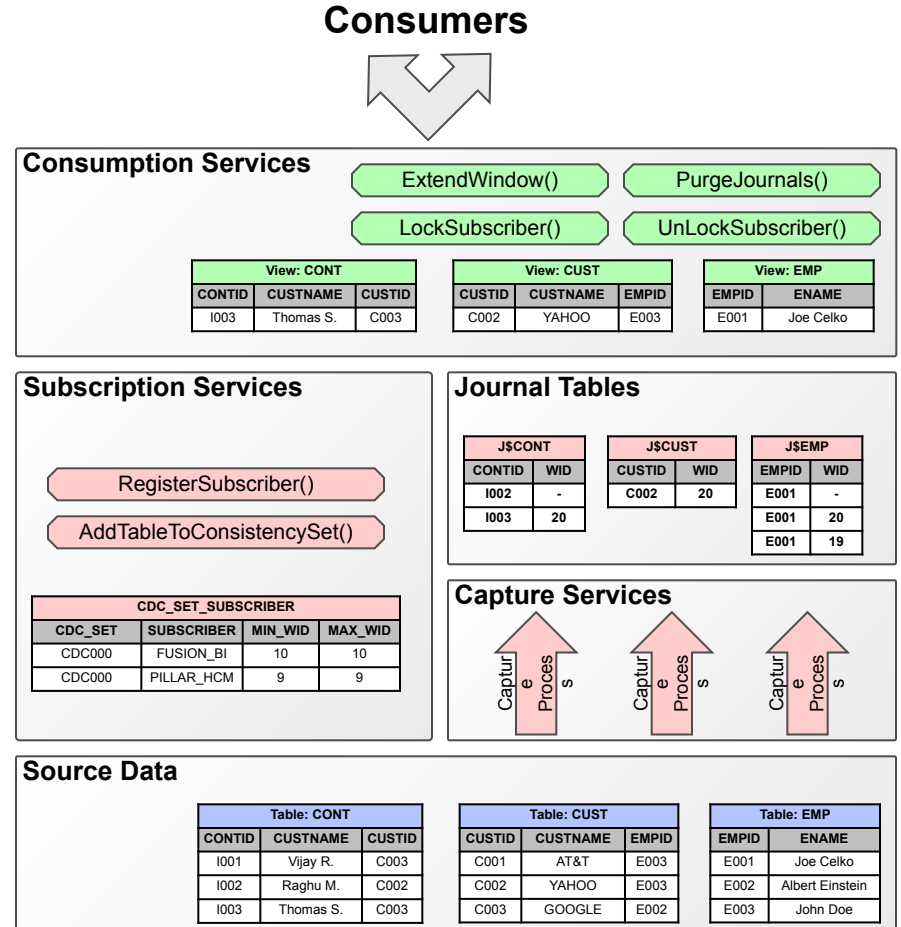
Pluggable Knowledge Module Architecture
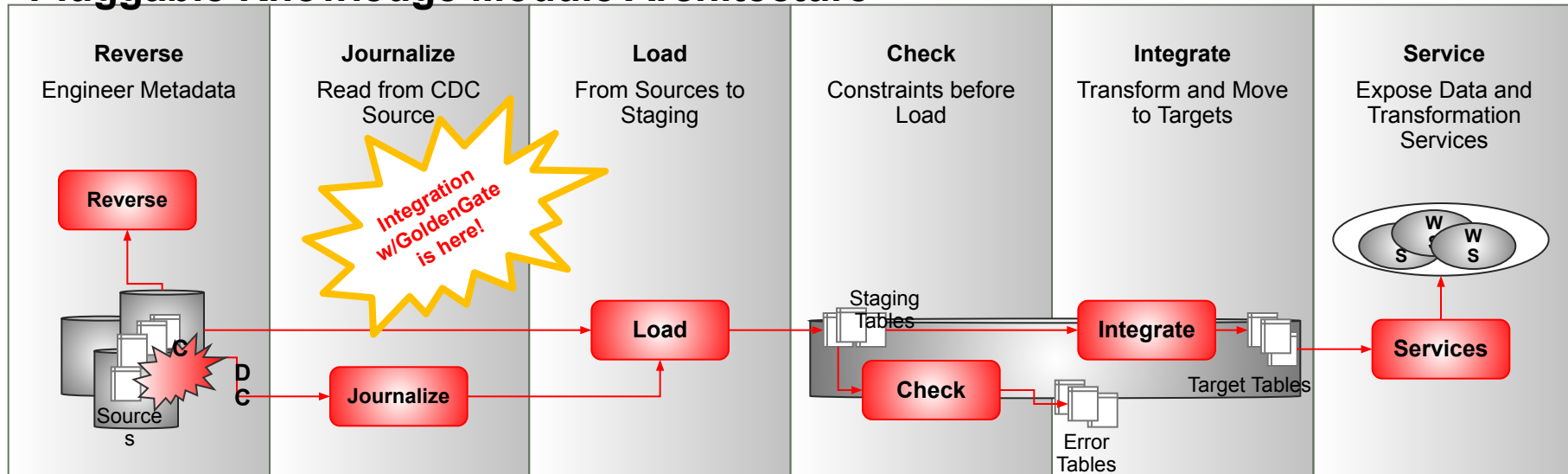
Benefits
- Leverage Database Optimizations:
  Native SQL; Native Functions; Native Loads; Native Journaling / CDC
- Tailor to an organization's existing best practices
- Ease administration work
- Reduce cost of ownership
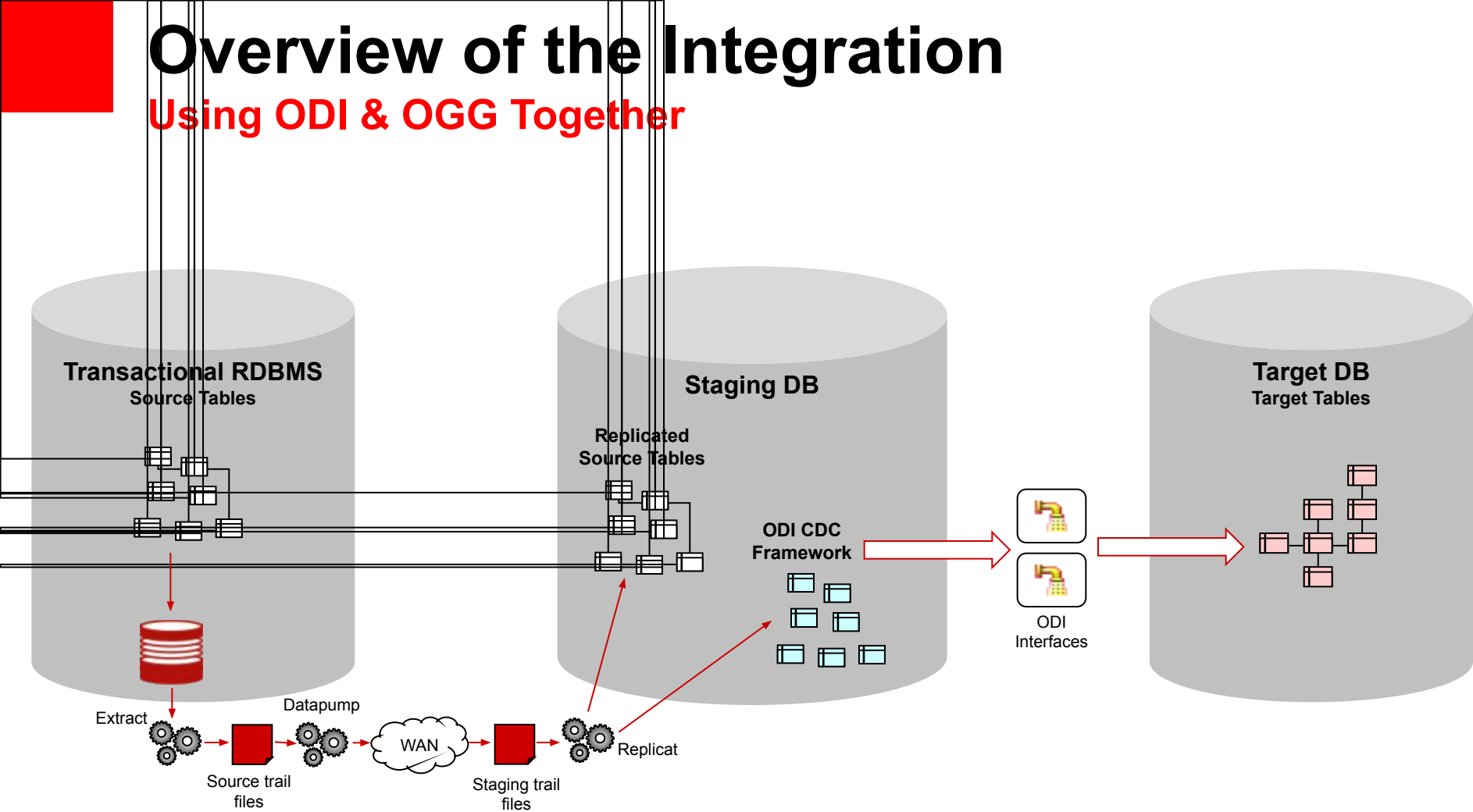
ORACLE

ODI CDC in a Nutshell
A General Framework for Change Capture on Source DBs

- Automatic w/JKMs
- Journal Tables
- Capture Services
  - Create Capture Process
  - Start/Stop Capture Process
- Subscription Services
  - Manage Consistency Sets
  - Register/Un-register Subscriber
- Consumption Services
  - Consumption Views
  - Consumption Operations
    - Extend Window
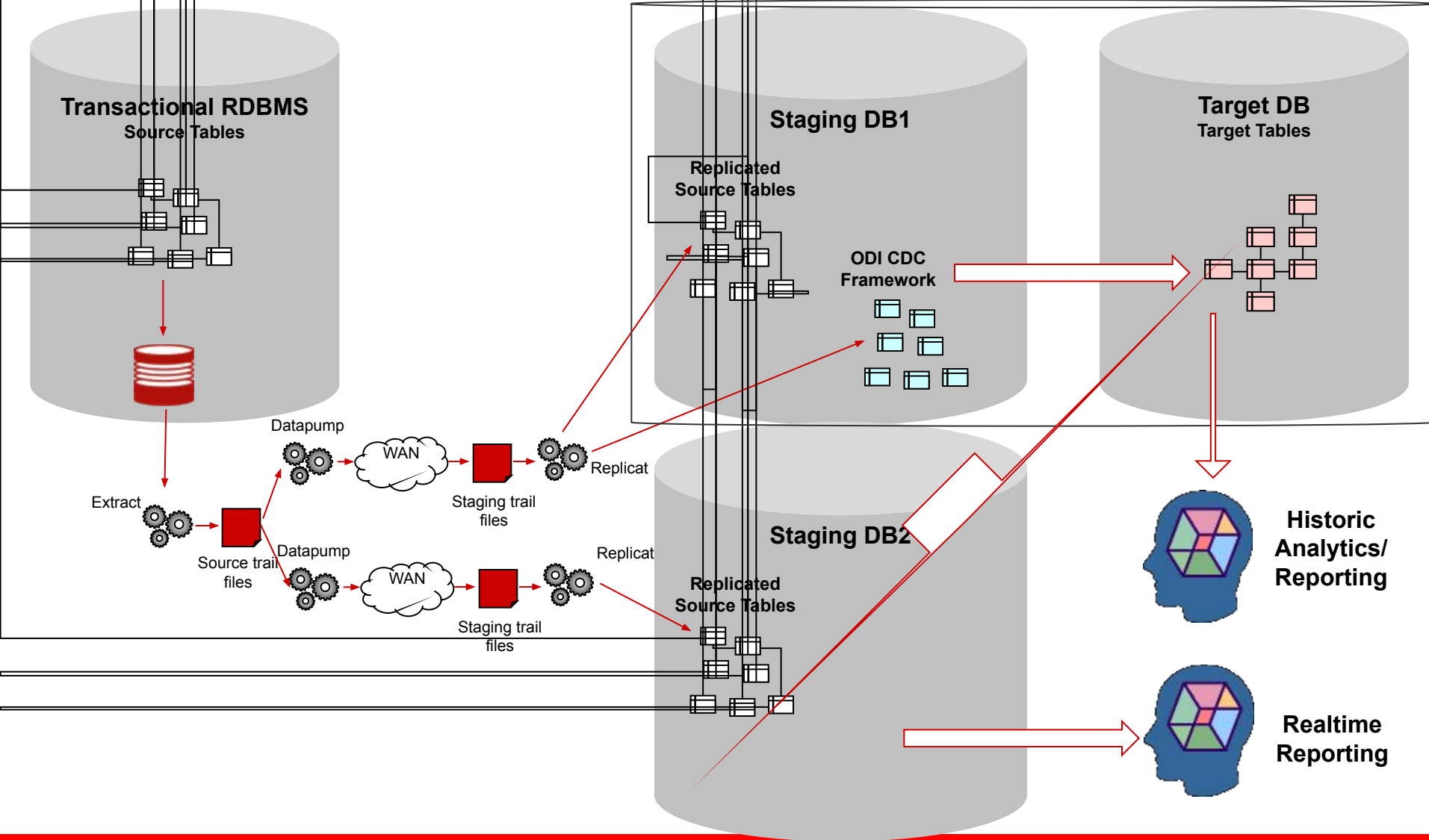   - Lock/Unlock Subscriber
  - Purge Operations

Consumers
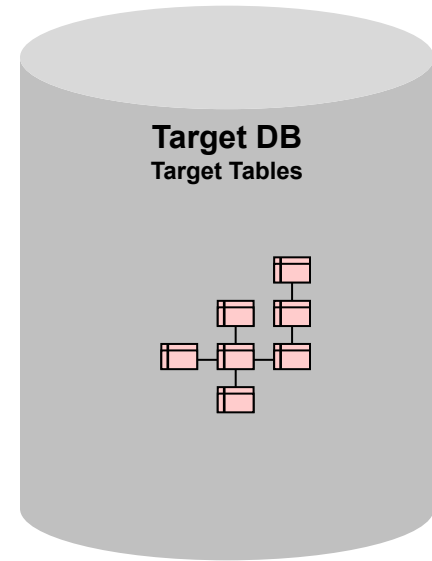
Consumption Services

ExtendWindow()   PurgeJournals()

LockSubscriber()   UnLockSubscriber()

View: CONT		
CONTID	CUSTNAME	CUSTID
I003	Thomas S.	C003

View: CUST		
CUSTID	CUSTNAME	EMPID
C002	YAHOO	E003

View: EMP	
EMPID	ENAME
E001	Joe Celko

Subscription Services

RegisterSubscriber()

AddTableToConsistencySet()

CDC_SET_SUBSCRIBER			
CDC_SET	SUBSCRIBER	MIN_WID	MAX_WID
CDC000	FUSION_BI	10	10
CDC000	PILLAR_HCM	9	9

Journal Tables

J$CONT	
CONTID	WID
I002	-
I003	20

J$CUST	
CUSTID	WID
C002	20

J$EMP	
EMPID	WID
E001	-
E001	20
E001	19

Capture Services

Capture Process   Capture Process   Capture Process

Source Data

Table: CONT		
CONTID	CUSTNAME	CUSTID
I001	Vijay R.	C003
I002	Raghu M.	C002
I003	Thomas S.	C003

Table: CUST		
CUSTID	CUSTNAME	EMPID
C001	AT&T	E003
C002	YAHOO	E003
C003	GOOGLE	E002

Table: EMP	
EMPID	ENAME
E001	Joe Celko
E002	Albert Einstein
E003	John Doe

Overview of the Integration
Using ODI & OGG Together
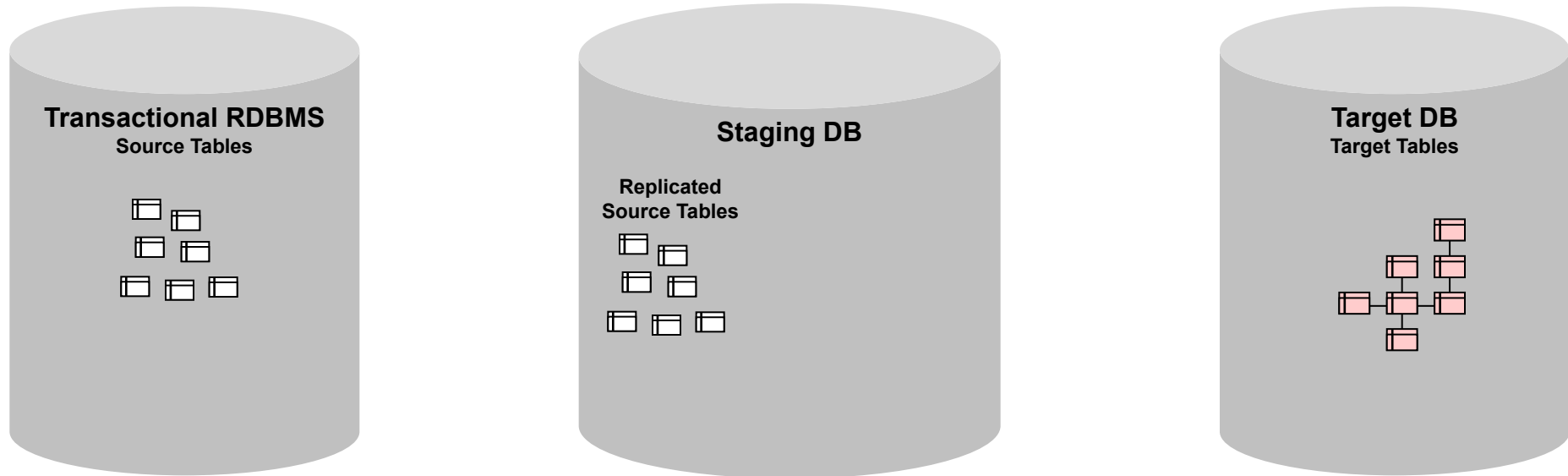

Transactional RDBMS
Source Tables

Staging DB

Replicated Source Tables

ODI CDC Framework

ODI Interfaces

Target DB
Target Tables

Extract

Datapump

WAN

Replicat

Source trail files

Staging trail files

ORACLE

Scenario: Analytics & Reporting

Transactional RDBMS
Source Tables

Staging DB1

Target DB
Target Tables

Replicated Source Tables

ODI CDC Framework

Datapump

WAN

Replicat

Staging trail files

Extract

Source trail files

Datapump

WAN

Replicat

Staging trail files

Staging DB2

Replicated Source Tables

Historic Analytics/ Reporting

Realtime Reporting

ORACLE

Step-by-step How It Works

Transactional RDBMS
Source Tables

Staging DB

Target DB
Target Tables

1. Replicated tables created with Common Format Designer (using ODI)

Transactional RDBMS
Source Tables

Staging DB

Replicated
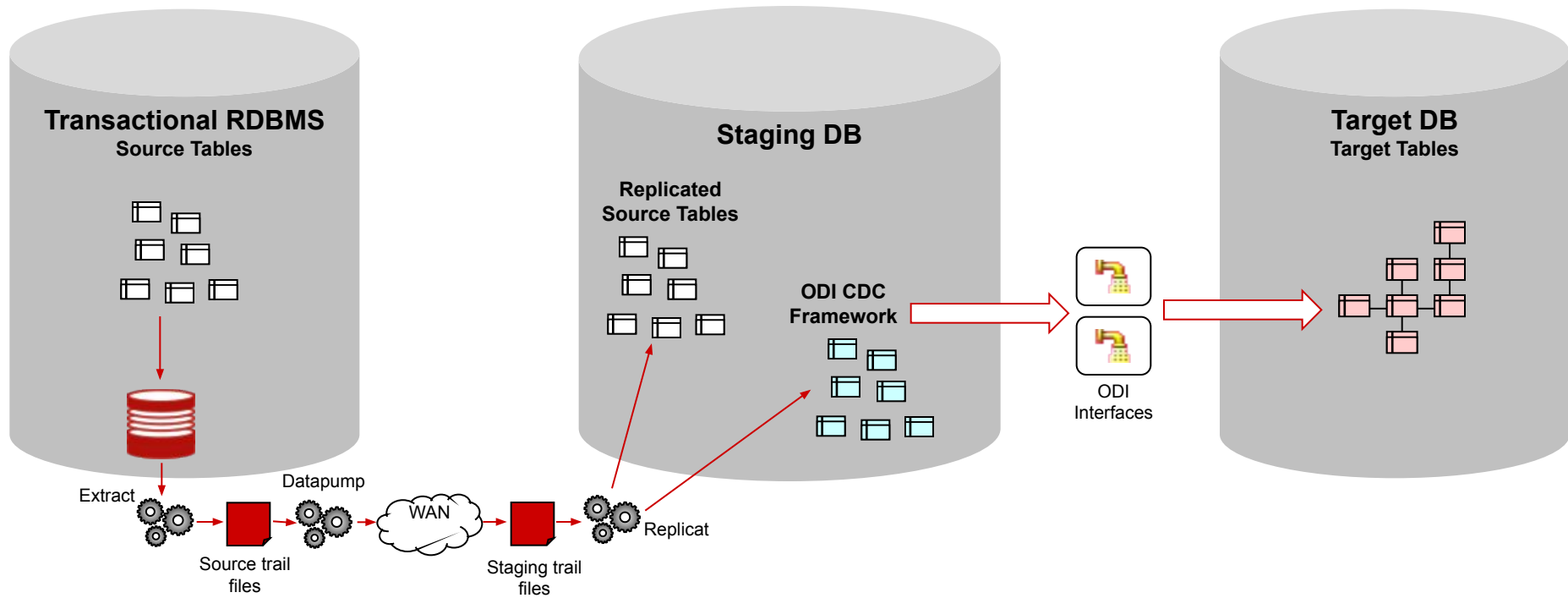Source Tables

Target DB
Target Tables

2. Start Capturing Changed Data in Source (OGG Extract process)
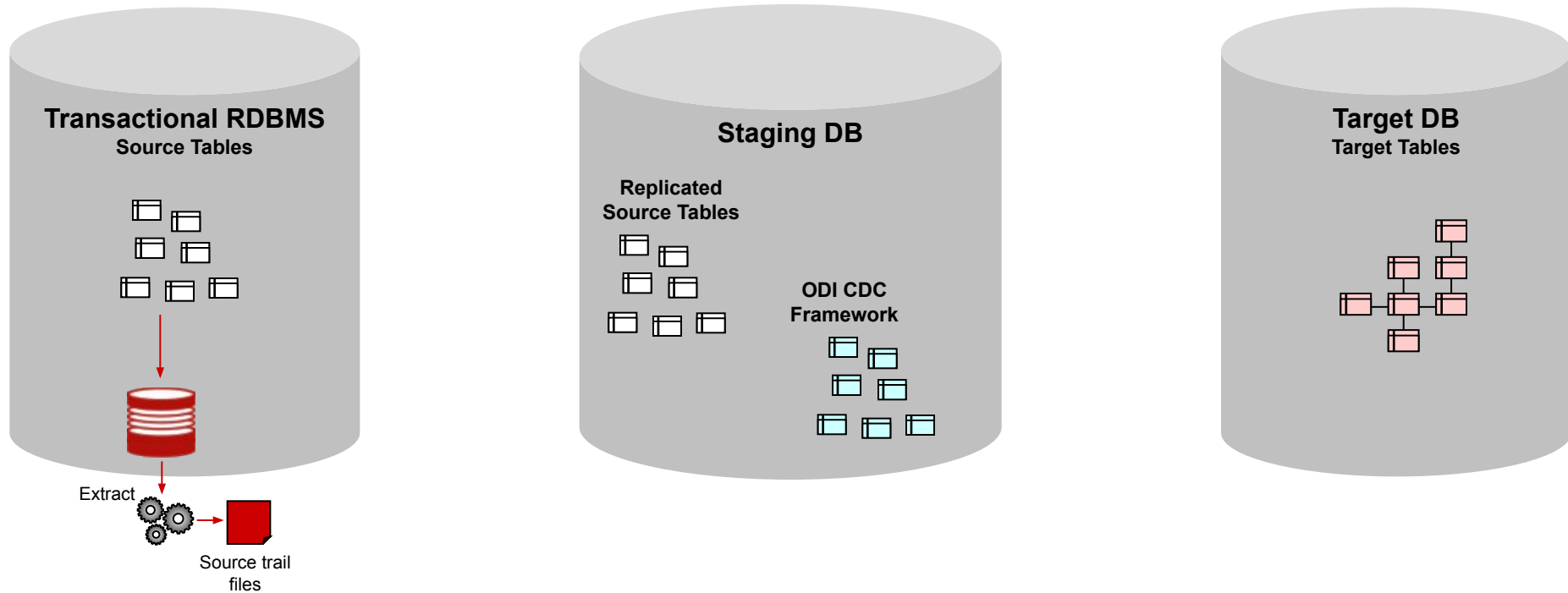
3. Initialize Staging and Target Data (with ODI or optionally, OGG)
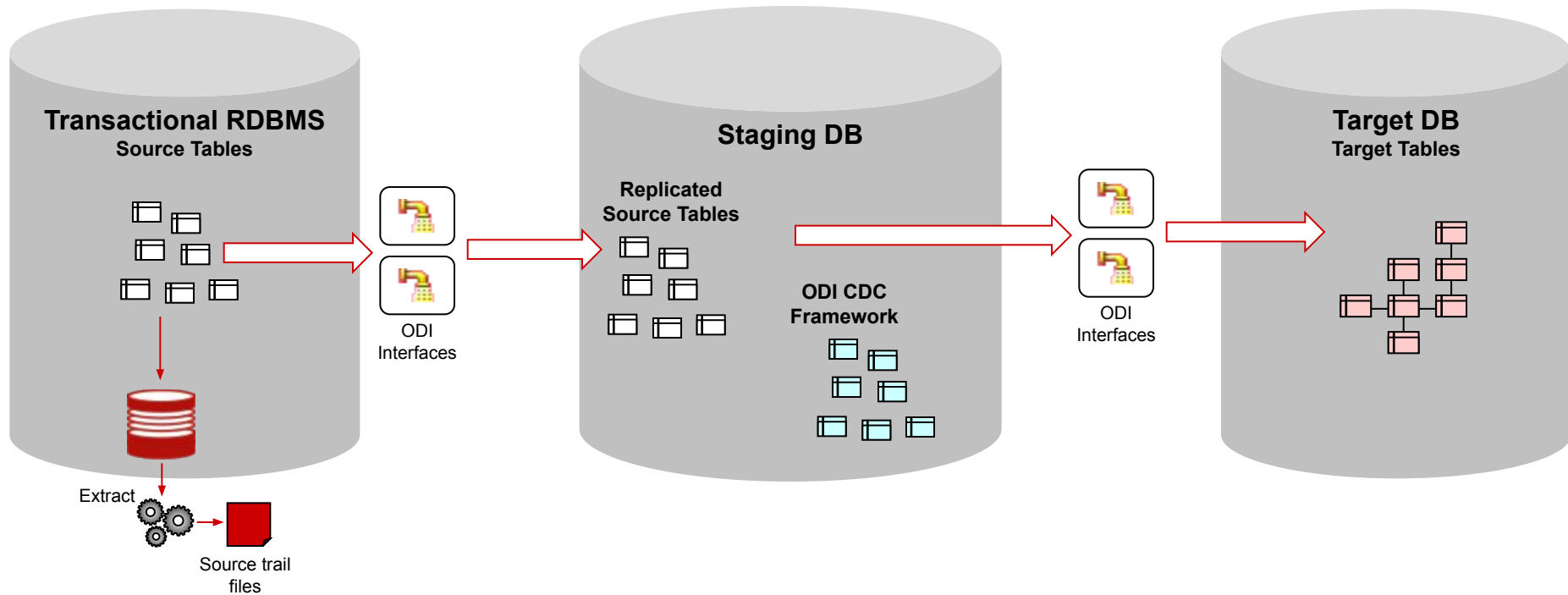
4. Start Replication / Propagate Changes to Target DB (OGG and ODI)

Simplified & Productive Development
DEMONSTRATION