ORACLE

Next Generation Analytics

Oracle Database In-Memory and Oracle Converged Database

Tirthankar Lahiri

Senior Vice President

Data and In-Memory Technologies

In-Memory Computing Summit 2020

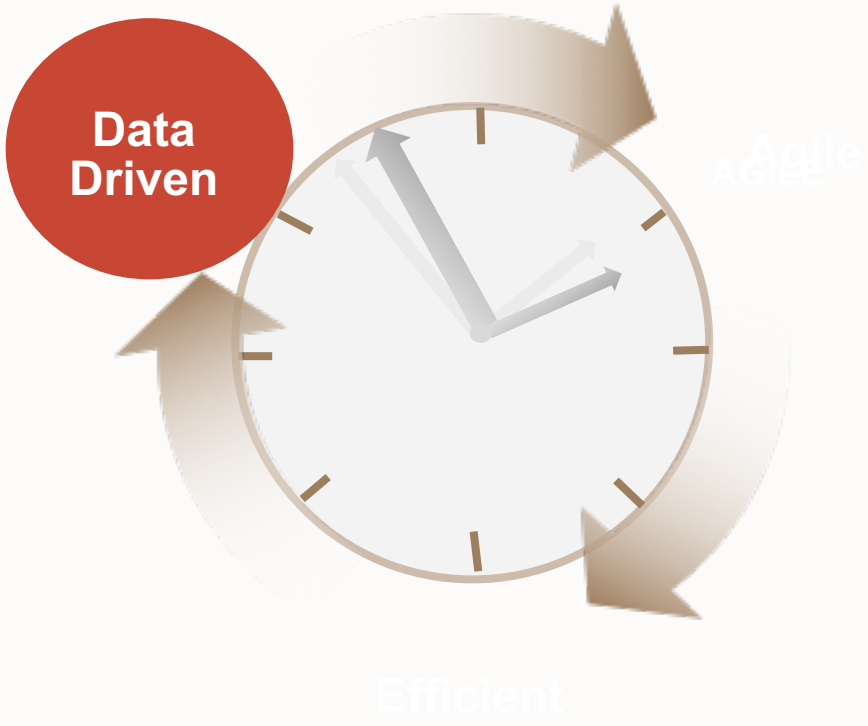
@dbinmemory

Safe harbor statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

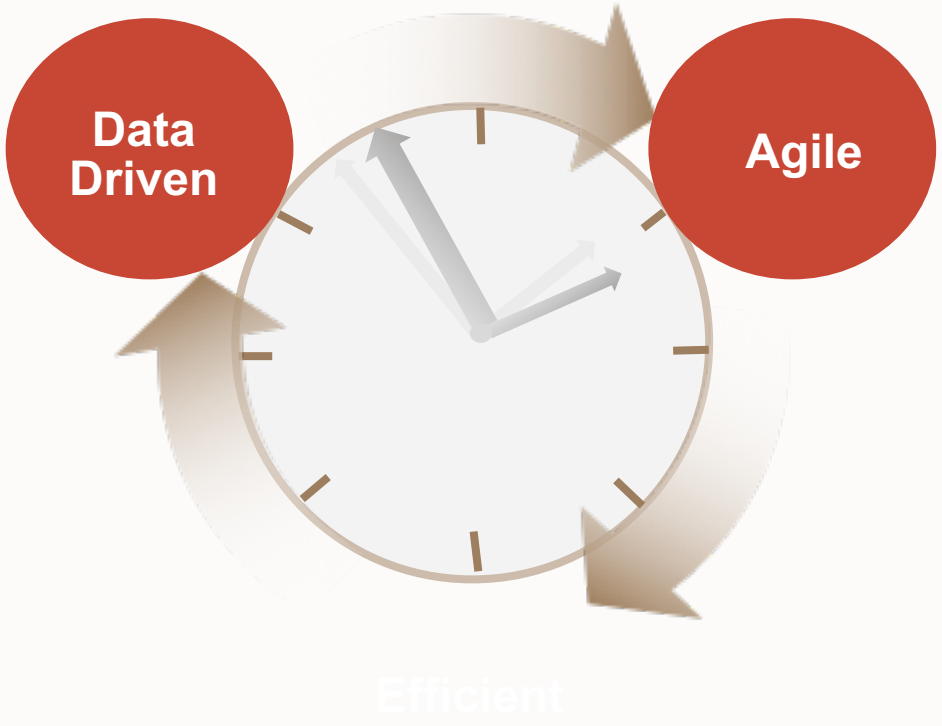
What is a Real-Time Enterprise?

What is a Real-Time Enterprise?



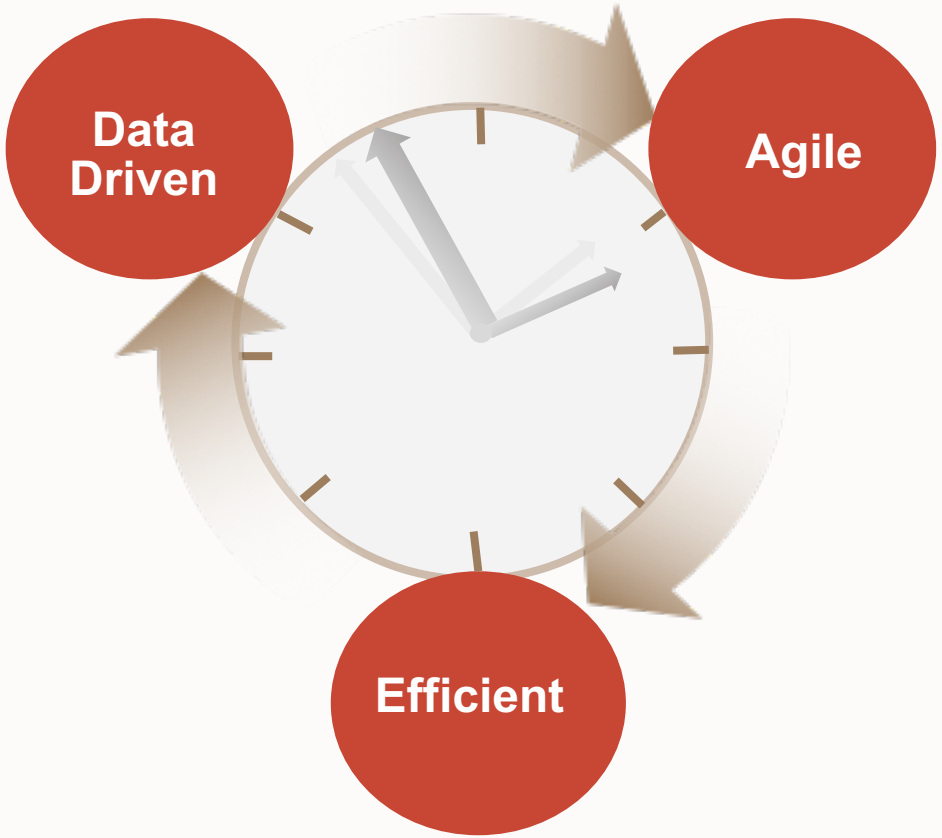
- **Data-Driven:** Uses metrics and calculations based on actual data

What is a Real-Time Enterprise?



- **Data-Driven:** Uses metrics and calculations based on actual data
- **Agile:** Rapidly adapts to changes in observed data

What is a Real-Time Enterprise?



- **Data-Driven:** Uses metrics and calculations based on actual data
- **Agile:** Rapidly adapts to changes in observed data
- **Efficient:** Continuously optimizes business processes to maximize revenue and profit

Real-Time Enterprise Examples

INSURANCE

Improve portfolios and reduce cost with real-time analytics for pricing

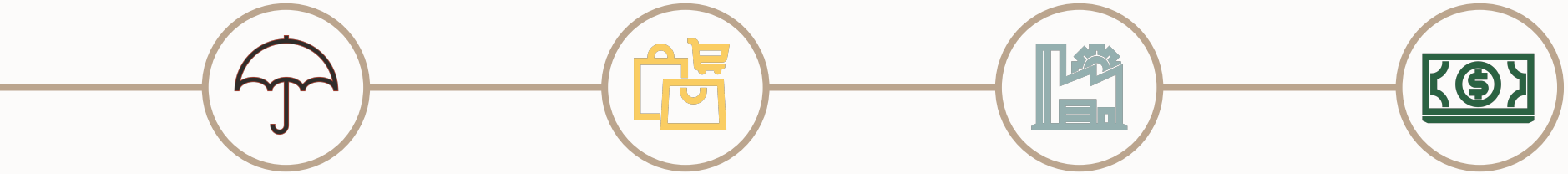
Real-Time Enterprise Examples

INSURANCE	RETAILERS
Improve portfolios and reduce cost with real-time analytics for pricing	Use location-based analytics to send personalized mobile coupons to customers

Real-Time Enterprise Examples

INSURANCE	RETAILERS	MANUFACTURING
Improve portfolios and reduce cost with real-time analytics for pricing	Use location-based analytics to send personalized mobile coupons to customers	Use real-time analytics to monitor quality and adjust parameters

Real-Time Enterprise Examples



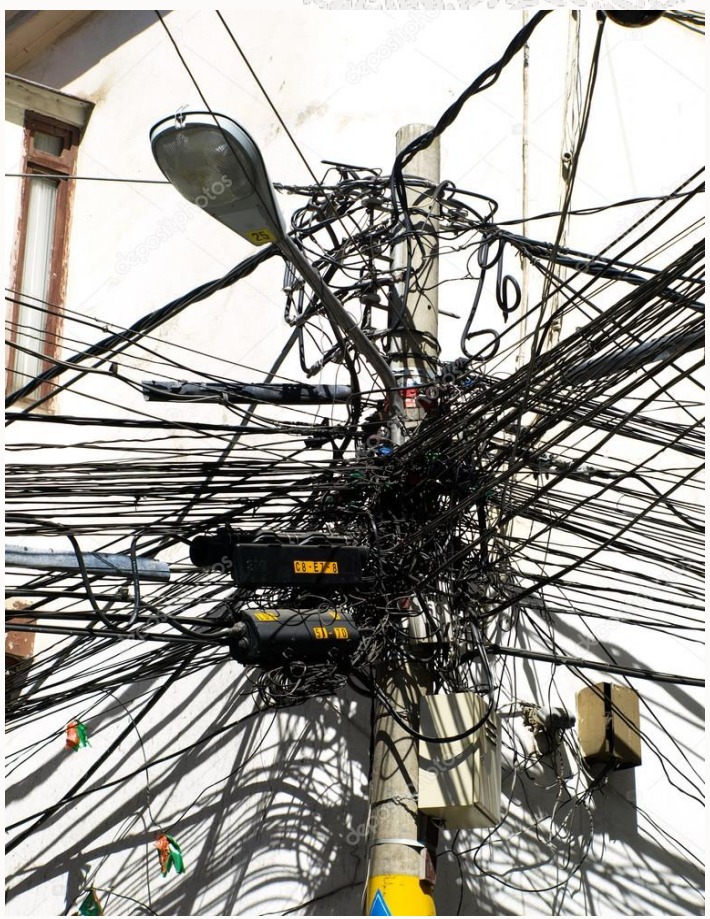
INSURANCE	RETAILERS	MANUFACTURING	FINTECH
Improve portfolios and reduce cost with real-time analytics for pricing	Use location-based analytics to send personalized mobile coupons to customers	Use real-time analytics to monitor quality and adjust parameters	Perform risk/fraud analysis across channels in real-time, not after the event occurs

Real-Time Enterprise Examples

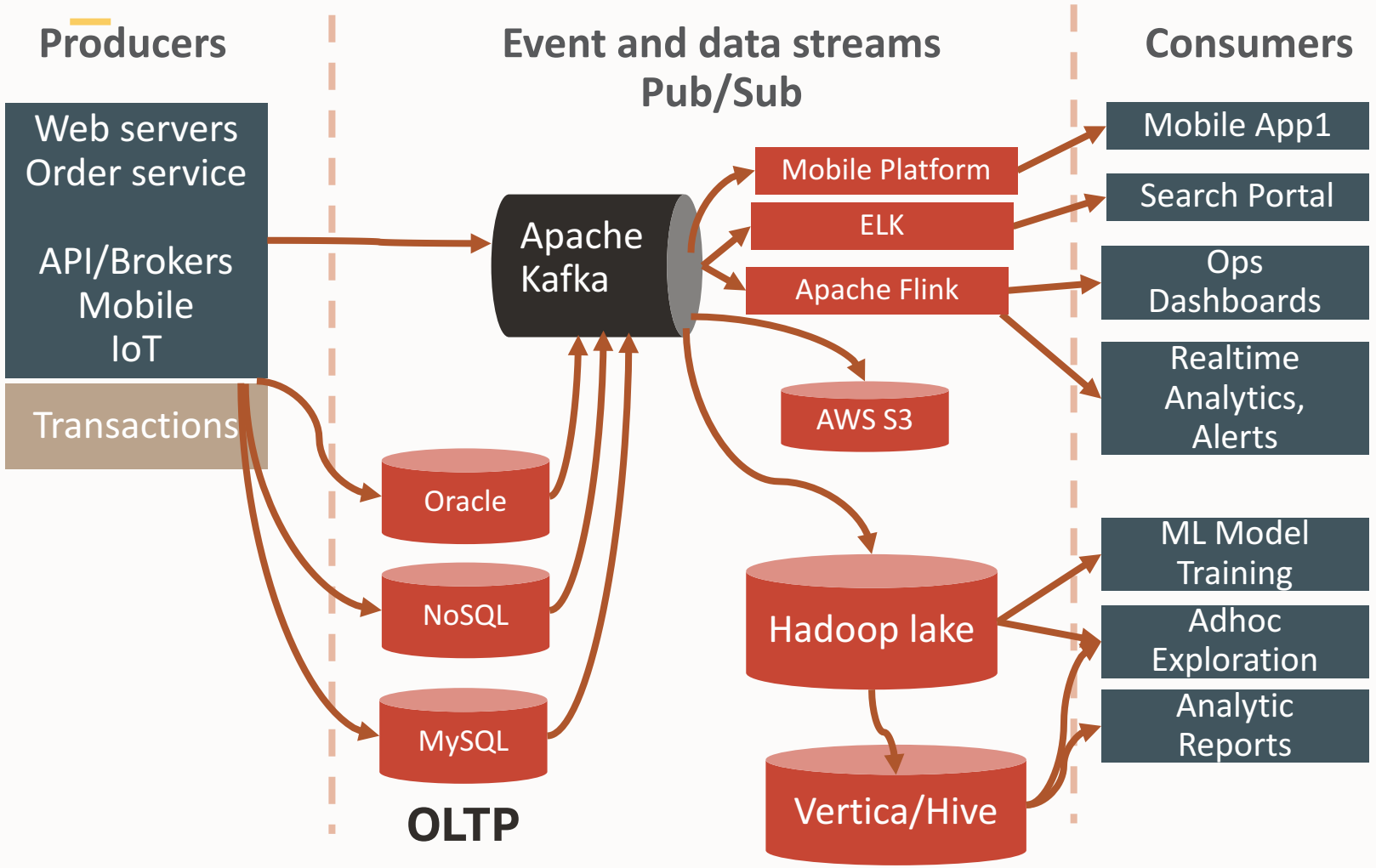
INSURANCE	RETAILERS	MANUFACTURING	FINTECH	TELECOM
Improve portfolios and reduce cost with real-time analytics for pricing	Use location-based analytics to send personalized mobile coupons to customers	Use real-time analytics to monitor quality and adjust parameters	Perform risk/fraud analysis across channels in real-time, not after the event occurs	Use real-time congestion metrics to optimize their networks

The Enemy of the Real Time Enterprise: Complexity

- Multiple systems, different technologies, different implementations, etc.
- Hurts security, administration, diagnosability
- Disrupts agility and efficiency



Real-World Macro-Complexity: Die By DIY

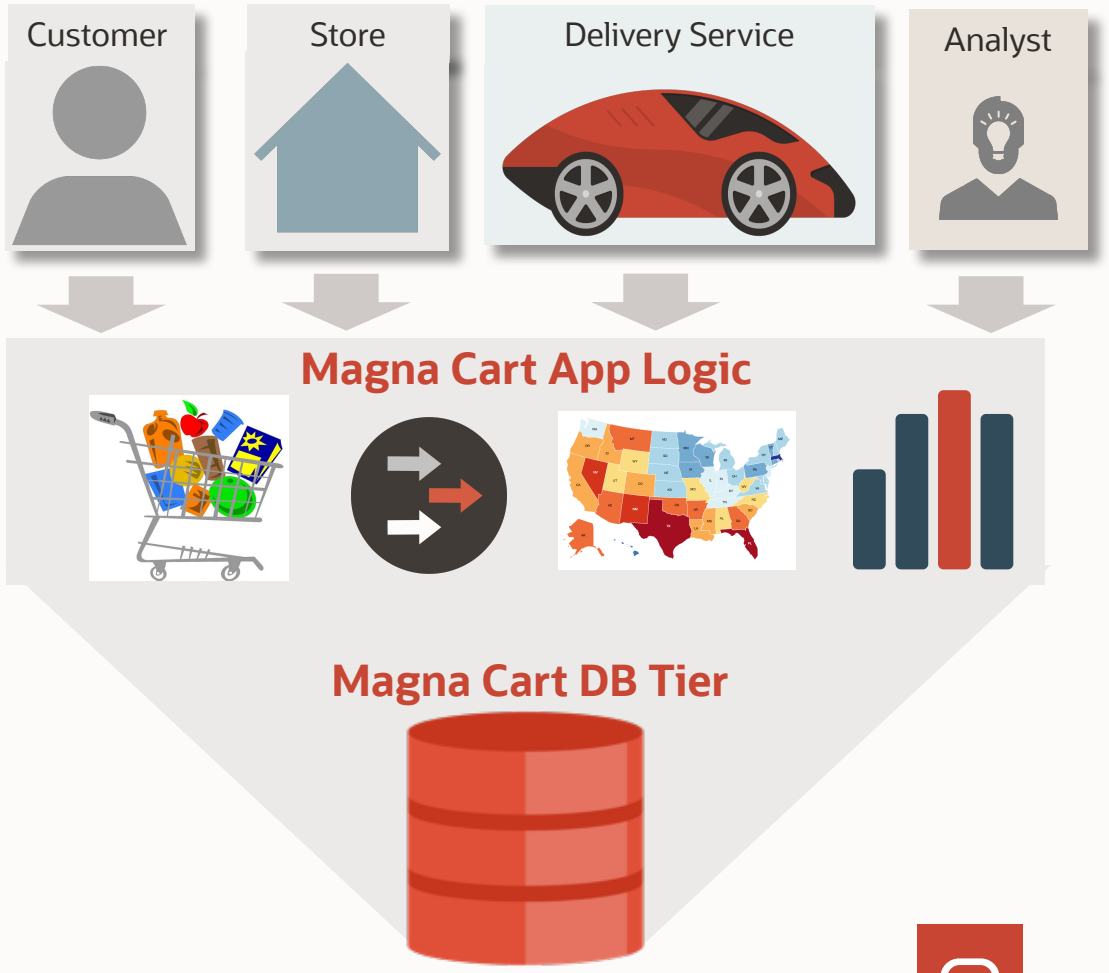


Macro-Complexity

- Multiple technologies
- Multiple data stores
- Data copied multiple times to do analytics
- Compromises security
- Compromises data consistency
- Complex to maintain
- Need highly skilled developers to build & keep running

Example of Application Evolution: Magna Cart

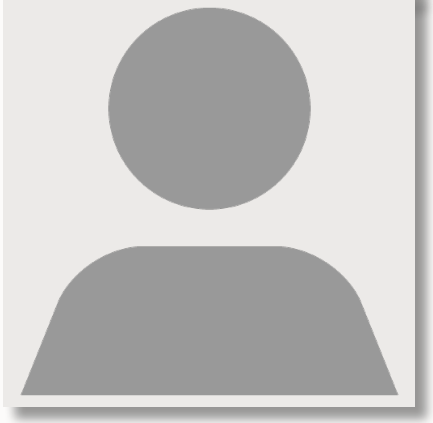
- Online Delivery App
  - Customers: place orders
  - Stores: process orders
  - Delivery Service: delivers orders
  - Analyst: Checks business metrics
- App logic includes:
  - Transactions for Orders and Delivery
  - Analytics for instant decisions and for continuous business optimization



Magna Cart: Transactional (OLTP) Workload

Browse Items,
Add Items to Order,
Place Order, …

Customer



Process Order,
Ship Order,
Restock Products, …

Store

Schedule Delivery
Pick up from Store,
Deliver to Customer, …

Delivery Service

Define OLTP Tables and Relationships

CUSTOMERS
id
email
name
address

ORDERS
id
customer_id
order_datetime
store_id
order_status

ITEMS
id
order_id
Shipment_id
product_id
quantity

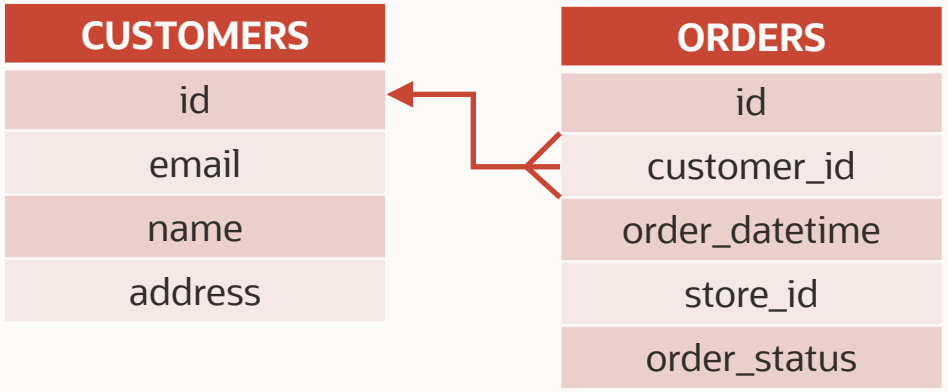
STORES
id
name
URL
address

DELIVERIES
id
store_id
customer_id
Delivery_address
Delivery_status

PRODUCTS
id
name
category
rating
unit_price

INVENTORY
store_id
product_id
product_inventory

Define OLTP Tables and Relationships



ITEMS
id
order_id
Shipment_id
product_id
quantity

DELIVERIES
id
store_id
customer_id
Delivery_address
Delivery_status

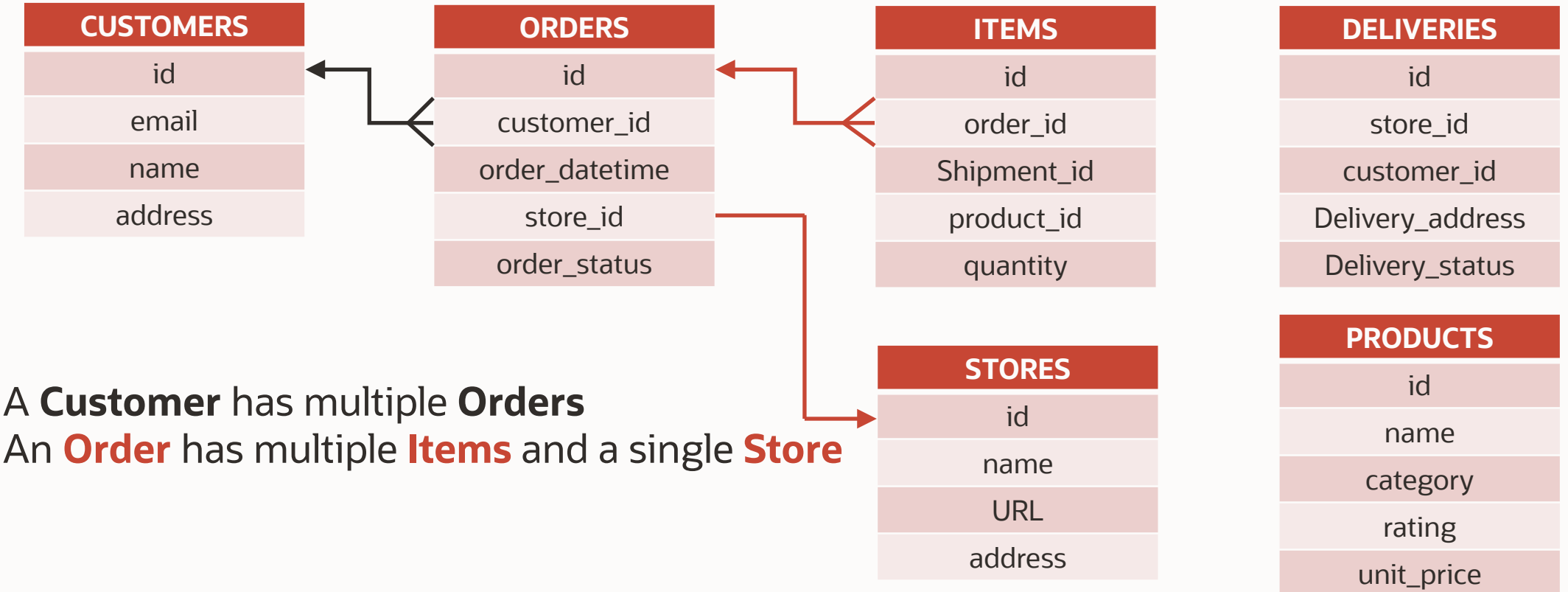
STORES
id
name
URL
address

PRODUCTS
id
name
category
rating
unit_price

INVENTORY
store_id
product_id
product_inventory

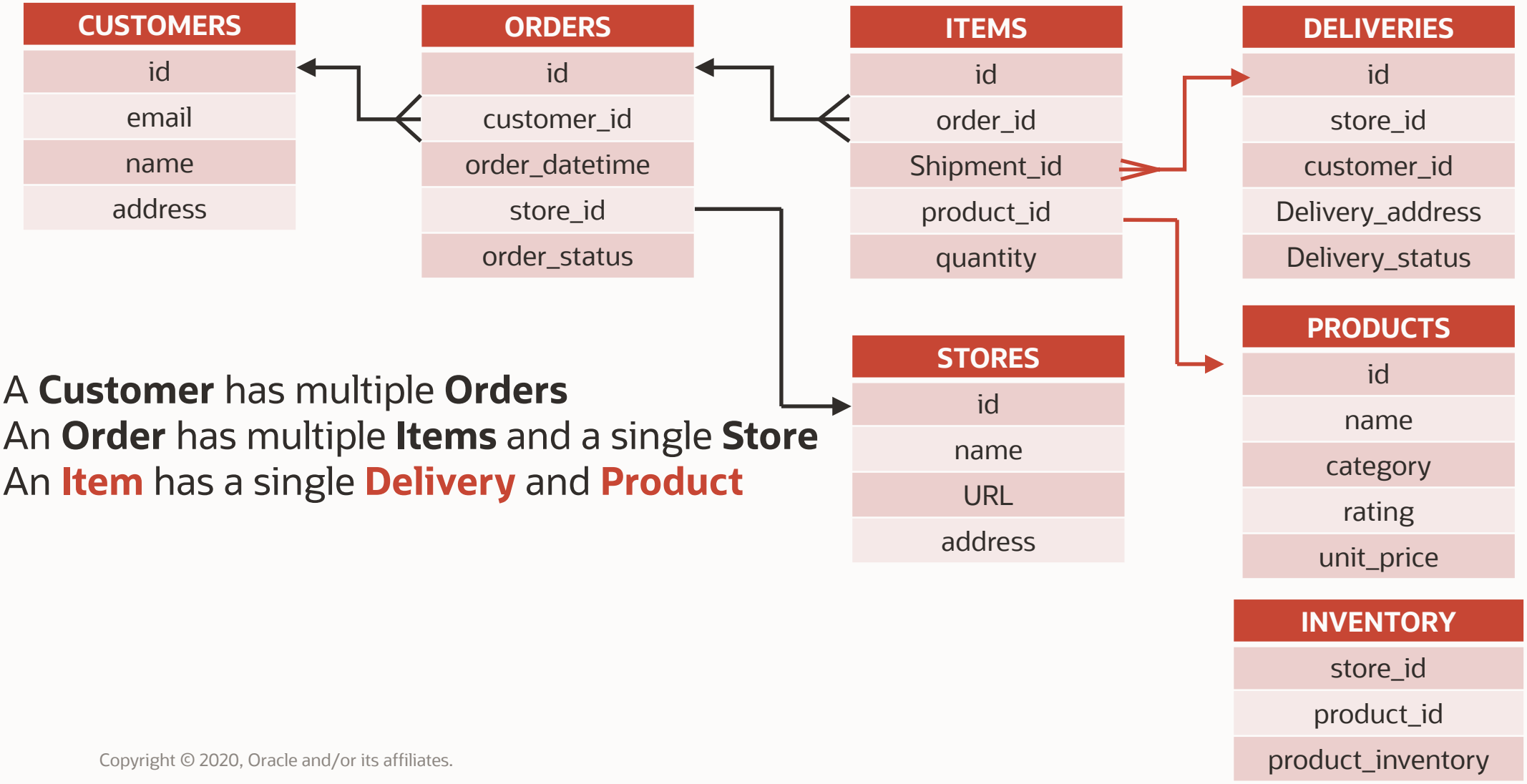
- A **Customer** has multiple **Orders**

Define OLTP Tables and Relationships



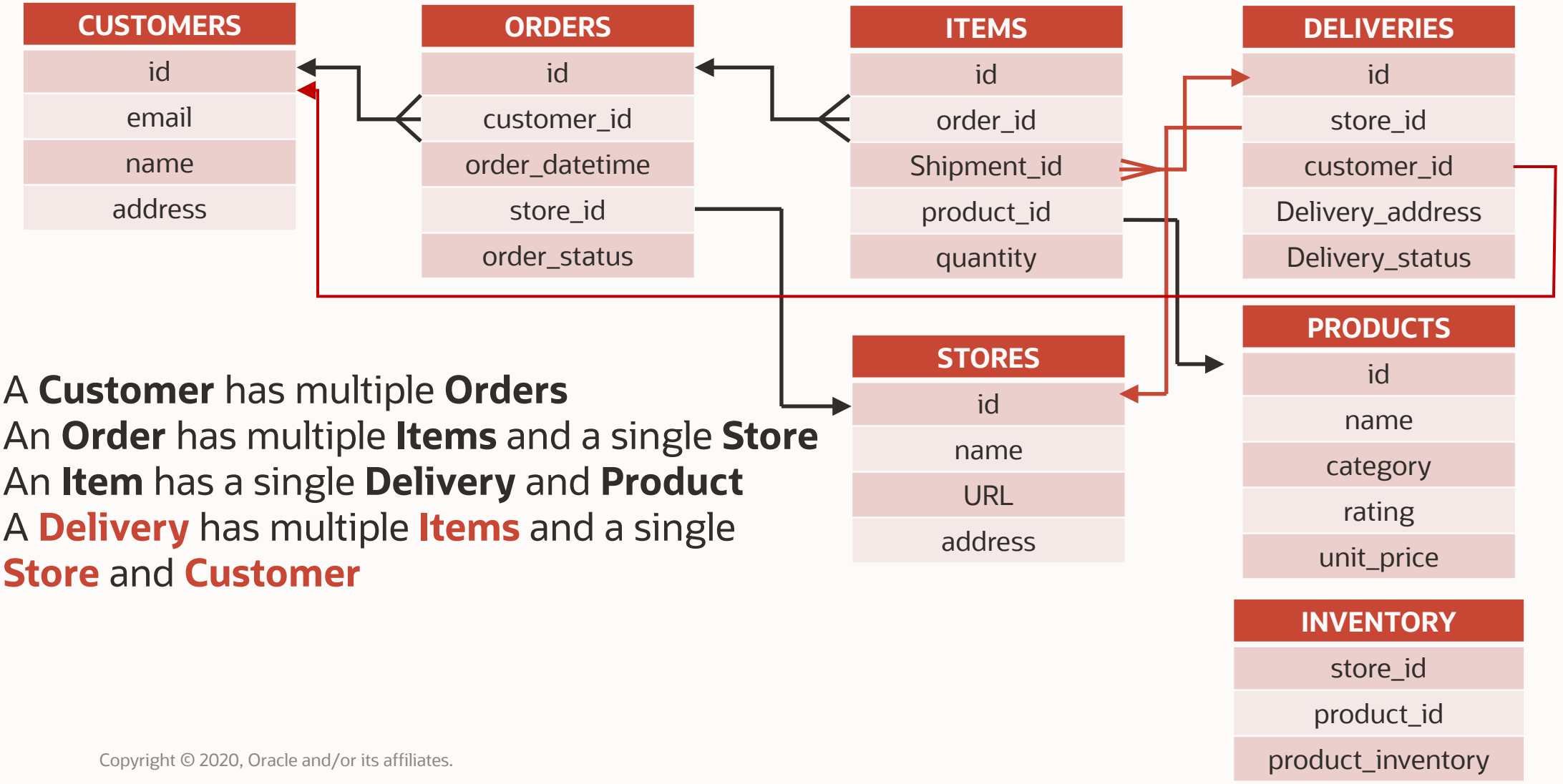
- A **Customer** has multiple **Orders**
- An **Order** has multiple **Items** and a single **Store**

Define OLTP Tables and Relationships



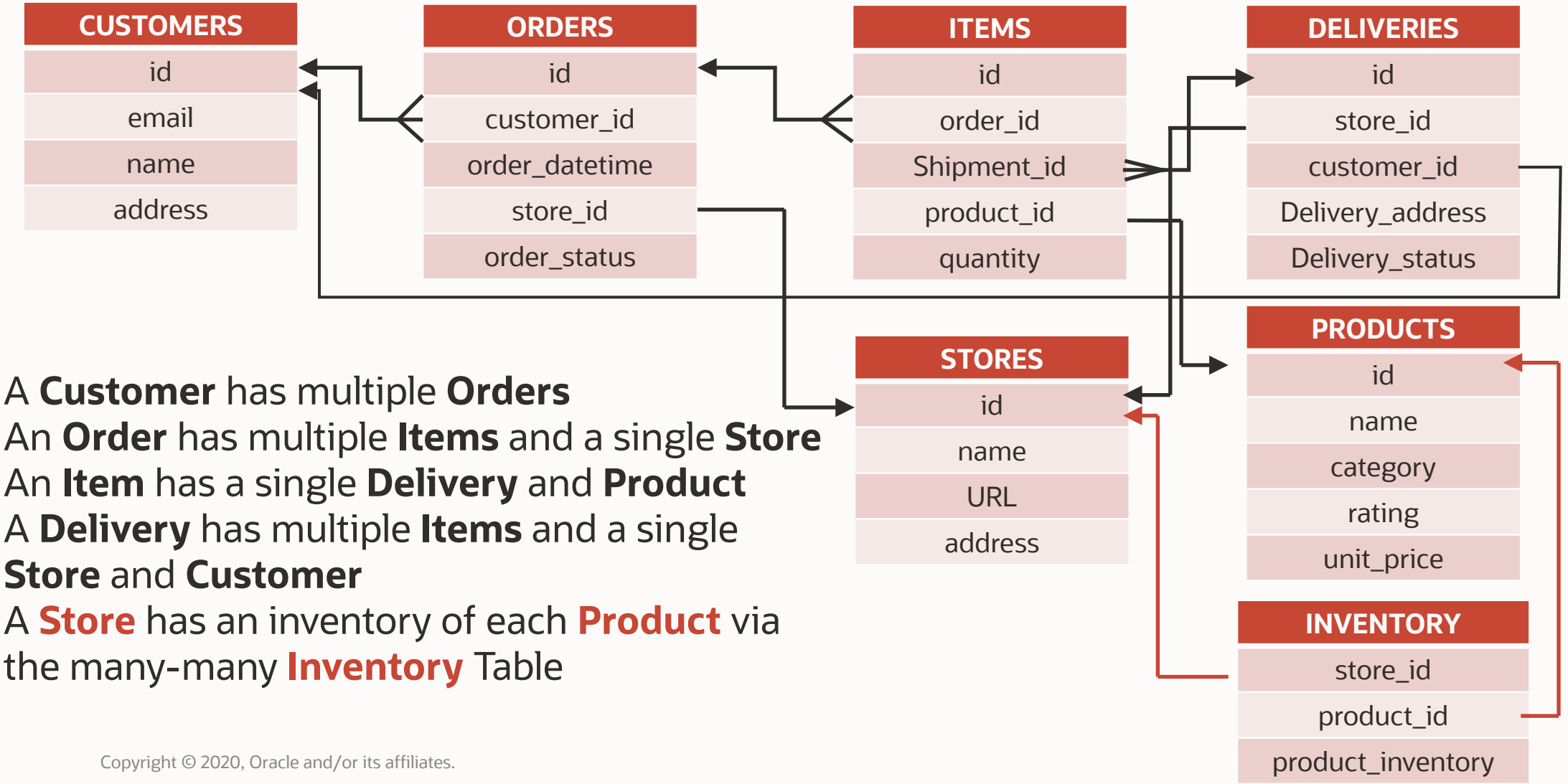
- A **Customer** has multiple **Orders**
- An **Order** has multiple **Items** and a single **Store**
- An **Item** has a single **Delivery** and **Product**

Define OLTP Tables and Relationships



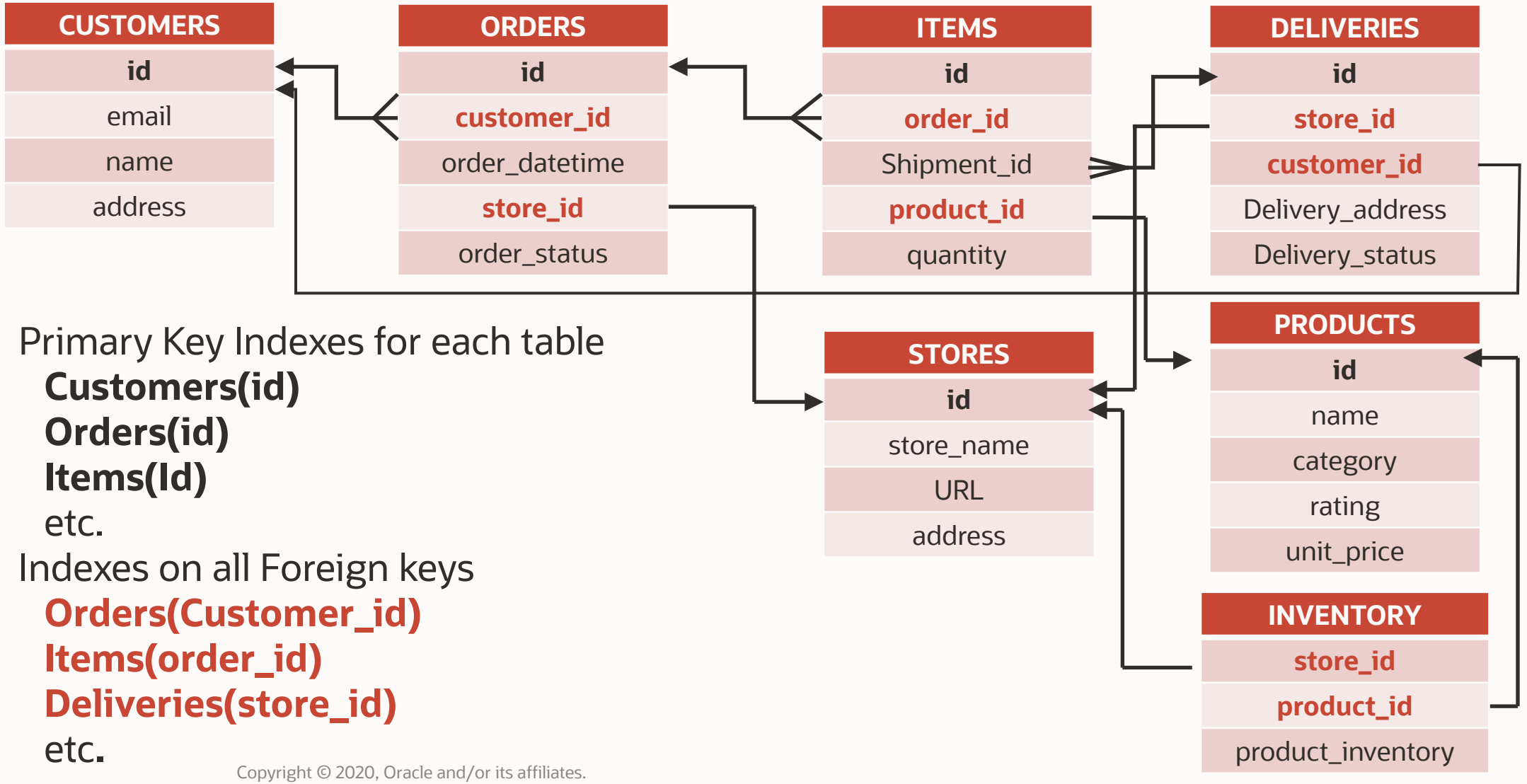
- A **Customer** has multiple **Orders**
- An **Order** has multiple **Items** and a single **Store**
- An **Item** has a single **Delivery** and **Product**
- A **Delivery** has multiple **Items** and a single **Store** and **Customer**

Define OLTP Tables and Relationships



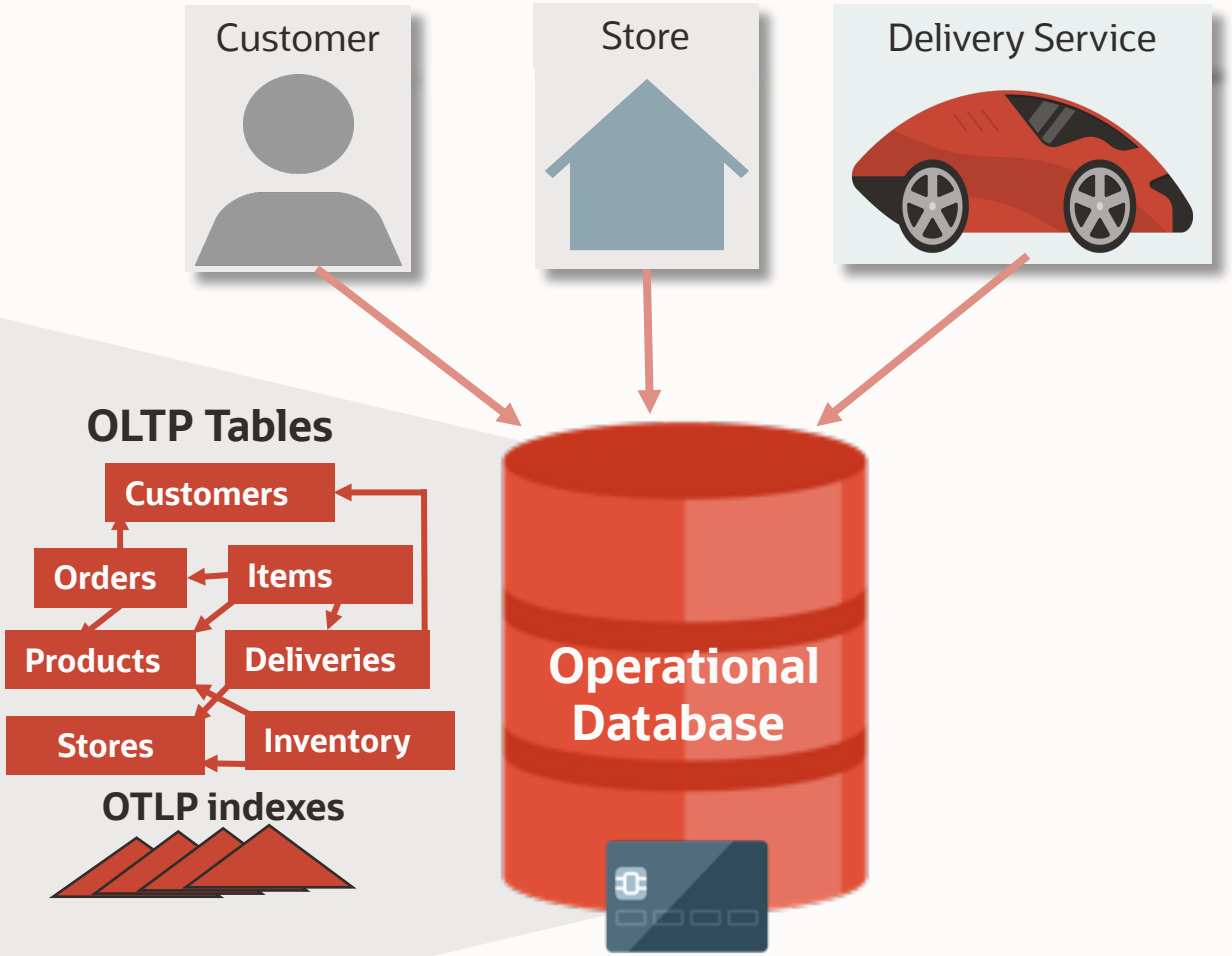
- A **Customer** has multiple **Orders**
- An **Order** has multiple **Items** and a single **Store**
- An **Item** has a single **Delivery** and **Product**
- A **Delivery** has multiple **Items** and a single **Store** and **Customer**
- A **Store** has an inventory of each **Product** via the many-many **Inventory** Table

Relationships Correspond to OLTP Indexes



- Primary Key Indexes for each table
  Customers(id)
  Orders(id)
  Items(Id)
  etc.
- Indexes on all Foreign keys
  Orders(Customer_id)
  Items(order_id)
  Deliveries(store_id)
  etc.

Magna Cart: Up and Running with Transactions!



- Single Database
- Simple Schema
- Simple Architecture
- Fast Transactions

Magna Cart: Real-Time Analytics

What is the top rated phone? Which stores carry it in my region? What other accessories do I need?

Customer

What other products and promotions should we offer a customer based on their current order items and location?

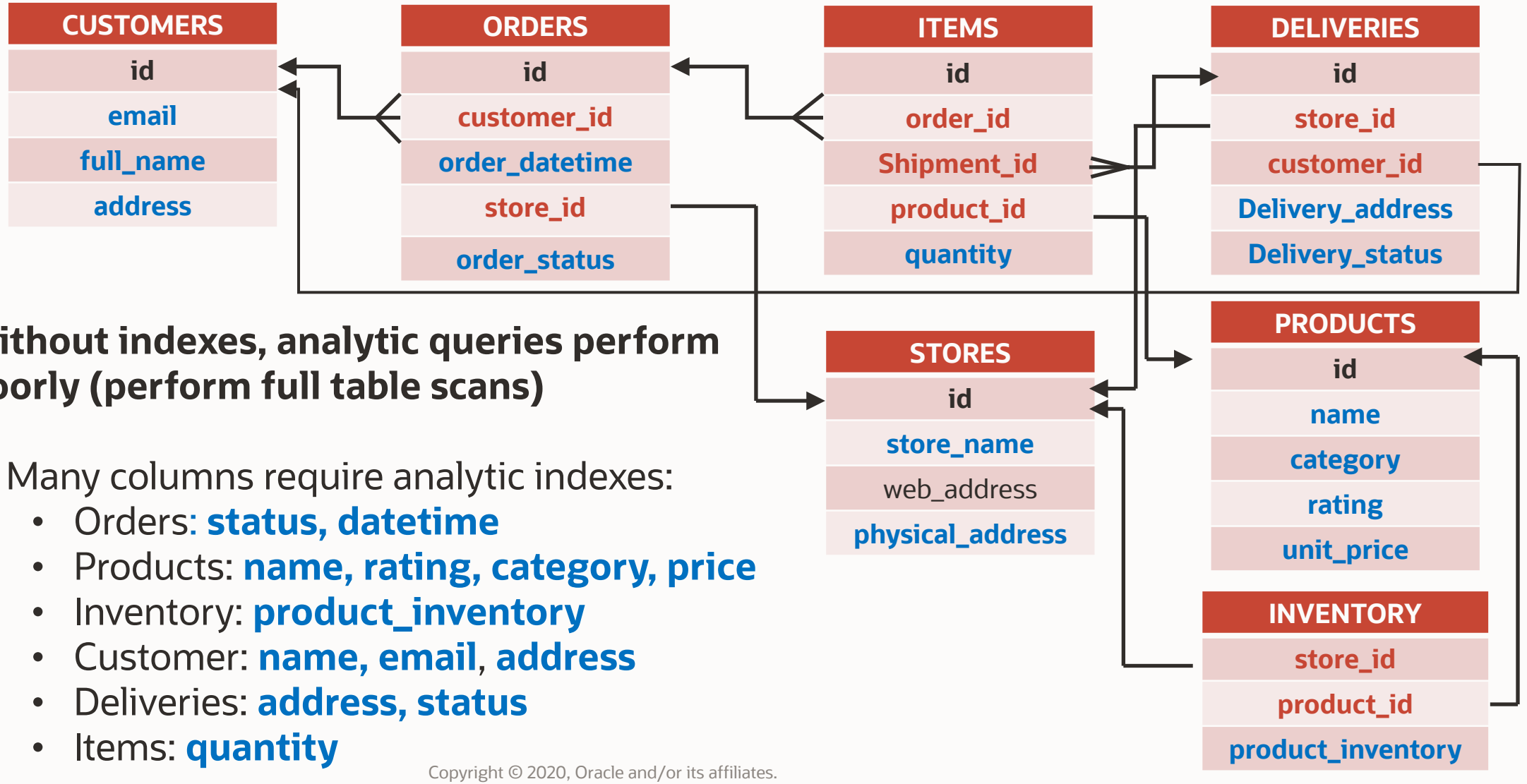
Store

What deliveries could be combined within a single trip? Which stores and customers have the most pending deliveries?

Delivery Service

These queries run slowly without analytic indexes

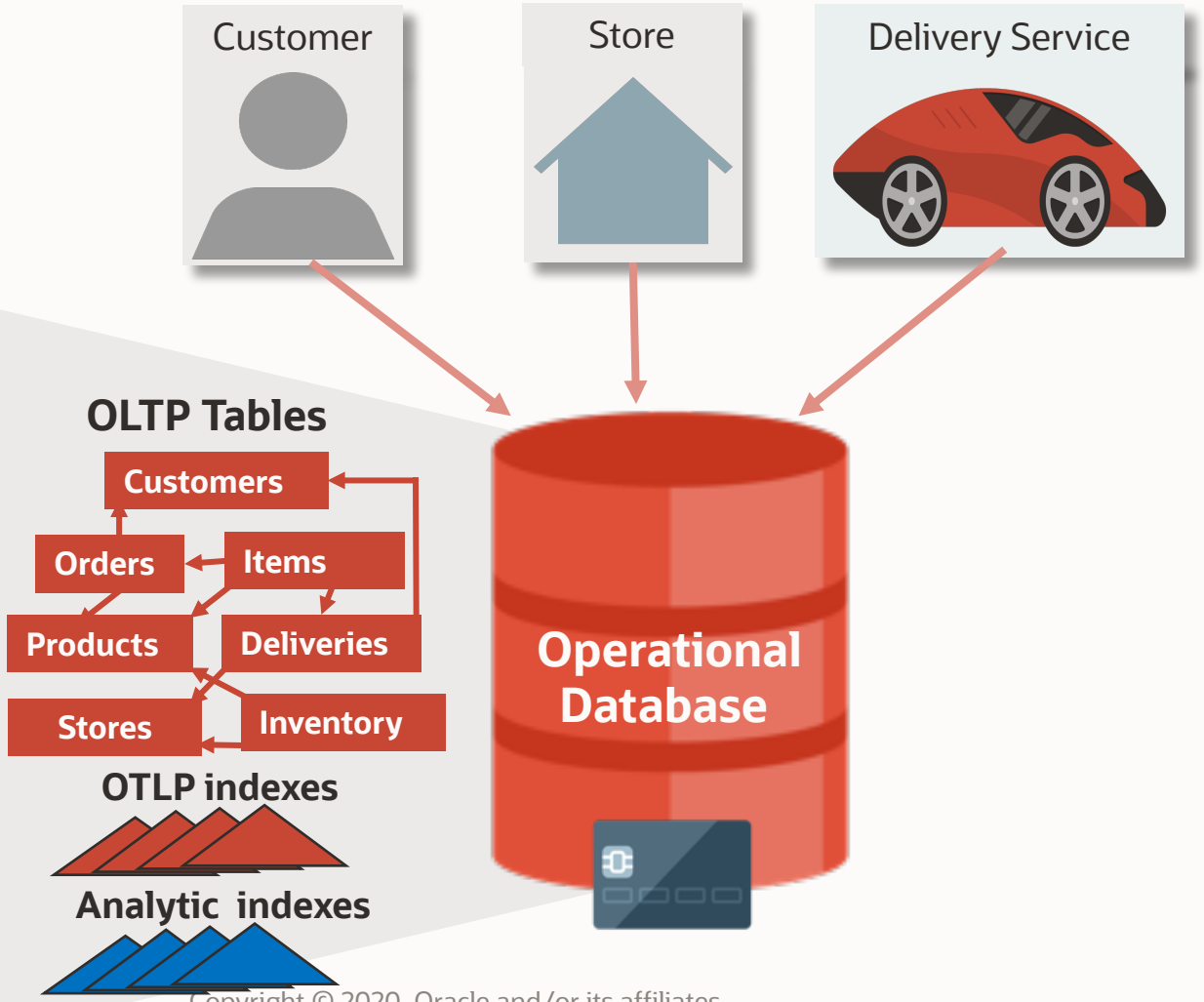
Define Indexes for Real-Time Analytics



Without indexes, analytic queries perform poorly (perform full table scans)

- Many columns require analytic indexes:
  - Orders: **status, datetime**
  - Products: **name, rating, category, price**
  - Inventory: **product_inventory**
  - Customer: **name, email**, **address**
  - Deliveries: **address, status**
  - Items: **quantity**

Magna Cart: Still Simple, But …



- Still a Single Database
- Still a Simple Schema
- But analytic indexes add overheads:
  - Slow down Transactions
  - Increase database size
  - Require knowledge of workload

Magna Cart: Longer Term Analytics

Which products give us our highest margins?

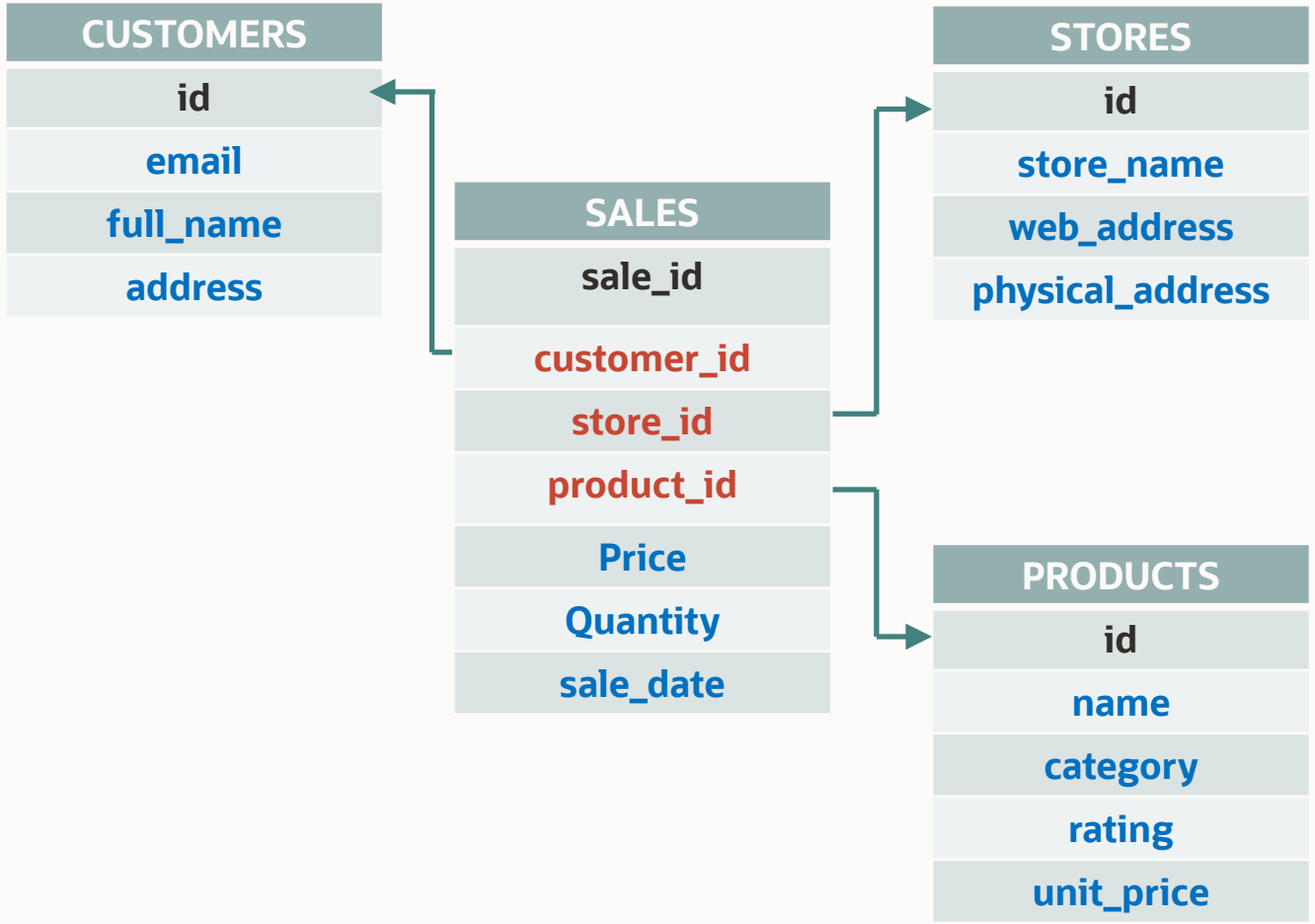
What are the top 10 stores in the north west region this month?

If we offer a 20% discount on flour, how much could our margins improve?

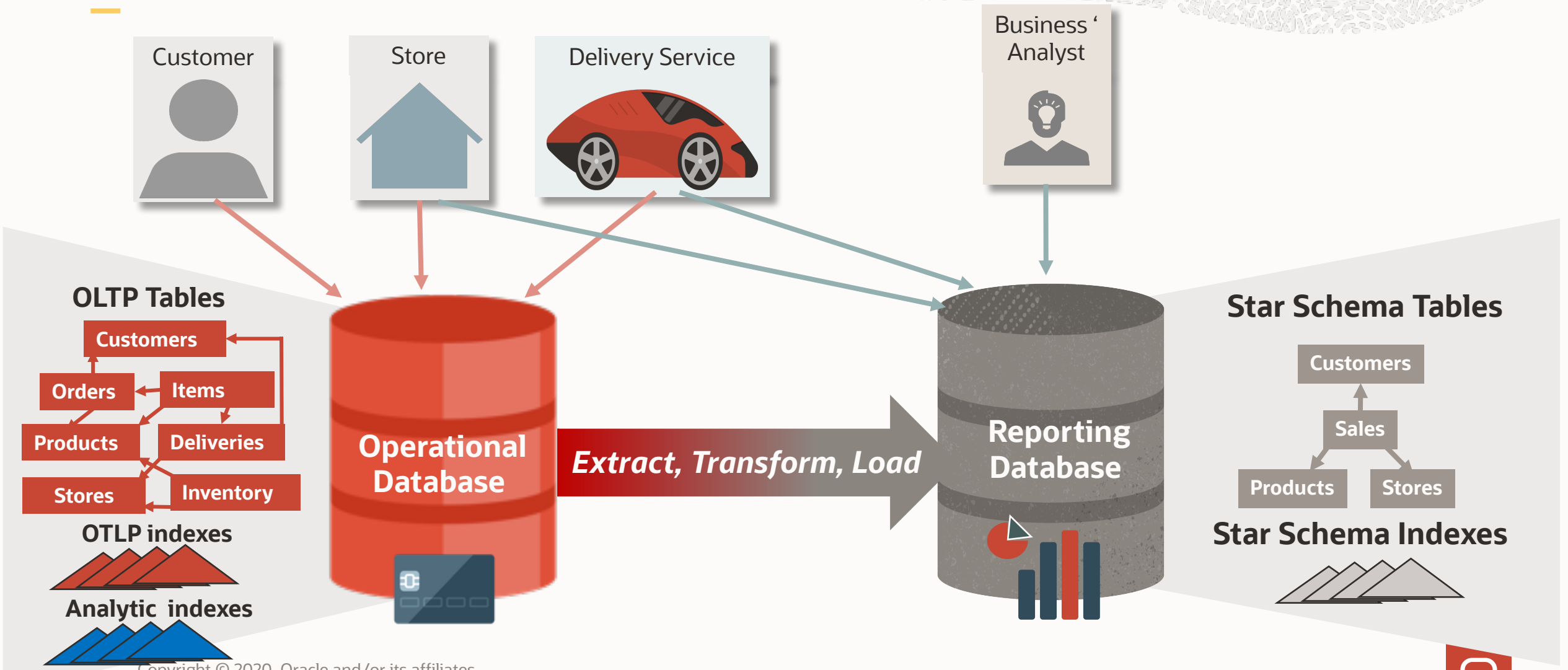
These queries typically access large amounts of historical data and are usually better offloaded to a Reporting Database

Define Reporting Database (Star) Schema

- Required for longer running queries on large historical data
- Maintains history of all sales, by customer, store, product
- Different schema from the OLTP database:
  - **Central “Fact" table:** SALES
  - **“Dimension” tables**: Customers, Products, Stores
  - Indexes on all key columns
  - Plus Analytic indexes on most columns



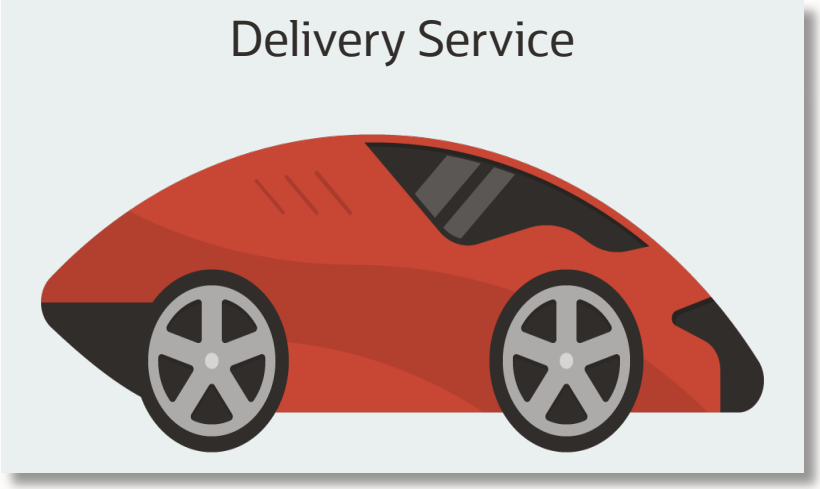
Magna Cart: Reporting Database *Doubles* Complexity



Magna Cart: Analytic "Dashboard" Reports

What is my YTD sales?
What is my YoY growth?
What are my top 10 products this month?

Store



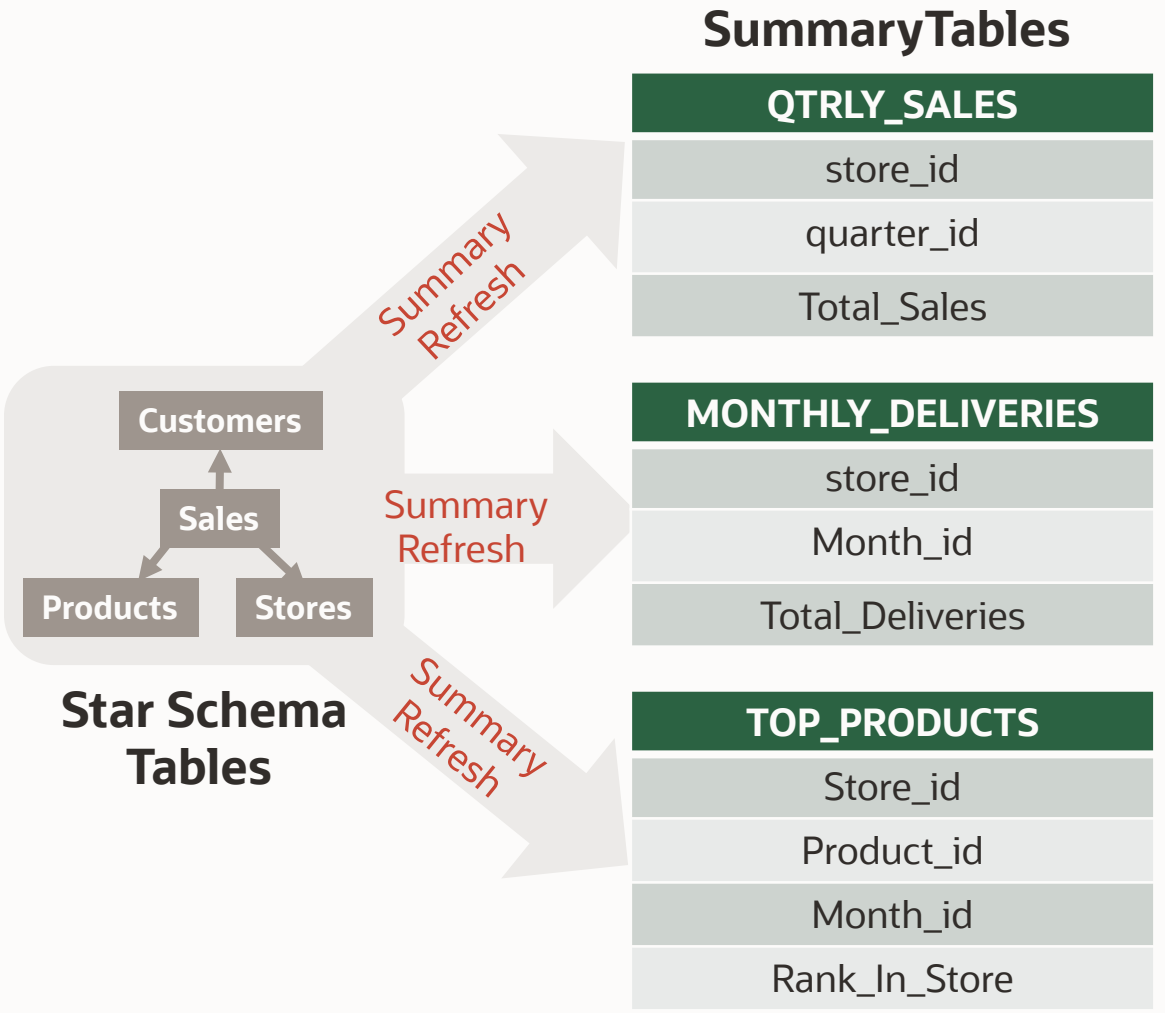
How many deliveries were made last week in the North East region?
How many customers were served last month?

Delivery Service

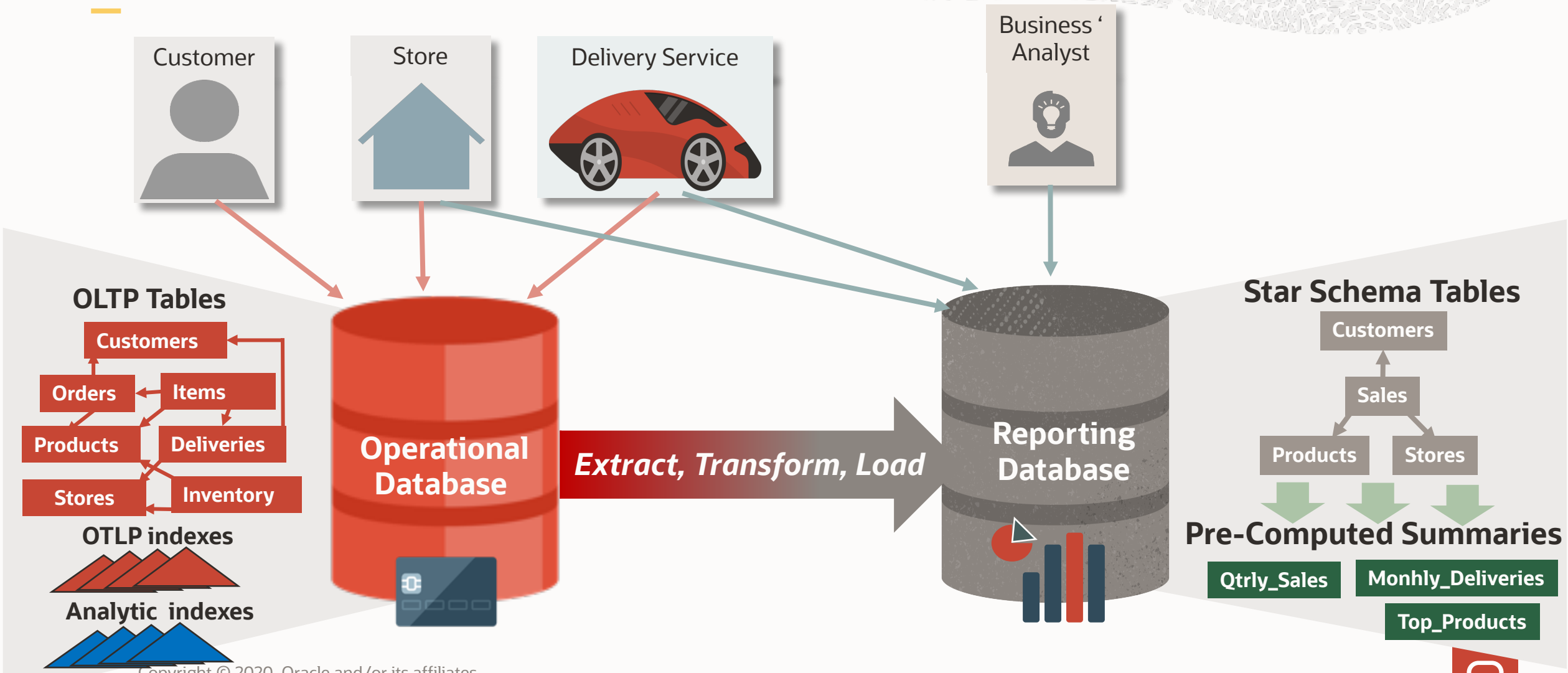
These queries are typically frequently run, resource intensive, and usually require pre-computed summaries for good performance

Define Pre-Computed Summaries

- Reports are of various types, from short term to long term
- Pre-computed summaries are often needed for frequent reports, e.g.
  - **Quarterly Sales by Store**
  - **Monthly Deliveries by Store**
  - **Top 10 Products in Store by Month**
- Typically maintained as Materialized Views that are periodically refreshed from the star schema

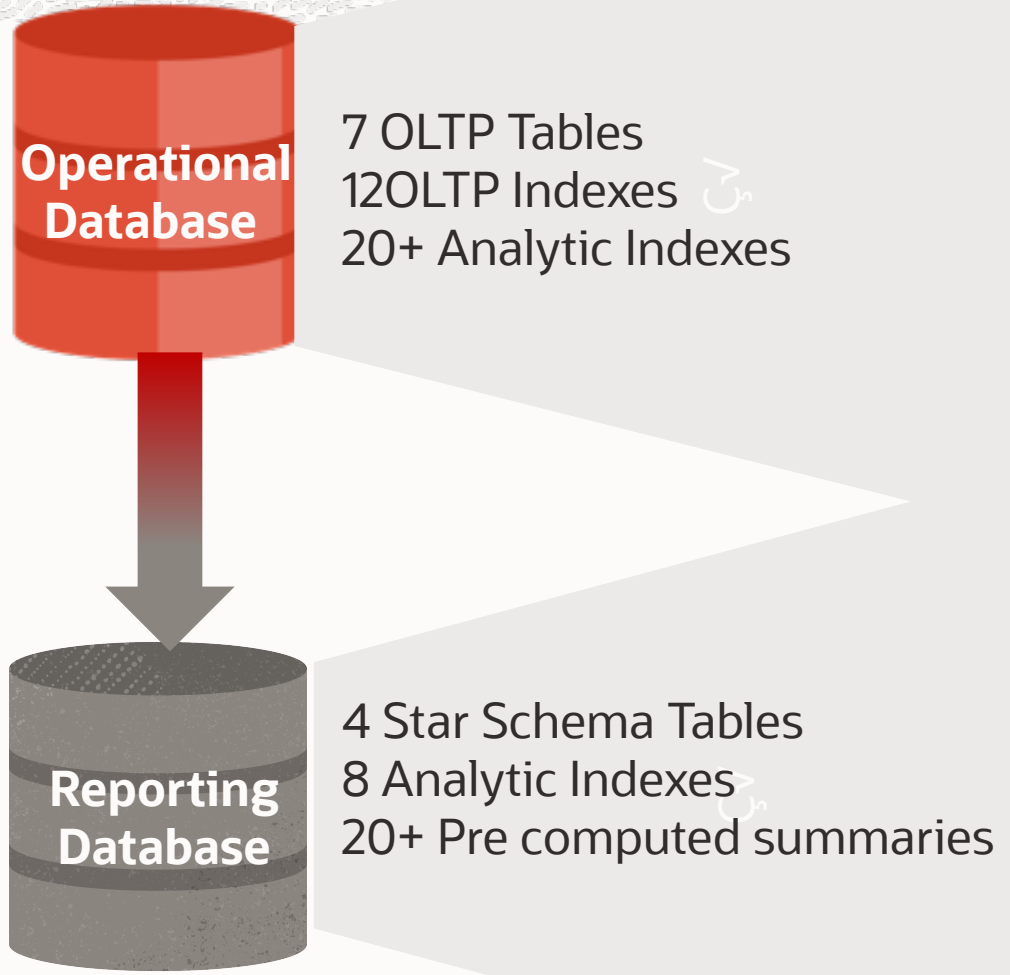


Magna Cart: Complexity Grows Further …

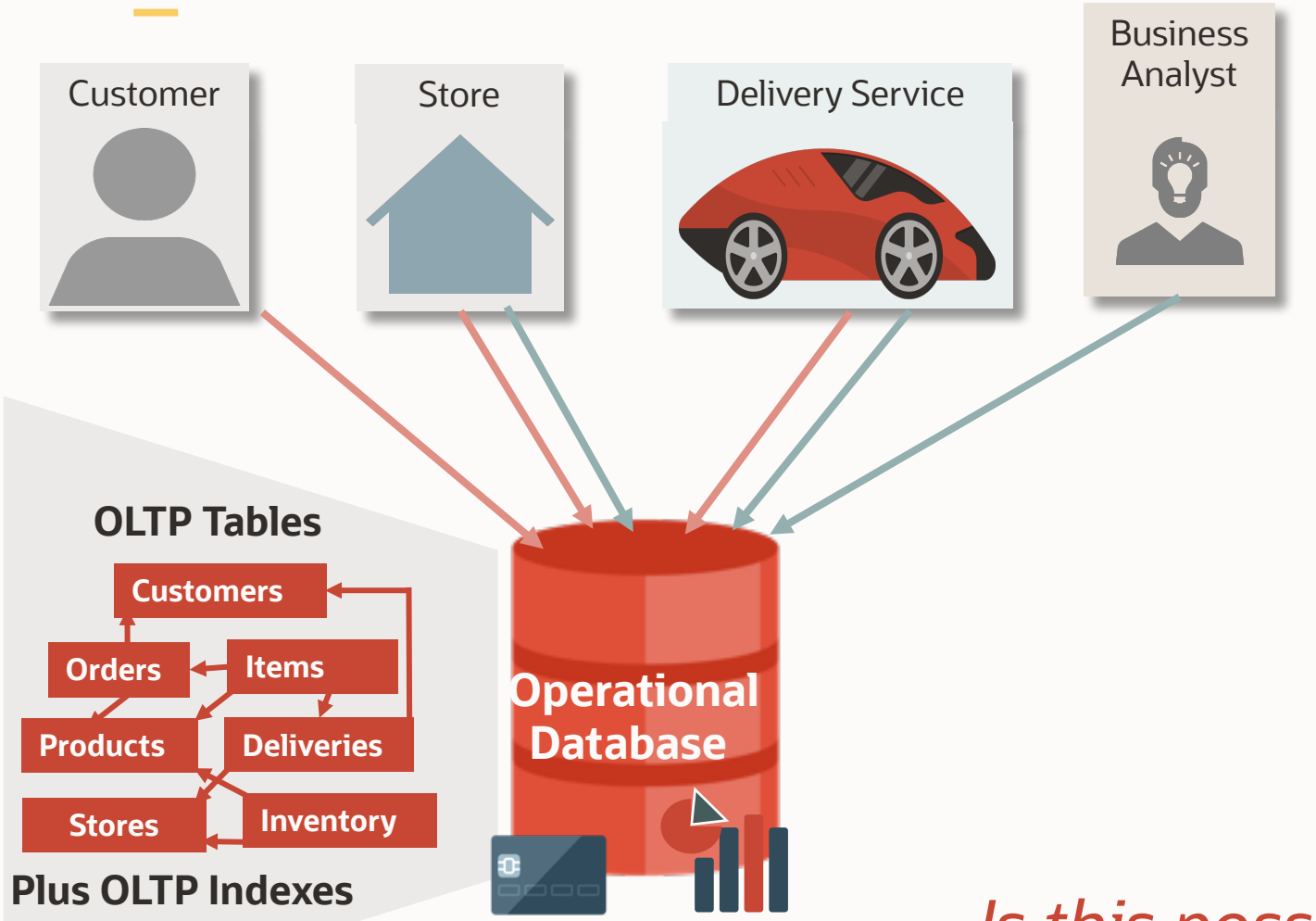


Sources of Complexity and Inefficiency

- **Reporting Databases** add major complexity:
  - Adds a new system to manage
  - Requires complex ETL logic
  - Usually lags Operational database
- **Analytic Indexes** add overheads:
  - Requires advance knowledge of workload
  - Slow down transactions and data ingest
  - Increases Database Size
- **Pre-computed summaries** add complexity:
  - Usually lag the source tables
  - Summaries adds additional load and DB size



What We Really Want is Simple



- A reporting database should not be required unless it is aggregating data from multiple sources
- No predefined indexes or precomputed summaries should be needed for Analytics
- Transactions and all Analytics should be able to run on the Operational Database

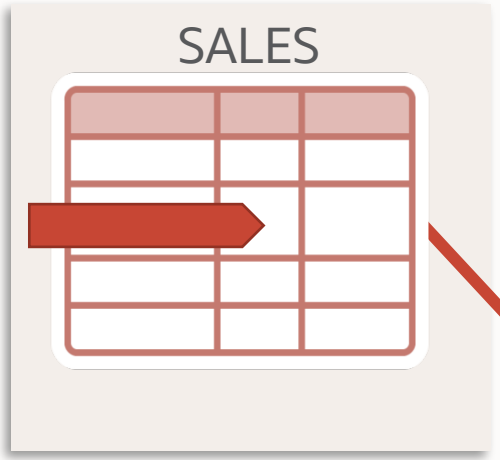
Is this possible?

YES

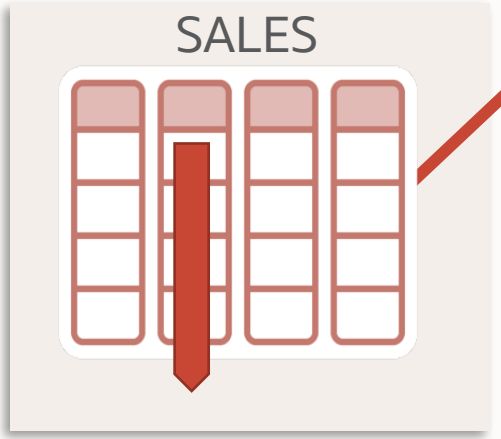
Introducing Database In-Memory

What's your favorite data format?

Row Format
Fast for OLTP!
Slower for Analytics

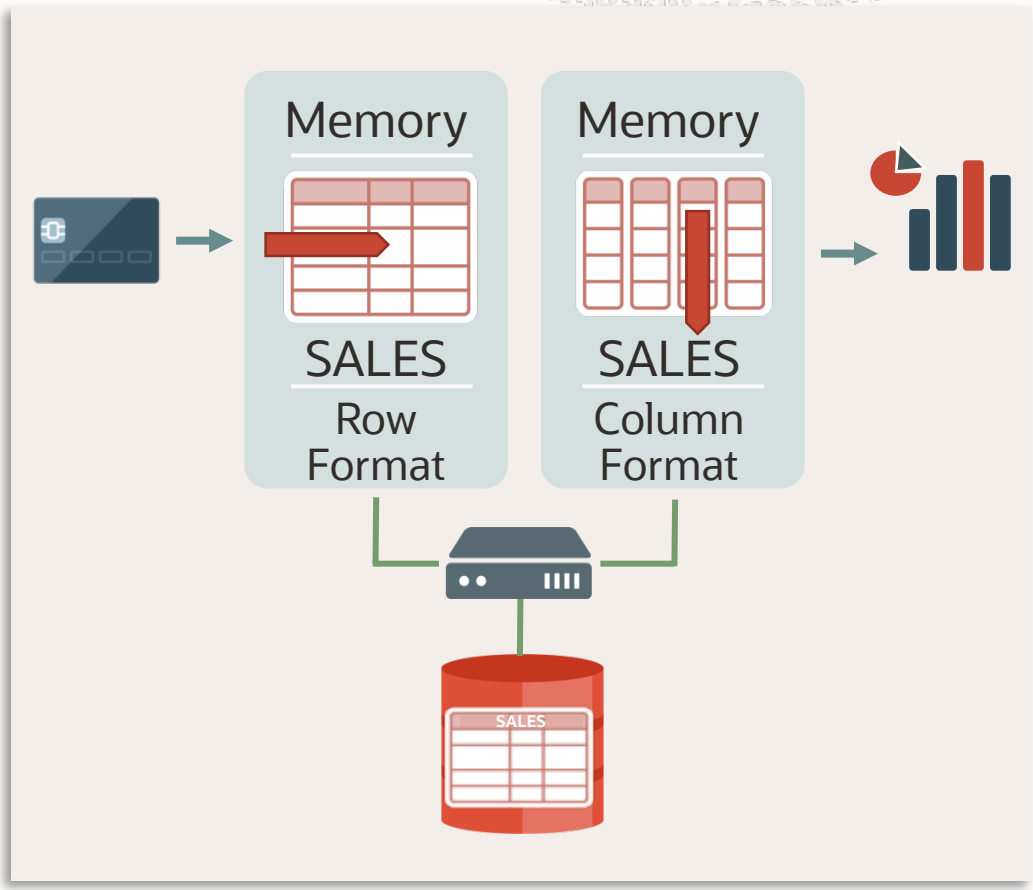


Column Format
Fast for Analytics!
REALLY slow for OLTP!

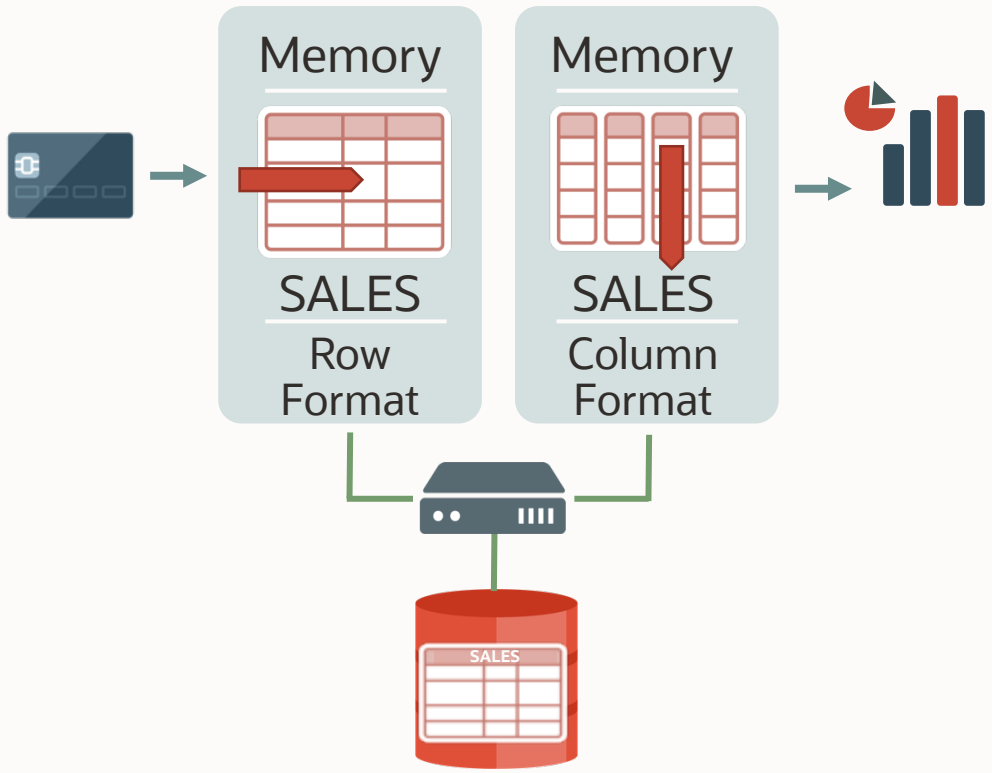


(+)

Dual Format
Best for both Fast Analytics and fast OLTP (No need for Analytic indexes)

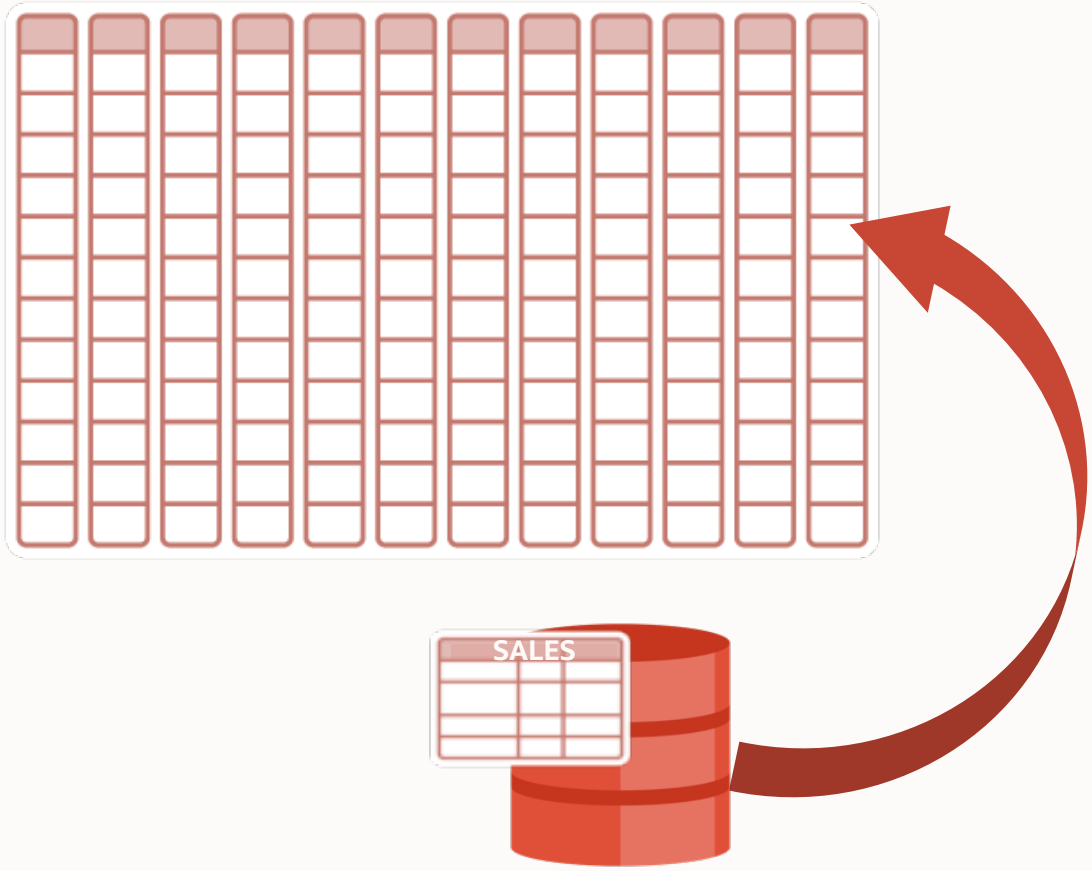


Database In-Memory: Real-Time Analytics with Fast OLTP



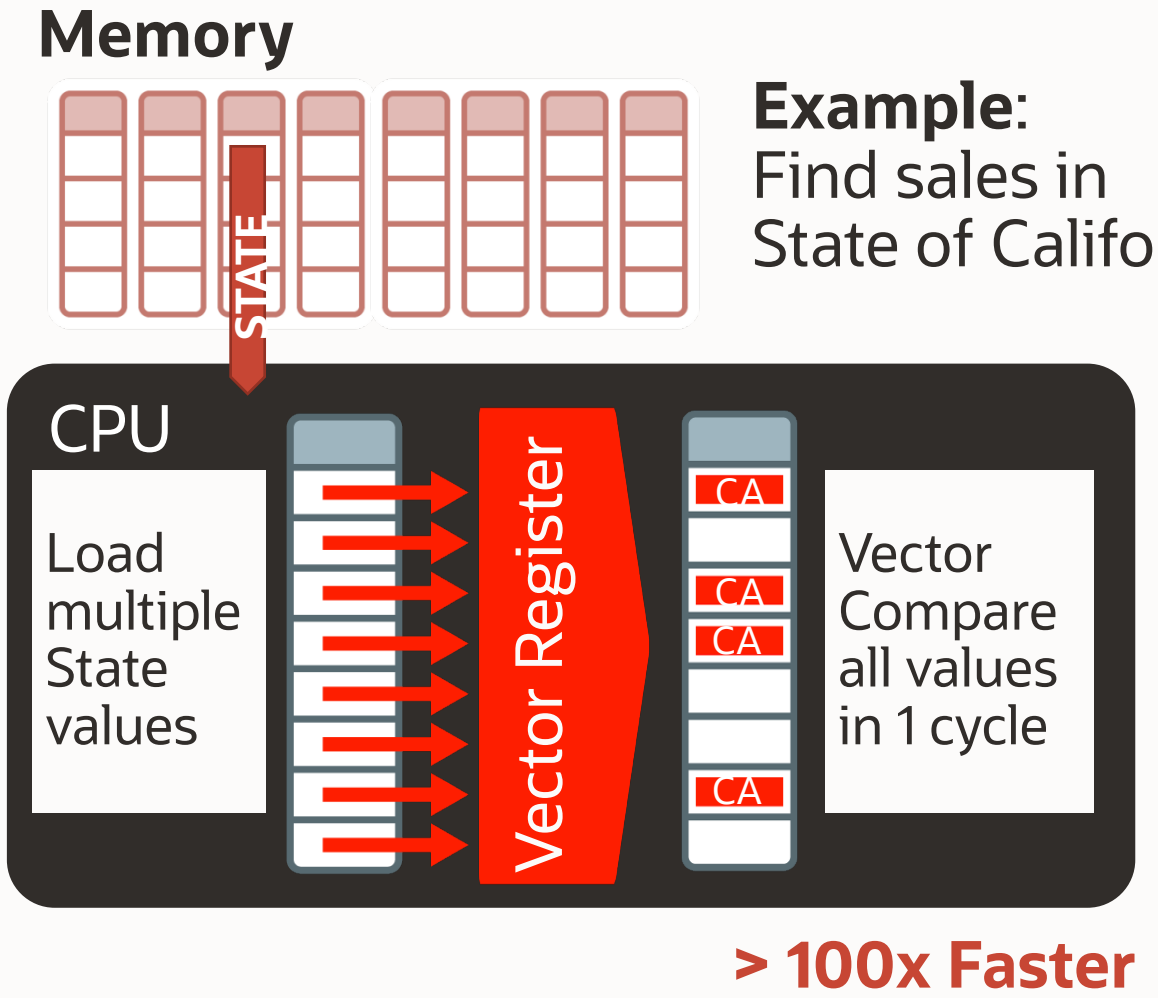
- Row format is ideal for OLTP and Column format for Analytics
- BOTH row and column format can be enabled for same table
  - Simultaneously active and consistent
- OLTP uses highly concurrent row format
  - Tens of millions of transactions per second
- Analytics uses in-Memory column format
  - SIMD vector instructions allow multiple values to be processed per instruction
  - Billions of Rows per second scan speeds

Database In-Memory: In-Memory Columnar Format



- Pure in-memory column format
  - In-memory maintenance: **Fast OLTP**
  - No changes to disk format
  - Available on all platforms
  - All enterprise features (security, availability) work transparently
- Does not require whole database to be in-memory
  - Can be enabled for hot data, at tablespace, table, partition, sub-partition level

In-Memory Enables SIMD Vector Processing



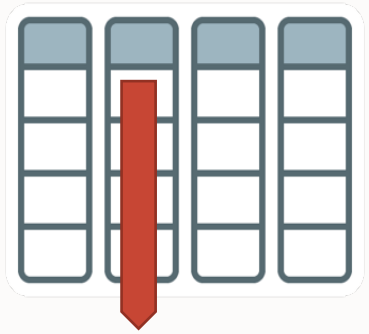
Example:
Find sales in State of California

- Column format benefit: Need to access only needed columns
- Process multiple values with a single SIMD Vector Instruction
- Billions of rows/sec scan rate per CPU core
  - Row format is millions/sec

In-Memory Technology Summary

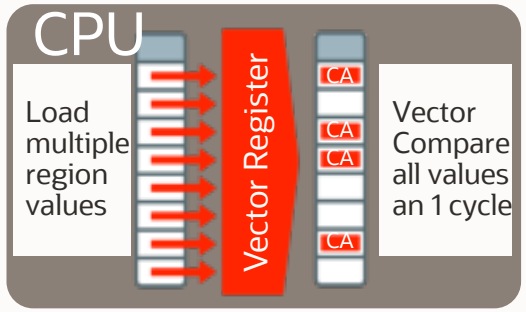
Greatly Accelerate Analytic Data Access

Columnar Format



Access only the columns you need

SIMD Vector Processing



Process multiple column values in a single CPU instruction

Storage Indexes

Prune out any unnecessary data from the column

Compression

Scan & filter data in compressed format

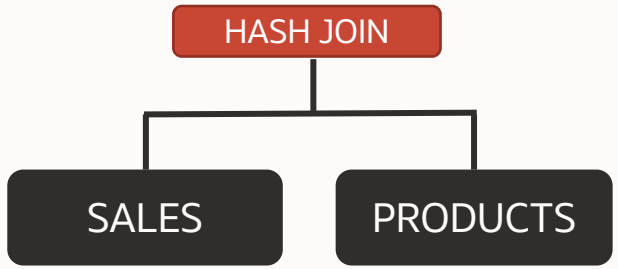
In-Memory Processing Summary

Greatly Accelerate all Aspects of Analytic Data Processing

In-Memory Scans

- Speed of memory
- Scan and Filter only the needed Columns
- Vector Instructions

In-Memory Joins



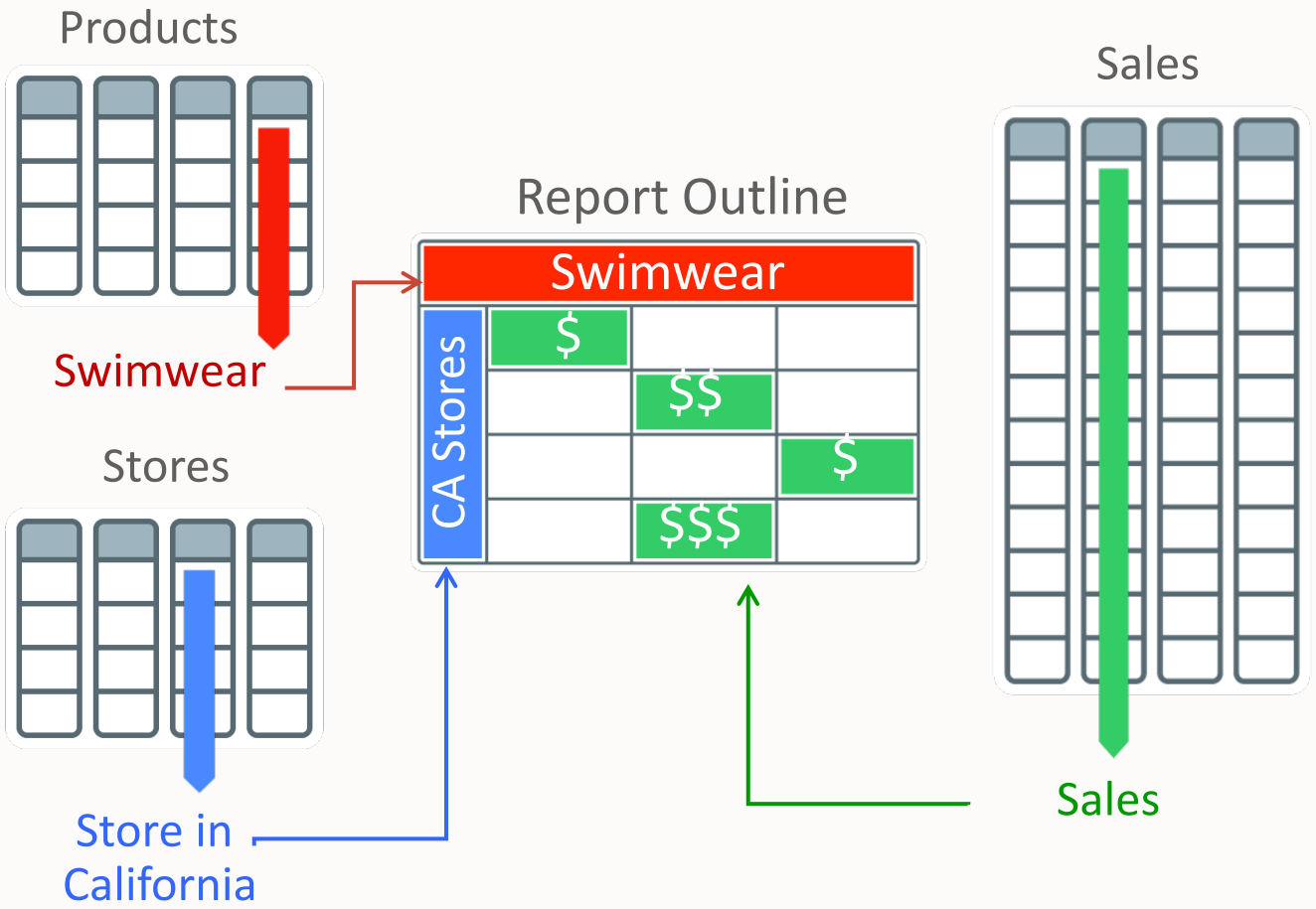
- Scan and Filter large table for values that match small table
- Convert Star Joins into 10X Faster Column Scans

In-Memory Reporting

- Create In-Memory Report Outline that is Populated during Fast Scan
- Runs complex reports 10x faster

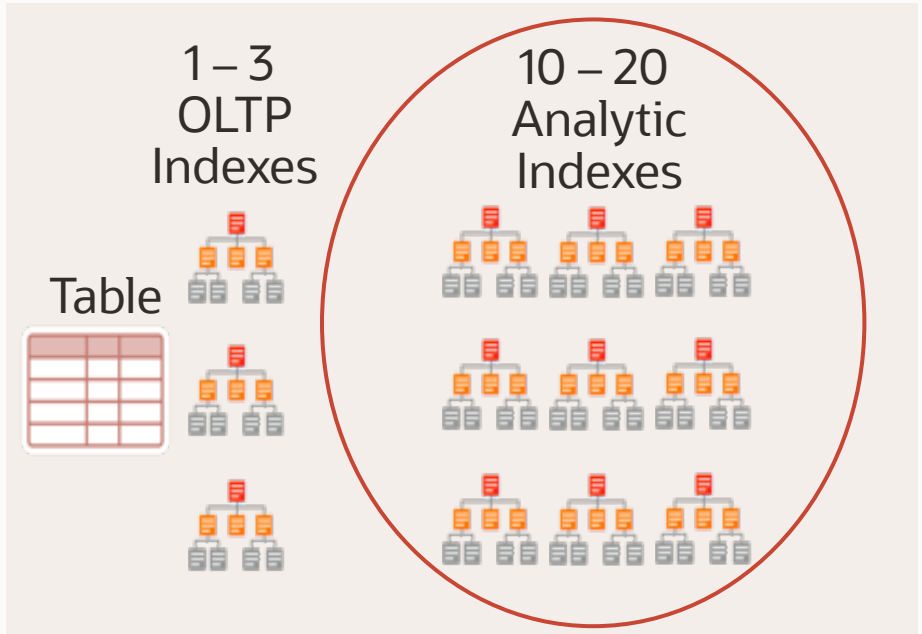
In-Memory Reporting

Example: Report sales of Swimwear in California Stores

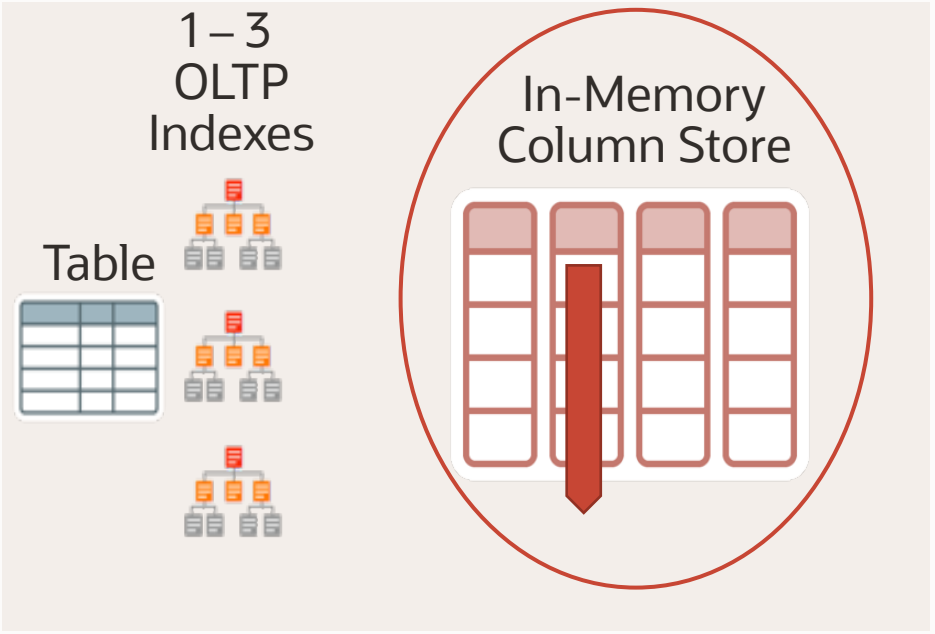


- First create in-memory report outline by scanning dimension tables
  - One cell per (California Store, Swimwear Product) combo
- Outline then populated by in-memory scan of fact table
- Report runs 10x faster

Accelerates Mixed Workloads like Magna Cart



REPLACE

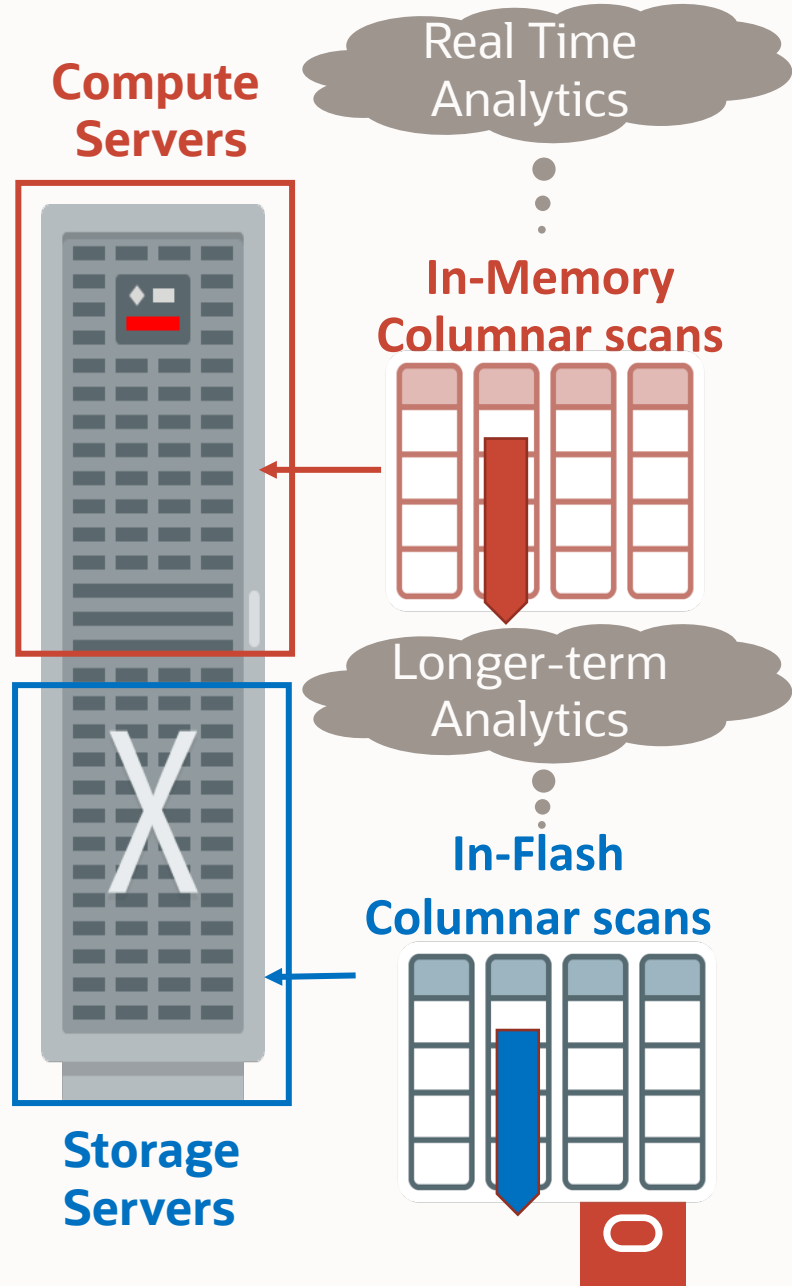


- Inserting one row into a table requires updating 10-20 analytic indexes: Slow!
- Fast analytics only on indexed columns
- Analytic indexes increase database size

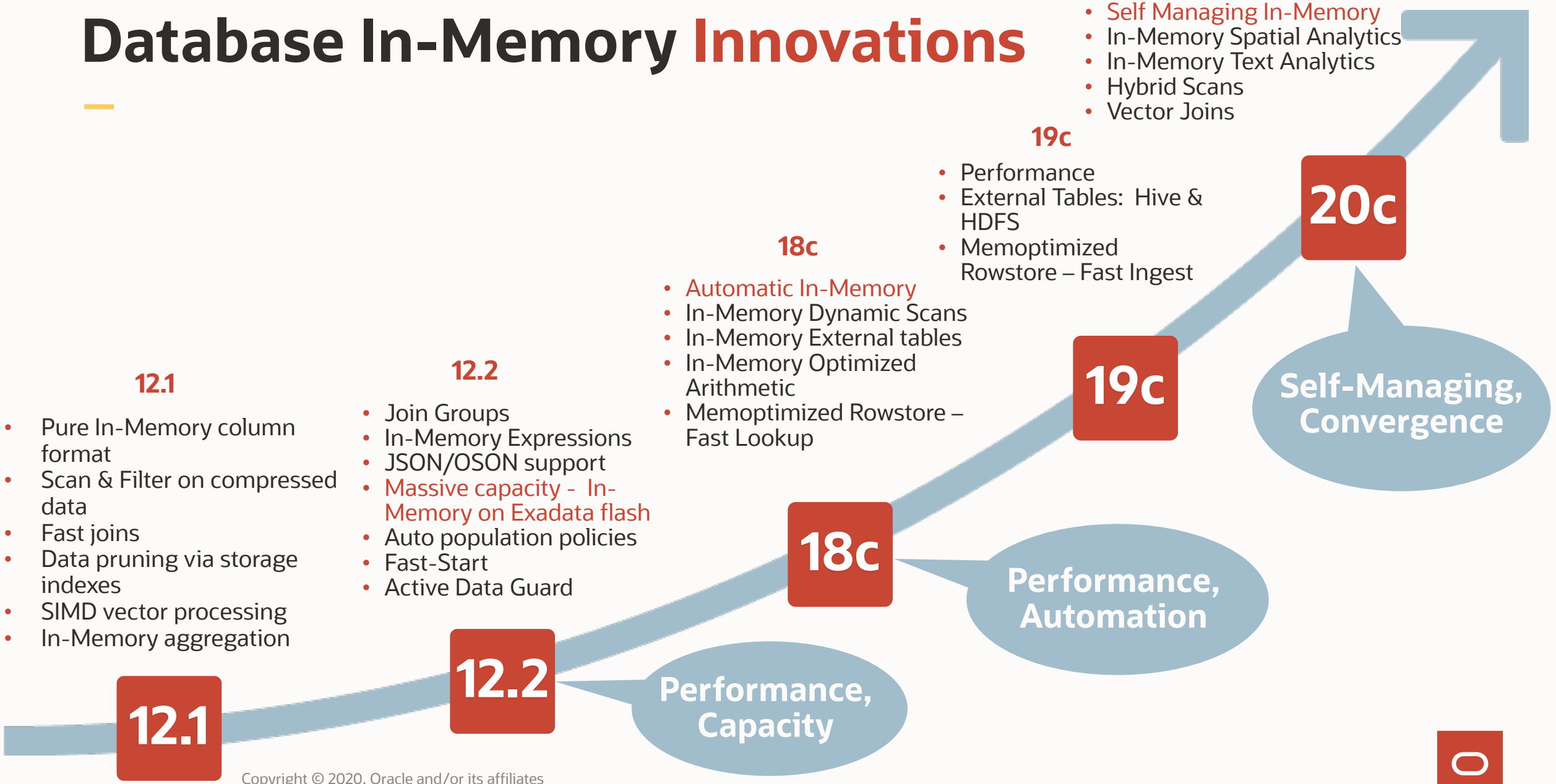
- Column Store not persistent so updates are: Fast!
- Fast analytics on any columns
- No analytic indexes: Reduces database size

In-Memory Extended to Flash

- In-Memory column format also available in in Exadata Smart Columnar Flash
  - Extends in-memory from DRAM in DB compute servers to Flash in storage servers
  - Enables SAME in-memory optimizations on data in Storage servers as on DB compute servers
- **Massive** increase in In-Memory Columnar capacity *500TB on full rack X8) for large tables that do not fit in DRAM
- Exadata smart query offload to storage replaces need to offload workload to reporting database

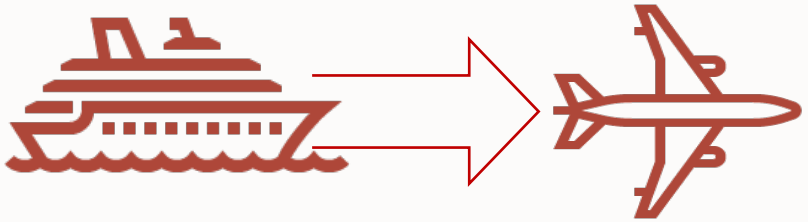


Database In-Memory Innovations



Database In-Memory Transforms Enterprise Architecture

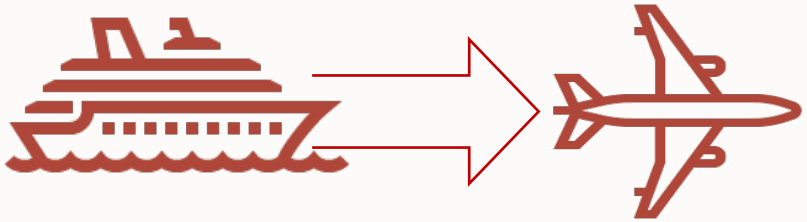
- Air travel was not just faster– it transformed society
  - Travel in hours not days
  - Changed business, trade, politics, education, media, sports, defense, tourism, employment …
  - No sea sickness (only jetlag) ☺

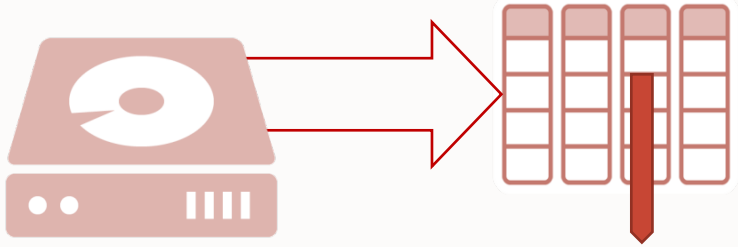


Database In-Memory Transforms Enterprise Architecture

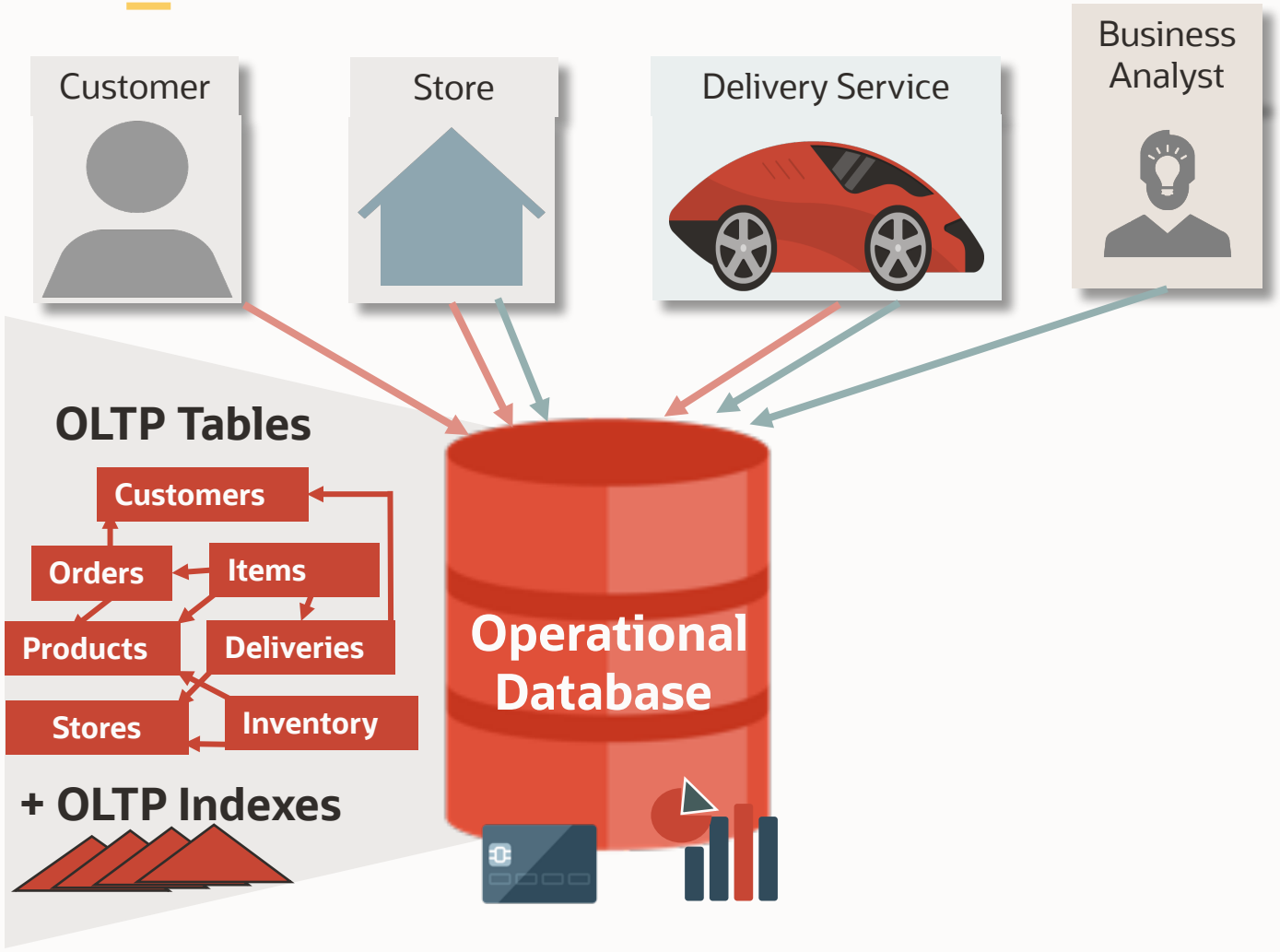
- Air travel was not just faster– it transformed society
  - Travel in hours not days
  - Changed business, trade, politics, education, media, sports, defense, tourism, employment …
  - No sea sickness (only jetlag) ☺
- Database In-Memory does not just improve performance

  - Instant business results → Faster decisions
  - No additional systems for reporting → Lower costs
  - Far simpler, more agile enterprise architecture



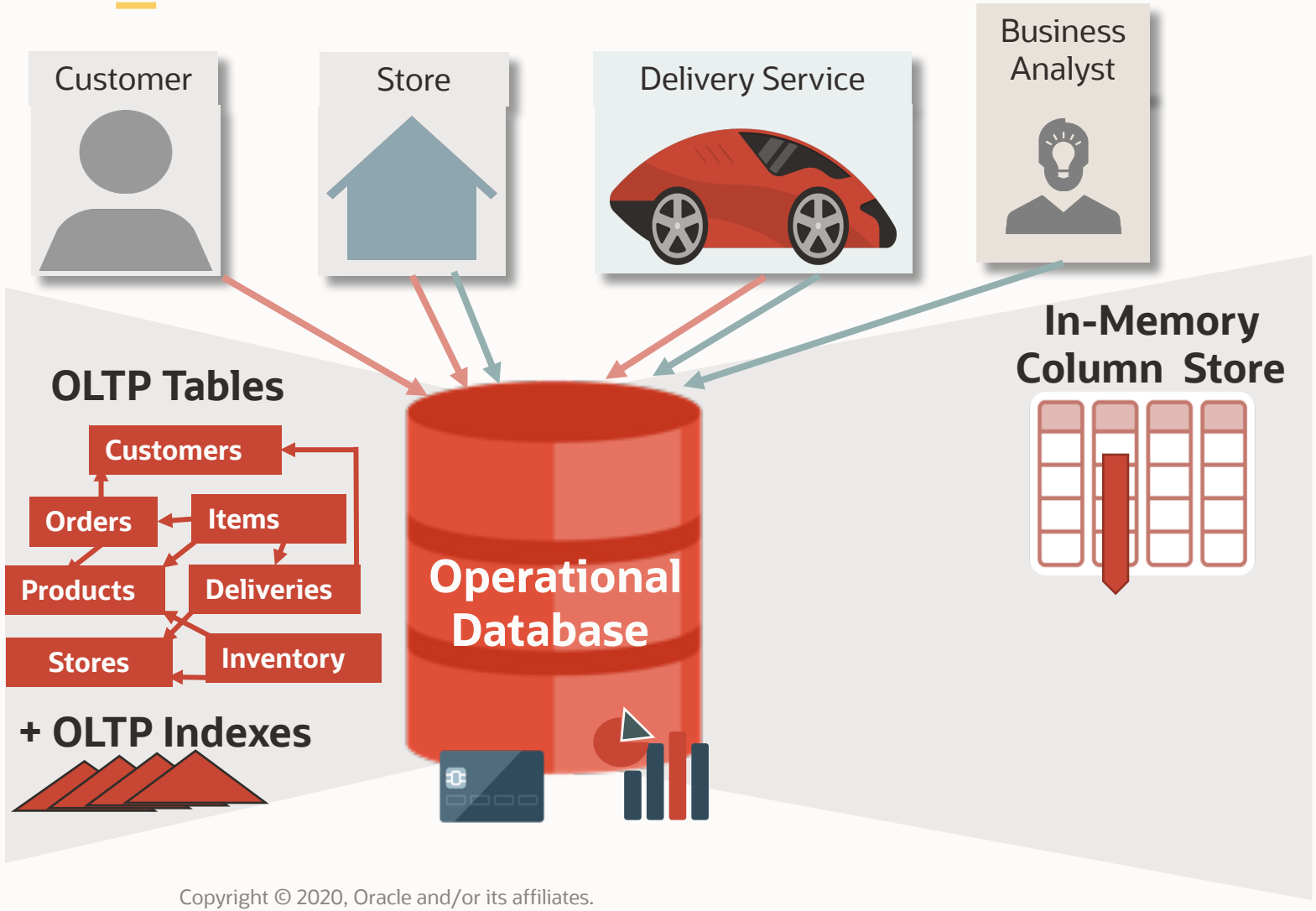


Magna Cart: Database In-Memory Brings Simplicity



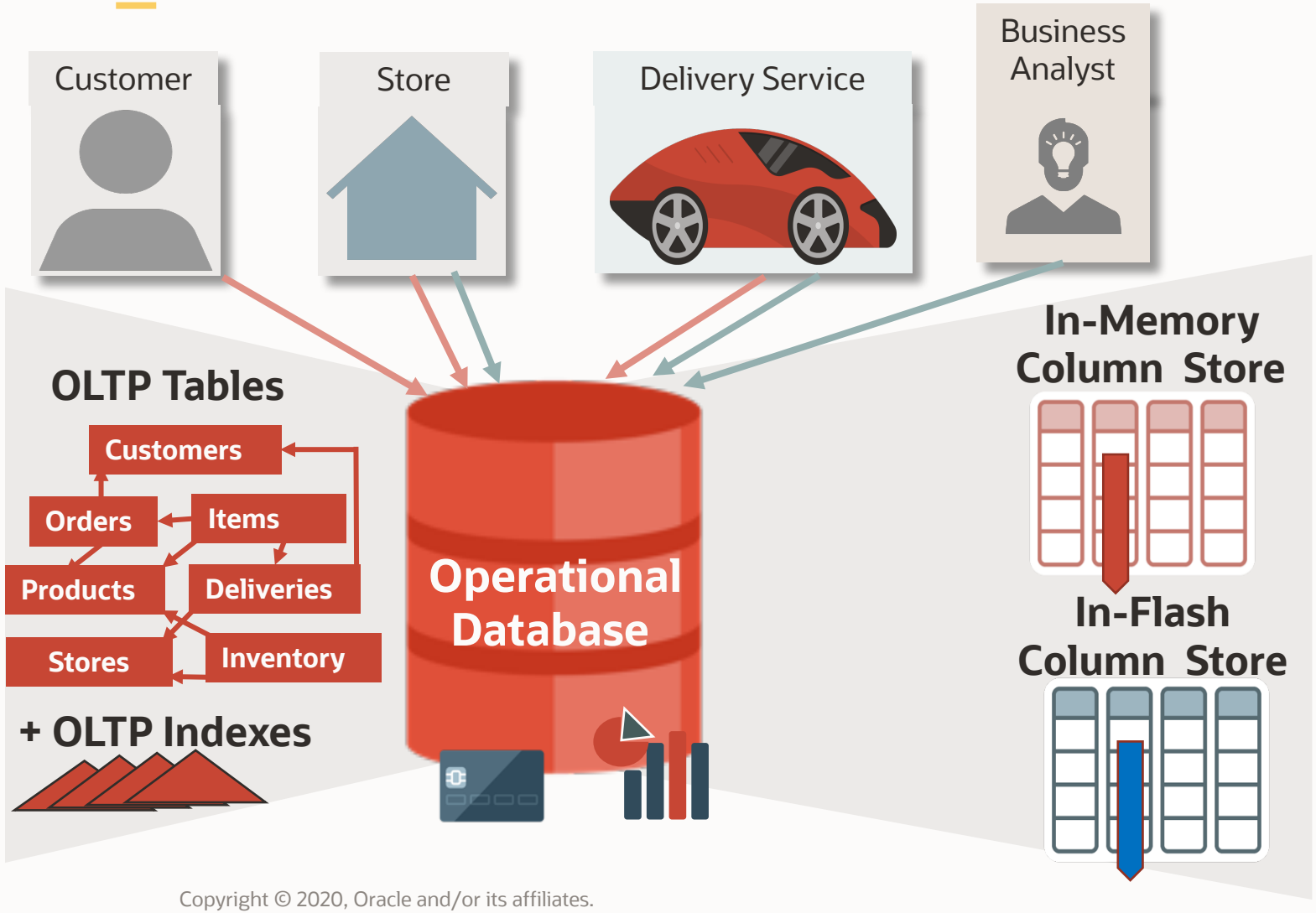
- **Entire workload can run on Operational database:**

Magna Cart: Database In-Memory Brings Simplicity



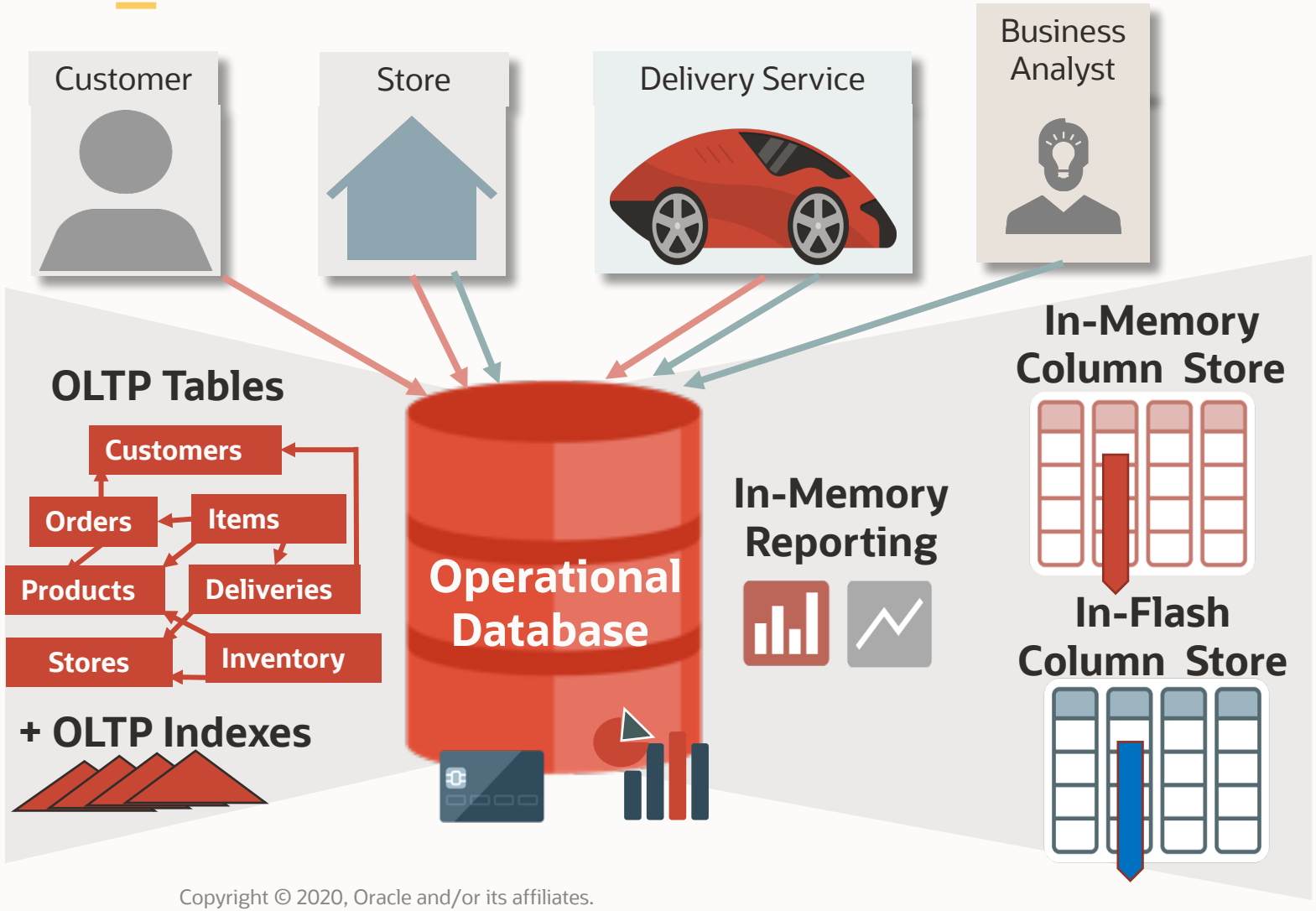
- **Entire workload can run on Operational database:**
  - In-Memory Column Store replaces Analytic Indexes

Magna Cart: Database In-Memory Brings Simplicity



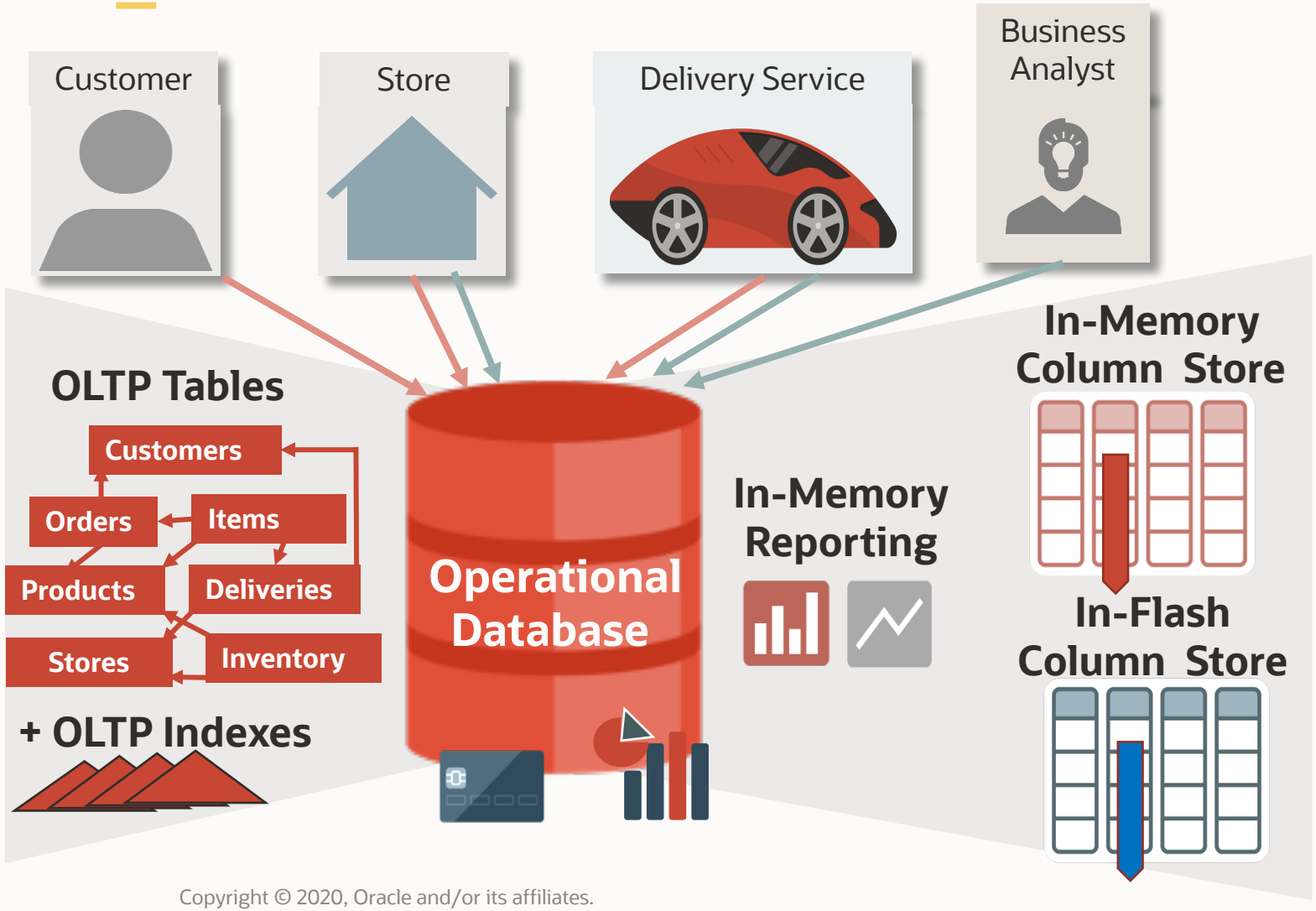
- **Entire workload can run on Operational database:**
  - In-Memory Column Store replaces Analytic Indexes
  - In-Flash Column Store replaces reporting database

Magna Cart: Database In-Memory Brings Simplicity



- **Entire workload can run on Operational database:**
  - In-Memory Column Store replaces Analytic Indexes
  - In-Flash Column Store replaces reporting database
  - In-Memory Reporting replaces Pre-Computed Summaries

Magna Cart: Database In-Memory Brings Simplicity



- **Entire workload can run on Operational database:**
  - In-Memory Column Store replaces Analytic Indexes
  - In-Flash Column Store replaces Reporting Database
  - In-Memory Reporting replaces Pre-Computed Summaries

Superfast transactions and analytics on all data, without prior knowledge of workload, all on a single database

How Customers Use Database In-Memory

Mankind Pharma –

Mixed Workload

- Analytical reports 11x faster
- Dropping indexes improved OLTP
- 90% reduction in database size

Lufthansa –

Reporting Application

- Analytic queries up to 100x faster
- Improved data ingest performance
- Reduction in database size

Shanghai Customs

Mixed Workload

- Processes Clearance 43x Faster
- Improves Declaration-Services Efficiency
- Reduced Costs

Die Mobiliar –

Mixed Workload

- Analytic queries 50-200X faster
- Database size reduced considerably
- Phase out of Netezza and mainframe systems

How Customers Use Database In-Memory

AT&T WiFi – reporting database

- Business Objects reports 100X faster
- ETL processes improved by 50% faster
- No changes to SAP Business Objects reports

Villeroy & Boch – SAP BW

- SAP BW COPA queries 30 – 33X faster
- SAP Transaction list queries 4 – 4,800X faster
- Avoided expensive & risky upgrade to S4/Hana

BOSCH – SAP CRM

- Dropped all custom indexes
- Analytic queries 2-20X faster, DML 2-3X faster
- No changes to application required

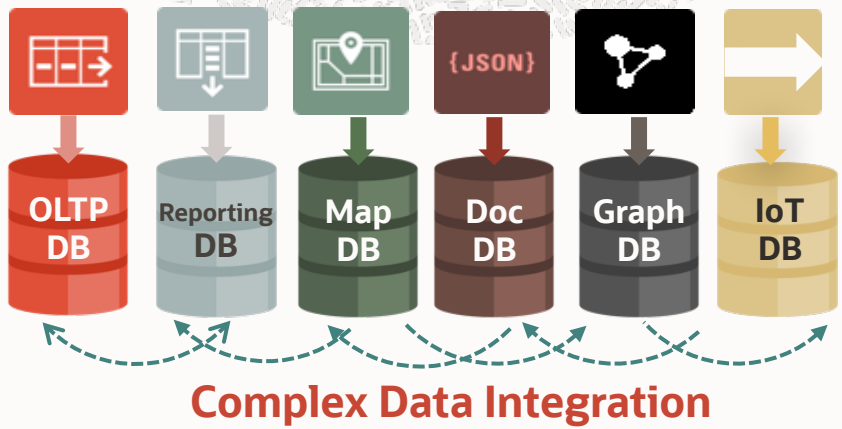
LION – SAP ERP

- Analytic queries 4X faster
- Transactions 2X faster
- Analytic queries now possible on 100s of Millions of Point-of-Sales Transactions

Oracle Converged Database

Path to Simpler Application Evolution

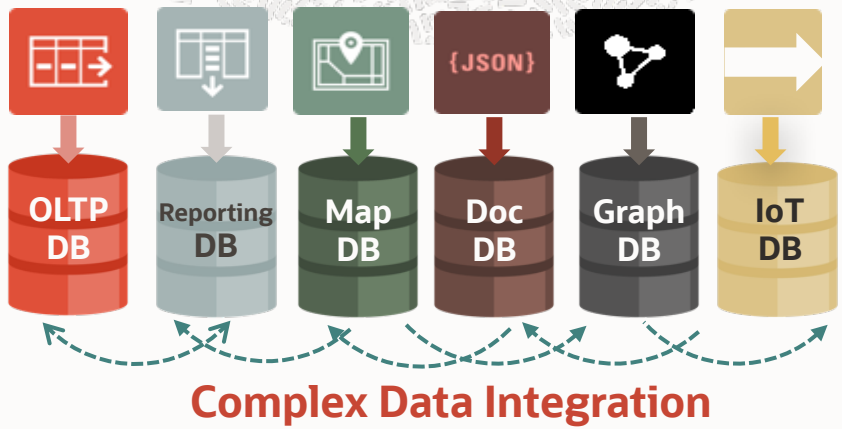
- As applications evolve, they often need other algorithms: Document, Graph, AI/ML. etc.



Oracle Converged Database

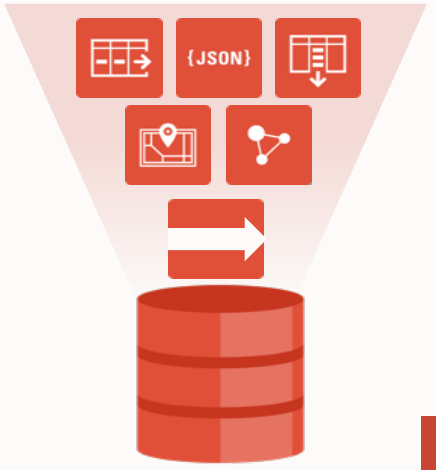
Path to Simpler Application Evolution

- As applications evolve, they often need other algorithms: Document, Graph, AI/ML. etc.
- Oracle converged database enables any type of workloads against any type of data
  - Relational, Document, Graph, Spatial, Files, etc.
  - Class leading Machine Learning portfolio
- Converged DB is: **Simpler**, **Lower Cost**, **More Reliable**
  - Same security management across all data
  - Prevents data fragmentation and copy contagion
  - Eliminates multi-database integration costs



v.s.

Converged DB for All Workloads

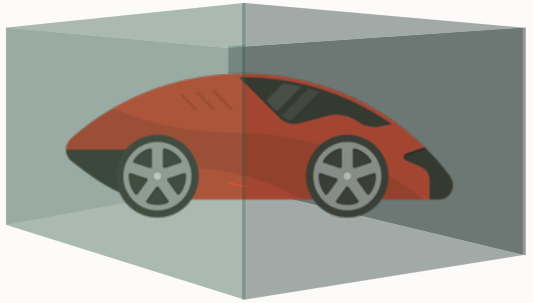


Autonomous Database

Ultimate Converged Platform

Self-Driving

- Scale-out database with fault-tolerance and DR
- Runs on enterprise-proven Exadata platform
- Full compatibility with existing enterprise databases



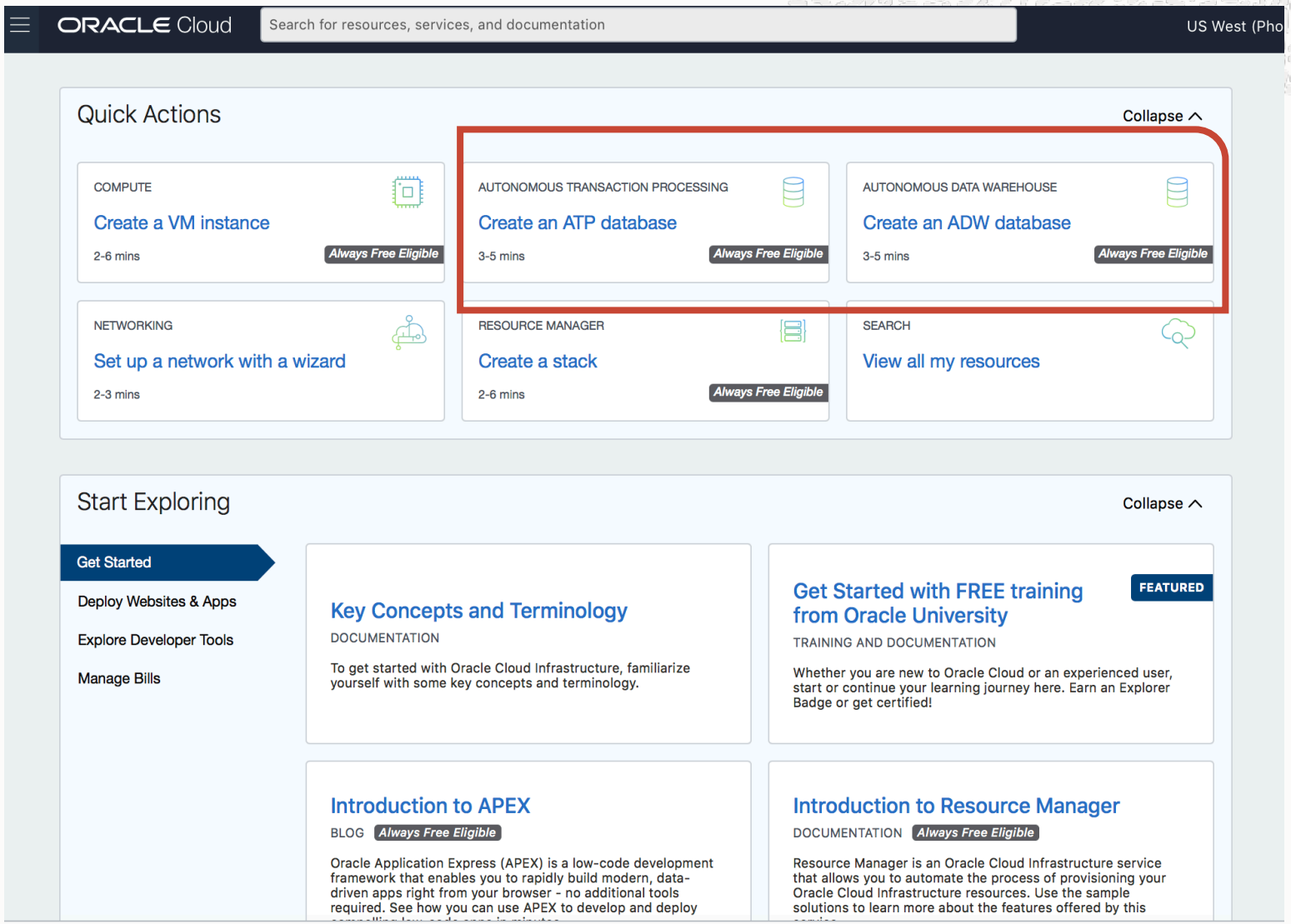
Self-Securing

- Automatically applies security updates online
- Secure configuration with full database encryption
- Sensitive data hidden from Oracle or customer admins

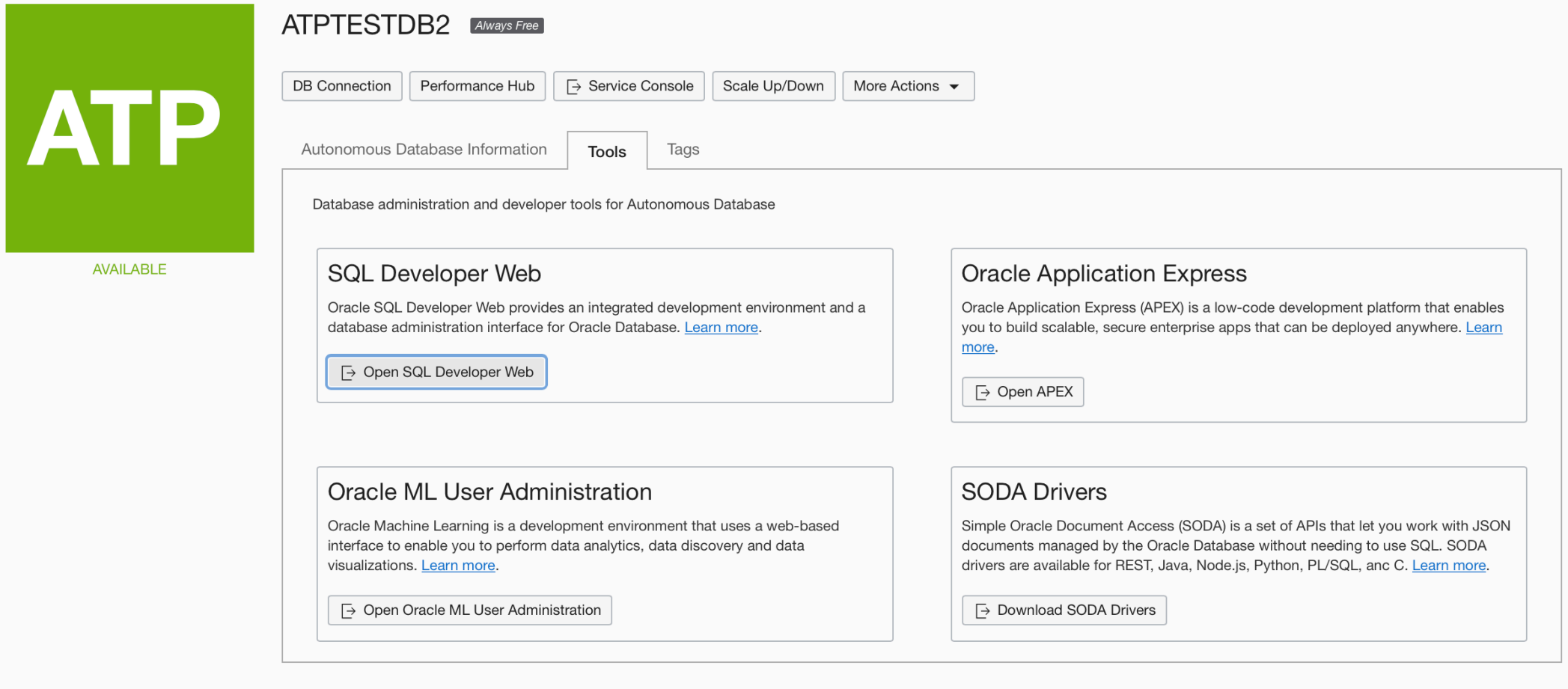
Self-Repairing

- Recovers automatically from any failure
- 99.995% uptime including maintenance
- Elastically scales compute or storage as needed

Oracle Cloud Free Tier: Easy On-Ramp to Autonomous DB

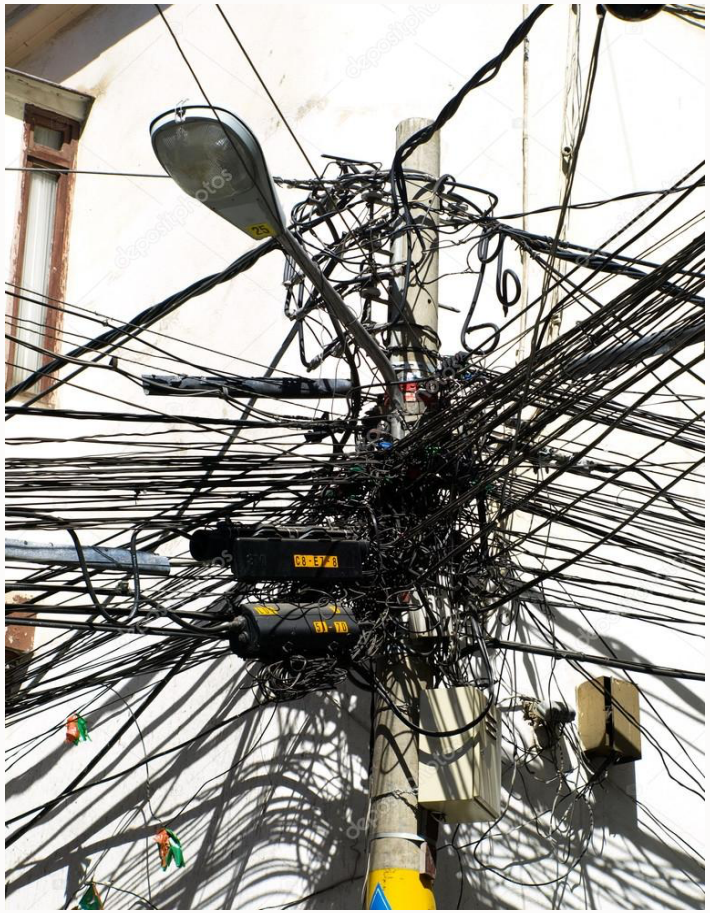


Autonomous Database Always Free Tier



Five Stages of Complexity

Denial: It's actually very simple …
Pain: It is broken!
Anxiety: When will it break next?
Depression: It always breaks …
Substance Abuse: Just once more, for this escalation …



Five Stages of Complexity

Denial: It's actually very simple …
Pain: It is broken!
Anxiety: When will it break next?
Depression: It always breaks …
Substance Abuse: Just once more, for this escalation …

"Simplicity is Complexity Resolved"

Getting Started With Database In-Memory

- **Download Free Oracle Express Edition**
  https://tinyurl.com/oracle18cXE
  Includes Database In-Memory with up to 2GB in-memory column store

- **Try out free Exadata Express on Oracle Cloud**
  https://tinyurl.com/exadataexpress
  Includes Database In-Memory with up to 5GB in-memory column store

- **Free In-Memory Base Level in Oracle Enterprise Edition 19.8 onwards**
  https://tinyurl.com/dbimbaselevel
  Enables an in-memory column store of up to 16GB per DB instance

For More Information

Visit the Database In-Memory Blog
https://blogs.oracle.com/in-memory/

Read the Database In-Memory Overview
https://www.oracle.com/database/technologies/in-memory.html

Try out Autonomous DB for Free!
https://www.oracle.com/cloud/free/

ORACLE