**1. Data Warehousing Introduction**

A **Data Warehouse (DWH)** is a central repository of integrated data collected from different sources. It supports **reporting**, **data analysis**, and **business intelligence**.

**Real-world Example:**  
Amazon uses data warehousing to analyze customer behavior, recommend products, and manage inventory across the globe.

**2. Data Store Vendors**

These are companies or platforms that offer data storage and management solutions for different purposes like warehousing, transactional systems, or big data.

**Popular Vendors:**

* Amazon Redshift
* Google BigQuery
* Snowflake
* Microsoft Azure Synapse
* Oracle Exadata

**Real-world Example:**  
Netflix uses Amazon Redshift to store viewing data and preferences to personalize recommendations.

**3. OLTP, OLAP Systems**

* **OLTP (Online Transaction Processing)** systems manage real-time transactional data (insert/update/delete).
* **OLAP (Online Analytical Processing)** systems are used for querying and analyzing large volumes of data.

**Real-world Example:**

* **OLTP**: Booking a flight on MakeMyTrip (handles thousands of transactions/sec).
* **OLAP**: Analysts at Flipkart running queries to identify sales trends over the last year.

**4. OLAP and OLTP Deep Dive**

**OLTP:**

* High concurrency
* Real-time updates
* Supports CRUD operations

**OLAP:**

* Complex queries
* Low concurrency
* Supports multi-dimensional analysis (e.g., sales by region, product, time)

**Real-world Example:**

* A retail company might use **OLTP** for daily sales billing and **OLAP** for analyzing seasonal buying behavior.

**5. Modern Data Architecture**

Modern data architecture supports **cloud-native**, **real-time analytics**, and **scalable** storage. It integrates **Data Lakes**, **Data Warehouses**, **Streaming Data**, and **AI/ML pipelines**.

**Key Components:**

* Data Ingestion tools (e.g., Apache Kafka)
* Storage (e.g., Data Lake + DWH)
* Transformation (ETL/ELT)
* Visualization (e.g., Power BI)

**Real-world Example:**  
Uber uses modern architecture with real-time data ingestion (Kafka), analytics (BigQuery), and dashboards to track rides, pricing, and user demand.

**6. DWH vs. Data Lake**

| **Feature** | **Data Warehouse** | **Data Lake** |
| --- | --- | --- |
| Data Type | Structured | Structured + Semi/Unstructured |
| Storage Cost | High | Low |
| Performance | Fast for analytics | Depends on processing engine |
| Schema | Schema-on-write | Schema-on-read |

**Real-world Example:**

* **Walmart** uses DWH for daily sales reports.
* **Netflix** uses a Data Lake to store and analyze user-generated logs, video content, and recommendation data.

**7. DWH Architecture**

Typical components:

1. **Data Sources** (e.g., ERP, CRM)
2. **ETL Layer** – Extract, Transform, Load
3. **Staging Area** – Temporary storage
4. **Data Warehouse** – Central repository
5. **Data Marts** – Subsets for specific departments
6. **BI Tools** – Dashboards and reports

**Real-world Example:**  
Banking companies like ICICI or HDFC use DWH architecture to process customer data from various systems and create financial reports and fraud detection insights.

**8. Data Mart**

A **Data Mart** is a focused subset of a data warehouse designed for a specific line of business (e.g., Sales, HR).

**Types:**

* **Dependent** (linked to main DWH)
* **Independent** (standalone)

**Real-world Example:**  
In an insurance company, the **Claims Department** uses a Claims Data Mart to analyze trends in policy claims.

**9. Operational Data Store (ODS)**

An **ODS** is used for operational reporting and is updated frequently with real-time or near real-time data.

**Key Features:**

* Combines data from multiple sources
* Low latency
* Supports operational decision-making

**Real-world Example:**  
Hospital Management Systems use an ODS to monitor patient admissions, discharges, and real-time occupancy rates.