Data Cloud

Crash Course : A Developer's Perspective

Learn by doing it!

SFBOLT

SF BOLT

Index

- Overview
- Data Flow DIagram
- Data Schema
- Data Cloud Introduction
- Data Stream
- Data Lake Object
- Data Model Object
- Data Ingestion
- Identity Resolution
- Unified Profiles
- Use Data Cloud Data in Salesforce
- Data Cloud Related List
- Giveaway

Overview

1. The Challenge

Rideswift Rental, a company that provides two and four-wheeler rentals, has all of its customer data in a private server. They recently acquired **Adventure Cloud**, a company whose data lives on a separate Salesforce server.

The problem? This data is in **separate silos**. It's impossible to easily see that a customer who rents a car from Rideswift might also be interested in a hiking trip from Adventure Cloud.

2. The Solution: Data Cloud

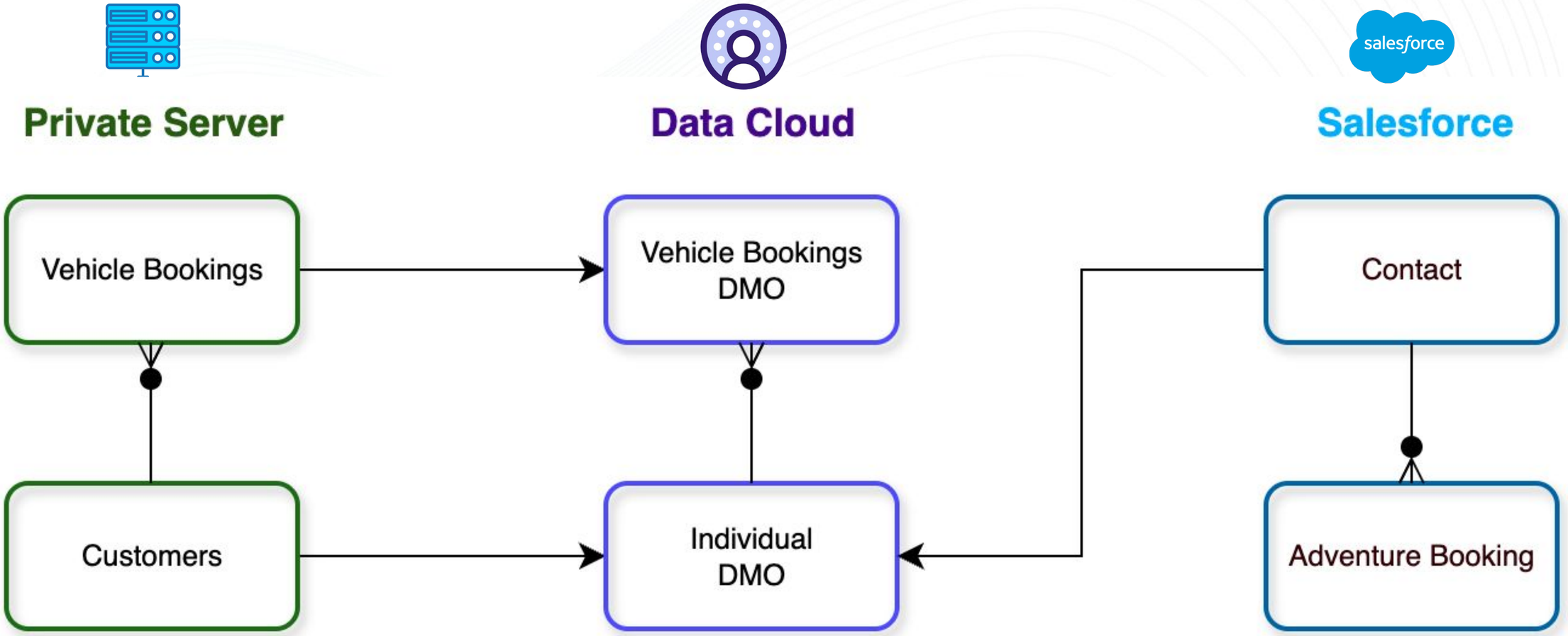
Data Cloud is the answer. Think of it as a central brain that connects all of your data sources. It takes scattered information and unifies it into a single, comprehensive view.

Data Cloud will:

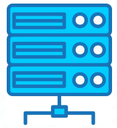
- **Ingest** data from Rideswift's server.
- **Ingest** data from Adventure Cloud's Salesforce server.
- **Unify** all the data to create a single, complete customer profile.

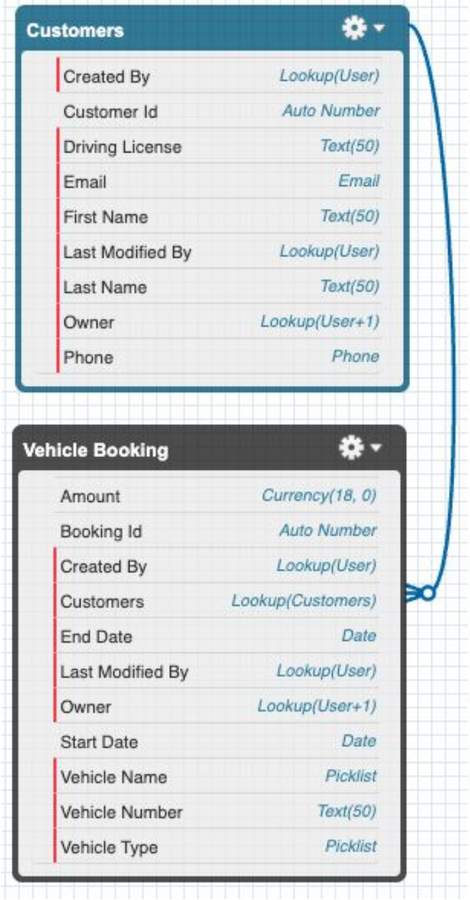
SF BOLT

Data Flow Diagram



Data Schema




Customers
Created By Lookup(User)
Customer Id Auto Number
Driving License Text(50)
Email Email
First Name Text(50)
Last Modified By Lookup(User)
Last Name Text(50)
Owner Lookup(User+1)
Phone Phone
Vehicle Booking
Amount Currency(18, 0)
Booking Id Auto Number
Created By Lookup(User)
Customers Lookup(Customers)
End Date Date
Last Modified By Lookup(User)
Owner Lookup(User+1)
Start Date Date
Vehicle Name Picklist
Vehicle Number Text(50)
Vehicle Type Picklist


salesforce


Contact
Account Name Lookup(Account)
Assistant Text(40)
Asst. Phone Phone
Birthdate Date
Buyer Attributes Picklist (Multi-Select)
Clean Status Picklist
Contact Owner Lookup(User)
Created By Lookup(User)
Creation Source Picklist
Data.com Key Text(20)
Department Text(80)
Description Long Text Area(32000)
Do Not Call Checkbox
Email Email
Email Opt Out Checkbox
Fax Phone
Fax Opt Out Checkbox
Gender Identity Picklist
Home Phone Phone
Individual Lookup(Individual)
Languages Text(100)
Last Modified By Lookup(User)
Last Stay-in-Touch Request Date
Adventure Booking
Adventure Number Auto Number
Amount Currency(18, 0)
Contact Lookup(Contact)
Created By Lookup(User)
End Date Date
Last Modified By Lookup(User)
Name Text(255)
Owner Lookup(User+1)
Start Date Date

What is Data Cloud?

Data Cloud is a centralized platform that **connects and unifies data from multiple, disconnected sources** into a single, comprehensive view. It essentially breaks down "data silos" and turns scattered information into a complete customer profile. This unified data can then be used to generate actionable insights and personalize customer experiences, helping a business grow.

Data Stream

Definition:
A data stream is a continuous flow of data generated in real-time from sources like sensors, applications, or user interactions, which is processed and analyzed as it arrives.

Key Points:
- Often used in real-time/ near real-time analytics.
- Common in use cases like financial transactions, or clickstream analysis.

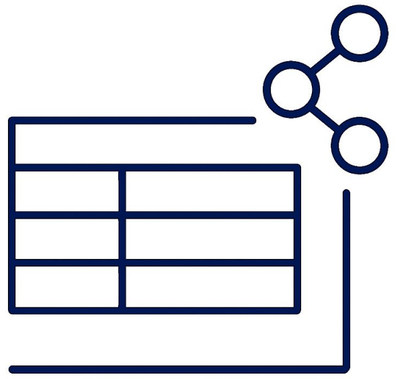
Example:
A fitness app sending live heart rate data from a wearable device to the cloud is a data stream.

Data Lake Object

Definition:
A data lake object refers to a unit of data stored in a data lake, which is a centralized repository that holds raw data in its actual format until needed.

Key Points:

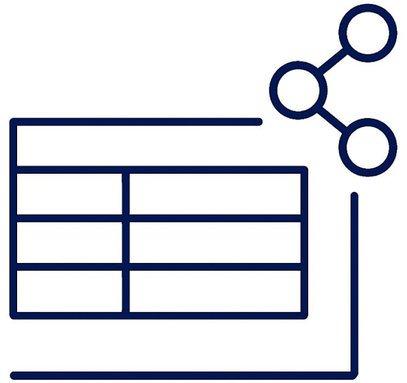
- Can be structured, semi-structured, or unstructured.
- Enables scalable storage and flexible access for analytics and machine learning.



Data Model Object

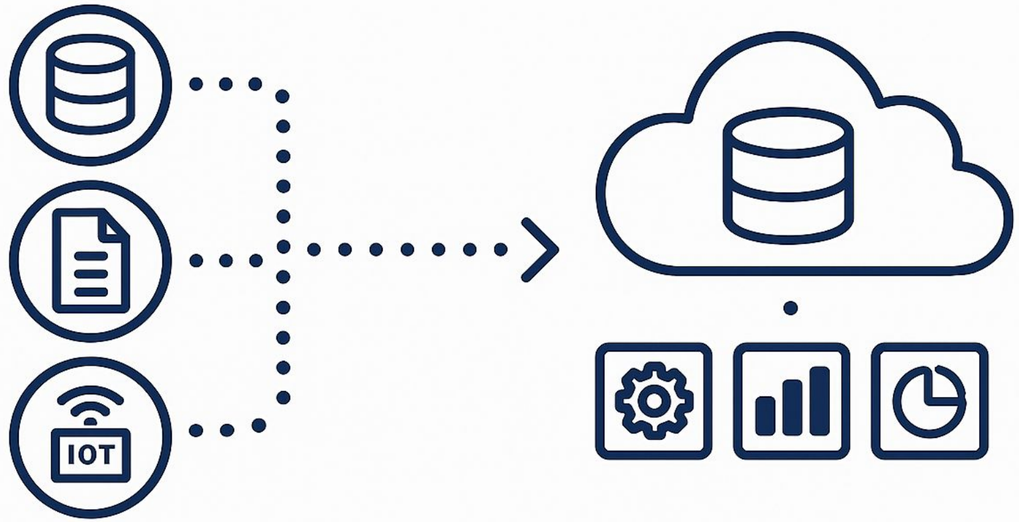
Definition:
A data model object represents a structured view of data, defining how data elements relate to each other used to organize, query, and analyze data effectively.

Key Points:
- Built using schemas, tables, relationships, and business logic.
- Helps in data governance, reporting, and semantic modeling.



Data Ingestion

Data ingestion is the process of collecting and importing data from various sources into a centralized cloud-based data platform, where it can be stored, processed, and analyzed.



Key Points:

- It’s the first step in building a data pipeline.
- Data can come from databases, files, APIs, IoT devices, streaming platforms, or other systems.
- Ingestion can be:
  - Batch-based (data is collected and loaded at scheduled intervals)
  - Real-time/streaming (data is continuously collected and processed as it arrives)
- The goal is to make data available and usable for analytics, reporting, machine learning, and other business needs.

Identity Resolution

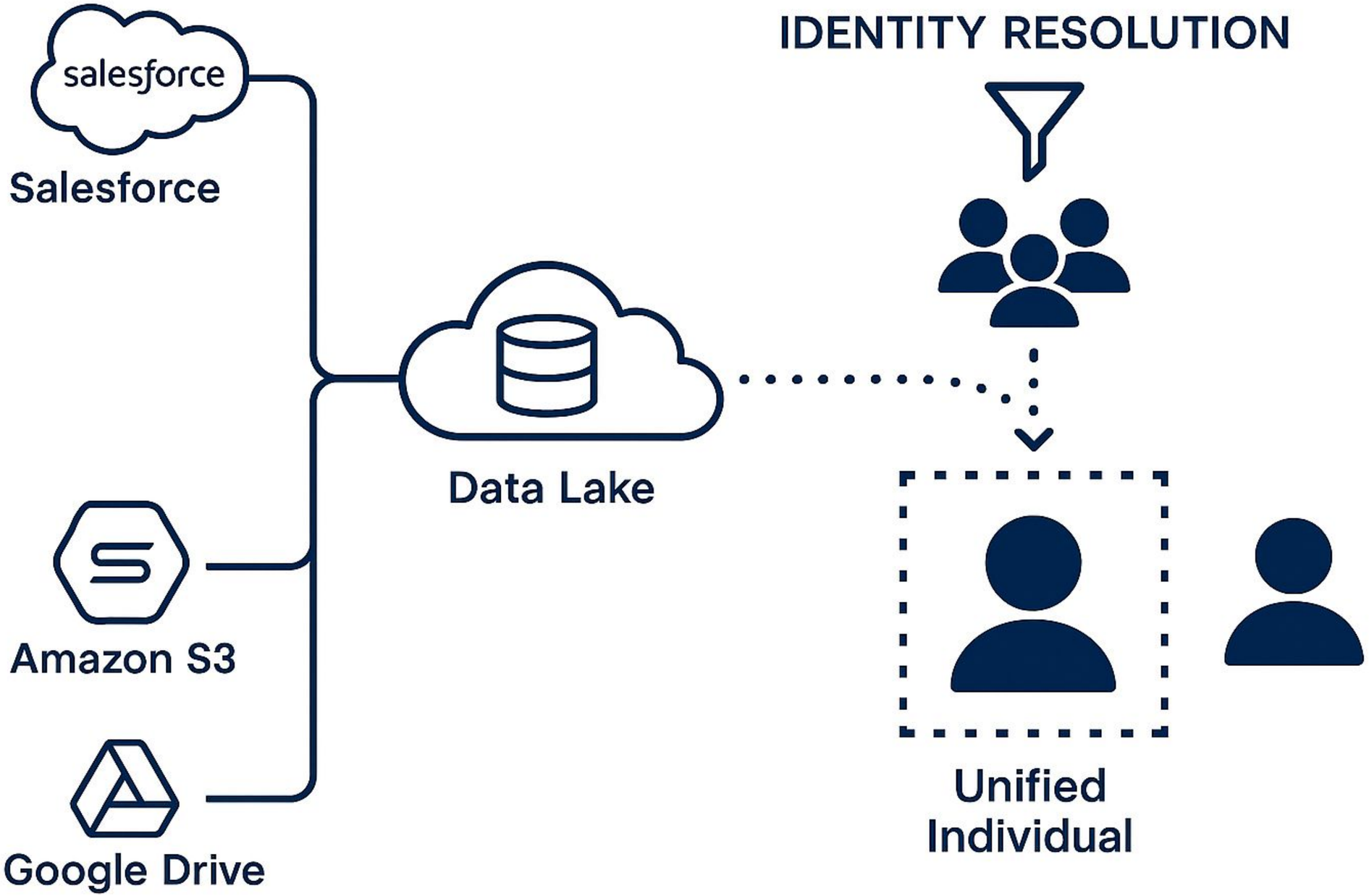
Identity Resolution is the process of accurately identifying and linking records that refer to the same entity - such as a customer or product across multiple data sources, even when the data is inconsistent, incomplete, or duplicated.

Key Characteristics:

- Uses techniques like matching algorithms, fuzzy logic, and machine learning.
- Resolves differences in names, addresses, emails, or IDs.
- Helps build Unified Profiles by merging fragmented data.
- Critical for personalization, fraud detection, and data quality.

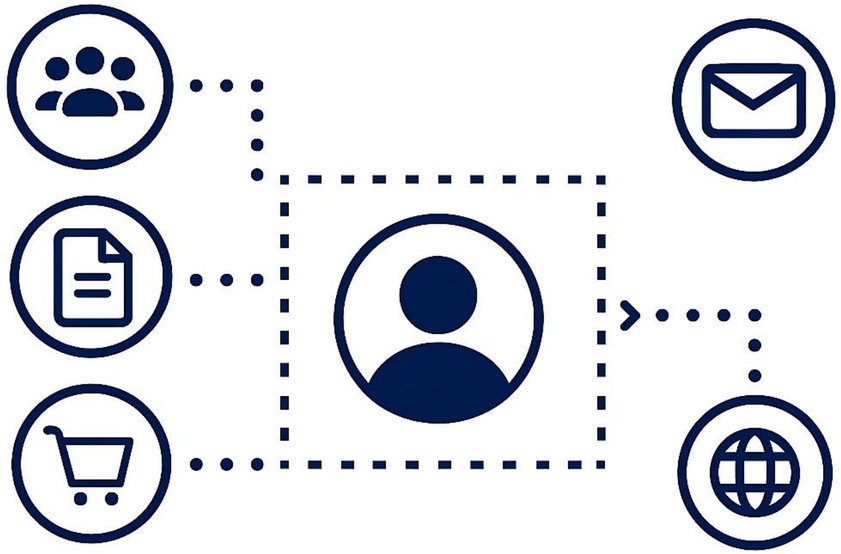
Example Use Case:

A customer might appear as “**Kapil Batra**,” “**K. Batra**,” and “**Kapil B**.” across different systems. Identity resolution links these records into one unified customer profile.


IDENTITY RESOLUTION
salesforce
Salesforce
Data Lake
Amazon S3
Google Drive
Unified
Individual

Unified Profiles

A Unified Profile is a comprehensive, 360-degree view of an entity - such as a customer, product, or asset. Created by integrating data from multiple sources into a single, consistent, and enriched record within the data cloud.



SF BOLT

Check the video description for Giveaway details!