

Evaluating Salesforce Data Migration Strategies: A Comparative Study of Traditional ETL and Salesforce Data Cloud

Ms. Priya Sunil Jagtap

SSMS's Institute of Management and Research, Aranyeshwar, Parvati, Pune

1priyajagtap6@gmail.com

Abstract

Salesforce data migration plays a vital role in enabling organizations to successfully adopt Salesforce as their primary Customer Relationship Management (CRM) platform. Selecting the right migration approach is crucial to ensure data accuracy, system performance, and longterm scalability. This paper presents a comparative study of **traditional Extract–Transform–Load (ETL) methods** and **Salesforce Data Cloud** as two distinct approaches to Salesforce data migration. Traditional ETL tools focus on batch-based data movement with extensive preprocessing outside Salesforce, while Salesforce Data Cloud provides a native, real-time, and unified data integration framework within the Salesforce ecosystem. The study evaluates both approaches based on key factors such as data complexity, transformation capabilities, implementation cost, scalability, and maintenance effort. Case examples from the **financial services, retail, and healthcare** sectors are used to demonstrate practical migration scenarios and decision-making considerations. The paper concludes with key insights to help organizations select an appropriate migration strategy aligned with their business needs, data volume, and digital transformation goals.

Introduction

Data migration is not merely a technical exercise but a foundational enabler of digital transformation. Organizations modernizing their CRM systems face challenges in ensuring data accuracy, integrity, compliance, and real-time availability. Traditionally, ETL tools such as Informatica, Talend, and IBM DataStage have been widely used for enterprise-scale migrations. However, Salesforce Data Cloud introduces a new paradigm, focusing on real-time harmonization, identity resolution, and cloud-native scalability within the Salesforce ecosystem.

This paper compares traditional ETL and Salesforce Data Cloud approaches to Salesforce data migration, providing a structured evaluation of their benefits, costs, and applicability across industries.

Problem Statement

Enterprises implementing Salesforce as their core CRM system face significant challenges in selecting an appropriate data migration approach. Traditional Extract–Transform–Load (ETL) tools provide strong data transformation and control capabilities but often require high infrastructure costs, skilled resources, and extended implementation timelines. In contrast, Salesforce Data Cloud enables real-time data ingestion and harmonization within the Salesforce ecosystem but may limit flexibility for organizations with diverse, non-Salesforce data architectures. The absence of a clear comparative framework makes it difficult for organizations to determine which approach best aligns with their data complexity, cost constraints, and long-term digital transformation goals. Therefore, a systematic comparison of traditional ETL methods and Salesforce Data Cloud is required to guide enterprises in making informed Salesforce data migration decisions.

Research Objectives

1. To analyze the differences between traditional ETL and Salesforce Data Cloud approaches to data migration.
2. To compare the cost, efficiency, and scalability of both methods.
3. To evaluate industry-specific use cases and outcomes.

Research Questions

1. What are the key differences between traditional ETL and Salesforce Data Cloud in terms of features and implementation?
2. How do the costs of migration differ across the two approaches?
3. Which industries benefit most from ETL, Data Cloud, or hybrid approaches?

Methodology

This research is based on secondary data analysis of Salesforce documentation, Gartner reports, peer-reviewed journals, and industry case studies. Comparative analysis and case-based evidence are used to draw insights into the strengths and trade-offs of ETL and Data Cloud.

Research Question 1:

1. What are the key differences between traditional ETL and Salesforce Data Cloud in terms of features and implementation?

Traditional Extract–Transform–Load (ETL) tools and Salesforce Data Cloud represent fundamentally different paradigms for Salesforce data migration and integration. Their differences can be observed across architecture, data handling capabilities, implementation complexity, and operational flexibility.

From an architectural perspective, traditional ETL tools operate as external systems that extract data from multiple source applications, apply transformations outside Salesforce, and load the processed

data into Salesforce objects. These tools are typically infrastructure-heavy and rely on scheduled batch jobs. In contrast, Salesforce Data Cloud is a native Salesforce platform designed to ingest, harmonize, and activate data in near real time. It leverages Salesforce's metadata-driven architecture and is tightly integrated with other Salesforce clouds, reducing the need for external middleware.

In terms of data processing and transformation, traditional ETL tools offer advanced transformation capabilities, including complex joins, aggregations, custom business rules, and multi-step workflows. This makes them well-suited for large-scale legacy migrations involving highly structured and unstructured data from heterogeneous systems. Salesforce Data Cloud, however, focuses on data harmonization rather than deep transformation. It standardizes data using predefined data models and identity resolution rules, which simplifies implementation but may limit customization for complex transformation scenarios.

Implementation effort and skill requirements also differ significantly. Traditional ETL implementations demand specialized technical expertise, extensive mapping documentation, infrastructure provisioning, and rigorous testing cycles. These projects often have longer timelines and higher upfront costs. Salesforce Data Cloud implementations are comparatively faster due to native connectors, declarative configuration, and reduced dependency on external systems. However, they require strong knowledge of Salesforce data models, identity resolution concepts, and licensing considerations.

Scalability and performance present another area of differentiation. Traditional ETL tools are optimized for bulk data movement and historical data migration, making them suitable for one-time or periodic large-volume loads. Salesforce Data Cloud is designed for continuous data ingestion and real-time unification, enabling organizations to maintain up-to-date customer profiles and analytics but potentially increasing operational costs at scale.

Finally, vendor dependency and ecosystem alignment influence adoption decisions. Traditional ETL tools are vendor-agnostic and can support multi-platform enterprise architectures. Salesforce Data Cloud, while highly efficient within the Salesforce ecosystem, introduces a higher degree of platform dependency, which may be a constraint for organizations with diverse technology stacks.

Overall, traditional ETL tools emphasize control, flexibility, and transformation depth, whereas Salesforce Data Cloud prioritizes speed, real-time data availability, and seamless Salesforce integration. The choice between the two depends on organizational data complexity, transformation requirements, and long-term platform strategy. **Comparative Analysis**

Parameter	Traditional ETL	Salesforce Data Cloud
Scope	Cross-platform, multiple source systems	Salesforce-centric, with native Salesforce integration
Transformation	Advanced, complex transformation logic	Limited, focused on harmonization and identity resolution

Scalability	Dependent infrastructure	on	Cloud-native, scalable
Real-time Capability	Mostly batch-oriented		Real-time ingestion and harmonization
Expertise Required	Skilled ETL developers		Salesforce administrators/architects
Compliance & Governance	Flexible but complex		Built-in Salesforce governance

Research Question 2:

2. How do the costs of migration differ across the two approaches?

The costing structure of traditional ETL-based migration and Salesforce Data Cloud can be analyzed across multiple cost dimensions, including licensing, infrastructure, implementation, and operational expenses.

Costing Structure of Traditional ETL Methods:

Traditional ETL-based Salesforce migration follows a **capital-intensive and project-driven cost model**, with most expenses incurred upfront.

1. Tool Licensing Costs

Enterprises typically incur annual or perpetual license fees for commercial ETL tools. These costs vary based on data volume, number of connectors, and environments (development, testing, production).

2. Infrastructure and Platform Costs

ETL tools require dedicated infrastructure such as cloud virtual machines or onpremises servers for data extraction, transformation, staging, and load processing. Additional costs include storage, compute scaling, and network bandwidth.

3. Implementation and Development Costs

Significant effort is required for data profiling, mapping design, transformation logic development, error handling, and reconciliation. Skilled ETL developers and data architects increase labor costs.

4. Testing and Validation Costs

Multiple testing cycles—including unit testing, system integration testing, and user acceptance testing—are necessary to ensure data accuracy and compliance, adding to project duration and cost.

5. Ongoing Maintenance Costs

Post-migration, organizations must maintain ETL workflows, update transformation rules when source systems change, and manage tool upgrades and monitoring.

Overall, traditional ETL costs are largely **fixed and front-loaded**, making them more predictable for one-time or phased migration initiatives.

Costing Structure of Salesforce Data Cloud

Salesforce Data Cloud follows a **subscription-based, consumption-driven cost model**, with costs distributed over time.

1. Platform Licensing and Subscription Costs

Data Cloud pricing is based on data ingestion volume, storage, harmonization, and activation usage. Costs increase with higher data frequency and real-time processing requirements.

2. Implementation and Configuration Costs

Initial setup involves data modeling, identity resolution configuration, and mapping to Salesforce objects. These activities are largely declarative, reducing development effort and implementation timelines.

3. Data Consumption and Storage Costs

Ongoing costs depend on the volume of data ingested, retained, and activated for analytics or personalization. High data velocity or long retention periods can significantly impact costs.

4. Operational and Usage-Based Costs

Continuous ingestion, real-time unification, and downstream activation (e.g., Marketing Cloud, Service Cloud) contribute to recurring operational expenses.

5. Training and Enablement Costs

Teams require training on Data Cloud concepts such as data streams, identity resolution, and harmonized data models, adding indirect costs.

Salesforce Data Cloud costs are **variable and scalable**, offering lower initial investment but potentially higher long-term expenses as data usage grows.

Comparative Cost Structure Summary

Cost Dimension	Traditional ETL Methods	Salesforce Data Cloud
----------------	-------------------------	-----------------------

Cost Model	Fixed, upfront (CAPEX-heavy)	Subscription & consumption-based (OPEX)
Licensing	External ETL tool licenses	Salesforce Data Cloud subscription
Infrastructure	Required (servers, storage, compute)	Minimal external infrastructure
Implementation Effort	High (custom development)	Moderate (declarative configuration)
Cost Predictability	High after implementation	Variable, usage-dependent
Long-term Cost Impact	Lower after migration completion	Increases with data volume & frequency

Research Question 3:

3. Which industries benefit most from ETL, Data Cloud, or hybrid approaches?

The suitability of traditional ETL methods, Salesforce Data Cloud, or a hybrid approach varies significantly across industries due to differences in data volume, regulatory requirements, integration complexity, and business objectives. This section presents realistic industry-based use cases to illustrate how organizations can align migration strategies with sector-specific needs.

Financial Services Industry

Financial services organizations, including banks and insurance providers, typically manage highly regulated, sensitive, and historically rich datasets originating from multiple legacy systems. These datasets often require extensive transformation, cleansing, and reconciliation before migration to Salesforce.

Traditional ETL approaches are particularly beneficial in this industry due to their strong transformation capabilities, auditability, and control over data processing logic. ETL tools enable batch-based migration of large historical datasets while ensuring compliance with regulatory standards and data lineage requirements. Salesforce Data Cloud can complement ETL by supporting near real-time ingestion of transactional or customer interaction data for analytics and personalization. As a result, a **hybrid approach** is commonly adopted, where ETL handles core migration and historical data, while Data Cloud supports ongoing data harmonization and real-time insights.

Retail and E-Commerce Industry

Retail organizations deal with high-volume, high-velocity customer data generated from point-of-sale systems, e-commerce platforms, loyalty programs, and digital channels. Realtime customer engagement and personalization are key competitive factors in this sector.

Salesforce Data Cloud provides significant advantages for retail organizations by enabling real-time ingestion and unification of customer data across multiple touchpoints. Its identity resolution and harmonized customer profiles support personalized marketing, dynamic offers, and customer journey analytics. Traditional ETL methods are typically used for migrating historical sales and customer data during initial Salesforce implementation. Postmigration, Data Cloud becomes the primary integration platform. Therefore, retail organizations benefit most from a **Data Cloud–centric or hybrid approach**, depending on the complexity of legacy systems.

Healthcare and Life Sciences Industry

Healthcare organizations manage patient, provider, and operational data across clinical, administrative, and regulatory systems. Data accuracy, security, and compliance with healthcare regulations are critical considerations.

Traditional ETL methods are well-suited for healthcare data migration due to their ability to perform deep transformations, anonymization, and validation of sensitive data. ETL tools support controlled, batch-based migration of electronic health records (EHRs) and operational data into Salesforce for patient engagement and care coordination use cases. Salesforce Data Cloud can be selectively used to integrate non-clinical data such as appointment interactions, patient communications, and engagement metrics. Consequently, healthcare organizations typically benefit from a **traditional ETL-led or hybrid approach**, with limited Data Cloud usage for real-time engagement scenarios.

Manufacturing and Industrial Sector

Manufacturing organizations rely on data from enterprise resource planning (ERP), supply chain management, and Internet of Things (IoT) systems. These data sources often involve complex schemas and batch-based processing.

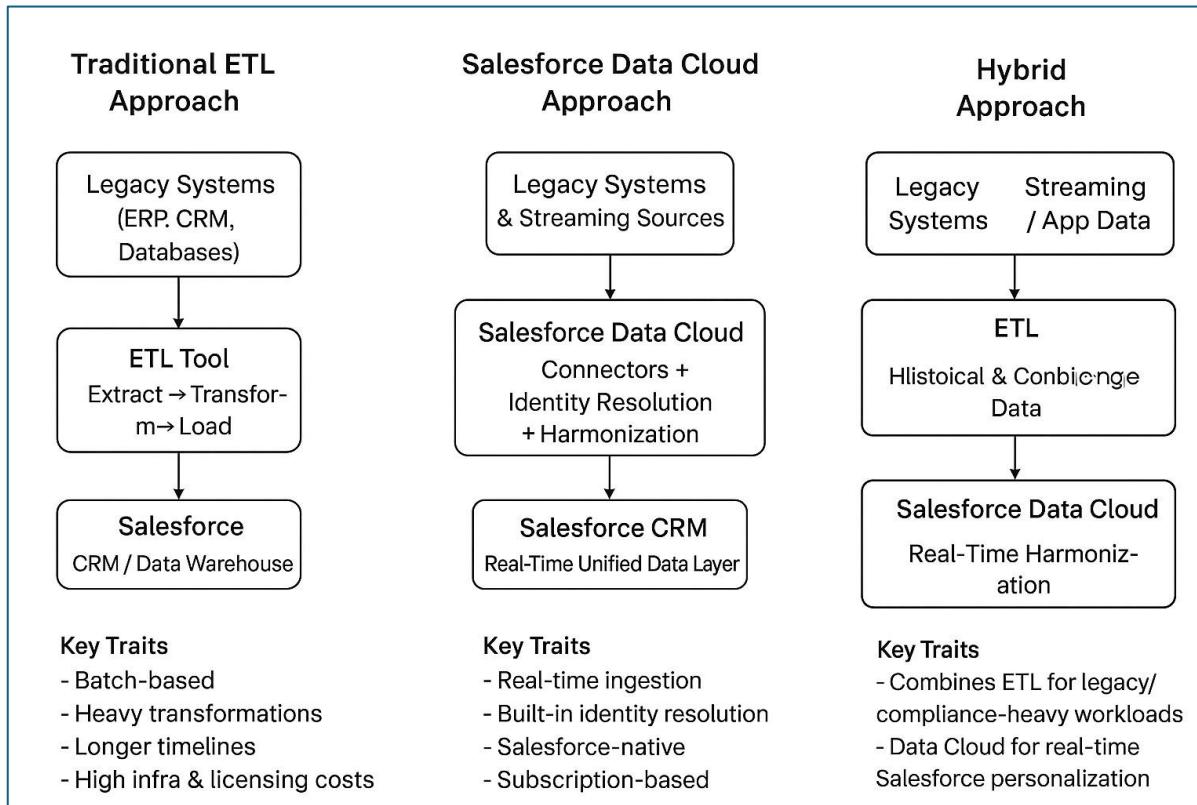
Traditional ETL tools are effective for migrating structured operational data from ERP and supply chain systems into Salesforce for account management and service analytics. Salesforce Data Cloud can add value by integrating customer interaction and service data for predictive maintenance and customer support use cases. A **hybrid approach** allows manufacturers to retain transformation control while leveraging Salesforce-native analytics capabilities.

Media, Telecom, and Digital Services

Media and telecom companies manage large-scale, event-driven datasets with frequent updates and customer interactions across digital channels.

Salesforce Data Cloud is particularly effective in this sector due to its real-time data ingestion and activation capabilities. It supports unified customer views, churn prediction, and personalized engagement

at scale. Traditional ETL is generally limited to initial migration of legacy billing or subscriber data. Therefore, these industries benefit most from a **Data Cloud-first approach** with minimal ETL dependency.



Conclusion

Data migration is a cornerstone of digital transformation, and organizations must balance cost, compliance, speed, and long-term value when selecting a strategy. Traditional ETL is effective for heterogeneous, compliance-heavy systems but comes at a high cost. Salesforce Data Cloud provides agility and real-time insights but is Salesforce-centric. Hybrid models, allow enterprises to combine the strengths of both approaches.

Ultimately, the right choice depends on industry context, existing IT landscape, and strategic priorities.

References

Inmon, W. H. (2005). *Building the data warehouse* (4th ed.). John Wiley & Sons.

Kimball, R., & Caserta, J. (2011). *The data warehouse ETL toolkit: Practical techniques for extracting, cleaning, conforming, and delivering data*. John Wiley & Sons.

Gartner. (2023). *Magic quadrant for data integration tools*. Gartner, Inc. Gartner. (2024). *Market guide for customer data platforms*. Gartner, Inc.

Forrester Research. (2023). *The Forrester Wave™: Enterprise data fabric*. Forrester Research, Inc.

McKinsey Global Institute. (2023). *Scaling advanced analytics in healthcare*. McKinsey & Company.

Deloitte. (2023). *Modernizing CRM data architecture in financial services*. Deloitte Insights.

Salesforce, Inc. (2024a). *Salesforce data cloud architecture and implementation guide*. Salesforce Documentation.

Salesforce, Inc. (2024b). *Salesforce data cloud pricing and consumption model overview*. Salesforce White Paper.

TM Forum. (2023). *Real-time data platforms in telecom and digital services*. TM Forum.

Joshi, Aarya. (2025). myresearchgo, Volume 1, October issue 7 , 2025. 10.64448/myresearchgo..vol.1.issue.7.00.
https://www.researchgate.net/publication/399022458_myresearchgo_Volume_1_October_issue_7_2025

Joshi, Aarya. (2025). Volume 1, September Issue 6, 2025. 1. 170.