PROFESSIONAL SERVICES

Snowflake Optimization with Iceberg

Reduce total cost of ownership with an open data lakehouse

As enterprises scale their analytics and AI workloads, they face growing challenges around cost efficiency, data interoperability, and architecture flexibility. Cloudera's Open Data Lakehouse provides a seamless way to optimize your total cost of ownership (TCO) while maintaining a unified, open architecture. By leveraging Apache Iceberg as a shared open table format, organizations can offload storage and data processing to Cloudera while continuing to use their existing tools (e.g. Snowflake environment) as the primary query interface - driving cost savings without disrupting user experience or workflows.

SmartLakehouse with Cloudera Professional Services

SmartLakehouse is a hands-on professional services engagement designed to help enterprises optimize their Snowflake environment, by extending it to an Open Data Lakehouse architecture on Cloudera and Iceberg to reduce costs and minimize data copies. Cloudera Professional Services will guide you in adopting Iceberg, offloading storage, data ingestion, and data processing to Cloudera while keeping Snowflake as your primary query interface – reducing costs without changing the end-user experience.

SmartLakehouse: Engagement Outcomes

Lakehouse foundation	Migrate use case	Cloudera education
Installation and setup of Cloudera Data Services (if required*)	Implement/migrate 1 use case (workload and data) to Cloudera	2x On Demand training seats for 1 year. Access self-paced learning

*for Public Cloud engagements, Cloudera's cloud specialist team will support the customer with PaaS onboarding before Professional Services engagement kick-off

SmartLakehouse bundle is an outcome-driven engagement that includes a reference implementation of one (1) Lakehouse data pipeline use case (to be jointly determined by the customer and Cloudera):

- Use case analysis & solution design
- Cloudera Data Services platform setup (if required)
- 1x Use case implementation – setting the foundation for a modern architecture
- Value analysis to determine estimated TCO reduction
- 2x On-Demand Training seats for Cloudera Education

Tailored packages are available for other Open Data Lakehouse use cases.

Why choose Cloudera Professional Services?

- **Extensive technical expertise —** Cloudera's global services team brings unmatched experience designing, architecting and implementing open source data & AI solutions for the enterprise.
- **Proven delivery accelerators —** Cloudera combines deep experience with a toolkit of accelerators to cut delivery from months to weeks — reducing risk, driving innovation, and faster ROI on data investments.
- **Comprehensive service offerings —** From on-premises to cloud, Cloudera supports your entire data journey — ranging from solution design, to implementation and production readiness across Data In Motion, Data Lakehouse, & AI.
- **Enterprise-grade design principals —** Helping build secure, portable, and well documented data applications — ensuring governance, hybrid cloud flexibility, and retained knowledge.
- **Driven by customer success —** Cloudera is built around the success of its customers — driving measurable outcomes and aligning every engagement to client goals and long-term impact.

Services engagement schedule

- **SmartLakehouse:** ~6weeks
- **Tailored packages** also available

Start the journey today

Book a free 1-day workshop with our solution experts. Contact your Cloudera representative for more details.

Learn more at cloudera.com

Services Approach & Methodology

A SmartLakehouse engagement follows a structured approach to help you establish an Open Data Lakehouse that works well with your existing Snowflake deployment (below).

SmartLakehouse is just the beginning. This foundational engagement jumpstarts your Open Data Lakehouse journey, providing a cost-efficient and interoperable architecture. As your data needs grow, Cloudera can help you scale, optimize, and expand your lakehouse to support broader analytics and AI workloads.

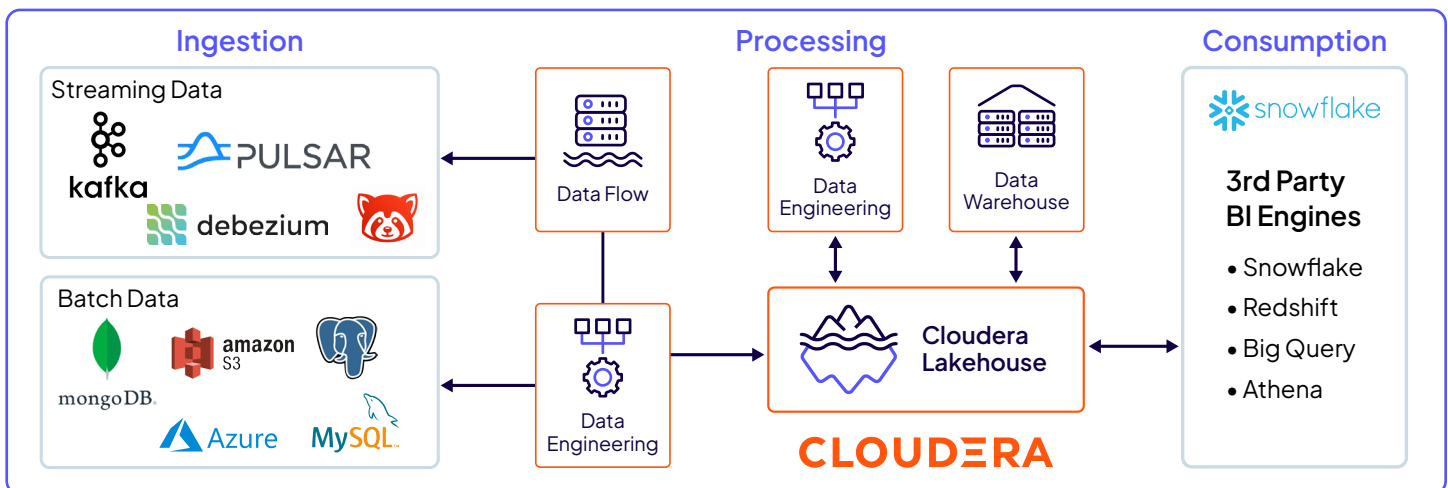| Use case analysis & solution design	Platform & use case setup	Use case implementation for Snowflake optimization		Hand-over & next steps
• Analysis of customer existing environment • Use case solution design and integration planning	• Install Cloudera Data Services (if required) • Lakehouse integration validation and testing	• Export snowflake storage to leverage Cloudera Iceberg • Data access and integration from Cloudera to Snowflake	• Build NiFi flow for use case to offload data ingestion • Build Spark app for use case - to offload data processing	• Customer enablement • Demo & customer handover • Documentation • Next steps planning

Optimizing Your Data Architecture With Cloudera

Cloudera's Open Data Lakehouse, built on Apache Iceberg, delivers a scalable, cost-efficient, and interoperable architecture that eliminates data silos and simplifies enterprise data management. By unifying storage, compute, security, and governance across all analytics in hybrid and multi-cloud environments, organizations can seamlessly support BI, AI/ML, and data engineering workloads with flexibility and control.

Key benefits of this architecture include:

- **Break Down Data Silos —** Eliminate redundant ETL processes and data copies by enabling direct access to Iceberg tables across multiple engines.

- **Cost Efficiency & Storage Optimization —** Reduce operational complexity and storage costs by up to 50% with a unified, open architecture.

- **Multi-Engine Flexibility —** Use the best tools for the job with Hive, Spark, Impala, Trino, Flink, and NiFi, or any 3rd party BI engine all operating seamlessly on Iceberg tables.

- **Enterprise-Grade Security & Governance —** Maintain unified access controls, data lineage, and compliance across distributed environments.

- **Hybrid & Multi-Cloud Ready —** Deploy and manage data across on-prem, public, and private clouds with a centralized control plane.

Leverage Cloudera for your snowflake data ingestion and processing while keeping end-user consumption in Snowflake.

CLOUDERA

Cloudera is the only true hybrid platform for data, analytics, and AI. With 100x more data under management than other cloud-only vendors, Cloudera empowers global enterprises to transform data of all types, on any public or private cloud, into valuable, trusted insights. Our open data lakehouse delivers scalable and secure data management with portable cloud-native analytics, enabling customers to bring GenAI models to their data while maintaining privacy and ensuring responsible, reliable AI deployments. The world's largest brands in financial services, insurance, media, manufacturing, and government rely on Cloudera to be able to use their data to solve the impossible—today and in the future.

To learn more, visit Cloudera.com and follow us on LinkedIn and X.