snowflake®

MANUFACTURING BEST PRACTICES GUIDE: UNLOCKING SAP DATA ANALYTICS IN SNOWFLAKE

CHAMPION GUIDES

EBOOK

TABLE OF CONTENTS

INTRODUCTION

Advanced analytics serve as a powerful tool for manufacturers to become more agile and efficient in their operations. But SAP customers in manufacturing often struggle with inefficiencies caused by multiple ERP and analytic server instances. The high rate of M&A activity in this industry creates a fragmented resource planning environment that causes data silos and makes enterprise-wide visibility virtually impossible. ERP consolidations often cause a loss of historical data from legacy systems, making it challenging to perform predictive analytics or develop ML models. In addition, without a flexible, multi-cloud solution, manufacturers can't fully harness the advantages of integrating ERP data with shop floor data, or enable robust business continuity.

With changes in SAP's warehousing strategy, SAP manufacturing customers must patch together multiple technologies for data warehousing, including SAP BW/4HANA, SAP BW Bridge, SAP Datasphere, SAP Analysis for Office and SAP Analytics Cloud for data reporting. This fragmented environment leads to increased complexity, data silos, higher costs and the need for specialized skills to manage and integrate different systems.

For all of these reasons, SAP customers in manufacturing are turning to Snowflake to leverage a modern, user-friendly data platform that can manage their data and analytics workloads in a simple, flexible and cost-effective manner. With Snowflake, manufacturers get a cloud-first data platform that integrates SAP and non-SAP ERP data in one governed location while delivering adaptability, scalability, security and secure data collaboration. The Snowflake Manufacturing Data Cloud provides native support for semi-structured and unstructured data, and Snowflake's unique architecture provides unlimited storage and computation and NoETL data sharing.

This paper outlines how manufacturers that use SAP can unlock advanced analytics by moving ERP data to Snowflake, and shares best practices for the migration process. SAP customers who follow best practices while moving to Snowflake can ensure their data is migrated accurately, effectively and promptly. Following best practices can also ensure user buy-in and help your organization make its data architecture future-ready.

FIVE BUSINESS PROCESSES THAT CAN BE IMPROVED USING SNOWFLAKE

Plan: Increase forecasting accuracy by integrating and enriching SAP data (such as historical sales data, inventory levels and other operational data) with third-party data from Snowflake Marketplace (such as market trends, demographic and weather data).

Source: Proactively manage supply chain risk by seamlessly merging data from SAP systems (such as purchase, inventory and schedule data) with data from external suppliers and logistics partners to achieve near real-time insights.

Make: Improve product yield and factory efficiency by combining and analyzing SAP, IoT, shop floor and other operational data in one place and applying machine learning (ML) to unlock advanced insights.

Deliver: Optimize timely customer fulfillment by uniting SAP delivery, inventory and cost data; data from various systems including planning and forecasting, warehouse and transportation management; and third-party logistics and freight rate data from Snowflake Marketplace.

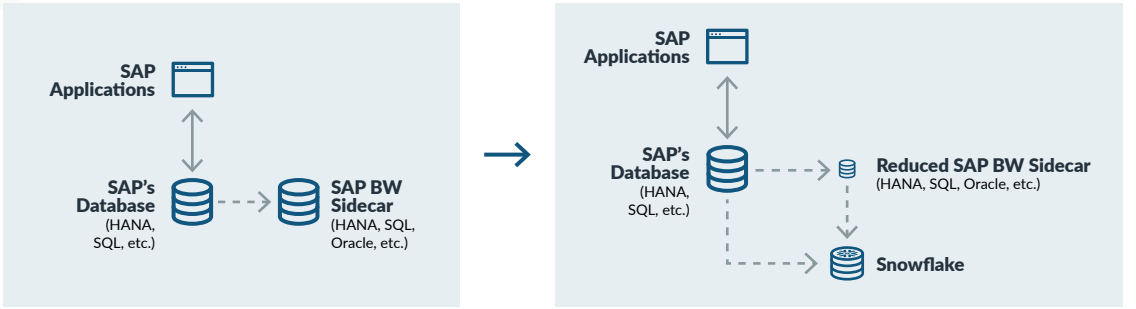
Service: Enhance the value and service of end products by integrating SAP supply chain data with IoT data and incorporating context (such as weather and geospatial data from Snowflake Marketplace) in Snowflake for ML-based predictive maintenance.

CHAPTER 1: MIGRATING DATA FROM SAP TO SNOWFLAKE DATA CLOUD

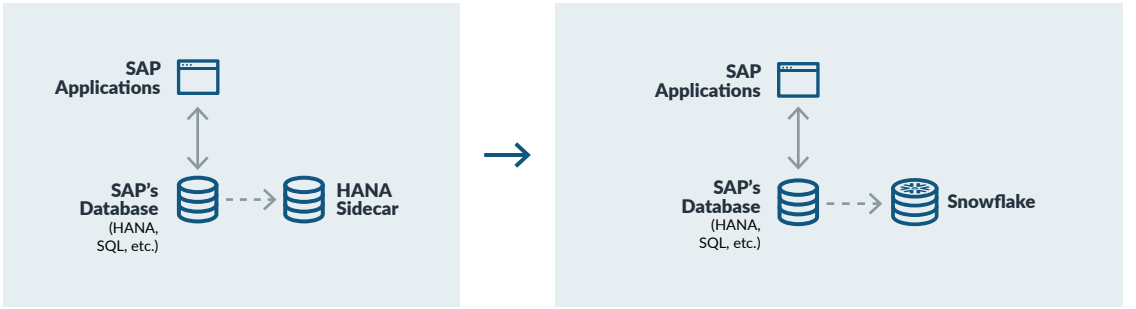
Manufacturers rely on systems that can accurately capture their processes, products and dependencies, efficiently handle transactions, scale well and provide cost-effective analytical capabilities. While SAP ERP systems are efficient at logging transactions, SAP analytical systems may fall short due to resource-intensive architectures. Similarly, SAP PaaS solutions, which utilize the same architecture, struggle with scalability and exhibit the same outage patterns as on-premises systems during upgrades and capacity increases.

As a result, manufacturers who initially chose SAP HANA data warehousing are transitioning to Snowflake, and those who depend on SAP Business Warehouse (BW) are shifting numerous workloads to Snowflake.

SAP BW Sidecar for Reporting and Analytics



SAP HANA Sidecar for Reporting and Analytics



Snowflake provides a robust, flexible, cloud-agnostic platform for handling and analyzing all enterprise and manufacturing data with a focus on scalability, data protection and a range of built-in features that enhance user experience and functionality.

Data management: Complex hierarchical data structures, such as bills of materials, can be difficult to manage and query. By using recursive common table expressions (CTEs) and storing these hierarchies as arrays, Snowflake simplifies the management of these structures, making it easier to work with data.

Data protection: Snowflake ensures data security through always-on encryption, row-level policies, masking, tagging and quality checks. Snowflake Time Travel keeps track of every change to data or data structure and includes features such as "UNDROP" that allow recovery from human error.

Data availability: Snowflake guarantees constant data availability. There are no maintenance windows when data is inaccessible. To ensure data durability and availability, Snowflake stores three copies of your data within a single availability zone. Snowflake can also create copies of your data in different geographical regions and across different cloud service providers—beneficial for disaster recovery, reducing latency and sharing data with business partners that may have standardized on different cloud providers.

Scalability: Snowflake's elastic architecture allows it to scale up and down on demand so that it can adjust its resources based on the workload, optimizing resource usage and cost. Snowflake also provides efficient data ingestion with Snowpipe that enables near real-time data loading into Snowflake's database.

Data sharing: Snowflake allows secure, real-time data sharing across ecosystems without the need for costly and inefficient ETL (extract, transform, load) processes.

Performance: Snowflake isolates compute clusters while leveraging one copy of the data, eliminating instability across the account and preventing bottlenecks caused by resource-intensive processes.

ML & AI workloads: Snowflake's single, integrated platform allows for real-time data sharing, efficient data management and scalability to allow data scientists to easily access and analyze large data sets, apply ML algorithms and generate valuable insights.

CHAPTER 2: SETUP

Setup is an important step in transitioning from an existing SAP data landscape to a Snowflake environment.

- **First,** analyze your current SAP data landscape, which includes understanding the existing data structures, schemas, code, security measures and governance policies and identifying processes that are no longer used.
- **Then,** set up the target environment in Snowflake, which includes creating equivalent schemas, converting code and setting up similar security and governance measures in the Snowflake environment.

SAP and Snowflake structure and handle data differently, so part of the transition process involves mapping or translating SAP data structures (objects) to equivalent structures in Snowflake. This process is easier than one might expect, as there's compatibility and straightforward conversion paths between the two systems.

Here's an overview of SAP to Snowflake object equivalence:

SAP	Snowflake
InfoAerea	Schema
InfoObject	Column
InfoObject with master data	Table
DSO / ODS	Table
Calculation / Analytic / Attribute Views	View
Composite Provider	View
Routine	Function
ABAP Program	Stored Procedure
Recordmode / Change Table	Stream
Process Chain	Task
InfoCube	CTAS / View
InfoCube with aggregate	Materialized View
PSA Table	Stage

Read more: An object-by-object migration guide from SAP to Snowflake.

Snowflake also simplifies environment management with **Zero-copy Cloning,** allowing you to create a copy of your data without duplicating it. This has several advantages: it eliminates the need for time-consuming data copying, saves costs by eliminating the need for additional storage, increases resilience and ensures your data remains secure while promoting collaboration.

With Zero-copy Cloning, you can create temporary environments as needed from the production environment instead of maintaining separate development (dev), testing (test) and quality assurance (quality) environments. Snowflake allows you to manage one environment (production) and create others as needed, rather than managing multiple environments in parallel. This can drastically simplify environment management and save resources.

CHAPTER 3: ACQUIRE DATA

Snowflake's solution partners make it simple for manufacturing companies to ingest data from SAP into Snowflake. Look specifically for partners who have native SAP connectors and use Snowpipe Streaming, a Snowflake service, to allow for continuous, near real-time data ingestion. Snowpipe Streaming can write directly to Snowflake tables without the need to stage files, reducing costs and latency. The Snowflake's Connector for Kafka also uses Snowpipe Streaming, providing another efficient method for data ingestion.

Partners that have native SAP connectors in Snowflake include:

- **Informatica:** Informatica's **Enterprise Data Integrator** (in public preview) offers a seamless and frictionless experience to integrate data from SAP into Snowflake. Informatica's Intelligent Data Management Cloud also has a feature called **SuperPipe** which uses Snowpipe Streaming for ingestion of SAP data, enabling 3.5x faster ingestion performance.
- **SNP:** **SNP Glue** brings a flexible, native, and optimized, SAP Integration solution to the Snowflake Data Cloud. Providing customers with a 10x cost reduction in Snowflake data processing associated with SAP data ingestion, and has empowered them to drastically reduce data latency while improving SAP data availability. Access SNP Glue on the **Snowflake Marketplace**.

Other partners that make it easy to ingest data from SAP into Snowflake include:

- **Fivetran:** Fivetran's SAP connectors are capable of moving large volumes of data in real time, replicating data from a variety of SAP sources quickly and securely.
- **Matillion:** Matillion's SAP Connector is built to simplify data teams' loading, transformation, synchronization and orchestration of data pipelines to deliver business-ready data faster, drive better decision-making, and improve operational efficiencies and overall business success.
- **Qlik:** Qlik extracts relevant data from SAP, ingests it into Snowflake in real time and transforms it to analytics-ready data. Qlik's solution automates the design, implementation and updates of data models while minimizing the manual, error-prone design processes of data modeling, ETL coding and scripting.

Find a Snowflake Solutions Partner

An **append-only, soft delete** data management strategy can improve performance and data consistency during data ingestion. Here, new data is only appended to the existing data set (never modified or deleted), and "deletions" are handled by marking records as deleted ("soft delete") rather than removing them. For high-churn tables with frequent changes, designing the ingestion process to only append data with soft deletes makes the data landing process faster and more efficient. Transformation and deduplication of the raw data layer is addressed in the transformation stage.

In addition, Snowflake stores data in **immutable micro-partitions,** meaning they cannot be modified once written, providing a reliable and consistent view of the data. They also allow for unlimited scalability and massive optimization through pruning (ignoring irrelevant micro-partitions during query processing).

Snowflake's native storage also features **Snowflake Time Travel,** which allows you to view data and metadata as it was at any point in the past. This feature enhances data protection, aids in compliance and auditing, and facilitates debugging and testing. If data is accidentally deleted or modified, you can use the Time Travel feature to restore the data to its previous state. Time Travel can be extremely useful for audit and compliance purposes as you can track changes over time, identifying who made changes and when. And if something goes wrong during the development and testing process, developers can use the feature to inspect the state of the data at different points in time.

CHAPTER 4: TRANSFORM

In modifying the source data from SAP to match the target format in Snowflake, manufacturing companies can use rules, merges, lookup tables or other conversion methods to ensure that data in Snowflake is accurate, consistent and in the right format for the company's needs. After replicating the data in Snowflake, you can work with and model the data however you prefer. For example, you can transform the replicated data into a star schema using views that encapsulate the business logic required, making data modeling more efficient and the data more meaningful for business analysis. Once these views have been created, they can be directly queried by business intelligence (BI) tools or materialized into tables that can be queried, improving the efficiency of BI processes and leading to more informed decision-making.

Here are some best practices for transforming SAP data to Snowflake:

USE STREAMS, TASKS, AND DYNAMIC TABLES FOR ROBUST TRANSFORMATIONS

When used together, Streams, Tasks, and Dynamic Tables (in public preview) can provide robust, asynchronous, multi-chained directed acyclic graphs (DAGs) for data transformations. You can set up complex sequences of transformations that handle dependencies correctly, execute in parallel where possible and provide robust error handling.

- **Streams** are a way of tracking changes to tables. They keep a record of all inserts, updates and deletions performed on a table, allowing you to see how the data in the table has changed over time.
- **Tasks** refer to scheduled SQL statements that Snowflake runs automatically at specified intervals. These can be used to automate various data transformation and data management operations.
- **Dynamic Tables** allow for more flexible and efficient data transformations. They can adjust structure based on the data that's being processed, which can be particularly useful when dealing with varying data formats or schemas.

INCREASE CONTROL WITH DATAOPS TOOLS

Enterprise DataOps tools such as dbt Cloud, managed by Snowflake partner **dbt Labs,** and Snowflake's data management features allow you to define the sequence of data operations explicitly, giving you more control over your data processes. They are used to manage and orchestrate data flows imperatively and provide a way to specify step by step how the data operations should be carried out. For example, dbt Cloud allows data analysts and engineers to standardize on a SQL-first workflow and promote data development best practices.

Partners such as **Coalesce** help data engineers transform their data at speed with minimal coding by auto-generating Snowflake SQL. It automates the creation of Snowflake features such as Dynamic Tables and Streams and Tasks, and helps data teams operationalize the use of ML-based functions in data pipelines.

OPTIMIZE HIERARCHICAL DATA

Analyzing hierarchical data is difficult and unreliable for legacy data warehouse systems. Snowflake provides a range of features and capabilities that can help optimize the handling and analysis of hierarchical data. For example, with **recursive common table expressions (CTE),** you can iterate through each level of your hierarchy and accumulate results. **Semi-structured data handling** lets you split semi-structured hierarchical data across multiple columns. And **data visualization** allows you to envision and analyze hierarchical data to more easily understand patterns and trends.

Read more: Optimize your hierarchical data with Snowflake

SIMPLIFY AND ENHANCE CURRENCY CONVERSION

There are two optimal ways to perform currency conversion: source the currency table from SAP to perform the conversion logic in SQL, or use Snowflake Marketplace to locate freely available and current currency exchange data. You can join your fact table on date and currency, and get currency rates instantly.

Read more: Currency conversion with Snowflake

CHAPTER 5: ANALYZE AND PREDICT

Through Snowflake's robust partner ecosystem, you can access solutions to optimize reporting and analytics with SAP data. Here are some of the cutting-edge partner tools manufacturers can use to transform, analyze, visualize and share data:

- **Tableau** is an enterprise BI platform that organizations use to build dashboards. Tableau provides interactive data visualization focused on business intelligence.
- **Dataiku, DataRobot** and **H2O.ai** provide tools and infrastructure to design, deploy and monitor ML models and develop AI solutions.

For organizations that prefer building open source solutions, Python, a popular programming language for data analysis, and Streamlit, an open source app framework, are natively available in Snowflake.

Learn more: Snowflake Partner Network

Enhance insights with the Snowflake partner ecosystem

The Snowflake partner ecosystem empowers manufacturers to gain more value from their SAP data in Snowflake, leading to improved operational efficiency, better decision-making and increased competitiveness. Here's how:

Snowflake Marketplace:
Access a wide range of third-party data sets that can be used to enhance insights derived from SAP data and improve forecasting and decision-making. Data sets include market trends, demographic data, weather data, pricing data and more.

Powered by Snowflake:
Use applications and services built on the Snowflake platform that can leverage the full capabilities of Snowflake — including its performance, scalability and security features — to gain powerful solutions for analyzing and utilizing SAP data.

Snowflake technology partners:
Leverage tools and services that are compatible with Snowflake to analyze SAP data more deeply and extract valuable insights. Gain additional capabilities such as advanced analytics, ML algorithms, data visualization tools and more.

System integrators:
Get expert help integrating SAP data with Snowflake and ensuring that data flows smoothly between different systems, which is crucial for gaining a comprehensive view of the business and making informed decisions.

CHAPTER 6:
RETIRE OLD INFRASTRUCTURE

As you migrate data and workloads from SAP to Snowflake, you can simplify your analytics landscape and reduce analytics spend by retiring old SAP infrastructure and reducing licensing and maintenance costs.

You can also leverage Snowflake's cost-efficient storage to archive historical ERP data from ECC/S4. This approach not only frees up HANA resources by relocating historical analysis to Snowflake, but also optimizes the use of your IT infrastructure.

In the context of a merger and acquisition, this strategy can be further extended. The historical data from the ERP systems of both companies can be combined into Snowflake, providing a unified view of the historical data and facilitating more comprehensive analysis and decision-making.

Once the historical data is safely archived in Snowflake, the unneeded ERP systems can be retired. This step can lead to significant cost savings by eliminating the maintenance and operational costs associated with these systems.

CONCLUSION

By adopting the best practices in this paper, manufacturers can migrate their ERP data from SAP to Snowflake in a way that ensures data consistency and integrity, optimizes advanced analytics and increases efficiencies in time and cost.

With Snowflake's unique capabilities, such as Zero-copy Cloning, Time Travel and near-instant scalability, manufacturers can leverage SAP data more strategically and enhance business processes from planning to servicing. And with the help of Snowflake's partners, manufacturers can optimize the value of their SAP data in Snowflake.

Discover how Snowflake Partners can help you unlock SAP data analytics in Snowflake.

ABOUT SNOWFLAKE

Snowflake enables every organization to mobilize their data with Snowflake's Data Cloud. Customers use the Data Cloud to unite siloed data, discover and securely share data, and execute diverse artificial intelligence (AI) / machine learning (ML) and analytic workloads. Wherever data or users live, Snowflake delivers a single data experience that spans multiple clouds and geographies. Thousands of customers across many industries, including 647 of the 2023 Forbes Global 2000 (G2K) as of October 31, 2023, use the Snowflake Data Cloud to power their businesses.

Learn more at **snowflake.com**