TECHNICAL GUIDE

snowflake®

SIMPLIFYING YOUR ANALYTICS LANDSCAPE BY MIGRATING FROM SAP TO SNOWFLAKE



CHAMPION GUIDES

TABLE OF CONTENTS

3 EXECUTIVE SUMMARY

4 PROBLEMS WITH SAP-TO-CLOUD MIGRATION

6 BENEFITS OF OFFLOADING SAP DATA INTO SNOWFLAKE

8 WORKLOAD #1: REPLICATE SAP BW DATA INTO SNOWFLAKE

10 WORKLOAD #2: REPLACE SAP BW

11 WORKLOAD #3: REPLACE SAP HANA DATA WAREHOUSE

12 WORKLOAD #4: KEEP SAP BUSINESSOBJECTS AND SAP ANALYTICS CLOUD

13 SNOWFLAKE ADDS VALUE TO SAP CLOUD MIGRATION

14 ABOUT SNOWFLAKE

EXECUTIVE SUMMARY

SAP products are ubiquitous—77% of the world's transactions revenue touches an SAP system at some point, according to SAP.[1] But 77% of the world's data is not in an SAP system and it shouldn't be. This paper explains why. As organizations increasingly transition critical transactional workloads to cloud-based iterations (often staying with SAP ERP or moving to SAP S/4HANA), SAP customers must, however, take pause and consider replacing their legacy SAP analytic infrastructure—SAP BW and SAP HANA data warehouse. This is because S/4HANA takes back the operational reporting making an SAP-centric warehouse such as SAP Business Warehouse (BW) and HANA obsolete. The value of a data warehouse is adding non-SAP, semi-structured data to a customer's SAP data. SAP cannot do this easily, if at all. Also, the migration of SAP analytics to the cloud often proves complex and rigid in terms of SAP's non-native cloud data architecture and provides no additional value. Snowflake can eliminate these stumbling blocks by providing a platform for SAP customers that is simple, elastic, secure, and cost-effective.



PROBLEMS WITH SAP-TO-CLOUD MIGRATION: COMPLEXITY AND RIGIDITY

Cloud migration helps customers reduce costs, outsource asset management, and alleviate in-house workload burdens. But according to TechReplublic, as SAP customers migrate to the cloud, a great many of them are having specific trouble with regard to their data.[2]

Many existing SAP customers are reviewing the formal move to a hosted, cloud-based SAP S/4HANA suite. But there has been a great deal of hesitation surrounding shifting on-premises SAP BW systems to SAP BW/4HANA because the latter will not deliver the modern computational elasticity, variety, and compute that cloud-native solutions provide. Most companies also have data outside of SAP, so there is no inherent advantage to keeping an on-premises or hosted SAP-centric BW system.

SAP's cloud answer—SAP DataSphere—is fractured. It does not have the maturity of SAP BW, and SAP's approach of using SAP DataSphere in combination with products such as SAP BW, SAP HANA Cloud, and SAP Hana Data Lake increases the complexity of the landscape rather than reducing it. Is there any migration path for customers to leverage their modelling work from SAP BW to SAP DataSphere? Yes; but according to SAP documentation, they must first migrate to SAP BW/4HANA 2.0 SPS07 or higher.[3]

Migrating to SAP BW/4Hana is not a simple matter for two main reasons:

COMPLEXITY

Today's SAP analytics environments have become too complicated, siloed, and expensive. Business users complain about long-running queries. Some organizations running an SAP BW 7.x system dread the jump to SAP BW/4HANA because of licensing (now split from the ERP), sizing, and migration effort. Many companies already offload data from SAP BW to an enterprise data warehouse (EDW) or even to a data lake because it is easier than getting non-SAP data into SAP. Per SAP documentation, those who do manage to load their non-SAP data into SAP BW may unwittingly lose the right to export it again if they don't have an SAP NetWeaver OpenHub license.[4] Business units may be using SAP Business Explorer (SAP BEx) or SAP BusinessObjects Platform (SBOP) to extract data out of SAP BW so they can use it in Excel or some other tool—effectively creating shadow IT instances that bypass governance efforts to prevent the duplication of data.

The variety, volume, and velocity of data being generated makes it too expensive to store data in SAP BW or an EDW, so a data lake becomes the seemingly cheapest immediate option. But now the data is in two different systems, a data warehouse and a lake. Getting data out of the lake for analytic and machine learning workflows requires separate computational resources, and users usually have to land the data in a database anyway for performance reasons.

Moving legacy SAP systems and hosting them in the cloud certainly shifts some of the infrastructure and software operational workload to someone else, but companies still pay for the complexity and the fragility of the data flows through increased costs, monitoring, and downtime. In addition, organizations then face vendor lock-in from both SAP and the cloud provider.

Here is an example of the migration steps from SAP BW 7.3 to BW/4HANA:

1. Upgrade from SAP BW 7.3 to 7.5.
2. Pay for SAP BW/4 license.
3. Pay for SAP HANA license.
4. Migrate all old SAP BW 3.x flows and objects to new advanced data store objects (ADSOs).
5. Rework all flows to take advantage of SAP HANA.
6. Develop a complex storage strategy with hot/warm/cold configuration.
7. Find teams to do the work (The SAP knowledge workforce is retiring or expensive in comparison to SQL developers.)
8. Find a new analytics front end, because SAP BEx tools are not delivered with SAP BW4/HANA.

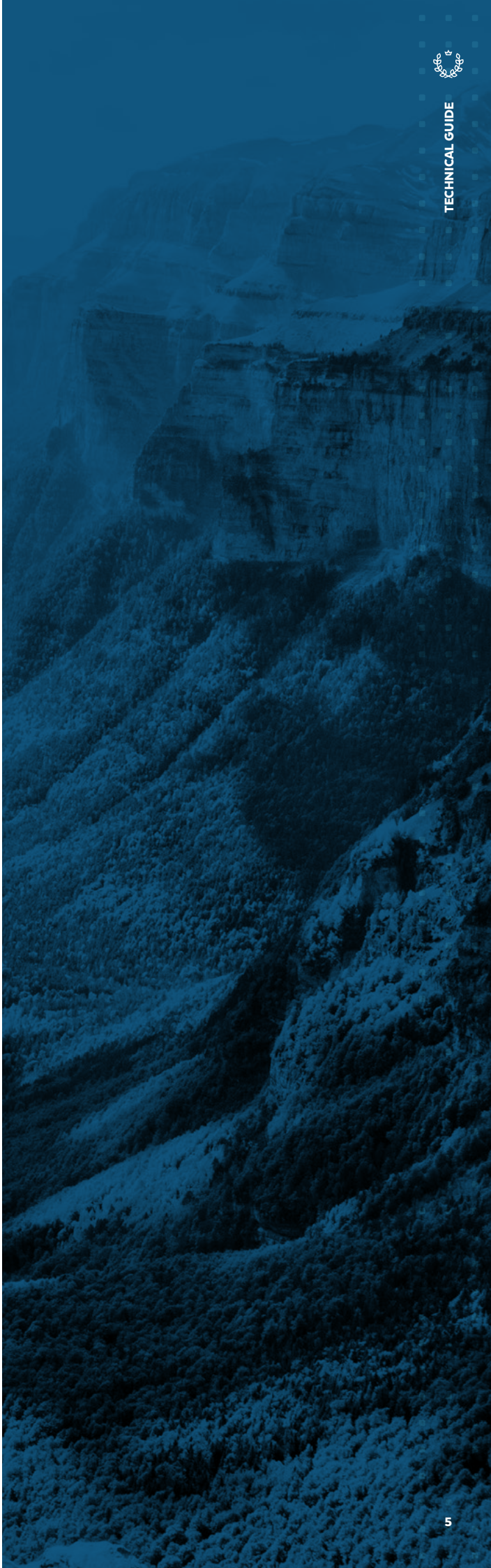
Once all the above steps are complete, the organization can partially migrate to SAP DataSphere—while still keeping an on-premises architecture that, at best, integrates a little better into its data lake.

RIGIDITY

Legacy data architectures are constrained by disk space and computational power. Legacy analytics platforms also require local deployment. Data is pushed up to regional and global machines with batch jobs, remote function calls (RFCs), and even through Microsoft Excel documents.

For example, SAP's on-premises resource planning solution—SAP ERP Central Component (SAP ECC)—feeds to a regional instance of SAP BW, which then feeds to a global instance of SAP BW. File extracts are then often used to share data out to partners, with the accordant governance implications.

Both of these key problems point toward the need for new technologies that can simplify critical SAP data workloads in the cloud at economies of scale.



THE BENEFITS OF OFFLOADING SAP DATA INTO SNOWFLAKE

Snowflake Data Cloud provides a simple and elastic alternative for SAP customers that simultaneously ensures an organization's critical information is protected. Snowflake was built from scratch to run in the cloud—whether it's Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), or a combination of all three. This frees companies from cloud provider lock-in and provides the following benefits:



ARCHITECTURAL SIMPLICITY

Snowflake improves the ease of use of SAP with greater architectural simplicity. The first benefit Snowflake offers SAP customers is making data accessible and actionable by putting it all in the same place. This enables companies to follow the FAIR principles (findable, accessible, interoperable, reusable) as defined in Wikipedia.[5]

Snowflake provides a solution that is cloud provider agnostic and that can ingest data from SAP operations systems (both on-premises and cloud), third-party systems, and signal data. Snowflake also robustly ingests the data—whether it's in structured or semi-structured formats or even if the data structure in the files changes. Gravitational centering of data provides a foundation for customers to build 360-degree views of customers, products, and the supply chain. This includes interweaving data sets from business partners and Snowflake Marketplace through collaboration.

Snowflake manages storage automatically in terms of capacity, compression, statistics, and performance. This eliminates the need for customers to build indexes or do housekeeping. Snowflake provides computational power that is elastic and isolated from other compute clusters used by other teams. This translates into data citizens being able to focus more of their time on tasks that add value to a company's top and bottom line.

In addition, Snowflake seamlessly streamlines the use of SAP data (in combination with other data sources) to enable trusted relevant business content to be accessible to all users. It is designed to handle the full volume, velocity, and variety of data. Users can load data and create reports at the same time. Snowflake also offers a system that has virtually no downtime—no four-hour window every week for upgrades and no system shutdown to increase or decrease compute.

CONVENIENT WORKLOAD ELASTICITY

Snowflake offers scalability to attach and detach services as needed. This delivers convenience, operational efficiency, and cost savings. Splitting compute from storage allows both to scale independently and elastically—leveraging the cloud providers' hardware and compute, without having to deal with their complexity and quirks.

With Snowflake, users just open a view of their data and others can mix it with their own data. They can start with 20 GB and grow to more than 20 PB and back down again. The amount of computational power organizations use scales in seconds and is isolated. For example, this means no more throttling of SAP BW queries and no more license negotiations because SAP HANA is throwing out-of-memory errors.

RELIABLE DATA SECURITY

With Snowflake, simplicity also means data security. All data is always encrypted—in storage or in transit—as a built-in feature of the platform. Data is landed once and views are shared out. This means one copy, one security model, and hundreds of elastic compute clusters with monitors on each one of them. Within the same account, users also have their test, development, and production environments. For local markets (instead of using FTP to transfer their data between systems), users simply open access for approximately 200 views to corporate, and their reporting is done. There is no latency, no auditing, and only one governance.

By alleviating the problems of complexity and rigidity while at the same time ensuring cloud data protection, Snowflake solves four of the most critical SAP workloads for businesses today:

- Replicating SAP BW data
- Replacing SAP BW
- Replacing SAP HANA data warehouse
- Keeping SAP BusinessObjects and SAP Analytics Cloud (SAC) but replacing the source

WORKLOAD #1: REPLICATING SAP BW DATA INTO SNOWFLAKE

You can quickly gain business value by first replicating SAP BW corporate data, bringing third-party or sensor data into Snowflake as well, leveraging Snowflake Marketplace, and then modelling all the data together.

The immediate benefit is the ability to have all your data available in one place for reporting and analysis: SAP, non-SAP, and machine-generated data (for example, SAP with Salesforce and clickstream data from Facebook and Twitter). Business users can then access all their data via their tool of choice (such as Tableau or Power BI) using their own isolated compute clusters, as shown in Figure 1.

This means your large machine learning query will not crash the whole system, nor will month-end processes slow down the systems. Snowflake's computational power elastically expands and contracts as needed. The business no longer has to do multiple extracts from SAP BW, nor does it have to stitch the data together.

All data can be queried from Snowflake. With the Snowflake Zero-Copy Cloning feature, you can take a zero-footprint copy of production for a machine learning training exercise, whether production is 100 gigabytes or 5 petabytes.

Snowflake allows you to instantly scale to literally tens of thousands of compute clusters and analyze petabytes of data without any impact on performance per user. No other analytics systems in the world can do this with the same ease, power, and governance that Snowflake offers. And if demand decreases, the system can automatically scale back down to one or zero compute clusters just as easily as it scaled up.

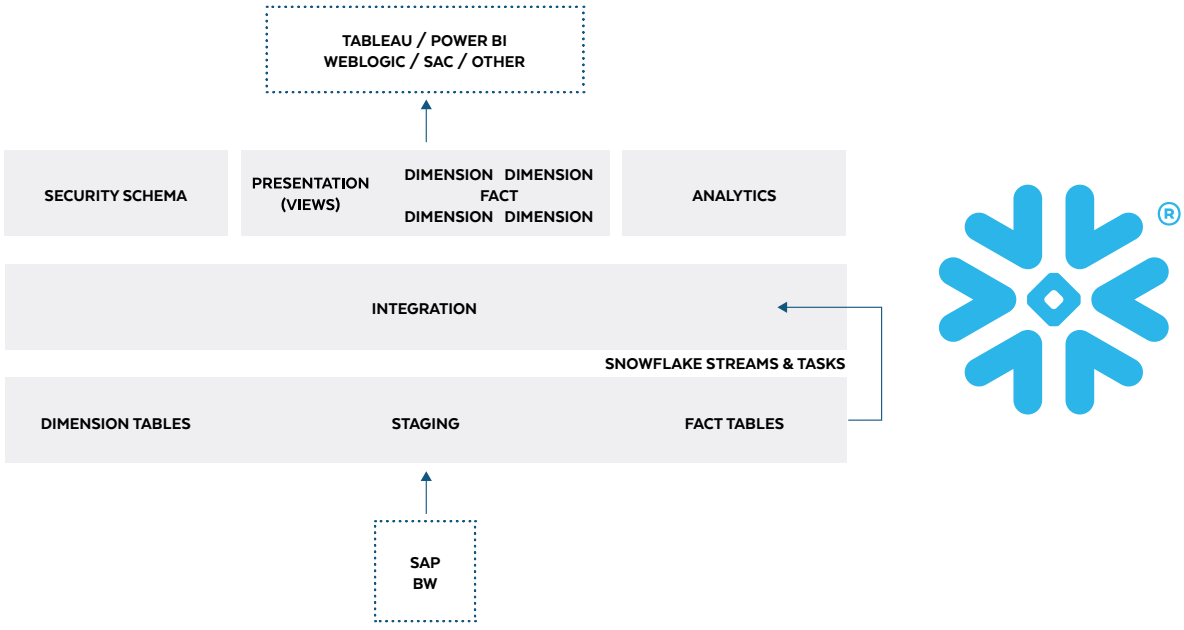


Figure 1: Replicating SAP BW data into Snowflake

Snowflake's many benefits with this particular workload include:

- **Replication:** Offloading data from SAP BW means the data modelling has already been done before the data goes into Snowflake. Then it is simply a matter of re-joining the facts and dimensions.
- **Bringing in other sources:** Customers can then easily combine the data (such as financials, sales, accounts receivable/accounts payable, or HR data) with other third-party data sets from Snowflake Marketplace—which offers a selection of over 100 different data providers specializing in finance, health, demographics, and weather. Snowflake ingests semi-structured data such as JSON, Avro, Parquet, and XML just as easily, and the data can be immediately queried. SAP just can't do this with such ease.
- **Optimization:** Customers can pull all the data they have in a data lake into Snowflake so it can be combined and instantly accessed without running through a data engineering application. At that point, they can reduce the size of the lake. If they still need traditional data lake functionality for some reason, they can easily copy data out of Snowflake on a case-by-case basis.
- **Modelling:** Snowflake supports all data models—from third-normal form, to dimensional star schema, snowflake schema, and Data Vault. There are modelling tools out there that will help, such as dbtvault, Qlik Compose, and WhereScape. Other more generic modelling tools include SqlDBM. SAP PowerDesigner also has Snowflake connectivity.



WORKLOAD #2: REPLACING SAP BW

SAP BW is a data warehousing solution with lots of built-in tools—such as tools for modelling, planning, consolidation, and process chains. Snowflake works differently, so most customers do not immediately replace SAP BW but rather leverage SAP BW in combination with Snowflake.

The first step in migration is to quickly provide value to the business by providing SAP data in an easy-to-consume fashion mixed with real-life data (see workload #1).

The second step is to start moving all those workflows off of SAP BW and onto Snowflake directly from the source systems. As SAP BW workflows transition onto Snowflake, superior performance will ramp up, thanks to Snowflake's massive parallel processing capabilities. Snowflake's robust built-in data engineering allows you to harmonize and transform your data, while external function calls allow you to pull logic from outside systems if necessary. Snowflake's Snowpark solution which can use Python or Scala can be used for highly complex transformations.

In more detail, integration tools offer three extraction patterns for classic SAP ECC, SAP S4/HANA, and SAP BW—ones that either replicate the base tables from SAP or go through the application layer. Users can choose standard SAP technology—such as SAP BW extractors, SAP LT Replication Server (SLT), Operational Data provisioning/Operational Delta queue (ODP/ODQ), SAP Data Services, or SAP BW OpenHub Service—to push and pull this data out, as shown in Figure 2. Alternatively, customers can use a wide variety of partner solutions that handle pool and cluster tables, replication, and differential loads, working through either log capture and SAP application-based triggers.

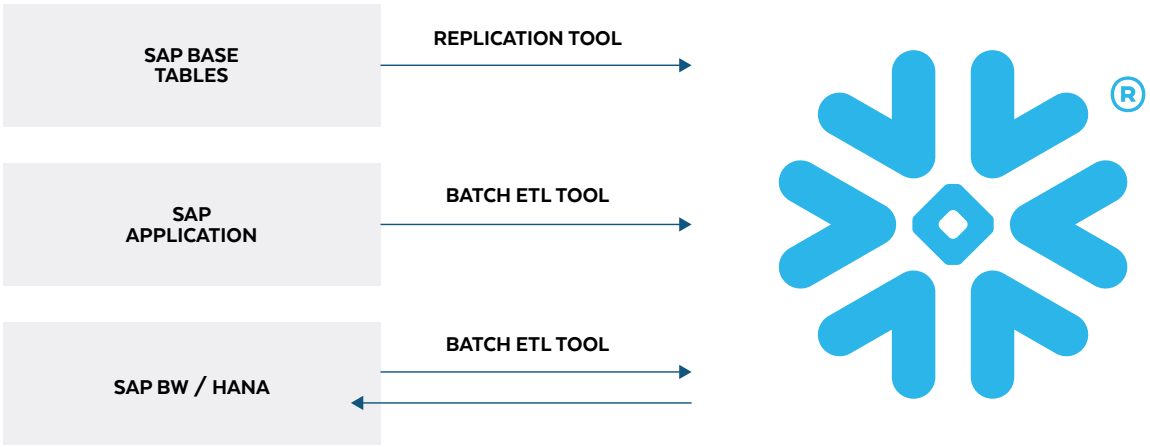


Figure 2: Replacing SAP BW

WORKLOAD #3: REPLACE SAP HANA DATA WAREHOUSE

Most customers are looking for faster, cheaper, and easier ways to get more value from their SAP data. They want to combine SAP and non-SAP data, run analytical workloads, run machine learning workloads, perform historical analysis, and mine for deeper insights. The biggest companies in the world took a gamble on SAP HANA as a data warehouse , but SAP HANA ran into scaling, concurrency, and cost issues. Customers who adopted the SAP HANA sidecar approach wanted both a real-time and SQL-based approach. Snowflake can do both at a much lower price point.

Regardless of the database underlying SAP applications, users cannot effectively run large, resource-intensive OLAP workloads in the same database running OLTP applications without having resource contention issues. It's great that SAP S/4HANA is taking back the operational reporting, making SAP BW and HANA EDW obsolete. This opens the door for a platform such as Snowflake, which offers unlimited computing power for analytic flows.

With this in mind, Snowflake provides advantages over SAP HANA with regard to price, elasticity, data sharing, isolated compute, variety, and volume. Plus Snowflake customers don't have to partition or manage indexes, and they don't have to upgrade software.

In addition, customers can run cloud-agnostic workloads and avoid vendor lock-in by cloud providers. Because Snowflake supports open standards, your data always belongs to you and it can easily be egressed at any time.

Replacing SAP HANA is straightforward at the table level because both SAP HANA and Snowflake are ANSI SQL databases and type conversion is minimal. Replacing SAP HANA Views does require rebuilding the logic in Snowflake. This is facilitated through the use of Snowflake tasks and streams, which automate data engineering harmonization and transformation. You can use common table expressions (CTEs) to capture recursive hierarchies from SAP tables and open them for analysis. Many Global Service Providers, such as Accenture, LTI, Deloitte, and Infosys, provide SAP HANA and CDS Views to Snowflake View conversion accelerators. Companies such at Atos, Price Waterhouse Cooper, and Qlik provide green-field business content templates.

WORKLOAD #4: LIFT AND SHIFT REPORTING TO SNOWFLAKE

Many companies hesitate to replace the underlying SAP HANA data source in their analytics stack because of the effort put into SAP BusinessObjects universes and the thousands of reports connected to them. Snowflake partners, for example, Wiiisdom, provide a clear path forward so users can reuse those universes and reports against a Snowflake database. Customers can also stay on SAP Analytics Cloud and leverage Snowflake's power in the back-end systems. For example, APOS provides the APOS Gateway that exposes Snowflake as a live data connection to SAP Analytics Cloud. Snowflake can also free users from the SAP stack—allowing them to effectively use their preferred analytics tool in a more performant manner. For example, some customers have moved off of SAP BW just so they can use Power BI or Tableau. For customers looking for a centralised semantic layer, AtScale can be a good fit. Snowflake opens companies to innovative solutions such as ThoughtSpot or DataRobot, which provide higher analytics to data citizens.

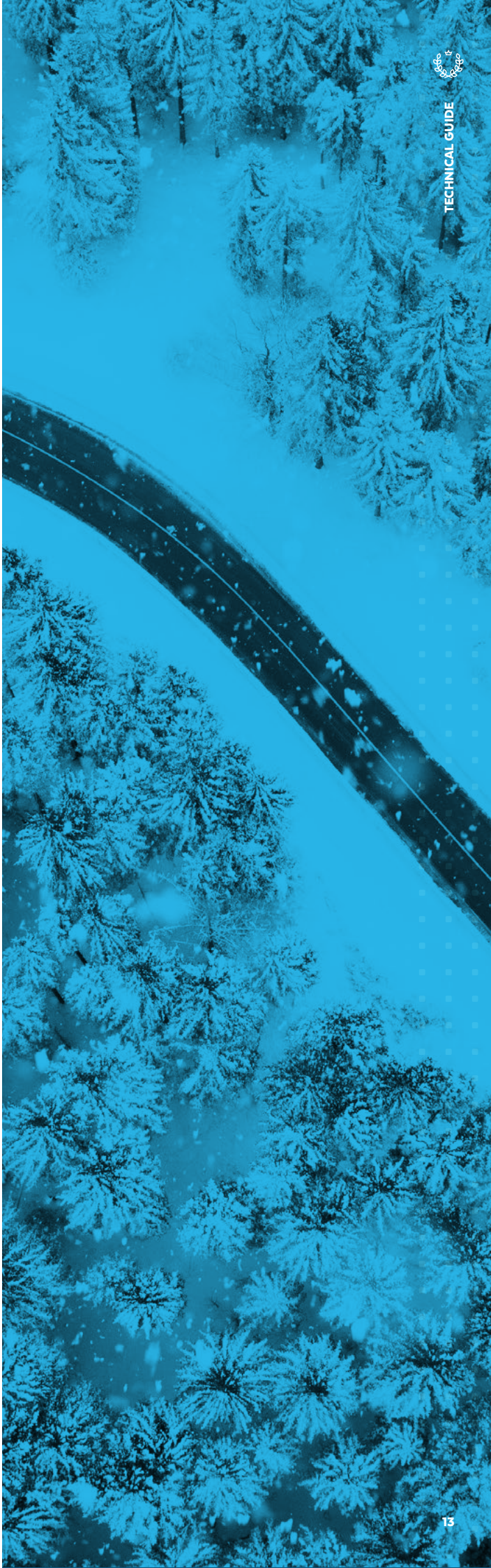


SNOWFLAKE ADDS VALUE TO SAP CLOUD MIGRATION

In addition to solving the aforementioned critical SAP workload cases, Snowflake's platform can provide massive operational efficiencies and growth to businesses—while reducing risk and accelerating time to market.

Cost is another consideration. Computer Weekly reports that migration to SAP BW/4HANA may also come with additional license and maintenance fees.[6] Snowflake provides "share what you want, when you want, with whom you want" cloud computing with instant elasticity, scalability, isolation, and a pay-as-you-go pricing structure. This simplifies an organization's ability to respond to business needs while helping reduce both CapEx and OpEx budgets.

Snowflake can be the foundation for your data—whether you are based locally, regionally, or globally. Whether you prefer AWS, Microsoft Azure, or GCP, Snowflake runs with and shares data across all three safely, securely, at scale. Snowflake even enables organizations to have disaster recovery across the cloud platforms—resilience that no cloud providers can offer on their own.



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 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 573 of the 2022 Forbes Global 2000 (G2K) as of January 31, 2023, use Snowflake Data Cloud to power their businesses.

Learn more at **snowflake.com**





© 2023 Snowflake. All rights reserved.

CITATIONS

[1] bit.ly/3g5118A

[2] tek.io/3uKERg5

[3] bit.ly/327Ei3B

[4] bit.ly/2OLTnF4

[5] bit.ly/32c62E4

[6] bit.ly/3mOuc1e