

The logo for Data Business, featuring the word "data" in a lowercase sans-serif font, followed by a stylized circular icon containing a white dot, and then the word "business" in a lowercase sans-serif font. The background of the entire page is a cityscape at sunset, with a blue sky and scattered clouds.

data  business

dbReplika

**High Volume Replication for SAP S/4HANA®,
SAP BW/4HANA®, Snowflake and Databricks**

Data Business GmbH

Mergenthalerallee 73-75

65760 Eschborn

Mobile: +49 6196 5860220

Mail: info@data-business.de

Website: www.data-business.de

dbReplika

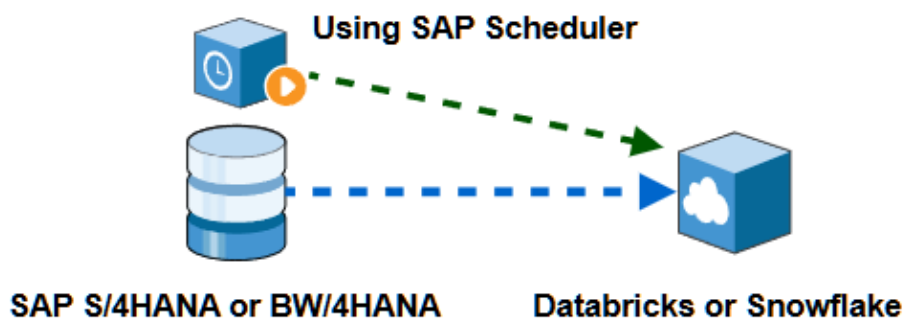
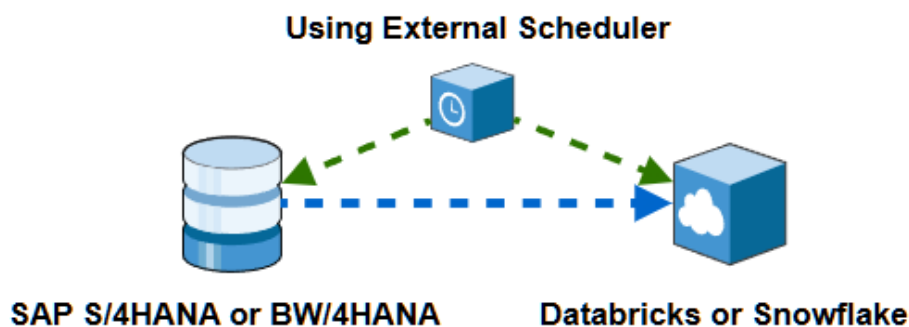
Unleash the Power of your SAP data

We specialize in building robust data platforms on SAP, cloud, and open-source ecosystems. Our innovative replication tool seamlessly integrates and transfers SAP data into modern cloud data platforms, empowering businesses to harness actionable insights with ease.

Leverage our expertise to transform your data landscape and drive smarter decision-making. Let us help you bridge the gap between SAP and the cloud.

"dbReplika enables you to replicate business rich SAP data sources and providers to your cloud platform of choice"

Architecture Options



Orchestration

Option 1 – Managed by External Scheduler

The triggering of a replication process is done through the execution of a docker image that accepts the login credentials to the SAP system, which triggers the data replication. This docker image can be run locally or in cloud and can be attached to most orchestration tools with ease. The docker image is part of the standard delivery package.

Option 2 – Managed by SAP BW Scheduler

If a customer would like to have an orchestration-free setup, we offer the possibility to directly transfer the data to targets like Snowflake and Databricks directly. Snowflake/Databricks will be configured to poll the new data.

Performance Metrics

Performance is influenced by the size of the records and the level of parallelization, which depends on the specific customer scenario.

For instance, with a **record size of 600 characters** and **5 parallel jobs**, it is possible to transfer approximately **100 million records in several minutes**.

Replication Features

Configuration

The dbReplika GUI allows users to activate a data source for replication in less than a minute. Each replication is represented by a artefact which when triggered transfers the data to the target cloud data warehouse.

Delta

dbReplika uses the standard BW Delta frameworks, thereby ensuring an SAP compliant delta by design. In cases a transfer fails, upon fixing of the root cause, the Delta queue is recovered and consistently transferred.

Filters

DTPs can be configured for partial replication of data by using the standard DTP filter capabilities.

SAP Licensing requirements

Although dbReplika does not technically require an OpenHub functionality, it is strongly recommended to purchase an OpenHub license, as this is the fundamental instrument that licenses SAP users for export of SAP data in batch mode. Customers may reach out to their SAP engagement managers to learn more about this.

Supported Systems

Source Systems	Source Types	Target Systems
SAP BWonHANA® >= 7.5	Datasources (ODP, SAPI, CDS)	Snowflake
SAP BW/4HANA® >= 1.0	Composite Providers, ADSO	Databricks
SAP S/4HANA® >=1709	Custom Tables via CDS Views	Other vendors

SAP Compliance

✔ **SAP Note 2814740** Database triggers in ABAP Dictionary
dbReplika doesn't implement any Database triggers to enable change data capturing, but rather uses what is offered in a standard and typical installation.

✔ **SAP Note 3255746** Unpermitted usage of ODP Data Replication APIs
dbReplika does not use unpermitted OPD APIs, but rather basis itself on Standard BW

✔ **SAP Note 2971304** SAP hasn't certified any redo log-based replication
dbReplika doesn't read any HANA redo logs for a delta capability or data replication

Data Security

dbReplika by design ensures that data never leaves the networks and systems of the customer. dbReplika runs as an SAP Add-on in the customer on-premise or SAP Private Cloud SAP system and doesn't need any cloud subscription.

Feature Matrix

	dbReplika	Other Vendors
1-Click setup	Yes	No
Low-Code / No-Code	Yes	No
Usage based pricing	Optional	Mandatory
Hidden follow up costs	No	Yes
Replication performance	Very Fast	Slow
Transfer method	Highly optimized	oData, RFC, JDBC, ODBC
S3 Support	Yes	No
CDS View Support	Yes	No
Custom Datasources	Yes	No
SAPI Datasources	Yes	No
ODP Datasources support	Yes	Violating SAP Note 3255746
HANA DB Log usage	Not needed	Violating SAP Note 2971304
Database trigger	Not needed	Violating SAP Note 2814740
Middleware needed	Not needed	Mandatory
SSH Connection needed	Not needed	Mandatory
SAP BTP Cloud	Not needed	Mandatory
SAP® Datasphere	Not needed	Mandatory
SAP Cloud Connector	Not needed	Mandatory
SAP JAVA Connector	Not needed	Mandatory
JDBC / ODBC Adapter	Not needed	Mandatory
External scheduler	Yes	No
BW scheduler support	Yes	Partially
Databricks ETL content	Yes	No or outdated
Databricks Notebook	Yes	No
Databricks Job support	Yes	No
Snowflake ETL content	Yes	No or outdated
Snowflake Notebook	Yes	No
Snowflake Stage	Yes	No
Snowflake Snowpipe	Yes	No

FAQ

How is our solution different from SAP OpenHub?

- No management view in S/4HANA, deltas cannot be repeated
- Code has many bugs and long-term strategy and support is unclear
- OpenHub has no end-to-end solution
- Large data transfers can't be split and parallelized
- Poor performance, export of millions of records can take hours
- Inflexible setup of CDS Views, problems with long fieldnames
- OpenHub can't be integrated to external orchestration tools

How does our solution support Snowflake and Databricks?

- OpenHub doesn't write to Snowflake and Databricks
- OpenHub has no out-of-box S3 Integration
- OpenHub has no Databricks Notebook support
- OpenHub has no Databricks Job support
- OpenHub has no Snowflake Notebook support
- OpenHub has no Snowflake Stage support
- OpenHub has no Snowflake Snowpipe support

Challenges in SAP Data Replication with Snowflake

Common Technical Challenges

Performance Bottlenecks

- SAP table locking during extraction
- Network bandwidth limitations during large data transfers
- Resource contention in production environments
- Slow processing of wide tables with numerous columns

Data Consistency Issues

- Handling complex SAP data types and conversions
- Maintaining referential integrity across tables
- Managing delta changes in clustered tables
- Synchronizing data across different time zones

Operational Complexities

- Complex SAP authorization requirements
- Limited extraction windows during business hours
- High memory consumption during full loads
- Monitoring and alerting across multiple systems

Integration Hurdles

- SAP module-specific extraction logic
- Custom ABAP code compatibility
- Pool and cluster table replication
- Handling of SAP buffer synchronization

Cost Management

- Snowflake compute costs during large loads
- Storage costs for historical data versions
- Network egress charges
- Development and testing environment expenses

Best Practices

- Implement incremental loading where possible
- Use parallel processing for large tables
- Schedule resource-intensive loads during off-peak hours
- Optimize table structures and indexes
- Regular monitoring and performance tuning

Challenges in SAP Data Replication with Databricks

Common Technical Challenges

Performance Issues

- SAP extractor performance limitations
- High latency during peak business hours
- Memory pressure during large table processing
- Slow processing of hierarchical data structures

Architecture Complexities

- Delta Lake table optimization challenges
- Cluster configuration for varying workloads
- Managing schema evolution
- Handling of concurrent write operations

Data Quality Concerns

- ABAP data type conversion challenges
- Maintaining data lineage
- Complex transformation logic validation
- Handling of SAP null values and special characters

Operational Challenges

- Job orchestration across environments
- Resource allocation for multiple workloads
- Managing compute costs for large datasets
- Monitoring distributed processing tasks

Integration Hurdles

- SAP connector stability issues
- Authentication and authorization complexity
- Network security configuration
- Managing CDC (Change Data Capture) failures

Best Practices

- Implement auto-scaling policies
- Use optimized file formats (Delta/Parquet)
- Set up proper partitioning strategies
- Deploy robust error handling mechanisms
- Establish clear SLAs for data freshness

Our Company



Data Business GmbH was founded in 2016 to serve global market with data engineering related services and products. Our focus is on solutions in the areas of data lakehouse architectures, data engineering, analytics and enterprise resource planning using SAP® products like SAP® S4/HANA, SAP® BW4/HANA and Open-Source technologies. As a project partner or interim service provider, we offer you customized and coordinated IT services and solutions with expertise and passion. Together we find the right solutions for your individual challenges. We deliver solutions for business rules and decision management with Low-Code & No-Code approach.

We believe that data is the key to unlocking hidden insights and transforming businesses. Our mission is to provide innovative software services and products that help businesses harness the power of data engineering, analytics, and AI. We are committed to delivering customized solutions that meet the unique needs of our clients. Data Business excels in four key areas of data-related services:

- **Data Engineering:** Data Business stands out as an industry leader in data engineering. Their commitment to developing state-of-the-art information systems is evident in their tailored solutions for clients. These solutions enable clients to make informed decisions and drive business growth, reflecting their dedication to delivering exceptional outcomes.
- **Data Analytics:** With over decades of experience, Data Business offers cutting-edge analytics services. Their team of skilled analysts and scientists collaborates with clients to craft data-driven strategies aligned with business objectives. Utilizing the latest tools and techniques, they extract actionable insights from data, empowering clients with real-time visibility into business performance through concise reports.
- **Data Science and AI:** Data Business is a frontrunner in data science, leveraging advanced technology and techniques to provide clients with actionable insights. Their scientific approach transforms vast datasets into meaningful information, supporting better decision-making. Whether it's data mining, machine learning, or big data analytics, their experts have the skills to help clients achieve their goals.
- **Data Platform Solutions (Cloud/OnPrem):** Data Business specializes in delivering cutting-edge data platforms, offering both on-premise and cloud options to cater to businesses of all sizes and industries. Their experienced team designs and implements data platforms tailored to clients' specific needs, ensuring data security, efficiency, and effectiveness.

Data Business Team



Johannes Iwanow

Chief Executive Officer

SAP S4/HANA® Integration and BI
25+ years cross industry experience
Touristic, Banking, Logistics, Consumer
M.Sc. in Business Administration

Mark Klein

Chief Operating Officer

SAP HANA® Integration and BI
25+ years cross industry experience
Banking, Logistic, Consumer
M.Sc. in Mathematics



Pralay Ahluwalia

Partner | Data Management & Analytics

Senior Business Architect
25+ years cross industry experience
Banking, Pharma, IT
M.Sc. in Finance & Risk / B.Sc. in Engineering



Roman Doubrava

Partner | Integration & Infrastructure

SAP HANA® Integration and BI
25+ years cross industry experience
IT Infrastructure, Banking, Energy



Microsoft, Excel® are trademarks of the Microsoft group of companies. SAP S/4HANA®, SAP BW/4HANA®, SAP® BWonHANA are the trademarks or registered trademark of SAP SE or its affiliates in Germany and in several other countries. Snowflake, Snowpipe are registered trademarks of Snowflake Computing, Inc. Databricks is a registered trademark of Databricks, Inc. Amazon S3, S3 are registered trademarks of Amazon Web Services, Inc.