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

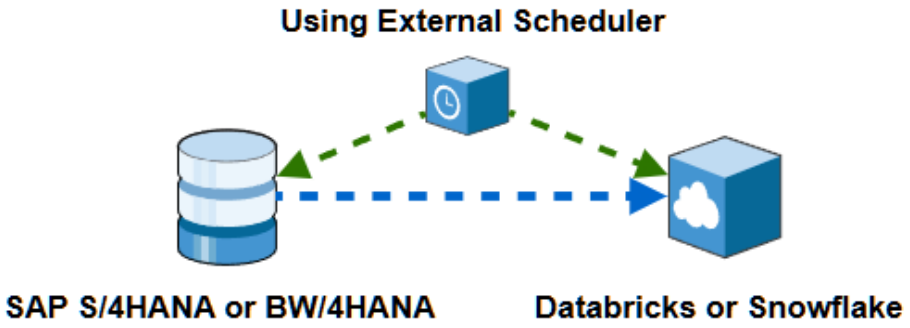
Unleash the Power of your SAP data

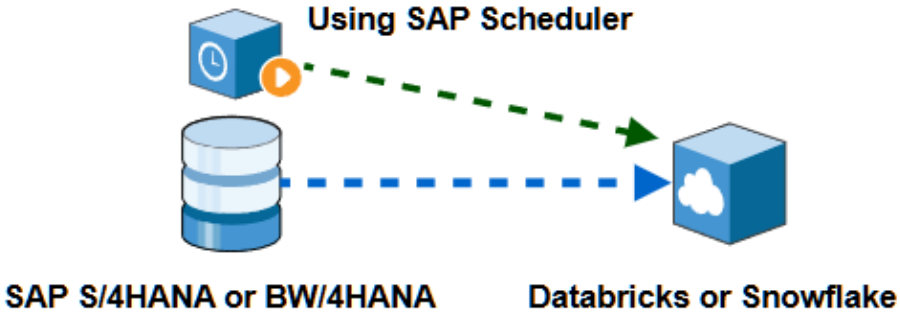
We specialize in building robust data platforms on SAP, cloud, and open-source ecosystems. Our innovative replication tool seamlessly integrates and transfers SAP data into modern cloud data platforms, empowering businesses to harness actionable insights with ease.

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 replication setup	**Yes**	**No**
Low-Code / No-Code	**Yes**	**Partially**
Cost efficient	**Yes**	**No**
Usage based pricing	**Low**	**High**
Hidden follow up costs	**No**	**Yes**
Replication performance	**Fast**	**Slow**
Transfer method	**Highly optimized**	**oData, RFC, JDBC, ODBC**
S3 Support	**Yes**	**Partially**
CDS View Support	**Yes**	**Partially**
Custom Datasources	**Yes**	**No**
SAPI Datasources	**Yes**	**No**
ODP Datasources support	**SAP compliant**	**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**	**Partially needed**
SSH Connection needed	**Not needed**	**Partially needed**
SAP BTP Cloud	**Not needed**	**Partially needed**
SAP® Datasphere	**Not needed**	**Partially needed**
SAP Cloud Connector	**Not needed**	**Partially needed**
SAP JAVA Connector	**Not needed**	**Partially needed**
JDBC / ODBC Adapter	**Not needed**	**Partially needed**
External scheduler	**Optional**	**Partially needed**
BW scheduler support	**Yes**	**No**
Databricks ETL content	**Inbound Stage**	**None or outdated**
Databricks Notebooks	**Supported**	**Partially**
Databricks Jobs	**Supported**	**Partially**
Snowflake ETL content	**Inbound Stage**	**None or outdated**
Snowflake Notebooks	**Supported**	**Partially**
Snowflake Stage	**Supported**	**Partially**
Snowflake Snowpipe	**Supported**	**Partially**

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 provide specialized consulting services in data engineering and related technologies. We focus on practical solutions involving 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 your consulting partner, we work closely with you to develop customized IT solutions that address your specific business requirements.

We understand that effective data management is crucial for informed business decisions. Our approach is to provide reliable consulting services that help organizations better utilize their data through sound engineering practices and appropriate analytical solutions. We're committed to delivering practical results that meet each client's unique needs and technical environment.
Data Business offers consulting expertise in four key service areas:

- Data Engineering: As a specialized SAP consulting firm, we focus on data engineering solutions that make a real difference. We work closely with our clients to develop tailored information systems that support better decision-making and drive meaningful business results. Our boutique approach ensures personalized attention and solutions designed specifically for each client's unique challenges.

- Data Analytics: We provide practical analytics services to help organizations better understand and utilize their data. Our consultants work closely with clients to identify key business questions and develop focused analytical approaches. We extract meaningful insights from your data and present them through clear, actionable reports that support informed decision-making and business planning.

- Data Science and AI: We help organizations apply data science techniques to better understand and utilize their business data. We work with clients to identify where methods like data mining, machine learning, and analytics can provide valuable insights. Our consultants focus on practical applications that align with your specific business needs.

- Data Platform Solutions (Cloud/OnPrem): We provide consulting services for data platform design and implementation. We help organizations develop practical solutions using on-premise, cloud, or hybrid approaches based on their specific requirements. Our consultants focus on delivering secure, reliable platforms that integrate effectively with existing systems and support your business objectives.

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
Co-Founder | AI & HPC Research
Hybrid Cloud & Data Warehousing
SAP® HANA & Infrastructure
Full Stack Architect
IT, Banking, Energy

Data Business GmbH
Mergenthalerallee 73-75
65760 Eschborn (Germany)

Tel: +49 6196 5860220
Web: www.data-business.de
Mail: info@data-business.de