

PayGlobal

Data-on-demand

Technical Overview

Prepared For
PayGlobal Customers

Produced By
PayGlobal Solutions & DevOps Team

14 July 2026

Version 3.0

PayGlobal Data-on-demand Technical Overview

Introductory Note

This document presents a technical overview of the PayGlobal Data-on-demand product. Please note that the features, capabilities, and functionalities described herein represent the current conceptual and developmental state of the solution. As the Early Access Program (EAP) progresses, the information is subject to change based on ongoing performance and security testing outcomes.

Product Summary

The Data-on-demand product is designed to automate, secure, and optimize data movement and management processes. It transforms business data from SQL Server databases into a centralized AWS cloud environment, unlocking user-friendly reporting and analytics capabilities.

Core Capabilities Overview

Data Platform Foundation

The system is built on the following foundational components:

Capability	Summary
Automated Data Collection	It automates the transfer of data from SQL Server databases to AWS cloud storage, enabling effortless collection. Users maintain control over table selection, ensuring relevance and efficiency.
Centralised Data Lake	The service organizes data in AWS S3, simplifying its management and accessibility for reporting and analytics.
Fast Data Access	Utilizing AWS's scalable resources, data transfers and processing are consistently fast and reliable.
Reliable Operations	Proactive monitoring includes auto-retries for failed tasks and notifications to DevOps via Slack/Zendesk for consistent operational robustness.
Usage Tracking	The service measures and tags usage for seamless cost management.
Effortless Updates	Modern DevOps tools enable smooth rollouts of updates across environments (development, testing, production).
Security Measures	Internal platform monitoring with strict access controls and audit logging will ensure data compliance and protection.

Development Features

Features currently active include:

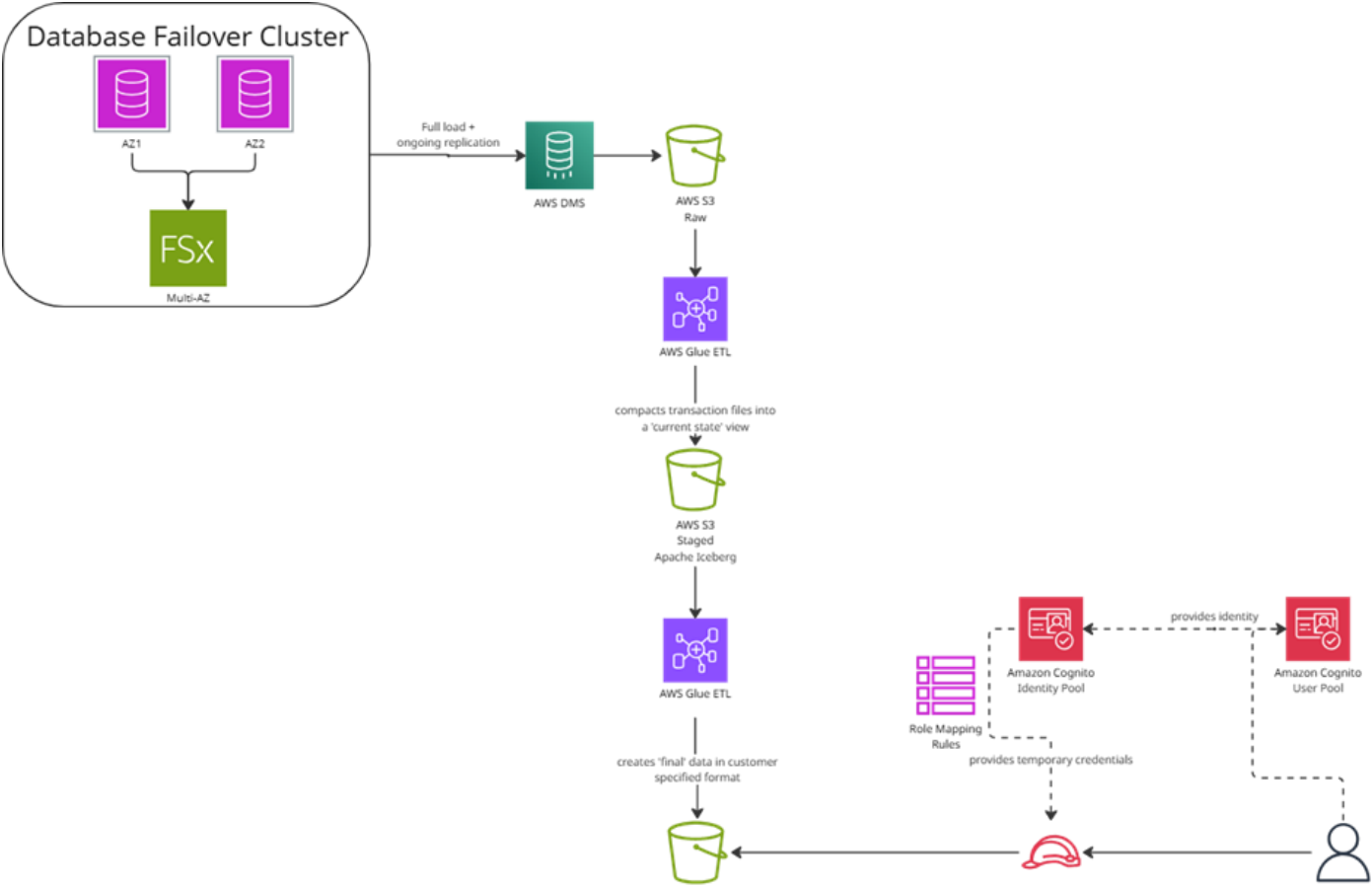
Capability	Summary
Full Snapshot	Snapshots of selected tables are exported as individual CSV files for comprehensive storage.
Incremental Changes (Delta Mode)	Customers can opt to receive updates, additions, and deletions (delta files in CSV formats) post-initial backup.
Improved Data Access Frequency	Current capabilities enable daily updates, with potential expansions allowing intervals as short as every 4 hours.
Secure Managed Access	A single secure customer login streamlines controlled access to the data lake.

Potential Future Enhancements

Features under exploration that may be added based on customer needs:

Capability	Summary
Direct Data Access	Limited access for AWS account holders to the Data Catalog for near real-time queries via AWS Athena.
Integration with Reporting Tools	Expanded usability of the Data Catalog with analytics and reporting platforms.
Expanded Data Platform Services	Further development towards integration with solutions like Snowflake for scalable historical data access.
Accelerated Data Availability	Continuous data transfer capabilities of AWS DMS may allow faster transformation schedules.

Deployment Diagram



Exported Data Flow

The process exports 2 flows of data as described below.

Snapshot Backup

- A snapshot of each table at the time of the execution
- Each execution overwrites the previous file
- There is a possible delay on the data due to the replication process

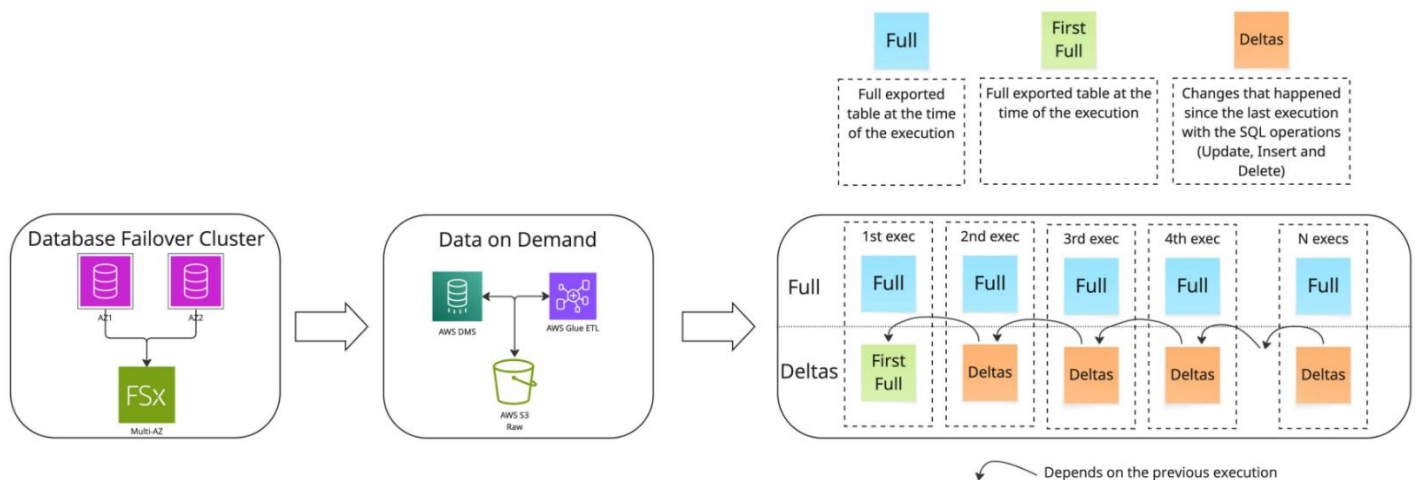
Deltas

- A continuous delta change that exports all changes on the table since the last execution of the previous deltas
- This file includes the SQL operation that happened on the table, if a line was INSERTED, UPDATED or DELETED, this will show a history of all changes on each table

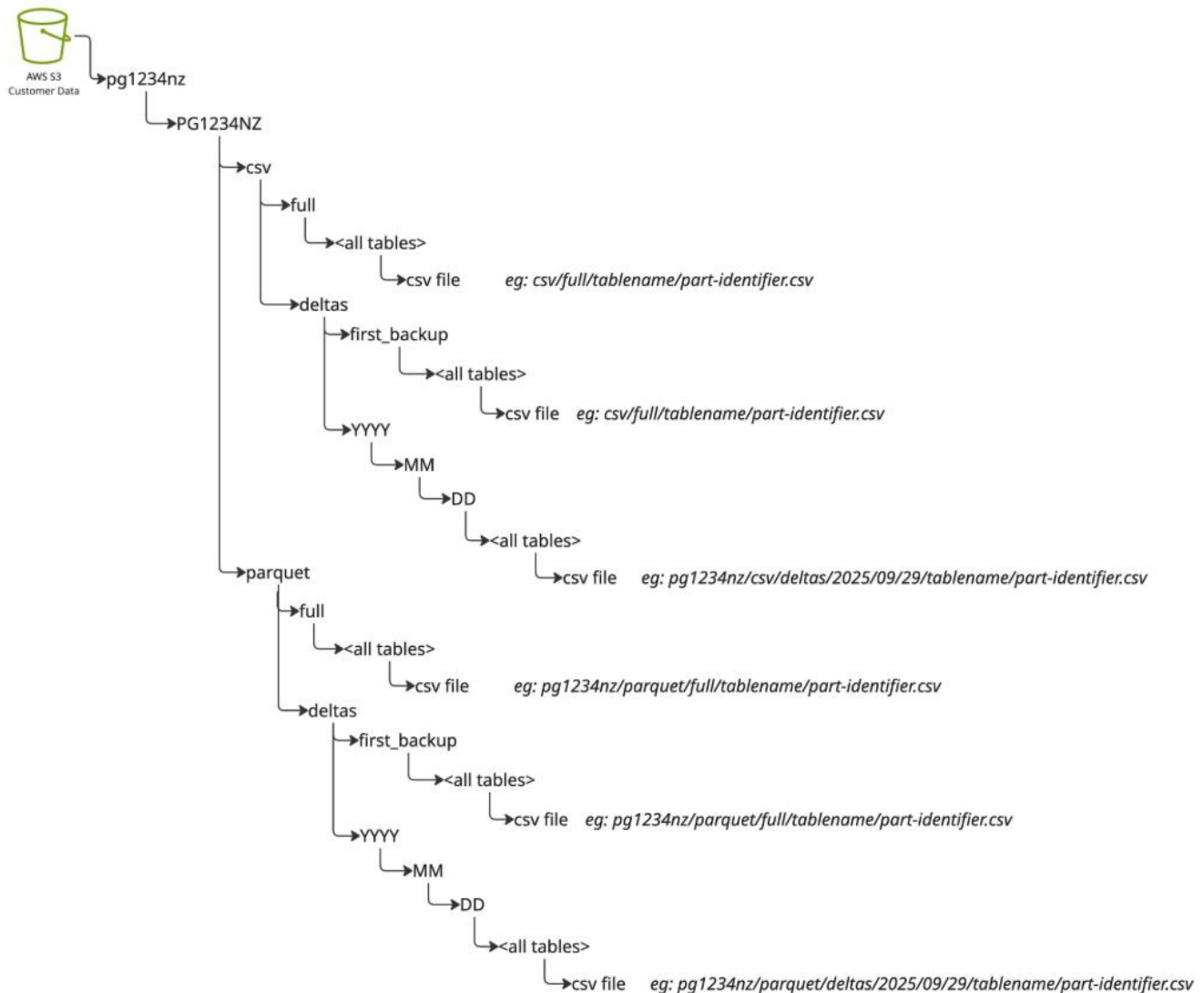
Important: The deltas are dependent on the previous execution, and all are dependent on the first full backup that is generated at the beginning of the replication process (customer onboarding to DoD)

**All files will be exported in both csv and parquet file types by default.*

The diagram below illustrates both processes:



File structure available to the customer:



Constraints and Limitations

There are a few important constraints and limitations to be aware of:

- All the delta files are dependent on the file from the previous execution, and all are kept in the storage, structured by date.
- In case something happens, and the customer loses track of one of the delta files, the table will need to be restarted from the first backup, and then all subsequent delta files applied.
- Creating a new replication (or consolidating existing ones) is not incremental — it is essentially **starting the data pipeline from zero**. Customers must be prepared to clear historical change data and re-ingest the new baseline.

Frequently Asked Questions (FAQs)

1. What is the primary purpose of PayGlobal Data-on-demand?

The solution automates secure data collection and storage, enabling effortless reporting and analytics.

2. How is my data protected?

Comprehensive security measures such as strict access controls, encryption, and audit logs ensure compliance and protection.

3. Can I control what data is included?

Yes, users retain selection rights for specific tables designated for extraction and storage.

4. What happens if a system operation fails?

The system automatically retries failed processes while notifying DevOps of any issues, ensuring reliability.

5. Are delta updates supported?

Yes, incremental changes are available after a full initial snapshot.

6. Does the system integrate with reporting tools?

Planned future integrations aim to support direct reporting through platforms connected to AWS.

7. What tools facilitate data transformation?

AWS Glue is the core tool used for seamless cleaning and transformation of data.

8. Can AWS account users access data directly?

Potential future enhancements may allow limited near real-time access to the Data Catalog.

9. How often is data deleted / archived?

The latest **full dataset** is kept as a snapshot backup, and each new execution overwrites the previous snapshot file (i.e., only the most recent snapshot is retained).

Delta files (incremental change logs) are generated continuously, depend on the prior run, and are kept in storage structured by date (i.e., they accumulate over time). There is no plan for removal/archival for the deltas at this stage, this will be reviewed as needed.