



DATA · 12 MIN READ

A Modern Data Ecosystem for Insurance: The Producer/Consumer Model on Snowflake

How to architect a simple, trusted, and scalable data ecosystem using Snowflake's platform — replacing legacy point-to-point integrations with a producer/consumer model that cuts latency from days to minutes.

By **StrataEdge AI** · Published November 1, 2025 · [strataedge.ai](#)

The current data landscape in most insurance organizations — centered on legacy SQL Server environments — is a critical bottleneck that directly impacts business operations and creates significant risk. The Bordereau (BDX) processing flow alone illustrates the scale of the problem:

- **Extreme Latency:** It takes a minimum of four days for BDX financial data to travel from initial receipt to posting in the D365 ledger, delaying cash flow recognition and business insight.
- **Manual Toil and High Error Rates:** The process requires a minimum of five manual touchpoints. This manual intervention contributes to approximately 30% of all BDX requiring rework.
- **Systemic Inefficiency:** Platforms are burdened by 33 distinct point-to-point integrations for BDX processing alone, creating a fragile and complex web.
- **Lack of Traceability:** This combination of latency, manual steps, and disparate integrations makes it nearly impossible to trace data provenance.

The Vision: Trusted, Real-Time Data Products

We construct a new data ecosystem based on a producer/consumer model. This architecture replaces the slow, opaque legacy system with a streamlined flow where each data product serves as a centralized, single source of truth for a specific business concept. The goal: reduce BDX processing time from four days to under one hour, eliminate manual touchpoints, and provide full, real-time visibility into the data pipeline.

Guiding Architectural Principles

Single Source of Truth via a Hub-and-Spoke Model: All data flows through the central Snowflake hub. Brittle point-to-point integrations are eliminated in favor of a model where producers publish to the hub and consumers subscribe from the hub.

Trust Through Governance and Snapshot Consistency: Every data product has a clear owner, a defined contract, and visible lineage managed in a central catalog. Consumption is based on immutable, time-stamped snapshots, guaranteeing that all related processes see the exact same version of the truth.

Automation and Event-Driven Flow: Manual touchpoints and overnight batch jobs are replaced with a fully automated, event-driven architecture that propagates data changes in minutes, not days.

Decentralized Ownership, Centralized Infrastructure: Business domains own their data logic and quality, while the technology team provides a robust, scalable, and secure central platform. This approach aligns with Data Mesh principles of domain ownership and data as a product.

Core Architecture

The architecture leverages the full power of Snowflake as the single source of truth for all analytical and reporting data. Its core features — separation of compute/storage, Time Travel, and unified processing with SQL/Snowpark — are foundational to this design.

Data Zones: A Staged Path to Quality

Landing Zone (Raw): Captures raw, unmodified data exactly as it arrives from source systems (raw BDX files, API JSON payloads), providing a permanent, auditable record for forensics and reprocessing.

Raw Vault Zone: Contains the integrated, auditable history of all source data, modeled using Data Vault 2.0 constructs. Components include Hubs (unique business keys), Links (relationships between Hubs), and Satellites (descriptive, contextual data with full non-destructive history).

Business Vault Zone: Sits on top of the Raw Vault and contains derived data and applied business rules. Allows domains to iterate on business logic without impacting downstream consumers.

Refined Zone (Data Products): The critical final layer where we expose data to the business. Data from the Raw and Business Vaults is transformed and denormalized into user-friendly dimensional models or flat, wide tables. These performant, easy-to-query views are our "Data Products."

The Producer/Consumer Model with Data Vault

The core of the central data platform is modeled using Data Vault 2.0 methodology. This approach integrates data from multiple source systems in a non-destructive way, tracking history and lineage by default.

Producers of the Raw Vault are typically technical data engineering pipelines. Their "product" is a well-formed, historically accurate representation of a source system loaded into Hubs, Links, and Satellites.

Producers of Data Products are the business-facing domain teams (Finance, Underwriting, etc.). They act as both consumers and producers — consuming data from the Vault and applying domain expertise to transform it into valuable, easy-to-use Data Products.

Domain Separation and Data Mesh Alignment

Each business domain (Finance, Underwriting, Claims, Operations) owns their data products end-to-end. They define transformation logic, establish quality standards, and maintain accountability. Data is treated as a product with explicit contracts, SLAs, and support models.

The self-serve data platform provides standardized capabilities for ingestion, transformation (dbt), quality testing, cataloging, and access control — eliminating redundant infrastructure while ensuring consistent governance.

Flexible Ingestion Framework

We classify source systems into three ingestion tiers:

Tier 1 — Direct Integrations: Leverages native Snowflake and AWS features. Includes custom services for complex business logic (like BDX Submission), S3 Event-driven Snowpipe Auto-Ingest for file-based flows, and Snowflake Data Sharing for partner data.

Tier 2 — Managed Connectors: For standard SaaS apps (Salesforce, D365, Rippling) using direct Snowflake connections or Snowflake's OpenFlow, providing direct integration within Snowflake's processing engine, reduced latency, and unified governance.

Tier 3 — Custom Functions: Reserved for sources requiring complex transformations or where custom business logic adds significant value.

Business Impact

This architecture transforms data from a cost center into a strategic asset — reducing processing time by 96%, eliminating manual touchpoints, and providing the governed, real-time data foundation that AI initiatives require. The producer/consumer model ensures every team has access to trusted, certified data products without needing to understand the complexity of the underlying platform.