

Sankar Sundaram

Senior Data Engineer & Cloud Data Architect · Data Warehousing Specialist · Azure · GCP · AWS

10+ years as a data engineer and cloud data architect building large-scale ELT/ETL and lakehouse platforms on Databricks — API ingestion, Delta Sharing, and orchestration across Airflow and dbt — with depth in enterprise data warehousing: dimensional modeling, T-SQL, stored procedures, and high-availability SQL Server. Proven record migrating legacy on-prem data warehouses to modern Delta Lake, with a focus on governance, data quality, and CI/CD-driven delivery.

[603-930-7755](tel:603-930-7755)

shankat2020@gmail.com

Jacksonville, FL

[LinkedIn ↗](#)

[GitHub ↗](#)

10+

Years in data engineering & cloud architecture

1.8B+

Rows ingested in a single retail sales pipeline

99.999%

SLA reliability across 30+ production pipelines

180+

ERP source tables migrated to Delta Lake

Core expertise

CLOUD & LAKEHOUSE PLATFORMS

Azure Databricks & Delta Lake

Google BigQuery & GCP

Azure Data Factory, Synapse & Microsoft Fabric

Azure SQL DB & SQL Server

DATA MODELING & SQL ENGINEERING

T-SQL, Stored Procs, Functions & Triggers

Performance Tuning & Query Optimization

Data Warehousing — Dimensional, Medallion &

ORCHESTRATION & TRANSFORMATION

Apache Airflow (Python DAGs)

dbt (Data Build Tool)

Spark SQL & PySpark

Kafka & Spark Structured Streaming

Databricks Asset Bundles

DELIVERY & PLATFORM OPS

CI/CD — Azure DevOps, GitHub, Bitbucket

Terraform, Docker & Kubernetes

Power BI, Tableau, SSRS & SSIS

Experience

○ Sr. Data Engineer / Architect (Azure) · Beaver Street Fisheries Apr 2025 – Present Jacksonville, FL

- Own end-to-end delivery of ERP migration, pricing analytics, retail sales integration, and product information initiatives on an **Azure Databricks Lakehouse (medallion / Delta Lake)** architecture.
- Led migration of **180+ ERP source tables to Delta Lake**, building Python DAGs orchestrated in Apache Airflow, replacing legacy on-prem batch processing with zero schema-change handling errors.
- Designed 30+ production Airflow pipelines covering 180+ tables, raising pipeline reliability to **99.999% SLA**.
- Engineered ingestion and modeling of **Walmart retail sales data at 1.8B+ row scale**, tuning partitioning, Z-ordering, and liquid clustering, and leveraging the Photon accelerator to speed up Spark query execution for 10× faster analytical queries.
- Authored and tuned T-SQL stored procedures, functions, and triggers in Azure SQL DB for ERP validation logic, applying performance tuning and query optimization.
- Configured Always On Availability Groups, replication, and clustering (WSFC) for Azure SQL DB high availability and disaster recovery.
- Implemented Databricks Asset Bundles (DABs) for CI/CD-deployed, versioned pipeline releases; enabled cross-org data sharing via Delta Sharing.
- Extended the Databricks lakehouse into Microsoft Fabric / OneLake to support enterprise reporting workflows for downstream BI consumers.
- Mentored engineering management and the broader team on the lakehouse architecture end-to-end — from design decisions through production implementation — building shared technical ownership across the organization.

○ Cloud Data Architect / Engineer (GCP) · Venus Jun 2023 – Apr 2025 Jacksonville, FL

- Migrated an on-prem **SQL Server enterprise data warehouse to Google BigQuery**, reducing report-refresh latency from 20 minutes to 2 minutes.

- Migrated legacy T-SQL stored procedures, functions, and triggers into BigQuery SQL and Databricks, applying database design and performance-tuning best practices.
- Developed dbt transformation pipelines structuring raw data into curated gold-layer models for self-service BI.
- Built Cloud Run services ingesting API data into BigQuery and Snowflake; established DevOps best practices in Azure DevOps.

○ **Data Engineer (Azure)** · Microsoft Inc.

Nov 2021 – Jun 2023

Redmond, WA

- Designed large-scale big-data pipelines for sustainability analytics on Azure Databricks and ADF/Synapse, improving reliability to **99.999% SLA**.
- Built ADF/Databricks pipelines across ADLS Gen2, Azure SQL, Event Hub, and Logic Apps; implemented governance and lineage with Microsoft Purview.
- Delivered interactive Power BI and Tableau dashboards using PySpark, Databricks notebooks, and Spark SQL.

○ **Technical Lead / Full Stack Data Engineer** · Pfizer Inc.

Dec 2020 – Nov 2021

Groton, CT

- Delivered big-data migration and API integrations on GKE, Cloud Dataflow, Cloud Run, and Cloud SQL; built an analytics platform on Snowflake and SnowSQL.
- Ingested data via AWS Glue, S3, and Lambda; ran CI/CD with Kubernetes, Docker, Jenkins, and Bitbucket.
- Built API integrations with Veeva Vault CRM and Veeva Link to ingest field-sales and medical-affairs data into the Snowflake analytics platform for commercial reporting.
- Supported regulatory and quality data flows by integrating pipelines with Veeva RIM and Veeva QMS, ensuring traceable, audit-ready data for compliance reporting.
- Led Veeva LIMS implementation end-to-end — requirements, workflow design, configuration, integration, and testing — partnering with Lab, Quality, IT, and Business Process Owners through UAT and go-live.
- Applied risk-based CSV/CSA and GxP principles across Veeva LIMS configuration and documentation, aligning with **21 CFR Part 11, EU Annex 11, GAMP 5, and ALCOA+**.

○ **Solutions Architect (Azure)** · Cirrus Aircraft / Daikin

Aug 2019 – Nov 2020

Minneapolis, MN

- Built ETL/ELT and ML pipelines using ADF, Azure Databricks, Spark, and Python for R&D and IoT/sensor streaming data.
- Designed star/snowflake schema models in Snowflake; automated IaC deployment with Terraform and Azure DevOps.

Education & certifications

M.S., Cyber Security

EC-Council University, Albuquerque, NM ·
GPA 3.80 / 4.0

Certifications

- ♦ Google Certified Professional Cloud Architect
- ♦ Microsoft Certified: Azure AI Engineer Associate
- ♦ Databricks Lakehouse Fundamentals
- ♦ Certified Gen AI for Data Engineers
- ♦ Google Cloud Digital Leader

Earlier experience

Data Analytics Consultant

2017 – 2019

Wells Fargo & US Bank, MN

Migrated SAS reporting to SQL/Azure Data Lake; built ETL (SSIS, ADF) and Power BI/Tableau analytics.

Technical Lead – IAM

2016 – 2017

Biogen, MA

Led enterprise systems integration; built Azure Data Lake, Informatica, and Oracle PL/SQL data ingestion.

Delivery Lead

2009 – 2016

Pfizer Inc., CT

Delivered BI/DW solutions (Informatica, OWB, Tableau, SAS) and MDM for life-sciences applications, integrating feeds from Veeva Vault (QualityDocs, eTMF).

Let's build what's next.

Open to roles spanning cloud data architecture & deployments, data engineering, data migration, lakehouse, data pipelines, data quality, and data modeling — plus API integrations and ERP-to-cloud migrations, BI & Data Warehousing projects. I'm equally driven by AI & ML: AI-driven development for new ideas, practical AI agent use cases, and responsible AI adoption with the right guardrails.

Sankar Sundaram — Jacksonville, FL

Visit analytics (approximate IP location) logged for traffic insights.

2026