

Vishal Kumar Sarvade

Staff Data Engineer | Data Architecture | AI Data Platforms

Santa Clara, CA | Phone: (727) 902-8486 | Email: sarvade1.vishalkumar@gmail.com

LinkedIn: linkedin.com/in/vishalkumarsarvade | GitHub: github.com/sarvadevishal

PROFESSIONAL SUMMARY

Data engineer and technical architect with 12+ years building cloud data warehouses, ETL/ELT data pipelines, and analytics for high-volume e-commerce and healthcare businesses.

Owns a 1 TB Amazon Redshift data platform end to end, from data architecture and dimensional modeling to change data capture (CDC) from Aurora MySQL and the 150+ Informatica Cloud and ETL Works jobs that feed it.

Strong SQL and Python. Cut a 12-hour workload to 30 minutes and saved about \$5,000 a month in platform costs.

Leads a 10-member offshore team, works directly with business stakeholders, and builds AI agents and an NL-to-SQL tool on an LLM that make the data easier to use.

TECHNICAL SKILLS

Languages & Engineering: SQL, Python, PySpark (POC), pandas, REST APIs, Git, Version Control, Code Reviews, Testing and Debugging, Root-Cause Analysis, Software Development Life Cycle (SDLC), Agile and Scrum, Jira, Documentation

Big Data & Real-Time Data: Amazon Redshift (MPP, RA3, Serverless), Redshift Zero-ETL, Redshift Data Sharing, Workload Management (WLM), Distribution and Sort Key Design, Apache Spark (POC), Batch and Near Real-Time Processing

Pipelines & Orchestration: ETL/ELT Pipelines, Change Data Capture (CDC), Workflow Orchestration, Data Ingestion, Data Integration, Informatica Cloud, ETL Works, AWS DMS, Apache Airflow (POC), Fivetran (POC), SSIS, Data Migration

Architecture, Modeling & Governance: Data Architecture, Data Modeling, Data Warehousing, Data Lake (S3), Snowflake (POC), Relational Databases (PostgreSQL, MySQL, SQL Server), Data Governance, Data Quality, PII, Data Privacy

Cloud & Reliability: AWS (S3, EC2, RDS, Lambda, IAM, Secrets Manager, API Gateway), Azure, Microsoft Fabric, Monitoring, Observability, Production Support, Troubleshooting, Performance Tuning, Cloud Cost Optimization, Deployments

AI & BI: AI Agents (n8n), LLMs and NL-to-SQL, Microsoft Copilot, Fabric Data Agents, Claude Code, Codex, Looker, Power BI, Semantic Models, KPI Dashboards, Self-Service Analytics, Executive and Operational Reporting, Data Visualization

PROFESSIONAL EXPERIENCE

Technical Architect | Grid Dynamics Inc, California, USA

April 2023 to Present

Project: Forever Living International Products

A global multi-level marketing and e-commerce platform across about 63 operating companies, built on Java, Spring Boot, Angular, MySQL, and Kafka microservices. Its 1 TB Redshift data mart of 200+ tables feeds every dashboard and report.

- **Query performance and cost:** rebuilt a critical aggregation workload on Redshift and cut its runtime from 12 hours to 30 minutes, a 24x gain. Consolidating an ETL Works instance and S3 storage saved about \$5,000 a month.
- **Change data capture:** replaced AWS DMS with a Zero-ETL pipeline that replicates Aurora MySQL into Redshift, which ended maintenance pauses, full reloads after failures, and binlog workarounds. The same pipeline now carries the move off AS400: Members run from Aurora MySQL for five operating companies and Orders for two.
- **Warehouse modernization:** moved a four-cluster Redshift estate (one producer, three consumers) from DC2 to RA3 with data sharing, retiring the sync jobs and giving consumers real-time access. Later traced a Serverless slowdown, with writes going from 10 seconds to 10 to 15 minutes, to a capacity setting, and recommended RA3 for writes with a Serverless workgroup for reporting.
- **Platform ownership:** run the data mart end to end, including the 100+ Informatica Cloud task flows and 50+ ETL Works jobs that handle data ingestion and transformation. Planning the job-by-job move of 110 Informatica jobs to ETL Works.
- **Architecture and data modeling:** own the star-schema data models, database design, and solution design across AWS and Azure. Presented the ingestion architecture to the client CTO and coordinate releases with the Bonus Engine, FLP360, Production Support, and AS400 teams.
- **Standards and governance:** set the team's standards for naming, keys, audit columns, documentation, and access grants, with every database change traced to its Jira ticket and Git commit. Built a data catalog portal from the business data dictionary, a Python tool that pulls Looker dashboard-to-column lineage, and a job metadata table that tracks every pipeline.
- **PII migration:** moved sensitive records (government ID, IBAN, retail, and customer service data) from CSV files and MySQL into Aurora MySQL, validating each batch for data accuracy and returning rejected rows to the customer as CSV.
- **AI and automation:** built an LLM-based NL-to-SQL tool with PII blocking and read-only guardrails, two n8n AI agents (FAQ and data analyst), an Informatica observability dashboard in Claude Code, and an Estimation app in Codex.
- **BI and Copilot:** delivered 40+ Looker dashboards and maintain 100+. Proved out Power BI as a Looker alternative with six dashboards, three of them built with Copilot, and stood up Microsoft Fabric. Now setting up Fabric data agents so Microsoft 365 Copilot can answer from Power BI semantic models.
- **Proofs of concept:** Snowflake, Apache Airflow, PySpark, and Fivetran. Wrote the BRD and TRD for a PySpark aggregation pipeline on Redshift.

- **Team leadership:** lead and mentor a 10-member offshore team through sprint planning, Scrum calls, and architecture reviews, and turn business requirements into delivered Jira tickets.

Environment: Amazon Redshift, Aurora MySQL, PostgreSQL, Amazon S3, Zero-ETL, AWS DMS, Informatica Cloud, ETL Works, SQL, Python, Looker, Power BI, Microsoft Fabric, n8n, Claude Code, Codex, Azure, Git, Jira

Associate Technical Architect | NextSphere Technologies Inc, Tampa, FL, USA

January 2022 to March 2023

Project: Forever Living International Products

- **Solution design:** gathered and refined business and technical requirements and turned them into practical solution designs and implementation plans.
- **Cross-functional delivery:** worked with application teams on design and release support so data integration stayed in step with platform changes.
- **Technical guidance and production support:** mentored team members, reviewed their work, resolved blockers, and coordinated issue resolution and maintenance to keep the platform stable.
- **Stakeholder communication:** worked directly with client teams to clarify requirements, discuss solution options, and explain technical decisions in plain terms.

Environment: AWS EC2, Amazon Redshift, Informatica Cloud, ETL Works, DB2, RDS, API Gateway, SQL, Git

Senior Software Engineer | NextSphere Technologies Inc, Clearwater, FL, USA

January 2020 to January 2022

Project: Forever Living International Products

- **Data pipeline development:** led a small team that designed and built end-to-end ETL pipelines moving data from PostgreSQL and AS400 into Amazon Redshift, using Informatica Cloud mappings, tasks, and task flows and ETL Works sync jobs.
- **Redshift performance tuning:** improved Amazon Redshift performance through distribution strategy, sort key design, and regular maintenance.
- **Data integration platform configuration:** tuned the Informatica Cloud Secure Agent for maximum task-flow capacity and set up connections to AS400, Amazon Redshift, and flat files for data migration.
- **Platform administration and deployments:** ran job setups, the data dictionary, connections, configuration changes, backups, and deployments.
- **Data quality and reporting automation:** automated the daily data-statistics report and audit checks in ETL Works.

Environment: AWS EC2, Redshift, PostgreSQL, AS400, Informatica Cloud, ETL Works, RDS, API Gateway, SQL, Git

Associate Technical Lead | NextSphere Technologies (India) Pvt Ltd, Hyderabad, India

March 2014 to January 2020

Project: Forever Living International Products

April 2018 to January 2020

- **Technical leadership and production support:** guided the offshore Production Support group, reviewed code to keep quality high, and led technical discussions.
- **Stakeholder communication:** reported progress to the Project Manager and worked through technical specifications with the customer.
- **Data pipeline development:** built ETL pipelines that load Amazon Redshift and created Informatica jobs with mapping designers, mapping tasks, and task flows.
- **Data quality audits:** developed ETL Works sync and data-quality audit jobs and maintained AS400 systems.
- **Troubleshooting and root-cause analysis:** earned repeated client appreciation for clarifying tricky business scenarios during requirements and pinpointing exact root causes during UAT.

Environment: AWS EC2, Informatica Cloud, RDS, API Gateway, SQL, Git

Project: Healthcare Data Migration (Client: Health Axis)

July 2014 to March 2018

- **Healthcare data migration:** led the move of members, providers, claims, authorizations, and eligibility records from legacy systems into the new platform using SQL Server stored procedures and SSIS, validating every business rule along the way.
- **Large-dataset processing:** built routines to process large CSV extracts (including 4 GB member loads), ran legacy-to-new gap analysis, and reviewed test scenarios with the client.

Environment: Visual Studio, SQL Server, SSIS, TFS, IIS, Bugzilla

EDUCATION

Bachelor of Technology (B.Tech), Electrical, Electronics and Communications Engineering

2009 to 2013

Vignana Bharathi Institute of Technology, India