Rajeev Jasti

PRINCIPAL DATA ENGINEER

(+1) 980-216-8513 · inboxtorj@gmail.com · www.rjasti.com · Charlotte, NC

SUMMARY

Hands-on Data Engineer with 8+ years of experience architecting distributed high-throughput microservices, enterprise OLAP/OLTP databases and middleware systems. Deep domain expertise in supply chain intelligence (ERP, WMS, TMS), with a proven track record of transforming multi-source operational data into a unified, scalable data warehouse. Deep technical expertise in database administration and optimizing execution plans, storage, logs, backup strategies, and transaction workloads across SQL Server, PostgreSQL and Redshift.

TECHNICAL SKILLS

Cloud & Serverless	Amazon Redshift, Amazon Bedrock, Amazon S3, AWS Glue, AWS Lambda, Amazon Athena, DynamoDB, API Gateway, CloudWatch, Amazon EKS, Amazon ECS, EC2, RDS, IAM, SNS, SQS, Glacier
Backend	FastAPI, boto3, SQLAlchemy, Pydantic, Multiprocessing, Multithreading
Data Engineering & BI	Matillion, PySpark, Looker, Pandas, NumPy
Databases & Middleware	PostgreSQL, SQL Server, Amazon Redshift, MySQL, Dell Boomi, Redis, Apache Kafka, RabbitMQ, Elasticsearch, DynamoDB, Alembic
Languages & Scripting	Python, T-SQL, PL/pgSQL, SQL, Bash / Linux
DevOps, Testing & Tooling	Git, Docker, Terraform, Kubernetes, GitHub Actions CI/CD, Pytest

EXPERIENCE

Nicholas and Company

Salt Lake City, Utah

Principal Data Engineer

Nov 2022 – Present

OLTP Architecture & Enterprise Pricing

- Own OLTP database development and administration for the systems running order-to-cash — pricing, rebates, commissions, and the integration layer connecting ERP, WMS, and TMS.
- Architected and delivered the enterprise pricing engine in SQL Server, resolving sales-rep overrides, customer contract pricing, commodity pricing, COGS, and discount layers into a real-time customer–item–DC price through ACID-compliant T-SQL procedures and functions.
- Scaled the pricing API to sustain 100,000–200,000 executions per minute at sub-100 ms latency with zero deadlocks under heavy concurrent load, serving internal e-commerce, external order entry, and ERP.
- Modeled pricing and customer master data as SCD Type 2, enabling point-in-time price reconstruction for any customer–item–DC combination across past, current, and future effective dates for audit and dispute resolution.
- Built PriceIQ, a statistical pricing subsystem that derives sales margins from three years of transaction history using IQR outlier fencing and skew-adjusted distribution analysis, intelligently pricing customer–item combinations that fall outside any custom pricing setup.
- Built the vendor rebate calculation and reporting platform, computing earned rebates across multiple program structures from purchasing and sales transaction history and giving Finance auditable accrual figures.
- Delivered commission-based payroll for the sales organization from transaction-level sales and margin data, and activity-based incentive payroll for warehouse and transportation teams sourced from WMS and TMS activity.
- Developed inbound and outbound feed processing for customers and suppliers covering sales orders, customer rebates, purchase orders, and receiving confirmations, with transactional ingestion and error handling.

Backend & Middleware Microservices

- Architected and deployed enterprise-grade serverless REST APIs using AWS API Gateway, Lambda, and Step Functions, serving high-throughput, sub-second reads and writes for core commerce entities across customers, items, orders, payments, balances, and invoices.

- Constructed an asynchronous, serverless order ingestion middleware using Lambda, SQS FIFO queues, and Step Functions (Saga pattern) to reliably process and fan out B2B/B2C orders to legacy ERP, WMS, and Inventory Distribution Systems.
- Engineered a distributed DynamoDB caching layer using SHA-256 payload fingerprinting and native TTL eviction, bypassing downstream relational database calls on identical requests and reducing primary database load by 30–50%.
- Implemented idempotent request handling and API Gateway throttling controls, eliminating duplicate order submissions across both internal e-commerce storefronts and high-volume third-party order entry channels.
- Designed fault-tolerant integration pipelines with automated Dead-Letter Queue redrive policies, exponential backoff, and circuit breakers, sustaining 99.9% transaction completion and insulating upstream checkout from downstream ERP batch-processing windows.
- Streamlined end-to-end order lifecycle workflows, transforming raw multi-channel order payloads into canonical distribution formats for automated picking, packing, and billing dispatch.
- Developed a high-concurrency inventory allocation and reservation microservice managing real-time Available-to-Promise stock reservations, preventing stockouts, phantom inventory, and overselling during peak transaction surges.
- Enforced strict transactional integrity and concurrency control using atomic conditional writes and distributed locks, ensuring multi-item, multi-warehouse order lines were reserved atomically prior to fulfillment handoff.
- Optimized allocation strategies for warehouse routing, factoring in stock proximity, fulfillment node priorities, and partial shipment rules to reduce fulfillment cycle times and split-shipment overhead.
- Implemented a dual-engine search strategy leveraging Elasticsearch for low-latency full-text catalog and order queries and SQL Server for complex relational filtering, with near-real-time indexing pipelines powering sub-second customer self-service search across millions of historical invoices and orders.

SQL Server Administration & Performance Engineering

- Eliminated high-frequency lock-conversion deadlocks and write skew on concurrent composite-key updates by pairing targeted locking hints (UPDLOCK, HOLDLOCK, ROWLOCK) with deterministic execution ordering in UPSERT and batch routines, removing index traversal races and PK violation errors.
- Enabled and tuned Read Committed Snapshot Isolation to remove reader-writer blocking and cut transaction contention across high-throughput OLTP workloads.
- Tuned execution plans to eliminate tempdb spills by correcting memory grant allocations, parameter sniffing behavior, and stale statistics.
- Redesigned composite clustered and covering non-clustered indexes to remove bookmark lookup bottlenecks and cross-index deadlock graphs; managed fill factors and selective index and heap rebuilds, and partitioned large tables to improve partition elimination as data volume grew.
- Automated enterprise database maintenance, index defragmentation, and statistics updates using Ola Hallengren's solution orchestrated through SQL Server Agent.
- Administered Always On Availability Groups and automated backup and recovery routines as part of the HA/DR strategy, and configured and tuned CDC to support low-latency replication into the enterprise data warehouse.
- Built monitoring and alerting in SolarWinds covering blocking, waits, and resource pressure; authored custom T-SQL administrative scripts auditing index fragmentation, connection pooling efficiency, and storage utilization; managed database roles, permissions, and least-privilege access.

Semantic Layer, Master Data & Application Platform

- Engineered enterprise LookML semantic models and Explores, defining standardized business logic and KPIs across core supply chain dimensions including inventory turns, dock-to-stock cycle times, order fill rates, and transportation freight spend.
- Set up Google's Model Context Protocol (MCP) server integration in Looker, exposing curated semantic definitions to generative AI agents to enable natural-language-to-SQL querying and prompt-driven dashboard generation.
- Implemented granular RBAC and user attribute-based row- and column-level security in Looker, restricting sensitive financial data and PII to authorized user groups, and administered the instance environment covering user provisioning, content and space curation, schedule delivery alerts, and usage monitoring.
- Optimized database workload and dashboard render speeds by configuring Looker datagroups, caching strategies, and persistent derived table lifecycle policies, eliminating redundant query execution against underlying database engines.
- Architected a scalable Master Data Management ecosystem across Zoho, establishing governed golden records for customer, item, and supplier entities with automated distribution to tier-1 ERP, WMS, and TMS systems.
- Designed an event-driven, decoupled integration pipeline using Zoho webhooks and resilient message queuing, eliminating data discrepancies and synchronization lag across warehouse fulfillment and freight routing networks.
- Engineered high-throughput, asynchronous FastAPI backend microservices with strict Pydantic schema validation, standardizing departmental custom data capture and preventing operational data corruption across downstream pipelines.
- Built and provisioned a fault-tolerant, auto-scaled internal application tier on AWS (EC2 in multi-AZ VPCs, Auto Scaling, hardened security groups, Apache HTTPD reverse proxies, Route 53) with automated open-source TLS rotation, unifying departmental data capture portals.

Data Engineer Feb 2018 – Oct 2022

Data Warehouse Development

- Architected multi-source ingestion pipelines using PySpark, Matillion, and Boomi to ingest structured and semi-structured data from relational databases (via CDC and watermark extracts), SaaS APIs, and third-party flat files into Amazon Redshift.
- Engineered hybrid batch and near-real-time streaming architectures decoupled via Amazon SQS, automating data staging, deduplication, cleansing, and complex business logic in SQL.
- Designed and maintained dimensional models (star and snowflake schemas), populating enterprise fact and dimension tables to power downstream BI analytics and executive reporting in Looker.
- Implemented reverse-ETL pipelines to operationalize curated warehouse data, pushing enriched analytical metrics back into operational SaaS platforms including Salesforce and Zoho.

Redshift Administration, Performance Tuning & Governance

- Optimized physical storage architecture by configuring optimal distribution styles (KEY, ALL, EVEN), compound sort keys, and column compression encodings (AZ64, ZSTD), reducing cluster storage footprint and network slice data redistribution (DS_DIST_BOTH).
- Configured Workload Management parameter groups, memory allocations, query monitoring rules, and user priority queues to prevent heavy batch ETL from bottlenecking high-priority BI dashboard queries.
- Accelerated query execution by implementing materialized views, refactoring complex query patterns, and scheduling automated maintenance procedures (ANALYZE, VACUUM) for zone-map refresh and disk space reclamation.
- Hardened cluster security and observability by enforcing PII data encryption, implementing role-based access control, and deploying CloudWatch dashboards and alarms for disk spill, CPU utilization, and concurrency metrics.

Analytics Delivery & Cross-Functional Partnership

- Partnered with supply chain stakeholders across Procurement, Receiving, Inventory Control, Warehouse Operations, Transportation, Sales, and Finance to identify operational bottlenecks and translate complex requirements into intuitive reporting suites.
- Delivered end-to-end Looker dashboards tracking daily order fulfillment, warehouse throughput, shipment status, and inventory variance, driving measurable efficiency improvements across distribution centers.
- Designed hybrid analytical architectures merging historical trend analysis with real-time operational reporting to give floor managers and executives unified visibility into daily logistics and distribution workflows.
- Standardized cross-departmental reporting definitions, eliminating data silos and KPI discrepancies across inventory accounting, sales performance, and logistics operations.

EDUCATION

University of Illinois Master's, Management Information Systems 2017

JNTU B.Tech, Electronics and Communication Engineering 2015