Data Engineering & Cloud Architecture Portfolio

GCP Solutions Delivering Measurable Business Impact

Wagner Silva | GCP Architect & Senior Data Engineer

Lisbon → Stockholm

Portfolio at a Glance

10

Production Projects

Enterprise-grade solutions

€48K

Annual Cost Savings

Infrastructure optimization

Current

Tech Skills

Sharpened & up-to-date

7+

Years Experience

Cloud-native expertise

These 10 projects are a sample of my work—
reliable, scalable solutions built on peer cooperation and cross-team collaboration.

Technology Stack & Capabilities

Cloud & Infrastructure

- Google Cloud Platform
- Kubernetes (GKE)
- Docker & Containerization
- Terraform (IaC)
- Linux Systems

Data & Databases

- BigQuery, Datastream, Pub/Sub
- Cloud Storage
- AlloyDB, Cloud SQL (PostgreSQL)
- Firestore, Bigtable, Memorystore
- Neo4j (Graph), Qdrant (Vector)

Dev & Integration

- Python
- TypeScript, SQL
- Cloud Run & Jobs, Functions
- APISIX, REST APIs, SFTP
- Vertex AI, Gemini, RAG

Security & Governance

Protecting Data at Every Layer

Project 1: Custom Row-Level Security

Secure Public Reporting without Login

CHALLENGE

- Public Looker Studio reports with row-level security but no user login
- Need to enforce data access permissions without authentication
- Scalability to thousands of anonymous users

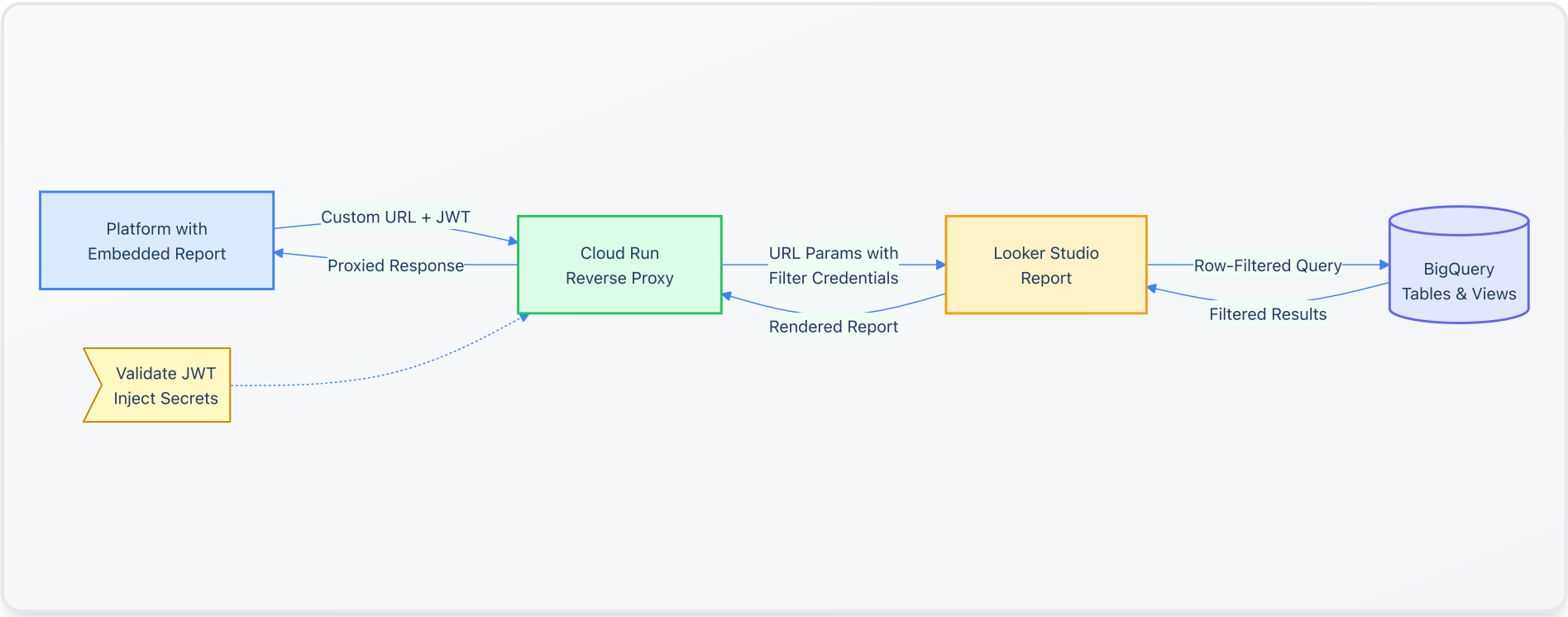
SOLUTION

- Python reverse proxy intercepts JWT tokens from public URLs
- Injects security credentials and user context into requests
- Filters BigQuery views based on embedded permissions

IMPACT

- Secure public reporting with zero data exposure
- Scalable to thousands of concurrent users
- No user authentication required while maintaining security

Python | Cloud Run | Looker Studio | BigQuery | JWT



Project 2: On-Demand Database Anonymization

Safe Development & Testing with Production-Quality Data

CHALLENGE

- Developers need fresh, realistic data on-demand
- Cannot expose production PII/sensitive data
- Need ability to refresh environments whenever needed

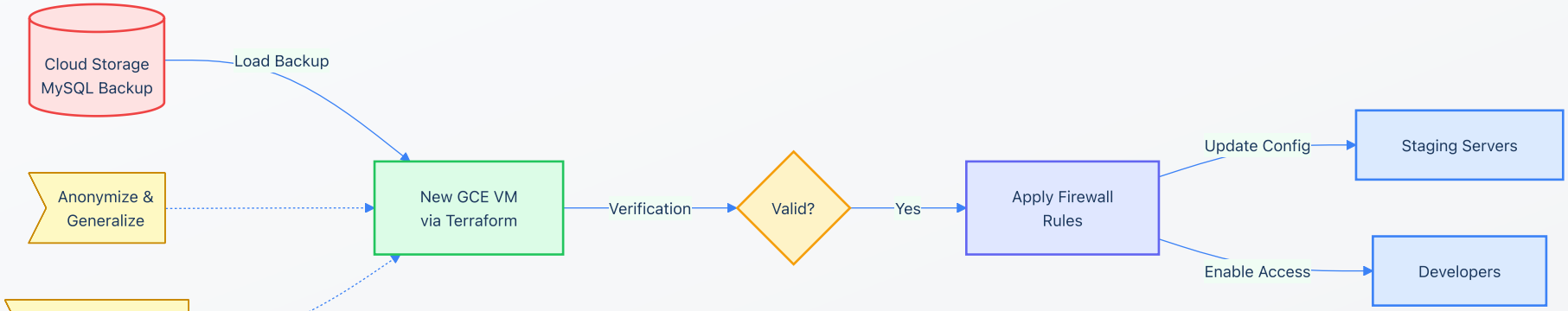
SOLUTION

- On-demand snapshot and anonymization tool (manual trigger)
- Python-based PII masking maintaining referential integrity
- Creates completely new database instances for dev/staging

IMPACT

- High-quality testing with realistic, fresh data
- Full data security compliance with zero PII exposure
- Fast environment provisioning (< 2 hours for full refresh)

Python | SQL | Cloud SQL | Cloud Storage | Firewall | GCE | Terraform



Project 3: Real-Time CDC Pipeline

ZNUNY/OTRS Support Analytics

CHALLENGE

- Real-time analytics required from operational MySQL database
- Data minimization requirements for compliance
- Cannot impact production database performance

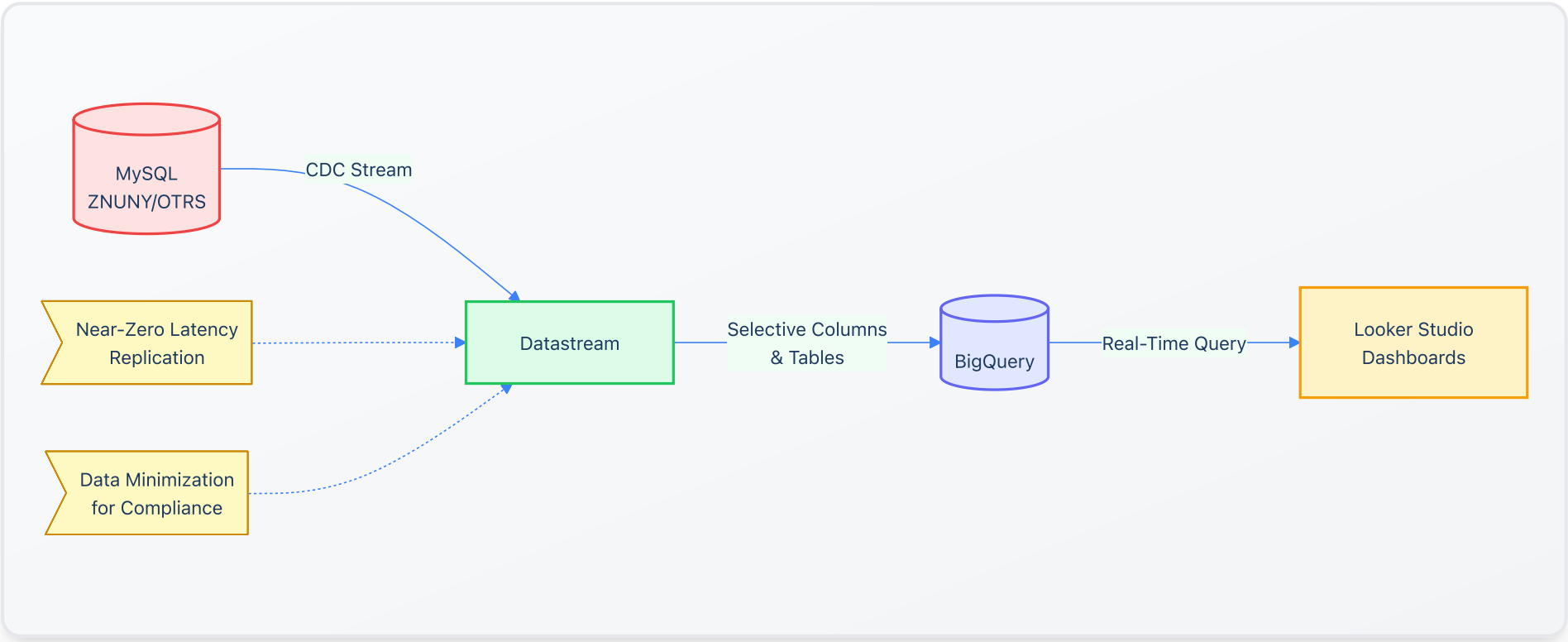
SOLUTION

- Datastream CDC for near-zero latency replication
- Selective column and table replication to BigQuery
- Real-time transformation pipeline for analytics

IMPACT

- Real-time support analytics dashboards in Looker Studio
- Reduced operational costs through selective replication
- Enhanced security with data minimization

MySQL | Datastream | BigQuery | Looker Studio



Data Integration & Pipelines

Scalable & Intelligent Data Flows

Project 4: Reusable SFTP Architecture

Zero Data Loss for Global Clients

CHALLENGE

- Sync product registration data with zero data loss tolerance
- Multiple global clients with similar SFTP requirements
- File format validation and error recovery needed

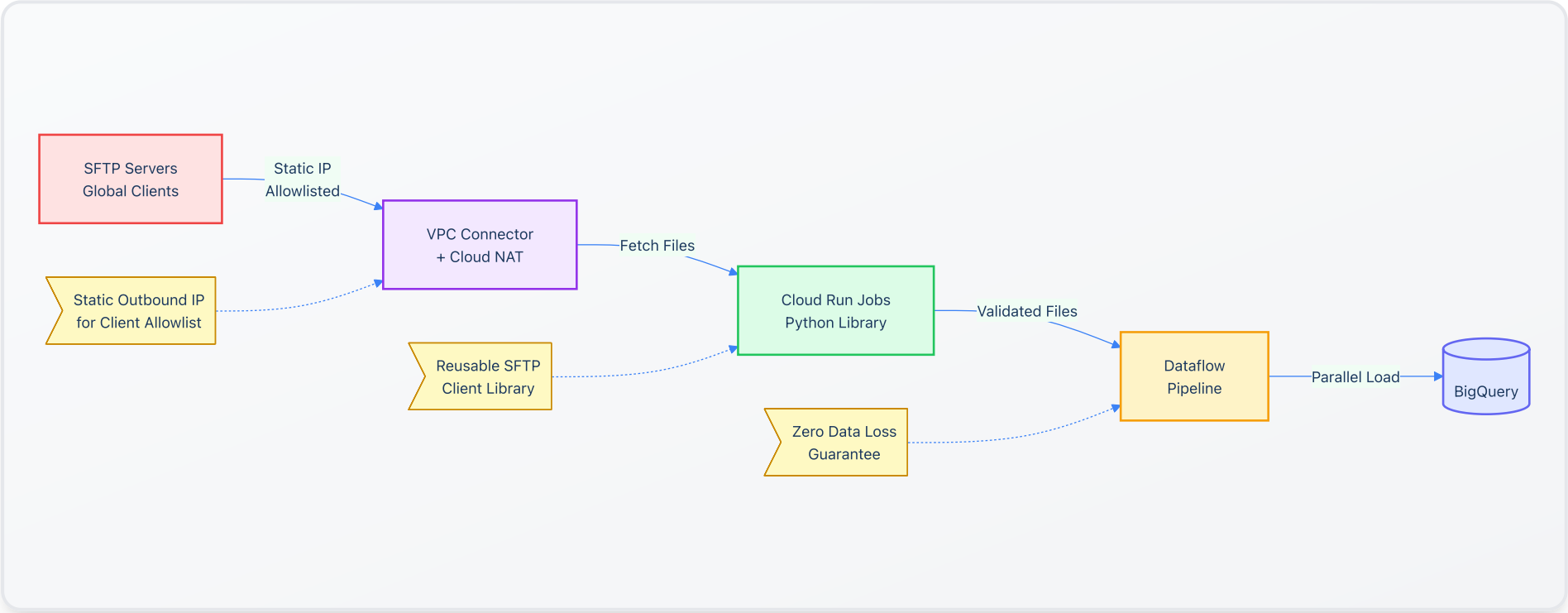
SOLUTION

- Reusable Python SFTP client library with comprehensive validation
- Atomic file processing with transactional guarantees
- Dataflow pipelines for parallel processing and BigQuery loading

IMPACT

- Zero data loss achieved across all client integrations
- 70% faster implementation time for second client
- Reusable architecture deployed to 3 different regions

Python | SFTP | Cloud Run Jobs | Dataflow | BigQuery | VPC Connector | Cloud NAT



Project 5: Smart API Integration

Pascom PBX Call Center Data Sync

CHALLENGE

- Pascom API has no filtering or pagination support
- Returns ALL data on every call (5-10K records)
- Need near real-time sync without duplicates

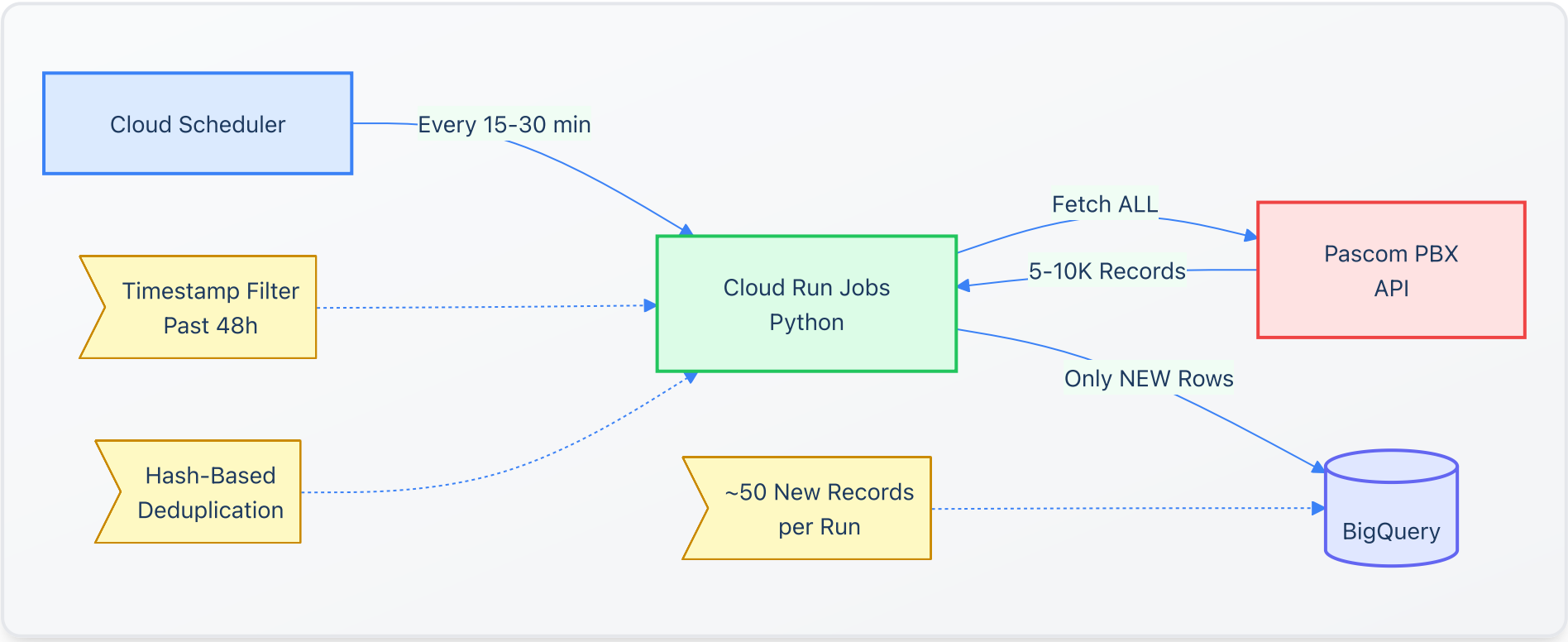
SOLUTION

- Python sync job on Cloud Run Jobs + Scheduler
- Timestamp filtering (past 48h) + hash-based deduplication
- Only loads NEW rows to BigQuery (no duplicates)

IMPACT

- Real-time call analytics operational
- No duplicate data in BigQuery ever
- Handles thousands of calls daily, runs every 15-30 min

Python | Cloud Run Jobs | Cloud Scheduler | BigQuery



Multi-System Orchestration

Synchronizing Global Operations

Project 6: Tri-Directional Sync

Salesforce, Internal DB, TalentLMS

CHALLENGE

- Keep 3 systems in sync with bidirectional data flow
- Different data models and update frequencies across systems
- Conflict resolution for concurrent updates

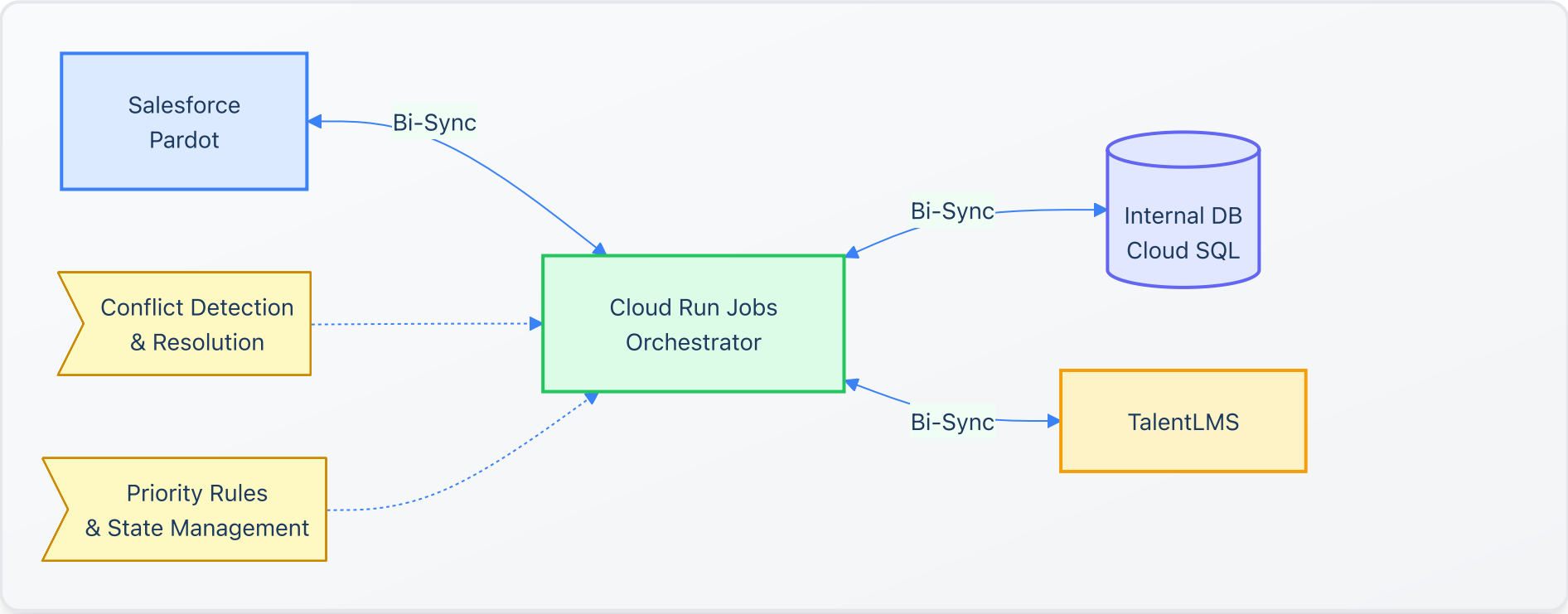
SOLUTION

- TypeScript orchestration engine with state management
- Conflict detection and resolution with priority rules
- Cloud Run Jobs with scheduled and event-driven execution

IMPACT

- Single source of truth across all platforms
- 20+ hours per week saved from manual data entry
- 99.9% synchronization success rate

TypeScript | Cloud Run Jobs | Salesforce API | TalentLMS API | REST



AI & Innovation

Enterprise AI Platform

Project 7: Enterprise RAG + AI Agents

Organization-Wide AI Platform

CHALLENGE

- AI search with role-based access and semantic understanding
- Need organization-wide AI without individual subscriptions
- SSO integration and cost control required

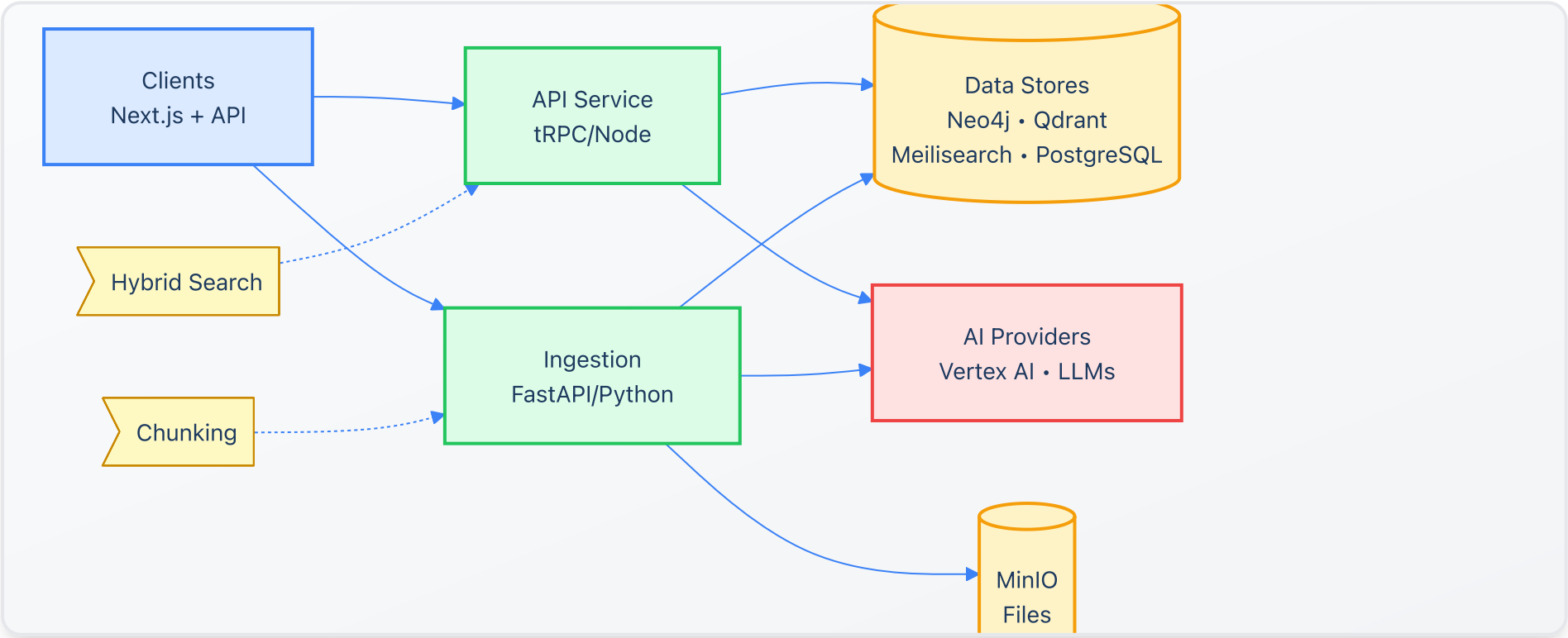
SOLUTION

- Multi-database RAG: Qdrant (vectors), Neo4j (graph), PostgreSQL
- Multi-level Vertex AI agents with Cloud IAP SSO
- Deployed on GKE for scalability

IMPACT

- 85% reduction in external AI subscription costs
- Semantic search with permission-based filtering
- Organization-wide AI access through single platform

Python | TypeScript | FastAPI | tRPC | Neo4j | PostgreSQL | Qdrant | Meilisearch | MinIO | Vertex AI | GKE



Enterprise Leadership

Cross-Functional Team Coordination

Project 8: Technical Leadership

SharePoint & Power BI for Global Conservation Foundation

CHALLENGE

- Complex stakeholder ecosystem (technical & non-technical)
- Microsoft 365 integration requirements
- Cross-functional team coordination needed

SOLUTION

- SharePoint site architecture + Power Automate workflows
- Power BI reporting infrastructure for executives
- Bridged technical and creative teams effectively

IMPACT

- Data-driven conservation decisions enabled
- 80% reduction in manual reporting effort
- Improved international collaboration

SharePoint | Power BI | Power Automate | Microsoft 365

ACTIONS TAKEN

Technical Delivery

- SharePoint hub with document libraries
- Power Automate data collection workflows
- Power BI executive dashboards
- Data integration from multiple sources

Stakeholder Management

- Foundation executive team
- Field staff across continents
- External designers and copywriters
- IT team coordination

Report Types Delivered

- Executive Dashboard (global metrics)
- Regional Reports (Americas, EMEA, APAC)
- Financial Reports (donor tracking)
- Impact visualization dashboards

Leadership Skills Applied

- Iterative feedback loops
- Clear documentation at all levels
- Consensus building across teams
- Flexible to changing requirements

Infrastructure & Cost Optimization

Smart Infrastructure Decisions

Project 9: GCP Cost Optimization

€15,000 Annual Savings

CHALLENGE

- Rising cloud costs threatening project sustainability
- No visibility into cost drivers and inefficiencies
- Need to optimize without impacting performance or reliability

SOLUTION

- Real-time cost monitoring dashboards and automated alerts
- Right-sizing analysis for compute, storage, and database resources
- Service optimization: Cloud Run vs Functions, Spot VMs, committed use discounts
- Code-level optimization: BigQuery query tuning and caching strategies

IMPACT

- €15,000 annual savings (30% cost reduction)
- Cost-conscious engineering culture established
- Automated cost monitoring prevents future overspend
- Optimized BigQuery usage patterns

BigQuery | Cloud Run | Cloud Monitoring | Terraform

ACTIONS TAKEN

Active Monitoring

- Real-time cost tracking dashboards
- Budget alerts and anomaly detection
- Cost allocation by project/team
- Weekly cost review meetings

Right-Sizing

- Cloud SQL: Reduced instance tiers
- GKE: Node pool right-sizing & autoscaling
- Storage: Moved archival to Coldline
- Eliminated over-provisioning

Service Optimization

- Cloud Run vs Functions selection
- Batch jobs → Cloud Run Jobs
- Dev environments: Spot VMs (60% off)
- Committed Use Discounts applied

Code & Queries

- BigQuery partitioning & clustering
- Redis caching layer for APIs
- Optimized data retention policies
- Query optimization: 70% less scanned

Project 10: APISIX + Crowdsec Gateway

Enterprise Security at Zero Cost

CHALLENGE

- Cloud Armor cost €36k/year unsustainable for budget
- Need enterprise-grade DDoS and threat protection
- Require better observability and custom rules

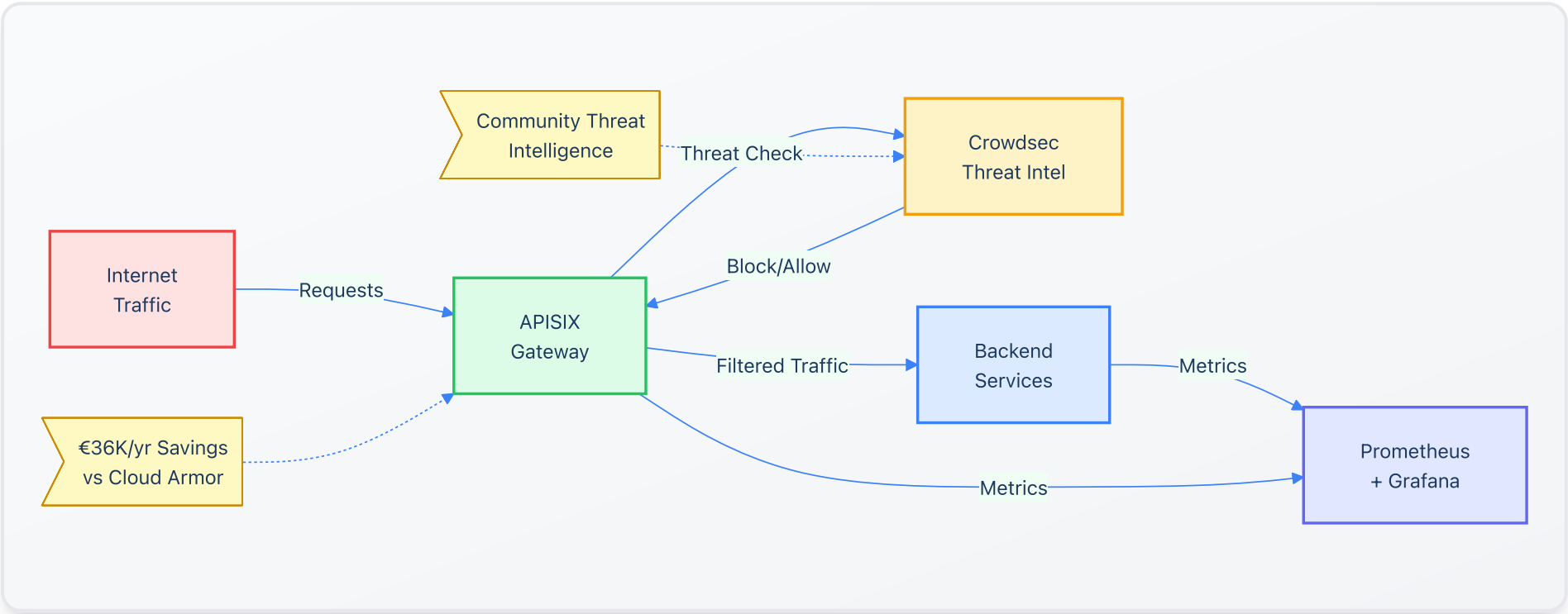
SOLUTION

- Open-source APISIX API Gateway deployed on GCE MIG
- Crowdsec threat intelligence integration for real-time protection
- Prometheus and Grafana for comprehensive monitoring

IMPACT

- €36k annual savings with zero compromise on security
- Better observability and custom rule configuration
- Enhanced security with community threat intelligence

APISIX | Crowdsec | GCE | MIG | Prometheus | Grafana



Total Business Impact

€15K

GCP Optimization

Infrastructure tuning

€33K

Cloud Armor Replacement

APISIX Implementation

85%

AI Cost Reduction

Scalable pay-as-you-go

€48K+

Total Annual Savings

Proven cost efficiency

Core Strengths & Highlights

Reusable

Architectures

Proven at scale

Zero

Data Loss

Financial data protection

€48K+

Savings Delivered

Cost efficiency

Security

First

Enterprise-grade

Global

Scale

Multi-continent ops

Innovation

AI/ML

RAG & Agents

Alignment with Role

Requirements

- 7+ years Data Engineer experience
- Strong GCP (BigQuery, Cloud Run, etc.)
- SQL proficiency & Python/TypeScript
- Terraform & CI/CD
- NoSQL & RDBMS experience

My Commitment

- With 7+ years delivering production data systems and a proven track record of impact, I bring the experience you need —and the curiosity to master whatever comes next.

Why Stockholm?

Career Motivation

- Specialization in GCP Data Engineering
- Modern data architectures
- Learning from Swedish tech excellence

Cultural Fit

- Appreciate Swedish values (honesty, equality)
- Cost-conscious mindset
- Quality over quantity philosophy

Relocation

- Targeting Stockholm specifically
- Ready to relocate (need Swedish work permit)
- Committed to long-term move

Logistics & Next Steps

Lisbon

Current Location

Ready to move

Visa

Status

Need Swedish work permit

ASAP

Start Date

Flexible / Remote start

Contact

Get in Touch

Email / LinkedIn

Thank You

Ready to bring data engineering excellence to Stockholm

Proven Impact

Wagner Silva | GCP Architect & Senior Data Engineer