

Secure Data Infrastructure: Case Study

1. Executive Summary

A specialized infrastructure project developed during a 2021-2022 internship focusing on bridging secure cloud storage with enterprise payment logic. The primary objective was to build a robust, high-availability pipeline to handle sensitive data migration and automated billing verification.

2. The Problem Statement

- **Manual Latency:** Existing data synchronization between local environments and cloud storage was manual and prone to human error.
- **Security Risks:** Data was being handled without a unified encryption standard during transit.
- **Billing Gaps:** No automated link existed between successful data processing and customer billing triggers.

3. Technical Architecture (The Logic)

Since this code is archived for security, the logic flow is outlined below:

- **Layer 1: Entry & Auth:** Utilized AWS IAM (Identity and Access Management) roles to ensure a "least-privilege" access model. No root accounts were used for daily operations.
- **Layer 2: Processing:** A backend service (Node/Python) monitored incoming data streams, validating file integrity before proceeding.
- **Layer 3: Encryption & Storage:** Implemented AES-256 encryption-at-rest within AWS S3 buckets. All transfers were forced over HTTPS/TLS.
- **Layer 4: Payment Bridge:** Integrated the **Stripe API** via webhooks. Upon successful data archival, a signal was sent to Stripe to verify the transaction and generate a receipt automatically.

4. Security & Compliance Highlights

- **Environment Isolation:** Strict separation of `.env` variables between Development, Staging, and Production.
- **Error Handling:** Implemented "Fail-Safe" logic where a payment failure would automatically roll back a data migration to prevent out-of-sync records.
- **Audit Logging:** Every interaction with the AWS infrastructure was logged for compliance and security monitoring.

5. Key Results & Impact

- **Automation:** Reduced manual data handling by approximately 100% through the new pipeline.
- **Reliability:** Achieved zero data loss during the final testing phase.
- **Scalability:** The infrastructure was designed to handle a 300% increase in data volume without requiring manual scaling.