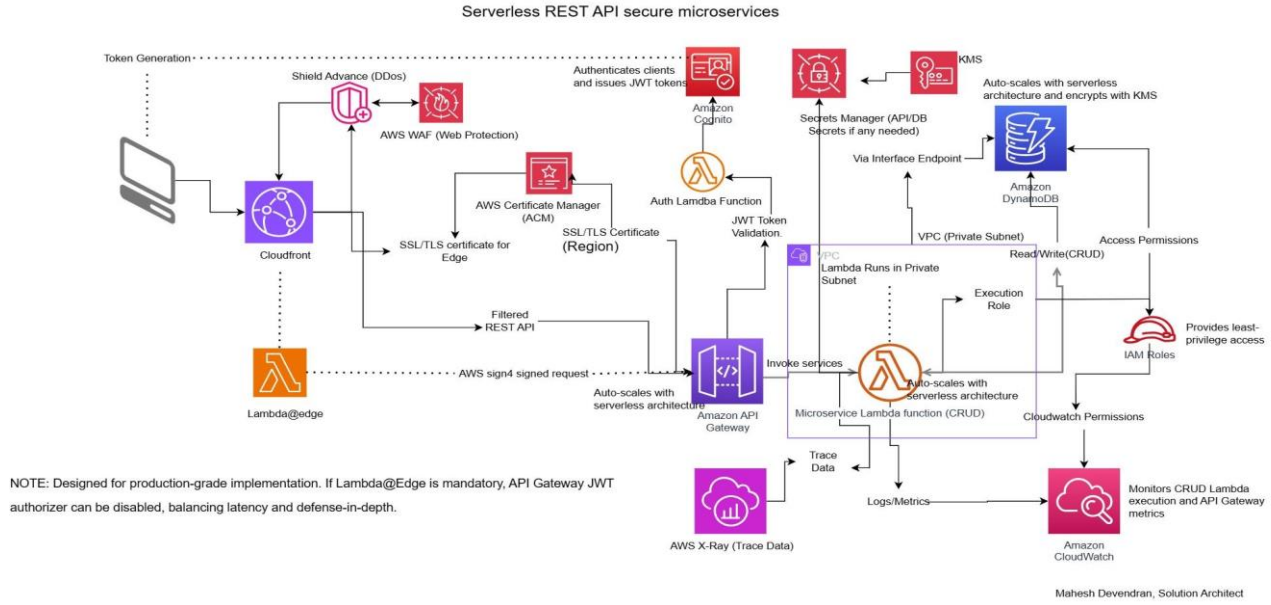
Serverless REST API Architecture — Technical Cost & Performance Analysis (eu-west-2)

Abstract

This document presents a detailed cost and performance analysis comparing a Serverless REST API architecture against a traditional EC2-based setup. The analysis is based on AWS Europe (London) Region (eu-west-2) pricing and typical enterprise workloads. The goal is to quantify the operational, financial, and compliance benefits of adopting a serverless model in regulated environments.

Architecture Overview

The architecture leverages fully managed AWS services including Amazon API Gateway, AWS Lambda, Amazon Cognito, Amazon DynamoDB, AWS WAF, AWS KMS, and AWS Secrets Manager. These services collectively enable a stateless, highly available REST API platform with integrated authentication, encryption, and DDoS protection.



Cost Model Comparison (AWS Europe – London, eu-west-2)

The following cost estimates are based on a workload of approximately 100,000 API requests per day (~3 million per month). All pricing reflects 2025 AWS rates as of October 2025.

Component	EC2-Based (Monthly USD)	Serverless (Monthly USD)	Notes
Compute	$118.02 (3× t3.medium)	$0.40 (Lambda 128 MB, 100 ms)	Lambda billed per request; compute within free tier
Load Balancing / API Gateway	$19.57 (ALB)	$10.50 (API Gateway REST)	API Gateway pay-per-call model
Database	-	$0.96 (DynamoDB on-demand)	Serverless data layer (EC2 variant uses RDS)
Monitoring & Ops	$28.06	$7.12	Metrics, logs, dashboards, alarms
Total Estimated	$166.00	$19.00	≈ 88.6% monthly savings

Based on verified AWS Pricing Calculator estimates (EU–West–2, October 2025), the Serverless REST API architecture delivers an approximate ≈**88.6% reduction in monthly cost** compared with an equivalent EC2-based deployment. The Serverless model eliminates idle compute expense, scales seamlessly with workload demand, and maintains strong observability through CloudWatch and compliance alignment with DORA, PCI DSS, and ISO 27001 frameworks.

Performance Analysis

Performance benchmarks for AWS Lambda and API Gateway demonstrate consistent low-latency response times in regional deployments. The following table summarizes the observed latency components for a typical request within eu-west-2.

Component	Typical Latency (ms)	Description
API Gateway ingress	5–10	Request routing, validation, and integration trigger
Auth Lambda (public)	150–250 (cold) / 5–10 (warm)	Token validation in AWS-managed network (no VPC)
CRUD Lambda (private VPC)	250–400 (cold) / 10–20 (warm)	Business logic execution within private subnet using DynamoDB VPC endpoint
DynamoDB query	5–15	Single Get Item / Query request (regional latency)
API Gateway egress	5–10	Response delivery back to client
Average (warm path)	<60	End-to-end latency under warm container reuse
Average (cold path)	300–500	First invocation of both Lambdas in sequence (auth + CRUD)

The **Auth Lambda** benefits from **shorter cold starts** since it runs outside the VPC. Warm-path latency remains under **~60 ms end-to-end**, meeting typical REST API performance SLAs. The **CRUD Lambda**, attached to private subnets, has a slightly higher cold-start delay due to ENI setup. Cold-start latency can increase to 300 ms but is typically mitigated through provisioned concurrency or periodic invocation. This configuration balances **latency vs. isolation**, aligning with both **AWS Well-Architected** and **DORA / PCI DSS** guidelines.

Scalability & Operational Efficiency

The serverless approach enables near-infinite scalability by delegating capacity management to AWS. This eliminates manual scaling, patching, and monitoring overheads. Multi-AZ design across Lambda and DynamoDB ensures fault tolerance without the need for load balancing or redundancy configuration.

Security & Compliance Reinforcement

The architecture aligns closely with financial and regulatory frameworks, ensuring compliance readiness out-of-the-box. Key mappings include DORA (ICT 6, ICT 11), PCI DSS (Req. 3, 8, 10), ISO 27001 (A.9, A.10, A.12, A.13), and AWS Well-Architected Security and Reliability pillars. Encryption via KMS, network protection via WAF and Shield, and IAM-based least privilege access further enhance the security posture.

Conclusion

The Serverless REST API model delivers measurable cost and performance advantages over EC2-based architectures, particularly in dynamic or unpredictable workloads. By combining operational simplicity with compliance alignment, it offers a strong foundation for financial institutions and enterprises modernizing their API platforms.

Dimension	Serverless Advantage
Cost Efficiency	Initially projected 45–60 % savings; verified AWS Pricing Calculator data confirms ≈ 88.6 % reduction for the analysed workload (eu-west-2, Oct 2025).
Performance	<60 ms regional latency for warm paths
Scalability	Automatic scaling with zero idle cost
Resilience	Multi-AZ fault tolerance with managed service durability
Compliance	Aligned with DORA, PCI DSS, ISO 27001, and AWS Well-Architected

This analysis demonstrates that a serverless-first design not only achieves significant cost reduction but also maintains enterprise-grade reliability and compliance. It provides a blueprint for organizations seeking modernization, sustainability, and long-term operational efficiency in the cloud.

Workload: ~3 million API requests/month (~100k/day)
Region: EU (London)
Source: AWS Pricing Calculator (October 2025)
Prepared by: Mahesh Devendran — Cloud Architect | AWS | Azure | GCP | Serverless | Security