

AWS Lambda

AWS Lambda is a serverless compute service from Amazon Web Services (AWS) that lets you run code (functions) without provisioning or managing servers, automatically handling scaling and infrastructure, allowing you to pay only for compute time used, ideal for event-driven tasks like data processing, web backends, and automation. Developers upload code in languages like Python, Node.js, Java, etc., and Lambda executes it in response to triggers (e.g., S3 file upload, API call).

Key Concepts

Serverless: Focus on code, not servers; AWS manages the underlying infrastructure.

Event-Driven: Functions run in response to events (triggers) from other AWS services or custom sources.

Automatic Scaling: Scales from zero to handle traffic spikes and scales down when idle.

Pay-Per-Use: You only pay for the milliseconds your code runs.

How it Works

Write Code: Create a function (e.g., in Python, Node.js, Java).

Upload: Package and upload your code as a ZIP file or container image.

Set Trigger: Configure an event source (e.g., S3, API Gateway, DynamoDB) to invoke your function.

Execute: Lambda automatically provisions resources, runs your code when triggered, and then shuts down.

Common Use Cases

File Processing: Resize images after S3 upload.

Data Processing: Real-time stream processing, batch jobs.

Web & Mobile Backends: APIs, IoT backends.

Scheduled Tasks: Run code on a timer.

Benefits

Reduced Operations: No server management, patching, or OS updates.

Cost-Effective: Pay for execution time, not idle servers.

High Availability: Built on fault-tolerant AWS infrastructure.

Scalability: Handles massive, unpredictable traffic automatically.

Revision #2

Created 29 October 2025 02:43:40 by AI API

Updated 11 December 2025 07:25:30 by AI Channel