



# Things you need to know to become a Serverless API Ninja

Allan A. Chua

API Days Singapore 2020



Allan Chua



Solutions & Integrations  
Architect @ FWD

1.5

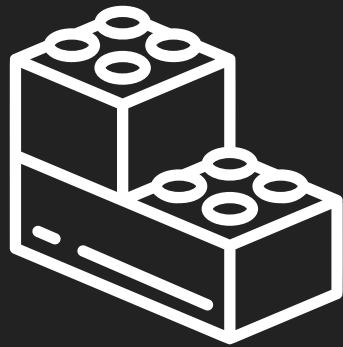
Years of Experience  
with Serverless



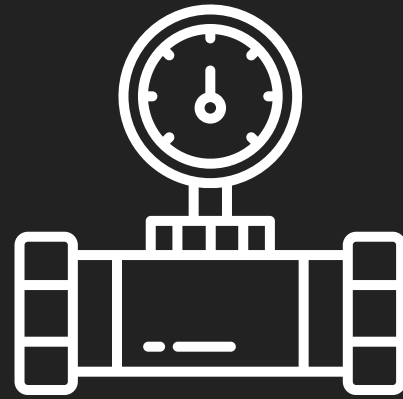
AWS Certified  
Solution Architect



Writing a Book  
about Serverless



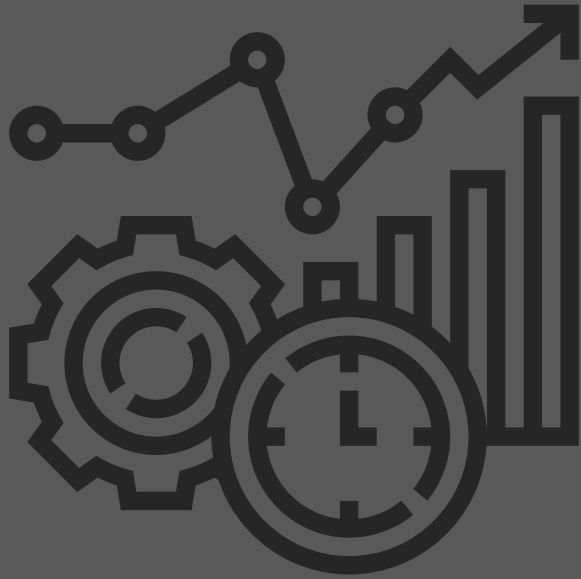
BASICS



INTEGRATIONS



CAUTION



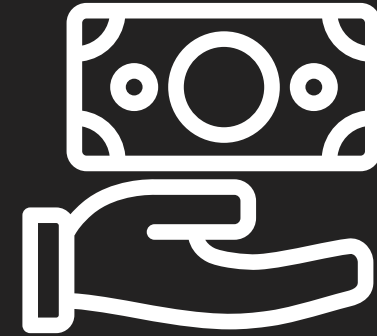
# Serverless Means Efficiency



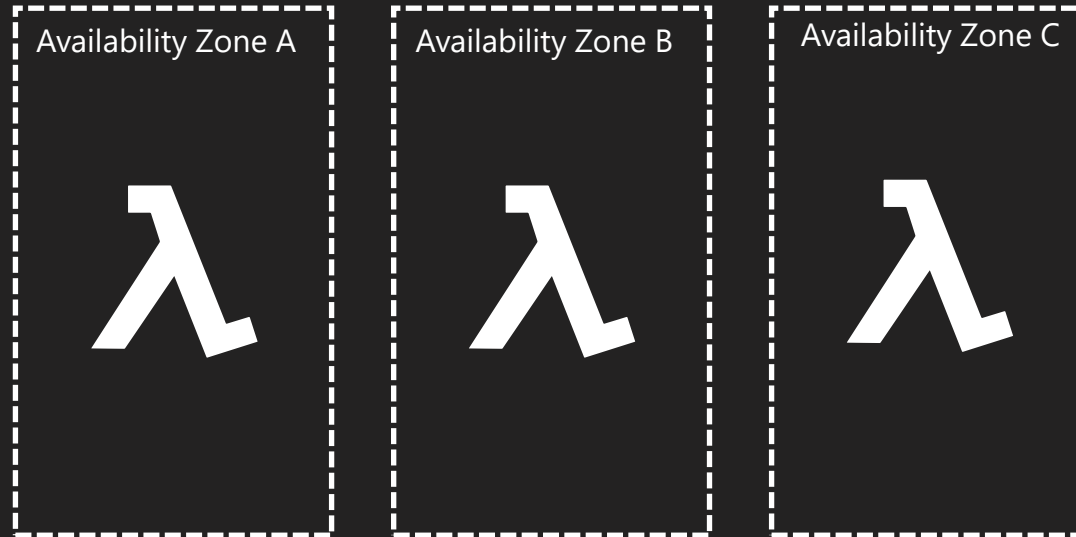
No Servers to  
Manage



Reduced Security  
Risk & Cost



Pay only for what is  
used



Built-in High Availability &  
Disaster Recovery

\$135.05

1 EC2 t2.xlarge  
\$0.185 / Hour

\$ 270.10

2 EC2 t2.xlarge  
\$0.370 / Hour  
High Availability

What traditional architectures will cost you per API Cluster

## Stacks (71)

▼ Lambda		\$0.00
▼ Asia Pacific (Singapore)		\$0.00
AWS Lambda APS1-Lambda-GB-Second		\$0.00
AWS Lambda - Compute Free Tier - 400,000 GB-Seconds - Asia Pacific (Singapore)	330,762.525 Lambda-GB-Second	\$0.00
AWS Lambda APS1-Request		\$0.00
AWS Lambda - Requests Free Tier - 1,000,000 Requests - Asia Pacific (Singapore)	92,060.000 Request	\$0.00
▼ Asia Pacific (Tokyo)		\$0.00
AWS Lambda APN1-Lambda-GB-Second		\$0.00
AWS Lambda - Compute Free Tier - 400,000 GB-Seconds - Asia Pacific (Tokyo)	1,868.013 seconds	\$0.00
AWS Lambda APN1-Request		\$0.00
AWS Lambda - Requests Free Tier - 1,000,000 Requests - Asia Pacific (Tokyo)	4,136.000 Requests	\$0.00

96K API calls worth 91 Hours of actual runtime  
for 71 CF Stacks in DEV environment for FREE



▼ Simple Queue Service		\$2.22
▼ Asia Pacific (Singapore)		\$2.22
Amazon Simple Queue Service APS1-Requests-Tier1		\$2.22
\$0.40 per 1,000,000 Amazon SQS Requests per month thereafter	5,559,154.000 Requests	\$2.22
First 1,000,000 Amazon SQS Requests per month are free	1,000,000.000 Requests	\$0.00

**6.6 Million** Queue Messages for **2\$**  
In DEV Environment

\$ 19,710

\$270.10 per cluster x 71

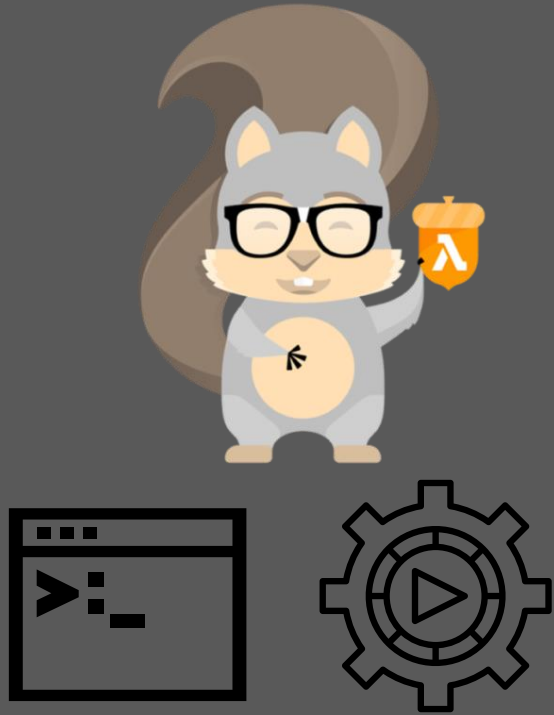
Traditional Architecture



\$ 0

Lambda based architecture

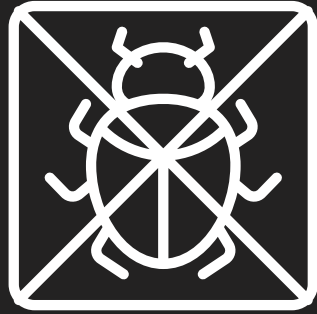
We drove the cost of dev environments



Working with AWS Lambda  
is way more convenient with  
**AUTOMATION**



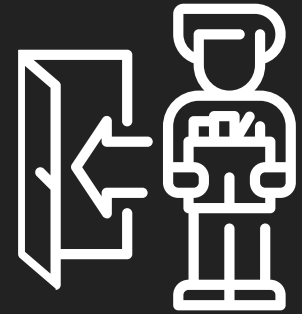
Reliable &  
Repeatable  
Builds



Reduced  
Human Errors

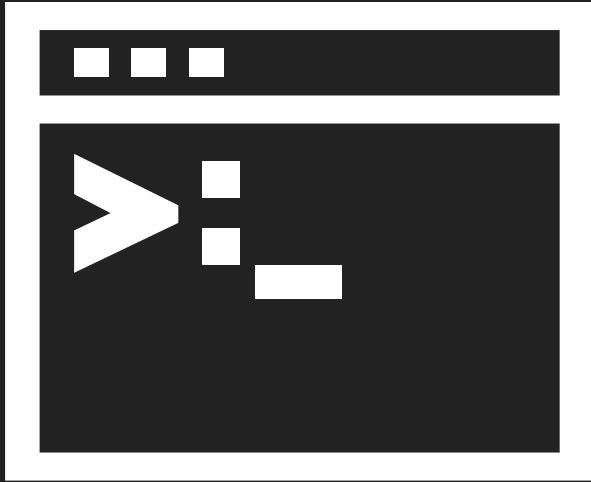


Maintainability



Protects businesses  
from developer  
turnover

Benefits of Automation in  
API Development



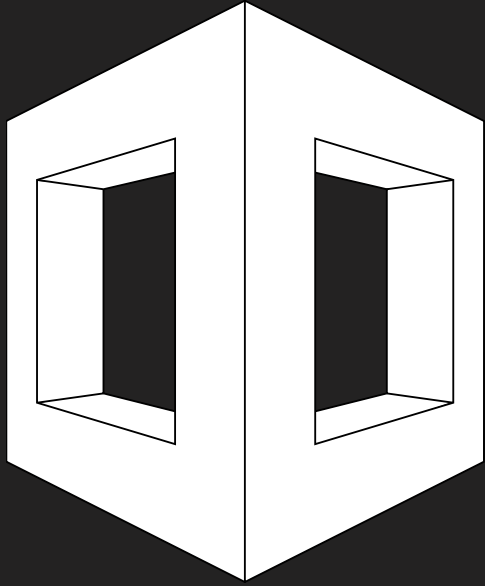
## Shell Scripting

- Deployment Automation (CI/CD)
- Dependency Management
- Decommissioning
- Validation Tasks



AWS SAM

- ✓ Definition of Serverless Resources
- ✓ Local testing of functions
- ✓ Shorthand Syntax for Cloud Formation



## Cloud Formation

- ✓ Definition of AWS resources
- ✓ Grouping of AWS resources
- ✓ Base syntax used by SAM and other frameworks



Automation  
Sample Code

- Shell Scripts
  - Install Dependencies
  - Release API + Tables
  - Decommissioning
- SAM Template
  - API Gateway
  - API Key
  - Usage Plan
  - Lambda APIs
- Cloud Formation
  - Dynamo DBs



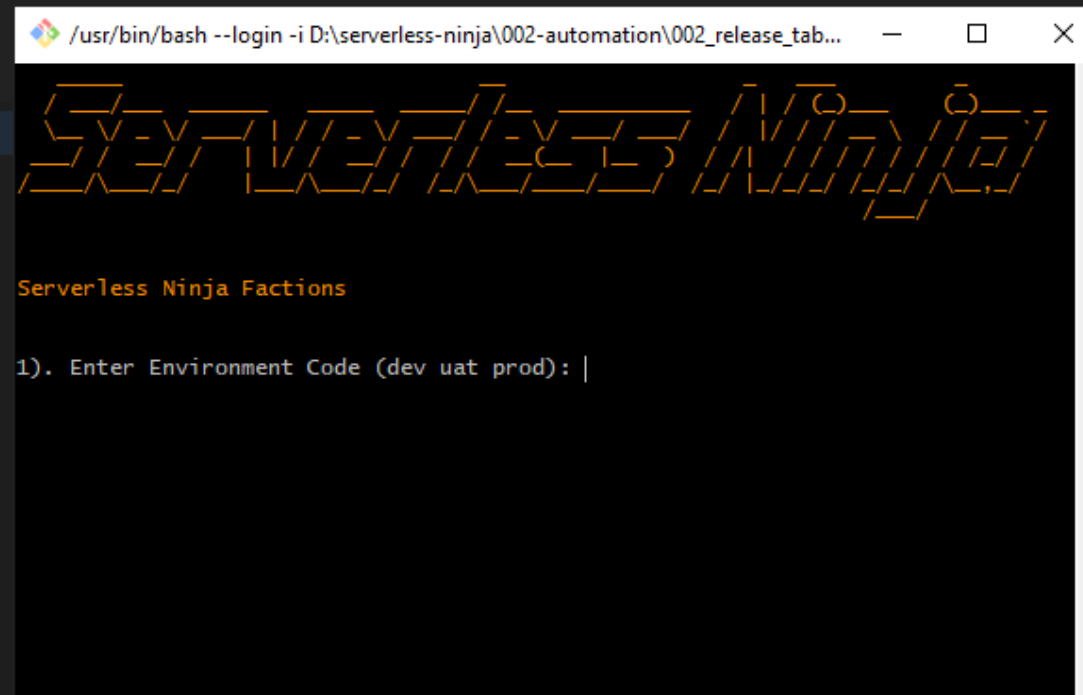
```
5 > AppName: ...
```

### Resources:

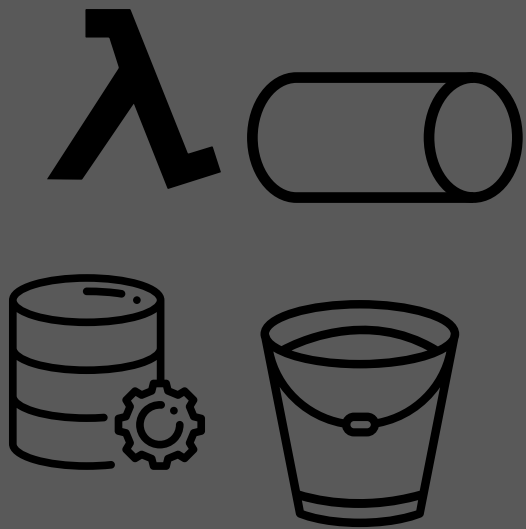
FactionsApiGateway: ...

FactionsApiKey: ...

FactionsUsagePlan: ...



Have fun playing with the IAC trainers.



Serverless  
is **not limited**  
to Lambdas

The most common  
**misconception** about  
serverless is:



**Serverless = AWS Lambda**



Database



Queues



File Buckets



CI / CD Tools

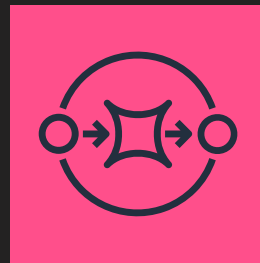
Any AWS service that **you don't manage**  
**at the OS level** is serverless



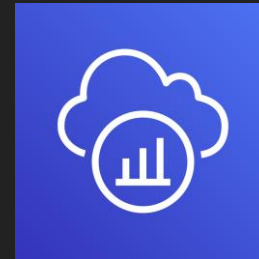
DynamoDB



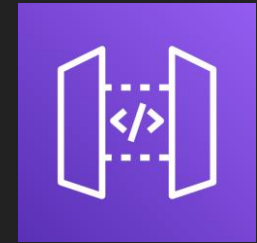
Lambda



SQS

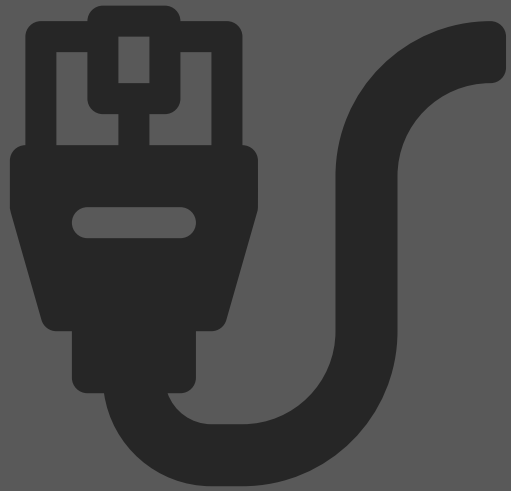


X Ray



API Gateway

The most popular AWS serverless  
services for building APIs



AWS Lambda have a  
good amount of  
**Integration Points**



Tons of people are **skeptical** about  
Lambda's integration capabilities.

Principal Engineer

## 3 Reasons AWS Lambda Is Not Ready for Prime Time

Does AWS Lambda keep its serverless marketing promise of continuous scaling?

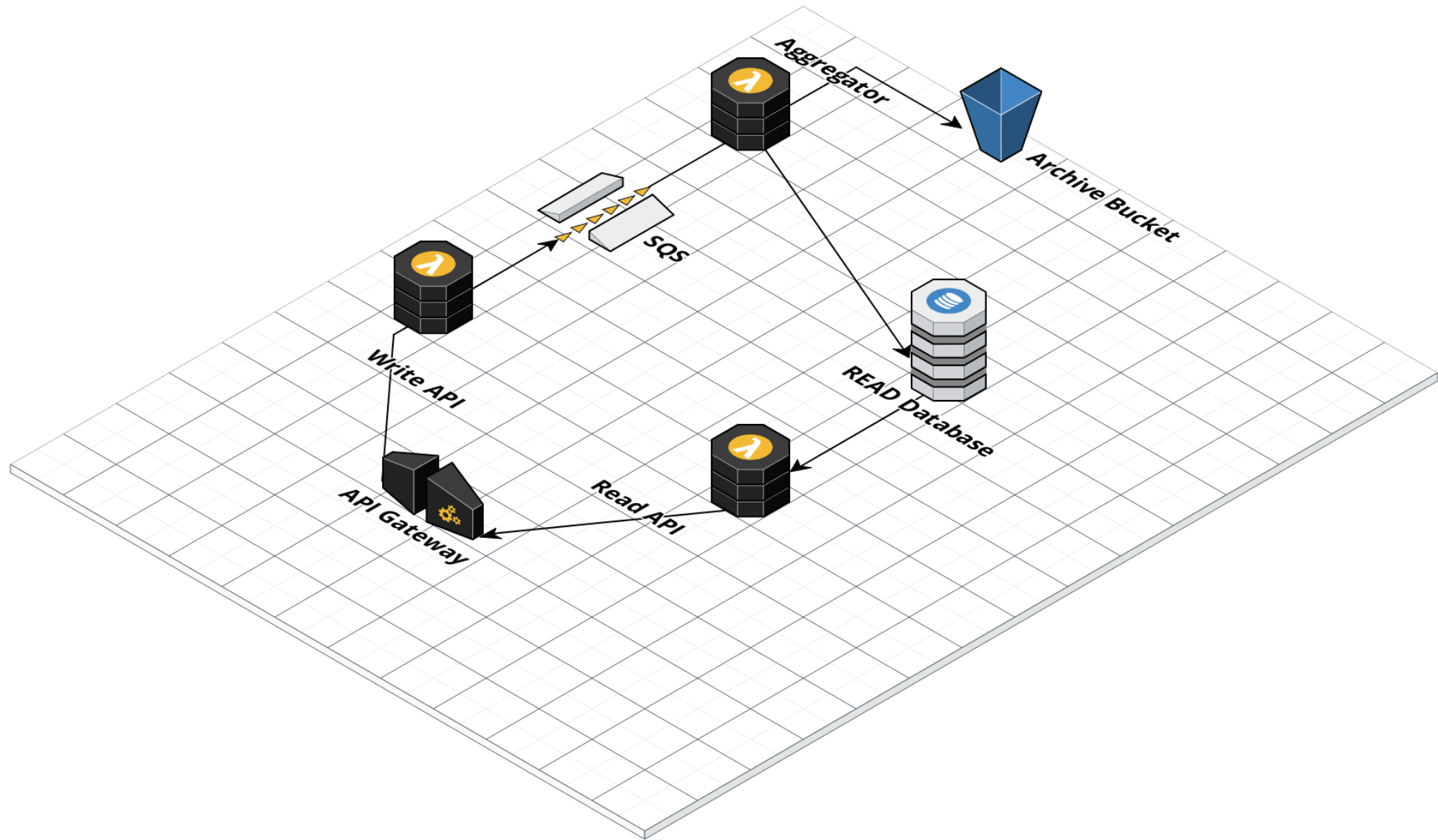
Back in 2016 people are bashing  
lambda big time



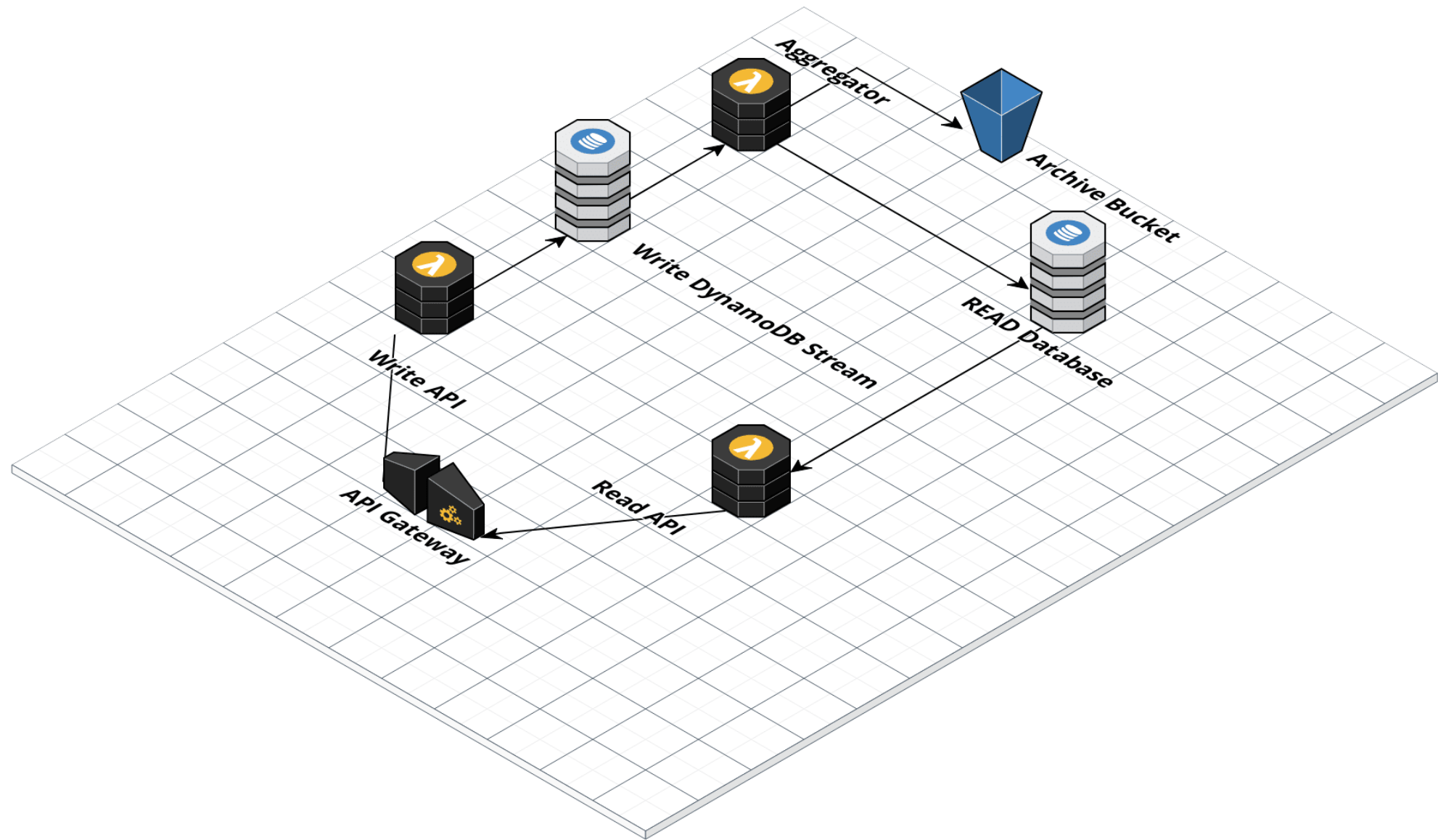
But hey, it has been almost

5 years since 2016.

AWS Introduced cool integrations since



# Basic CQRS using Lambda



# CQRS & DynamoDB Streaming



## Lambda Integrations

<https://docs.aws.amazon.com/lambda/latest/dg/lambda-services.html>



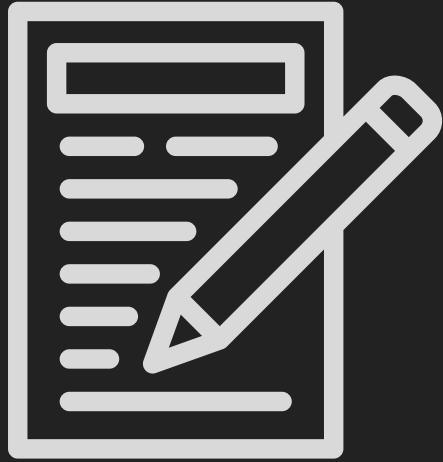
Run shared code using  
**Lambda Layers**



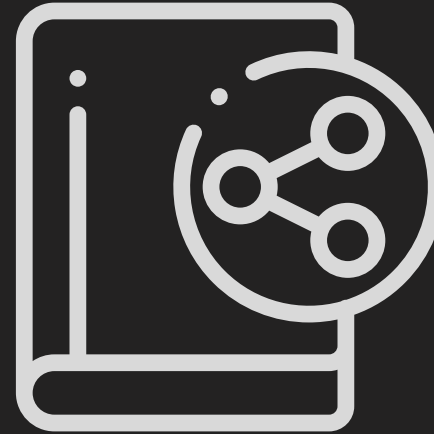
Lambda-based APIs require each  
endpoint to be packaged in isolation



Which often leads to  
**CODE DUPLICATION**



Shared Code



Dependencies

# Notorious Duplicates





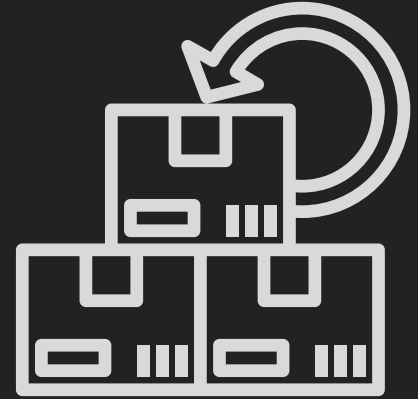
To solve duplication issue:  
**Use Lambda Layers**



Centralized  
Dependencies



Easier Code  
Updates

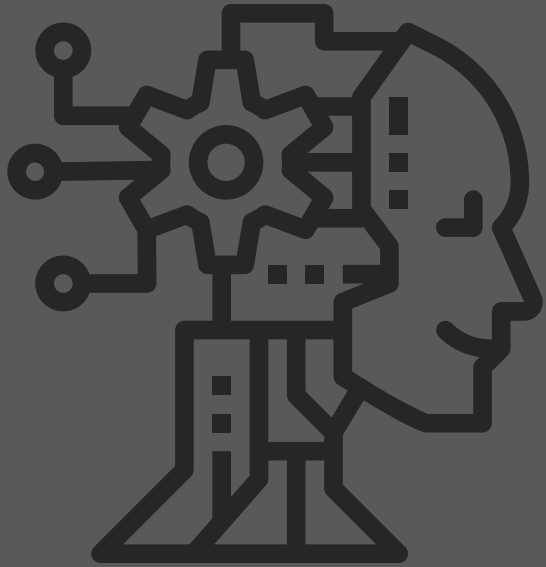


Smaller  
Deployments

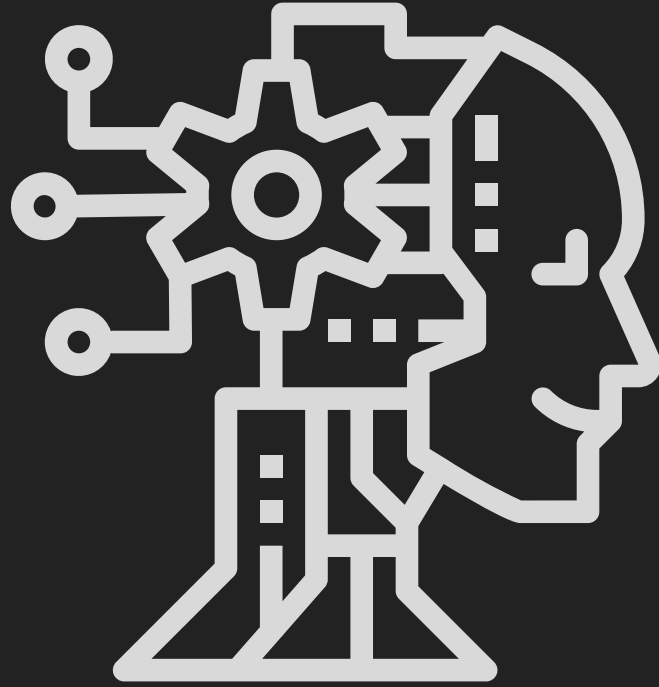
Benefits



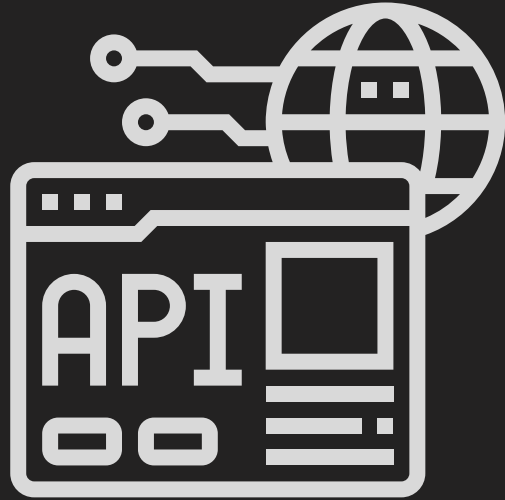
# Sample API Implementation using Lambda Layers



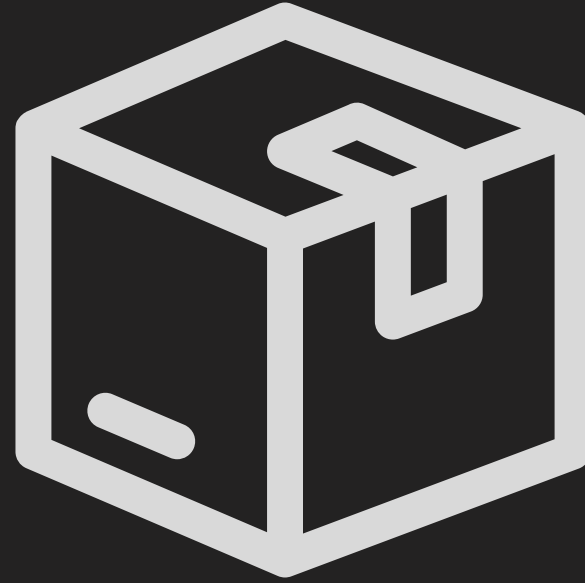
# Integrating vs Infusing Artificial Intelligence in your APIs



Digital Transformation have recently  
gained tons of traction and AI is riding  
this wave



Integrating with  
AI APIs



Infusing AI to  
your APIs

Lucky for us, we can either **infuse** or  
**integrate** AI Models to Lambda APIs



Integration with a base  
API powered by AI



amazon Rekognition



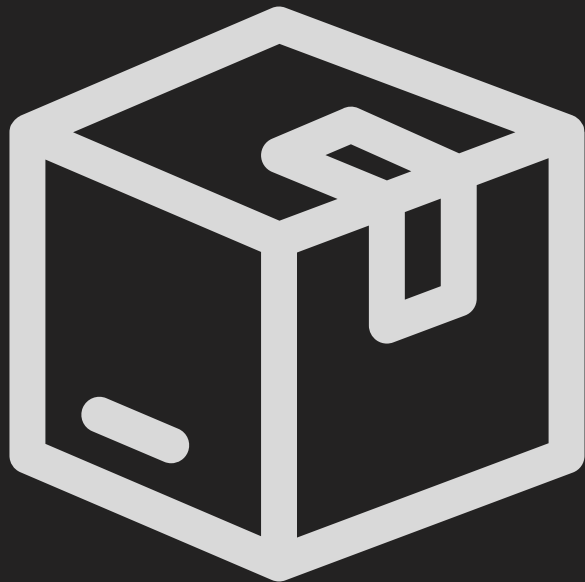
Cloud Vision API



Blog used for integrating Google Cloud Vision using NodeJS and Lambda by **Alex DeBrie**

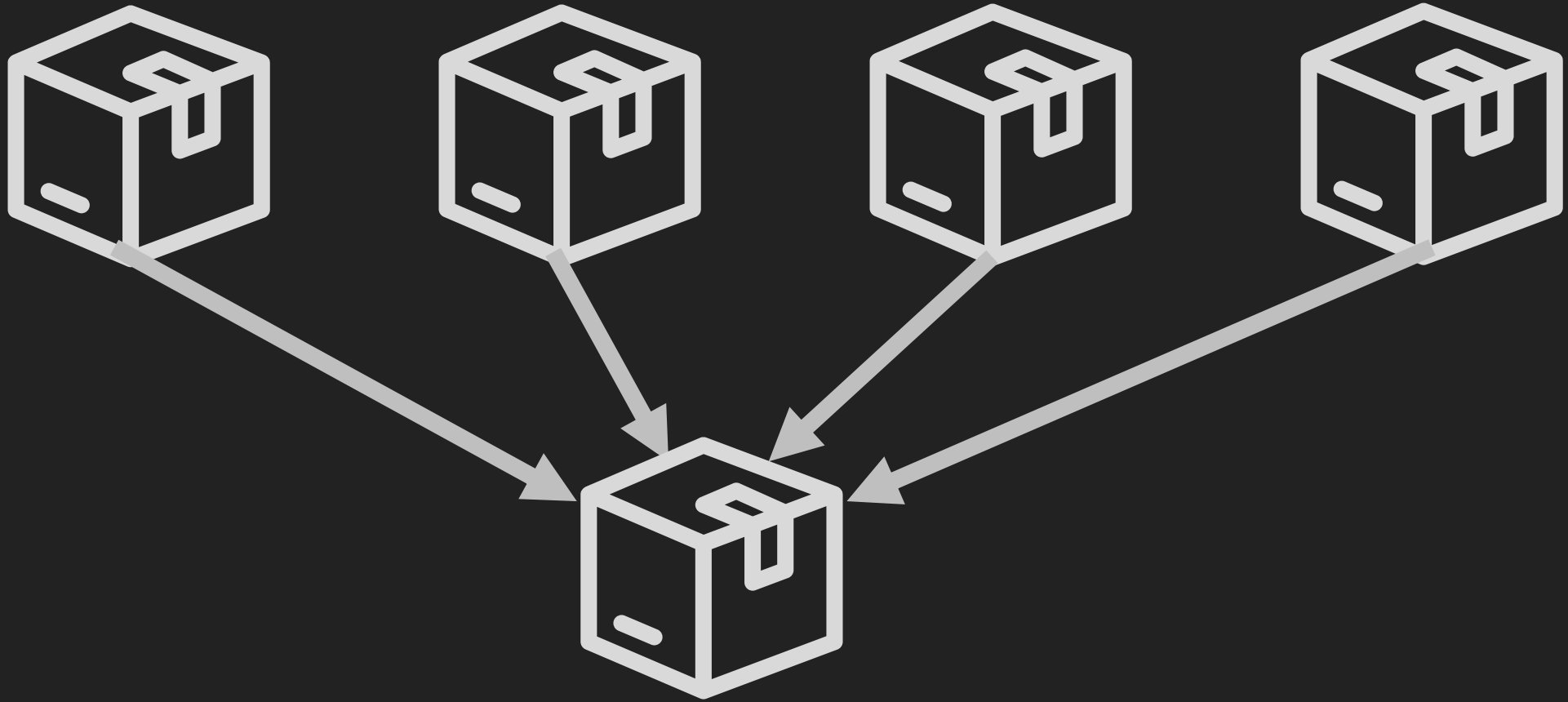
<https://www.serverless.com/blog/google-cloud-functions-application>





Infusing / Packaging  
Custom Built AI Models to  
your APIs



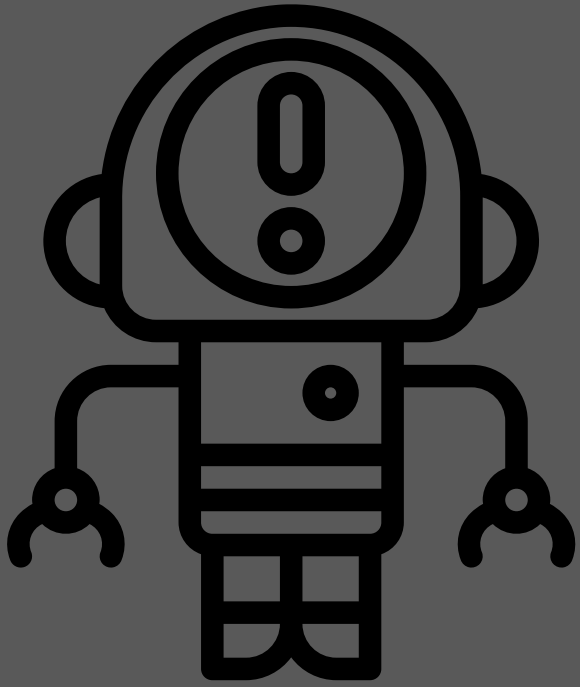


An unorthodox approach on how can you achieve high accuracy ML is by combining multiple models in an election system inside the API



## How to embed Scikit Learn Models inside Lambda Functions by Will High

<https://www.highonscience.com/blog/2017/09/29/ml-scoring-service-on-aws-lambda/>



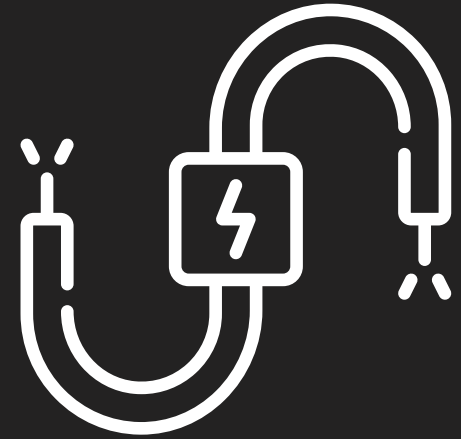
# Integrating RPA in your Serverless APIs



Manual Process  
Automation

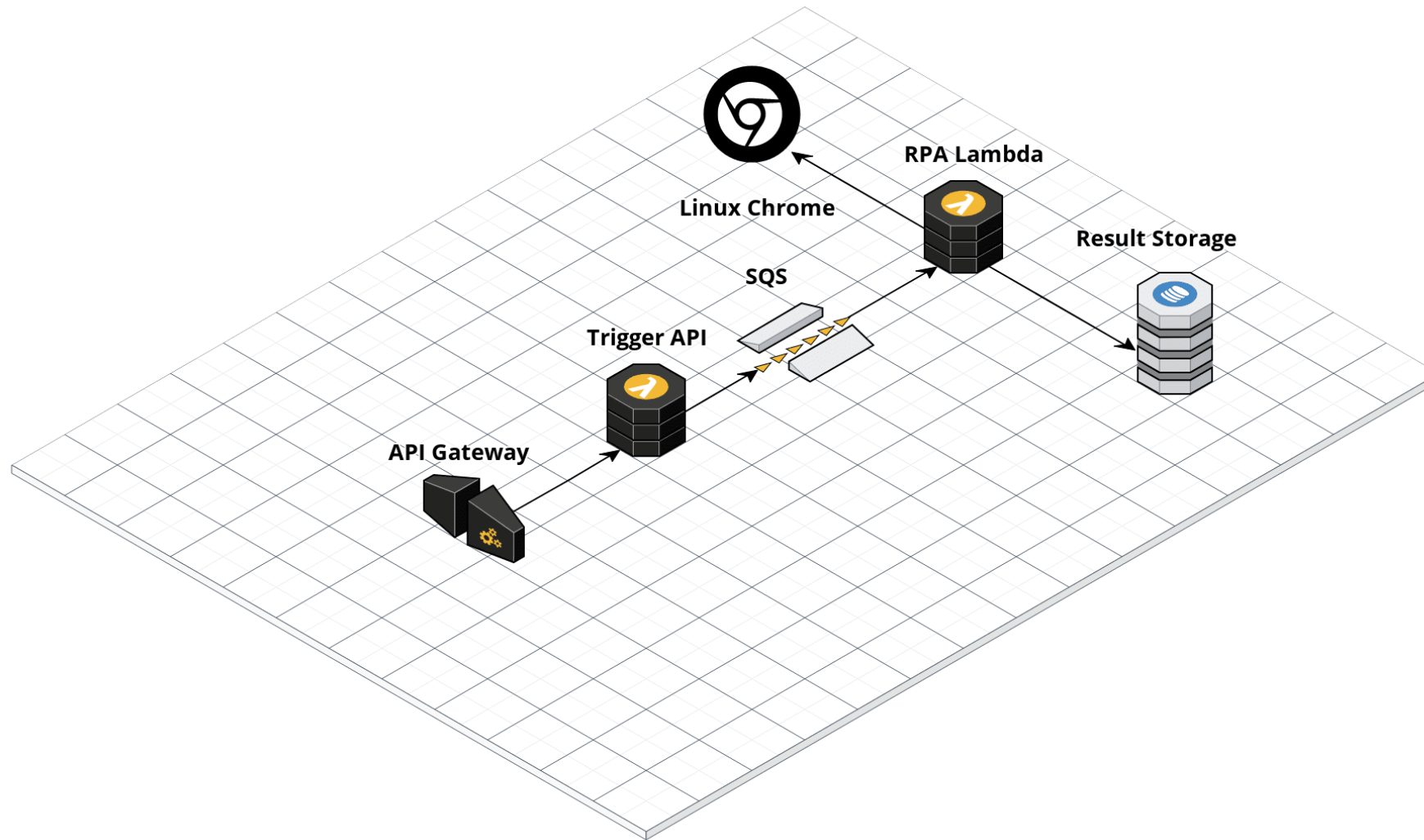


Data Scraping &  
Market Intelligence



Integration with  
Legacy Systems

RPA is a huge advantage in the Insurtech  
World.

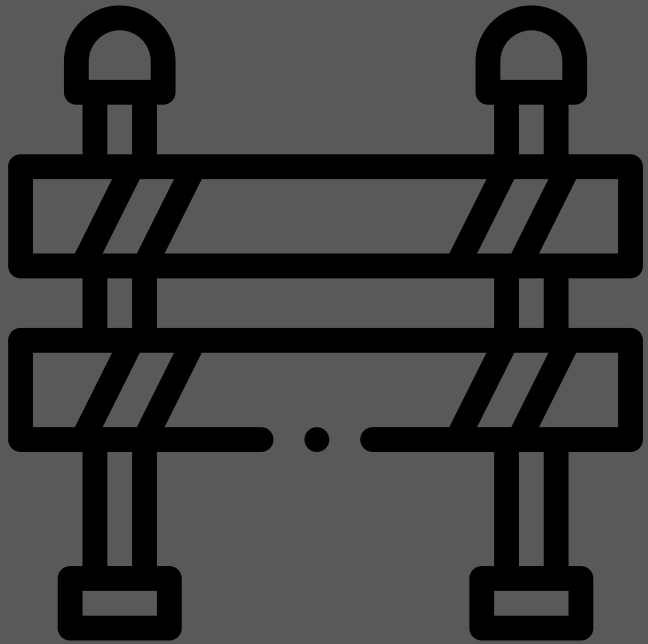


Typical Serverless RPA Integration



Asynchronous API triggering  
RPA worker using AWS  
Lambda

<https://bit.ly/3kW9sTM>



Serverless have  
**Limitations**





Its rewarding to be  
**mindful** about serverless  
**limitations** when  
proposing to your **boss**

3 GB

Memory Limitation

6 MB

Invocation Payload

15 Mins

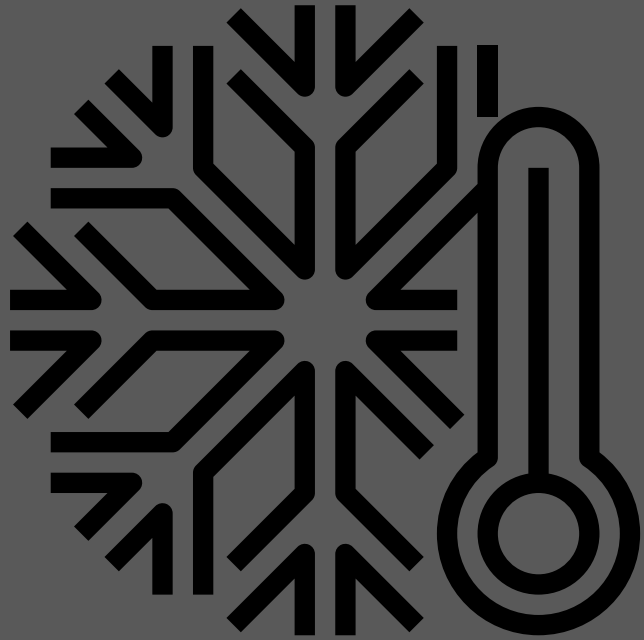
Max Execution Time

200

Max Resource per Cloud  
Formation Stack

30 Secs

API Gateway Max  
REST Timeout



# Wrestling with Coldstart



Poor Choice of Language



Heavy Dependencies

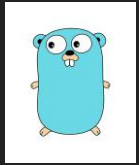


Too much code

**VPC**

VPC Integration

# Common Causes of Coldstart



Pick GO, JS and Python



Use Lambda Layers

VPC

Minimize VPC

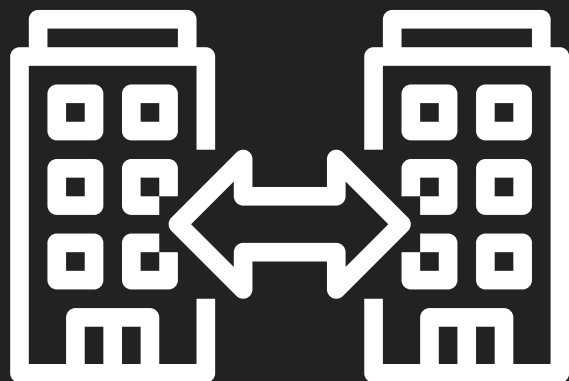


Keep Warm

# Coldstart Busters



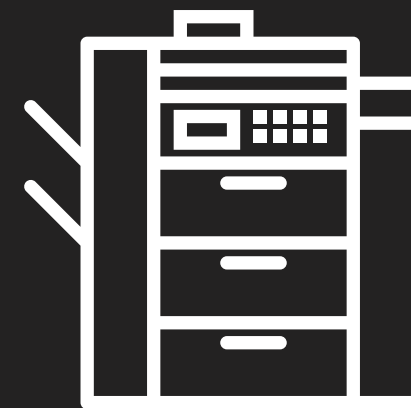
Serverless is  
NOT for everything



B2B Services



State Law vs Foreign  
Data Storage



On-Premise

It can be **rewarding** to utilize **Docker**  
and **K8s** on certain cases



# Questions & **Answers** Section



# Credits to the following:

- Joel Khoo (Ninja Captain)
- Michel Do (Chatbot Master)
- Nattavit Kamoltham (CI/CD & AI)
- Jon Scheele for the invite

# Searching Serverless Codebase is easier with Quod AI

The image is a screenshot of the Quod AI website, which is a search engine for codebases. The main heading on the page is "Search engine. For your codebase." Below this, there are three bullet points with checkmarks:

- ✓ A search engine working on **your own source code**
- ✓ Ask questions in **natural language**, get live answers from your source code
- ✓ AI-generated **documentation**

Below the bullet points is a form labeled "Your work email" and a purple button labeled "Request Demo".

On the right side of the page, there is a large purple banner with a "Request Demo" button. Below the banner, there is a search bar with the text "Search in vbudhram / fantasy-football-io". Below the search bar, there is a section titled "vbudhram / fantasy-football-io" with a "Source" button. The section contains the following text:

Fantasy Football News, Player and League API. This documentation was AI generated and is up-to-date

364 questions 56 tags Javascript 106 GitHub stars

Below this, there are tabs for "Overview", "Questions", and "Tags". The "Overview" tab is selected. The "Overview" section contains the following text:

Technical overview

This project is a web project with [front-end UI](#) and [back-end HTTP APIs](#). It uses [SQL Query](#), [ORM Query](#) and [Message Queue](#).

Featured tags

Onboard onto this repository using tags below or [view all](#).

article Socket Events HTTP API

At the bottom of the page, there is a purple button labeled "Request Demo" and a purple button with a white envelope icon.

How companies are using Quod AI



Thanks for listening to my propositions