



AWS API SERVICE FOR POLQA3

Relatore: Francesco Galatioto

Looking for...

..How to use POLQA as a service via API

Constraints:

- POLQA's license is hardware binded
- POLQA instance needs params and .wav file payload to perform
- Large amount of data/instances per minute (about 11GB/d and 2000/d API calls)
- Async design to manage possible delay queue

Success cases with No-Code tools:



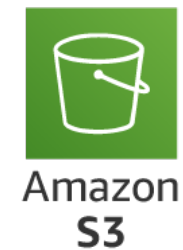


Our AWS solution

Based on AWS Lambda that is an event-driven serverless computing platform.

Scalable | Low-Code | Multi language (C#/.Net, JAVA, Node.JS , Python, Ruby)

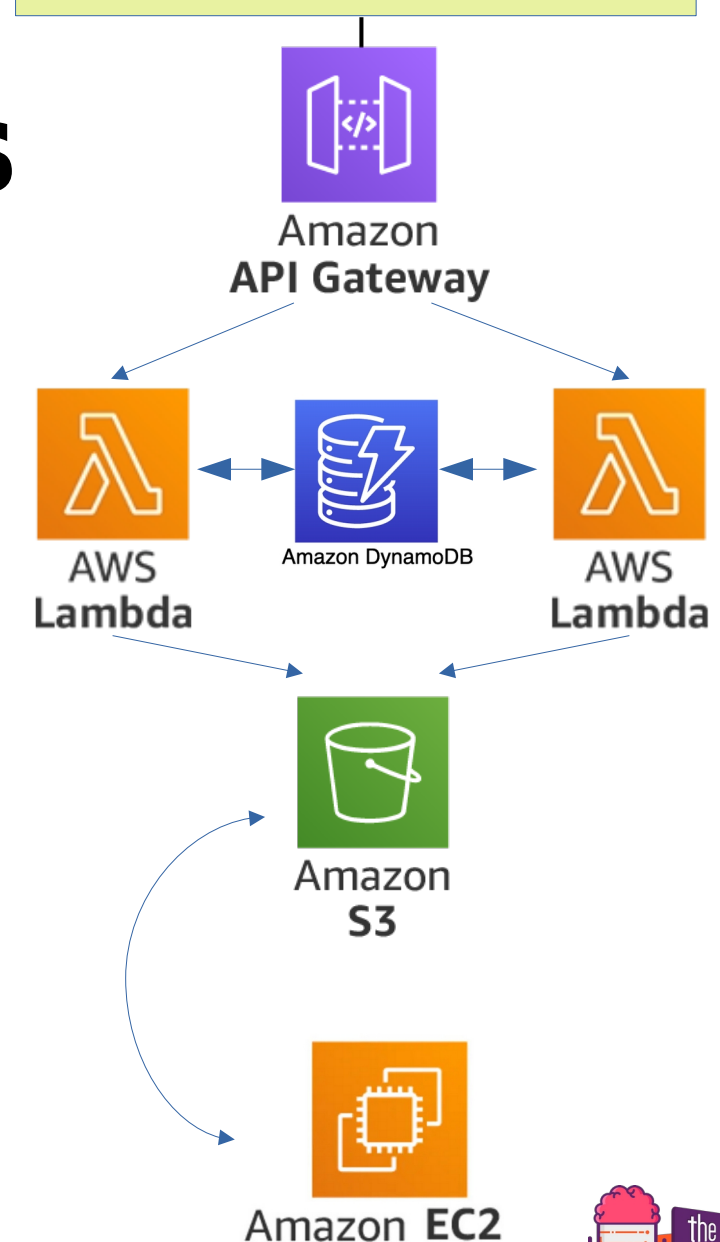
POLQA®



Integrated services

- Trigger/auth by APIGateway (REST API)
- Storage by S3 (exposed)
- Calculation by EC2 with POLQA onboard (segregated)
- Code by Lambdas (async)
- Database by DynamoDB (Lambda connection)

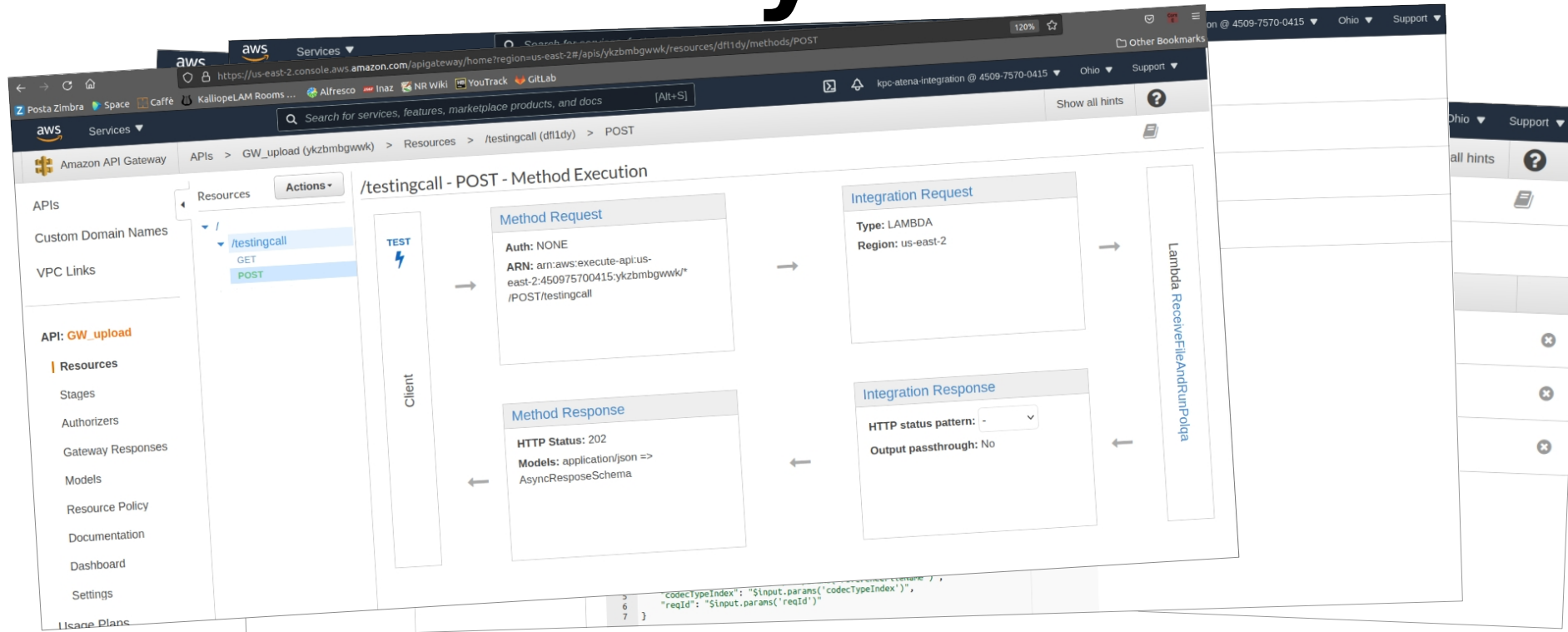
1.POST (with payload) / 2.GET



REST Gateway



Amazon
API Gateway



- 1 resource 2 methods (POST & GET that will call 2 lambdas)
- Only audio/wave binary media type accepted (POST) + Auth key
- Integration Request mapped (to manage blob payload and params)
- Response mapped for custom errors by **Lambda Error Regex**



EC2 & S3



EC2

- T3.xlarge
 - 2 cores (4vCPU, 16GiB, network performance up to 5Gigabit)
- CentOS8 with POLQA onboard
- Dynamic public IPv4 (issue solved by Lambda)

S3

- All public access blocked (security)





Lambdas

- Python 3.6
 - Paramiko to run SSH EC2 client
(We used as “layer” to manage function's dependencies independently and keep deployment package small)
 - Boto3 to manage S3 bucket & DynamoDB table
 - AWS CLI (to move and delete S3 files)
- 2 Lambdas each one for POST & GET call
 - I) **Receive** data payload into S3; **Name** by Context `aws_request_id` as UID; **Move** it from S3 to EC2; **Run** Polqa; **Write** reqId into db; **Give back** 202 StatusCode with your own reqId (or errors)
 - II) **Verify** if reqId exists by db, **Find** POLQA result file on EC (if ready); **Move** it into S3; **Parse** it to give you back the JSON 200 answer; **Delete** db item. Otherwise, if file **is not ready yet**, give you back a 202 with a “retry after” x seconds message (or errors)



DynamoDB

No-SQL database that permits GET call to know if a reqId already exists

- Managed by Lambda Boto3
- Fast and easy
- On-Demand/Provisioned billing plan

- **TTL *TimeToLive* (60minutes)**
- **Stream events (delete)**



CLEANING FILES ISSUE

(by dedicated Lambda that remove S3&EC2 wav/txt files when triggered)

- 1) Not need to implement all files remove exceptions
- 2)Easier maintenance

A little bit of security

- Used S3 bucket as data swap
- Gateway accepts as binary audio/wave files only
- APIKey (to limit access, throttling and quota)
- ReqId validation by DynamoDB



TO DO

- Environment variables are always encrypted at REST, and can be encrypted client-side as well
- Firewall
- Client certificate
- Improve management on IAM roles



Features



- Implement 408 **request timeout/too many requests** to stop possible GET call loop (at the moment limited only by DynamoDB TTL option)
- Better fit hardware request for Lambda and EC2 (up or down)
- Increase security as before



Seeing the code..





Biblio&Links

-How to upload files to Amazon s3 using Nodejs, Lambda and API Gateway

<https://olotintemitope.medium.com/how-to-upload-files-to-amazon-s3-using-nodejs-lambda-and-api-gateway-bae665127907>

-Boto3 documentation

<https://boto3.amazonaws.com/v1/documentation/api/latest/index.html>

-Working with Amazon S3 buckets with Boto3

<https://towardsdatascience.com/working-with-amazon-s3-buckets-with-boto3-785252ea22e0>

-Amazon S3 objects overview

<https://docs.aws.amazon.com/AmazonS3/latest/userguide/UsingObjects.html>

-Paramiko package documentation:

<http://docs.paramiko.org>

-Using an IAM role to grant permissions to applications running on Amazon EC2 instances

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_use_switch-role-ec2.html

-SSH to EC2 instances via AWS Lambda

<https://www.transposit.com/blog/2019.12.18-using-lambda-as-an-ssh-proxy/>

-Paramiko : Error reading SSH protocol banner

<https://stackoverflow.com/questions/25609153/paramiko-error-reading-ssh-protocol-banner>



<https://gitlab.netresults.dev:10443/galatioto/polqa3>