

Academic Year 2019/2020

Semester I

CST 333 - Distributed & Grid Computing

Assignment 2 - Distributed Cloud Computing

No.	Name	Matric	Email
1	Muhammad Iqrar Amin	140659	iqraramin@student.usm.my
2	Tan Lee Sing	141097	tanleesing1998@student.usm.my

Due Date: 26th January

Table of Contents

1	Pro	ject Overview	. 3		
	1.1	Services Used	3		
	1.2	System Architecture	. 4		
2	Pro	ject Environment Configuration and Settings	. 5		
	2.1	Amazon S3 Bucket	. 5		
	2.2	Amazon Route 53	. 8		
	2.3	Amazon Certificate Manager	10		
	2.4	Amazon CloudFront	11		
	2.5	Lambda Function	13		
	2.6	Amazon EventBridge	14		
3	Les	sons Learned	16		
4	Tas	k Distribution	17		
5	Source Code				
6	Proof of Concept (Screenshots)				
R	References 23				

1 Project Overview

We have hosted a static HTML website in an S3 bucket. The website uses CloudFront for its Content Delivery Network (CDN), which caches its contents to optimize performance. Whenever we update an object in the bucket, we will need to invalidate (purge) the object's cache so that visitors to our website will see the updated content.

1.1 Services Used

Table 1.1 shows all the services that our project is using to create the static application that will automatically invalidate the CDN cache.

Table 1.1: AWS Services Description

No.	Logo	Name	Description	
1	53	Route 53	Route 53 is used to automatically handle the requests that go in from Internet to the internal resources of AWS in our case, CloudFront.	
2		CloudFront	Amazon CloudFront is a worldwide content delivery network (CDN) service that is used to delivers data securely and with low latency.	
3		S3	An S3 bucket is utilized to store all the code for our static website which is essentially a resume/portfolio website.	
4	益	Certification Manager	By using Certificate Manager, we will create an SSL certification for our static website, so the traffic is server using HTTPS protocol.	
5	λ	Lambda	Lambda functions provide the necessary backend code that will be used to automatically invalidate the cache by sending a request to CloudFront.	
6	र्दुटी	EventBridge	EventBridge is used to create an event which will trigger the lambda function whenever a new file is added or deleted S3 bucket.	

1.2 System Architecture

Figure 1.1 shows the overall architecture our project. There are two main use cases in this project. The first one is when the user of the static website wants to access it. In this case, the request from users will travel through the Route 53 first which will direct the request to the CloudFront where all the resource for our static website is cached for all unchanged requests. In case the contents are not cached, they are directly taken from S3 bucket and served to the user. Moreover, the connection from client to the CloudFront and ultimately to the S3 bucket is encrypted using SSL Certificate from AWS Certification Manager (ACM). The second use case is when a developer/programmer modifies the static website in the S3 bucket, it will trigger an event defined in the EventBridge which is essentially a Lambda function that will automatically invalidates the cache in the CloudFront CDN so that the latest content in the S3 bucket is served to the users of static website.

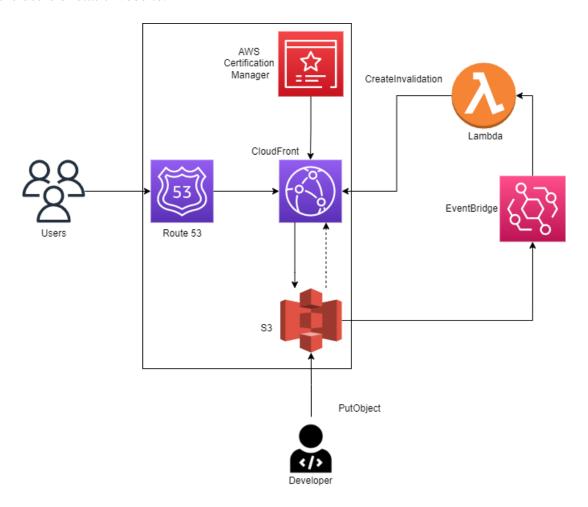
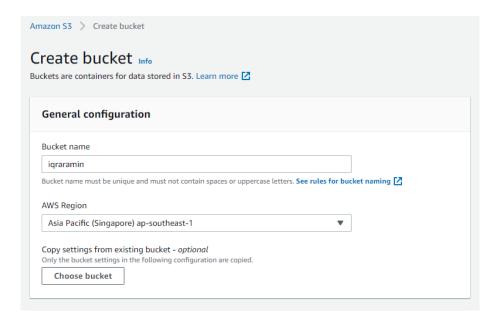


Figure 1.1: System Architecture Diagram

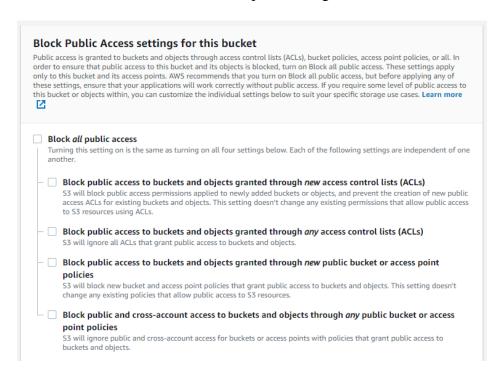
2 Project Environment Configuration and Settings

2.1 Amazon S3 Bucket

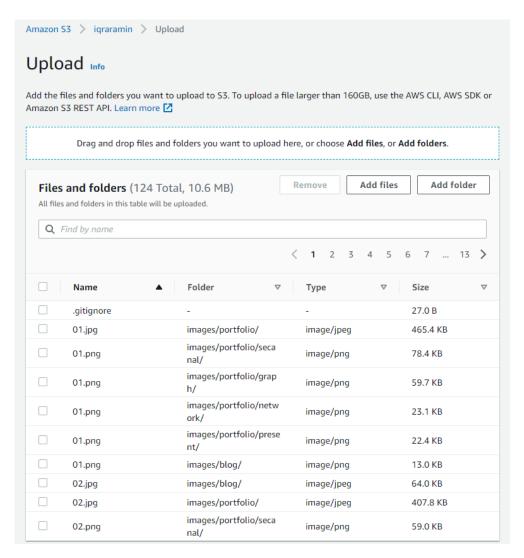
For our project, we first started with configuring the Amazon S3 bucket, we used the general configuration shown in the following image:



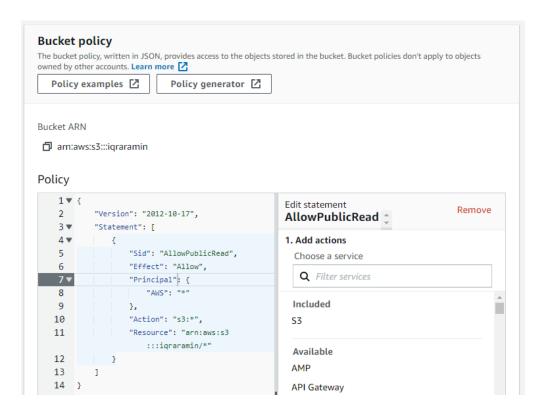
then we choose to allow all the public access on this bucket since this will be distributed globally to everyone on the internet, with the rest of the options being default:



Once we created the bucket, we now upload all our code for the static website in the S3 bucket, the following is the configuration we choose to upload our content:

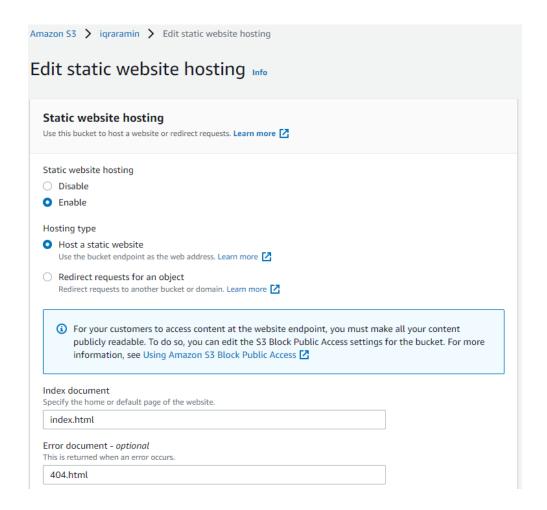


Now that we have uploaded our content to the S3 bucket, we need to give appropriate permissions in the Bucket Policy, the following is the config for newly added policy:



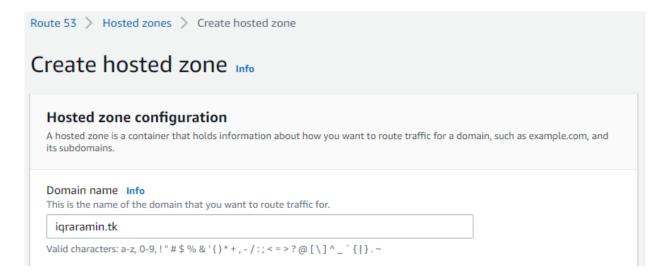
The policy essentially means that all the public traffic trying to access this resource will be granted access to view the content of the bucket, here is the JSON used to configure it:

Lastly, we will enable the static website hosting in the Properties of the S3 bucket, we used *index.html* as the home page and *404.html* as error page, the following is the configuration used:



2.2 Amazon Route 53

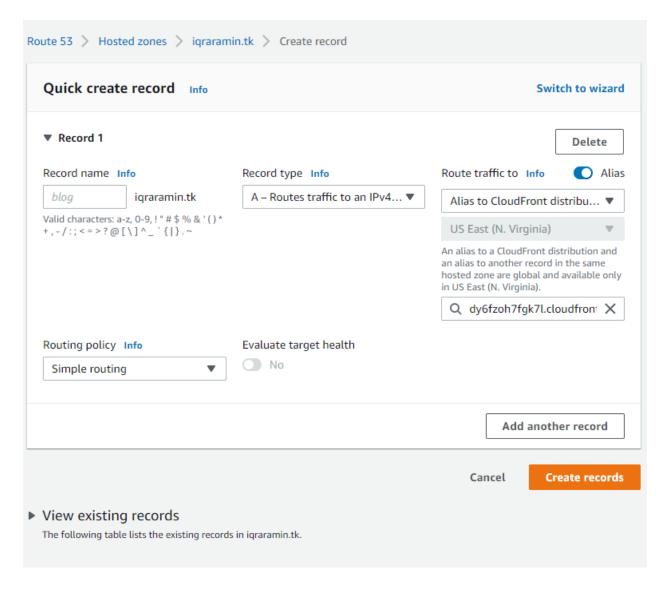
Now that the initial setup for S3 bucket is complete, we will create a hosted zone for our domain name in the Route 53, the following is the hosted zone configuration:



we will then use the nameservers created in this hosted zone in the domain name manager so that both domain name and our resources are connected, the following are the nameservers:

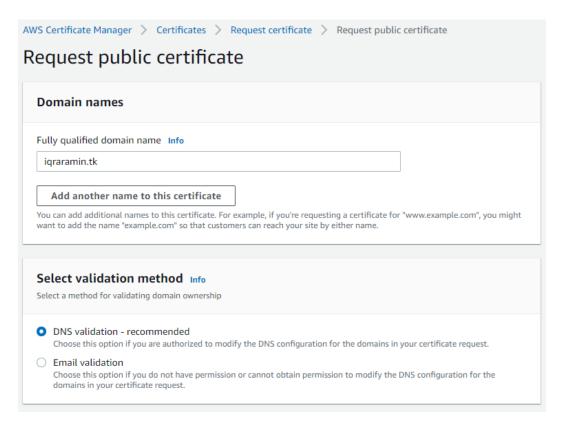
	iqraramin.tk	NS	Simple	-	ns-1292.awsdns-33.org. ns-1976.awsdns-55.co.uk. ns-362.awsdns-45.com. ns-875.awsdns-45.net.
--	--------------	----	--------	---	------------------------------------------------------------------------------------------------------

Once we create a CloudFront distribution for our static website, we will need to add another record in hosted zone so that the traffic from the distribution will be directed to the domain, the following is the config used to create the record:

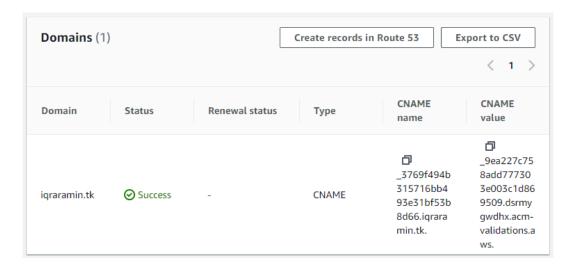


2.3 Amazon Certificate Manager

After setting up the appropriate routes for the static website, we will secure the all the connection to the site using SSL certificate provided by Amazon Service. We will create a public certificate as the website is publicly hosted, the following is the config used:

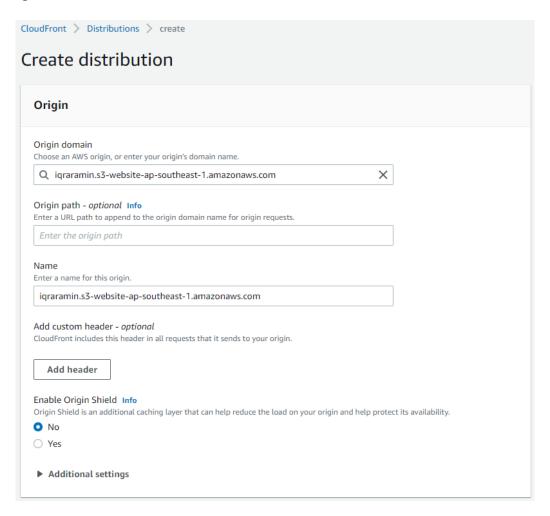


Once the certificate is created, we need to create a record in the Route 53 for it to validate the content of our website, the following settings were used to create a CNAME record in Route 53:

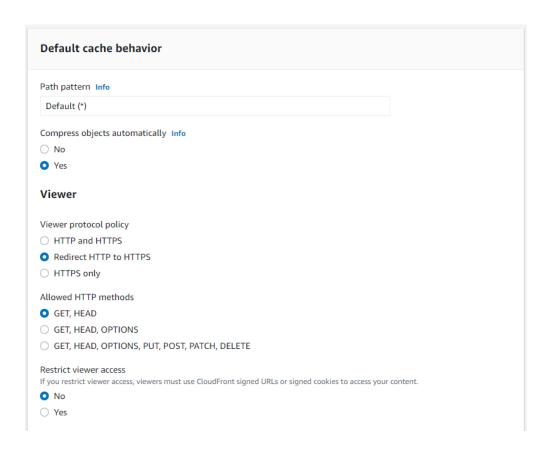


2.4 Amazon CloudFront

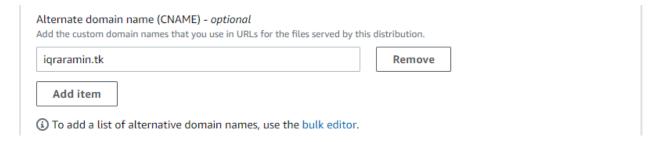
After creating an SSL certificate using Certificate Manager, we will create a CloudFront distribution. For "Origin", we choose the domain same as the URL for our static website on S3 bucket, and the Name was auto generated by AWS. The following are the configuration for "Origin" part of the dist.:



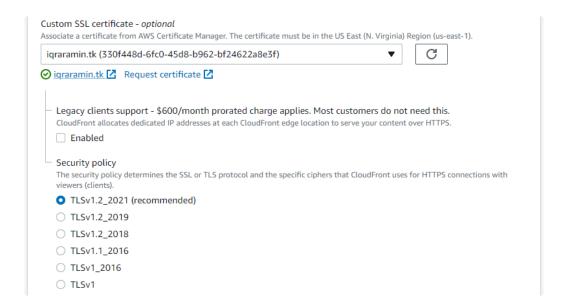
Moving on the "Default cache behavior" part of the configuration, we only require changing on part and the remaining part will be left as the default. We only change the *Viewer protocol policy* from "HTTP and HTTPS" to "Redirect HTTP to HTTPS" since we only want a to establish secure connection, so we redirect all the unsecure connection to the secure channel. The following are the configuration for "Default cache behavior":



In the "Settings" section of the distribution configuration, we only need to change two things first is the *Alternative domain name (CNAME)* which is used to add the custom domain names that we use in URLs for the files served by this distribution. The following is the config used:

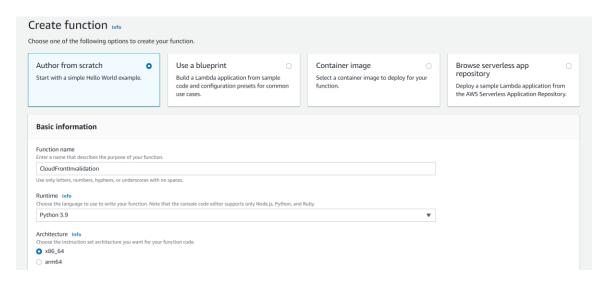


Secondly, we need to add a SSL certificate that we created previously using Amazon Certificate Manager, this will associate a certificate from AWS Certificate Manager. The certificate must be in the US East (N. Virginia) Region (us-east-1) hence we create the previous certificate in the same Region. The following are the configs for this:



2.5 Lambda Function

After setting up Amazon S3 bucket and Amazon CloudFront, we will create a Lambda function for automatic invalidation whenever a new object is created or deleted. In creating function, we choose author from scratch and the programming language that we choose is Python.



Then, in the code source, we will type in our code and deploy the code.

```
Code source Info

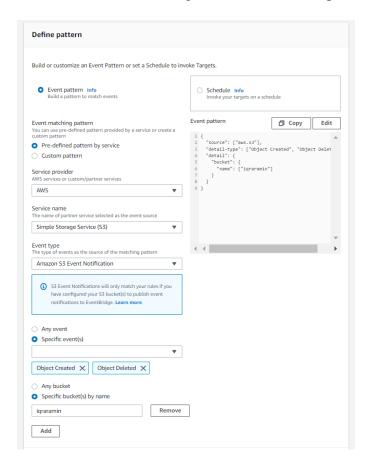
A File Edit Find View Go Tools Window Test Tools Changes not doployed

O Go to Anything (CREP)

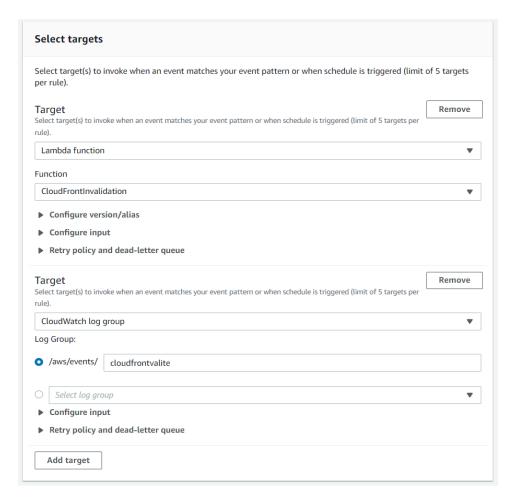
I from _future__import print_function
| Sambda_function.py | Immbda_function |
```

2.6 Amazon EventBridge

After the function is created, we create a new rule in Amazon EventBridge to trigger the function in Lambda. For define pattern, we choose event pattern and not schedule as we want to trigger function in Lambda every time a new object is created or deleted. Then, we will choose the predefined pattern by service and below is the configuration for the event pattern.



Moving on to select target, the first target, we choose Lambda function and the second target, we choose CloudWatch log group, so that we can track the log. Below is the configuration for the select targets. After that, we can create the rule.



3 Lessons Learned

The following are the lessons we learned from creating a static website using Amazon Web Services, we learned how to:

- 1. Host static website on Amazon Bucket.
- 2. Route the traffic of CloudFront CDN to the domain name using Route 53.
- 3. Create SSL certificate using Amazon Certificate Manager.
- 4. Configure SSL certificate in CloudFront to use HTTPS.
- 5. Configure different services in Amazon Web Service
- 6. Configure a domain name with the CloudFront resources.
- 7. Use different methods to trigger Lambda function from S3.
- 8. Create manual and automatic invalidation in CloudFront.
- 9. Configure Amazon EventBridge to receive notification from S3 and target to the destination that we want.

4 Task Distribution

The following are the task distribution for this project between Muhammad Iqrar Amin and Tan Lee Sing:

Muhammad Iqrar Amin

- Research possible ways to host static website on the Amazon Web Services
- Create and configure S3 Bucket to host static website
- Create and configure Route 53 to route traffic from CDN to domain
- Create and configure SSL Certificate using Certificate Manager
- Create and configure CloudFront to act as a CDN
- Write Report section 1, 2 (2.1, 2.2, 2.3, 2.4), 3, 4
- Prepare Slides

Tan Lee Sing

- Research possible ways to automate cache invalidation in CloudFront from S3 Bucket
- Create and configure Lambda function to invalidate cache in CloudFront
- Create and configure event to trigger previously create Lambda function
- Write Report section 2 (2.5, 2.6), 3, 5, 6

5 Source Code

```
from __future__ import print_function
import boto3 as b3
import time
def lambda_handler(event, context):
    path = []
    #CloudFront will invalidate "/" if "index.html" has changed
    file = event["detail"]["object"]["key"]
    if file == "index.html":
        path.append("/")
    else:
        path.append("/" + file)
    # Get bucket name for S3
    bucket_name = event["detail"]["bucket"]["name"]
    client = b3.client('s3')
    # Get CloudFront distribution ID from S3 tag
    tags = client.get bucket tagging(Bucket=bucket name)
    for tag in tags['TagSet']:
        if tag["Key"] == "distributionId":
            distribution_id = tag["Value"]
            break
    # Perform invalidation
    client = b3.client('cloudfront')
    invalidation = client.create_invalidation(DistributionId=distribution_id,
        InvalidationBatch={
            'Paths': {
                'Quantity': 1,
                'Items': path
        'CallerReference': str(time.time())
```

The code above is the code to invalidate CloudFront whenever a new object is created or deleted. At line 9, the code is to get the file name of the object. Then, we will check if the file name is "index.html" or another file. If the file is "index.html", we will append "/" to the path.

At line 15, we will get the bucket name for the use to get distribution ID for our distribution in CloudFront. Then, we will invalidate CloudFront using the distribution ID that we get [1].

6 Proof of Concept (Screenshots)

The pictures above are examples of changes that we make before and after.

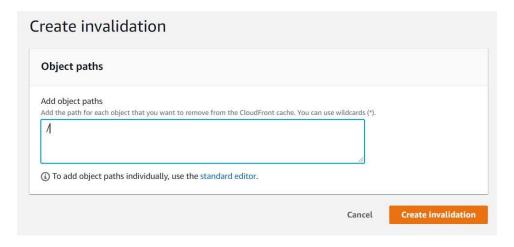


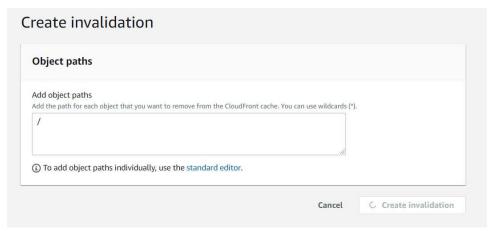
Figure above show the webpage before change.



Figure above show the result that we expect.

Before automatic invalidation, when we do changes like update the "index.html", the website will not update, and we need to wait 24 hours for the website to update. If we want to update the website, we need to do manual invalidation in CloudFront. Below, we show the step to do manual invalidation in CloudFront.

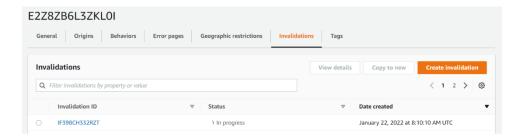






We can see from the step above; we need to put the object path where we make change and create the invalidation.

After we deploy the Lambda function for automatic invalidation, the Invalidation process will happen automatically.



It will show that the invalidation is in progress in the invalidation tab in CloudFront.

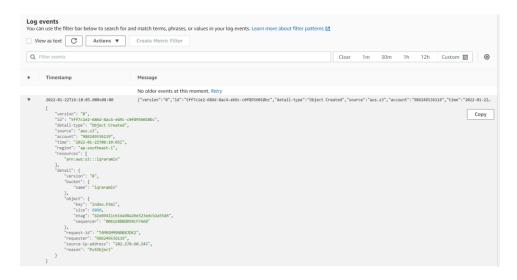
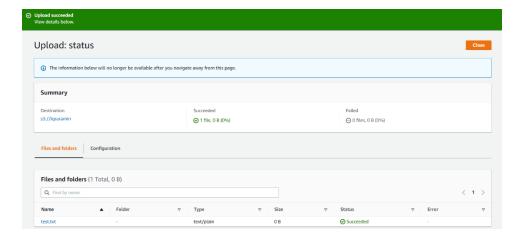
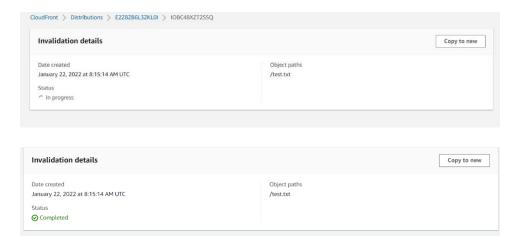


Figure above show the log event when the object is created or deleted in CloudWatch.

Next, we will show the example when we upload another file. The file that we will upload is a text document named "test.txt". Below is the example of the result that show in AWS.



The pictures above show the result when we upload "test.txt" into Amazon S3.



The pictures above show that the cache invalidation automatically got triggered and then changed the cache in the CloudFront.

References

[1] LewisCraik. [Online]. Available: https://github.com/LewisCraik/invalidate-cloudfront-on-s3-update.