AWS CloudFront

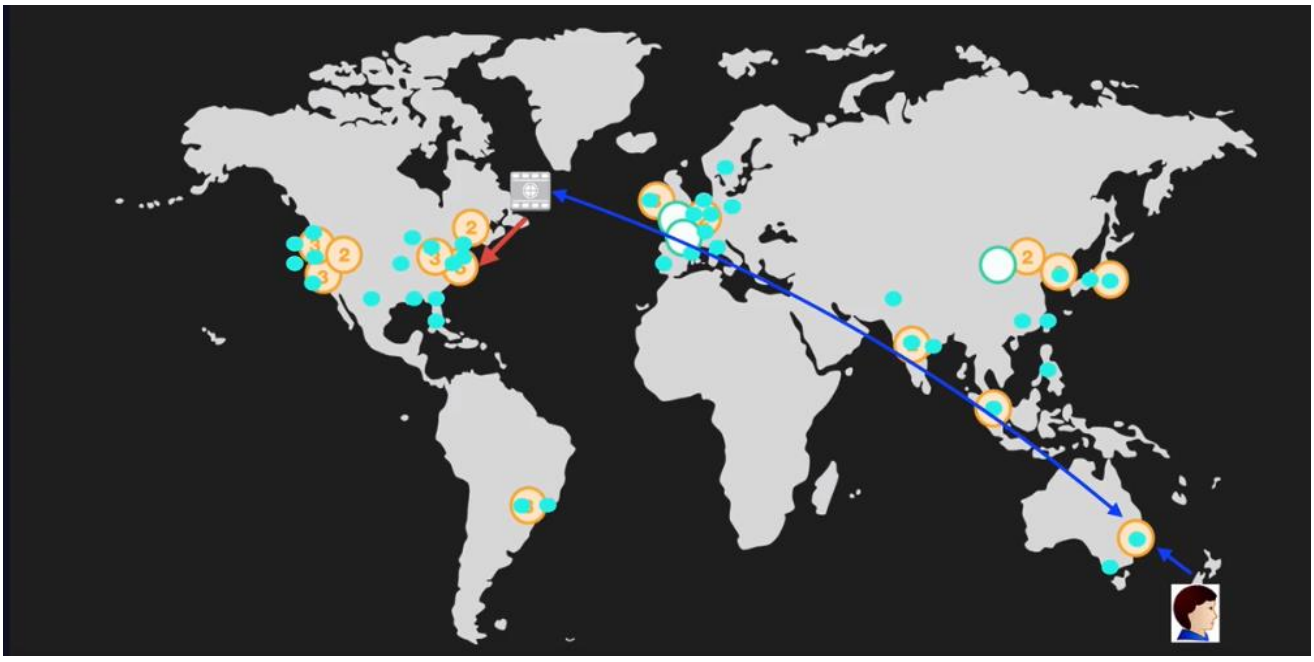
Amazon CloudFront is a web service that speeds up distribution of your static and dynamic web content, for example, .html, .css, .php, image, and media files, to end users. CloudFront delivers your content through a worldwide network of edge locations. When an end user requests content that you're serving with CloudFront, the user is routed to the edge location that provides the lowest latency, so content is delivered with the best possible performance. If the content is already in that edge location, CloudFront delivers it immediately. If the content is not currently in that edge location, CloudFront retrieves it from an Amazon S3 bucket or an HTTP server (for example, a web server) that you have identified as the source for the definitive version of your content.

AWS CloudFront can serve two types of content:

1. **Web content:** Web files like .html .css .php and images
2. **RTMP content:** media files like videos

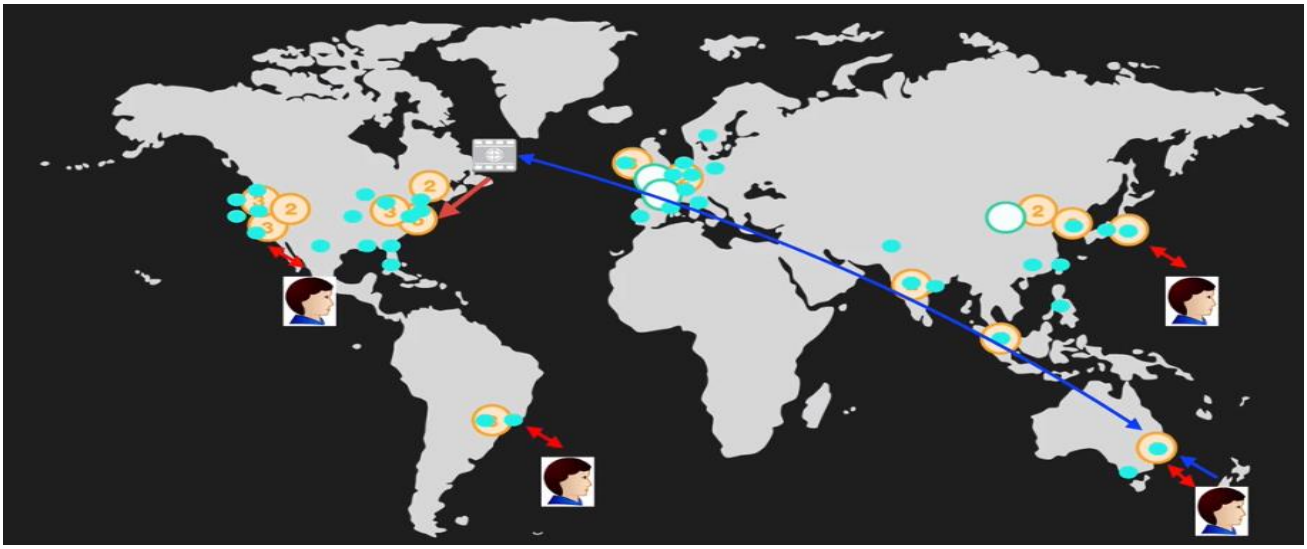
How Amazon CloudFront works?

As you can see an user in New Zealand requests for web content which is hosted on data centers in Northern Virginia. Ideally, the user will experience some latency as the network goes through a zig-zag route to fetch and deliver the content back to the user.



Prepared by Sriganesh Pera

Amazon CloudFront reduces the latency with the help of **edge locations** spread all around the world where the network from the user routes to its nearest edge location to get the content that is cached and stored in there.



CloudFront Concepts:

1. **Origin:** The location where original data is stored. Either an S3 Bucket,an EC2 instance, an Elastic Load Balancer, a Route53 Endpoint or an external system.

2. **Edge Location:** The location where your content is cached. There are 2 types of edge locations:

   a) **Global edge location**: Global edge locations is located close to user and this location stores the content as cache and is delivered to the user. **Data is not stored forever, it has a set time after which the data expires**.
   Popular content is cached in Global edge location.There are 68 global edge locations around the world.

   b) **Regional edge location:** It actually sit between global edge location and regional edge location. When the user requests for the content through

CloudFront link the request first checks the global edge location whether content is cached or not if content is not available in global edge location then it turns to regional edge location and if regional edge location has the content cached it will return the content to global edge location for caching and in turn the global edge location delivers the content to the user.If regional edge location didn't cached the content then the request is routed to the origin of the content i.e datacenters of AWS.

Not so popular content/infrequently accessed content is cached in regional edge location.

3. **Distribution:** The name given to CDN which consists of a collection of Edge locations

CloudFront features:

1. Objects are cached according to TTL(time to live) and then removed when TTL is expired.

2. Objects can be removed from the cache(invalidated) before the TTL but you will incur charge.

3. CloudFront supports both static(i.e images) and dynamic (php,ruby,and python scripts) content.

Lab Objective: Create a web content distribution using S3 as an origin and understand Caching behavior

Step 1: Creating Origin for CloudFront in S3.

Create a bucket in S3, save and upload .html file and make that bucket as well as the object publicly accessible.

<html>

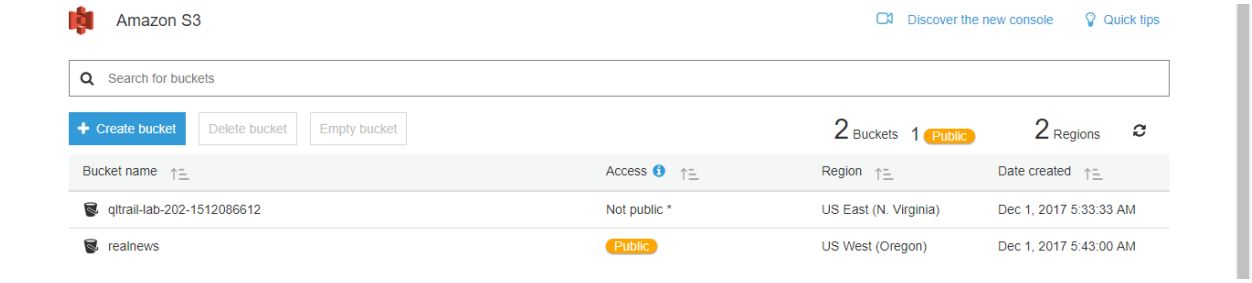
<head>Today's News</head>

<body>

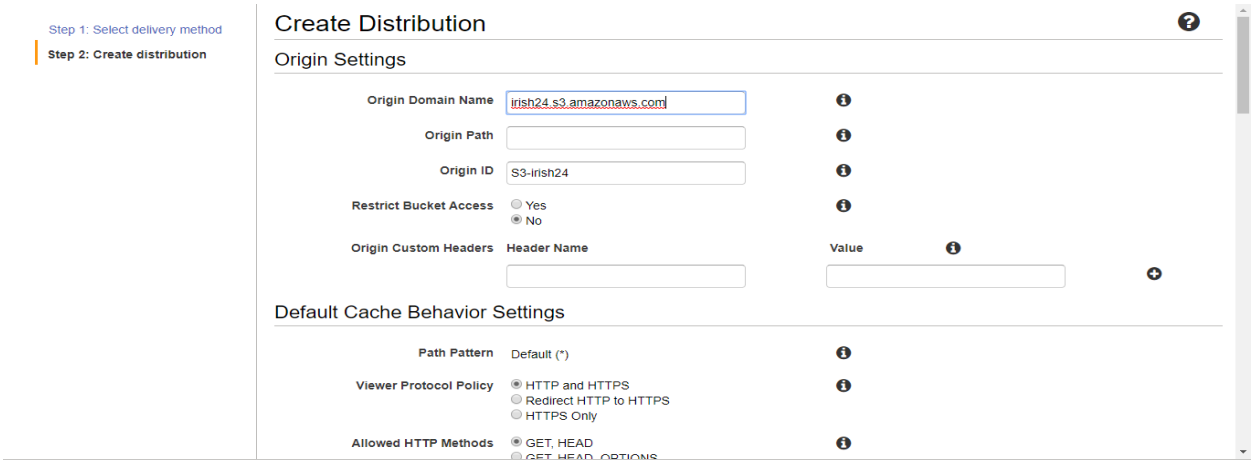
Pirates of Caribbean new poster

</body>

</html>



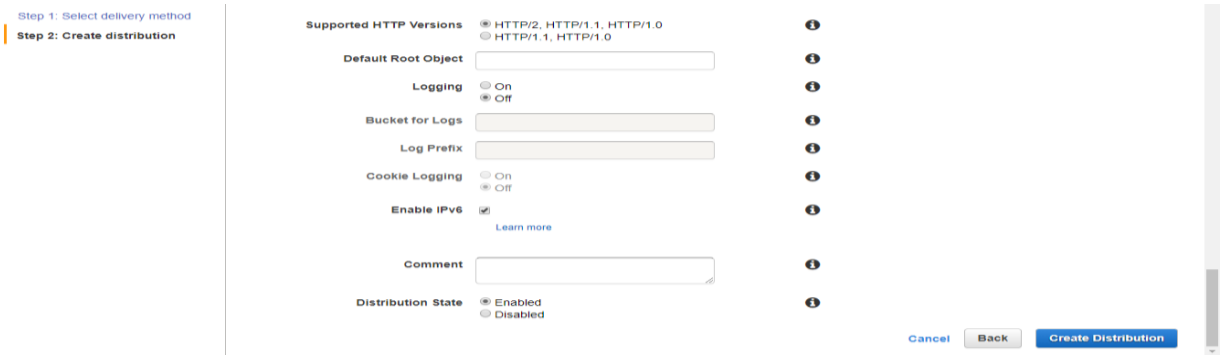
Step 2: Create Web distribution in CloudFront



Prepared by Sriganesh Pera

Now go to CloudFront and click on web distribution

And in the Original Domain name select your S3 link and accept defaults and click create distribution. Before, you click create please carefully go through each option and try to understand the settings



As you can see in the status it is shown as 'In Progress'. It takes approximately 15 mins to create the distribution.

And after 15 mins you will see the status as 'Deployed'

As you notice a domain name is already provided by CloudFront.

Amazon CloudFront now knows where your Amazon S3 original server is located.

Step 3: Open the CloudFront link using the domain name assigned

Now if you copy the domain link and add sublink to it you will see that S3 content is delivered to you from your nearest edge location.

Your domain link with your html sublink looks like this:

Prepared by Sriganesh Pera

http://dm186xf1uxq2.cloudfront.net/<name of your html file>

You will see this:

Your html file is cached in your nearest edge location and that is how CloudFront serves your content real fast.

Step 4: Changing the content of web files and understanding cache behavior.

Somehow you felt that the content that you're serving to your customers is wrong and you want to change it.

Now try to change it:

Before you delete your.html link. Copy the link in a text file. Now, go to your S3 Bucket and delete your .html file. Now try to load the same CloudFront domain

http://dm186xf1uxq2.cloudfront.net/<name of your html file>

You will still notice that same html file is loaded in spite of deleting it from its origin S3.

Wondered why?

Because CloudFront caches the data to its nearest edge location and it wont go away even you delete the source.The default time is 24 hrs/86400 seconds that means it will **go away after 1 day which is the default TTL.**

Now update the .html file using this content

```
<html>

<head>Today's News</head>

<h1> Sorry for previous post </h1>

<body>

<p>Real Pirates of Carribean new poster</p>

<p><img
src="https://thecinematicexperiance.files.wordpress.com/2017/01/pirates-of-the-caribbean-3.jpg?w=547"/></p>

</body>

</html>
```

Now save the same content using the same file name and upload it to S3 bucket and make it public and check your CloudFront domain link

You will see that update is not reflected.

This is because the edge location won't get that updated version from your origin as it still cached with the same name.

Then how to change the content?

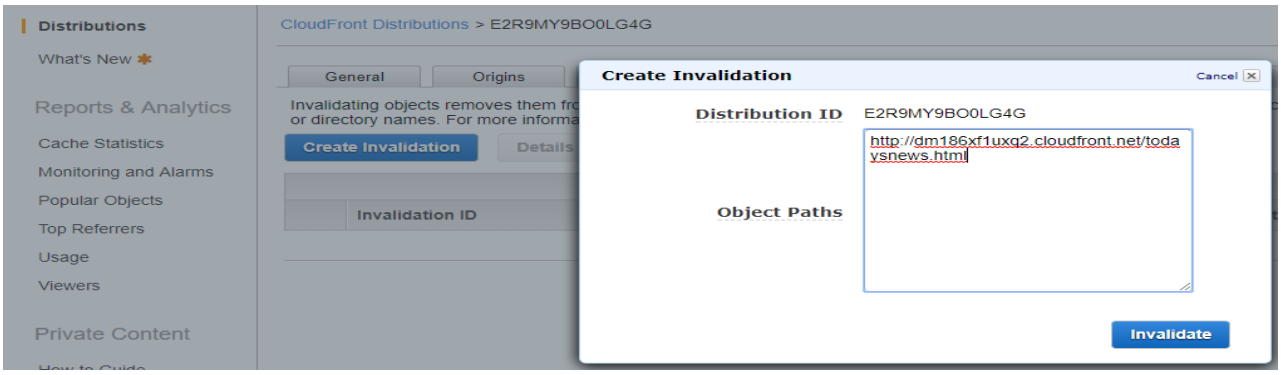
Simply , rename the filename(you can rename inside S3;no need to delete and reupload) to some other name and try to load the CloudFront domain. **You will see the change is reflected.**

Step 5: Delete the cache content

Now in order to remove an object from CloudFront edge caches before it expires:

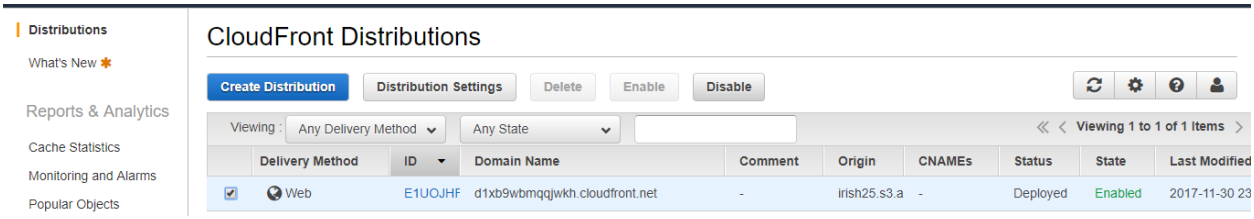
Select your web content in CloudFront dashboard and then click on 'Distribution settings' then click on 'invalidations' and then click on 'Create invalidations' and in the object path give the CloudFront domain like that you saved in a separate file before deleting. Then click create invalidation

to



Step 5:Clean up!

Select the distribution and click disable



Once it is disabled. Delete the distribution.

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~End of Lab ~~~~~~~~~~~~~~~~~~~~~~~~~~~

Prepared by Sriganesh Pera
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~