

Amazon CloudFront (Content Delivery Network)

Amazon CloudFront is AWS's fault-tolerant and globally scalable content delivery network service. It provides seamless integration with other Amazon Web Services to provide an easy way to distribute content.

Amazon CloudFront speeds up distribution of your static and dynamic content through its worldwide network of **edge locations**. Normally, when a user requests content that you're hosting without a CDN, the request is routed back to the source web server which could reside in a different continent to the user initiating the request. However, if you're using CloudFront, the request is instead routed to the closest edge to the user's location which provides the lowest latency to deliver the best performance through cached data.

So essentially Amazon CloudFront acts as a content delivery network service, which provides a means of distributing the source data of your web traffic closer to the end-user requesting the content via **AWS edge locations as cached data**. As this data is cached, after a set period, this cached data will expire and so AWS CloudFront doesn't provide durability of your data. Instead, it distributes the source data which resides on durable storage, such as Amazon S3.

AWS edge locations are sites deployed in major cities and highly populated areas across the globe. While edge locations are not used to deploy your main infrastructure, such as EC2 instances or EBS storage, they are used by AWS services such as AWS CloudFront to cache data and reduce latency for end user access. For **example**, you may have your website hosted on EC2 instances or S3 within the Ohio region, with an associated CloudFront distribution. When a user accesses your website from Europe, they would then be redirected to their closest edge location in Europe, where cached data could be read off your website. This significantly reduces **latency**.

CloudFront uses distributions to control which source data it needs to redistribute and to where. These distributions can be configured as one of two different delivery methods. Firstly, a web distribution and this type of distribution is used if you want to speed up distribution of static and dynamic content, for example, .html, .css, .php, and graphics files, distribute media files using HTTP or HTTPS, add, update, or delete objects, and submit data from web forms, and use live streaming to stream an event in real-time. Alternatively, you can create an RTMP distribution, which is used if you want to distribute streaming media with the Adobe Flash media service RTMP protocol. The benefit of using RTMP distribution is that your end user can start viewing the media before the complete file has been downloaded from the edge location. The source data for an RTMP distribution can only exist within an S3 bucket and not an EC2 web server.

When configuring your distributions, you will be required to enter your origin information, this is essentially where the distribution is going to get the data to distribute across edge locations and it will be the DNS name of the S3 bucket or the HTTP server. If the origin is an S3 bucket, then it can be selected from a drop-down list. If you are using S3 as a static website, you must enter the static hosting website endpoint.

If using an S3 bucket as your origin, then for additional security you can create a CloudFront user called an origin access identity, known as OAI, which can be associated with your newly created distribution. This simply ensures that only this OAI can access and serve content from your bucket and therefore preventing anyone circumventing your CloudFront distribution by accessing the files directly in the bucket using the object URL.

You will also be required to select a host of different caching behavior options, defining how you want the data at the edge location to be cached via various methods and policies. Lastly, you will define the distribution settings themselves, and this will look at which edge locations you want your data to be distributed to, which can either be US, Canada, and Europe, US, Canada, Europe, and Asia, or all edge locations for the best performance. You can also define if you want your distribution to be associated to a web application firewall access control list for additional security and web application protection. For more information on AWS WAF, please see the following course. In addition to using a web application firewall access control list, you can also implement additional encryption security by specifying an SSL certificate that must be used with a distribution.

Once your distribution is configured, you simply enable the distribution for it to be created. When content from your website is accessed, the end-user will be directed to their closest edge location in terms of latency, to see if the content is cached by CloudFront at that edge location. If the content is there, the user will access the content from the edge location instead of the origin, therefore reducing latency. If the content is not there, or the cache has expired for that content at the edge location, then CloudFront will request the content from the source origin again. This content will then be used to maintain a fresh cache for any future request until it again expires.