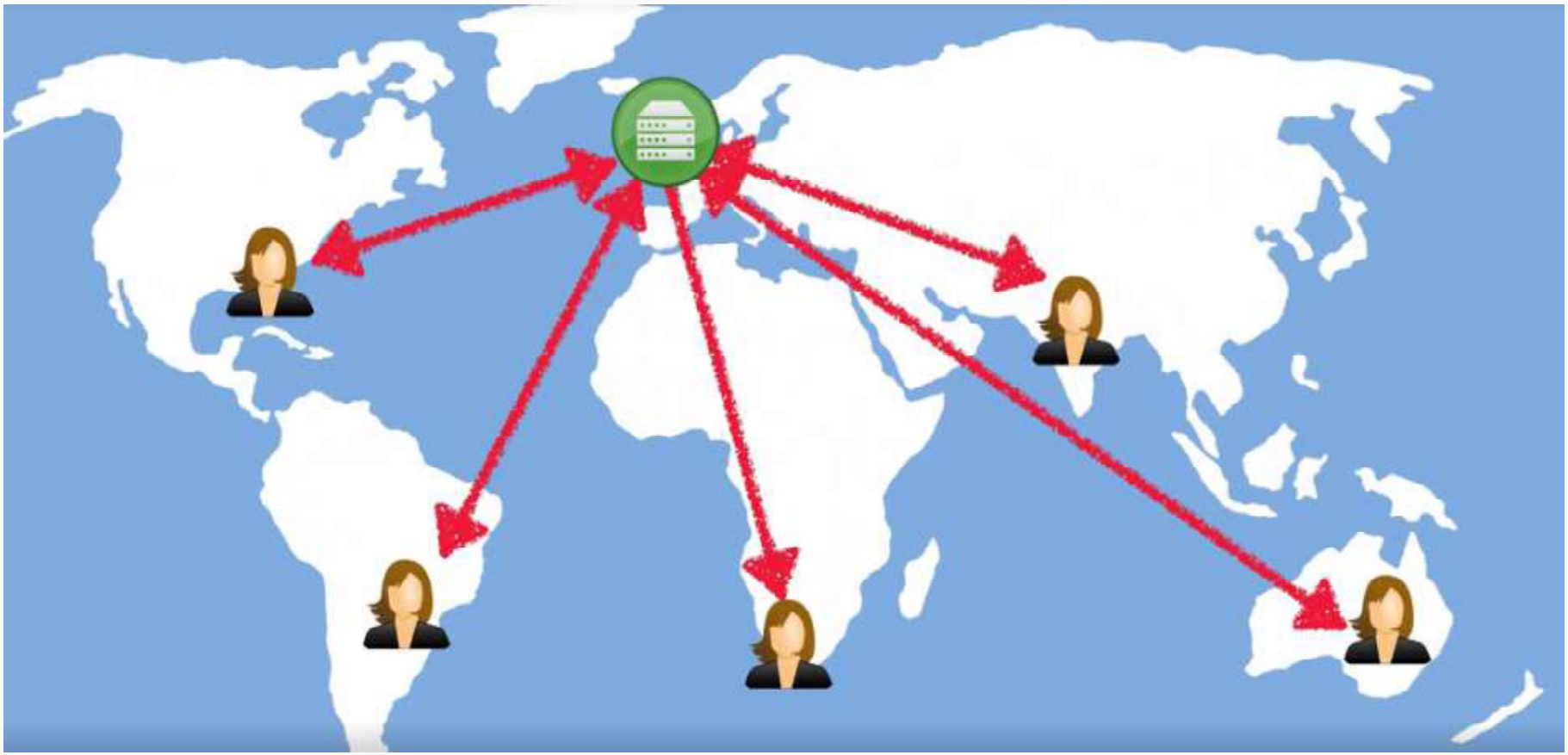
Amazon Web Services

CloudFront

Why Content Delivery Network



Content Delivery Network (CDN)

➢A *content delivery network* is a system of distributed servers that deliver pages and other Web content to a user, based on the geographic locations of the user, the origin of the webpage and the content delivery server

➢This service is effective in speeding the delivery of content of websites with high traffic and websites that have global reach.

➢The closer the CDN server is to the user geographically, the faster the content will be delivered to the user.

➢CDNs also provide protection from large surges in traffic

How CDN Works

➢Servers nearest to the website visitor respond to the request.

➢The content delivery network copies the pages of a website to a network of servers that are dispersed at geographically different locations, caching the contents of the page.

➢When a user requests a webpage that is part of a content delivery network, the CDN will redirect the request from the originating site's server to a server in the CDN that is closest to the user and deliver the cached content.

➢CDNs will also communicate with the originating server to deliver any content that has not been previously cached

➢When delivering large scale websites to a global audience, CDNs can reduce latency, accelerate site load times, reduce bandwidth consumption secure applications and even block data scrappers and other forms of spammers hitting your server.

How CDN Works

CDN : Key Concepts

- System of Distributed Servers or Network to deliver Web Content to a user based on Geographic location of the user, the origin of web page and the Content Delivery Server.
- **Edge Location:** Location where the content will be Cached. This different from AWS Region/AZ's.
- **Origin:** Origin of files that CDN will distribute. This can either be an S3 bucket, an EC2 instance, an Elastic load balancer or Route53.
- **Distribution:** CDN consisting of collection of Edge Locations.
- **TTL or Time to Live:** The amount of time that your data would be cached on edge locations.
- **Web Distribution :** Used for Websites
- **RTMP:** Used for media streaming. (Adobe Flash Content)

CDN : Key Concepts

- System of Distributed Servers or Network to deliver Web Content to a user based on Geographic location of the user, the origin of web page and the Content Delivery Server.
- **Edge Location:** Location where the content will be Cached. This different from AWS Region/AZ's.
- **Origin:** Origin of files that CDN will distribute. This can either be an S3 bucket, an EC2 instance, an Elastic load balancer or Route53.
- **Distribution:** CDN consisting of collection of Edge Locations.
- **TTL or Time to Live:** The amount of time that your data would be cached on edge locations.
- **Web Distribution :** Used for Websites
- **RTMP:** Used for media streaming. (Adobe Flash Content)

CloudFront Key Concepts

- Amazon CloudFront can be used to deliver an entire website , including the dynamic, static, streaming and interactive content using a global network of Edge Locations.
- Request are automatically routed to the nearest edge location so the content is delivered with the best possible performance.
- Integrates with other Amazon Services like S3, EC2 and Load balancers etc.
- Amazon CloudFront works seamlessly with any non-AWS origin server, that stores the original versions of the files.
- Edge location are not for READ only, we can write to these Edge locations as well.
- Cached objects may be cleared but there are charges for it.

Setting up CloudFront to Deliver Your Content

➢You specify ***origin servers***, like an Amazon S3 bucket or your own HTTP server, from which CloudFront gets your files which will then be distributed from CloudFront edge locations all over the world. An origin server stores the original of your objects.

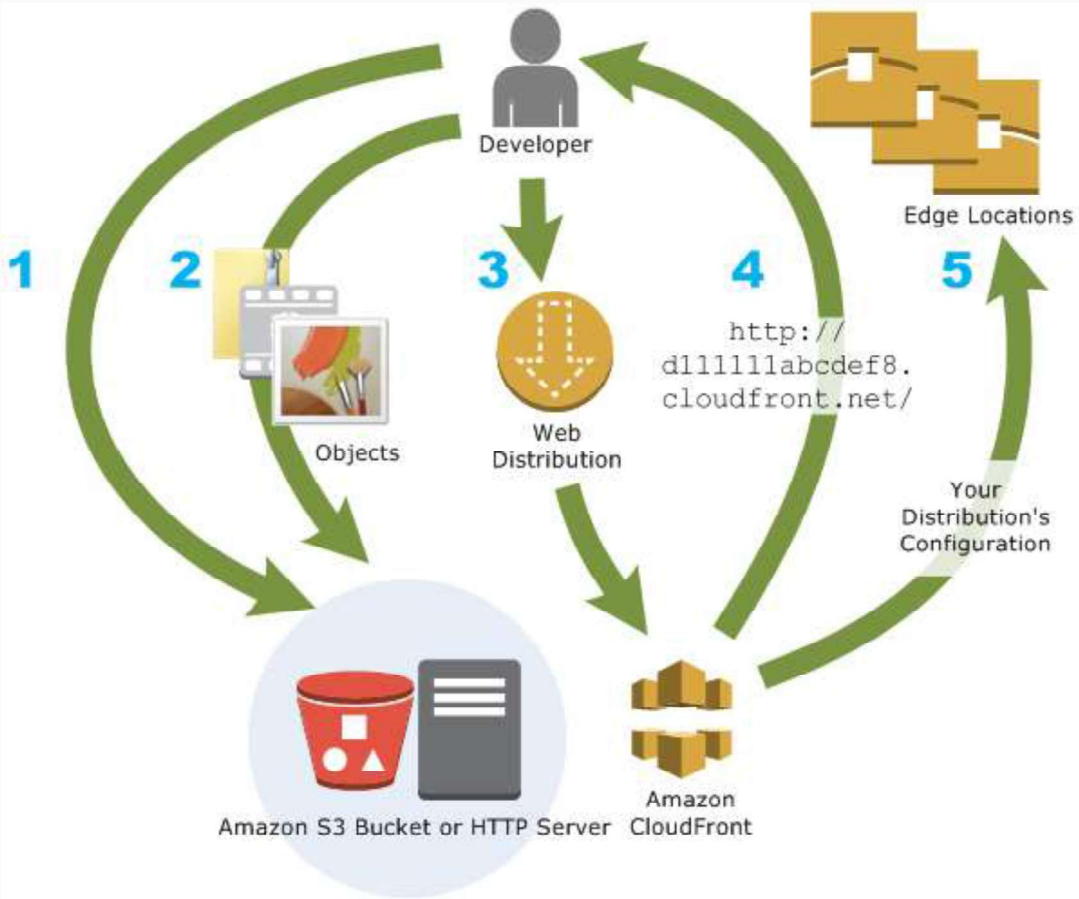
➢You **upload your files** to your origin servers. Your files include web pages, images, and media files etc

➢You **create a CloudFront *distribution***, which tells CloudFront which origin servers to get your files from when users request the files through your web site or application.

➢CloudFront **assigns a domain name** to your new distribution

➢CloudFront sends your distribution's configuration (but not your content) to all of its **edge locations**

Setting up CloudFront to Deliver Your Content



How CloudFront Delivers Content to Your Users

➢A user accesses your website or application and requests one or more objects, such as an image file and an HTML file.

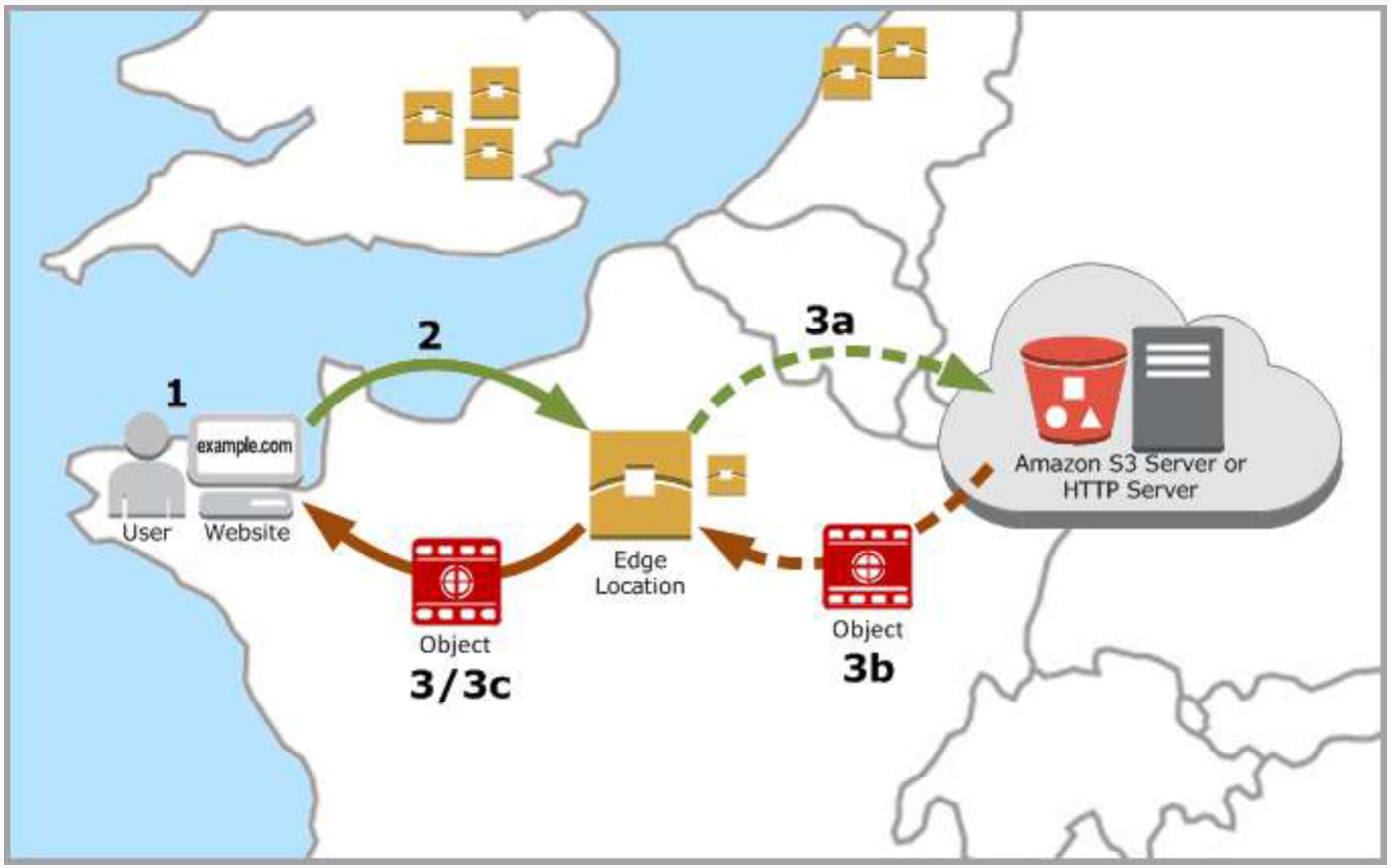
➢DNS routes the request to nearest CloudFront edge location in terms of latency

➢In the edge location, CloudFront checks its cache for the requested files. If the files are in the cache, CloudFront returns them to the user.

➢ If the files are *not* in the cache,

1. CloudFront compares the request with the specifications in your distribution and forwards the request for the files to the applicable origin server for the corresponding file type.
2. The origin servers send the files back to the CloudFront edge location.
3. As soon as the first byte arrives from the origin, CloudFront begins to forward the files to user.
4. CloudFront also adds the files to the cache in the edge location for the next time someone requests those files.

How CloudFront Delivers Content to Your Users



Accessing CloudFront

- AWS Management Console
- AWS SDKs
- CloudFront API
- AWS Command Line Interface
- AWS Tools for Windows PowerShell

CloudFront Reports

- CloudFront Cache Statistics Reports
  - Total number of requests for all HTTP status codes and all methods
  - Percentage of Viewer Requests by Result Type
  - Bytes Transferred to Viewers
  - Percentage of GET Requests that Didn't Finish Downloading

- CloudFront Popular Objects Report
  - Number of requests for the object, hits and misses, the hit ratio, bytes served for misses, the total bytes served, incomplete downloads

- CloudFront Top Referrers Report
  - Number of requests from a referrer, and the number of requests from a referrer as a percentage of the total number of requests during the specified period.

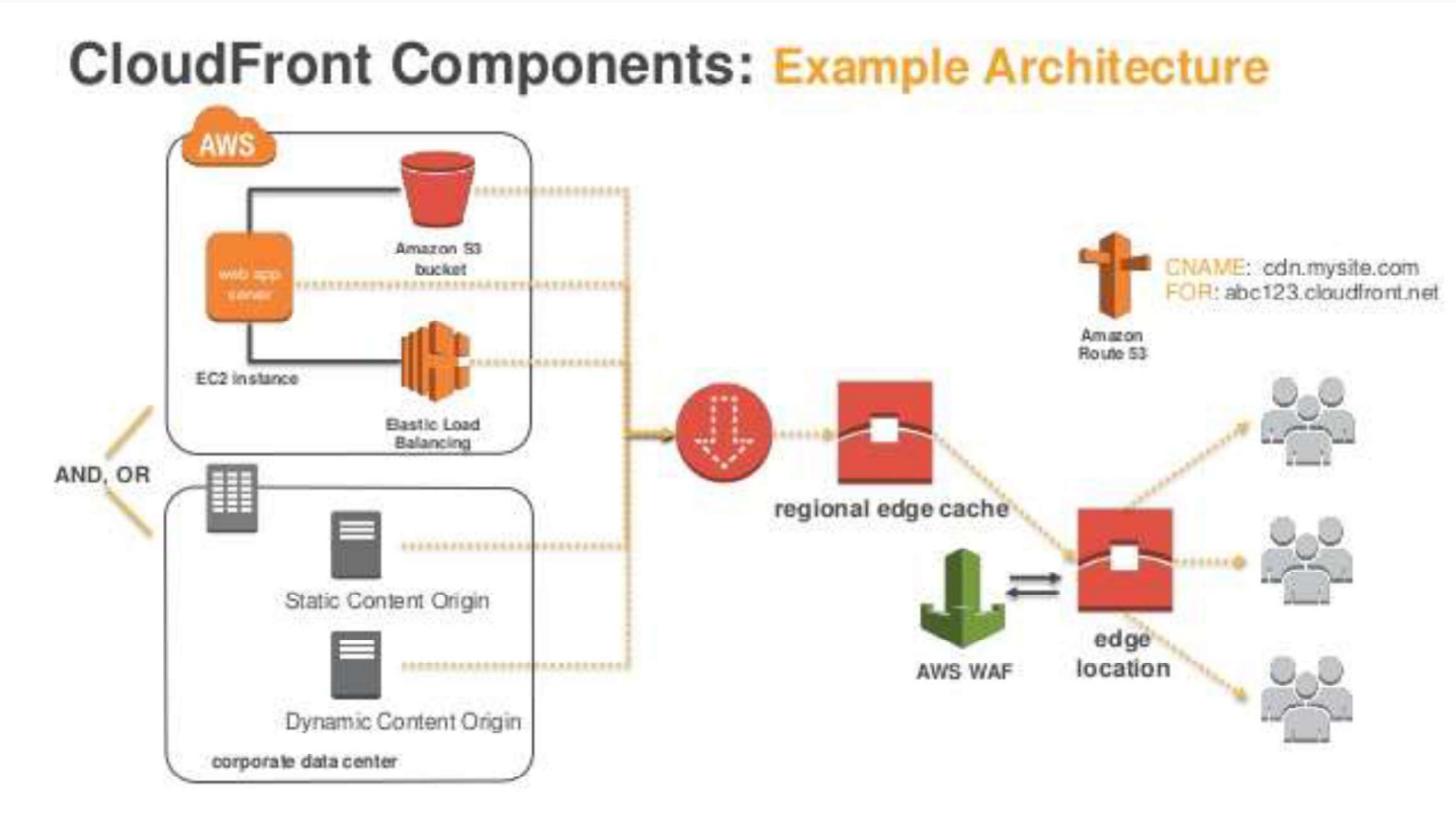
- CloudFront Usage Reports
  - Number of Requests, Data Transferred by Protocol, Data Transferred by Destination

- CloudFront Viewers Reports
  - Devices , Browsers, Operating Systems and Locations

CloudFront Service Components

- Distributions
- Origins
- Behaviors
- Restrictions, Error Pages, Tags
- AWS WAF Web ACL
- Edge Locations
- Price Classes

CloudFront Components



Web and RTMP Distributions

➢You can use web distributions to serve the following content over HTTP or HTTPS

1. Static and dynamic download content
2. Multimedia content. You can't serve Adobe Flash multimedia content over HTTP or HTTPS
3. A live event, such as a meeting, conference, or concert, in real time.

➢For web distributions, your origin can be either an Amazon S3 bucket or an HTTP server, for example, a web server.

➢RTMP distributions stream media files using Adobe Media Server and the Adobe Real-Time Messaging Protocol (RTMP)

➢An RTMP distribution must use an Amazon S3 bucket as the origin.

Origins

➢When you create a web distribution, you specify where CloudFront sends requests for the files that it distributes to edge locations.

➢CloudFront supports using Amazon S3 buckets and HTTP servers as origins.

➢When you use Amazon S3 as an origin for your distribution, you place any objects that you want CloudFront to deliver in an Amazon S3 bucket.

➢Configure CloudFront to communicate with your Amazon S3 bucket using SSL.

➢Use an origin access identity to require that your users access your content using CloudFront URLs

➢Update the content of your bucket by submitting POST and PUT requests to CloudFront.

Restrictions, Errors

➢Restrictions is a property of the CloudFrontDistribution DistributionConfig property type that lets you limit which viewers can access your content.

Eg: Geographical Restrictions

➢When CloudFront requests an object from your Amazon S3 bucket or custom origin server, your origin sometimes returns an HTTP 4xx or 5xx status code, which indicates an error has occurred. For 5xx status codes, whether the requested object is currently in the CloudFront edge cache.

Web Access Control List (ACL)

➢A web access control list (web ACL) gives you fine-grained control over the web requests that your Amazon CloudFront distributions or Application Load Balancers respond to.

➢You can allow or block the following types of requests:

Originate from an IP address or a range of IP addresses
Originate from a specific country or countries
Contain a specified string or match a regular expression in a particular part of requests
Exceed a specified length
Appear to contain malicious SQL code (known as SQL injection)
Appear to contain malicious scripts (known as cross-site scripting)

Edge Location

➢An edge location is where end users access services located at AWS.

➢They are located in most of the major cities around the world and are specifically used by CloudFront (CDN) to distribute content to end user to reduce latency.

➢It is like frontend for the service we access which are located in AWS cloud.

➢Amazon CloudFront has added several regional edge cache locations globally, at close proximity to your viewers.

➢They are located between your origin web server and the global edge locations that serve content directly to your viewers.

Price Classes

➢CloudFront has edge locations all over the world. Cost for each edge location varies and, as a result, the price varies depending on the edge location from which CloudFront serves your requests.

➢CloudFront edge locations are grouped into geographic regions, and regions are grouped into price classes.

➢Best Coverage – All regions

➢Lowest Cost - North America and Europe

➢Great Coverage & Optimized cost – North America, Europe, East and South East Asia

Benefits Of using CloudFront

Time Saving

It reduces effort and ensures productivity enhancement.

Content Privacy

We can enable Amazon CloudFront's private feature to regulate who can access our content and who can not.

Geo Targeting

Detects the geographical area or the country where the end users are accessing the content. This feature may help any business to promote it's products and services in the country where most end users are.

Quick Content Delivery

It uses HTTP or HTTPS protocols to deliver content and recognizes the origin servers that contain the original version of our content by using URL rules that we configure which results in Quick delivery of the content that we want to search.

Less Expensive

Since, we only pay for the data transfer and requests that we actually use it saves a lot of money for us.