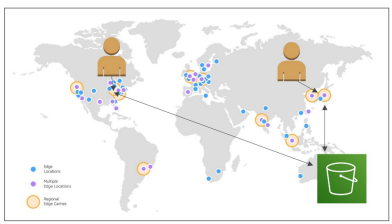
§ CloudFront & AWS Global Accelerator

Amazon CloudFront

- Content Delivery Network (CDN)
- Improves read performance, content is cached at the edge
- Improves users experience
- 216 Point of Presence globally (edge locations)
- DDoS protection (because worldwide), integration with Shield, AWS Web Application Firewall

→ store distributed

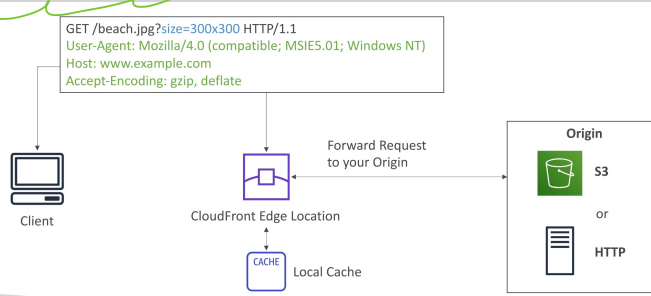


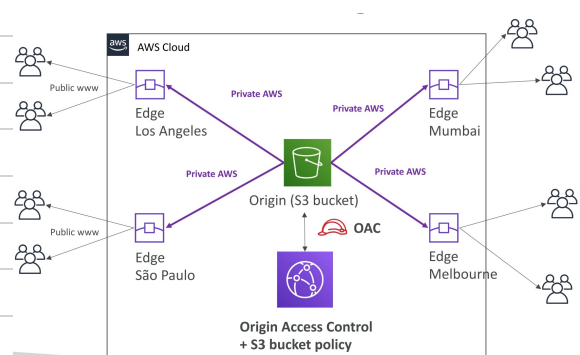
Source: https://aws.amazon.com/cloudfront/features/?nc=sn&loc=2

S3 bucket
- for distributing & caching
- origin access control (OAC)
- used as an ingress to upload files to S3

cloudfront origins

Custom origin
- ALB
- EC2
- S3 static website
- Any http backend





CloudFront vs S3 Cross Region Replication

- CloudFront:
  - Global Edge network
  - Files are cached for a TTL (maybe a day)
  - Great for static content that must be available everywhere
- S3 Cross Region Replication:
  - Must be setup for each region you want replication to happen
  - Files are updated in near real-time
  - Read only
  - Great for dynamic content that needs to be available at low-latency in few regions

→ We can use cloudfront easy without making the bucket public, by using origin access control

go to cloudfront > distribution > create distribution

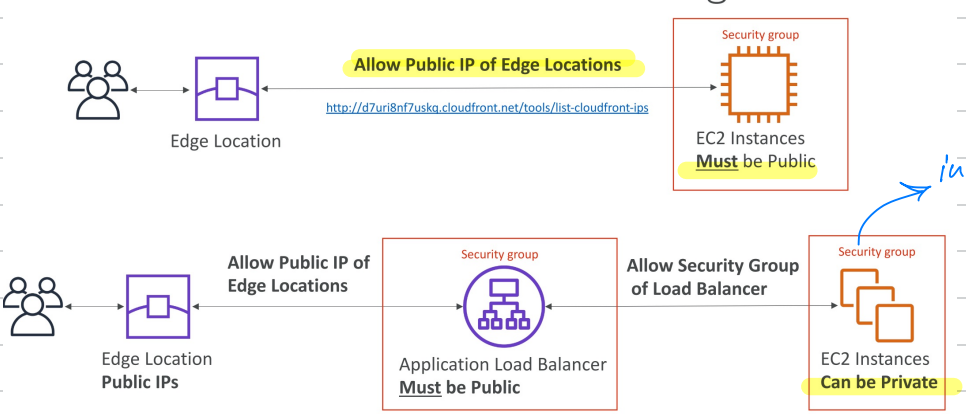
need to update bucket policy. < Create control < Origin access cont < name < setting

→ Adding bucket policy for cloudfront ... just go to edit distribution & copy policy ! paste S3 ...

dynamic object = index.html

once deployed, just open the domain in cloudfront.

CloudFront – ALB or EC2 as an origin



CloudFront Geo Restriction

distribution > security > cloudfront geo restrictions.

- You can restrict who can access your distribution
  - **Allowlist:** Allow your users to access your content only if they're in one of the countries on a list of approved countries.
  - **Blocklist:** Prevent your users from accessing your content if they're in one of the countries on a list of banned countries.
- The "country" is determined using a 3rd party Geo-IP database
- Use case: Copyright Laws to control access to content

→ pricing is cloudfront is different all around the world.

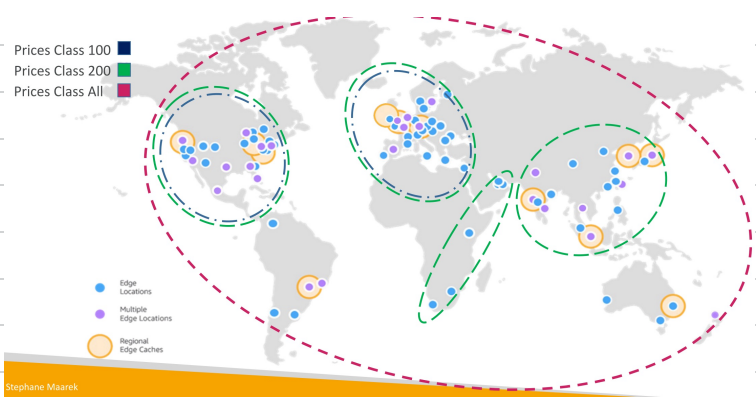
more the data out of cloudfront lower the price

min → US, Mexico & Canada
max → India

CloudFront – Price Classes

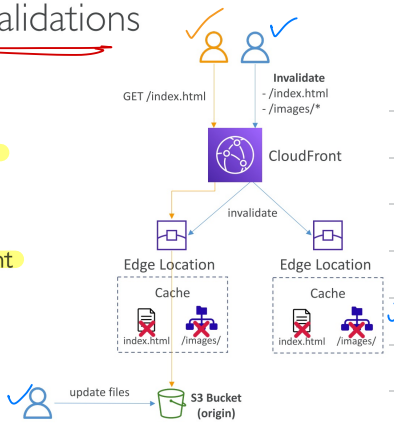
- You can reduce the number of edge locations for **cost reduction**
- Three price classes:
  1. **Price Class All:** all regions – best performance
  2. **Price Class 200:** most regions, but excludes the most expensive regions
  3. **Price Class 100:** only the least expensive regions

Edge Locations Included Within	United States, Mexico, & Canada	Europe & Israel	South Africa, Kenya, & Middle East	South America	Japan	Australia & New Zealand	Hong Kong, Philippines, Singapore, South Korea, Taiwan, & Thailand	India
Price Class All	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price Class 200	Yes	Yes	Yes	x	Yes	x	Yes	Yes
Price Class 100	Yes	Yes	x	x	x	x	x	x



CloudFront – Cache Invalidations

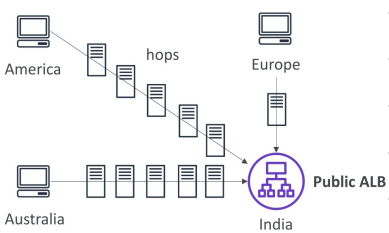
- In case you update the back-end origin, CloudFront doesn't know about it and will only get the refreshed content after the TTL has expired
- However, you can force an entire or partial cache refresh (thus bypassing the TTL) by performing a CloudFront Invalidation
- You can invalidate all files (*) or a special path (/images/*)



this file I will be removed after invalidation

Global users for our application

- You have deployed an application and have global users who want to access it directly.
- They go over the public internet, which can add a lot of latency due to many hops
- We wish to go as fast as possible through AWS network to minimize latency

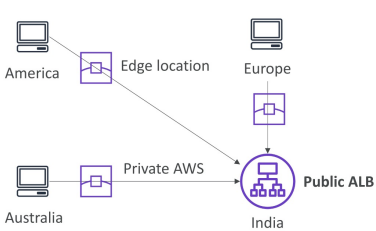


- **Unicast IP:** one server holds one IP address
- **Anycast IP:** all servers hold the same IP address and the client is routed to the nearest one

this concept is used by global accelerator.

AWS Global Accelerator

- Leverage the AWS internal network to route to your application
- 2 Anycast IP are created for your application
- The Anycast IP send traffic directly to Edge Locations
- The Edge locations send the traffic to your application



→ works with EC2, Elastic IPs, ALB, NLB, public or private

- Consistent Performance
  - Intelligent routing to lowest latency and fast regional failover
  - No issue with client cache (because the IP doesn't change)
  - Internal AWS network
- Health Checks
  - Global Accelerator performs a health check of your applications
  - Helps make your application global (failover less than 1 minute for unhealthy)
  - Great for disaster recovery (thanks to the health checks)
- Security
  - only 2 external IP need to be whitelisted
  - DDoS protection thanks to AWS Shield

→ global service but must specify US West (Oregon)

→ if endpoint is ALB then will also leverage its health checks.

AWS Global Accelerator vs CloudFront

- They both use the AWS global network and its edge locations around the world
- Both services integrate with AWS Shield for DDoS protection.

- CloudFront
  - Improves performance for both cacheable content (such as images and videos)
  - Dynamic content (such as API acceleration and dynamic site delivery)
  - Content is served at the edge

→ content is cached at edge

- Global Accelerator
  - Improves performance for a wide range of applications over TCP or UDP
  - Proxying packets at the edge to applications running in one or more AWS Regions.
  - Good fit for non-HTTP use cases, such as gaming (UDP), IoT (MQTT), or Voice over IP
  - Good for HTTP use cases that require static IP addresses
  - Good for HTTP use cases that required deterministic, fast regional failover

→ the packets reach it to the application but by using internal network

create a application on EC2 instance which exposes as http endpoint. in 2 regions.

Security group allow 3rd for http

create global accelerator.

Add listner
- port
- protocol
- client affinity (like sticky session)

Add endpoint groups (Add 2 endpoints for the EC2 instances