

Why Global Application?

Why make a global application?



- A global application is an application deployed in multiple geographies
- On AWS: this could be Regions and / or Edge Locations
- Decreased Latency
 - Latency is the time it takes for a network packet to reach a server
 - It takes time for a packet from Asia to reach the US
 - Deploy your applications closer to your users to decrease latency, better experience
- Disaster Recovery (DR)
 - If an AWS region goes down (earthquake, storms, power shutdown, politics)...
 - You can fail-over to another region and have your application still working
 - A DR plan is important to increase the availability of your application
- Attack protection: distributed global infrastructure is harder to attack

1.

Global AWS Infrastructure

- Regions: For deploying applications and infrastructure
- Availability Zones: Made of multiple data centers
- Edge Locations (Points of Presence): for content delivery as close as possible to users
- More at:
<https://infrastructure.aws/>



2.

Global Applications in AWS

- Global DNS: Route 53
 - Great to route users to the closest deployment with least latency
 - Great for disaster recovery strategies
 - Global Content Delivery Network (CDN): CloudFront
 - Replicate part of your application to AWS Edge Locations – decrease latency
 - Cache common requests – improved user experience and decreased latency
 - S3 Transfer Acceleration
 - Accelerate global uploads & downloads into Amazon S3
 - AWS Global Accelerator:
 - Improve global application availability and performance using the AWS global network
- 3.

Route 53 Overview:

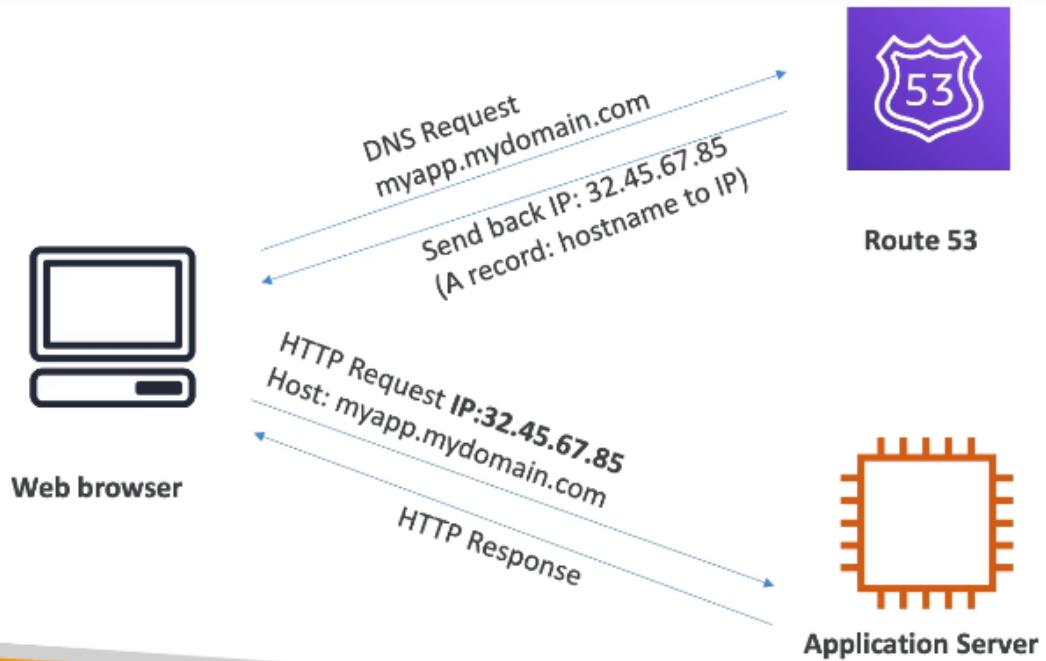
Amazon Route 53 Overview



- Route53 is a Managed DNS (Domain Name System)
 - DNS is a collection of rules and records which helps clients understand how to reach a server through URLs.
-
- In AWS, the most common records are:
 - www.google.com => 12.34.56.78 == A record (IPv4)
 - www.google.com => 2001:0db8:85a3:0000:0000:8a2e:0370:7334 == AAAA IPv6
 - search.google.com => www.google.com == CNAME: hostname to hostname
 - example.com => AWS resource == Alias (ex: ELB, CloudFront, S3, RDS, etc...)

- 1.
2. Lets suppose we have a user who wants to access our application server which has a public IPv4..Now we want to be able to access our application server using a normal URL.
3. For this..first we go into route 53 and we create a “A record” ...so when a browser request access to my application..then dns will reply back with its IP.
4. Then this can be used to send request to our application server..and application server return s the http response

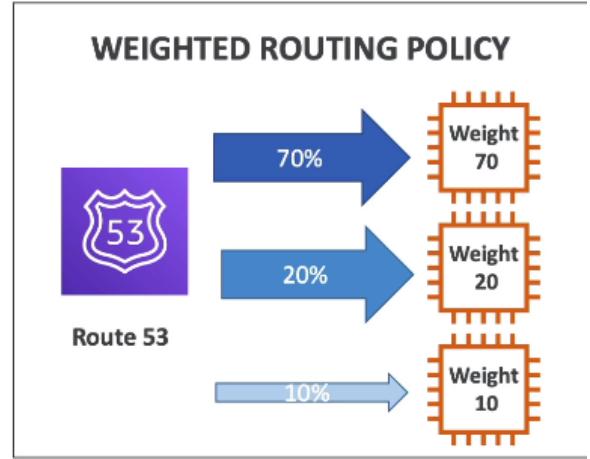
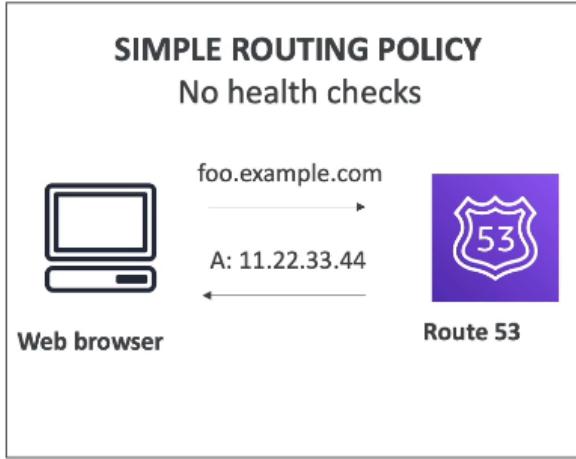
Route 53 – Diagram for A Record



- 5.
6. Route 53 Routing policies
7. The simple routing policy ..does not perform any health checks..and here browser requests Route 53 and it sends back the ip address of our application
8. Weighted Routing policy...Here we will be having multiple instances and each instance will allocated with the weights..
9. So here 70% of traffics goes to instance 1...20% to instance 2 and 10% to instance3..It is similar to the load balancer

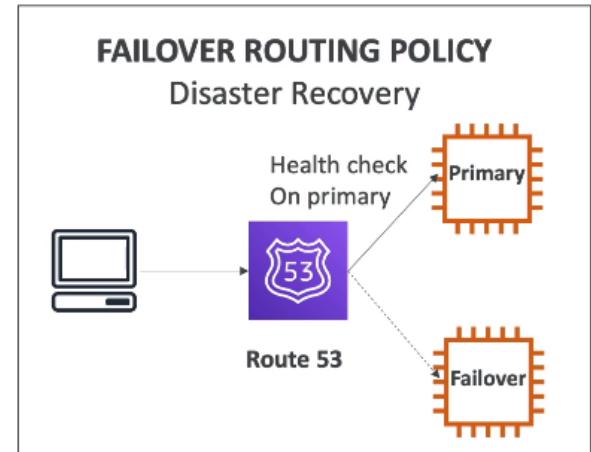
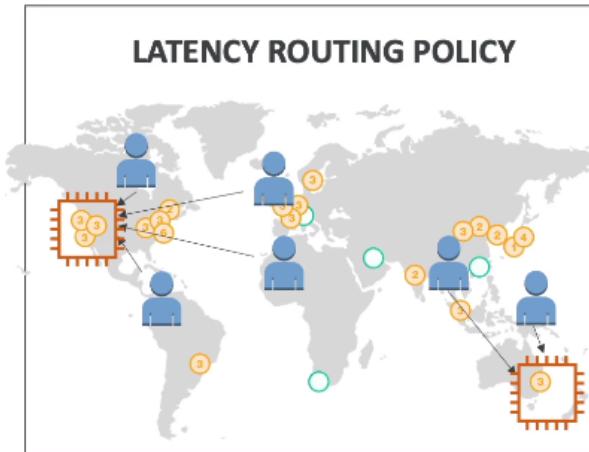
Route 53 Routing Policies

- Need to know them at a high-level for the Cloud Practitioner Exam



- 10.
11. Latency Routing policy
12. Lets suppose we have 2 instances one in california and one in australia...So this policy routes our users to the nearest located instance ..so as to decrease the latency
13. So, in this example,
14. Route 53 will be used to minimize the latency between the users and the servers by making the users connect to the server. That is the closest to them.
15. And we have FailOver Routing policy....
16. Here we have a primary instance and failover(backup) instance...our route 53 will health check on primary...If in case route 53 detects our primary instance is unhealthy,,it redirects requests to failover instance

Route 53 Routing Policies



17.

Route 53 HandsOn:

1. First we will click on route 53 console

1: Domain Search Choose a domain name
2: Contact Details stephane-ccp .com - \$12.00 Check
3: Verify & Purchase To register a domain name, start by finding one that's available. Enter the first part of the name (such as example in example.com), choose an extension (such as .com or .org), and click Check. We'll tell you whether it's available and whether you can get it with other extensions. Learn more.

2.

Note: using Route 53 costs \$12 a year for the domain
And \$0.50 per month for the hosted zone
Don't follow along with me if you want to remain within the free tier

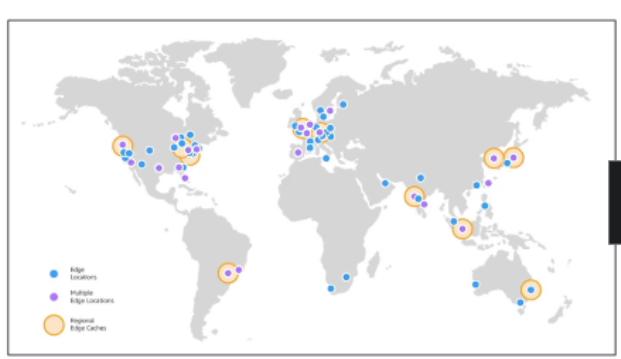
CloudFront Overview

1. It is a Content Delivery Network
2. CloudFront is a content delivery network, or CDN, so anytime you see CDN at the exam, think CloudFront.
3. It improves the read performance by caching the content of your website at the different edge locations. And because your content is cached all around the world, then your users all around the world will have a lower latency,
4. we are getting DDoS protection. So DDoS is a sort of attack where all your servers around the world are getting attacked at the same time,

AWS 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



The map displays the global distribution of CloudFront edge locations and edge caches. Edge locations are marked with blue dots, while edge caches are marked with orange circles. The locations are concentrated in North America, Europe, and Asia, with smaller clusters in South America, Australia, and Africa.

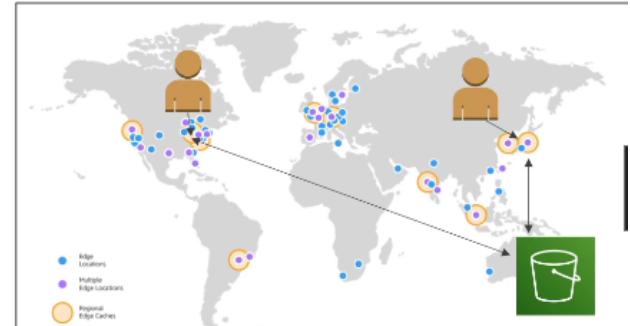
Source: <https://aws.amazon.com/cloudfront/features/?nc=sn&loc=2>

- 5.
6. Lets suppose we have a s3 bucket in australia..and a user from us region request the content from american edge location using cloud front..and cloud front will fetch the content from australia
7. Now, if another user in the US will be requesting the same content, then it will be served directly from the edge and it will not go all the way to Australia to serve that content.

AWS 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



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

8.

what is cloudfront origin in aws

In AWS, a CloudFront origin is the location where your content is stored. CloudFront uses origins to deliver your content to users around the world with high performance and low latency.

9.

CloudFront – Origins

- S3 bucket
 - For distributing files and caching them at the edge
 - Enhanced security with CloudFront Origin Access Control (OAC)
 - OAC is replacing Origin Access Identity (OAI)
 - CloudFront can be used as an ingress (to upload files to S3)
- Custom Origin (HTTP)
 - Application Load Balancer
 - EC2 instance
 - S3 website (must first enable the bucket as a static S3 website)
 - Any HTTP backend you want

10.

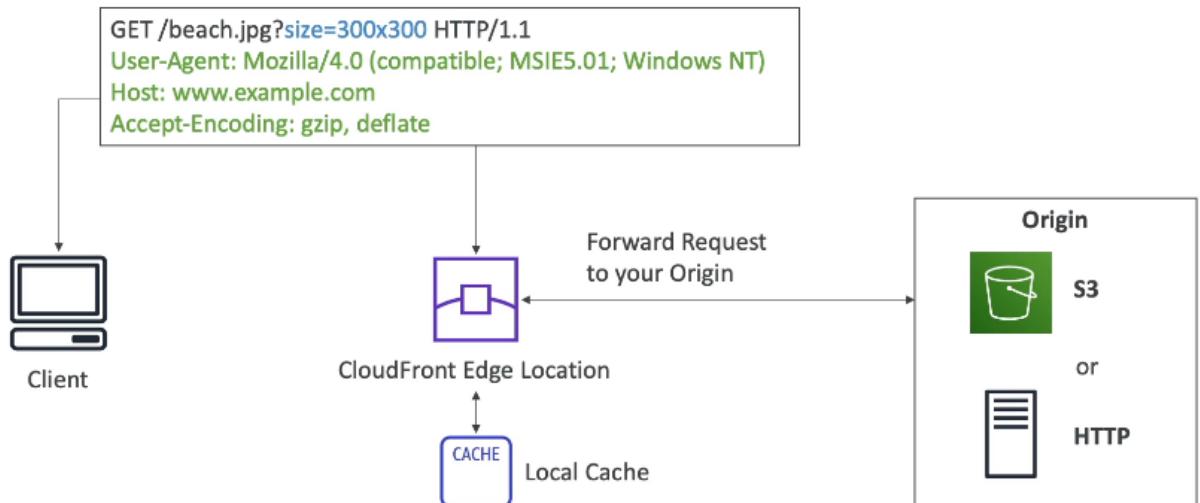
Sure. Here is an example of a CloudFront origin:

- **Amazon S3 bucket:** An Amazon S3 bucket is a simple storage service that offers high availability, durability, scalability, and security. You can use an Amazon S3 bucket as a CloudFront origin by configuring CloudFront to deliver content from your Amazon S3 bucket.

For example, suppose you have an Amazon S3 bucket named **my-bucket** that contains a website. You can configure CloudFront to deliver content from your Amazon S3 bucket by specifying the **my-bucket** bucket in the CloudFront origin configuration. When a user requests a file from your website, CloudFront will fetch the file from your Amazon S3 bucket and deliver it to the user.

- 11.
12. Cloudfront at high level
13. Lets suppose we have an cloudfront edge location..and it is connected to the origin
14. Now the client request to access the data...in our example its beach.jpg
15. Now our cloud front checks..if its is in the cache...if its nor der..then it requests it from the origin and stores it in cache
16. Now If other user trying to access the same data from same edge location...Then it need not go to the origin..as it is having it in its cache

CloudFront at a high level



- 17.
18. S3 as an Origin

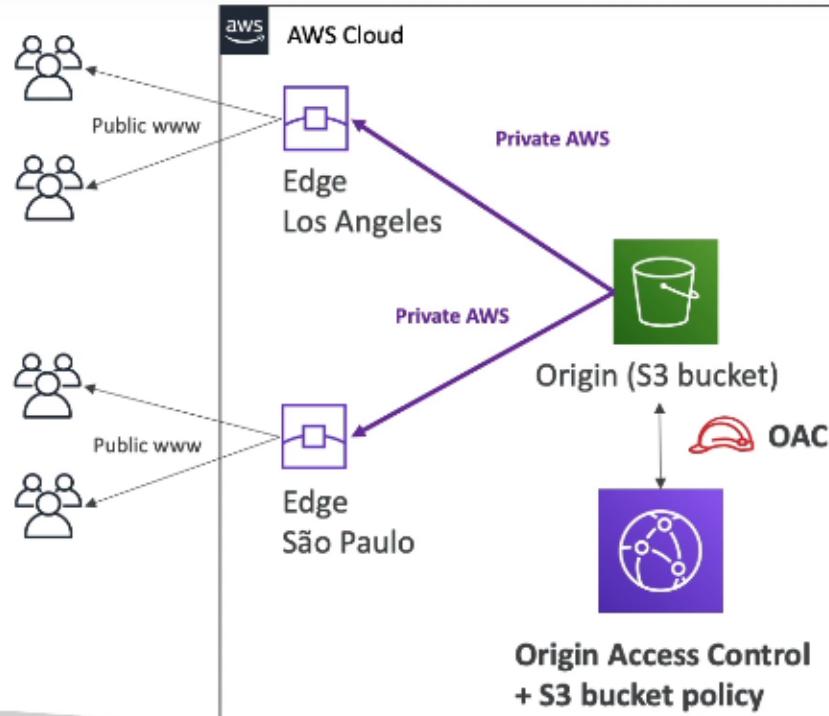
Sure. In AWS, an Amazon S3 bucket origin is a location where your content is stored. CloudFront uses origins to deliver your content to users around the world with high performance and low latency.

To use an Amazon S3 bucket as a CloudFront origin, you need to specify the bucket name and the region in which the bucket is located. CloudFront will then fetch the content from your S3 bucket and deliver it to users.

Here are some of the benefits of using an Amazon S3 bucket as a CloudFront origin:

- **High availability:** Amazon S3 buckets are highly available and durable. This means that your content is always available to users, even if one of the S3 servers goes down.
- **Scalability:** Amazon S3 buckets can scale to handle large amounts of traffic. This means that your website or application can handle more users without impacting performance.
- **Security:** Amazon S3 buckets are protected by a variety of security measures, including HTTPS, bucket policies, and object ACLs. This helps to protect your content from unauthorized access and malicious attacks.

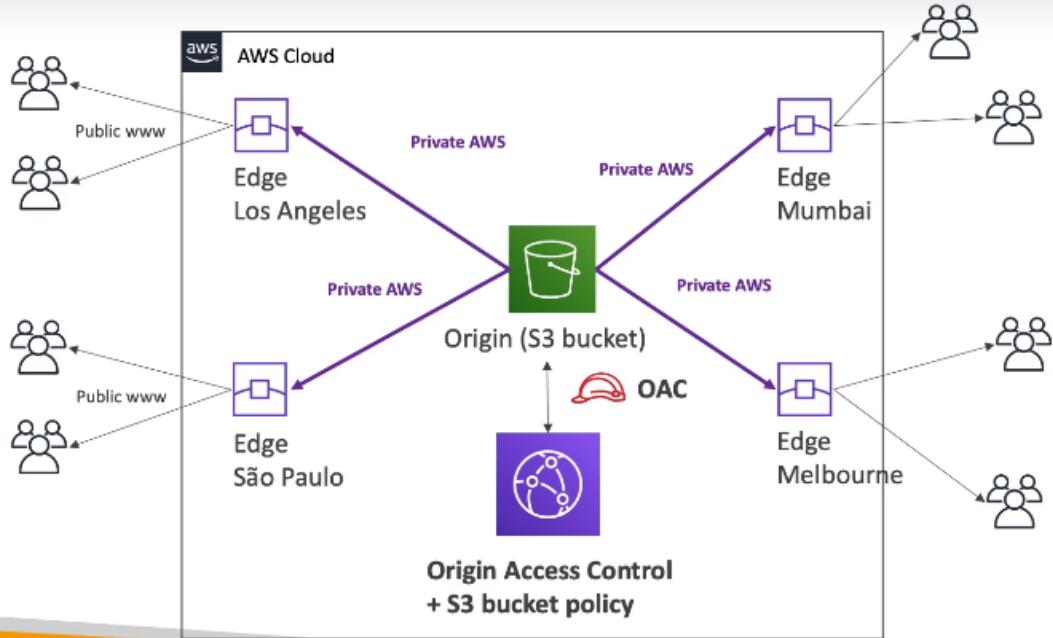
CloudFront – S3 as an Origin



20.

21. Here we have an s3 bucket as an origin in some region..and we have edge location all over the world
22. And your users accessing the edge location in Los Angeles will get their content directly served through the edge location, but first the edge location will get it from the origin S3 bucket through the private network ..and later edge location will store them..and later delivers it in low latency

CloudFront – S3 as an Origin



23.

24. So we can see content from our s3 bucket...is delivering across the world..using edge locations

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

25.

26. So they serve very different purposes, CloudFront is a CDN, which is to cache content all around the world,

27. whereas S3 Cross-Region Replication is to really replicate an entire bucket into another region. So hopefully that makes sense about CloudFront.

CloudFront HandsOn:

1. Here first we have to create an S3 bucket to hold our files for our distribution. So, let's create a bucket
2. After creating the we will upload files into it
3. Now we will see how we use cloudfront and make these files accessible without making them public
4. Now go to cloudfront and click on create a global distribution..next we will give a name
5. Then we have to create an OAC ...
6. And at last we have to provide default root object..and click on create distribution
7. After successfully creating we get a warning to update s3 bucket policy..and it also gives us the policy
8. Then go to s3 bucket policy and update the policies
9. Now if we open our cloudfront domain..it approaches s3 bucket for data and then store the data in its cache
10. Later if we open the same url..our data will be loaded from the cloudfront
11. What is OAC?

Origin access control (OAC) in AWS CloudFront is a feature that allows you to control which users have access to your content. You can use OAC to restrict access to your content to specific IP addresses or ranges, or to specific CloudFront distributions.

OAC can be used to protect your content from unauthorized access, such as from malicious actors or from users who are not supposed to have access to your content.

To use OAC, you need to create an origin access identity (OAI). An OAI is a unique identifier that you can use to grant access to your content. You can then associate the OAI with one or more CloudFront distributions. When a user requests content from a CloudFront distribution that is associated with an OAI, CloudFront will check the user's IP address against the list of allowed IP addresses or ranges. If the user's IP address is not allowed, CloudFront will deny the request.

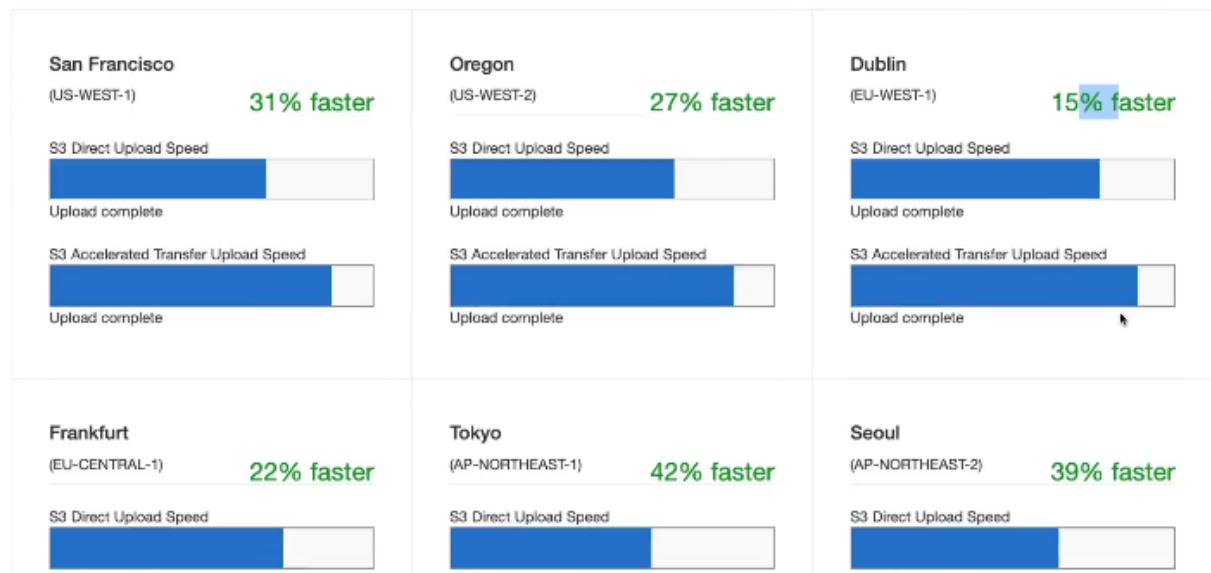
OAC is a powerful tool that can be used to protect your content from unauthorized access. If you are concerned about the security of your content, you should consider using OAC.

- 12.

S3 Transfer Acceleration

1. If a user(who is in USA) wants to upload data to the s3 bucket which is in australia..
2. Then first we'll upload our file into an edge location in USA...And edge location using its internal private network..will send the data to the s3 bucket in reliable and fast way
3. This is a sample of s3 acceleration
4. And this is only used when you want to upload or download a file from this bucket. That is far away from you.
5. We can check the transfer speed using a tool

Upload speed comparison in other regions

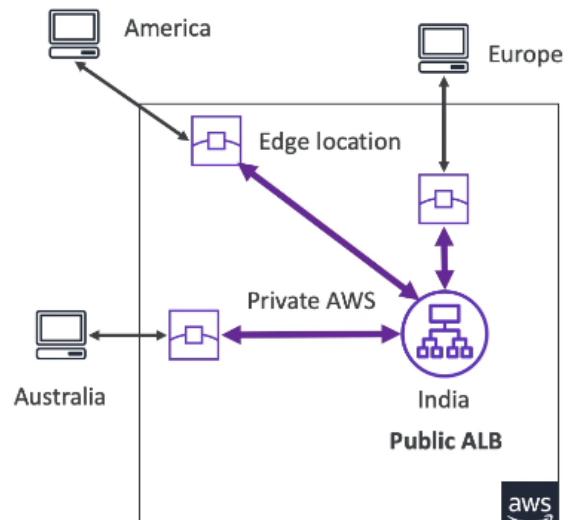


AWS Global Accelerator

AWS Global Accelerator

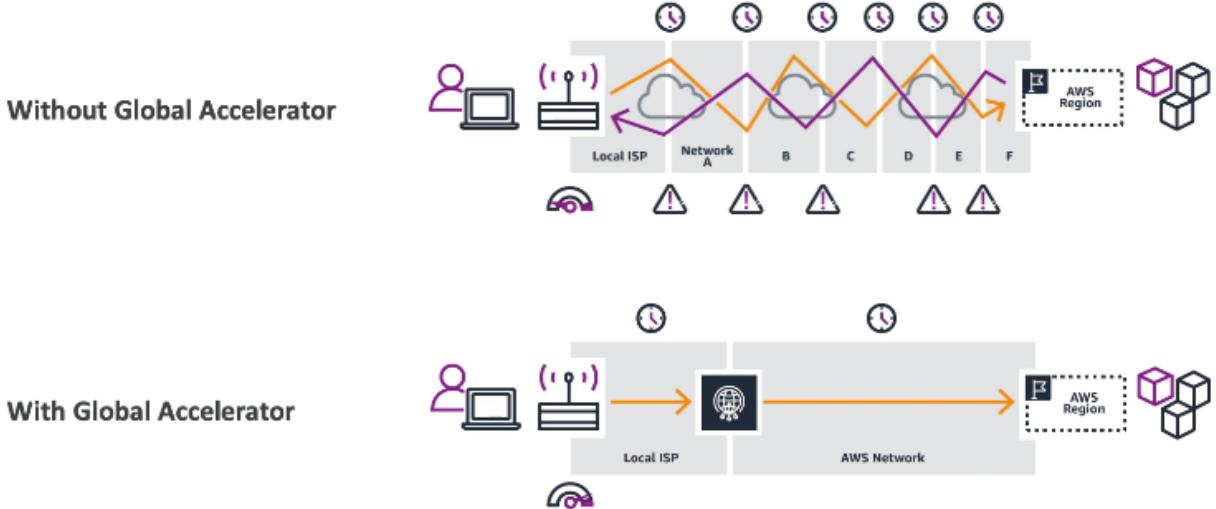


- Improve global application availability and performance using the AWS global network
- Leverage the AWS internal network to optimize the route to your application (60% improvement)



- 1.
2. Here for example..the application is in India..and users from US,AUS and UK trying to access it..
3. So to make things faster ...Edge locations will take request in the region and through aws internal network ..this edge location will access application and revert back to users in US,AUS,Uk
 - 2 Anycast IP are created for your application and traffic is sent through Edge Locations
 - The Edge locations send the traffic to your application
- 4.

AWS Global Accelerator



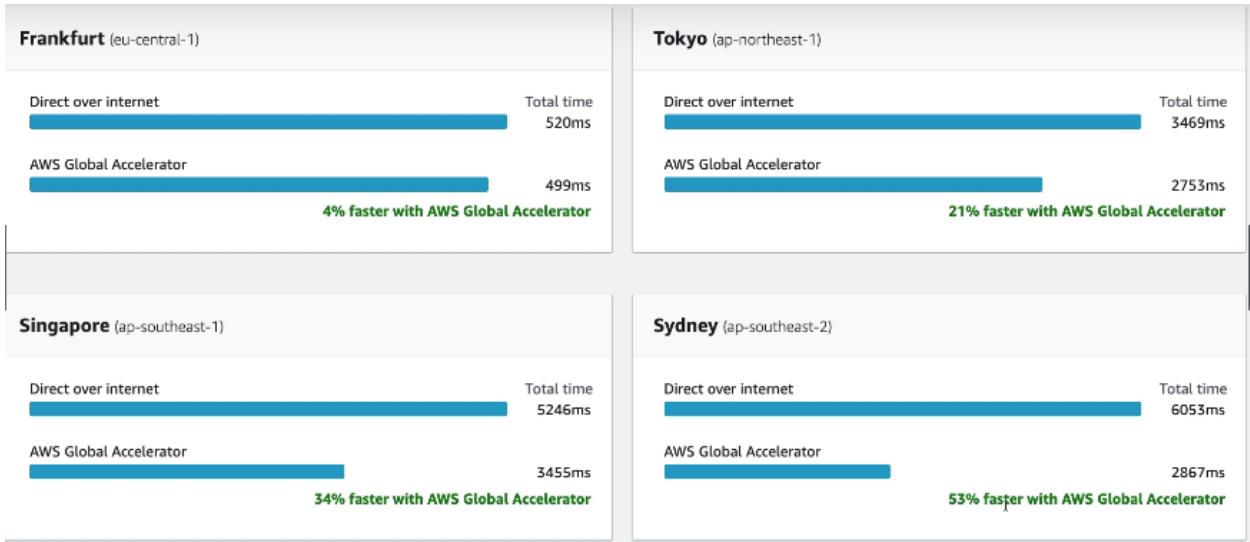
5.

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 – Content Delivery Network**
 - Improves performance for your cacheable content (such as images and videos)
 - Content is served at the edge
- **Global Accelerator**
 - No caching, proxying packets at the edge to applications running in one or more AWS Regions.
 - Improves performance for a wide range of applications over TCP or UDP
 - Good for HTTP use cases that require static IP addresses
 - Good for HTTP use cases that required deterministic, fast regional failover

6.

7. Speed comparisons using global accelerators



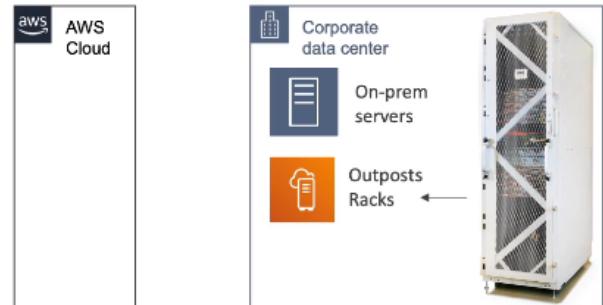
8.

AWS Outposts

AWS Outposts



- **Hybrid Cloud:** businesses that keep an on-premises infrastructure alongside a cloud infrastructure
- Therefore, two ways of dealing with IT systems:
 - One for the AWS cloud (using the AWS console, CLI, and AWS APIs)
 - One for their on-premises infrastructure
- **AWS Outposts** are “server racks” that offers the same AWS infrastructure, services, APIs & tools to build your own applications on-premises just as in the cloud
- **AWS will setup and manage “Outposts Racks”** within your on-premises infrastructure and you can start leveraging AWS services on-premises



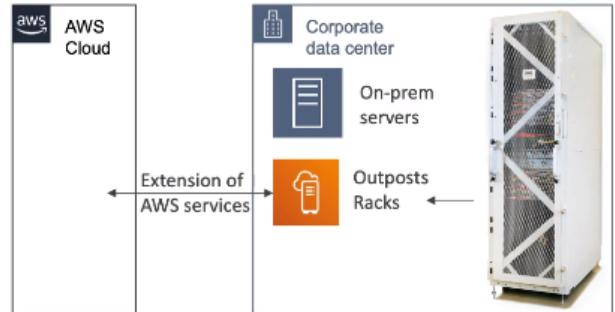
1.

- 1.
2. But a difference between an EC2 instance running on the cloud and an EC2 instance running in your own data center is that now you are responsible for their security, the physical security of the rack itself because that rack is within your own data center.

AWS Outposts



- Hybrid Cloud: businesses that keep an on-premises infrastructure alongside a cloud infrastructure
- Therefore, two ways of dealing with IT systems:
 - One for the AWS cloud (using the AWS console, CLI, and AWS APIs)
 - One for their on-premises infrastructure
- AWS Outposts are “server racks” that offers the same AWS infrastructure, services, APIs & tools to build your own applications on-premises just as in the cloud
- AWS will setup and manage “Outposts Racks” within your on-premises infrastructure and you can start leveraging AWS services on-premises
- You are responsible for the Outposts Rack physical security



3.

AWS Outposts



- Benefits:
 - Low-latency access to on-premises systems
 - Local data processing
 - Data residency
 - Easier migration from on-premises to the cloud
 - Fully managed service
- Some services that work on Outposts:



Amazon EC2



Amazon EBS



Amazon S3



Amazon EKS



Amazon ECS



Amazon RDS



Amazon EMR

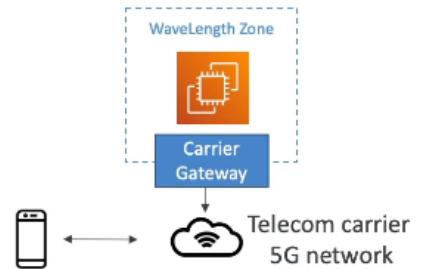
4.

AWS WaveLength

AWS WaveLength



- WaveLength Zones are infrastructure deployments embedded within the telecommunications providers' datacenters at the edge of the 5G networks
- Brings AWS services to the edge of the 5G networks
- Example: EC2, EBS, VPC...
- Ultra-low latency applications through 5G networks
- Traffic doesn't leave the Communication Service Provider's (CSP) network

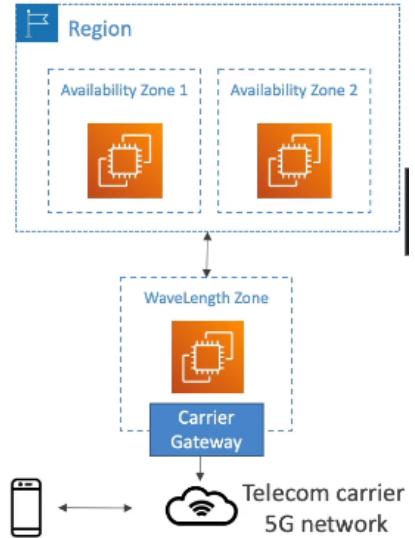


1. (This step is highlighted with a yellow bar)
2. Lets suppose there an telecom carrier network ..they have their own wavelength and bandwidth ..Here we can create an ec2 instance..which is controlled by the telecom network
3. If we want secure connection and more bandwidth..we can connect it to..nearest aws region

AWS WaveLength



- WaveLength Zones are infrastructure deployments embedded within the telecommunications providers' datacenters at the edge of the 5G networks
- Brings AWS services to the edge of the 5G networks
- Example: EC2, EBS, VPC...
- Ultra-low latency applications through 5G networks
- Traffic doesn't leave the Communication Service Provider's (CSP) network
- High-bandwidth and secure connection to the parent AWS Region
- No additional charges or service agreements
- Use cases: Smart Cities, ML-assisted diagnostics, Connected Vehicles, Interactive Live Video Streams, AR/VR, Real-time Gaming, ...



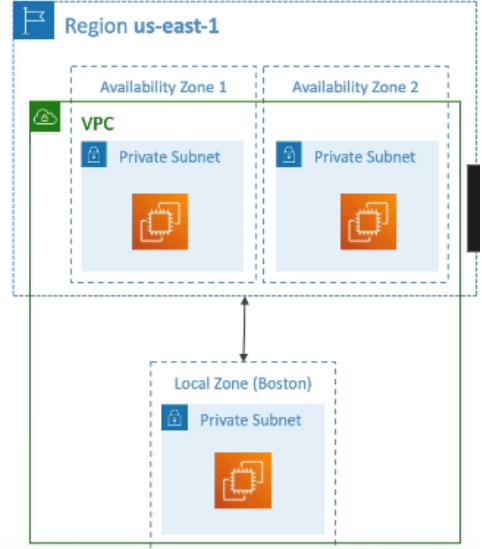
4. (This step is highlighted with a yellow bar)

AWS Local Zones

AWS Local Zones



- Places AWS compute, storage, database, and other selected AWS services closer to end users to run latency-sensitive applications
- Extend your VPC to more locations – “Extension of an AWS Region”
- Compatible with EC2, RDS, ECS, EBS, ElastiCache, Direct Connect ...
- Example:
 - AWS Region: N.Virginia (us-east-1)
 - AWS Local Zones: Boston, Chicago, Dallas, Houston, Miami, ...



- 1.
2. Here we have a region with 2 available zones
3. So how does that work? Well, here's your region,
4. US-east-1 with two AZs, but it's possible for you to define a local zone in Boston and then extend your VPC to be going across these AZs and local zones. And then you would be able to launch an EC2 instance into this local zone.
5. If we go to zone settings of N.virginia..we have many locals zones available...like boston,chicago etc
6. So if we have more users in chicago..we can enable our local chicago zone..which provides low latency to users

Global Applications Architecture

Global Applications Architecture

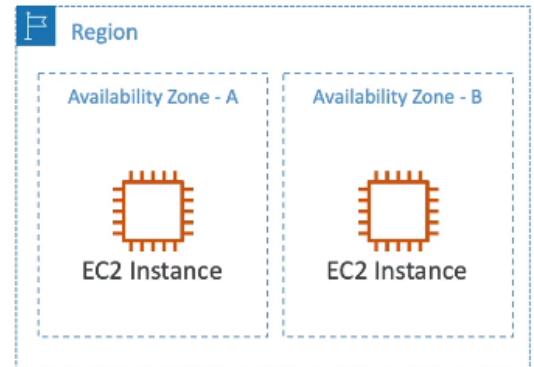
Single Region, Single AZ

- ✖ High Availability
- ✖ Global Latency
- ⌚ Difficulty



Single Region, Multi AZ

- ✓ High Availability
- ✖ Global Latency
- ⌚ Difficulty

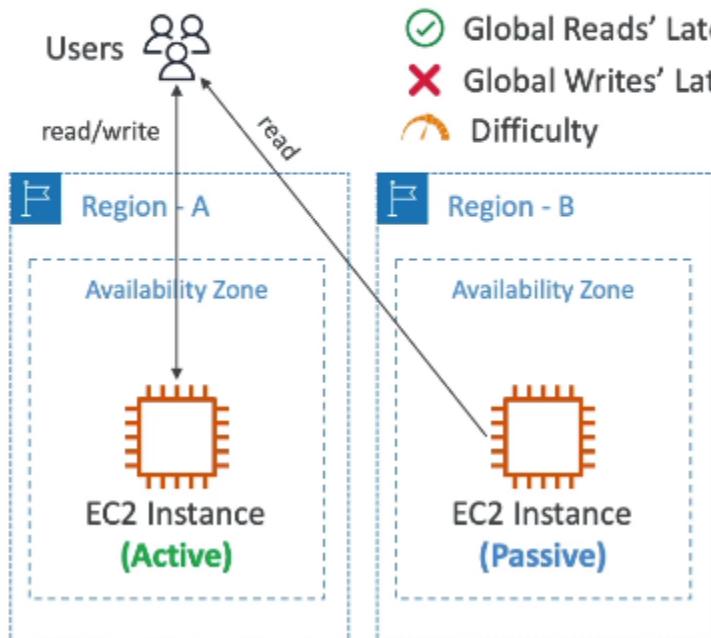


- 1.
2. Next we have multi region

Global Applications Architecture

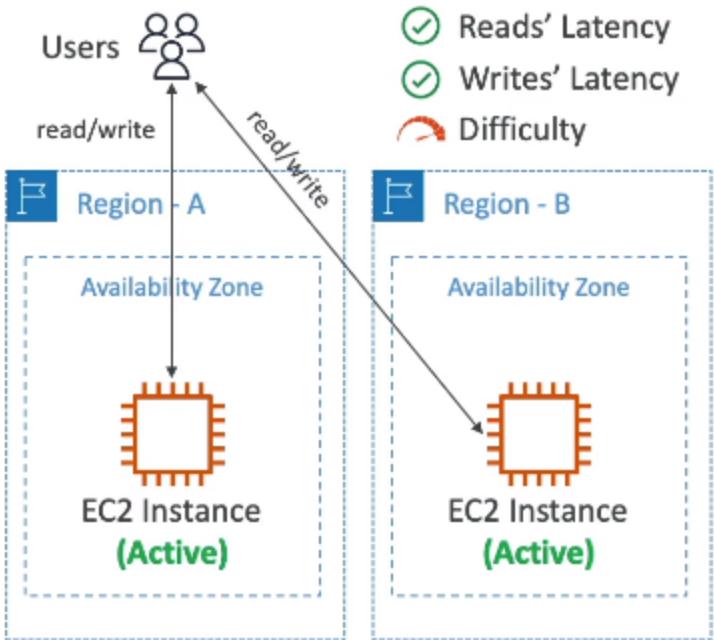
Multi Region, Active-Passive

- ✓ Global Reads' Latency
- ✗ Global Writes' Latency
- ⌚ Difficulty



- 3.

Multi Region, Active-Active



4.

Summary

Global Applications in AWS - Summary



- Global DNS: Route 53
 - Great to route users to the closest deployment with least latency
 - Great for disaster recovery strategies



- Global Content Delivery Network (CDN): CloudFront
 - Replicate part of your application to AWS Edge Locations – decrease latency
 - Cache common requests – improved user experience and decreased latency



- S3 Transfer Acceleration
 - Accelerate global uploads & downloads into Amazon S3



- AWS Global Accelerator
 - Improve global application availability and performance using the AWS global network

1.

Global Applications in AWS - Summary



- AWS Outposts
 - Deploy Outposts Racks in your own Data Centers to extend AWS services
- AWS WaveLength
 - Brings AWS services to the edge of the 5G networks
 - Ultra-low latency applications
- AWS Local Zones
 - Bring AWS resources (compute, database, storage, ...) closer to your users
 - Good for latency-sensitive applications

2.