Caching

>Overview

Amazon ElastiCache is a web service that makes it easy to deploy, operate, and scale an in-memory data store and cache in the cloud.

Caching allows you to efficiently reuse previously retrieved or computed data.

This improves the performance of web applications by allowing you to retrieve information from fast, managed, in-memory data stores, instead of relying entirely on slower disk-based database

We can cache internal and external

- Internal
    - Caching is an important tool we can use to speed up our databases, the less we talk to the database, the better
- External
    - We can cache data that's going to be returned to our users

There are 4 different caching options

- CloudFront
- ElastiCache
- DAX
- Global Accelerator

>Caching with CloudFront

- Content Delivery (Cloud Front)
    - CloudFront is Amazon's content delivery network that you can use to serve cached copies of your content to remote users at low latency rates.
    - Amazon CloudFront is the content delivery network (CDN) you can use to lower access latency for your application users by storing cached versions/content of frequently requested data at AWS edge locations.
    - Amazon CloudFront is a content delivery network (CDN) that helps deliver static and dynamic web content to users faster than just serving it out of an AWS Region.
    - CloudFront caches objects in edge locations around the world and automatically directs users to the edge location that will give them the best performance at any given time.
    - Remember with edge locations you're not provisioning EC2 Instances
    - Origin
        - Origin is the original source of the content
    - CloudFront used in the real world
        - S3 Static websites
            - Used with S3 to deploy content globally
        - Prevent attacks
            - CloudFront can stop certain web attacks, like DDoS
        - IP Address blocking
            - Prevents users certain countries from access content

    - The farther a user is away from that region, the more network latency they'll encounter when accessing it. CloudFront solves this problem by caching your content in a number of data centers called *edge locations*. There are more than 150 edge locations spread out around the world on six continents.

You can deploy AWS WAF on Amazon CloudFront as part of your CDN solution, the Application Load Balancer that fronts your web servers or origin servers running on EC2, Amazon API Gateway for your REST APIs, or AWS AppSync for your GraphQL APIs.

- Important CloudFront settings
    - Security – Defaults to HTTPS connections with the ability to add custom SSL cert
- Global Distribution
    - Pick general areas of the globe
        - Also allow or deny list for particular countries but if you do have to block or allow particular countries… Don't use CloudFront you want to use the Web Application Firewall (WAF)

CloudFront supports AWS endpoints

CloudFront is the only option to add HTTPS to a static website being hosted in a S3 Bucket

>Caching data with ElasticCache and DAX

Amazon ElastiCache is a fully managed, in-memory caching service supporting flexible, real-time use cases.

Memcached vs. Redis

https://aws.amazon.com/elasticache/redis-vs-memcached/

- Memcached - Memcached is an easy-to-use, high-performance, in-memory data store. It offers a mature, scalable, open-source solution for delivering sub-millisecond response times making it useful as a cache or session store. Memcached is a popular choice for powering real-time applications in Web, Mobile Apps, Gaming, Ad-Tech, and E-Commerce.
    - https://aws.amazon.com/memcached/
- Redis - Redis, which stands for Remote Dictionary Server, is a fast, open source, in-memory, key-value data store. All Redis data resides in memory, which enables low latency and high throughput data access.
    - https://aws.amazon.com/redis/

DynamoDB Accelerator (DAX)

Amazon DynamoDB Accelerator (DAX) is a fully managed, highly available, in-memory cache for Amazon DynamoDB that delivers up to a 10 times performance improvement—from milliseconds to microseconds—even at millions of requests per second.

DAX is designed to run within an Amazon Virtual Private Cloud (Amazon VPC) environment. Amazon VPC defines a virtual network that closely resembles a traditional data center.

>Fixing IP Caching with Global Accelerator

AWS Global Accelerator is a networking service that improves the performance of your users' traffic by up to 60% using Amazon Web Services' global network infrastructure. When the internet is congested, AWS Global Accelerator optimizes the path to your application to keep packet loss, jitter, and latency consistently low.

AWS Global Accelerator helps bridge the gap between single and multiple Region deployments by improving the network routing for local and global user traffic. If traffic to your application's single Region is left on the public internet, it can be negatively impacted by internet congestion and local outages.

Global Accelerator allows you to create weights for your application endpoints (As well as Route 53)

Exam Tips

1. Caches
    a. Pick an answer that includes a caching solution. Which AWS loves to put caches in every possible location
    b. Latency is low, caches go in front of everything
2. CloudFront
    a. Fixes all connection issues
    b. Use CloudFront to cache your content externally
    c. Fast, hence the cache content to speed up delivery of data
    d. CDN works for both AWS and on-site
    e. Blocks connections for certain countries, but WAF is a better tool for it
3. Caching data with ElastiCache and DAX
    a. On the exam favor answers that include a database caching solution
        i. Which you need to know when each one is applicable. Generally applied in about every situation.
    b. Redis can be more than just a cache – it can be a standalone database as well
        i. Support from mult-AZ, failover and backups to where the other solutions don't
        ii. Has more features than Memcached
4. Global Accelerator
    a. Any scenario-based questions about IP caching think about Global Accelerator
    b. Uses edge locations to help speed everything up
    c. 2 Static IPs – Which you're provided 2 IPs that don't change, also bring your own.

Backups are not supported on any solutions besides Redis. Memcached and DAX are not a source of truth for your data.