*Secrets Manager*

1. Programmatically retrieve encrypted secret values at runtime (Avoid hardcode)

2. Store different types of secrets

3. Encrypt secret data with KMS (encrypt at rest)

4. Automatically rotate your secrets

5. Control access with IAM

*AuroraDB - Fastest database*

1. Relates to MySQL or PostgreSQL

*DynamoDB*

- Read/write Capacity Mode

Read Request Unit (handle read request per second)

1 RRU - 1 strongly consistent read or 2 eventually consistent read for 4KB item

2 RRU - 1 Transactional read for 4KB item

Write Request Unit (handle write request per second)

1 WRU - 1KB item

2 WRU - Transactional write for 1KB item

- Data Distributed by Partition Key (primary key), sorted by Sort Key

- Global table - deploying a multi-region, multi-master database

- DynamoDB accelerator - a compatible caching service, speed up access time for eventually consistent data.

- monitored by CloudWatch

- overcome ProvisionedThroughputExceededException error with exponential backoff

- For Query operation, specify the ReturnConsumedCapacity parameter with

* NONE—non return (This is the default.)
* TOTAL—return number of read capacity units consumed.
* INDEXES—return read capacity units consumed in details

*ElasticCache - putting before database layer*

- Strategies

- Lazy Loading - loads data into the cache only when necessary

- Write-through - adds data or updates data in the cache whenever data is written to the database

- Adding TTL - adding a TTL value to each write (combine TTL with Lazy Loading, write-through )

*CodeBuild*

- [Error Pattern](https://docs.aws.amazon.com/zh_cn/codebuild/latest/userguide/troubleshooting.html" \l "troubleshooting-dependency-caching)

*Kinesis Data Analytics*

- analyze streaming data using standard SQL to

- Generate time-series Analytics

- Feed real-time dashboards

- Create real-time metrics

*Kinesis Data Firehose -* delivering real-time streaming data to AWS services or third-party services

Amazon Kinesis Data Streams

- Reading Data from Amazon Kinesis Data Streams

|  |  |  |
| --- | --- | --- |
| Characteristics | without Enhanced Fan-Out | with Enhanced Fan-Out |
| Shard Read Throughput | Fixed at a total of 2 MB/sec per shard | up to 2 MB/sec with each consumer |

*Layer* -contain additional code, libraries, configuration files or data

*X-ray*

- trace requests through API Gateway REST APIs to the underlying services

- daemon

- Install and run the X-ray daemon instance

- Giving the daemon permission to send data to X-Ray

抽样公式: Reservoir Size + Fixed Rate % [(Total Requests – Reservoir Size)]

默认比例: One request per second & five percent of any additional request per host.

*Cloudformation* - setup model with AWS resources

- Features

- Simplify infrastructure management

- Quickly replicate your infrastructure

- Easily control and track changes to your infrastructure

- Concepts

- Templates

- save as .json, .yaml, .template, .txt file

- use as as blueprints for building your AWS resources

- stacksets - enable crud stack across multiple accounts

- In yaml / json file, Parameters have different types:

- String

- Number

- List

- CommaDelimitedList

- AWS-Specific Parameter Types

- SSM Parameter Types

- Best Practices

- Do not embed credentials in your templates, use Parameter (dynamic reference) instead

- Use IAM to control access

- Use AWS CloudTrail to log AWS CloudFormation calls

*Simple Storage Service(S3)*

- Increases in HTTP 503 Responses with version Enabled bucket, check with Amazon S3 Inventory

-文件过大时，通过分段传输 multipart upload

- ensure that all objects are deleted prior to deletion of S3 buckets by DeletionPolicy of CloudFormation Template.

*API Gateway*

*- Optimize*

*- Enabling API caching to enhance responsiveness*

- multiple mechanisms for controlling and managing access to your API

- Resource policies using with IP address or VPC endpoints

- Standard AWS IAM roles and policies,IAM tags

- Lambda authorizers

- Amazon Cognito user pools

- API keys

*Elastic beanstalk* - deploy and manage applications in the AWS Cloud without caring infrastructure

- Auto Scaling

- Modifying launch configuration, environment manifest files to change the instance type

- adding RDS to EB environment, retention options

- Snapshot — create a snapshot to backup

- Delete — Elastic Beanstalk terminates the database. the database instance is no longer available

- Retain — The database instance isn't terminated. It remains available and operational

- Deployment policy

- All at once – Deploy the new version to all instances simultaneously. All instances in your environment are out of service for a short time while the deployment occurs.

- Rolling – Deploy the new version in batches. Each batch is taken out of service during the deployment phase, reducing your environment's capacity by the number of instances in a batch.

- Rolling with additional batch – Deploy the new version in batches, but first launch a new batch of instances to ensure full capacity during the deployment process.

- Immutable – Deploy the new version to a fresh group of instances by performing an immutable update

- Traffic splitting – Deploy the new version to a fresh group of instances and temporarily split incoming client traffic between the existing application version and the new one.

- Elastic File System (EFS)

- create a shared directory that stores files uploaded or modified by users of your application

*CodeCommit* - a version control service

- simplest way to connect to CodeCommit repositories is to

- configure Git credentials in the IAM console

- use those credentials for HTTPS connections

Lambda

- Alias

- a pointer to specific function version

Elastic Cloud Service

- A management service that makes it easy to run, stop, and manage Docker containers on a cluster.

- Running Daemon On ECS

- create a Docker image that runs the X-Ray daemon

- deploy the docker container to your Amazon ECS cluster

CloudWatch

- alarm 标准是1min，如果需要1秒以内的，可以通过high resolution metrics

Instances

|  |  |
| --- | --- |
| M5 instance | support Enhance Networking  network performance with 10 Gbps to 25 Gbps  Placement Groups and NVMe EBS volumes. |
| T2 instance | Do not support Enhance Networking  network performance only up to 1 Gbps |

OpsWorks

- a configuration management service to configure and operate applications in a cloud enterprise by using Puppet or Chef.

Simple Queue Service (SQS)

- a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications.

- Using SQS, you can send, store, and receive messages between software components at any volume, without losing messages or requiring other services to be available.

Types:

- standard type

1. support a nearly unlimited number of API calls per second
2. at-least-once message delivery
3. out of order

- FIFO type

1. exactly-once processing / non-duplicated message

2. ordered messageges between software components at any volume

Short and long polling

- With short polling, thes ReceiveMessage request queries only a subset of the servers, send it back even empty

- With long polling, the ReceiveMessage request queries all of the servers for messages. send an empty response only if the polling wait time expires. The maximum long polling wait time is 20 seconds

Step Functions

- a serverless orchestration service that combine Lambda functions and other services to build applications

- Best practices

Use timeouts to avoid stuck executions

Use Amazon S3 ARNs instead of passing large payloads

Avoid Reaching the history quota

Handle Lambda service exceptions

Avoid latency when polling for activity tasks

Choosing Standard or Express Workflows

Amazon CloudWatch Logs resource policy size restrictions

杂项

通过 Cache-Control: max-age=0 header, 确保个别请求越过缓存，获取最新数据

On Premise server 不能放置在EC2上, 不能添加到 IAM group中

Cloudwatch alarm 不会被log stream触发，会被metric触发

CloudWatch event 会识别各种system events （例如EC2重启和state changes）

In-place 不跟 lambda和 ECS合作

Canary和Linear不是ECS的部署方式 how u shift ur traffic

Blue/green 是 ecs的部署方式

.config后缀的文件都是在.ebextensions 文件夹里

Cors header: Access-Control-Allow-Headers / Origin

403报错信息，可能是加密过后的，需要用 \*\*decode-authorization-message\*\*

Serverless Aplication model 简称 SAM,有自己的SAM CLI负责打包和部署

当accountA管理资源,accountB管理管道，需要KMS和IAM的授权才能处理好

ElasticBeanstalk needs a lifecyle policy to avoid hitting the version limit

SQS queue default visibility timeout is 30 seconds.

Lambda functions failed asynchronously, could use dead letter queue to check the reason