AWS Other Services

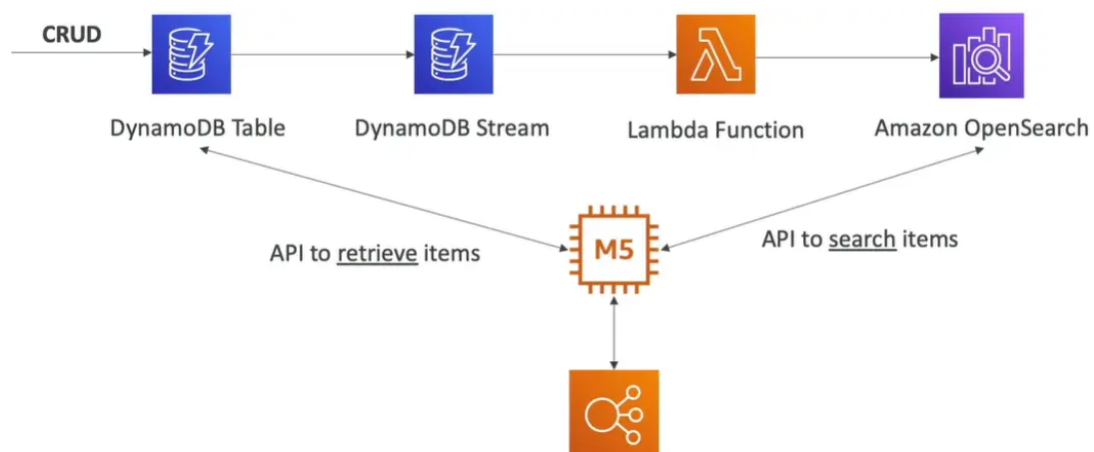## AWS SES – Simple Email Service

- send emails to people using
  - SMTP interface
  - or AWS SDK
- ability to receive email.Integrate with
  - S3
  - SNS
  - Lambda
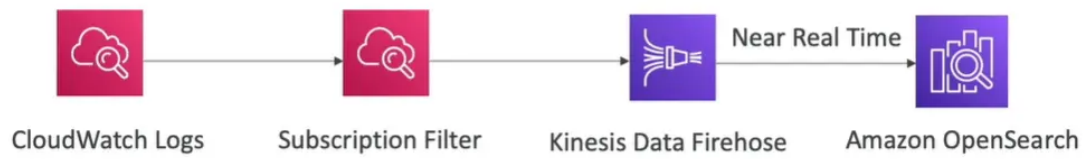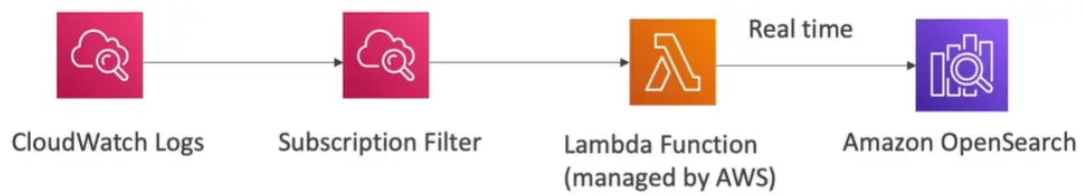- Integrated with IAM for allowing to send emails

Amazon OpenSearch Service

- Amazon OpenSearch is successor to Amazon ElasticSearch
- In DynamoDB, queries only exist by primary key or indexes...
- With OpenSearch, you can search an field, even partially matches
- it's common to use OpenSearch as a complement [补充] to another database
- two modes : managed cluster or serverless cluster
- does not natively support SQL (can be enabled via a plugin)
- ingestion [摄入] from Kinesis Data Firehose, AWS IoT, and CloudWatch Logs
- security through Cognito & IAM, KMS encryption, TLS
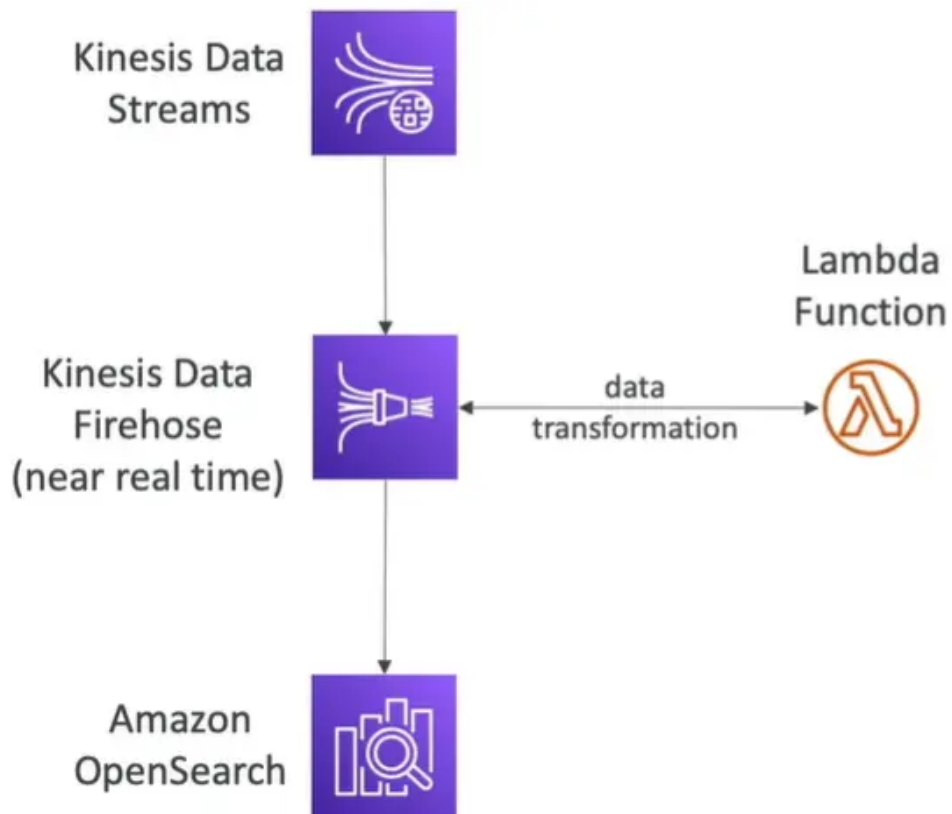- comes with OpenSearch Dashboards (visualization)
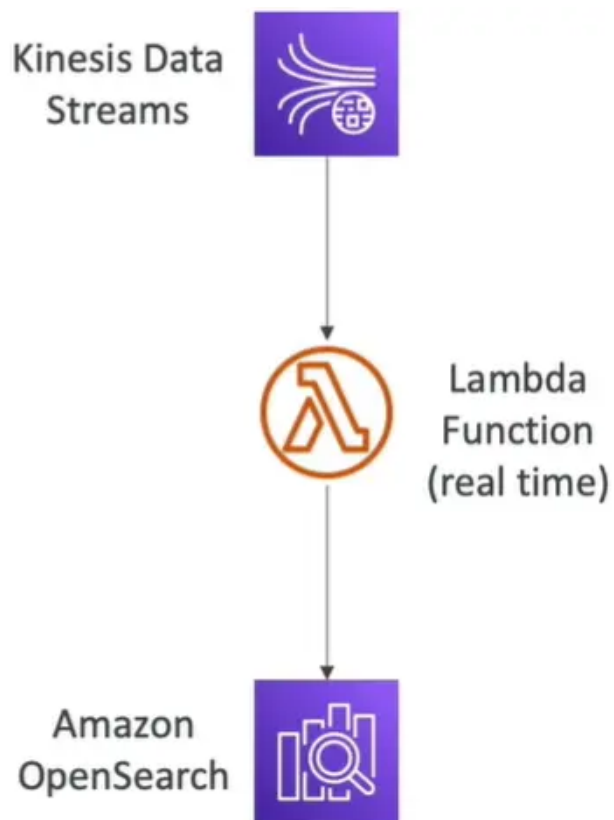
Open Search patterns DynamoDB

OpenSearch patterns CloudWatch Logs

CloudWatch Logs → Subscription Filter → Lambda Function (managed by AWS) — Real time → Amazon OpenSearch

CloudWatch Logs → Subscription Filter → Kinesis Data Firehose — Near Real Time → Amazon OpenSearch

OpenSearch patterns Kinesis Data Stream & Kinesis Data Firehose

Kinesis Data Streams → Kinesis Data Firehose (near real time) ← data transformation → Lambda Function → Amazon OpenSearch

Kinesis Data Streams

Lambda Function (real time)

Amazon OpenSearch

Amazon Athena

load data

S3 Bucket

Query & Analyze

Amazon Athena

Reporting & Dashboards

Amazon QuickSight

- **serverless** query service to analyze data stored in Amazon S3
- uses standard SQL language to query the files (built on Presto)
- supports CSV, JSONj,ORC,Avro and Parquet
- Pricing: $ 5.00per TB of data scanned

- commonly used with Amazon Quicksight for reporting / dashboards
- **use cases** : business intelligence / analytics / reporting , analyze & query VPC Flow Logs, ELB logs, CloudTrail trails,etc ...
- Exam Tip : <u>analyze data in S3 using serverless SQL, use Athena</u>

Performance improvement

- use **columnar [柱状] data** for cost-savings (less scan)
  - apache parquet or ORC is recommended
  - huge performance improvement
  - use Glue to convert your data to Parquet or ORC
- **Compress data [压缩数据]**for smaller retrievals (bzip2, gzip, lz4, snappy,zlip, zstd....)
- Partition datasets in S3 for easy querying on virtual colums

  ```
  s3://yourBucket/pathToTable
          /<PARTITION_COLUMN_NAME>=<VALUE>
          /<PARTITION_COLUMN_NAME>=<VALUE>
          /<PARTITION_COLUMN_NAME>=<VALUE>
          /etc...
  ```
  - Example: s3://athena-examples/flight/parquet/year=1991/month=1/day=1/
- **Use larger files (>128MB)** to minimize overhead [开销]

Federated Query [联合查询]
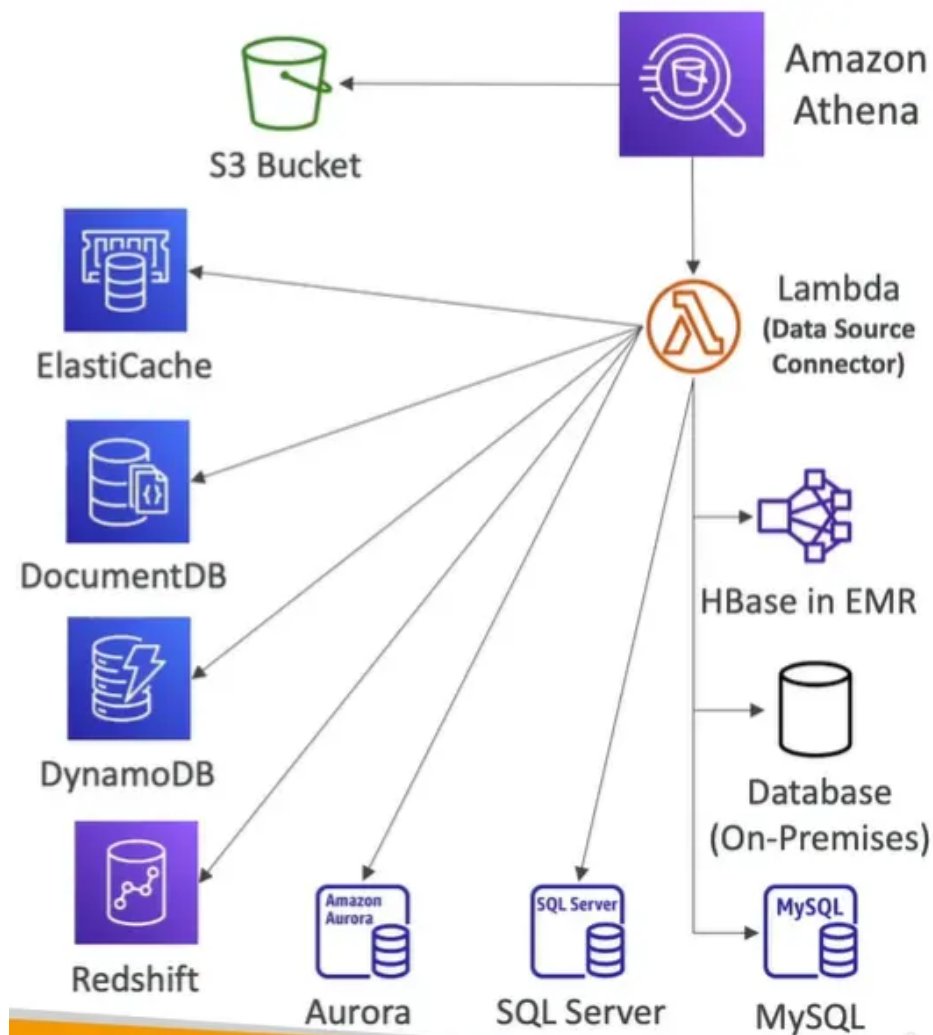
- allows you to run SQL queries across data stored in relational, non-relational, object, and custom data sources (AWS or on-premises)

- use Data Source Connectors that run on AWS Lambda to run Federated Queries (eg, CloudWatch Logs, DynamosDB, RDS,...)
- store the results back in Amazon S3

Amazon Managed Streaming for Apache Kafka (Amzon MSK)

- Alternative to Amazon Kinesis
- Fully managed Apache Kafka on AWS
  - allow you to create ,update, delete clusters
  - MSK creates & manages Kafka brokers nodes & Zookeeper nodes for you
  - Deploy the MSK cluster in your VPC, multi-AZ (up to 3 for HA)
  - Automatic recovery from common Apache Kafka failures
  - data is stored on EBS volumes for as long as you want
- MSK Serverless
  - run Apache Kafka on MSK without managing the capacity
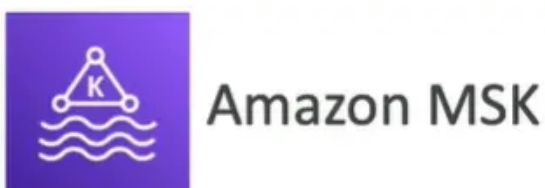 - MSK automatically provisions resources and scales compute & storage

Apache Kafka at a high level
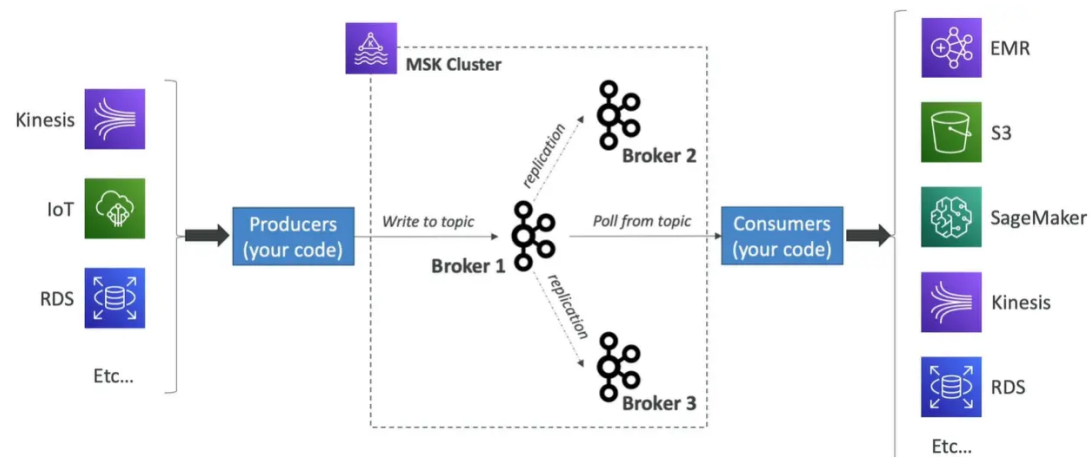
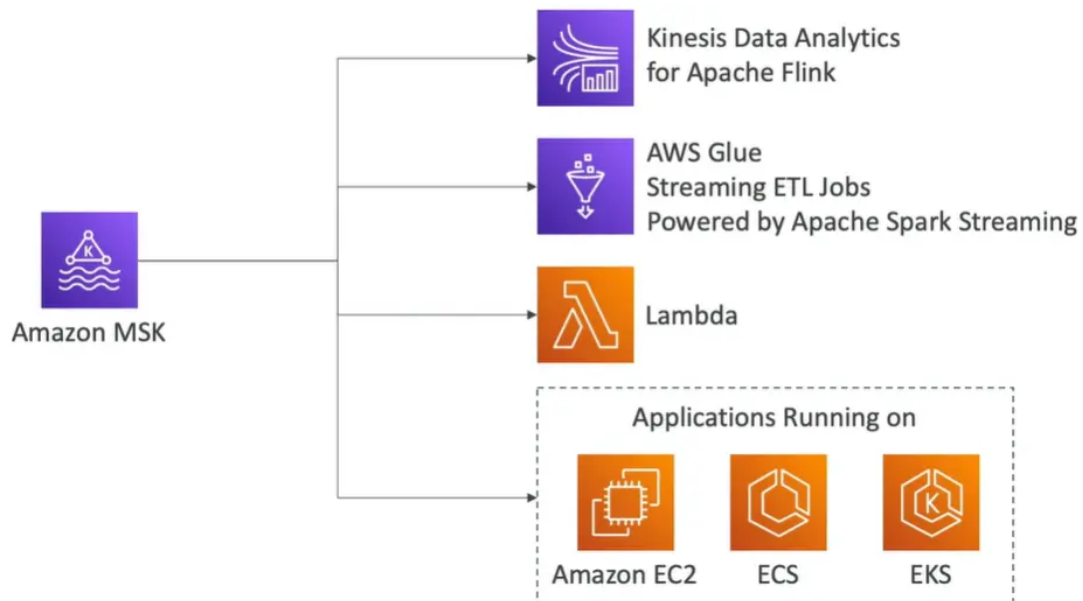## Kinesis Data Stream vs, Amazon MSK
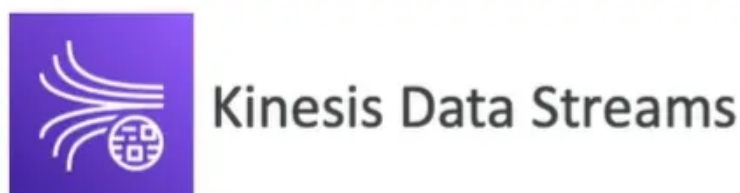
- 1MB message size limit
- data streams with shards
- shard splitting & merging
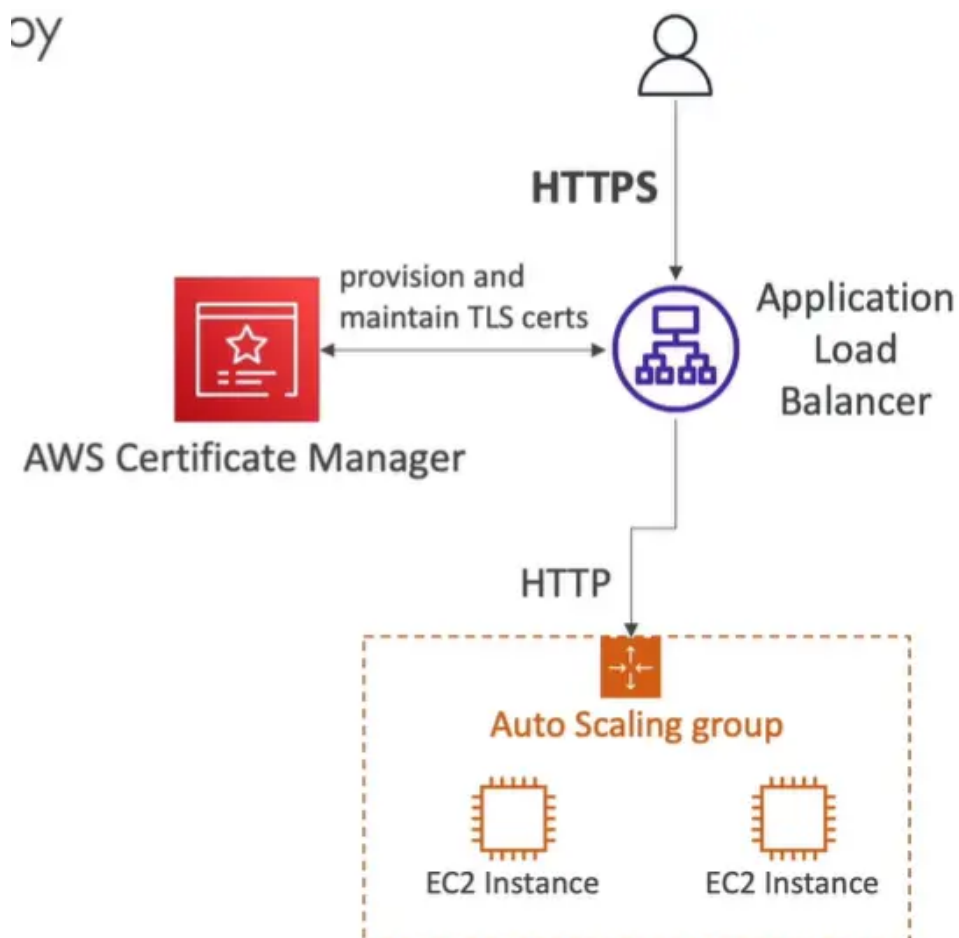- TLS in-flight encryption
- KMS at-rest encryption

- 1MB default, configure for higher (ex,10MB)
- Kafka Topics with partitions

- can only add partitions to a topic
- PLAINTEXT or TLS In–flight encryption
- KMS at–rest encryption

Amazon MSK Consumers

AWS Certificate Manager (ACM)

- let'a you easily provision, manage , and deploy SSL / TLS Certificates
- used to provide in–filght encryption for websites (HTTPS)
- supports both public and private TLS certificates
- free of charge for public TLS certificates

- Automatic TLS certificate renewal
- integrations with (load TLS certificate on)
    - Elastic Load Balancers
   - CloudFront Distributions
    - APIs on API Gateway

## AWS Private Certificate Authority (CA)
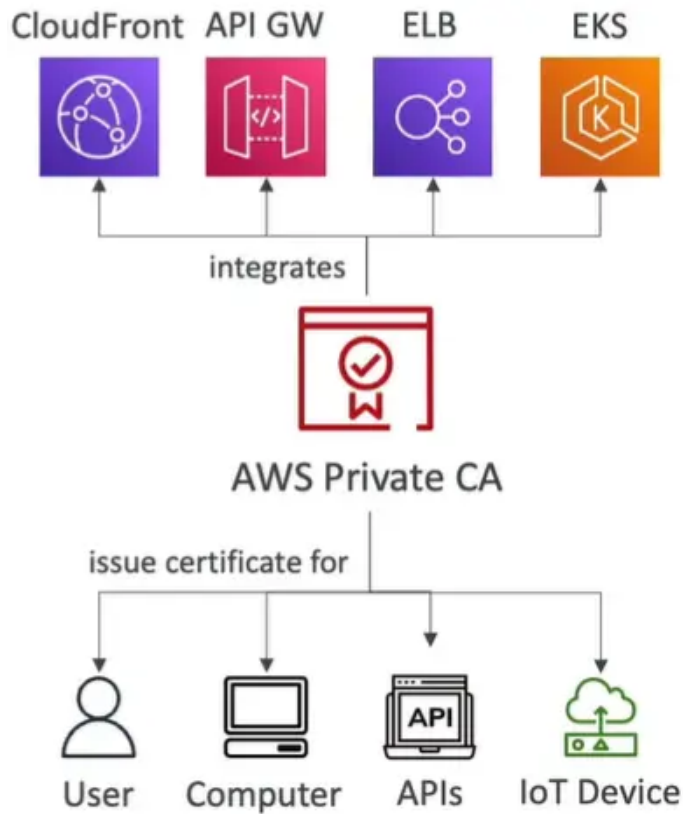
- managed service allows you to create private certificate Authorities (CA), including root and subordinaries CAs
- can issue and deploy end–entity X.509 certificates
- certificates are trusted only by your Organization (not the public internet)
- works for AWS service that are integrated with ACM
- use cases:
    - encrypted TLS communication, Cryptographically signing code
    - authenticate users, computers, API endpoints, and IoT devices
    - enterprise customers building a public key infrastructure (PKI)

Amazon Macie

- Amazon Macie is a fully managed data security and data privacy service that uses **machine learning and pattern matching to discover and protect your sensitive data in AWS**
- Macie help identify and alert you to **sensitive data, such as personally identifiable information (PII)**
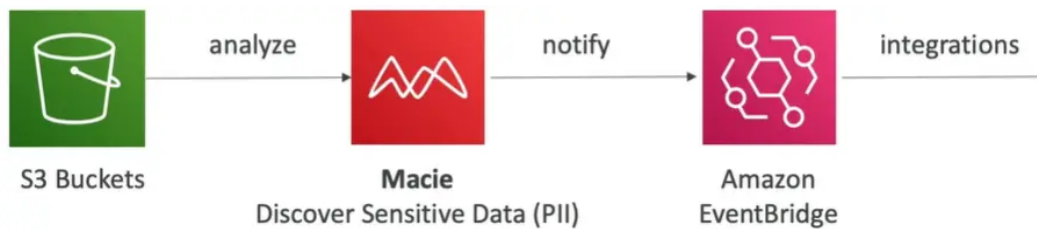
AWS AppConfig
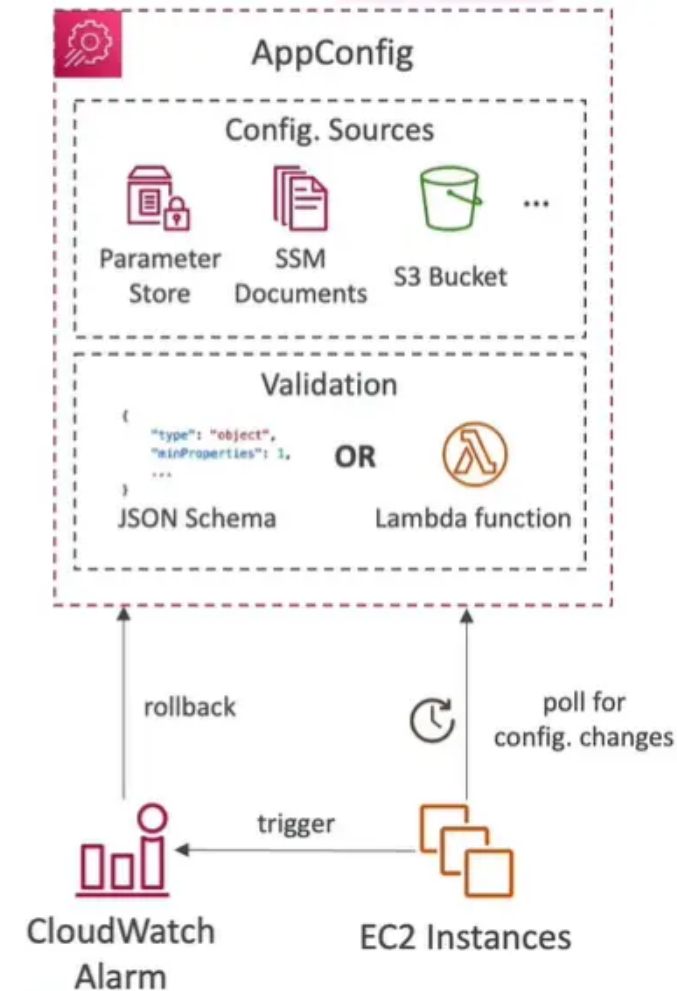
- configure, validate, and deploy dynamic configurations
- deploy dynamic configuration changes to your applications independently of any code deployments
  - you don't need to restart the application
- feature flags ,application tuning, allow/block listing..
- use with apps on EC2 instances, Lambda, ECS, EKS..
- gradually deploy the configuration changes and rollback if issues occur
- validate configuration changes before deployment using
  - JSON Schema (syntactic check) or
  - Lambda Function – run code to perform validation (semantic check)

CloudWatch Evidently
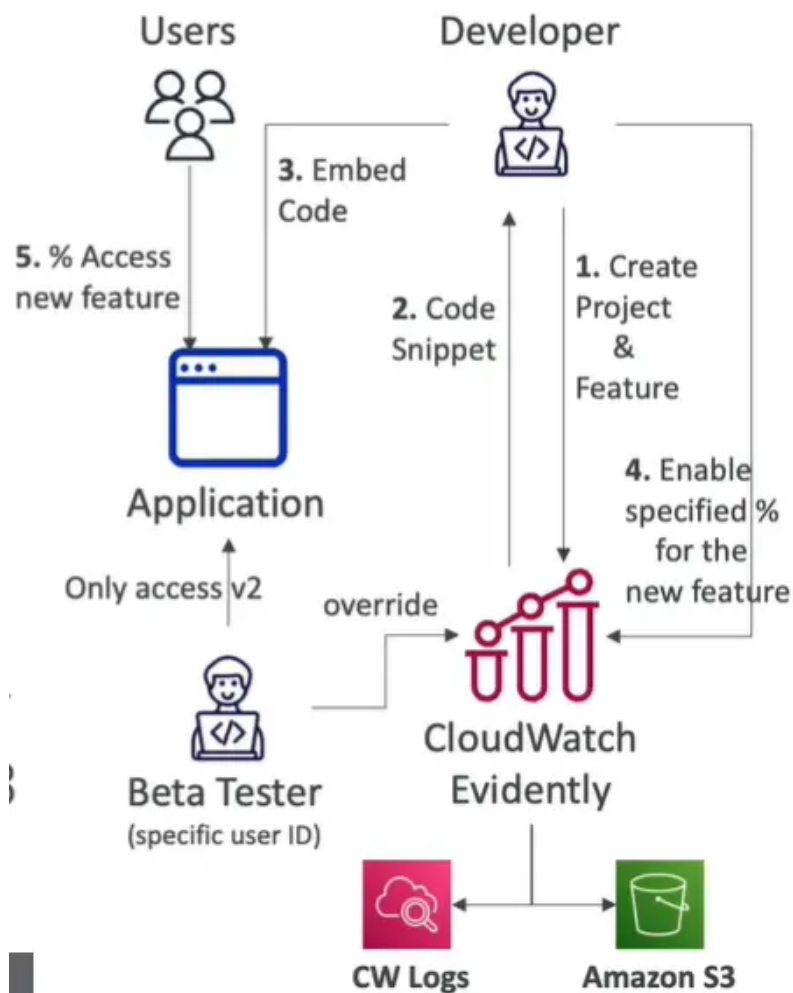
- safely validate new features by serving them to a specified % of your usrs
  - reduce risk and identify unintended consequences
  - collect experiment data, analyze using stats, monitor performance
- Launches (= feature flags) : enable and disable features for a subset of users
- experiments (= A/B testing) : compare multiple versions of the same feature
- overrides : pre–define a variation for a specific usre
- store evaluation events in CloudWatch Logs or S3