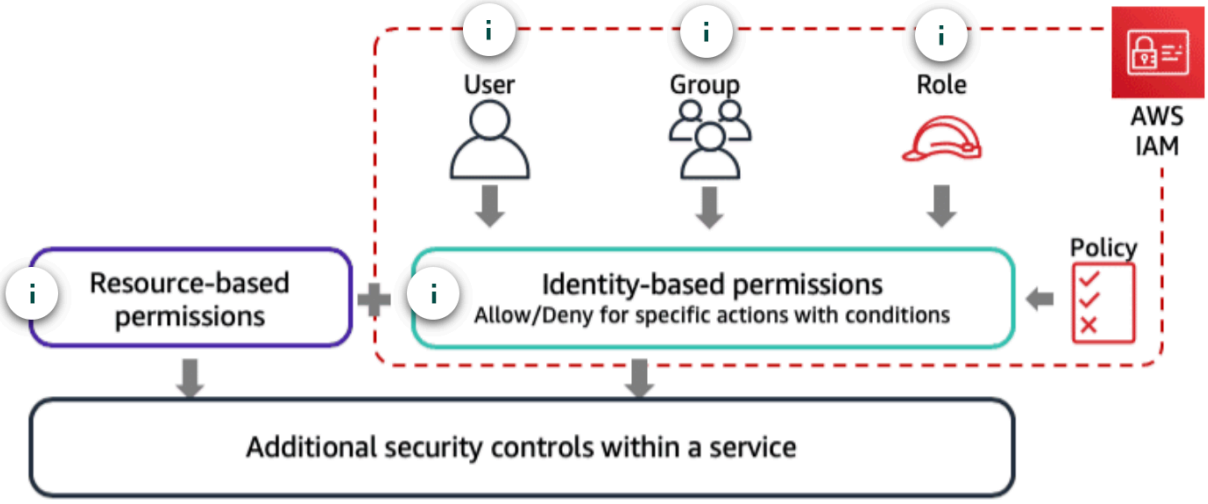
AWS DAS-C01 readiness domain5

Wednesday, 26 May 2021 7:37 PM

AWS Identity and Access Management (IAM)

IAM allows you to authenticate yourself through users, groups, and roles.



Identity-based resources are attached to an IAM entity and indicate what the user or group is permitted to do.
Resource-based permissions are attached to a resource and indicate what a specified user (or group of users) is permitted to do with it.

Identity based policy

The example below is a policy attached to a user who is attempting to work with Amazon EMR. In the policy, the StringEquals conditional operator tries to match the department "dev" with the value for the tag department. If the tag department hasn't been added to the Amazon EMR cluster, or doesn't contain the value "dev", the policy doesn't apply, and the actions aren't allowed by this policy. If no other policy statements allow the actions, the user can only work with Amazon EMR clusters that have this tag with this value.

```
"Effect": "Allow",
"Action": [
    "elasticmapreduce:DescribeCluster",
   "elasticmapreduce:ListSteps",
    "elasticmapreduce:TerminateJobFlows",
```

```
    "elasticmapreduce:SetTerminationProtection",
    "elasticmapreduce:ListInstances",
    "elasticmapreduce:ListInstanceGroups",
    "elasticmapreduce:ListBootstrapActions",
    "elasticmapreduce:DescribeStep"
],
"Resource": [
    "*"
 ],
"Condition": {
    "StringEquals": {
        "elasticmapreduce:ResourceTag/department": "dev"
    }
```

Resource-base policy

The example below is a policy attached to an Amazon S3 bucket. This policy grants permissions to any user to perform any Amazon S3 operations on objects in the specified Amazon S3 bucket. However, the request must not originate from the range of IP addresses specified in the condition. The condition in this statement identifies the 54.240.143.* as the range of allowed IPv4 IP addresses.

```
"Version": "2012-10-17",
"Id": "S3PolicyId1",
"Statement": [
  {
      "Sid": "IPAllow",
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::examplebucket/*",
      "Condition": {
  "NotIpAddress": {"aws:SourceIp":"54.240.143.0/24"}
       }
   }
]
```

IAM with Amazon redshift

- Amazon Redshift **does not support resource-based policies**.
- Amazon Redshift provides service-linked roles, which include all the permissions that the service needs to call other AWS services on behalf of your Amazon Redshift cluster
- As an alternative to maintaining user names and passwords in your Amazon Redshift database, you can configure the system to permit

users to use IAM credentials to log on to the database.

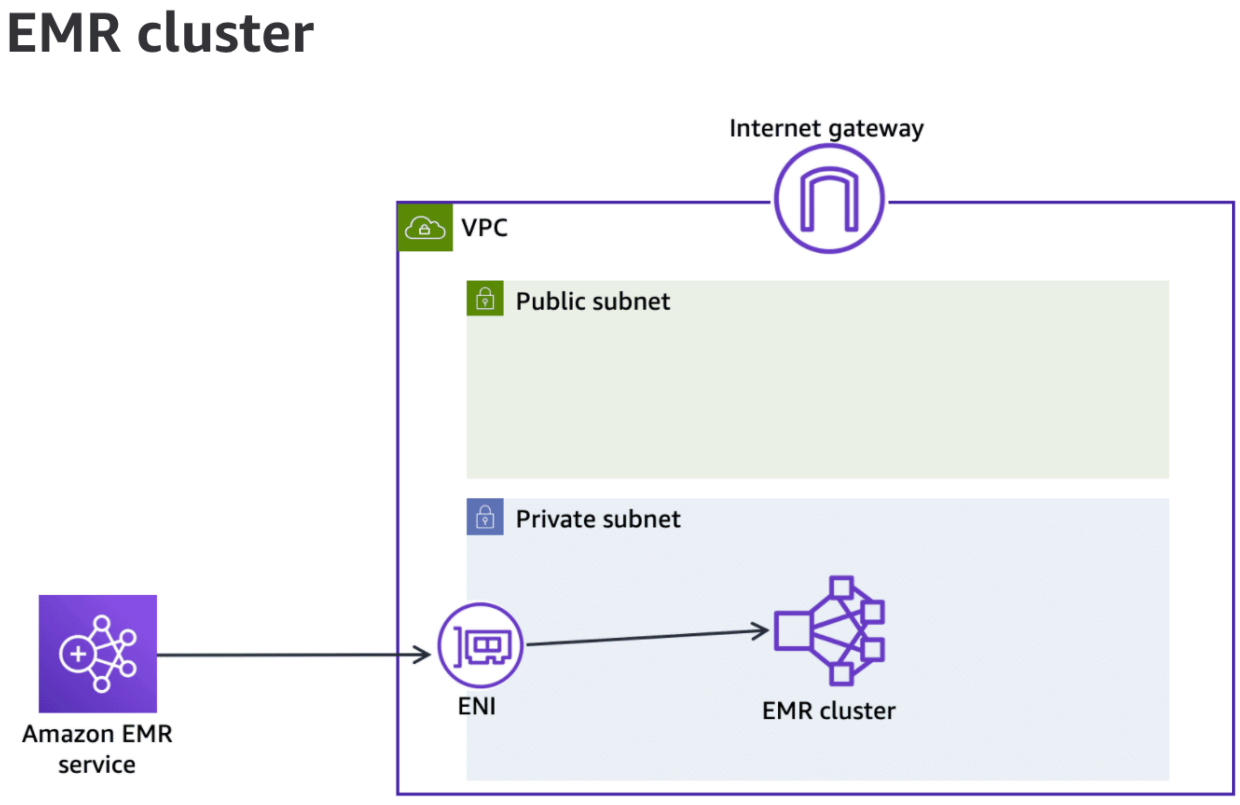
- You can configure your system to let users access the database by using federated single-sign-on (SSO) through a SAML 2.0 identity provider.
- Amazon Redshift assumes an IAM role when data is loaded into a cluster using the COPY command or exported data from a cluster using the UNLOAD command. This eliminates the need to embed AWS access credentials within SQL commands.

IAM with Amazon DynamoDB

- Restrict access at the **identity level**. For example, you can grant permissions to users, groups, or roles to provide read-only access or write-only access to certain items or attributes in a table or secondary index.
- Restrict access to **individual data items and attributes**. For example, you can grant permissions to a table, but restrict access to **specific items based on certain primary key values**.

Network security and securing the physical boundary

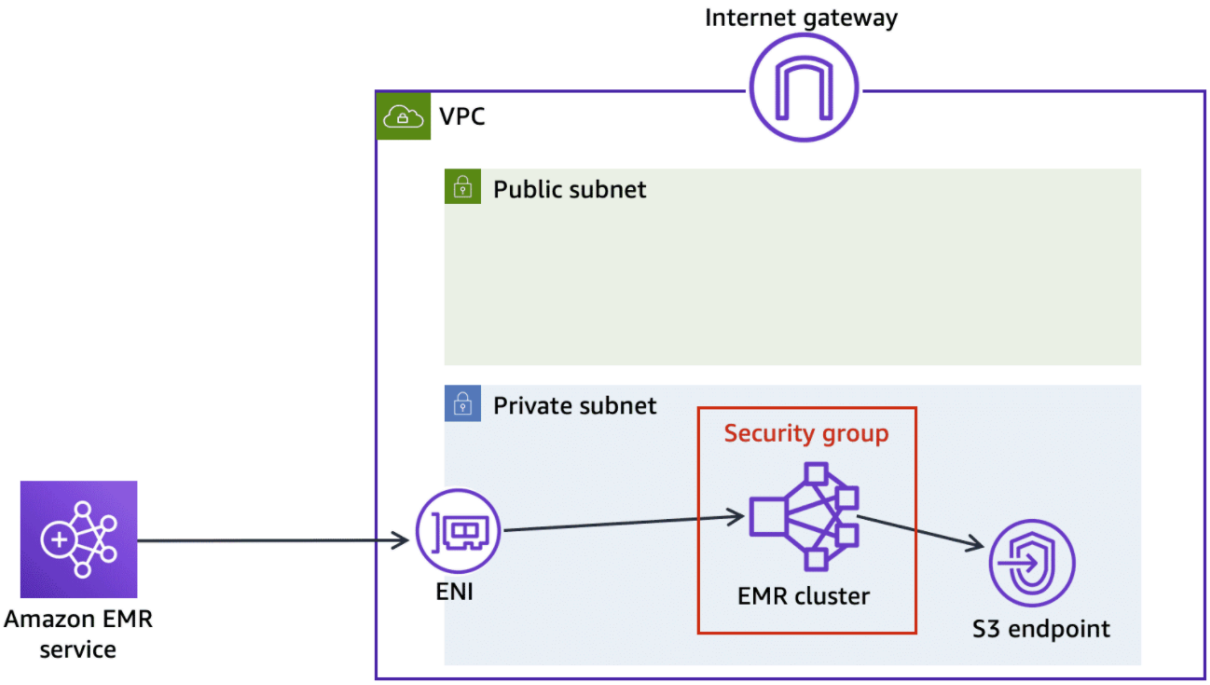
To properly isolate your resources, you will need to configure an Amazon Virtual Private Cloud (Amazon VPC).



When you launch an Amazon EMR cluster, you must specify one Amazon

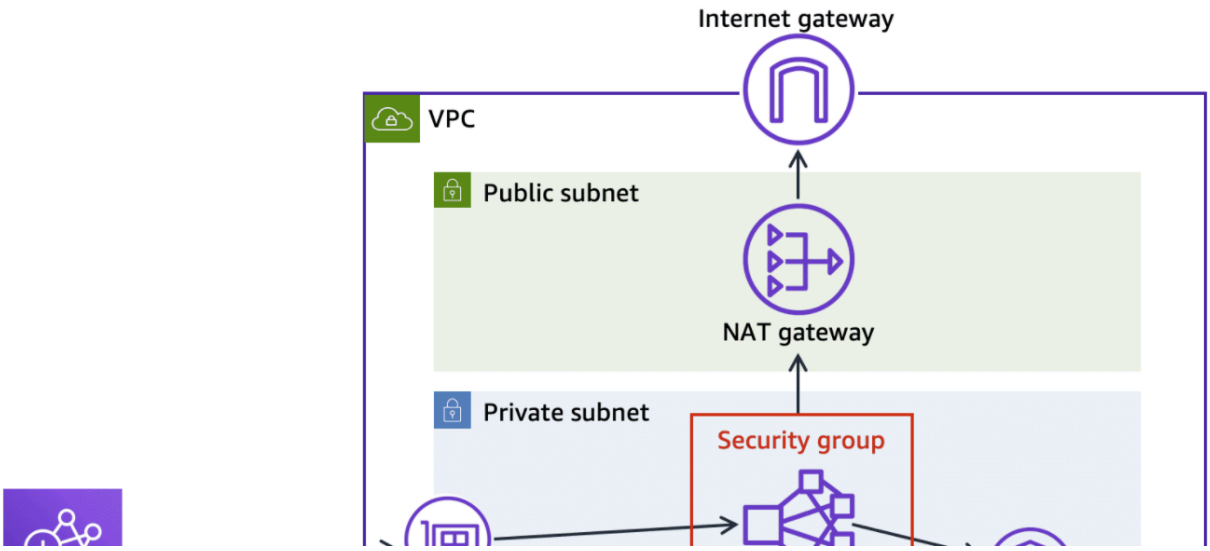
When you launch an Amazon EMR cluster, you must specify one Amazon EMR–managed security group for the master instance, one for Amazon EMR–managed security group for the core/task instances, and optionally, one for the Amazon EMR resources used to manage clusters in private subnets. The core/task instances share the same security group.

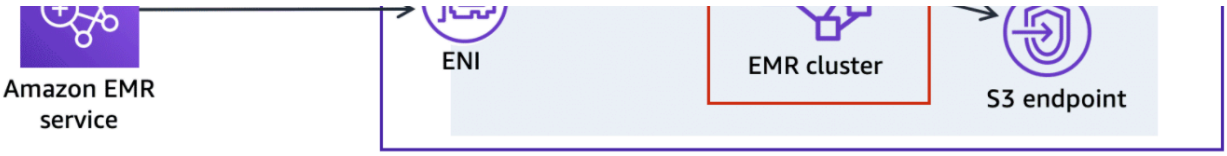
S3 endpoint



You can create a private endpoint for Amazon S3 in your subnet to give your Amazon EMR cluster direct access to data in Amazon S3.

NAT Gateway





You can optionally create a Network Address Translation (NAT) Gateway for your cluster to interact with other AWS services, like Amazon DynamoDB and AWS Key Management Service (AWS KMS).

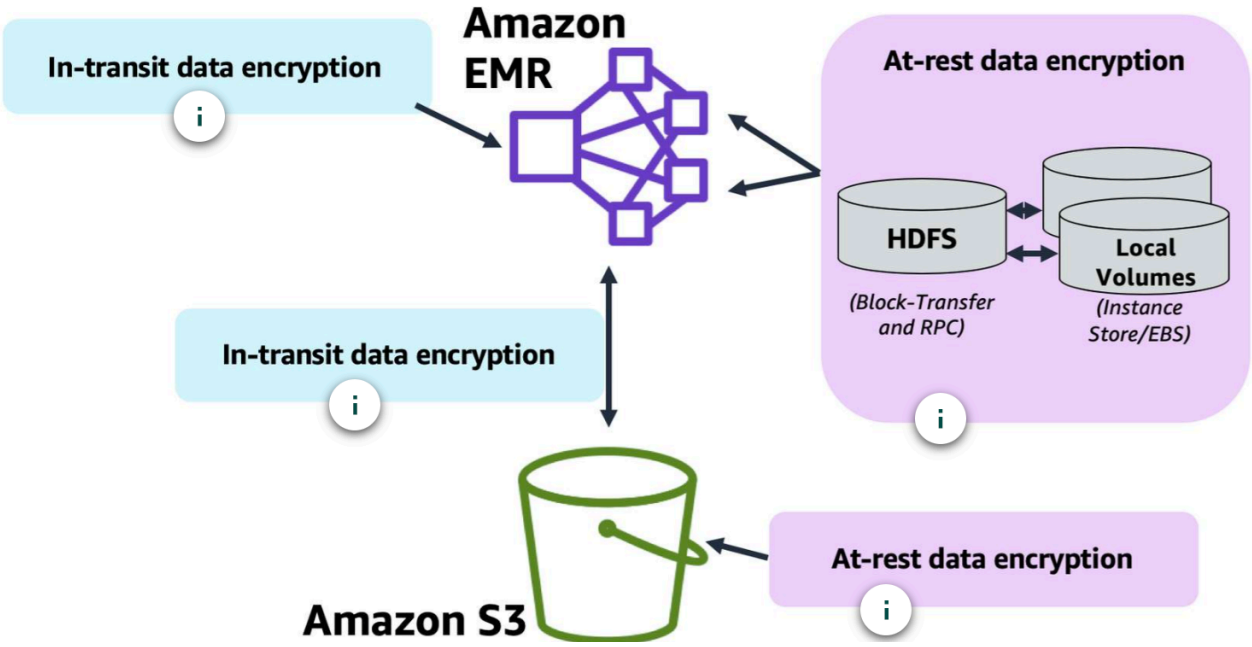
Additional security controls

Apache Ranger on Amazon EMR: Apache Ranger is a role-based access control framework to enable, monitor, and manage comprehensive data security across the Hadoop platform.

Amazon Redshift stored procedures: For fine-grained access control, you can create stored procedures to perform functions without giving a user access to the underlying tables. For example, only the owner or a superuser can truncate a table, and a user needs write permission to insert data into a table. Instead of granting a user permissions on the underlying tables, you can create a stored procedure that performs the task. You then give the user permission to run the stored procedure.

Data Protection and Encryption

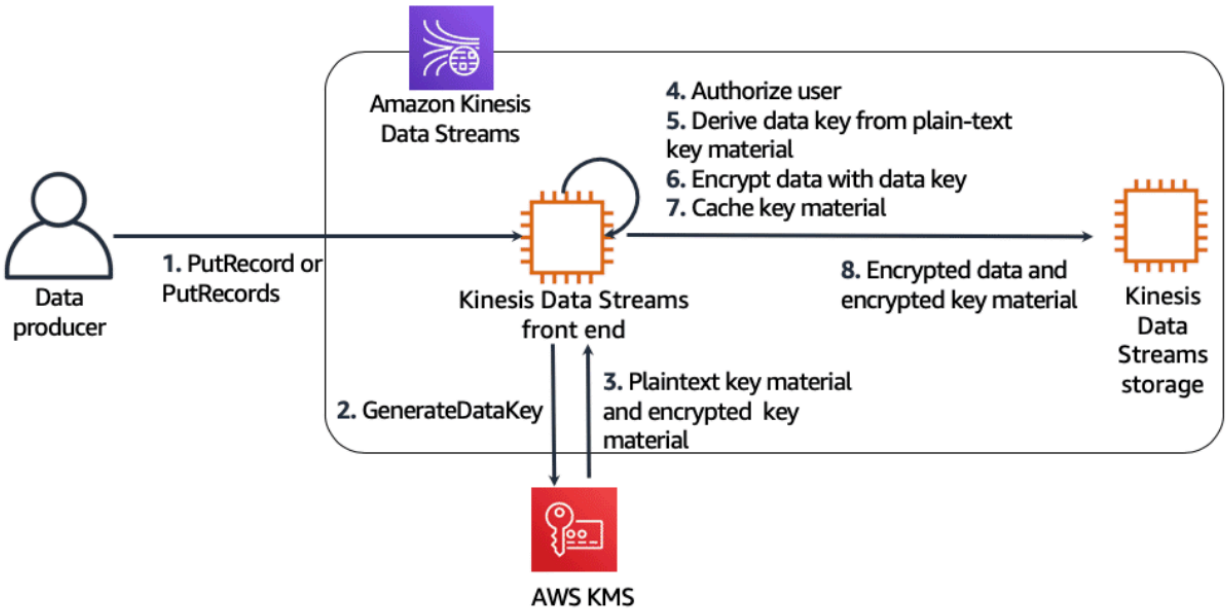
Encryption at rest and in transit



Amazon S3 encryption method	Description
SSE-S3	Amazon S3 manages the data and master encryption keys.
SSE-KMS	AWS manages the data key but you manage the customer master key (CMK) in AWS KMS.
SSE-C	You manage the encryption key.
CSE-KMS	Objects are encrypted before being uploaded to Amazon S3 and the client uses keys provided by AWS KMS
CSE-Custom	Objects are encrypted before being uploaded to Amazon S3 and the client uses a custom Java class that provides the client-side master key.

Kinesis data stream encryption

Amazon Kinesis Data Streams can automatically encrypt sensitive data as a producer enters it into a stream. Kinesis Data Streams uses AWS KMS master keys for encryption. To read from or write to an encrypted stream, producer and consumer applications must have permission to access the master key.



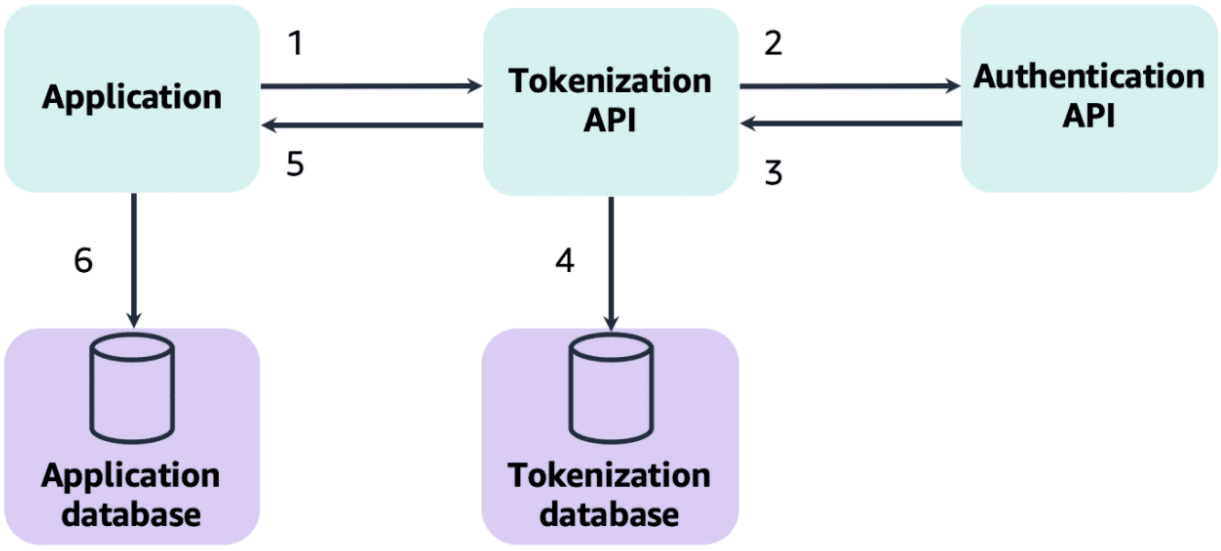
Redshift encryption

With Amazon Redshift, encryption is an optional, immutable cluster configuration. You can encrypt data at rest using either server-side or client-side encryption, or you can encrypt your data in transit using SSL.

DynamoDB encryption
DynamoDB offers encryption for user data at rest and also encryption for data in transit between on-premises clients and DynamoDB, and between DynamoDB and other AWS resources within the same AWS Region.

Tokenization

Tokenization helps protect certain parts of the data that has high sensitivity or a specific regulatory compliance requirement, such as PCI. Separating the sensitive data into its own dedicated, secured data store and using tokens in its place helps you avoid the potential cost and complexity of end-to-end encryption. It also allows you to reduce risk with temporary, one-time-use tokens.



1. The application presents sensitive data such as a credit card number to the tokenization API.
2. The tokenization API requests authentication of the application user.
3. The authentication API authenticates the user.
4. The tokenization API generates a token for the credit card number and stores and links the token and the credit card number in the tokenization database.
5. The tokenization API returns the token to the application.
6. The application stores the token in the application database in place of the credit card number.

Data governance

A good data governance and compliance program encompasses identifying

A good data governance and compliance program encompasses identifying the required compliance frameworks (such as HIPAA or PCI) and understanding the contract and agreement obligations, as well as setting up a system to monitor policies, standards, and security controls to be able to respond to events and changes in risk.

AWS CloudTrail and Amazon CloudWatch

While CloudTrail allows you to track who is doing what in your AWS account by recording API activity, CloudWatch monitors how your resources are performing by collecting metrics and log information across your application landscape.

CloudTrail

CloudTrail allows you to **audit and review API calls** and detect security anomalies inside your AWS account. For example, in Amazon RDS you can use CloudTrail to log all calls to Amazon RDS API, such as CreateDBInstance, StartDBInstance and StopDBInstance.

CloudWatch

You can use CloudWatch to monitor performance and resource utilization. In addition, you can **set up alert rules that trigger Amazon SNS notifications or Lambda functions** to run in response to a security or risk event.

Enforce compliance with AWS Config

AWS Config reports the configuration state of your AWS resources and allows you to define custom rules to track desirable or undesirable configuration conditions.