AWS SysOps Administrator – Associate Level

TECHNOLOGY

Storage and Data Management

```
#include <iostream>
#include <cstdlib>
using namespace std;
int main()
{
    cout << "Hello, world!" << endl;
    return 0;
}
```

Learning Objectives

By the end of this lesson, you will be able to:

- Check S3 versioning on AWS Console
- Work with default encryption and bucket policies
- Upgrade and change an EC2 volume
- Create a query in Athena to perform operations on a specific bucket in S3
- Customize a file system and access it using a specific EC2 instance

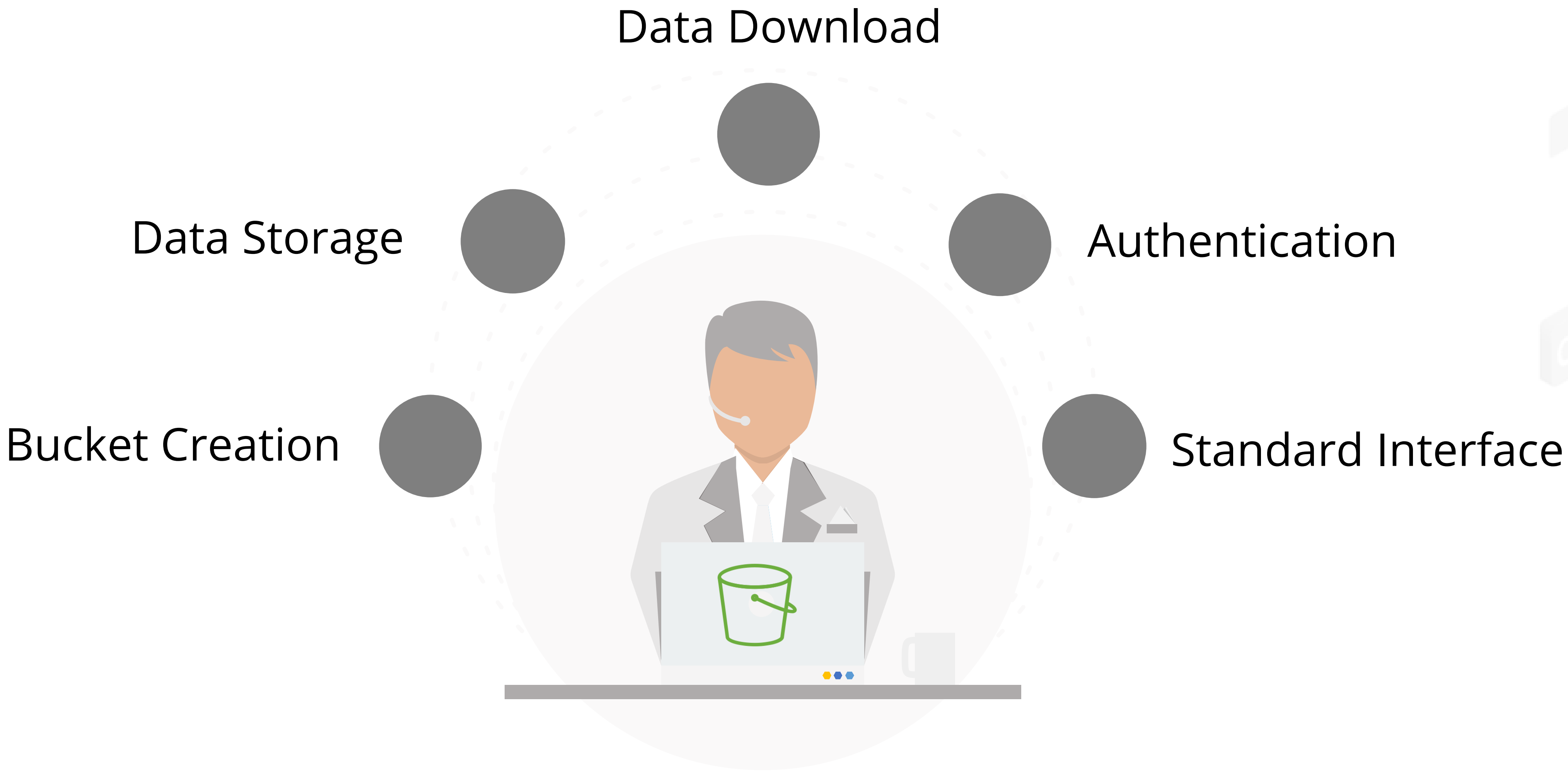
Introduction to S3

What Is S3?

S3 is defined as the storage to make web-scale computing easier for developers.

It has a web service interface that can be used to store, get data from anywhere and anytime on the web.

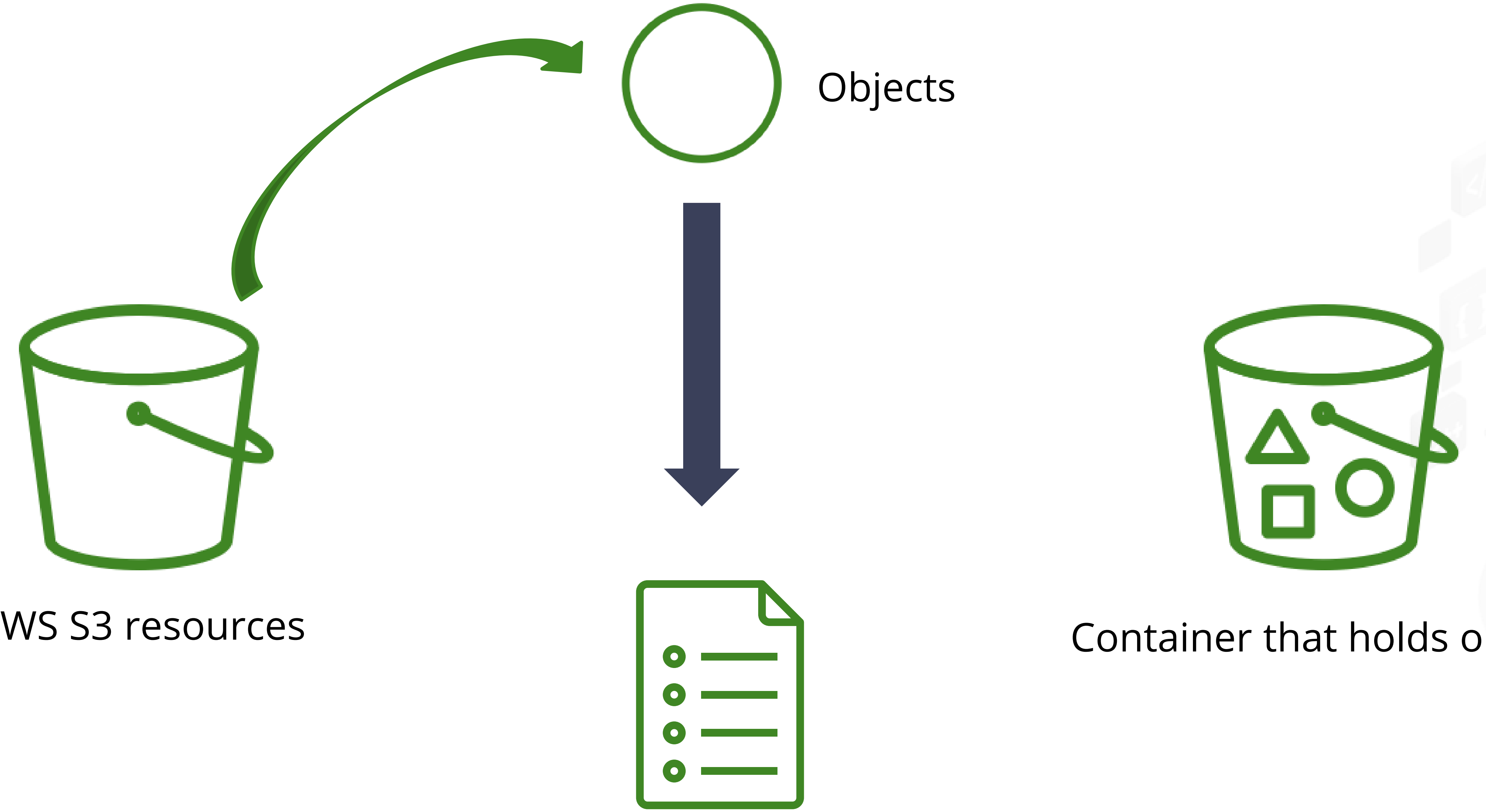
S3: Advantages



Concepts of S3

Buckets	Objects	Keys	Regions	Amazon S3 Data Consistency Model
They are defined as containers for objects stored in Amazon S3.	They are entities stored in S3 and consist of object data and metadata.	They are unique identifiers for objects that are stored in a bucket. Every object has a unique key.	They are geographical AWS regions where the buckets created in S3 are stored.	It provides the read-after-write consistency for the PUTS of new objects in the S3 bucket in all regions with a warning.

Concepts of S3



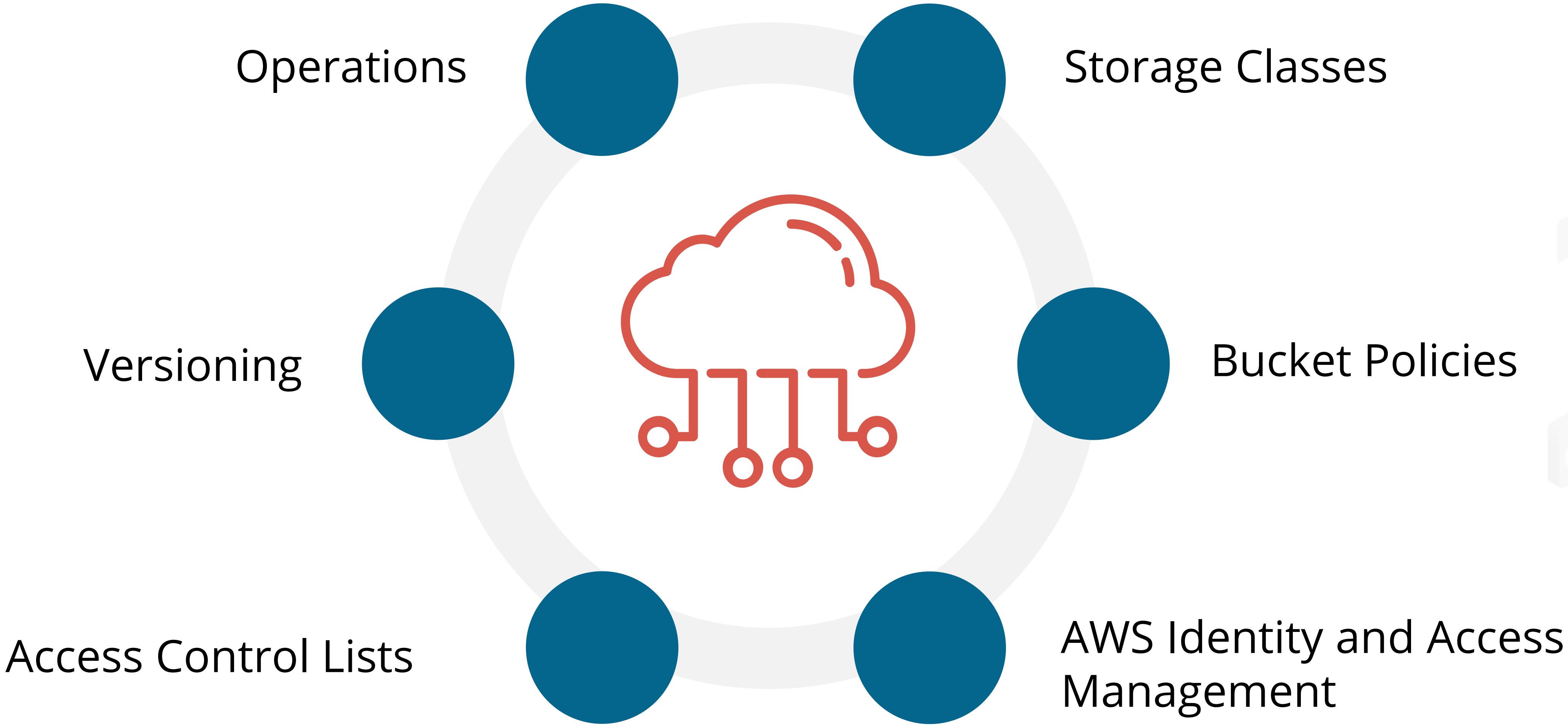
Objects

AWS S3 resources

Container that holds objects

Data with key and version ID

Features of S3



Operations

Storage Classes

Versioning

Bucket Policies

Access Control Lists

AWS Identity and Access Management

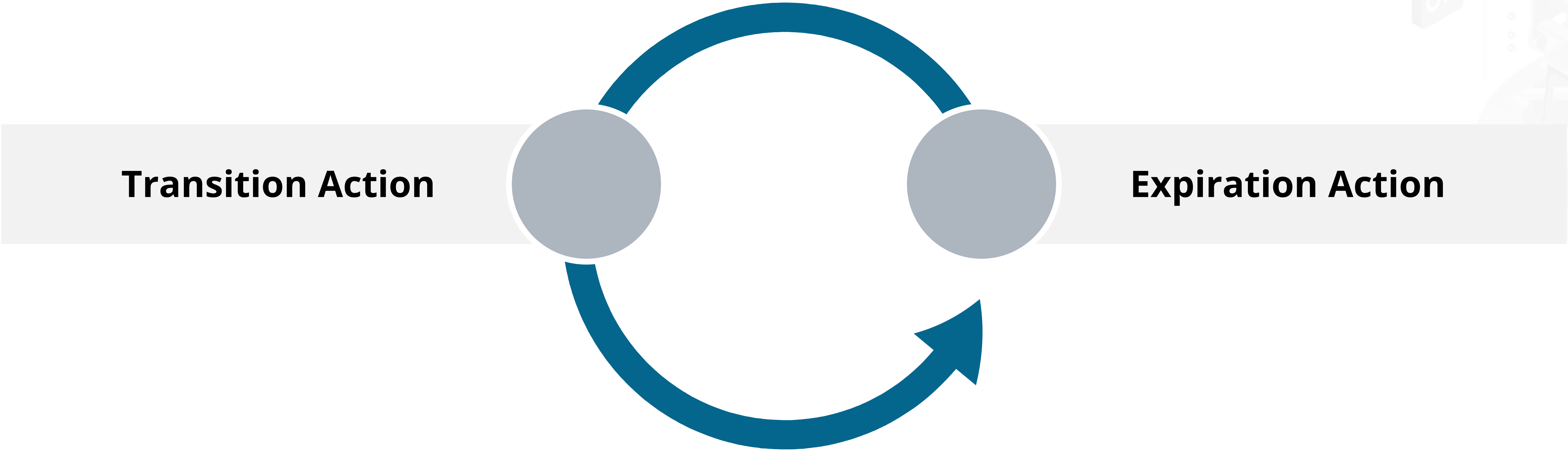
TECHNOLOGY

S3 Lifecycle Policies

S3 Lifecycle Policies

Amazon S3 Lifecycle policies are configured to manage the objects to be stored effectively throughout the lifecycle.

Below are the two types of actions:



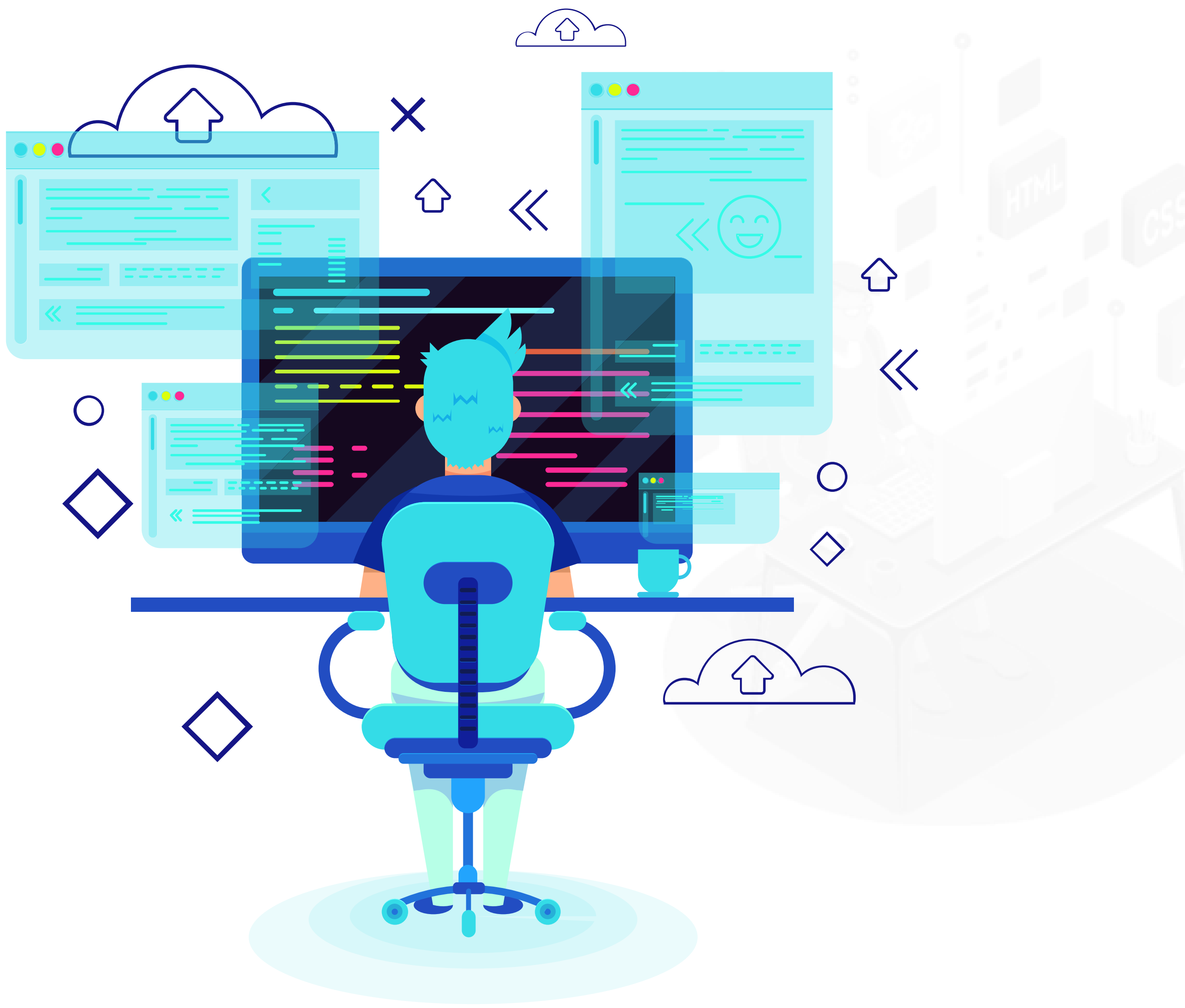
Transition Action

Expiration Action

S3 Lifecycle Policies: Example

Below is an example of lifecycle configuration to abort multipart uploads API that is used to upload large objects in parts.

```
<LifecycleConfiguration>
    <Rule>
        <ID>sample-rule</ID>
        <Filter>
           <Prefix>SomeKeyPrefix/</Prefix>
        </Filter>
        <Status>rule-status</Status>
        <AbortIncompleteMultipartUpload>
          <DaysAfterInitiation>7</DaysAfterInitiation>
        </AbortIncompleteMultipartUpload>
    </Rule>
</LifecycleConfiguration>
```



Default Encryption and Bucket Policies

Duration: 10 Min.

Problem Statement:

Demonstrate how default encryption and bucket policies work in AWS.

ASSISTED PRACTICE

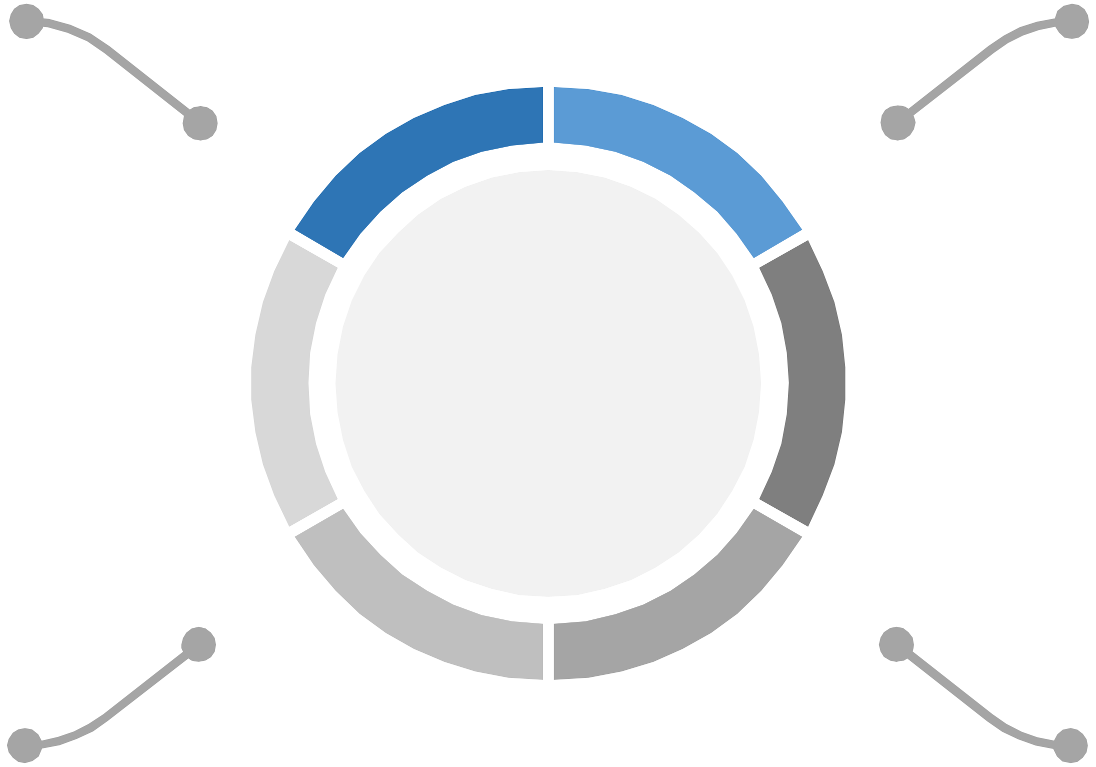
Assisted Practice: Guidelines

Steps to work with default encryption:

1. Log in to your AWS lab
2. Click on **Amazon S3**, and create a bucket with all the relevant details
3. Click on the **Automatically encrypt objects when stored in S3** checkbox
4. Verify the working of default encryption
5. Upload the file and check the encryption

S3 Versioning

Versioning refers to storing multiple variants of an object in the same bucket. Below are the benefits of versioning:



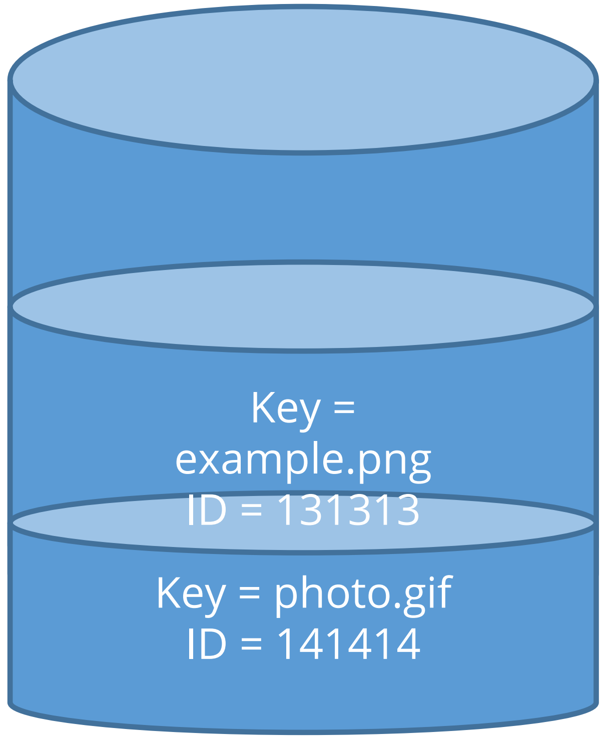
Enables us to recover objects from accidental deletion or overwrite

Used to preserve and restore every version of each object stored

Helps Amazon S3 to automatically generate the unique version ID for the object to be stored

Used to recover from application failures and unintended user actions

S3 Versioning: Example



We can have two objects in a bucket with the same key but different version IDs, such as example.png (version 131313) and photo.gif (version 141414).

S3 Versioning

Duration: 10 Min.

Problem Statement:

Check S3 versioning on the AWS Console.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to check S3 versioning:

1. Log in to your AWS lab
2. Go to the Amazon dashboard and choose Amazon S3
3. Create a bucket
4. Configure the S3 bucket to enable versioning
5. Upload the files, verify the versioning, and reupload the bucket to check consistency

TECHNOLOGY

MFA Delete

MFA Delete

Multi-Factor Authentication (MFA) is used to add another layer of security by configuring the bucket.

MFA requires the authentication for the below operations:

Changing the versioning state of the bucket

Permanently deleting the version of an object

MFA Delete

MFA Delete requires the below two forms of authentications together:

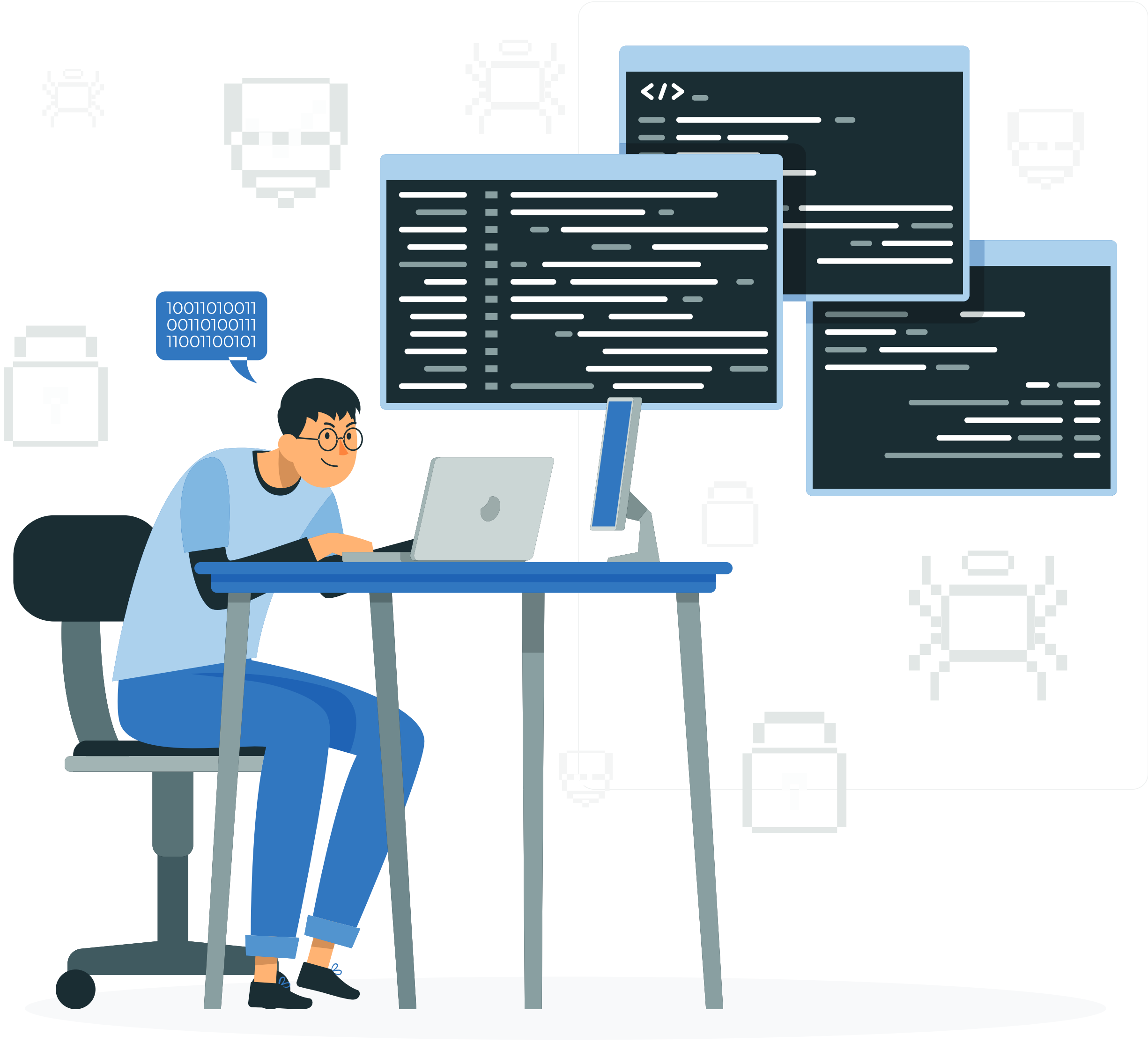
- User security credentials
- Concatenation of a serial number, space, and the six-digit code displayed on an approved authentication device

MFA Delete

MFA Delete can help prevent accidental bucket deletions by doing the following:

- Adding a user to initiate the delete action to prove the physical possession of an MFA device
- Adding an extra layer of security and friction to the delete action

MFA Delete



```
<VersioningConfiguration
xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
  <Status>VersioningState</Status>
  <MfaDelete>MfaDeleteState</MfaDelete>
</VersioningConfiguration>
```

S3 Encryption

S3 Encryption

Amazon S3 default encryption is used to set the default encryption behavior for an S3 bucket. This is done so that all the new objects are encrypted when they are stored in the bucket.

The objects are encrypted using server-side encryption with either Amazon S3-managed keys (SSE-S3) or customer master keys (CMKs) stored in AWS Key Management Service (AWS KMS).

S3 Encryption

When you use server-side encryption, Amazon S3 encrypts an object before saving it to the disk and decrypts it when you download the objects.

TECHNOLOGY

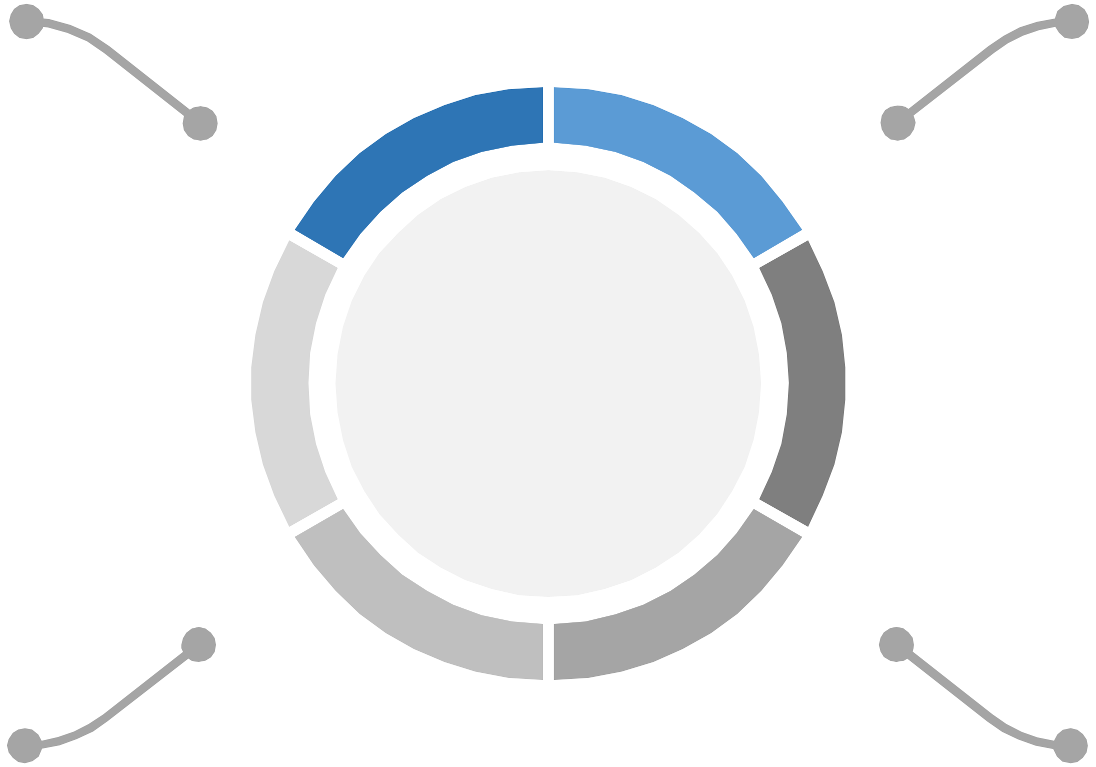
EC2 Volume Types

Volume Types

Characteristics				
Volume Types	General Purpose SSD (gp2)	Provisioned IOPS SSD (io2)	Throughput Optimized HDD (st1)	Cold HDD (sc1)
Description	It balances price and performance for workloads	It is meant for mission-critical, low-latency, or high-throughput workloads	It is a low-cost HDD volume that is designed for frequently accessed workloads	It is the lowest cost HDD volume that is designed for less frequently accessed workloads
Durability	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.999% durability (0.001% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)
Volume size	1 GiB - 16 TiB	4 GiB - 16 TiB	500 GiB - 16 TiB	500 GiB - 16 TiB
Max IOPS per volume	16,000 (16 KiB I/O) *	64,000 (16 KiB I/O)	500 (1 MiB I/O)	250 (1 MiB I/O)

Terminating an Instance

When an instance is terminated:



The data on instance store volumes linked with that instance is deleted.

It automatically gets deleted from the console after a short while.

Amazon EBS root device volumes are automatically deleted.

All the resources get disassociated from the instance.

Amazon EC2 Instance Store

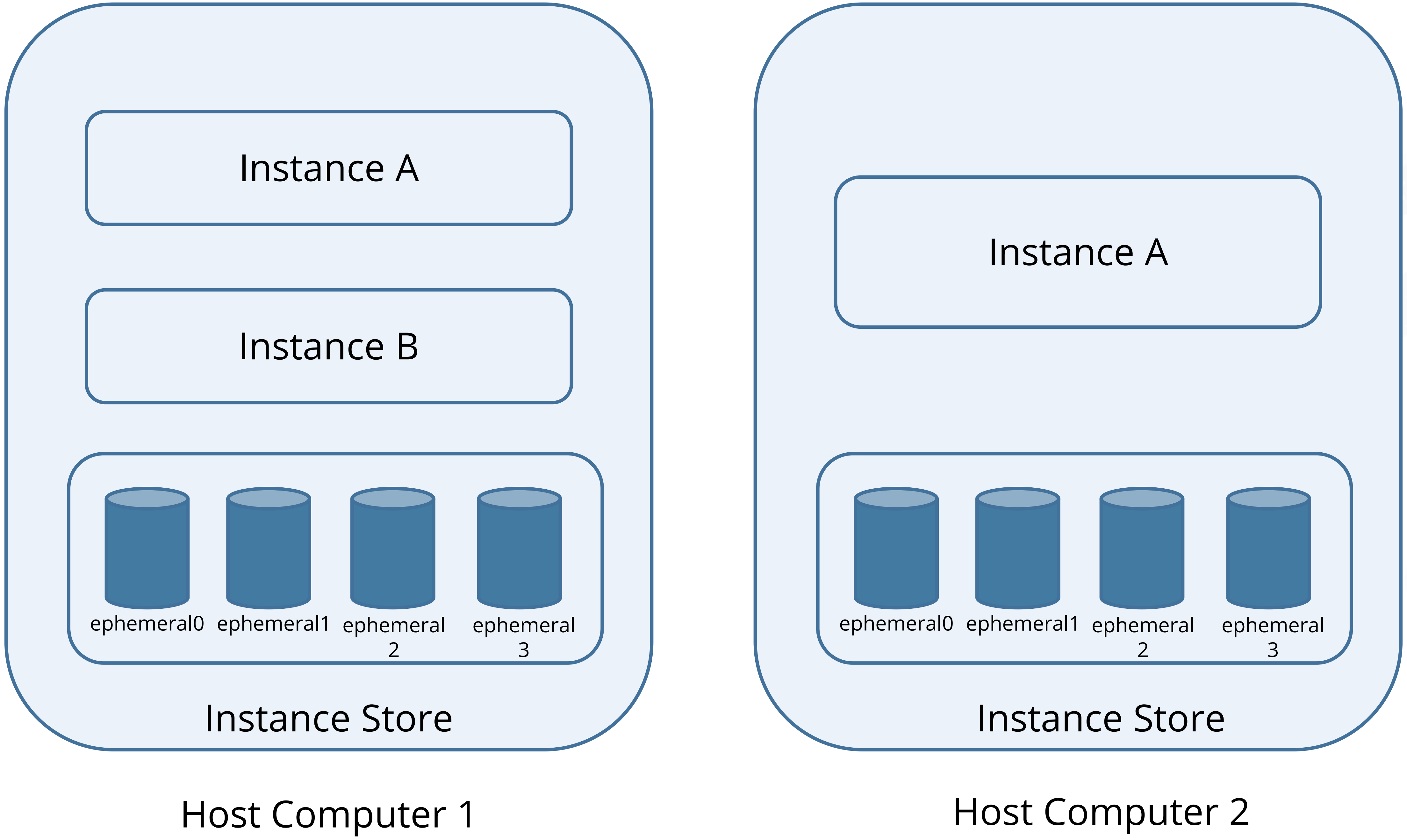
An instance store is used to provide temporary block-level storage for the instance that is located on disks attached to the computer.

03 Size increases with the increase in the number of devices available

02 Consists of one or more instance store volumes exposed as block devices

01 Ideal for the temporary storage of information that changes frequently

Amazon EC2 Instance Store



Upgrading EC2 Volume and Changing Volume Types

Duration: 15 Min.

Problem Statement:

Upgrade the volume of an EC2 instance running on AWS, and change the volume types.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to upgrade volume of an EC2 instance:

1. Log in to your AWS lab
2. Go to the Amazon dashboard, and click on EC2 instance
3. Choose the AMI machine type, and select **t2.micro** as the instance type
4. Add storage, and configure **Security Group**
5. Launch instance, and then configure EBS storage and storage space
6. Create a volume and attach the instance

TECHNOLOGY

KMS and CloudHSM

KMS

1 Key Management System (KMS) is used to create and manage cryptographic keys.

2 It is used to control their use across AWS services.

3 KMS protects the keys by using hardware security modules that are validated.

KMS

Below are the benefits of KMS:

- Centralized Key Management
- Fully Managed
- Managed Encryption
- Digitally Signed Data
- Compliance
- Built-in Auditing
- Cost-Effective
- Highly Secured

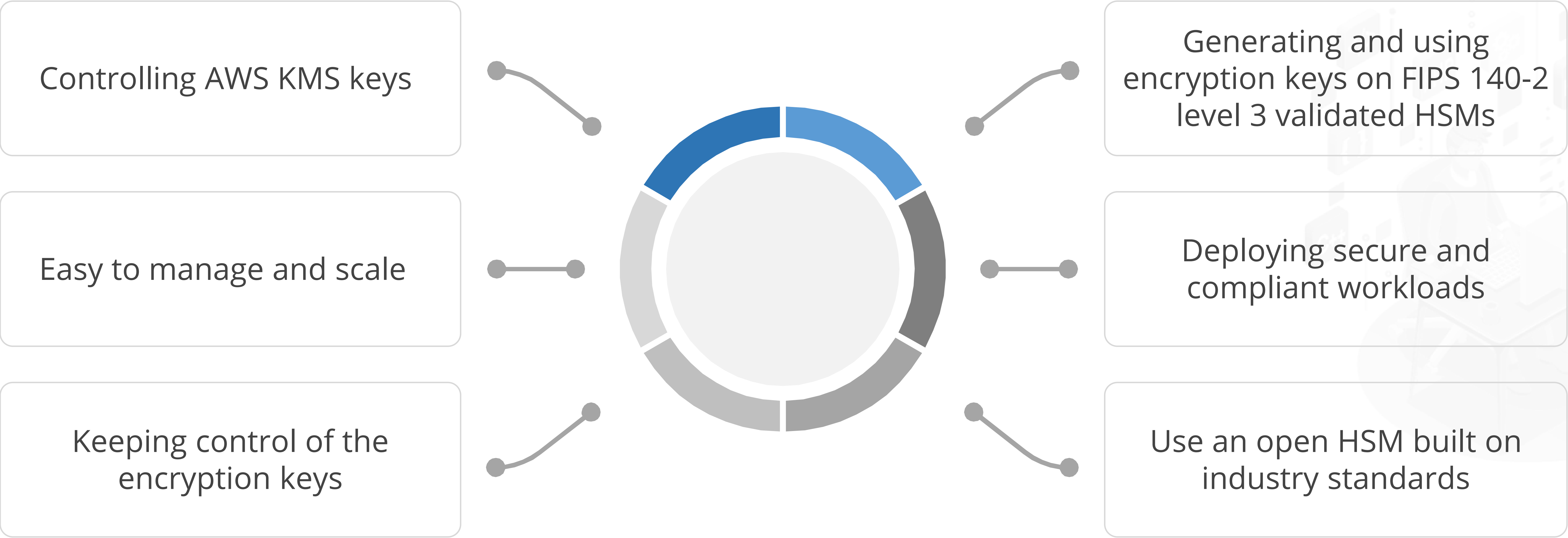
CloudHSM

AWS CloudHSM is defined as a Cloud-based Hardware Security Module (HSM) that is used to generate and use encryption keys on the cloud.

It helps us to manage encryption keys using FIPS 140-2 Level 3 validated HSMs.

CloudHSM

Below are the benefits of KMS:



Controlling AWS KMS keys

Easy to manage and scale

Keeping control of the encryption keys

Generating and using encryption keys on FIPS 140-2 level 3 validated HSMs

Deploying secure and compliant workloads

Use an open HSM built on industry standards

CloudHSM vs. KMS

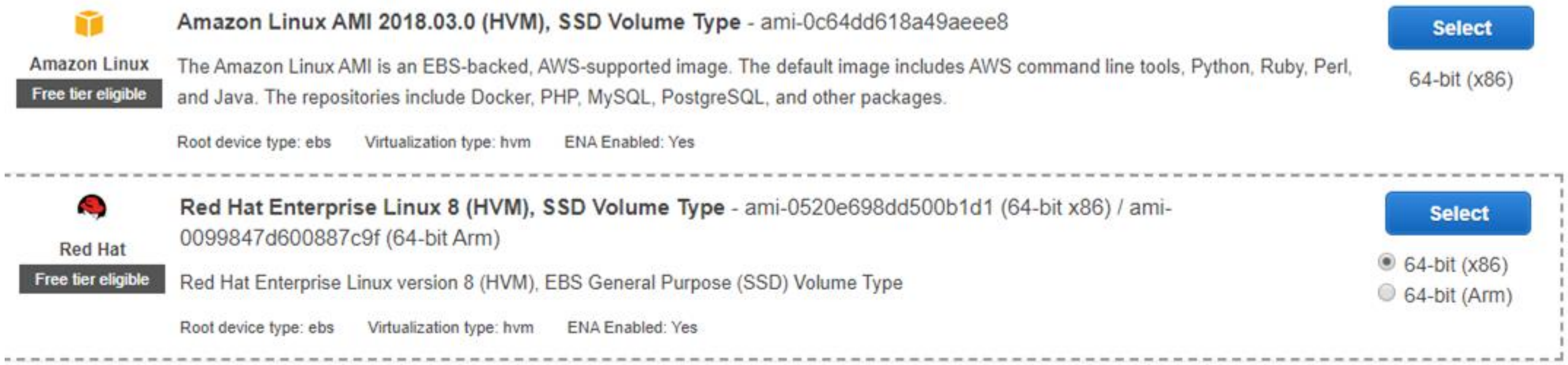
Properties	AWS CloudHSM	AWS KMS
Tenant	Single-Tenant	Multi-Tenant
Standard	FIPS 140-2 Level 3 Common Criteria EAL4+	FIPS 140-2 Level 2
Access Authentication/Policy	Quorum-based K of N principle	AWS IAM Policy
Key Accessibility	Accessed and shared across multiple VPCs	Accessible in multiple regions
High Availability	ADD HSM in different Availability Zones	AWS-Managed service
Master Keys	Master Key HSM	Customer-owned master key AWS-managed master Key AWS-owned master key

TECHNOLOGY

Amazon Machine Image

What Is an AMI?

Amazon Machine Image (AMI) is used to provide the information required to launch an instance. A single AMI can be used to launch multiple instances with the same configuration.



Amazon Linux

Free tier eligible

Amazon Linux AMI 2018.03.0 (HVM), SSD Volume Type - ami-0c64dd618a49aeee8

The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.

Root device type: ebs   Virtualization type: hvm   ENA Enabled: Yes

64-bit (x86)

Red Hat

Free tier eligible

Red Hat Enterprise Linux 8 (HVM), SSD Volume Type - ami-0520e698dd500b1d1 (64-bit x86) / ami-0099847d600887c9f (64-bit Arm)

Red Hat Enterprise Linux version 8 (HVM), EBS General Purpose (SSD) Volume Type

Root device type: ebs   Virtualization type: hvm   ENA Enabled: Yes

● 64-bit (x86)
○ 64-bit (Arm)

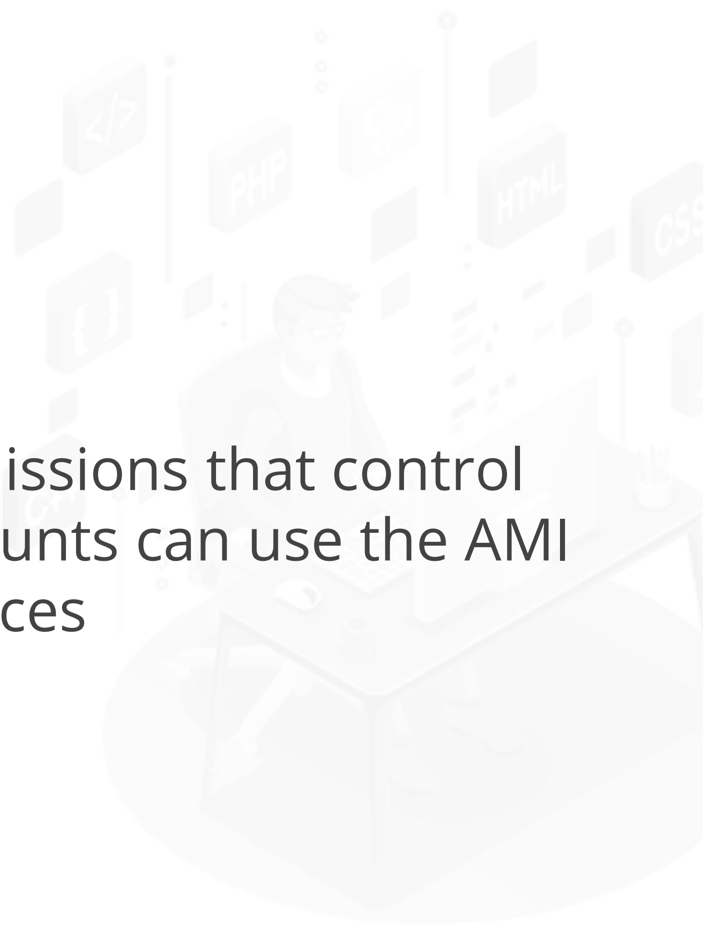
What Is an AMI?

An AMI includes the following:

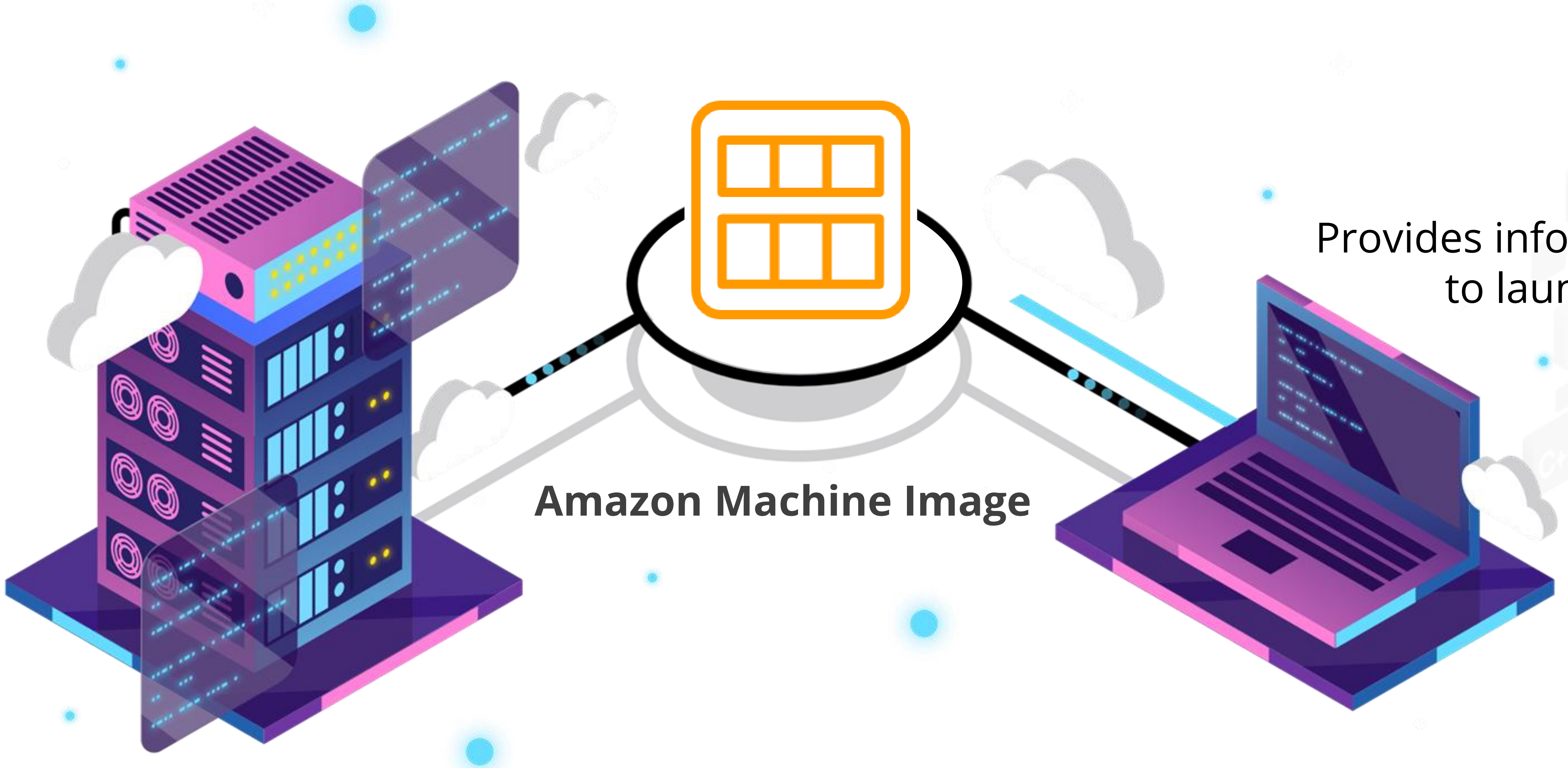
One or more EBS snapshots for instance-store-backed AMIs

Launching permissions that control which AWS accounts can use the AMI to launch instances

A block device mapping that specifies the volumes to attach to the instance when it's launched



What Is an AMI?



Amazon Machine Image

Provides information required to launch EC2 instances

Comprises preconfigured templates for the creation of virtual servers (EC2 instances) in the AWS environment

AMI Types

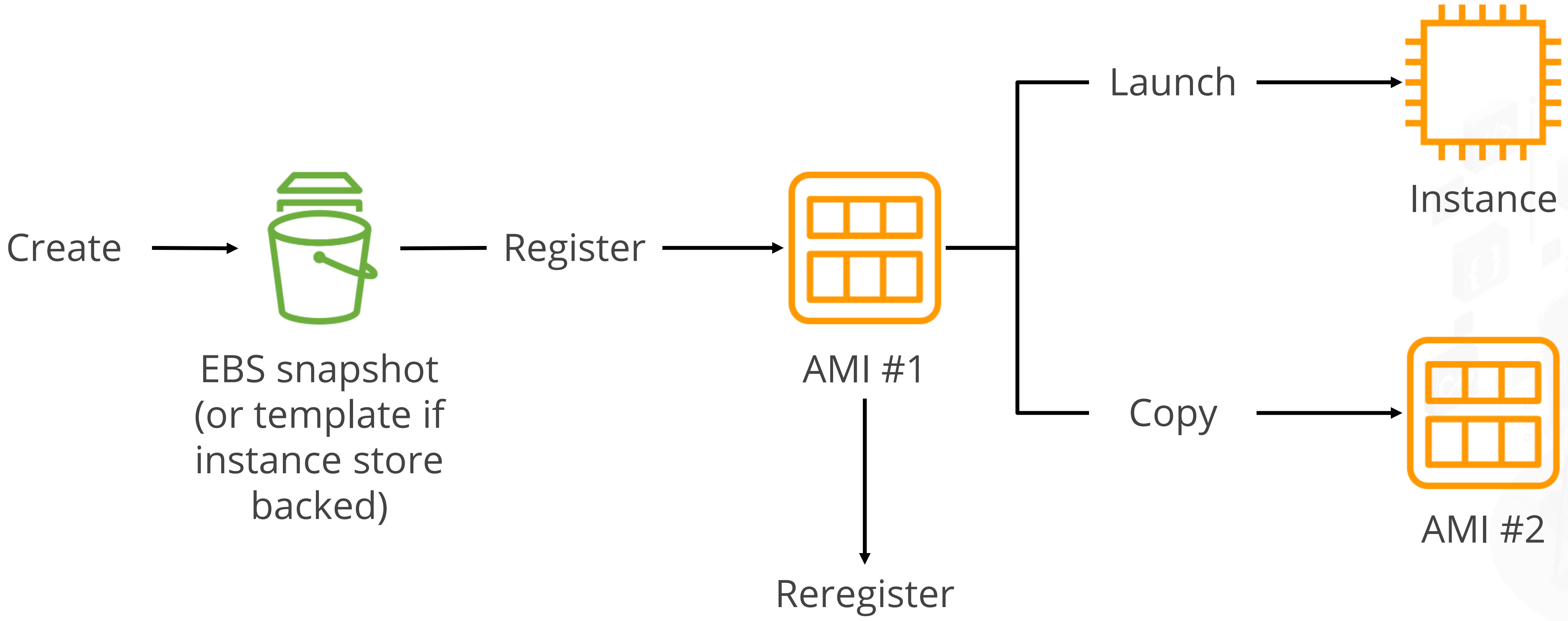
AWS-managed AMIs

Private and custom AMIs created by users

Public custom AMIs: Provided by community

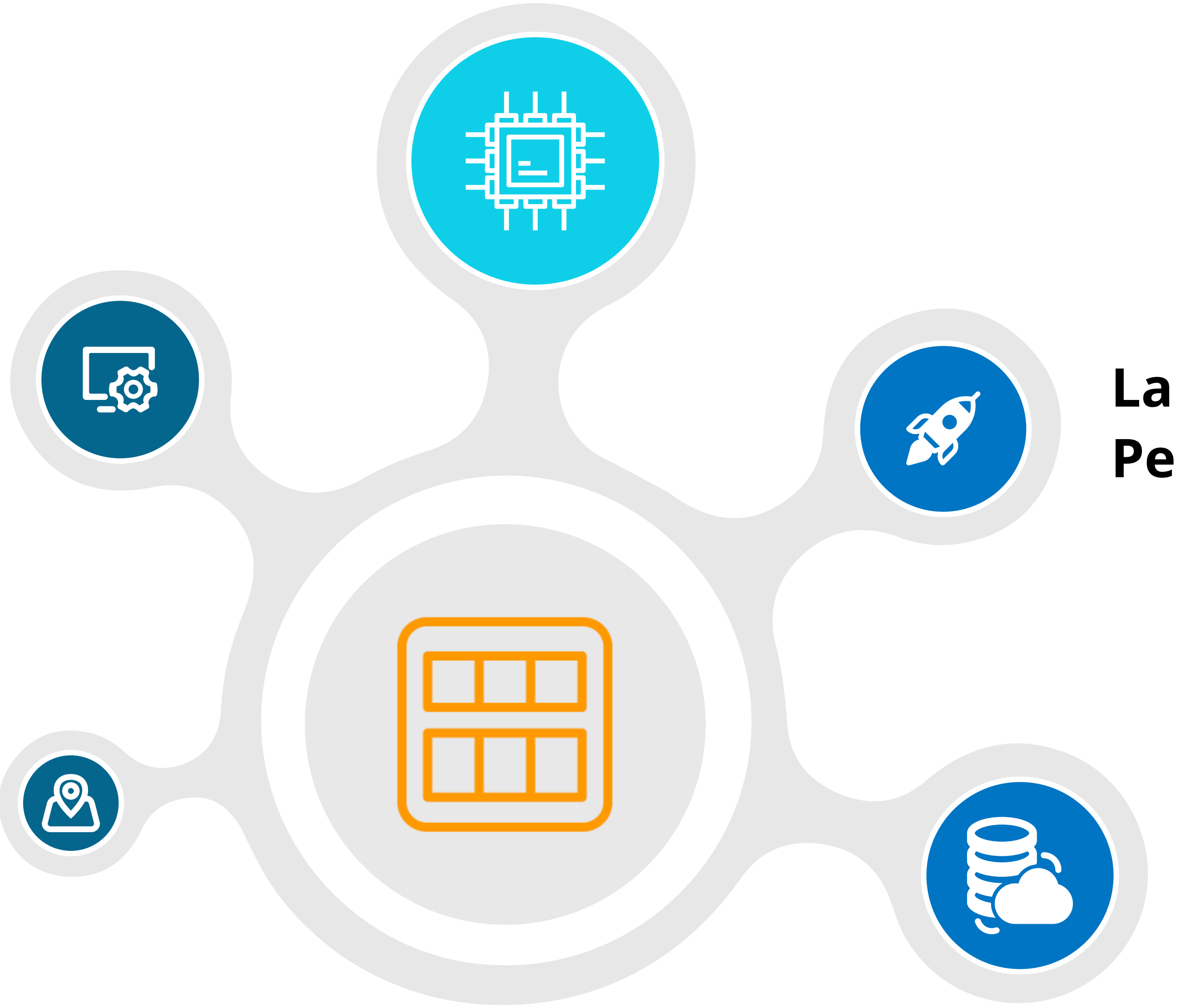
Private AMIs: Shared with you but are created by other AWS accounts

AMI Lifecycle



AMI: Characteristics

An AMI can be selected based on the following characteristics:



Architecture (32-bit or 64-bit)

Operating Systems: Linux, Windows, and Ubuntu

Launch Permissions: Define who has access to the AMI

Region: AMIs are region specific

Storage: For the root device

Customizing and Configuring AMIs

It can improve provisioning when an instance is launched.

Configuration files can be used to configure and customize the environment quickly and consistently.

It is possible to create a custom AMI instead of the standard AMI included in the platform.

Customizing and Configuring AMIs

Below are the benefits of custom AMI:

Allows us to make changes in the low-level components

Improves the provisioning time

Reduces the time taken for the configuration server

Shared AMIs

- AMIs created by a developer and made available for another developer are called shared AMIs.

- The best approach for using EC2 is to use shared AMI that has components and custom content required.

- AMIs can be created and shared with others.

Shared AMIs: Characteristics

Risk

1. Amazon is not responsible for its integrity and security.

1. Users must deploy the shared AMI in the data center.

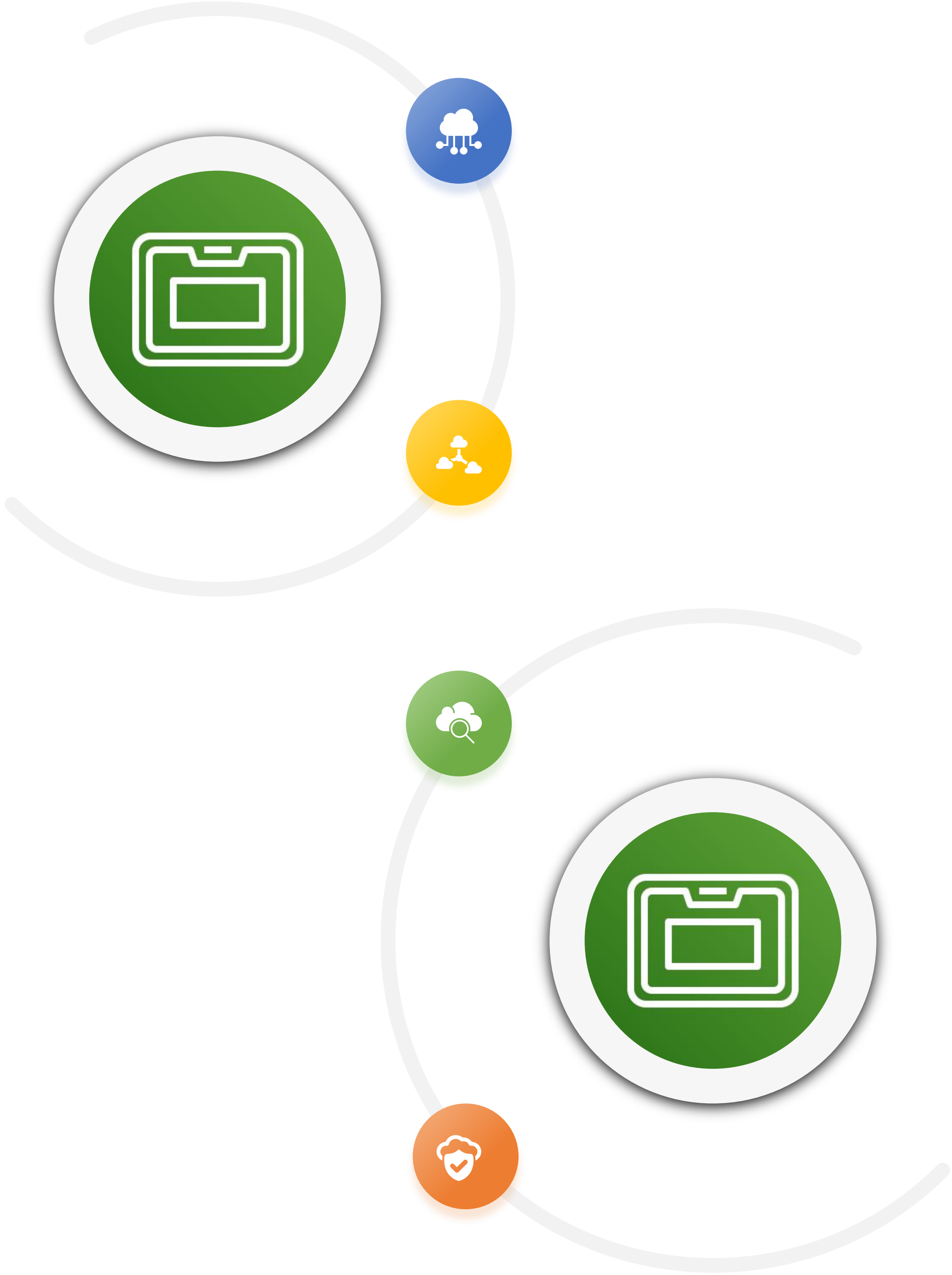
1. Use shared AMI from a trusted source.

Snowball

What is Snowball?

An edge computing and storage device

It is used to transfer a large amount of data to on-site data storage and S3 faster.

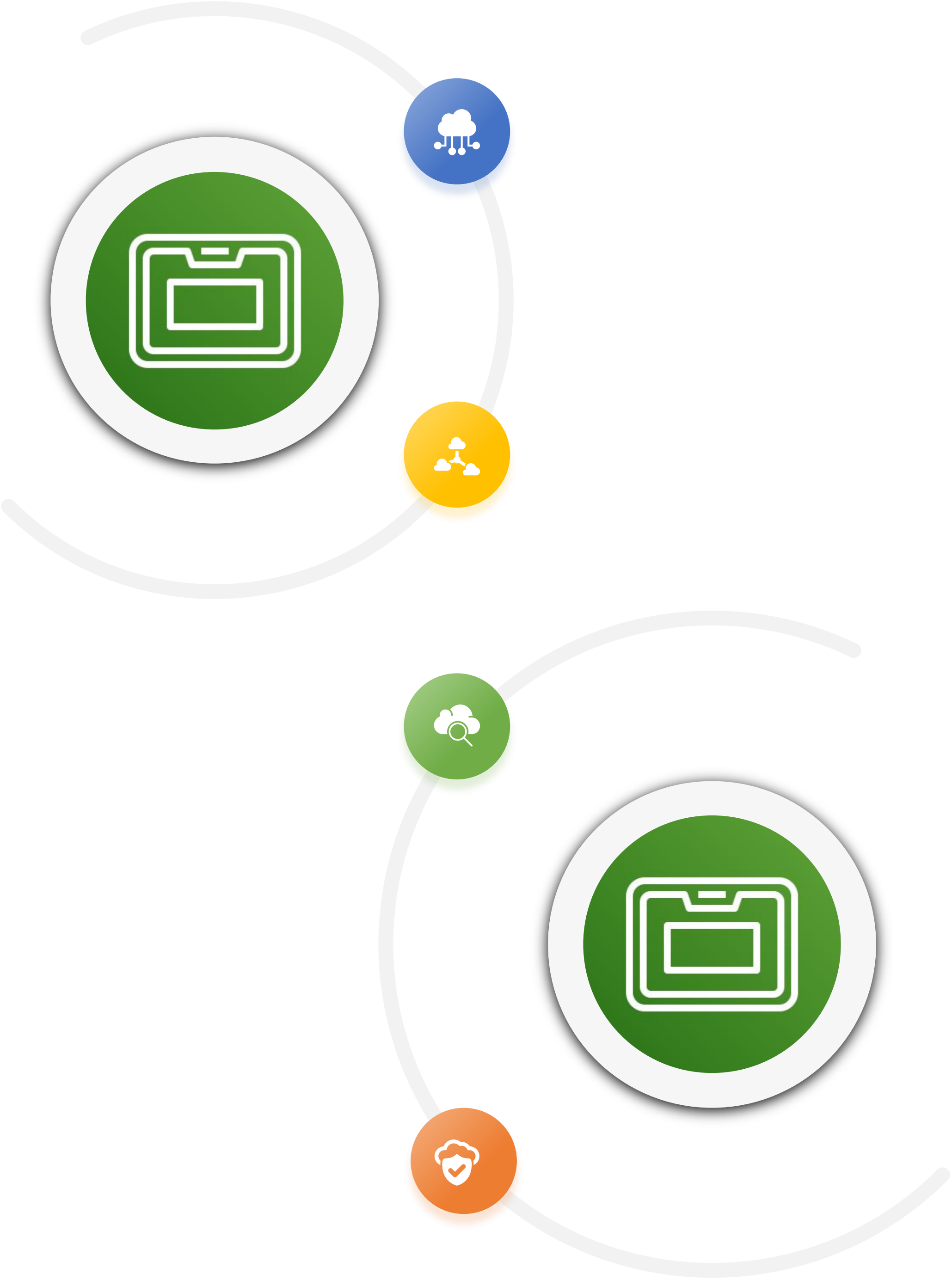


It provides an interface to create jobs, track data, and track the status of the jobs.

AWS KMS protects snowball, and it is used to secure and protect data in transit.

What is Snowball?

Below are the features of snowball:



Intended for transferring large amounts of data

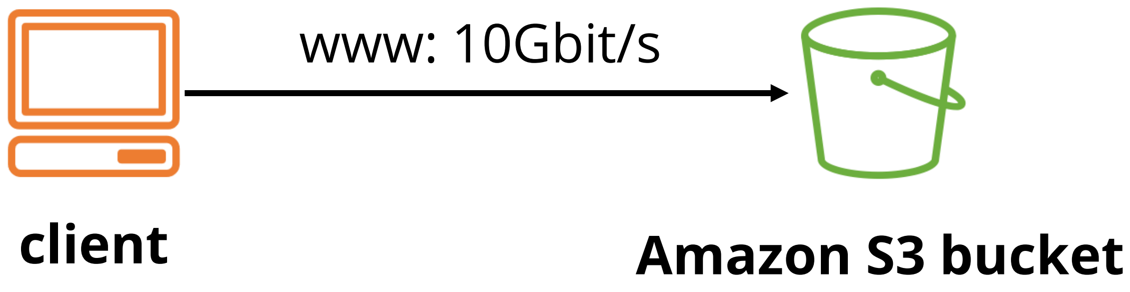
Protects the data at rest and in physical transit

No need to buy or maintain your own hardware devices

Allows you to perform local data transfers between your on-premises data center and a snowball

Working of Snowball

- **Direct upload to S3:**

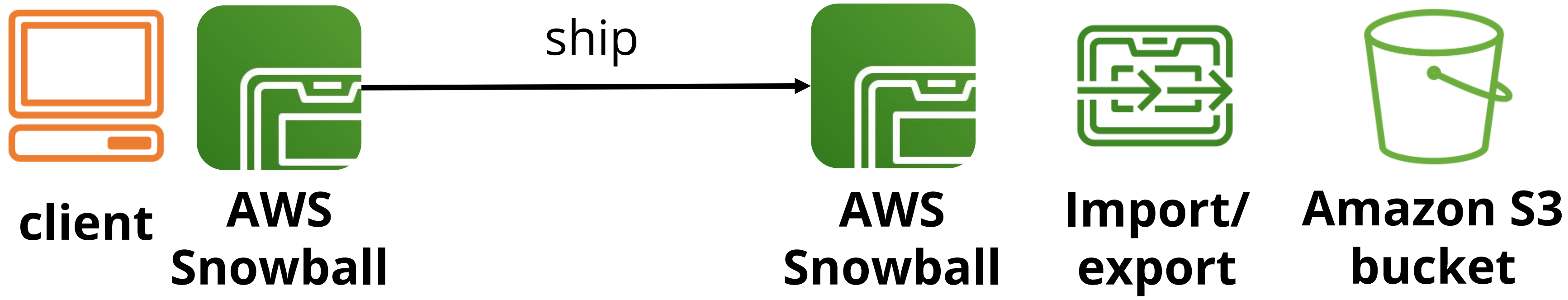


www: 10Gbit/s

client

Amazon S3 bucket

- **With snowball**



ship

client

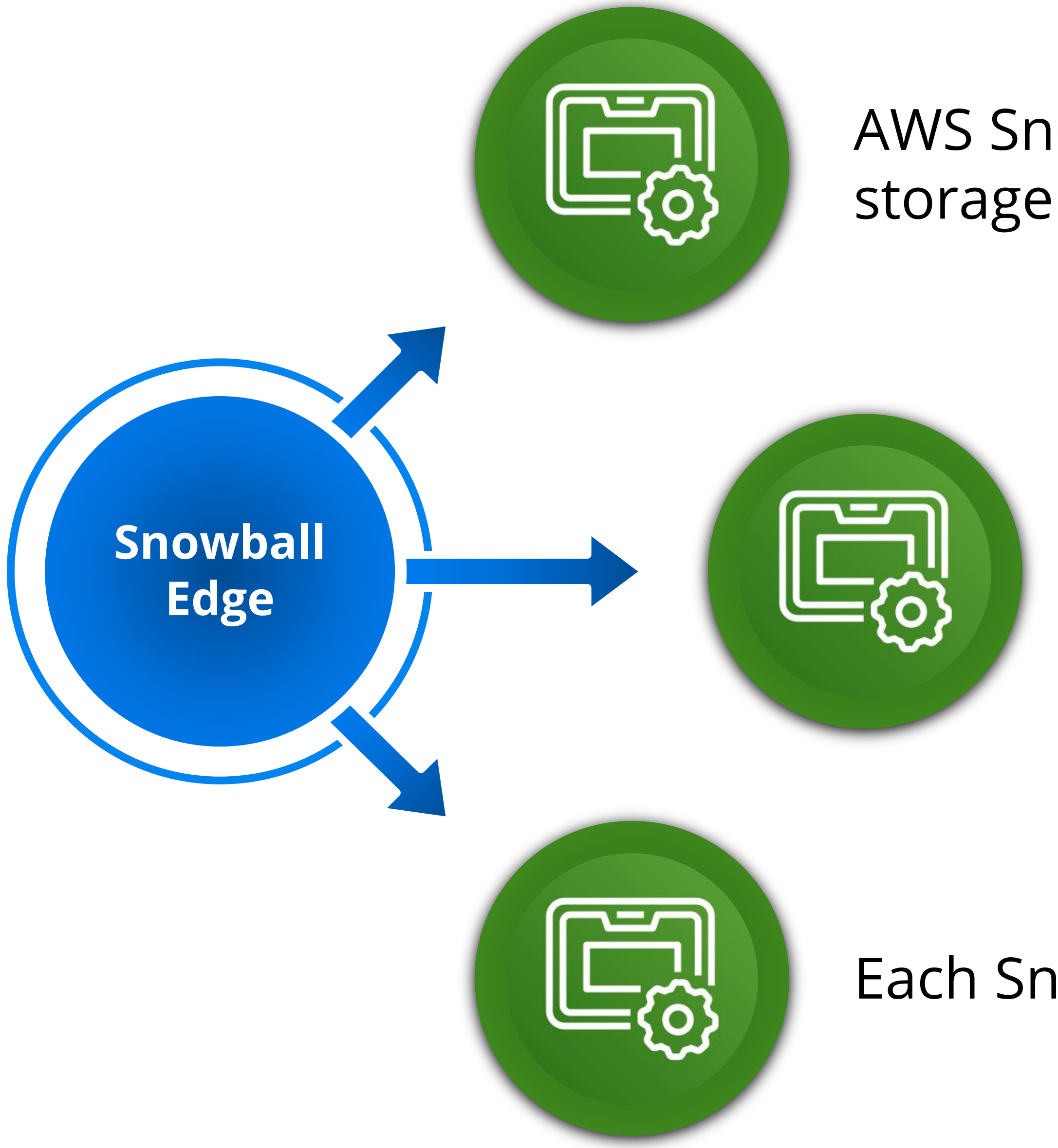
AWS Snowball

AWS Snowball

Import/ export

Amazon S3 bucket

What is Snowball Edge?



Snowball Edge

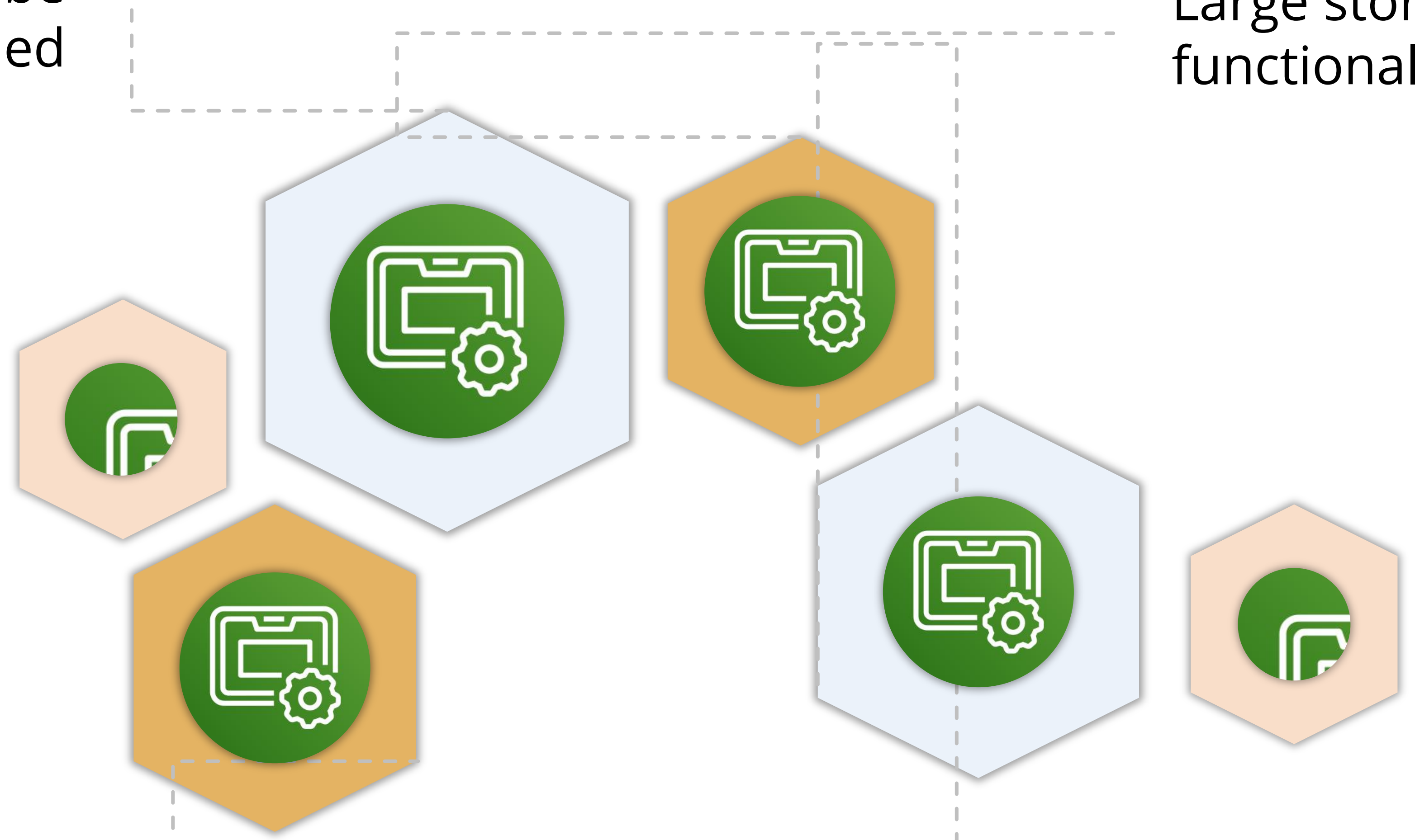
AWS Snowball Edge is a Snowball device with on-board storage and compute power for selecting AWS capabilities.

In addition to data transfer, it undertakes local processing and edge-computing workloads between your environment and the AWS Cloud.

Each Snowball Edge device transports data at speeds faster than the internet.

What is Snowball Edge?

Below are a few features of Snowball Edge:



Supports a custom EC2 AMI, so that processing can be performed

Large storage capacity or compute functionality for devices

Network adapters with transfer speeds of up to 100 GB/second

Supports custom Lambda function

Useful for pre-processing the data while moving

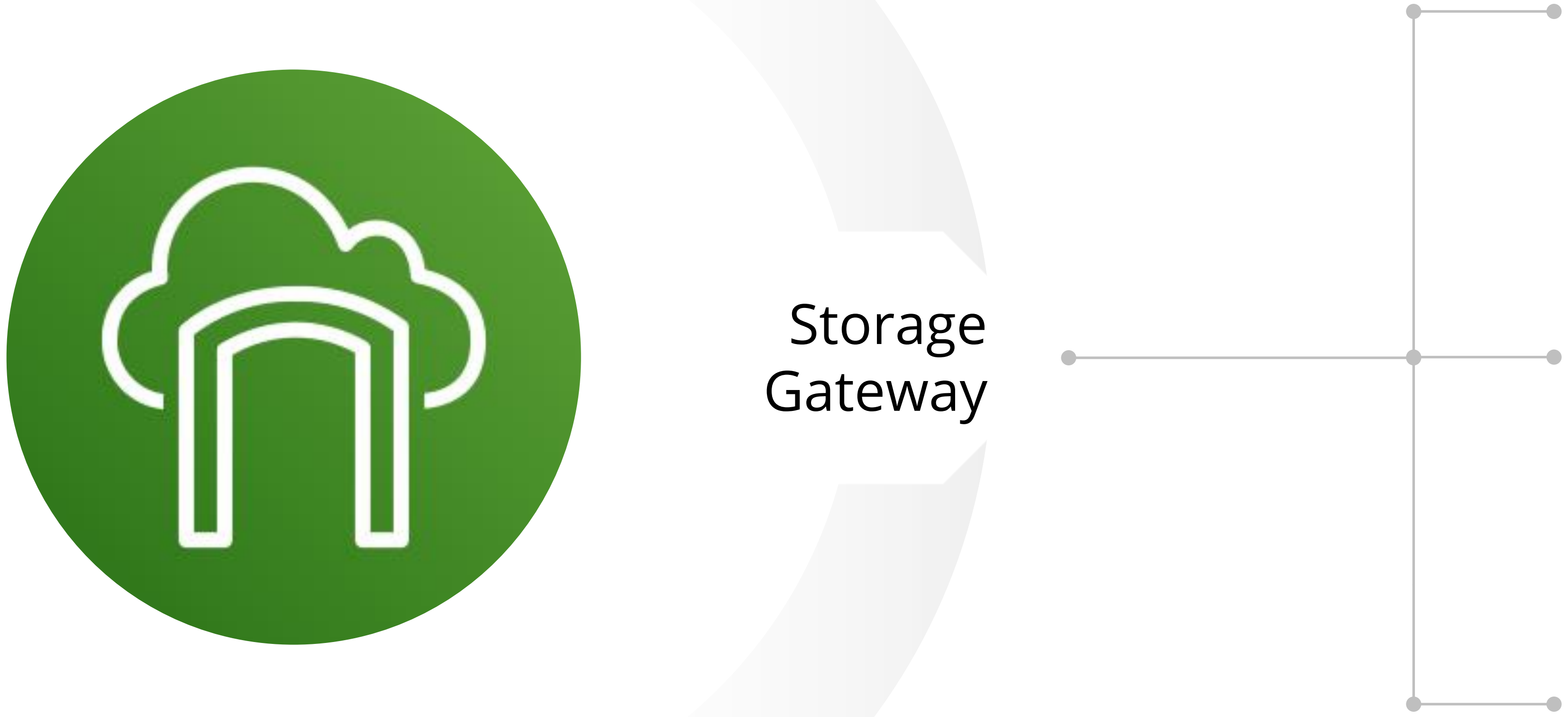
Has Amazon S3 and Amazon EC2 compatible endpoints available, enabling programmatic use cases

Snowball Vs. Snowball Edge

Properties	Snowball	Snowball Edge
Import and Export of data from AWS S3	Yes	Yes
Durable local storage	No	Yes
Local compute with AWS Lambda	No	Yes
Amazon EC2 compute instances	No	Yes
Use in a cluster of devices	No	Yes
Use with AWS IoT Greengrass (IoT)	No	Yes
Transfer files through NFS with a GUI	No	Yes
50 TB (42 TB usable) - US regions only	Yes	No
80 TB (72 TB 72 usable)	Yes	No
100 TB (83 TB usable)	No	Yes
100 TB Clustered (45 TB per node)	No	Yes

Storage Gateway

What is Storage Gateway?



Storage Gateway

AWS Storage Gateway is a bridge between the on-premise data and the cloud data in S3.

It provides seamless integration with data security features between an on-premise environment and AWS environment.

It can be used to store data in the AWS Cloud for scalable and cost-effective storage.

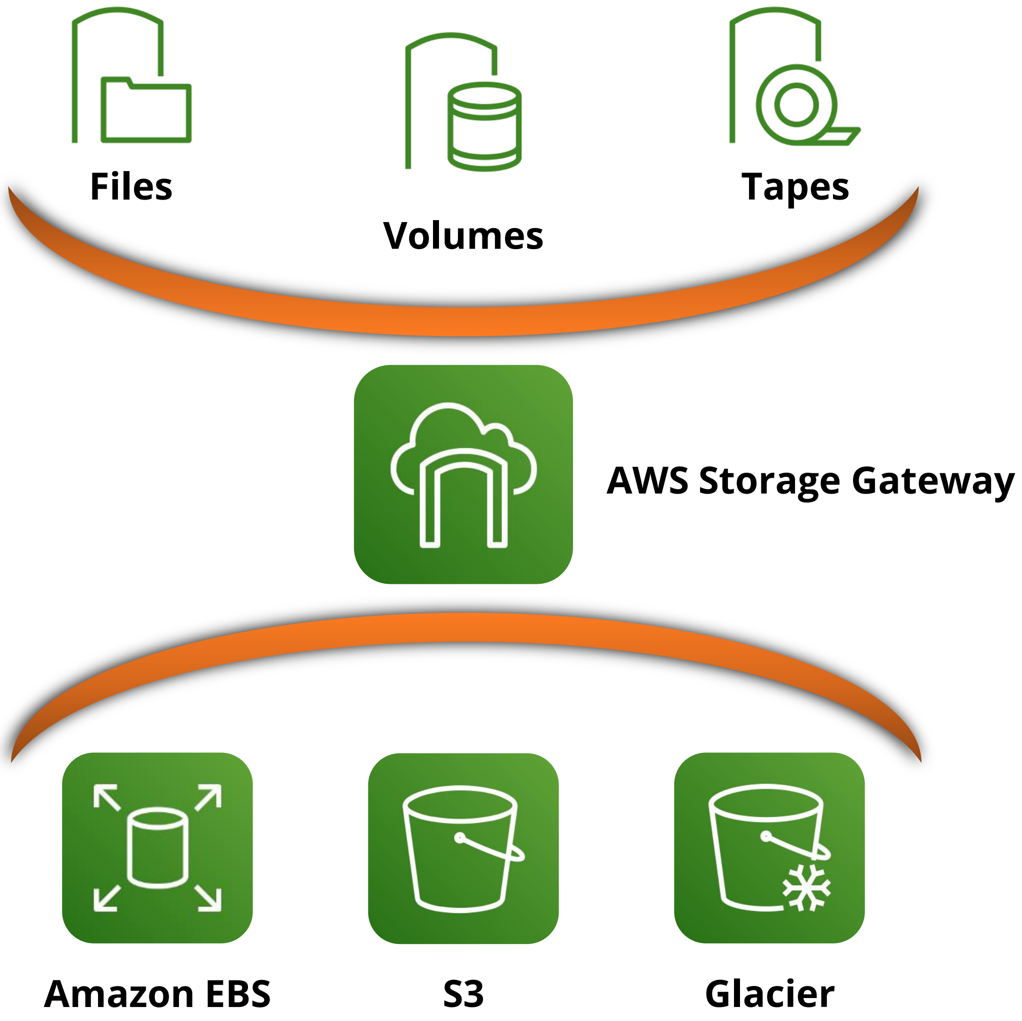
What is Storage Gateway?

Below are the three types of storage gateways:

1 File Gateway

2 Volume Gateway

3 Tape Gateway



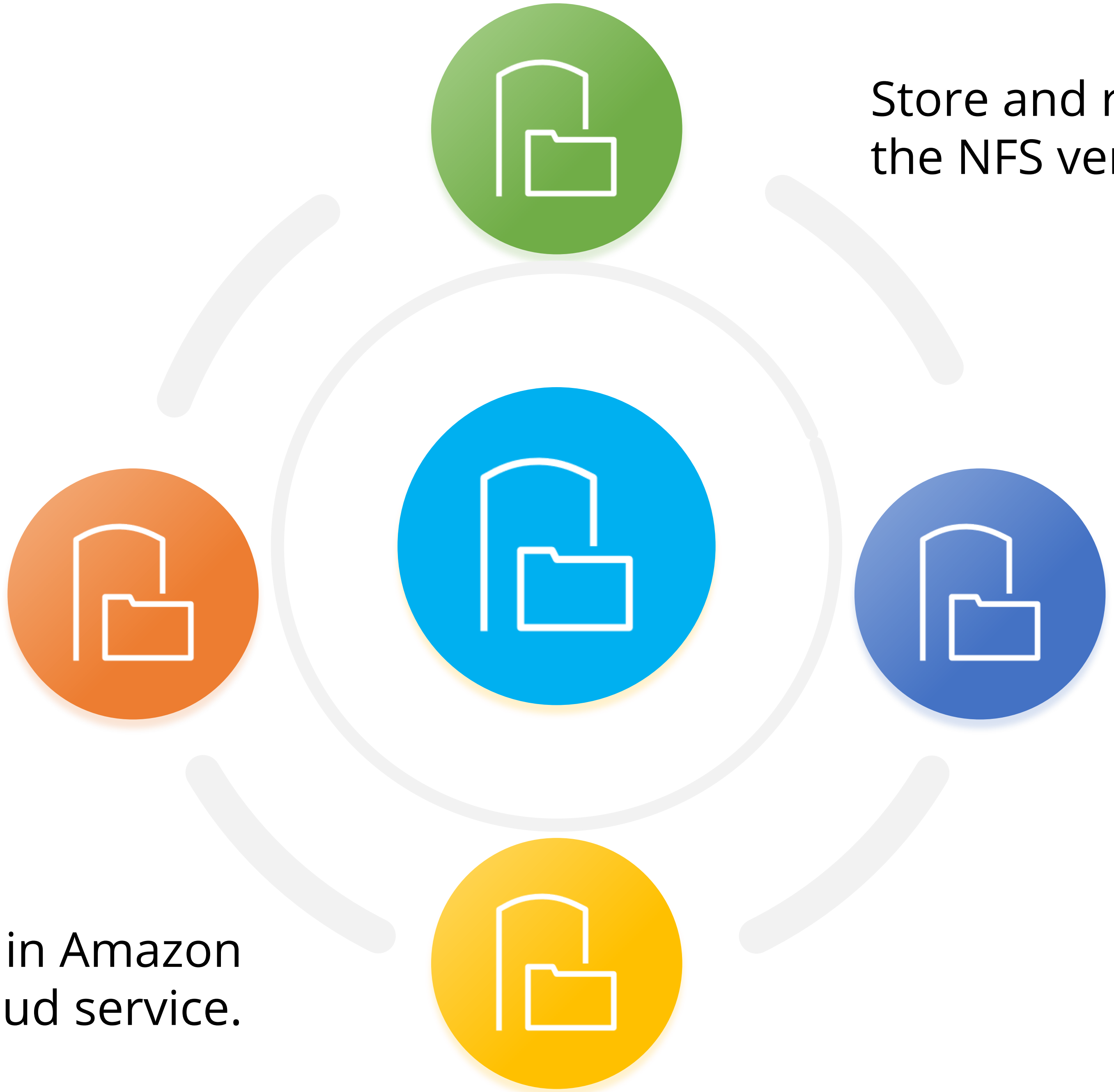
File Gateway

In File Gateway, configured S3 buckets are accessible using the NFS and SMB protocol.

A file gateway is used to support a file interface into Amazon S3 and combines service and a software appliance.

File Gateway

Using File Gateway, we can do the following:



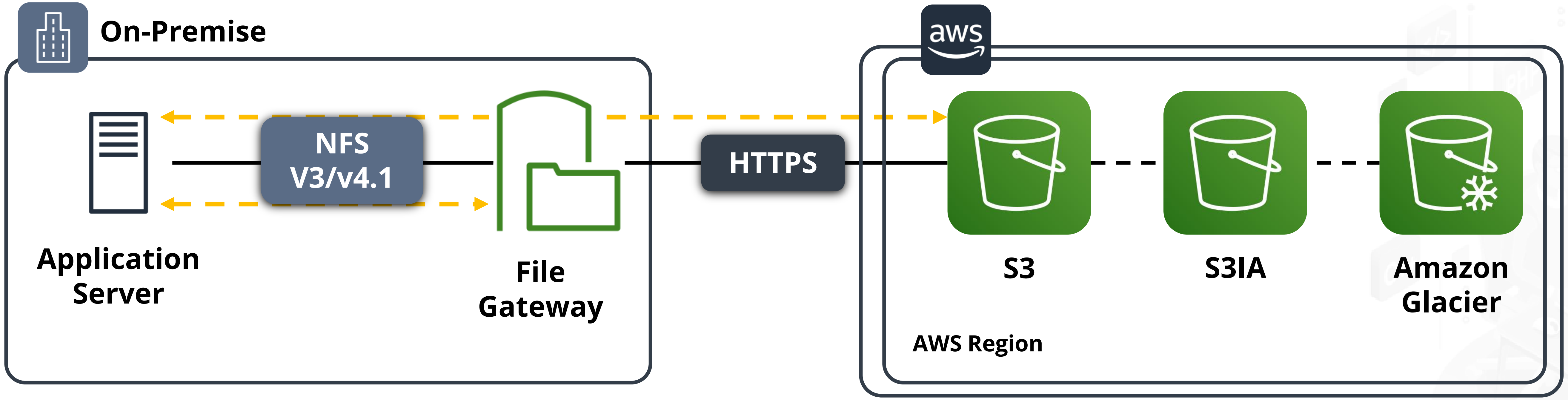
Store and retrieve files directly using the NFS version 3 or 4.1 protocol.

Store and retrieve files directly using the SMB file system version 2 and 3 protocol.

Access the data directly in Amazon S3 from any AWS cloud service.

Manage the Amazon S3 data using lifecycle policies.

File Gateway



Volume Gateway

A volume gateway is used to provide cloud-based storage volumes that can be mounted as Internet Small Computer Interface (ISCSI) device from the on-premise application servers.

The gateway supports the following volume configurations:

01

Cached Volumes:

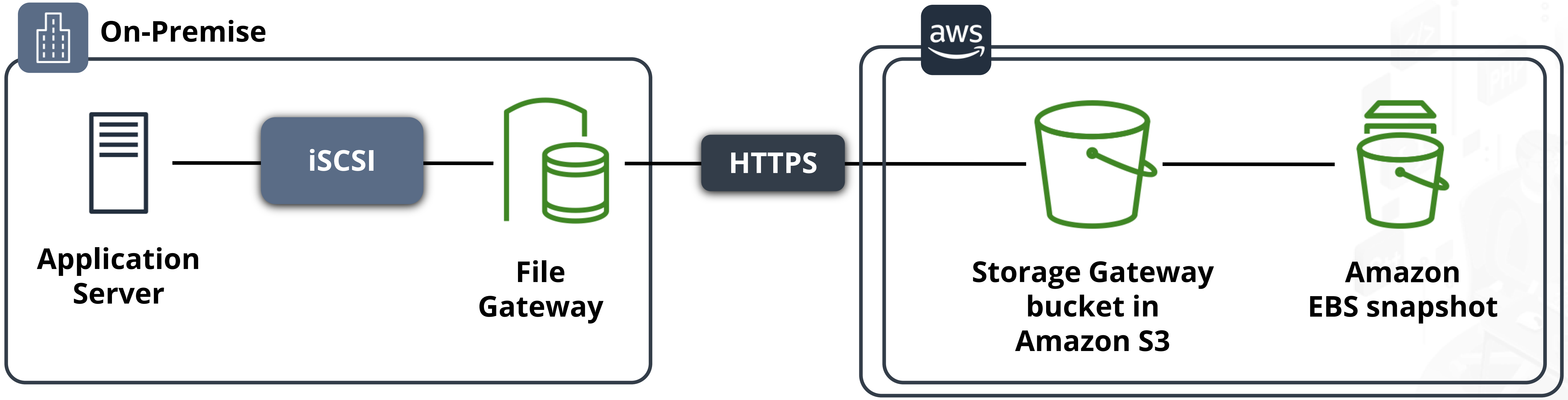
It provides substantial cost savings on primary storage and minimizes the need to scale the storage on-premises.

02

Stored Volumes:

It provides a durable and inexpensive off-site backup that can be recovered to the local data center or Amazon Elastic Compute Cloud (Amazon EC2).

Volume Gateway



Tape Gateway

A tape gateway is used to provide cloud-backed virtual tape storage.

It is deployed into the on-premise environment as a VM.

It is used to backup data cost-effectively and durably.

Tape Gateway

Below are the characteristics of tape gateway:

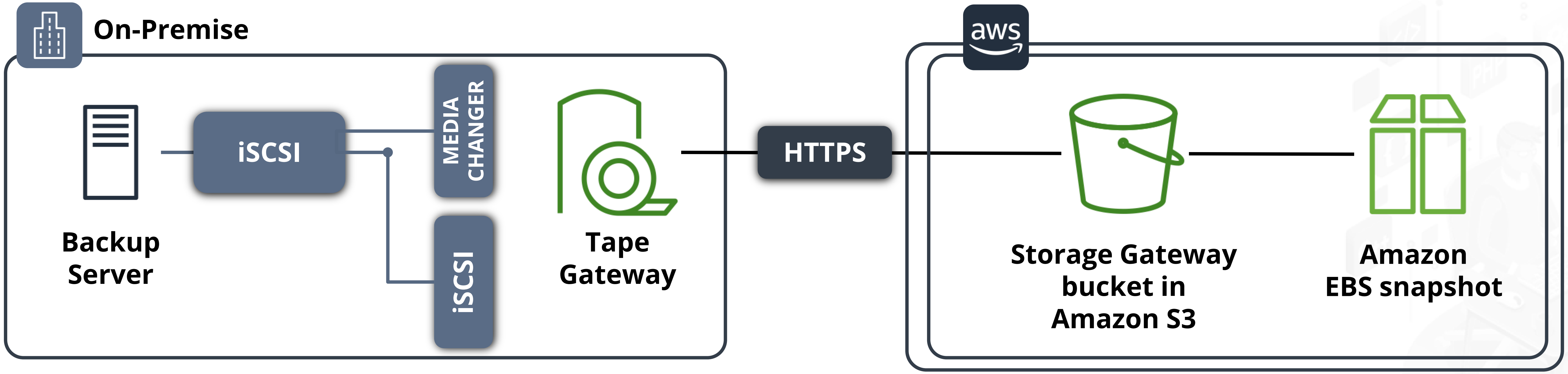
01. Companies use tape gateway to implement the same process, but on the cloud.

02. Backup data using the existing tape-based processes

03. Works with leading backup software vendors

04. Provides a virtual tape infrastructure that scales with the business needs

Tape Gateway



TECHNOLOGY

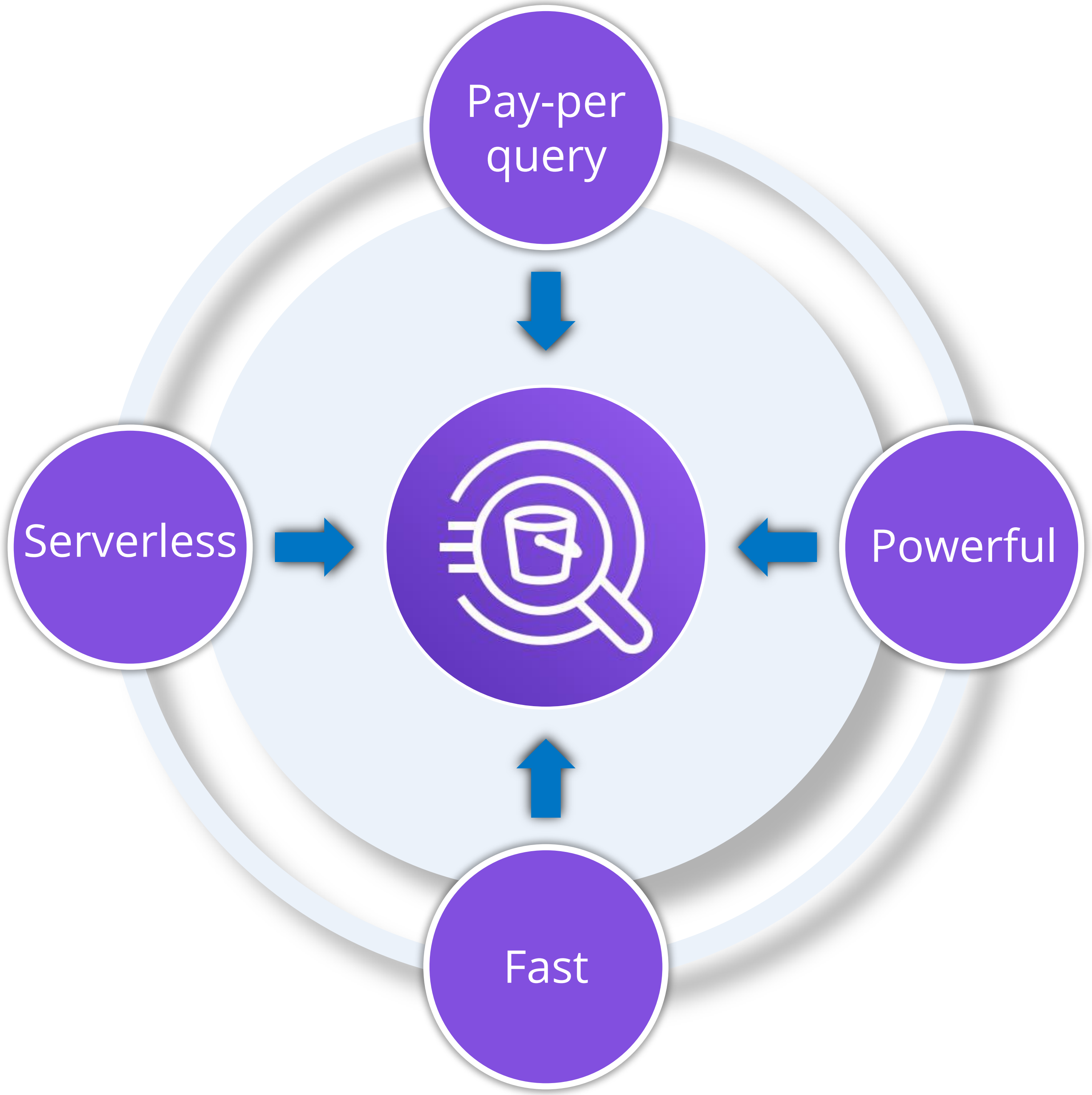
Introduction to Athena

What is Athena?

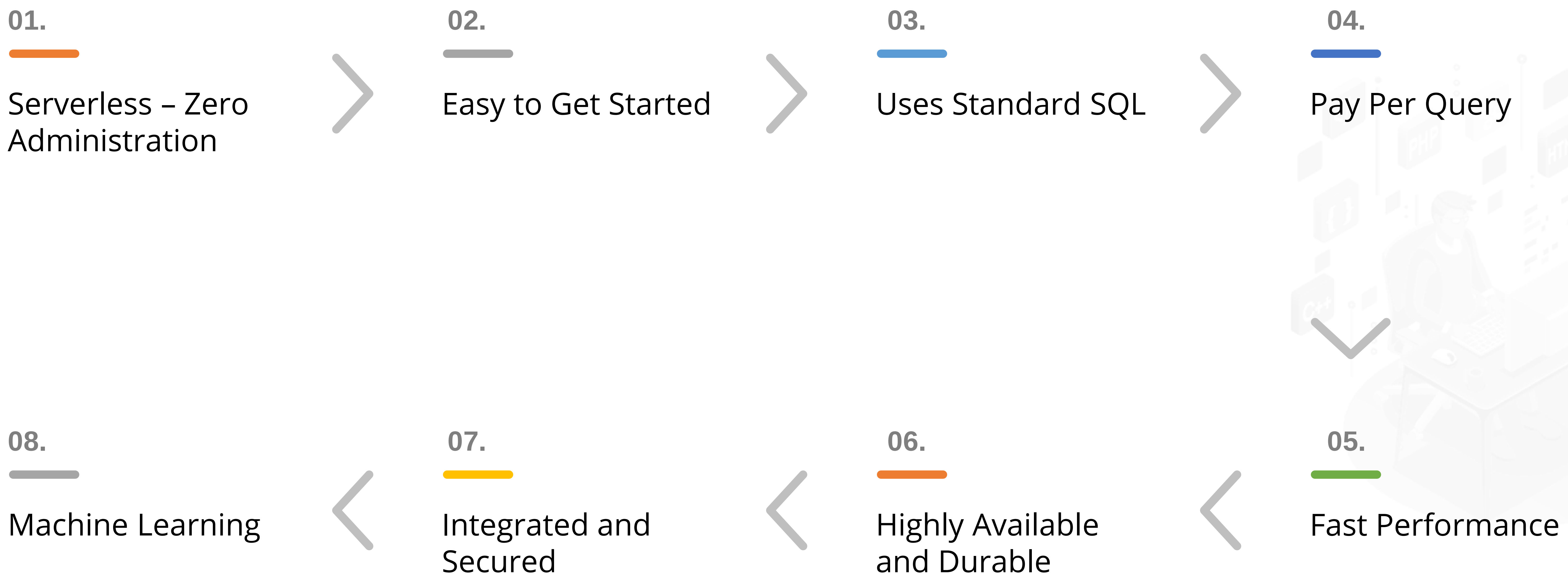
It is an interactive query service that is used to analyze data directly in Amazon S3 using the standard SQL.

Athena can be pointed to the data stored in S3 using a few actions from the AWS Management Console, and standard SQL can be used to run ad-hoc queries to get results in seconds.

Benefits of Athena



Features of Athena



01.

Serverless – Zero Administration

02.

Easy to Get Started

03.

Uses Standard SQL

04.

Pay Per Query

05.

Fast Performance

06.

Highly Available and Durable

07.

Integrated and Secured

08.

Machine Learning

Working with Athena

Duration: 20 Min.

Problem Statement:

Create a query in Athena to perform operations on a specific bucket in S3.

Assisted Practice: Guidelines

Steps to create a query in Athena:

1. Log in to your AWS lab

1. Create two S3 buckets

1. Add logging to the bucket

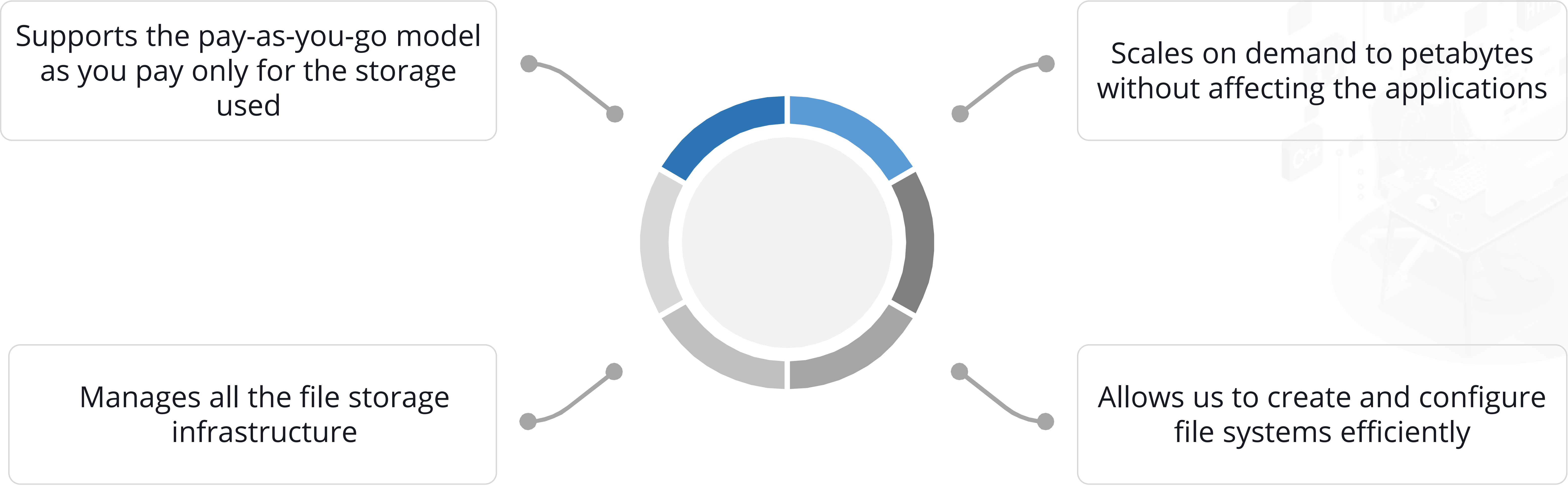
1. Create and execute Athena queries

TECHNOLOGY

Elastic File System

What is EFS?

Amazon Elastic File System (Amazon EFS) is a service that provides a scalable and fully managed elastic NFS file system that can be used with AWS Cloud services and on-premises resources.



Supports the pay-as-you-go model as you pay only for the storage used

Scales on demand to petabytes without affecting the applications

Manages all the file storage infrastructure

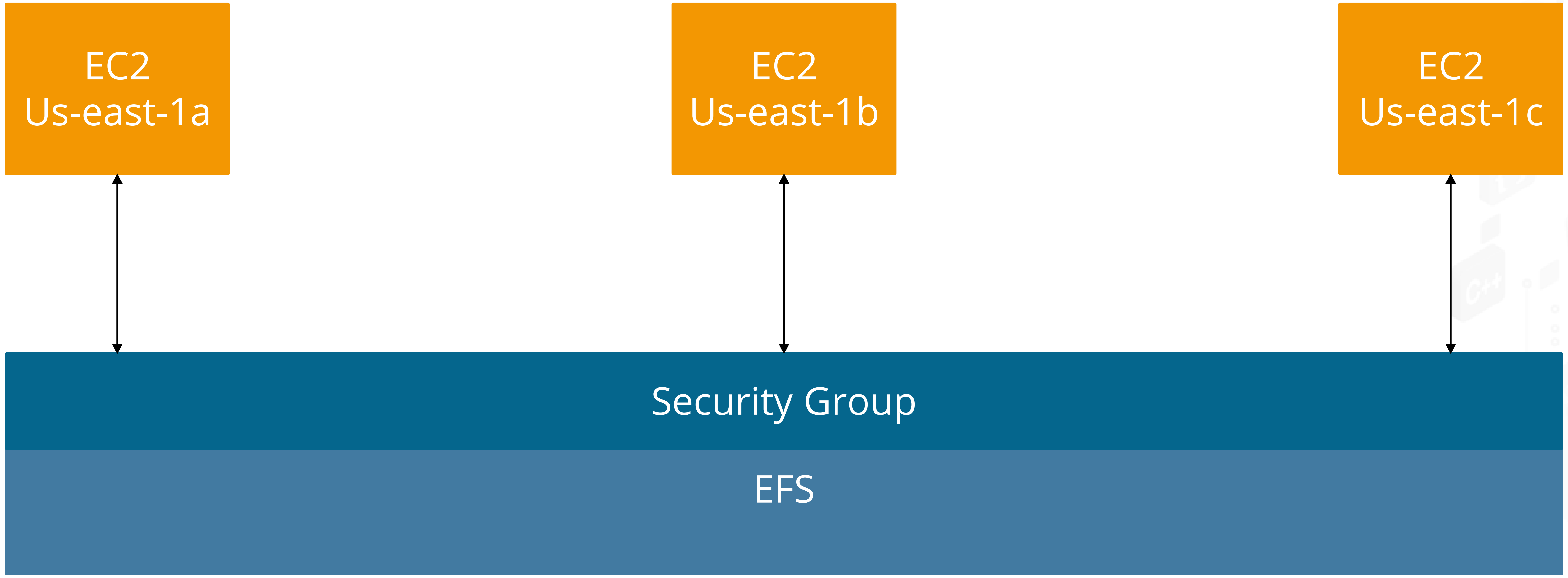
Allows us to create and configure file systems efficiently

What is EFS?

EFS

- Managed NFS, which can be mounted on many EC2s
- Works with EC2 instance in Multi-AZ
- Highly scalable and available service

What is EFS?



Working with EFS

Duration: 20 Min.

Problem Statement:

Customize a file system and access it using a specific EC2 instance. Perform each step and create a report.

Assisted Practice: Guidelines

Steps to customize a file system:

1. Log in to your AWS lab
2. Go to **AWS Console** and select **EFS service**
3. Create a file system and customize it
4. Configure the security credentials
5. Add a key pair
6. Create and generate reports

Key Takeaways

- Amazon S3 lifecycle policies are configured to store objects effectively throughout the lifecycle.
- Key Management System (KMS) is used to create and manage cryptographic keys.
- AMIs created by a developer and made available for another developer are called shared AMIs.
- Athena can be pointed to the data using a few actions from the AWS Management Console, and standard SQL can be used to run ad-hoc queries to get results in seconds.
- EFS manages the file-storage infrastructure.

Lesson-End Project

Problem Statement:

Use S3 Batch Operations to encrypt the existing data in the existing S3 bucket, and use Athena to check the bytes of data uploaded and downloaded from the monitored bucket.

Background of the problem statement:

Your company has got a contract for storage and data management from a client. As you are a senior SysOps engineer, this task has been assigned to you.

LESSON-END PROJECT