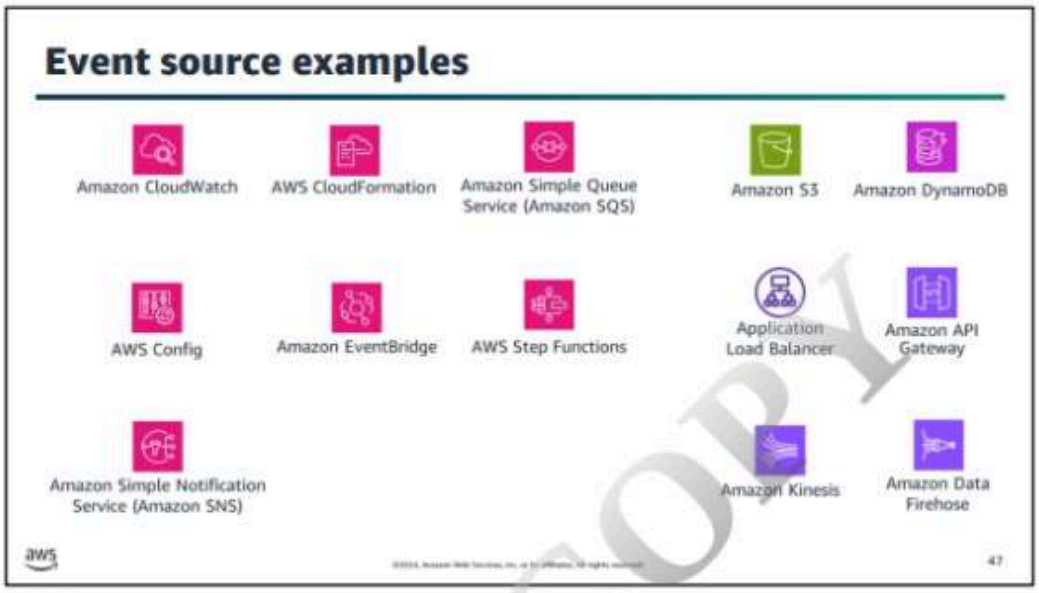
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



You can manually invoke a Lambda function, or it can be invoked by an AWS service. Some of these services are covered in more detail later in this course.

For more information about Lambda event sources, see "Services that can invoke Lambda functions" in the *AWS Lambda Developer Guide* at https://docs.aws.amazon.com/lambda/latest/dg/lambda-services.html#listing-of-services-and-links-to-more-information.

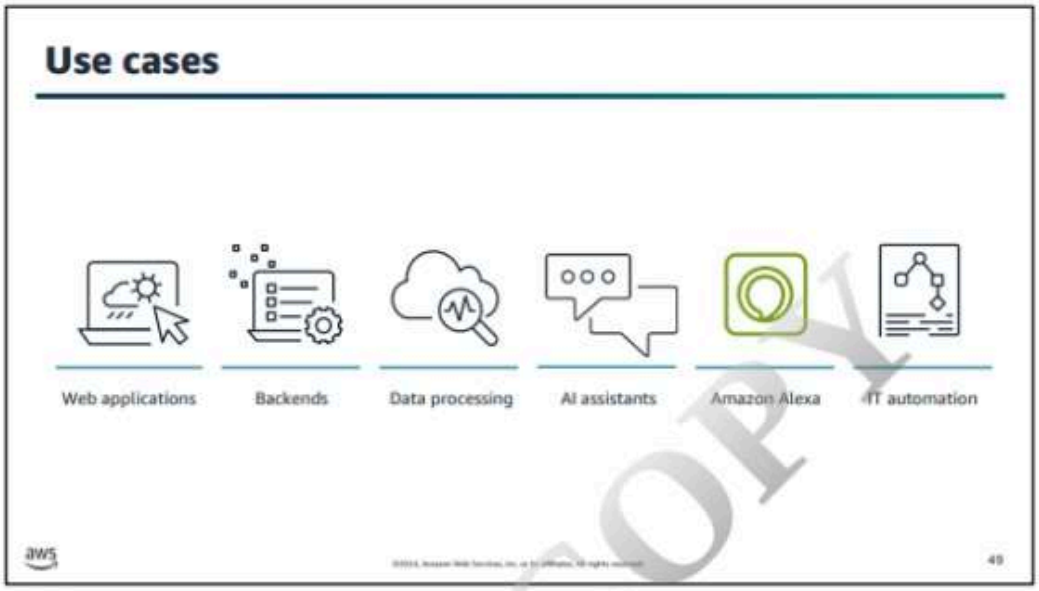
If you want a way to configure an HTTPS endpoint in front of your function, consider using AWS Lambda Function URLs.

For more information about Lambda Function URLs, see "Announcing AWS Lambda Function URLs: Built-in HTTPS Endpoints for Single-Function Microservices" in the *AWS News Blog* at https://aws.amazon.com/blogs/aws/announcing-aws-lambda-function-urls-built-in-https-endpoints-for-single-function-microservices/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



AWS Lambda is a cost-effective solution for a number of tasks. Consider the following use cases:

Web applications
- Static websites
- Complex web applications
- Packages for Flask and Express

Backends
- Applications and services
- Mobile
- Internet of things (IoT)

Data processing
- Real-time processing
- Amazon EMR
- AWS Batch

AI assistants
- Powering chatbot logic

Amazon Alexa
- Powering voice-activated applications
- Alexa Skills Kit

IT automation:
- Policy engines
- Extending AWS services
- Infrastructure management

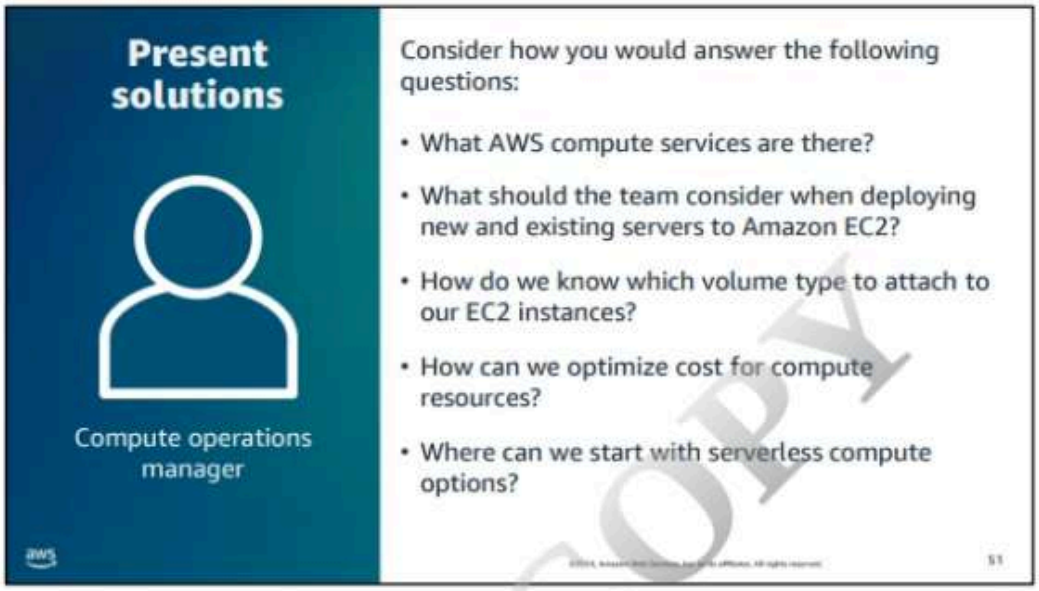
DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Imagine that you are now ready to talk to the compute operations manager and present solutions that meet their architectural needs. Think about how you would answer the questions from the beginning of the lesson about compute.

Your answers should include the following solutions:

- Identify AWS services that are used to create compute capacity in your accounts.
- Consider things like AMIs, instance types and sizes, authentication, placement, tenancy, scripting, and metadata in Amazon EC2.
- Build and attach Amazon EBS and instance store volumes based on your needs.
- Plan for the future cost of compute by using a combination of Savings Plans, Spot Instances, and On-Demand Instances.
- Use AWS Lambda to reduce manual work and manage cost for developers.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Module review

In this module, you learned about the following:

- ✓ Compute services
- ✓ EC2 instances
- ✓ Instance storage
- ✓ Amazon EC2 pricing options
- ✓ AWS Lambda

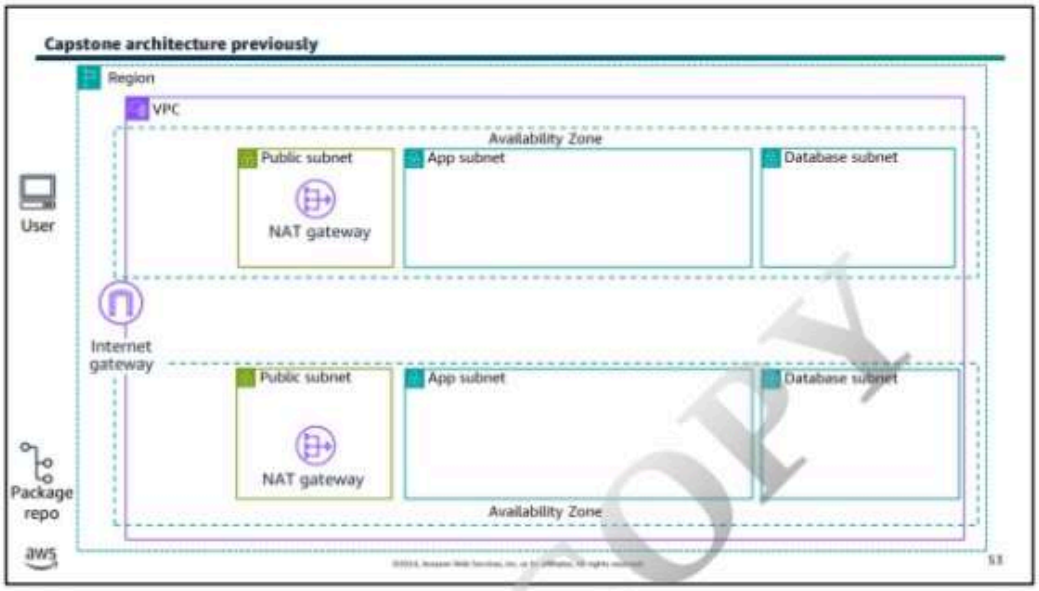
Next, you review the following:

- Capstone check-in
- Lab introduction
- Knowledge check

52

DO NOT COPY
javierartiga.espublico@gmail.com

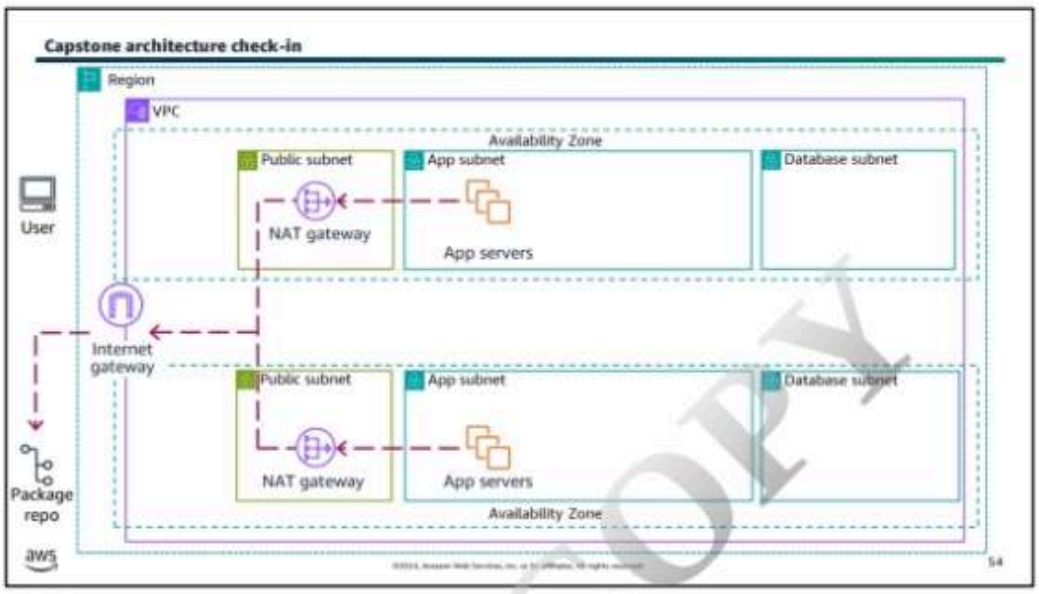
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



In this module, you explored AWS compute services such as Amazon EC2 and AWS Lambda.

Review the capstone architecture to explore some of the design decisions. This architecture helps you provide the following benefits:

- EC2 instances can quickly create your WordPress website. You can use an AWS Marketplace AMI for WordPress, or you can use a script to install WordPress with user data at EC2 launch time.
- NAT gateways allow outbound traffic from your EC2 instances in private subnets. In this way, your private EC2 instances can connect to the internet through the internet gateway. It stops external sources from initiating connections to your private EC2 instances.

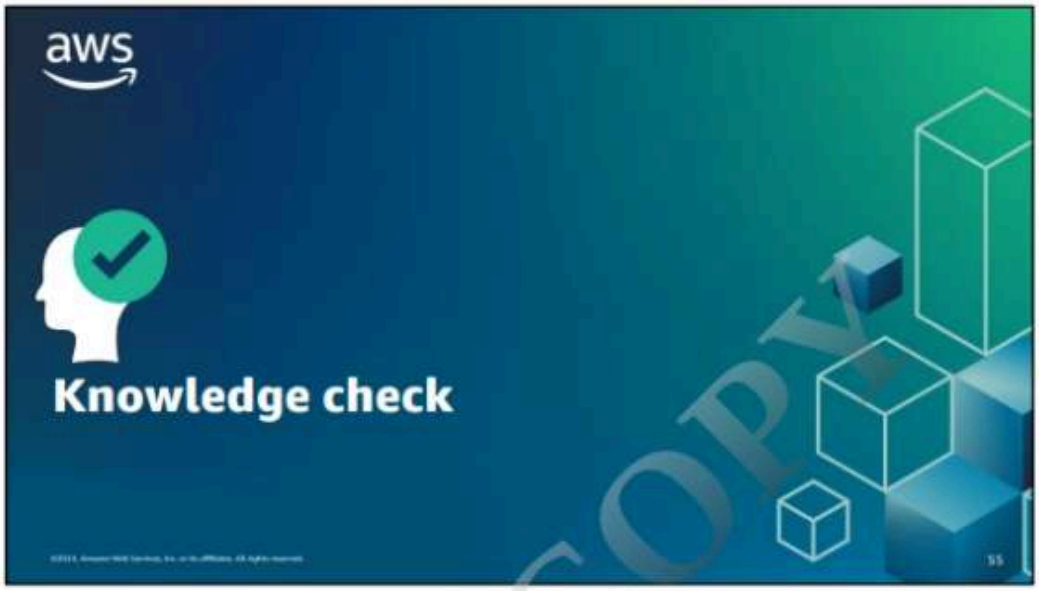
You place the EC2 instances into your application subnet for the three-tier architecture. Think about the kind of manual effort that will be required if your team is managing up to four EC2 instances at once.

What if you want to scale your application in or out to save on cost without sacrificing performance? What can you do to automate the creation of new EC2 instances? You'll explore scaling and monitoring tools, load balancers, and EC2 launch templates later in this course.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 1

Which statements are true of Amazon Machine Images (AMIs)? (Select TWO.)

Choice	Response
A	AMIs can specify the subnets for launch.
B	AMIs can include block device mapping that specifies the volumes to attach to the Amazon EC2 instance when it is launched.
C	AMIs can only be obtained from the AWS Marketplace.
D	You can launch multiple instances from a single AMI.
E	AMIs can only be used by users within a single account.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 1 and answer

Which statements are true of Amazon Machine Images (AMIs)? (Select TWO.)

Choice	Response
A	AMIs can specify the subnets for launch.
B correct	AMIs can include block device mapping that specifies the volumes to attach to the Amazon EC2 instance when it is launched.
C	AMIs can only be obtained from the AWS Marketplace.
D correct	You can launch multiple instances from a single AMI.
E	AMIs can only be used by users within a single account.

The correct answers are B and D.

AMIs include block device mapping that specifies the volumes to attach to the Amazon EC2 instance when it is launched. You can launch multiple instances from a single AMI.

An Amazon Machine Image (AMI) provides the information that is required to launch an instance. You must specify an AMI when you launch an instance.

You can launch multiple instances from a single AMI when you need multiple instances with the same configuration. You can use different AMIs to launch instances when you need instances with different configurations.

For more information about AMIs, see "Amazon Machine Images (AMI)" in the *Amazon EC2 User Guide* at https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html.

DO NOT COPY javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 2

In the instance type name m8g.2xlarge, which aspect of the name indicates the generation of the instance family?

Choice	Response
A	m
B	g
C	2xlarge
D	8

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 2 and answer

In the instance type name m8g.2xlarge, which aspect of the name indicates the generation of the instance family?

Choice	Response
A	m
B	g
C	2xlarge
D correct	8

The correct answer is D, the number 8.

The second component of the instance family is the generation.

For more information about the series and options for instance type see "Amazon EC2 instance type naming conventions" at https://docs.aws.amazon.com/ec2/latest/instancetypes/instance-type-names.html.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 3

Which statements are true regarding Lambda? (Select TWO.)

Choice	Response
A	Functions currently only support Python.
B	You are responsible for updating and patching Lambda servers.
C	Functions can be allocated up to 10 GB of memory.
D	Functions can run for a maximum of 15 minutes.
E	Functions require a security group.

60

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 3 and answer

Which statements are true regarding Lambda? (Select TWO.)

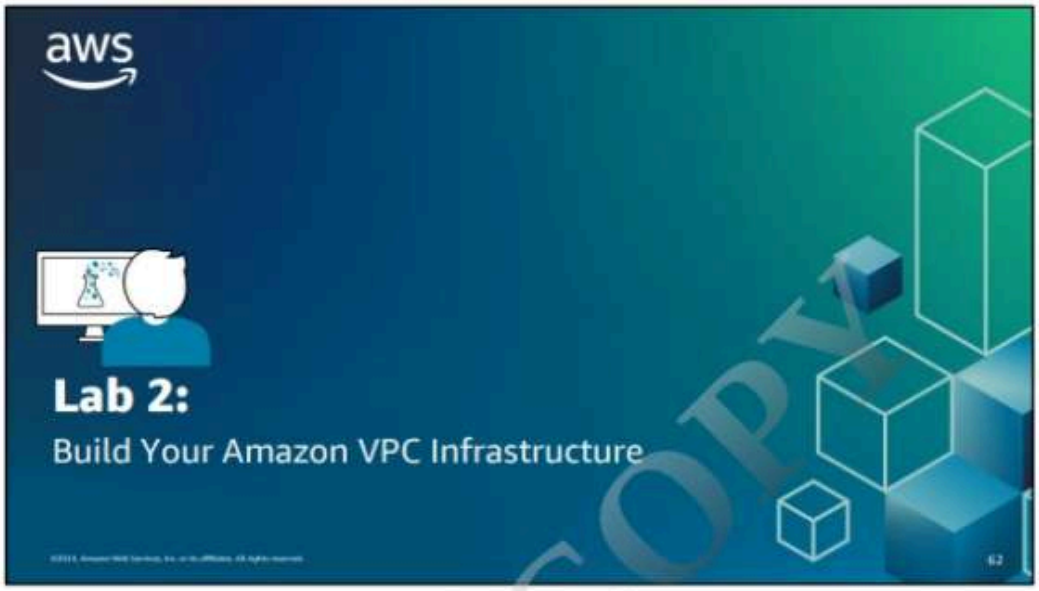
Choice	Response
A	Functions currently only support Python.
B	You are responsible for updating and patching Lambda servers.
C correct	Functions can be allocated up to 10 GB of memory.
D correct	Functions can run for a maximum of 15 minutes.
E	Functions require a security group.

The correct answers are C and D.

Lambda functions can be allocated up to 10 GB of memory, and they can run for up to 15 minutes.

DO NOT COPY
javierartiga.espublico@gmail.com

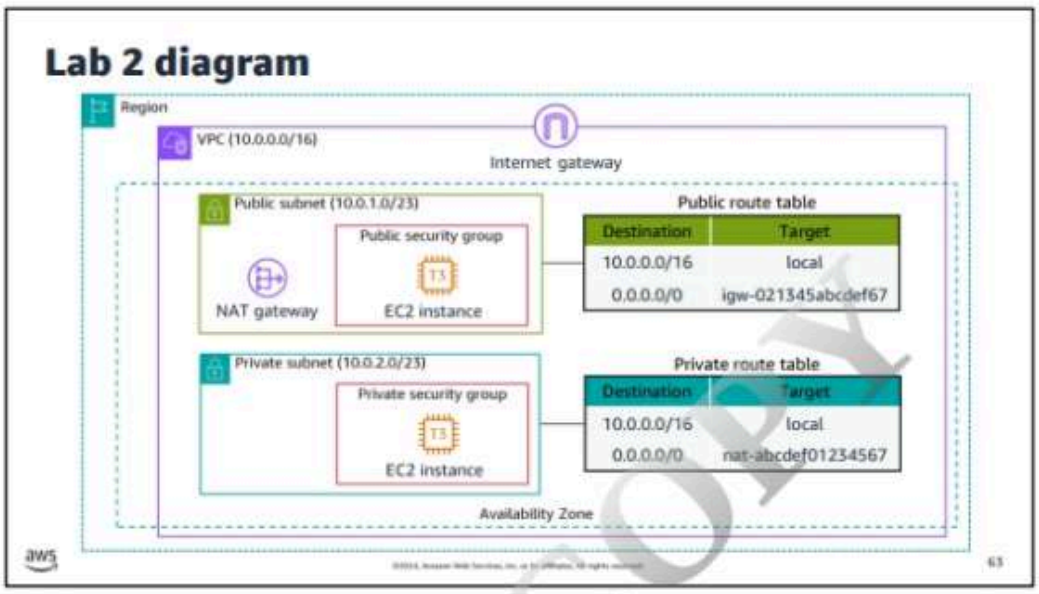
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Lab 2 uses a VPC across one Availability Zone with a public and private subnet. Each subnet contains a T3 EC2 instance. Separate security groups control the EC2 instances. You manage route tables separately for the public and private subnet. A NAT gateway in the public subnet receives traffic from the private subnet. An internet gateway receives traffic from the public subnet.

Review the diagram here, and then read the lab tasks on the following page.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

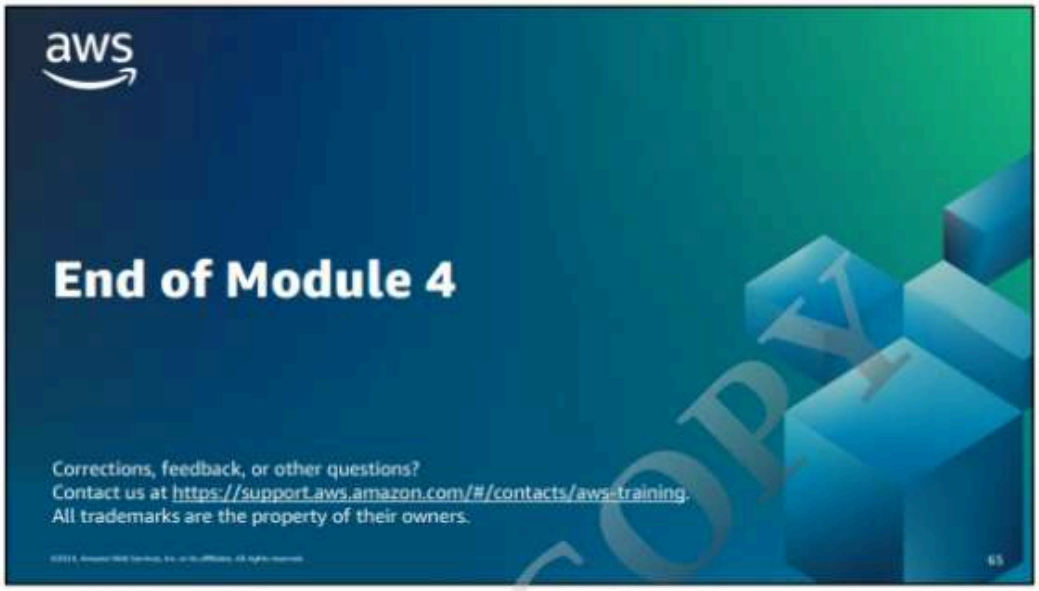
Lab tasks

Task Step	Action
1	Create an Amazon VPC in a Region.
2	Create public and private subnets.
3	Create an internet gateway.
4	Route internet traffic in the public subnet to the internet gateway.
5	Create a public security group.
6	Launch an Amazon EC2 instance into a public subnet.
7	Connect to a public instance through HTTP.
8	Connect to the Amazon EC2 instance in the public subnet.
9	Create a NAT gateway and configure routing in the private subnet.
10	Create a security group for private resources.
11	Launch an Amazon EC2 instance into a private subnet.
12	Connect to the Amazon EC2 instance in the private subnet.
Optional 1	Test connectivity to the private instance from the public instance.
Optional 2	Retrieve instance metadata.

64

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

**Accessibility note for screen reader users: To make code examples in this presentation more accessible, Start of code and End of code labels have been added before and after blocks of code on slides in this presentation.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Poll question

What type of storage are you or your teams using in your environments? Choose all that apply.

A. Block storage

B. File storage

C. Object storage

D. I don't know

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Module overview

- Business request
- Storage services
- Amazon Simple Storage Service (Amazon S3)
- Shared file systems
- Data migration tools
- Present solutions
- Capstone check-in
- Knowledge check

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Imagine that your storage team lead meets with you so that they can understand how storage concepts translate to Amazon Web Services (AWS). Here are some questions that they are asking you.

At the end of this module, you meet with the storage team lead and present some solutions.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

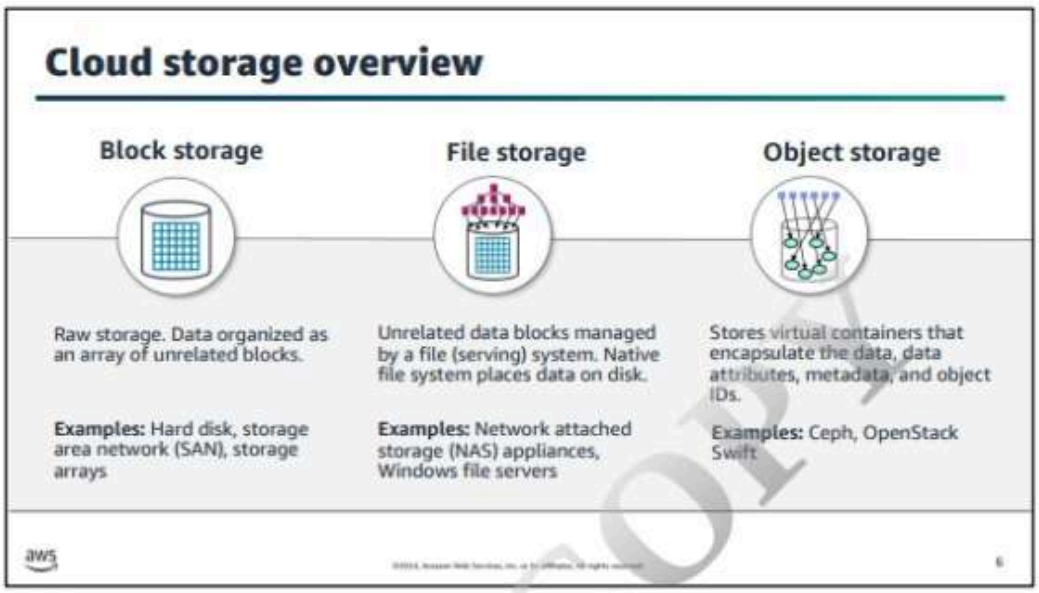
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

The storage team lead asks, "What are some services to consider when looking at block, file, and object storage?" The storage team must compare and contrast AWS storage products and services, based on business scenarios.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



There are three types of cloud storage: object, file, and block. Each storage option has a unique combination of performance, durability, cost, and interface.

Block storage – Enterprise applications such as databases or enterprise resource planning (ERP) systems often require dedicated, low-latency storage for each host. This storage is similar to direct-attached storage (DAS) or a storage area network (SAN). Block-based cloud storage solutions such as Amazon Elastic Block Store (Amazon EBS) are provisioned with each virtual server and offer the ultra-low latency required for high-performance workloads.

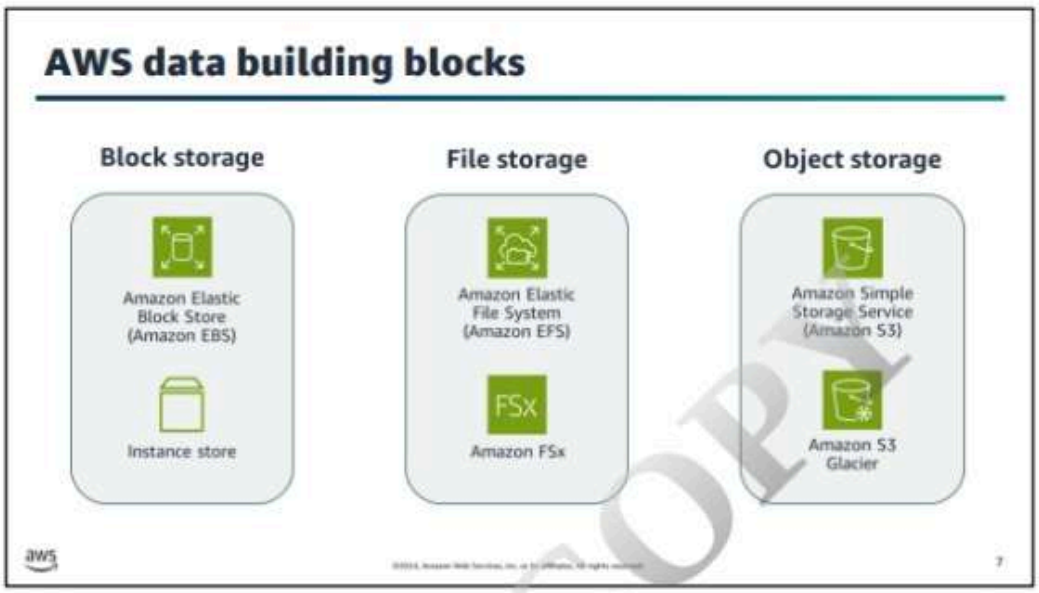
File storage – Many applications must access shared files and require a file system. This type of storage is often supported with a network attached storage (NAS) server. File storage solutions such as Amazon Elastic File System (Amazon EFS) are ideal for use cases such as large content repositories, development environments, media stores, or user home directories.

Object storage – Applications developed in the cloud need the vast scalability and metadata of object storage. Object storage solutions such as Amazon Simple Storage Service (Amazon S3) are ideal for building modern applications. Amazon S3 provides scale and flexibility. You can use it to import existing data stores for analytics, backup, or archive.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



AWS has solutions for your block, file, and object storage needs. Block storage has already been discussed in the Compute module. In this module, you will learn about the following services:

- Amazon EFS and Amazon FSx for file storage
- Amazon S3 and Amazon S3 Glacier for object storage

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

The storage team lead asks, "How do we choose the right object storage solution for my use case?" The team needs to identify object storage classes to meet business needs.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Amazon S3 is *object-level storage*. An object includes file data, metadata, and a unique identifier. Object storage does not use a traditional file and folder structure.

Amazon S3 storage tiers are all designed to provide 99.999999999 percent (11 9's) of data durability of objects over a given year. By default, data in Amazon S3 is stored redundantly across multiple facilities and multiple devices in each facility. Amazon S3 can be accessed through the web-based AWS Management Console, programmatically through the API and SDKs, or with third-party solutions (which use the API and SDKs).

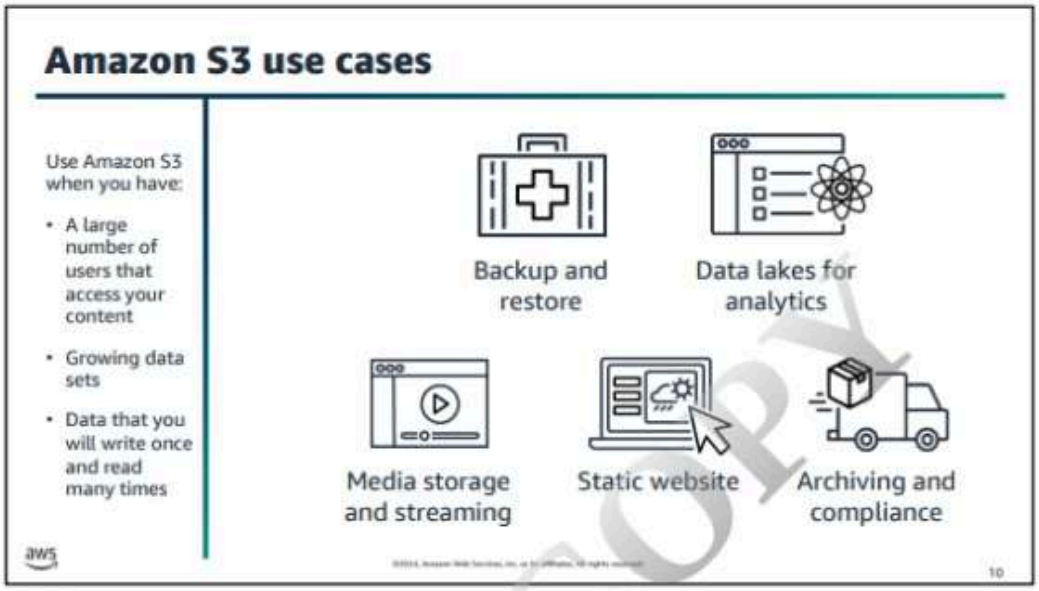
With Amazon S3, you can do the following:

- **Accelerate innovation** – Integrate S3 buckets as storage solutions for static files and rely less on traditional file systems.
- **Increase agility** – With hosted object storage, you won't need to expand your storage as the quantity and size of data grows. Individual objects cannot be larger than 5 terabytes (TB). However, you can store as much total data as you need.
- **Reduce cost** – Use the variety of storage tiers in Amazon S3 to spend less on infrequently accessed data. Archive data in Amazon S3 for your long-term storage needs.
- **Strengthen security** – Store your data in Amazon S3 and secure it from unauthorized access with encryption features and access management tools. Amazon S3 maintains compliance programs that help you to meet regulatory requirements, such as the following:
    - Payment Card Industry Data Security Standards (PCI-DSS)
    - Health Insurance Portability and Accountability Act (HIPAA) and Health Information Technology for Economic and Clinical Health (HITECH)
    - Federal Risk and Authorization Management Program (FedRAMP)
    - European Union (EU) Data Protection Directive
    - Federal Information Security Management Act (FISMA)

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 provides you with flexible, low-cost object storage solutions. Some use cases for Amazon S3 include the following:

- **Backup and restore** – You can use Amazon S3 to store and retrieve any amount of data, at any time. You can use Amazon S3 as the durable store for your application data and file-level backup and restore processes. Amazon S3 is designed for 99.999999999 percent durability, or 11 9's of durability.
- **Data lakes for analytics** – Run big data analytics, AI, machine learning (ML), and high-performance computing (HPC) applications to unlock data insights.
- **Media storage and streaming** – You can use Amazon S3 with Amazon CloudFront's edge locations to host videos for on-demand viewing in a secure and scalable way. Video on demand (VOD) streaming means that your video content is stored on a server, and viewers can watch it at any time. You learn more about Amazon CloudFront later in this course.
- **Static website** – You can use Amazon S3 to host a static website. On a *static* website, individual webpages include static content. They might also contain client-side scripts. Object storage in Amazon S3 helps you to manage data access, replications, and data protection for static files.
- **Archiving and compliance** – Replace your tape with low-cost cloud backup workflows, while maintaining corporate, contractual, and regulatory compliance requirements.

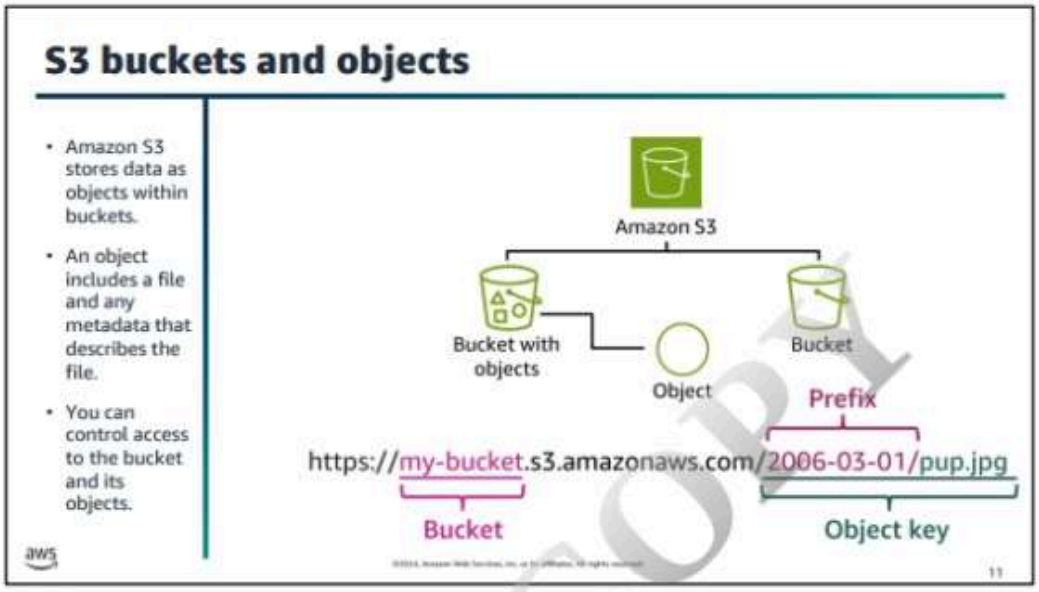
Consider other storage solutions when you are frequently changing the data, or when you have block storage requirements.

Learn more about how businesses are solving their storage challenges. For more information, see "How AWS Partners Are Utilizing Amazon S3 to Help Customers Solve for Scale" in the *AWS Partner Network (APN) Blog* at https://aws.amazon.com/blogs/apn/how-aws-partners-are-utilizing-amazon-s3-to-help-customers-solve-for-scale/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 stores data as objects within buckets. An *object* is composed of a file and any metadata that describes that file. To store an object in Amazon S3, you upload the file into a *bucket*.

When you upload a file, you can set permissions on the object and add metadata. You can have one or more buckets in your account. For each bucket, you control who can create, delete, and list objects in the bucket. Choose the geographical AWS Region where Amazon S3 will store the bucket and its contents. You can access logs for the bucket and its objects. With Amazon S3, you can have up to 100 buckets in each account.

The diagram contains a virtual-hosted-style access URL made up of a bucket and an object key. An object key is the unique identifier for an object in a bucket. The combination of a bucket, key, and version ID uniquely identifies each object. Every object can be uniquely addressed through the combination of the web service endpoint, bucket name, key, and optionally, a version.

For example, in the URL https://my-bucket.s3.amazonaws.com/2006-03-01/pup.jpg, *my-bucket* is the name of the bucket and *2006-03-01/pup.jpg* is the key. The 2006-03-01/ portion of the object key is called the *prefix*. For more information about creating object key names, see "Creating object key names" in the *Amazon Simple Storage Service User Guide* at https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-keys.html.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Think about how physical objects in an office building or home are protected. You have more than one option for how to secure your items:

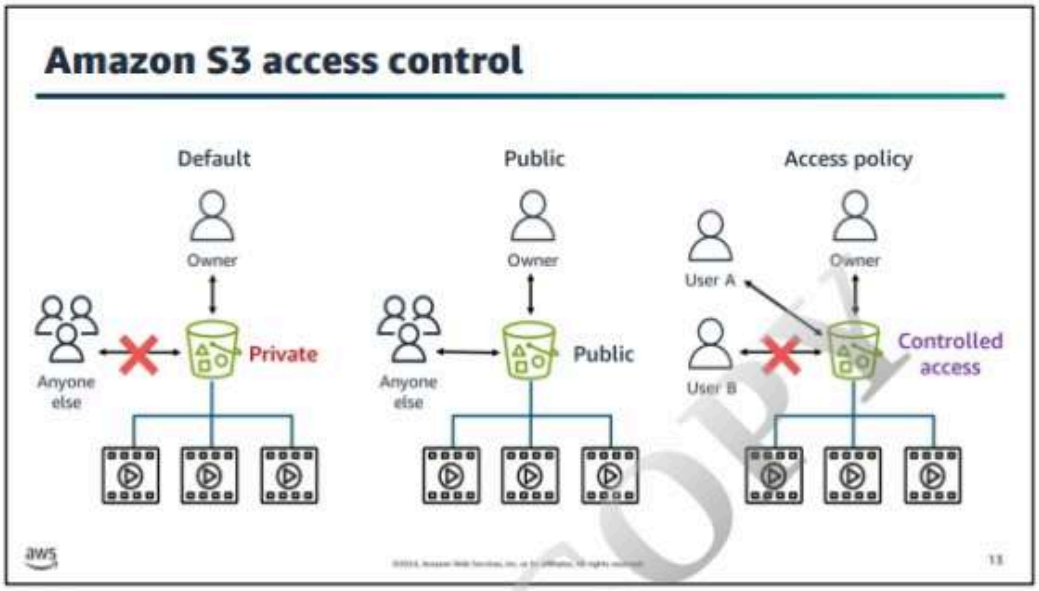
- Security guards
- A safe or a locker
- A key or key card system
- Rules for entry
- Employees only area

You use some or all of these options depending on the type of objects you are caring for.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



By default, all Amazon S3 resources—buckets, objects, and related resources (for example, lifecycle configuration and website configuration)—are private. Only the resource owner, an AWS account that created it, can access the resource. The resource owner can grant access permissions to others by writing access policies.

You can make a resource in Amazon S3 public, which allows access to anyone. However, most Amazon S3 use cases do not require public access. Amazon S3 usually stores data from other applications. Public access is not recommended for these types of buckets. Amazon S3 includes a *block public access* feature. This feature acts as an additional layer of protection to prevent accidental exposure of customer data.

The resource owner can also provide controlled access to a resource. You can grant access permissions to others by writing access policies.

For more information about Amazon S3 access control, see "Blocking public access to your Amazon S3 storage" in *Amazon Simple Storage Service User Guide* at https://docs.aws.amazon.com/AmazonS3/latest/userguide/access-control-block-public-access.html.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Bucket policies

- Resource-based policy for an S3 bucket
- Controls access to a bucket without managing permissions in AWS Identity and Access Management (IAM)

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": "*",
      "Action": [
      "s3:ListBucket",
    "s3:GetObject"
    ],
    "Resource": [
     "arn:aws:s3:::doc-example-bucket",
     "arn:aws:s3:::doc-example-bucket/*"
      ]
    }
  ]
}
```

Bucket policy in JSON format

15

You can create and configure bucket policies to grant permission to your Amazon S3 buckets and objects.

Bucket policies are resource-based policies for your S3 buckets. Access control for your data is based on policies, such as IAM policies, S3 bucket policies, and AWS Organizations service control policies (SCPs).

Use JSON-based access policy language to write your bucket policy. You can use it to add or deny permissions for the objects in your bucket.

In the example, the bucket policy allows any principal to list the bucket and get any object from the bucket. You should consider limiting public access to buckets and objects in this way. Amazon S3 has tools that help you to prevent overly-permissive public buckets.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Amazon S3 Block Public Access

Block public ...

- Access to buckets and objects granted through *new* ACLs
- Access to buckets and objects granted through *any* ACLs
- Access to buckets and objects granted through a *new* public bucket or access point policies
- *Cross-account access* to buckets and objects through *any* public bucket or access point policies

16

Sometimes you want to make sure that no matter what, a bucket will never see public access. Newly created Amazon S3 buckets and objects are private and protected by default. With Amazon S3 access control, you can stop all access to the bucket from ACLs, bucket polices, and cross-origin resource sharing (CORS) settings.

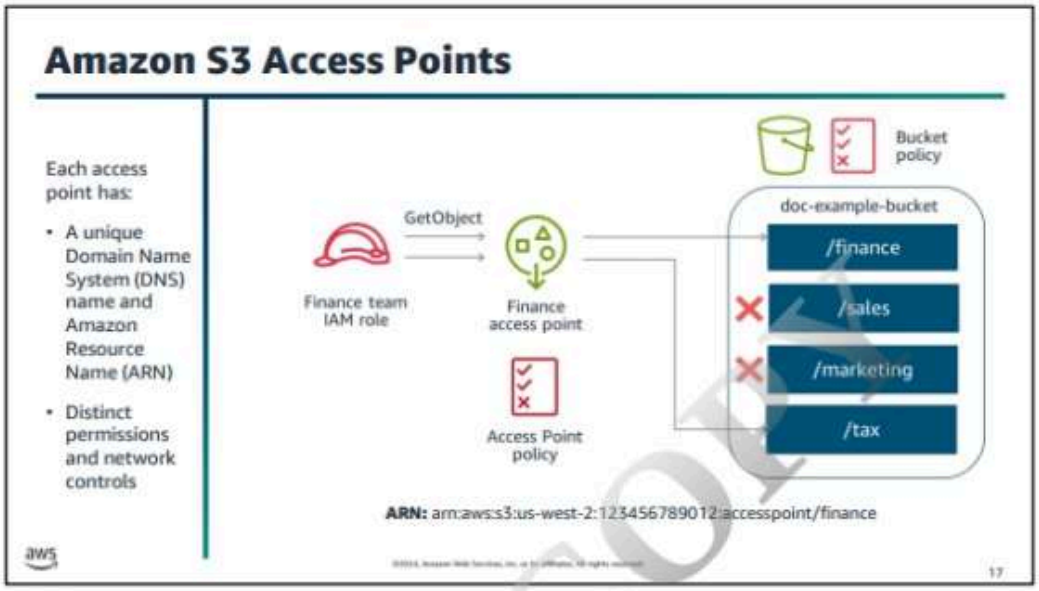
Public access is granted to buckets and objects through ACLs, bucket policies, or both. To avoid public access to all of your S3 buckets and objects, turn on *block all public access* at the account level. These settings apply account-wide for all current and future buckets.

You can protect yourself and your organization from information leaks by using S3 Block Public Access settings. Turn it on to prevent your operators from accidentally opening your buckets to the public.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 Access Points simplify managing data access at scale for shared datasets in Amazon S3. S3 Access Points are named network endpoints that you can use to perform S3 object operations, such as GetObject and PutObject. Access points are attached to buckets. Each access point has distinct permissions and network controls that Amazon S3 applies for any request that is made through that access point.

In the example, a finance employee assumes the finance team IAM role and sends a GetObject request to your *finance* access point. With the access point policy, the finance role can get objects in doc-example-bucket with the prefixes /finance and /tax. The finance role does not have access to your sales and marketing prefixed objects or any other objects in your S3 bucket. The S3 bucket policy allows the finance access point to have access to your bucket.

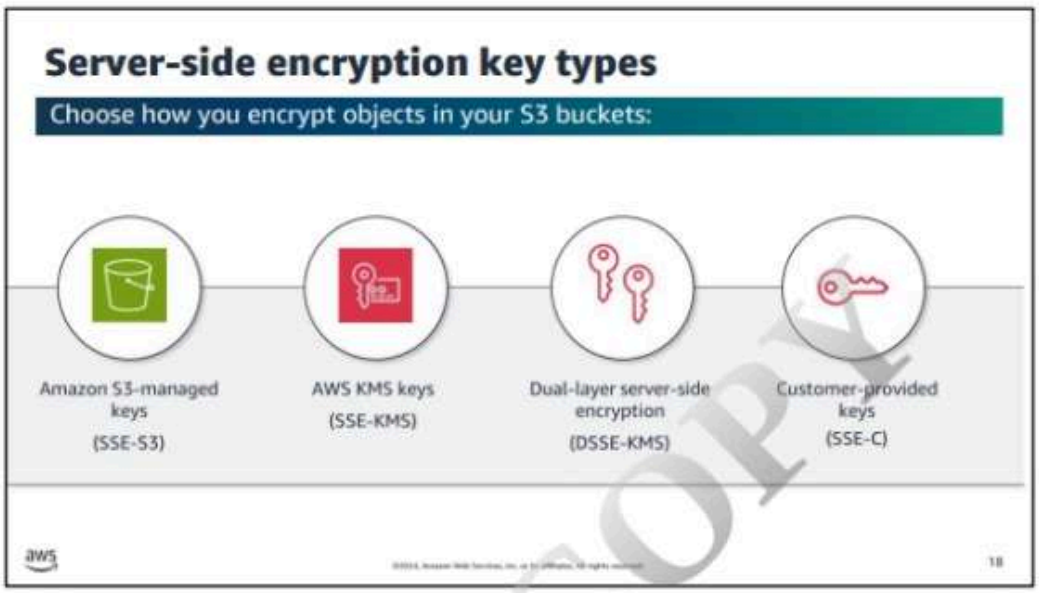
Each access point enforces a customized access point policy that works in conjunction with the bucket policy attached to the underlying bucket. To restrict Amazon S3 data access to a private network, you can configure an access point to accept requests only from a virtual private cloud (VPC). You can also configure custom block public access settings for each access point.

You can only use access points to perform operations on objects. You can't use access points to perform other Amazon S3 operations, such as modifying or deleting buckets. S3 Access Points work with some, but not all, AWS services and features. For example, you can't configure S3 Cross-Region Replication to operate through an access point.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Cryptographic keys are used to encrypt your data at rest. Amazon S3 offers three options for encrypting your objects:

- **Server-side encryption (SSE) with Amazon S3-managed keys (SSE-S3)** – When you use SSE-S3, each object is encrypted with a unique key. As an additional safeguard, it encrypts the key itself with a primary key that it regularly rotates. Amazon S3 server-side encryption uses 256-bit Advanced Encryption Standard (AES-256) to encrypt your data.

- **Server-side encryption with AWS KMS keys stored in AWS Key Management Service (AWS KMS) (SSE-KMS)** – KMS keys stored in SSE-KMS are similar to SSE-S3, but with some additional benefits and charges. There are separate permissions for the use of a KMS key that provides added protection against unauthorized access of your objects in Amazon S3. SSE-KMS also provides you an audit trail that shows when your KMS key was used, and by whom.

- **Dual-layer server-side encryption with AWS KMS keys (DSSE-KMS)** – Using DSSE-KMS applies two individual layers of object-level encryption instead of one layer. Each layer of encryption uses a separate cryptographic implementation library with individual data encryption keys.

- **Server-side Encryption with Customer-Provided Keys (SSE-C)** – With SSE-C, you manage the encryption keys and Amazon S3 manages the encryption as it writes to disks. Also, Amazon S3 manages decryption when you access your objects.

For more information about server-side encryption, see "Protecting data using server-side encryption" in the *Amazon Simple Storage Service User Guide* at https://docs.aws.amazon.com/AmazonS3/latest/userguide/serv-side-encryption.html.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Think about where you keep important objects in your office or home. There is more than one place to keep each thing.

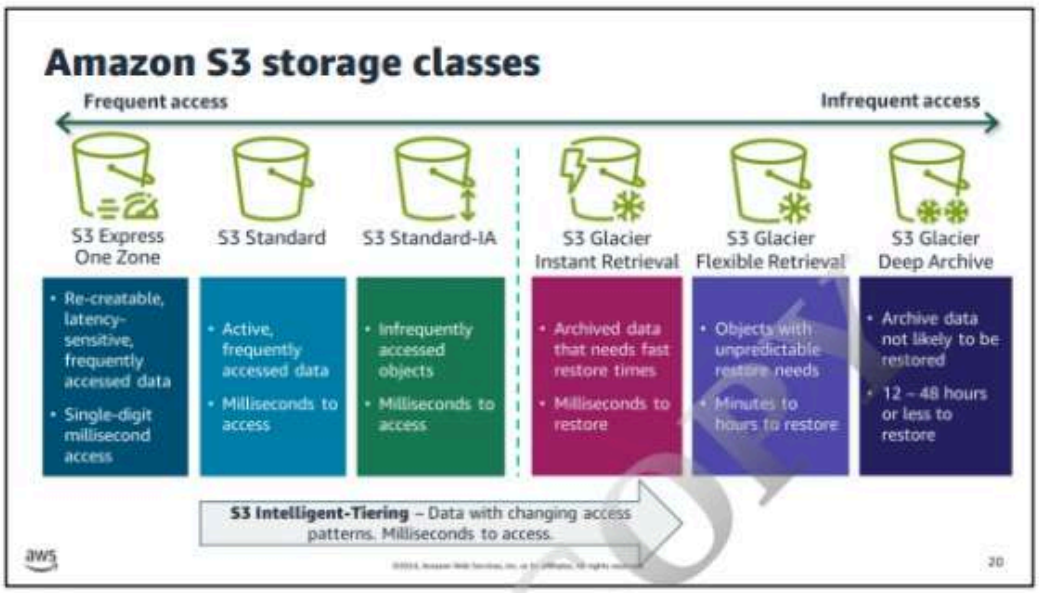
- At your workplace
- Another room on the same floor
- The basement of the same building
- A different building near yours
- A storage unit in another location

You make decisions based on how often you use that object, and that decision can change over time. You choose Amazon S3 storage classes with similar considerations.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Each object in Amazon S3 has a storage class associated with it. All storage classes offer high durability (99.999999999 percent durability).

Choose a class depending on your use case scenario and performance access requirements:

- **S3 Express One Zone** for data that requires single-digit millisecond access and that can be re-created as it's stored only in a single Availability Zone.
- **S3 Standard** for general-purpose storage of frequently accessed data.
- **S3 Standard-Infrequent Access (S3 Standard-IA)** for long-lived, but less frequently accessed data.
- **S3 Glacier Instant Retrieval** for archive data that is rarely accessed but requires a restore in milliseconds.
- **S3 Glacier Flexible Retrieval** for the most flexible retrieval options that balance cost with access times ranging from minutes to hours. Your retrieval options permit you to access all the archives you need, when you need them, for one low storage price. This storage class comes with multiple retrieval options:
    - Expedited retrievals (restore in 1–5 minutes).
    - Standard retrievals (restore in 3–5 hours).
    - Bulk retrievals (restore in 5–12 hours). Bulk retrievals are available at no additional charge.
- **S3 Glacier Deep Archive** for long-term cold storage archive and digital preservation. There are two retrieval tiers that are available:
    - Standard retrieval (restore in 12 hours or within 9-12 hours when S3 Batch Operations are used).
    - Bulk retrieval (restore within 48 hours) at a fraction of the cost of the Standard tier.

Another option is **S3 One Zone-Infrequent Access (S3 One Zone-IA)** for long-lived, less frequently accessed data that can be stored in a single Availability Zone. S3 Express One Zone is optimized for performance and lower request costs, making it a better choice for frequently accessed data that requires low latency. However, for infrequently accessed data, S3 Standard-IA offers higher availability across multiple Availability Zones, at a lower storage cost compared to S3 Express One Zone.

S3 Intelligent-Tiering is an additional storage class that provides flexibility for data with unknown or changing access patterns. It automates the movement of your objects between storage classes to optimize cost.

- To save more on data that does not require immediate retrieval, you can enable the optional **Deep**

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Archive Access tier, which moves objects not accessed for 180 days to lower-cost storage. Alternatively, you can activate both the **Archive Access** and **Deep Archive Access tiers**, moving objects not accessed for 90 days to the Archive Access tier and then to the Deep Archive Access tier after 180 days.

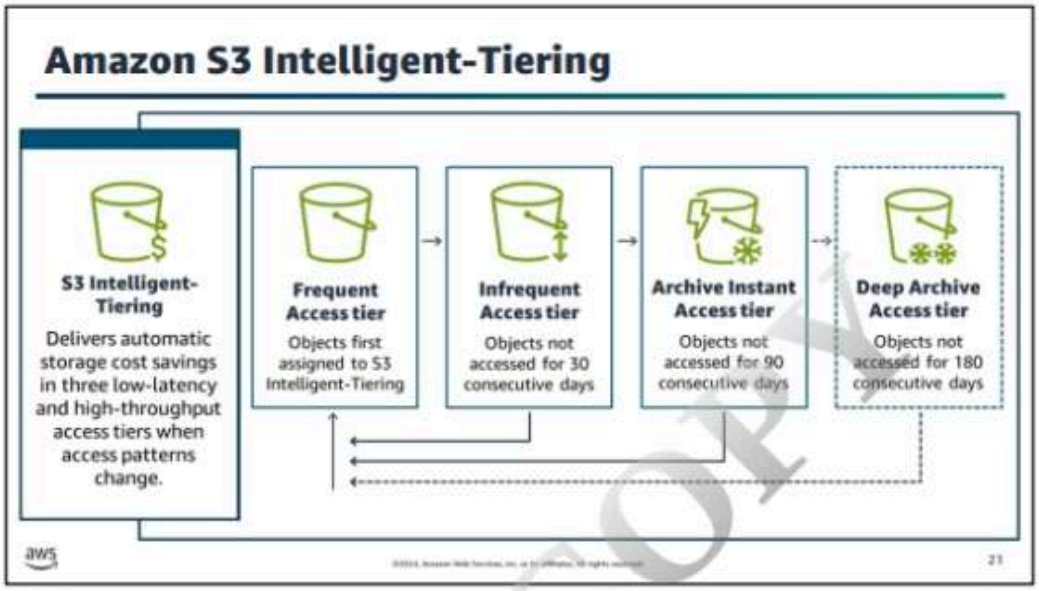
S3 on Outposts (not pictured) delivers object storage to your on-premises AWS Outposts environment. Outposts will be discussed later in this course.

For more information about choosing the right storage class based on your use and cost requirements, see "Amazon S3 pricing" at https://aws.amazon.com/s3/pricing/.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 Intelligent-Tiering is the only storage class that delivers automatic storage cost savings when data access patterns change, without performance impact or operational overhead. Your data moves between access tiers as usage patterns change.

When you assign an object to S3 Intelligent-Tiering, it is placed in the Frequent Access tier which has the same storage cost as S3 Standard. Objects not accessed for 30 days are then moved to the Infrequent Access tier where the storage cost is the same as S3 Standard-IA. After 90 days of no access, an object is moved to the Archive Instant Access tier, which has the same cost as S3 Glacier Instant Retrieval.

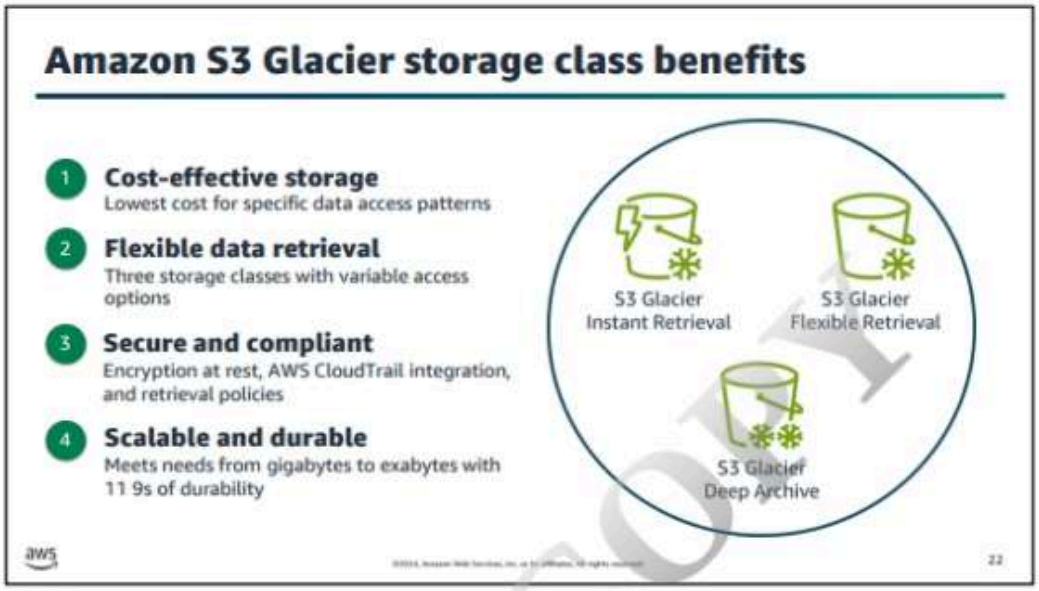
To save more on data that doesn't require immediate retrieval, you can activate the optional asynchronous Deep Archive Access tier. When turned on, objects not accessed for 180 days are moved to the Deep Archive Access tier.

S3 Intelligent-Tiering is the ideal storage class for data with unknown, changing, or unpredictable access patterns, independent of object size or retention period. You can use S3 Intelligent-Tiering as the default storage class for virtually any workload, especially data lakes, data analytics, new applications, and user-generated content.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 Glacier is a service that gives you extremely-low cost, powerful, and flexible data storage solutions. The storage is purpose-built for your archived data.

Three storage class options help optimize your cost for what you need to retrieve and how fast.

Your data in Amazon S3 Glacier can be encrypted at rest. In addition, Amazon S3 Glacier products offer you a suite of features including compliance, audit logging, and cost management. Those features might not be available to you with a traditional archiving solution.

Like other Amazon S3 storage classes, Amazon S3 Glacier products can scale from gigabytes to exabytes of your data. Amazon S3 Glacier storage classes have 99.999999999 percent durability.

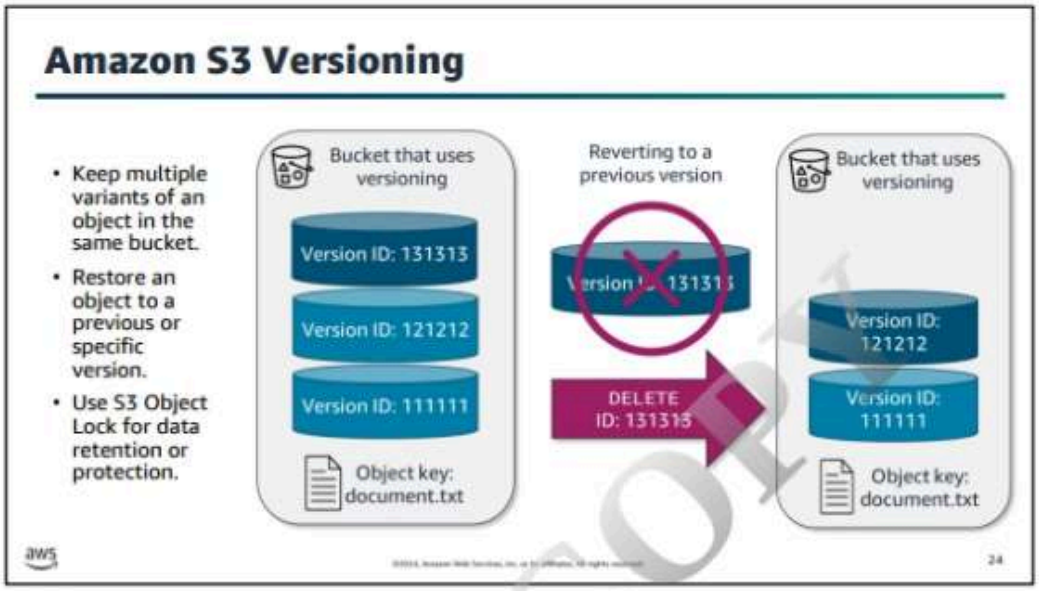
In mathematical terms, 11 nines of durability means that out of 10,000 objects, you might expect to lose one every 10 million years. We asked a large Hollywood studio to run the same Markov model to determine the number of nines for two copies of data on tape. They came back with 5–6 nines. Having six more nines means that the durability of Amazon S3 Glacier is six orders of magnitude more durable than two copies of your data on tape.

For more information, see "Amazon S3 Glacier storage classes" at https://aws.amazon.com/glacier/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 uses object storage. Therefore, if you want to change a part of a file, you must make the change and then reupload the entire modified file.

Buckets that use versioning can help you recover objects from accidental deletion or overwrite.

- If you delete an object, instead of removing it permanently, Amazon S3 inserts a *delete marker*, which becomes the current object version.
- If you overwrite an object, it results in a new object version in the bucket.

When S3 Versioning is turned on, you can restore the previous version of the object to correct the mistake.

The diagram shows a bucket that uses versioning with an object that has three versions. The current version ID is 131313. You can restore to the previous version ID of 121212. To delete versioned objects permanently, you must use DELETE Object versionId. This example deletes version ID 131313.

You can use S3 Object Lock for data retention or protection. By using the write once, read many (WORM) model, you can prevent accidental overwrites or deletions within Amazon S3 storage. Use retention periods for locking an object for a fixed period of time, or a legal hold for a lock until explicitly removed.

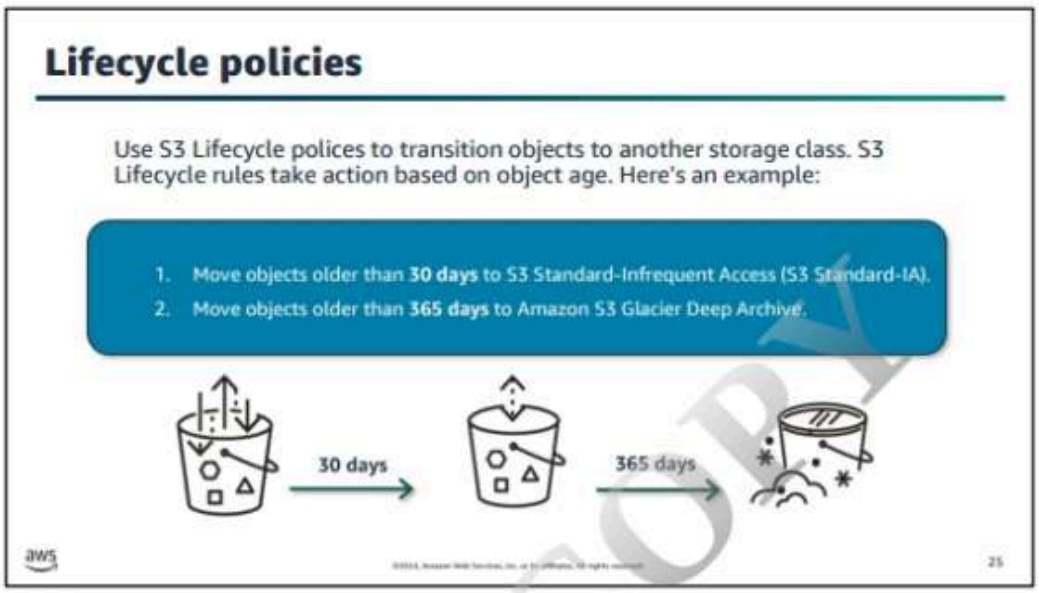
For more information about versioning, see "Retaining multiple versions of objects with S3 Versioning" at https://docs.aws.amazon.com/AmazonS3/latest/dev/Versioning.html.

For more information about S3 Object Lock, see "Locking objects with Object Lock" at https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lock.html.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



With S3 Lifecycle policies, you can delete or move objects based on age. You should automate the lifecycle of your data that is stored in Amazon S3. Using S3 Lifecycle policies, you can have data cycled at regular intervals between different Amazon S3 storage types. In this way, you reduce your overall cost because you are paying less for data as it becomes less important with time.

In addition to being able to set lifecycle rules *per object*, you can also set lifecycle rules *per bucket*. Amazon S3 supports a waterfall model for transitioning between storage classes. *Lifecycle configuration* automatically changes data storage tiers.

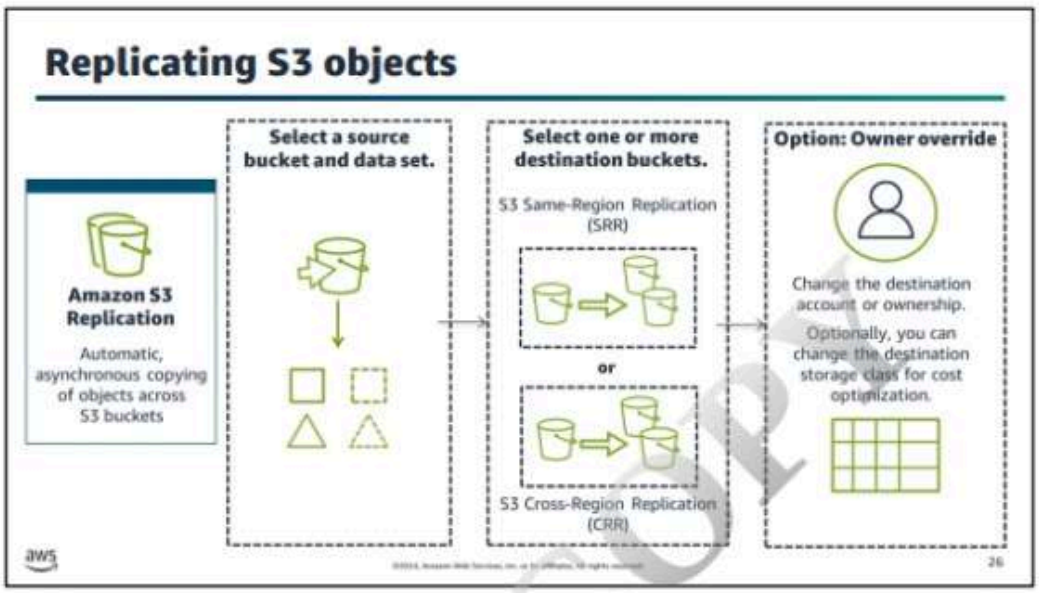
For more information about transitioning objects and managing your storage lifecycle by using S3 Lifecycle, see the following in the *Amazon Simple Storage Service User Guide*:

- "Transitioning objects using Amazon S3 Lifecycle" at https://docs.aws.amazon.com/AmazonS3/latest/userguide/lifecycle-transition-general-considerations.html
- "Managing the lifecycle of objects" at https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lifecycle-mgmt.html

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



With Amazon S3, customers get a high level of availability and durability for their data in every AWS Region. Data that is stored in any Amazon S3 storage class is stored across a minimum of three Availability Zones. Each Availability Zone is separated by miles within a Region. For this reason, many AWS customers choose Amazon S3 to store their business-critical and application-critical data. The only exception is S3 One Zone-Infrequent Access (S3 One Zone-IA), which, as its name indicates, is a one-zone service.

Replication can help you do the following tasks:

- **Replicate objects while retaining metadata** – Make sure that your replica is identical to the source object if it is necessary.
- **Replicate objects into different storage classes** – Use replication to put objects directly into S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, or another storage class in the destination buckets.
- **Maintain object copies under different ownership** – Tell Amazon S3 to change replica ownership to the AWS account that owns the destination bucket.
- **Keep objects stored over multiple AWS Regions** – Meet compliance requirements by replicating data to another AWS Region.

First, identify the source bucket and data set. Then, select one or more destination buckets. You can use S3 Same-Region Replication (SRR) for things like log aggregation or replicating between development and production accounts. If you must meet compliance or other requirements to have data in more than one geography, choose S3 Cross-Region Replication (CRR).

Optionally, you can change the destination ownership of objects with an owner override. You might do this, for example, to restrict access to object replicas.
Learn how to meet business data resiliency. For more information, see "How to meet business data resiliency with Amazon S3 cross-Region replication" in the *AWS Public Sector Blog* at https://aws.amazon.com/blogs/publicsector/how-to-meet-business-data-resiliency-s3-cross-region-replication/.
To learn more about replication, see "Amazon S3 Replication" at https://aws.amazon.com/s3/features/replication/.

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Earlier, you were provided with some examples that related to securing and storing objects. Next, think about how you might place objects into your S3 buckets. You want to be efficient with how you upload the object and use features that help you manage what happens when the object is there.

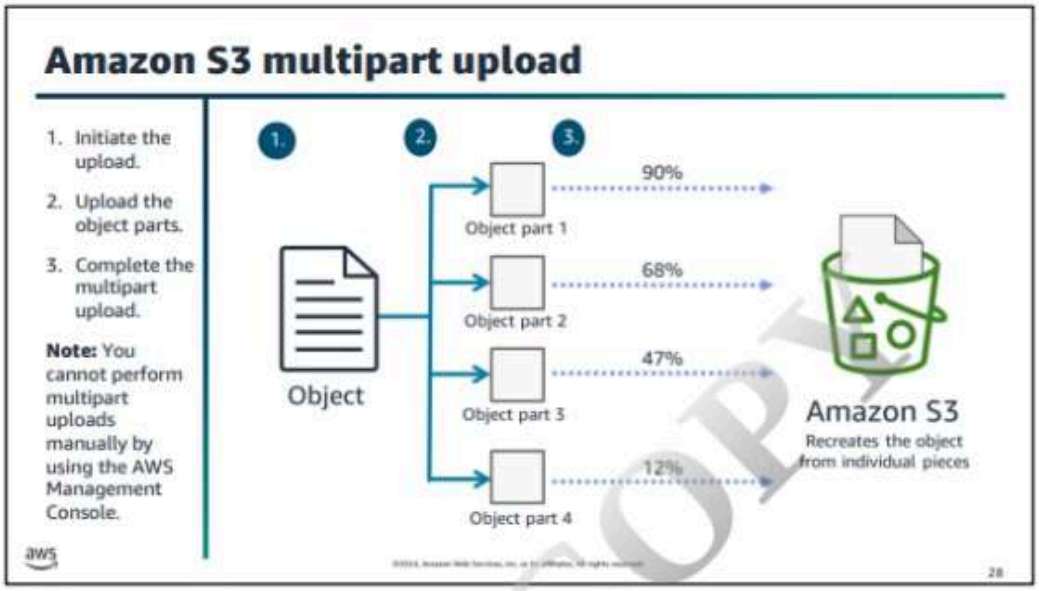
Consider the following questions:

- What if you need to upload large objects? How do you upload them, and how does it work?
- How can you increase the speed at which objects are uploaded to AWS Regions that are not close to you?
- Can actions be automated based on events like when you upload an object?

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



With a multipart upload, you can consistently upload large objects in manageable parts. This process involves three steps:

1. Initiate the upload.
2. Upload the object parts.
3. Complete the multipart upload.

When the multipart upload request is completed, Amazon S3 will recreate the full object from the individual pieces.

You can improve the upload process of larger objects with the following features:

- **Improved throughput** – You can upload parts in parallel to improve throughput.
- **Quick recovery from any network issues** – Smaller part sizes minimize the impact of restarting a failed upload due to a network error.
- **Pausing and resuming object uploads** – You can upload object parts over time. When you have initiated a multipart upload, there is no expiration. You must explicitly complete or cancel the multipart upload.
- **Beginning an upload before you know the final object size** – You can upload an object as you are creating it.
- **Uploading large objects** – Using the multipart upload API, you can upload large objects, up to 5 TB.

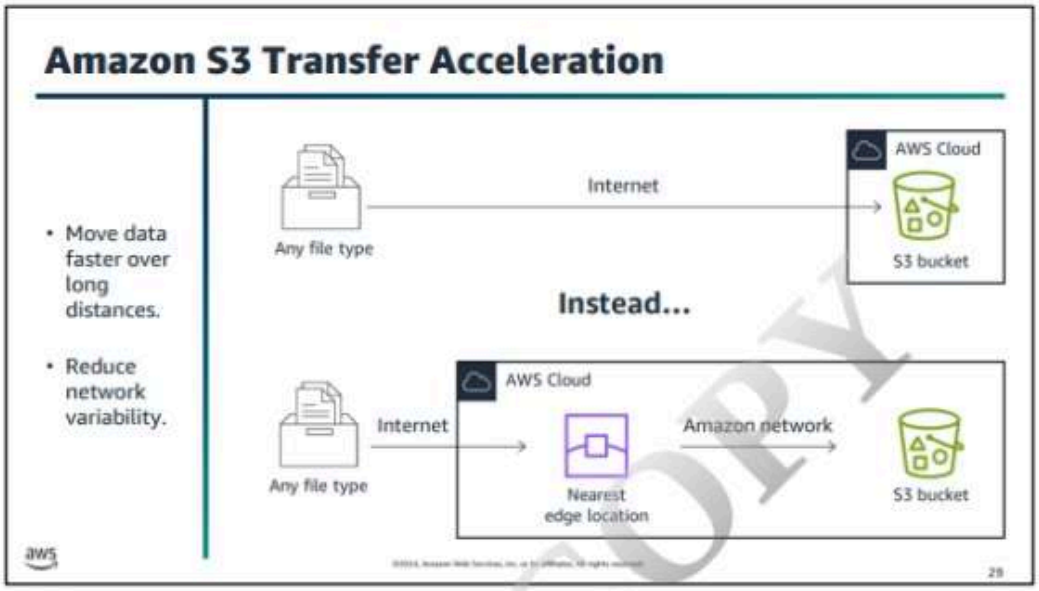
Note: You cannot perform multipart uploads manually by using the AWS Management Console.

For more information about multipart uploads, see "Uploading and copying objects using multipart upload" in the *Amazon Simple Storage Service User Guide* at https://docs.aws.amazon.com/AmazonS3/latest/userguide/mpuoverview.html.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon S3 Transfer Acceleration uses AWS globally distributed edge locations to facilitate fast data transfer into an S3 bucket. The data is routed to Amazon S3 over an optimized network path.

Use Transfer Acceleration for the following:
- When you have customers all over the world who upload to a centralized bucket
- When you transfer gigabytes or terabytes of data across continents on a regular basis
- When you underutilize the available bandwidth when uploading to Amazon S3 over the internet

S3 Transfer Acceleration shortens the distance between client applications and AWS servers that acknowledge PUTS and GETS operations to Amazon S3. Transfer Acceleration uses a global network of hundreds of edge locations. AWS automatically routes your uploads and downloads through the closest edge locations to your application.

You can use the Amazon S3 Transfer Acceleration Speed Comparison tool to compare accelerated and non-accelerated upload speeds across Amazon S3 Regions.

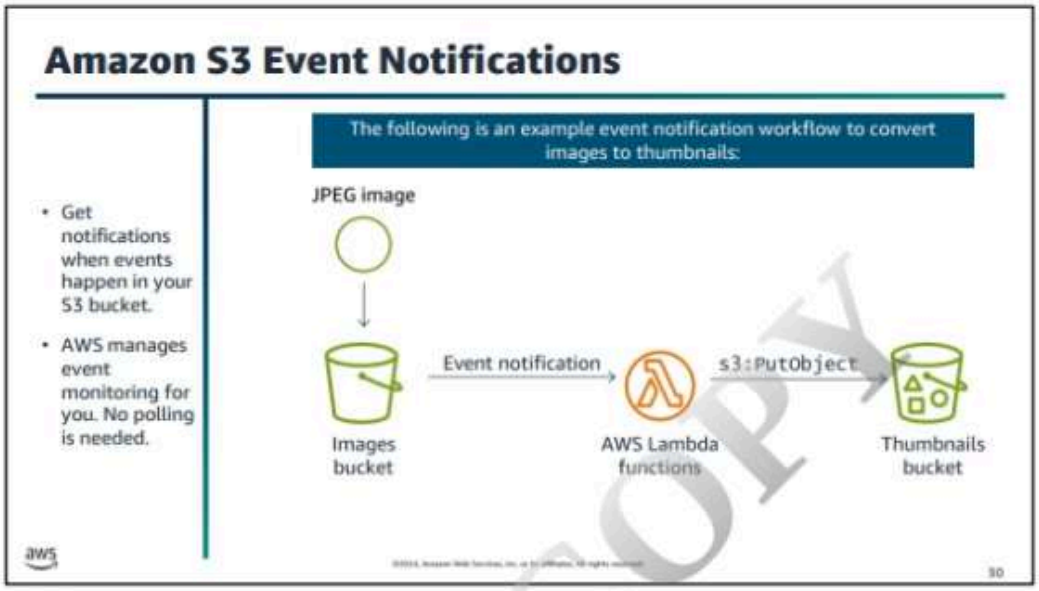
For more information about Transfer Acceleration, see S3 Transfer Acceleration at https://aws.amazon.com/s3/transfer-acceleration/.

Learn more about the speed comparison tool. For more information, see "Using the Amazon S3 Transfer Acceleration Speed Comparison tool" in the *Amazon Simple Storage Service User Guide* at https://docs.aws.amazon.com/AmazonS3/latest/userguide/transfer-acceleration-speed-comparison.html.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



With Amazon S3 Event Notifications, you can receive notifications when certain object events happen in your bucket. Event-driven models such as S3 Event Notifications mean that you no longer need to build or maintain server-based polling infrastructure to check for object changes. You also don't pay for idle time of that infrastructure when there are no changes to process.

Amazon S3 can send event notification messages to the following destinations:
- Amazon Simple Notification Service (Amazon SNS) topics
- Amazon Simple Queue Service (Amazon SQS) queues
- AWS Lambda functions

You specify the Amazon Resource Name (ARN) value of these destinations in the notification configuration.

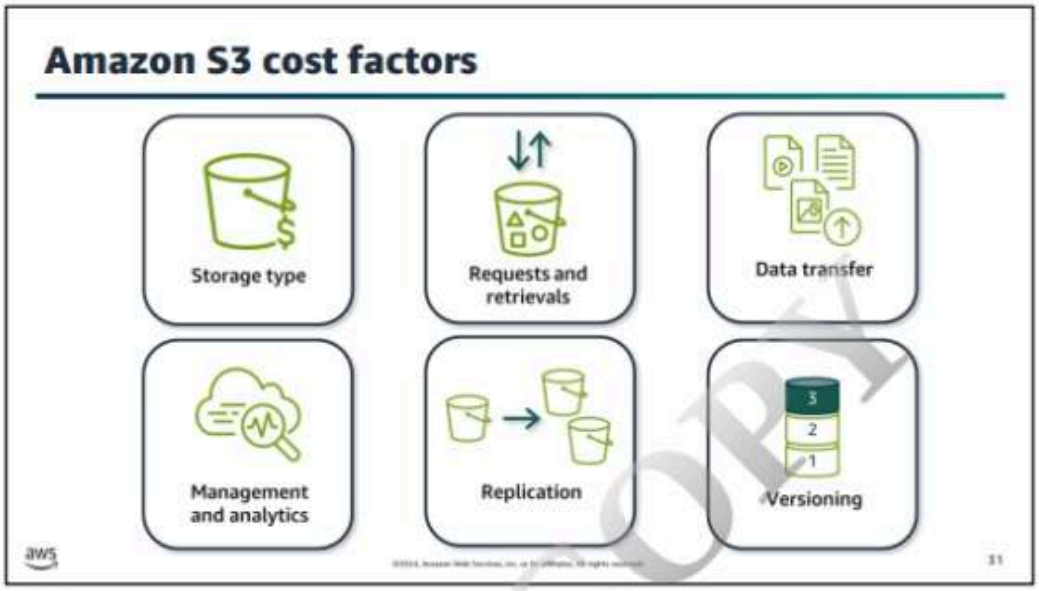
In the example, you have a JPEG image uploaded to the images bucket that your website uses. Your website needs to be able to show smaller *thumbnail* preview images of each uploaded file. When the image object is added to the S3 bucket, an event notification is sent to invoke a series of AWS Lambda functions. The output of your Lambda functions is a smaller version of the original JPEG image and this object is put in your thumbnails bucket. S3 Event Notifications manages the activity in the bucket for you and automates the creation of your thumbnail.

For more information about Amazon S3 Event Notifications, see "Reliable event processing with Amazon S3 Event Notifications" on the *AWS Storage Blog* at https://aws.amazon.com/blogs/storage/reliable-event-processing-with-amazon-s3-event-notifications/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Cost is an important part of choosing the right Amazon S3 storage solution.

Some of the Amazon S3 cost factors to consider include the following:

- **Storage** – Per-gigabyte cost to hold your objects. You pay for storing objects in your S3 buckets. The rate that you're charged depends on your objects' size, how long you stored the objects during the month, and the storage class. You incur per-request ingest charges when using PUT, COPY, or lifecycle rules to move data into any S3 storage class.
- **Requests and retrievals** – The number of API calls—PUT and GET requests. You pay for requests that are made against your S3 buckets and objects. Costs for S3 requests are based on the request type, and are charged on the quantity of requests. When you use the Amazon S3 console to browse your storage, you incur charges for GET, LIST, and other requests that are made to facilitate browsing.
- **Data transfer** – Usually no transfer fee for data-in from the internet and, depending on the requester location and medium of data transfer, different charges for data-out.
- **Management and analytics** – You pay for the storage management features and analytics that are activated on your account's buckets. These features are not discussed in detail in this course.

S3 Replication and S3 Versioning can have a big impact on your AWS bill. These services both create multiple copies of your objects, and you pay for each PUT request in addition to the storage tier charge. S3 Cross-Region Replication also requires data transfer between AWS Regions.

For more information about potential costs in Amazon S3, use the AWS Pricing Calculator at https://calculator.aws/#/.

For more information, see Amazon S3 pricing at https://aws.amazon.com/s3/pricing/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

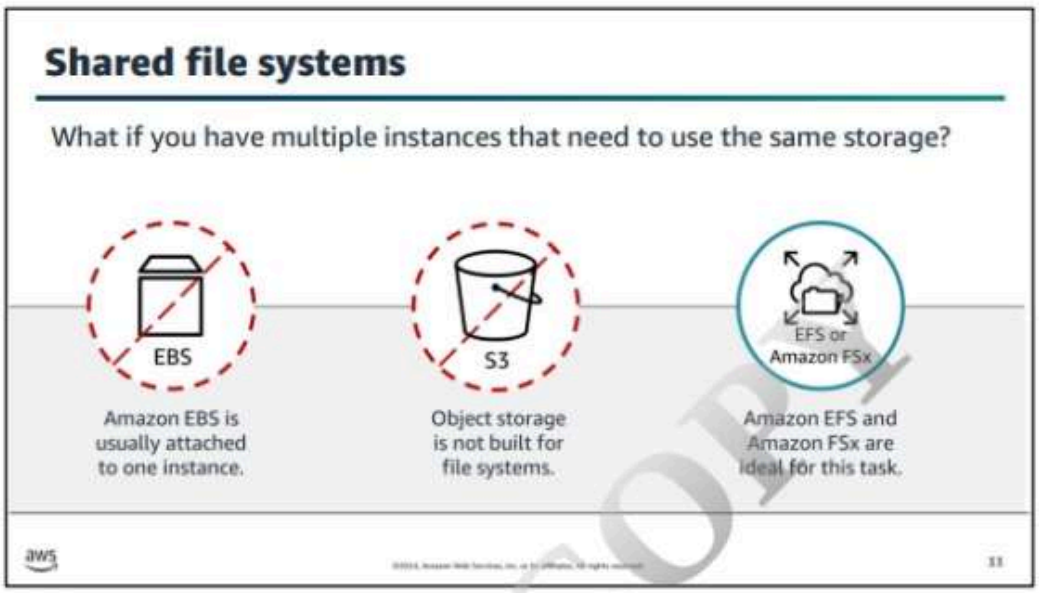
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

The storage team lead asks, "What are some file-based options for building secure and scalable storage in the AWS Cloud?" The storage team must know about AWS services that provide traditional file server solutions. The company wants your advice for how to choose the right service for their use case.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



How do you handle an application that is running on multiple instances and must use the same file system?

Amazon EBS provides block storage, so you could use it as the underlying storage component of a self-managed file storage solution. Amazon EBS supports Multi-Attach for up to 16 Linux Amazon Elastic Compute Cloud (Amazon EC2) instance attachments, but it is a very specialized use case. In most cases, an EBS volume is attached to one Amazon EC2 instance. This limit makes it difficult to have the scalability, availability, and affordability of a fully managed file storage solution.

Amazon S3 is an option, but what if you need the performance and read/write capacity of a network file system? S3 is an object store system, not a block store, so changes overwrite entire files, not blocks of characters within files.

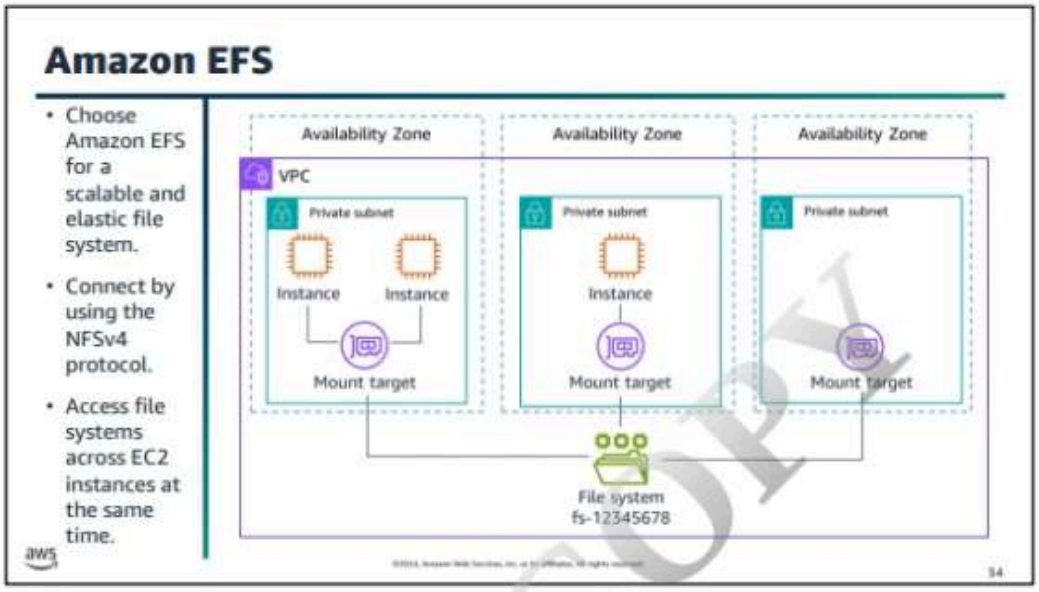
For high throughput changes to files of varying sizes, a file system is superior to an object store system. Amazon Elastic File System (Amazon EFS) and Amazon FSx are ideal for this use case. Using a fully managed file storage solution in the cloud removes complexities, reduces costs, and simplifies management. You continue to learn about shared file systems in this section of the module.

Learn about the considerations and limitations of multi-attach. For more information, see "Attach a volume to multiple instances with Amazon EBS Multi-Attach" in the *Amazon Elastic Compute Cloud User Guide for Linux Instances* at https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volumes-multi.html.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon EFS provides a scalable, elastic file system for Linux-based workloads for use with AWS Cloud services and on-premises resources. You can create a file system, mount it on an Amazon EC2 instance, and then read and write data to and from your file system. You can mount an Amazon EFS file system in your VPC through the Network File System (NFS) versions 4.0 and 4.1 (NFSv4) protocol. You do not need to take action to expand the file system as your storage needs grow. EC2 instances in your VPC can access Amazon EFS file systems concurrently, so applications that scale beyond a single connection can access a file system.

Availability and durability refer to the redundancy with which an Amazon EFS file system stores data within an AWS Region. You have the following choices for your file system's availability and durability:

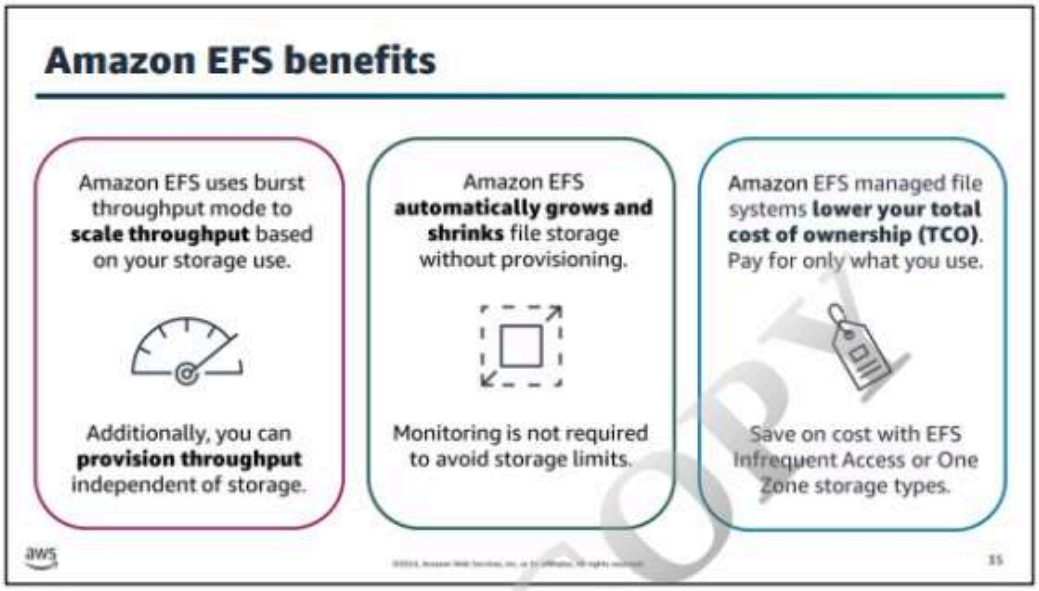
- Choosing an EFS Standard storage class creates a file system that stores file system data and metadata redundantly across all Availability Zones within an AWS Region. You can also create mount targets in each Availability Zone in the AWS Region. EFS Standard storage classes offer the highest levels of availability and durability.
- Choosing an EFS One Zone file system creates a file system that stores file system data and metadata redundantly within a single Availability Zone. File systems that use EFS One Zone storage classes can have only a single mount target. This mount target is located in the Availability Zone in which the file system is created.

In this example, a VPC uses Amazon EFS Standard storage class. The VPC has three private subnets, each in a different Availability Zone. With EFS Standard storage class, each subnet can have its own mount target. The EC2 instances in each subnet can access the file system through the mount target located in its Availability Zone.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon EFS provides a shared, persistent layer that provides stateful applications with the ability to elastically scale up and down. Examples include DevOps, web serving, web content systems, media processing, machine learning, analytics, search index, and stateful microservices applications. Amazon EFS can support a petabyte-scale file system, and the throughput of the file system also scales with the capacity of the file system.

Because Amazon EFS is serverless, you don't need to provision or manage the infrastructure or capacity. Amazon EFS file systems can be shared with up to tens of thousands of concurrent clients, regardless of the type. These clients could be traditional EC2 instances or containers that run in one of your self-managed clusters. They might run in one of the AWS container services: Amazon Elastic Container Service (Amazon ECS), Amazon Elastic Kubernetes Service (Amazon EKS), and AWS Fargate. Or, they might run in a serverless function that runs in Lambda.

Use Amazon EFS to lower your total cost of ownership (TCO) for shared file storage. Choose Amazon EFS One Zone for data that does not require replication across multiple Availability Zones, and save on storage costs. Amazon EFS Standard-Infrequent Access (EFS Standard-IA) and Amazon EFS One Zone-Infrequent Access (EFS One Zone-IA) are storage classes for files not accessed every day. They provide cost-optimized price and performance for these files.

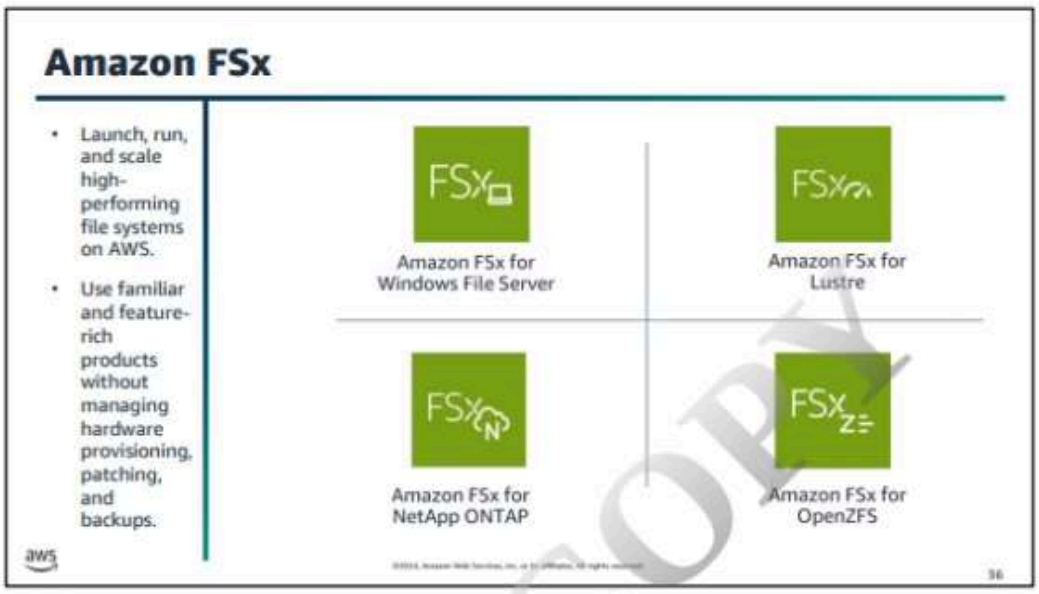
Use Amazon EFS scaling and automation to save on management costs, and pay for only what you use.

For more information about Amazon EFS use cases, see File Systems for Enterprise Applications at https://aws.amazon.com/efs/enterprise-applications/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



With Amazon FSx, you can quickly launch and run feature-rich and high-performing file systems. The service provides you with four file systems to choose from. This choice is based on your familiarity with a given file system or matching the feature sets, performance profiles, and data management capabilities to your needs.

FSx for Windows File Server provides fully managed Microsoft Windows file servers that are backed by a native Windows file system. Built on Windows Server, Amazon FSx delivers a wide range of administrative features such as data deduplication, end-user file restore, and Microsoft Active Directory.

FSx for Lustre is a fully managed service that provides high-performance, cost-effective storage. FSx for Lustre is compatible with the most popular Linux-based Amazon Machine Images (AMIs). These AMIs include Amazon Linux, Amazon Linux 2, Red Hat Enterprise Linux (RHEL), CentOS, SUSE Linux, and Ubuntu.

FSx for NetApp ONTAP provides fully managed shared storage in the AWS Cloud with the popular data access and management capabilities of ONTAP.

FSx for OpenZFS provides fully managed shared file storage built on the OpenZFS file system. It is powered by the AWS Graviton family of processors, and accessible through the NFS protocol (v3, v4, v4.1, v4.2).

For more information about Amazon FSx file system options, see Choosing an Amazon FSx File System at https://aws.amazon.com/fsx/when-to-choose-fsx/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

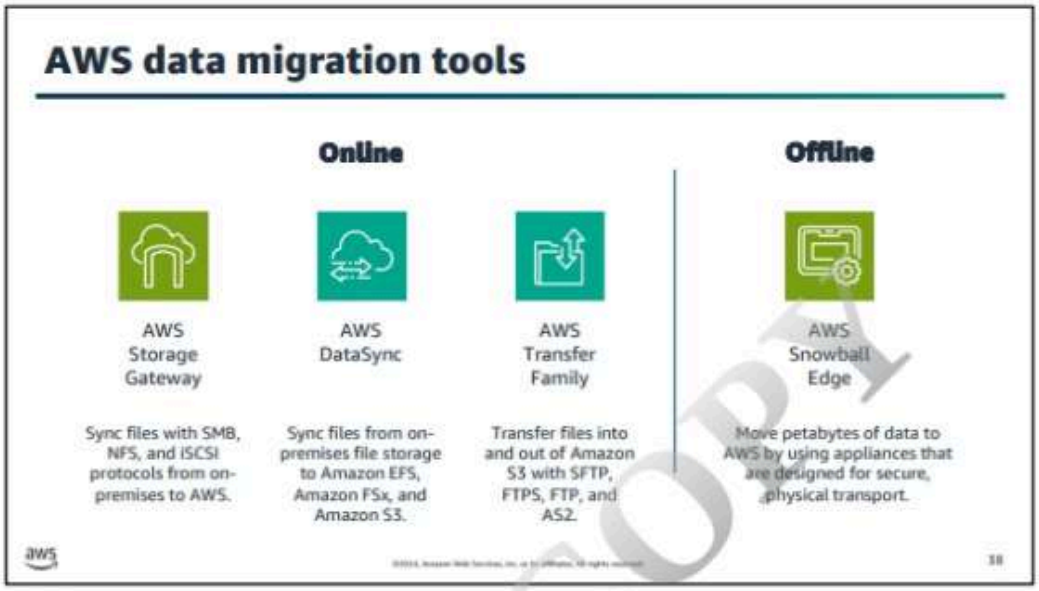
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

The storage team lead asks, "How can we move lots of data to the cloud in a relatively short time period?" The storage team must plan data migrations from on-premises data centers to the AWS Cloud. The company wants your advice to choose the right tools.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



AWS Snowball Edge is a type of Snowball device with on-board storage and compute power for select AWS capabilities. Snowball Edge can process data locally, run edge-computing workloads, and transfer data to or from the AWS Cloud.

Before choosing which tool to use, you should know the following information:

- Where you are moving data
- What your use cases are
- The types of data that you are moving
- The network resources available

AWS offers a wide variety of services and AWS Partner tools to help you migrate your datasets (files, databases, machine images, block volumes, or tape backups). In this module, you learn about the following tools:

- *AWS Storage Gateway* simplifies on-premises adoption of AWS storage. You can use Storage Gateway to seamlessly connect and extend your on-premises applications to AWS storage. It supports multiple file transfer protocols: Server Message Block (SMB), Network File System (NFS), and Internet Small Computer Systems Interface (iSCSI).
- *AWS DataSync* is a data transfer service that facilitates moving data between on-premises storage and Amazon EFS, Amazon FSx, and Amazon S3.
- *AWS Transfer Family* permits the transfer of files into and out of Amazon S3 or Amazon EFS over the following protocols:
    - Secure Shell File Transfer Protocol (SFTP)
    - File Transfer Protocol Secure (FTPS)
    - File Transfer Protocol (FTP)
    - Applicability Statement 2 (AS2)

    Note: The Transfer Family is not covered in this course.
- *AWS Snowball Edge* is a device that can process data locally, run edge-computing workloads, and transfer data to or from the AWS Cloud.

To learn more about on-premises AWS storage services, see "Comparing your on-premises storage patterns with AWS Storage services" in the *AWS Storage Blog* at https://aws.amazon.com/blogs/storage/comparing-your-on-premises-storage-patterns-with-aws-storage-services/.

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

AWS Storage Gateway

AWS Storage Gateway

AWS Storage Gateway is a service that gives your applications seamless and secure integration between on-premises environments and AWS storage.

It provides you with low-latency access to cloud data with a Storage Gateway appliance.

19

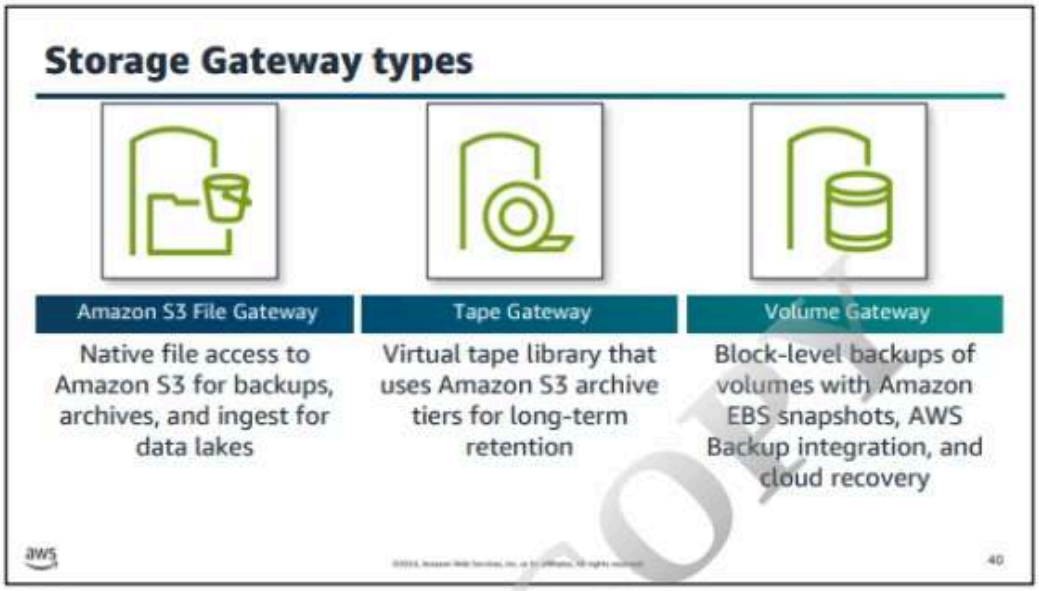
AWS Storage Gateway connects an on-premises software appliance with cloud-based storage. It provides seamless integration with data security features between your on-premises IT environment and the AWS storage infrastructure.

You can use Storage Gateway to store data in the AWS Cloud for scalable and cost-effective storage that helps maintain data security. Storage Gateway offers file-based, volume-based, and tape-based storage solutions, which are integrated with AWS Identity and Access Management (IAM), AWS Key Management Service (AWS KMS), AWS CloudTrail, and Amazon CloudWatch.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Choose a Storage Gateway type that is the best fit for your workload.

Amazon S3 File Gateway presents a file interface that you can use to store files as objects in Amazon S3. You use the industry-standard NFS and SMB file protocols. Access your files through NFS and SMB from your data center or Amazon EC2, or access those files as objects directly in Amazon S3.

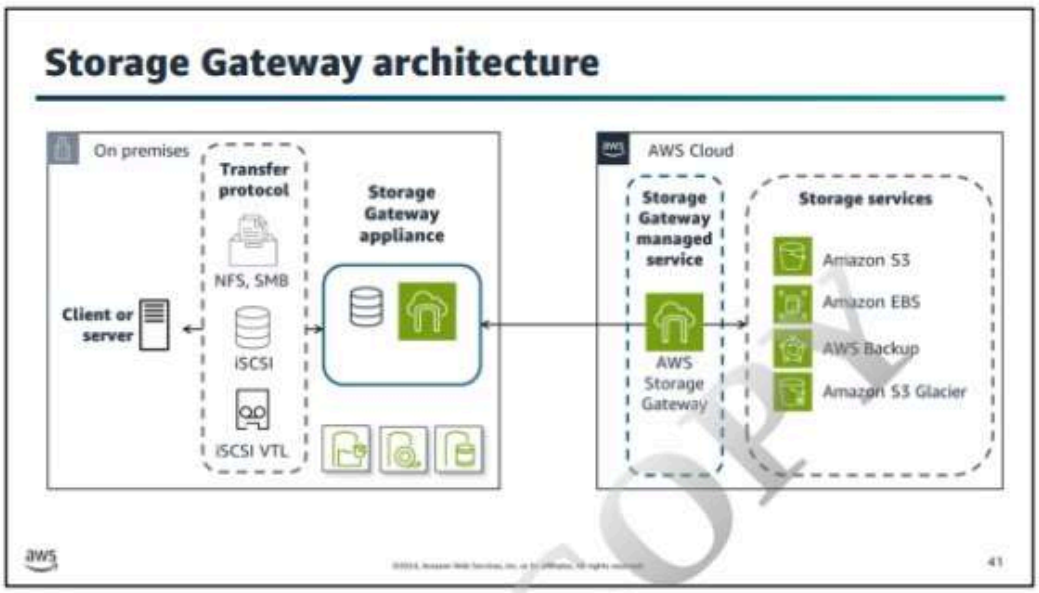
Tape Gateway presents an iSCSI-based virtual tape library (VTL) of virtual tape drives and a virtual media changer to your on-premises backup application. Tape Gateway stores your virtual tapes in Amazon S3 and creates new ones automatically, which helps you with management and your transition to AWS.

Volume Gateway presents block storage volumes of your applications by using the iSCSI protocol. You can asynchronously back up data that is written to these volumes as point-in-time snapshots of your volumes. Then, you can store it in the cloud as Amazon EBS snapshots. You can back up your on-premises Volume Gateway volumes by using the service's snapshot scheduler or by using the AWS Backup service.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



The Storage Gateway Appliance supports the following protocols to connect to your local data:

- NFS or SMB for files
- iSCSI for volumes
- iSCSI VTL for tapes

Your storage gateway appliance runs in one of three modes: Amazon S3 File Gateway, Tape Gateway, or Volume Gateway.

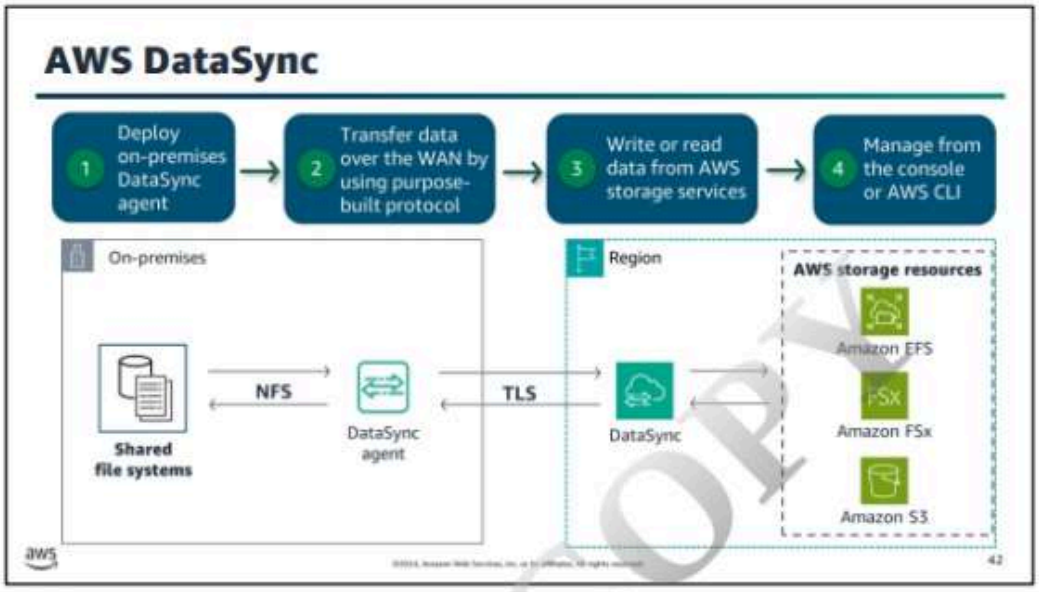
You can send data that is moved to AWS by using Storage Gateway to the following destinations through the Storage Gateway managed service:

- Amazon S3 (Amazon S3 File Gateway, Tape Gateway)
- Amazon S3 Glacier (Amazon S3 File Gateway, Tape Gateway)
- Amazon EBS (Volume Gateway)

You can use AWS Backup to schedule volume snapshots with Volume Gateway.

DO NOT COPY javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Manual tasks that are related to data transfers can slow down migrations and burden IT operations. DataSync facilitates moving large amounts of data between on-premises storage and Amazon EFS, Amazon FSx, and Amazon S3. By default, data is encrypted in transit by using TLS 1.2. DataSync automatically handles scripting copy jobs, scheduling and monitoring transfers, validating data, and optimizing network usage.

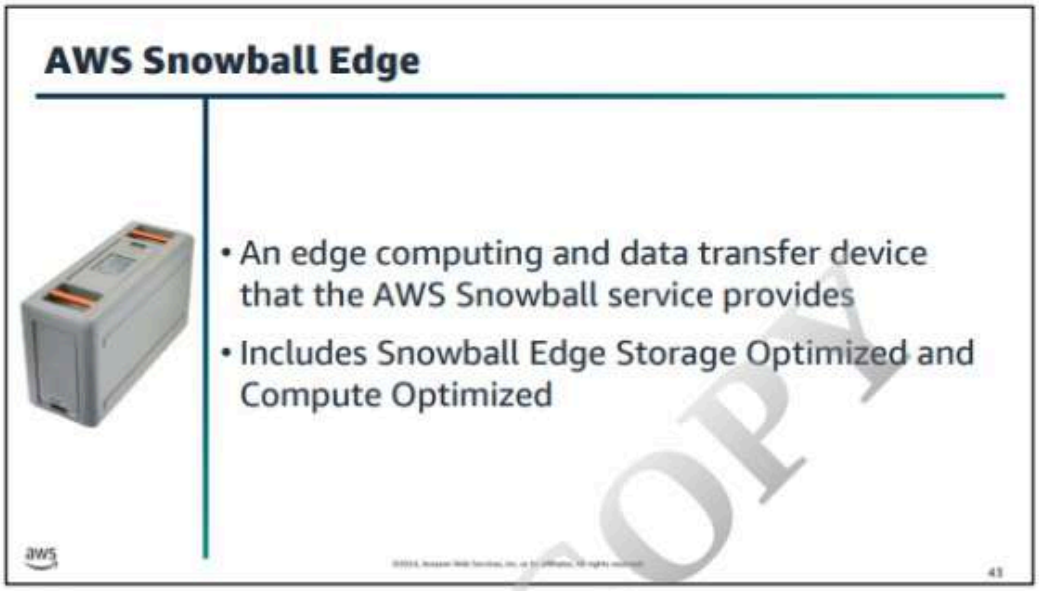
Reduce on-premises storage infrastructure by shifting SMB-based data stores and content repositories from file servers and NAS arrays to Amazon S3 and Amazon EFS for analytics.

DataSync deploys as a single software agent that can connect to multiple shared file systems and run multiple tasks. The software agent is typically deployed on premises through a virtual machine to handle the transfer of data over the wide area network (WAN) to AWS. On the AWS side, the agent connects to the DataSync service infrastructure. Because DataSync is a service, there is no infrastructure for customers to set up or maintain in the cloud. DataSync configuration is managed directly from the console.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



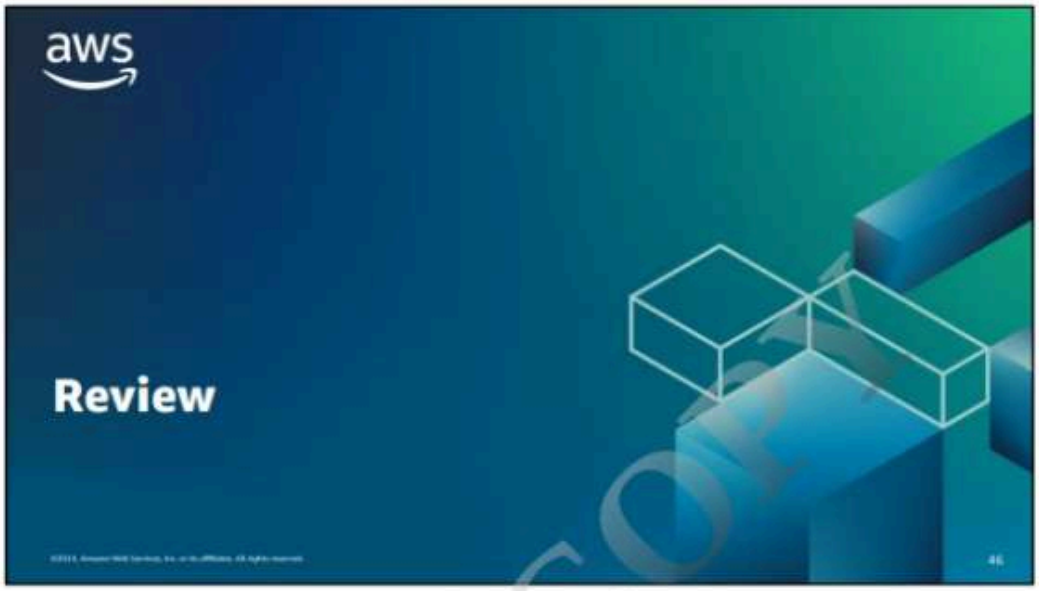
Snowball Edge is a petabyte-scale data transport option that doesn't require you to write code or purchase hardware to transfer data. All that you need to do is create a job in the console, and a Snowball appliance will be shipped to you. Attach the appliance into your local network and transfer the files directly onto it. When the device is ready to be returned, the E Ink shipping label will automatically update the return address so that the device is delivered to the correct AWS facility. For more information, see AWS Snowball FAQs at https://aws.amazon.com/snowball/faqs/.

Snowball Edge is ideal for edge processing use cases that require additional computing, memory, and storage power in remote, disconnected, or harsh environments. For more information, see "AWS Snowball Edge device hardware information" in the *AWS Snowball Edge Developer Guide* at https://docs.aws.amazon.com/snowball/latest/developer-guide/device-differences.html.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

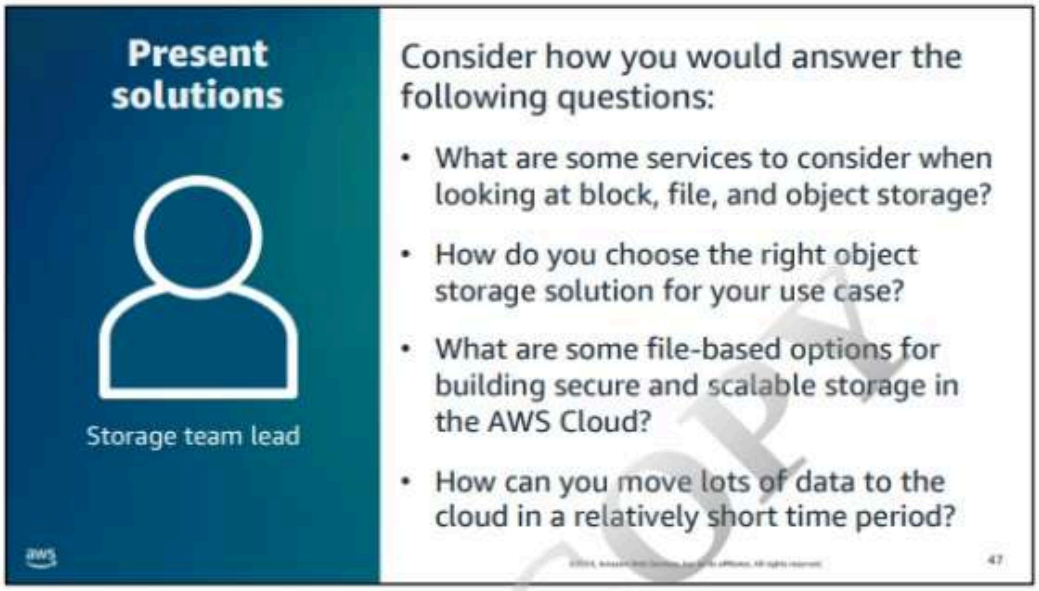
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Imagine that you are now ready to talk to the storage team lead and present solutions that meet their architectural needs.

Your answers should include the following solutions:

- Use Amazon EBS for block-level storage. Amazon S3 and Amazon S3 Glacier have object-level storage options to review. For file-level storage, choose Amazon EFS or Amazon FSx.
- Review the Amazon S3 storage classes. Think about the objects that you are storing and how often they will be accessed. Use that information to choose a cost-effective solution. Consider using Amazon S3 Glacier for long-term archive data.
- Amazon EFS is a quick, straightforward, and scalable file storage solution that supports the NFS protocol. For more specialized solutions, learn more about the Amazon FSx family of services.
- To move data from one source to another destination, use AWS DataSync. For hybrid storage solutions, use AWS Storage Gateway. For an offline way to move data from on-premises to AWS Cloud destinations, use the AWS Snow Family.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Module review

In this module, you learned about the following:

- ✓ Storage services
- ✓ Amazon S3
- ✓ Shared file systems
- ✓ Data migration tools

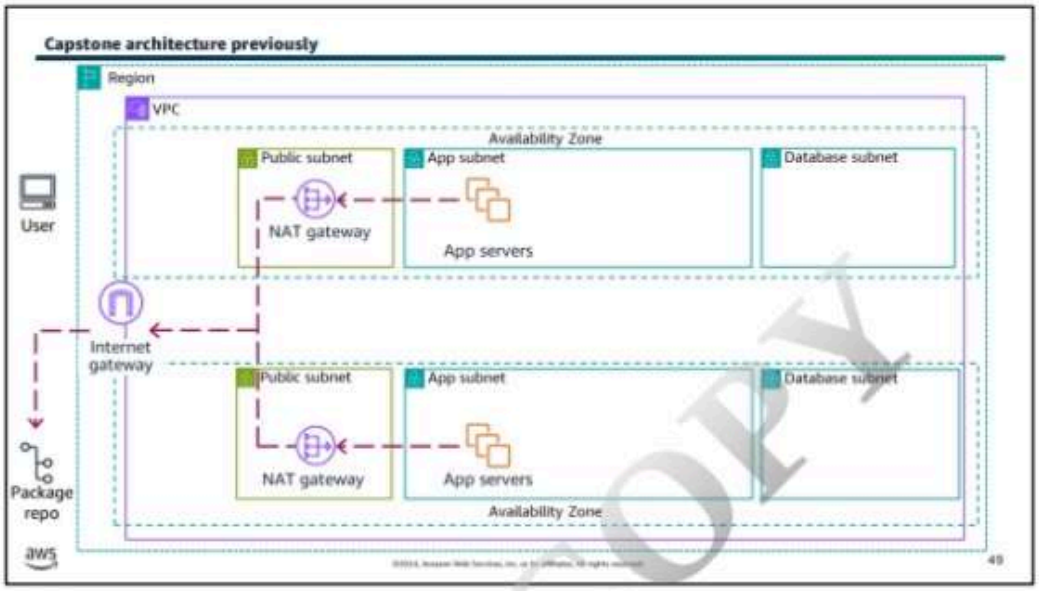
Next, you review the following:

- Capstone check-in
- Knowledge check

48

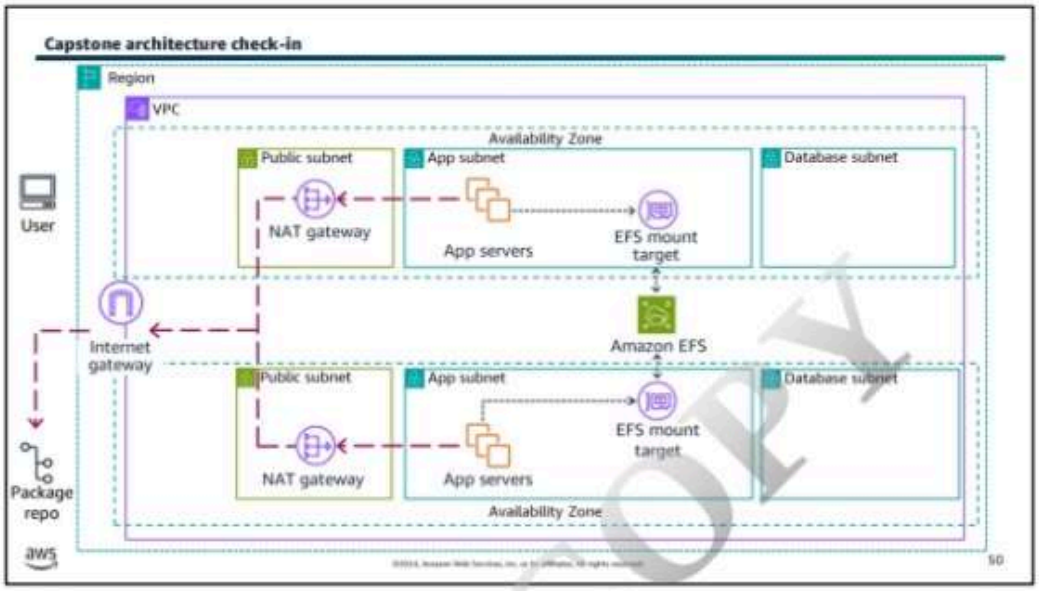
DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



In this module, you explored multiple storage solutions. For the capstone project, you are now able to build a shared file storage system for your WordPress application that uses Amazon EFS Standard.

This solution automatically creates one EFS mount target for you in each Availability Zone. The WordPress installation directory is placed on the EFS mount so that multiple EC2 instances can read it. If one instance fails and is replaced, the files already exist in an accessible and redundant file share.

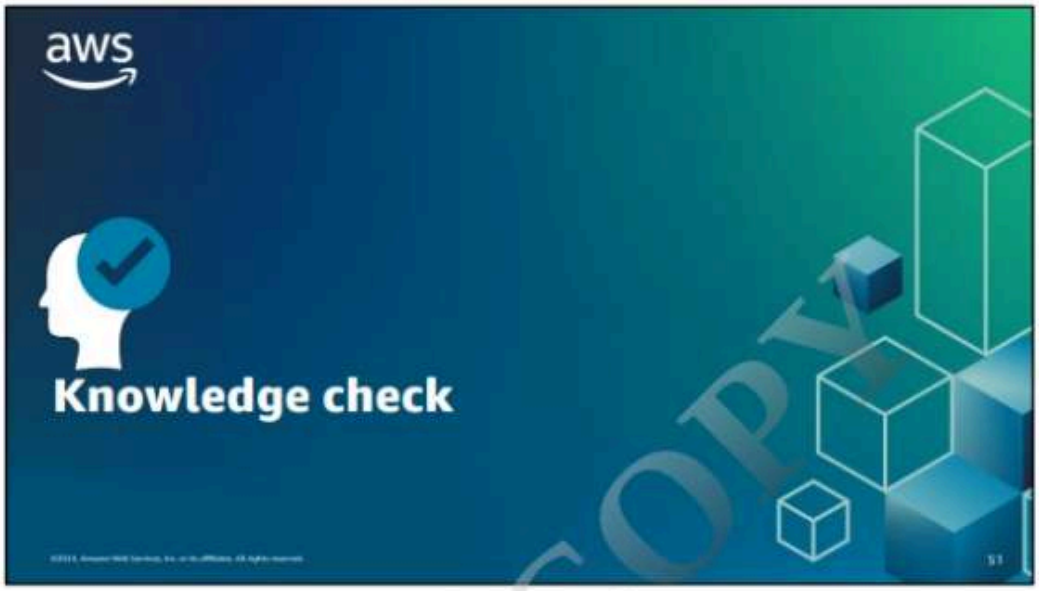
The WordPress website also requires you to build a database to use for the three-tier architecture. You should think about what options there are in AWS for database services. You will explore this topic in the next module.

[Image description: Partial selection of capstone architecture with one VPC in one Region. Two Availability Zones are shown. Each has a public subnet, app subnet, and database subnet. An internet gateway is placed at the edge of the VPC. Each Availability Zone has a NAT gateway in a public subnet. App servers and an EFS mount target are placed in each app subnet. An Amazon EFS file system is placed in the VPC but is connected to EFS mount targets in each app subnet. App servers communicate with the EFS mount targets in their own subnets to reach the EFS file system. App servers send requests to the public internet through a NAT gateway in their Availability Zone. **End description.]**

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 1

Which Amazon S3 feature would be used to automatically copy new objects to a bucket in a different AWS Region?

Choice	Response
A	Same-Region Replication (SRR)
B	Amazon S3 Versioning
C	AWS DataSync
D	Cross-Region Replication (CRR)

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 1 and answer

Which Amazon S3 feature would be used to automatically copy new objects to a bucket in a different AWS Region?

Choice	Response
A	Same-Region Replication (SRR)
B	Amazon S3 Versioning
C	AWS DataSync
D correct	Cross-Region Replication (CRR)

The answer is D, Cross-Region Replication.

You use S3 Cross-Region Replication (CRR) to copy objects across Amazon S3 buckets in different AWS Regions. CRR can help you do the following tasks:

- **Meet compliance requirements** – Although Amazon S3 stores your data across multiple geographically distant Availability Zones by default, compliance requirements might dictate that you store data at greater distances. To satisfy these requirements, use CRR to replicate data between distant AWS Regions.
- **Minimize latency** – Imagine that your customers are in two geographic locations. You can minimize latency in accessing objects by maintaining object copies in AWS Regions that are geographically closer to your users.
- **Increase operational efficiency** – If you have compute clusters in two different AWS Regions that analyze the same set of objects, you might choose to maintain object copies in those Regions.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 2

Which Amazon S3 feature can force an action to occur after an event takes place within a bucket?

Choice	Response
A	Invoking
B	Event Notifications
C	Lambda
D	Alarm

54

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 2 and answer

Which Amazon S3 feature can force an action to occur after an event takes place within a bucket?

Choice	Response
A	Invoking
B correct	Event Notifications
C	Lambda
D	Alarm

The answer is B, event notification.

With Amazon S3 Event Notifications, you can receive notifications when certain object events happen in your bucket. These notifications can invoke actions in other AWS services such as AWS Lambda.

For more information about Amazon S3 Event Notifications, see "Reliable event processing with Amazon S3 Event Notifications" in the *AWS Storage Blog* at https://aws.amazon.com/blogs/storage/reliable-event-processing-with-amazon-s3-event-notifications/.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 3

A developer has two Linux applications in different Availability Zones that must share a common file system. Which solution is best for this use case?

Choice	Response
A	AWS Storage Gateway
B	Amazon FSx for Windows File Server
C	Amazon S3
D	Amazon Elastic File System (Amazon EFS)

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 3 and answer

A developer has two Linux applications in different Availability Zones that must share a common file system. Which solution is best for this use case?

Choice	Response
A	AWS Storage Gateway
B	Amazon FSx for Windows File Server
C	Amazon S3
D correct	Amazon Elastic File System (Amazon EFS)

The answer is D, Amazon EFS.

Other file sharing systems are available to choose from, such as Amazon FSx for Lustre or Amazon FSx for OpenZFS. However, neither of these solutions is presented as a possible answer. Amazon EFS presents a scalable file system that can be mounted with the NFS protocol by multiple Linux EC2 instances.

AWS Storage Gateway supports hybrid environments, which is not mentioned here. FSx for Windows File Server would permit a Linux application to mount the file system through SMB. However, it can be a complicated solution to implement. Amazon S3 is an object storage solution, and it does not natively present you with a mountable file system.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 4

Which modes are available in the AWS Storage Gateway appliance? (Select THREE.)

Choice	Response
A	Memory Gateway
B	Tape Gateway
C	Volume Gateway
D	Amazon EBS File Gateway
E	Amazon S3 File Gateway
F	Amazon S3 Glacier File Gateway

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Knowledge check question 4 and answer

Which modes are available in the AWS Storage Gateway appliance? (Select THREE.)

Choice	Response
A	Memory Gateway
B correct	Tape Gateway
C correct	Volume Gateway
D	Amazon EBS File Gateway
E correct	Amazon S3 File Gateway
F	Amazon S3 Glacier File Gateway

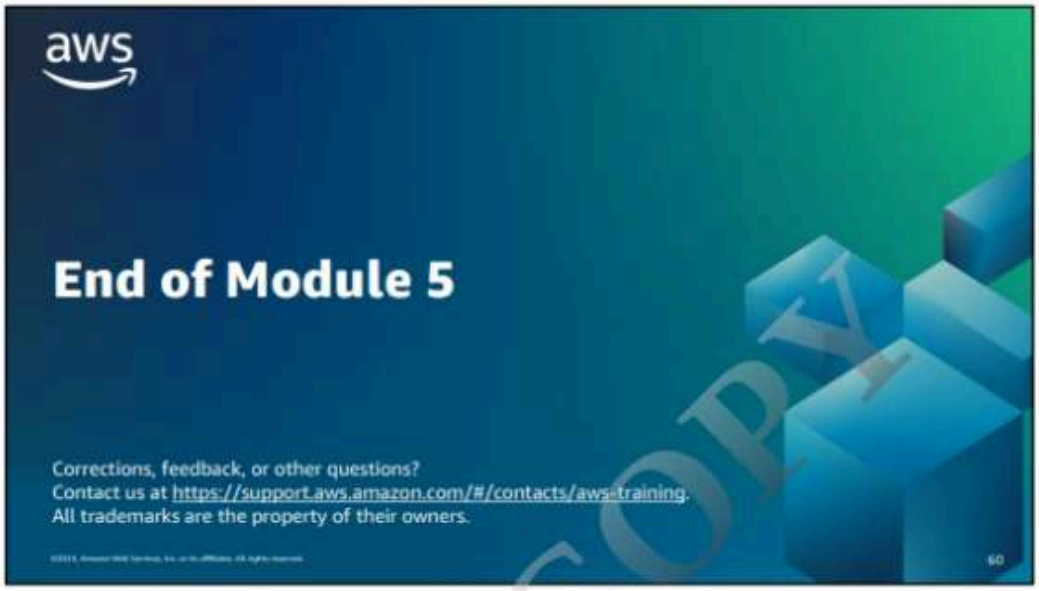
The correct answers are B, C, and E: Tape Gateway, Volume Gateway, and File Gateway.

Tape Gateway, Volume Gateway, and Amazon S3 File Gateway are the three modes available in the AWS Storage Gateway appliance. The Storage Gateway appliance acts as a bridge between on-premises environments and AWS Cloud storage services, so that you can to securely transfer data to and from AWS.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Poll question

Which database technologies have you used in your workloads? (Select all that apply.)

A. Relational databases

B. Nonrelational databases

C. Database caching

D. Database migration tools

E. None of these

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Module overview

- Business request
- Database services
- Amazon Relational Database Service (Amazon RDS)
- Amazon DynamoDB
- Database caching
- Database migration tools
- Present solutions
- Capstone check-in
- Knowledge check
- Lab 3: Create a Database Layer in Your Amazon VPC Infrastructure

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Imagine that your database services manager meets with you to discuss how to manage databases in the cloud. Here are some questions that they are asking. At the end of this module, you meet with the database services manager and present some solutions.

DO NOT COPY
javierartiga.espublico@gmail.com

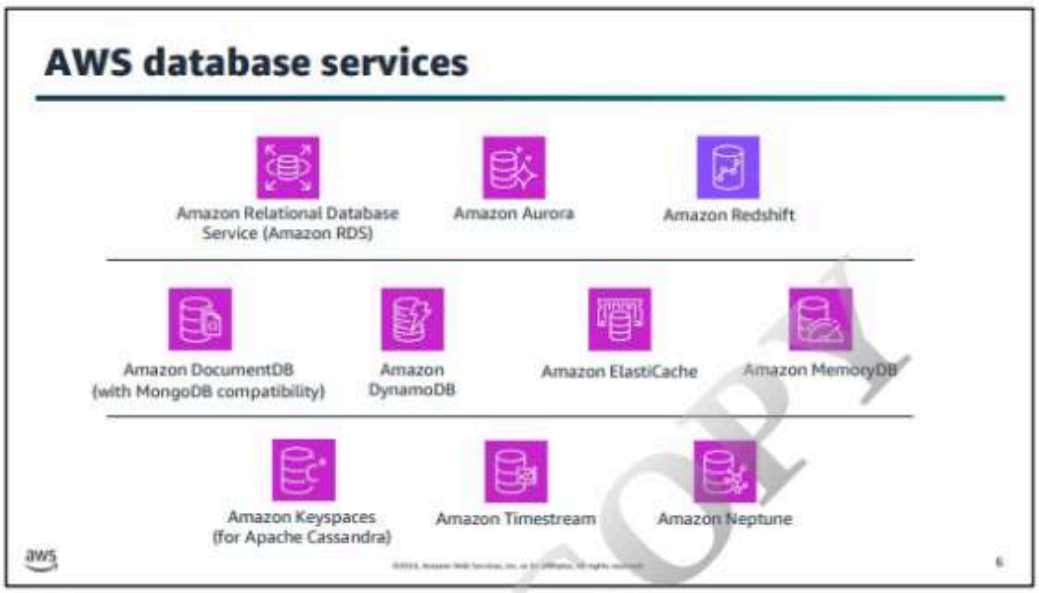
Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

The database services manager asks, "What are the AWS database solutions?" The database team is beginning to investigate opportunities for running databases in the cloud. The company wants you to identify which AWS database solutions they should consider.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



AWS database engines are purpose-built and include relational, key-value, document, in-memory, graph, time series, and wide-column databases. You'll learn how to pick the right database for your needs.

For more information, see "AWS Cloud Databases" at https://aws.amazon.com/products/databases/.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Relational and nonrelational databases

Characteristic	Relational (SQL) databases	Nonrelational (NoSQL) databases
Data storage	Tables with rows and columns	Key-value, wide-column, graph, document, or other models
Schemas	Fixed	Dynamic
Example database services	Amazon RDS, Aurora	DynamoDB, ElastiCache

For decades, the predominant data model that was used for application development was the relational data model. Relational databases such as Oracle, IBM DB2, SQL Server, MySQL, and PostgreSQL used this model. It wasn't until the mid- to late 2000s that other data models began to gain significant adoption and usage. To differentiate and categorize these new classes of databases and data models, the term *NoSQL* was coined. Often the term NoSQL is used interchangeably with *nonrelational*.

A *relational database* is a collection of data items with predefined relationships between them. These items are organized as a set of tables with columns and rows. Each column in a table holds a certain kind of data, and a field stores the actual value of an attribute. The rows in the table represent a collection of related values of one object or entity. This data can be accessed in many different ways without reorganizing the database tables themselves. This module focuses on two SQL databases services: Amazon Relational Database Service (Amazon RDS) and Amazon Aurora.

NoSQL is a term that is used to describe nonrelational database systems that are highly available, scalable, and optimized for high performance. Compared to the relational model, NoSQL databases use alternative models for data management, such as key-value pairs or document storage. This module focuses on two NoSQL databases services: Amazon DynamoDB and Amazon ElastiCache.

For more information about relational databases, see "What Is a Relational Database?" at https://aws.amazon.com/relational-database/.

For more information about nonrelational databases, see "What Is NoSQL?" at https://aws.amazon.com/nosql/.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

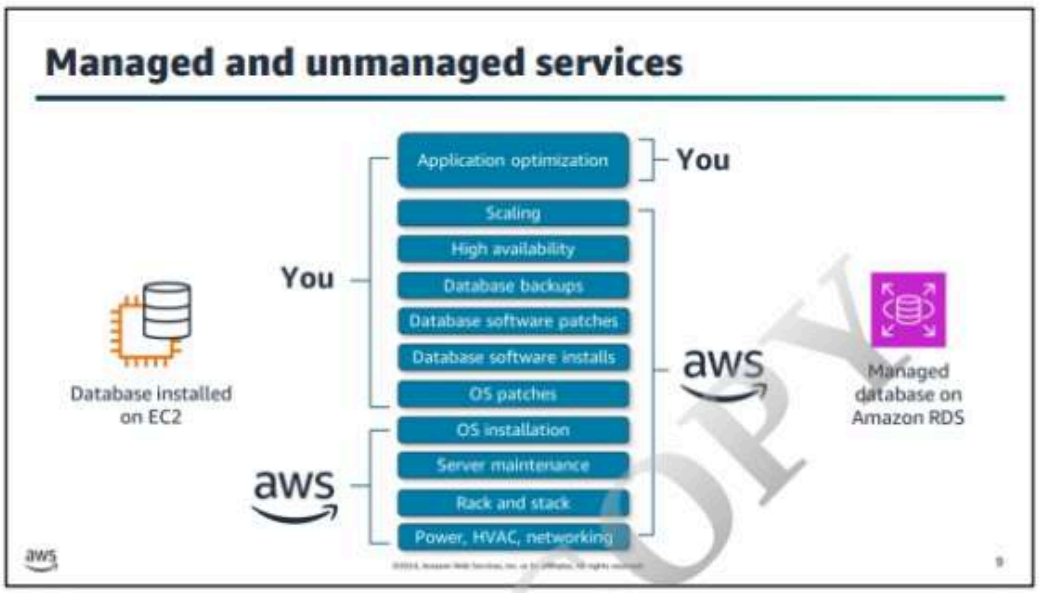
Choosing the right database

Relational database	Nonrelational (NoSQL) database
You require strict schema rules and data quality enforcement.	You need your database to scale horizontally.
Your database doesn't need extreme read/write capacity.	Your data does not lend itself well to traditional schemas.
If you have a relational dataset that does not require extreme performance, a relational database management system can be the best, lowest effort solution.	Your read/write rates exceed the rates that can be economically supported through a traditional SQL database.

Though many types of databases exist with varying features, this table shows some of the differences between SQL (relational) and NoSQL (nonrelational) databases.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



When building resources in the cloud, consider the level of control that you need and the resources that you have for managing that resource.

For example, when running a database in the cloud, you can install a database on an Amazon Elastic Compute Cloud (Amazon EC2) instance. Or you can choose a managed database option, such as Amazon RDS:

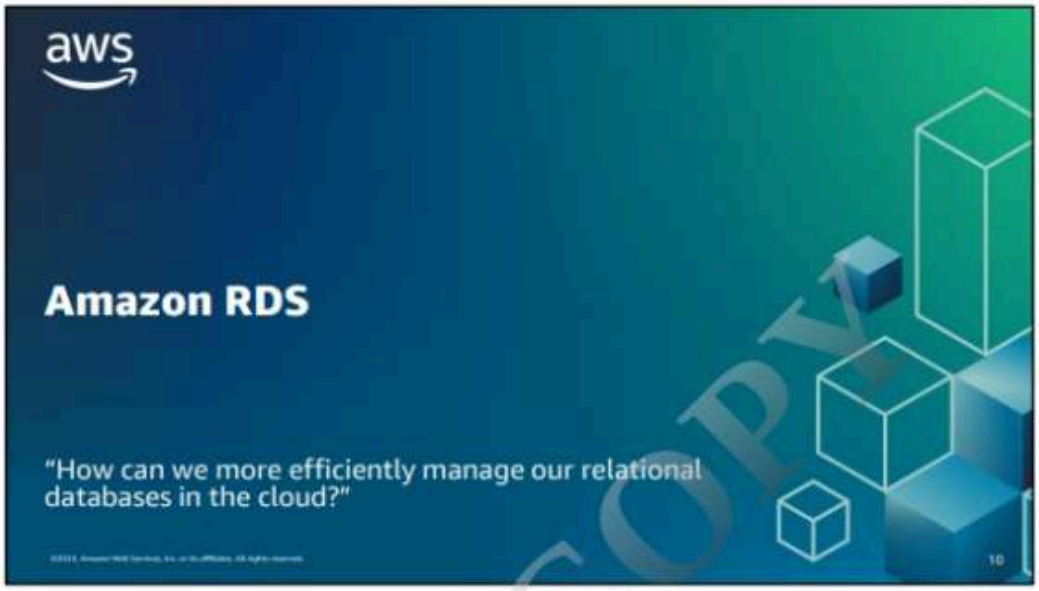
- Installing a database on an EC2 instance gives you complete control over all aspects of the database except the hardware. The trade-off is that this increased amount of control requires more resources and expertise to manage.
- Using a managed database service removes the undifferentiated work of managing your databases.

Databases that AWS manages provide systems for high availability, scalability, and backups. You can choose to use scaling, high availability, database backups, database software patches, database software installs, and operating system (OS) patches. In general, you are responsible only for optimizing your applications to make sure that the database layer works as well as possible with your application.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

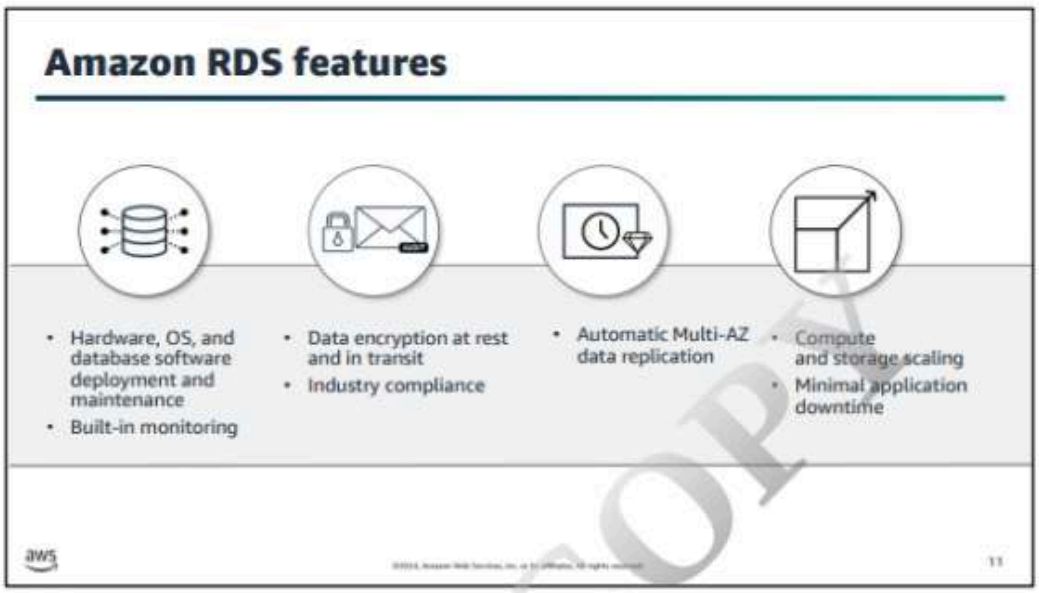


The database services manager asks, "How can we more efficiently manage our relational databases in the cloud?" The database team has relational databases and is considering Amazon RDS. The company wants you to explain the benefits of running a database in a managed service and examine the features of Amazon RDS.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



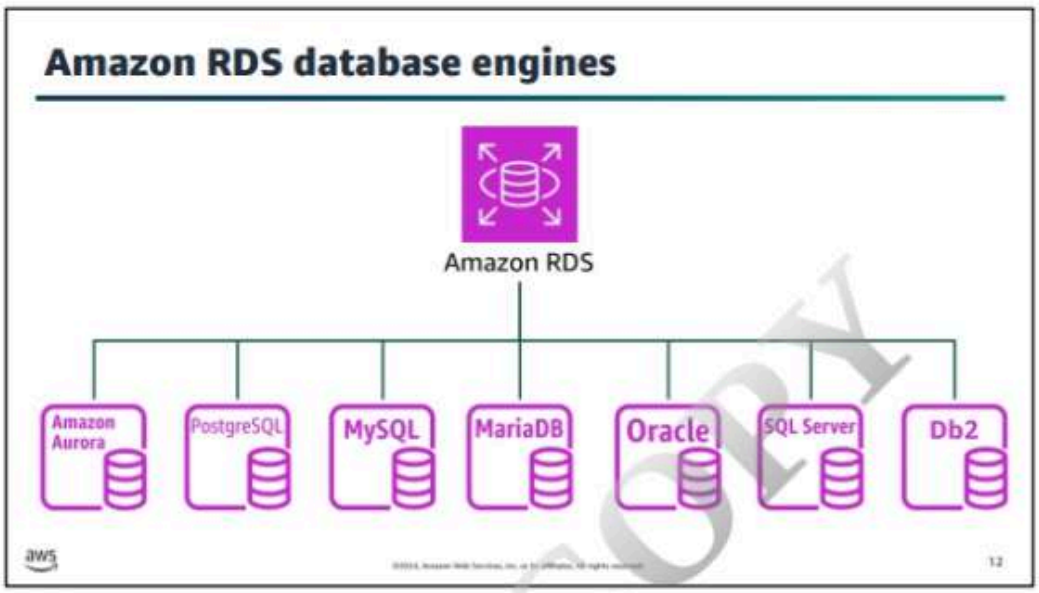
Amazon RDS is a web service that helps you to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while managing time-consuming database administration tasks. By using Amazon RDS, you can focus on your applications and business. Amazon RDS provides you with seven familiar database engines to choose from, including Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle Database, and Microsoft SQL Server. Therefore, most of the code, applications, and tools that you already use with your existing databases can be used with Amazon RDS.

Amazon RDS automatically patches the database software and backs up your database. It stores the backups for a user-defined retention period and provides point-in-time recovery. You benefit from the flexibility of scaling the compute resources or storage capacity associated with your relational DB instance with a single API call.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Amazon RDS is available on seven database engines, which optimize for memory, performance, or I/O. The database engines include the following:

- Amazon Aurora
- PostgreSQL
- MySQL
- MariaDB
- Oracle Database
- SQL Server
- IBM Db2

DO NOT COPY
javierartiga.espublico@gmail.com

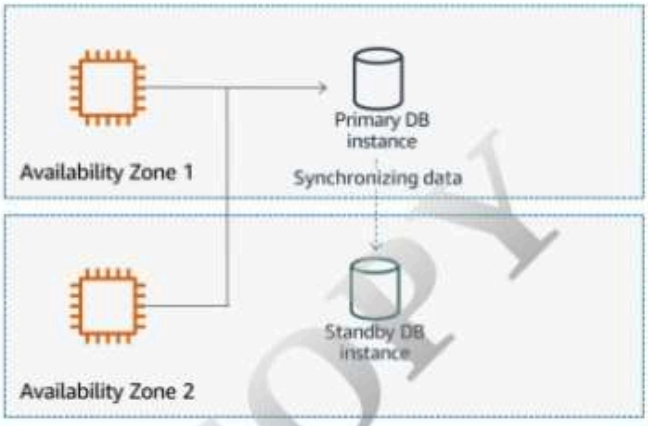
© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Amazon RDS Multi-AZ deployments

Multi-AZ deployments:

- Replicate data to a standby DB instance in another Availability Zone
- Are not used for read-only scenarios



Amazon RDS Multi-AZ deployments provide enhanced availability and durability for database (DB) instances, which makes them a natural fit for production database workloads. When you provision a Multi-AZ DB instance, Amazon RDS synchronously replicates the data to a standby instance in a different Availability Zone.

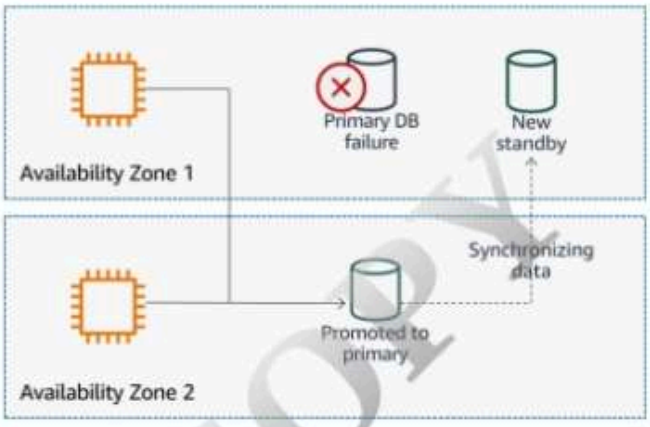
You can modify your environment from Single-AZ to Multi-AZ at any time. Each Availability Zone runs on its own physically distinct, independent infrastructure and is engineered to be highly reliable.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Amazon RDS Multi-AZ failover

Upon failure, the standby DB instance picks up the load.

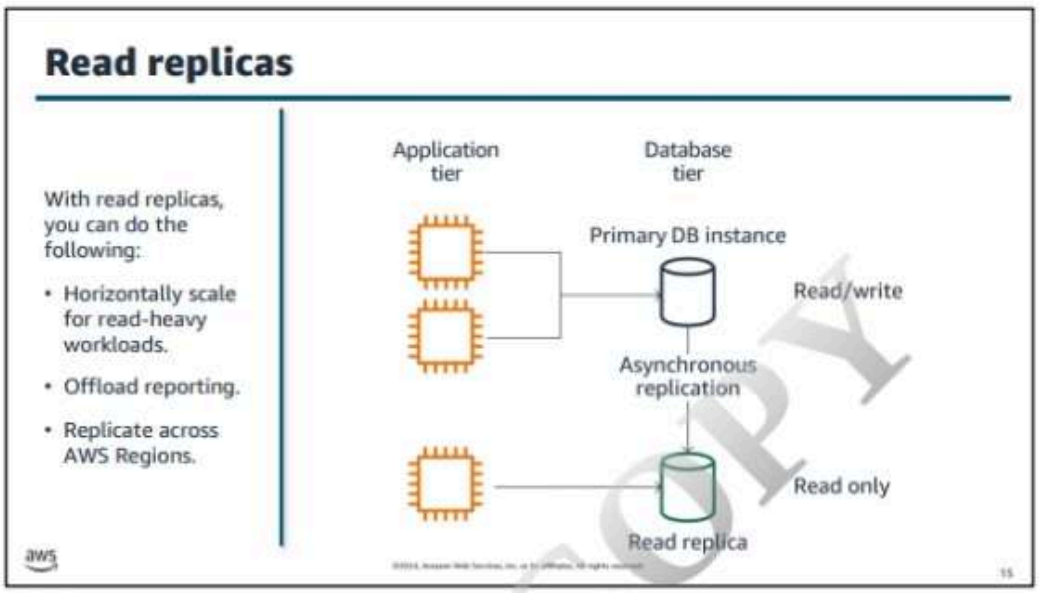


In case of the primary instance failure, Amazon RDS performs an automatic failover to the standby instance.

In this example, two EC2 instances in separate Availability Zones are connected to the primary database in one Availability Zone. A standby database is hosted in the other Availability Zone. When the primary database fails, Amazon RDS promotes the secondary database to primary. Because it assumes the primary databases endpoint, the EC2 instances can resume traffic with the new primary database. Meanwhile, a new standby database is created in the other Availability Zone.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



With Amazon RDS, you can create read replicas of your database. Amazon automatically keeps them in sync with the primary DB instance. Read replicas are available in Amazon RDS for Aurora, MySQL, MariaDB, PostgreSQL, Oracle, and Microsoft SQL Server. Read replicas can help you do the following tasks:

- Relieve pressure on your primary node with additional read capacity.
- Bring data close to your applications in different AWS Regions.
- Promote a read replica to a standalone instance as a disaster recovery (DR) solution if the primary DB instance fails.

You can add read replicas to handle read workloads so that your primary database doesn't become overloaded with read requests. Depending on the database engine, you can also place your read replica in a different Region from your primary database. This placement gives you the ability to have a read replica closer to a particular location.

You can configure a source database as Multi-AZ for high availability and create a read replica in Single-AZ for read scalability. With Amazon RDS for MySQL and MariaDB, you can also set the read replica as Multi-AZ and as a DR target. When you promote the read replica to be a standalone database, it will be replicated to multiple Availability Zones.

For more information about read replicas, see "Working with DB Instance Read Replicas" in the *Amazon Relational Database Service User Guide* at https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/USER_ReadRepl.html#USER_ReadRepl.PostgreSQL.

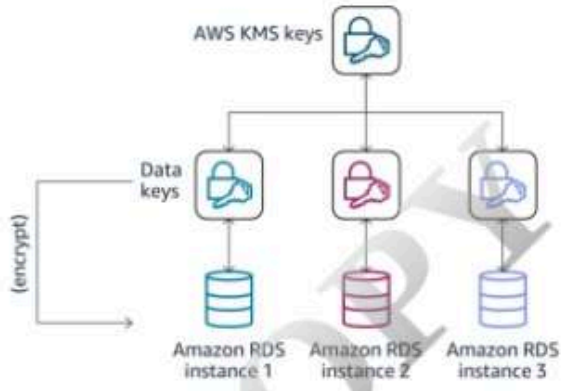
[**Image description:** Two instances in the application tier connect to a primary database instance in the database tier. This database can perform read/write operations. Amazon RDS uses asynchronous replication to create a read replica of this database. A third instance in the application tier connects to the read replica, which can perform only read operations. **End Description.**]

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Data encryption at rest

- Encryption of data at rest is managed by AWS KMS.
- A unique data key encrypts your data.
- An AWS KMS key encrypts data keys.
- Data encryption is available for all Amazon RDS engines.

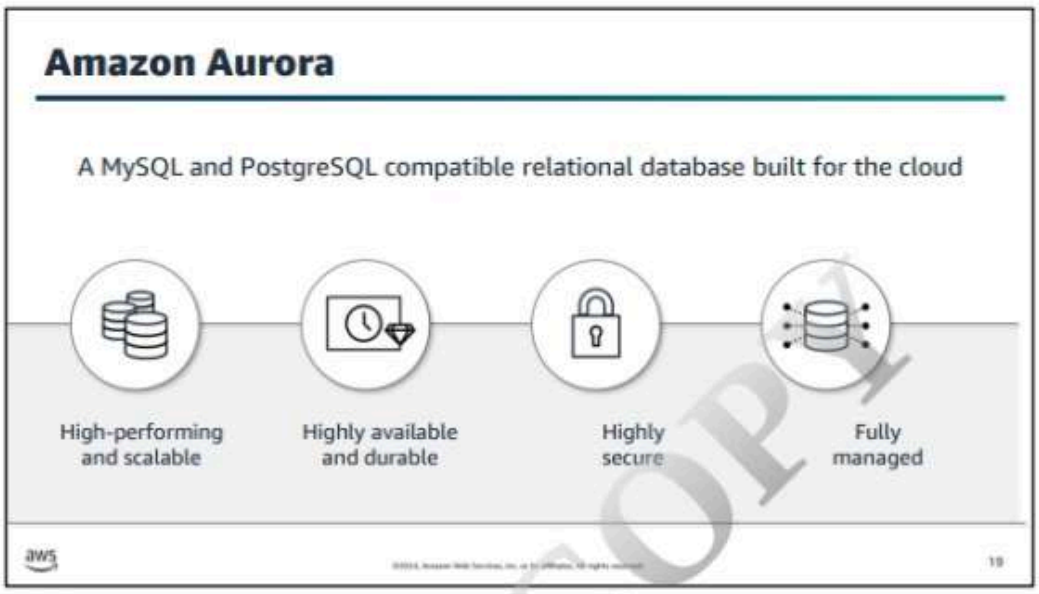


Amazon RDS provides encryption of data at rest by using the AWS Key Management Service (AWS KMS). AWS KMS is a managed service that provides the ability to create and manage encryption keys and then encrypt and decrypt your data with those keys. All of these keys are tied to your AWS account, and you fully manage them. AWS KMS provides an additional layer of protection against unauthorized access to the underlying storage of your Amazon RDS instance. AWS KMS uses industry-standard AES-256 encryption to protect data that is stored on the underlying host that your Amazon RDS instance is running on.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



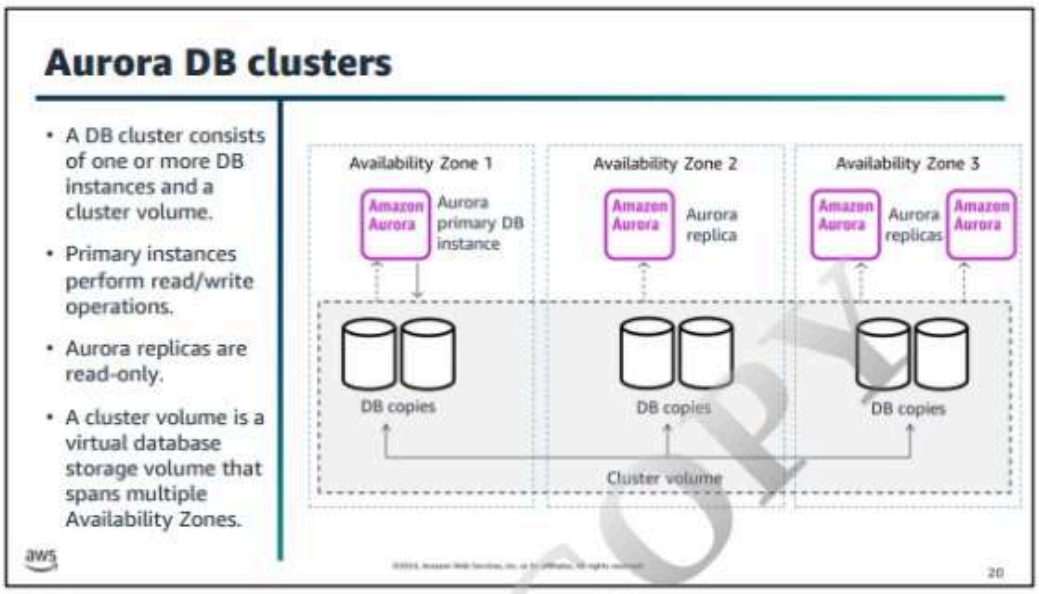
Amazon Aurora is an enterprise-class relational database. It is compatible with MySQL and PostgreSQL relational databases. It is up to five times faster than standard MySQL databases and up to three times faster than standard PostgreSQL databases. Aurora helps to reduce your database costs by reducing unnecessary I/O operations while helping ensure that your database resources remain reliable and available. Consider Aurora if your workloads require high availability. It replicates six copies of your data across three Availability Zones and continuously backs up your data to Amazon Simple Storage Service (Amazon S3).

Aurora supports network isolation, encryption at rest and in transit, and compliance and assurance programs. Aurora is managed by Amazon RDS, so it requires no server provisioning, software patching, setup, configuration, or backups.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



An Amazon Aurora DB cluster consists of one or more DB instances and a cluster volume that manages the data for those DB instances. The instances perform the compute functions of the database while the cluster volume stores the actual data.

Aurora offers two instance types:

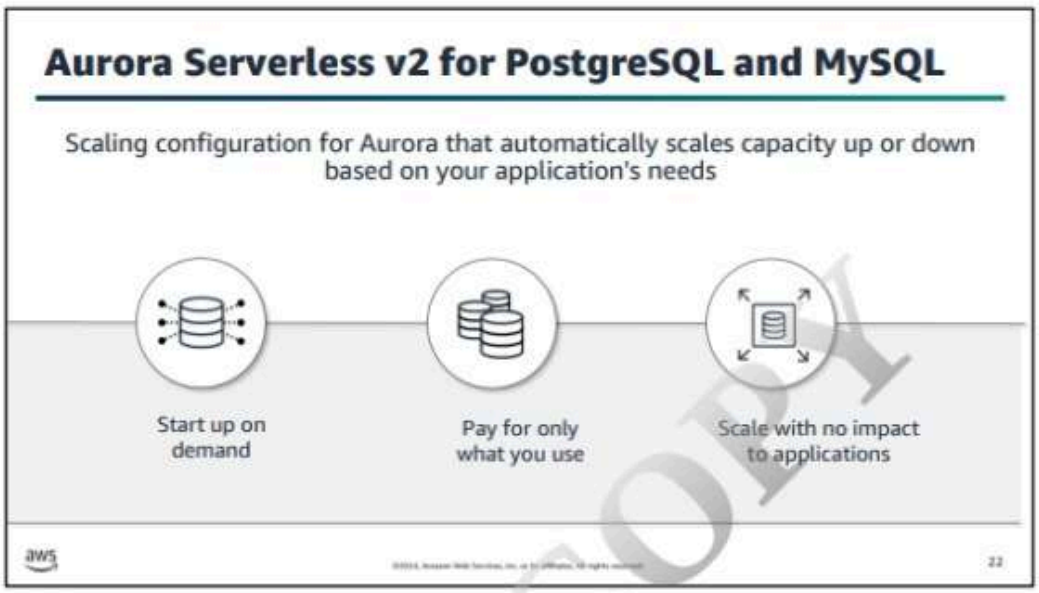
- **Primary instance:** This instance type supports read and write operations and performs all the data modifications to the cluster volume. Each Aurora DB cluster has one primary instance.
- **Aurora replica:** This instance type supports read operations only. Each Aurora DB cluster can have up to 15 Aurora replicas in addition to the primary instance. Multiple Aurora replicas distribute the read workload. You can increase availability by locating Aurora replicas in separate Availability Zones. You can have a read replica in the same Region as the primary instance.

An Aurora cluster volume is a virtual database storage volume that spans multiple Availability Zones. Each Availability Zone has a copy of the DB cluster data. Storage in the cluster volume is replicated across hundreds of storage nodes. Aurora presents the cluster volume as a single, logical volume to the primary instance and to Aurora replicas in the DB cluster. Write operations to the cluster volume are often available to Aurora replicas in less than 100 milliseconds.

DO NOT COPY javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Aurora Serverless v2 is an on-demand, auto scaling configuration for Amazon Aurora. Aurora Serverless v2 helps to automate the processes of monitoring the workload and adjusting the capacity for your databases. Capacity is adjusted automatically based on application demand. You're charged for only the resources that your DB clusters consume. Thus, Aurora Serverless v2 can help you to stay within budget and avoid paying for computer resources that you don't use.

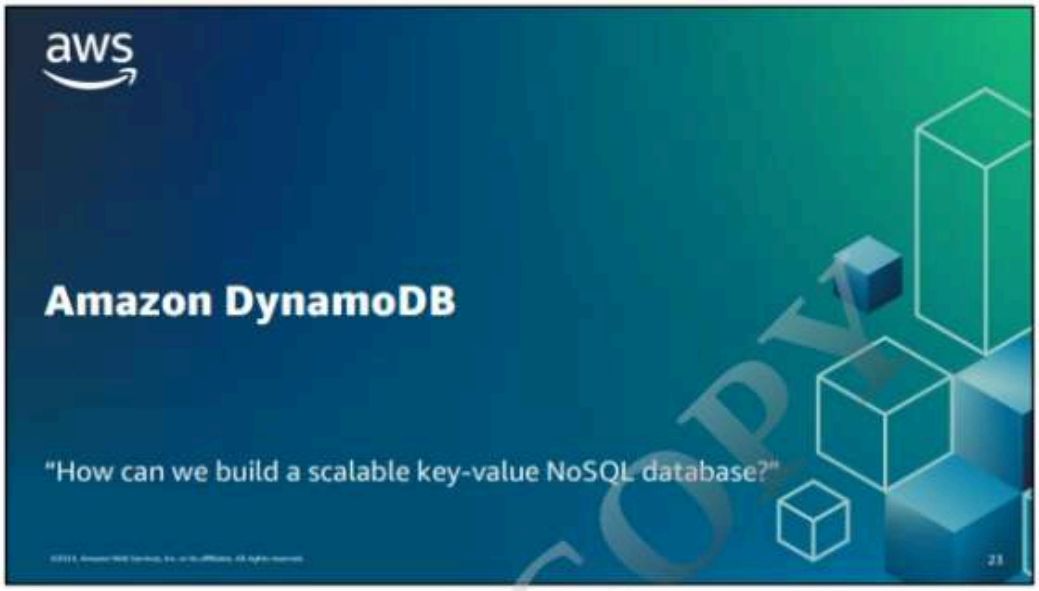
With Aurora Serverless v2, your database automatically scales capacity to accommodate the application's peak load and scales back down when the surge of activity ceases. With Aurora Serverless v2, you no longer need to provision for peak or average capacity. You can specify an upper capacity limit to handle your most demanding workloads, and that capacity isn't used unless it's needed.

You also set your minimum capacity setting. The scaling rate for an Aurora Serverless v2 DB instance depends on its current capacity. The higher the current capacity, the faster it can scale up. You might need the DB instance to quickly scale up to a very high capacity. If so, consider setting the minimum capacity to a value where the scaling rate meets your requirement. This type of automation is especially valuable for multitenant databases, distributed databases, development and test systems, and other environments with highly variable and unpredictable workloads.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

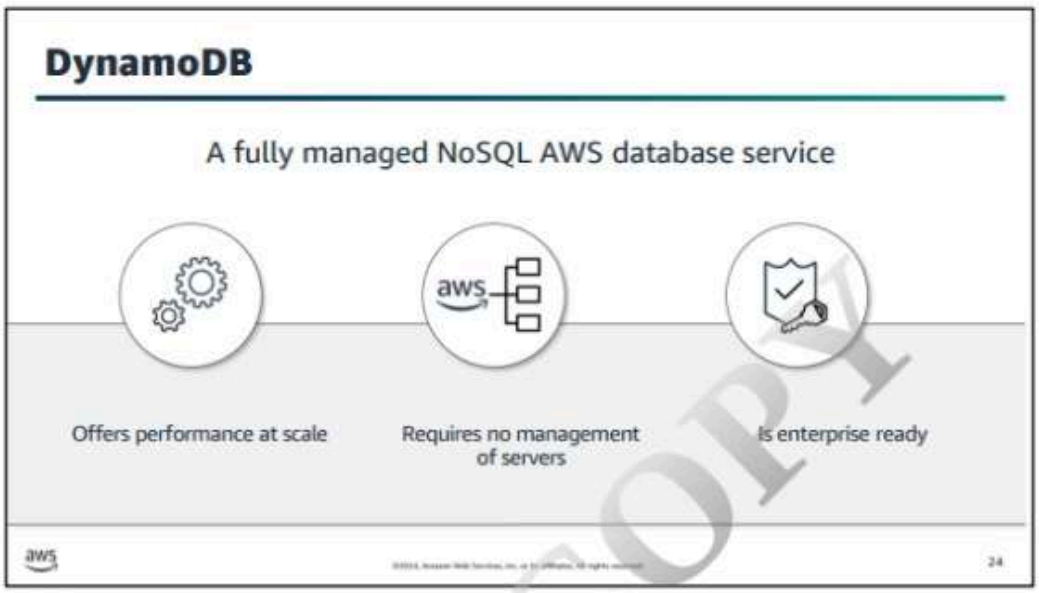


The database services manager asks, "How can we build a scalable key-value NoSQL database?" The database team is considering using NoSQL databases in some of their workloads. The company wants you to identify a high-performance key-value database solution.

DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



DynamoDB is a fully managed NoSQL database service. The service itself manages the complexity of running this massively scalable, distributed NoSQL database. The software developers can focus on building applications instead of managing infrastructure.

NoSQL databases are designed for scale, but their architectures are sophisticated, and significant operational overhead can be a part of running a large NoSQL cluster. DynamoDB removes the need to become an expert in advanced distributed computing concepts. You only need to learn the straightforward API of DynamoDB by using the SDK for the programming language of choice.

DynamoDB is cost effective. You pay for the storage that you are consuming and the I/O throughput that you have provisioned. It is designed to scale elastically while maintaining high performance. You can choose to provision a small amount of capacity when the storage and throughput requirements of an application are low. If you choose auto scaling, additional capacity is provisioned when the required I/O throughput increases within limits that you set. The on-demand choice permits an application to seamlessly grow to support millions of users that make thousands of concurrent requests to the database every second.

DynamoDB supports end-to-end encryption and fine-grained access control.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Key-value data

- Is structured in key-value pairs with a flexible schema
- Is ideal for uses where the needed data can be mapped to a primary key
- Partitions data by key
- Delivers high-throughput, low-latency reads and writes

Database name: Gamers

Primary key

Attributes

GamerTag	UserId	Level	Points	TopScore	Plays
"Hammer57"	107	21	4,050	483,610	1,722
"FluffyDuffy"	285	5	1,123	10,863	43

25

NoSQL databases use a variety of data models to access and manage data. These types of databases are optimized specifically for applications that require large data volume, low latency, and flexible data models. NoSQL databases are optimized by relaxing some of the data consistency restrictions of other databases. One common model is key-value data.

Key-value databases are good for use cases where the requested data can be associated with a single primary key. Consider this example of a video game's user profile database. Each item has a collection of key-value pairs, including keys such as *TopScore*, *UserID*, and *Level*. All key-value pairs for an item are associated with the primary key, *GamerTag*. You can rapidly retrieve any of these key-value pairs by locating their *GamerTag*.

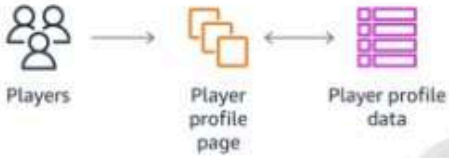
DO NOT COPY
javierartiga.espublico@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DynamoDB use case 1

Player profile page



GamerTag	UserId	TopScore	MemberSince	SubscriptionType
"HammerS7"	101	5,842	"2021-09-15:17:24:31"	"Gold"
"FluffyDuffy"	243	1,024	"2021-10-22:23:18:01"	"Platinum"
"NewPlayer"	623	687	"2021-10-22:23:22:01"	"Free"

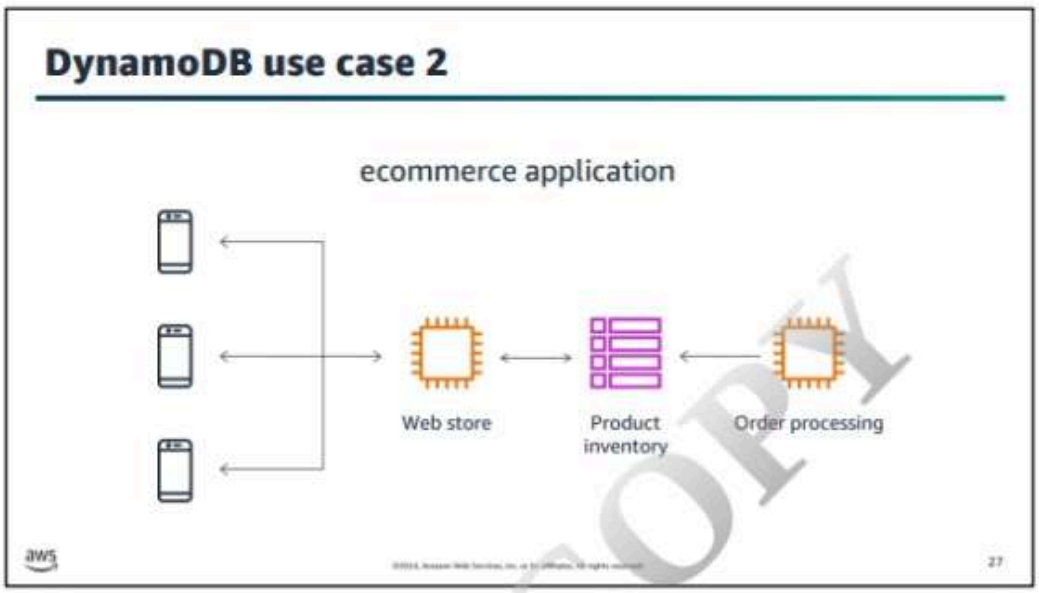
26

Game makers can support basic player profile pages by using DynamoDB. You can store user profile data in DynamoDB by using *GamerTag* as the key. When a profile page loads, the application needs to make only a single read request to DynamoDB by using the GamerTag value. In addition, when a new game launches, DynamoDB can scale rapidly to provide enough storage and throughput to support spikes in traffic.

For more information about gaming use cases, see "Amazon DynamoDB: Gaming Use Cases and Design Patterns" in the *AWS Database Blog* at https://aws.amazon.com/blogs/database/amazon-dynamodb-gaming-use-cases-and-design-patterns/.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.



Imagine that you have an ecommerce application to sell products to your customers. You need the web store to display an accurate inventory. This task becomes a significant challenge during peak traffic and holiday sales. DynamoDB supports a flexible schema to accommodate data for a variety of products. It can also manage many concurrent read and write operations while maintaining the accuracy of stored data.

DO NOT COPY
javierartiga.espublico@gmail.com

Printed by: javierartiga.espublico@gmail.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DynamoDB tables

- There is a mandatory key-value access pattern.
- The partition key determines the data distribution.
- The sort key permits rich query capabilities.

Composite primary key

Partition key | Sort key (optional)

Table name: Gamers

GamerTag	GameID	TopScore	Genre	Subscription
"Hammer57"	"PuzzleGame"	483,610	"puzzle"	
"FluffyDuffy"	"MysteryGame"	10,863		true

Item

Attributes

28

DynamoDB stores data in tables. When creating a table, you must specify a table name and a partition key, which are the only two required entities.

DynamoDB uses *primary keys* to uniquely identify each item in a table and *secondary indexes* to provide more query flexibility. Two types of primary keys are supported:

- **Simple primary key:** A simple primary key consists of just one attribute, which is designated as the *partition key*. If you use only the partition key, no two items can have the same value.
- **Composite primary key:** A composite primary key consists of both a partition key and a *sort key*. In this case, the partition key value for multiple items can be the same, but their sort key values must be different.

You work with the core components: tables, items, and attributes. A *table* is a collection of *items*, and each item is a collection of *attributes*. In this example, the table includes two items, with simple primary keys *Hammer57* and *FluffyDuffy*. The item with the primary key *Hammer57* includes three attributes: *GameID*, *TopScore*, and *Genre*. The primary key for *FluffyDuffy* includes a *Subscription* attribute, and it does not include the *Genre* attribute.

For more information about the components of DynamoDB, see "Core Components of Amazon DynamoDB" in the *Amazon DynamoDB Developer Guide* at https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/HowItWorks.CoreComponents.html.

DO NOT COPY javierartiga.esplier@gmail.com

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.