Passing the AWS Cloud Practitioner exam

Study Group #4: S3, Databases, and more

Presented By:
Christophe Limpalair,
Linux Academy

Linux Academy

Homework Quiz

1) Why is AWS more economical than traditional data centers for applications with varying compute workloads?

A) **Amazon Elastic Compute Cloud (Amazon EC2) costs are billed on a monthly basis.**

B) **Customers retain full administrative access to their Amazon EC2 instances.**

C) **Amazon EC2 instances can be launched on-demand when needed.**

D) **Customers can permanently run enough instances to handle peak workloads**

Homework Quiz

1) Why is AWS more economical than traditional data centers for applications with varying compute workloads?

A) **Amazon Elastic Compute Cloud (Amazon EC2) costs are billed on a monthly basis.**

B) **Customers retain full administrative access to their Amazon EC2 instances.**

C) **Amazon EC2 instances can be launched on-demand when needed.**

D) **Customers can permanently run enough instances to handle peak workloads**

Homework Quiz

3) Which AWS offering enables customers to find, buy, and immediately start using software solutions in their AWS environment?

A) **AWS Config**
B) **AWS OpsWorks**
C) **AWS SDK**
D) **AWS Marketplace**

Homework Quiz

3) Which AWS offering enables customers to find, buy, and immediately start using software solutions in their AWS environment?

A) **AWS Config**

B) **AWS OpsWorks**

C) **AWS SDK**

D) **AWS Marketplace**

Homework Quiz

4) Which AWS networking service enables a company to create a virtual network within AWS?

A) **AWS Config**

B) **Amazon Route 53**

C) **AWS Direct Connect**

D) **Amazon Virtual Private Cloud (Amazon VPC)**

Homework Quiz

4) Which AWS networking service enables a company to create a virtual network within AWS?

A) **AWS Config**

B) **Amazon Route 53**

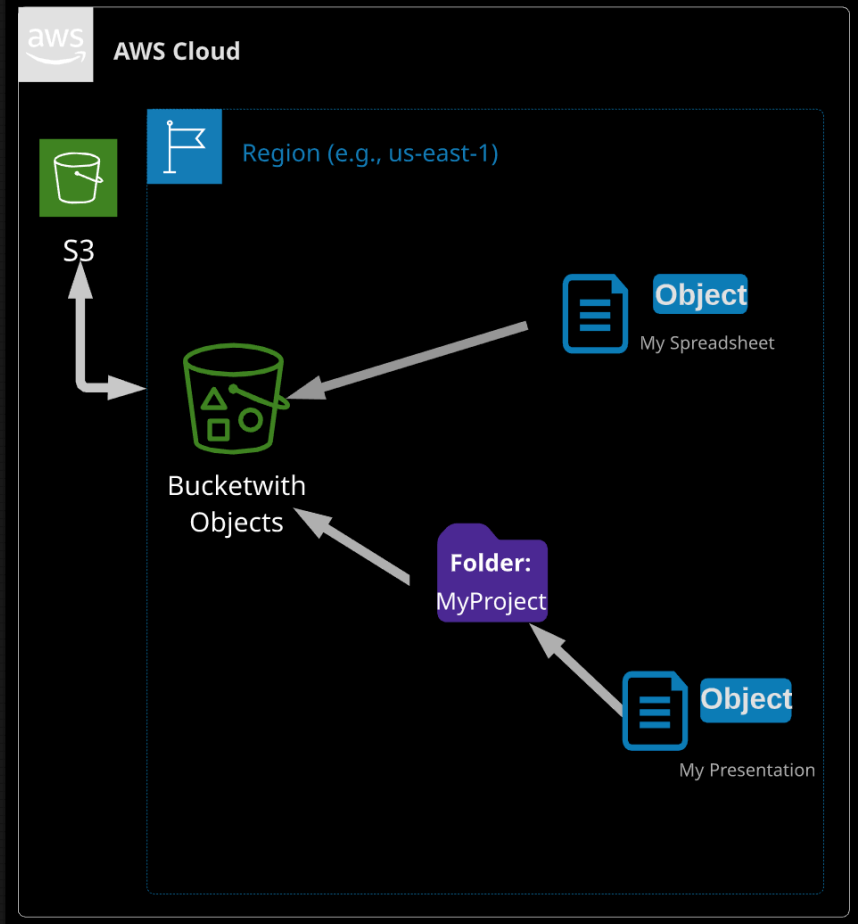
C) **AWS Direct Connect**

D) **Amazon Virtual Private Cloud (Amazon VPC)**

Amazon S3

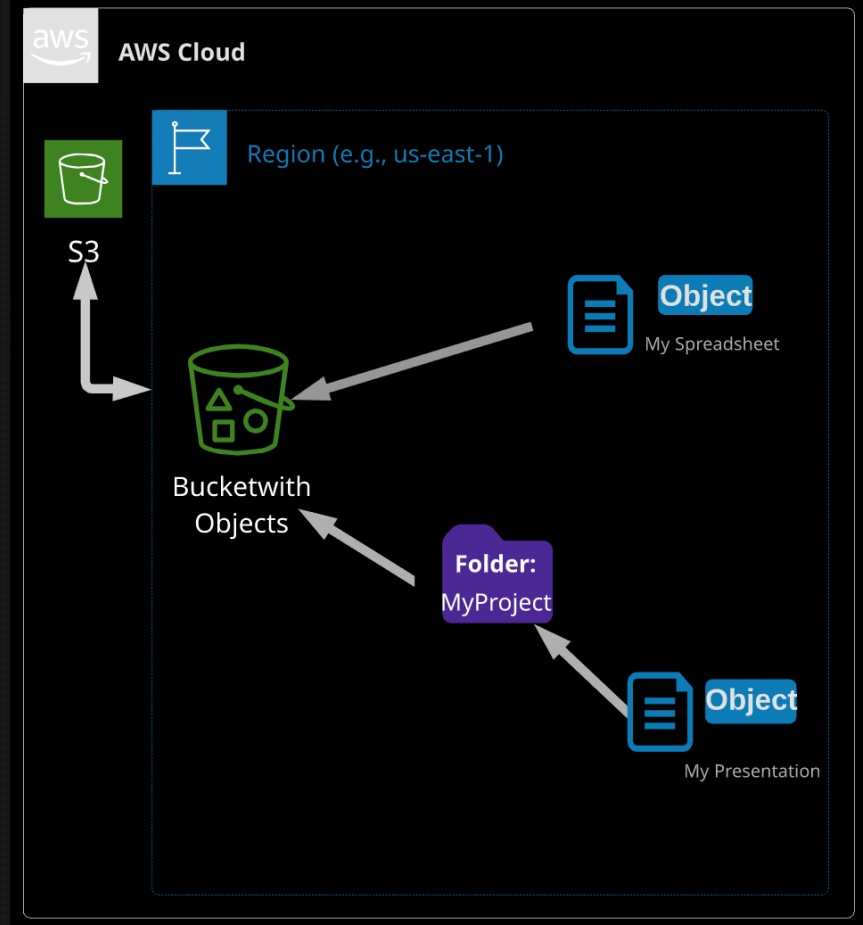
Amazon S3 is an **object storage service** that is built to be scalable, high available, secure, and performant. Customers of all sizes and industries can use it to store and protect any amount of data for a range of use cases.

- Use cases:
  - Websites, mobile apps, backup and restore, applications, big data analytics, etc
- S3 is structured with:
  - Buckets
  - Folders
  - Objects



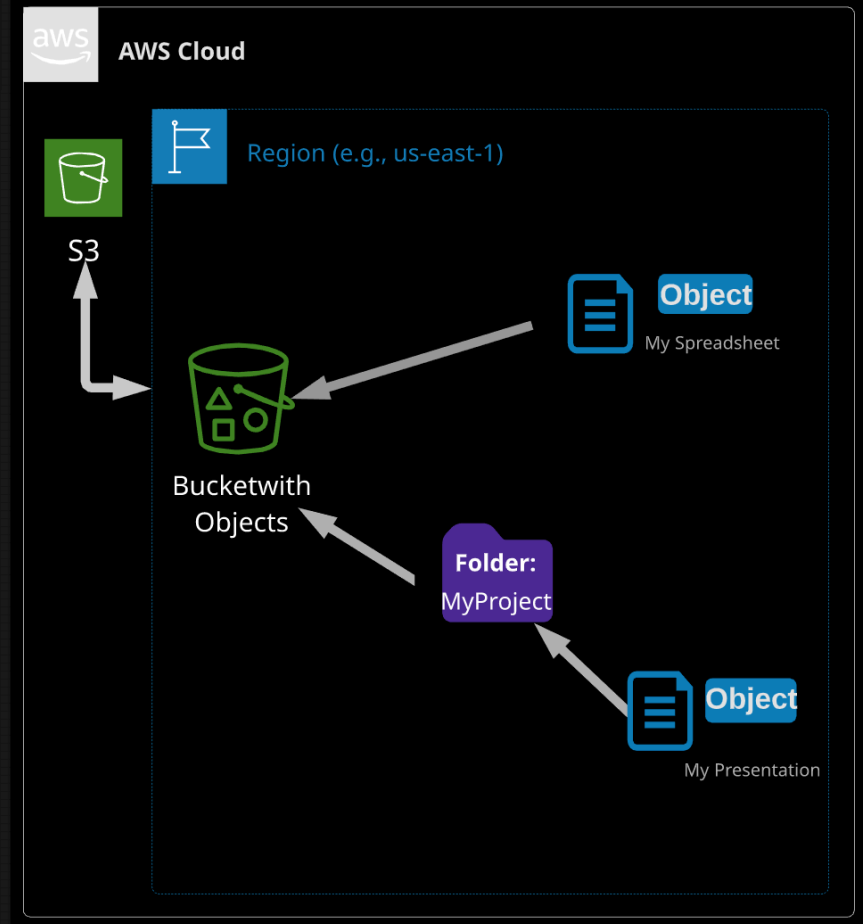
Amazon S3 - Buckets

- Buckets
  - Highest level used to group folders and objects together
  - Apply **bucket-wide policies**
  - Can be used to host static websites
  - Created in regions
- Bucket names:
  - Must be unique across all of AWS
  - 3 to 63 characters in length
  - Can only contain lower case letters, numbers, hyphens
  - Must not be formatted as an ip address (ie 192.168.1.1)



Amazon S3 – Folders & Objects

- Folders
  - Subfolders in buckets, to help organize
- Objects
  - Any file stored
  - Supports a wide range of formats
  - Once stored in S3, you or your resources can access these objects if you have the correct permissions
- Versioning
  - Can be enabled to preserve, retrieve, or restore every version of objects stored
  - Enabled at the bucket level



Amazon S3 – Storage Classes

- Standard
  - Most expensive storage class
  - Designed for general, all-purpose storage
  - Default storage option
  - 99.999999999% (“11 nines”) object durability
  - 99.99% object availability
- Standard-IA
  - Designed for objects that you do not access frequently but must be immediately available when accessed (uses multiple AZs)
  - 99.999999999% object durability
  - 99.90% object availability
  - Less expensive than the standard storage class
  - Great for disaster recovery, backups

Amazon S3 – Storage Classes

- One Zone-IA
  - Designed for objects that you do not access frequently but must be quickly available when accessed (uses only one AZ)
  - 99.99% object durability
  - 99.50% object availability
  - ~20% less expensive than the Standard-IA storage class

- Intelligent-Tiering
  - Designed to optimize costs by automatically moving data to the most cost-effective tier based on your usage
  - Monitors file for usage and moves them after 30 days of inactivity
  - 99.999999999% (“11 nines”) object durability
  - 99.90% object availability
  - Pricing depends on the assigned storage class

Amazon S3 – Storage Classes

- Glacier & Deep Archive
  - Designed for long-term archival storage
  - May take several hours to retrieve objects
  - 99.999999999% object durability
  - Cheapest storage class in AWS
  - Deep Archive has a longer storage minimum and longer retrieval time (only for 1-2x access/year)

- Object Lifecycles
  - Transition objects to different storage classes after a certain period of time
  - Delete objects after a certain period of time
  - Use cases:
    - Delete log files
    - Move files you don’t need frequently anymore but have to keep for compliance

Pop Quiz

You are storing financial logs that need to be retained for 6 years but that are extremely unlikely to be accessed during that time period. Which storage option will be most cost effective?

A) **Standard-IA (Infrequent Access)**
B) **Glacier Archive**
C) **Intelligent-Tiering**
D) **Standard**

Pop Quiz

You are storing financial logs that need to be retained for 6 years but that are extremely unlikely to be accessed during that time period. Which storage option will be most cost effective?

A) **Standard-IA (Infrequent Access)**

B) **Glacier Archive**

C) **Intelligent-Tiering**

D) **Standard**

Pop Quiz

You are storing objects in S3 and aren't sure which storage class will be best for cost effectiveness since you don't know how often the objects will be accessed. You do know they will be accessed a few times a year, in random patterns. Which storage class should you select?

A) Standard

B) One Zone-IA

C) Intelligent-Tiering

D) Glacier

Pop Quiz

You are storing objects in S3 and aren't sure which storage class will be best for cost effectiveness since you don't know how often the objects will be accessed. You do know they will be accessed a few times a year, in random patterns. Which storage class should you select?

A) **Standard**

B) **One Zone-IA**

C) **Intelligent-Tiering**

D) **Glacier**

AWS Storage Gateway

Connect on-prem applications to cloud storage, caching data locally for low-latency access.

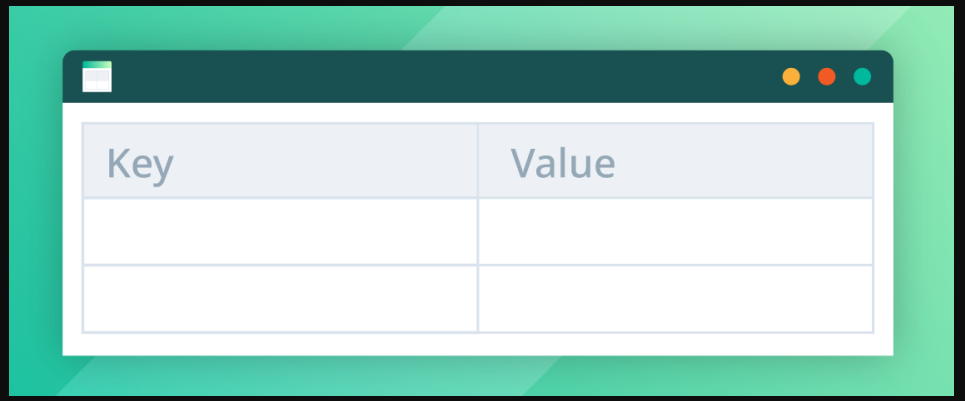
- Tape Gateway
  - Cost-effective, long-term, off-site data archiving
- File Gateway
  - Data uploaded to S3 for object-based workloads
- Volume Gateway
  - Data backed up as point-in-time snapshots of volumes, and stored in the cloud as Amazon EBS snapshots.
- Use cases:
  - Hybrid cloud
  - Backup & restore, Disaster Recovery on AWS
  - Manage capacity fluctuations with storage needs

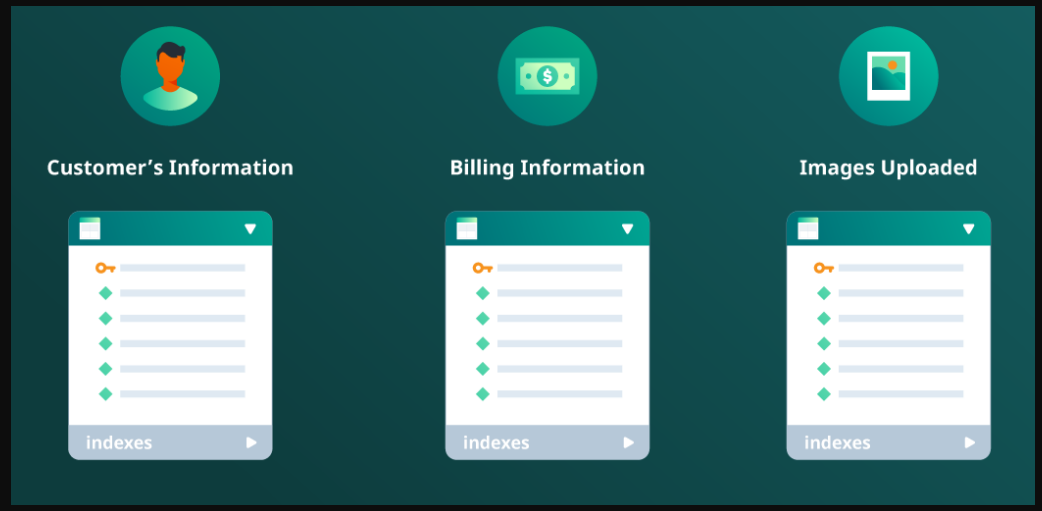
Amazon RDS and DynamoDB

Amazon RDS is a service for Relational Databases (aka SQL)

Amazon DynamoDB is a service for Non-Relational Databases (aka NoSQL)

Article and video on the differences






```
> db.users.find()
{ "_id" : ObjectId("5cd2a65bed7d8cede075fa63"), "name" : "steve", "status" : "active" }
{ "_id" : ObjectId("5cd2a663ed7d8cede075fa64"), "name" : "dave", "status" : "active" }
{ "_id" : ObjectId("5cd2a66eed7d8cede075fa65"), "name" : "kathy", "status" : "active" }
{ "_id" : ObjectId("5cd2a677ed7d8cede075fa66"), "name" : "jill", "status" : "active" }
{ "_id" : ObjectId("5cd2a681ed7d8cede075fa67"), "name" : "bill", "status" : "retired" }
{ "_id" : ObjectId("5cd2a688ed7d8cede075fa68"), "name" : "donna", "status" : "retired" }
```

Linux Academy

Amazon RDS

“Amazon RDS makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while automating time-consuming administration tasks such as hardware provisioning, database setup, patching and backups.”

RDS offers different database engine options:

- Amazon Aurora
- MySQL
- MariaDB
- PostgreSQL
- Oracle
- Microsoft SQL Server

Amazon RDS

- You can select the instance type to use:
  - Memory optimized
  - Performance or I/O optimized

- Main RDS benefits:
  - Scalability with a few clicks, resizable
  - Cost efficient (“1/10th the cost of other commercial DBs”)
  - High availability → easily create read replicas across AZs
  - Apply patches and backups without downtime
  - Less up-front effort → you specify options, they provision it

Amazon Aurora

Amazon Aurora features a distributed, fault-tolerant, self-healing storage system that auto-scales up to 64TB per database instance. It delivers high performance and availability with up to 15 low-latency read replicas, point-in-time recovery, continuous backup to Amazon S3, and replication across three Availability Zones (AZs).

- Benefits of Aurora
    - Fully managed by RDS
    - High scalability and high performance

MySQL and PostgreSQL-compatible for enterprise-scale databases.

Up to 5x faster than MySQL, 3x faster than PostgreSQL

Amazon DynamoDB

DynamoDB is a key-value and document database with very high performance (single-digit millisecond performance). NoSQL database engine, similar to MongoDB, CassandraDB.

- Benefits of DynamoDB
  - Great for unstructured data (NoSQL)
  - Fully managed, multi-region, multi-master
  - Built-in security, backup and restore, and in-memory caching

- Common Use Cases:
  - Mobile, web, gaming, ad tech, IoT

Amazon RDS vs DynamoDB

- RDS
  - For when you need a SQL database option
  - Used when data is clearly defined and structured
  - Easy to set up, highly available, fault tolerant, scalable
  - Common use cases include online stores and banking systems

- DynamoDB
  - For when you need a NoSQL option
  - Used when data is fluid and can change - doesn’t require a schema
  - Fast, highly scalable, fully managed
  - Common use cases include social media, web analytics

AWS-Managed vs Customer-Managed

When it comes to running databases on AWS, unless you need absolute control over how the database is configured and managed, it’s recommended to choose RDS or DynamoDB options as it will make configuration and management much easier and automated. This means you can spend more time on your business solution instead of the technology behind it.

Pop Quiz

Your startup company has limited engineering resources and needs to launch an application as fast as possible in order to monetize. This application will require high availability, fault tolerance, and scalability since it expects a ramp up in traffic shortly after launch. This means your database needs to be resizable, and if the master node fails, it needs to have a read replica as backup. Which of these options would make the most sense for your startup?

A) Oracle database

B) Microsoft SQL Server

C) Database on EC2 and EBS-backed

D) Amazon RDS

Pop Quiz

Your startup company has limited engineering resources and needs to launch an application as fast as possible in order to monetize. This application will require high availability, fault tolerance, and scalability since it expects a ramp up in traffic shortly after launch. This means your database needs to be resizable, and if the master node fails, it needs to have a read replica as backup. Which of these options would make the most sense for your startup?

A) Oracle database

B) Microsoft SQL Server

C) Database on EC2 and EBS-backed

D) Amazon RDS

Pop Quiz

Your organization is running a massive-scale web application that requires an extremely fast and scalable database option. The data to be stored doesn't have a rigid structure and can change and be accessed frequently. Which of these options is likely to be the best fit for your needs?

A) Amazon Aurora

B) Amazon RDS (PostgreSQL or MySQL)

C) Amazon DynamoDB

D) Amazon EC2, EBS-backed

Pop Quiz

Your organization is running a massive-scale web application that requires an extremely fast and scalable database option. The data to be stored doesn't have a rigid structure and can change and be accessed frequently. Which of these options is likely to be the best fit for your needs?

A) Amazon Aurora

B) Amazon RDS (PostgreSQL or MySQL)

C) Amazon DynamoDB

D) Amazon EC2, EBS-backed

AWS Database Migration Service (DMS)

Helps you **migrate databases to AWS (and within AWS).** Once the migration has started, DMS manages all the complexities of the migration process including **automatically replicating data changes** that occur in the source database during the migration process. You can also use this service for **continuous data replication** with the same simplicity.

- Main DMS benefits:
  - No downtime
  - Continuous data replication
  - Supports migrating to Aurora, Redshift, DynamoDB, S3
  - Supports migration between AWS services as well

Amazon ElastiCache

Caching services can be used to **help improve speed and performance of web applications**.

ElastiCache offers fully managed **Redis** and **Memcached** options, which lets you deploy, run, and scale these **in-memory data stores** with little effort.

Downside is it’s short-term storage and ephemeral, and not meant as a database engine.

Amazon CloudFront

CloudFront is a **content delivery network (CDN)** that securely delivers data, videos, applications, and APIs to customers on a global scale and with low latency and high transfer speeds.

CloudFront is directly integrated in the AWS global infrastructure which provides **low latency**.

- It also works directly with services like:
  - AWS Shield for DDoS mitigation
  - Amazon S3 to serve data with less latency
  - Elastic Load Balancing & Amazon EC2 to serve applications
  - Lambda@Edge to run custom code closer to customers' users
  - Route53

Amazon CloudFront

- Common Use Cases:
  - Video delivery
  - Static asset caching (JavaScript, CSS, HTML files, images)
  - API acceleration and DDoS mitigation

Amazon Redshift

Redshift is a **data warehouse** service designed to handle petabytes of data for analysis. You can use standard SQL and existing Business Intelligence (BI) tools to analyze all of your data from a data lake.

- Benefits:
  - Run complex queries against massive amounts of data using sophisticated query optimization and massive parallel query execution
  - You can do this without having to move the data or transform it into a set schema.
  - You can use Redshift to query data from your **data lake** alone, or also in combination with your data warehouse.
- Common Use Cases:
  - Business intelligence
  - Predictive analytics (machine learning)
  - Real-time streaming analytics

Pop Quiz

Your engineering team is in charge of running a video streaming service to deliver content to your customers, but they've started receiving complaints that certain parts of the world have longer buffering time than other parts of the world. Which of these services would best help solve this issue?

A) Amazon ElastiCache

B) Amazon Aurora

C) Amazon RDS

D) Amazon CloudFront

Pop Quiz

Your engineering team is in charge of running a video streaming service to deliver content to your customers, but they've started receiving complaints that certain parts of the world have longer buffering time than other parts of the world. Which of these services would best help solve this issue?

A) Amazon ElastiCache

B) Amazon Aurora

C) Amazon RDS

D) Amazon CloudFront

Pop Quiz

Your engineers have created a large data storage pipeline to send structured and unstructured data to Amazon S3. Now, they need to run complex queries to analyze and make sense of this data. What should they use?

A) Amazon Redshift

B) Amazon Aurora

C) Amazon DynamoDB

D) Amazon RDS

Pop Quiz

Your engineers have created a large data storage pipeline to send structured and unstructured data to Amazon S3. Now, they need to run complex queries to analyze and make sense of this data. What should they use?

A) Amazon Redshift

B) Amazon Aurora

C) Amazon DynamoDB

D) Amazon RDS

Homework

Review key terms & additional information

- ✓ **Download presentation slides and review**
- ✓ **Download sample AWS exam questions and answer:**
  - ✓ **#2**
- ✓ **Practice with these Hands-On Labs:**
  - ✓ **Creating a Basic Amazon S3 Lifecycle Policy**
  - ✓ **Creating Amazon S3 buckets, managing objects, and enabling versioning**

Read up on the following:

- ✓ **CloudTrail -** https://aws.amazon.com/cloudtrail/
- ✓ **CloudWatch -** https://aws.amazon.com/cloudwatch/
- ✓ **SNS -** https://aws.amazon.com/sns/
- ✓ **Lambda -** https://aws.amazon.com/lambda/

Additional Info

Final meeting – September 25th:

CloudTrail, CloudWatch, and more

Links / Resources / Info
https://github.com/Ellopunk/Cloud_Practitioner

Linux Academy Community
https://linuxacademy.com/join/community

Linux Academy Slack Channel
https://linuxacademy-community-slack.herokuapp.com/

Thank you

If you found this helpful, please invite your colleagues!

Linux Academy