

Certified Solutions Architect Professional

Introduction



A CLOUD GURU

So You Want To Be a Pro...

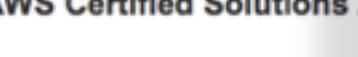


A CLOUD GURU

[Menu](#)  Contact Sales Products Solutions More ▾ English ▾ My Account ▾ Create an AWS Account

AWS Certified Solutions Architect - Professional

Schedule an exam

Training and Certification	Training Overview	AWS Certification	Recertification	Partner Training	FAQs
<p>The AWS Certified Solutions Architect – Professional exam validates advanced technical experience in designing distributed applications and systems on the AWS platform. Exam concepts you should understand for this exam include:</p> <ul style="list-style-type: none"> Designing and deploying dynamically scalable, highly available, fault-tolerant, and reliable applications on AWS Selecting appropriate AWS services to design and deploy an application based on given requirements Migrating complex, multi-tier applications on AWS Designing and deploying enterprise-wide scalable operations on AWS Implementing cost-control strategies 	<p> Training and Certification</p> <p>Your company's on-premises content management system has the following architecture: <ul style="list-style-type: none"> Application Tier – Java code on a JBoss application server Database Tier – Oracle database regularly backed up to Amazon Simple Storage Service (Amazon S3) using the Oracle RMAN backup utility Static Content – stored on a 512GB gateway stored Storage Gateway volume attached to an Oracle database server via the iSCSI interface </p> <p>Which AWS based disaster recovery strategy will give you the best RTO?</p> <ol style="list-style-type: none"> Deploy the Oracle database and the JBoss app server on EC2. Restore the Oracle database from backups stored in Amazon S3. Generate an EBS volume of static content from the Storage Gateway and attach it to the JBoss EC2 server. Deploy the Oracle database on RDS. Deploy the JBoss app server on EC2. Restore the Oracle database from backups from Amazon Glacier. Generate an EBS volume of static content from the Storage Gateway and attach it to the JBoss EC2 server. Deploy the Oracle database and the JBoss app server on EC2. Restore the Oracle database from backups stored in Amazon S3. Restore the static content by attaching an AWS Storage Gateway volume as an iSCSI volume to the JBoss EC2 server. Deploy the Oracle database and the JBoss app server on EC2. Restore the Oracle database from backups stored in Amazon S3. Restore the static content from an AWS Storage Gateway-VTL and attach it to the JBoss EC2 server. <p>An ERP application is deployed in multiple Availability Zones in a single region. The customer's RTO must be less than 3 hours, and the RPO is 15 minutes. The customer realized a failure occurred roughly 1.5 hours ago. Which DR strategy can be used to achieve this event of this kind of failure?</p> <ol style="list-style-type: none"> Take 15-minute DB backups stored in Amazon Glacier, with transaction logs every 5 minutes. Use synchronous database master-slave replication between two Availability Zones. Take hourly DB backups to Amazon S3, with transaction logs stored in S3 every hour. Take hourly DB backups to an Amazon EC2 instance store volume, with transaction logs stored in Amazon S3 every 5 minutes. 	<p> AWS Certified Solutions Architect</p>			



AWS Certified Solutions Architect–Professional (SAP-C01) Exam Guide

Introduction

The AWS Certified Solutions Architect–Professional (SAP-C01) exam is intended for individuals who perform a Solutions Architect–Professional role. This exam validates advanced technical skills and experience in designing distributed applications and systems on the AWS platform.

It validates an examinee's ability to:

- Design and deploy dynamically scalable, highly available, fault-tolerant, and reliable applications on AWS
- Select appropriate AWS services to design and deploy an application based on given requirements
- Migrate complex, multi-tier applications on AWS
- Design and deploy enterprise-wide scalable operations on AWS
- Implement cost-control strategies

Recommended AWS and General IT Knowledge and Experience

- Two or more years of hands-on experience designing and deploying cloud architecture on AWS
- Ability to evaluate cloud application requirements and make architectural recommendations for implementation, deployment, and provisioning applications on AWS
- Ability to provide best practice guidance on the architectural design across multiple applications and projects of the enterprise
- Familiarity with a scripting language
- Familiarity with Windows and Linux environments
- Familiarity with AWS CLI, AWS APIs, AWS CloudFormation templates, the AWS Billing Console, and the AWS Management Console
- Explain and apply the five pillars of the AWS Well-Architected Framework
- Map business objectives to application/architecture requirements
- Design a hybrid architecture using key AWS technologies (e.g., VPN, AWS Direct Connect)
- Architect a continuous integration and deployment process

Exam Preparation

These training courses and materials may be helpful for examination preparation:

AWS Training: (aws.amazon.com/training)

- Advanced Architecting on AWS: 3-day instructor-led live or virtual course
- AWS Certification Exam Readiness Workshops: AWS Certified Solutions Architect–Professional: 1-day live course
- AWS Security Fundamentals: 1-day instructor-led live course
- AWS Well-Architected Training: 2-hour online training course
- EC2 Systems Manager: 2-hour online training course
- Migrating to AWS: 2-day instructor-led live or virtual course
- Preview Course: Deep Dive into Amazon Elastic Block Store (EBS): 1-hour online training course
- Preview Course: Deep Dive into Elastic File System (EFS): 65-minute online training course
- Preview Course: Migrating and Tiering Storage to AWS: 1-hour online training course

Suggested AWS Whitepapers (aws.amazon.com/whitepapers) Kindle and .pdf, and Other Materials

- AWS Security Best Practices whitepaper, August 2016

AWS Bad-ass-ery Levels



A CLOUD GURU



AWS Bad-ass-ery Levels



"S3 is an object store with 99.9999999% data durability."

"S3 can be secured with a bucket policy and/or IAM rules applied to that bucket."

"When heavy I/O is required, storage accessed natively via the OS is preferred. In these cases, instance store, EBS or EFS make better choices (in that order) than S3 because of its consistency model, lack of object locking and API-only access."

AWS Bad-ass-ery Levels



Chuck Norris'
fist as a baby. →



Disclaimer



WARNING

Viewing this course will not guarantee
that you will pass the exam.

Disclaimer

There is no fast track.

There are no shortcuts.

There is no “exam hack”.

There is only hard work and preparation.

You will be rewarded.



Holistic Preparation Plan

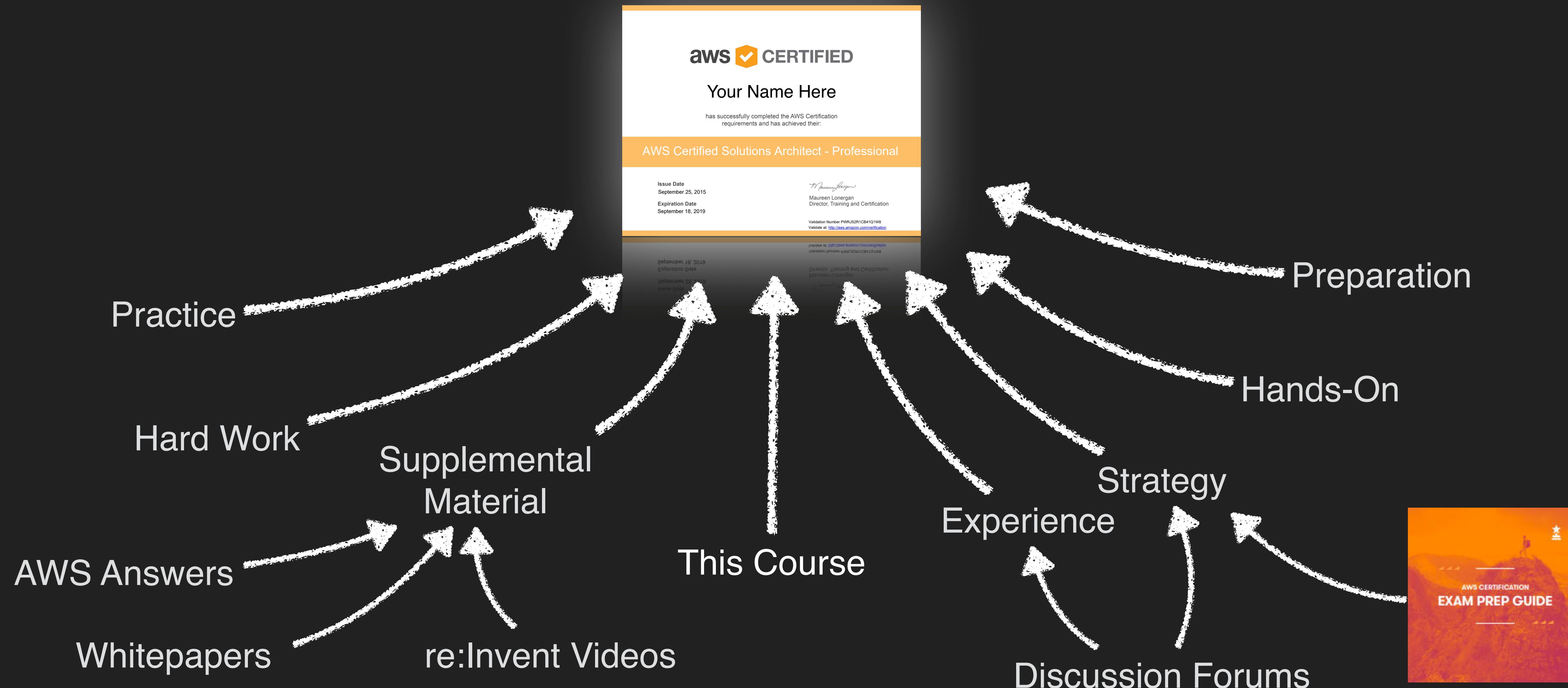


This Course

Holistic Preparation Plan



A CLOUD GURU



A dark, low-light photograph of a person from the chest up. They are wearing a light-colored button-down shirt over a dark t-shirt. Their hands are clasped in front of them, holding a thick, dark book or folder. The background is dark and out of focus.

Introduction

About The Exam

Exam Prerequisites



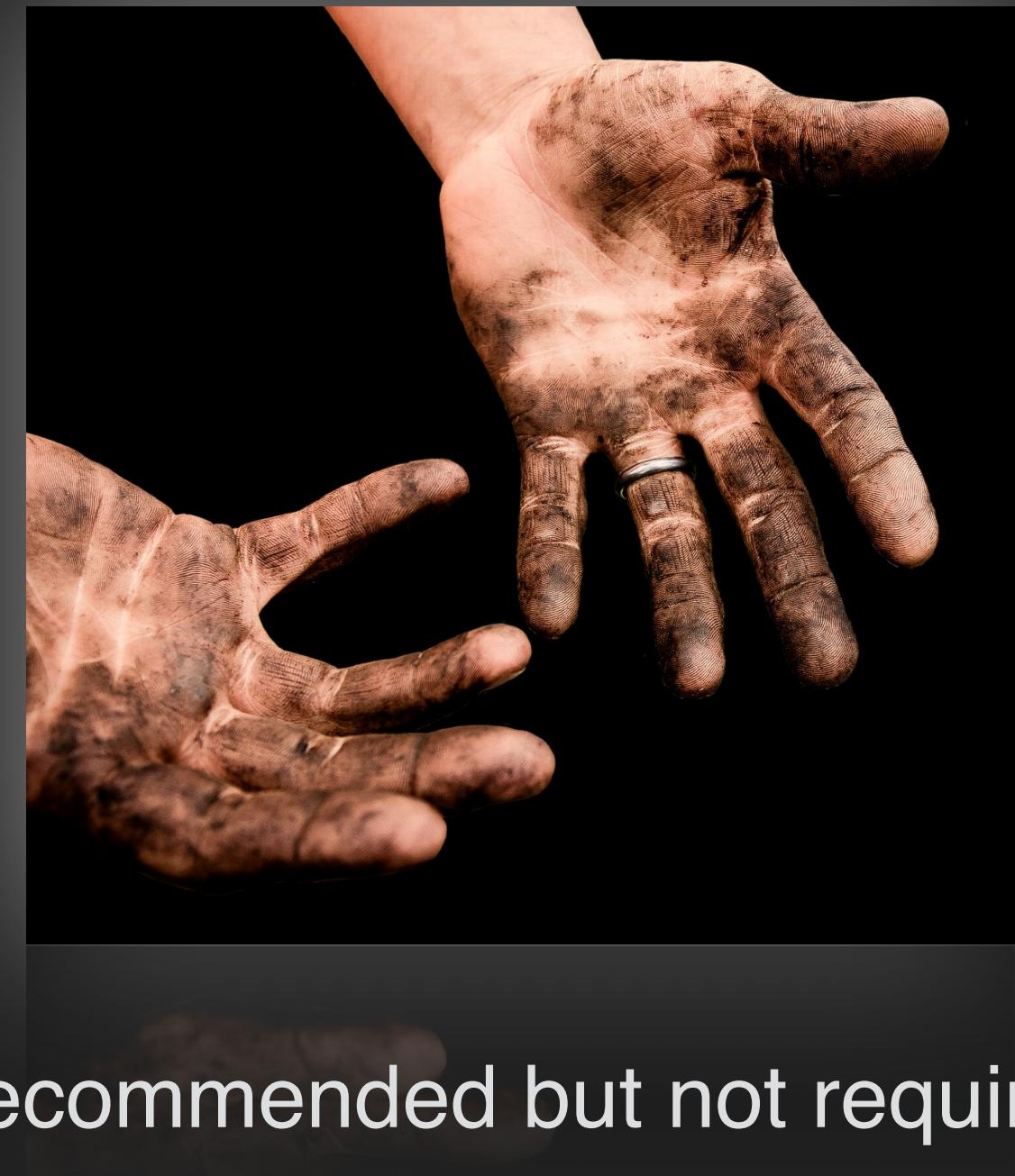


A CLOUD GURU

Exam Prerequisites



Recommended but not required



Recommended but not required

Certified Solutions Architect Pro Exam



- 170 minutes, ~74 questions
- Multiple-choice (pick best answer) and multiple-answer (pick 3 of 8)
- No partial credit for questions
- You can mark questions and go back to them within the time limit.
- No points for unanswered questions
- Score between 100 and 1000 with minimum passing score of 750. Scaled scoring models are used.

Exam Blueprint (Update October 2018)



A CLOUD GURU

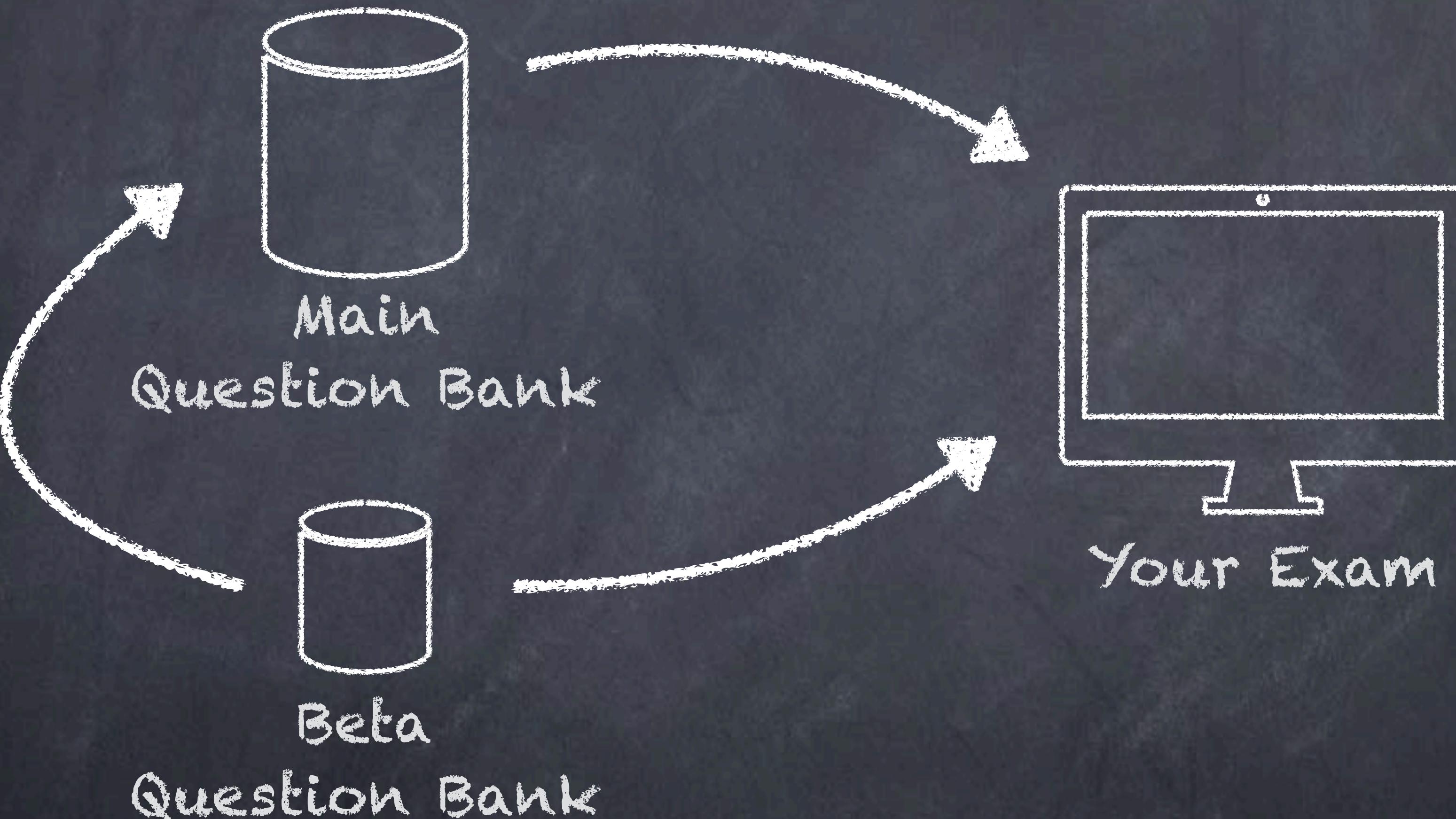
Domain	% of Examination
1. Design for Organizational Complexity	12.5%
2. Design for New Solutions	31%
3. Migration Planning	15%
4. Cost Control	12.5%
5. Continuous Improvement for Existing Solutions	29%
TOTAL	100%

100%	100%
21 COURSES IN 100+ SOLUTIONS	21 COURSES IN 100+ SOLUTIONS

Every Exam is Unique



A CLOUD GURU



Certified Solutions Architect Pro Exam



- A new product, service, or feature must be generally available (GA) for six months before appearing on a certification exam.
- Special accommodations can be requested once an exam is scheduled
 - but it must be reviewed and might require supporting documentation.
 - Extra time for English-as-a-Second-Language
 - Medical or physical needs such as screen magnification, screen reader, trackball, etc. are available.



A CLOUD GURU

Be Careful!

Select one or more responses that best complete the statement or answer the question. Distractors, or incorrect answers, are response options that an examinee with incomplete knowledge or skill would likely choose. However, they are generally plausible responses that fit in the content area defined by the test objective.

they are generally plausible responses that fit in the content area defined by the test objective.



A CLOUD GURU

Be Careful!

Select one or more responses that best complete the statement or answer the question. Distractors, or incorrect answers, are response options that an examinee with incomplete knowledge or skill would likely choose. However, they are generally plausible responses that fit in the content area defined by the test objective.

they are generally plausible responses that fit in the content area defined by the test objective.

Translation: AWS will try to trick you to expose gaps in your knowledge!

A dark, low-light photograph of a person from the chest up. The person is wearing a dark button-down shirt over a white collared shirt and a dark tie. Their hands are clasped together in their lap. They appear to be sitting at a desk, with a dark surface visible in the lower foreground.

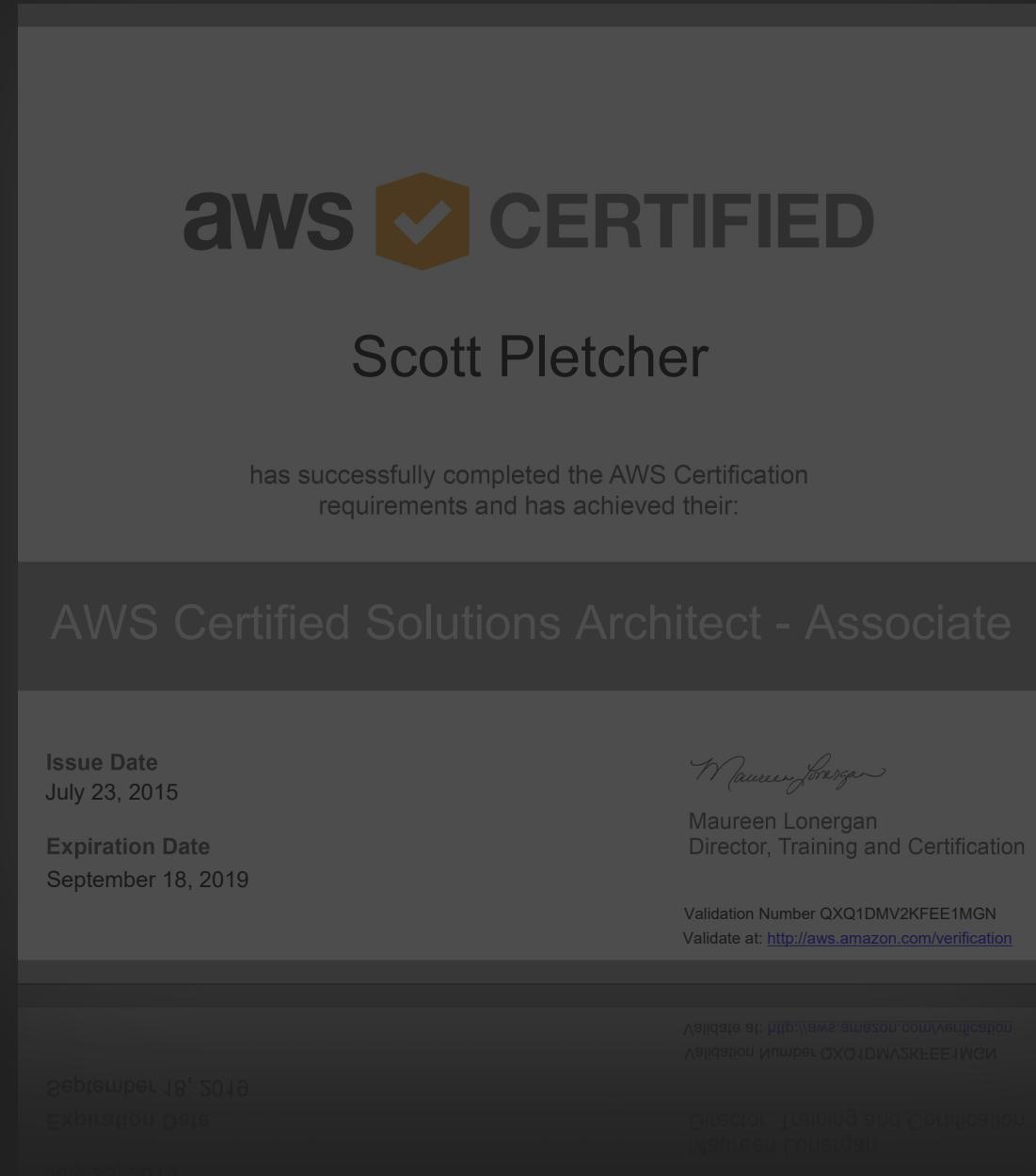
Introduction

About This Course

Certification Prerequisites



1



2



3



This course assumes you are an experienced practitioner!

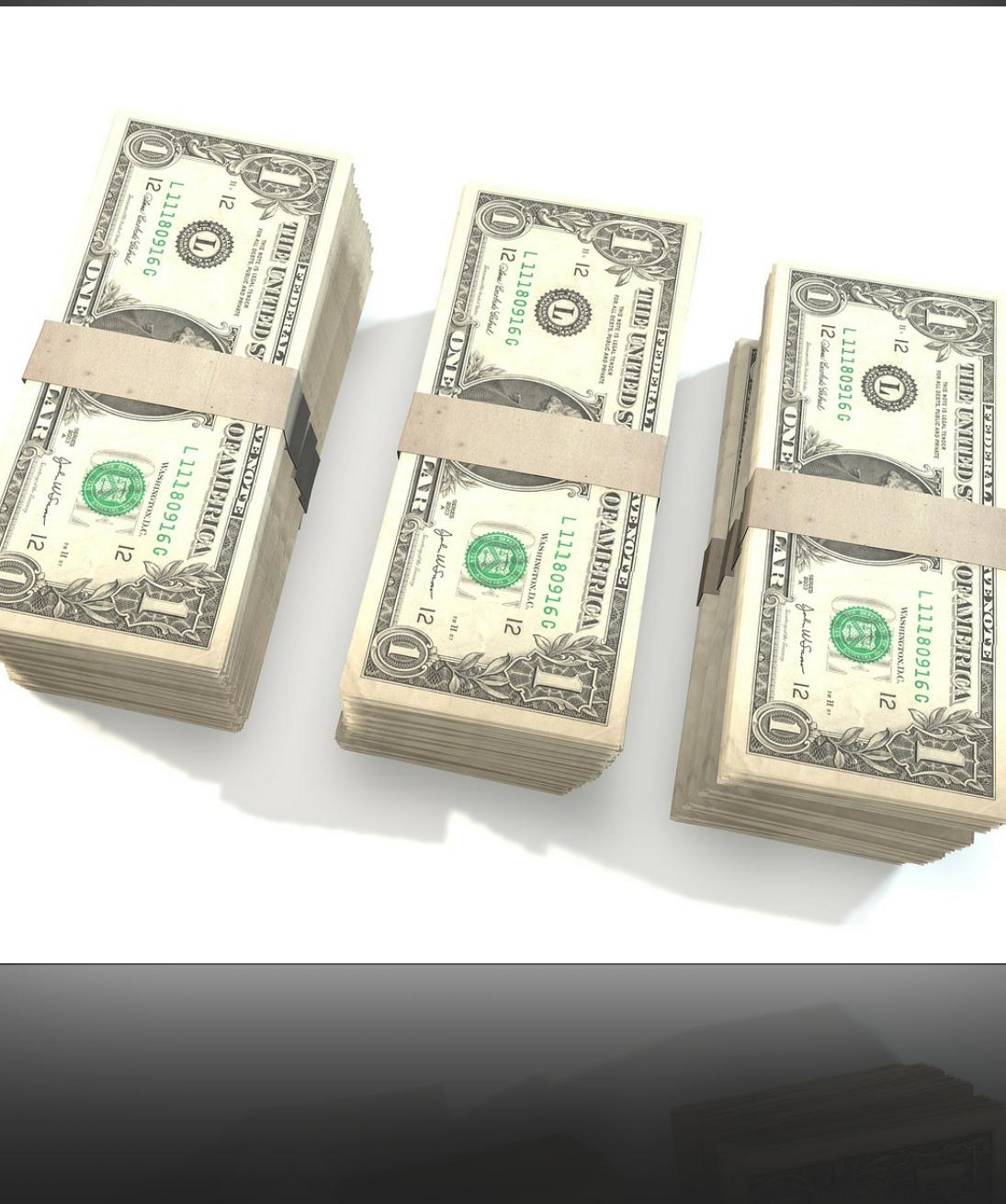
Certification Prerequisites



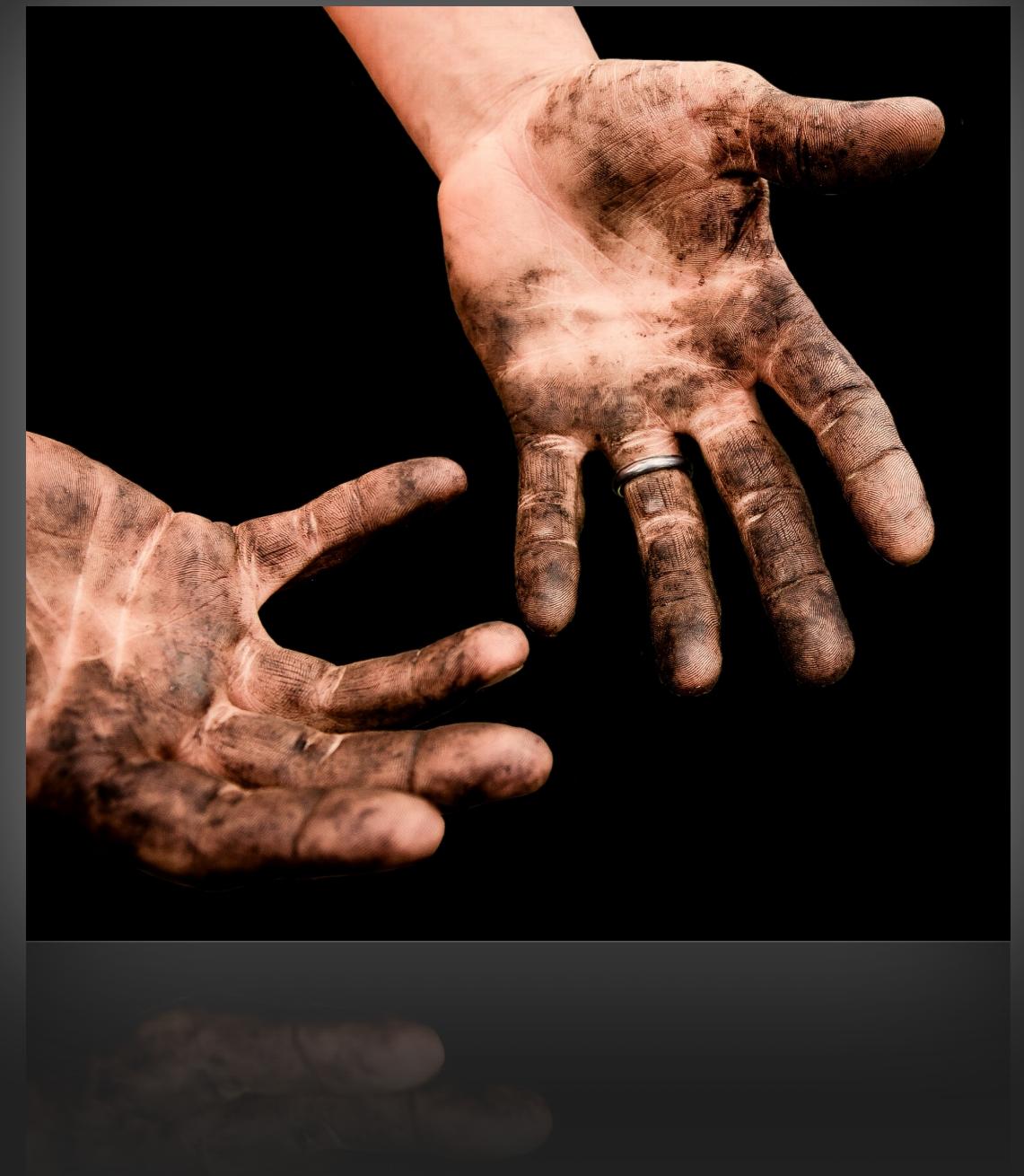
1



2



3



This course assumes you are an experienced practitioner!

Blooms Taxonomy



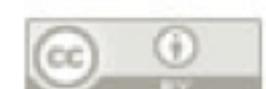
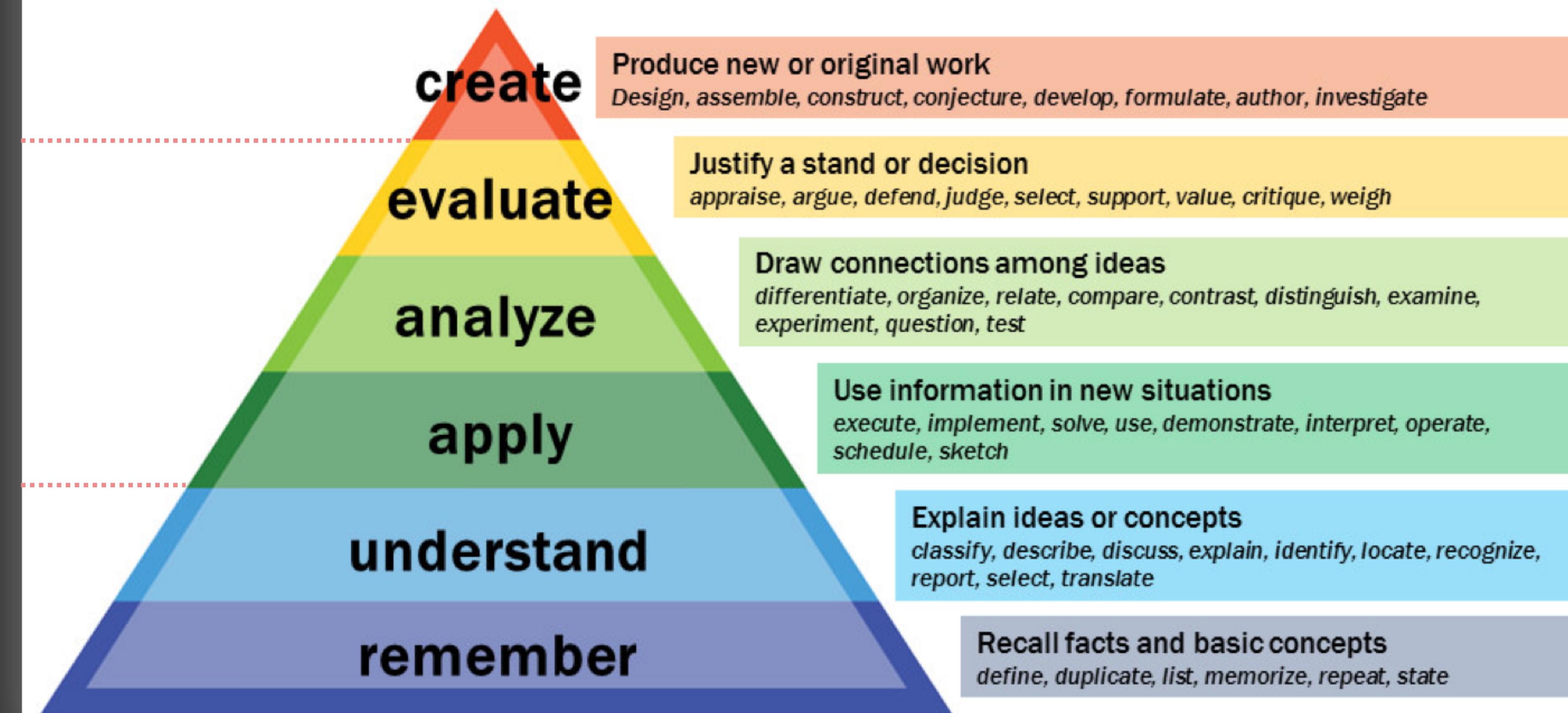
A CLOUD GURU

AWS
Objective

Professional
Exams

Associate
Exams

Bloom's Taxonomy



Vanderbilt University Center for Teaching

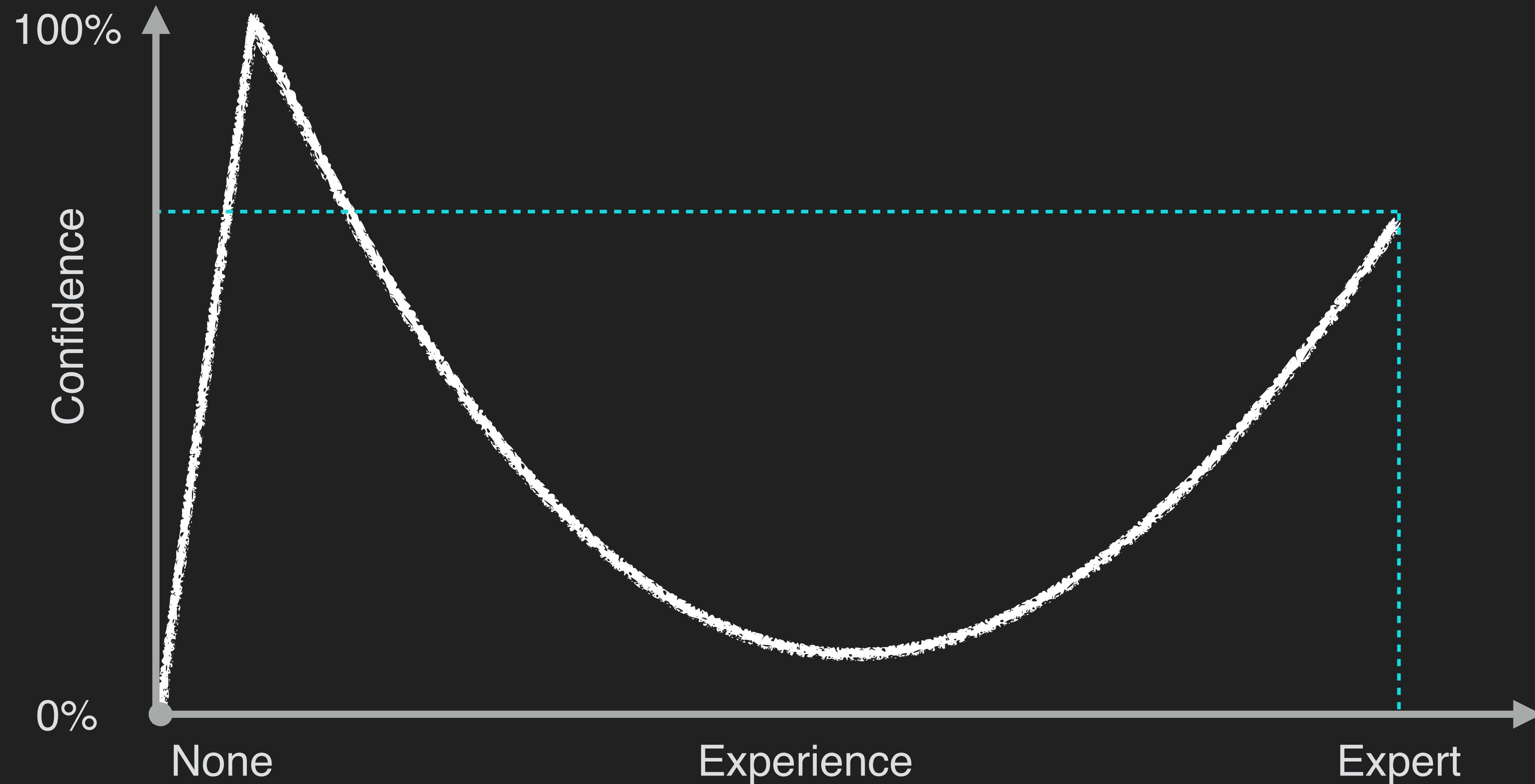


Vanderbilt University Center for Teaching

Dunning-Kruger Effect



A CLOUD GURU

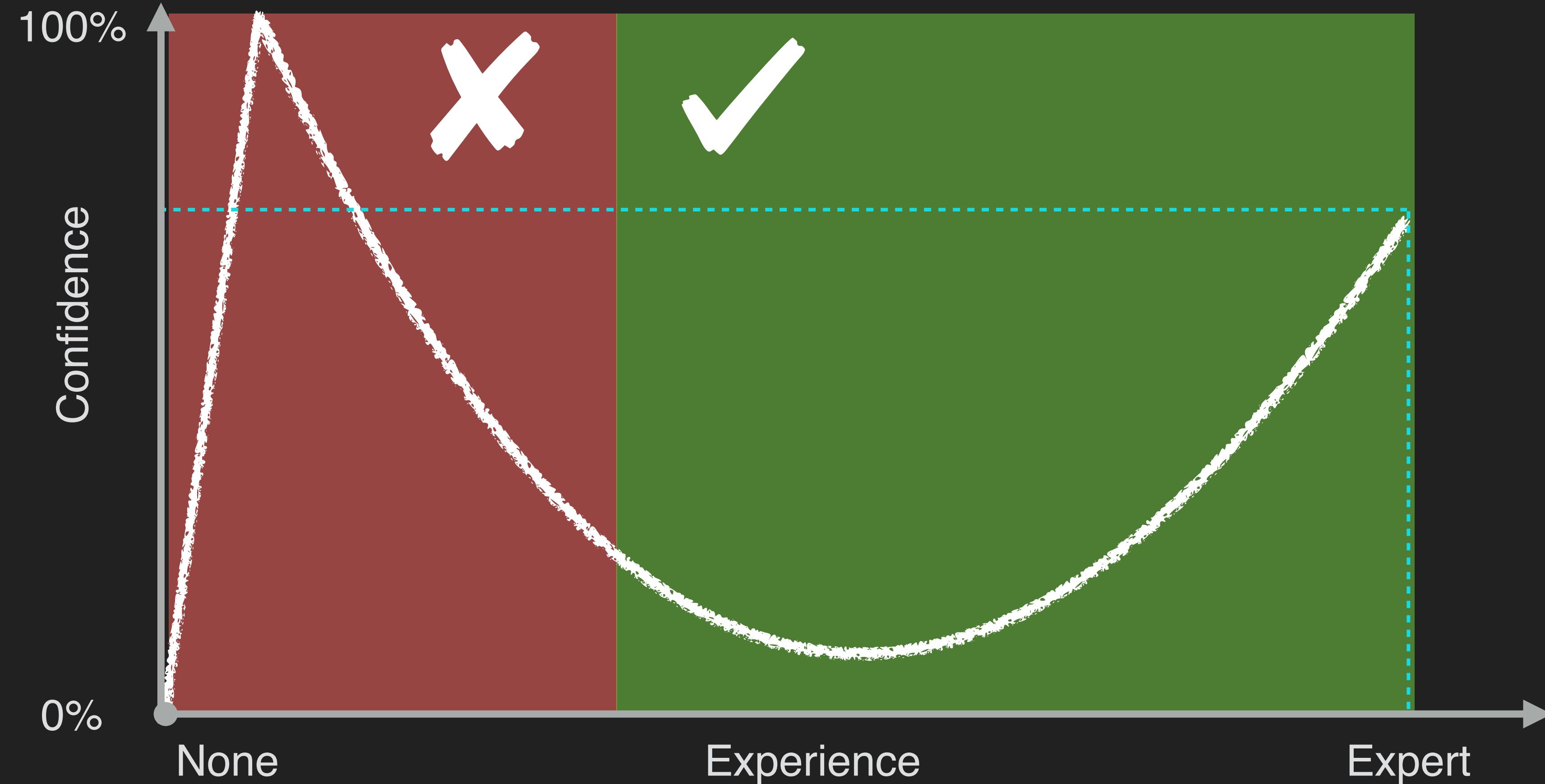


"Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments",
Kruger, J. & Dunning, D., Journal of Personality and Social Psychology, Vol 77(6), Dec 1999, 1121-1134.

Dunning-Kruger Effect



A CLOUD GURU

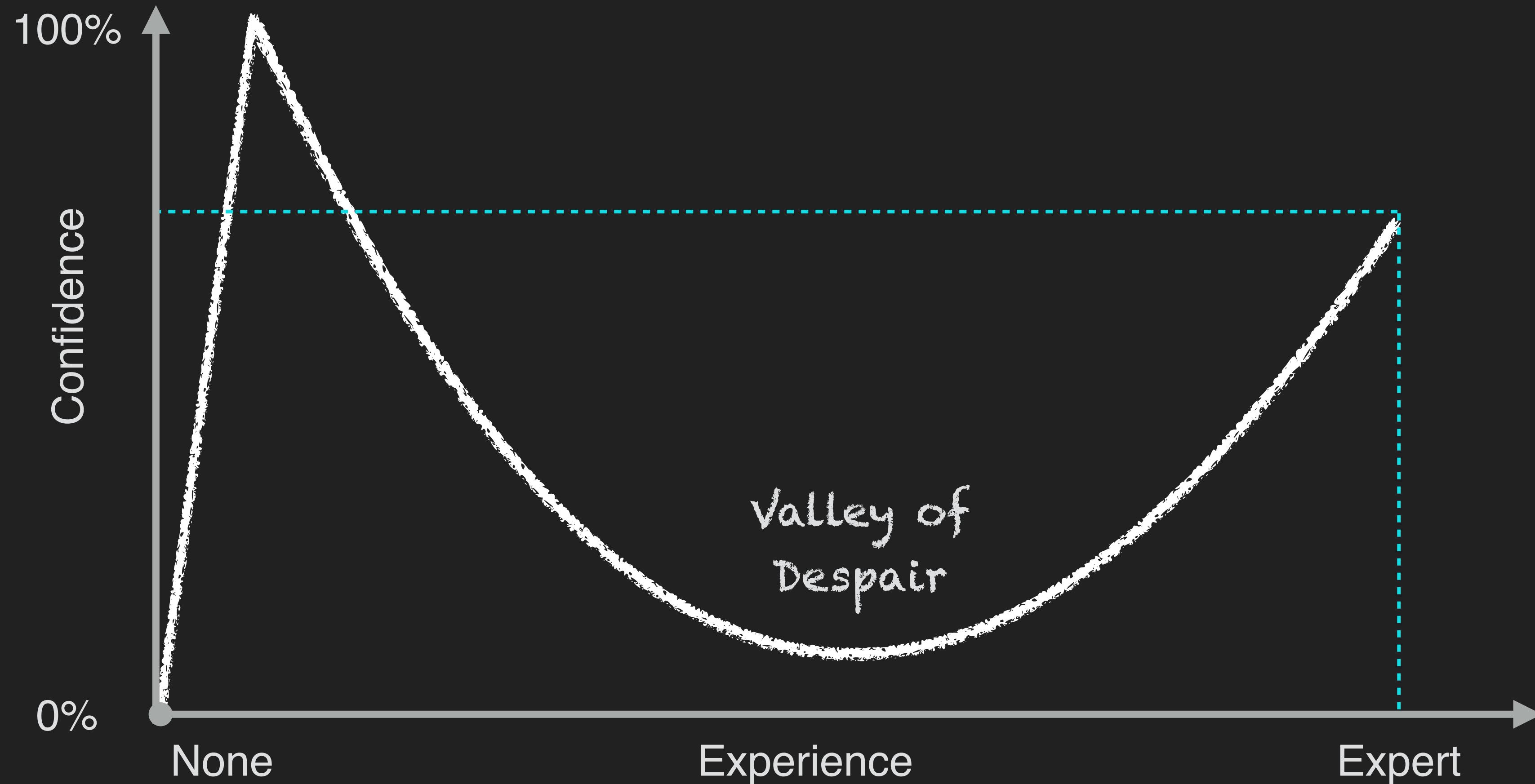


"Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments",
Kruger, J. & Dunning, D., Journal of Personality and Social Psychology, Vol 77(6), Dec 1999, 1121-1134.

Dunning-Kruger Effect



A CLOUD GURU



"Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments",
Kruger, J. & Dunning, D., Journal of Personality and Social Psychology, Vol 77(6), Dec 1999, 1121-1134.

Exam Aspects



A CLOUD GURU



Knowledge
of AWS



Contextual
Reasoning



Time
Management

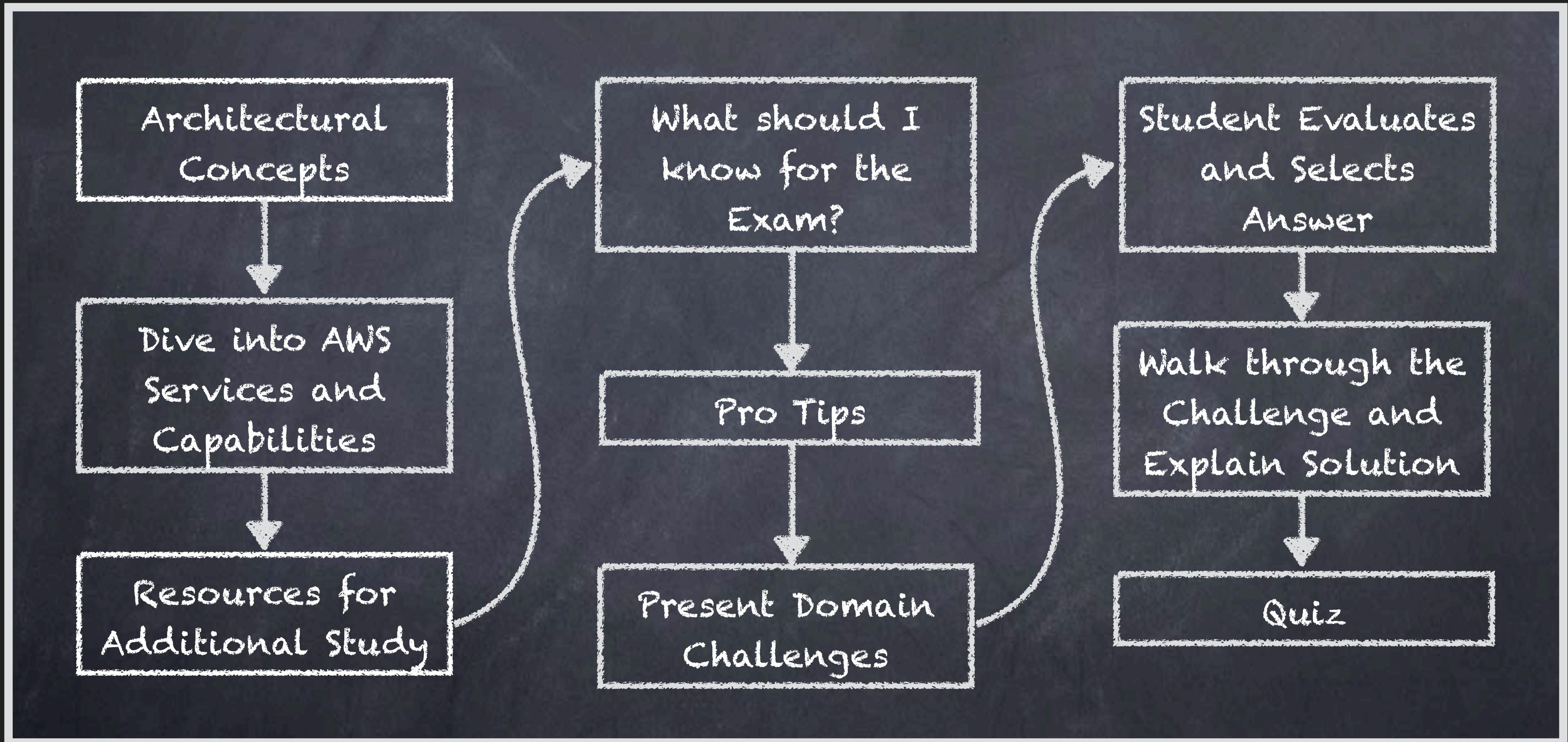


Comprehension Skills



Coping Under Pressure

Chapter Format

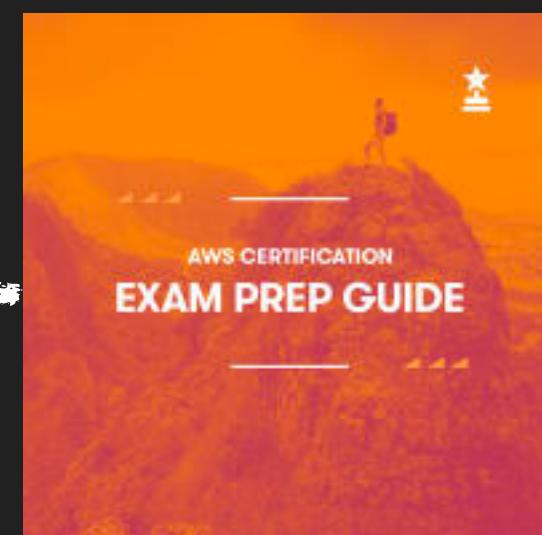
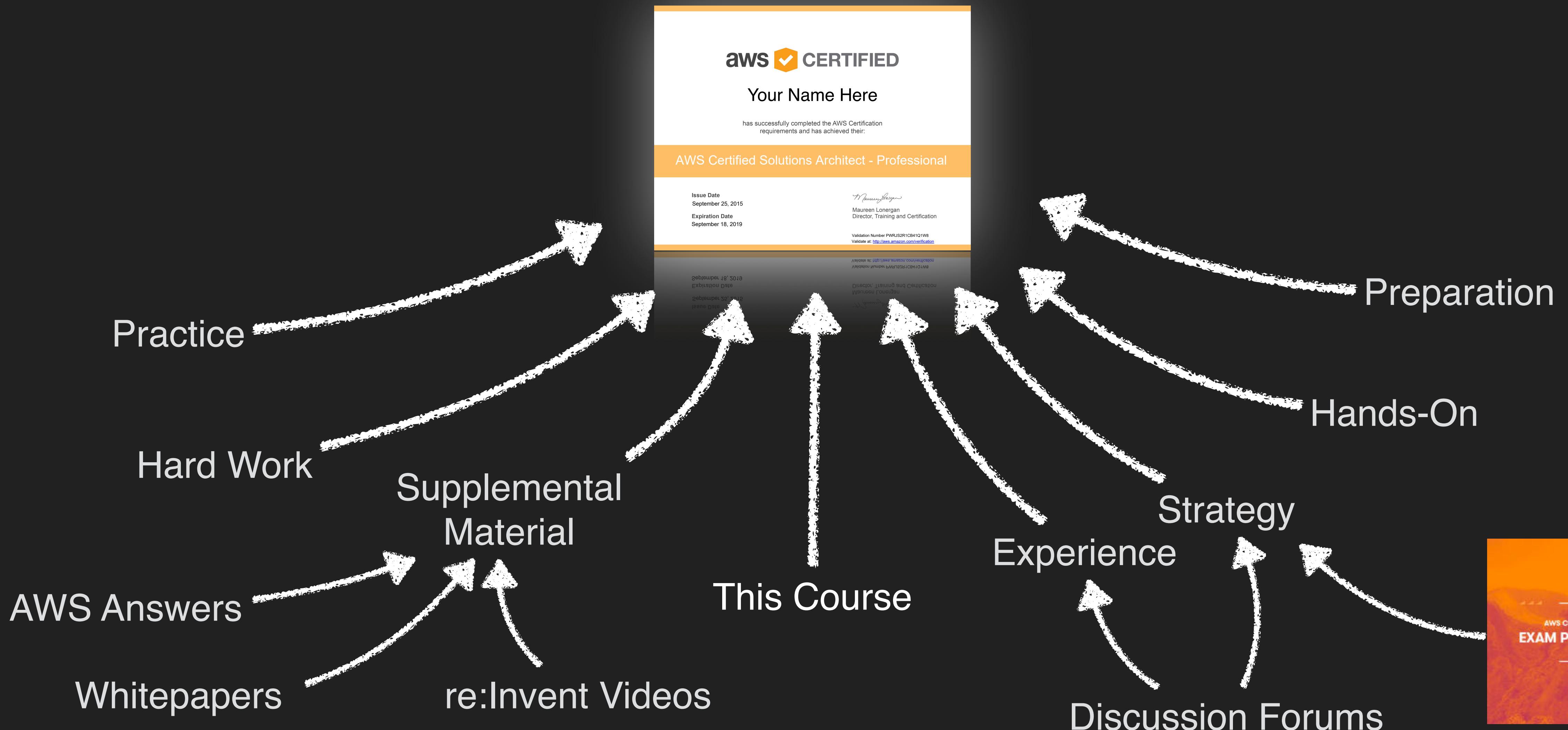


Do This, Not That.



Do This	Not That
Focus on service capabilities and what problems they solve.	Memorize service limits, prices, version numbers.
Study to become an Expert.	Study to pass the Exam.
Make it your own.	Expect learning through osmosis.
Review the Supplemental Material.	Ignore the Supplemental Material.

Holistic Preparation Plan



AWS Certification Preparation Guide



A CLOUD GURU

AWS Certification Preparation Guide

The must-have course for anyone
serious about passing the AWS
certification exams

certification exams

A dark, out-of-focus photograph of a person from the chest up. They are wearing a dark, button-down shirt over a light-colored, possibly white, t-shirt. Their hands are clasped together in front of them. The background is dark and indistinct.

Introduction

About the Labs

Geocaching



A CLOUD GURU



No Easy Way



A CLOUD GURU



Here Be Dragons!



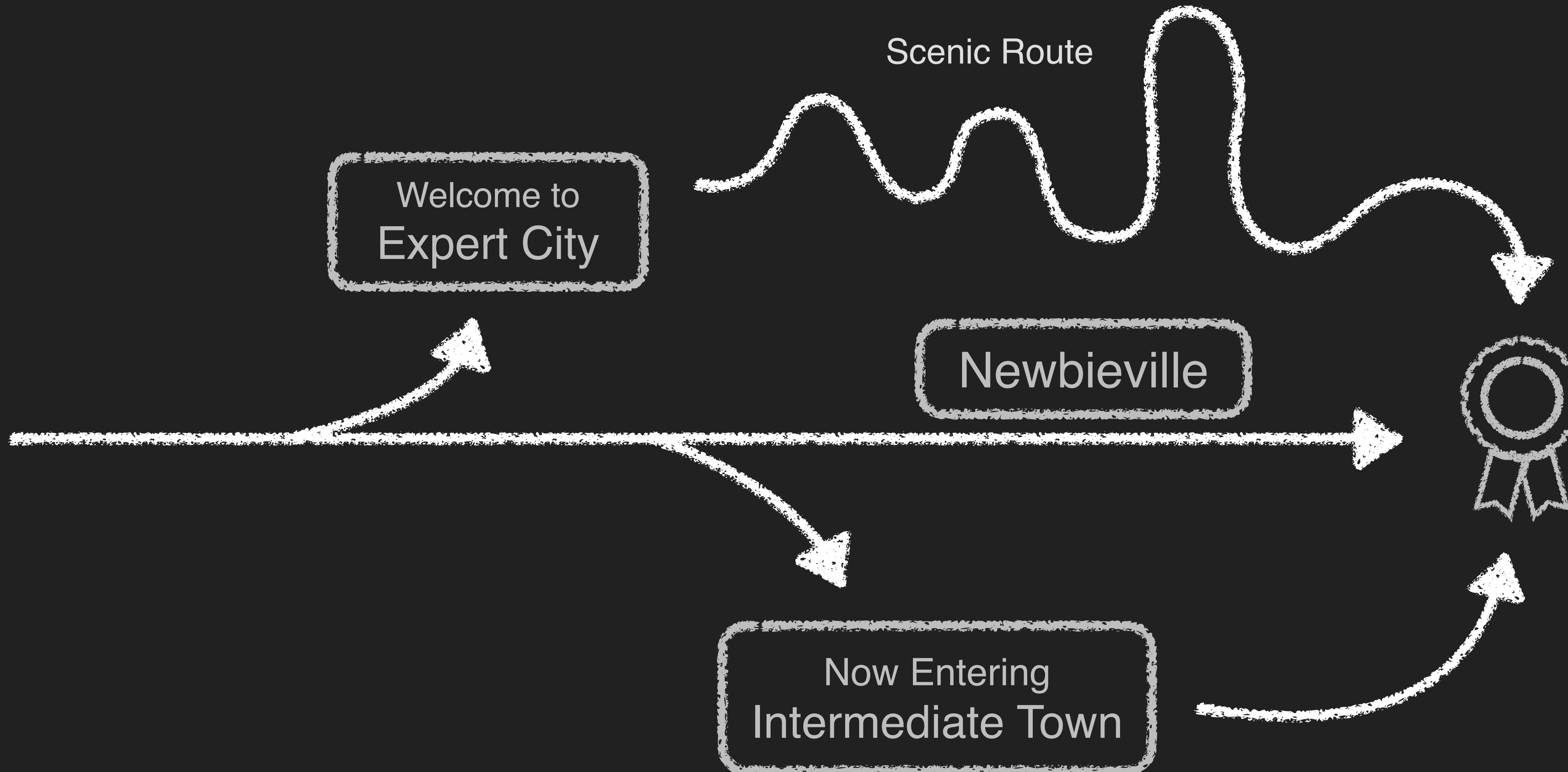
A CLOUD GURU



You Choose Your Own Adventure!



A CLOUD GURU



Data Stores

Introduction



A CLOUD GURU



Exam Blueprint

Domain 1: Design for Organizational Complexity

- 1.1. Determine cross-account authentication and access strategy for complex organizations (for example, an organization with varying compliance requirements, multiple business units, and varying scalability requirements).
- 1.2. Determine how to design networks for complex organizations (for example, an organization with varying compliance requirements, multiple business units, and varying scalability requirements).
- 1.3. Determine how to design a multi-account AWS environment for complex organizations (for example, an organization with varying compliance requirements, multiple business units, and varying scalability requirements).

Domain 2: Design for New Solutions

- 2.1. Determine security requirements and controls when designing and implementing a solution.
- 2.2. Determine a solution design and implementation strategy to meet reliability requirements.
- 2.3. Determine a solution design to ensure business continuity.
- 2.4. Determine a solution design to meet performance objectives.
- 2.5. Determine a deployment strategy to meet business requirements when designing and implementing a solution.

olutions.

- 2.6. Determine a deployment strategy to meet business requirements when designing and implementing a solution to support a new service or application.



Exam Blueprint

Domain 3: Migration Planning

- 3.1. Select existing workloads and processes for potential migration to the cloud.
- 3.2. Select migration tools and/or services for new and migrated solutions based on detailed AWS knowledge.
- 3.3. Determine a new cloud architecture for an existing solution.
- 3.4. Determine a strategy for migrating existing on-premises workloads to the cloud.

Domain 4: Cost Control

- 4.1. Select a cost-effective pricing model for a solution.
- 4.2. Determine which controls to design and implement that will ensure cost optimization.
- 4.3. Identify opportunities to reduce cost in an existing solution.

Domain 5: Continuous Improvement for Existing Solutions

- 5.1. Troubleshoot solution architectures.
- 5.2. Determine a strategy to improve an existing solution for operational excellence.
- 5.3. Determine a strategy to improve the reliability of an existing solution.
- 5.4. Determine a strategy to improve the performance of an existing solution.
- 5.5. Determine a strategy to improve the security of an existing solution.
- 5.6. Determine how to improve the deployment of an existing solution.

A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a smartphone. The phone's screen is visible but mostly black, suggesting it is off or in sleep mode.

Data Stores

Concepts

Data Persistence



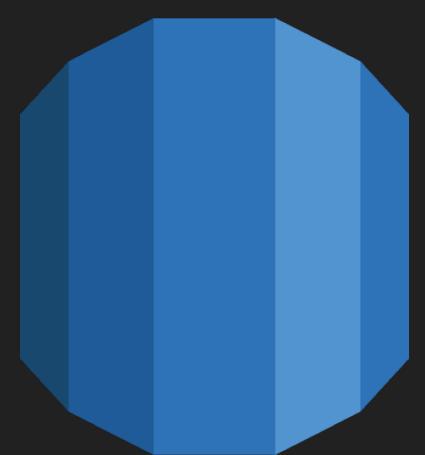
A CLOUD GURU

Persistent Data Store

Data is durable and sticks around after reboots, restarts, or power cycles.



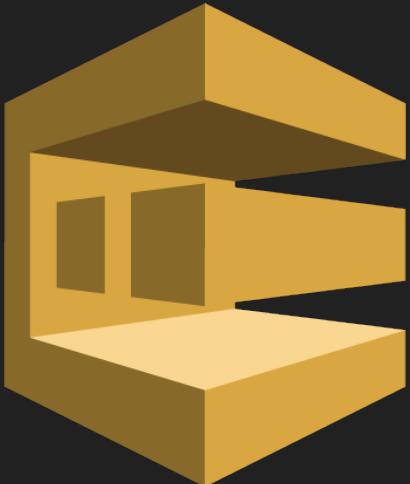
Glacier



RDS

Transient Data Store

Data is just temporarily stored and passed along to another process or persistent store.



SQS



SNS

Ephemeral Data Store

Data is lost when stopped.



EC2 Instance



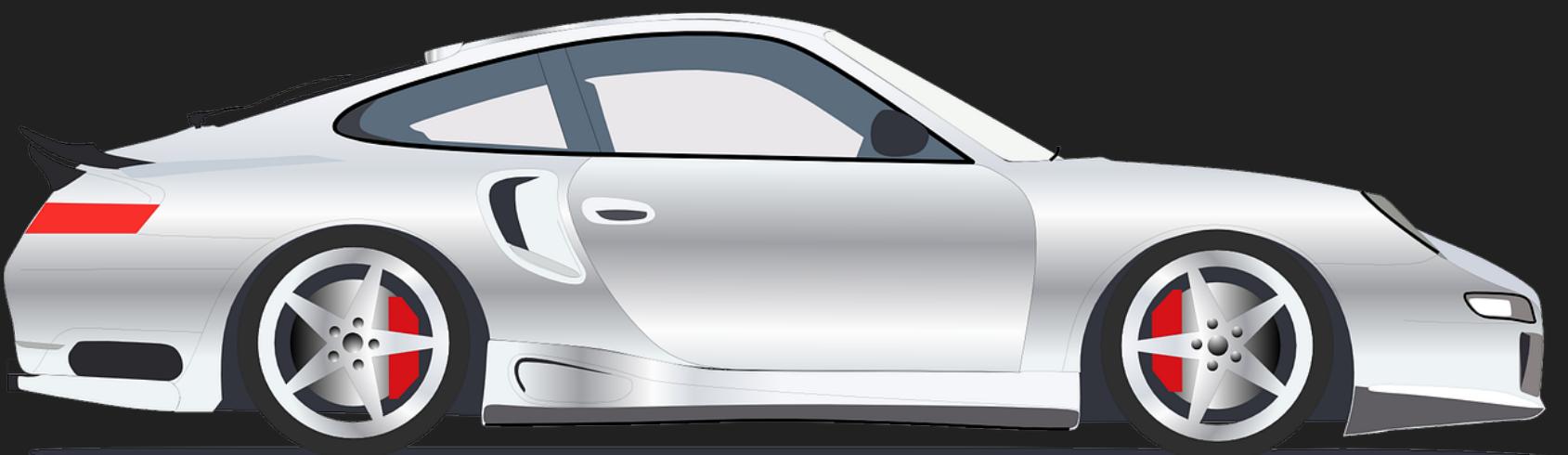
Memcached Store



IOPS vs. Throughput

Input/Output Operations per Second (IOPS)

Measure of how fast
we can read and write
to a device



Throughput

Measure of how much
data can be moved at
a time



Consistency



A CLOUD GURU

“Consistency is far better
than rare moments of
greatness.”

– Scott Ginsberg

Consistency Models - ACID & BASE



A CLOUD GURU

ACID

Atomic - Transactions are “all or nothing”.

Consistent - Transactions must be valid.

Isolated - Transactions can’t mess with one another.

Durable - Completed transaction must stick around.

BASE

Basic Availability - values availability even if stale

Soft-state - might not be instantly consistent across stores

Eventual Consistency - will achieve consistency at some point

A dark, low-light photograph of a person from the chest up. They are wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a small, bright screen of a smartphone, which is the focal point of the image. The background is dark and indistinct.

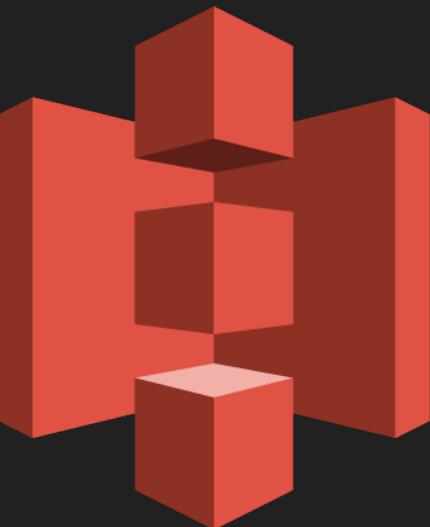
Amazon S3

Amazon S3



A CLOUD GURU

- One of the first AWS service introduced back in 2006
- S3 is an Object Store.
- Used in other AWS services — directly and behind-the-scenes
- Maximum object size is 5TB; largest object in a single PUT is 5GB.
- Recommended to use multi-part uploads if larger than 100MB



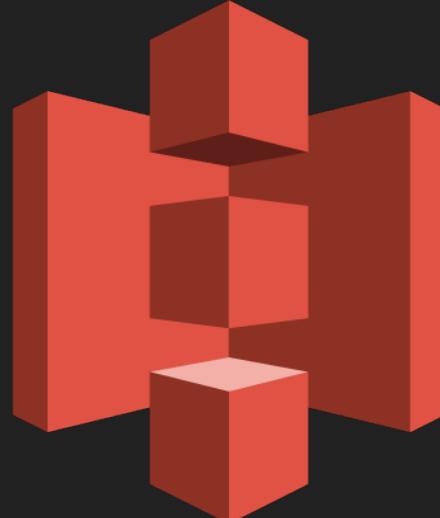
Amazon
S3

S3 Object Store



A CLOUD GURU

s3://bucket/finance/april/16/invoice_45675.pdf = a **KEY**, not a file path



Amazon
S3

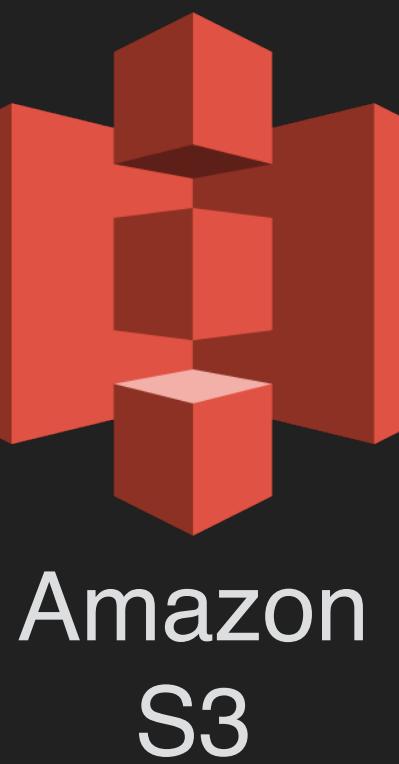
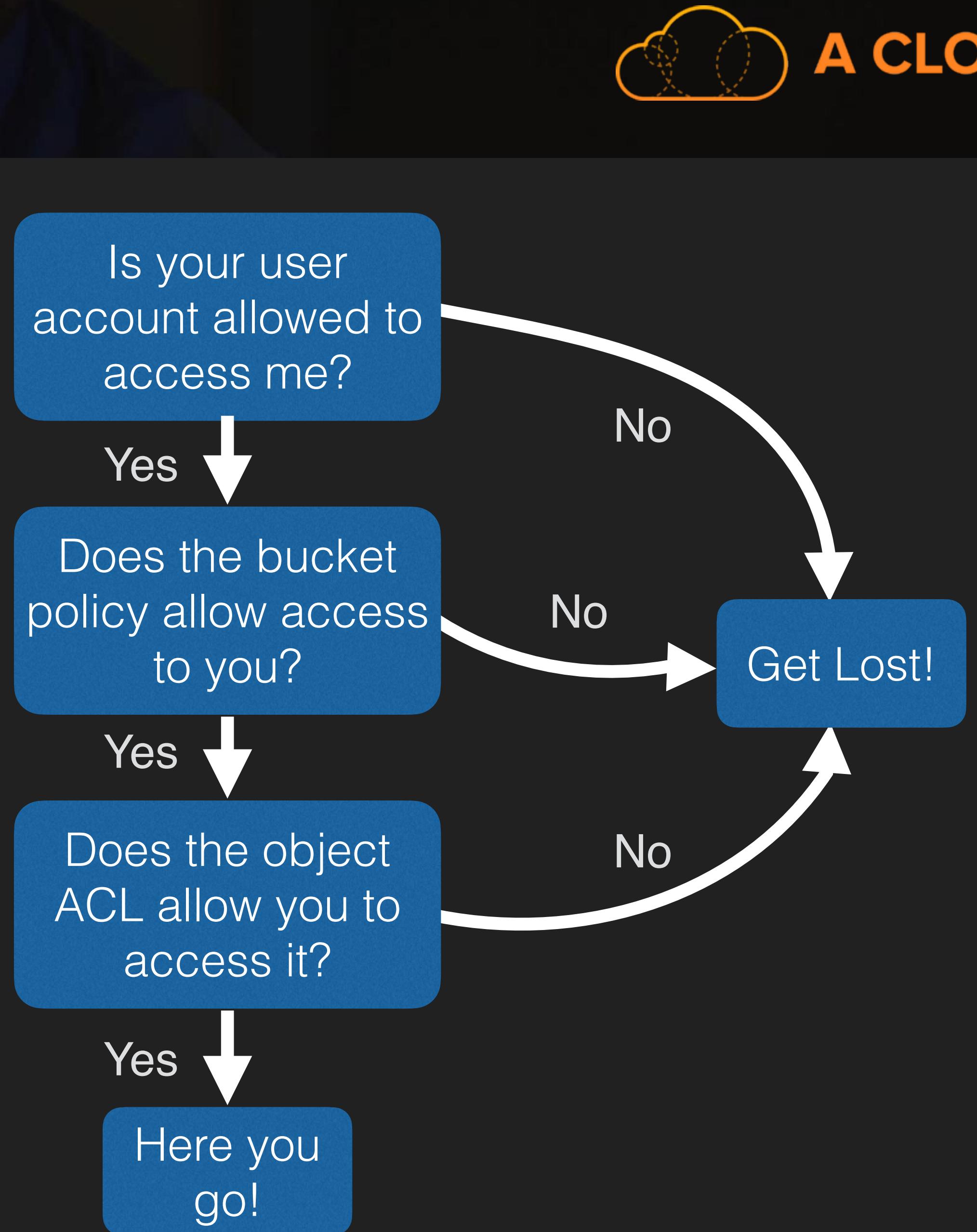


S3 Consistency

AWS Documentation Statement	What S3 Is Thinking
“S3 provides read-after-write consistency for PUTs of new objects.”	Cool, I’ve never seen this object, and no-one has asked about it before. Welcome aboard, new object — and you can read it immediately.
“HEAD or GET requests of the key before an object exists will result in eventual consistency.”	Wait a second, someone already asked about this key, and I told them, “never saw it”. I remember that, and need to honor that response until I completely write this new object and fully replicate it. So, I’ll let you read it eventually.
“S3 offers eventual consistency for overwrite PUTs and DELETEs.”	Ok, so you want to update or delete an object. Let’s make sure we get that update or delete completed locally, then we can replicate it to other places. Until then, I have to serve up the current file. I’ll serve up the update/delete once its fully replicated - eventually.
“Updates to a single key are atomic.”	Whoa, there. Only one person can update this object at a time. If I get two requests, I’ll process them in order of their timestamp and you’ll see the updates as soon as I replicate elsewhere.

S3 Security

- Resource-based (Object ACL, Bucket Policy)
- User-based (IAM policies)
- Optional Multi-factor Authentication before Delete



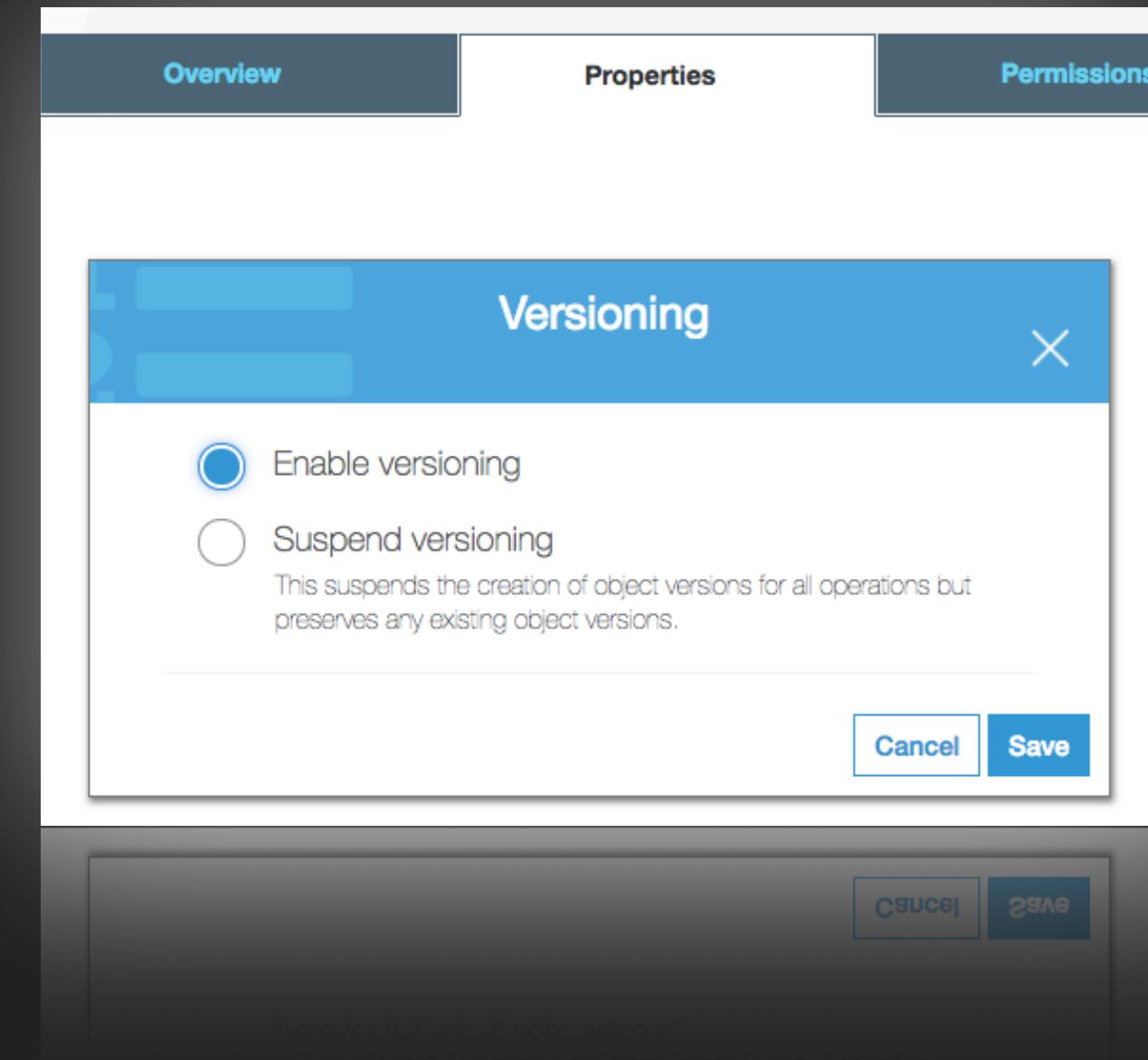
S3 Data Protection



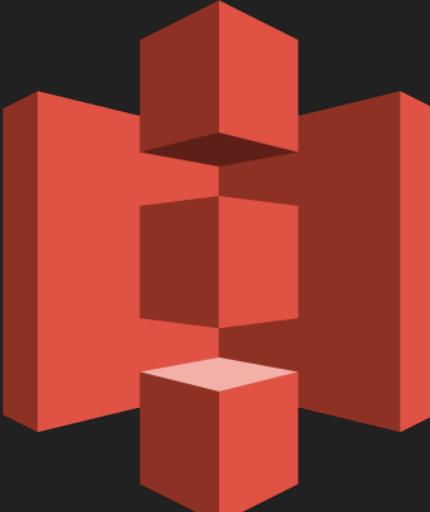
A CLOUD GURU

Versioning:

- New version with each write
- Enables “roll-back” and “un-delete” capabilities
- Old versions count as billable size until they are permanently deleted.
- Integrated with Lifecycle Management



secure.zip	Jun 6, 2018 10:51:40 PM
<input type="checkbox"/> Jun 6, 2018 10:51:40 PM (Delete marker)	M7d_muZZ8A2BpwR64yoxUc...
<input type="checkbox"/> Jun 6, 2018 10:40:31 PM	SeBOeoVH2AqV47zGW4ZUU...
<input type="checkbox"/> M9 RE:0P:0R 8f05 9 nUL	...UUD2W2T4PVASHV09QD9S



Amazon
S3

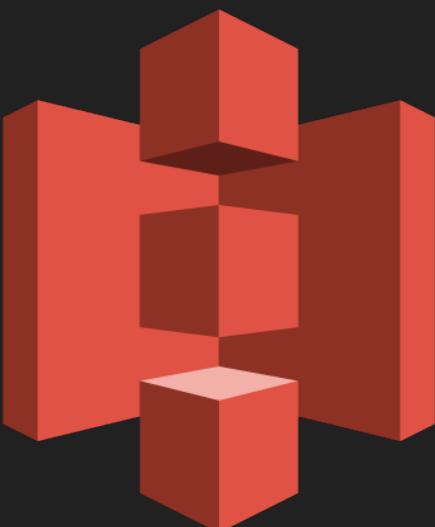
S3 Data Protection



A CLOUD GURU

Optionally require Multi-factor Authentication:

- Safeguard against accidental deletion of an object
- Change the versioning state of your bucket

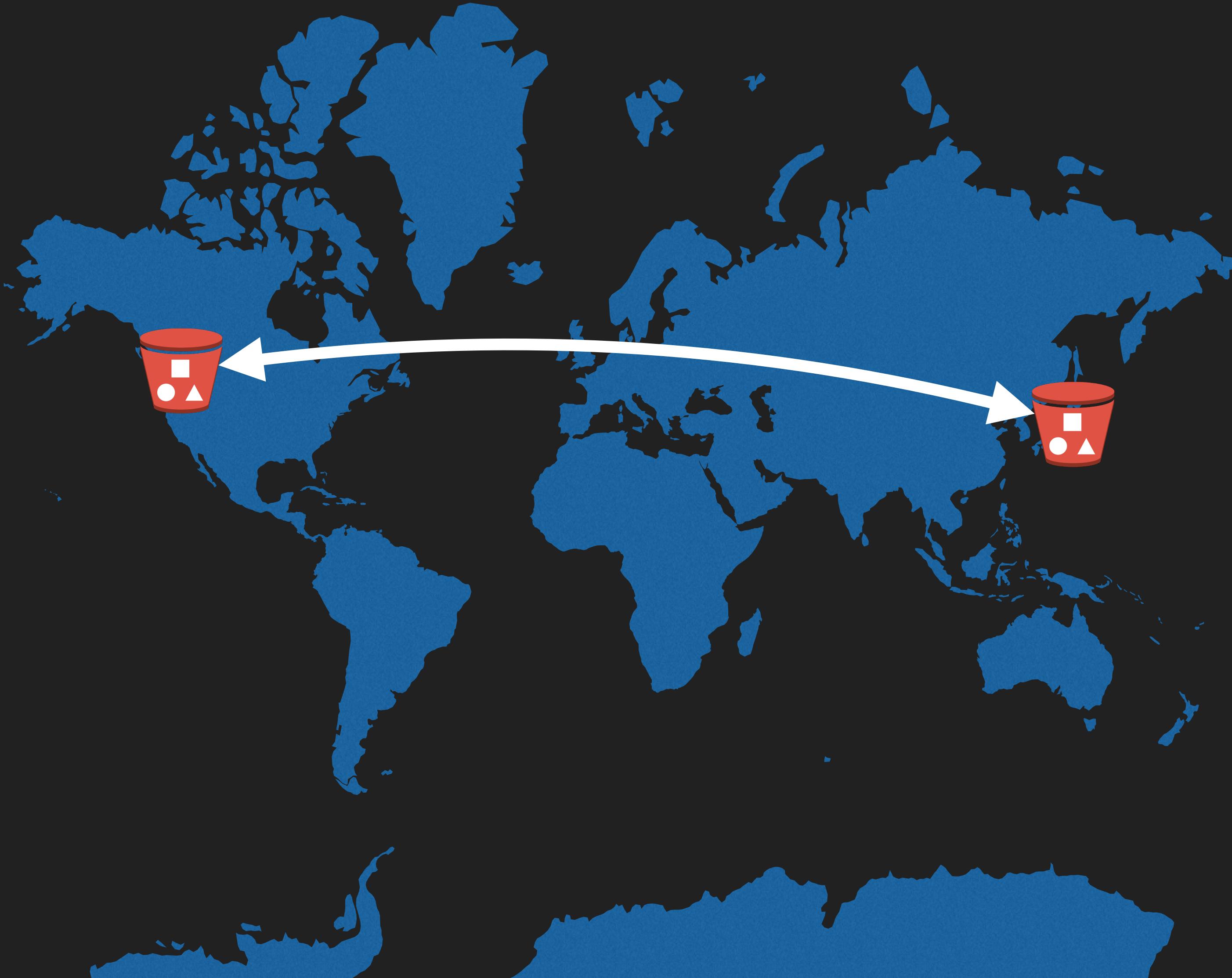


Amazon
S3

S3 Data Protection

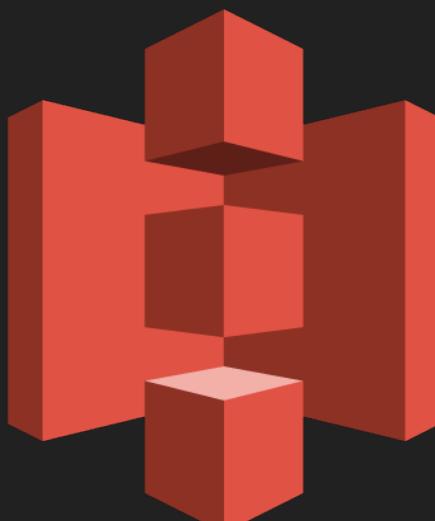


A CLOUD GURU



Cross-Region Replication

- Security
- Compliance
- Latency



Amazon
S3

S3 Lifecycle Management



A CLOUD GURU

- Optimize storage costs
- Adhere to data retention policies
- Keep S3 volumes well-maintained

Lifecycle rule

① Name and scope ② Transitions ③ Expiration ④ Review

Configure transition i

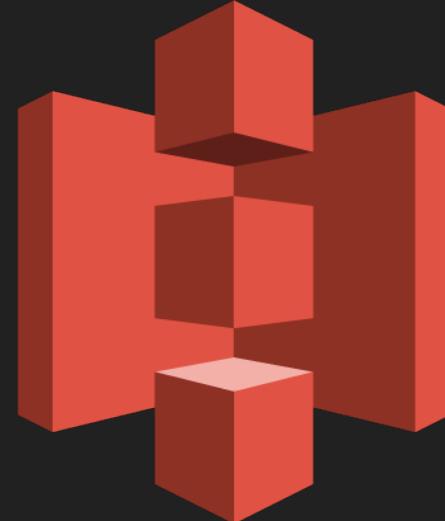
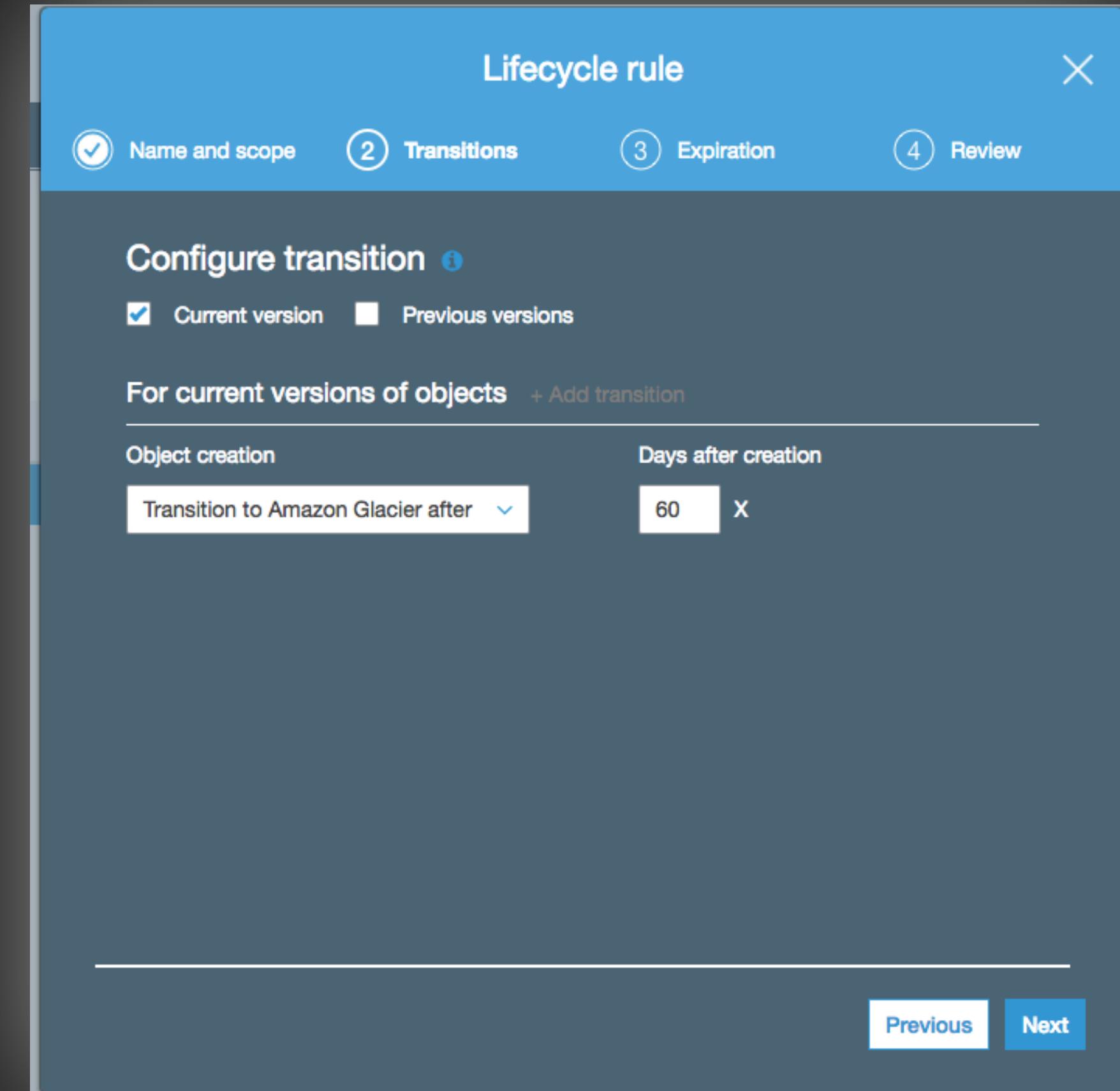
Current version Previous versions

For current versions of objects + Add transition

Object creation Days after creation

Transition to Amazon Glacier after ▼ 60 X

Previous Next



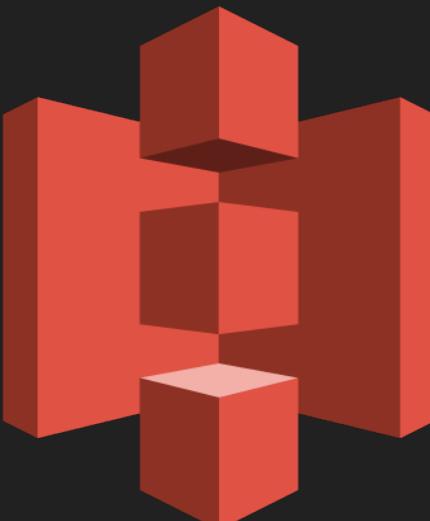
Amazon
S3

S3 Analytics



A CLOUD GURU

Data Lake Concept	Athena, Redshift Spectrum, QuickSight
IoT Streaming Data Repository	Kinesis Firehose
Machine Learning and AI Storage	Rekognition, Lex, MXNet,
Storage Class Analysis	S3 Management Analytics



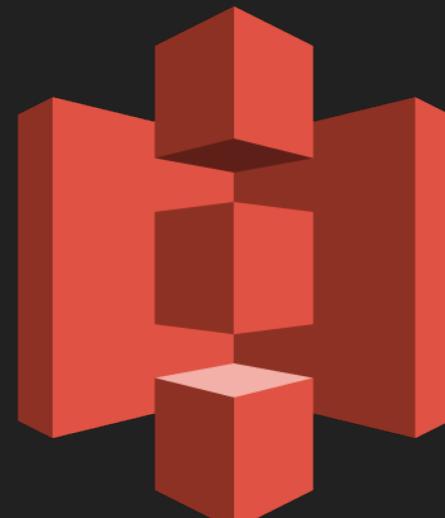
Amazon
S3

S3 Encryption at Rest



A CLOUD GURU

Encryption Option	Meaning
SSE-S3	Use S3's existing encryption key for AES-256
SSE-C	Upload your own AES-256 encryption key which S3 will use when it writes the objects
SSE-KMS	Use a key generated and managed by AWS Key Management Service
Client-Side	Encrypt objects using your own local encryption process before uploading to S3 (i.e. PGP, GPG, etc.)

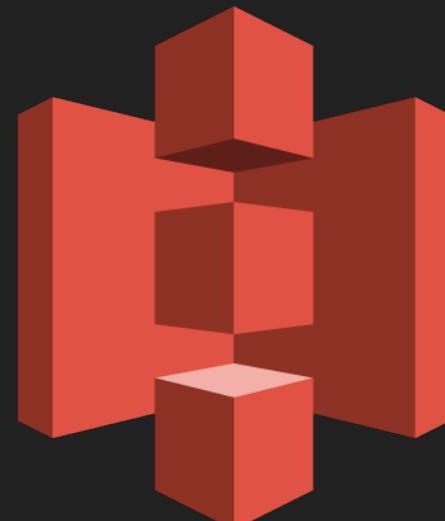


Amazon
S3



More Nifty S3 Tricks

Transfer Acceleration	Speed up data uploads using CloudFront in reverse
Requester Pays	The requester rather than the bucket owner pays for requests and data transfer.
Tags	Assign tags to objects for use in costing, billing, security, etc.
Events	Trigger notifications to SNS, SQS or Lambda when certain events happen in your bucket
Static Web Hosting	Simple and massively scalable static website hosting
BitTorrent	Use the BitTorrent protocol to retrieve any publicly available object by automatically generating a .torrent file



A dark, moody photograph of a person from the chest up. They are wearing a dark-colored button-down shirt over a blue collared shirt and a patterned tie. Their hands are clasped together in front of them. The background is dark and out of focus.

Amazon Glacier

Amazon Glacier



A CLOUD GURU



Amazon Glacier



A CLOUD GURU

- Cheap, slow to respond, seldom accessed
- “Cold Storage”
- Used by AWS Storage Gateway Virtual Tape Library
- Integrated with AWS S3 via Lifecycle Management
- Faster retrieval speed options if you pay more

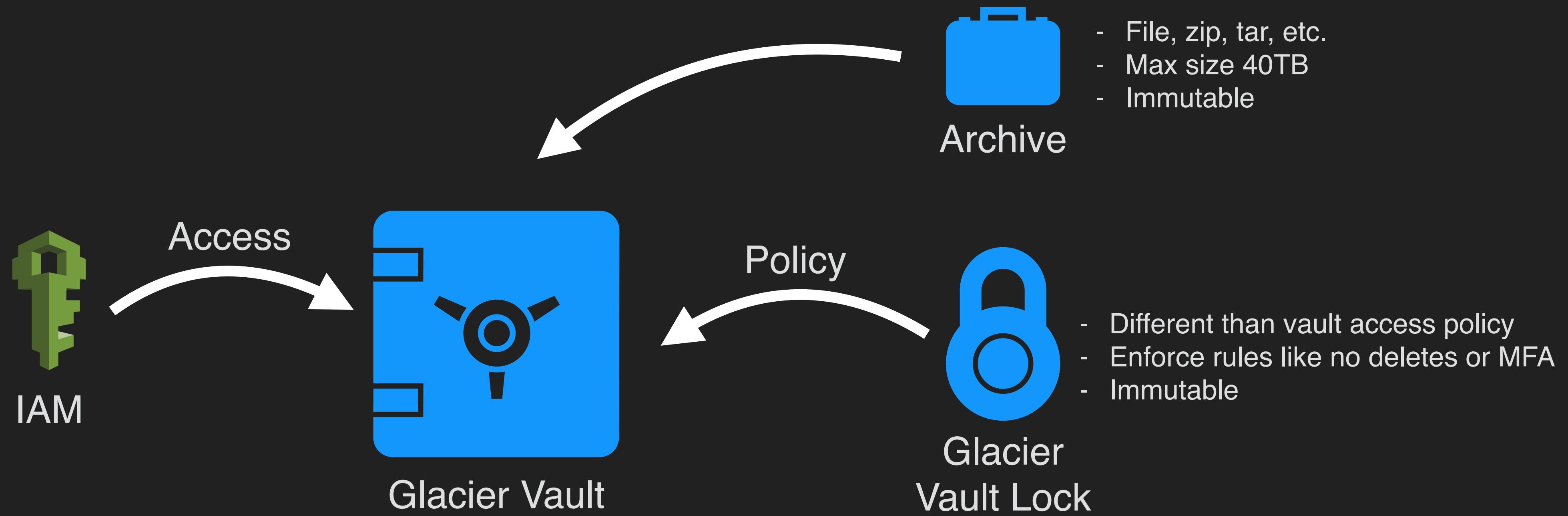


Amazon
Glacier

Amazon Glacier



A CLOUD GURU



Amazon
Glacier

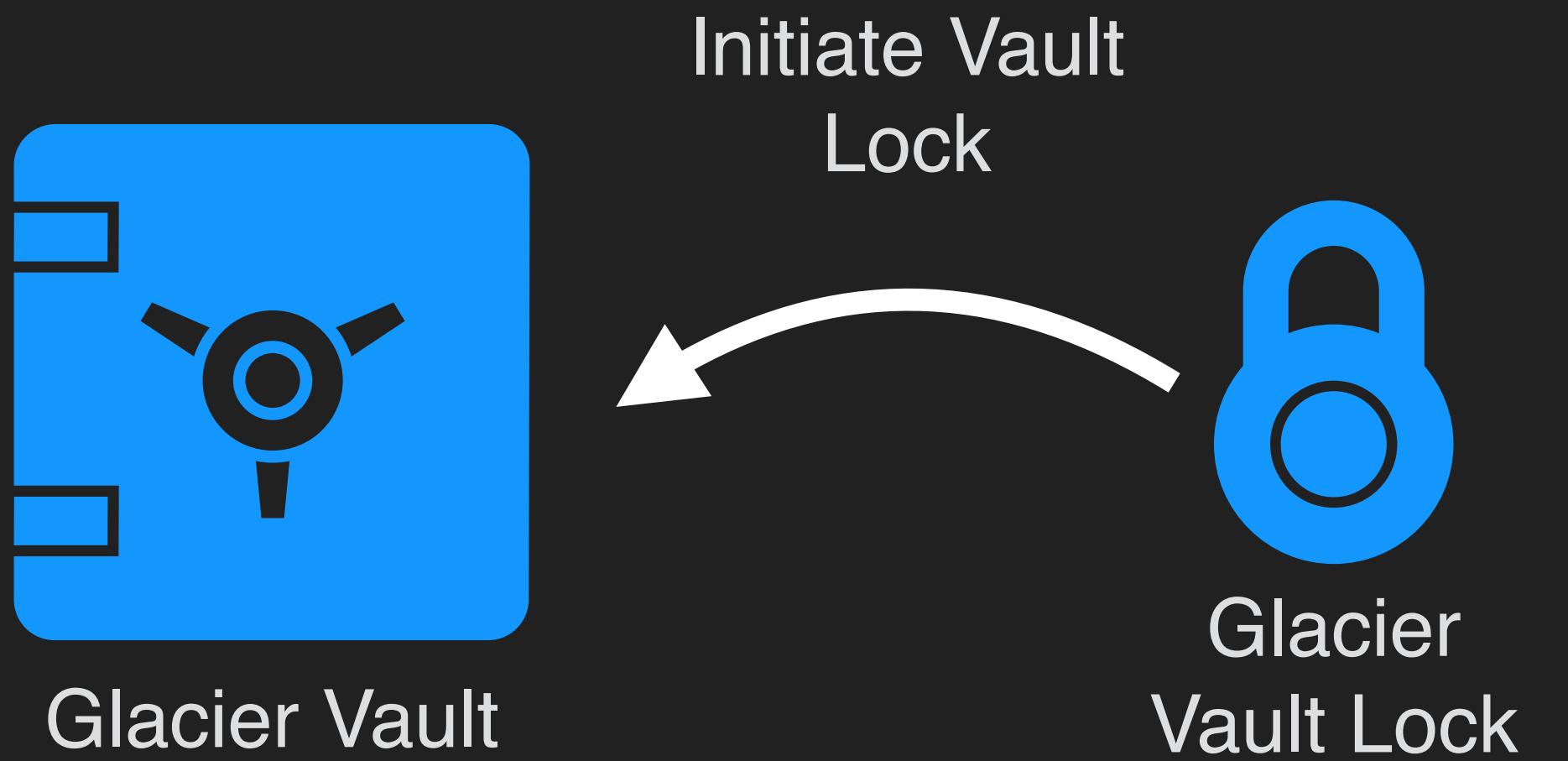
Amazon Glacier Vault Locks



A CLOUD GURU



24 Hour
Timeout



Amazon
Glacier

Amazon Glacier Vault Locks



A CLOUD GURU

Abort
Vault Lock

OR

Complete
Vault Lock



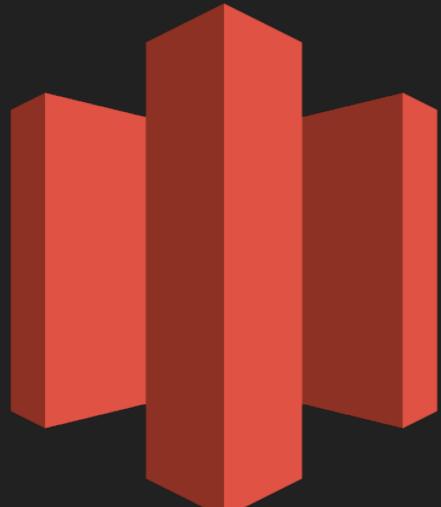
24 Hours



Glacier
Vault



Glacier
Vault Lock



Amazon
Glacier

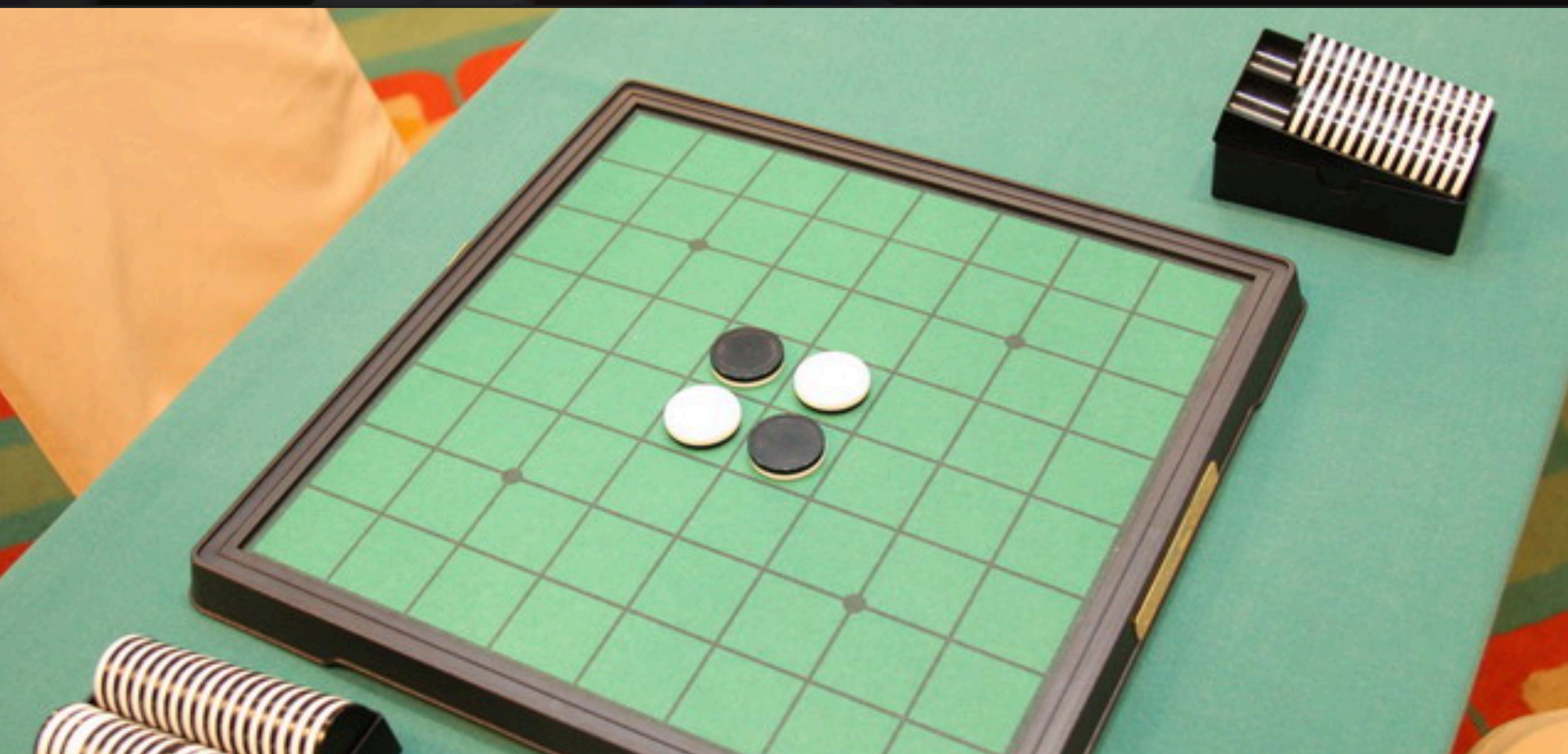
A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a small, bright screen of a smartphone, which is the primary light source in the image.

Amazon EBS

Amazon Elastic Block Storage



A CLOUD GURU

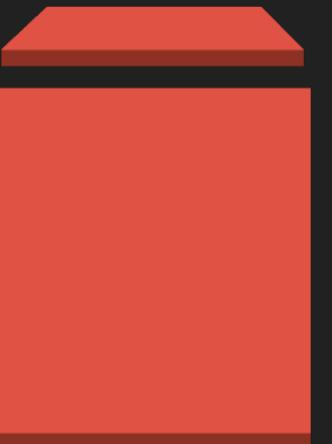


Amazon Elastic Block Storage



A CLOUD GURU

- Think “virtual hard drives”
- Can only be used with EC2
- Tied to a single AZ
- Variety of Optimized choices for IOPS, Throughput and Cost
- Snapshots are great!

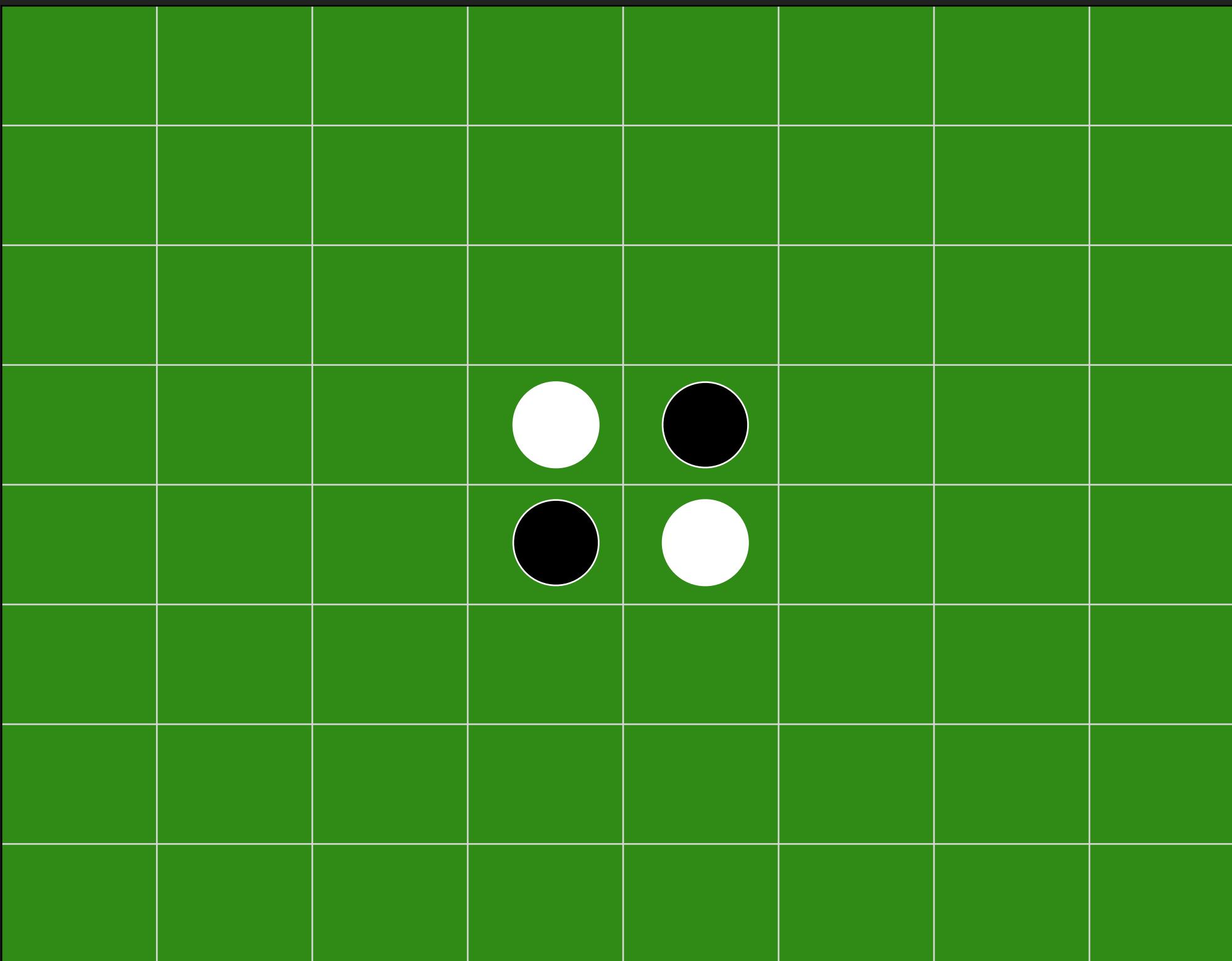


Amazon
EBS

Amazon Elastic Block Storage



A CLOUD GURU

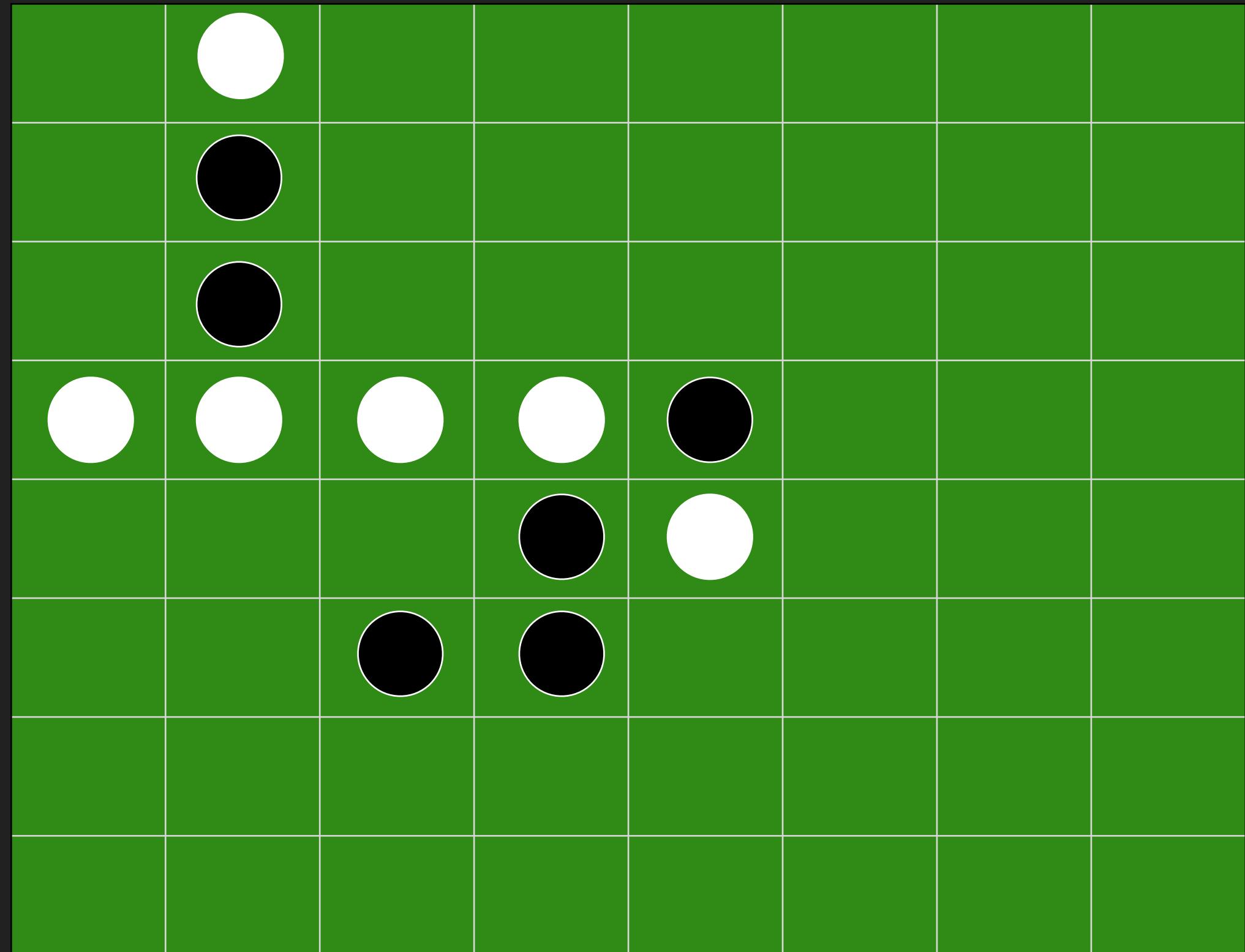


Pay for 64 blocks
Only using 4 at the moment...

Amazon Elastic Block Storage



A CLOUD GURU

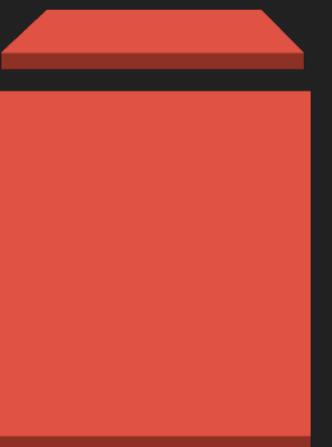
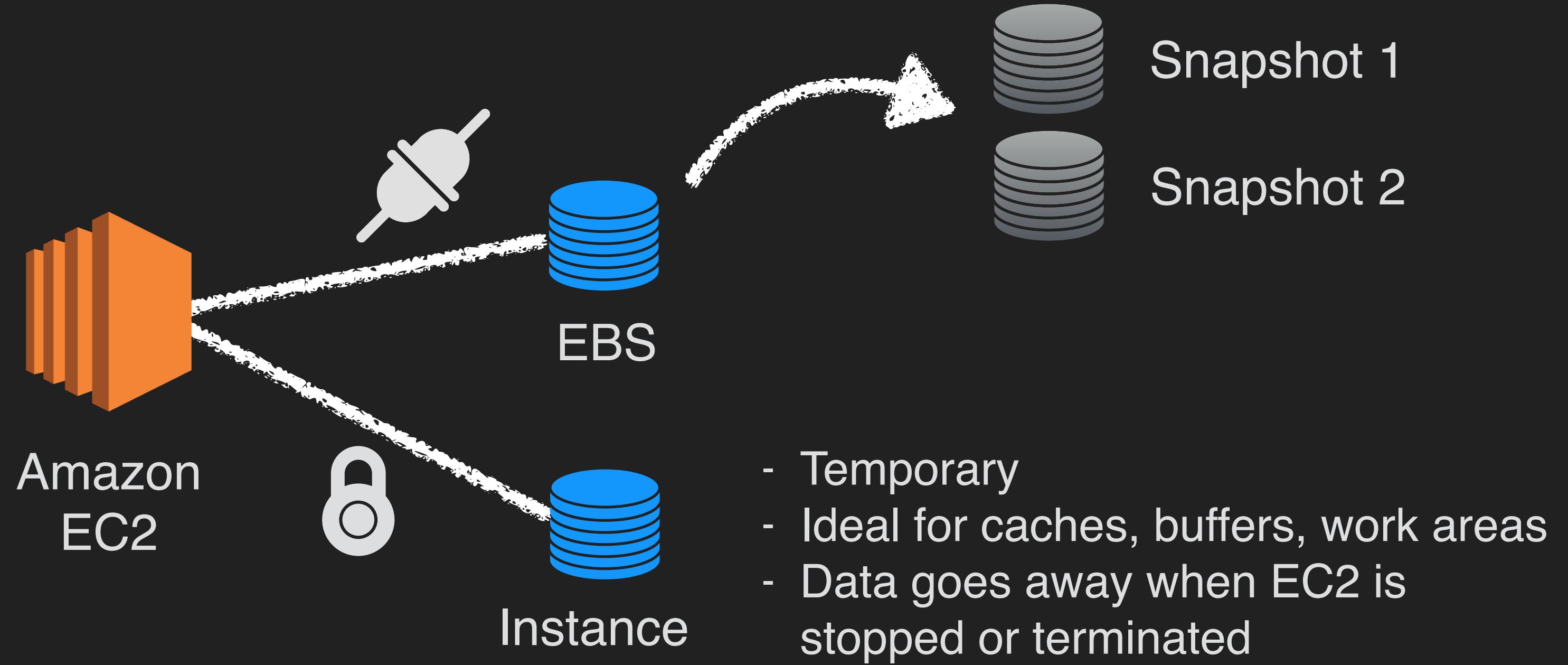


Pay for 64 blocks
Using 12 now...

Amazon Elastic Block Storage



A CLOUD GURU



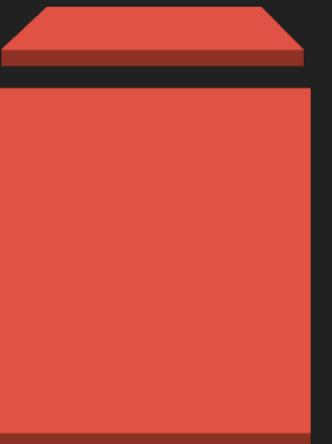
Amazon
EBS

Amazon EBS Snapshots



A CLOUD GURU

- Cost-effective and easy backup strategy
- Share data sets with other users or accounts
- Migrate a system to a new AZ or Region
- Convert unencrypted volume to an encrypted volume

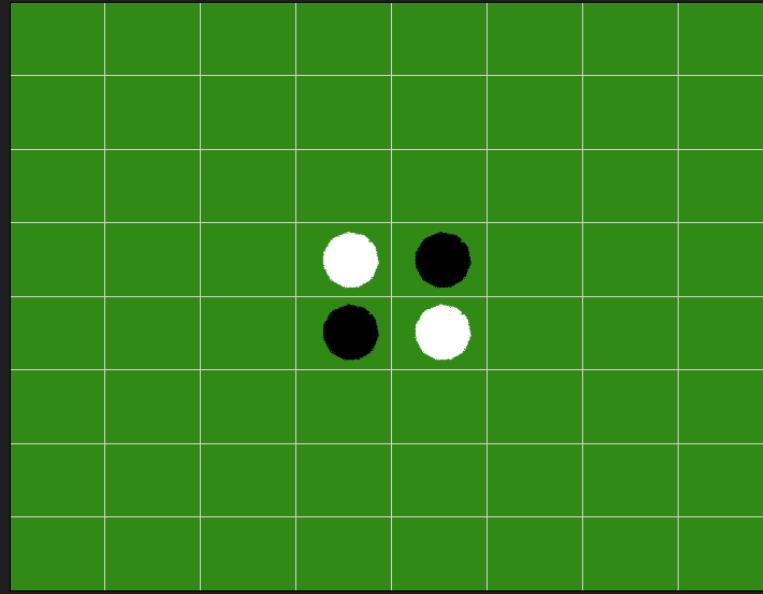
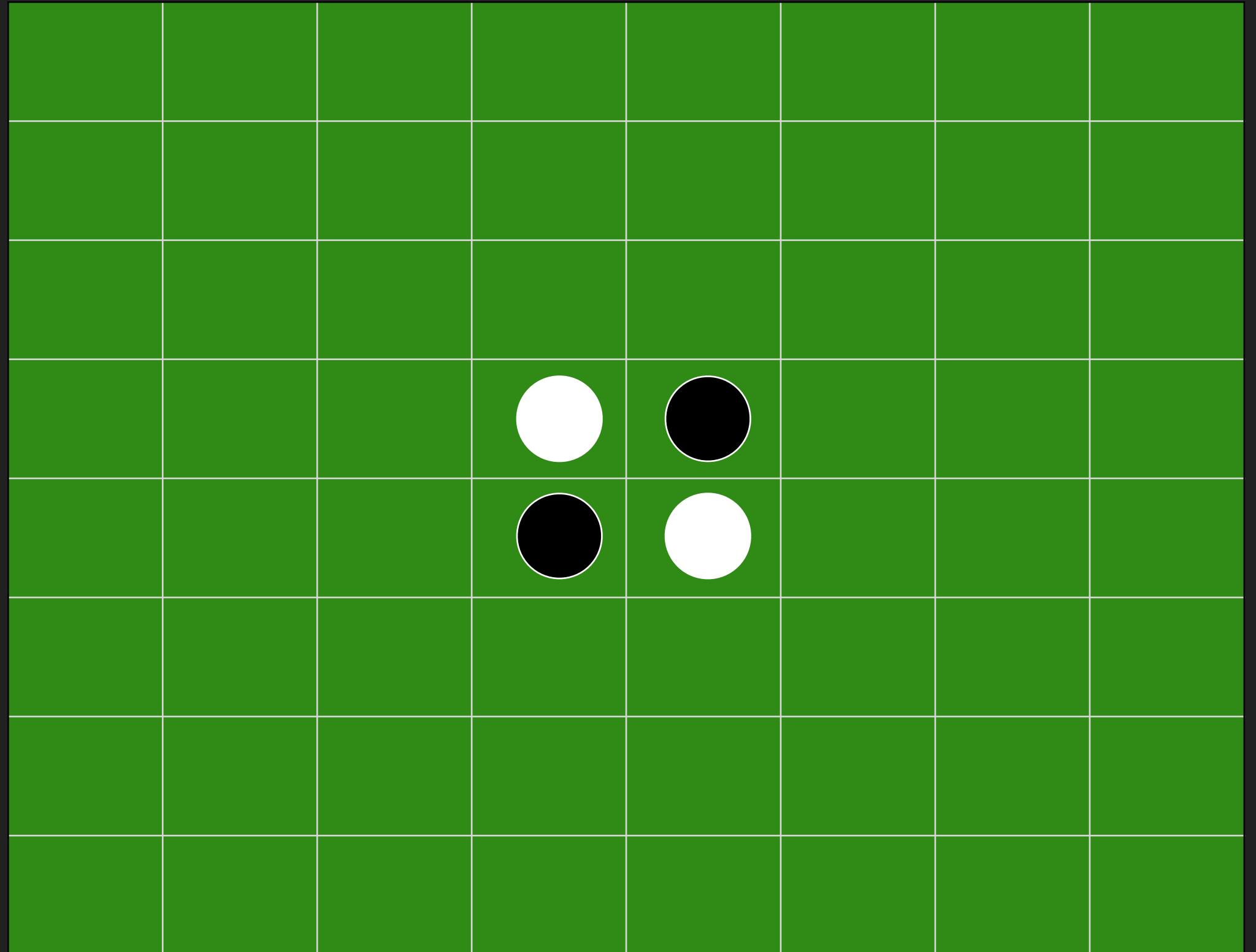


Amazon
EBS

Amazon EBS Snapshots



A CLOUD GURU

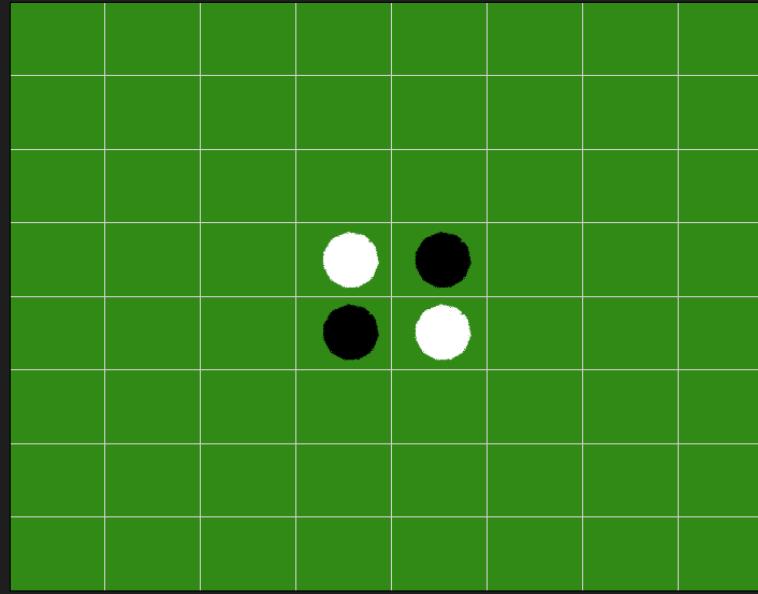
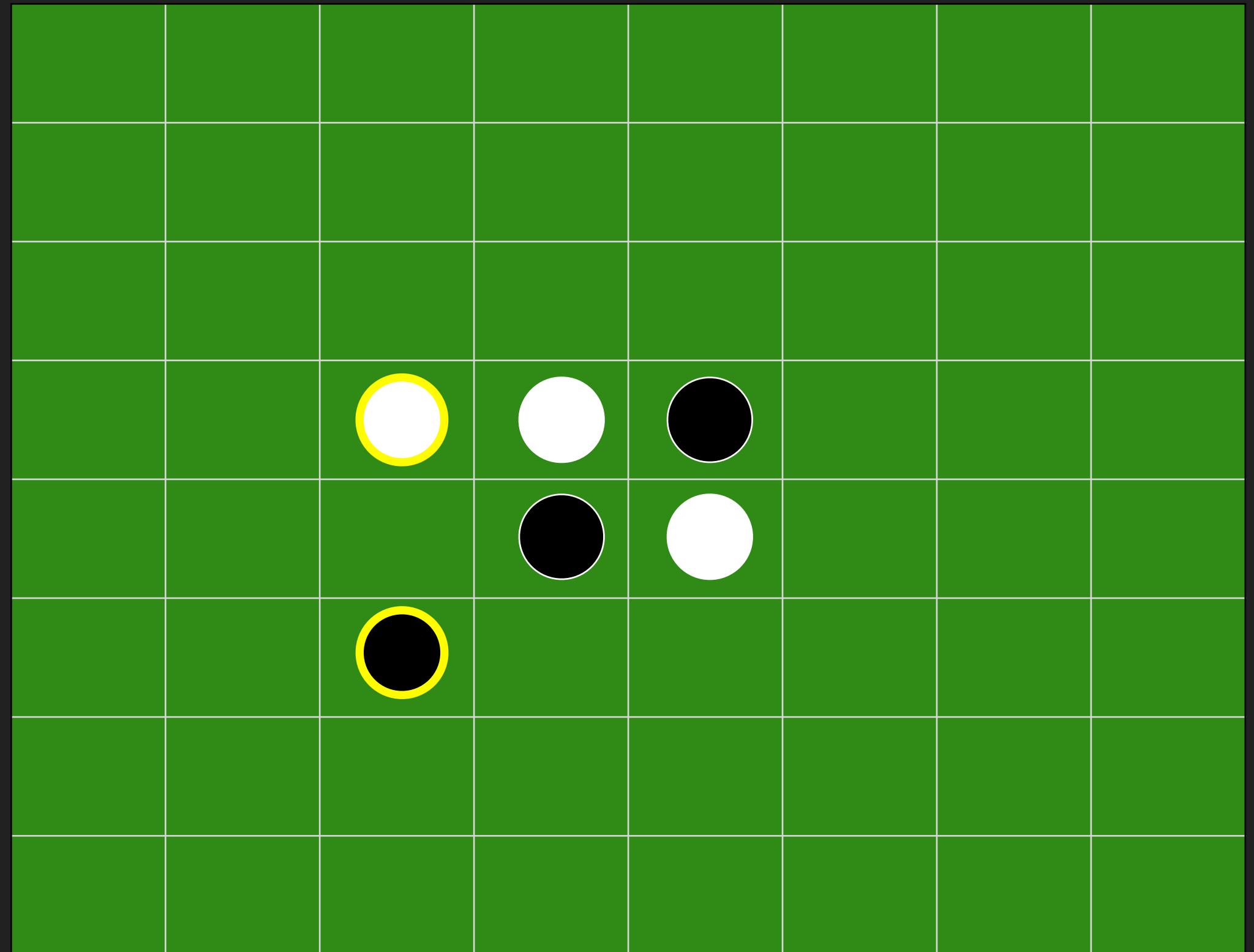


Snapshot 1

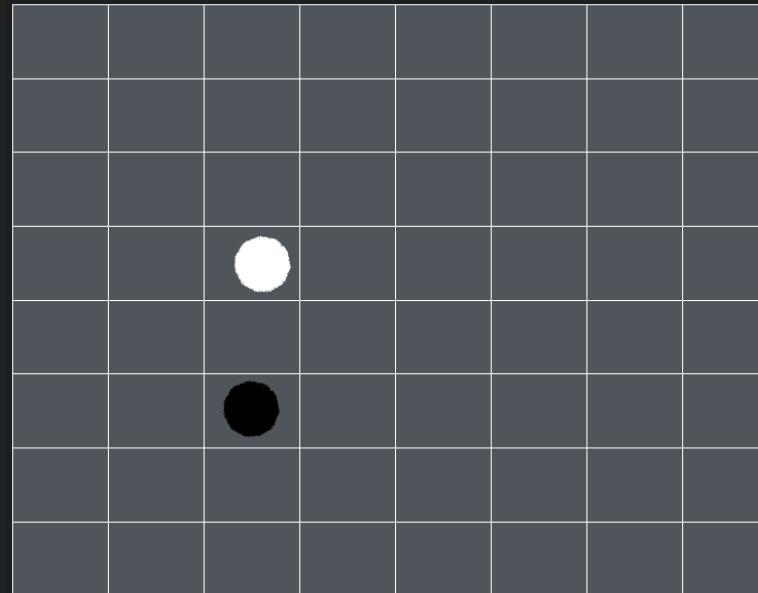
Amazon EBS Snapshots



A CLOUD GURU



Snapshot 1

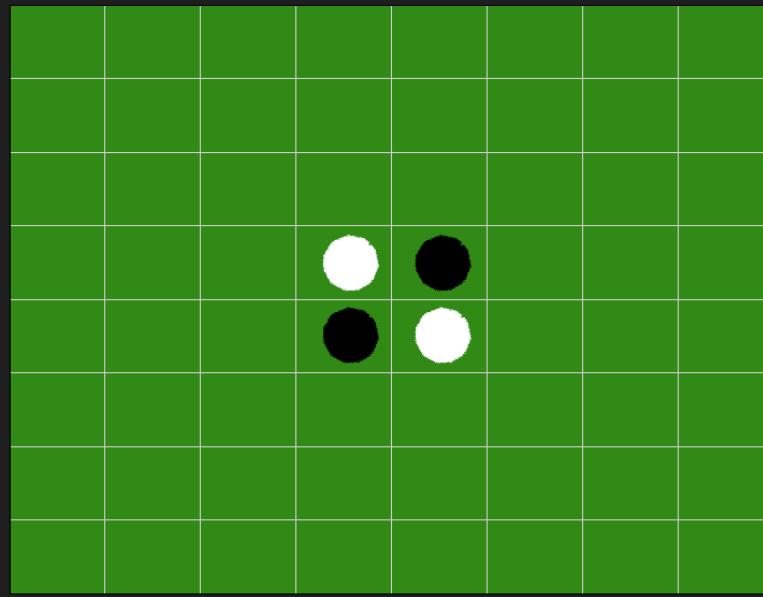
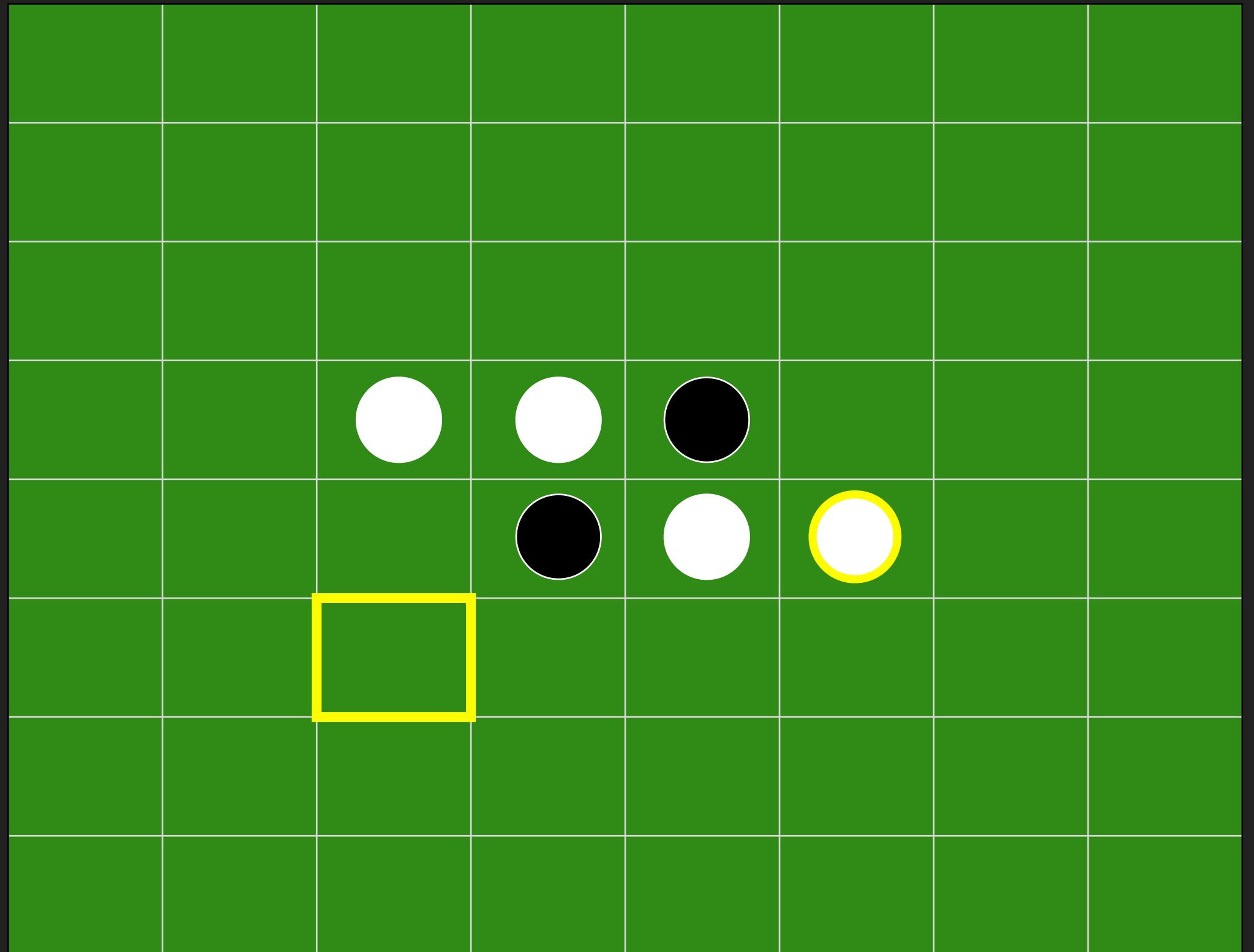


Snapshot 2

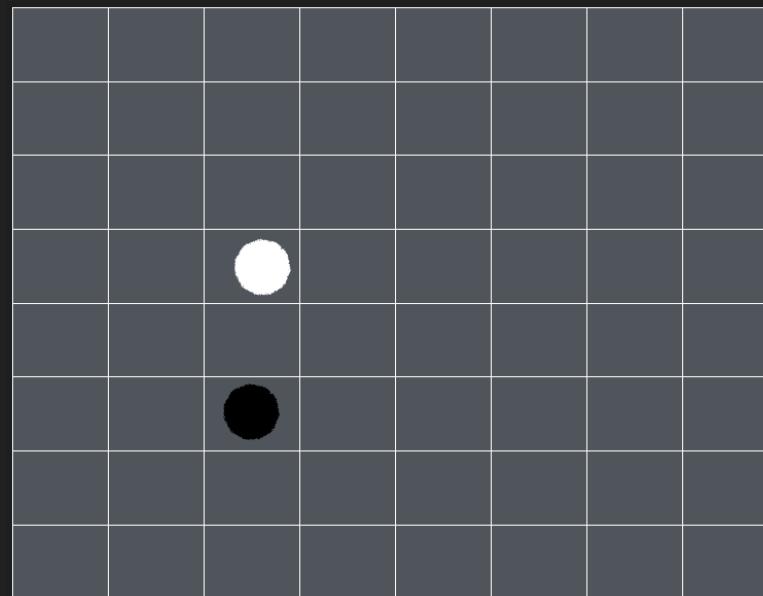
Amazon EBS Snapshots



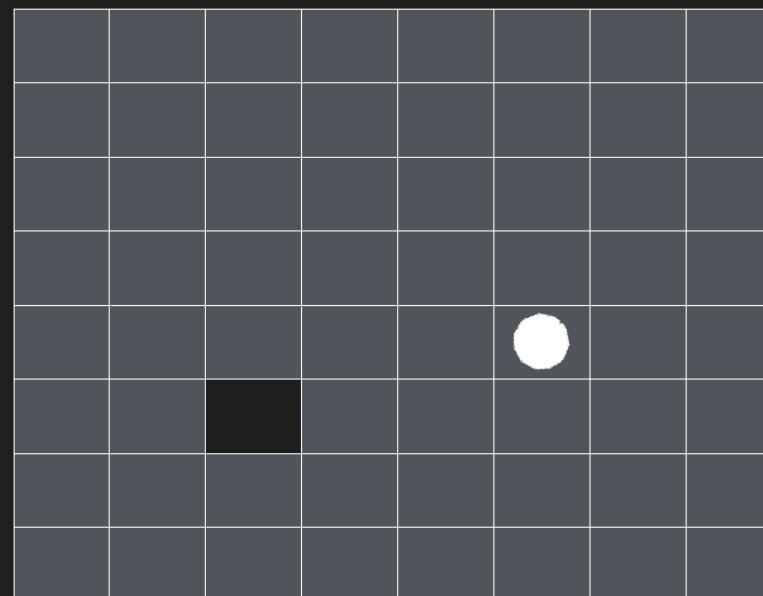
A CLOUD GURU



Snapshot 1



Snapshot 2

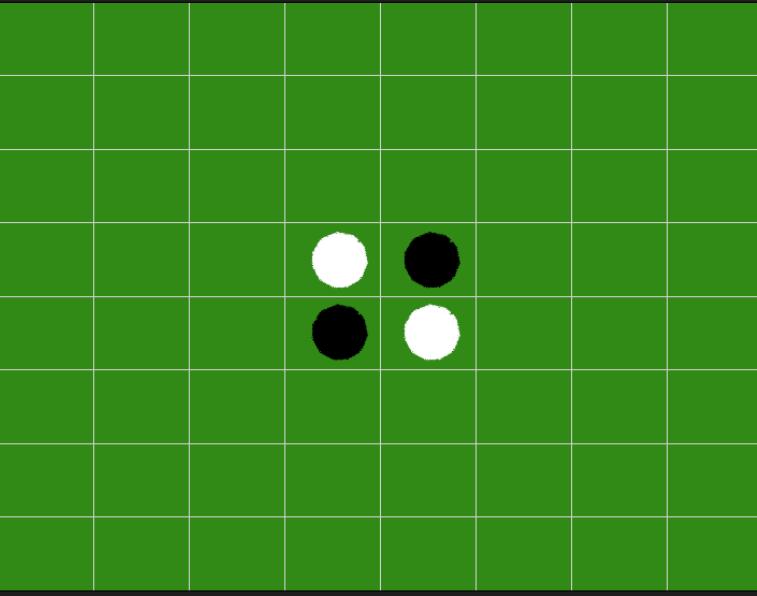
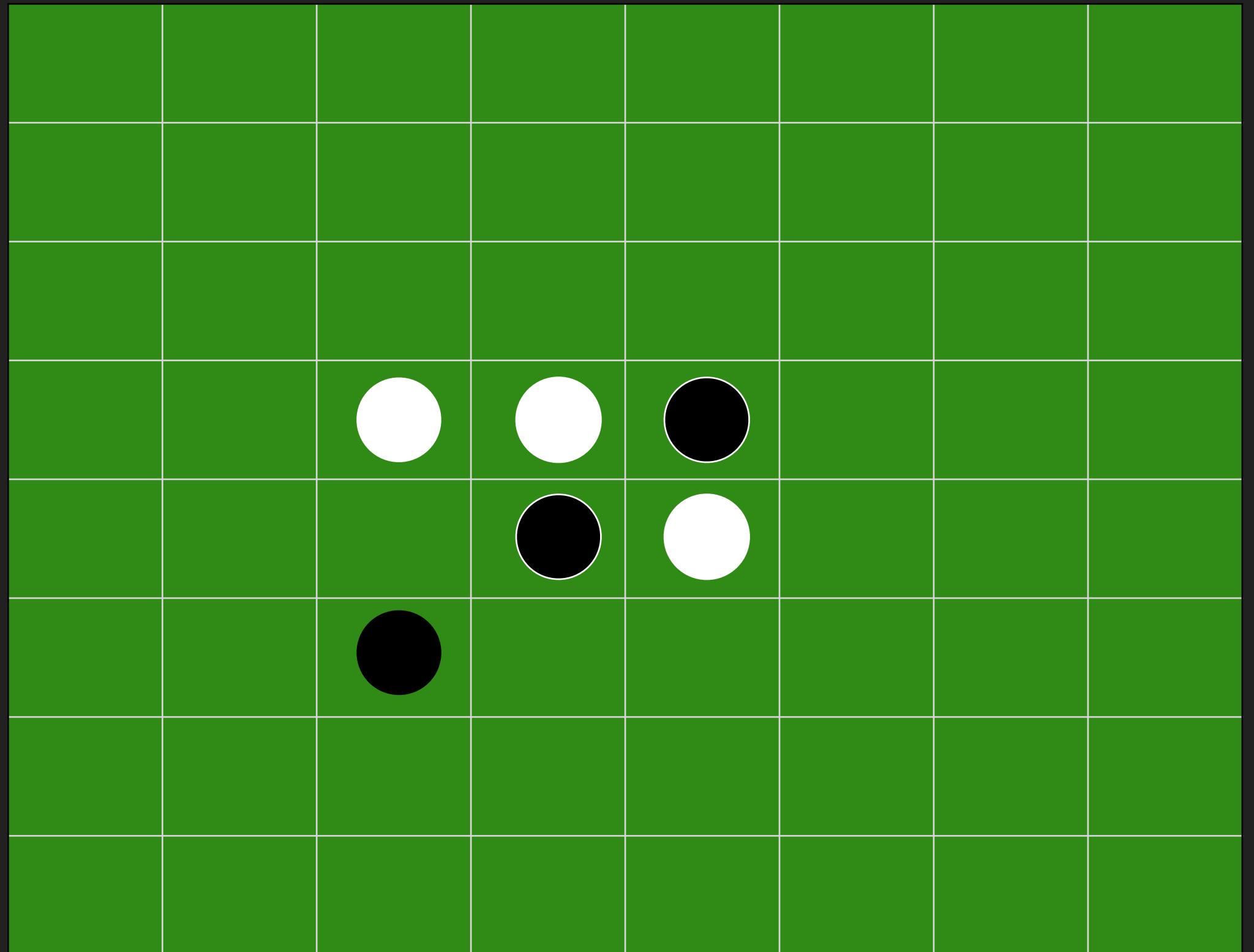


Snapshot 3

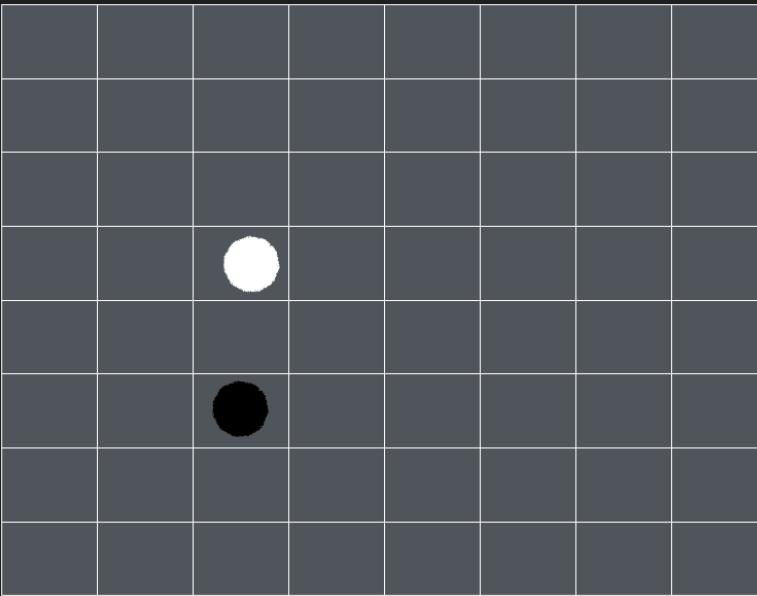
Amazon EBS Snapshots



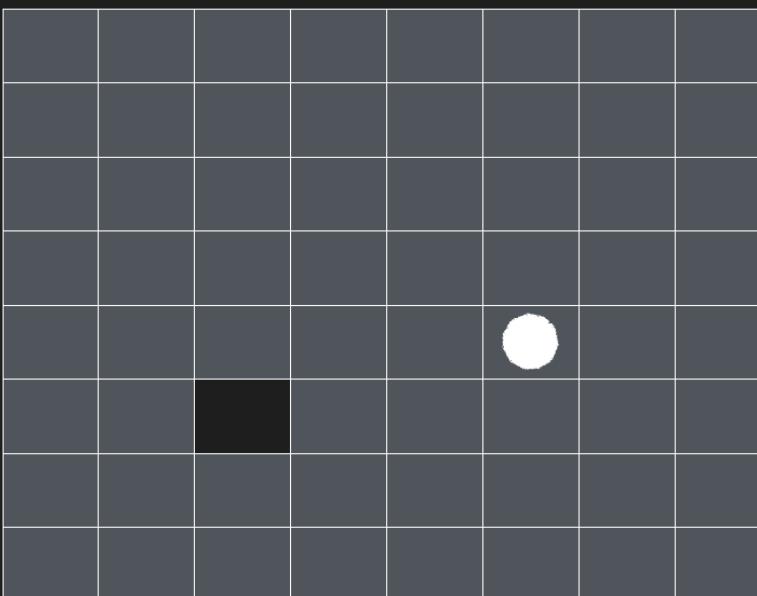
A CLOUD GURU



Snapshot 1



Snapshot 2



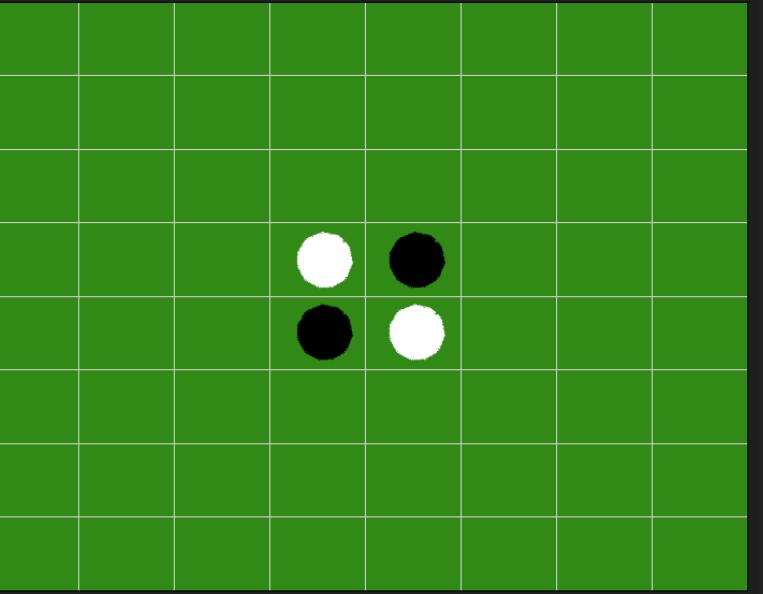
Snapshot 3



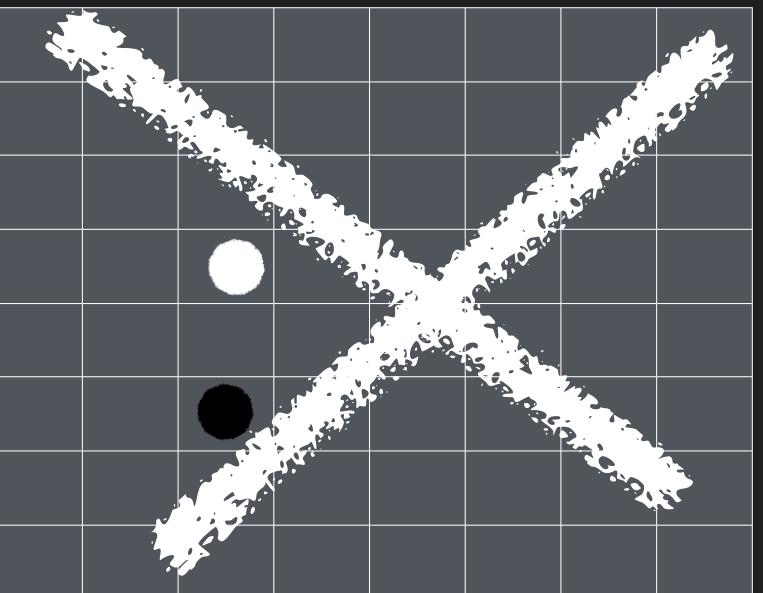
Amazon EBS Snapshots



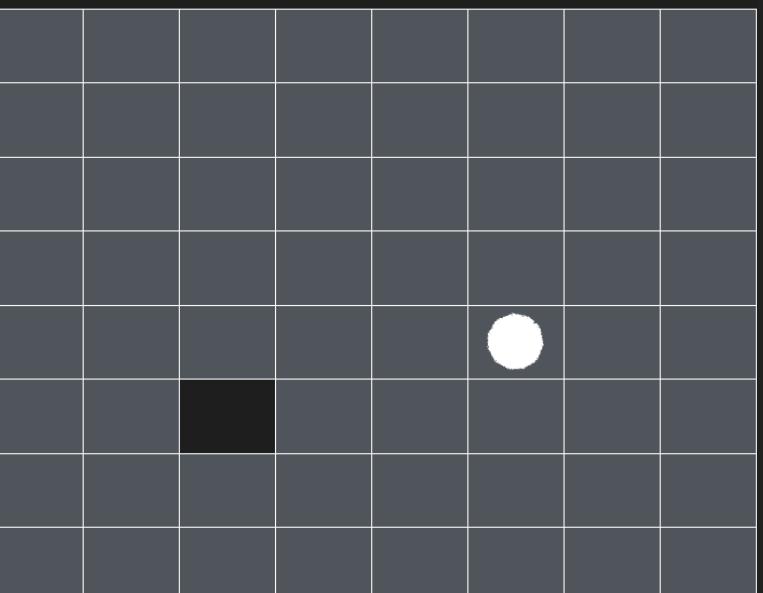
A CLOUD GURU



Snapshot 1



Snapshot 2

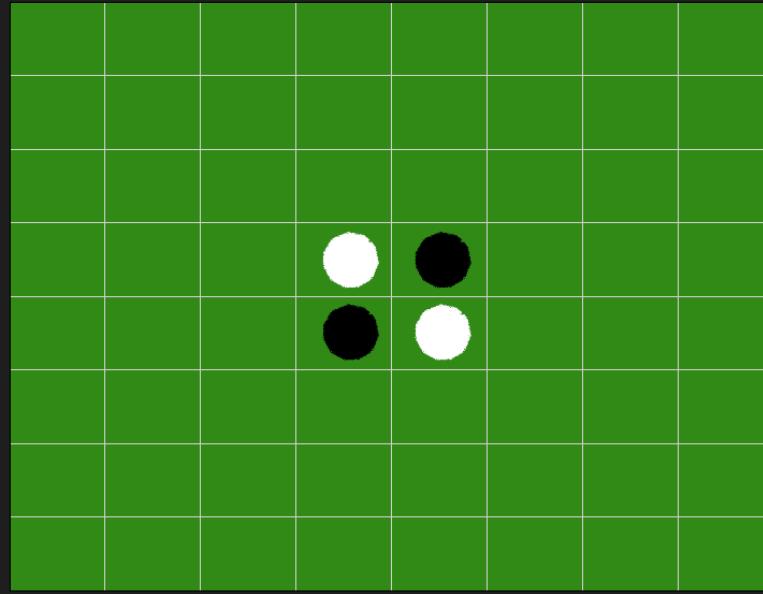
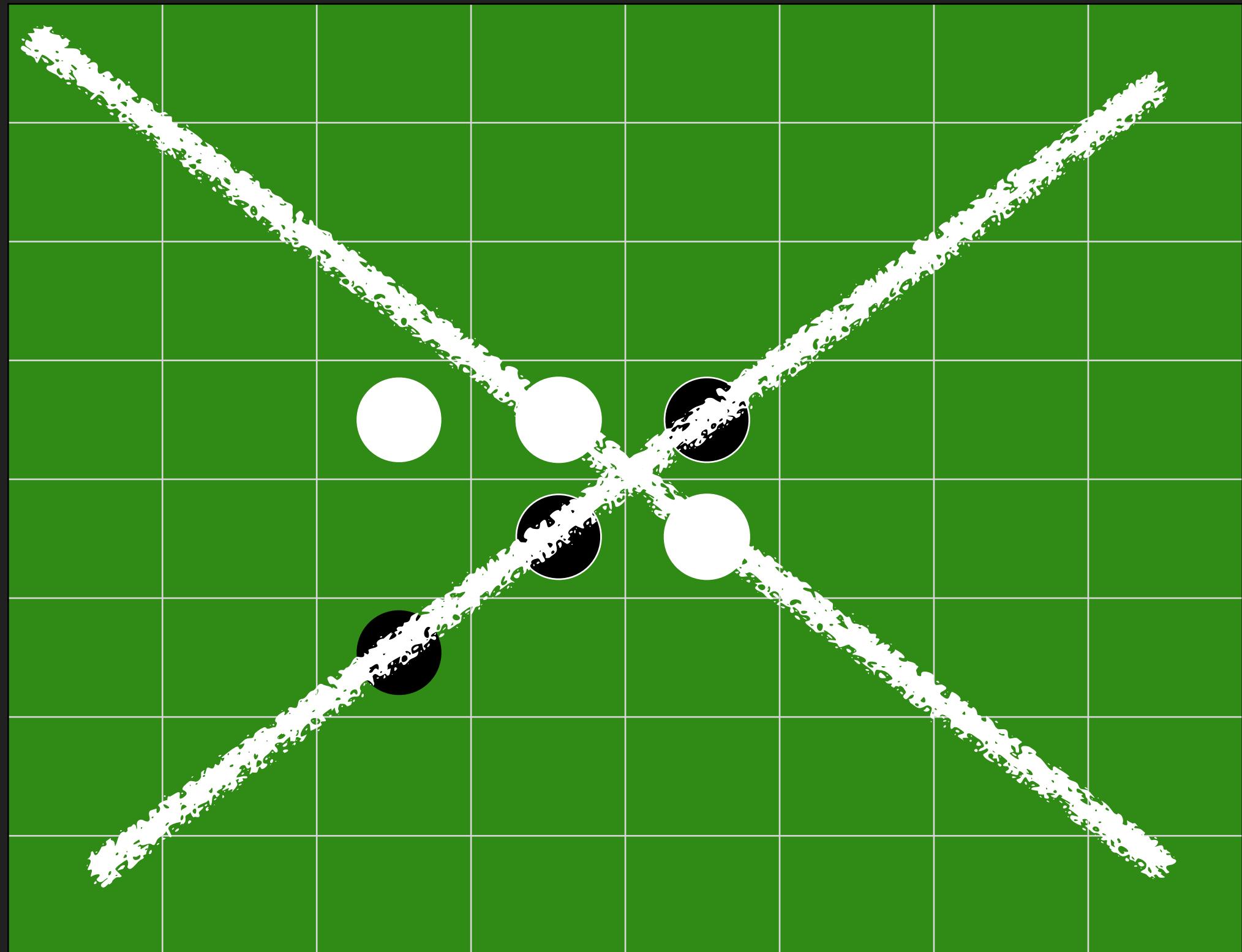


Snapshot 3

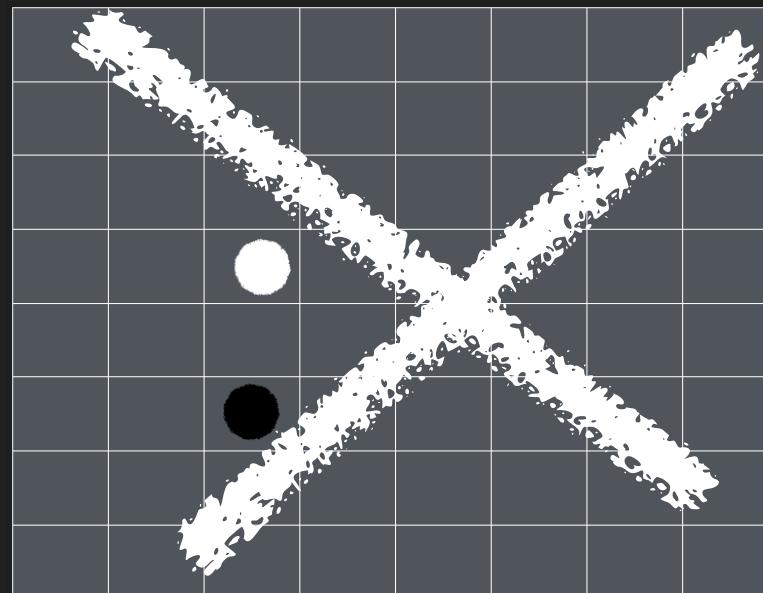
Amazon EBS Snapshots



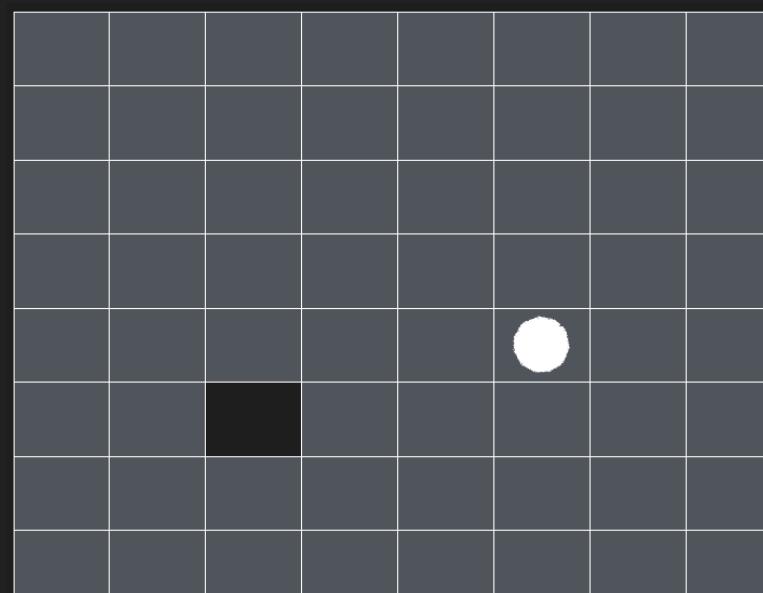
A CLOUD GURU



Snapshot 1



Snapshot 2

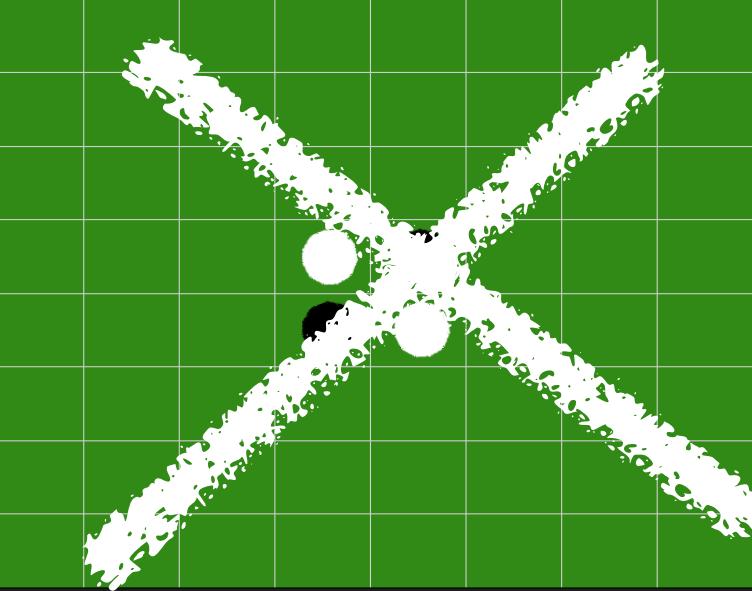


Snapshot 3

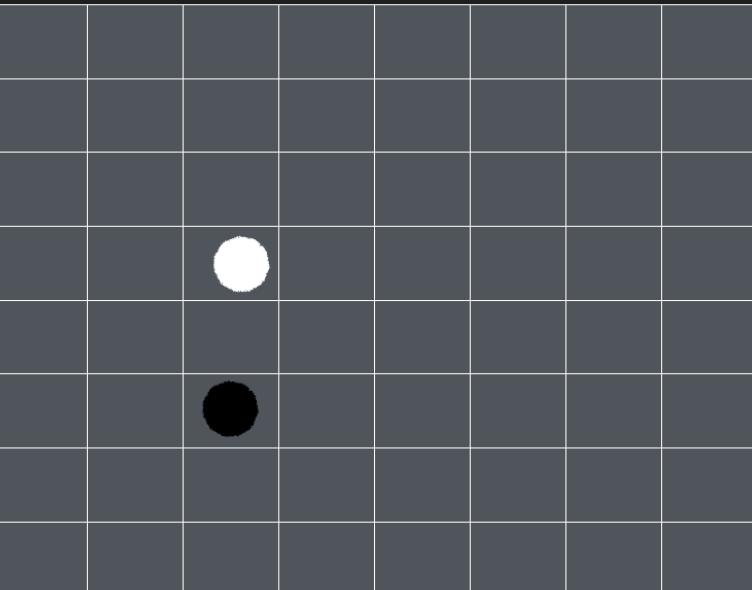
Amazon EBS Snapshots



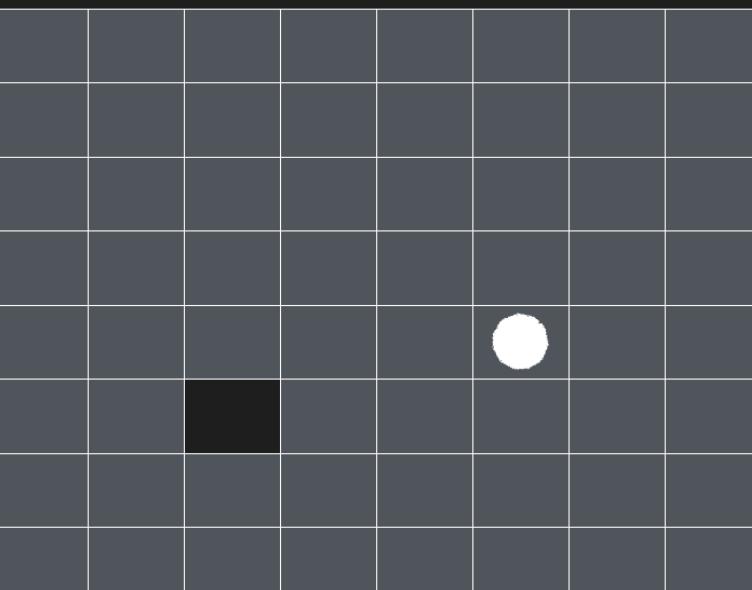
A CLOUD GURU



Snapshot 1

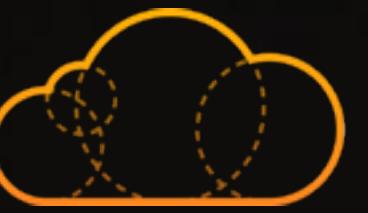


Snapshot 2

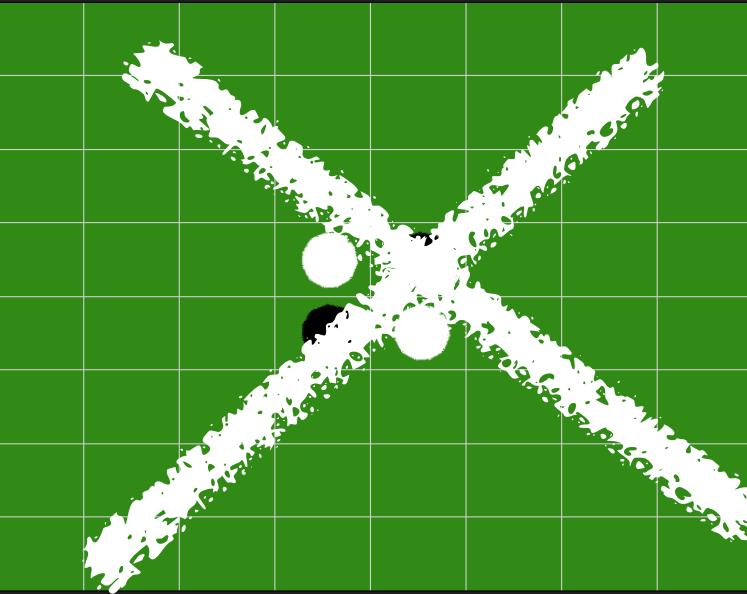


Snapshot 3

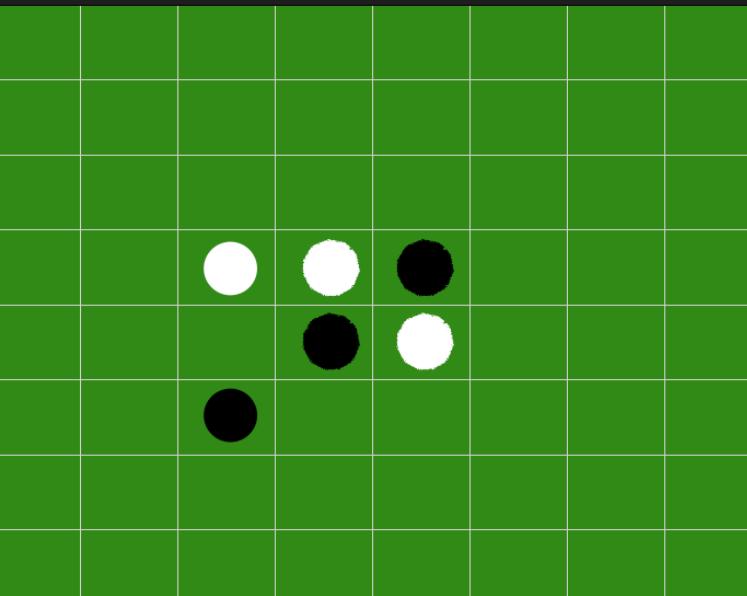
Amazon EBS Snapshots



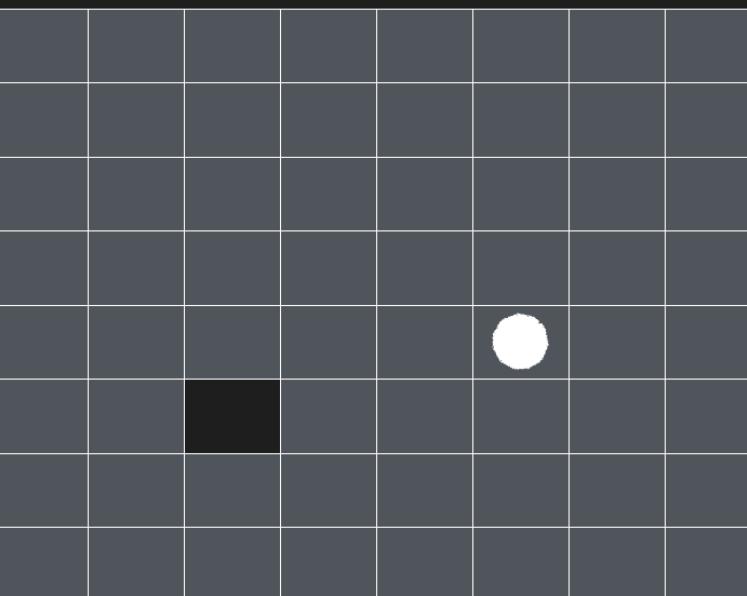
A CLOUD GURU



Snapshot 1



Snapshot 2

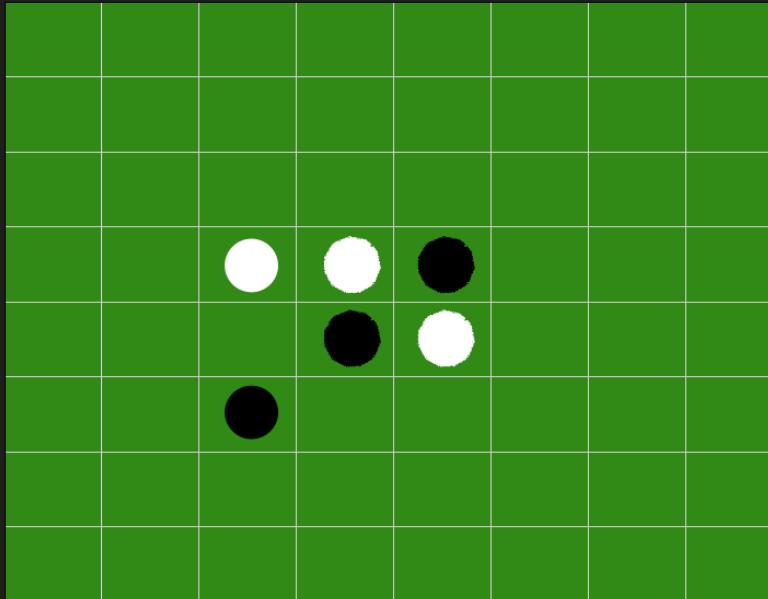
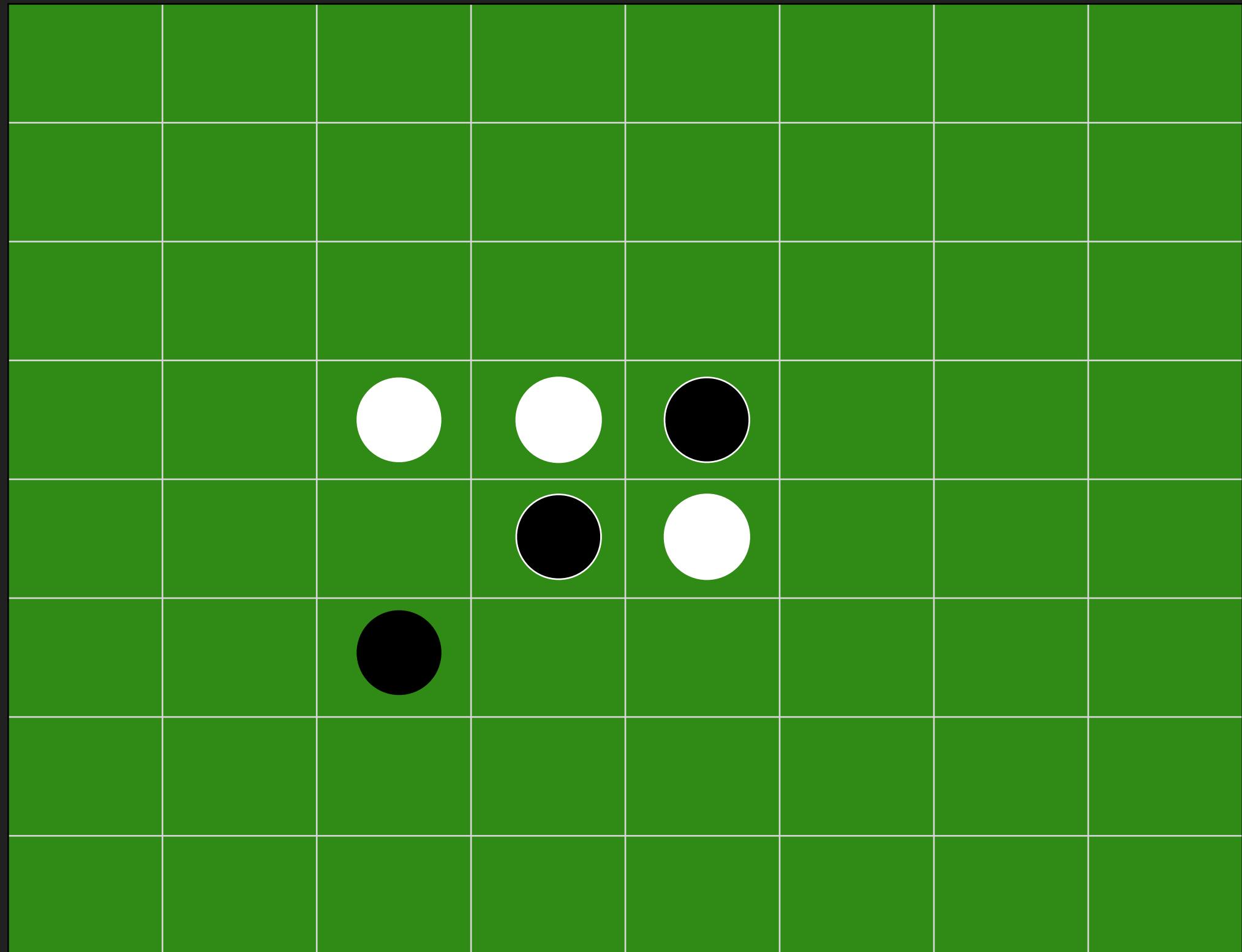


Snapshot 3

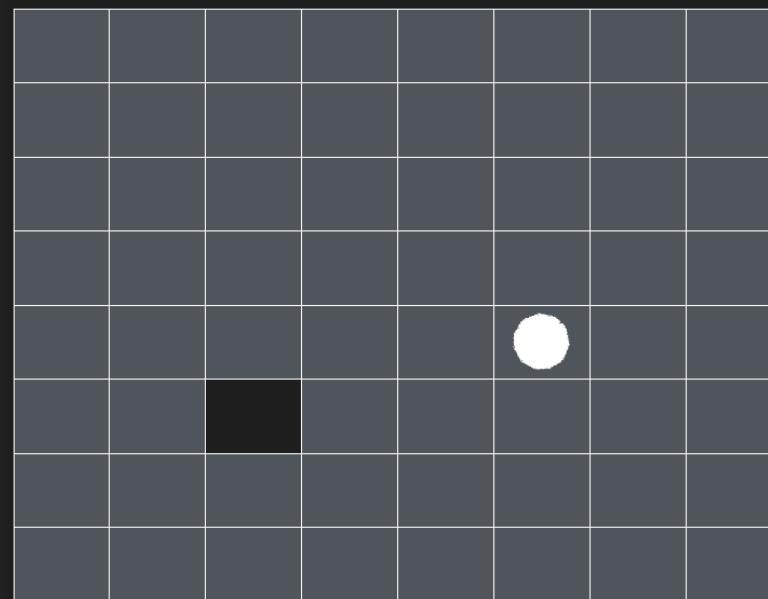
Amazon EBS Snapshots



A CLOUD GURU



Snapshot 2



Snapshot 3

Amazon Data Lifecycle Manager



A CLOUD GURU

- Schedule snapshots for volumes or instances every X hours
- Retention rules to remove stale snapshots

Create Snapshot Lifecycle Policy

Data Lifecycle Manager for EBS Snapshots will help you automate the creation and deletion of EBS snapshots based on a schedule. Volumes are targeted by tags

Description* i

Select resource type Volume Instance

Target with these tags This policy will be applied to volumes with **any** of the following tags.

You cannot use tags that are in use by another enabled or disabled lifecycle policy.

i

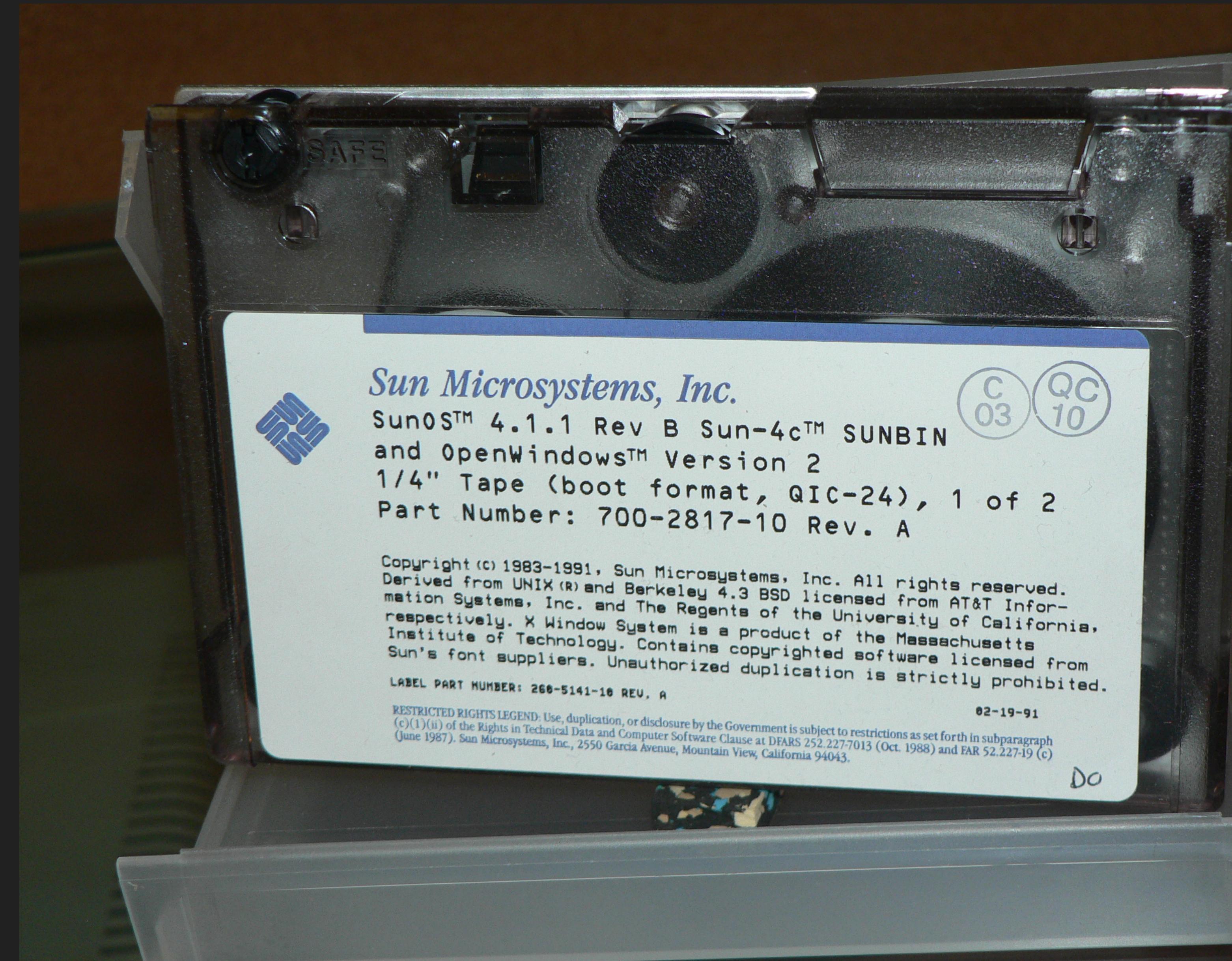
A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a small, bright screen of a smartphone, which is the primary light source in the scene.

Amazon EFS

Amazon Elastic File Service



A CLOUD GURU



Amazon Elastic File Service



A CLOUD GURU

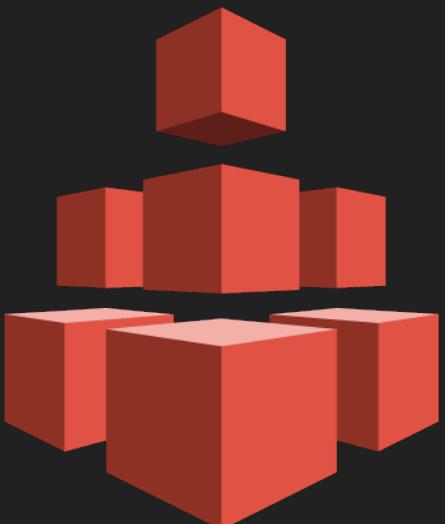


Amazon Elastic File Service



A CLOUD GURU

- Implementation of NFS file share
- Elastic storage capacity, and pay for only what you use (in contrast to EBS)
- Multi-AZ metadata and data storage
- Configure mount-points in one, or many, AZs.
- Can be mounted from on-premises systems (Caution here...)
- Alternatively, use Amazon DataSync.

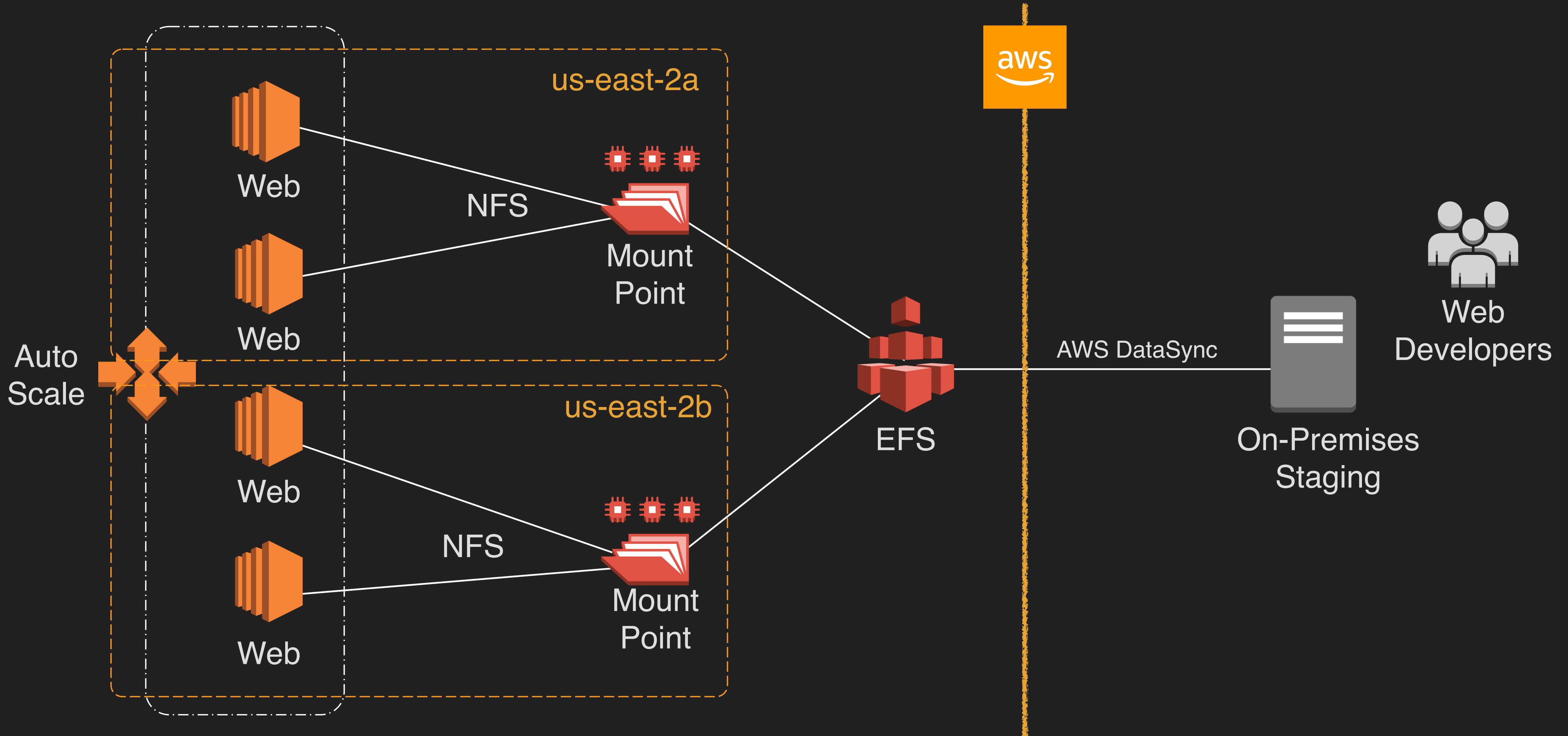


Amazon
EFS

Amazon Elastic File Service



A CLOUD GURU



Amazon Storage Gateway

Amazon Storage Gateway



A CLOUD GURU

- Virtual machine that you run on-premises with VMWare or HyperV OR via a specially configured Dell hardware appliance
- Provides local storage resources backed by AWS S3 and Glacier
- Often used in disaster recovery preparedness to sync to AWS
- Useful in cloud migrations



AWS Storage
Gateway

Amazon Storage Gateway



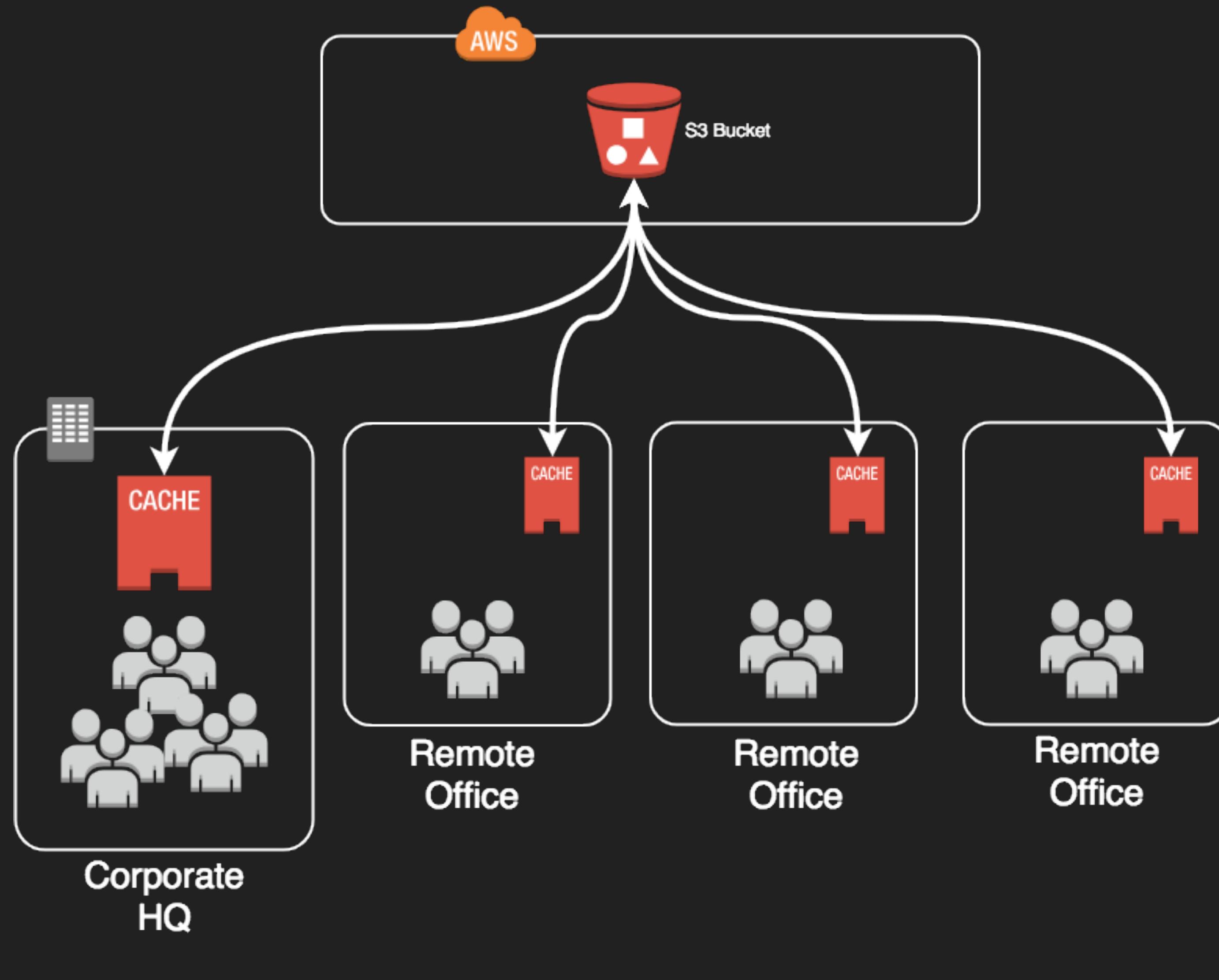
A CLOUD GURU

New Name	Old Name	Interface	Function
File Gateway	None	NFS, SMB	Allow on-prem or EC2 instances to store objects in S3 via NFS or SMB mount point
Volume Gateway Stored Mode	Gateway-stored Volumes	iSCSI	Async replication of on-prem data to S3
Volume Gateway Cached Mode	Gateway-cached Volumes	iSCSI	Primary data stored in S3 with frequently access data cached locally on-prem
Tape Gateway	Gateway-Virtual Tape Library	iSCSI	Virtual media changer and tape library for use with existing backup software

Amazon Storage Gateway



A CLOUD GURU



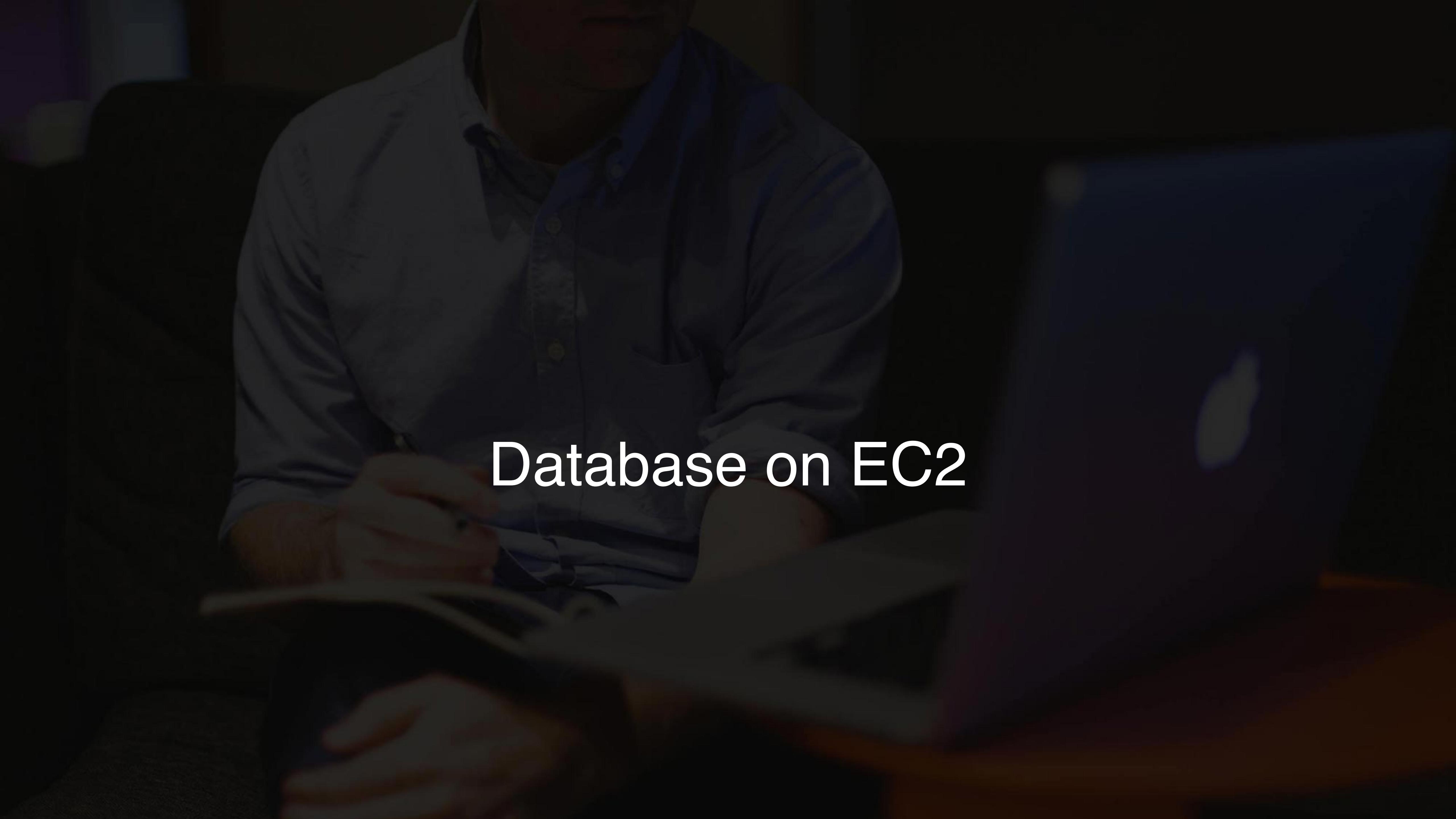
A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, and they appear to be holding a small device, likely a smartphone, which is partially visible at the bottom left.

Amazon WorkDocs



- Secure, fully managed file collaboration service
- Can integrate with AD for SSO
- Web, mobile and native clients (no Linux client)
- HIPAA, PCI DSS and ISO compliance requirements
- Available SDK for creating complementary apps



A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a smartphone. The phone's screen is visible, though its content is not readable.

Database on EC2



Database on EC2

- Run any database with full control and ultimate flexibility.
- Must manage everything like backups, redundancy, patching, scale
- Good option if you require a database not yet supported by RDS, such as IBM DB2 or SAP HANA.
- Good option if it is not feasible to migrate to AWS-managed database.

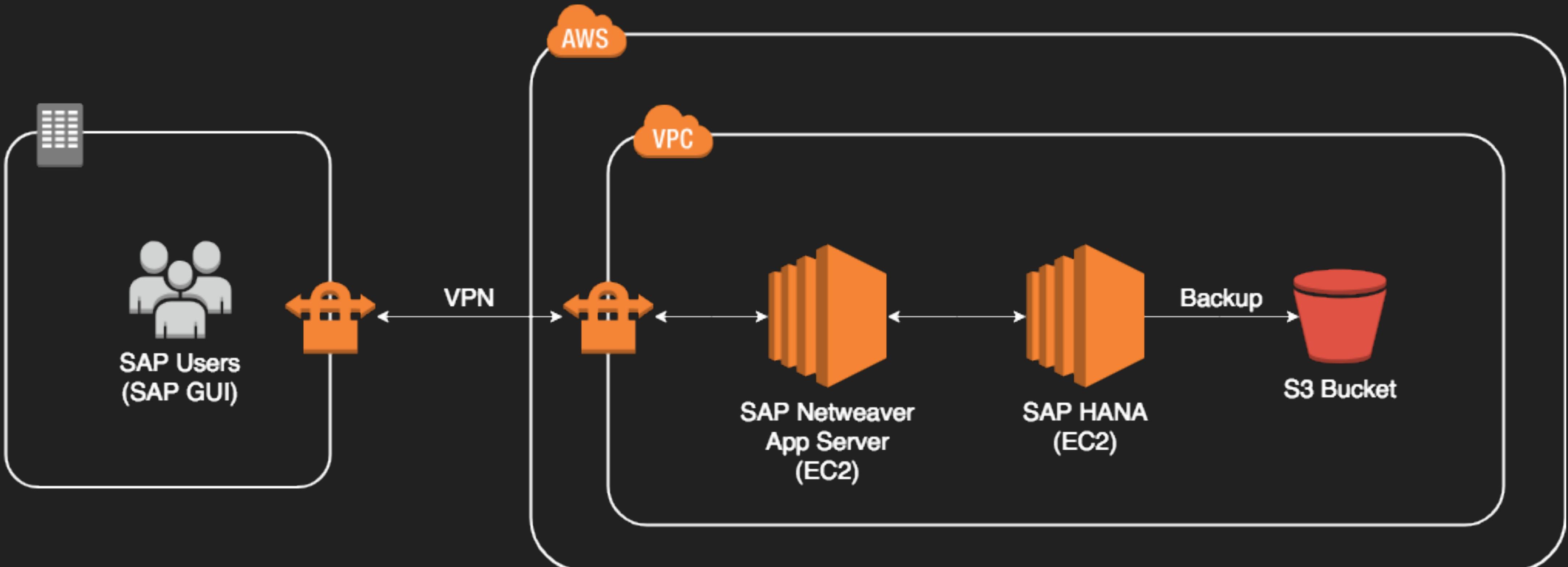


Database on EC2



A CLOUD GURU

Simple SAP S/4 HANA Deployment on AWS



A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a smartphone horizontally. The screen of the phone is visible but appears mostly black or very dimly lit.

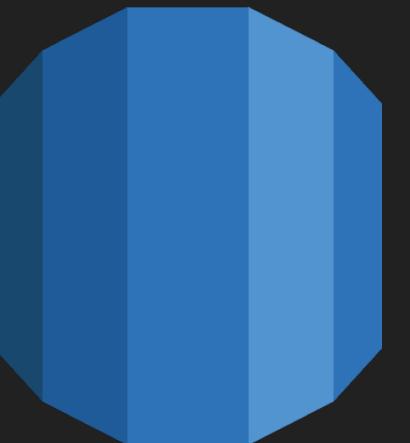
Amazon RDS

Amazon RDS



A CLOUD GURU

- Managed database option for MySQL, Maria, PostgreSQL, Microsoft SQL Server, Oracle and MySQL-compatible Aurora
- Best for structured, relational data store needs
- Aims to be drop-in replacement for existing on-prem instances of same databases
- Automated backups and patching in customer-defined maintenance windows
- Push-button scaling, replication and redundancy



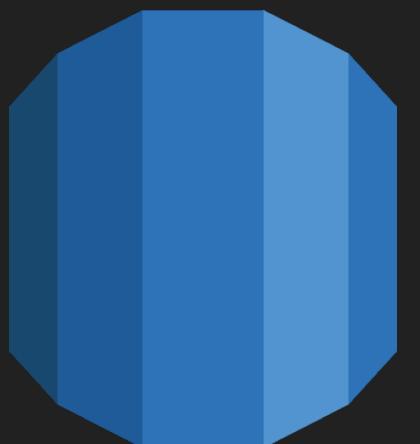
Amazon
RDS

Amazon RDS Anti-Patterns



A CLOUD GURU

If you need...	Don't use RDS, instead use..
Lots of large binary objects (BLOBs)	S3
Automated scalability	DynamoDB
Name/Value Data Structure	DynamoDB
Data is not well structured or unpredictable	DynamoDB
Other database platforms like IBM DB2 or SAP HANA	EC2
Complete control over the database	EC2

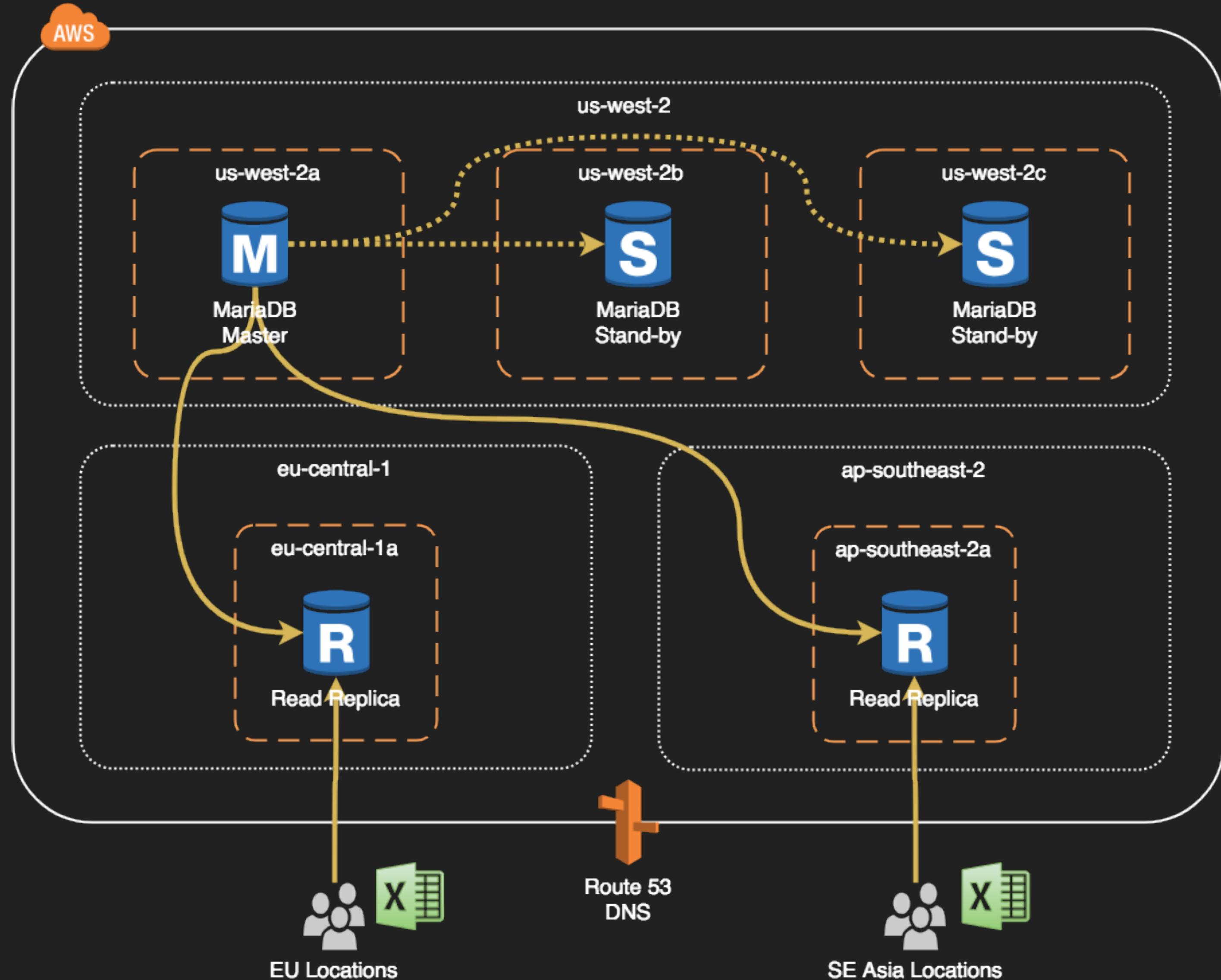


Amazon
RDS

Amazon RDS



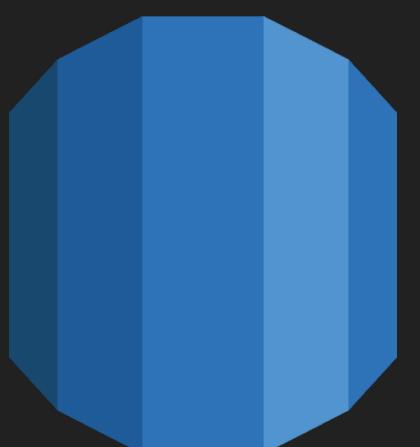
A CLOUD GURU



- Multi-AZ RDS
- Read-Replicas service regional users

NOTE for MySQL: non-transactional storage engines like MyISAM don't support replication; you must use InnoDB (or XtraDB on Maria)

MariaDB is an open-open-source fork of MySQL.
MySQL behaves similar.

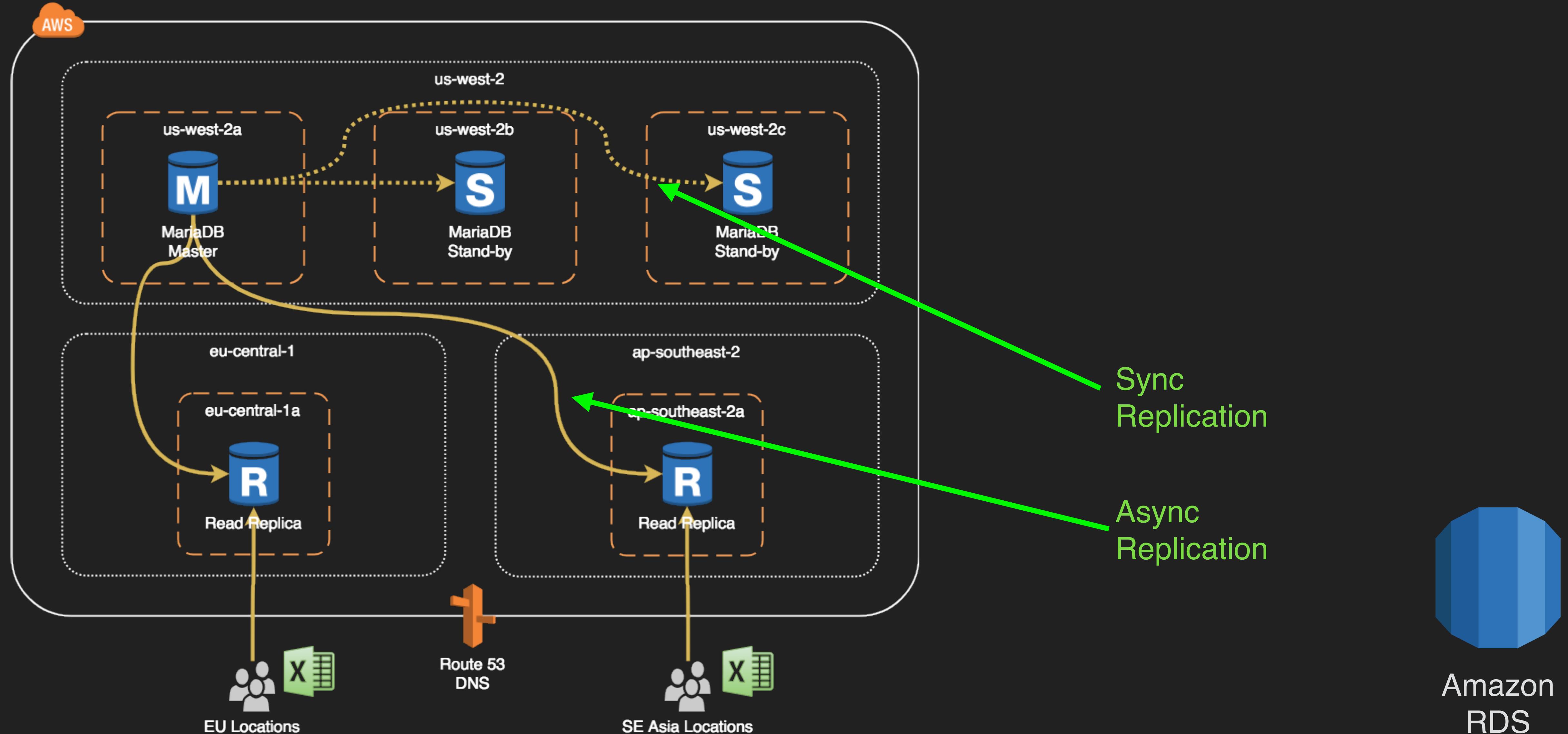


Amazon
RDS

Amazon RDS

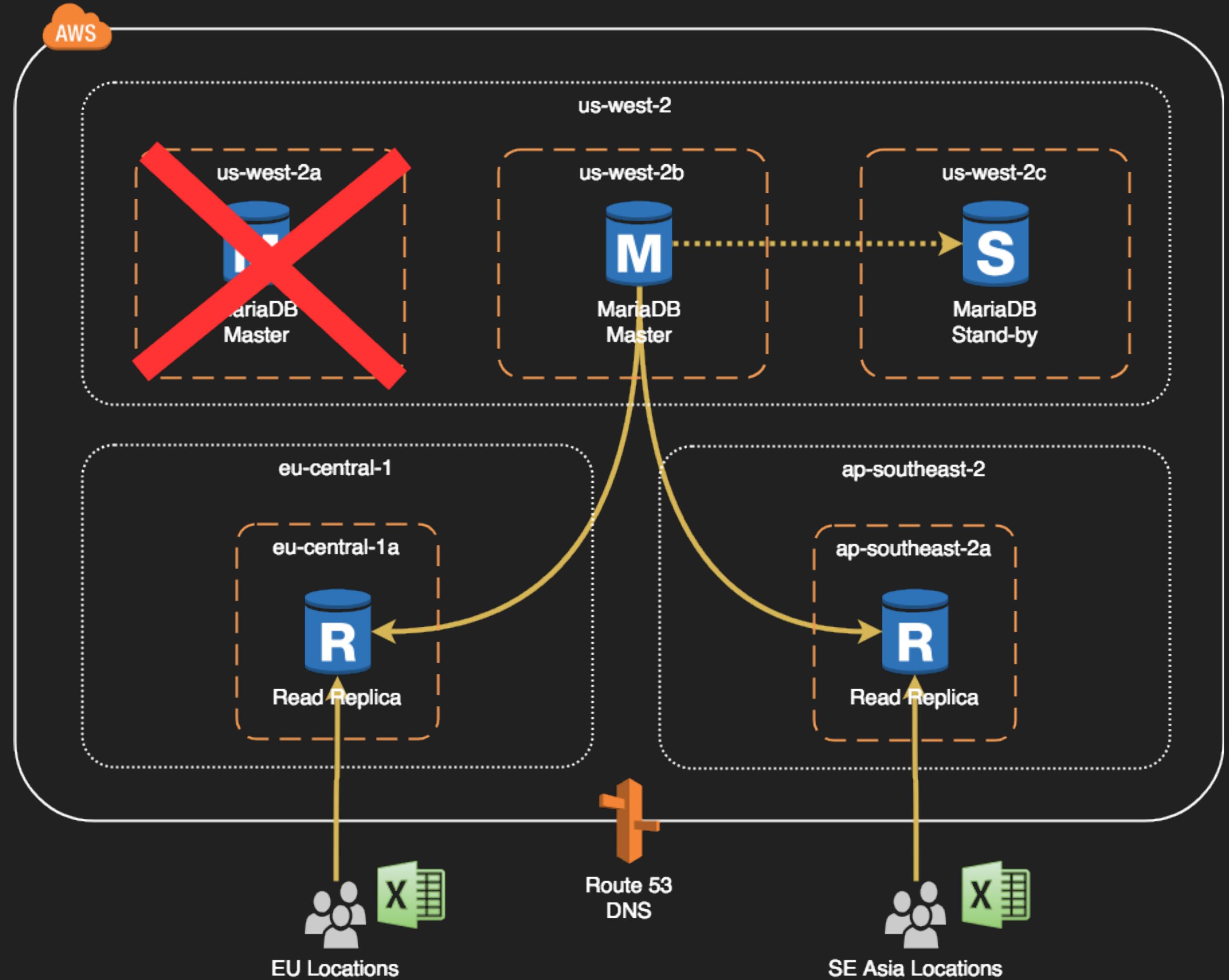


A CLOUD GURU

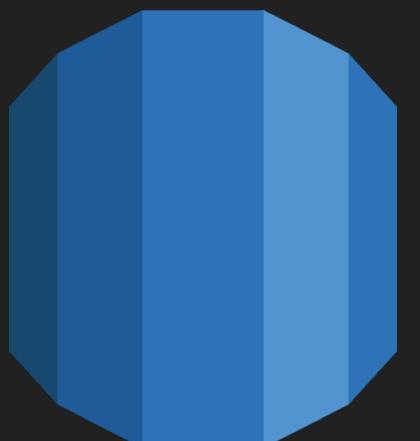




Amazon RDS



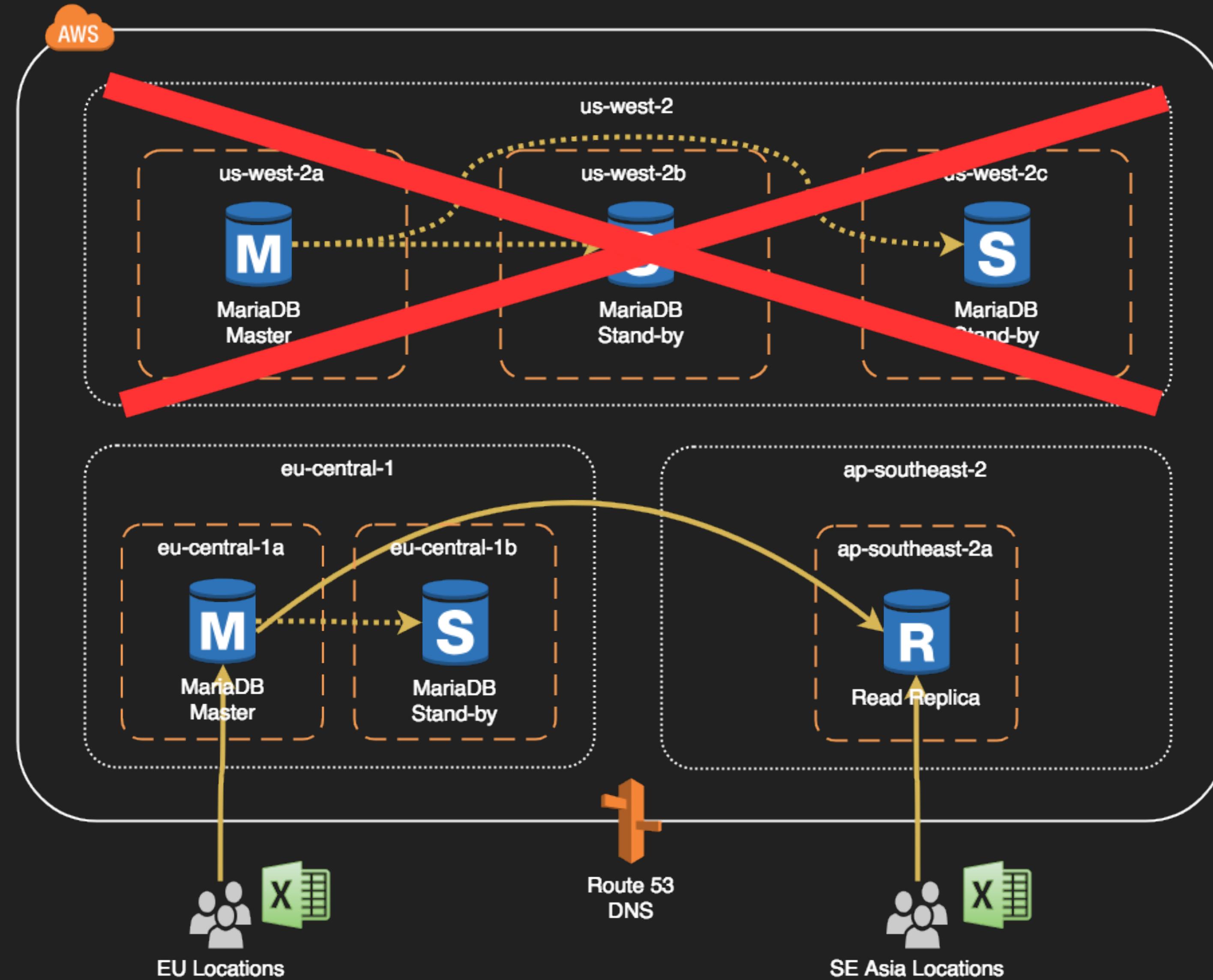
- One AZ fails
- Stand-by in another AZ assumes role of Master
- Read Replicas keep on keeping on



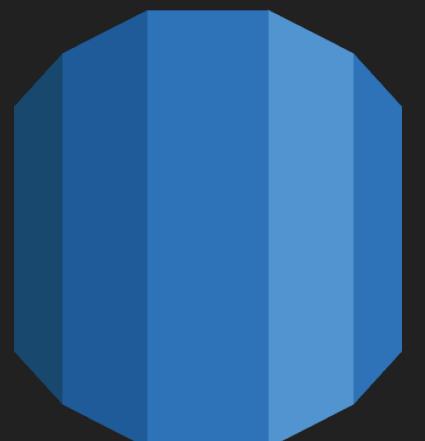
Amazon
RDS



Amazon RDS



- Whole region failed
- Read Replica promoted to Stand-Alone (Single-AZ)
- Single-AZ reconfigured to Multi-AZ



Amazon
RDS

A dark, low-light photograph of a person from the chest up. The person is wearing a dark button-down shirt over a blue collared shirt and a dark tie. They are holding a smartphone in their hands, which are positioned in the lower-left foreground. The background is dark and indistinct.

Amazon DynamoDB

Amazon DynamoDB



A CLOUD GURU

Dynamo: Amazon's Highly Available Key-value Store

Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati,
Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall
and Werner Vogels

Amazon.com

ABSTRACT

Reliability at massive scale is one of the biggest challenges we face at Amazon.com, one of the largest e-commerce operations in the world; even the slightest outage has significant financial consequences and impacts customer trust. The Amazon.com platform, which provides services for many web sites worldwide, is implemented on top of an infrastructure of tens of thousands of servers and network components located in many datacenters around the world. At this scale, small and large components fail continuously and the way persistent state is managed in the face of these failures drives the reliability and scalability of the software systems.

This paper presents the design and implementation of Dynamo, a highly available key-value storage system that some of Amazon's core services use to provide an "always-on" experience. To achieve this level of availability, Dynamo sacrifices consistency under certain failure scenarios. It makes extensive use of object versioning and application-assisted conflict resolution in a manner that provides a novel interface for developers to use.

Categories and Subject Descriptors

D.4.2 [Operating Systems]: Storage Management; D.4.5 [Operating Systems]: Reliability; D.4.2 [Operating Systems]: Performance;

General Terms

Algorithms, Management, Measurement, Performance, Design, Reliability

1. INTRODUCTION

Amazon runs a world-wide e-commerce platform that serves tens of millions customers at peak times using tens of thousands of servers located in many data centers around the world. There are

One of the lessons our organization has learned from operating Amazon's platform is that the reliability and scalability of a system is dependent on how its application state is managed. Amazon uses a highly decentralized, loosely coupled, service oriented architecture consisting of hundreds of services. In this environment there is a particular need for storage technologies that are always available. For example, customers should be able to view and add items to their shopping cart even if disks are failing, network routes are flapping, or data centers are being destroyed by tornados. Therefore, the service responsible for managing shopping carts requires that it can always write to and read from its data store, and that its data needs to be available across multiple data centers.

Dealing with failures in an infrastructure comprised of millions of components is our standard mode of operation; there are always a small but significant number of server and network components that are failing at any given time. As such Amazon's software systems need to be constructed in a manner that treats failure handling as the normal case without impacting availability or performance.

To meet the reliability and scaling needs, Amazon has developed a number of storage technologies, of which the Amazon Simple Storage Service (also available outside of Amazon and known as Amazon S3), is probably the best known. This paper presents the design and implementation of Dynamo, another highly available and scalable distributed data store built for Amazon's platform. Dynamo is used to manage the state of services that have very high reliability requirements and need tight control over the tradeoffs between availability, consistency, cost-effectiveness and performance. Amazon's platform has a very diverse set of applications with different storage requirements. A select set of applications requires a storage technology that is flexible enough to let application designers configure their data store appropriately.

Amazon DynamoDB



A CLOUD GURU

- Managed, multi-AZ NoSQL data store with Cross-Region Replication option
- Defaults to eventual consistency reads but can request strongly consistent read via SDK parameter
- Priced on throughput, rather than compute
- Provision read and write capacity in anticipation of need.
- Autoscale capacity adjusts per configured min/max levels.
- On-Demand Capacity for flexible capacity at a small premium cost.
- Achieve ACID compliance with DynamoDB Transactions

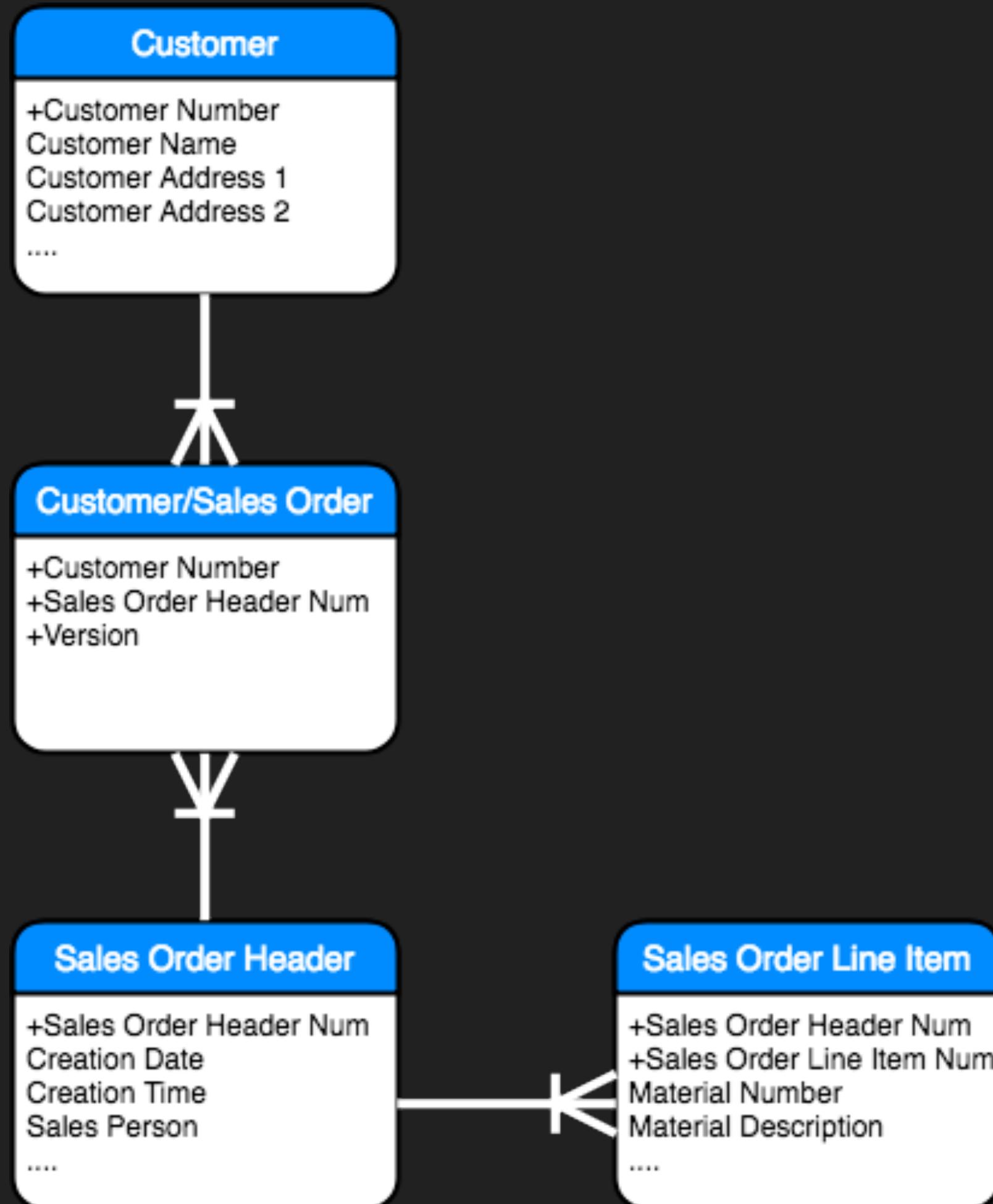


Amazon
DynamoDB

Relational vs NoSQL



A CLOUD GURU



```
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
```

```
"salesOrderNum" : "34234324",
"creationDate" : "3/1/2018",
"salesOrder" : {
    "salesOrderType" : "schedule agreement",
    "salesOrderLine" : [
        {
            "lineItem" : "1",
            "material" : {
                "materialNumber" : "HYDJF234",
                "materialDescription" : "Flange, 8cm, Iron"
            }
        },
        "customer" : {
            "customerNum" : "343535",
            "customerName" : "ABC Company",
            "customerAddress1" : "564 Main Street"
        }
    ]
}
```

Relational vs NoSQL



A CLOUD GURU

Name

```
1  [
2      "salesOrderNum" : "34234324",
3      "timestamp" : "2018-06-11T20:13:47Z",
4      "salesOrder" : {
5          "salesOrderType" : "schedule agreement",
6          "salesOrderLine" : [
7              {
8                  "lineItem" : "1",
9                  "material" : {
10                      "materialNumber" : "HYDJF234",
11                      "materialDescription" : "Flange, 8cm, Iron"
12                  }
13              }
14          ],
15          "customer" : {
16              "customerNum" : "343535",
17              "customerName" : "ABC Company",
18              "customerAddress1" : "564 Main Street"
19          }
20      }
21  ]
```

Value

Amazon DynamoDB



A CLOUD GURU

```
1  [
2      "salesOrderNum" : "34234324",
3      "timestamp" : "2018-06-11T20:13:47Z",
4      "salesOrder" : {
5          "salesOrderType" : "schedule agreement",
6          "salesOrderLine" : [
7              {
8                  "lineItem" : "1",
9                  "material" : {
10                      "materialNumber" : "HYDJF234",
11                      "materialDescription" : "Flange, 8cm, Iron"
12                  }
13              },
14          ],
15          "customer" : {
16              "customerNum" : "343535",
17              "customerName" : "ABC Company",
18              "customerAddress1" : "564 Main Street"
19          }
20      }
21  ]
```

Item

Amazon DynamoDB



A CLOUD GURU

```
1  {
2      "salesOrderNum" : "34234324",
3      "timestamp" : "2018-06-11T20:13:47Z",
4      "salesOrder" : {
5          "salesOrderType" : "schedule agreement", Attribute
6          "salesOrderLine" : [
7              {
8                  "lineItem" : "1",
9                  "material" : {
10                      "materialNumber" : "HYDJF234",
11                      "materialDescription" : "Flange, 8cm, Iron"
12                  }
13              },
14          ],
15          "customer" : {
16              "customerNum" : "343535",
17              "customerName" : "ABC Company",
18              "customerAddress1" : "564 Main Street"
19          }
20      }
21  }
```

Amazon DynamoDB



A CLOUD GURU

Table

```
1  [
2    "salesOrderNum" : "34234324",
3    "creationDate" : "3/1/2018",
4    "salesOrder" : {
5      "salesOrderType" : "schedule agreement",
6      1 [
7        "salesOrderNum" : "34234324",
8        "creationDate" : "3/1/2018",
9        "salesOrder" : {
10          5 [
11            "salesOrderType" : "schedule agreement",
12            1 [
13              "salesOrderNum" : "34234324",
14              "creationDate" : "3/1/2018",
15              "salesOrder" : {
16                10 [
17                  "salesOrderType" : "schedule agreement",
18                  2 [
19                    "salesOrderNum" : "34234324",
20                    "creationDate" : "3/1/2018",
21                    "salesOrder" : {
22                      15 [
23                        "salesOrderType" : "schedule agreement",
24                        "salesOrderLine" : [
25                          1 [
26                            "salesOrderNum" : "34234324",
27                            "creationDate" : "3/1/2018",
28                            "salesOrder" : {
29                              "salesOrderType" : "schedule agreement",
30                              "salesOrderLine" : [
31                                1 [
32                                  "lineItem" : "1",
33                                  "material" : {
34                                    "materialNumber" : "HYDGF234",
35                                    "materialDescription" : "Flange, 8cm, Iron"
36                                  }
37                                ]
38                              ]
39                            ]
40                          ]
41                        ]
42                      ]
43                    ]
44                  ]
45                ]
46              ]
47            ]
48          ]
49        ]
50      ]
51    ]
52  ]
```



Partition Key

```
1  [
2      "salesOrderNum" : "34234324", Primary Key
3      "timestamp" : "2018-06-11T20:13:47Z",
4      "salesOrder" : {
5          "salesOrderType" : "schedule agreement",
6          "salesOrderLine" : [
7              {
8                  "lineItem" : "1",
9                  "material" : {
10                      "materialNumber" : "HYDJF234",
11                      "materialDescription" : "Flange, 8cm, Iron"
12                  }
13              }
14          ],
15          "customer" : {
16              "customerNum" : "343535",
17              "customerName" : "ABC Company",
18              "customerAddress1" : "564 Main Street"
19          }
20      }
21  ]
```

In this case, we're choosing a simple primary key known as the partition key.

IT MUST BE UNIQUE!

To access, you only need to know the **salesOrderNum**.

Partition Key and Sort Key



A CLOUD GURU

```
1  [
2      "salesOrderNum" : "34234324",
3      "timestamp" : "2018-06-11T20:13:47Z", Primary Key
4      "salesOrder" : {
5          "salesOrderType" : "schedule agreement",
6          "salesOrderLine" : [
7              {
8                  "lineItem" : "1",
9                  "material" : {
10                      "materialNumber" : "HYDJF234",
11                      "materialDescription" : "Flange, 8cm, Iron"
12                  }
13              }
14          ],
15          "customer" : {
16              "customerNum" : "343535",
17              "customerName" : "ABC Company",
18              "customerAddress1" : "564 Main Street"
19          }
20      }
21  ]
```

In this case, we're choosing a composite primary key known as a partition key and sort key.

We can have occurrences of the same Partition Key so long as the sort key is different.



Secondary Indexes

Index Type	Description	How to Remember
Global Secondary Index	Partition key and sort key can be different from those on the table.	I'm not restricted to just the partitioning set forth by the partition key. I'm GLOBAL BABY!
Local Secondary Index	Same partition key as the table but different sort key	I have to stay local and respect the table's partition key, but I can choose whatever sort key I want.

- There is a limit to the number of indexes and attributes per index
- Indexes take up storage space



Secondary Indexes

Index Type	When to Use	Example
Global Secondary Index	<p>When you want a fast query of attributes outside the primary key – without having to do a table scan (read everything sequentially)</p>	<p>“I’d like to query Sales Orders by Customer number rather than Sales Order Number.”</p>
Local Secondary Index	<p>When you already know the partition key and want to quickly query on some other attribute</p>	<p>“I have the Sales Order Number, but I’d like to retrieve only those records with a certain Material Number.”</p>

Global Secondary Index



A CLOUD GURU

```
1  [
2    "salesOrderNum" : "34234324",
3    "timestamp" : "2018-06-11T20:13:47Z",
4    "salesOrder" : {
5      "salesOrderType" : "schedule agreement",
6      "salesOrderLine" : [
7        {
8          "lineItem" : "1",
9          "material" : {
10            "materialNumber" : "HYDJF234",
11            "materialDescription" : "Flange, 8cm, Iron"
12          }
13        },
14      ],
15      "customer" : {
16        "customerNum" : "343535", // Highlighted field
17        "customerName" : "ABC Company",
18        "customerAddress1" : "564 Main Street"
19      }
20    }
21  ]
```

Suppose we created a Global Secondary Index using `customerNum`.

We could query by Customer Number at light-speed.



Local Secondary Index

```
1  [
2      "salesOrderNum" : "34234324",
3      "timestamp" : "2018-06-11T20:13:47Z",
4      "salesOrder" : {
5          "salesOrderType" : "schedule agreement",
6          "salesOrderLine" : [
7              {
8                  "lineItem" : "1",
9                  "material" : {
10                      "materialNumber" : "HYDJF234",
11                      "materialDescription" : "Flange, 8cm, Iron"
12                  }
13              },
14          ],
15          "customer" : {
16              "customerNum" : "343535",
17              "customerName" : "ABC Company",
18              "customerAddress1" : "564 Main Street"
19          }
20      }
21  ]
```

Suppose we created a Local Secondary Index using **materialNumber**.

We could query by Sales Order Number and Material Number at light-speed.



Attribute Projections

```
1
2   "salesOrderNum" : "34234324",
3   "timestamp" : "2018-06-11T20:13:47Z",
4   "salesOrder" : {
5     "salesOrderType" : "schedule agreement",
6     "salesOrderLine" : [
7       {
8         "lineItem" : "1",
9         "material" : {
10           "materialNumber" : "HYDJF234",
11           "materialDescription" : "Flange, 8cm, Iron"
12         }
13       },
14     ],
15     "customer" : {
16       "customerNum" : "343535",
17       "customerName" : "ABC Company",
18       "customerAddress1" : "564 Main Street"
19     }
20   }
21 }
```

Projection

customerNum (Key)
customerName
customerAddress1
salesOrderNum
timestamp



Secondary Indexes

If you need to...	Consider...	Cost	Benefit
...access just a few attributes the fastest way possible	Projecting just those few attributes in a global secondary index	Minimal	Lowest possible latency access for non-key items
...frequently access some non-key attributes	Projecting those attributes in a global secondary index	Moderate; aim to offset cost of table scans	Lowest possible latency access for non-key items
...frequently access most non-key attributes	Projecting those attributes or even the entire table in a global secondary index	Up to Double	Maximum flexibility
...rarely query but write or update frequently	Projecting keys only for the global secondary index	Minimal	Very fast write or updates for non-partition-key items



Design Best Practices

```
1  {
2      "CustomerId" : "123",
3      "SortKey" : "Details",
4      "FirstName" : "Pumpkin",
5      "Surname" : "Escobar",
6      "Contact" : "pumpkin.escobar@gmail.com"
7  }
8  {
9      "CustomerId" : "123",
10     "SortKey" : "Purchases-2019-03",
11     "Period" : "2019-03",
12     "TotalPurchases" : "45000.00",
13     "Currency" : "USD"
14 },
15 [
16     "CustomerId" : "123",
17     "SortKey" : "order-45670",
18     "IsOpen" : "True",
19     "TotalValue" : "2300.00",
20     "Currency" : "USD"
21 ]
```

Design Best Practices



A CLOUD GURU

```
1  {
2    "CustomerId" : "123",
3    "SortKey" : "Details",
4    "FirstName" : "Pumpkin",
5    "Surname" : "Escobar",
6    "Contact" : "pumpkin.escobar@gmail.com"
7  }
8  {
9    "CustomerId" : "123",
10   "SortKey" : "Purchases-2019-03",
11   "Period" : "2019-03",
12   "TotalPurchases" : "45000.00",
13   "Currency" : "USD"
14 },
15 [
16   "CustomerId" : "123",
17   "SortKey" : "order-45670",
18   "isOpen" : "True",
19   "TotalValue" : "2300.00",
20   "Currency" : "USD"
21 ]
```

Design Best Practices



A CLOUD GURU

Partition Key	Sort Key	Attribute1	Attribute2
"CustomerId": "123"	"SortKey" : "Details"	"FirstName" : "Pumpkin"	"Surname" : "Escobar"
	"SortKey" : "Purchases-2019-03"	"Period" : "2019-03"	"TotalPurchases" : "45000.00"
	"SortKey" : "order-45670"	"IsOpen" : "True"	"TotalValue" : "2300.00"

Design Best Practices



A CLOUD GURU

Partition Key	Sort Key	Attribute1	Attribute2
"CustomerId": "123"	"SortKey" : "Details"	"FirstName" : "Pumpkin"	"Surname" : "Escobar"
	"SortKey" : "Purchases-2019-03"	"Period" : "2019-03"	"TotalPurchases" : "45000.00"
	"SortKey" : "order-45670"	"IsOpen" : "True"	"TotalValue" : "2300.00"

Aggregations



A CLOUD GURU

Partition Key	Sort Key	Attribute1	Attribute2
"CustomerId": "123"	"SortKey" : "Details"	"FirstName" : "Pumpkin"	"Surname" : "Escobar"
	"SortKey" : "Purchases-2019-03"	"Period" : "2019-03" Global Secondary Index	"TotalPurchases" : "45000.00"
	"SortKey" : "order-45670"	"IsOpen" : "True"	"TotalValue" : "2300.00"



Sparse Indexes

Partition Key	Sort Key	Attribute1	Attribute2
		“Period” : “2019-03”	“TotalPurchases” : “45000.00”
		“Period” : “2019-03”	“TotalPurchases” : “23000.00”
		“Period” : “2019-03”	“TotalPurchases” : “5600.00”



Sparse Indexes

Partition Key	Sort Key	Attribute1	Attribute2
		“Period” : “2019-03”	“TotalPurchases” : “45000.00”
		“Period” : “2019-03”	“TotalPurchases” : “45000.00”
		“Period” : “2019-03”	“TotalPurchases” : “45000.00”

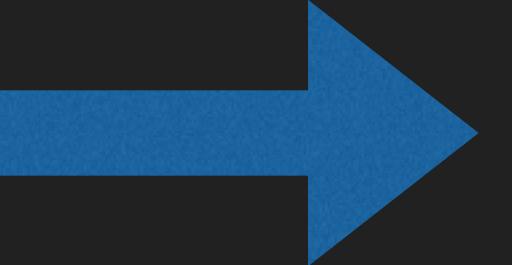
Replicas via Secondary Indexes



A CLOUD GURU

DynamoDB Table

CustID	Seq	Attrb1	Attrb2



Global Secondary Index

CustID	Seq	Attrb1	Attrb2

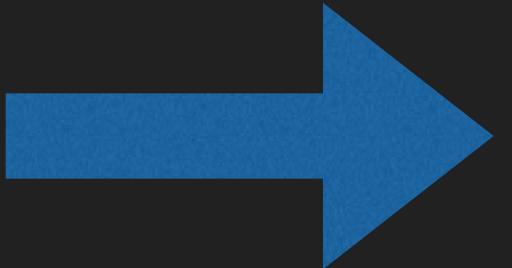
Replicas via Secondary Indexes



A CLOUD GURU

DynamoDB Table

CustID	Seq	Attrb1	Attrb2



Global Secondary Index

CustID	Seq	Attrb1	Attrb2

Higher RCU/WCU Limits



Premium Customers

Lower RCU/WCU Limits



Free-Tier Customers

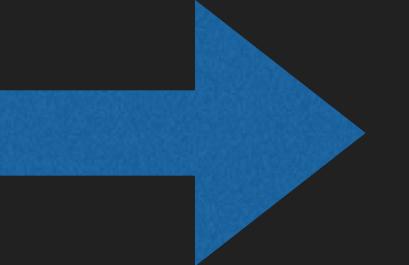
Replicas via Secondary Indexes



A CLOUD GURU

DynamoDB Table

CustID	Seq	Attrb1	Attrb2



Global Secondary Index

CustID	Seq	Attrb1	Attrb2

High Write Capacity Limits



Writes Only from Data Stream

High Read Capacity Limits

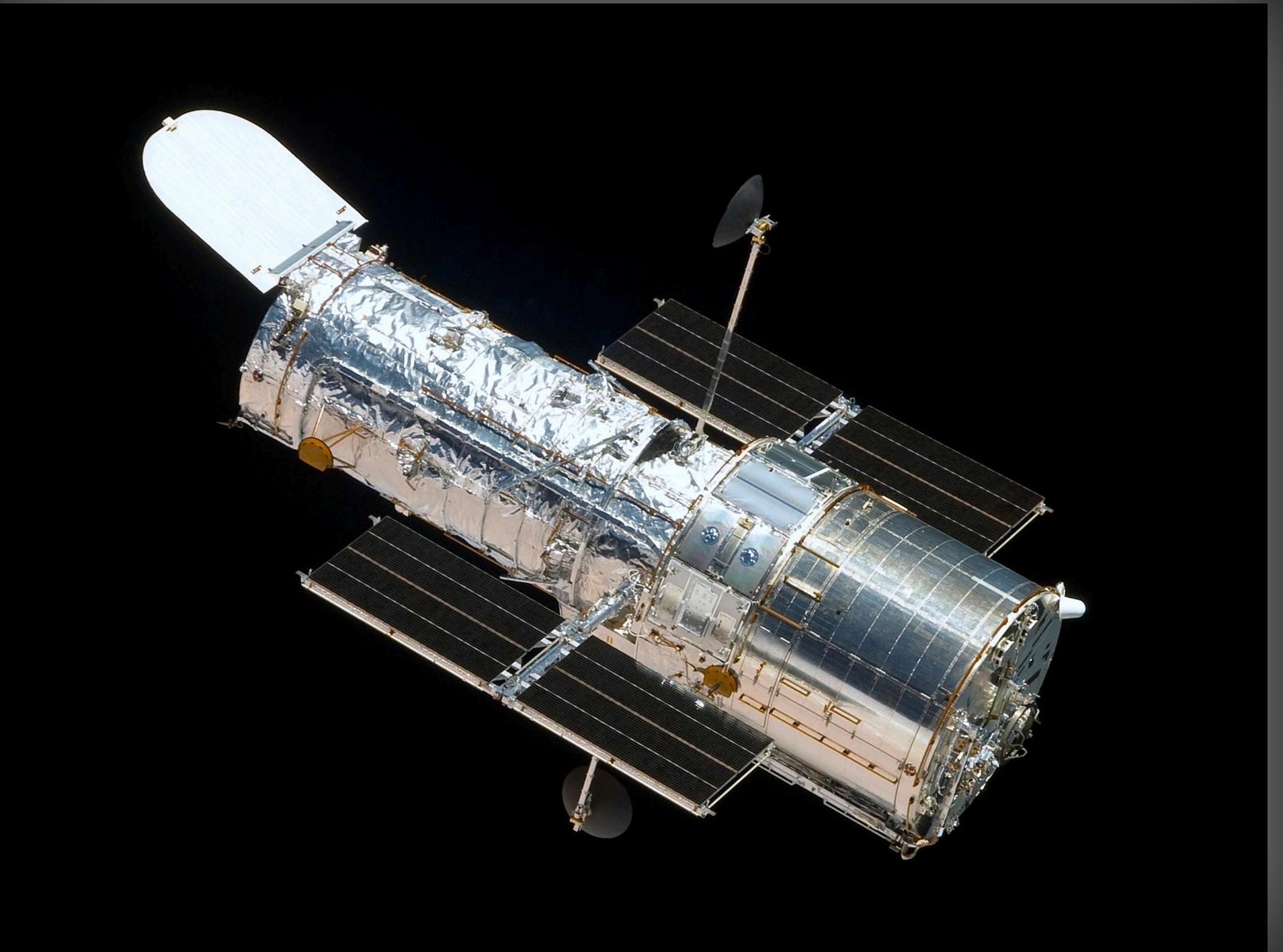
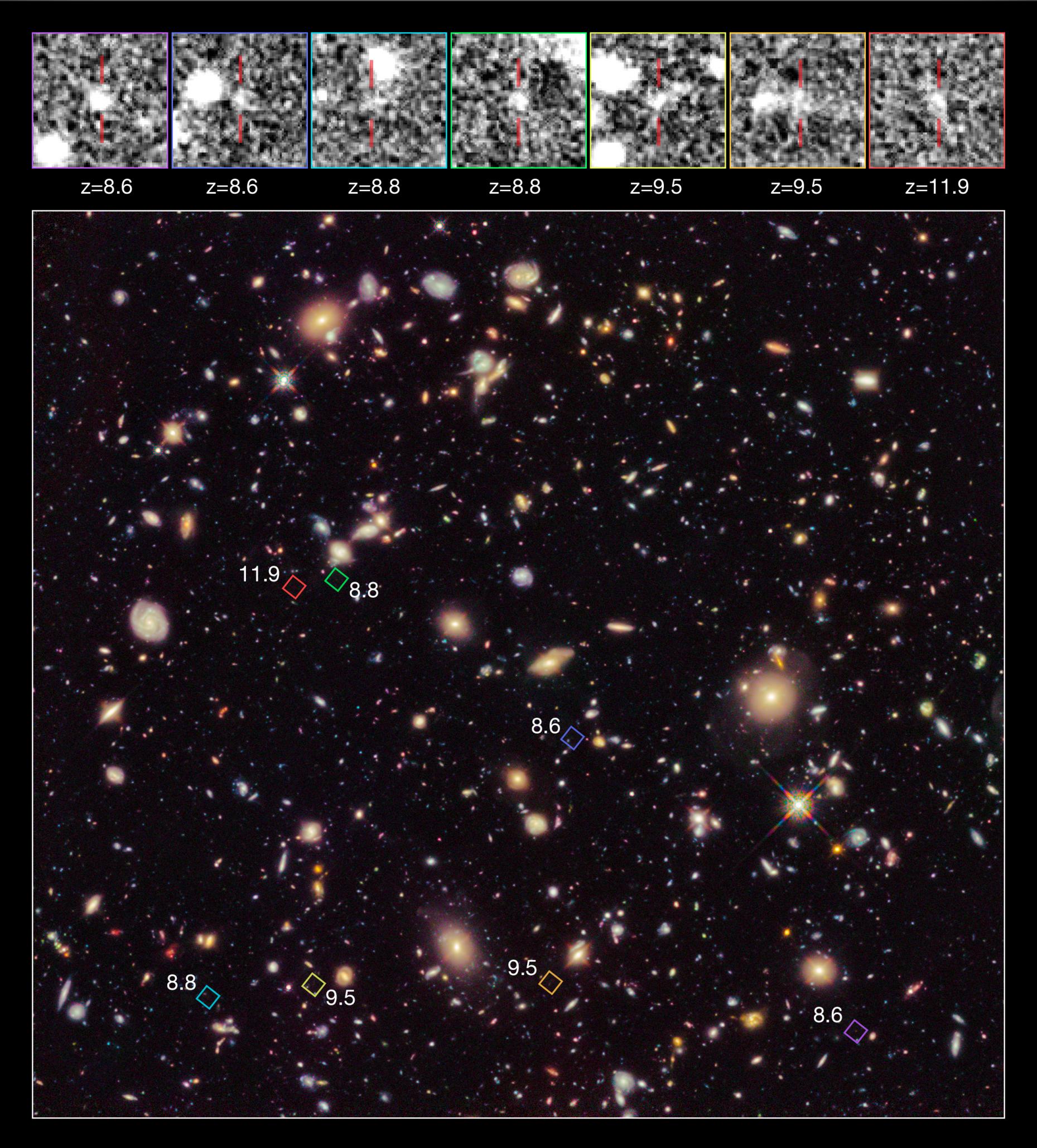


Read Only for Analytics

A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, and they appear to be holding a small object, likely a smartphone, which is partially visible at the bottom left.

Amazon Redshift

Amazon Redshift



Amazon Redshift



A CLOUD GURU

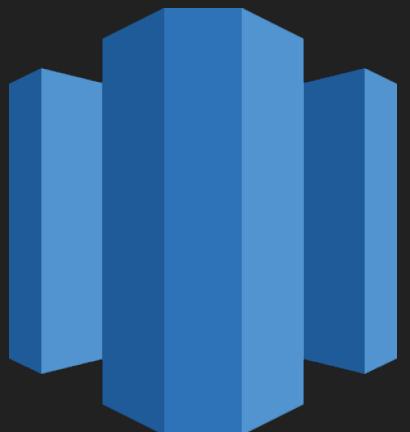
ORACLE®

Amazon Redshift



A CLOUD GURU

- Fully managed, clustered peta-byte scale data warehouse
- Extremely cost-effective as compared to some other on-premises data warehouse platforms
- PostgreSQL compatible with JDBC and ODBC drivers available; compatible with most BI tools out of the box
- Features parallel processing and columnar data stores which are optimized for complex queries
- Option to query directly from data files on S3 via Redshift Spectrum



Amazon
Redshift

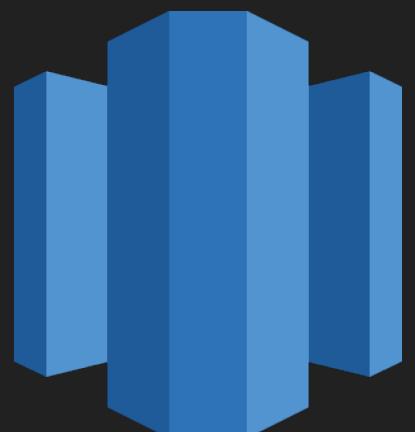
Data Lake



A CLOUD GURU



- Query raw data without extensive pre-processing
- Lessen time from data collection to data value
- Identify correlations between disparate data sets

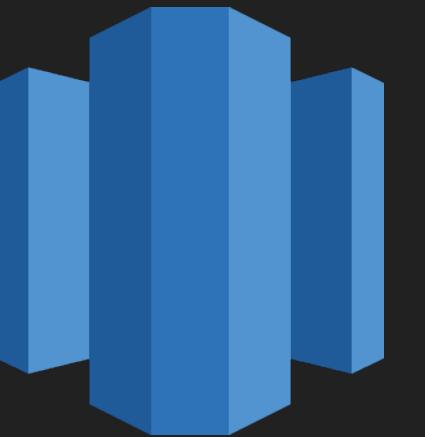


Amazon
Redshift

Data Lake



A CLOUD GURU



Amazon
Redshift
Spectrum



Analytics
Tools

A dark, low-light photograph showing a person from the chest up. They are wearing a dark button-down shirt over a dark tie. Their hands are clasped in front of them, holding a small, indistinct object. The background is dark and out of focus.

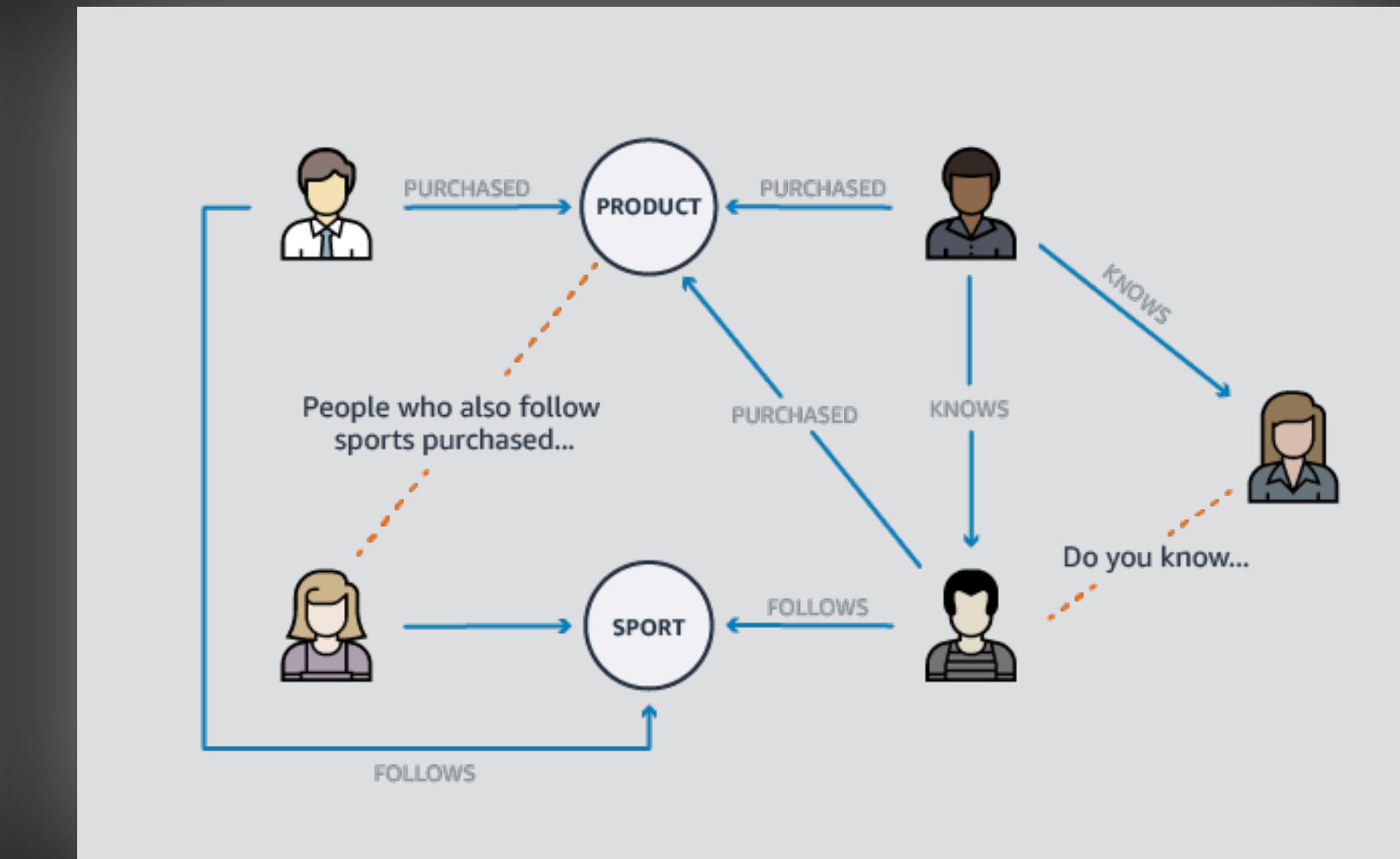
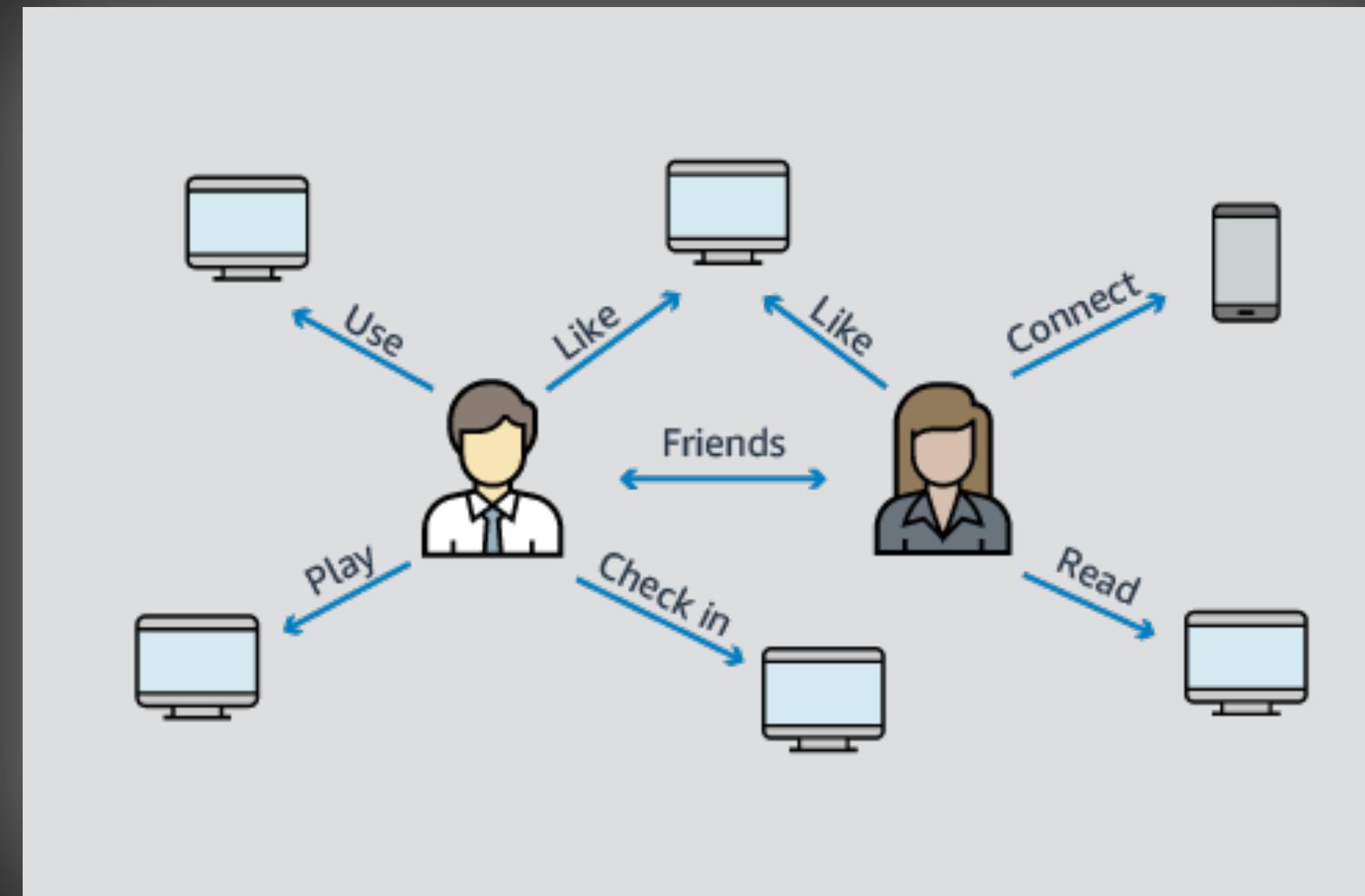
Amazon Neptune

Amazon Neptune



A CLOUD GURU

- Fully-managed graph database
- Supports open graph APIs for both Gremlin and SPARQL



<https://aws.amazon.com/neptune/>

A dark, low-light photograph of a person from the chest up. The person is wearing a dark button-down shirt over a blue collared shirt and a dark tie. They are holding a smartphone in their hands, which are positioned in the lower-left foreground. The background is dark and indistinct.

Amazon ElastiCache

Amazon ElastiCache



A CLOUD GURU

Consonants

b c d f g h j k l m n
b k d f g h dʒ k l m n
bib cake dad fife gag hat judge kick lull mime noon

ŋ p r s ʂ t v w y z ʒ
ŋ p r s z t v w j z ʒ
sing pipe roar sauce is tot valve will yes zoo vision

Joined consonants

ch sh th wh a e i o u ʊ
tʃ ʃ θ ð ʍ æ ɛ ɪ ə ʌ ʊ
church shush thin then whale at egg in odd up book

Long vowels / diphthongs

aæauɛœœɔœœuœœiœœoœœu
ɑ: eɪ ɔ: ɪ ʊ ɔ: u: ju: aɪ ɔɪ aʊ
father ape all eat oak ooze use ice oil owl

Also, **A** is used following a vowel letter to write the sound in “earn” etc

Amazon ElastiCache



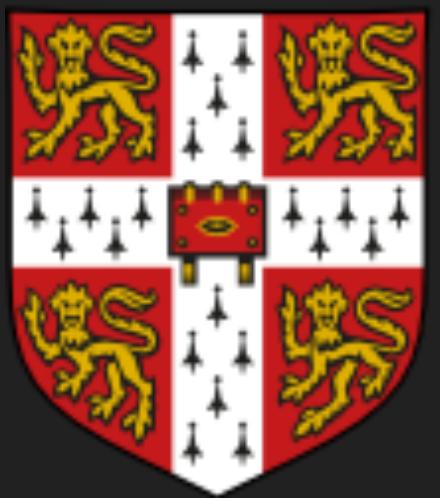
A CLOUD GURU

cache

Amazon ElastiCache

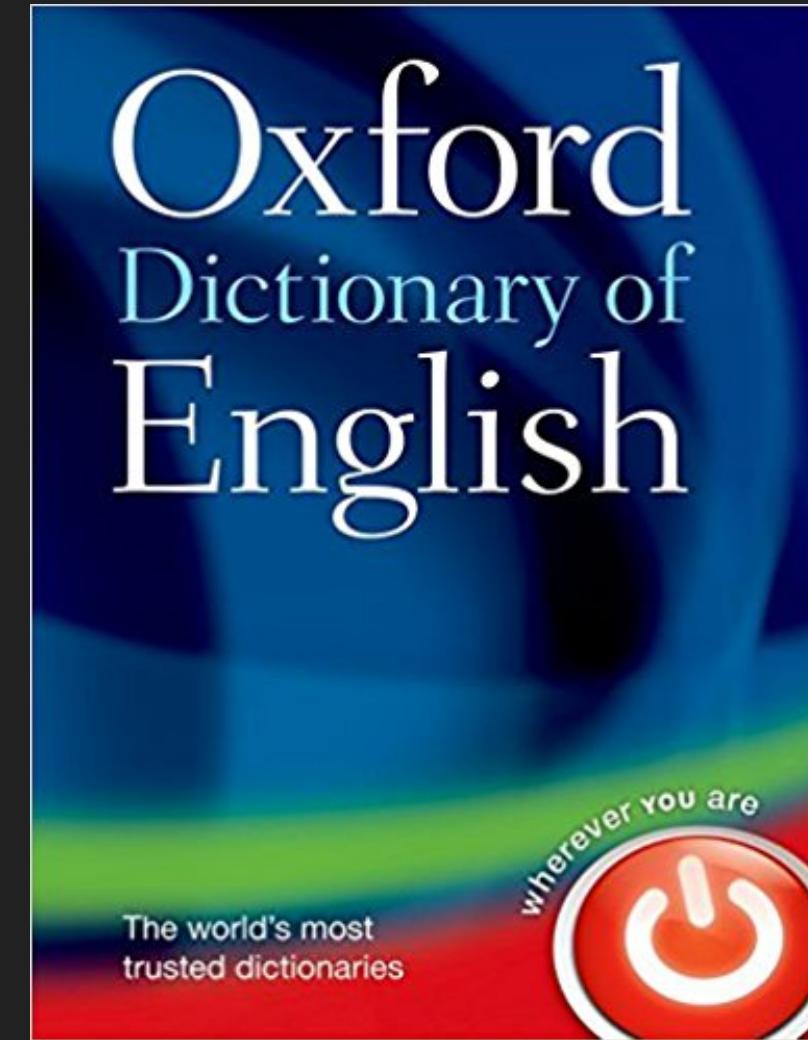


A CLOUD GURU



Cambridge
Dictionary

kæʃ



Amazon ElastiCache



A CLOUD GURU

kæʃ

k = “kick”

æ = “at”

ʃ = “hush”

Amazon ElastiCache



A CLOUD GURU

Insert
Clip of
Ryan
Here

Amazon ElastiCache



A CLOUD GURU

- Fully managed implementations of two popular in-memory data stores — Redis and Memcached
- Push-button scalability for memory, writes and reads
- In Memory key/value store—not persistent in the traditional sense...
- Billed by node size and hours of use



Amazon
ElastiCache

Amazon ElastiCache



A CLOUD GURU

Use	Benefit
Web Session Store	In cases with load-balanced web servers, store web session information in Redis so if a server is lost, the session info is not lost and another web server can pick-up.
Database Caching	Use Memcache in front of AWS RDS to cache popular queries to offload work from RDS and return results faster to users.
Leaderboards	Use Redis to provide a live leaderboard for millions of users of your mobile app.
Streaming Data Dashboards	Provide a landing spot for streaming sensor data on the factory floor, providing live real-time dashboard displays.



Amazon
ElastiCache

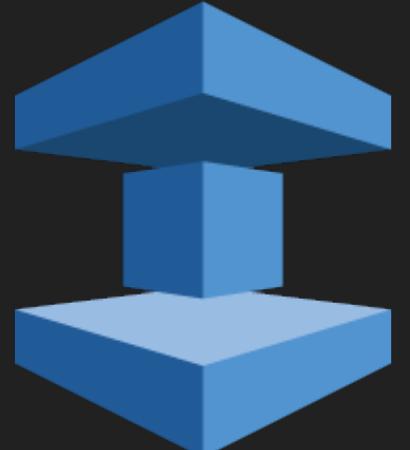
Amazon ElastiCache



A CLOUD GURU

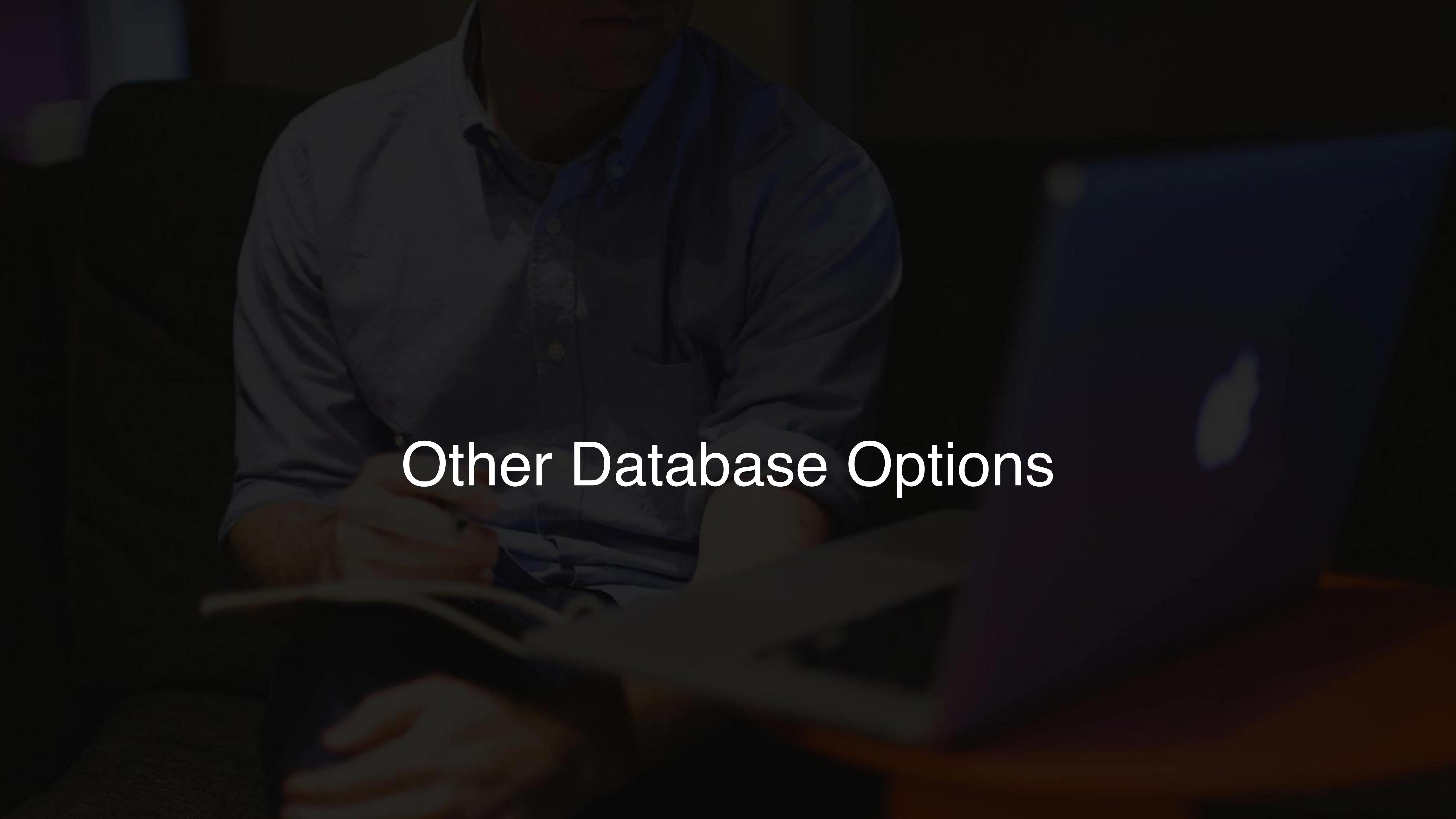


- | | |
|--|---|
| <ul style="list-style-type: none">• Simple, no-frills, straight-forward• You need to scale out and in as demand changes• You need to run multiple CPU cores and threads• You need to cache objects (i.e. like database queries) | <ul style="list-style-type: none">• You need encryption• You need HIPAA compliance• Support for clustering• You need complex data types• You need high-availability (replication)• Pub/Sub capability• Geospatial Indexing• Backup and Restore |
|--|---|



A Cache is a Cache...use the right tool for the job.

Amazon
ElastiCache

A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a smartphone horizontally. The screen of the phone is visible but appears mostly black or very dimly lit.

Other Database Options



- SQL Engine overlaid on S3 base on Presto
- Query raw data objects as they sit in an S3 bucket
- Use or convert your data to Parquet format if possible for a big performance jump
- Similar in concept to Redshift Spectrum but...

Use Amazon Athena	Use Redshift Spectrum
Data lives mostly on S3 without the need to perform joins with other data sources	Want to join S3 data with existing RedShift tables or create union products



Amazon
Athena

Amazon Quantum Ledger Database



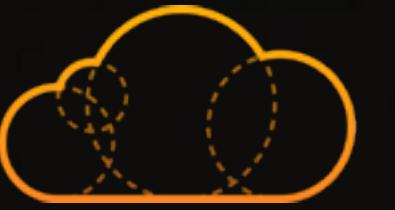
A CLOUD GURU

- Based on blockchain concepts
- Provides an immutable and transparent journal as a service without having to setup and maintain an entire blockchain framework
- Centralized design (as opposed to decentralized consensus-based design for common blockchain frameworks) allows for higher performance and scalability.
- Append-only concept where each record contributes to the integrity of the chain.

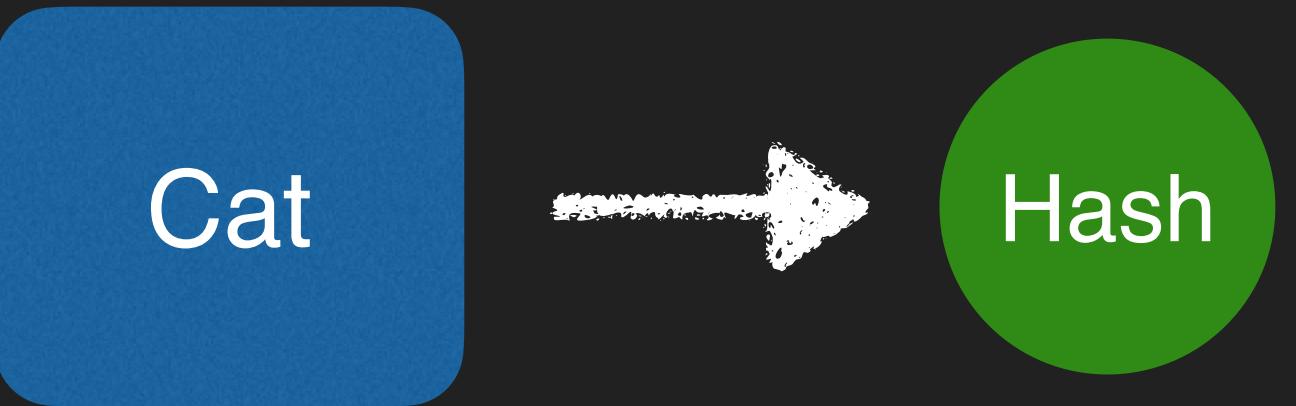


Amazon
QLDB

Amazon Quantum Ledger Database



A CLOUD GURU

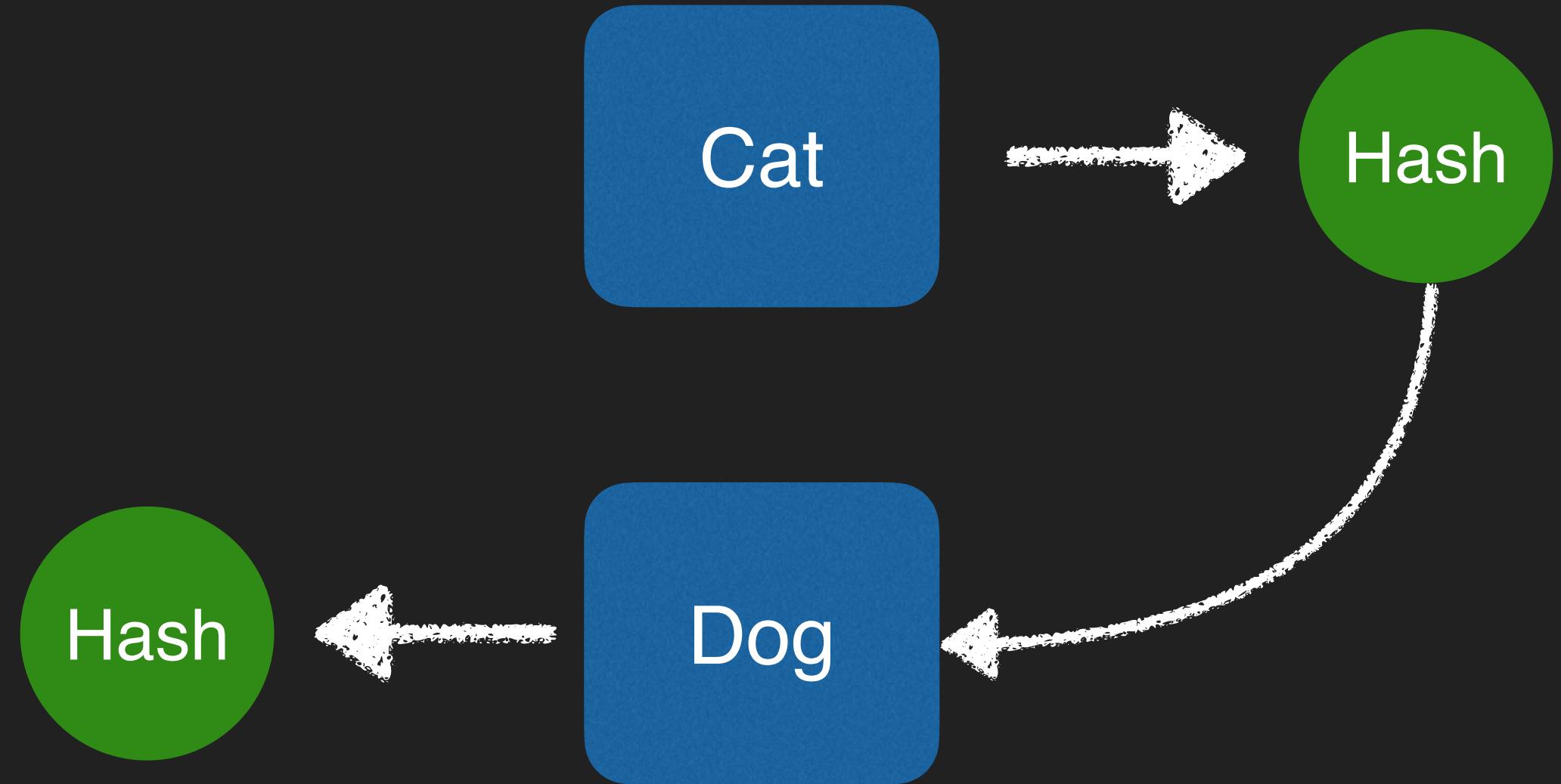


Amazon
QLDB

Amazon Quantum Ledger Database



A CLOUD GURU

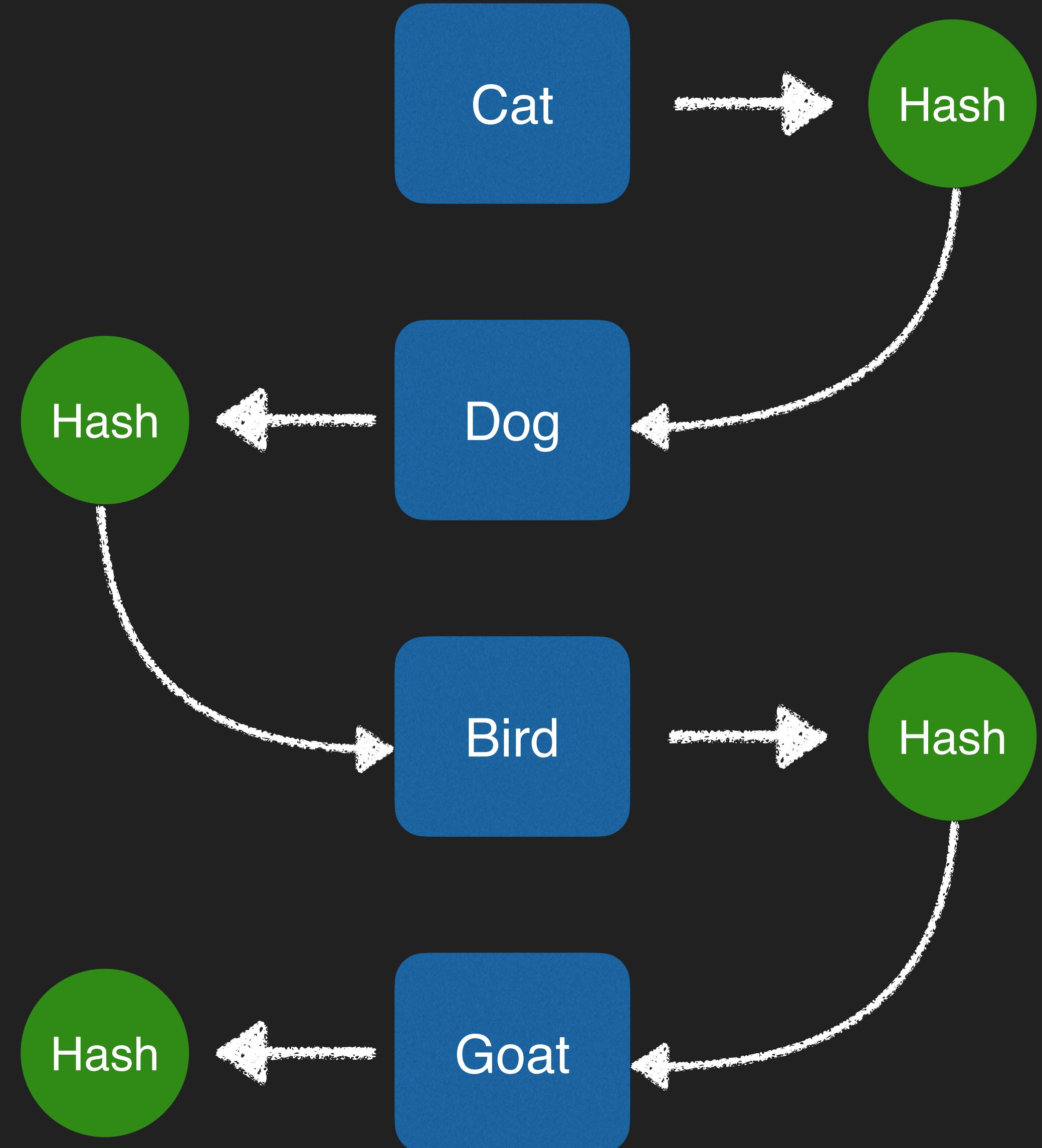


Amazon
QLDB

Amazon Quantum Ledger Database



A CLOUD GURU

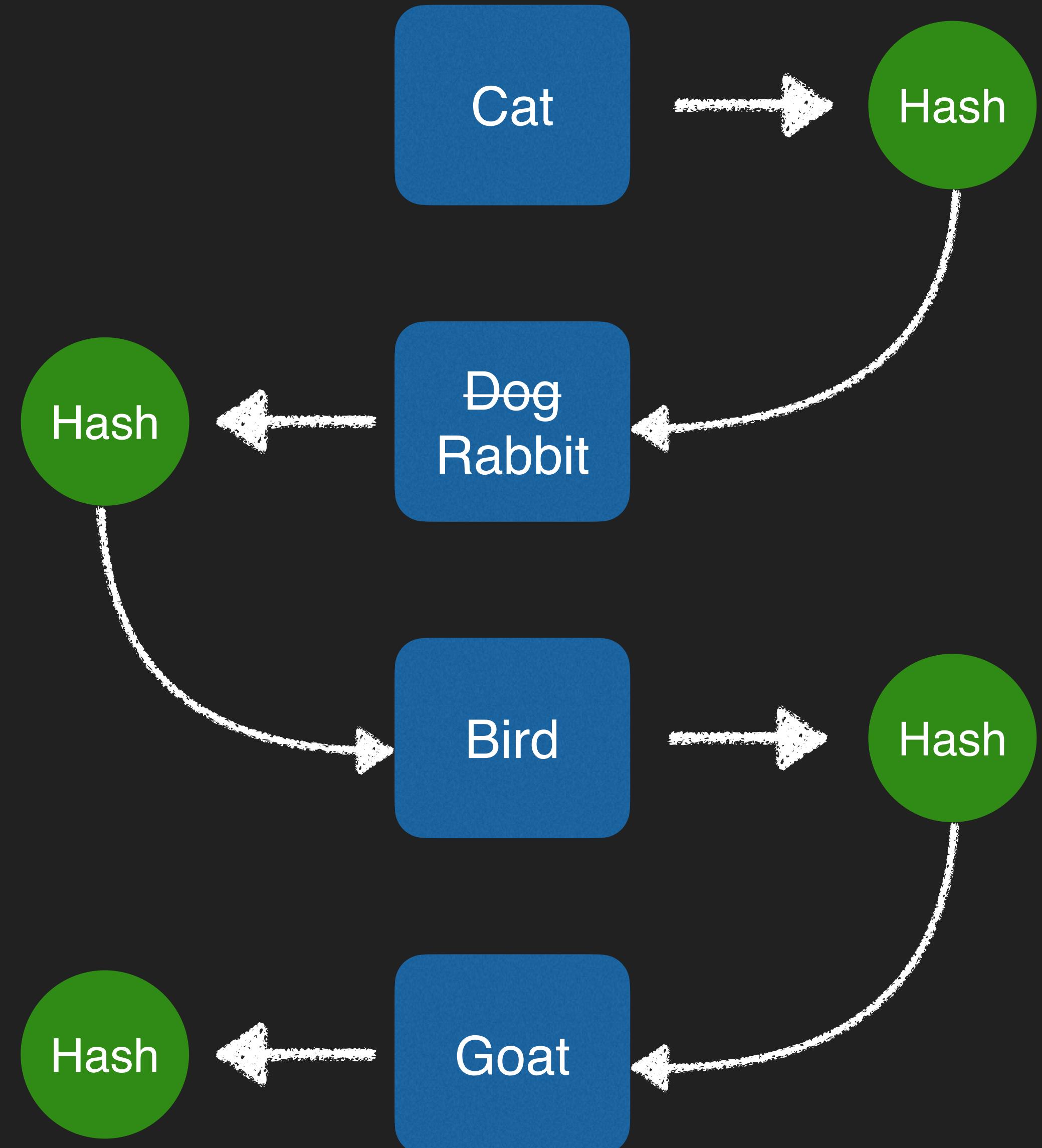


Amazon
QLDB

Amazon Quantum Ledger Database



A CLOUD GURU

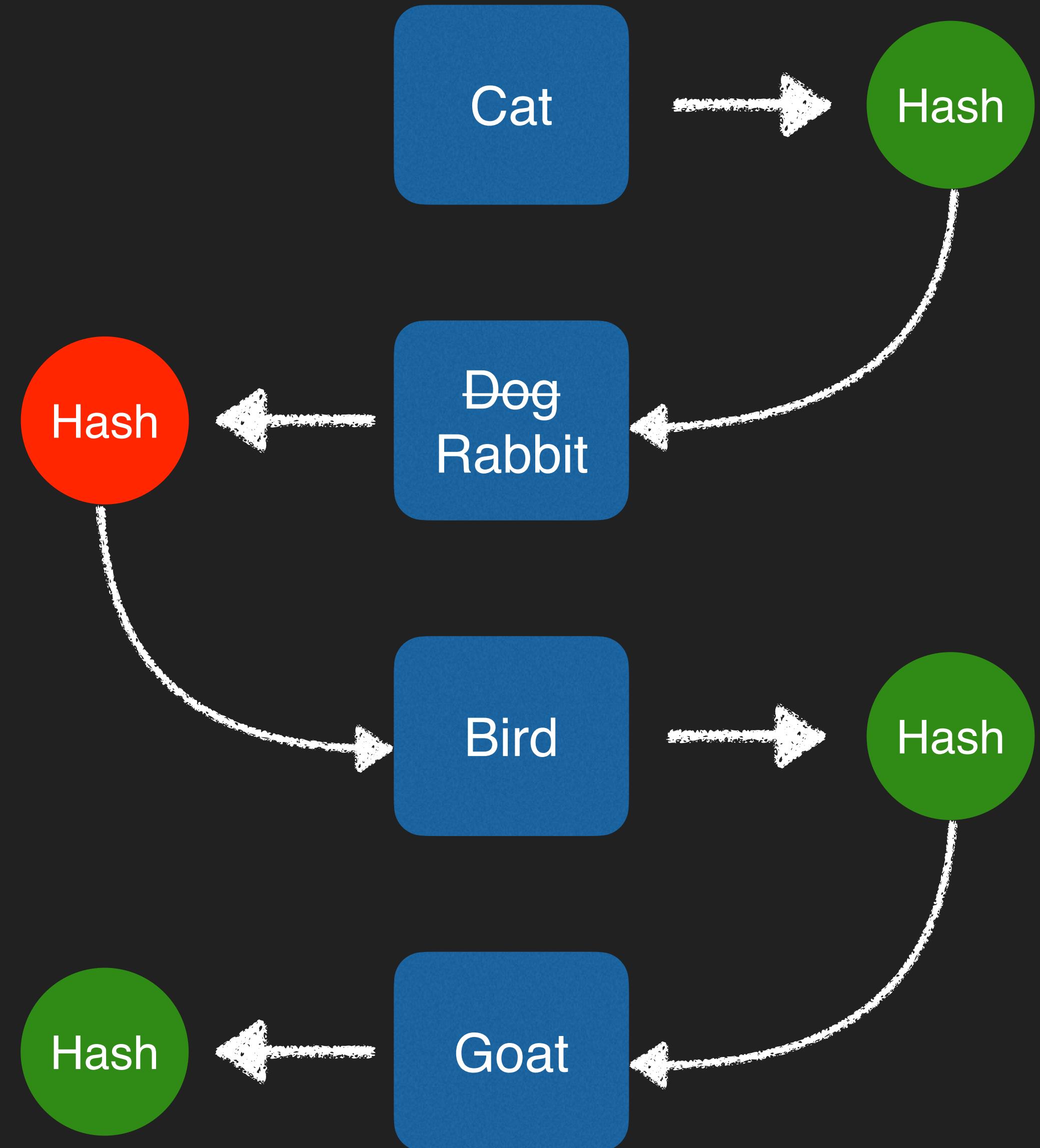


Amazon
QLDB

Amazon Quantum Ledger Database



A CLOUD GURU

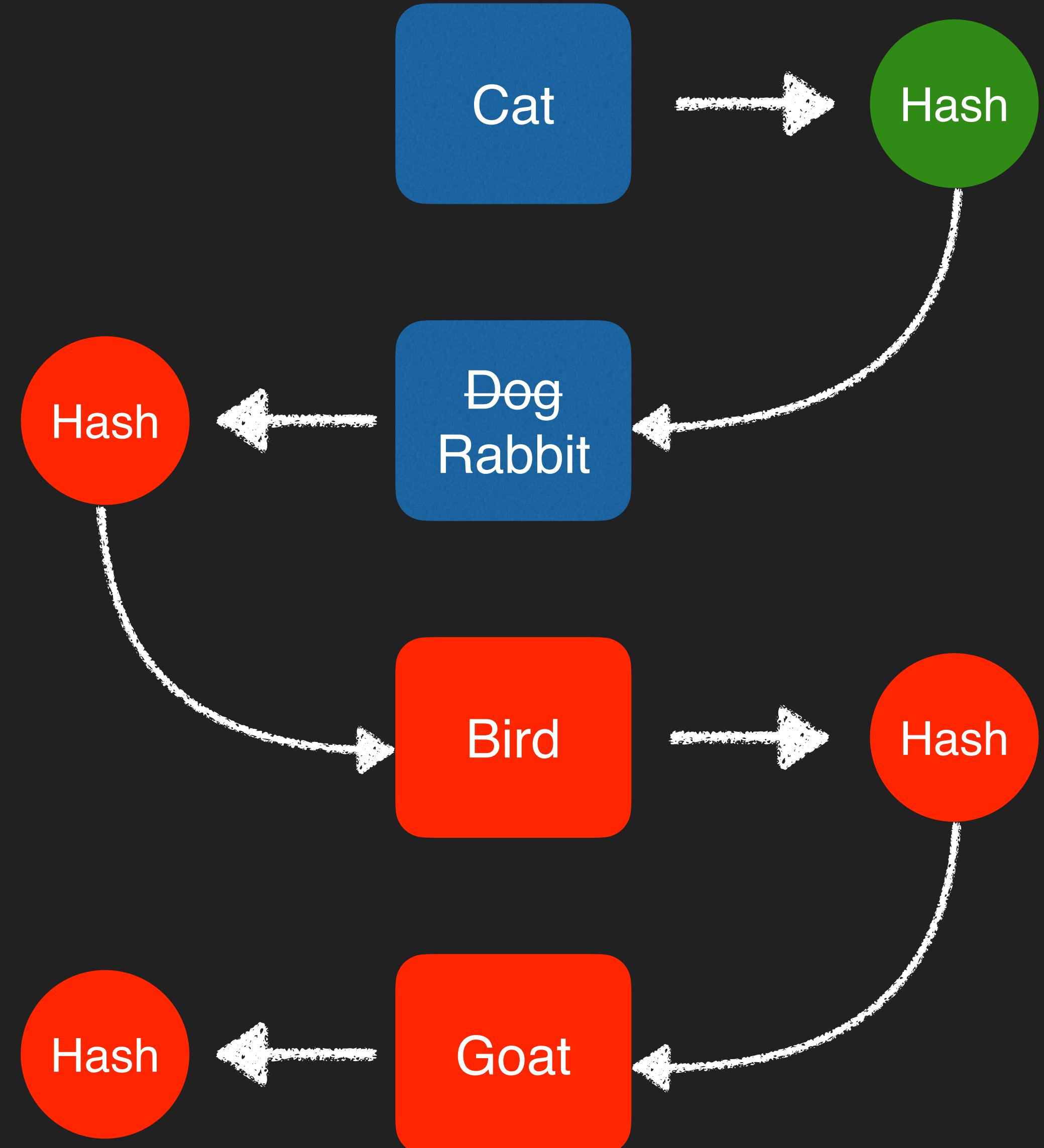


Amazon
QLDB

Amazon Quantum Ledger Database



A CLOUD GURU



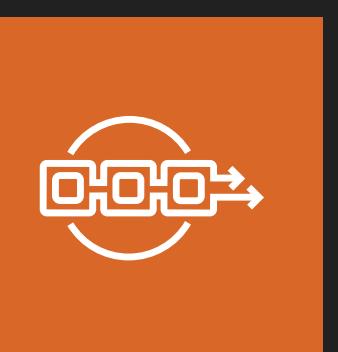
Amazon
QLDB

Amazon Managed Blockchain



A CLOUD GURU

- Fully managed blockchain framework supporting open source frameworks of Hyperledger Fabric and Ethereum
- Distributed consensus-based concept consisting of a network, members (other AWS accounts), nodes (instances) and potentially applications.
- Uses the Amazon QLDB ordering service to maintain complete history of all transactions.



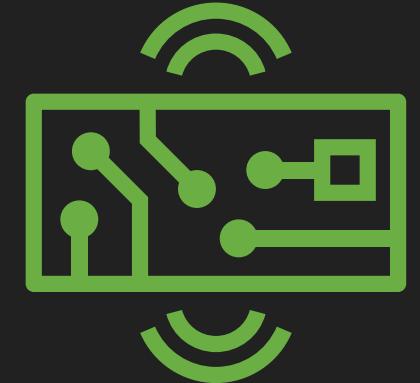
Amazon
Managed
Blockchain

Amazon Timestream Database



A CLOUD GURU

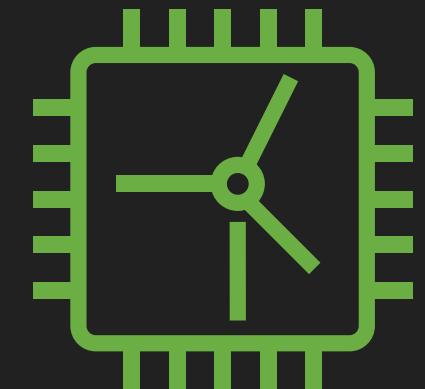
- Fully managed database service specifically built for storing and analyzing time-series data
- Alternative to DynamoDB or RedShift and includes some built-in analytics like interpolation and smoothing



Industrial
Machinery



Sensor
Networks



Equipment
Telemetry



Amazon
Timestream

Amazon DocumentDB



A CLOUD GURU

- “with MongoDB compatibility”
- AWS’s invention that emulates the MongoDB API so it acts like MongoDB to existing clients and drivers.
- Fully managed with all the good stuff (multi-AZ HA, scalable, integrated with KMS, backed up to S3)
- An option if you currently use MongoDB and want to get out of the server management business.



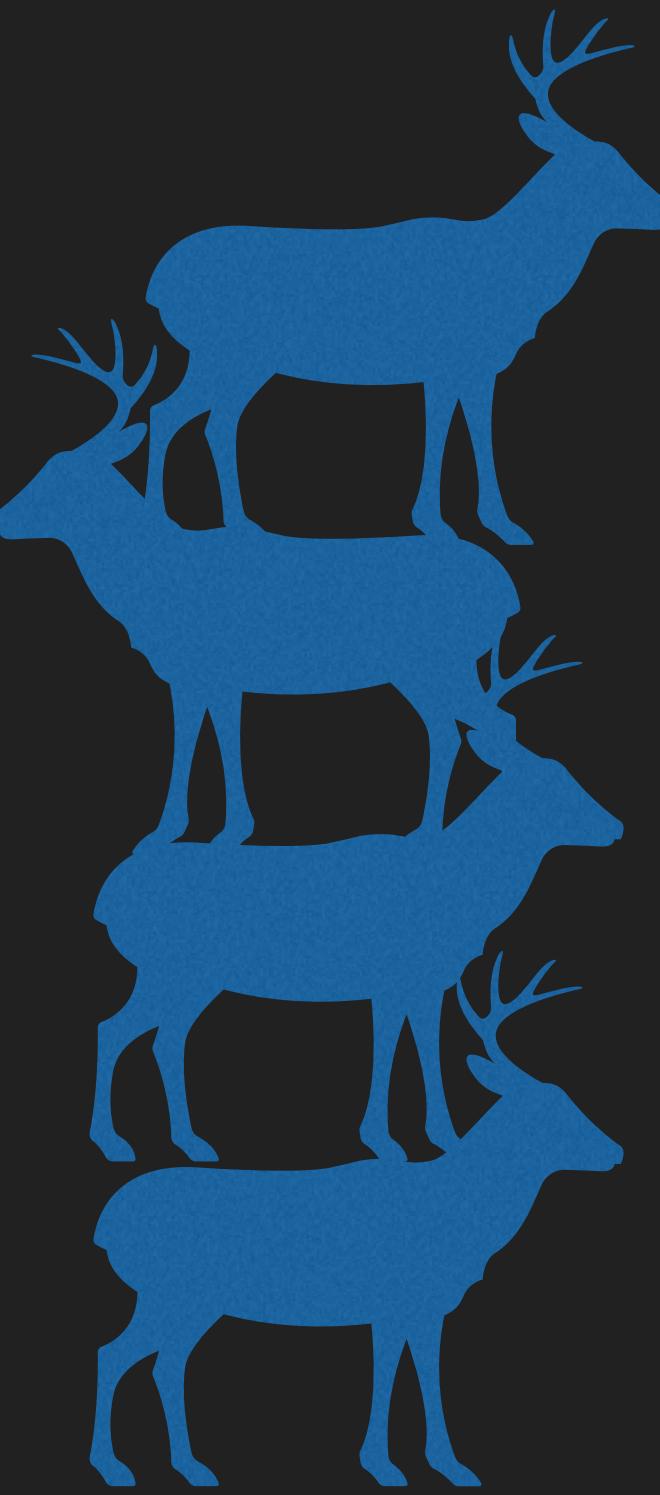
Amazon
DocumentDB

Amazon ElasticSearch



A CLOUD GURU

- Not to be confused with ElastiCache
- Mostly a search engine but also a document store (caution here)
- Amazon ElasticSearch Service components are sometimes referred to as an ELK stack.



Amazon ElasticSearch

- Not to be confused with ElastiCache
- Mostly a search engine but also a document store (caution here)
- Amazon ElasticSearch Service components are sometimes referred to as an ELK stack.

Analytics

Kibana

Intake

LogStash

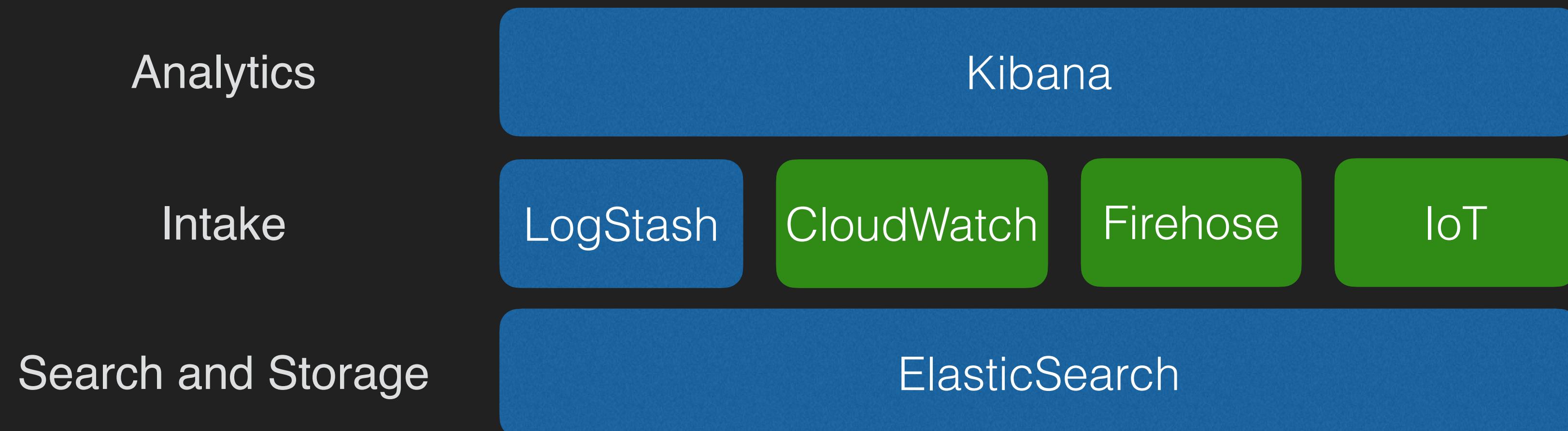
Search and Storage

ElasticSearch



Amazon ElasticSearch

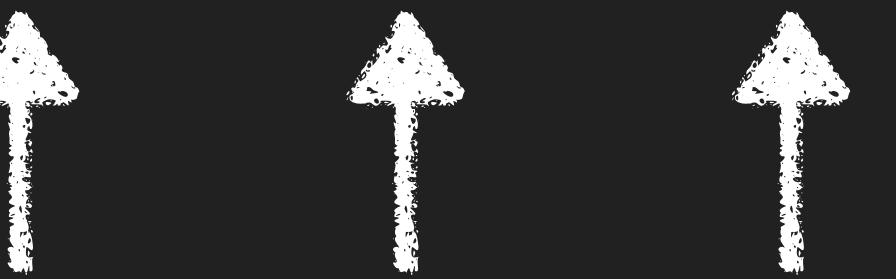
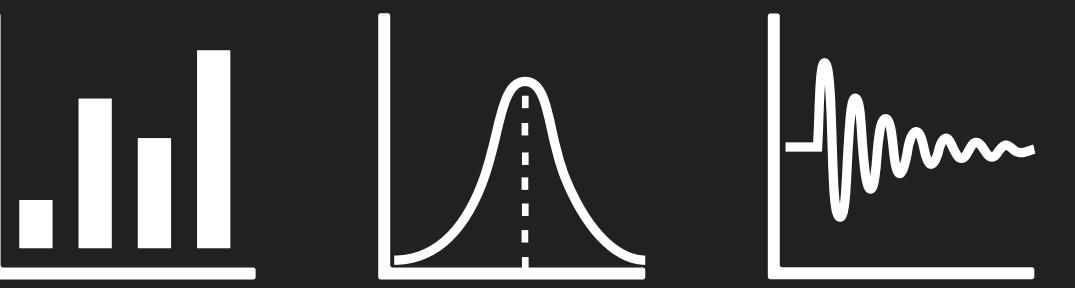
- Not to be confused with ElastiCache
- Mostly a search engine but also a document store (caution here)
- Amazon ElasticSearch Service components are sometimes referred to as an ELK stack.



Amazon ElasticSearch



A CLOUD GURU



Analytics

Kibana

Intake

IoT

Search and Storage

ElasticSearch



A dark, low-light photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a smartphone horizontally. The screen of the phone is visible but appears mostly black or very dimly lit.

Database Comparisons

Amazon Database Options



A CLOUD GURU

Database on EC2	<ul style="list-style-type: none">• Ultimate control over database• Preferred DB not available under RDS
Amazon RDS	<ul style="list-style-type: none">• Need traditional relational database for OLTP• Your data is well-formed and structured
Amazon DynamoDB	<ul style="list-style-type: none">• Name/value pair data or unpredictable data structure• In-memory performance with persistence
Amazon Redshift	<ul style="list-style-type: none">• Massive amounts of data• Primarily OLAP workloads
Amazon Neptune	<ul style="list-style-type: none">• Relationships between objects a major portion of data value
Amazon Elasicache	<ul style="list-style-type: none">• Fast temporary storage for small amounts of data• Highly volatile data

Amazon Database Options



A CLOUD GURU

Amazon QLDB	<ul style="list-style-type: none">• Need an immutable fully-managed ledger only
Amazon Managed Blockchain	<ul style="list-style-type: none">• Full blockchain framework for custom applications
Amazon Timestream	<ul style="list-style-type: none">• Optimized for time-series streaming data; Nice alternative to Redshift or DynamoDB for streaming sources
Amazon DocumentDB	<ul style="list-style-type: none">• Want full MongoDB compatibility for existing applications
Amazon Athena	<ul style="list-style-type: none">• Quickly query data on S3 without the need for extensive ETL operations

A dark, low-light photograph of a person from the chest up. They are wearing a dark-colored button-down shirt. Their hands are clasped in front of them, and they appear to be holding a small object, possibly a smartphone, which is partially visible at the bottom left.

Data Stores

Exam Tips



Exam Tips

- Read the [AWS Storage Options](#) white paper and note anti-patterns
- Know when to use various data stores:

RDS	<ul style="list-style-type: none">• Traditional relational data models• Existing apps requiring RDBMS• OLTP, ACID-compliant
DynamoDB	<ul style="list-style-type: none">• High I/O needs• Scale dynamically
S3	<ul style="list-style-type: none">• BLOBs
EC2	<ul style="list-style-type: none">• Database not supported under RDS• Need complete control
Redshift	<ul style="list-style-type: none">• OLAP



A CLOUD GURU

For Further Study



For Further Study

AWS Whitepapers	Required
re:Invent Videos	Optional but Recommended



A CLOUD GURU

For Further Study

The screenshot shows a video player interface with the following details:

- Top Left:** A small thumbnail for the AWS Certified Solutions Architect - Professional exam.
- Top Center:** The text "NOW VIEWING" followed by the lesson title "Exam Tips".
- Top Right:** A play button icon with a checkmark, indicating the lesson is current.
- Middle Left:** A sidebar titled "ALL LESSONS" with "CURRENT LESSON" highlighted. Below it is a "RESOURCES" section containing the following items:
 - Whitepaper: Storage Options in the Cloud
 - Whitepaper: Multi-Tenant SaaS Storage Strategies
 - Whitepaper: Performance at Scale with Amazon ElastiCache
 - Video: Deep Dive on Amazon S3 and Amazon Glacier Storage Management
 - Video: Deep Dive on Amazon Relational Database Service
 - Video: ElastiCache Deep Dive: Best Practices and Usage Patterns
 - Video: Deep Dive: Using Hybrid Storage with AWS Storage Gateway to Solve On-Premises Storage Problems
- Center:** The main video area displays the title "Data Stores" above "Exam Tips". A play button icon is overlaid on the video frame.
- Bottom:** A control bar with playback controls (rewind, fast forward, volume, etc.) and two smaller video preview windows below it.

For Further Study



A CLOUD GURU

2:24 PM Thu Dec 6 81%

Data Stores Exam Tips

Exam Tips CHAPTER 2.16

AWS Certified Solutions Architect - Professional 2019 (Beta exam)

OPEN

LINKS & RESOURCES UP NEXT

- 1 Whitepaper: Storage Options in the Cloud <https://d1.awsstatic.com/whitepapers/Storage/AWS%20Storage%20Services%20Whitepaper-v9.pdf>
- 2 Whitepaper: Multi-Tenant SaaS Storage Strategies https://d1.awsstatic.com/whitepapers/Multi_Tenant_SaaS_Storage_Strategies.pdf
- 3 Whitepaper: Performance at Scale with Amazon ElastiCache <https://d0.awsstatic.com/whitepapers/performance-at-scale-with-amazon-elasticsearch.pdf>
- 4 Video: Deep Dive on Amazon S3 and Amazon Glacier Storage Management <https://www.youtube.com/watch?v=SUWqDOnXeDw>
- 5 Video: Deep Dive on Amazon Relational Database Service <https://www.youtube.com/watch?v=TjxC-B9Q9tQ>
- 6 Video: ElastiCache Deep Dive: Best Practices and Usage Patterns https://www.youtube.com/watch?v=_YBdsuUq2M
- 7 Video: Deep Dive: Using Hybrid Storage with AWS Storage Gateway to Solve On-Premises Storage P... <https://www.youtube.com/watch?v=9wgaV70FeaM>

2:23 PM 84%

Exam Tips

A Cloud Guru

- Read the [AWS Storage Options](#) white paper and note anti-patterns
- Know when to use various data stores:

RDS	<ul style="list-style-type: none">• Traditional relational data models• Existing apps requiring RDBMS• OLTP, ACID-compliant
DynamoDB	<ul style="list-style-type: none">• High I/O needs• Scale dynamically
S3	<ul style="list-style-type: none">• BLOBs
EC2	<ul style="list-style-type: none">• Database not supported under RDS• Need complete control
Redshift	<ul style="list-style-type: none">• OLAP

3 MB

Exam Tips

CHAPTER 2.16

AWS Certified Solutions Architect - Professional 2019 (Beta exam) OPEN

LINKS & RESOURCES UP NEXT

- 1 Whitepaper: Storage Options in the... <https://d1.awsstatic.com/whitepapers/Storage/AWS%20Storage%20Services%20Whitepaper-v9.pdf>
- 2 Whitepaper: Multi-Tenant SaaS Stor... https://d1.awsstatic.com/whitepapers/Multi_Tenant_SaaS_Storage_Strategies.pdf
- 3 Whitepaper: Performance at Scale... <https://d0.awsstatic.com/whitepapers/performance-at-scale-with-amazon-elasticsearch.pdf>
- 4 Video: Deep Dive on Amazon S3 and... <https://www.youtube.com/watch?v=SUWqDOnXeDw>
- 5 Video: Deep Dive on Amazon Relatio... <https://www.youtube.com/watch?v=TjxC-B9Q9tQ>
- 6 Video: ElastiCache Deep Dive: Best... https://www.youtube.com/watch?v=_YBdsuUq2M
- 7 Video: Deep Dive: Using Hybrid St... <https://www.youtube.com/watch?v=9wgaV70FeaM>



For Further Study

AWS Whitepapers	Required
re:Invent Videos	Optional but Recommended



A CLOUD GURU

For Further Study

AWS Whitepapers	Required
re:Invent Videos	Optional but Recommended

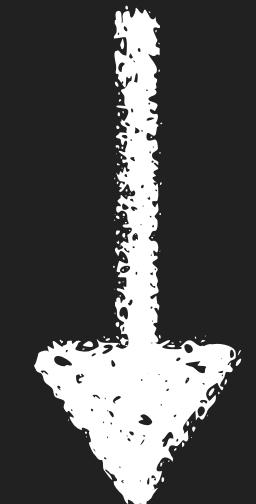




A CLOUD GURU

For Further Study

AWS Whitepapers	Required
re:Invent Videos	Optional but Recommended



A dark, low-light photograph of a person from the chest up. They are wearing a dark, button-down shirt. Their hands are clasped in front of them, and they appear to be holding a small object, possibly a smartphone. The background is dark and out of focus.

Data Stores

Pro Tips

Pro Tips - Storage

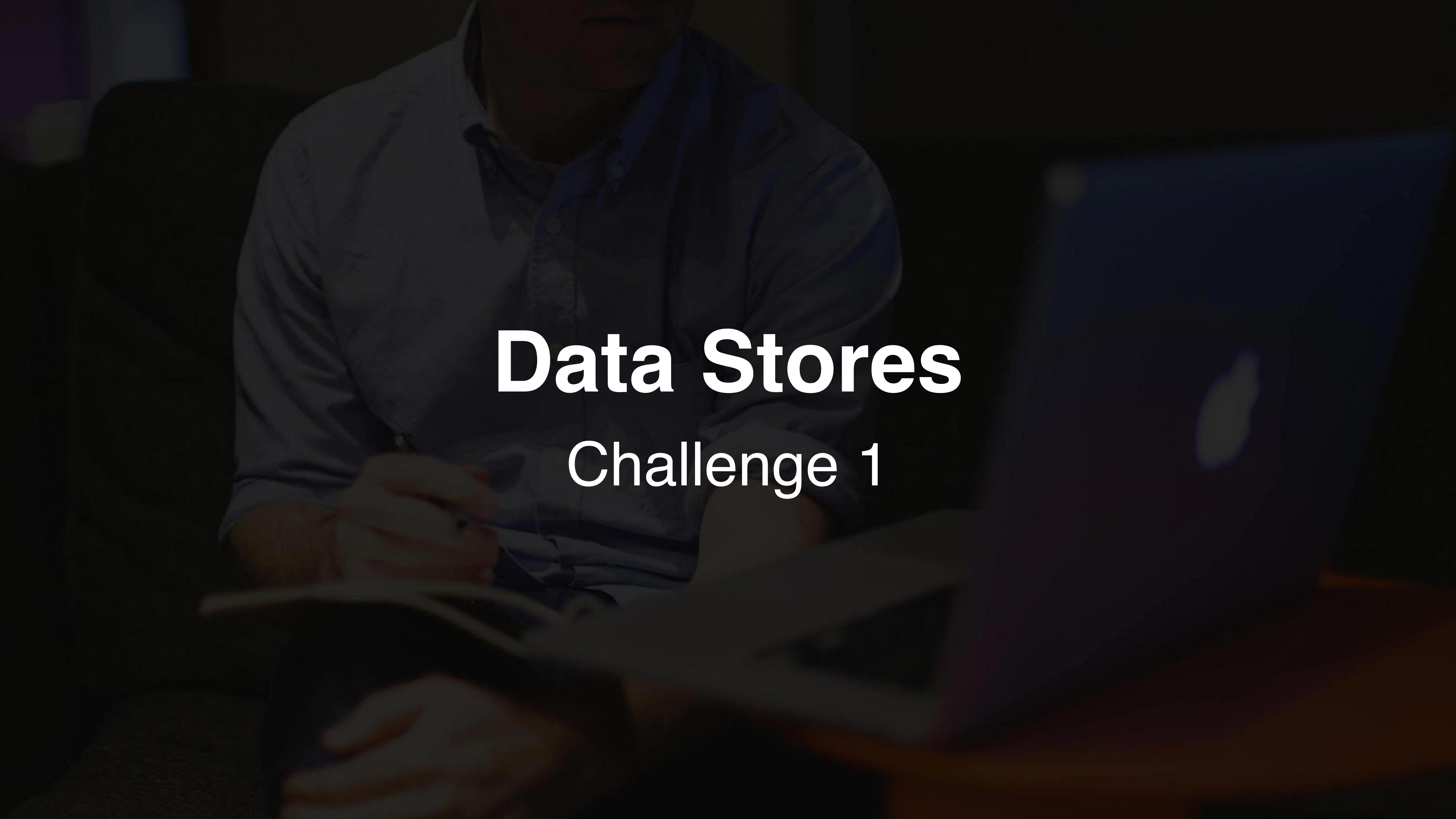
- Archiving and Backup often a great “pilot” to build AWS business case
- Make use of the S3 endpoints within your VPC
- Learn how to properly secure your S3 bucket
- Encrypt, Encrypt, Encrypt

Pro Tips - Databases



A CLOUD GURU

- Consider Aurora for your production MySQL/Maria or PostgreSQL needs.
- Consider NoSQL if you don't need relational database features.
- Databases on EC2 cost less on the surface than RDS, but remember to factor in management (backups, patching, OS-level hardening).
- There can be a performance hit when RDS backups run if you have only a single AZ instance.

A dark, out-of-focus photograph of a person from the chest up. The person is wearing a dark-colored button-down shirt. Their hands are clasped in front of them, holding a small, light-colored object, likely a smartphone. The background is dark and indistinct.

Data Stores

Challenge 1



Challenge 1

You work for a small independent PC game maker who is readying for the release of their newest game. It's going to be released as a freemium model where gamers can download and play the first portion of the game for free and then must pay for additional game modules.

The game's base data files are 20GB with the potential of another 50GB if the customers purchase all the game modules. You've been asked to architect a method to allow the public to efficiently download the game. The company expects huge demand in the first few days so the solution should be scalable. Additionally, the company wants to distribute the game for the lowest possible AWS cost since capital is extremely tight.

Order these ideas by best to worst.

- A. Create an EC2 instance with EBS volumes and configure as an FTP server. Build a small application containing an FTP client that customers can download and run.
- B. Upload the files to an S3 bucket and grant it public read-only access. Configure Requester Pays for the S3 bucket and distribute via URL published on the company website.
- C. Upload the files to an S3 bucket using the Infrequent Access Storage Class. Configure the bucket to act as a static web-server and distribute via URL published on the company website.
- D. Upload the game files to an S3 bucket and grant public read-only access. Build a small application containing a BitTorrent client that customers can download and run.



Challenge 1

You work for a small independent PC game maker who is readying for the release of their newest game. It's going to be released as a freemium model where gamers can download and play the first portion of the game for free and then must pay for additional game modules.

The game's base data files are **20GB** with the potential of another **50GB** if the customers purchase all the game modules. You've been asked to architect a method to **allow the public** to efficiently download the game. The company expects huge demand in the first few days so the solution **should be scalable**. Additionally, the company wants to distribute the game for the **lowest possible AWS cost** since capital is extremely tight.

Order these ideas by best to worst.

- Large files
- Must be accessible to the public
- Scalable
- Lowest cost option
- There may be better solutions, but we have to work with these four
- No other constraints - don't invent external constraints



Challenge 1

You work for a small independent PC game maker who is readying for the release of their newest game. It's going to be released as a freemium model where gamers can download and play the first portion of the game for free and then must pay for additional game modules.

The game's base data files are 20GB with the potential of another 50GB if the customers purchase all the game modules (no single file is over 5 GB). You've been asked to architect a method to allow the public to efficiently download the game. The company expects huge demand in the first few days so the solution should be scalable. Additionally, the company wants to distribute the game for the lowest possible AWS cost since capital is extremely tight.

2

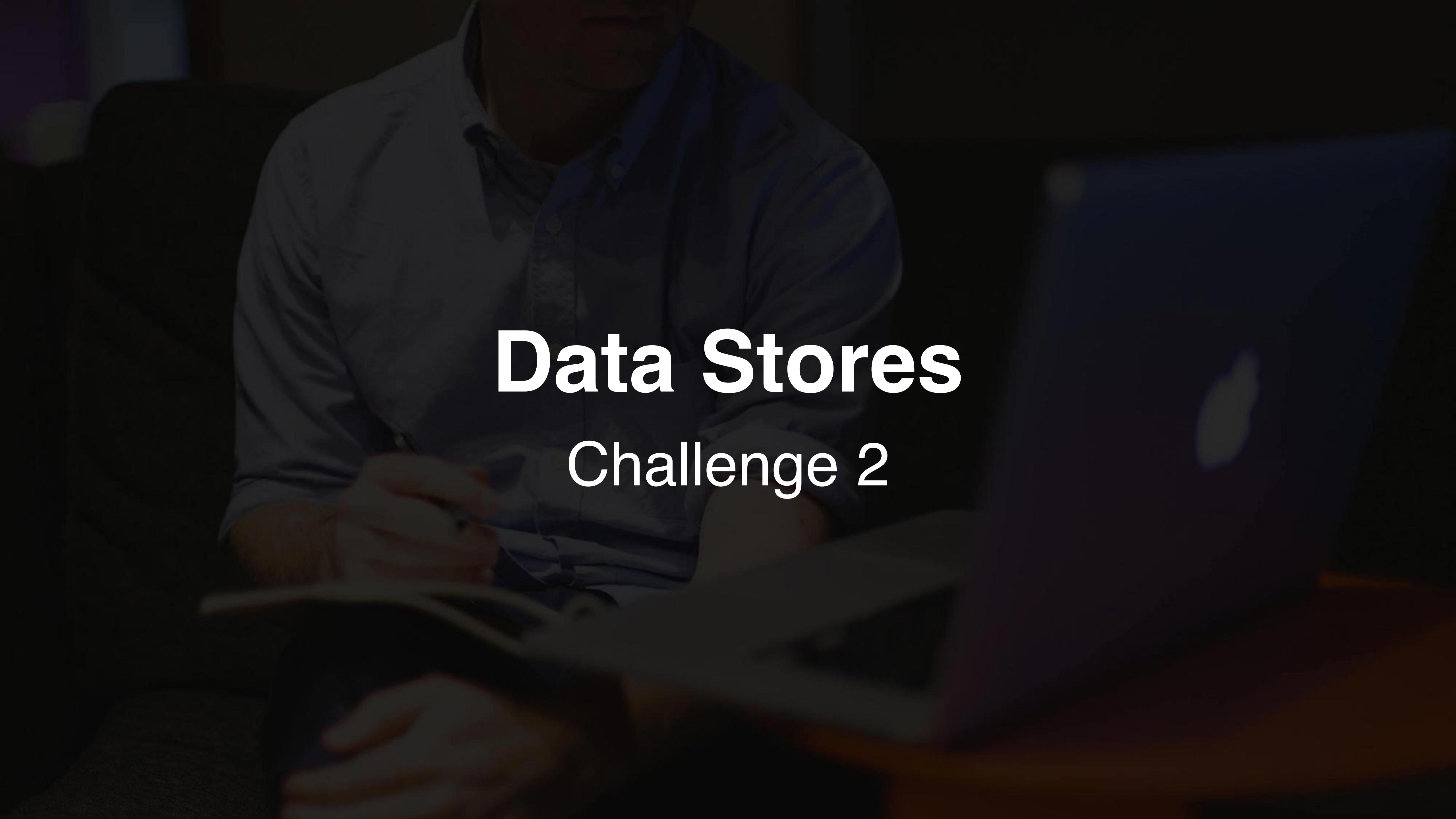
4

3

1

- A. Create an EC2 instance with EBS volumes and configure as an FTP server. Build a small application containing an FTP client that customers can download and run.
- B. Upload the files to an S3 bucket and grant it public read-only access. Configure Requester Pays for the S3 bucket and distribute via URL published on the company website.
- C. Upload the files to an S3 bucket using the Infrequent Access Storage Class. Configure the bucket to act as a static web-server and distribute via URL published on the company website.
- D. Upload the game files to an S3 bucket and grant public read-only access. Build a small application containing a BitTorrent client that customers can download and run.

Order these ideas by best to worst.

A dark, low-light photograph of a person from the chest up. The person is wearing a dark, button-down shirt. Their hands are clasped in front of them, holding a smartphone. The phone's screen is visible but mostly black.

Data Stores

Challenge 2



Challenge 2

Your client is a small government department which has decided to migrate their internal document storage from a hosted environment to AWS. Due to organizational policy they do not archive data, everything must be 'online' but older data is rarely accessed.

On SAN storage, they have accumulated 30 TB of 'unmanaged' user documents and other files which grows at a rate of approximately 1TB per month. A cluster of Linux systems are attached to the SAN storage via iSCSI and only provide NFS shares of the SAN storage.

Propose a design solution that will use AWS services, meet the stipulated requirements, be easy to manage, and reduce costs where possible.

Choose One:

- A. Use EC2 & EBS capabilities to create large files server instances. Use EBS provisioned storage to achieve the necessary I/O. Use Direct Connect to connect to the Head Office. Setup snapshots per volume to achieve increased resiliency and fast recovery.
- B. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Present the File Gateway shares via NFS for on-premises systems. Configure the backing S3 buckets with lifecycle policies to take advantage of Standard - IA & Glacier to reduce costs
- C. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Present the Gateway shares via NFS for on-premises systems. Configure the backing S3 buckets with lifecycle policies to take advantage of Standard - IA to reduce costs
- D. Make use of the EFS to create multiple volumes based on volume and I/O demand. Establish a Direct Connect connection and allow consumer systems to directly mount the EFS shares via NFS. Configure EFS to create snapshots to S3, Then use S3 lifecycle policies to take advantage of Standard – IA.
- E. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Mount the Gateway volumes via iSCSI on the Linux systems to present as NFS shares. Configure the backing S3 buckets with lifecycle policies to take advantage of



Challenge 2

Your client is a small government department which has decided to migrate their internal document storage from a hosted environment to AWS. Due to **organizational policy** they **do not archive data, everything must be 'online'** but older data is rarely accessed.

On SAN storage, they have accumulated 30 TB of 'unmanaged' user documents and other files which grows at a rate of approximately 1TB per month. A cluster of Linux systems are attached to the SAN storage via iSCSI and only provide NFS shares of the SAN storage.

Propose a design solution that will use AWS services, **meet the stipulated requirements, be easy to manage, and reduce costs where possible.**

Choose One:

- A. Use EC2 & EBS capabilities to create large files server instances. Use EBS provisioned storage to achieve the necessary I/O. Use Direct Connect to connect to the Head Office. Setup snapshots per volume to achieve increased resiliency and fast recovery.
- B. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Present the File Gateway shares via NFS for on-premises systems. Configure the backing S3 buckets with lifecycle policies to take advantage of Standard - IA & Glacier to reduce costs
- C. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Present the Gateway shares via NFS for on-premises systems. Configure the backing S3 buckets with lifecycle policies to take advantage of Standard - IA to reduce costs
- D. Make use of the EFS to create multiple volumes based on volume and I/O demand. Establish a Direct Connect connection and allow consumer systems to directly mount the EFS shares via NFS. Configure EFS to create snapshots to S3, Then use S3 lifecycle policies to take advantage of Standard – IA.
- E. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Mount the Gateway volumes via iSCSI on the Linux systems to present as NFS shares. Configure the backing S3 buckets with lifecycle policies to take advantage of



Challenge 2

Your client is a small government department which has decided to migrate their internal document storage from a hosted environment to AWS. Due to organizational policy they do not archive data, everything must be 'online' but older data is rarely accessed.

On SAN storage, they have accumulated 30 TB of 'unmanaged' user documents and other files which grows at a rate of approximately 1TB per month. A cluster of Linux systems are attached to the SAN storage via iSCSI and only provide NFS shares of the SAN storage.

Propose a design solution that will use AWS services, meet the stipulated requirements, be easy to manage, and reduce costs where possible.

Choose One:

- A. Use EC2 & EBS capabilities to create large files server instances. Use EBS provisioned storage to achieve the necessary I/O. Use Direct Connect to connect to the Head Office. Setup snapshots per volume to achieve increased resiliency and fast recovery.
- B. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Present the File Gateway shares via NFS for on-premises systems. Configure the backing S3 buckets with lifecycle policies to take advantage of Standard - IA & Glacier to reduce costs
- C. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Present the Gateway shares via NFS for on-premises systems. Configure the backing S3 buckets with lifecycle policies to take advantage of Standard - IA to reduce costs
- D. Make use of the EFS to create multiple volumes based on volume and I/O demand. Establish a Direct Connect connection and allow consumer systems to directly mount the EFS shares via NFS. Configure EFS to create snapshots to S3, Then use S3 lifecycle policies to take advantage of Standard – IA.
- E. Make use of the File-Gateway service to create multiple buckets and gateways in Head Office based on volume and I/O demand. Mount the Gateway volumes via iSCSI on the Linux systems to present as NFS shares. Configure the backing S3 buckets with lifecycle policies to take advantage of

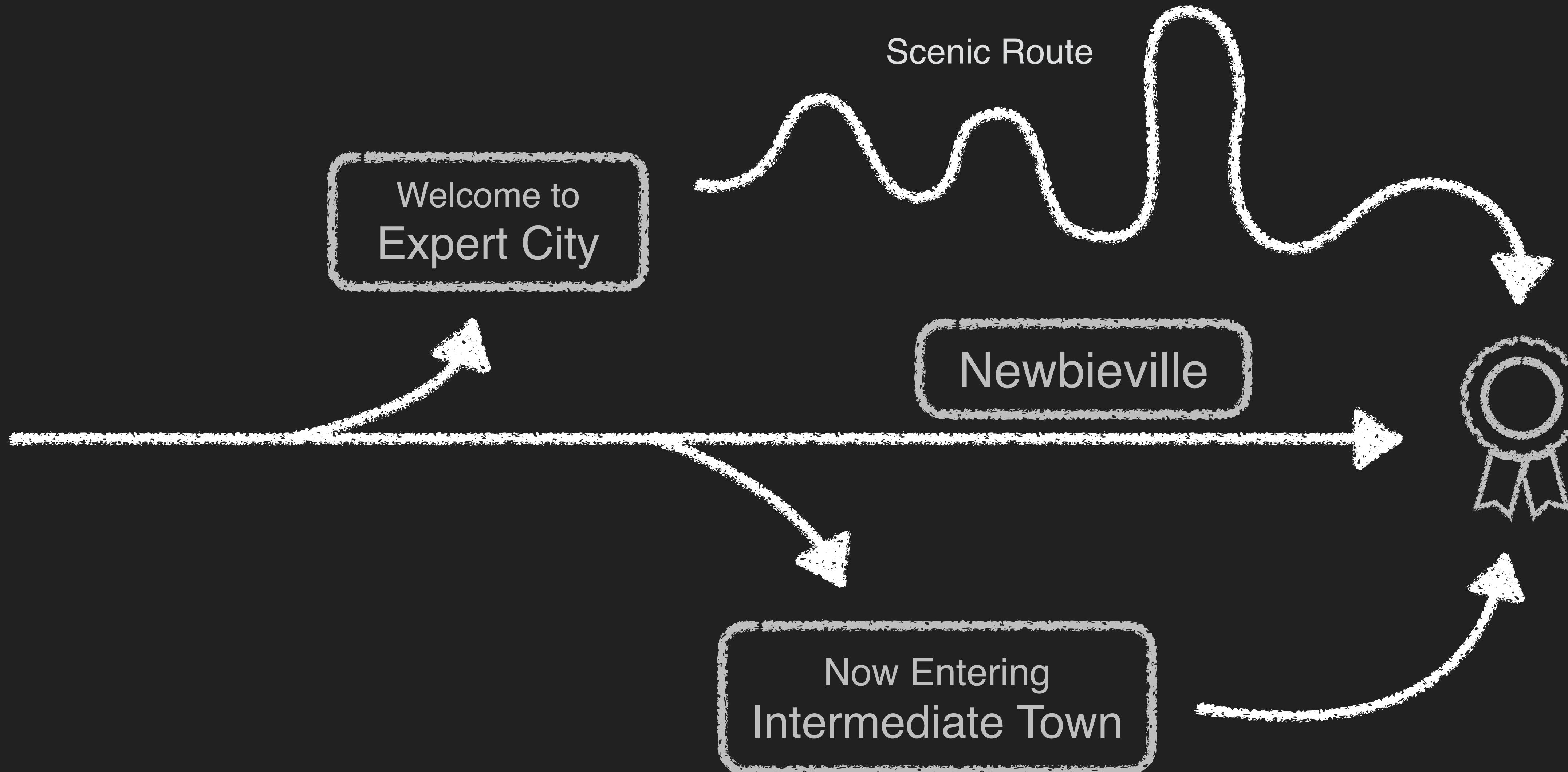
A dark, low-light photograph of a person from the chest up. They are wearing a light-colored lab coat over a dark shirt. Their hands are clasped in front of them, holding a small, clear test tube. The background is dark and out of focus.

Data Storage Lab Air Quality Analysis

You Choose Your Own Adventure!



A CLOUD GURU



Air Quality Analysis

You work for a US-based non-profit organization that promotes better environmental conditions in communities most in need.

Currently, your organization is preparing a campaign around awareness of Ozone as a pollutant.

You have been asked to identify the top regions in the US with the highest average Ozone levels for use in a regionally targeted social media awareness campaign.



Air Quality Analysis



Sequence	Task	Hint
1	Identify an open source data repository for air quality readings.	
2	Setup a way to query that data via AWS services and tools.	
3	What city had the highest average ozone (O ₃) reading on October 9, 2018?	

You will need the answer to #3 for the quiz at the end of the Chapter.

Air Quality Analysis



Expert Level

Pause now and go give it a try without hints.

Air Quality Analysis



Intermediate Level

Keep going for some hints to get you started.

Air Quality Analysis



Sequence	Task	Hint
1	Identify an open source data repository for air quality readings.	<i>registry.opendata.aws</i>
2	Setup a way to query that data via AWS services and tools.	Athena works well against JSON files
3	What city had the highest average ozone (O3) reading on October 9, 2018?	You'll have to find out!

You will need the answer to #3 for the quiz at the end of the Chapter.

Air Quality Analysis



Intermediate Level

Pause now and go give it a try.
Keep going to see how I did it.

Air Quality Analysis



A CLOUD GURU

Walkthrough