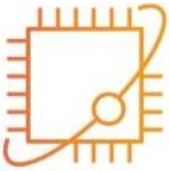
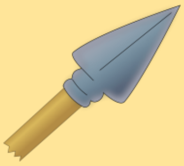
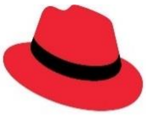
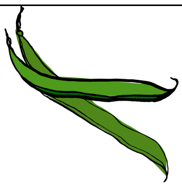
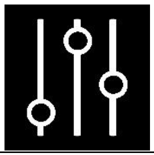
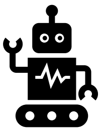
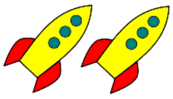
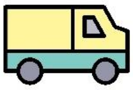
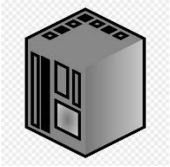
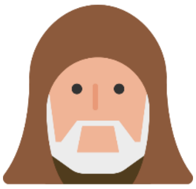
Present Period

Jan 2019 -

Contents

 Data Exchange	 Launch Wizard	Savings Plans	
December 2019			
 DeepComposer			
 Braket			
3rd December 2019			
 Augmented AI	 CodeGuru	 Compute Optimizer	 Detective
 Fraud Detector	 Kendra	 Keyspaces (for Apache Cassandra)	 Wavelength
2020			
CloudEndure Disaster Recovery	AppFlow	ChatBot	CodeArtifact

Honeycode	App2Container	CoPilot	Interactive Video Service
Migration Evaluator	 Network Firewall	MWAA	Monitron
Lookout for Equipment	Panorama	 **Proton**	HealthLake
Audit Manager	Lookout for Metrics	IoT EduKit	**Managed Grafana**
CloudShell	Managed Service for Prometheus		
2021			
 Fault Injection Simulator	 Red Hat OpenShift Service	 Nimble Studio	 Amazon FinSpace
 DevOps Guru	 App Runner	 Amazon Location Service	 MemoryDB for Redis
			

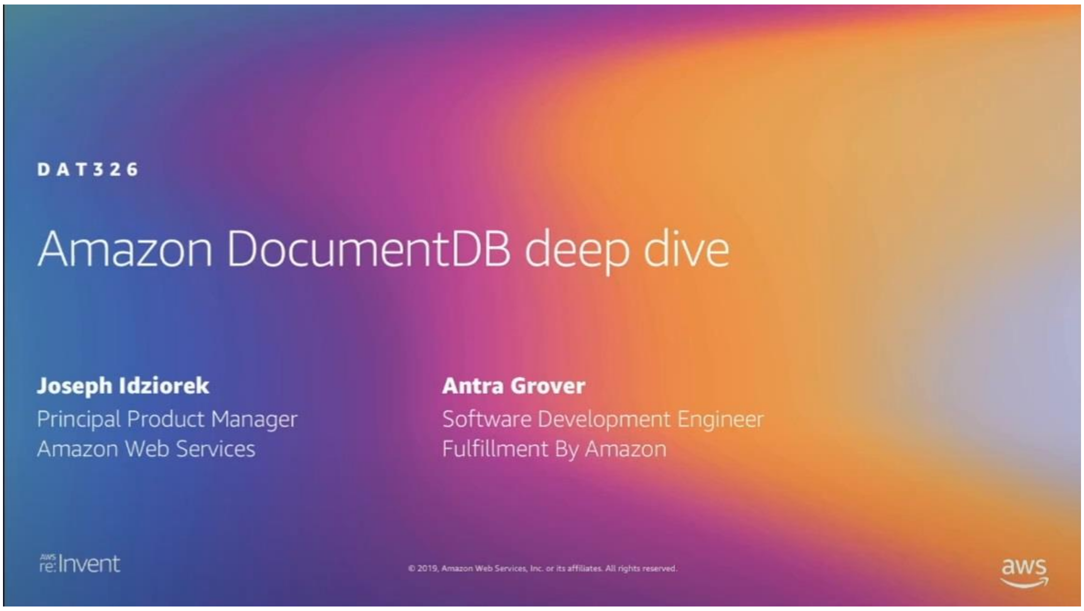
OpenSearch Service	Cloud Control API	AWS Resilience Hub	IoT RoboRunner
 AWS Private 5G	 IoT TwinMaker	 AWS FleetWise	 Mainframe Modernisation
 Cloud WAN	Microservice Extractor	AWS Re Post	Application Migration Service

DocumentDB

DocumentDB is plainly a database for documents. But its existence raises a number of questions:

1. Why might I want to place documents into a database? Why not, for instance, just place the information contained within the documents in plain form, the way we normally do with databases?
2. What do we mean by "document"?
3. Why was there a need to introduce a database specifically for documents?
4. What are the advantages of DocumentDB?
5. What are the disadvantages of DocumentDB?

We start with a presentation from 2019 by Joseph Idziorek and Antra Grover.



DAT326

Amazon DocumentDB deep dive

Joseph Idziorek
Principal Product Manager
Amazon Web Services

Antra Grover
Software Development Engineer
Fulfillment By Amazon

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

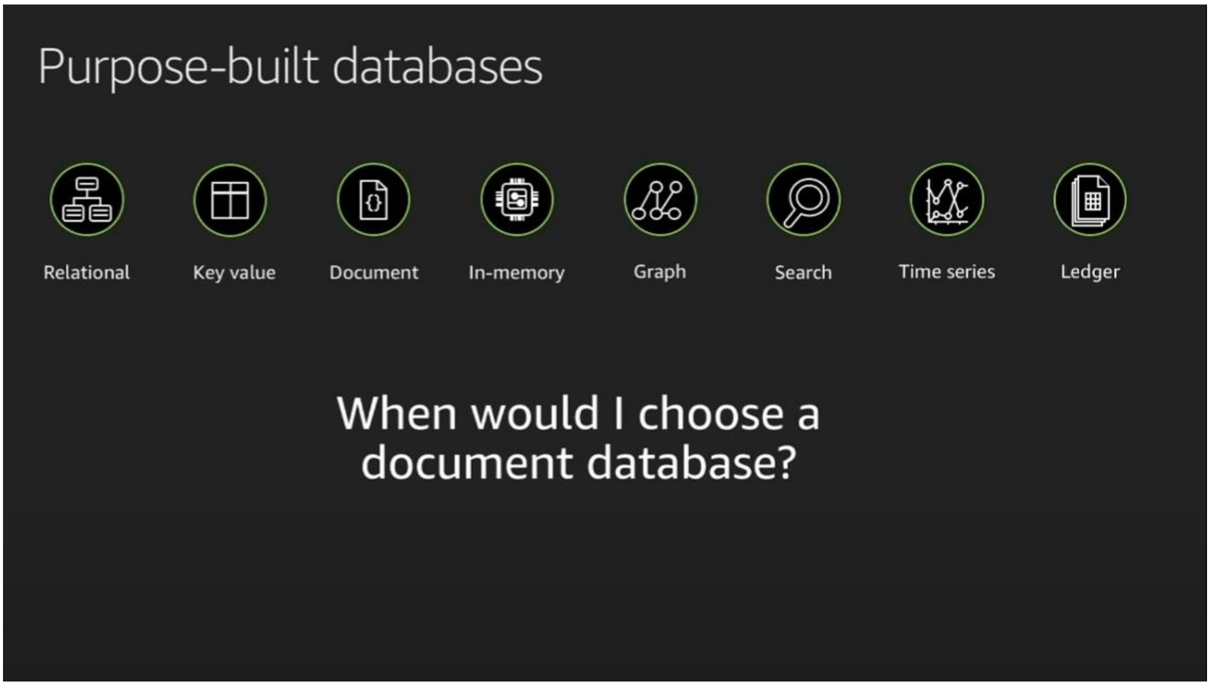
Joseph Idziorek's name is among fifteen associated with the paper that laid out the architectural design of DynamoDB. This was the AWS non-relational database announced in 2012. Here we are, seven years later, seeing the release of DocumentDB.

Joseph Idziorek [Idziorek 2010]

Straightaway in this presentation, we are given an answer to the question of why we need a database specifically for documents:

> As you folks well know, we have many databases at AWS. The reason for this is that we really don't believe there is a one-size-fits-all database.
>
> [Idziorek 2019]

A slide shows the array of databases which AWS offers. Indeed, Relational Database Service (RDS) is the Amazon offering for relational databases, Amazon ElastiCache is for in-memory databases, Amazon Neptune is for with graphical databases, Amazon Timestream is for time series data, Amazon Quantum Ledger Database is for ledger data. The slide says "search" – this is probably referring to Amazon Open Search, but there is also Amazon Kendra for building enterprise search systems and CloudSearch, released in 2012.



Purpose-built databases

Relational

Key value

Document

In-memory

Graph

Search

Time series

Ledger

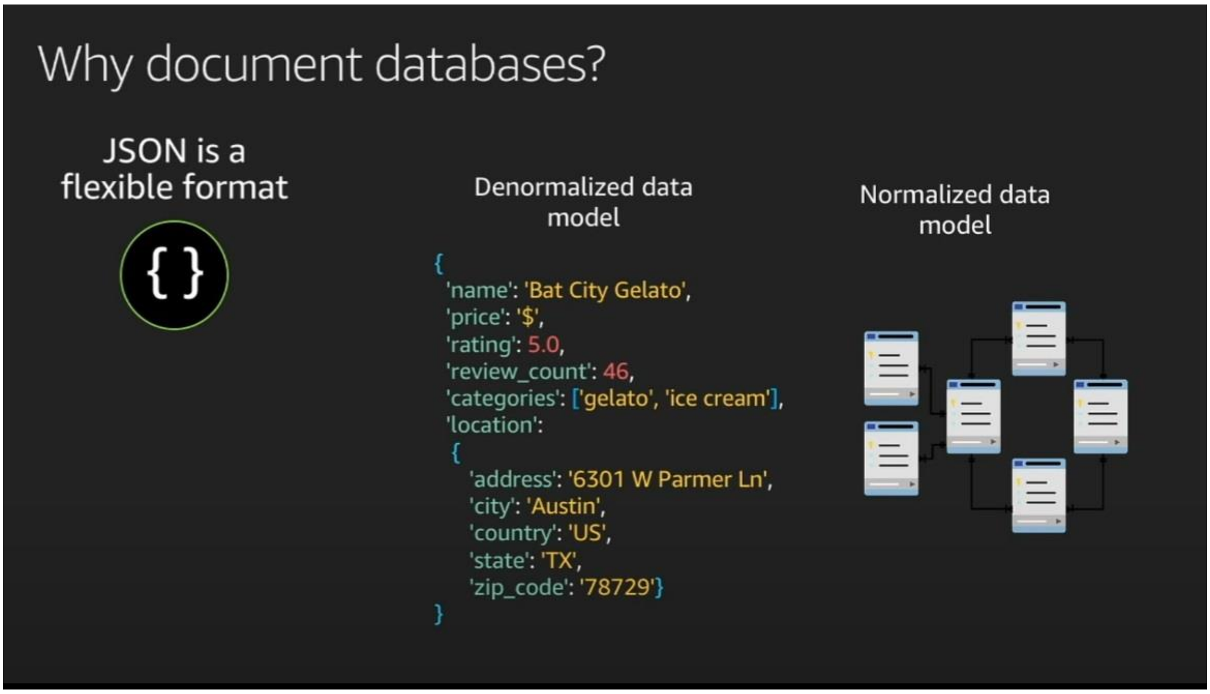
When would I choose a document database?

Idziorek and AWS's statement is completely unsatisfactory, however. He says "we really don't believe there is a one-size-fits-all database." What we want to know, however is the reason AWS hold this belief. Why is it that we cannot use as database to fit everything? What will happen if the correct database is not used? Surely data is data, after all.

Idziorek suggests some reasons for using the document model starting with the fact that

> if you think about the document model, this is how humans intuitively model data.
>
> We think about hierarchical lists. For folks taking notes—you know, back in university when you made notes—you know, you write things in hierarchical lists. You tab and you indent and you have lists. That's how you conceptually think about modelling data – and JSON lends itself really nicely to that data model.
>
> We don't typically think in normalized tables and rows. That's not a data model that we have ingrained in our head.
>
> So, it's very flexible from that perspective.
>
> [Idziorek 2019]

In this informal passage, we are given hints about what a document is. It looks to be a textual representation, which allows you to **indent** lines of text and represent **hierarchies**. Idziorek also suggests a reason to employ this document entity. The reason is that "this is how humans intuitively model data". Idziorek has in mind somebody listening to a lecture, writing out notes. Presumably, the argument is that because humans use the document format in this way, when the document format is used in DocumentDB, it is easier for humans to input data into it. Whether this holds is not immediately clear. Documents, do, after all seem to be nothing but agglomerations of smaller data types such as Booleans, integers, strings and so on.



Why document databases?

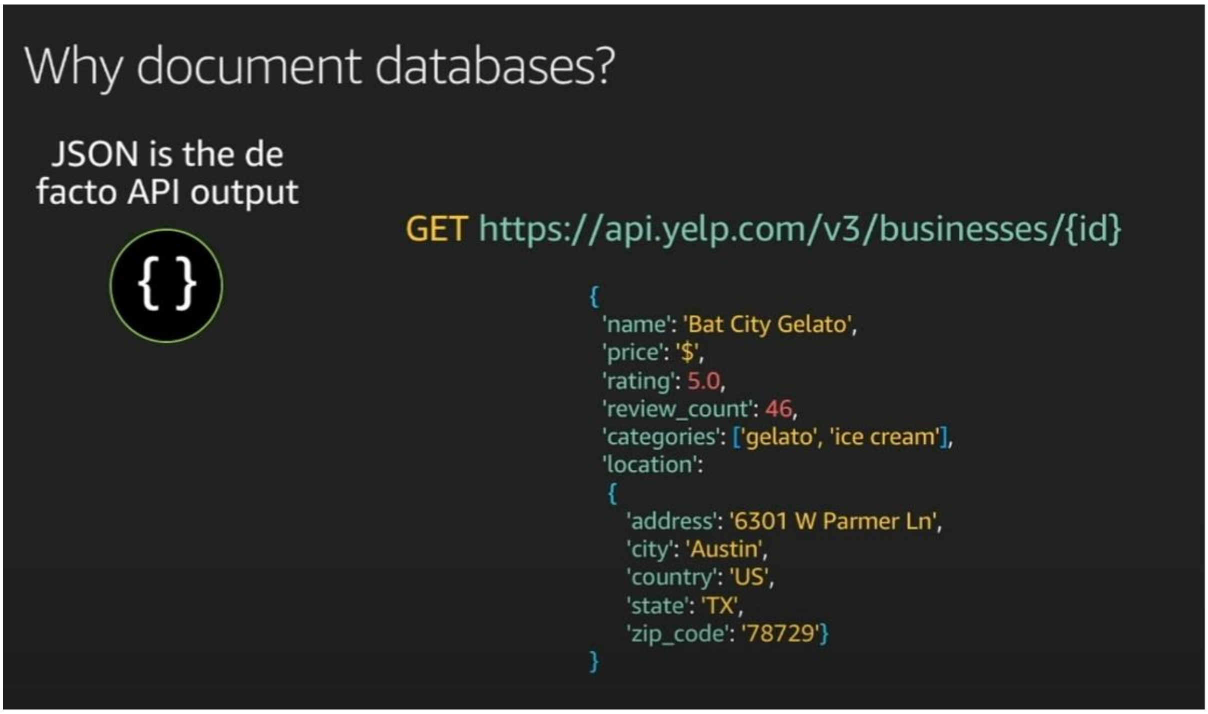
JSON is a flexible format

{ }

Denormalized data model

```
{
 'name': 'Bat City Gelato',
 'price': '$',
 'rating': 5.0,
 'review_count': 46,
 'categories': ['gelato', 'ice cream'],
 'location':
  {
    'address': '6301 W Parmer Ln',
    'city': 'Austin',
    'country': 'US',
    'state': 'TX',
    'zip_code': '78729'}
}
```

Normalized data model



Why document databases?

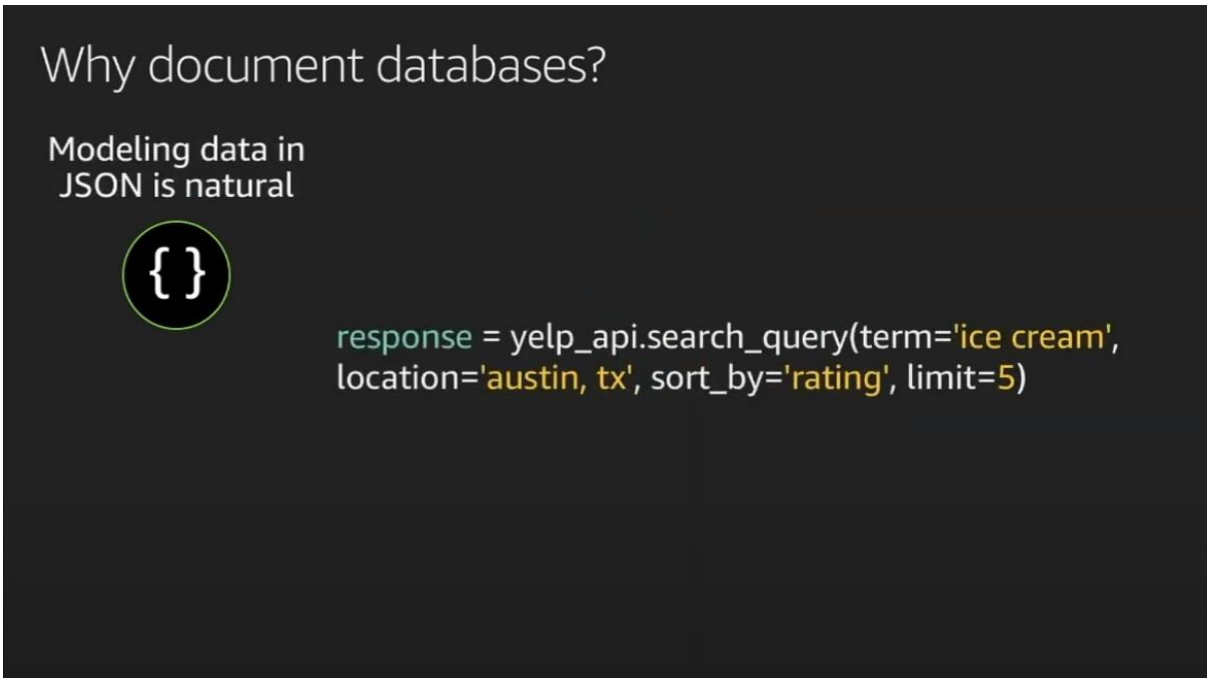
JSON is the de facto API output

{ }

GET https://api.yelp.com/v3/businesses/{id}

```
{
 'name': 'Bat City Gelato',
 'price': '$',
 'rating': 5.0,
 'review_count': 46,
 'categories': ['gelato', 'ice cream'],
 'location':
  {
    'address': '6301 W Parmer Ln',
    'city': 'Austin',
    'country': 'US',
    'state': 'TX',
    'zip_code': '78729'}
}
```

The slides above describe a distinction between normalized and denormalized data. Documents consist of denormalized data. We see some keys, in JSON format (name, price, rating). The second slide highlights a second reason for using the document format. Namely, (1) documents can be written in JSON, which stands for JavaScript Object Notation and (2) JSON is the de facto API output. We want to use APIs, so we should use the document format.

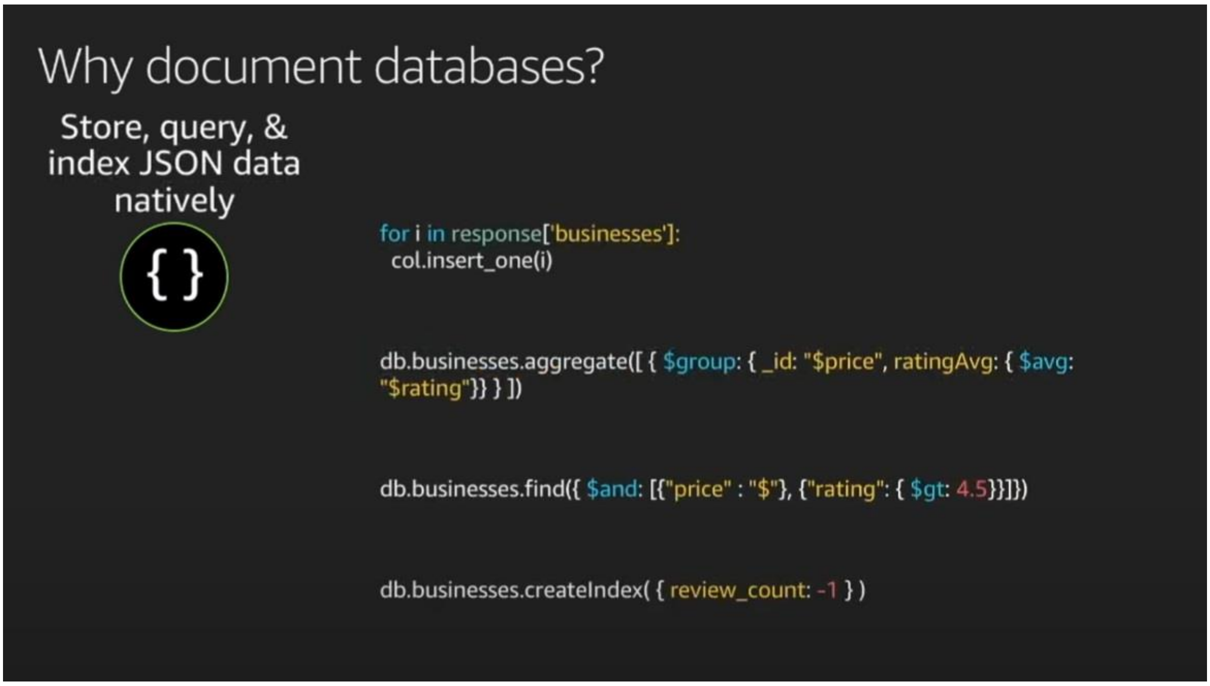


Why document databases?

Modeling data in JSON is natural

```
response = yelp_api.search_query(term='ice cream',
location='austin, tx', sort_by='rating', limit=5)
```

Idziorek gives another reason for using document databases. It is “natural” to model data in JSON. Above we see an equation. The name of the object is yelp_api. The method, belonging to the object is search_query. Presumably, Idziorek is referring to the fact that JSON is object-oriented and so tries to mimic the way things are in the world. The text above, for example, is human-readable and it is plain what the data is doing. We have a search query; the term search for is “ice cream” and there are three more parameters (location, sort_by and limit).



Why document databases?

Store, query, & index JSON data natively

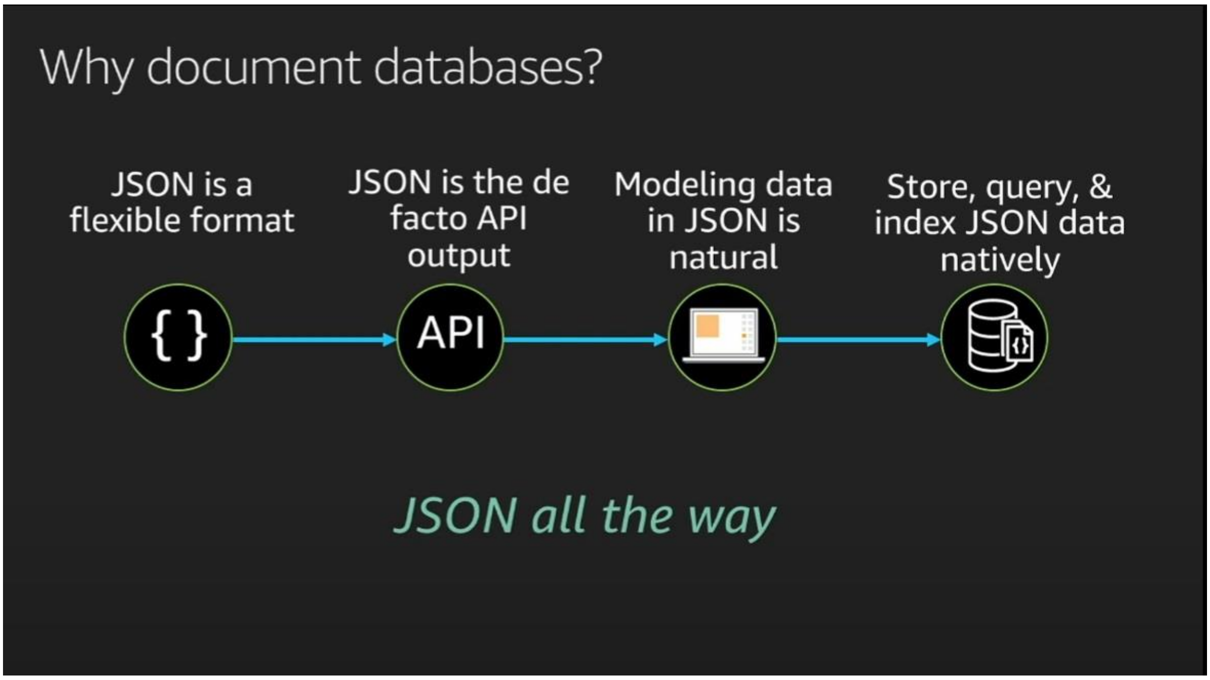
```
for i in response['businesses']:
  col.insert_one(i)

db.businesses.aggregate([ { $group: { _id: "$price", ratingAvg: { $avg:
"$rating"}} } ])

db.businesses.find({ $and: [{"price" : "$"}, {"rating": { $gt: 4.5}}]})

db.businesses.createIndex( { review_count: -1 } )
```

The final reason for using document databases is that they allow you to "store, query and index JSON data natively". Idziorek says:

> It's a really easy, natural way for us to model data in our application.

In the slide above, we can see a for… in structure. Presumably, this slide shows JSON being used within JavaScript. The JSON is being queried *natively*. In other words, we can simply refer to the words in the JSON document, and do not need to perform any transformation on the data.

We are given a summary of the four reasons for using a document database:



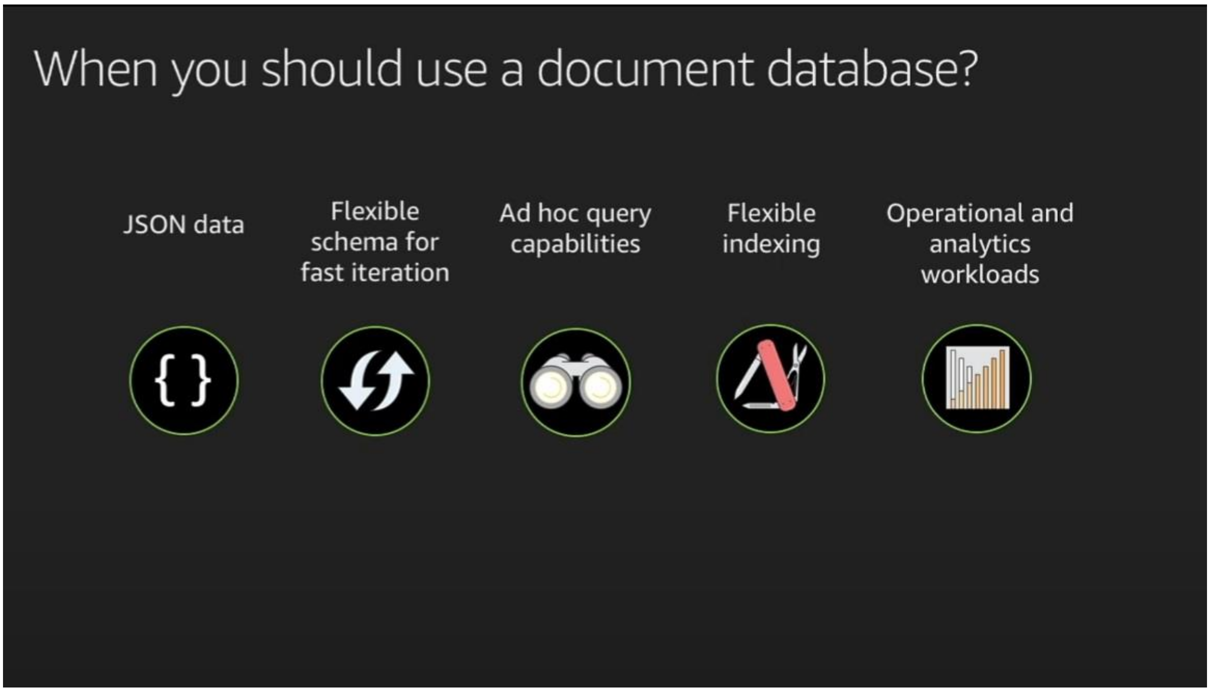
Why document databases?

JSON is a flexible format
{ }
JSON is the de facto API output
API
Modeling data in JSON is natural
Store, query, & index JSON data natively

JSON all the way

As a listener, I still have some questions. In what sense do we still have a database? If a database consists of rows and columns of cells, do we find documents placed in particular cells? This must be the way things are, as I cannot envisage it being another other way. Perhaps a single JSON document is to be considered a database. The lines of text are to be regarded as rows etc. It is hard to formulate why this is not the way it is – after all, an indentation could be translated to a skipping over of a column; blank lines are blank rows. None of these fundamental issues—the sense in which we have a document and a database—are touched upon by Idziorek.

Idziorek then seemingly turns to a new discussion topic – when you should use a document database. This discussion *could* involve a discussion of when the document model is appropriate (drawing limits on the situations in which we should use DocumentDB) but in reality, similar points are made as were in the previous topic.



When you should use a document database?

JSON data

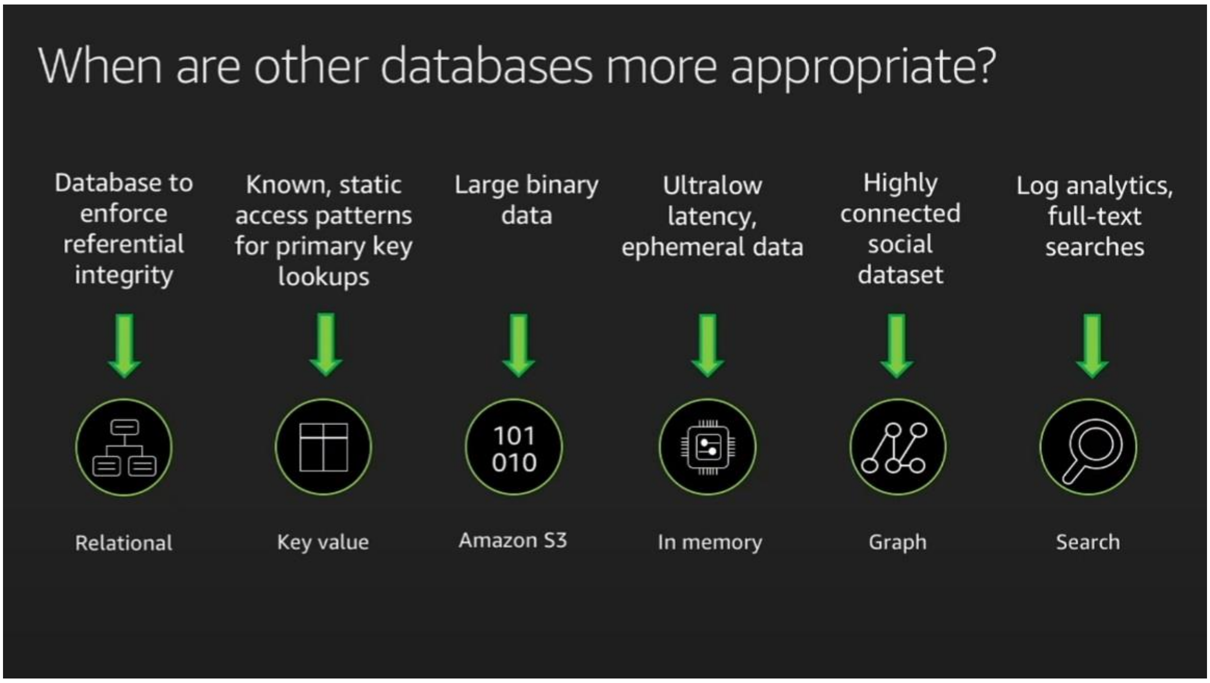
Flexible schema for fast iteration

Ad hoc query capabilities

Flexible indexing

Operational and analytics workloads

The slide above shows five items, but these are things anybody would want. *Anybody* would want a flexible schema; *anybody* would want ad hoc query capabilities; *anybody* would want flexible indexing. So, this is not helpful as a discussion of situations in which document databases are appropriate – it is more of the rolling off of benefits which Amazon love to do. Anyhow, it is clear that the next questions we need to think about are that of why, exactly, a document database allows flexible indexing. Furthermore, *Why does a document database allow* ad hoc *query capabilities*?



When are other databases more appropriate?

Database to enforce referential integrity → Relational

Known, static access patterns for primary key lookups → Key value

Large binary data → Amazon S3

Ultralow latency, ephemeral data → In memory

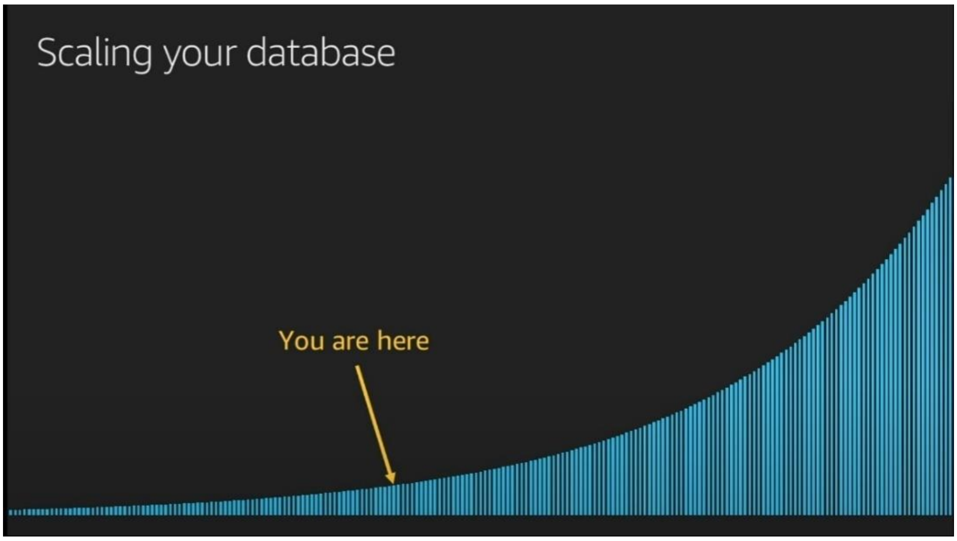
Highly connected social dataset → Graph

Log analytics, full-text searches → Search

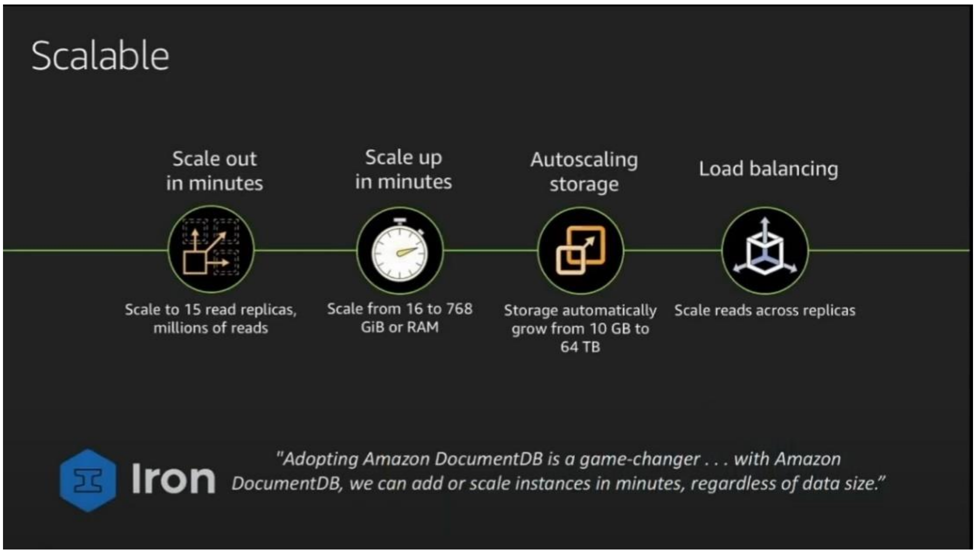
Idziorek does point out that there are many situations in which a document database would not be appropriate.

Architectural details

The presentation then turns towards a number of architectural details of AWS DocumentDB. They fall into the following sections:



Scaling your database

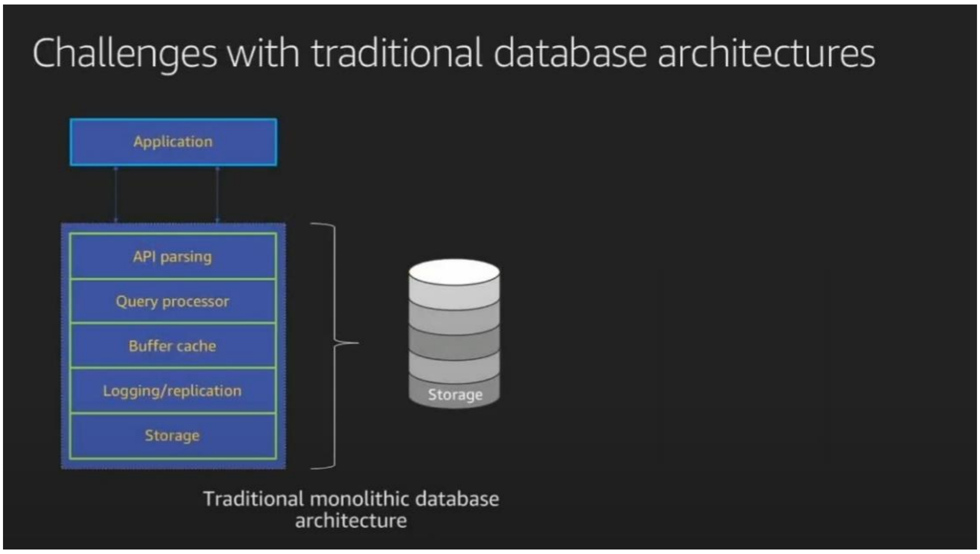


Scalable

Scale out in minutes	Scale up in minutes	Autoscaling storage	Load balancing
Scale to 15 read replicas, millions of reads	Scale from 16 to 768 GiB or RAM	Storage automatically grow from 10 GB to 64 TB	Scale reads across replicas

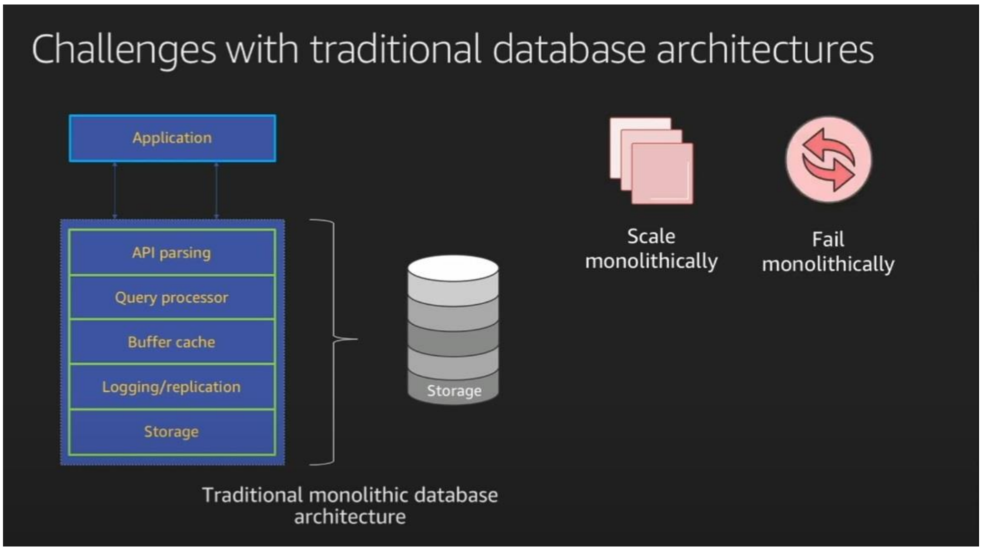
"Adopting Amazon DocumentDB is a game-changer . . . with Amazon DocumentDB, we can add or scale instances in minutes, regardless of data size."

Sometimes storage is distinct from databases, so it is interesting that above they talk about autoscaling storage.



Challenges with traditional database architectures

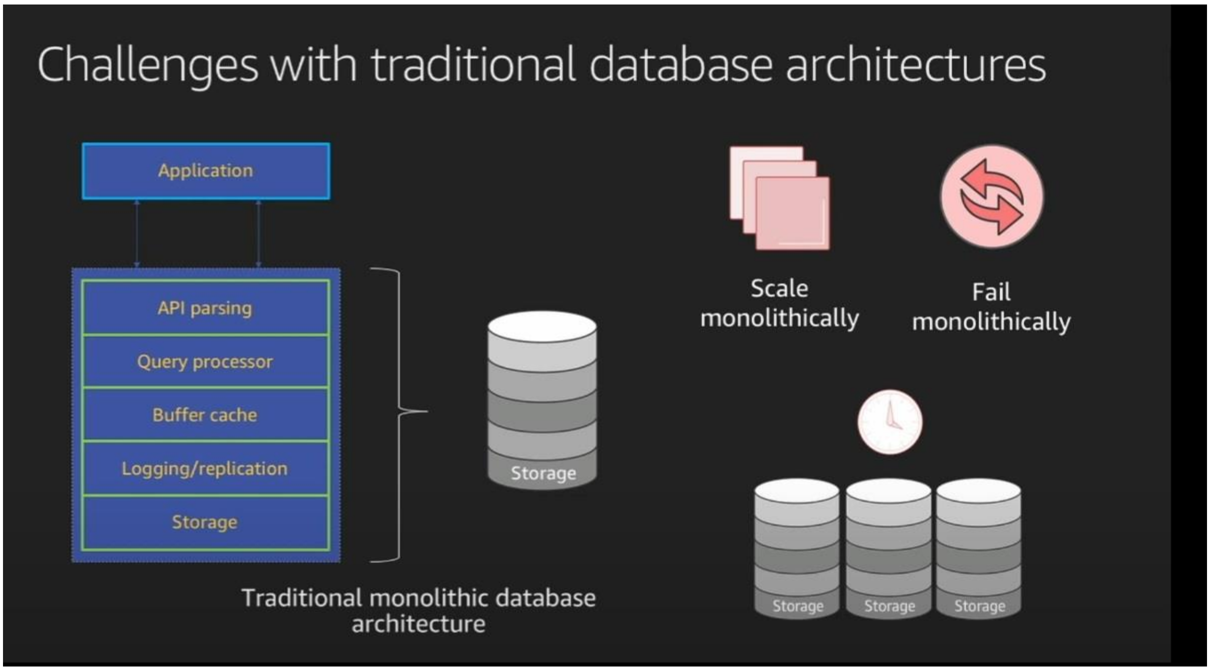
Traditional monolithic database architecture



Challenges with traditional database architectures

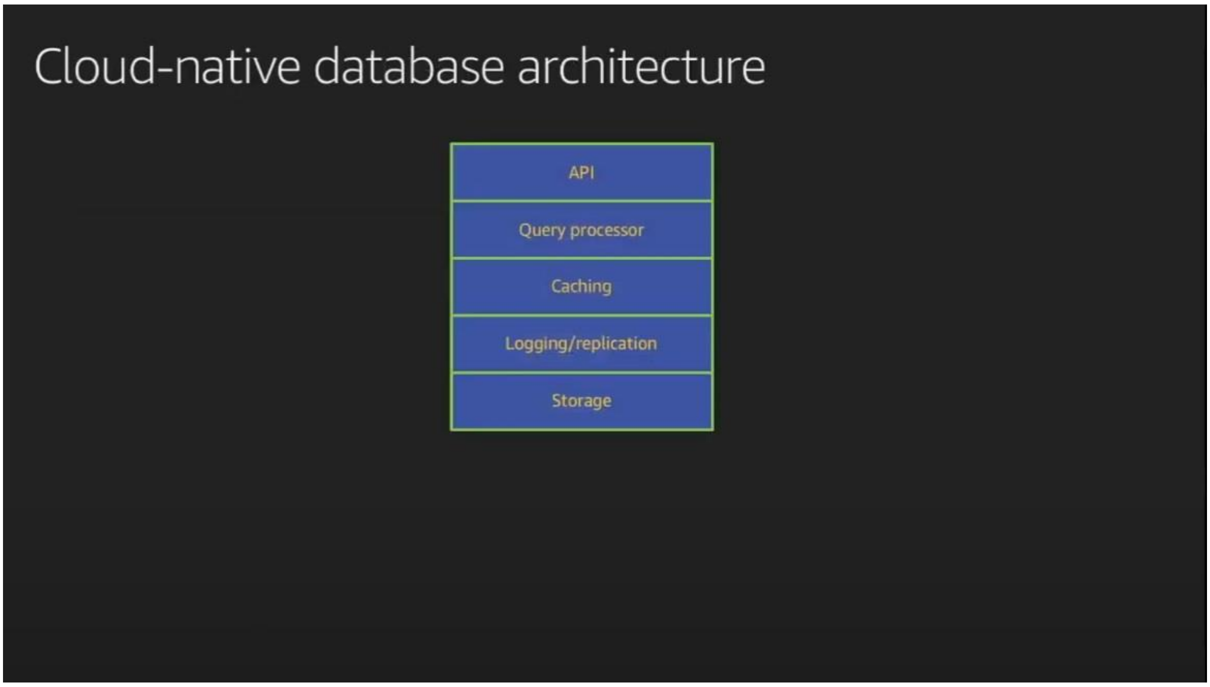
Traditional monolithic database architecture

What does it mean to scale monolithically? Perhaps it means that if one of the six components on the lefthand side scales, then they must all scale.



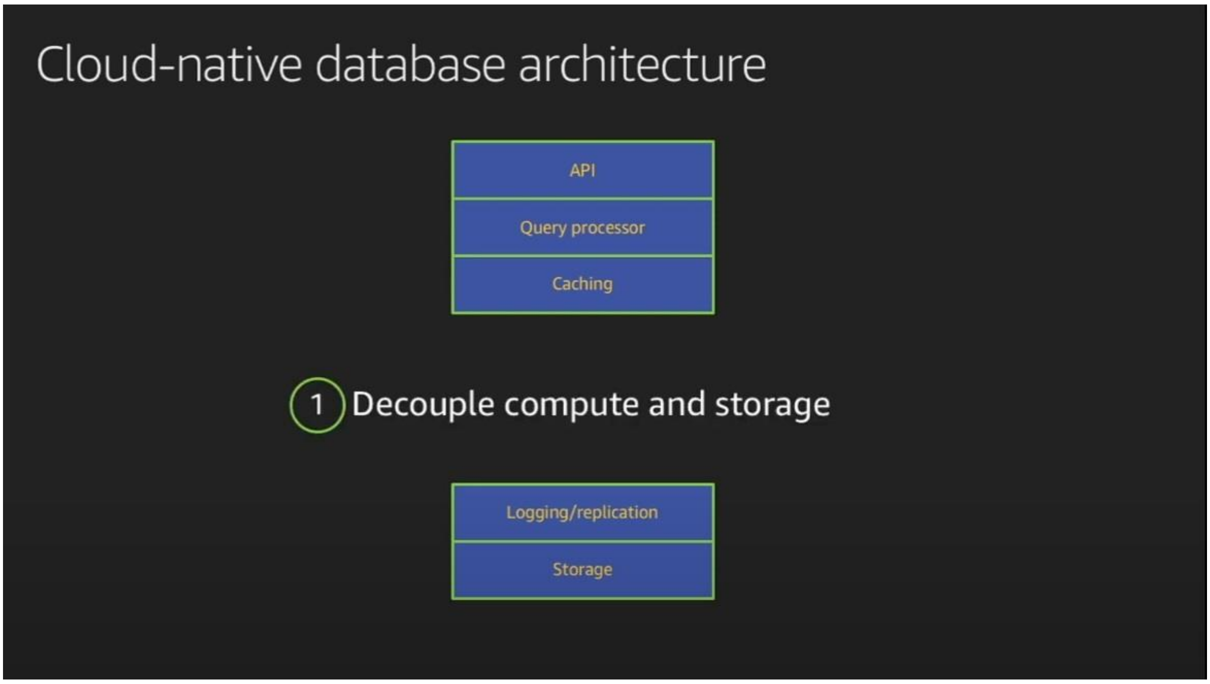
Challenges with traditional database architectures

A traditional database architecture is distinguished from a cloud-native database architecture.



Cloud-native database architecture

In the image above, things have changed. There is no longer a rectangle which says "application". It no longer says "buffer cache" and instead just says "caching". There are five components in total.



Cloud-native database architecture

API

Query processor

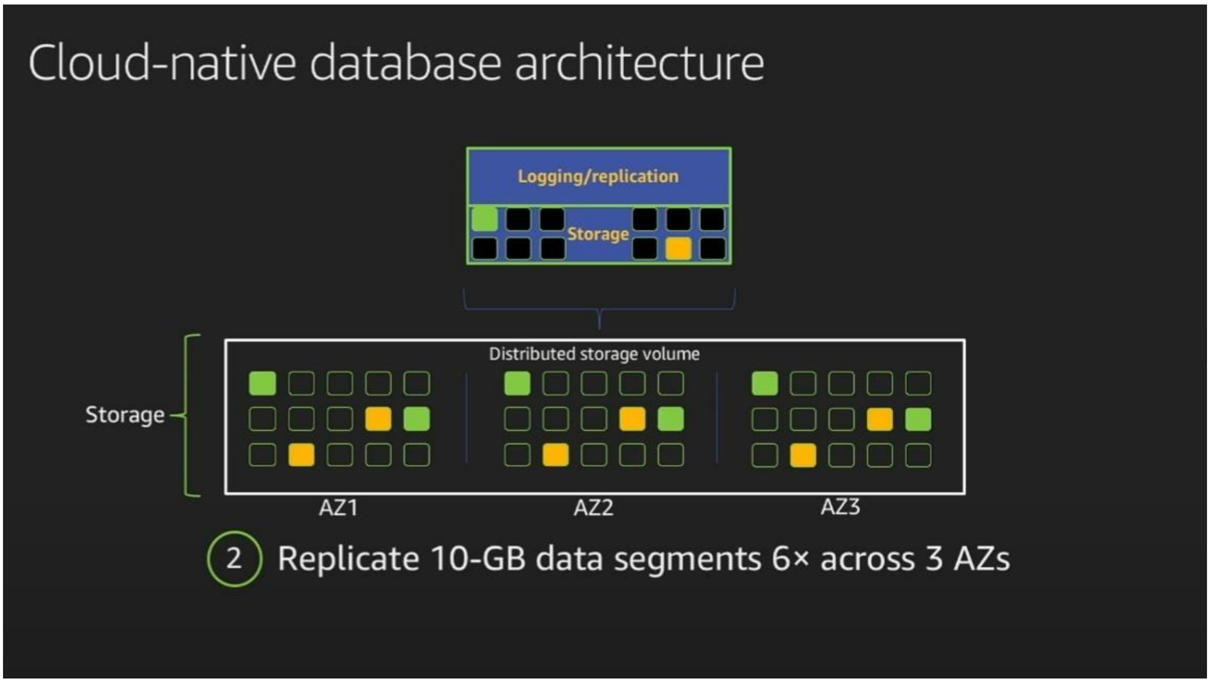
Caching

1 Decouple compute and storage

Logging/replication

Storage

With a cloud-native database architecture, the storage is decoupled from the compute. The block of rectangles at the top is the compute, and the block at the bottom is the storage.



Cloud-native database architecture

Logging/replication

Storage

Distributed storage volume

Storage

AZ1

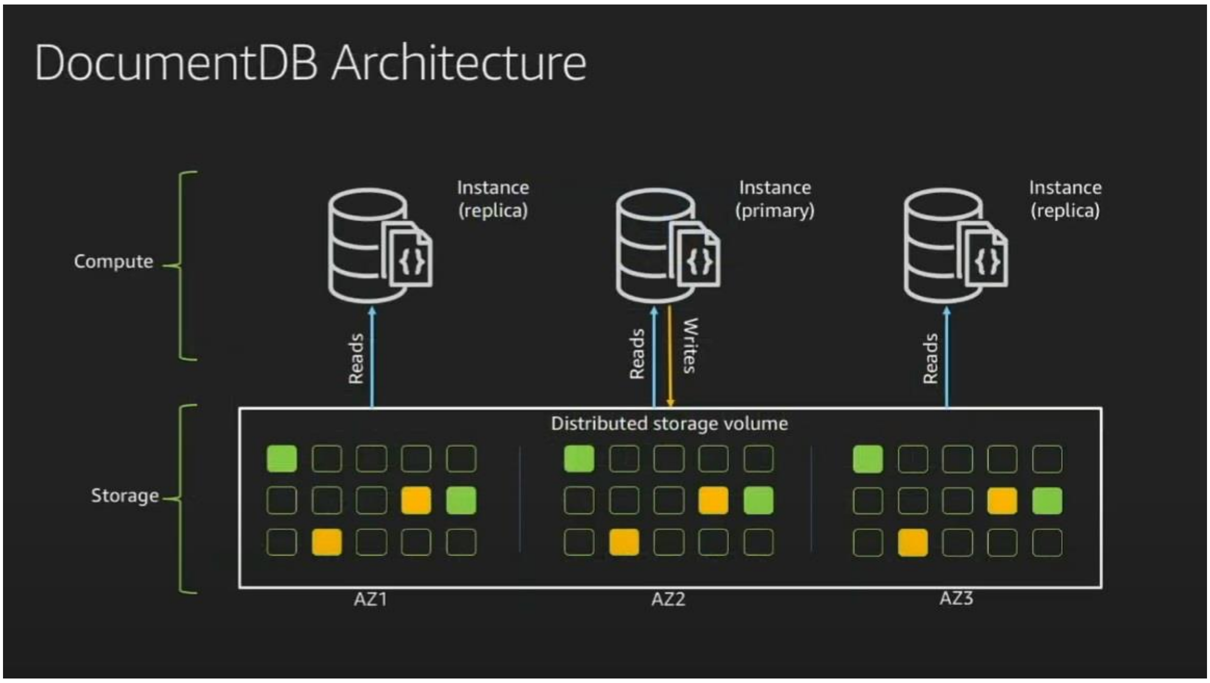
AZ2

AZ3

2 Replicate 10-GB data segments 6× across 3 AZs

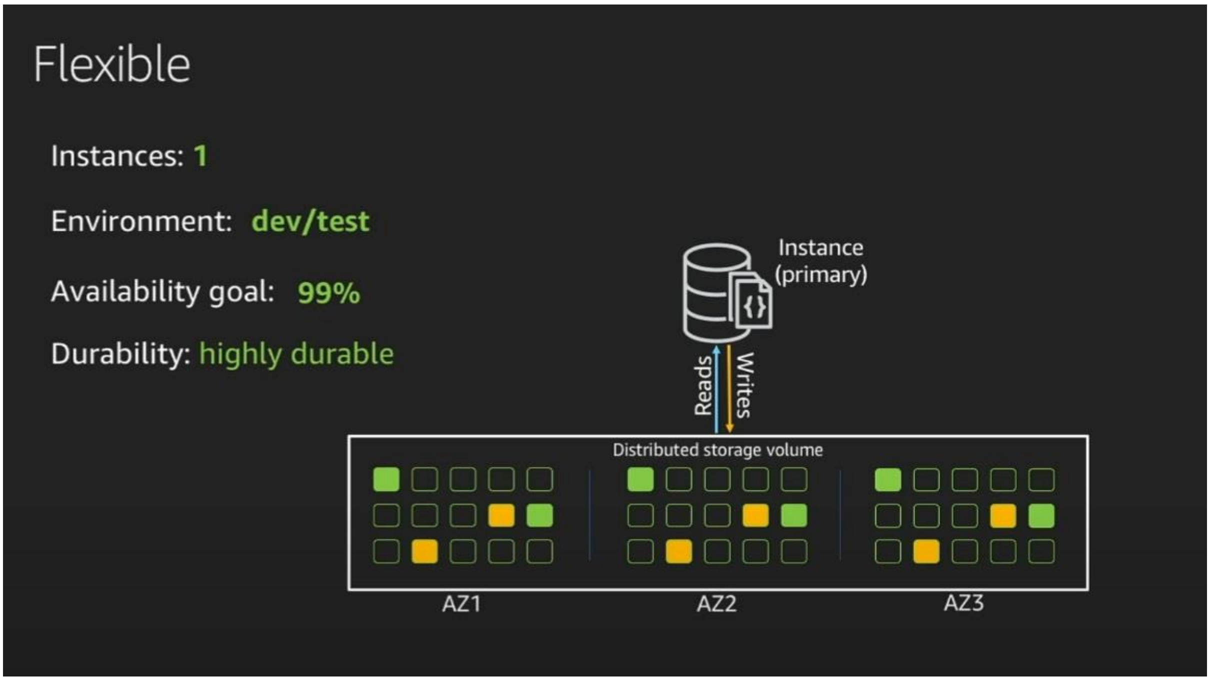
We can see that the two, horizontal blue rectangles have been preserved from the previous slide.

You will notice arrays of rectangles in the image above. Perhaps these represent storage volumes. Each array contains fifteen rectangles. Some of the rectangles are coloured in, having a solid green colour. Others have a solid orange colour. Why are there two colours being used? In each availability zone, the rectangles which are coloured in are the same. For example, in all three availability zones, the top left rectangle is solid green. In all three zones, on the bottom row, there is a box coloured in orange which is second from the left.



DocumentDB Architecture

The slide above shows that one instance, of the three, is performing both reads and writes. Two of the compute instances are only reading the storage. We designate the instance that is both reading and writing as the *primary* instance. The other two compute instances are designates as *replica* instances.



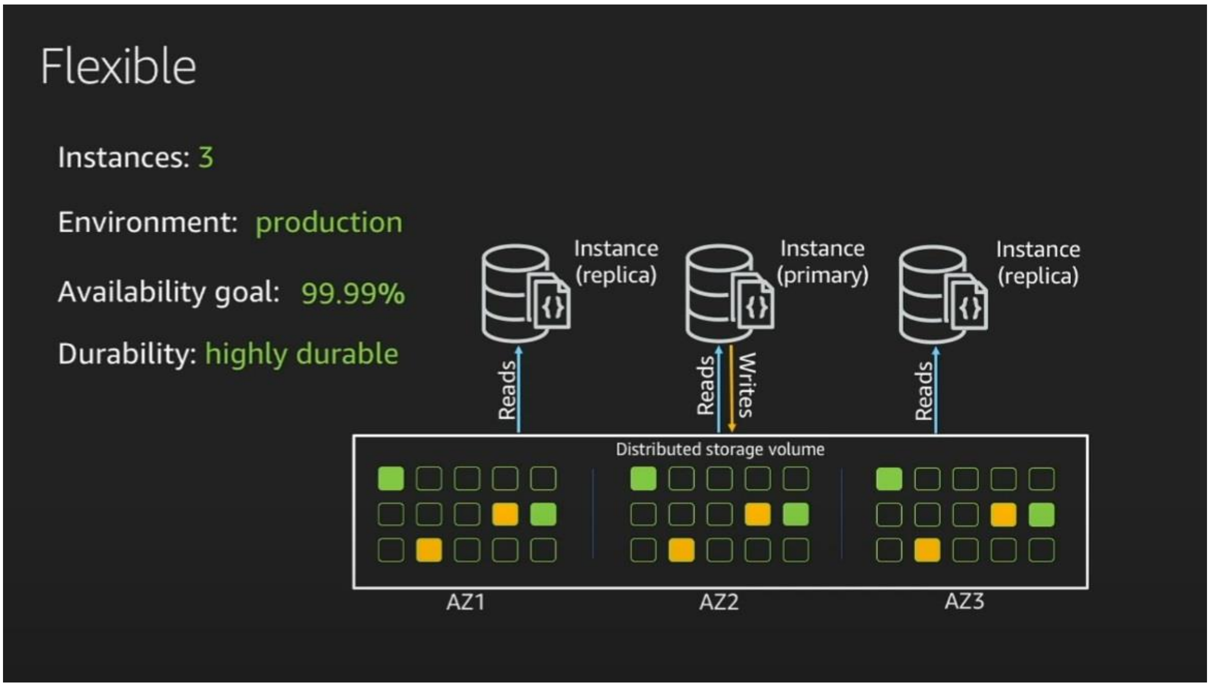
Flexible

Instances: 1

Environment: dev/test

Availability goal: 99%

Durability: highly durable



Flexible

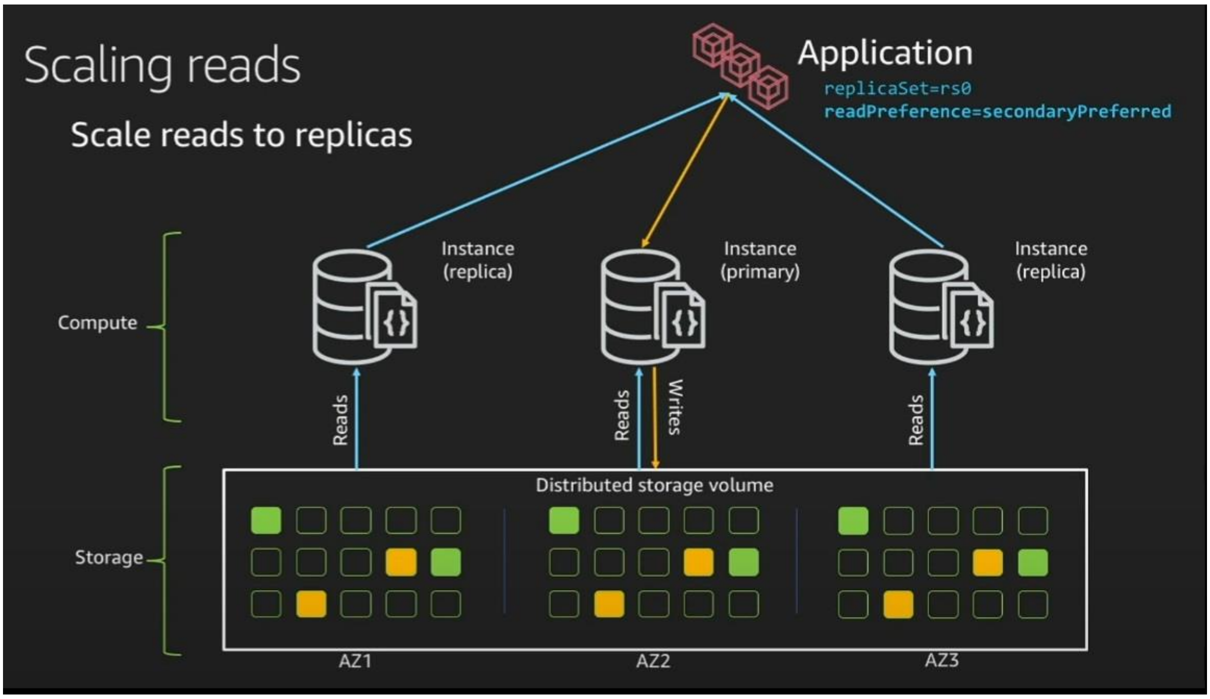
Instances: 3

Environment: production

Availability goal: 99.99%

Durability: highly durable

In the slide above, we can see that the availability goal is much higher, being *four* nines. To achieve this, we need to have three compute instances. Notice how not all of the compute instances need to have the capability to perform writes.



Scaling reads

Scale reads to replicas

Application

```
replicaSet=rs0
readPreference=secondaryPreferred
```

While displaying the slide above, Idziorek says:

> One of the other advantages that we do is—because we push the replication and durability down into the storage layer, the replica instances don't actually participate in the durability of your writes. If you think about

traditional database architectures, what typically happens is you will write to the primary, the primary will send that over to the replica. Both of them kind of do this race to write down to disk on both instances, right?

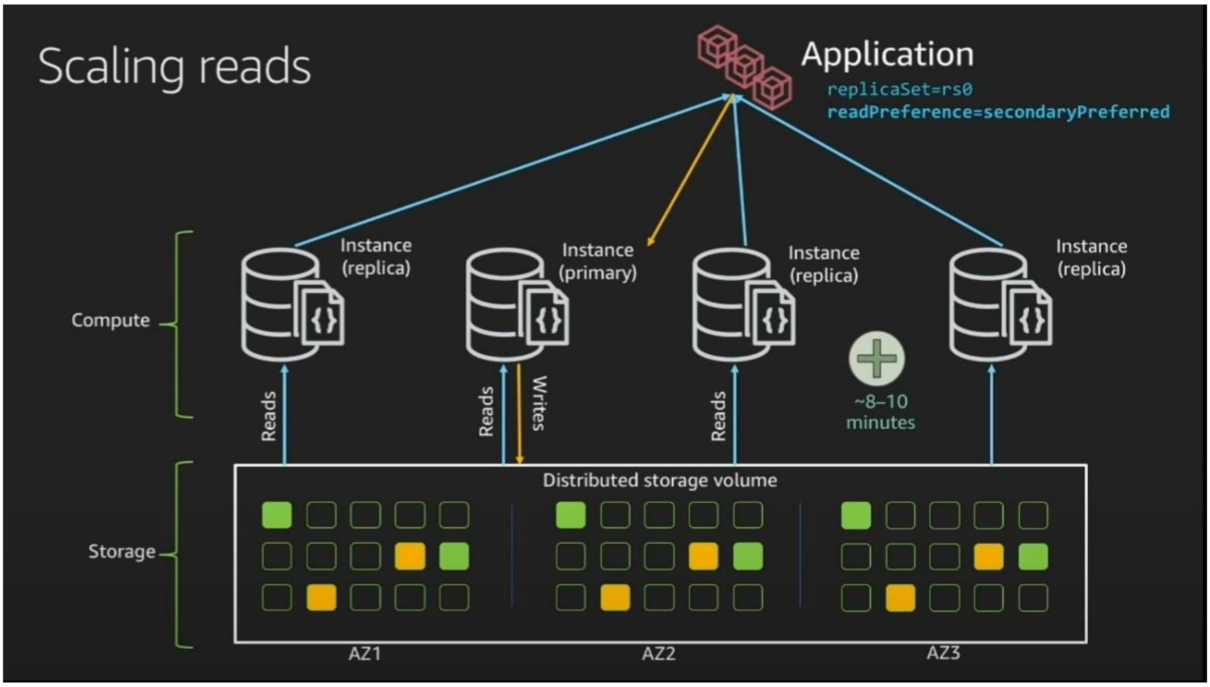
They both come up at some time—coalesce at the primary, say "yes, we both wrote to disk". And then you return back to the client. That's typically how it works with traditional database architectures.

With DocumentDB, the storage layer handles the durability and replication and writes go **through the primary**.

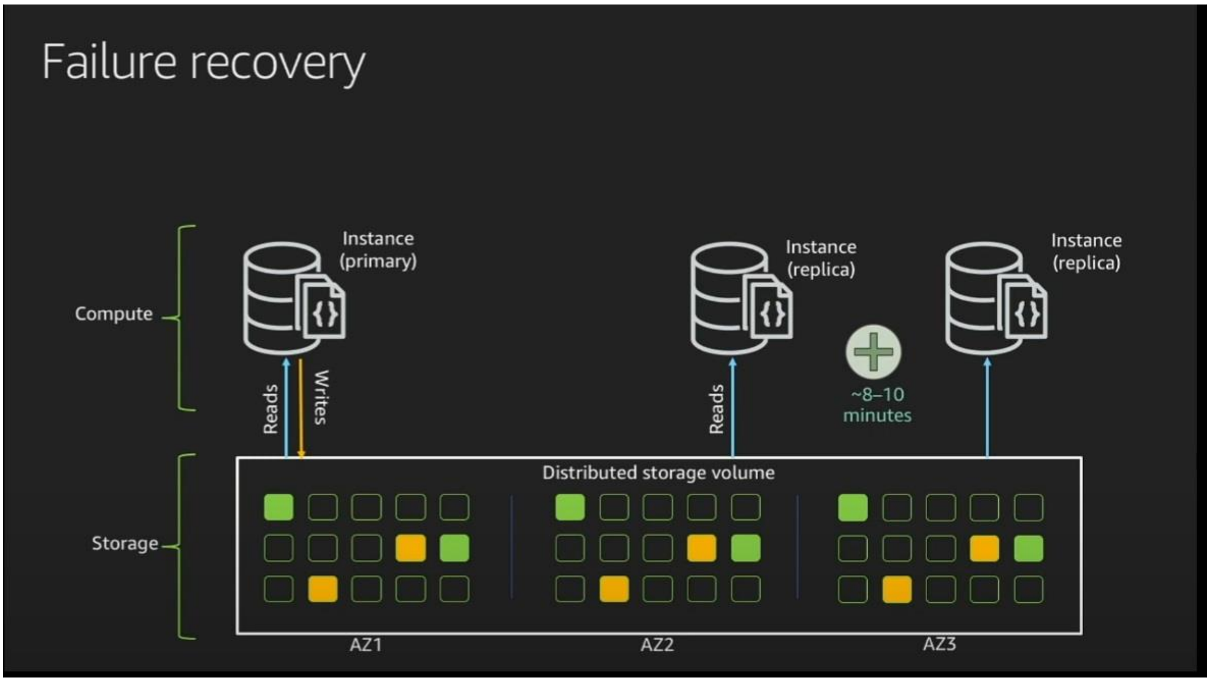
That means the replicas don't participate in the durability of the data, and that frees them up to do a lot more work for you. You're not concerned about writes blocking reads and reads blocking writes, and this confluence of traffic happening on the replicas. They basically become isolated targets for you to be able to have other read workloads work on them. That gives you a lot more bang for the buck, because you can use those as read targets.

So, we actually **highly encourage** our customers to connect to their clusters as a Replica Set, which means that our client is aware of the cluster topography, and then be able to choose a read preference of "Secondary Preferred" – which means that if there are secondaries, read from them [and] if not, read from the primary.

[Idziorek 2019: 18:28]



Scaling reads



Failure recovery

Joseph Idziorek says:

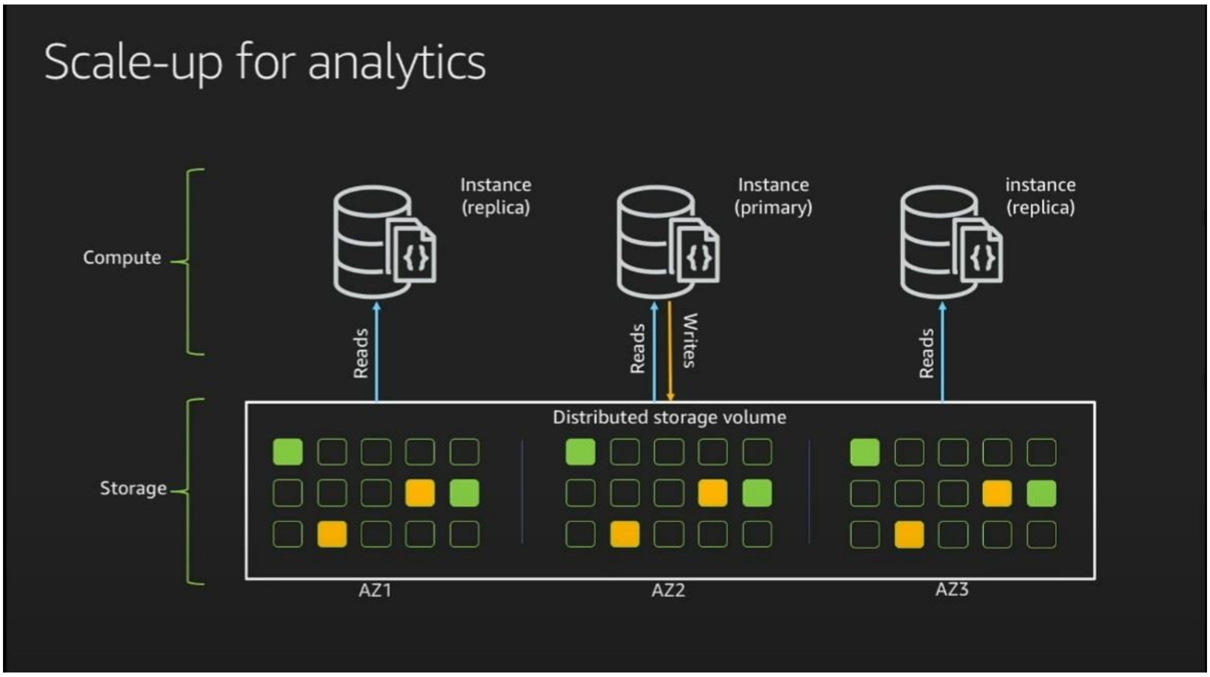
> Failure recovery is the same way. What we usually talk about in databases is how long it takes to do a failover. With DocumentDB, it's typically 30 seconds. What we often *don't* talk about, though, is how long it takes your

database to recover [that thing you lost]. Conceivably, that database is there because you want it or need it.

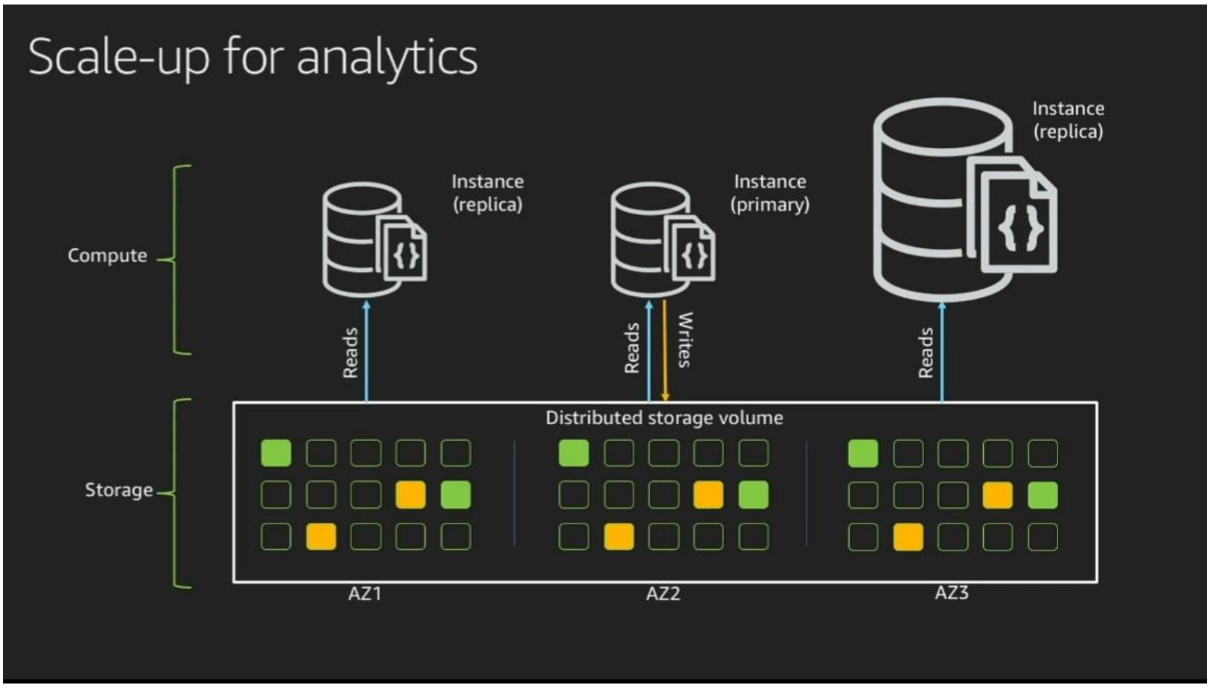
Those are things that you want to be there, because you paid for it, and if you paid for it, then you want it there. So, How long does it take to recover that?

Again, it's a function of eight to ten minutes to spin up that EC2 instance.

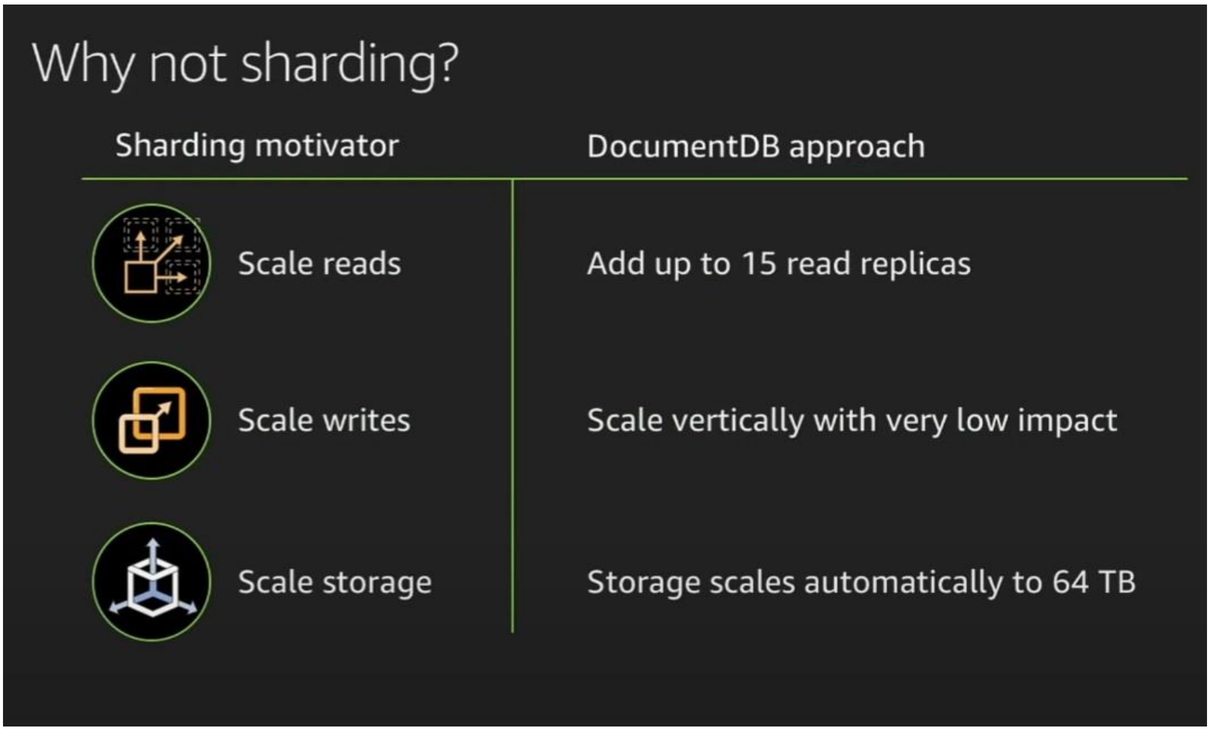
[Idziorek 2019: 20:41]



Scale-up for analytics



Scale-up for analytics



Why not sharding?

Sharding motivator	DocumentDB approach
Scale reads	Add up to 15 read replicas
Scale writes	Scale vertically with very low impact
Scale storage	Storage scales automatically to 64 TB

Glossary

Start-stop cluster – a feature found in DocumentDB.

MongoDB –

ReplicaSet -

Bibliography

Image Credits

[Idziorek 2010]

"Idziorek uses IBM PhD Fellowship to research cloud computing security". Joseph Idziorek has found his place in cloud computing, and IBM has noticed. Apr 5, 2010.

[Insider]

US President Donald Trump stood next to a large pile of documents. *Business Insider*. Source: https://www.businessinsider.com/trump-stood-next-to-a-huge-pile-of-paper-showing-big-government-2017-12?r=US&IR=T

[Idziorek 2019]

Idziorek, Joseph and Antra Grover (2019). Amazon DocumentDB deep dive. *ReInvent* [Conference]. Available at: <https://www.youtube.com/watch?v=D3_hWN9C9iE&ab_channel=AWSEvents>

https://cacm.acm.org/blogs/blog-cacm/252974-understanding-nosql-database-types-document/fulltext

AWS Backup

1. What are the important issues when it comes to backing up data?
2. Why do we need AWS Backup?
3. What are the disadvantages of AWS Backup?

Before the launch of AWS Backup, customers had to separately schedule backups from native service consoles. This overhead was also present when there was a need to change backup schedules, or initiate a restore across multiple AWS services. AWS Backup solves this problem by providing customers with a single pane of glass to create/maintain backup schedules, perform restores, and monitor backup/restore jobs.

[Fiore 2019]

Additional Tips

Here are some additional hints and tips for working with AWS Backup:

Resource Relationships

There may be cases where, for audit or compliance purposes, you must identify the relationship of a deleted resource with a recovery point in AWS Backup. For example, you may have EBS volume snapshots going back a number of years after an underlying Amazon EC2 instance was terminated. You must be able to provide evidence to an auditor that the EBS volume for your snapshot was associated with the terminated EC2 instance.

For situations like these, I recommend enabling AWS Config configuration recording of your AWS resources. This helps identify and track AWS resource relationships (including deleted resources) for up to seven years.

Backup Overlaps

If you use any scripts or AWS Lambda functions to take snapshots of AWS resources that are also being protected by AWS Backup, I recommend ensuring that there is no overlap between AWS Backup and your scripts/Lambda functions, as this can lead to backup job failures.

Regional Resource Assignments

AWS Backup supports resource assignments within the same Region. Separate backup plans must be created to back up AWS resources within that Region. For an up-to-date listing of Regions currently supported by AWS Backup, see the AWS Region table.

Restore Validation

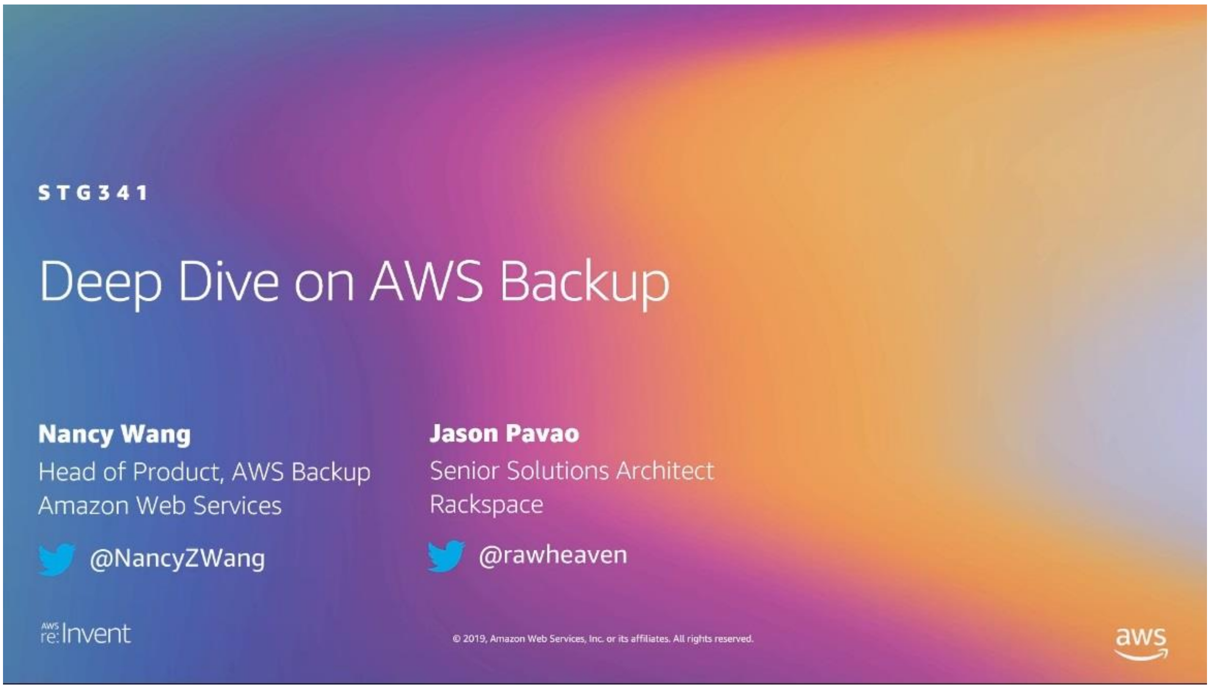
Generally, the most comprehensive data-protection strategies include regular testing and validation of your restore procedures before you need them. Testing your restores also helps in preparing and maintaining recovery runbooks. That, in turn, ensures operational readiness during a disaster recovery exercise, or an actual data loss scenario.

Encryption Permissions

When using AWS Backup with encrypted resources, such as EBS volumes, the AWS Backup IAM service role must be granted permissions to the AWS KMS keys used to encrypt your resources. For more information about adding the default service role *AWSBackupDefaultServiceRole* as a new KMS key user, see Editing Keys.

Snapshot Limits

As your AWS footprint grows over time, the number of snapshots in your account will also grow. You will want to review your service limits on a regular basis to ensure you aren’t in danger of getting close to snapshot-related service limits, which can cause your backups to fail. An easy way to keep track of these and other service limits is to utilize AWS Trusted Advisor , as it will report on major service limits regardless of which support plan you subscribe to. In the event that you notice any service limits reaching the Yellow or Red Criteria, you can open a Service Limit Increase case for the protected service (i.e. RDS, EBS, etc.) through the AWS Support Center. You can find more information on managing service limits and requesting service limit increases in our AWS Support Knowledge Center



STG341

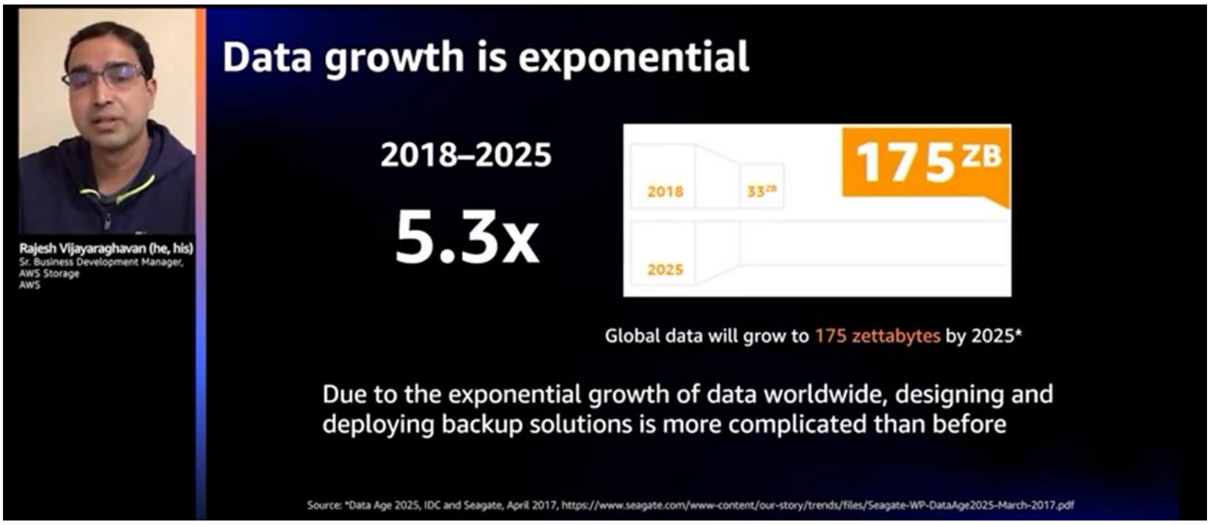
Deep Dive on AWS Backup

Nancy Wang
Head of Product, AWS Backup
Amazon Web Services
@NancyZWang

Jason Pavao
Senior Solutions Architect
Rackspace
@rawheaven

AWS re:Invent

aws



Rajesh Vijayaraghavan (he, his)
Sr. Business Development Manager,
AWS Storage
AWS

Data growth is exponential

2018–2025

5.3x

Global data will grow to 175 zettabytes by 2025*

Due to the exponential growth of data worldwide, designing and deploying backup solutions is more complicated than before

Source: *Data Age 2025, IDC and Seagate, April 2017, https://www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf

Presentation by Rajesh Vijayaraghavan

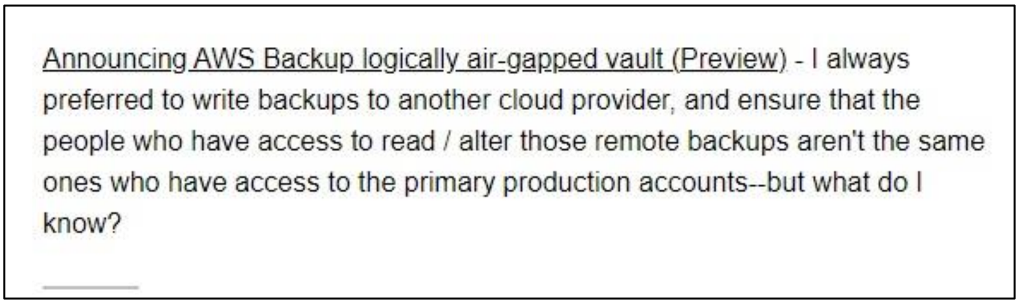
Announcing AWS Backup logically air-gapped vault (Preview)

Posted On: Aug 9, 2023

Today, AWS Backup announces the preview of logically air-gapped vault, a new type of AWS Backup Vault that allows secure sharing of backups across accounts and organizations, supporting direct restore to help reduce recovery time from a data loss event. AWS Backup is a fully managed service that centralizes and automates data protection across AWS services and hybrid workloads. Logically air-gapped vault stores immutable backup copies that are locked by default, and isolated with encryption using AWS owned keys.

Get started with logically air-gapped vault using the AWS Backup console, API, or CLI. Target backups to a logically air-gapped vault by specifying it as a copy destination in your backup plan. The vault can be shared for recovery with other accounts using AWS Resource Access Manager (RAM). Once you share the vault with a recovery or test account, you can initiate direct restore jobs from that account, eliminating the overhead of copying backups first.

AWS Backup logically air-gapped vault (Preview) is available in US East (Virginia) Region. It currently supports backup and restore of Amazon Elastic Compute Cloud (Amazon EC2), Amazon Elastic Block Store (Amazon EBS), Amazon Relational Database Service (Amazon RDS), Amazon Simple Storage Service (Amazon S3), and Amazon Elastic File System (Amazon EFS). For more information on this preview, visit the AWS Backup product page, documentation, and the launch blog. To enroll in this preview, send a request via email to aws-backup-vault-preview@amazon.com.



Announcing AWS Backup logically air-gapped vault (Preview) - I always preferred to write backups to another cloud provider, and ensure that the people who have access to read / alter those remote backups aren't the same ones who have access to the primary production accounts--but what do I know?

Corey Quinn in an email on 14th August 2023

Glossary

Bibliography

[AWS 2023]

Announcing AWS Backup Logically Air-gapped vault. [Announcement]. Available at: https://aws.amazon.com/about-aws/whats-new/2023/08/aws-backup-logically-air-gapped-vault-preview/?ck_subscriber_id=1560524742

[Wang 2019]

Wang, Nancy and Jason Pavao (2019). Deep Dive on AWS Backup. *ReInvent 2019* [Conference]. Available at: <https://www.youtube.com/watch?v=av8DpL0uFjc&ab_channel=AWSEvents>

[Vijayaraghavan 2021]

Vijayaraghavan, Rajesh and Aditya Kalyanakrishnan (2021). Backup, disaster recovery, and ransomware protection with AWS. Available at: <https://www.youtube.com/watch?v=Ru4jxh9qazc&ab_channel=AWSEvents>

[Fiore 2019]

Fiore, Anthony (2019). Protecting your data with AWS Backup. *AWS Storage Blog*. Available at: <https://aws.amazon.com/blogs/storage/protecting-your-data-with-aws-backup/>

[AWS 2019]

Introducing AWS Backup. YouTube Channel: Amazon Web Services. Available at: <https://www.youtube.com/watch?v=QDiXzFx2iMU&ab_channel=AmazonWebServices>

Amazon WorkLink

Bibliography

Amazon Deep Learning Containers

Bibliography

CloudEndure®

Bibliography

AWS Cloud Development Kit

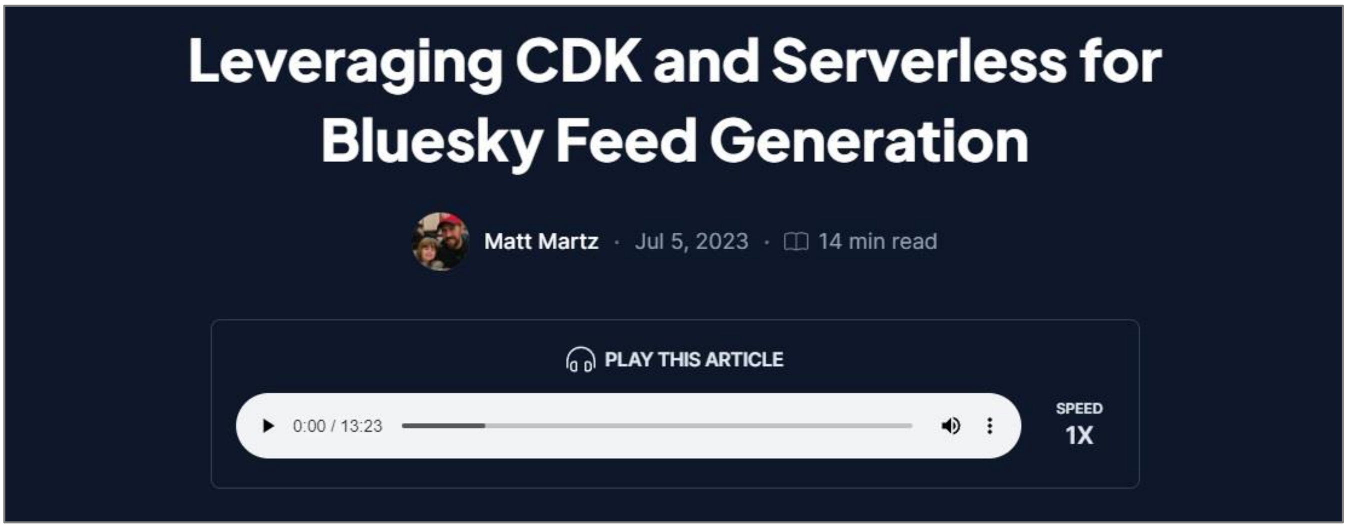
Dixit R Jain
Posted on 2 Jul

💖 3

Lost in AWS CDK? Here's Your Map for Debugging Like a Pro

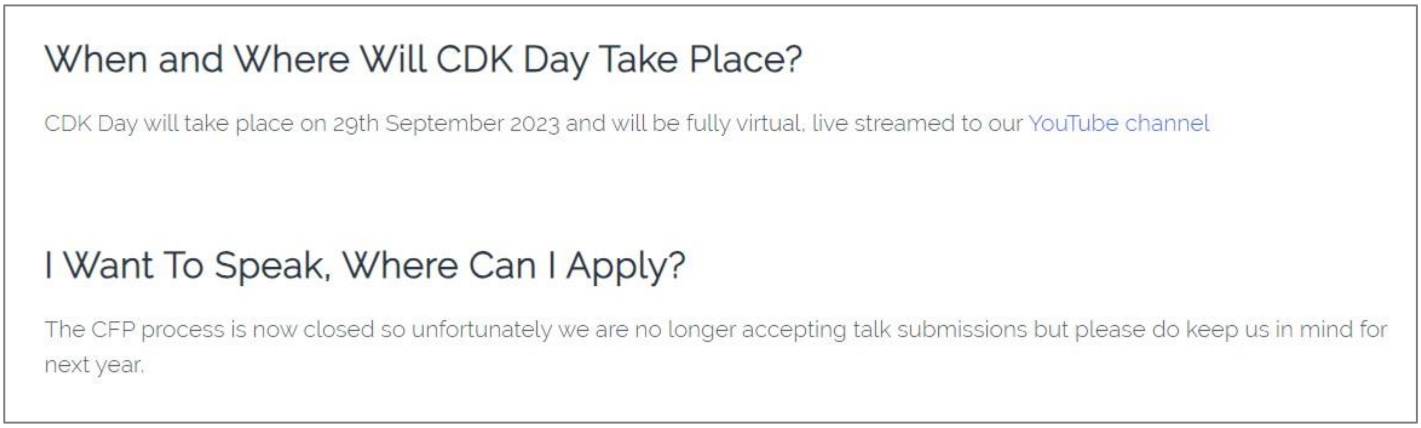
#aws #cdk #debugging #infrastructureascode

Greetings, fellow AWS adventurers! We all know that embarking on a journey through the vast landscape of the Amazon Web Services (AWS) can sometimes feel like navigating through a dense jungle. The AWS Cloud Development Kit (CDK), a powerful Infrastructure as Code tool, is like our trusty machete, slashing through the undergrowth and making the path clearer. But what happens when our machete hits a snag? That's right, we debug!



Leveraging CDK and Serverless for Bluesky Feed Generation

Matt Martz · Jul 5, 2023 · 14 min read

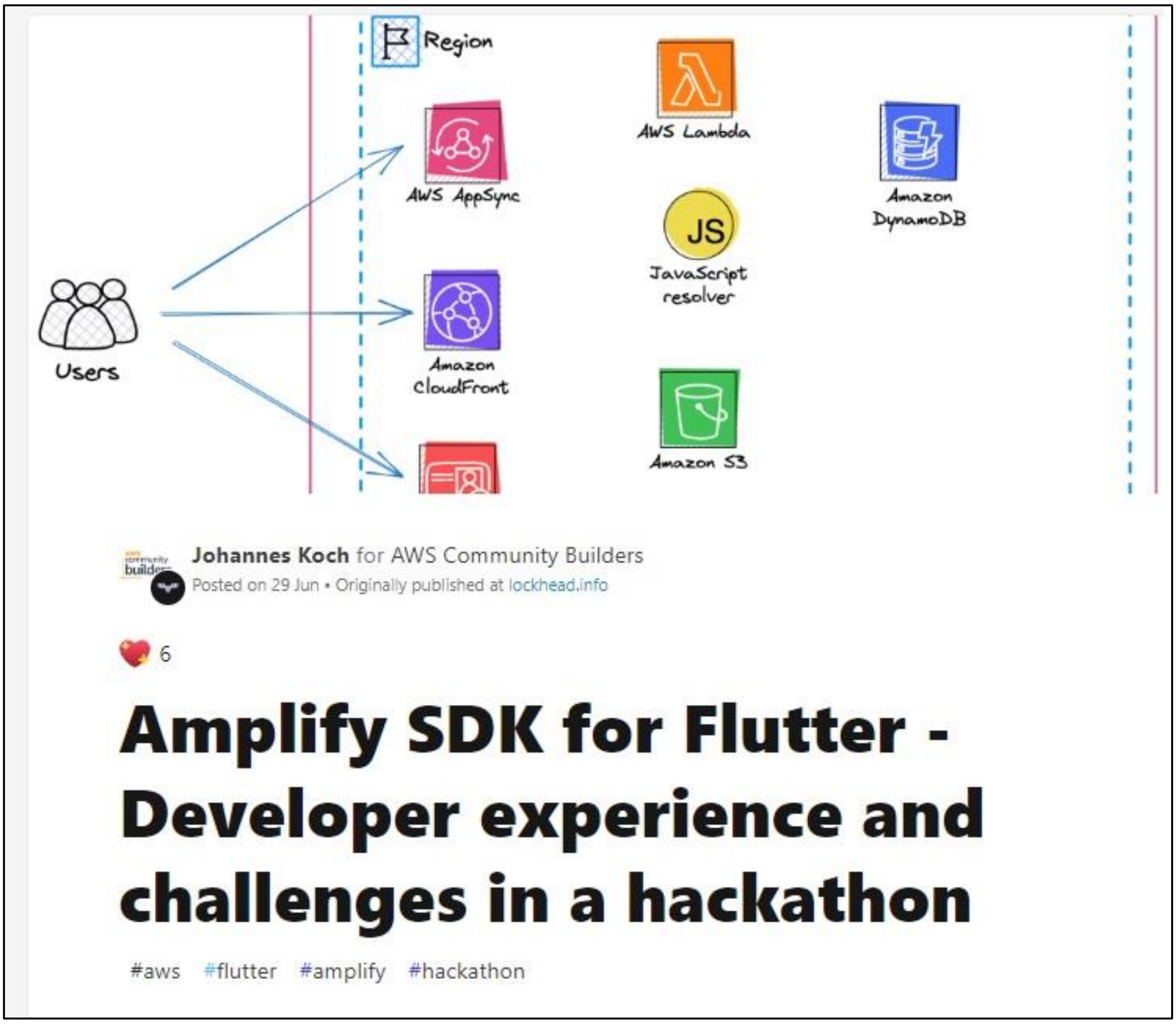


When and Where Will CDK Day Take Place?

CDK Day will take place on 29th September 2023 and will be fully virtual, live streamed to our YouTube channel

I Want To Speak, Where Can I Apply?

The CFP process is now closed so unfortunately we are no longer accepting talk submissions but please do keep us in mind for next year.



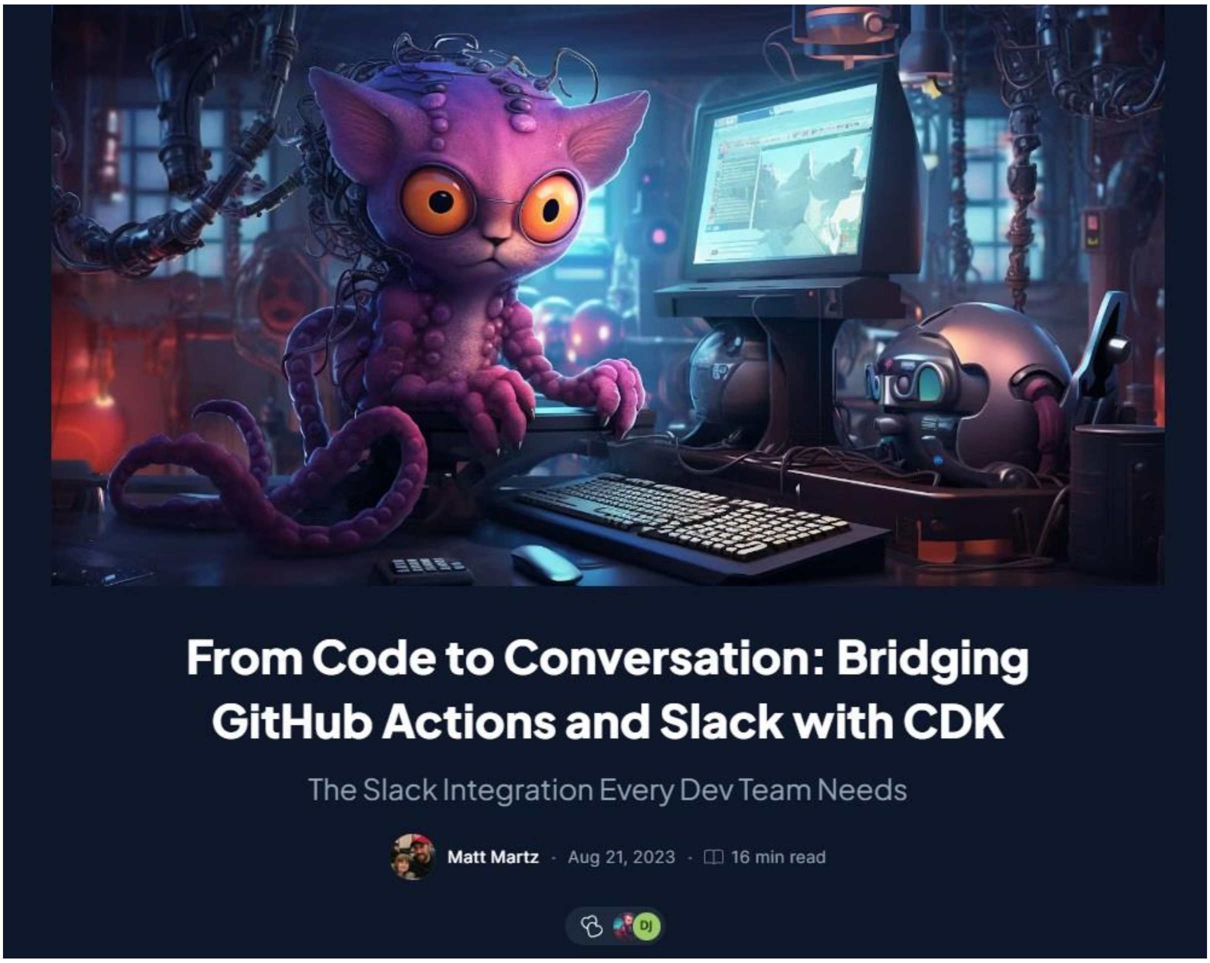
Johannes Koch for AWS Community Builders

Posted on 29 Jun • Originally published at lockhead.info

❤️ 6

Amplify SDK for Flutter - Developer experience and challenges in a hackathon

#aws #flutter #amplify #hackathon



From Code to Conversation: Bridging GitHub Actions and Slack with CDK

The Slack Integration Every Dev Team Needs

Matt Martz · Aug 21, 2023 · 16 min read

Bibliography

https://garden.io/blog/aws-security-issue?ck_subscriber_id=1560524742

Millican, Peter (2011). Abstraction and Idealism. Available at: <https://www.youtube.com/watch?v=B2OgmlMlL90&ab_channel=UniversityofOxford>

Werner Vogels on the AWS Cloud Development Kit (AWS CDK). Available at: <https://www.youtube.com/watch?v=AYYTrDaEwLs&ab_channel=AWSEvents>

[Martz 2023]

Martz, Matt (2023). From Code to Conversation: Bridging GitHub Actions and Slack with CDK. Available at: <https://matt.martz.codes/cdk-slackbot?ck_subscriber_id=1560524742>

EventBridge

Introducing Amazon EventBridge

Posted On: Jul 11, 2019

Amazon EventBridge is a serverless event bus that makes it easy to connect applications together using data from your own applications, Software-as-a-Service (SaaS) applications, and AWS services. EventBridge delivers a stream of real-time data from event sources, such Zendesk, Datadog, or Pagerduty, and routes that data to targets like AWS Lambda. You can set up routing rules to determine where to send your data to build application architectures that react in real time to all of your data sources. EventBridge allows you to build event-driven architectures, which are loosely coupled and distributed. This improves developer agility as well as application resiliency. EventBridge makes it easy to build event-driven applications because it takes care of event ingestion and delivery, security, authorization, and error-handling for you. EventBridge leverages the CloudWatch Events API, so CloudWatch Events users can access their existing default bus, rules, and events in the new EventBridge console, as well as in the CloudWatch Events console.

EventBridge supports 10 SaaS application partners and 90+ AWS Services as event sources. To become an EventBridge SaaS integration partner, please visit the EventBridge partner page.

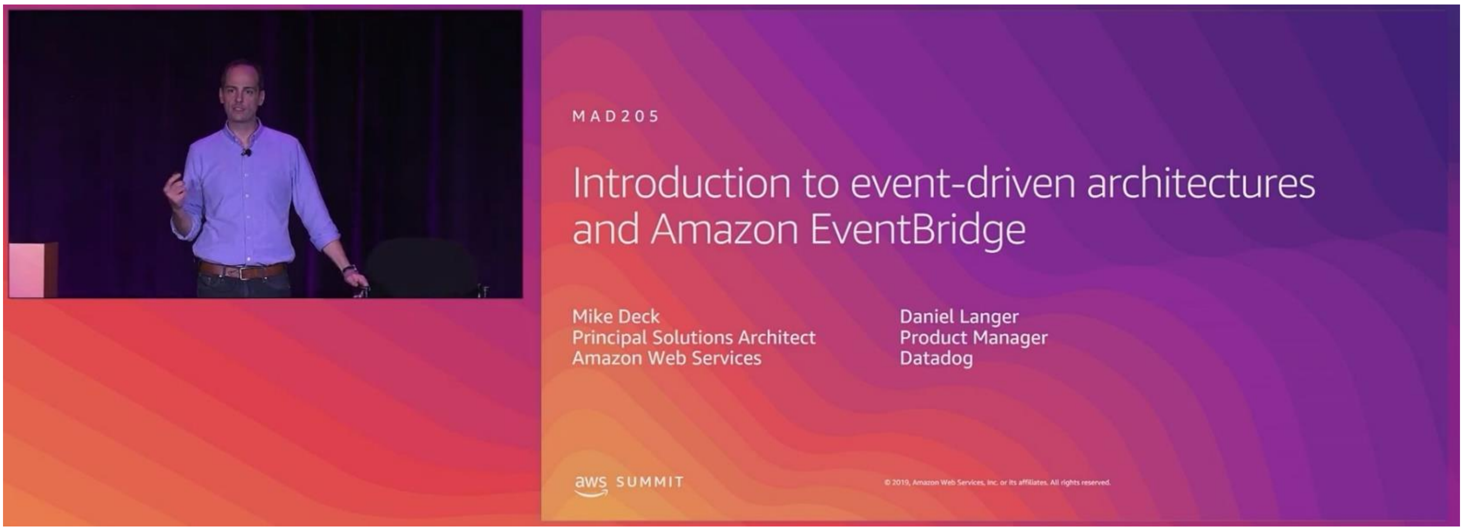
There is no up-front cost commitment or minimum fee for using EventBridge, and you pay only for the events published to your event bus. All state change events published by AWS services are free.

You can get started with EventBridge using the AWS Management Console, Command Line Interface (CLI), or SDK. You can create a new event bus and receive events from SaaS applications in minutes; then simply create a rule to match events from a list of AWS services or SaaS applications and proceed to set up targets for your events.

EventBridge is generally available today in the following regions: US East (Ohio and N. Virginia), US West (Oregon and N. California), Canada (Central), EU (Stockholm, Paris, Ireland, Frankfurt, and London), Asia Pacific (Mumbai, Tokyo, Hong Kong, Seoul, Singapore, and Sydney), and South America (Sao Paulo).

To learn more about EventBridge:

- Read the AWS News blog post
- Visit the Amazon EventBridge product page
- Read the Amazon EventBridge Developer Guide



MAD205

Introduction to event-driven architectures and Amazon EventBridge

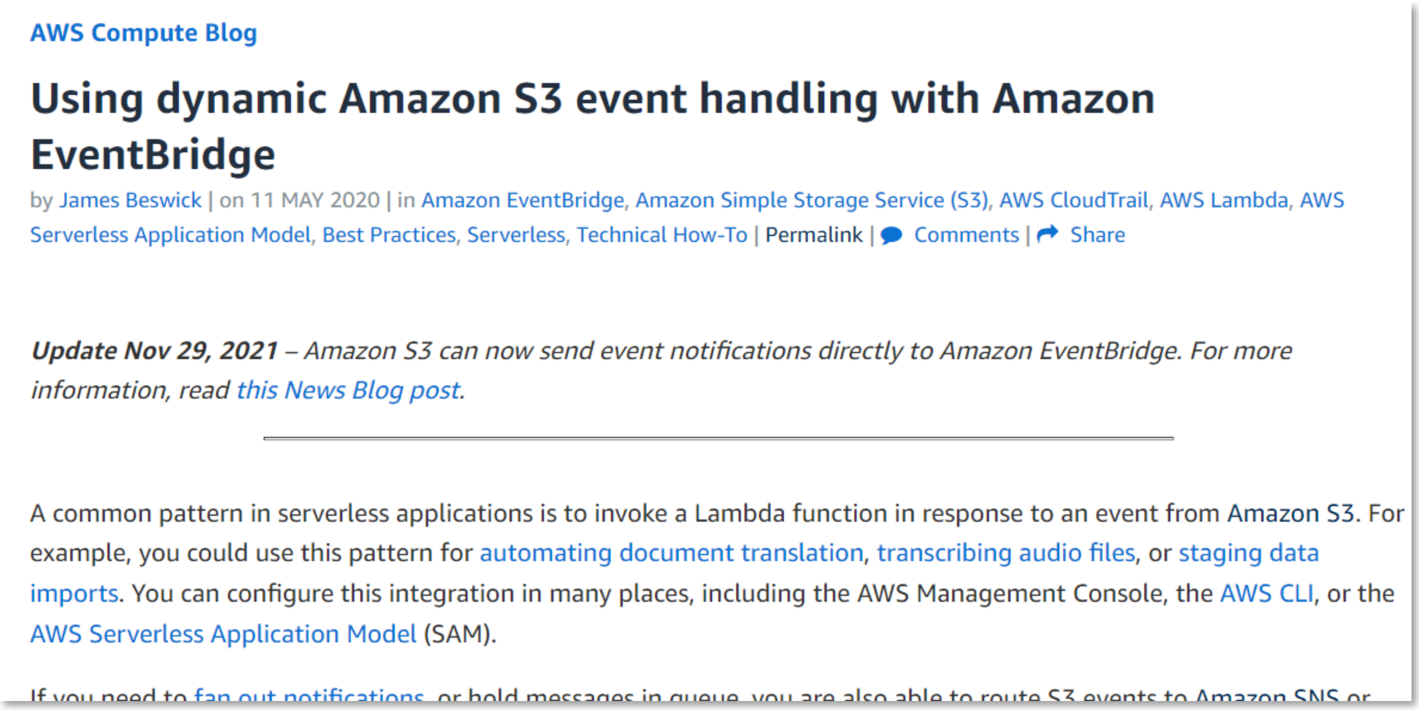
Mike Deck
Principal Solutions Architect
Amazon Web Services

Daniel Langer
Product Manager
Datadog

aws SUMMIT

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What happens in the presentation?



AWS Compute Blog

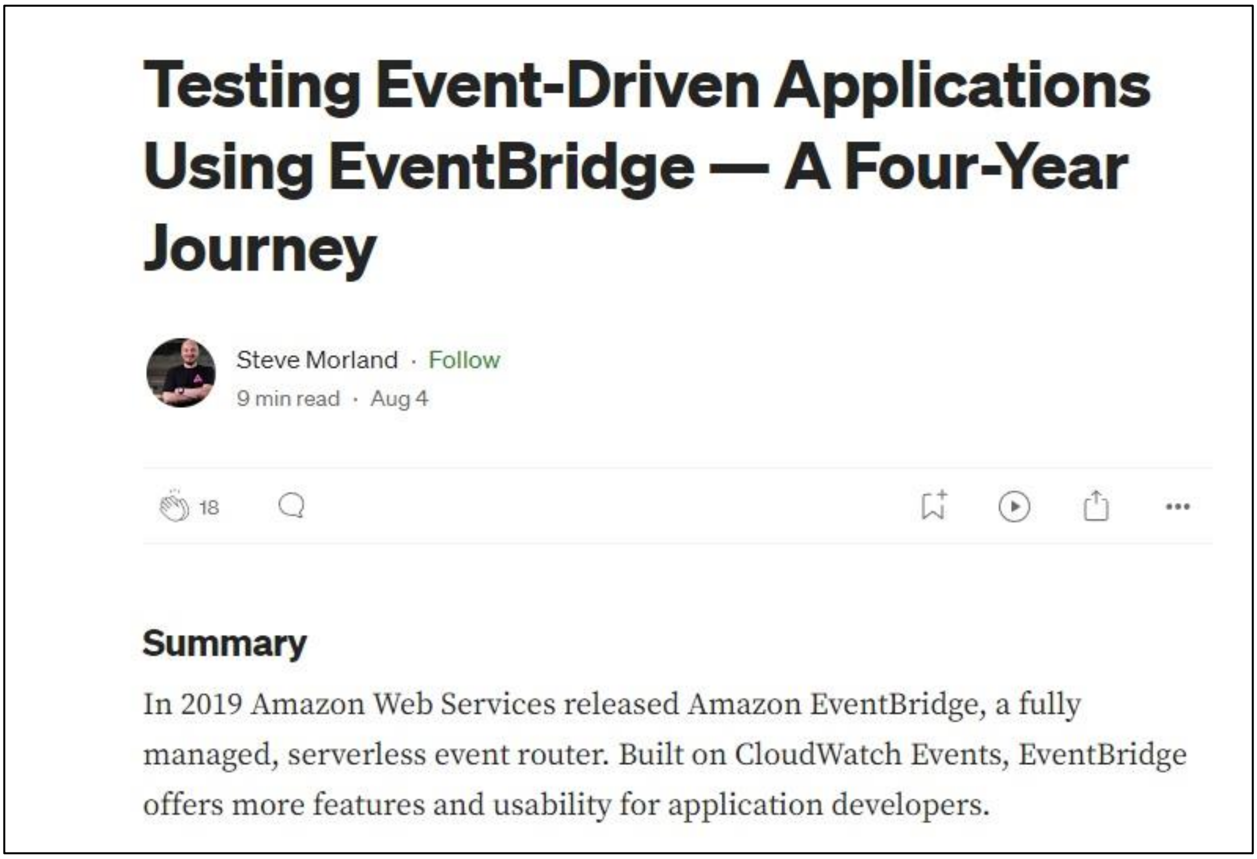
Using dynamic Amazon S3 event handling with Amazon EventBridge

by James Beswick | on 11 MAY 2020 | in Amazon EventBridge, Amazon Simple Storage Service (S3), AWS CloudTrail, AWS Lambda, AWS Serverless Application Model, Best Practices, Serverless, Technical How-To | Permalink | Comments | Share

***Update Nov 29, 2021** – Amazon S3 can now send event notifications directly to Amazon EventBridge. For more information, read this News Blog post.*

A common pattern in serverless applications is to invoke a Lambda function in response to an event from Amazon S3. For example, you could use this pattern for automating document translation, transcribing audio files, or staging data imports. You can configure this integration in many places, including the AWS Management Console, the AWS CLI, or the AWS Serverless Application Model (SAM).

If you need to fan out notifications, or hold messages in queue, you are also able to route S3 events to Amazon SNS or



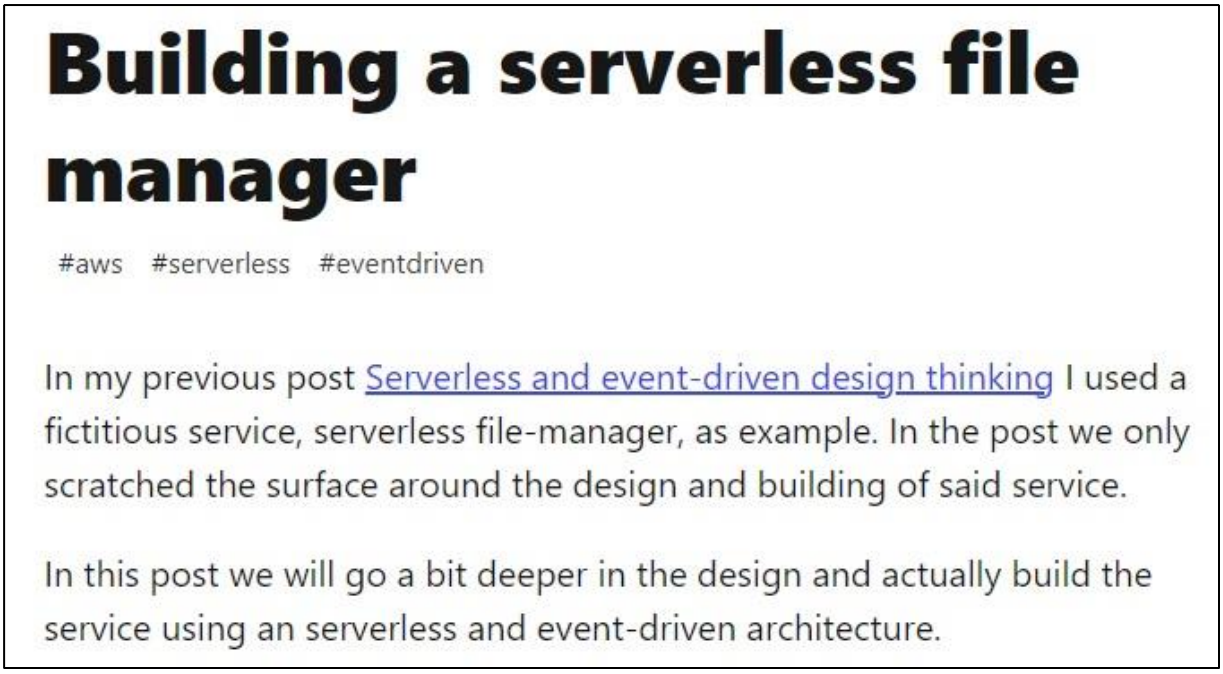
Testing Event-Driven Applications Using EventBridge — A Four-Year Journey

Steve Morland · Follow

9 min read · Aug 4

Summary

In 2019 Amazon Web Services released Amazon EventBridge, a fully managed, serverless event router. Built on CloudWatch Events, EventBridge offers more features and usability for application developers.

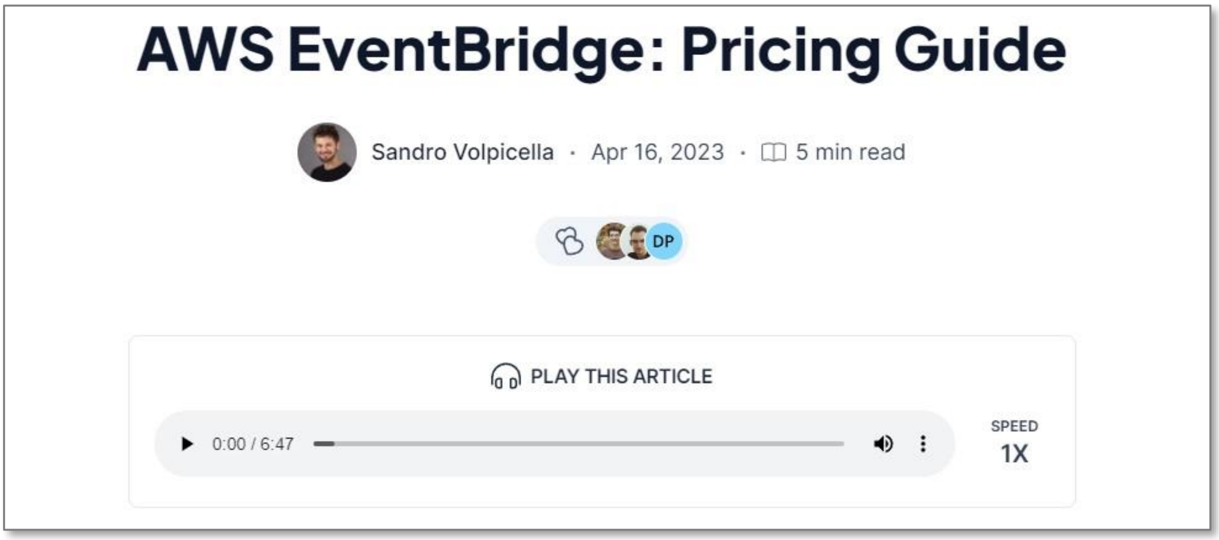


Building a serverless file manager

#aws #serverless #eventdriven

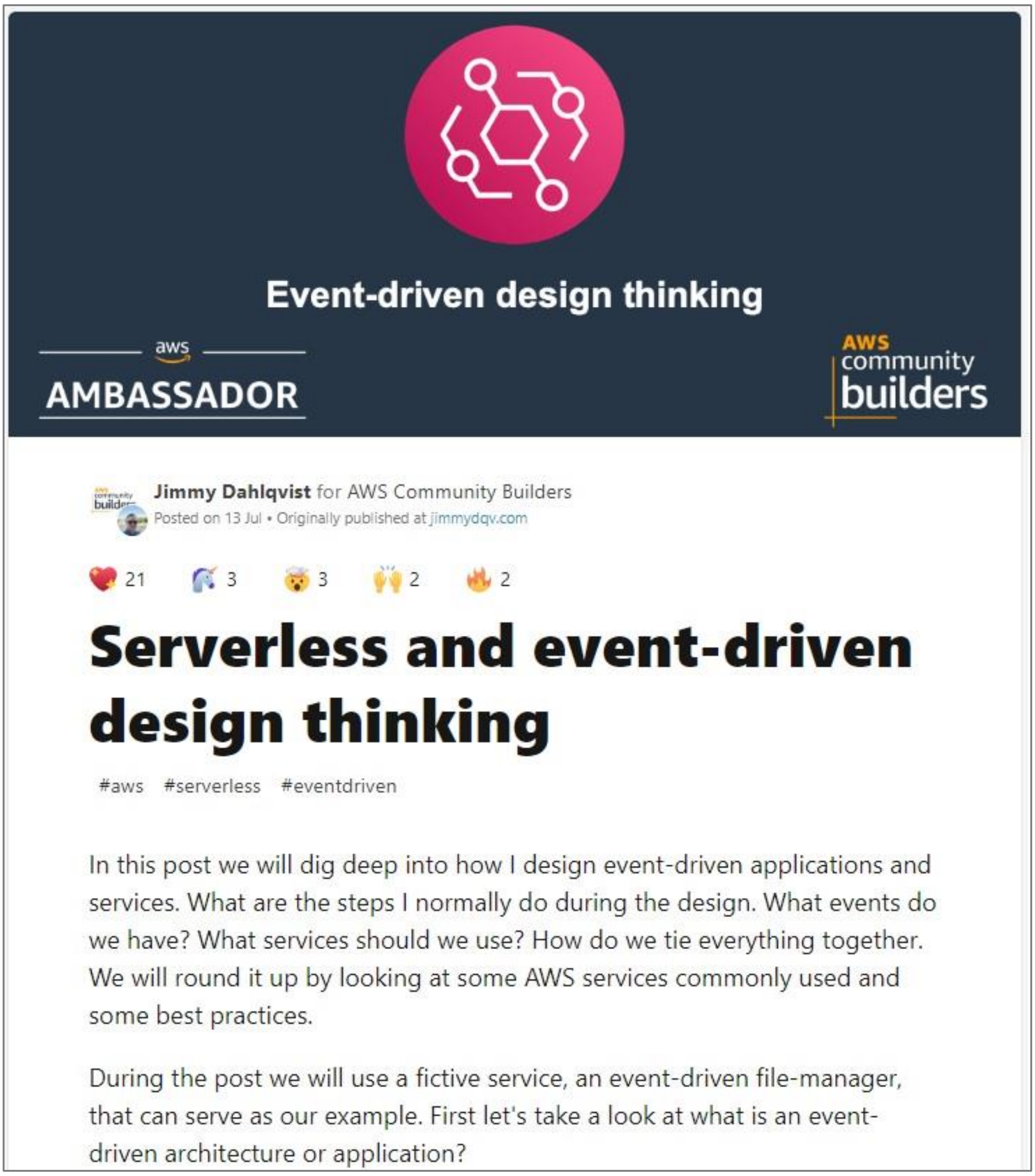
In my previous post Serverless and event-driven design thinking I used a fictitious service, serverless file-manager, as example. In the post we only scratched the surface around the design and building of said service.

In this post we will go a bit deeper in the design and actually build the service using an serverless and event-driven architecture.



AWS EventBridge: Pricing Guide

Sandro Volpicella · Apr 16, 2023 · 5 min read



Jimmy Dahlqvist for AWS Community Builders
Posted on 13 Jul • Originally published at jimmydqv.com

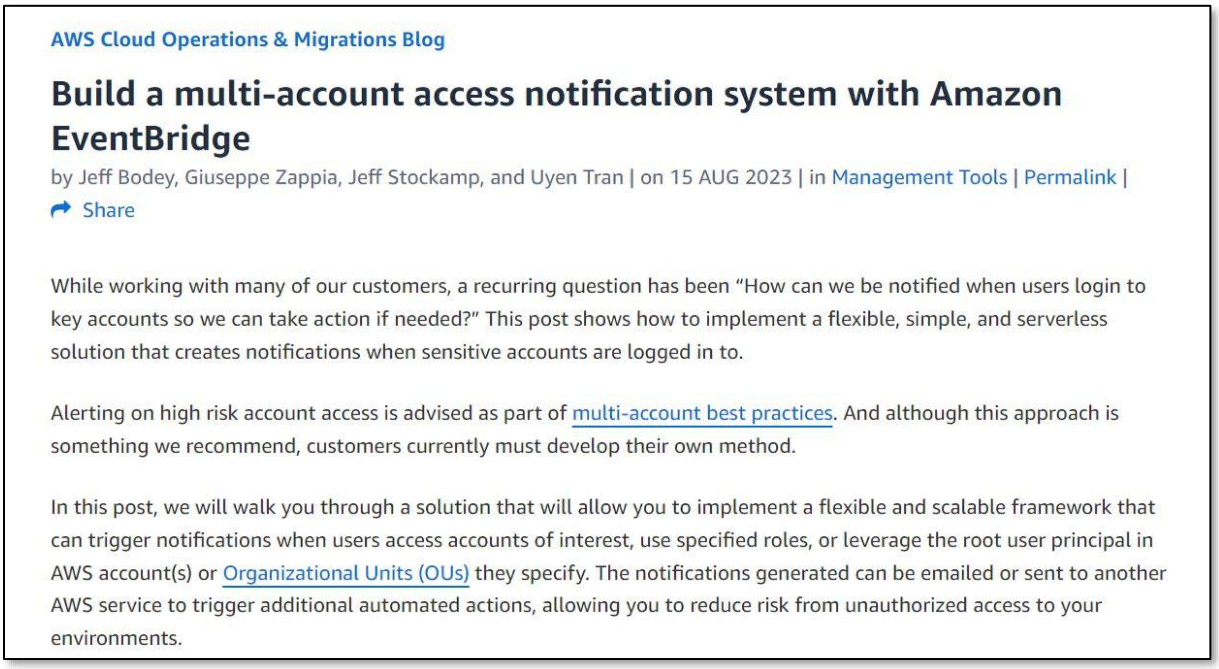
❤️ 21 🦄 3 🤯 3 🙌 2 🔥 2

Serverless and event-driven design thinking

#aws #serverless #eventdriven

In this post we will dig deep into how I design event-driven applications and services. What are the steps I normally do during the design. What events do we have? What services should we use? How do we tie everything together. We will round it up by looking at some AWS services commonly used and some best practices.

During the post we will use a fictive service, an event-driven file-manager, that can serve as our example. First let's take a look at what is an event-driven architecture or application?



AWS Cloud Operations & Migrations Blog

Build a multi-account access notification system with Amazon EventBridge

by Jeff Bodey, Giuseppe Zappia, Jeff Stockamp, and Uyen Tran | on 15 AUG 2023 | in Management Tools | Permalink |

Share

While working with many of our customers, a recurring question has been "How can we be notified when users login to key accounts so we can take action if needed?" This post shows how to implement a flexible, simple, and serverless solution that creates notifications when sensitive accounts are logged in to.

Alerting on high risk account access is advised as part of multi-account best practices. And although this approach is something we recommend, customers currently must develop their own method.

In this post, we will walk you through a solution that will allow you to implement a flexible and scalable framework that can trigger notifications when users access accounts of interest, use specified roles, or leverage the root user principal in AWS account(s) or Organizational Units (OUs) they specify. The notifications generated can be emailed or sent to another AWS service to trigger additional automated actions, allowing you to reduce risk from unauthorized access to your environments.

Glossary

Golden AMI - A *golden AMI* is an AMI that contains the latest security patches, software, configuration, and software agents that you need to install for logging, security maintenance, and performance monitoring. [Waikar 2017]

Bibliography

I. Official

[AWS 2024]

Amazon EventBridge Pipes now supports event delivery through AWS PrivateLink [Announcement]. Available at: < https://aws.amazon.com/about-aws/whats-new/2024/04/amazon-eventbridge-pipes-event-delivery-aws-privatelink/>

[Brisals 2023]

Brisals, Sheen (2023). Advanced event-driven patterns with Amazon EventBridge. Reinvent 2023 [Conference]. Available at: <https://www.youtube.com/watch?v=6X4lSPkn4ps&t=2224s&ab_channel=AWSEvents>

II. Unofficial

Morland, Steve (2023). Testing Event-Driven Applications Using EventBridge – A Four-Year Journey. Available at: <https://medium.com/@smorland/testing-event-driven-applications-using-eventbridge-a-four-year-journey-8ed0c80fc4ca>

Building a serverless file manager. Available at: <https://dev.to/aws-builders/building-a-serverless-file-manager-1k01?utm_source=substack&utm_medium=email>

https://blog.awsfundamentals.com/aws-eventbridge-pricing-guide?utm_source=substack&utm_medium=email

Dahlqvist, Jimmy (2023). Serverless and event-driven design thinking. Available at: <https://dev.to/aws-builders/serverless-and-event-driven-design-thinking-24jf?utm_source=substack&utm_medium=email>

[Waikar 2017]

Waikar, Kanchan and David Aiken (2017). How to set up Continuous Golden AMI Vulnerability Assessments with Amazon Inspector. *AWS Security Blog*. Available at: <https://aws.amazon.com/blogs/security/how-to-set-up-continuous-golden-ami-vulnerability-assessments-with-amazon-inspector/>

[Sodkiewicz 2023]

Sodkiewicz, Marcin (2023). Distributed Circuit Breakers in Event-Driven Architectures on AWS. April 30th 2024. *Medium*. Available at: <https://sodkiewiczm.medium.com/distributed-circuit-breakers-in-event-driven-architectures-on-aws-95774da2ce7e>

Data Exchange

Bibliography

Launch Wizard ←

Bibliography

Savings Plans ←

Saving Plans

Kenneth Clarke is acknowledged by those on both sides of the political spectrum to be one of the United Kingdom's greatest post-war Chancellors of the Exchequer

Introducing Savings Plans

Posted On: Nov 6, 2019

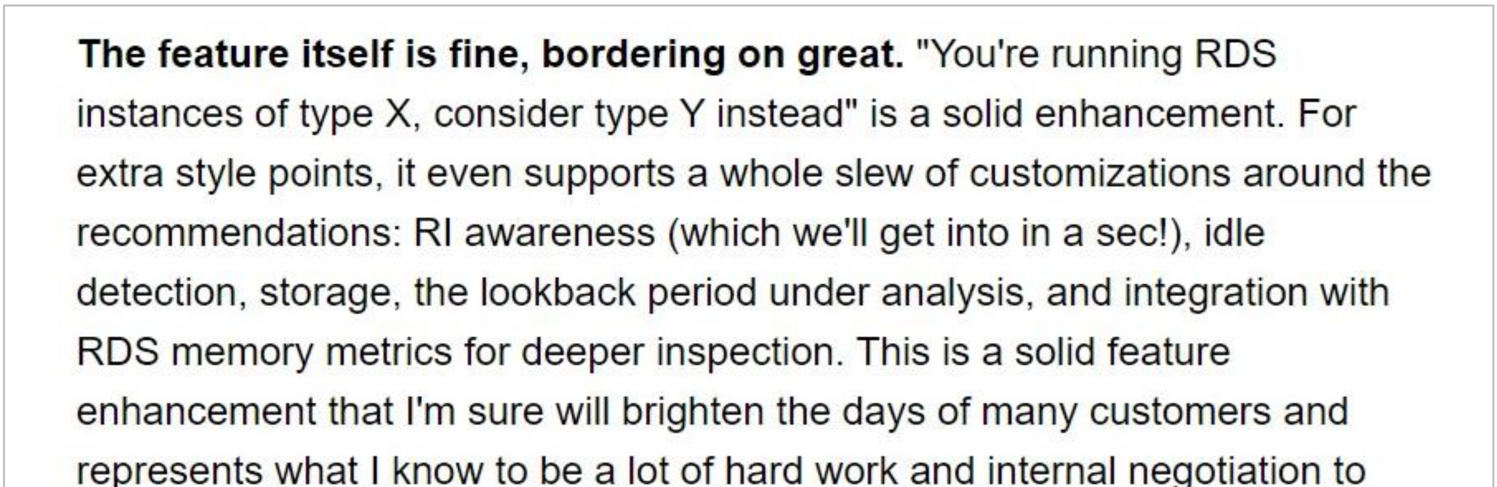
Today we are announcing Savings Plans, a new flexible pricing model that allows customers to save up to 72% on Amazon EC2 and AWS Fargate in exchange for making a commitment to a consistent amount of compute usage (e.g. $10/hour) for a 1 or 3 year term. With a Savings Plan, usage up to the commitment will be charged at the discounted Savings Plan rates and any usage beyond the commitment will be charged at regular On Demand rates. Savings Plans offer significant savings over On Demand, just like Reserved Instances, but automatically reduce customers' bills on compute usage across any AWS region, even as usage changes. This provides customers the flexibility to use the compute option that best suits their needs and continue to save money, all without having to perform exchanges or modifications.

Starting today, Savings Plans are available in all AWS regions outside of China. Click here to learn more about Savings Plans.

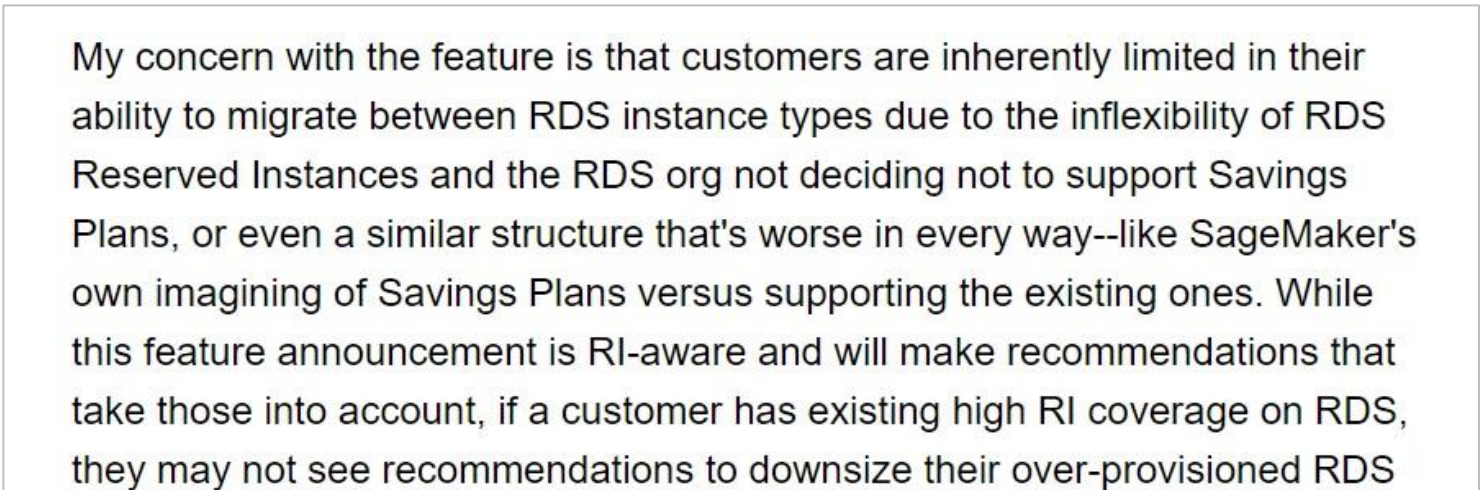
nn Unsubscribe 26 Jun 2024, 20:50 (2 days ago) ☆

There have been some noises about this week's newsletter issue in which I criticized the release of AWS Compute Optimizer offering RDS recommendations thusly:

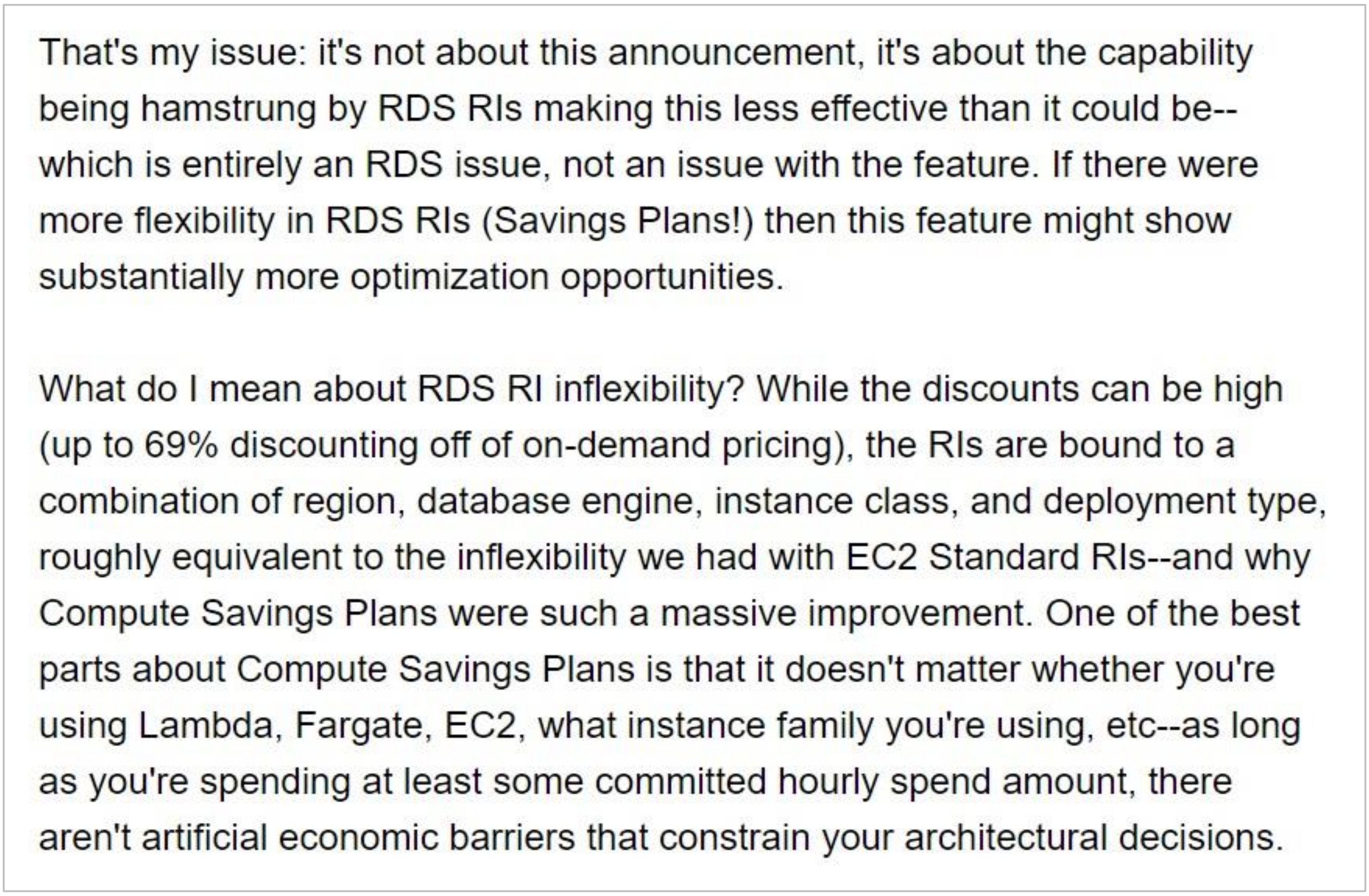
> Too bad it's completely useless for most customers, because RDS only has its own bespoke Reserved Instances, which are wildly inflexible. The fact that Savings Plans don't extend to cover RDS is one of the more customer-hostile things AWS does, and a



The feature itself is fine, bordering on great. "You're running RDS instances of type X, consider type Y instead" is a solid enhancement. For extra style points, it even supports a whole slew of customizations around the recommendations: RI awareness (which we'll get into in a sec!), idle detection, storage, the lookback period under analysis, and integration with RDS memory metrics for deeper inspection. This is a solid feature enhancement that I'm sure will brighten the days of many customers and represents what I know to be a lot of hard work and internal negotiation to

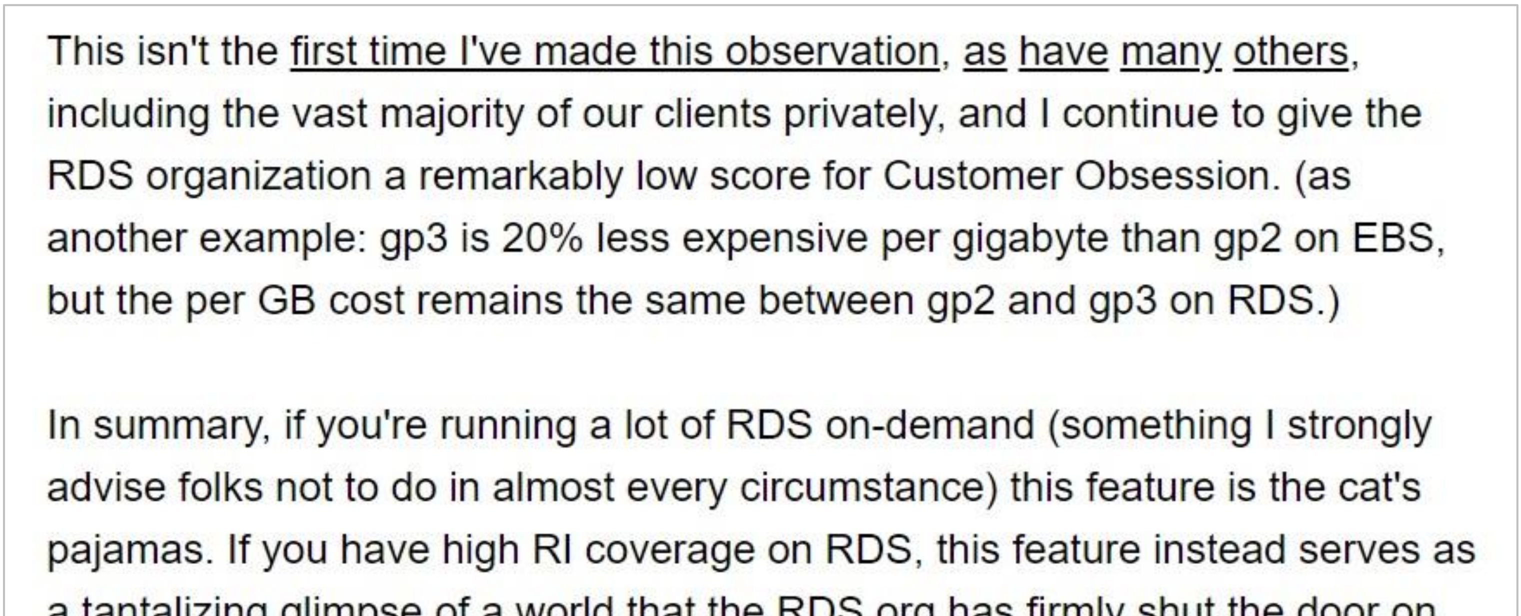


My concern with the feature is that customers are inherently limited in their ability to migrate between RDS instance types due to the inflexibility of RDS Reserved Instances and the RDS org not deciding not to support Savings Plans, or even a similar structure that's worse in every way--like SageMaker's own imagining of Savings Plans versus supporting the existing ones. While this feature announcement is RI-aware and will make recommendations that take those into account, if a customer has existing high RI coverage on RDS, they may not see recommendations to downsize their over-provisioned RDS



That's my issue: it's not about this announcement, it's about the capability being hamstrung by RDS RIs making this less effective than it could be--which is entirely an RDS issue, not an issue with the feature. If there were more flexibility in RDS RIs (Savings Plans!) then this feature might show substantially more optimization opportunities.

What do I mean about RDS RI inflexibility? While the discounts can be high (up to 69% discounting off of on-demand pricing), the RIs are bound to a combination of region, database engine, instance class, and deployment type, roughly equivalent to the inflexibility we had with EC2 Standard RIs--and why Compute Savings Plans were such a massive improvement. One of the best parts about Compute Savings Plans is that it doesn't matter whether you're using Lambda, Fargate, EC2, what instance family you're using, etc--as long as you're spending at least some committed hourly spend amount, there aren't artificial economic barriers that constrain your architectural decisions.



This isn't the first time I've made this observation, as have many others, including the vast majority of our clients privately, and I continue to give the RDS organization a remarkably low score for Customer Obsession. (as another example: gp3 is 20% less expensive per gigabyte than gp2 on EBS, but the per GB cost remains the same between gp2 and gp3 on RDS.)

In summary, if you're running a lot of RDS on-demand (something I strongly advise folks not to do in almost every circumstance) this feature is the cat's pajamas. If you have high RI coverage on RDS, this feature instead serves as a tantalizing glimpse of a world that the RDS org has firmly shut the door on

Bibliography

I. Official

[AWS 2019]

Introducing Savings Plans. Available at: <https://aws.amazon.com/about-aws/whats-new/2019/11/introducing-savings-plans/>

II. Unofficial

[Quinn 2024]

Quinn, Corey (2024). Apparently I Stuttered: A Compute Optimizer Clarification. [Email newsletter]. *Last Week in AWS Extras*. June 26th 2024.

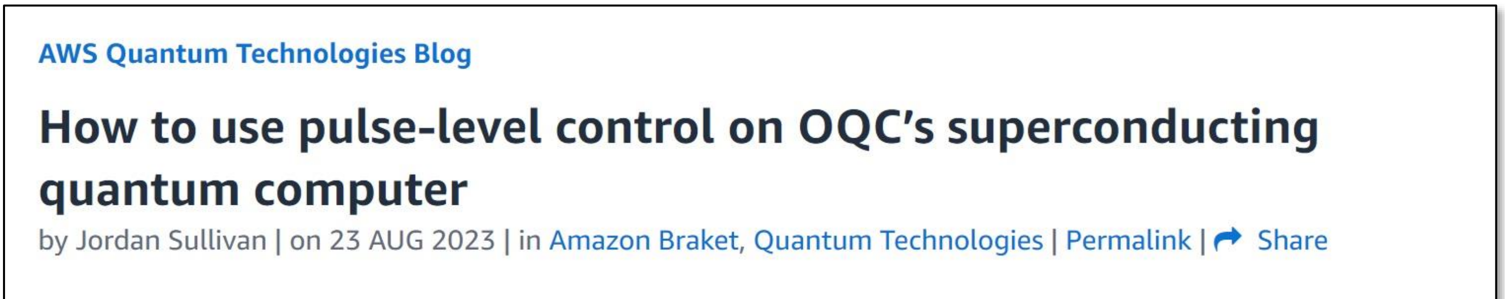
III. Critical

IV. General

AWS DeepComposer

Bibliography

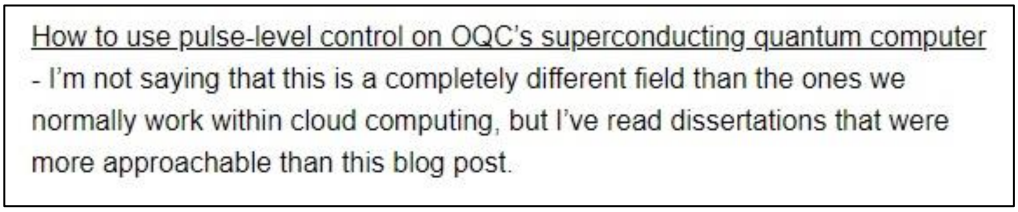
Braket



AWS Quantum Technologies Blog

How to use pulse-level control on OQC's superconducting quantum computer

by Jordan Sullivan | on 23 AUG 2023 | in Amazon Braket, Quantum Technologies | Permalink | Share



How to use pulse-level control on OQC's superconducting quantum computer - I'm not saying that this is a completely different field than the ones we normally work within cloud computing, but I've read dissertations that were more approachable than this blog post.

Quinn in an email on Aug 28th 2023

Bibliography

[Sullivan 2023]

How to use pulse-level control on OQC's superconducting quantum computer. *AWS Quantum Technologies Blog*. Aug 23rd 2023. Available at: <https://aws.amazon.com/blogs/quantum-computing/how-to-use-pulse-level-control-on-oqcs-superconducting-quantum-computer/?ck_subscriber_id=1560524742>

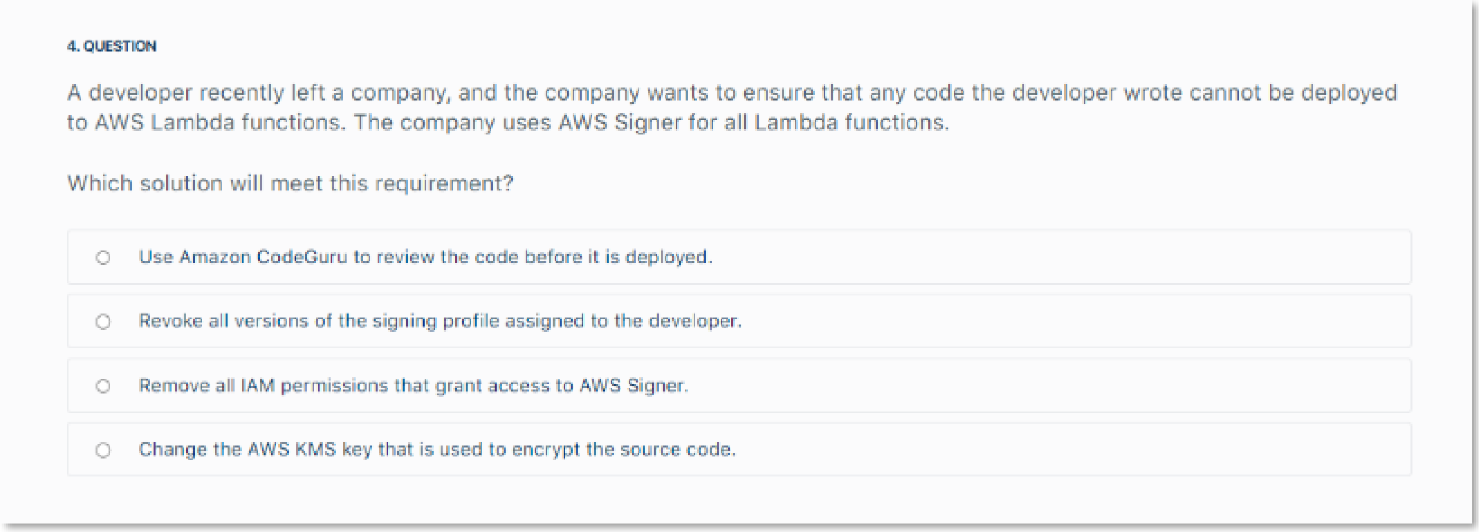
Augmented AI

←

Bibliography

CodeGuru

Srinivasan Sengamedu and Daniela Tzvetkova presenting in 2019, the year Amazon CodeGuru was announced



4. QUESTION

A developer recently left a company, and the company wants to ensure that any code the developer wrote cannot be deployed to AWS Lambda functions. The company uses AWS Signer for all Lambda functions.

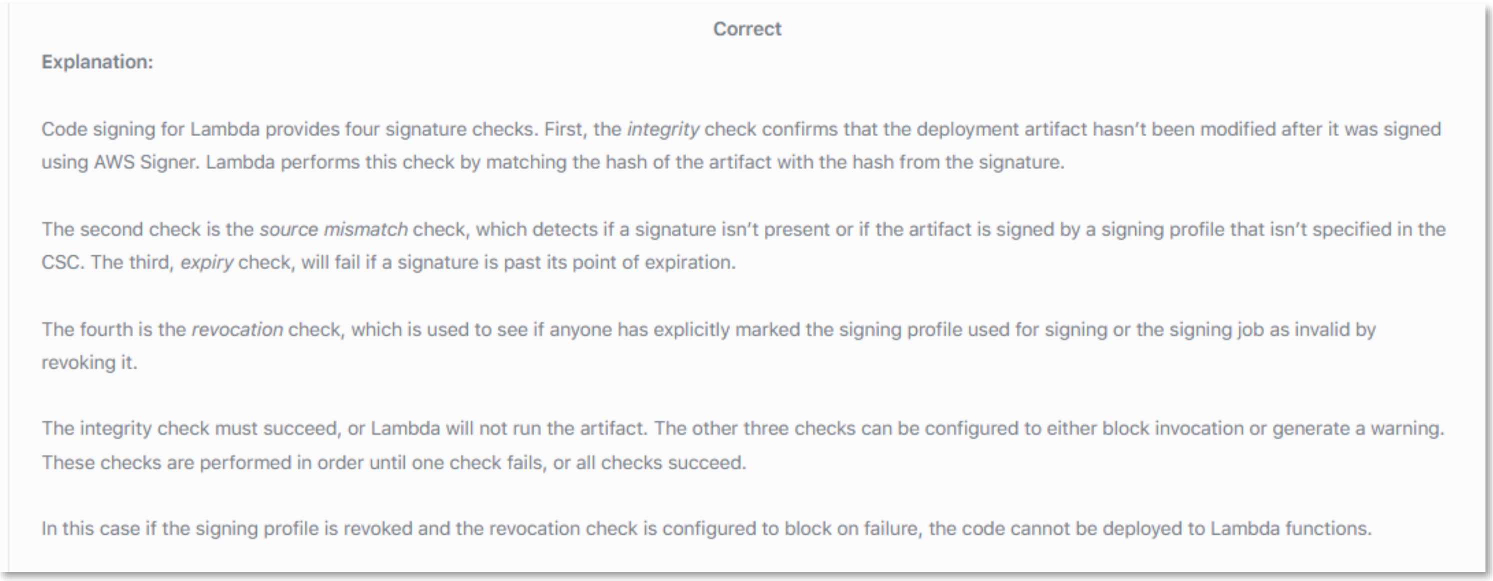
Which solution will meet this requirement?

○ Use Amazon CodeGuru to review the code before it is deployed.

○ Revoke all versions of the signing profile assigned to the developer.

○ Remove all IAM permissions that grant access to AWS Signer.

○ Change the AWS KMS key that is used to encrypt the source code.



Correct

Explanation:

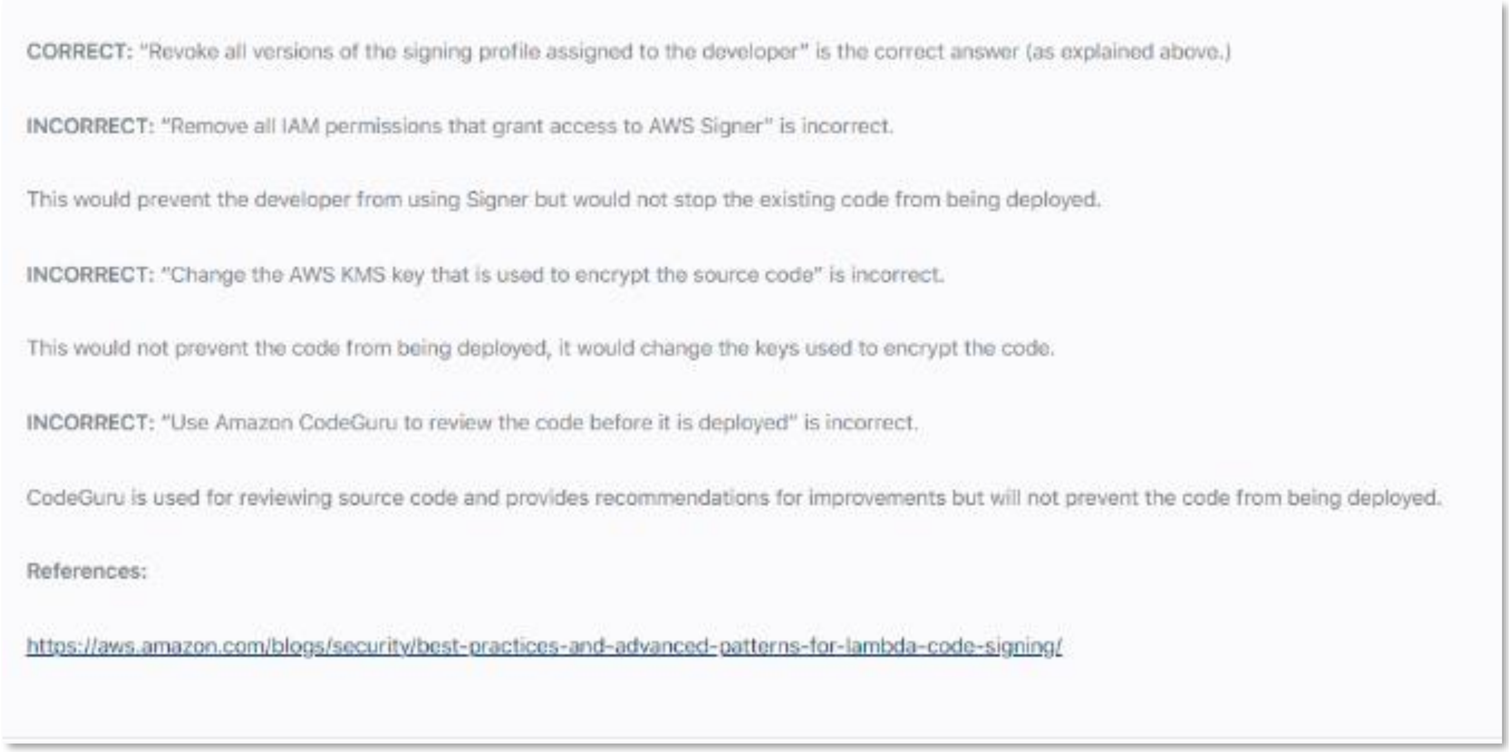
Code signing for Lambda provides four signature checks. First, the *integrity* check confirms that the deployment artifact hasn't been modified after it was signed using AWS Signer. Lambda performs this check by matching the hash of the artifact with the hash from the signature.

The second check is the *source mismatch* check, which detects if a signature isn't present or if the artifact is signed by a signing profile that isn't specified in the CSC. The third, *expiry* check, will fail if a signature is past its point of expiration.

The fourth is the *revocation* check, which is used to see if anyone has explicitly marked the signing profile used for signing or the signing job as invalid by revoking it.

The integrity check must succeed, or Lambda will not run the artifact. The other three checks can be configured to either block invocation or generate a warning. These checks are performed in order until one check fails, or all checks succeed.

In this case if the signing profile is revoked and the revocation check is configured to block on failure, the code cannot be deployed to Lambda functions.



CORRECT: "Revoke all versions of the signing profile assigned to the developer" is the correct answer (as explained above.)

INCORRECT: "Remove all IAM permissions that grant access to AWS Signer" is incorrect.

This would prevent the developer from using Signer but would not stop the existing code from being deployed.

INCORRECT: "Change the AWS KMS key that is used to encrypt the source code" is incorrect.

This would not prevent the code from being deployed, it would change the keys used to encrypt the code.

INCORRECT: "Use Amazon CodeGuru to review the code before it is deployed" is incorrect.

CodeGuru is used for reviewing source code and provides recommendations for improvements but will not prevent the code from being deployed.

References:

https://aws.amazon.com/blogs/security/best-practices-and-advanced-patterns-for-lambda-code-signing/

Bibliography

I. Official
II. Unofficial
III. Critical
IV. General

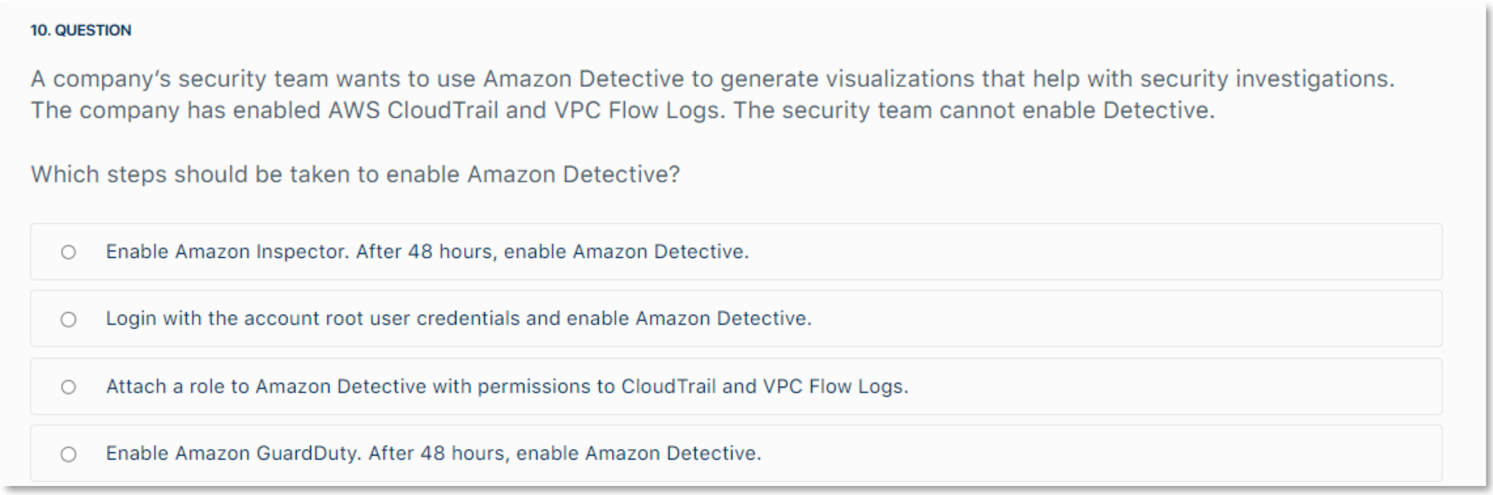
I. Official
II. Unofficial
III. Critical
IV. General

AWS Compute Optimizer ←

Bibliography

Amazon Detective ←

GREATER LONDON POLICE
ENTER AND SEARCH PREMISES
THERESA MAY
THERESA MAY
VOLUNTEER
ARMIES
"COMMUNITY CAMP FIGHTERS"
3
1.
INCIDENT
MANAGEMENT LOG
BOOK 198
CCTV
FOOTAGE
2
POLICE
26.09
RIVER
THAMES
LONDON
SERIAL
PODCAST
SARAH
KOENIG
04.
ADNAN
EXHIBIT
14
HAMMER
BENSON



10. QUESTION

A company's security team wants to use Amazon Detective to generate visualizations that help with security investigations. The company has enabled AWS CloudTrail and VPC Flow Logs. The security team cannot enable Detective.

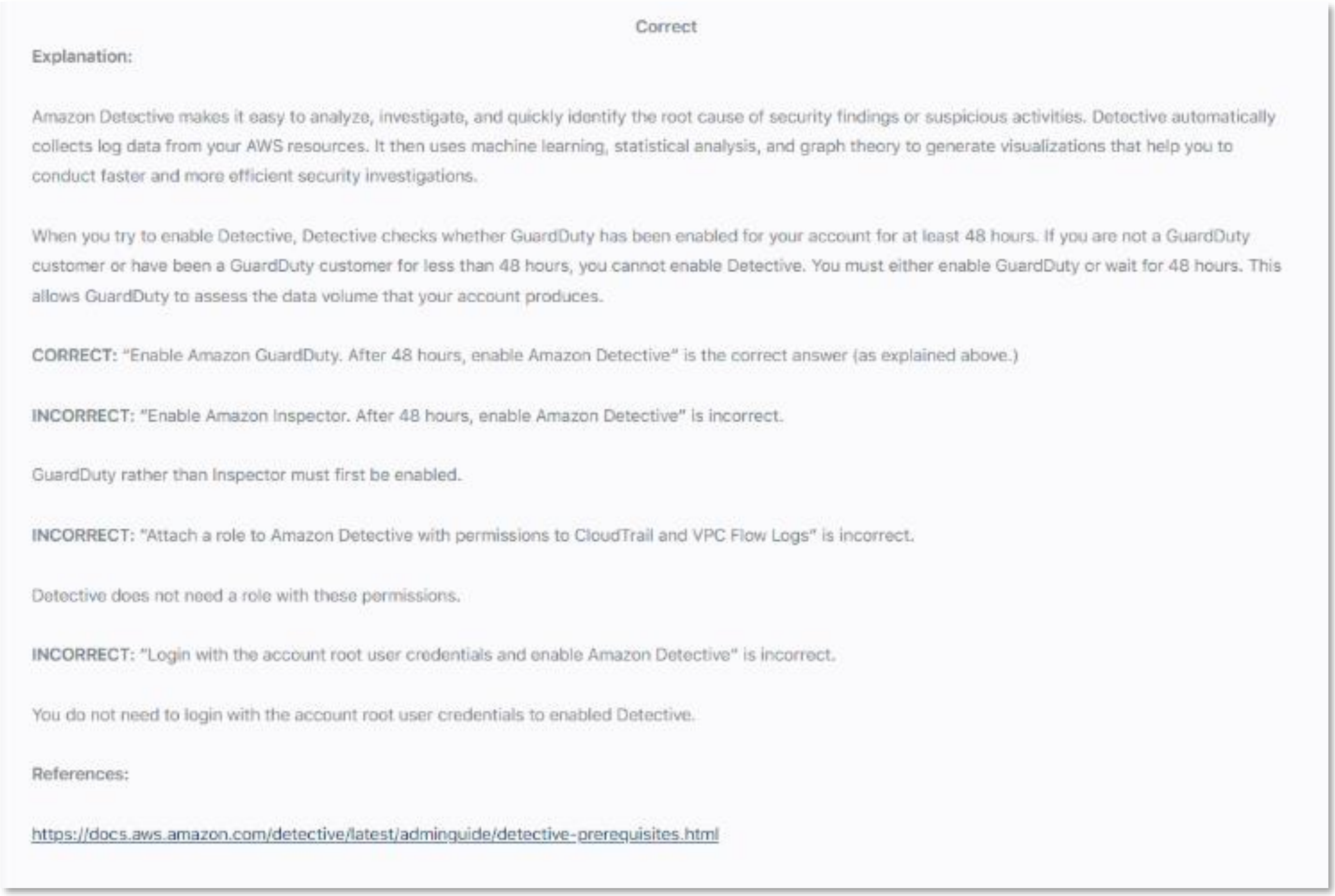
Which steps should be taken to enable Amazon Detective?

○ Enable Amazon Inspector. After 48 hours, enable Amazon Detective.

○ Login with the account root user credentials and enable Amazon Detective.

○ Attach a role to Amazon Detective with permissions to CloudTrail and VPC Flow Logs.

○ Enable Amazon GuardDuty. After 48 hours, enable Amazon Detective.



Correct

Explanation:

Amazon Detective makes it easy to analyze, investigate, and quickly identify the root cause of security findings or suspicious activities. Detective automatically collects log data from your AWS resources. It then uses machine learning, statistical analysis, and graph theory to generate visualizations that help you to conduct faster and more efficient security investigations.

When you try to enable Detective, Detective checks whether GuardDuty has been enabled for your account for at least 48 hours. If you are not a GuardDuty customer or have been a GuardDuty customer for less than 48 hours, you cannot enable Detective. You must either enable GuardDuty or wait for 48 hours. This allows GuardDuty to assess the data volume that your account produces.

CORRECT: "Enable Amazon GuardDuty. After 48 hours, enable Amazon Detective" is the correct answer (as explained above.)

INCORRECT: "Enable Amazon Inspector. After 48 hours, enable Amazon Detective" is incorrect.

GuardDuty rather than Inspector must first be enabled.

INCORRECT: "Attach a role to Amazon Detective with permissions to CloudTrail and VPC Flow Logs" is incorrect.

Detective does not need a role with these permissions.

INCORRECT: "Login with the account root user credentials and enable Amazon Detective" is incorrect.

You do not need to login with the account root user credentials to enabled Detective.

References:

https://docs.aws.amazon.com/detective/latest/adminguide/detective-prerequisites.html

Waiting 48 hours

Fraud Detector

Let’s explore this service, entitled Amazon Fraud Detector. Initially, we need to answer a number of questions:

1. What is fraud?
2. What is Amazon Fraud Detector?
3. What are the advantages of the product?
4. What are the disadvantages of the product?
5. Are there any disadvantages to using ML on fraud?
6. What special considerations need to be taken when dealing with fraud carried out using computers?
7. Are there any disadvantages to preventing fraud in real time?

Let’s begin with the first question, *What is fraud?* Well, [O’Neill] tells us that:

> A philosophical conception of fraud, inspired by Kant, defines it as denying to the weaker party in a financial transaction (such as a consumer or investor) information that is necessary to make a rational (or autonomous) decision.
>
> [O’Neill]

We know intuitively what fraud is. When you feel that you have been the victim of fraud, you feel *cheated*. The verbose definition above is compatible with this. It tells us that fraud involves information being withheld. The “weaker party” is misled, because they are not in full possession of the facts. Here are some examples of fraud in everyday life:

1. Somebody uses a counterfeit bank note in a pub.
2. Somebody uses a false identity (such as a false name) while engaging in a personal relationship with another.
3. An educational company promises textbooks to students but never provides them.

We can generalise that last example. Any situation involving a company which promises some service, takes money, and then fails to provide the service, might be said to have committed fraud. The second example, involving a personal relationship, would be contested by some as a true example of fraud. This is because it involves a personal relationship, and not necessarily any financial transaction. Such a situation might involve an intelligence agent initiating a sexual relationship with somebody, motivated perhaps by love of country but with no particular penchant for money (similar to the plot in [Lester 2016] and [Patterson 2022]). This alerts us to the fact that fraud is usually associated with deception *for financial gain*, as opposed to deception more generally.

But some might argue that the above example, in which the victim of the fraud is used for their information—and not primarily their money—is in fact an example of fraud. Consider academia, an industry supposedly devoted to harvesting knowledge. If somebody were to publish an academic paper which contained false results, then this would be considered fraud. Yet here, the victims of the fraud (the readers of the paper) have no cash leave their bank account. Liam Kofi Bright has refined a theory about why fraudulent scientific investigations occur. The consensus is that when fraud occurs, scientists seek credit for their findings, and place this desire for self-promotion above a concern for the truth. It follows that in order to tackle fraud, scientists must simply *more vigorously* aim for the truth, while they meanwhile (justly) aim to take credit for good work. Bright argues that this is not an effective strategy:

> Simply making truth compete on equal footing with credit only changes the situations in which scientists are incentivised to commit fraud, but does nothing to suggest that scientists will be incentivised to commit fraud less often.
>
> [Bright 2016]

In other words, we cannot simply ask scientists to aim for both credit and truth – or, rather, this will do nothing to reduce fraud. What reasons does Bright have for suggesting this? Well, he points out that there are in fact ways in which aggressively demanding that somebody spread truths can cause them—*incentivise* them—to spread lies. Bright outlines "noble lies", which are slight distortions of the finer details brought about so that people arrive at truths. You commit a "noble lie" because you want others to arrive at the truth and in fact

> The desire to see the truth propagated can make scholars willing to pay the cost associated with fraud, taking on a risk to their personal career for the sake of what they believe to be the communal good.

Bright refers to scientific atlas compilers who would leave out certain features, which they predict would mislead the reader. "Scientists would avoid reproducing exactly the observed sample in their image, since they feared prompting readers to mistake what was peculiar to this sample for a typical property of the object" [Bright 2016]. Bright also points out that for scholars,

> Truth seeking is liable to make the most epistemically powerful, those most able to sway the opinion of others, more likely to commit fraud.

This is because if a scholar is influential, their paper—false though it might be—may represent a well-defined falsehood, whose consideration will help others to arrive at the truth. It prods and pokes the wider community. Thus, publishing the false paper will help the community to arrive at the truth.

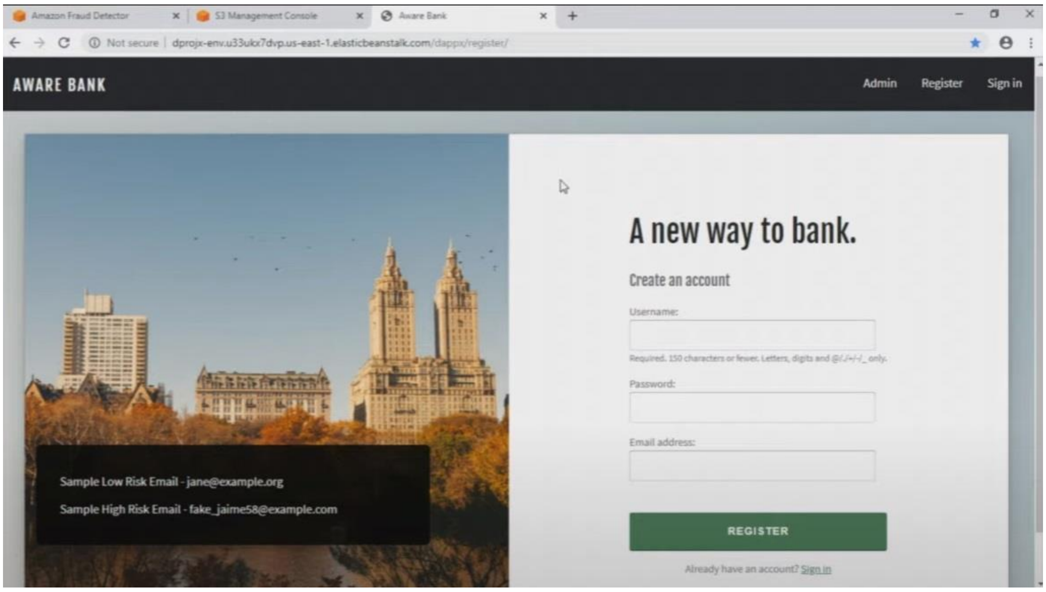
All these considerations lead Bright to conclude that in order to reduce fraudulent papers, it is not sufficient to ask scholars to aim for both credit for their work as well as the truth. This latter project (asking scholars to maintain both these motivations) is referred to by Bright as the "motive modification tradition".

This concludes the discussion of what fraud is. Fraud is withholding information or lying. It is often a crime.

Amazon Fraud Detector

At the ReInvent conference in 2019, two AWS representatives delivered a presentation about the new product. In the low-resolution image below, Ryan Schmiedl is on the left, and Nick Tostenrude on the right. The individual in the centre is Kara Suro, a representative from the financial services company Charles Schwab.

Quick Demo



A new way to bank.

Create an account

Username:

Required. 150 characters or fewer. Letters, digits and @/./+/-/_ only.

Password:

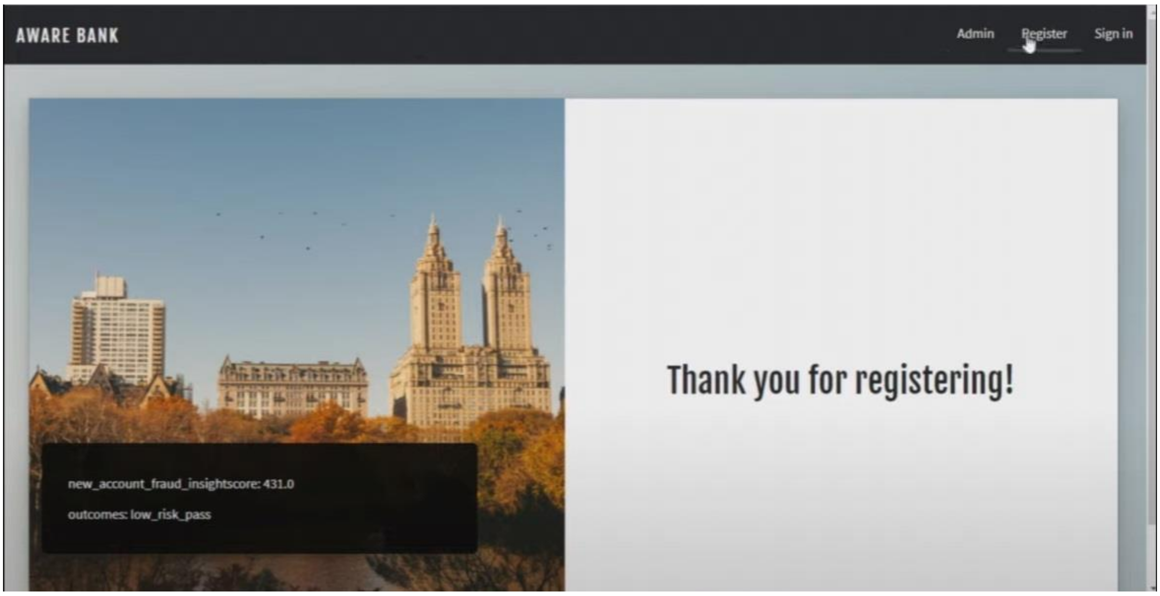
Email address:

Already have an account? Sign in

Sample Low Risk Email - jane@example.org

Sample High Risk Email - fake_jaime58@example.com

The presentation begins with a quick demonstration. In the image above, we see a field for a username, password, and email address to be inputted. This is a sign-up form. Tostenrude inputs an example user, password and email address. The screen then shows “thank you for registering”. Why? Because the email address generated an fraud score of 431. This is considered to be a low score. In fact, the web page for the demo was configured to display the fraud score.

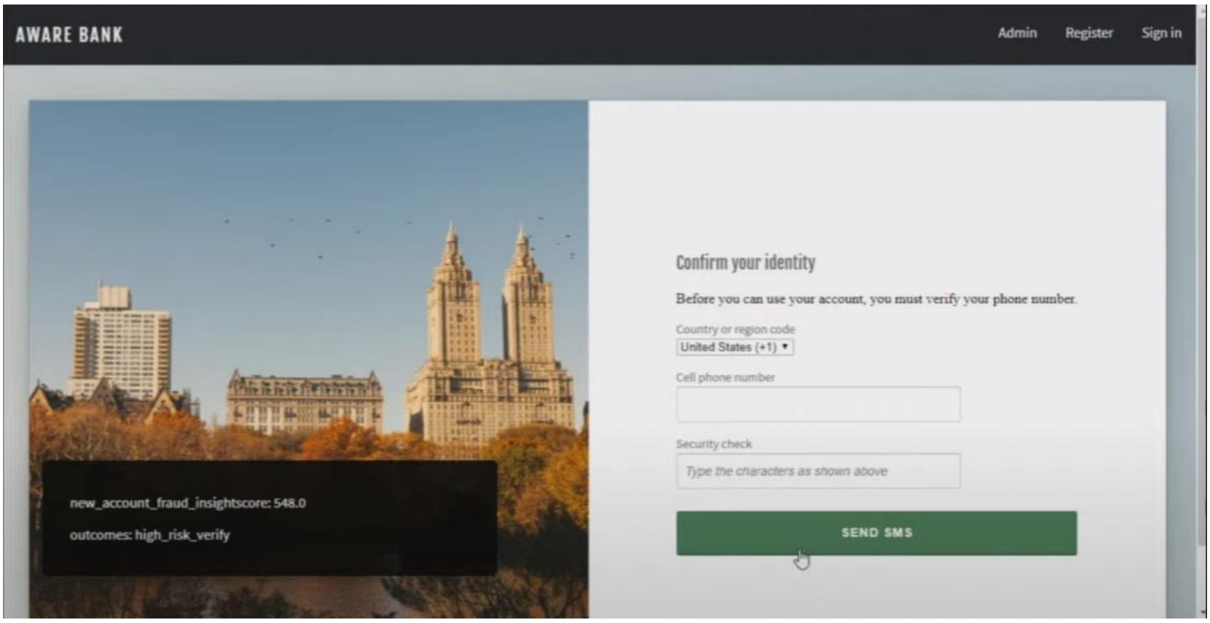


Thank you for registering!

new_account_fraud_insightscore: 431.0

outcomes: low_risk_pass

Tostenrude then inputs a different “high risk” email address. This time, instead of being told “thank you for registering” immediately, the user is asked to input a phone number. This did happen when the email address was low risk.



Confirm your identity

Before you can use your account, you must verify your phone number.

Country or region code
United States (+1) ▾

Cell phone number

Security check
Type the characters as shown above

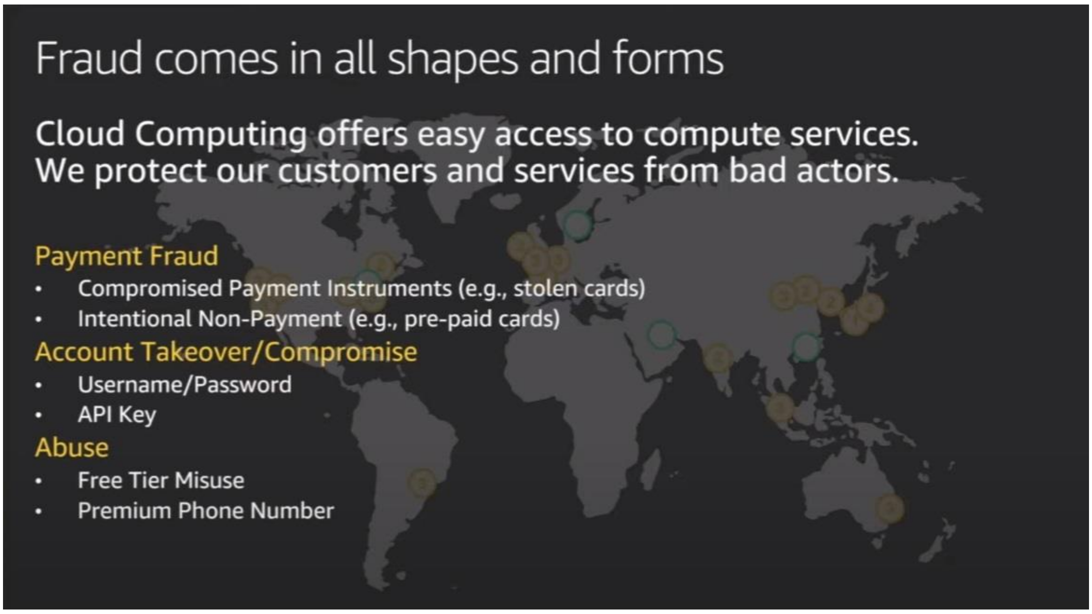
new_account_fraud_insightscore: 548.0

outcomes: high_risk_verify

This email address generated a risk score of 548. This was higher than 431. All this serves to demonstrate the way two email addresses were treated differently. “Amazon Fraud Detector can enable you to dynamically adjust the customer experience based on risk” Toldenrude tells us.

Schmiedl makes the point that the customer experience of the latter, suspect user is slightly diminished. It is more of a hassle for this user, because they must input their phone number. He suggests that the virtue of Amazon Fraud Detector (AFD) is the ability to tailor the experience to the user—to minimise the users who are hassled.

Types of Fraud



Fraud comes in all shapes and forms

Cloud Computing offers easy access to compute services.
We protect our customers and services from bad actors.

Payment Fraud
- Compromised Payment Instruments (e.g., stolen cards)
- Intentional Non-Payment (e.g., pre-paid cards)

Account Takeover/Compromise
- Username/Password
- API Key

Abuse
- Free Tier Misuse
- Premium Phone Number

Fraud at AWS, at banks, and online, comes into a number of recognisable types. There is payment fraud, account takeover, and abuse. Schmiedl tells us that the payment model used by AWS is particularly prone to fraud. AWS is a "post-pay service". You do not provide cash immediately – when you sign up to use AWS, you provide the *form* for payment (the details of a credit card) but not the *matter* (not the actual cash). Sometimes people pay for AWS products which they cannot in fact afford. Schmiedl says:

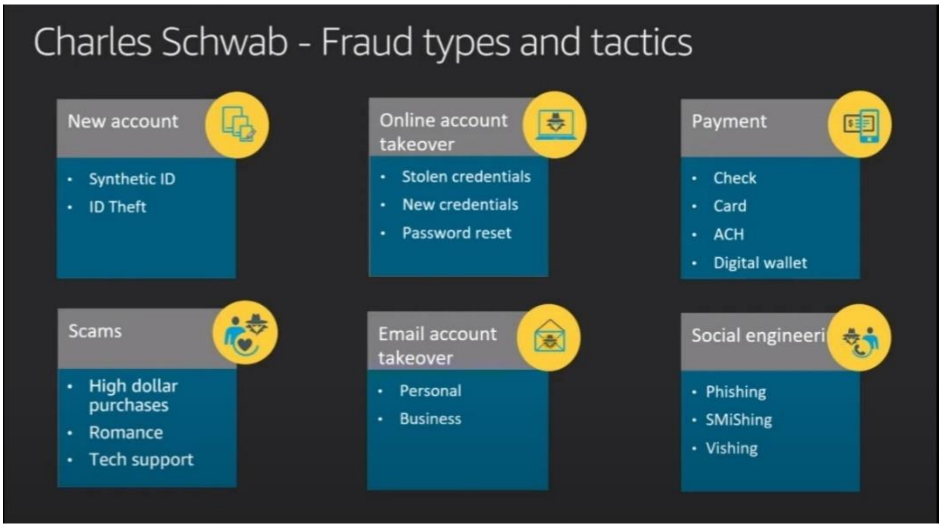
> We accept pre-paid cards. With a pre-paid card, you do not know how much they're going to use. You can do an auth, it can go through, and if they start using and consuming services beyond what's on that card, you're left holding the bag.
>
> [Schmiedl 2019]

To "do an auth" must be to do an authentication (an establishing of the identity of the person holding the card).

What about **abuse**? Schmiedl says "abuse is defined from the position of your organization. What types of activities do you want to prevent?" We are told that for AWS in particular, there is free-tier abuse, promotional abuse, folks that are trying to use some kind of promotion, and create hundreds of accounts, trying to figure out how to either sell the accounts or chain the accounts together to do something for free".

Sometime, people set up a premium phone number. They associate an account with a phone number, which costs money to call. AWS then call the number, and are charged money.

Fraud at Schwab



Charles Schwab - Fraud types and tactics

New account
- Synthetic ID
- ID Theft

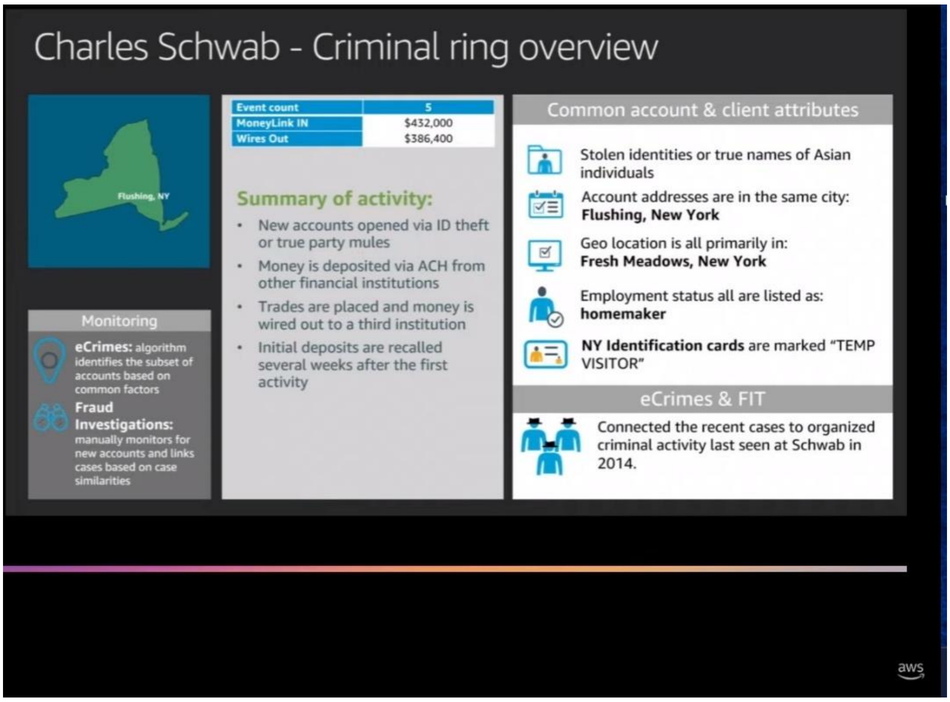
Online account takeover
- Stolen credentials
- New credentials
- Password reset

Payment
- Check
- Card
- ACH
- Digital wallet

Scams
- High dollar purchases
- Romance
- Tech support

Email account takeover
- Personal
- Business

Social engineeri
- Phishing
- SMiShing
- Vishing



Charles Schwab - Criminal ring overview

Event count	5
MoneyLink IN	$432,000
Wires Out	$386,400

Summary of activity:

- New accounts opened via ID theft or true party mules
- Money is deposited via ACH from other financial institutions
- Trades are placed and money is wired out to a third institution
- Initial deposits are recalled several weeks after the first activity

Monitoring

eCrimes: algorithm identifies the subset of accounts based on common factors

Fraud Investigations: manually monitors for new accounts and links cases based on case similarities

Common account & client attributes

Stolen identities or true names of Asian individuals

Account addresses are in the same city: **Flushing, New York**

Geo location is all primarily in: **Fresh Meadows, New York**

Employment status all are listed as: **homemaker**

NY Identification cards are marked "TEMP VISITOR"

eCrimes & FIT

Connected the recent cases to organized criminal activity last seen at Schwab in 2014.

Glossary

Fraud -

Abuse –

Synthetic ID -

True party -

Bibliography

[Lester 2016]

Lester, Adrian and Sophie Ookonedo (2016). *Undercover*. Aired on the BBC Apr 3rd 2016. See Lester discussing the drama here: <https://www.youtube.com/watch?v=mPdySMyFAEI&ab_channel=BarkingattheMoon>

[Patterson 2022]

Patterson, Adam (2022). *Rogue Agent* [film]. Directed by Adam Patterson. Starring James Norton and Gemma Arterton.

[Bright 2016]

Bright, Liam Kofi (2016). On Fraud. *Philosophical Studies*.

Schmiedl, Ryan and Nick Tostenrude and Kara Suro (2019). Introducing Amazon Fraud Detector: Detect more online fraud faster. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=ZkPQtjETrBw&ab_channel=AWSEvents>

Tadinada, Karthik. Building and Deploying a Real-time Banking Fraud Detection System. *London Data Science Summit*. Available at: <https://www.youtube.com/watch?v=cOzLO37mkwk&ab_channel=CambridgeSpark>

Turner, Steve (2018). Amazon SageMaker for Fraud Detection. Available at: <https://www.youtube.com/watch?v=wzwkLV9gDXk&ab_channel=AmazonWebServices>

Greene, Eric (2020). Fraud Detection: Using ML to Identify and Manage Fraudulent Activities. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=INO3TNT2ECk&ab_channel=AWSEvents>

Catch more potential online fraud faster with Amazon Fraud Detector. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=5QSXbgbvleo&ab_channel=AWSEvents>

Kurt Kufeld announces Fraud Detector is Generally Available. YouTube Channel: Amazon Web Services. Available at: <https://www.youtube.com/watch?v=0mvrl6AXcb8&ab_channel=AmazonWebServices>

Frost, Mike (2021). Proactively Detect and Prevent Online Fraud with Amazon SageMaker and Amazon Fraud Detector. YouTube Channel: AWS Online Tech Talks. Available at: <https://www.youtube.com/watch?v=viih7LpB1gg&ab_channel=AWSOnlineTechTalks>

[O’Neil]

de Bruin, Boudewijn, Lisa Herzog, Martin O’Neill, and Joakim Sandberg, "Philosophy of Money and Finance", The Stanford Encyclopedia of Philosophy (Spring 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/spr2023/entries/money-finance/>.

Kendra

What does the word “kendra” mean? It is hard to find reliable sources on this. I know it’s a popular name in the USA: Kendra Wilkinson appeared on the British television programme *I’m a Celebrity Get me Out of Here*; Kendra Brooks is an activist working on education in Philadelphia; and Kendra Smith is a bassist, active in the eighties, who went to live in a cabin in the woods.

I also know that in India there are things called KVKs. Starting in the 1970s, the government set up KVKs, which are centres of agricultural research. They’re a sort of bridge between the scientific research of the universities and the farmers “on the ground”. (Here is Modi visiting a KVK in the state of Maharashtra; here is a video on a KVK further south, in Kerala, and their work on bananas). In July 2020, there were 721 KVKs across India.

Incidentally, Werner Vogels recently visited Bengaluru (the "Silicon Valley of India") where he learnt about technology empowering artisans who handcraft things such as lampshades (see *thomu bommalata*). Note that Vogels really focusses on farmers in his episode on Jakarta.

KVK stands for **Krishi Vigyan Kendra**. This means something like "farm science centre". There are also forest science centres in India known as *Van* Vigyan Kendra. The Indian government has a centre devoted to malware called the "Cyber Swachhta Kendra". Finally, there is a conference centre in New Delhi called the Vigyan Bhawan. The Prime Minister often speaks here. "Vigyan Bhawan" is often transliterated as "*science* building". This all makes clear that it's *Vigyan* that roughly means science.

As for the "kendra" in KVK, this means "centre". In many Ancient Greek works you can find a word *similar* to "kendra". It always denotes some sort of **sharp point**. Homer uses it to refer to the pointed stick used to prod horses; Plato uses it to refer to the centre of a circle; Aristophanes uses it to refer to the sting of a wasp; Polybius for the pointed part of a spear. So, we can see that "kendra" has roots in Proto-Indo-European language.

What's the point of Amazon Kendra? This is Amazon's contribution to the domain of enterprise search. What on earth is meant by *enterprise search*?

Well, we can start by contrasting an *enterprise search* with other types of search. When you search for something on google.com, a search engine systematically goes through documents available on the web. You've performed a web search. You might also perform a desktop search. However, what we are interested in is internal employees within a company searching the vast amounts of information held by the company. This is enterprise search.

Do yourself a favour and read the first chapter of Martin White's book entitled Enterprise Search. It has a swallow (bird) on the cover. Another of White's books is called *Making Search Work* (2007). White is in fact 'a chemist by training', as a result, he's been able to help many pharmaceutical companies with complex *information discovery challenges*. He is a fellow of the Royal Society of Chemistry. Learn more on his page on Sheffield University, England (he is a Visiting Professor).

"Search Patterns" (2010) can be okay for beginners. Again, this is O'Reilly and there is a butterfly on the front cover. Some reviews are not great. Discover more books here. This area is not boring.

Amazon Kendra uses machine learning to improve the accuracy of searches within an enterprise. This marketing video from AWS goes through the problems faced in implementing enterprise search:

- the majority of data today is unstructured and siloed.
- Key word searches lack context
- Search results are often inaccurate

Amazon Kendra to the rescue! It alleviates these problems with machine learning. It uses new methods as well: "going beyond keyword matching, Kendra uses natural language understanding, and **reading comprehension**". Getting machines to really comprehend a text they've just read is harder than you might imagine, as this Amazon Science article proves.

Kendra allows you to use things called **relevance sliders**. These allow you to tune your results. "Tune" here seems to mean alter the way Kendra regards certain features, in prioritizing search items. Determine whether the *popularity* of a result should be considered; the *authority* of the source of an item; or the *freshness* of the item.

We know that searching for things helps you *do* things. So presumably, the service is called Kendra because it allows you to find *precisely* what you're looking for, giving you a higher degree of **control** over your environments. This is in line with centres in India—for helping farmer control the natural world—and sharp sticks used in Ancient Greece. Searching for things and not finding what you’re looking for is a massive problem. Jean-Pierre Dodel cites a study that found that $5,700 can be lost per employee with bad search tools. This is because so much time is spent retrieving the correct information.

This presentation by Jean-Pierre Dodel is an extremely clear introduction to Amazon Kendra. Dodel talks about “the combination of three models”. The models are (1) Reading comprehension (2) FAQ matching (3) document ranking. Dodel gives us three highly specific examples of queries:

1. Where is the IT support desk?
2. How long is maternity leave?
3. How do I configure my VPN?

We are also told that Kendra is optimized for **14 major domains** such as IT, finance, insurance, Pharma, and industrial.

There is also a presentation on Kendra from Rajagopalan.

I’ve recently come across this article from someone who knows his stuff: https://capiche.com/q/all-you-wanted-to-know-about-amazon-kendra

Keyspaces ←

Amazon Wavelength

CloudEndure Disaster Recovery

What is CloudEndure?

CloudEndure was founded in 2012 by Ofer Gadish, Gil Shai, Ofir Ehrlich and Leonid Feinberg. The CEO of CloudEndure is shown below:

Ofer Gadish, founder of CloudEndure. Source: Twitter

CloudEndure, as its name implies, offer services which increase the ability of cloud-based workloads to endure:

> CloudEndure's backup technology includes block-level replication that runs in the background and works in real time, with recovery point objectives of less than a second, according to the company. It also has automated virtual

> machine conversion capabilities to move VMs from one target infrastructure to another, as well as a workload orchestration engine.
>
> [Kanaracus 2019]

However, it is also concerned with helping customers to migrate to the cloud. It seems to have a particular focus on public clouds:

> CloudEndure's big value proposition has been their low/no downtime migration into public cloud, making it easy for enterprises to start adopting public cloud in a backup and failover manner

It seems that many enterprises use a combination of cloud providers (Azure, AWS, GCP). Given that CloudEndure helped a lot of companies to establish a hybrid cloud in this manner, a question is now raised following AWS's acquisition of this company. Kanaracus wrote in 2019 that:

> CloudEndure's website proclaims it is now an AWS company, but continues to list Microsoft Azure, Google Cloud Platform and VMware as partners. VMware is also an investor in the company.
>
> [Kanaracus 2019]

So, AWS have acquired CloudEndure but VMWare provide investment to it.

Kanaracus suggests that one of the reasons AWS may have wanted to acquire CloudEndure is that it would provide AWS with insights into how other clouds:

> AWS ultimately wants to move more workloads to its own infrastructure, rather than act as a middleman for multiple clouds. As AWS works on deeper integrations between its portfolio and CloudEndure, it could gain insights into how CloudEndure customers run cloud workloads elsewhere, and could illuminate ways to improve AWS' cloud migration competitiveness
>
> [Kanaracus 2019]

Bibliography

[Kanaracus 2019]

Kanaracus, Chris (2019). CloudEndure acquisition expands AWS' cloud migration tools. *TechTarget*. Available at: <https://www.techtarget.com/searchaws/news/252455858/CloudEndure-acquisition-expands-AWS-cloud-migration-tools>

[Lu 2021]

Lu, Wally (2021). CloudEndure Migration Factory Best Practices. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=is7cOcNUHlw>

[Jaeger 2019]

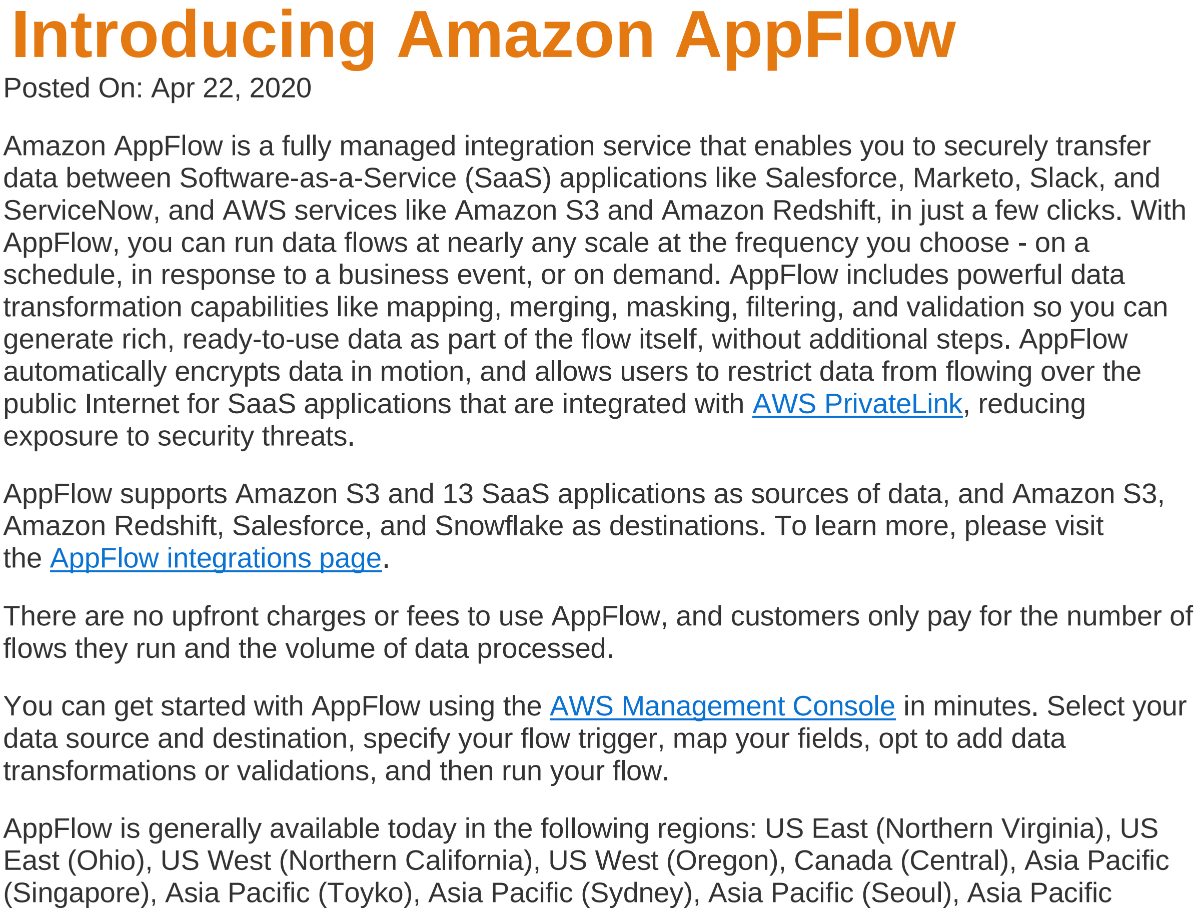
Jaeger, Ryan (2019). Automated Disaster Recovery using CloudEndure. Oct 15th 2019. *AWS Architecture Blog*. Available at: <https://aws.amazon.com/blogs/architecture/automated-disaster-recovery-using-cloudendure/>

AppFlow

←

⚠ Clusters of Confusable Concepts

AppFlow sounds quite similar to a number of other AWS products, namely AppRunner, AppStream and AppSync. To help with this, observe that AppFlow is the only name with "low" in its name. The product allows data to be transferred between AWS and a number of smaller, commercial SaaS entities (Salesforce, Zendesk etc). It is thus about the megalith that is Amazon stooping down—coming down lower—to provide integration with these smaller companies.



Introducing Amazon AppFlow

Posted On: Apr 22, 2020

Amazon AppFlow is a fully managed integration service that enables you to securely transfer data between Software-as-a-Service (SaaS) applications like Salesforce, Marketo, Slack, and ServiceNow, and AWS services like Amazon S3 and Amazon Redshift, in just a few clicks. With AppFlow, you can run data flows at nearly any scale at the frequency you choose - on a schedule, in response to a business event, or on demand. AppFlow includes powerful data transformation capabilities like mapping, merging, masking, filtering, and validation so you can generate rich, ready-to-use data as part of the flow itself, without additional steps. AppFlow automatically encrypts data in motion, and allows users to restrict data from flowing over the public Internet for SaaS applications that are integrated with AWS PrivateLink, reducing exposure to security threats.

AppFlow supports Amazon S3 and 13 SaaS applications as sources of data, and Amazon S3, Amazon Redshift, Salesforce, and Snowflake as destinations. To learn more, please visit the AppFlow integrations page.

There are no upfront charges or fees to use AppFlow, and customers only pay for the number of flows they run and the volume of data processed.

You can get started with AppFlow using the AWS Management Console in minutes. Select your data source and destination, specify your flow trigger, map your fields, opt to add data transformations or validations, and then run your flow.

AppFlow is generally available today in the following regions: US East (Northern Virginia), US East (Ohio), US West (Northern California), US West (Oregon), Canada (Central), Asia Pacific (Singapore), Asia Pacific (Toyko), Asia Pacific (Sydney), Asia Pacific (Seoul), Asia Pacific

(Mumbai), Europe (Paris), Europe (Ireland), Europe (Frankfurt), Europe (London), and South America (São Paulo).

Bibliography

Introducing Amazon AppFlow. Apr 22nd 2020. Available at: <https://aws.amazon.com/about-aws/whats-new/2020/04/introducing-amazon-appflow/>

ChatBot

AWS Chatbot Now Generally Available

Posted On: Apr 22, 2020

AWS Chatbot is now generally available for all customers. AWS Chatbot is an interactive agent for “ChatOps” that makes it easy to monitor and interact with your AWS resources in your Slack channels and Amazon Chime chat rooms. With AWS Chatbot you can receive alerts, run commands to return diagnostic information, invoke AWS Lambda functions, and create AWS support cases. Previously, AWS Chatbot was in public preview.

To get started with AWS Chatbot, go to the AWS console, perform an authorization with Slack or Amazon Chime, and add AWS Chatbot to your channel or chat room in just a few steps. For additional information on how AWS Chatbot can help improve your team’s collaboration, see today’s post in the AWS News blog. For additional information or to get started using AWS Chatbot, please visit the AWS Chatbot homepage.

There is no additional charge for AWS Chatbot. You pay for only the underlying services that you use, like Amazon Simple Notification Service or AWS Lambda, as if you were using the services without AWS Chatbot. You can use AWS Chatbot in any commercial AWS Region. Please refer to the Regional Product and Services table for details about AWS Chatbot availability.

Ilya Bezdelev

Bibliography

[Stone 2019]

Stone, Nicki and Ilya Bezdelev and AWS Reinvent 2019 Launchpad. Available at: <https://www.youtube.com/watch?v=u-1SbynvVkA>

CodeArtifact

Introducing AWS CodeArtifact: A fully managed software artifact repository service

Posted On: Jun 10, 2020

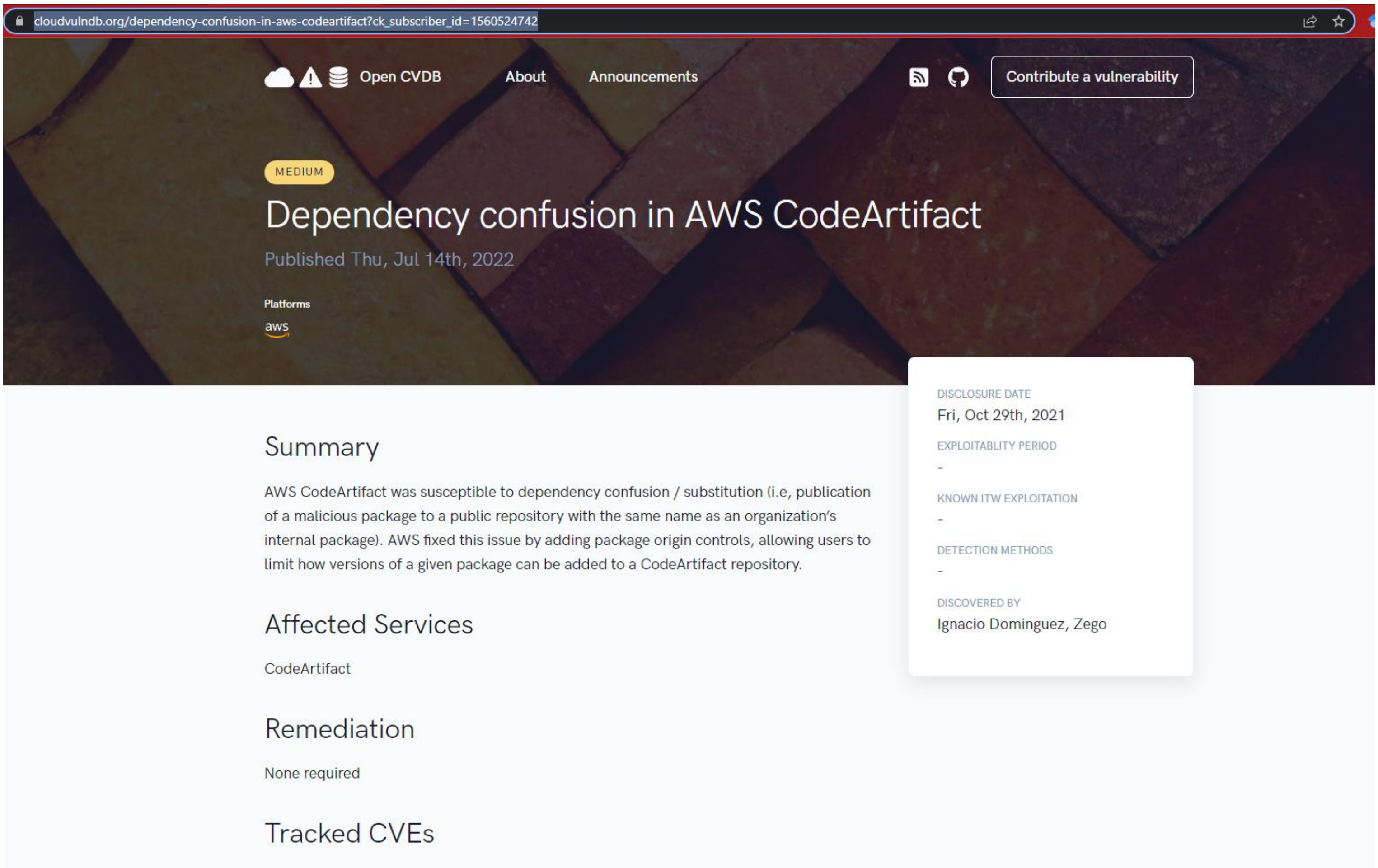
AWS CodeArtifact is a fully managed software artifact repository service that makes it easy for organizations of any size to securely store, publish, and share packages used in their software development process. CodeArtifact eliminates the need for you to set up, operate, and scale the infrastructure required for artifact management so you can focus on software development. With CodeArtifact, you only pay for what you use and there are no license fees or upfront commitments.

AWS CodeArtifact works with commonly used package managers and build tools such as Maven and Gradle (Java), npm and yarn (JavaScript), pip and twine (Python), making it easy to integrate CodeArtifact into your existing development workflows. CodeArtifact can be configured to automatically fetch software packages from public artifact repositories such as npm public registry, Maven Central, and Python Package Index (PyPI), ensuring teams have reliable access to the most up-to-date packages.

IT leaders can use AWS CodeArtifact to create centralized repositories for sharing software packages approved for use across their development teams. CodeArtifact's integration with AWS Identity and Access Management (IAM) provides them with the ability to control who has access to the packages. Further, CodeArtifact's support for AWS CloudTrail gives leaders visibility into which packages are in use and where, making it easy to identify packages that need to be updated or removed. CodeArtifact also supports encryption with AWS Key Management Service so customers can control the keys used to encrypt their packages.

To learn more about AWS CodeArtifact check out the blog, visit the CodeArtifact product page or refer to CodeArtifact documentation.

You can see a full list of AWS Regions where AWS CodeArtifact is available here.


cloudvulndb.org/dependency-confusion-in-aws-codeartifact?ck_subscriber_id=1560524742
Open CVDB
About
Announcements
Contribute a vulnerability

MEDIUM

Dependency confusion in AWS CodeArtifact

Published Thu, Jul 14th, 2022

Platforms

aws

DISCLOSURE DATE

Fri, Oct 29th, 2021

EXPLOITABLITY PERIOD

-

KNOWN ITW EXPLOITATION

-

DETECTION METHODS

-

DISCOVERED BY

Ignacio Dominguez, Zego

Summary

AWS CodeArtifact was susceptible to dependency confusion / substitution (i.e, publication of a malicious package to a public repository with the same name as an organization's internal package). AWS fixed this issue by adding package origin controls, allowing users to limit how versions of a given package can be added to a CodeArtifact repository.

Affected Services

CodeArtifact

Remediation

None required

Tracked CVEs

Bibliography

Dependency Confusion in AWS CodeArtifact. Available at: <https://www.cloudvulndb.org/dependency-confusion-in-aws-codeartifact?ck_subscriber_id=1560524742>

[Standish 2020]

Standish, John (2020). Sharing code at arm's length with AWS CodeArtifact. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=vEwv2cBS-MQ>

Amazon Honeycode

Today we are launching Amazon Honeycode in beta form. This new fully-managed AWS service gives you the power to build powerful mobile & web applications without writing any code. It uses the familiar spreadsheet model and lets you get started in minutes. If you or your teammates are already familiar with spreadsheets and formulas, you'll be happy to hear that just about everything you know about sheets, tables, values, and formulas still applies.

[Barr 2020]

Bibliography

[Barr 2020]

Barr, Jeff (2020). Introducing Amazon Honeycode – Build Web and Mobile Apps Without Writing Code. *AWS News Blog*. June 24th 2020. Available at: <https://aws.amazon.com/blogs/aws/introducing-amazon-honeycode-build-web-mobile-apps-without-writing-code/>

App2Container

Bibliography

Copilot

Amazon ECS announces AWS Copilot, a new CLI to deploy and operate containers in AWS

Posted On: Jul 9, 2020

Today, Amazon Elastic Container Service (ECS) announced AWS Copilot, a command line interface tool that helps customers develop, release, and operate containerized applications on AWS. With a single command, AWS Copilot creates all infrastructure and artifacts required to run production-ready service on Amazon ECS and AWS Fargate, including task definitions, image repositories, and AWS resources like load balancers, deployment pipelines. With AWS Copilot, users can focus on developing their applications rather than setting up infrastructure.

With AWS Copilot you can launch your service using one of the built in service patterns providing you secure and scalable infrastructure. Once your service is up and running, AWS Copilot has built in commands to easily add more services, infrastructure like databases or Amazon S3 Buckets, or even new deployment environments in different accounts and regions. While getting started and growing your application is important, AWS Copilot wants to make it easy for you to release and operate your application too. AWS Copilot lets you set up a fully functional CI/CD AWS CodePipeline with just two commands to safely deploy through multiple deployment stages. Additionally, you can view all of your services' logs and health in seconds without having to leave your terminal. Finally, with AWS Copilot, you can choose a collaboration model that fits your team. When using AWS Copilot you can work alone or share your application with other users in your AWS account.

To get started, see the following resources:

- Copilot can deploy applications to AWS Fargate supported regions and is available for download on Mac and Linux here.
- Connect with us and learn more about the project on our open source GitHub repository.
- Read about AWS Copilot on the AWS Containers blog.
- See tutorials with Copilot in our documentation.

Bibliography

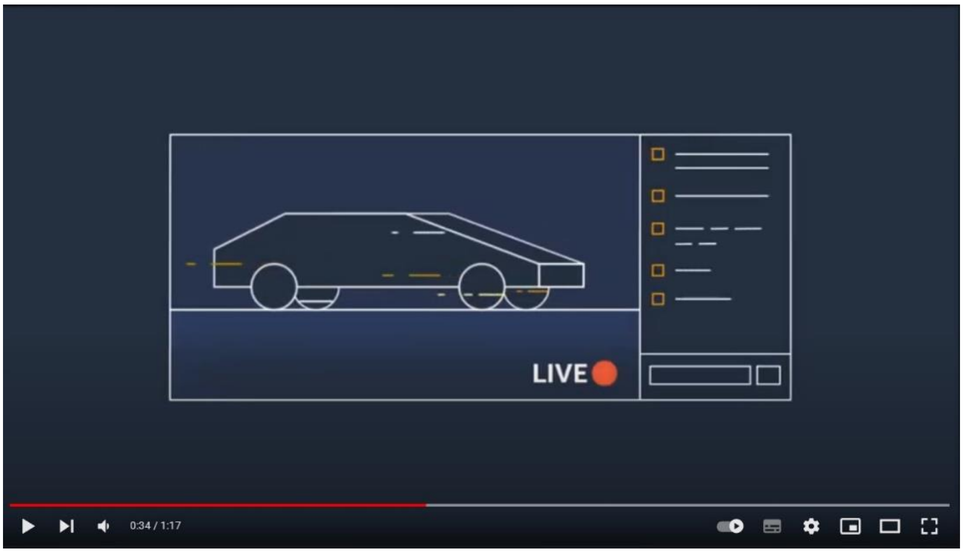
[Peck 2020]

Peck, Nathan (2020). Introducing AWS Copilot. *Containers* [Blog]. Available at: https://aws.amazon.com/blogs/containers/introducing-aws-copilot/

Interactive Video Service

This product did not feature in Andy Jassy’s keynote at ReInvent.

AWS produced a video which purports to explain what, exactly, Amazon Interactive Video Service is.



Introducing Amazon Interactive Video Service (Amazon IVS)

Posted On: Jul 15, 2020

Amazon Interactive Video Service (Amazon IVS) is a managed live streaming solution that is quick and easy to set up, and ideal for creating interactive video experiences. Send your live streams to Amazon IVS using standard streaming software like Open Broadcaster Software (OBS) and the service does everything you need to make low-latency live video available to any viewer around the world, letting you focus on building interactive experiences alongside the live video.

You can easily customize and enhance the audience experience through the Amazon IVS player SDK and a Timed Metadata API. This allows you to build a more valuable relationship with your viewers on your own websites and applications.

Today, it takes customers months to build interactive applications with video workflows for content ingestion, processing, and distribution, and then they still need to configure transcoders for adaptive bitrate (ABR) streaming to support multiple device types, select appropriate streaming protocols, set up content delivery networks (CDNs), and integrate video players.

Even after all this work, live-streamed interactive content requires low latency for a good user experience. Amazon IVS ingests the video, then automatically transcodes and optimizes it, making live video streams available within seconds across an AWS-managed global infrastructure using the same video technology Twitch uses for its live streaming service.

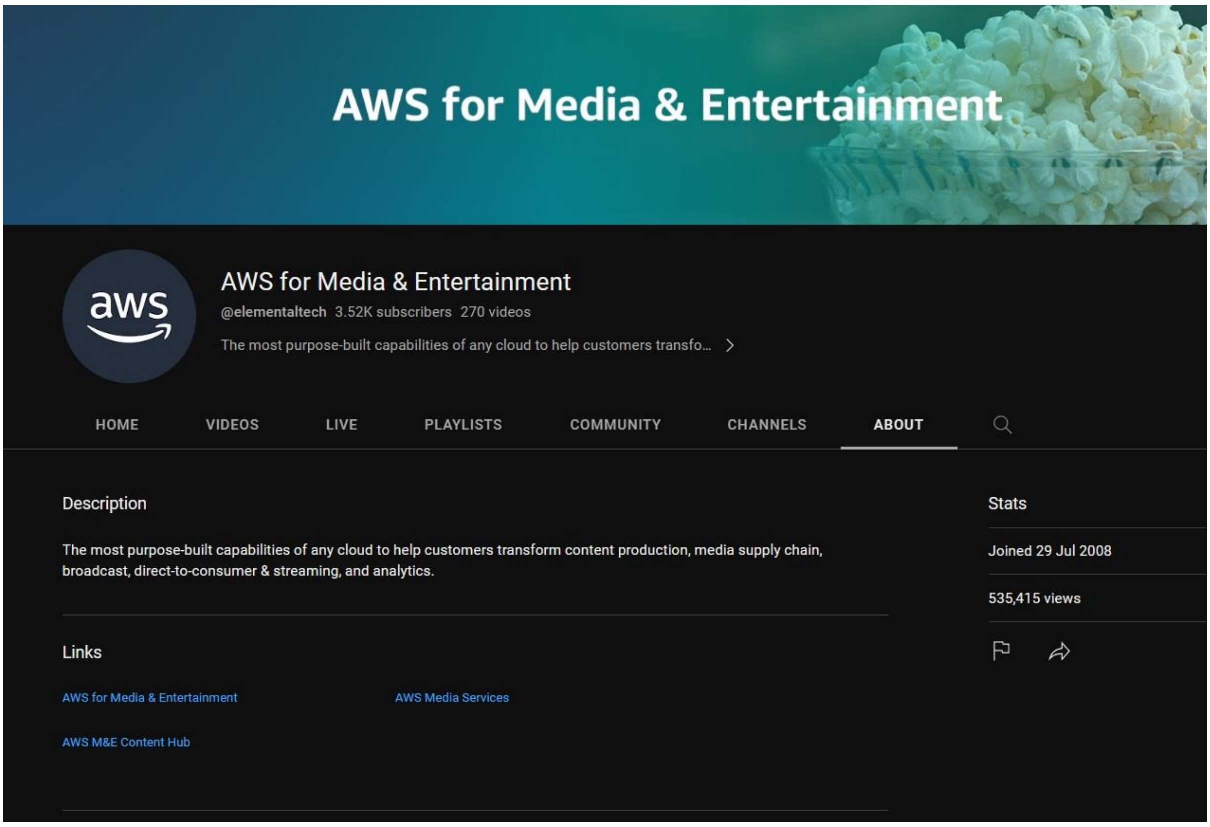
There are no upfront commitments required to use Amazon IVS, and customers pay only for the duration of video input to Amazon IVS and the duration of video output delivered to viewers.

The Amazon IVS console and APIs for control and creation of video streams are available in the US East (N. Virginia), US West (Oregon), and Europe (Ireland) regions. Video ingest and delivery are available around the world over a managed network of infrastructure optimized for live video.

To learn more about Amazon Interactive Video Service:

- [Read the AWS News blog post]
- [Visit the Amazon Interactive Video Service product page]
- [Read the Amazon Interactive Video Service Documentation]

There is a YouTube Channel devoted to the Interactive Video Service:



AWS for Media & Entertainment

@elementaltech 3.52K subscribers 270 videos

The most purpose-built capabilities of any cloud to help customers transfo...

Description

The most purpose-built capabilities of any cloud to help customers transform content production, media supply chain, broadcast, direct-to-consumer & streaming, and analytics.

Links

AWS for Media & Entertainment

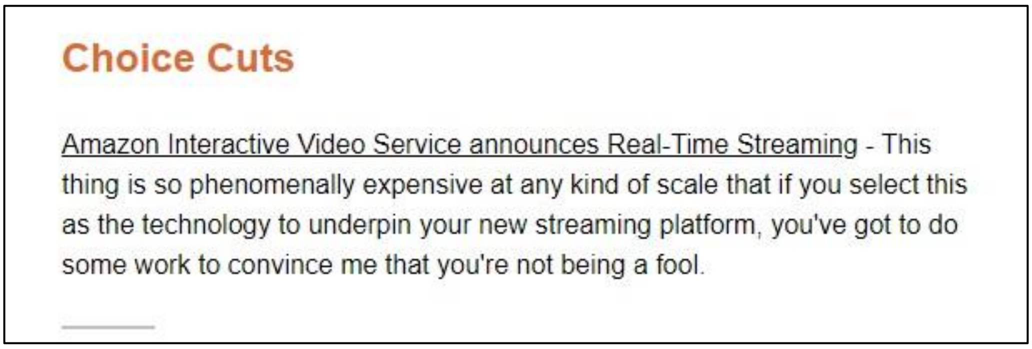
AWS Media Services

AWS M&E Content Hub

Stats

Joined 29 Jul 2008

535,415 views



Choice Cuts

Amazon Interactive Video Service announces Real-Time Streaming - This thing is so phenomenally expensive at any kind of scale that if you select this as the technology to underpin your new streaming platform, you've got to do some work to convince me that you're not being a fool.

Email from Corey Quinn on 14th Aug 2023

Bibliography

I. Official

[AWS 2020]

Amazon Interactive Video Service Explained. YouTube Channel: AWS for Media & Entertainment. Available at: <https://www.youtube.com/watch?v=Nhx6yAtk0ec>

[AWS 2020b]

Introducing Amazon Interactive Video Service (Amazon IVS). Available at: <https://aws.amazon.com/about-aws/whats-new/2020/07/introducing-amazon-ivs/>

[AWS 2023]

Amazon Interactive Video Service announces Real-Time Streaming. [Announcement]. Aug 7th 2023. Available at: <https://aws.amazon.com/about-aws/whats-new/2023/08/amazon-interactive-video-service-real-time-streaming/?ck_subscriber_id=1560524742>

Migration Evaluator

AWS produced a YouTube video [AWS 2020]. It explained what AWS Migration Evaluator is.

aws migration evaluator

Formerly TSO Logic

Fast Track your Business Case for AWS

We are told:

> In this video, you'll see how to fast-track your business case for AWS with AWS Migration Evaluator (formerly TSO Logic).

Bibliography

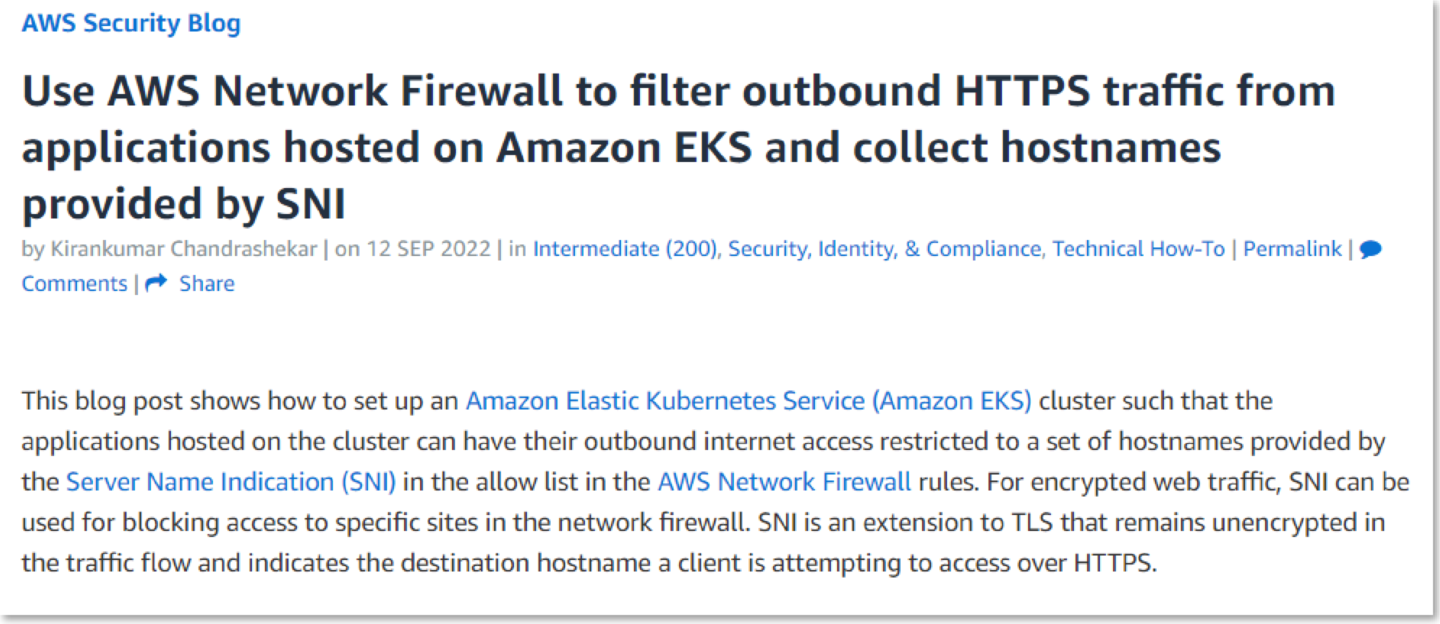
[Faraz 2022]

Shafiq, Faraz (2022). How to migrate, modernize and grow using the AWS MAP. *ReInvent 2022* [Conference]. Available at: <https://www.youtube.com/watch?v=yXCjuAaXACo>

[Miller 2019]

Miller, Ron (2019). AWS Makes another acquisition, grabbing TSO Logic. *TechCrunch*. Available at: <https://techcrunch.com/2019/01/15/aws-makes-another-acquisition-grabbing-tso-logic/>

Network Firewall

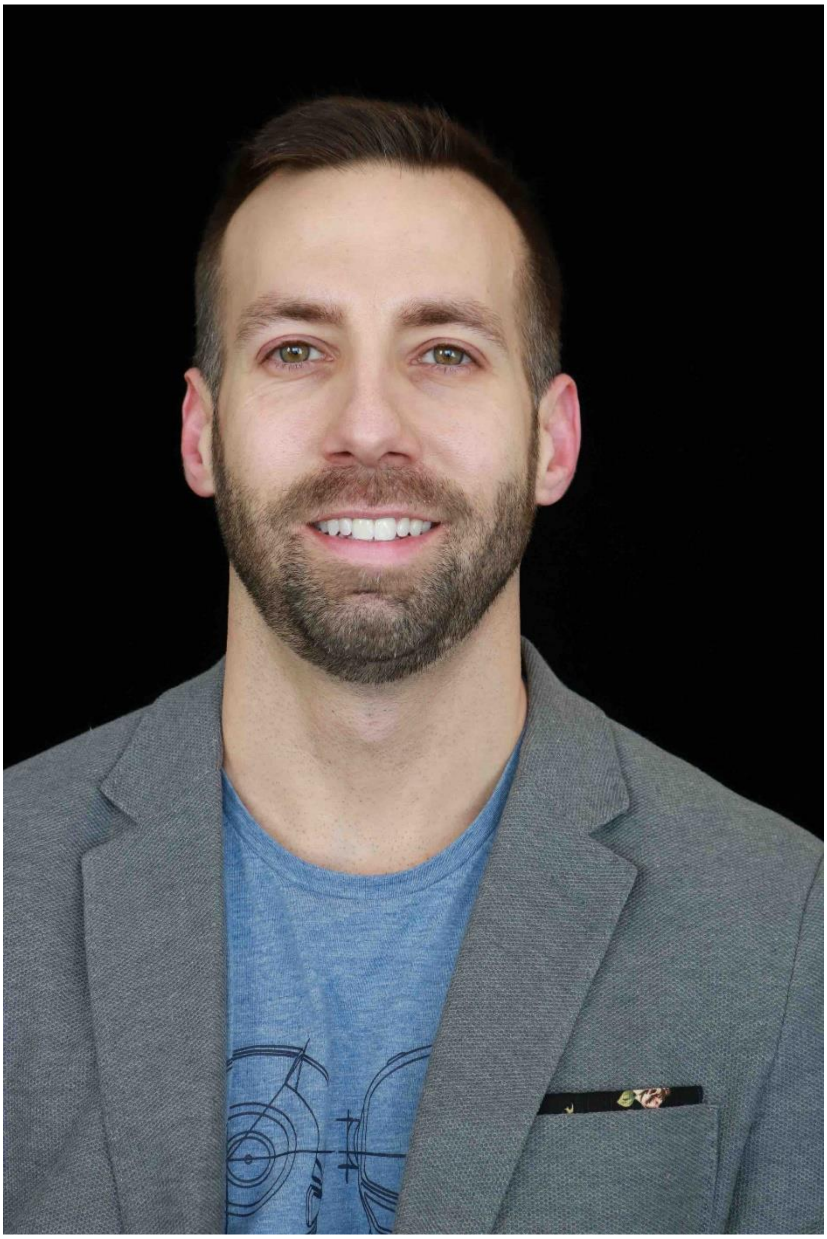


AWS Security Blog

Use AWS Network Firewall to filter outbound HTTPS traffic from applications hosted on Amazon EKS and collect hostnames provided by SNI

by Kirankumar Chandrashekar | on 12 SEP 2022 | in Intermediate (200), Security, Identity, & Compliance, Technical How-To | Permalink | Comments | Share

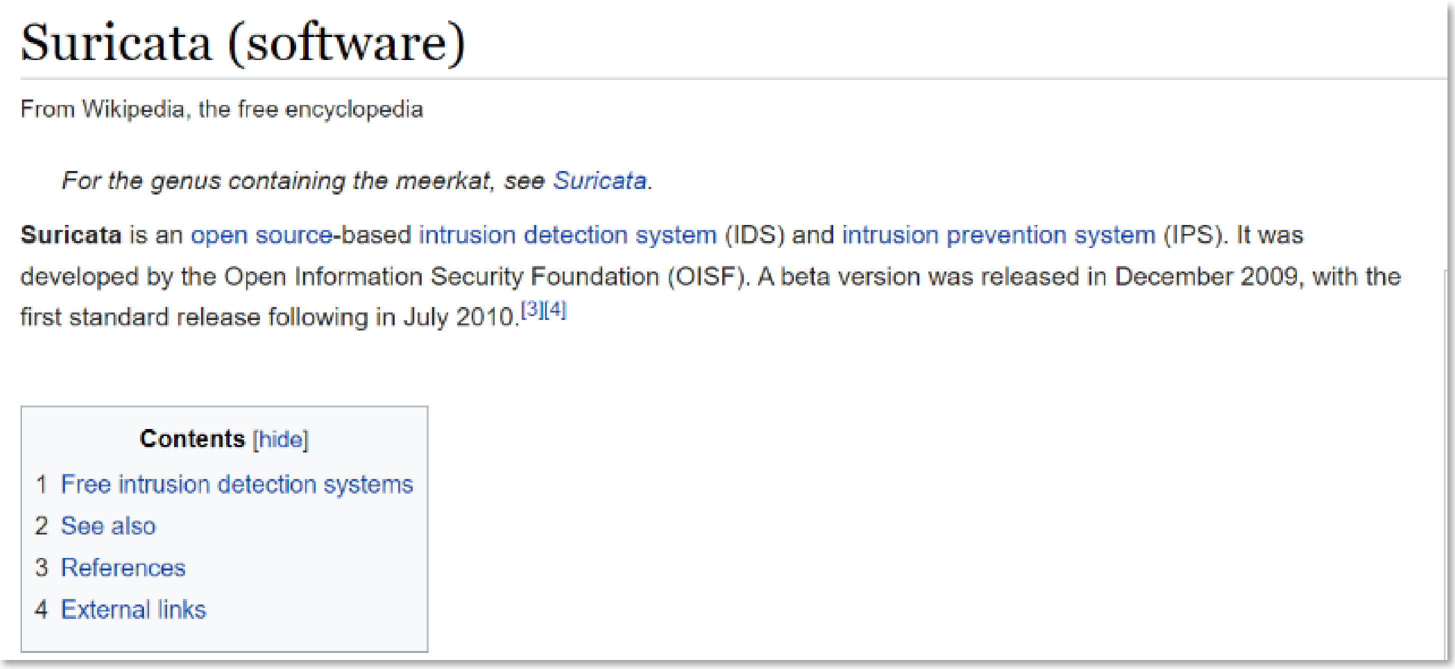
This blog post shows how to set up an Amazon Elastic Kubernetes Service (Amazon EKS) cluster such that the applications hosted on the cluster can have their outbound internet access restricted to a set of hostnames provided by the Server Name Indication (SNI) in the allow list in the AWS Network Firewall rules. For encrypted web traffic, SNI can be used for blocking access to specific sites in the network firewall. SNI is an extension to TLS that remains unencrypted in the traffic flow and indicates the destination hostname a client is attempting to access over HTTPS.



Joel Desaulniers

Patrick Duffy (2021) mentions Suricata:

> Network Firewall helps customers protect their VPCs by protecting the workload at the network layer. Network Firewall is an automatically scaling, highly available service that simplifies deployment and management for network administrators.
>
> With Network Firewall, you can perform inspection for inbound traffic, outbound traffic, traffic between VPCs, and traffic between VPCs and AWS Direct Connect or AWS VPN traffic. You can deploy stateless rules to allow or deny traffic based on the protocol, source and destination ports, and source and destination IP addresses. Additionally, you can deploy stateful rules that allow or block traffic based on domain lists, standard rule groups, or Suricata compatible intrusion prevention system (IPS) rules.



Suricata (software)

From Wikipedia, the free encyclopedia

For the genus containing the meerkat, see Suricata.

Suricata is an open source-based intrusion detection system (IDS) and intrusion prevention system (IPS). It was developed by the Open Information Security Foundation (OISF). A beta version was released in December 2009, with the first standard release following in July 2010.[3][4]

Contents [hide]

What is Suricata?

Suricata is indeed a genus of mammals. They are part of the family called The most famous species within the family is perhaps the meerkat. The meerkat is a small mongoose found in Southern Africa.

Meerkats can be helpful: they've been known to be used to kill rodents in South African rural households. There are ways in which meerkats are economically important for crop protection. For example, they hold back the intrusion of disease by hounding out yellow mongooses which are ridden with rabies. They're generally tame, but ultimately unsuitable as pets because sometimes they become slightly too aggressive.

Suricata the software application, is an example of an intrusion detection system (IDS). It was first released in 2010, and so it about a decade old now. Its conventional to refer to the particular version of Suricata you are interested in. For example, Suricata 5 was released in October 2019, the year before the AWS Network Firewall was announced. Suricata 6 was released in October 2020, a mere month before AWS Network Firewall was announced in December 2020.

Suricata 1.0 was published in July 2010, after about two years of development. It was building upon Snort, which is an Intrusion Detection System (IDS) which had itself already been around for about a decade. One of distinguishing features of the first version of Suricata was that when it looked for keywords in the protocol fields of an HTTP request, the exact protocol or port did not matter. "For the first time, a signature could ask for a specific protocol field without having to do the protocol parsing itself" [Leblond 2021].

The unique ability possessed by Suricata can be summed up as a "port-agnostic approach". The port of the traffic did not matter. Leblond tells us that Suricata 1.0 also

> broke with the previously strict definition of what an IDS was: logging HTTP requests to a file. This was not in the initial specifications, but it turned out to not be too complex to build and did not have a major impact on performance. This opportunistic approach continues to uniquely define all Suricata development.

[Leblond 2021]

The Suricata project received public funding, allowing it to continue. Now, it seems that the major funding model is that private organisations, which are members of a consortium, contribute. It is guided by the Open Information Security Foundation (OISF).

However, it remains correct to describe the project as "open source". It is also more than an intrusion detection system. It is now apt to describe it as an intrusion *prevention* system. It achieves network security monitoring and full packet capture (FPC) [Manev 2021].

Recap of basic AWS concepts

A person's ability to understand AWS Network Firewall would probably be enhanced were they to be the kind of person who has just been given a refresher on the basic AWS concepts.

Four Container Concepts

AWS makes use of subnetworks, VPCs, Availability Zone and Regions. Now, two of those four container concepts are logical constructions and two have a physical grounding.

AZs are nothing but

(i) units of infrastructure, such that

(ii) the deterioration of that unit
(iii) is impossible to cause the deterioration of another AZ in the Region.

Since the deterioration is to include physical deterioration (through earthquake or flood etc), AZ must—by definition—be physically distinct. Strictly, this could be a very small distance, but of course many AZs are considerable distances from one another. This is enough to establish that the concept of an AZ has a physical grounding.

The concept of an AZ is not the same as the concept of a data centre. An Availability Zone might consist of *multiple* data centres, or just one data centre.

The relation between VPCs and Subnets is easy to understand. A VPC is, roughly, a virtual network. The segments which this is carved up into are your sub-networks (or subnets). Just as every fish must be in water to survive

FISHBOWL 1

EVERY EC2 INSTANCES MUST BE IN A SUBNET.

Since subnets are nothing but segments of your VPC (network), it follows that an EC2 instance *will* live within a VPC.

FISHBOWL 2

EVERY EC2 INSTANCE WILL BE IN A VPC.

A VPC will have a CIDR block. A CIDR block represents a range of **possible IP addresses**.

If I ask somebody to write down a CIDR block on a napkin, how will I know that what they've written down is a CIDR block? There will be four numbers, separated by periods. At the end will be a forward slash character, with some number after it. The numbers between the periods must be between 0 and 255 (inclusive). The number after the slash must be between 0 and 32 (inclusive). This is one way you can tell you are looking at a CIDR block.

An IP address is, if nothing else, 32 bit positions. Each bit-positions can have a value of either 1 or 0. However, each bit-position plays a different role. Some bit positions are used for identifying the network on which the host resides. Other bit positions really do concentrate on identifying the host. This means that the 32 bit-positions can be thought of in terms of a network-ID region, followed by a host ID region. The number of bit-positions used to identify the network is not always the same. Network engineers developed ways of communicating exactly how many bit-positions they were using to identify the network.

You're amazing, Alan

They'd pronounce things like:

> "I'm Alan. Acknowledge me.
>
> I've amassed an absolutely enormous number of hosts into a network. I work for a big company.

- Alan

Turns out Alan had a very tiny network ID portion. He only needed eight bit positions (the first byte) for identifying the network.

These sorts of IP addresses, with just the first 8 bit-positions used for the network ID, became known as Class A addresses.

Carol's Cartel

Of course, it's also sometimes necessary to perceive, or ***see***, a greater number of networks. Carol desires to see more networks, by using a greater number of bit-positions for the network ID portion. *With more bit-positions, you have more permutations*. Some people carefully carve out the first 24 bit positions for the network ID. This is three times more than Alan used (he used just 8 bit positions). Such IP addresses—where the first 24 bits are used for network identification—are known as Class-C addresses.

Eventually, people grew tired with being limited to fixed numbers. The choice was essentially threefold: Be like Alan, using 8 bits for a few large networks. Or be like Carol, using 24 bits for the network ID portion.

The character we haven't mentioned is this 16 year old called Ben. You could, of course, be like boring Ben. He used 16 bit-positions for the network ID portion. This is more than Alan, and fewer than Carol, who uses 24 positions. Anyhow, the choice is threefold: the choice is between Class A, B, or C.

Eventually, the switch was made to Class-less routing. We would no longer limit ourselves to using either 8 or 16 or 24 bit-positions for the network ID. Instead, you can use any number of bits-positions for the network ID. You could choose to use 8 or 9 or 10 (and so on) bit positions for the network ID portion. This is the birth of class-less routing of traffic between domains.

When you create a VPC on AWS, you essentially create a network. Like a family surname which will be latched on to the end of *any* child born in the family home, your VPC has some a network ID, which will form the first part of *any* IP address allocated to a resource born in the VPC. Recall that "network ID" denotes the bit-positions at the beginning of an IP address. These bit-positions are *not* for identifying hosts.

As well as establishing an actual network ID for the VPC, we need to communicate to people how many bit-positions we've decided to use for the network-ID portion. Have we decided to be similar to Alan, using about 8 bits for our network ID region? Or will we use a similar number of bit positions as Carol, who used 24 bit positions? We tell people what we've settled on by printing a forward slash at the end of the actual network ID. The forward slash is followed by a number. The number denotes the number of bit-positions we're using to describe the network ID.

Now, remember that there are not necessarily any *actual* resources yet. When we create a VPC, we are creating the portion of the address which will go in front of *any* resource born on our patch.

We need a name for this whole combination. That is, the combination of the actual network ID and the number which tells people how many bit-positions we're using to describe the network ID. We *could* call it a *RANGE OF IP ADDRESSES*. One disadvantage of this expression is that is could denote any collection of IP addresses, perhaps ones possessed by IP instances.

Ideally, we'd use a term that implies that the set of IP addresses are "up-for-grabs". We'd use a term that does not suggest that the IP addresses are already being used by EC2 instances. At the same time, we don't want to suggest unbounded possibilities! The term should make clear that the IP address must begin with the network IP. And it must abide by the number we've printed after the forward slash.

The term which most readily suggests **POTENTIAL**, but within bounds, is *block*. Consider someone carving a statue from a block of marble. It is often said that the statue is present in potential. Note how the sculptor:

(1) must obtain the block from something larger

(2) cannot expand the size of his block once he has it;

(3) *can* create things by using the block (e.g. statue)

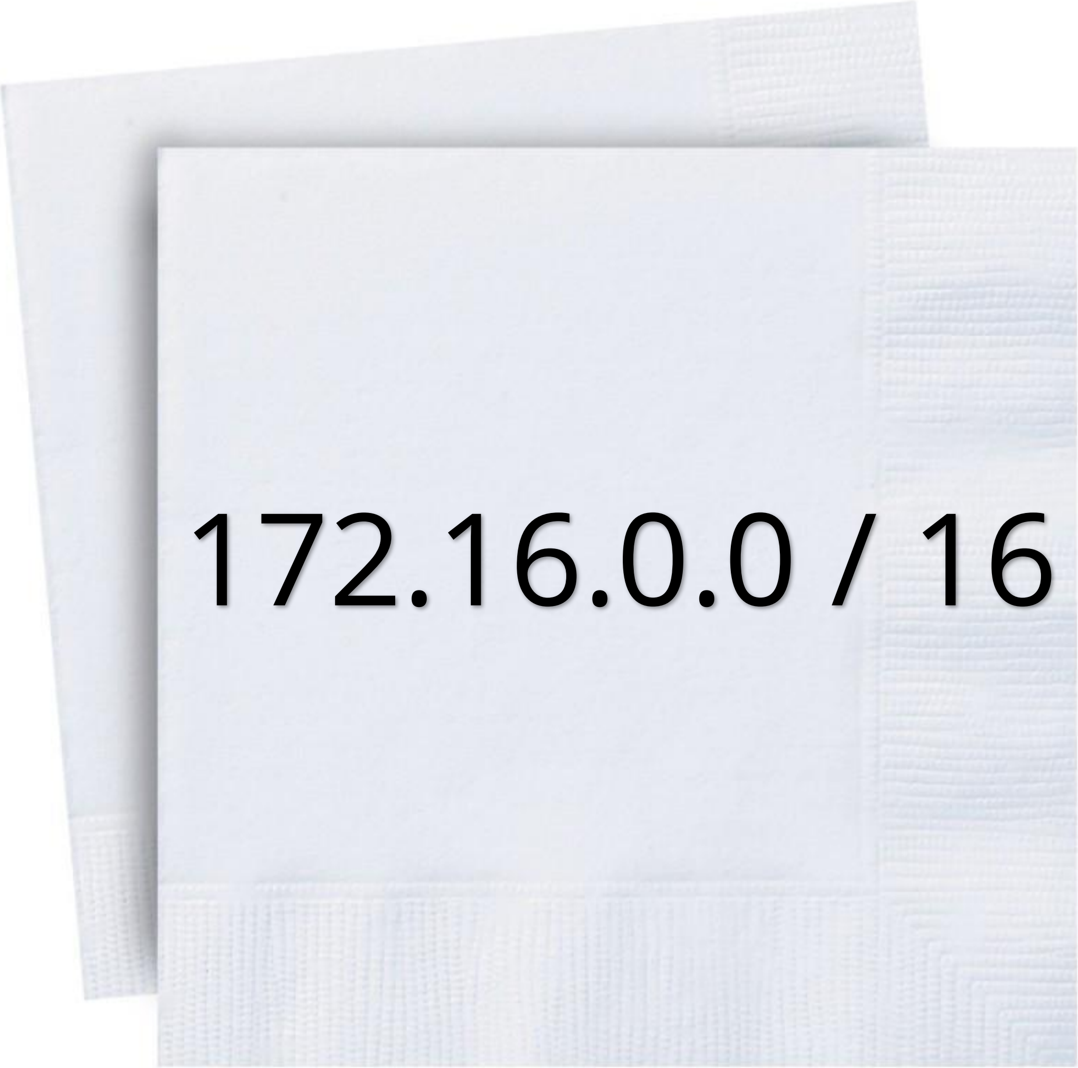
(4) does not have nothing.

Even though he doesn't have a statue yet, it doesn't mean he has nothing! Similarly, the block you're given with your VPC is *not* an actual IP address. This is despite it bearing a superficial similarity to one (there will be four numbers separated by periods). Even though your block is not an IP address, it's not nothing.

Your VPC block depicts a range of potential IP addresses. The IP addresses are made *actual* when EC2 instances, or other resources, are blessed with them. Just as giving children surnames allows them to be uniquely referred to in public spaces, assigning IP address to resources enables *communication*. IP addresses help us to route packets from one domain to another domain.

Recall that we decided to abandon the strict classes. The classes constrained us in this way: "Your surname must be 8 characters" (Class A); "you're surname must be 16 characters" (Class B) and so on. We decided that the 'surname'—or, the initial portion of the IP address used for identifying the network—could be *any* number of bit positions. It's for this reason that the block—charged with potential IP addresses—is described as the *Classless* Inter-Domain Routing Block. The expression "inter domain" simply means "between domains". This might be contrasted with communication *within* a domain (intra-domain). *The IP addresses derived from the CIDR block really are unique.*

If I ask somebody to write down on a napkin the **CIDR block** they are using, they might write this on the napkin:



A CIDR block can communicate to me three things:

1. The bit positions possibly used for expressing a network ID.

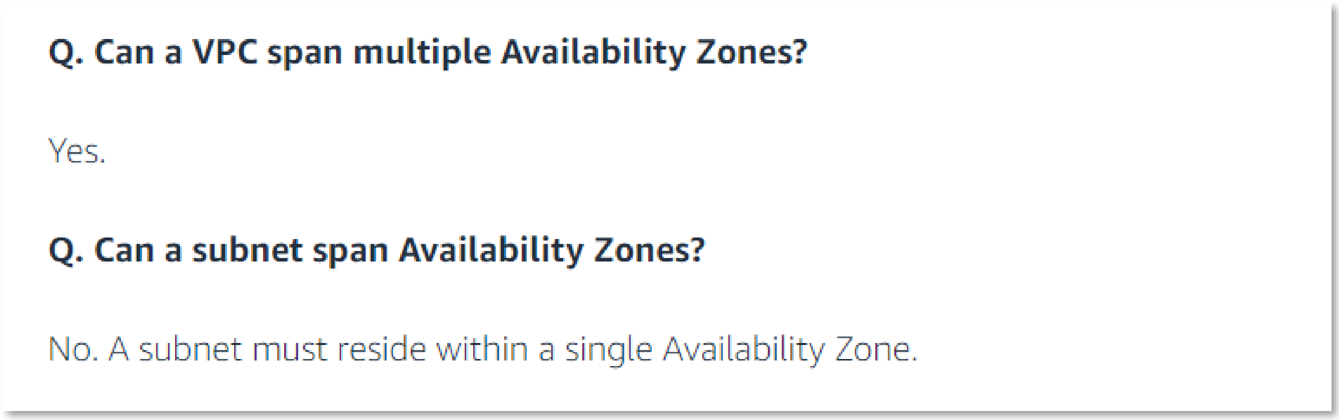
2. An actual network ID

3. An actual host ID (IP address).

The number after the forward slash only designates (1). Usually, the CIDR block communicates (1) and (2) to us. Occasionally, it communicates all three things.

The VPC (your network) is divided into smaller chunks (your sub-networks).

Now I want to talk about the relationship between these logical concepts (VPC, sub-networks), and the physically-grounded concept of an Availability Zone. AWS tell us that a subnet *cannot* in fact split its soul into multiple pieces, and sprawl across availability zones (See [Overflow 2020] for speculation as to why). A sub-network must live with the fact that if its home Availability Zone is engulfed in flames, it will die. Interestingly, however, the VPC ("network"), out of which the subnet is carved, *can* split its soul. The VPC *can* sprawl across AZs.



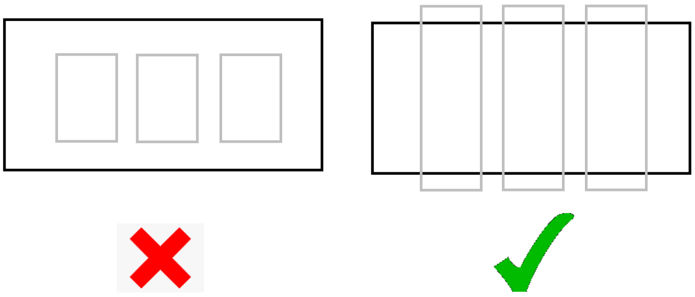
Q. Can a VPC span multiple Availability Zones?

Yes.

Q. Can a subnet span Availability Zones?

No. A subnet must reside within a single Availability Zone.

Sometimes, people draw architectural diagrams with the AZs inside the VPC. This suggests that a VPC contains an availability Zone, which is of course absurd. Really, the VPC transcends the AZs, or *spans* the AZs.



It is perhaps better to say that while the VPC *transcends* AZs, the very subnets which comprise it, are AZ-bound. (This fact itself is somewhat puzzling)

Recall that EC2 instances *must* reside within a subnet. The more general claim is of course that EC2 instances must be in your VPC (since sub-networks are simply portions of your VPC network).

VPC transcend Availability Zones but *are* bound to exactly one AWS Region. Every AZ exists in exactly one AWS Region. Like AZs, Regions are independent of one another, such that if one fails, this will not cause a failure in another. The AWS Region also has a physical grounding; there are Regions all over the world.

Regions *contain* AZs. Every AZ sits in exactly one AWS Region.

Region-bound Services

Most AWS services are bound to a particular Region. This means two things. First, that for some AWS Services, they simply are not available for you to use in certain Regions. Second, that that you use the AWS Service, all its resources inhabit one region only (despite, perhaps, the service being available in all Regions). Famously, about three services **transcend** the AWS Regions. They include **CloudFront**, the content-delivery network announced at the end of AWS's Foundational period, **IAM**, the access-control service announced in 2010, when AWS was furnishing its firmament, and **Route 53**, a DNS service, capable of, for example, mapping friendly names for the content-delivery networks announced two months earlier (for discussion of which AWS services are global, see [Overflow 2021])

The distinction between AWS services which are global and those which are Region seems to be an informal characterisation. The tenuous nature of the distinction is brought out in a service such as S3. The page displayed on the Management Console for S3 prohibits the user from selecting a Region in the bar at the top, stating:

> S3 does not require region selection.

Also, bucket names are required to be globally unique. That is, there cannot be two AWS buckets *on earth* with the same name. We have two reasons to characterise S3 as a global service.

On the other hand, creating a bucket on the console involves selecting an AWS Region from a drop-down list. Buckets are region-bound. This leads people to characterise S3 as Region-bound.

There is no incoherence in S3; it merely puts pressure on the distinction between Global and Region-bound services. S3 buckets themselves are Region-bound. The Management Console does not see Region boundaries, listing buckets from across all Regions. The utility of the Cross-Region Replication feature remains.

Routing Principles

Routing on AWS is guided by certain principles. One such principle is known as LPM. Scientists will be tempted to attest that this stands for Litres per minute; some computer scientists might know it as Load Program Memory; Lines per minute can be a measure of the efficiency of printers. But here, LPM stands for Longest Prefix Match.

LPM

Longest **P**refix **M**atch

We find the principle of Longest Prefix Match stated in the VPC documentation.

> We use the most specific route that matches either IPv4 traffic or IPv6 traffic to determine how we route the traffic.

Don’t be startled by the fact that the LPM principle is here stated here in terms of the “more specific route”.

One route, A, is more specific than another route, B, if and only if:

> A has a longer prefix than B.

What does it mean for A to have a longer prefix than B? Well, all IP addresses consists of 32 bit positions. *Some* of these are bit positions are requisitioned for the task of identifying the network; the rest of the bit positions devote themselves to identifying hosts on the network. It is conventional to communicate precisely how *many* positions are used for network identification – in any given case—by writing a forward slash, followed by a number. For example, /8 communicates to readers that the first *eight* bit-positions have been requisitioned for identifying the network.

The name for the set of bit-positions that devote themselves to identifying the network, is the **prefix**.

Generally, a prefix is some group of characters affixed to the beginning of lots of different things.

An example of a prefix is “re”. Words such as heat, route, and capitulate can all have this prefix attached. They become re-heat, re-route and recapitulate, respectively. A second example of a prefix in the English language is “para”. It is observable in *paranormal*, *paratrooper* and *parallel*. Notice how these two prefixes (re- and para-) have a differing number of letters.

The Ambiguity in “Prefix”

With IP addresses, we call the bit-positions devoted to identifying the network the *prefix*. The use of the term—*prefix*—is ambiguous. Sometimes it refers merely to the number of bit positions allocated for the network ID (staying silent on what the network ID actually *is*). At other times, we refer also to what the network ID actually *is*. We might call these two senses the formal and substantive senses.

The formal sense is very common. I might ask “what is the prefix?” and someone would respond “/16”. And this would be seen as perfectly respectable. All this tells me, however, is the *form* of the prefix. I’m merely told that the prefix sprawls across 16 bit-positions. I have not been told what the value adopted by each position actually *is*.

The formal sense of “prefix” is counterintuitive. A prefix is something that is common to lots of things. It is the portion of the IP address common to all the hosts on the network. However, if I ask for a prefix and you tell me **/16**, then you have not told me what the thing which is common to all the hosts on the network actually *is*. This can be quite counterintuitive for novices.

We might say that A has a longer prefix than B. This means that in A, more bit-positions are devoted for the network ID than in B. If A has /16 and B has /8, then A has a longer prefix than B. This is because more bit positions are devoted to the network ID (as opposed to the host ID). This, then, is what we mean when we say that A has a longer prefix than B.

If A has a longer prefix, it is more specific. Why is this? Why is it that a longer prefix (/16 rather than /8) is described as “more specific”?

Well, with a longer prefix, you have more bit-positions which you can use to specify networks. This allows more permutations. The number of networks which you can single out is greater. If you used 3 bit-positions, you’d be able to specify 2^3 networks. If you used 10 bit-positions, you’re able to 2^{10} networks. So, in this sense, larger prefixes have a greater *specifying power*. Longer prefixes are more specific.

To recapitulate, let us formulate the principle of Longest Prefix Match (LPM):

Longest Prefix Match (LPM)

The IP address which devotes a **greater** number of bit-positions to network identification is given priority over those IP addresses which devote fewer bit positions

E.g., a /16 CIDR range will have priority over /8

There are some over principles relating to routing which we need to cover. We need to consider what might be meant by “route propagation”, for example. Generally, *to propagate* means to spread out. This raises some questions. Why is the AWS VPC documentation talking about propagation? What, exactly, is spreading out? What is its origin and what is it spreading out *to*?

To begin the answer, let’s look at what the documentation says. It states:

> If you've attached a virtual private gateway to your VPC and enabled route propagation on your subnet route table, routes representing your Site-to-Site VPN connection automatically appear as propagated routes in your route table.
>
> AWS VPC

So, routes representing a Site-to-site VPN appear as propagated routes in our route table. We can ask what this means. What does it mean for a route to appear in the route table *as a propagated route*?

In computer networking, there is a distinction between static routing and dynamic routing. Static routing is associated with the idea of manually entering routes into a route table. Wikipedia tells us:

> In many cases, static routes are manually configured by a network administrator by adding in entries into a routing table, though this may not always be the case.

The Wikipedia article is clear, however, that manual configuration is not an essential property of static routes. The word "static" is a hint that:

> static routes are fixed and do not change if the network is changed or reconfigured.

Dynamic routing is, roughly, the opposite of static routing. Wikipedia tells us that:

> **Dynamic routing**, also called **adaptive routing**,[1][2] is a process where a router can forward data via a different route for a given destination based on the current conditions of the communication circuits within a system.

The AWS VPC documentation talks about route propagation. I think propagation is the means by which dynamic routing is achieved. Some sort of origin updates various outposts with the route required to reach the origin. In the case of AWS, the origin always seems to be a Virtual Private Gateway. The AWS documentation writes:

> Route propagation allows a Virtual Private Gateway to automatically propagate routes to the route tables. The means that you don't need to manually enter VPN routes to your route tables.

AWS do not tend to talk about "static routes" and "dynamic routes". Instead, they talk about "static routes" and "*propagated* routes". It seems to be the case that here we have two types of routes. The entries in a route table are either STATIC ROUTES or PROPAGATED ROUTES.

As we have seen, routes can overlap. To use the terminology we have developed, Concurrent Coverage applies with route tables. That is, a single item of traffic can be "covered" by more than one route.

The question raised, therefore, is how we resolve Concurrent Coverage. If an entity if covered by multiple rules, how do we determine which rule is effective? NACLs resolve using the Priority Principle; IAM policies resolve using the Dagger Deny. AWS are clear that the type of route is pertinent here. Whether the route is a STATIC ROUTE or a PROPAGATED ROUTE is important. AWS are clear that if a STATIC and PROPAGATED route compete, static routes have priority.

STATIC SUPREMACY

IF THERE IS A STATIC ROUTE AND A DYNAMIC ROUTE, THEN THE STATIC ROUTE IS GIVEN PRIORITY.

We see support for the principle of Static Supremacy in the AWS VPC documentation. It states:

> If the destination of a propagated route overlaps a static route, the static route takes priority.

[AWS VPC]

So, we have now explored two principles, Longest Prefix Match (LPM) and Static Supremacy. A higher-level question is now raised: What if some situation is concurrently covered by both LPM and Static Supremacy? How would we resolve this?

The Specific Propagated Route

For example, suppose that there are two routes, one static and one propagated. Suppose that the propagated route is "more specific". In other words, the propagated route has a longer prefix.

Which route should be given priority? According to LPM, it should be the propagated route, because it has a longer prefix. Yet Static Supremacy, implores us to give the static route priority. How do we resolve this disagreement?

The AWS documentation states:

> Please note that for routes that overlap, more specific routes always take priority irrespective of whether they are propagated routes, static routes, or routes that reference prefix lists.
>
> VPC guide

This makes it clear that even the **STATIC SUPREMACY** principle bows down in the face of **LPM**. More specific routes, which have longer prefixes, are given absolute priority (even if they *are* propagated).

Target as "LOCAL"

Sometimes, in route tables, the TARGET column will simply have the word LOCAL printed in it. When this occurs, the DESTINATION column tends to contain the CIDR block for the VPC.

This route is distinguished from other routes, owing to the fact that it does not concern a single resource outside of the VPC. Instead, it is concerned with resources within the VPC. Its concern is not global, but local. AWS tell us that:

> there is a local route that allows the resources in the VPC to communicate with each other
>
> VPC Guide (Get Started)

Once again, the term "default" can cause trouble here. Sometimes, when people talk about a "default route", they are concerned with a route whose destination is "0.0.0.0/0". For example, a popular textbook for solutions architects has a section called THE DEFAULT ROUTE and writes:

> To enable Internet access for your instances, you must create a default route pointing to the Internet Gateway.
>
> Piper and Clinton 2021, p.95

The VPC user guide states that the local route is:

> A default route for communication within the VPC.

However, the local route is **not** a catch-all route, and is **not** to be identified by looking for an entry containing 0.0.0.0/0. So, why do AWS describe the local route as a default route. I suspect that it is because this route does not need to be added manually:

> Every route table contains a local route for communication within the VPC. This route is added by default to all route tables.

So, in the context of routing, the expression "default route" can either denote:

1. An innate route
   A route which exists without ever having been manually created

2. A catch-all route
   A route which covers traffic in situations in which more-specific routes do not cover the traffic.

And (1) does not tell us anything about whether the route is indestructible.

In fact, the local route is innate and also indestructible. In this respect, it is similar to the Default Deny, the rule marked by an asterisk, which lurks at the bottom of NACLs.

Strictly speaking, local routes do not belong to a VPC but rather to a CIDR block. For most intents and purposes, this is pure pedantry, since every VPC must have a CIDR block. This subtlety came to light on 29th August in 2017, a Tuesday on which North Korea fired a missile above Japan and Dua Lipa's song "New Rules" was no.1 in the charts. AWS announced that you could now add more than one CIDR range to your VPC. It's conventional to denote the CIDR range that accompanied the creation of the VPC as the "primary CIDR" and to denote the one added later the "secondary CIDR range".

Crucially, if you have two CIDR ranges, you can in fact have a route table with *two* entries with a target as LOCAL.

This is all summarised in this paragraph:

> Every route table contains a local route for communication within the VPC. This route is added by default to all route tables.
>
> If your VPC has more than one IPv4 CIDR block, your route tables contain a local route for each IPv4 CIDR block. If you've associated an IPv6 CIDR

> block with your VPC, your route tables contain a local route for the IPv6 CIDR block.
>
> You cannot modify or delete these routes in a subnet route table or in the main route table.

VPC documentation

Implicit Action

We have already familiar with the distinction between *implicit* and *explicit*. In IAM, all principals are implicitly denied from performing actions; the Security Group guarding the inner sanctum which is the ENI implicitly denies all traffic. As for NACLs, they have an indestructible rule, which denies everything. (Because a rule accounts for this denial, it is questionable whether this denial is really an *implicit* denial.) If an implicit action is one that applies to the majority of entities, covering things not encountered before, then you can use rules to toggle the implicit action. Expose the Default Deny and the unknowns will be denied. Lay an Apron Allow over the Default Deny and the unknowns will be allowed. WACLs certainly allow you to toggle the implicit action, but using radio buttons, rather than installing a rule.

This distinction between implicit and explicit comes to the fore once again when we think about routing within the VPC. We are told that:

> When you create a VPC, it automatically has a main route table.
>
> When a subnet does not have an **explicit routing table** associated with it, the main routing table is used by default.

[VPC (Route Tables)]

This **MAIN ROUTE TABLE** is a character we should acquaint ourselves with.

Figure 1. A Main Route Table, made out of solid oak, sourced in Northeastern USA.

It is true that every subnet must be associated with a route table. The **Main Route Table**, however, belongs to the VPC. It is innate:

> When you create a VPC, it automatically has a main route table.
>
> AWS VPC

As well as being innate (present from the birth of the VPC), the **Main Route Table** is indestructible. The **Main Route Table** plays a very important role in the successful running of your VPC, stepping in to provide guidance for the sub-networks who need it. We're told:

> When a subnet does not have an explicit routing table associated with it, the main routing table is used.

So, we might say that the rules contained in the Main Route Table are constitute the implicit action for the subnets in the VPC. The main route table *will be used* if no explicit route table is created for individual subnets. You can alter the rules contained in the Main Route Table:

> You can add, remove, and modify routes in the main route table.

At the moment, it seems that every subnet has just two options. They can be (1) implicitly associated with the route table, or (2) explicitly associated with their own route table. In fact, though, subnets can make their association with the Main Route Table more intentional. We're told:

> You can explicitly associate a subnet with the main route table, even if it's already implicitly associated.

Now, why would you want to do this? An association is an association, after all. Well, it's possible to replace the Main Route Table. And when a sitting Main Table is deposed in this way, we might be sympathetic enough to allow certain subnets to retain their association with the outgoing Main Table. This is achieved by making the association explicit. Note that:

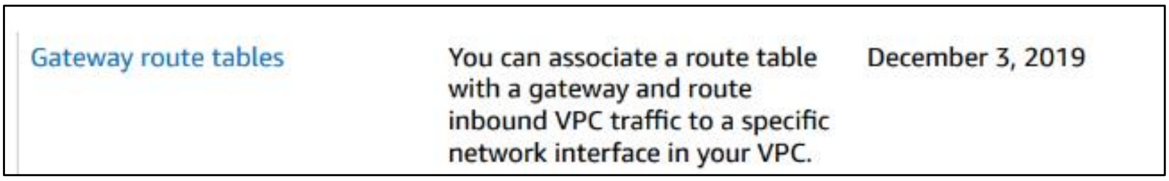
> When you change which table is the Main Route Table, it also changes the default for additional new subnets, or for any subnets that are not explicitly associated with any other route table.

VPC Guide

We can think of the Main Route Table as a special kind of route table. There can only be one per VPC, after all. Another special kind of route table is the Gateway Route Table.

Gateway Route Table

If we look at the document history, at the bottom of the VPC user guide, we can see that Gateway Route Tables became possible in December 2019:



Gateway route tables	You can associate a route table with a gateway and route inbound VPC traffic to a specific network interface in your VPC.	December 3, 2019

A Gateway Route Table is nothing but a route table which is associated with a gateway, such as an Internet Gateway or a Virtual Private Gateway (as opposed to, say, a subnet). AWS tell us that:

> You can create a **gateway route table** for fine-grain control over the routing path of traffic entering your VPC.
>
> For example, you can intercept the traffic that enters your VPC through an internet gateway by redirecting that traffic to a middlebox appliance (such as a security appliance) in your VPC.

So, it is clear that by attaching a route table to the IGW, we can direct traffic to an instance to be inspected. For some context, here is Sebastian Stormacq writing in December 2019:

> When I was delivering the Architecting on AWS class, customers often asked me how to configure an Amazon Virtual Private Cloud to enforce the same network security policies in the cloud as they have on-premises. For example, to scan all ingress traffic with an Intrusion Detection System (IDS) appliance or to use the same firewall in the cloud as on-premises. Until today, the only answer I could provide was to route all traffic back from their VPC to an on-premises appliance or firewall in order to inspect the traffic with their usual networking gear before routing it back to the cloud. This is obviously not an ideal configuration, it adds latency and complexity.
>
> Today, we announce new VPC networking routing primitives to allow to route all incoming and outgoing traffic to/from an Internet Gateway (IGW) or Virtual Private Gateway (VGW) to a specific EC2 instance's Elastic Network Interface. It means you can now configure your Amazon VPC to send all traffic to an EC2 instance before the traffic reaches your business workloads. The instance typically runs network security tools to inspect or to block suspicious network traffic (such as IDS/IPS or Firewall) or to perform any other network traffic inspection before relaying the traffic to other EC2 instances.
>
> Stormacq 2019

There are some constraints here. AWS list three things which would prevent us from associating a route table with a gateway:

1. The route table contains existing routes with targets other than a network interface, Gateway Load Balancer endpoint, or the default local route.

2. The route table contains existing routes to CIDR blocks outside of the ranges in your VPC.

3. Route propagation is enabled for the route table.

So, this advancement was announced in 2019, which made it easier to route traffic to an appliance. Then, two years later (in 2021) another advancement was announced, which *also* aimed to make it easier to use appliances.

While the 2019 announcement was about route tables attached to gateways, which receive traffic from *outside* the VPC, the 2021 announcement allowed you to route traffic between subnets *within* your VPC. This is achieved by what AWS decided to call, in 2021: **routing enhancements**.

Sébastien Stormacq, writing in 2021, provides some context:

> Since December 2019, Amazon Virtual Private Cloud (Amazon VPC) has allowed you to route all ingress traffic (also known as north – south traffic) to a specific network interface. You might use this capability for a number of reasons. For example, to inspect incoming traffic using an intrusion detection system (IDS) appliance or to route ingress traffic to a firewall.
>
> Since we launched this feature, many of you asked us to provide a similar capability to analyze traffic flowing from one subnet to another inside your VPC, also known as east – west traffic. Until today, it was not possible because a route in a routing table cannot be more specific than the default local route (check the VPC documentation for more details). In plain English, it means that no route can have a destination using a smaller CIDR range than the default local route (which is the CIDR range of the whole VPC). For example, when the VPC range is `10.0.0/16` and a subnet has `10.0.1.0/24`, a route to `10.0.1.0/24` is more specific than a route to `10.0.0/16`.

Stormacq 2021

The local route

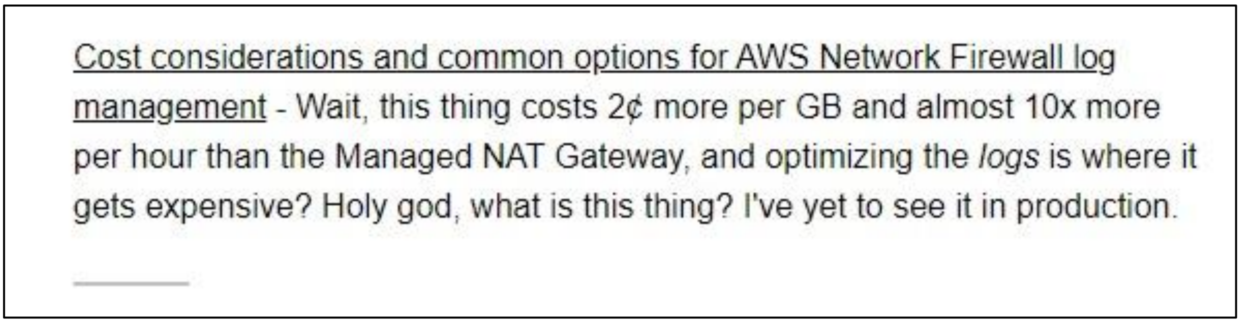
So, now, in 2022, the VPC documentation has this to say:

> You **can** add a route to your route tables that is more specific than the local route.
>
> The destination must match the entire IPv4 or IPv6 CIDR block of a subnet in your VPC.
>
> The target must be a NAT Gateway, network interface, or Gateway Load Balancer endpoint.

VPC documentation

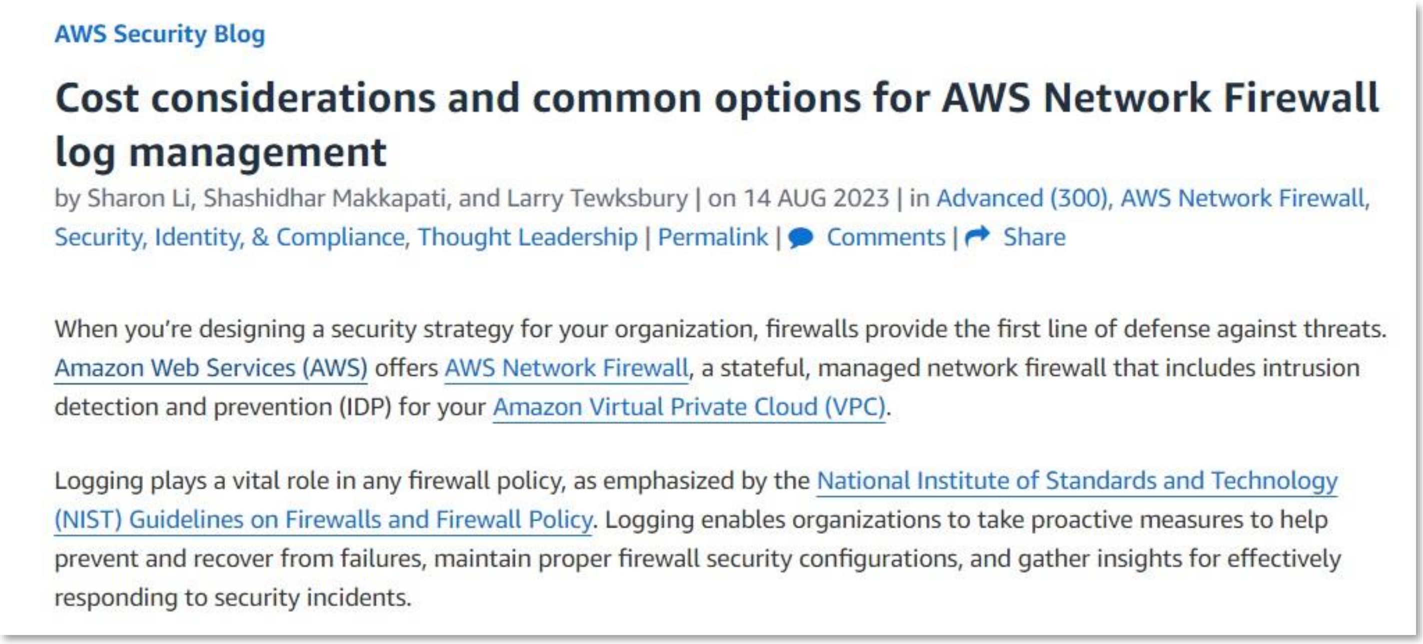
END OF RECAP

AWS Network Firewall



Cost considerations and common options for AWS Network Firewall log management - Wait, this thing costs 2¢ more per GB and almost 10x more per hour than the Managed NAT Gateway, and optimizing the *logs* is where it gets expensive? Holy god, what is this thing? I've yet to see it in production.

Email from Corey Quinn on August 21st 2023



AWS Security Blog

Cost considerations and common options for AWS Network Firewall log management

by Sharon Li, Shashidhar Makkapati, and Larry Tewksbury | on 14 AUG 2023 | in Advanced (300), AWS Network Firewall, Security, Identity, & Compliance, Thought Leadership | Permalink | Comments | Share

When you're designing a security strategy for your organization, firewalls provide the first line of defense against threats. Amazon Web Services (AWS) offers AWS Network Firewall, a stateful, managed network firewall that includes intrusion detection and prevention (IDP) for your Amazon Virtual Private Cloud (VPC).

Logging plays a vital role in any firewall policy, as emphasized by the National Institute of Standards and Technology (NIST) Guidelines on Firewalls and Firewall Policy. Logging enables organizations to take proactive measures to help prevent and recover from failures, maintain proper firewall security configurations, and gather insights for effectively responding to security incidents.

Why is this blogpost interesting? It is interesting because it refers to a document produced by NIST. Specifically this is their Guidelines on Firewalls and Firewall Policies. Before this blogpost, I was not aware that logging was important when it came to firewalls. All discussions have focussed on the rules within firewalls. Admittedly, I have come across some exam questions on sending logs from WAF to an S3 bucket. This question seemed to mention Kinesis Firehose (either as a correct or incorrect option).

The blogpost discusses three architectural patterns for logging. Stateless rules do not support logging – this we are told from the offset. Now, why might this be? Why can we not have logging for stateless rules? Stateless rules are uni-directional rules. They do not remember that a request has been made and allow the corresponding response (or vice versa). Perhaps it is the fact that with stateful rules, we are already taking in facts about the traffic. To log something is to take in information. So, both ideas (stateful rules and logging) involve the same idea.

Network Firewall supports three log destinations. They are S3, Kinesis Firehose and CloudWatch Logs. This blogpost is excellent because it goes into some of the higher-level issues. That is, it considers some of the trade-offs which are relevant to logging solutions. The article assumes that you generate 15 terabytes of data in us-east-1 per month,

An important thing to note is this term "vended log delivery". It has a significance for costs:

> CloudWatch vended log pricing can influence overall costs significantly in this pattern, depending on the amount of logs generated by Network Firewall, so it's recommended that your team be aware of the charges described in Amazon CloudWatch Pricing – Amazon Web Services (AWS). From the CloudWatch pricing page, navigate to Paid Tier, choose the Logs tab, select your Region and then under Vended Logs, see the information for Delivery to S3.

What does this mean? We are sending logs to S3 (not to CloudWatch Logs), so why are they talking about CloudWatch (vended) logs)? This is puzzling to the reader.

I find the following statement interesting:

> In our example, 15 TB is converted and compressed to approximately 380 GB in the S3 bucket.

How are they allowed to make this statement? Even if we grant it, the result is staggering. We are told:

> The total monthly cost in the us-east-1 Region is approximately $3800.

This summary table is helpful:

Pattern	Log delivery and storage cost as a multiple of the baseline cost	Functionalities	Dependencies
Amazon S3, Athena, QuickSight	1	The most economical option for log analysis.	The solution requires security engineers to have a good analytics skillset. Familiarity with Athena query and query running time will impact the incident response time and the cost.
Amazon CloudWatch	1.8	Log analysis, dashboards, and reporting can be implemented from the CloudWatch console. No additional service is needed.	The solution requires security engineers to be comfortable with CloudWatch Logs Insights query syntax. The CloudWatch Logs Insights query will impact the incident response time and the cost.
Amazon Kinesis Data Firehose, OpenSearch	2.7+	Investigate, detect, analyze, and respond to security threats quickly with OpenSearch.	The solution requires you to invest in managing the OpenSearch cluster.

Bibliography

I. Official

[AWS 1]

Solving with AWS Solutions: Network Firewall. Oct 6th 2021. YouTube Channel: Amazon Web Services. Available at: <https://www.youtube.com/watch?v=prPmk35v9ps&t=204s&ab_channel=AmazonWebServices>

[Baer 2022]

Baer, Aaron and Mike McGinnis and Dave Desroches (2022). AWS Re:Inforce 2022 [Conference]. YouTube Channel: AWS Events. Available at: https://www.youtube.com/watch?v=VMVeTvX4OLw&ab_channel=AWSEvents

[Yun 2020]

Yun, Channie (2020). AWS Network Firewall – New Managed Firewall Service in VPC. *AWS News Blog*. Nov 17th 2020. Available at: https://aws.amazon.com/blogs/aws/aws-network-firewall-new-managed-firewall-service-in-vpc/

[Tomic 2021]

Tomic, Alex and Cameron Worrell (2021). Automatically block suspicious traffic with AWS Network Firewall and Amazon GuardDuty. *AWS Security Blog*. March 16th 2021. < https://aws.amazon.com/blogs/security/automatically-block-suspicious-traffic-with-aws-network-firewall-and-amazon-guardduty/>

[Desaulniers 2021]

Desaulniers, Joel and Maritza Mills (2021). Hands-on walkthrough of the AWS Network Firewall flexible rules engine – Part 1. *AWS Security Blog*. April 27th 2021.

< https://aws.amazon.com/blogs/security/hands-on-walkthrough-of-the-aws-network-firewall-flexible-rules-engine/>

[Duffy 2021]

Duffy, Patrick (2021). Protect your remote workforce by using a managed DNS firewall and network firewall. *AWS Security Blog*. Sep 13th 2021.
< https://aws.amazon.com/blogs/security/protect-your-remote-workforce-by-using-a-managed-dns-firewall-and-network-firewall/>

[Vaidyanathan 2021]

Vaidyanathan, Shiva and Brian Lazear and Lakshmikanth Pandre (2021). Hands-on walkthrough of the AWS Network Firewall flexible rules engine – Part 2. *AWS Security Blog*. Nov 17th 2021.
< https://aws.amazon.com/blogs/security/hands-on-walkthrough-of-the-aws-network-firewall-flexible-rules-engine-part-2/>

[Jones 2022]

Jones, Marshall and Jesse Lepich and Syed Shareef (2022). Using AWS security services to protect against, detect, and respond to the Log4j vulnerability. *AWS Security Blog*. Jan 7th 2022. Available at: https://aws.amazon.com/blogs/security/using-aws-security-services-to-protect-against-detect-and-respond-to-the-log4j-vulnerability/

[Puthiyavettle 2022]

Puthiyavettle, Ajit (2022). How to deploy AWS Network Firewall to help protect your network from malware. *AWS Security Blog*. Jan 27th 2022. Available at: https://aws.amazon.com/blogs/security/how-to-deploy-aws-network-firewall-to-help-protect-your-network-from-malware/

[Taggart 2022]

Taggart, Marta and Mark Ryland (2022). Top 2021 AWS service launches security professionals should review – Part 2. *AWS Security Blog*. Available at: https://aws.amazon.com/blogs/security/top-2021-aws-service-launches-security-professionals-should-review-part-2/

[Gaddamanugu 2022]

Gaddamanugu, Harith (2022). How to deploy AWS Network Firewall by using AWS Firewall Manager. *AWS Security Blog*. Aug 26th 2022. Available at: https://aws.amazon.com/blogs/security/how-to-deploy-aws-network-firewall-by-using-aws-firewall-manager/

[Ahmad 2020]

Ahmad, Shakeel and Evgeny Vaganov (2020). Deployment models for AWS Network Firewall. *Networking & Content Delivery* [Blog] Nov 17th 2020. Available at: https://aws.amazon.com/blogs/networking-and-content-delivery/deployment-models-for-aws-network-firewall/

[Ahmad 2021]

Ahmad, Shakeel and Evgeny Vaganov (2021). Deployment models for AWS Network Firewall with VPC routing enhancements. *Networking and Content Delivery*. Sep 10th 2021. Available at: https://aws.amazon.com/blogs/networking-and-content-delivery/deployment-models-for-aws-network-firewall-with-vpc-routing-enhancements/

[Applebaum 2022]

Applebaum, Tyler and Akshay Karanth (2022). Introducing Prefix Lists in AWS Network Firewall Stateful Rule Groups. *Networking & Content Delivery* [Blog] Sept 30th 2022. Available at: https://aws.amazon.com/blogs/networking-and-content-delivery/introducing-prefix-lists-in-aws-network-firewall-stateful-rule-groups/

[Trunnel 2021]

Trunnel, Benjamin (2021). Deploy centralized traffic filtering using AWS Network Firewall. *Networking & Content Delivery* [Blog] June 15th 2021. Available at: https://aws.amazon.com/blogs/networking-and-content-delivery/deploy-centralized-traffic-filtering-using-aws-network-firewall/

[Vaidyanathan 2021]

Vaidyanathan, Shiva and Rashpal Kler (2021). Migrating from Squid Web Proxy to AWS Network Firewall. *Networking & Content Delivery* [Blog] Aug 23rd 2021. Available at: https://aws.amazon.com/blogs/networking-and-content-delivery/migrating-from-squid-web-proxy-to-aws-network-firewall/

[Kiem 2021]

https://aws.amazon.com/blogs/architecture/integrate-aws-network-firewall-with-your-isv-firewall-rulesets/

[Mills 2021]

Mills, Maritza (2021). Introducing AWS Network Firewall. [Online presentation]. YouTube Channel: AWS Events. Available at: https://www.youtube.com/watch?v=CISgqpVn75Q&t=199s&ab_channel=AWSEvents

[Vaganov 2021]

Vaganov, Evgeny (2021). AWS Network Firewall and network security in Amazon VPC. AWS Summit Online ASEAN. Available at: https://www.youtube.com/watch?v=J6ewybE4yP8&ab_channel=AWSEvents

[Vaganov 2022]

Vaganov, Evgeny (2022). AWS Network Firewall and network security in Amazon VPC. Available at: https://www.youtube.com/watch?v=WNFknf9zyZg&ab_channel=AWSEvents

[Woodard 2021]

AWS Network Firewall. YouTube Channel: Angelbeat Seminars. May 8th 2021. Available at: <https://www.youtube.com/watch?v=yJ5bYwS6kvY&t=496s&ab_channel=AngelbeatSeminars>

[Lehwess 2021]

Lehwess, Matt and Alexandre Huides. Available at: https://www.youtube.com/watch?v=fi3vcenH6UY&ab_channel=AWSEvents

https://aws.amazon.com/blogs/networking-and-content-delivery/vpc-sharing-key-considerations-and-best-practices/

II. Unofficial

[Hocking 2021]

AWS Network Firewall for Ingress/Egress. YouTube Channel: Phillip Hocking. Available at: https://www.youtube.com/watch?v=-cpkB-QZlTg&ab_channel=PhillipHocking

[Tutorial 2021 a]

How to build an AWS Network Firewall Environment. YouTube Channel: AWS Tutorials – Giuseppe Borgese. Jan 19th 2021. Available at: <https://www.youtube.com/watch?v=Xb-matrBNOs&ab_channel=AWSTutorials-GiuseppeBorgese>

[Tutorial 2021 b]

AWS Network Firewall with Demo. March 29th 2021. YouTube Channel: CloudDeepDive. Available at: <https://www.youtube.com/watch?v=V6bRE6ggYaU&t=1022s&ab_channel=CloudDeepDive>

[Tutorial 2021 c]

AWS Network Firewall. YouTube Channel: Cloud Ops. Oct 1st 2021. Available at: <https://www.youtube.com/watch?v=RXp1jteof2Y&ab_channel=CloudOps>

[Tutorial 2021 d]

Fundamentals of AWS Network Firewall. Oct 2nd 2021. YouTube Channel: Public Cloud Tips and Tricks. Available at: <https://www.youtube.com/watch?v=aef2IpRVJM0&ab_channel=PublicCloudDesignTipsandTricks>

[Tamil 2022 a]

AWS Network Firewall **Theory**. YouTube Channel: Tamil Cloud. Jan 1st 2022. Available at**:** <https://www.youtube.com/watch?v=nh6nNO4D0-4&ab_channel=TamilCloud>

[Tamil 2022 b]

AWS Network Firewall Architecture. YouTube Channel: Tamil Cloud. Jan 1st 2022. Available at: <https://www.youtube.com/watch?v=sRyspJw4NDE&ab_channel=TamilCloud>

[Tamil 2022 c]

AWS Network Firewall **Walkthrough**. YouTube Channel: Tamil Cloud. Jan 1st 2022. Available at: https://www.youtube.com/watch?v=2gXdq3S2N4E&ab_channel=TamilCloud

[Musonza 2020]

AWS Network Firewall. YouTube Channel: Tendai Musonza. Available at: https://www.youtube.com/watch?v=r6mEnUTTkMQ&ab_channel=TendaiMusonza

[Duvall 2021]

AWS in 5 minutes: AWS Network Firewall. March 26th 2021. YouTube Channel: Paul Duvall. Available at: https://www.youtube.com/watch?v=uz4OZqRGCgM&ab_channel=PaulDuvall

[Fortinet 1]
Feb 12th 2021. Available at: <https://www.youtube.com/watch?v=1CL69tWkfEc&ab_channel=Fortinet>

[Wikipedia 1]
https://en.wikipedia.org/wiki/Meerkat#Interactions_with_humans

III. Critical

[Surname1]
Smith, David (year). Title of Work Here. 1st Jan 2022. City: Publisher. Available at:
<URL here>.

IV. General

[Surname1]
Smith, David (year). Title of Work Here. 1st Jan 2022. City: Publisher. Available at:
<URL here>.

[Cooke 2020]
Cooke, Alastair and Jeffrey Powers. AWS VPC Networking Basics. YouTube Channel: Build Day Live.
Available at: https://www.youtube.com/watch?v=UhqJY0P27uI&ab_channel=BuildDayLive

[Leblond 2021]
Leblond, Eric (2021). June 7th 2021. Suricata: The First 12 Years of Innovation. Available at:
<https://www.stamus-networks.com/blog/suricata-the-first-12-years-of-innovation?hsLang=en>

[Manev 2021]
Manev, Peter (2021). The Other Side of Suricata. Available at: <https://www.stamus-networks.com/blog/the-other-side-of-suricata#:~:text=Suricata%20started%20as%20an%20open,continuously%20evolved%20to%20stay%20relevant>

[Overflow 2021]
AWS Global Services. *Stack Overflow* [Forum]. Question asked Aug 17th 2021. Available at:
https://stackoverflow.com/questions/68811957/aws-global-services

[Overflow 2016]

Are S3 buckets region specific? Stack Overflow [Forum]. Question asked Apr 20th 2016. Available at:
https://stackoverflow.com/questions/36754094/are-s3-buckets-region-specific

[Overflow 2020]

Why can't a subnet span availability zones in AWS? Stack Overflow [Forum].Question asked: June 8th 2020. Available at: https://stackoverflow.com/questions/62256096/why-cant-a-subnet-span-availability-zones-in-aws

[Exchange 2019]

Route propagation from virtual private gateway in AWS. ServerFault [Forum] Question asked May 26th 2019. Available at: https://serverfault.com/questions/968867/route-propagation-from-virtual-private-gateway-in-aws

[Wikipedia 2]

"Static Routing". Available at: https://en.wikipedia.org/wiki/Static_routing

MWAA

MWAA stands for Managed Workflows for Apache Airflow.

What on earth is Apache?

Bibliography

AWS fixes 1-click Apache Airflow session hijack flaw. Available at: https://www.scmagazine.com/news/aws-fixes-1-click-apache-airflow-session-hijack-flaw?ck_subscriber_id=1560524742

Monitron

Lookout for Equipment

Panorama ←

Proton ←

HealthLake

Lake Powell, Utah

Ten Facts about Amazon HealthLake

1. It was announced by Matt Wood

Many services are announced by the CEO of AWS, but HealthLake was announced by Dr Matt Wood at Re:Invent. It was 8th Dec 2020—the second UK vaccine was administered, to a man called William Shakespeare, from Warwickshire.

Wood initially trained in medicine 60 miles further north, in Nottingham. He explained to a virtual audience that HealthLake "transforms data seamlessly, to automatically understand and extract meaningful medical information from raw, disparate data such as prescriptions, procedures, and diagnoses".

You might be wondering how a service can be designed specifically for *medical* information. Information is just information, after all. I will respond to this over the course of the article.

2. Data is stored in the FHIR format

FHIR (pronounced "fire") is a standard that is widely used in the medical industry. FHIR stands for Fast Healthcare Interoperability Resource. The non-profit responsible for FHIR has an office in the Great Lake State (Michigan). It's opposite a Denny's, in Ann Arbor.

Known as Health Level Seven International (HL7), they helped to standardize information transferred within and between hospitals. HL7 refers to the protocol developed by the organisation's founders. They were academics at the University of California, San Francisco (UCSF), sick of inefficiencies in the late 1970s.

The *medical record* is an interesting entity. As Spronk notes, it's "an assemblage of information from various sources: the clinical lab, the radiology department, the consultant, the nurse". Hospitals had mainframes quite early on, but no backend.

Clement McDonald, one of the designers of the HL7 protocol, argued in 1990 that the medical industry could learn from supermarkets. They'd recently implemented barcodes. He wrote "we need similar labels for our informational goods to facilitate future use of computerized record systems in medicine". His editorial, calling for standardized interfaces, was essentially mocked, one reviewer noting "standards would be of no use because physicians did not have computers in their office and never would" (See Spronk).

L7 *is* a reference to the OSI model, but they didn't know about OSI at the time, as Don Simborg notes:

> This was the first Level 7 protocol ever used in healthcare to my knowledge.
>
> It was only later that I learned about the OSI model and realized that what we were doing was at its level 7. We continued to add computers to the UCSF network and refined the protocol over the years."
>
> [Spronk]

Amazon HealthLake stores data in the FHIR format.

3. It's not AWS Lake Formation

Do not confuse AWS HealthLake with other services. There are two ways to mess up here: thinking of other services containing *lake* and others containing *health*.

When Wood announced HealthLake in Dec 2020, Lake Formation was two years old. It allows you to create *data lakes*—repositories of data kept in a more 'raw' state than usual. James Dixon coined the term *data lake* to stress that the stored data is still imbued with potentiality. It hasn't necessarily adopted a particular form yet. Dixon contrasts water (representing data) *in a lake* with water *in a bottle*. The latter has been heavily processed.

Lake Formation (pronounced Lake **pause** Formation). does a lot. You can ingest data from various sources, clean it using ML algorithms, and even control access to particular rows and columns. You can then use analytics and query tools. For example, Redshift, Athena, and QuickSight can be used (rack your brain, RAQ). Furthermore, "Lake Formation builds on the capabilities available in AWS Glue." says the page currently.

There was also a service called "AWS Personal Health Dashboard". Currently, it's styled as just the AWS Health Dashboard. It was in the "Management and Governance" category. First mentioned on 1st December 2016—the day a spacecraft resupplying the ISS was lost 382 seconds after launch—this was a dashboard that showed notifications related to the services you were using. AWS could tell you when they were tinkering with (updating, retiring) infrastructure.

You could even have notifications "delivered to you via email or mobile for quick visibility". Ram Atur explains how the *Service* Health Dashboard had been around since 2008. It was a page that listed all the services, saying "operating normally" with green ticks. Then things got personal in 2016. Oh, and as Indiana Jones has to say when asked if there's a doctor onboard, the *PHD is unrelated to medicine*.

The tendency to associate 'health' with the upkeep of humans (medicine) rather than the state of computers has been noted. Corey Quinn wrote that the name will 'confuse the living hell out of any healthcare AWS customers'.

The word "healthy" is a classic example of ambiguity as it is (healthy foods *produce* health, healthy people *possess* it *etc*) So, it doesn't help that the IT world uses it in a further distinct sense. Servers are not alive (a truth we'll increasingly have to *argue* is the case). So *healthy,* in IT, denotes a thing successfully functioning—a thing not displaying signs of imminent demise (akin to a *healthy economy,* or *automobile*). Arguably, it's among the metaphors that abound in IT: firewall, desktop, lake… cloud. Still, it could have been called the Status Dashboard.

Quinn was in fact commenting on the new "Health Dashboard". Announced on 28th Feb 2022, this is intended to consolidate the Service- and personal- health dashboard. Scientists predict the next dashboard will land in 2026. The *dashboard* is often dropped, producing "AWS Health". Do not confuse it with "AWS *for* Health", which *does* concern healthcare. This was announced in July 2022.

Use an integral upper-case. It's Health*L*ake, similar to GuardDuty, CloudWatch, WorkSpaces, QuickSight and SageMaker. (Ignore Lake Formation here. And Auto Scaling, Control Tower, Cloud Map). Because we're trying to be… what's the word, hop?

4. It's compliant with HIPAA

HIPAA (pronounced HIP-uh) is a law in the USA, created in the third year of Bill Clinton's presidency. If you're storing medical data, you must comply with HIPAA. A nice summary is provided by the National Cancer Institute.

You might not know that it's also known as the Kassebaum-Kennedy bill, named after the two Senators that published it. The Edward Kennedy in question was the younger brother of JFK. Republican Nancy Kassebaum (now 90 years old) worked for the family business, which operated several radio stations, before entering politics. Reagan chose Kassebaum to organise the 1980 conference, held 60 miles to the east of the Denny's. Her appointment was seen as a 'nod to the more liberal in the party' (say Wikipedia).

HIPAA doesn't just deal with the electronic transfer of medical information. It was quite a dramatic reform to the way citizens of the USA maintained their "health coverage" (The USA has no NHS.) It was perhaps the biggest change since the Medicare and Medicaid programmes were introduced in 1965. See an early, incomplete version of the bill here; a more complete version here.

HIPAA "provides tax incentives to purchase long-term care insurance" (Atchinson and Fox 1997). Let me make things clear, especially for UK readers (we have the NHS). HIPAA gave rights to self-employed people and small employers. We're talking about the *right to be offered health insurance* (by the health insurer). And HIPAA said this right to full coverage is *portable*: it moves with you as you move to a new job, or as your employer moves to a new health plan. HIPAA stands for Health Insurance *Portability* and Accountability Act.

What does the act have to do with "accountability"? The whole law has 5 parts (or *Titles*). It's in Title 2 that we find various rules about how personal health information must be handled. Its rules about keeping data private have come to be called **"the HIPAA Privacy Rule"**. If you are a healthcare provider (or any sort of **Covered Entity**) you must abide by these rules, not disclosing personal health information (PHI). Examples of Covered Entities are hospitals and care homes. There is also **the Security Rule** which

> complements the Privacy Rule. While the Privacy Rule pertains to all Protected Health Information, the Security Rule is limited to Electronic Protected Health Information. It lays out 3 types of security safeguards: administrative, physical, and technical.
>
> Edemekong 2022

So, if you manage electronic PHI, you need to be thinking about what you must do to meet those three safeguards (some further research will be necessary).

HIPAA functions to single out certain individuals, who work in Covered Entities, and states that if these individuals do not protect PHI in the appropriate way, then they should be sanctioned. In this way, HIPAA makes certain individuals accountable. It's the HI Portability *and Accountability* Act.

Sybex, who, *still today,* are taking students' cash, promising to prepare them for the Comptia Cloud+ exam, (and who print the name of a copy editor in the front) repeatedly write HIPPA in this textbook. If I took a shot every time someone wrote HIPPA, you'd need to alter my medical records. Now that we understand that it deals with *portability* as well as *accountability*, in one *act* (law), it is clear that it simply must be HI – PA – A.

Amazon HealthLake is compliant with HIPAA.

5. It's been used in Chicago hospitals

We've visited Michigan twice already (HL7 has an office in Ann Arbor; the 1980 Rep conference was in Detroit). Now we're flying West, into the US state of Illinois, where we're going to read about an early adopter of Amazon HealthLake. (*Chicago*. Here is Obama in Chicago, precisely 3 weeks before RedShift is announced. He'd just secured another four-year term, to begin Jan 2013.).

The Rush Medical Center has a long history. It goes back to Rush Medical College, established in 1837 (the year Michigan is established, and Queen Vic gets on the throne.) The namesake is Benjamin Rush, who held views of his time. At one point, he tried his hand at classifying mental illnesses (you know, creating a taxonomy.)

The medical centre is associated with Rush University. A teaching hospital (or… several), it often makes it onto the US News list of the best 20 hospitals in the whole of the USA. It's a leader in orthopaedics. The Washington Post filmed inside the hospital's "command centre" activated for COVID-19.

On 30th April 2021, (about four months after Wood's announcement) a guest blog post appears on the AWS website:

> Rush University Medical Center creates COVID-19 analytics hub on AWS

The article suggests that the hub is used for hospitals within Chicago which are not part of the Rush University Medical Centre. "Seventeen of the 32 hospitals in Chicago are currently submitting data, and Rush plans to integrate all 32 hospitals by summer 2021." The post explains how the solution parses things called CCDs (Continuity of Care Documents), which can be imported. Other HL7 data is also imported. The solution involves converting it to FHIR Release 4. It is stated that Amazon QuickSight—a business analytics services—is used to visualize which ICU beds are occupied.

We get some other nice details. It uses 'representational state transfer (REST) application programming interfaces (APIs) for create, read, update, and delete functions.' We are told that 'Amazon HealthLake currently provides search capabilities on 71 FHIR resource types for its data stores' (July 2021).

6. Compute resources? Consider it done.

7. HealthLake uses ontology mapping

In his announcement, Matt Wood tells us that Health Lake allows you to "bring together all of this data in minutes, with natural language understanding, ontology mapping, and medical comprehension". What on earth is ontology mapping?

The fancy word *ontology* can take a while to get used to. It is the study of what exists. For example, when the Scottish philosopher David Hume said that there are no causes, merely one thing following another, he was making an ontological claim. When Margaret Thatcher said "there is no such thing as society", this was an ontological claim. Philosophers who study what exists are ontologists (not onco-logists, which are doctors who study cancer).

But information scientists are also concerned with ontology. Or really, specific *ontologies*. "The philosophical term 'ontology' has been adopted by the information-science community to refer to an automated representation (taxonomy, controlled vocabulary) of a given domain" Katherine Munn (2008) writes.

Welcome to the world of *applied ontology*. Stored data needs to settle on certain categories of things. Barry Smith, one of leading writers on this field tells us that "an ontology can be seen, roughly, as a taxonomy of entities in a given domain, complete with formal rules that govern the taxonomy" (Smith 2008:22). Put simply, an ontology is a terminology or controlled vocabulary.

(Apparently an ontology of Barry Smith's is needed… This one was supervised at Manchester by Wolfe Mays, who proudly tracked down the blueprints of a jet engine designed by his friend Wittgenstein, and who pondered machines simulating a mental activity, with colleague 'Alan').

So, there's a need to develop ontologies in the domain of healthcare. We had to classify diseases, prescriptions, and drugs. This area—part of medical informatics—is essential if our computers are to make inferences useful to the clinician. They're certainly necessary for AI. If HealthLake has a medical ontology built in, that's an immense advantage.

For further reading, consider the introduction to Martin Berzell's thesis "Electronic Healthcare Ontologies"; Barry's Smith's *Applied Ontology: An Introduction* (2008); and Smith's YouTube playlist on Biomedical Ontology; see his critique of HL7.

8. HealthLake uses machine learning

A blog post on HealthLake which I must recommend is Julien Simon's (2021). I have read Simon's work before, when I looked at SageMaker, the AWS service that let's you build and run machine learning models. HealthLake is an application of machine learning

> As data is uploaded, HealthLake uses integrated natural language processing to extract entities present in your documents and stores the corresponding metadata. These entities include anatomy, medical conditions, medication, protected health information, test, treatments, and procedures. They are also matched to industry-standard ICD-10-CM and RxNorm entities.

ICD stands for "International Classification of Diseases". CM stands for "Clinical Modification". Apologies if you're experiencing an acronym attack.

ICD-10-CM is a medical classification list devised by the World Health Organisation (WHO). It's a modification of ICD-10, specially designed for the USA. The USA tends to be quite slow to adopt newer versions of ICD. The key point is that Covered Entities (of HIPAA fame) *must* use ICD-10 codes [source]. The code for being struck by a turtle is W5922XD. There are more fun ones here.

Julien Simon tells us how HealthLake uses natural language processing. Natural language processing usually refers to the ability of machines to "comprehend" human language (as opposed to code and machine language). It's quite difficult to get computers to appreciate all the nuances of language. HealthLake can:

> extract entities present in your documents and stores the corresponding metadata.

Examples of these "entities" include anatomy, medical conditions, medication, protected health information, test, treatments, and procedures.

9. Kass Hout helped to design it

One of 'the scientists behind the new service' is Taha Kass-Hout, according to this Amazon Science article. Kass-Hout joined AWS in 2017 but previously worked for the FDA (Food and Drug Administration).

In his 2015 State of the Union Address, Barack Obama announced the Precision Medicine Initiative.

Before that, he worked for the US Centres for Disease Control and Prevention (CDC). Kass-Hout emphasises that FHIR is helpful because it standardizes *structured* data. The problem, however, is that a lot of data related to a patient's health is unstructured. It is

> clinical notes, PDF laboratory reports, insurance claims, X-ray and MRI images, recorded conversations, heart ECG or brain EEG traces, and more

Kass-Hout 2020

Amazon HealthLake helps with this. It ingests the data and then manipulates it in quite an involved way. It then does three things. (1) It normalizes the information. (2) It tags any dates. (3) It tags any key descriptions of events. Such key descriptions might be procedures, medications, diagnoses etc. It then creates an index. That is really important, because it enables the data to be *searched*. This cannot be underestimated: it's not enough to have data, it must be searchable. Whole industries are built upon enabling organisations to search what they have (see Amazon Kendra).

10. It's part of something big

Google Cloud have developed something called the Healthcare Data Engine. Elise Reuter writes that it:

> stores patient data in a FHIR format, a standard backed by interoperability rules.
>
> Google developed it based on its work with Mayo Clinic, which struck a 10-year partnership with Google earlier this year in hopes of creating an "AI factory."

So, the Healthcare Data Engine stores data and uses machine learning. This is something that Amazon HealthLake does.

Amazon intends to do more in healthcare. Lee (2021) writes:

> Amazon, long assumed to be a sleeping giant in healthcare, is finally awakening.

The technology behind Alexa is being used to help surgeons in operating theatres. Amazon is not *quite* the same as the other big tech companies, however. Its network of warehouses and delivery drivers are unique. These are potential medicine delivery infrastructures (as noted by Nicholas Desai). Therefore, in the next decades, we might not just see new toys for health providers, but the emergence of A Health Provider.

In September 2019, we saw that Amazon had launched Amazon Care. This was a virtual clinic, allowing a select group to speak to doctors over video and phone. (This expanding area is called telemedicine or telehealth.) It was trialled on employees of Amazon in Seattle. Home visits could also be arranged.

On 25th August 2022, it was reported that Amazon intends to close Amazon Care. It's not within the scope of this article to provide a prognosis, but certainly, spinning things down doesn't exactly signify failure at Amazon's scale.

See this

https://www.youtube.com/watch?v=wNB5gu7klUo&ab_channel=AmazonWebServices

Information retrieval: a health and biomedical perspective.

Audit Manager

←

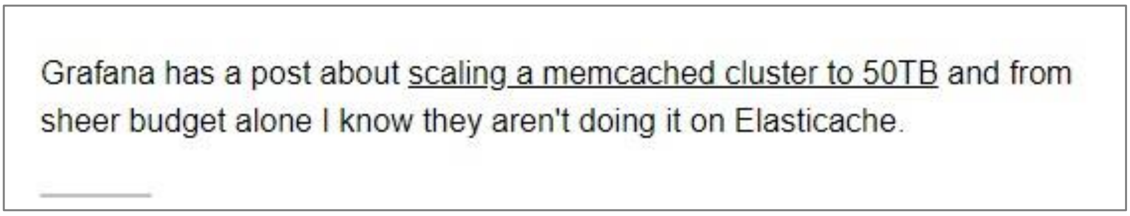
Lookout for Metrics

←

IoT EduKit

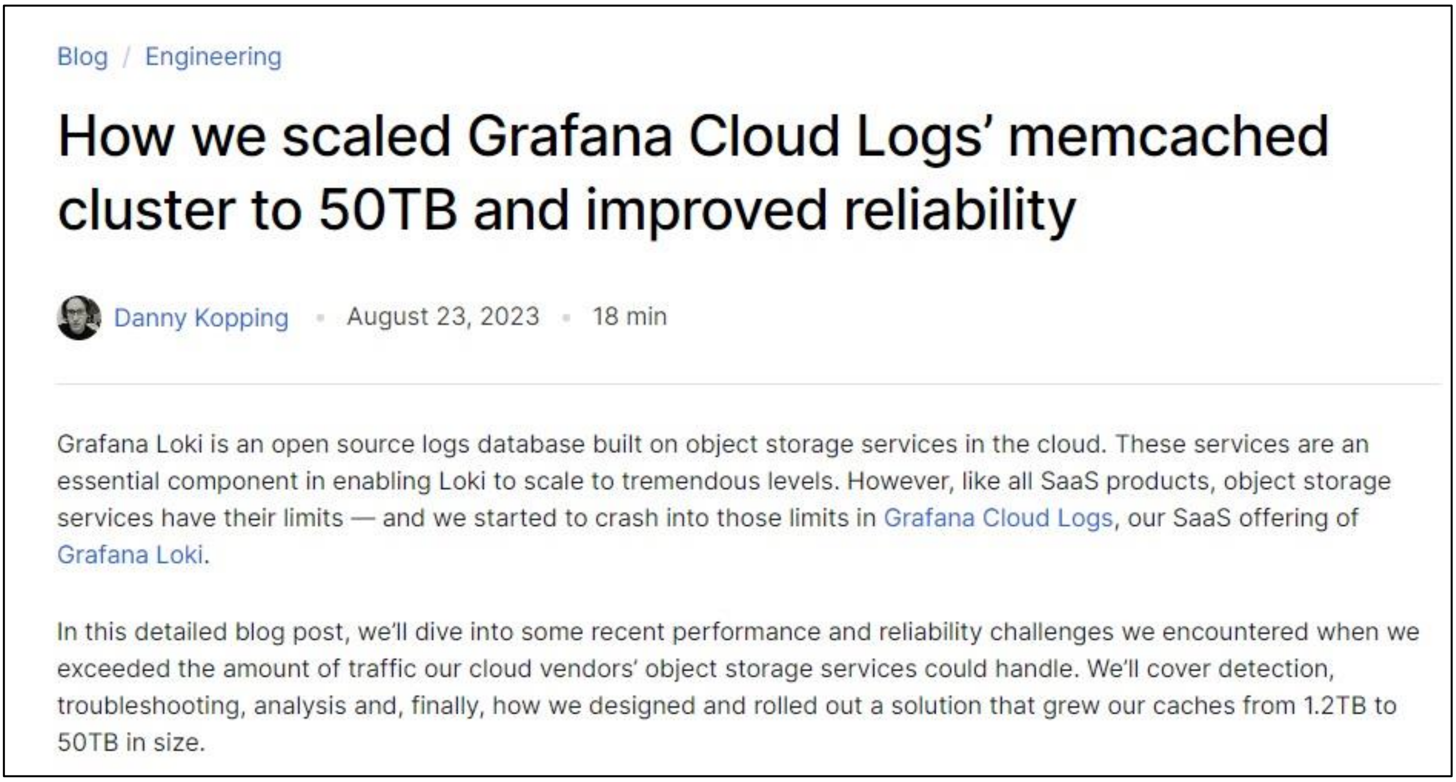
←

Managed Grafana



Grafana has a post about scaling a memcached cluster to 50TB and from sheer budget alone I know they aren't doing it on Elasticache.

Email from Corey Quinn on Aug 28th 2023



Blog / Engineering

How we scaled Grafana Cloud Logs' memcached cluster to 50TB and improved reliability

Danny Kopping • August 23, 2023 • 18 min

Grafana Loki is an open source logs database built on object storage services in the cloud. These services are an essential component in enabling Loki to scale to tremendous levels. However, like all SaaS products, object storage services have their limits — and we started to crash into those limits in Grafana Cloud Logs, our SaaS offering of Grafana Loki.

In this detailed blog post, we'll dive into some recent performance and reliability challenges we encountered when we exceeded the amount of traffic our cloud vendors' object storage services could handle. We'll cover detection, troubleshooting, analysis and, finally, how we designed and rolled out a solution that grew our caches from 1.2TB to 50TB in size.

Bibliography

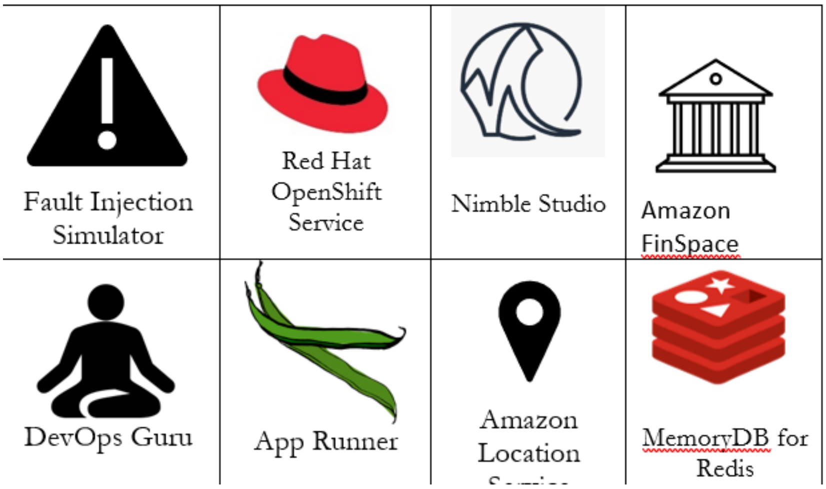
[Kopping 2023]

Kopping, Danny (2023). How we scaled Grafana Cloud Logs' memcachd cluster to 50 TB and improved reliability. *Grafana Labs* [Blog]. Available at: <https://grafana.com/blog/2023/08/23/how-we-scaled-grafana-cloud-logs-memcached-cluster-to-50tb-and-improved-reliability/?ck_subscriber_id=1560524742>

CloudShell

Managed Service for Prometheus

Fault Injection Simulator



Amazon DynamoDB now supports an AWS FIS action to pause global table replication

Posted On: Apr 30, 2024

Amazon DynamoDB now supports an AWS Fault Injection Service action to pause replication for global tables. FIS is a fully managed service for running controlled fault injection experiments to improve an application’s performance, observability, and resilience. Global tables replicate your Amazon DynamoDB tables automatically across your choice of AWS Regions to achieve fast, local read and write performance. Customers can use the new FIS action to observe how their application responds to a pause in regional replication, and tune their monitoring and recovery process to improve resiliency and application availability.

Global tables are designed to meet the needs of high availability applications, providing you 99.999% availability, increased application resiliency, and improved business continuity. This new FIS action reproduces the real-world behavior when replication to a global table replica is interrupted and resumed. This lets customers test and build confidence that their application responds as intended when resources in a Region are not accessible. Customers can create an experiment template in FIS to integrate the experiment with continuous integration and release testing and to combine with other FIS actions. For example, DynamoDB Pause Replication is combined with other actions in the Cross-Region: Connectivity scenario to isolate a Region.

DynamoDB Pause Replication is now available in all AWS commercial Regions where FIS is available. To learn more, visit the DynamoDB FIS actions documentation.

Bibliography

I. Official

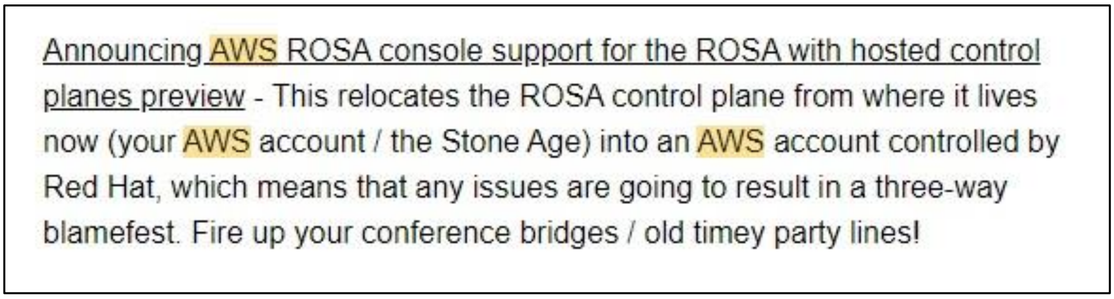
[Hornsby 2023]

Hornsby, Adrian and Iris Sheu (2023). Improve application resilience with AWS Fault Injection Service. YouTube Channel: AWS Events. *Reinvent 2023* [Conference]. Available at: <https://www.youtube.com/watch?v=N0aZZVVZiUw&ab_channel=AWSEvents>

[AWS 2024]

Amazon DynamoDB now supports an AWS FIS action to pause global table replication [Announcement]. Apr 30th 2024. Available at: <https://aws.amazon.com/about-aws/whats-new/2024/04/amazon-dynamodb-fis-action-pause-global-table-replication/>.

RedHat OpenShift Service



Announcing AWS ROSA console support for the ROSA with hosted control planes preview - This relocates the ROSA control plane from where it lives now (your AWS account / the Stone Age) into an AWS account controlled by Red Hat, which means that any issues are going to result in a three-way blamefest. Fire up your conference bridges / old timey party lines!

Corey Quinn in an email from Aug 28th 2023

Announcing AWS ROSA console support for the ROSA with hosted control planes preview

Posted On: Aug 24, 2023

In May, Red Hat announced the preview of Red Hat OpenShift Service on AWS (ROSA) with hosted control planes (HCP), a new deployment model for ROSA clusters. Today, we are introducing an AWS account configuration workflow for ROSA with HCP on the AWS Management Console. Under the original ROSA deployment model, now called ROSA classic, all AWS infrastructure required to run the ROSA control plane is hosted on your AWS account. Now, you can create ROSA clusters with the control plane hosted and managed on a service account. During the preview, ROSA with HCP clusters will not incur ROSA service fees and should not be used for production workloads.

ROSA with HCP helps reduce the AWS infrastructure cost of running ROSA clusters, as the ROSA control planes are hosted and managed by Red Hat on AWS. The new deployment model reduces the ROSA cluster create time, and you can separately schedule OpenShift version upgrades for the control plane and the worker node machine pools. By moving control plane infrastructure out of your AWS account, ROSA with hosted control planes mitigates the risk that actions taken on your account lead to a degraded ROSA control plane.

ROSA with HCP is available in preview today in the following AWS regions: US East (N. Virginia), US East (Ohio), US West (Oregon), Asia Pacific (Jakarta), Europe (Frankfurt), Europe (Ireland).

To get started with ROSA with HCP, see [Getting started with ROSA with HCP in the ROSA User Guide](). Log in with your Red Hat account to learn more about Red Hat Technology Preview Features Support Scope.

Nimble Studio ←

FinSpace ←

Bibliography

[Gopalakrishnan 2023]

Gopalakrishnan, Balaji Kumar and Pradeep Misra (2023). How Amazon Finance Technologies built an event-driven and scalable remittance service using Amazon DynamoDB. Aug 22nd 2023. Available at: <https://aws.amazon.com/blogs/database/how-amazon-finance-technologies-built-an-event-driven-and-scalable-remittance-service-using-amazon-dynamodb/?ck_subscriber_id=1560524742>

DevOps Guru

App Runner

AWS Announces AWS App Runner

Posted On: May 18, 2021

Amazon Web Services, Inc. (AWS), an Amazon.com (NASDAQ: AMZN) company, announces the general availability of AWS App Runner, a fully managed container application service that makes it easy for customers without any prior containers or infrastructure experience to build, deploy, and run containerized web applications and APIs in just a few clicks. Customers simply provide source code, a container image, or deployment pipeline, and App Runner automatically builds and deploys the web application, load balances traffic, scales on demand, and monitors application health.

Developed with AWS operational, security, and configuration best practices, App Runner eliminates the need for customers to manage infrastructure, servers, or container orchestrators. AWS maintains and manages the infrastructure running customer applications, which means that teams get to production quickly and efficiently. Out of the box, App Runner is built for web scale, so there's no need to re-platform or re-architect as the business grows. It makes it simpler for customers to rapidly deliver innovative solutions and business value. App Runner is the easiest way for customers to run their web applications (including APIs, backend web services, and websites) on AWS.

AWS App Runner is now generally available, and you can use it in the following AWS Regions: US East (N. Virginia), US East (Ohio), US West (Oregon), Europe (Ireland), and Asia Pacific (Tokyo).

You can use AWS App Runner with the AWS console, AWS CLI, or AWS Copilot. To get started, visit our product page.

Justin | @rothgar
Adam Keller | @realadamjkeller
CONTAINERS from the couch
aws
Brent Langston | @brentContained

Screenshot from [Garrison 2021]

AWS App Runner adds support to update and rebuild a failed service

Posted On: Jun 30, 2023

AWS App Runner adds support for editing and rebuilding a failed service. App Runner makes it easier for developers to quickly deploy containerized web applications and APIs to the cloud, at scale, and minimizing the need to manage infrastructure.

You may face a failure scenario while creating an App Runner service, where the service is left in CREATE_FAILED state. Until now, App Runner did not support updating configurations or rebuilding a failed service. You needed to delete the failed service and create a new service to build with necessary changes to your source code or service configuration. Starting today, you can update and rebuild a failed service with or without any changes to the source code or configuration. This saves you cycles of creation and deletion in order to get the service to a successful state.

You can update a failed service from the service configuration page in the App Runner console, or by using the UpdateService API. To learn more about App Runner service creation and troubleshooting guide for service create failures, see AWS App Runner service creation guide. To learn more about App Runner, see the AWS App Runner Developer Guide.

Bibliography

[AWS 2023]

AWS App Runner adds support to update and rebuild a failed service. June 30th 2023. Available at: <https://aws.amazon.com/about-aws/whats-new/2023/06/aws-app-runner-update-rebuild-failed-service/?utm_source=substack&utm_medium=email>

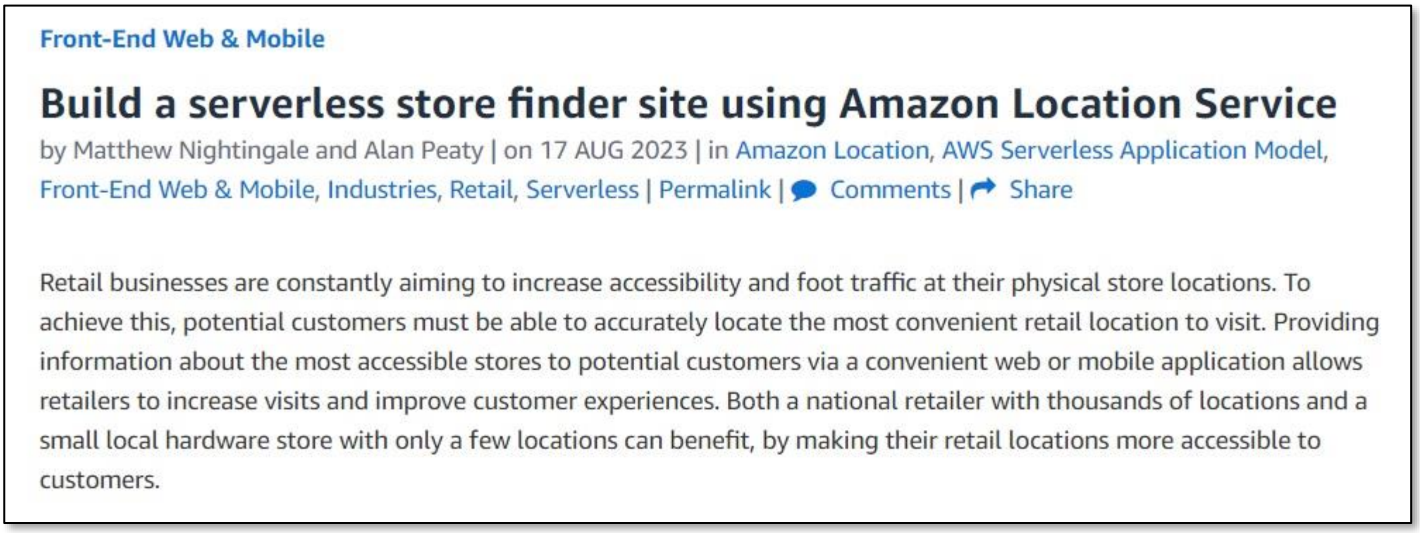
[Garrison 2021]

Garrison, Justin and Adam Keller and Brent Langston and Akshay Ram (2021). YouTube Channel: Containers from the Couch. Available at: https://www.youtube.com/watch?v=97Ua6Gv_HSo&ab_channel=ContainersfromtheCouch

[Davis 2021]

Davis, Neal (2021). YouTube Channel: Digital Cloud Training. Nov 1st 2022. Available at: https://www.youtube.com/watch?v=ycdo9UyNs98&ab_channel=DigitalCloudTraining

Amazon Location Service



Front-End Web & Mobile

Build a serverless store finder site using Amazon Location Service

by Matthew Nightingale and Alan Peaty | on 17 AUG 2023 | in Amazon Location, AWS Serverless Application Model, Front-End Web & Mobile, Industries, Retail, Serverless | Permalink | Comments | Share

Retail businesses are constantly aiming to increase accessibility and foot traffic at their physical store locations. To achieve this, potential customers must be able to accurately locate the most convenient retail location to visit. Providing information about the most accessible stores to potential customers via a convenient web or mobile application allows retailers to increase visits and improve customer experiences. Both a national retailer with thousands of locations and a small local hardware store with only a few locations can benefit, by making their retail locations more accessible to customers.

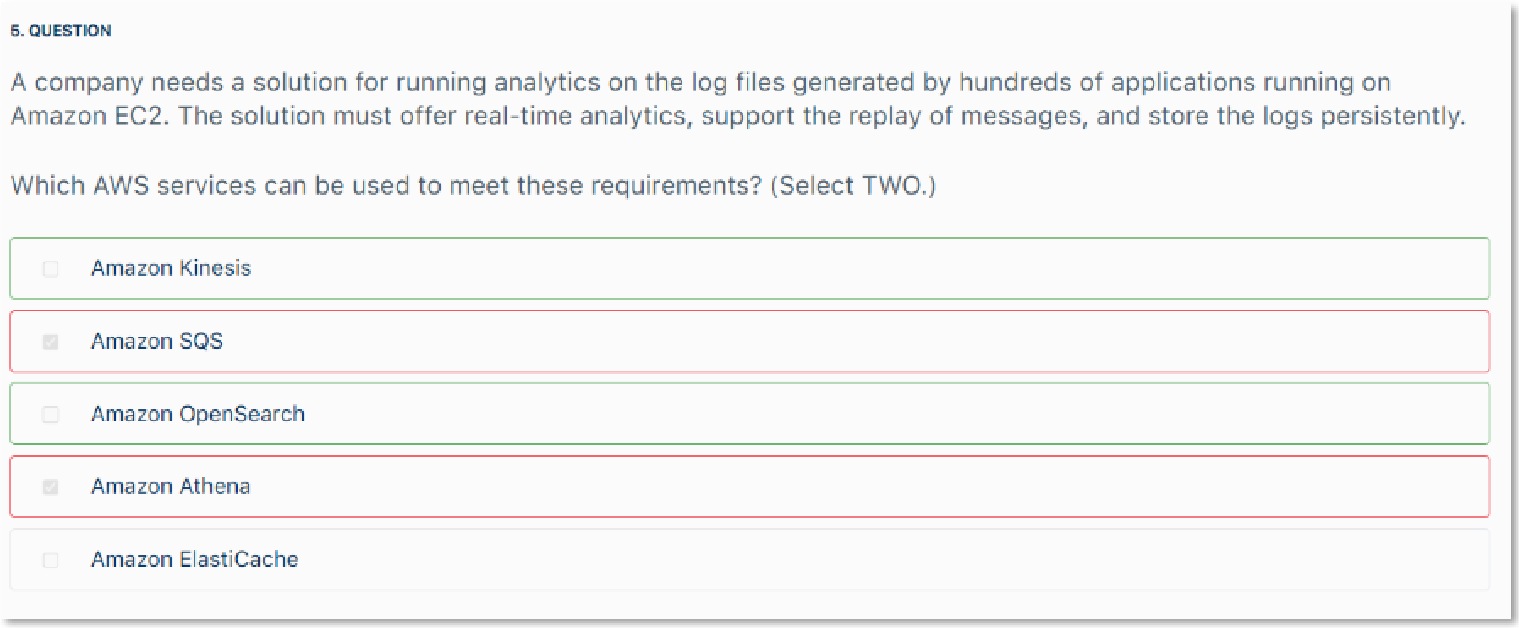
Bibliography

[Nightingale 2023]

Nightingale, Matthew and Alan Peaty (2023). Build a serverless store finder site using Amazon Location Service. *Front-End Web and Mobile* [blog]. Available at: <https://aws.amazon.com/blogs/mobile/build-a-serverless-store-finder-site-using-amazon-location-service/?ck_subscriber_id=1560524742>

MemoryDB for Redis

Amazon OpenSearch Service

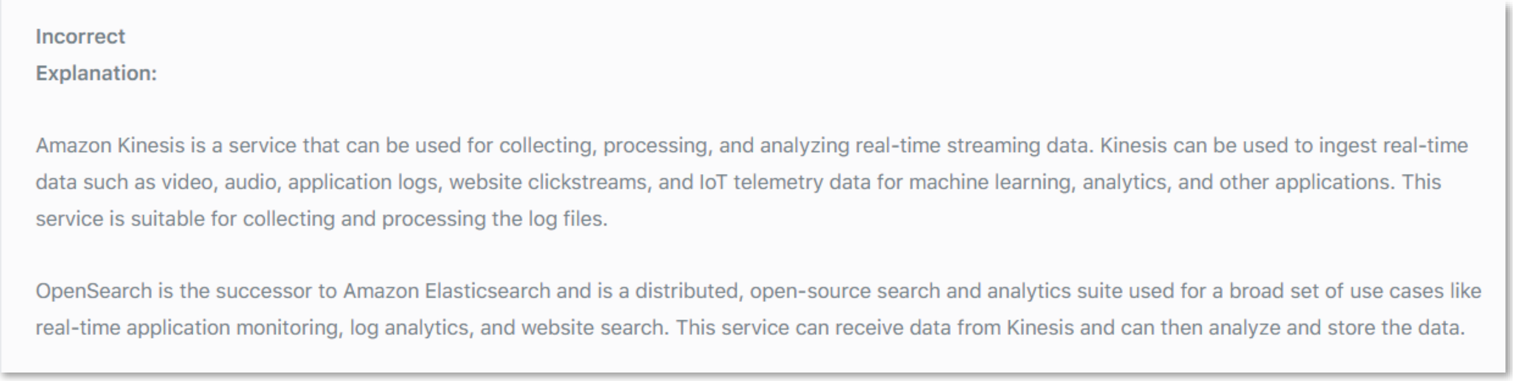


5. QUESTION

A company needs a solution for running analytics on the log files generated by hundreds of applications running on Amazon EC2. The solution must offer real-time analytics, support the replay of messages, and store the logs persistently.

Which AWS services can be used to meet these requirements? (Select TWO.)

- □ Amazon Kinesis
- ☑ Amazon SQS
- □ Amazon OpenSearch
- ☑ Amazon Athena
- □ Amazon ElastiCache

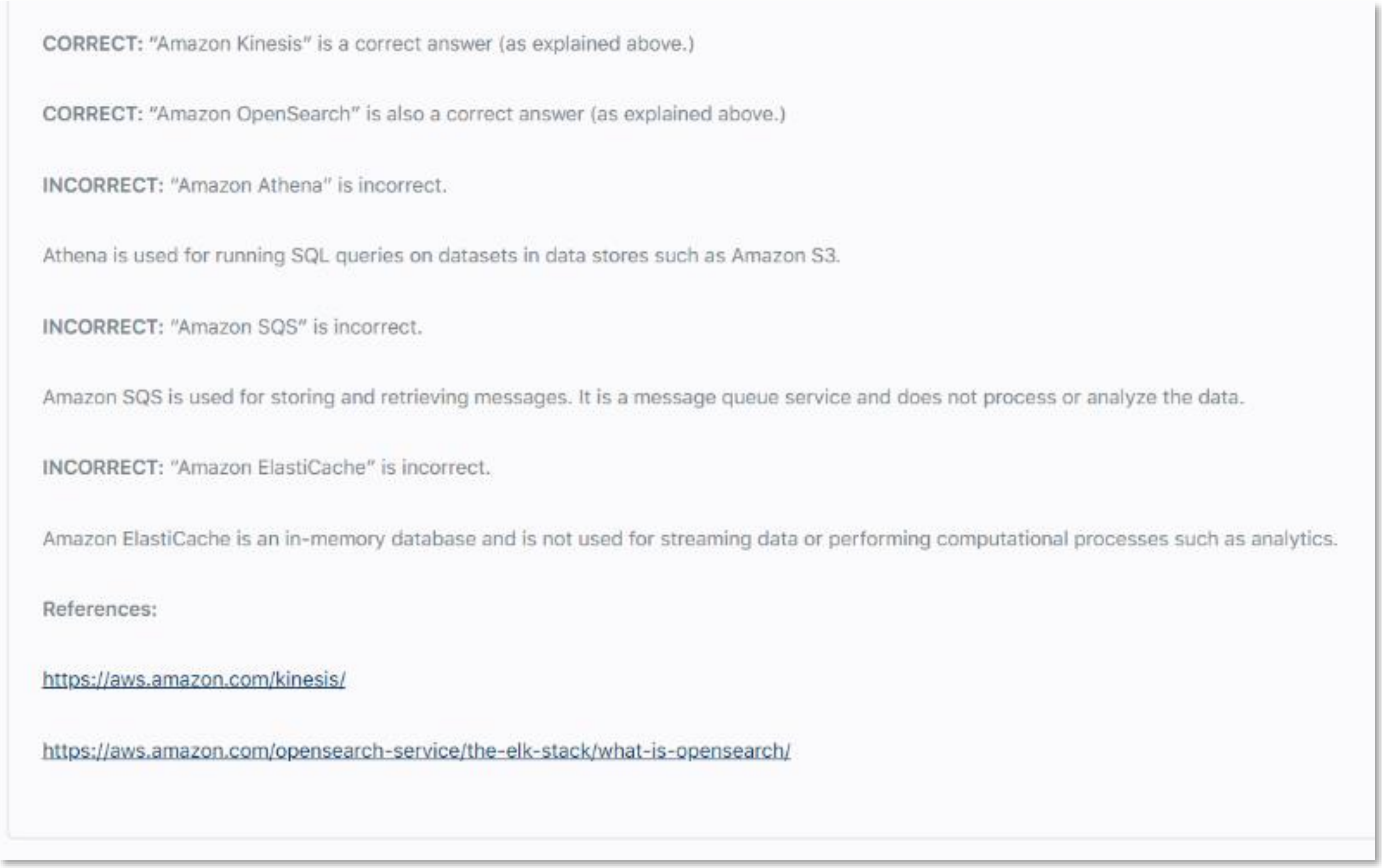


Incorrect

Explanation:

Amazon Kinesis is a service that can be used for collecting, processing, and analyzing real-time streaming data. Kinesis can be used to ingest real-time data such as video, audio, application logs, website clickstreams, and IoT telemetry data for machine learning, analytics, and other applications. This service is suitable for collecting and processing the log files.

OpenSearch is the successor to Amazon Elasticsearch and is a distributed, open-source search and analytics suite used for a broad set of use cases like real-time application monitoring, log analytics, and website search. This service can receive data from Kinesis and can then analyze and store the data.



CORRECT: "Amazon Kinesis" is a correct answer (as explained above.)

CORRECT: "Amazon OpenSearch" is also a correct answer (as explained above.)

INCORRECT: "Amazon Athena" is incorrect.

Athena is used for running SQL queries on datasets in data stores such as Amazon S3.

INCORRECT: "Amazon SQS" is incorrect.

Amazon SQS is used for storing and retrieving messages. It is a message queue service and does not process or analyze the data.

INCORRECT: "Amazon ElastiCache" is incorrect.

Amazon ElastiCache is an in-memory database and is not used for streaming data or performing computational processes such as analytics.

References:

https://aws.amazon.com/kinesis/

https://aws.amazon.com/opensearch-service/the-elk-stack/what-is-opensearch/

Cloud Control API

AWS announces the general availability of AWS Cloud Control API

Posted On: Sep 30, 2021

AWS announces the general availability of AWS Cloud Control API, a set of common application programming interfaces (APIs) that is designed to make it easy for developers to manage their cloud infrastructure in a consistent manner and leverage the latest AWS capabilities faster. Using Cloud Control API, developers can manage the lifecycle of hundreds of AWS resources and over a dozen third-party resources with five consistent APIs instead of using distinct service-specific APIs. With this launch, AWS Partner Network (APN) Partners can now automate how their solutions integrate with existing and future AWS features and services through a one-time integration, instead of spending weeks of custom development work as new resources become available. Terraform by HashiCorp and Pulumi have integrated their solutions as part of this launch.

Cloud Control API enable developers to create, read, update, delete, and list (CRUDL) AWS and third-party service resources with consistent APIs. Resources include schema (properties and handler permissions) and handlers that control API interactions with the underlying services. Using Cloud Control API, developers have a uniform method to manage supported services throughout their lifecycle, so there are fewer APIs to learn as developers add services to their infrastructure. For instance, developers can create supported cloud resources using Cloud Control API's CreateResource API, be it an AWS Lambda function, an Amazon Elastic Container Service (ECS) cluster, or hundreds of other AWS resources along with over a dozen third-party solutions available on the CloudFormation Registry spanning monitoring, databases, or security management resources. Developers can move faster by removing the need to author, maintain, and set up custom code across distinct service-specific APIs. Furthermore, Cloud Control API is up-to-date with the latest AWS resources as soon as they are available on the CloudFormation Registry, enabling APN partners to integrate their own solutions with Cloud Control API just once, and then automatically access new AWS resources without additional integration work.

Cloud Control API is generally available in the following AWS Regions: US East (N. Virginia, Ohio), US West (Oregon, N. California), Canada (Central), Europe (Ireland, Frankfurt, London, Stockholm, Paris, Milan), Asia Pacific (Hong Kong, Mumbai, Osaka, Singapore, Sydney, Seoul, Tokyo), South America (Sao Paulo), Middle East (Bahrain), Africa (Cape Town), and AWS GovCloud (US).

You can use the AWS CLI or AWS SDKs to get started with Cloud Control API. To learn more:

- Visit the AWS Cloud Control API product page
- Check out the AWS News Blog post
- Refer to the User Guide and API reference

Bibliography

I. Official

AWS Resilience Hub

Announcing general availability of AWS Resilience Hub

Posted On: Nov 10, 2021

Amazon Web Services (AWS) has announced the general availability of AWS Resilience Hub, a new service that provides you with a single place to define, validate, and track the resilience of your applications so that you can avoid unnecessary downtime caused by software, infrastructure, or operational disruptions.

The resilience of an application refers to its ability to maintain availability and recover from software and operational disruption within a specified target measured in terms of Recovery Time Objective (RTO) and Recovery Point Objective (RPO). Using AWS Resilience Hub, you can define your applications' resilience targets (RTO and RPO) and help validate that these targets can be met prior to deployment. AWS Resilience Hub provides automated assessments that identify resilience weaknesses and provide recommended remediation. AWS Resilience Hub also integrates with AWS Fault Injection Simulator to test that resilience targets can be met under different conditions (e.g., database disruptions). When integrated into customers' CI/CD pipelines, AWS Resilience Hub provides continuous resilience assessments and testing.

AWS Resilience Hub provides a comprehensive view of your overall application portfolio resilience status through its dashboard. To help you track the resilience of applications, AWS Resilience Hub aggregates and organizes resilience events (e.g., unavailable database or failed resilience validation), alerts, and insights from services like Amazon CloudWatch and AWS Fault Injection Simulator. AWS Resilience Hub also generates a resilience score, a scale that indicates the level of implementation for recommended resilience tests, alarms, and recovery SOPs. This score can be used to measure resilience improvements over time.

AWS Resilience Hub is available today in US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Tokyo), Europe (Ireland), and Europe (Frankfurt). Additional regions will be added in the future.

You can try AWS Resilience Hub free for 6 months – up to 3 applications. After that, the AWS Resilience Hub price is $15.00 per application per month. Metering begins once you run the first resilience assessment in AWS Resilience Hub. You will incur charges for any AWS service provisioned by AWS Resilience Hub. Consult your AWS pricing plan for more details on additional charges and visit the AWS Resilience Hub pricing page.

To learn more about AWS Resilience Hub, visit our product page.

Bibliography

I. Official

[AWS 2021]

Announcing general availability of AWS Resilience Hub. [Announcement]. Nov 10th 2021. Available at: <https://aws.amazon.com/about-aws/whats-new/2021/11/aws-resilience-hub-general-availability/>

[Stormacq 2021]

Stormacq, Sebastian (2021). Measure and Improve Your Application Resilience with AWS Resilience Hub.. *AWS News Blog*. Nov 10th 2021. Available at: <https://aws.amazon.com/blogs/aws/monitor-and-improve-your-application-resiliency-with-resilience-hub/>

AWS IoT RoboRunner

Bibliography

Bibliography

What is 5G? Available at:
<https://www.youtube.com/watch?v=2DG3pMcNNlw&ab_channel=CNBCInternational>

AWS IoT TwinMaker

NEW

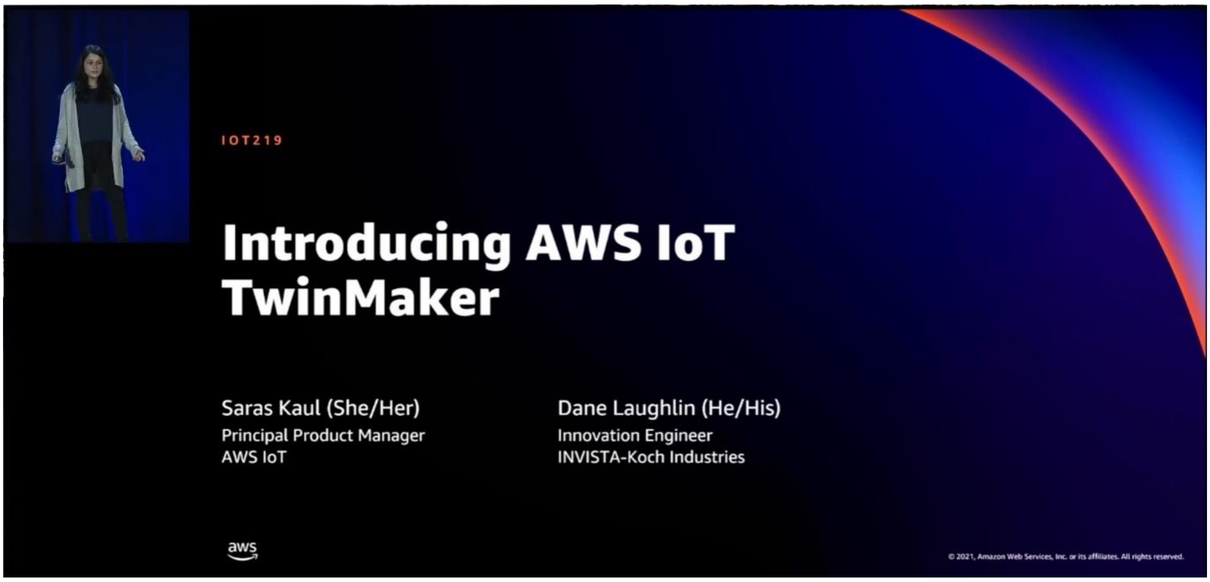
AWS IoT TwinMaker

A new service that makes it easy to create and use digital twins of real-world systems

PREVIEW AVAILABLE TODAY

Questions:

1. Might we one day have a digital twin of a river? Or a rat? A bacterium?
2.



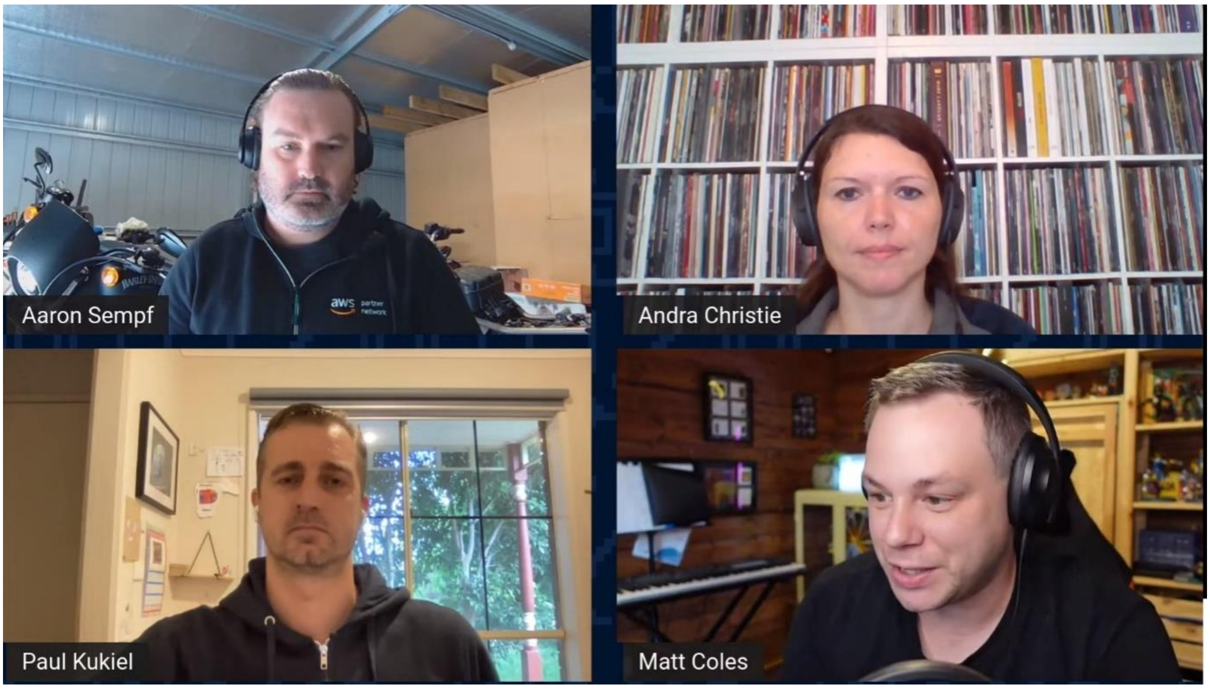
IOT219

Introducing AWS IoT TwinMaker

Saras Kaul (She/Her)
Principal Product Manager
AWS IoT

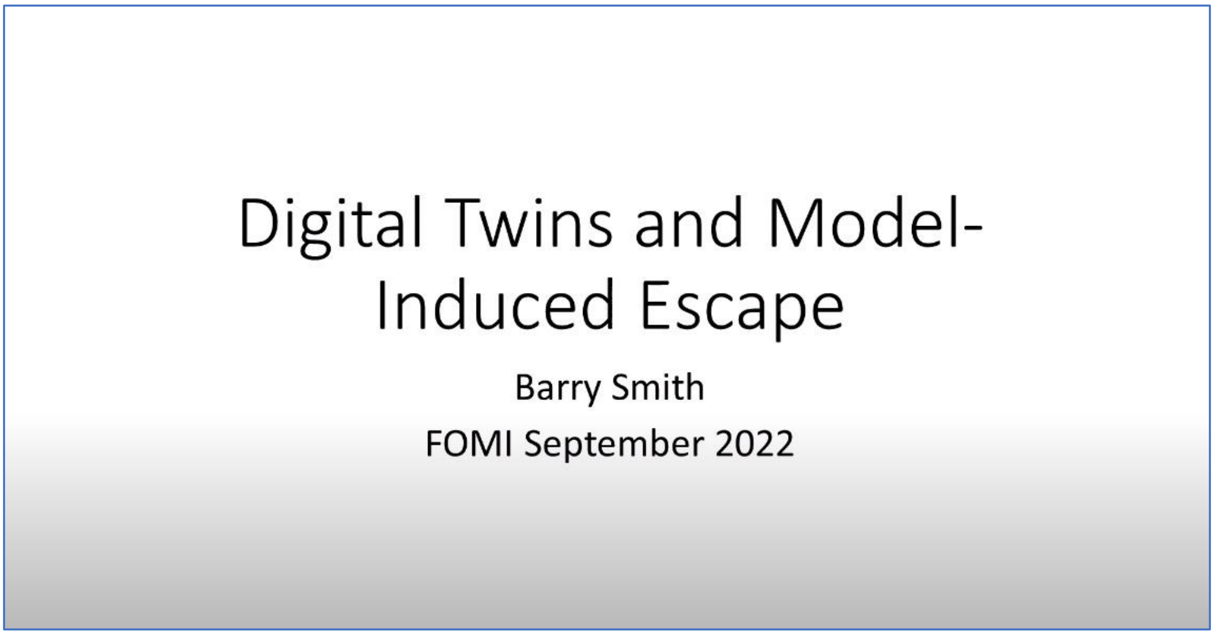
Dane Laughlin (He/His)
Innovation Engineer
INVISTA-Koch Industries

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Above, we can see an image from a discussion of AWS IoT TwinMaker. The discussion addresses a number of questions:

What is a digital twin?



Digital Twins and Model-Induced Escape

Barry Smith

FOMI September 2022

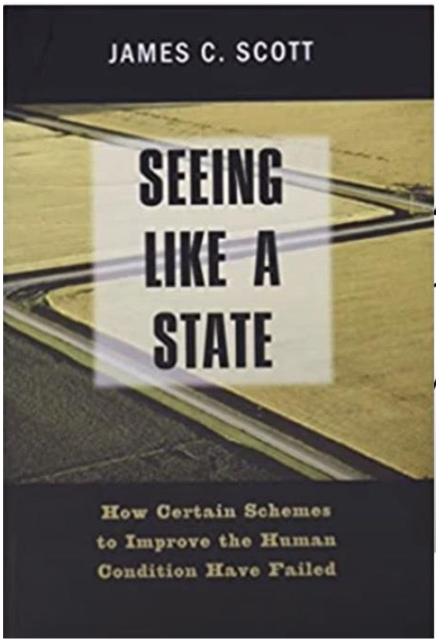
“Model induced escape: the model itself creates the reasons for the model failing. Something similar happens with spam filters. If you have a really good spam filter then the authors of spam will use that spam filter in order to create new and wonderful types of spam, and thereby render your spam filter no longer operative.

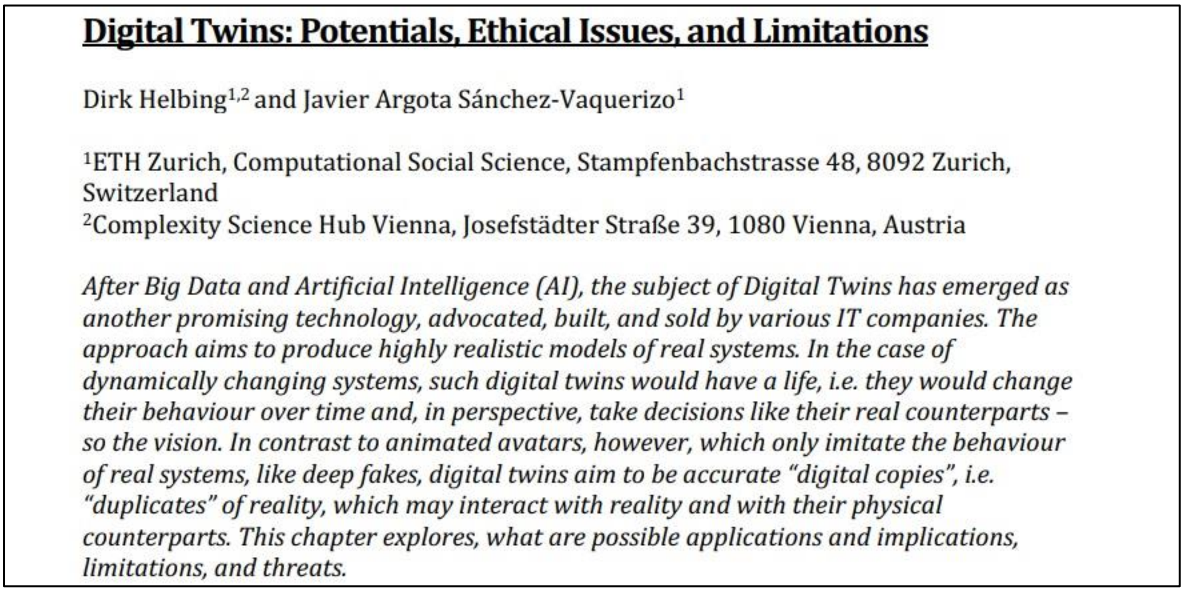
This has a consequence for digital twins, particularly where digital twins are designed to reflect the patterns of behavior of people (this would apply in other areas too).

So, the idea is that the **very existence** of a digital twin, operating in relation to people's behaviour patterns, will lead to changes in those behaviour patterns. There will now be a digital twin **inside** that system and so the digital twin will no longer be accurate because the digital twin does not accommodate this new feature of the system which it is supposed to be a model of, namely that the digital twin is no longer a part of the system.

Now, this is a very beautiful part of the system – a digital twin must always be no longer be a twin once it is installed. But now, there are more substantial ways in which the digital twin will change behaviour. There will be some who will react to the knowledge that they are being monitored by a digital twin by protesting that their behaviour is being monitored.

These attitudes are nicely summarised in this book from 1999:


JAMES C. SCOTT
SEEING LIKE A STATE
How Certain Schemes to Improve the Human Condition Have Failed



Digital Twins: Potentials, Ethical Issues, and Limitations

Dirk Helbing[1,2] and Javier Argota Sánchez-Vaquerizo[1]

[1]ETH Zurich, Computational Social Science, Stampfenbachstrasse 48, 8092 Zurich, Switzerland
[2]Complexity Science Hub Vienna, Josefstädter Straße 39, 1080 Vienna, Austria

After Big Data and Artificial Intelligence (AI), the subject of Digital Twins has emerged as another promising technology, advocated, built, and sold by various IT companies. The approach aims to produce highly realistic models of real systems. In the case of dynamically changing systems, such digital twins would have a life, i.e. they would change their behaviour over time and, in perspective, take decisions like their real counterparts – so the vision. In contrast to animated avatars, however, which only imitate the behaviour of real systems, like deep fakes, digital twins aim to be accurate "digital copies", i.e. "duplicates" of reality, which may interact with reality and with their physical counterparts. This chapter explores, what are possible applications and implications, limitations, and threats.

Glossary

Digital Twin -

Bibliography

I. Official
II. Unofficial
III. Critical
IV. General

I. Official

[Kaul 2021]

Kaul, Saras and Dane Laughlin (2021). Introducing AWS IoT TwinMaker. *ReInvent 2021* [Conference]. Available at: <https://www.youtube.com/watch?v=fdBKRyhC9Yk&ab_channel=AWSEvents>

[Christie 2022]

Christie, Andra (2022). YouTube Channel: Devs in The Shed. Available at: https://www.youtube.com/watch?v=SRT5QczaQj8&ab_channel=DevsintheShed

II. Unofficial

III. Critical

Digital Twins and the Problem of Model-Induced Escape. Available at: <https://www.youtube.com/watch?v=7M3B9v7Ylhc&ab_channel=BarrySmith>

[Dirk 2022]

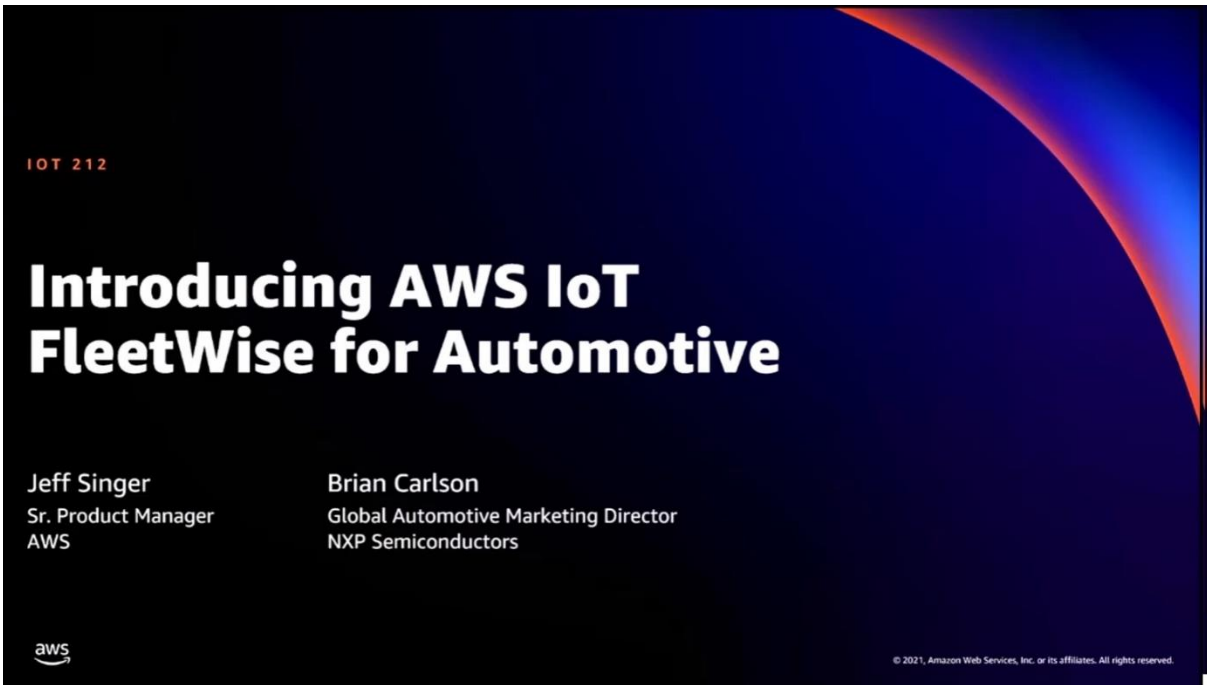
Helbing, Dirk and Javier Argota Sánchez-Vaquerizo (2022). Digital Twins: Potentials, Ethical Issues, and Limitations.

IV. General

Scott, James (1999). Seeing like a state.

AWS FleetWise

What is AWS IoT FleetWise?



IOT 212

Introducing AWS IoT FleetWise for Automotive

Jeff Singer
Sr. Product Manager
AWS

Brian Carlson
Global Automotive Marketing Director
NXP Semiconductors

NXP

Jeff Singer, the Senior Product Manager for FleetWise, introduces Brian Carlson onto the stage. Carlson is from NXP, a partner of AWS. It is explained that NXP have:

> decades of experience. A culmination, going back to Motorola, FreeScale, and Philips semiconductors. So, decades of experience in automotive processing, security.
>
> Your passport probably has our chip in their, for security—or your bank cards; enterprise networking. What we basically did here is converge all of that expertise and knowledge from our side into a single chip that really matches well with what AWS are doing in automotive, specifically here with FleetWise.

[Singer 2021]

NXP S32G
Vehicle Processing
Platforms

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Glossary

OEM – stands for Original Equipment Manufacturer.

Event-based collection -

Rule-based collection –

Bibliography

I. Official
II. Unofficial
III. Critical
IV. General

I. Official

[AWS 2021]

AWS IoT FleetWise. YouTube Channel: Amazon Web Services. Nov 30th 2021. Available at: <https://www.youtube.com/watch?v=oNr2-chK-bc&ab_channel=AmazonWebServices>

[Singer 2021]

Singer, Jeff and Brian Carlson (2021). Introducing AWS IoT FleetWise for Automotive. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=YWoexR5lN3o&ab_channel=AWSEvents>

II. Unofficial

III. Critical

IV. General

[Deloitte]

Software-Defined Vehicles – A Forthcoming Industrial Evolution. *Deloitte.* Available at: <https://www2.deloitte.com/cn/en/pages/consumer-business/articles/software-defined-cars-industrial-revolution-on-the-arrow.html>

Mainframe Modernization

NEW

AWS Mainframe Modernization

Migrate, modernize, and run mainframe workloads on AWS

PREVIEW AVAILABLE TODAY

“So you can see, we continue to seek ways to make the cloud more powerful and more cost-effective for every workload.

But do we really mean *every* workload? What about mainframes?

A lot of companies, of course, have been running applications on mainframes for many decades. Customers from every industries still rely on them.

But mainframes are expensive. complicated. And there are fewer and fewer people who are learning to program COBOL these days. So, maintaining a mainframe is like trying to shoot a basketball with two feet on the ground – you know you can do it, but you also know there’s got to be a better way.

This is why so many of our customers are trying to get off their mainframe as fast as they possibly can, to gain the agility and elasticity of the cloud.

We've seen customers reduce their costs by up to 70% (or more) after migrating.

Now there are a couple of different paths that customers take. Some start with a "lift and shift" approach and bring the application pretty much "as is". Others refactor and break the application down into microservices in the cloud.

Neither road is as easy as customers would like, and in fact whichever way you go can take months, even years. You have to evaluate the complexity of the application's source code, understand the dependencies on other systems, convert or recompile the code, and then you still have to test it.

All of this is before you've actually moved anything. It can be a **messy business**, and it involves a lot of moving pieces.

It isn't something that people really want to do on their own and while we have lots of partners that can help, even then it's a lot of time to be standing with two feet on the ground while your competitors are **shooting jumpshots**.

So today I'm really pleased to announce AWS Mainframe Modernization, which is a new service to make it faster to migrate, modernize and run mainframe applications on AWS.

[Selipsky 2021: 31:03]

AWS Mainframe Modernization is now generally available

Posted On: Jun 8, 2022

Introduced at re:Invent in November 2021, AWS Mainframe Modernization is now generally available for customer and partner use. Mainframe Modernization is a unique platform that allows you to migrate and modernize your on-premises mainframe workloads to a managed and highly available runtime environment on AWS. The service currently supports two main migration patterns—replatforming and automated refactoring—so you can select your best-fit migration path and associated toolchains based on your migration assessment results.

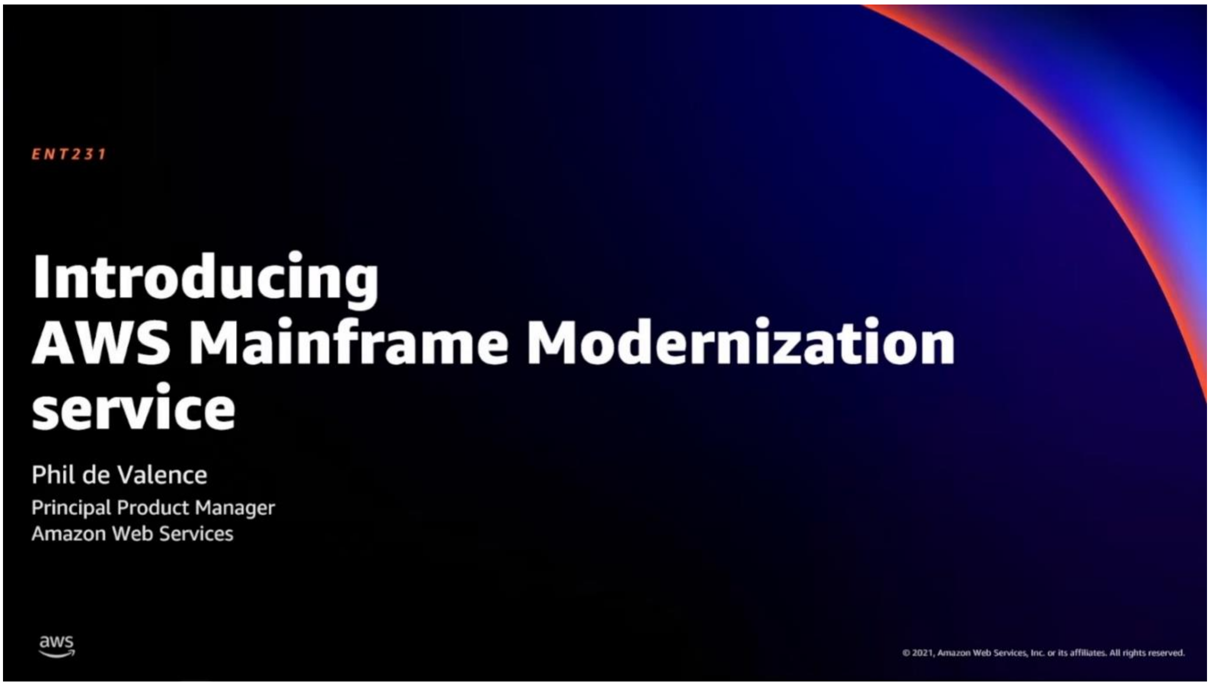
Use Mainframe Modernization to easily migrate and modernize your mainframe applications, increasing agility and reducing costs. You can break up and manage your complete migration with infrastructure, software, and tools to refactor or replatform legacy applications. Deploy, test, run, maintain, operate, and evolve migrated applications in the runtime environment with no upfront costs.

To learn more and start planning your mainframe migration and modernization to AWS today, see the AWS Mainframe Modernization web page and the AWS Mainframe Modernization User Guide.

Glossary

Mainframe –

COBOL -



ENT231

Introducing AWS Mainframe Modernization service

Phil de Valence

Principal Product Manager
Amazon Web Services

Since the introduction of System/360 on April 7 1964, mainframe computers have enabled many industries to transform themselves. The mainframe has revolutionized the way people buy things, how people book and purchase travel, and how governments manage taxes or deliver social services. Two thirds of the Fortune 100 companies have their core businesses located on a mainframe. And according to a 2018 estimate, $3 trillion ($3 x 10^12) in daily commerce flows through mainframes.

Mainframes are using their very own set of technologies: programming languages such as COBOL, PL/1, and Natural, to name a few, or databases and data files such as VSAM, DB2, IMS DB, or Adabas. They also run “application servers” (or transaction managers as we call them) such as CICS or IMS TM. Recent IBM mainframes also run applications developed in the Java programming language deployed on WebSphere Application Server.

[Stormacq 2022]

ENT230

Executing a large-scale migration & modernization

Jonathan Allen (he/him)
Director, Enterprise Strategy
AWS

Mary McDaniel (she/her)
Global Practice Manager Migrations
AWS

Steve Steuart (he/him)
WW Mainframe Business Development
AWS

Bibliography

[AWS 2021]

Introducing AWS Mainframe Modernization – Preview. [Announcement]. Nov 30th 2021. Available at: <https://aws.amazon.com/about-aws/whats-new/2021/11/introducing-aws-mainframe-modernization/>

[AWS 2022]

AWS Mainframe Modernization is now generally available [Announcement]. Available at: https://aws.amazon.com/about-aws/whats-new/2022/06/aws-mainframe-modernization-generally-available/#:~:text=Introduced%20at%20re%3AInvent%20in,available%20runtime%20environment%20on%20AWS.

[Perez 2021]

Perez, Sarah (2021). AWS Introduces a solution to get customers off their mainframes more quickly. *TechCrunch*. Available at: <https://techcrunch.com/2021/11/30/aws-introduces-a-solution-to-get-customers-off-their-mainframes-more-quickly/>

[Valence 2021]

Valence, Phil de (2021). Introducing AWS Mainframe Modernization service. ReInvent 2021 [Conference]. Available at: <https://www.youtube.com/watch?v=07u4ANDTZt8&ab_channel=AWSEvents>

[Stormacq 2022]

Stormacq, Sebastien (2022). Modernize Your Mainframe Applications & Deploy Them In The Cloud. *AWS News Blog*. June 8th 2022. Available at: <https://aws.amazon.com/blogs/aws/modernize-your-mainframe-applications-deploy-them-in-the-cloud/>

[Wikipedia]

Mainframe Computer. *Wikipedia*. Available at: <https://en.wikipedia.org/wiki/Mainframe_computer>

Executing a large-scale migration and modernization. YouTube Channel: AWS Events. Available at: <https://www.youtube.com/watch?v=TqdjwcFGuCw&ab_channel=AWSEvents>

Tang, Lewis (2022). Augmentation patterns to modernize a mainframe on AWS. *AWS Architecture Blog*. July 25th 2022. Available at: <https://aws.amazon.com/blogs/architecture/augmentation-patterns-to-modernize-a-mainframe-on-aws/>

Schwartz, Mark (2018). Yes, You Should Modernize Your Mainframe with the Cloud. AWS Cloud Enterprise Strategy Blog. July 5th 2018. Available at:< https://aws.amazon.com/blogs/enterprise-strategy/yes-you-should-modernize-your-mainframe-with-the-cloud/>

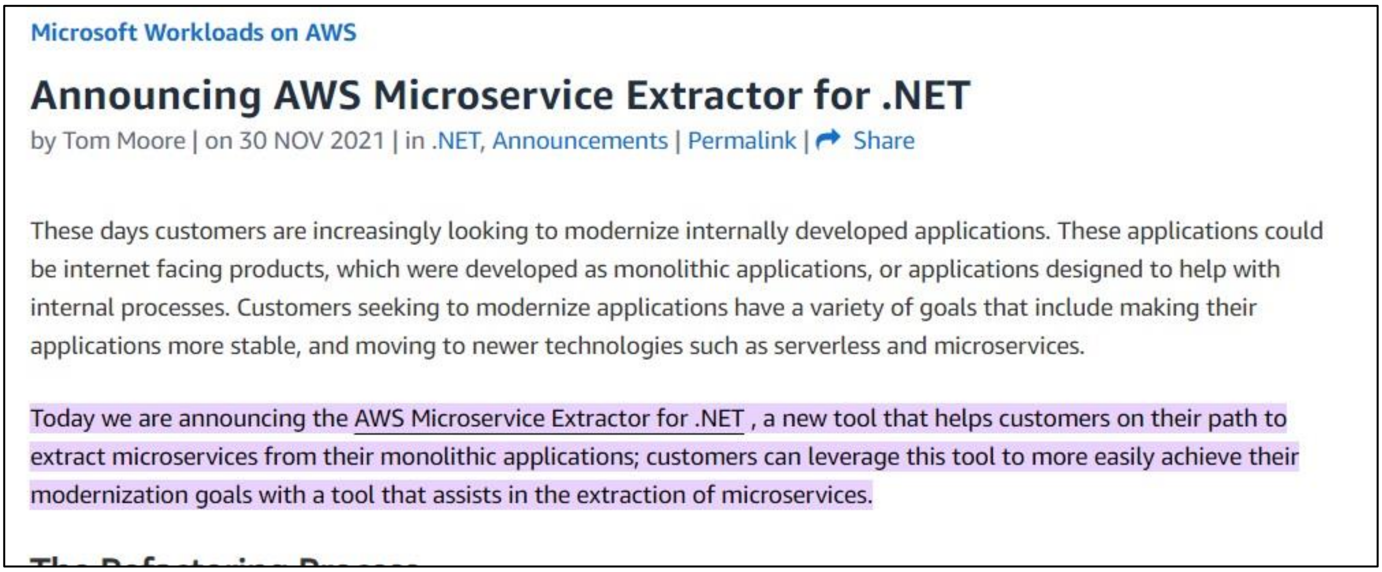
[Cantrill 2020]

Cantrill, Bryan (2020). The Soul of a New Machine: Rethinking the Computer. YouTube Channel: Stanford Online. Available at: https://www.youtube.com/watch?v=vvZA9n3e5pc&ab_channel=StanfordOnline

←

Bibliography

Microservice Extractor



Microsoft Workloads on AWS

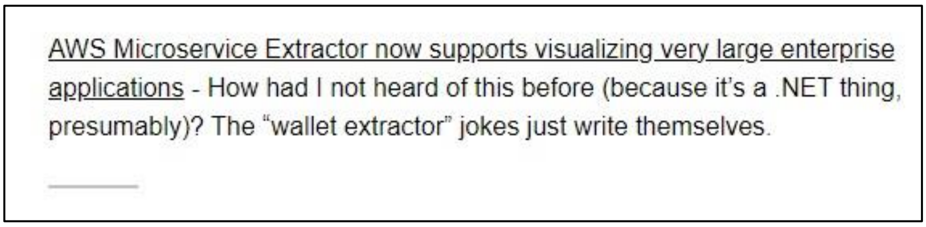
Announcing AWS Microservice Extractor for .NET

by Tom Moore | on 30 NOV 2021 | in .NET, Announcements | Permalink | Share

These days customers are increasingly looking to modernize internally developed applications. These applications could be internet facing products, which were developed as monolithic applications, or applications designed to help with internal processes. Customers seeking to modernize applications have a variety of goals that include making their applications more stable, and moving to newer technologies such as serverless and microservices.

Today we are announcing the AWS Microservice Extractor for .NET , a new tool that helps customers on their path to extract microservices from their monolithic applications; customers can leverage this tool to more easily achieve their modernization goals with a tool that assists in the extraction of microservices.

The Refactoring Process



AWS Microservice Extractor now supports visualizing very large enterprise applications - How had I not heard of this before (because it's a .NET thing, presumably)? The "wallet extractor" jokes just write themselves.

Corey Quinn on the microservice extractor, in an email from August 28th 2023

Bibliography

AWS RePost

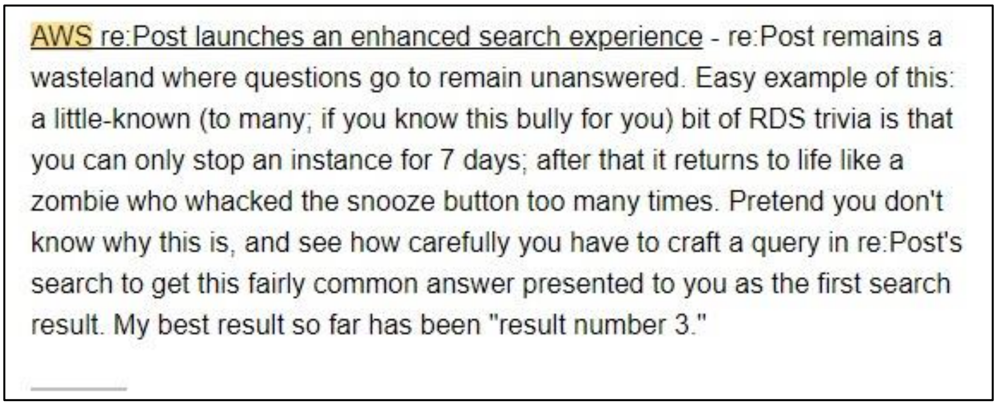
Introducing AWS re:Post, a new, community-driven, questions-and-answers service

Posted On: Dec 2, 2021

Amazon Web Services (AWS) announces the availability of AWS re:Post (re:Post), a new, community-driven, questions-and-answers service to help AWS customers remove technical roadblocks, accelerate innovation, and enhance operation. AWS re:Post enables you to ask questions about anything related to designing, building, deploying, and operating workloads on AWS, and get answers from community experts, including AWS customers, Partners, and employees.

AWS re:Post replaces AWS Forums and introduces new ways to improve the accuracy of answers provided, as well as the likelihood of receiving an answer from the community. AWS re:Post automatically connects your question with subject matter experts, and is also integrated with AWS Support. Customers with AWS Premium Support subscriptions receive responses from AWS employees for questions that are not answered by the community.

AWS re:Post is part of the AWS Free Tier and is available to anyone with an AWS account at https://repost.aws.



AWS re:Post launches an enhanced search experience - re:Post remains a wasteland where questions go to remain unanswered. Easy example of this: a little-known (to many; if you know this bully for you) bit of RDS trivia is that you can only stop an instance for 7 days; after that it returns to life like a zombie who whacked the snooze button too many times. Pretend you don't know why this is, and see how carefully you have to craft a query in re:Post's search to get this fairly common answer presented to you as the first search result. My best result so far has been "result number 3."

Corey Quinn on re:Post, in an email sent on August 28th 2023

Bibliography

Application Migration Service (AWS MGN)

Announcing general availability of AWS Application Migration Service

Posted On: May 18, 2021

Amazon Web Services (AWS) has announced the general availability of AWS Application Migration Service (AWS MGN), a new service that enables organizations to move applications to AWS without making changes to the applications, their architecture, or the migrated servers. AWS Application Migration Service is the primary migration service recommended for lift-and-shift migrations to AWS. Customers currently using CloudEndure Migration or Server Migration Service (SMS) are encouraged to switch to AWS Application Migration Service for future migrations. Visit our product comparison page for specific reasons to use CloudEndure Migration or SMS.

AWS Application Migration Service minimizes time-intensive, error-prone manual processes by automatically converting source servers from physical, virtual, and cloud infrastructure to run natively on AWS. This helps to simplify the migration of all types of applications to AWS, including enterprise applications such as SAP CRM, Oracle E-Business Suite, Microsoft SharePoint, and commercial databases.

AWS Application Migration Service continuously replicates your source servers without interfering with the normal operation of the servers, enables non-disruptive testing prior to cutover, and allows for cutover windows measured in minutes. With AWS Application Migration Service, you use the same straightforward process, regardless of source infrastructure, operating system, or the application being migrated. It does not require server reboots or changes to the source infrastructure.

AWS Application Migration Service is based on the technology of CloudEndure Migration and provides similar capabilities, but is available on the AWS Management Console. This enables seamless integration with other AWS services, such as AWS CloudTrail, Amazon CloudWatch, and AWS Identity and Access Management (IAM).

AWS Application Migration Service is currently available in US East (N. Virginia), US West (Oregon), US East (Ohio), Asia Pacific (Tokyo), Asia Pacific (Sydney), Asia Pacific (Singapore), Europe (Ireland), Europe (Frankfurt), and Europe (Stockholm).

You can use AWS Application Migration Service free for 90 days for each server you want to migrate. You pay only for AWS infrastructure that is provisioned during migration and after cutover, such as compute (Amazon EC2) and storage (Amazon EBS) resources. Visit the pricing page for additional details.

To learn more about AWS Application Migration Service, see our documentation. To get started, sign in to the AWS Application Migration Service Console.

Bibliography

[AWS 2021]

Announcing general availability of the Application Migration Service. Available at: https://aws.amazon.com/about-aws/whats-new/2021/05/announcing-general-availability-of-aws-application-migration-service/

[Shah 2021]

Namrata Shah. Aug 17th 2021. YouTube Channel: NamrataHShah. Available at: <https://www.youtube.com/watch?v=VbD2cS3uZy0&ab_channel=NamrataHShah>