
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

STG221-L

AWS storage innovation at exabyte scale

Mai-Lan Tomsen Bukovec (she/her)
VP of Foundational Data Services
AWS

Andy Warfield (he/him)
VP and Distinguished Engineer
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

TELEMETRY

LOGS

EVENTS

THIRD PARTY

101 ZB

of data created and replicated in the world in 2022

IOT

WEB

APPS

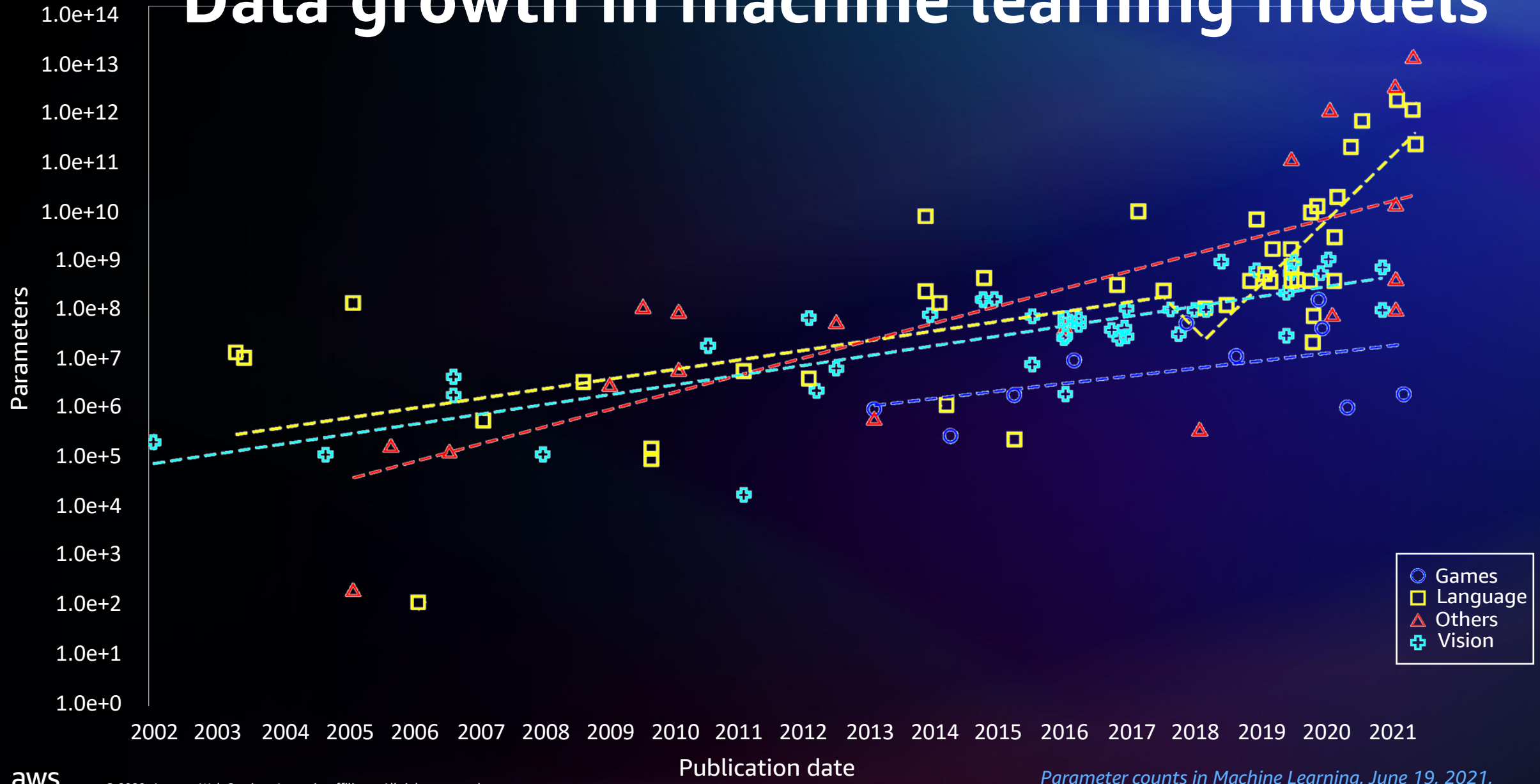
SENSORS

MOBILE

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Source: *Worldwide IDC Global DataSphere Forecast, 2022–2026: Enterprise Organizations Driving Most of the Data Growth*, Doc. #US49018922

Data growth in machine learning models



Parameter counts in Machine Learning, June 19, 2021, AI ALIGNMENT FORUM

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data protection

AWS BACKUP

AWS ELASTIC DISASTER RECOVERY

AMAZON EBS SNAPSHOTS

Data storage

BLOCK

AMAZON EBS

OBJECT

AMAZON S3 & S3 GLACIER

FILE

AMAZON FSX FAMILY

AMAZON EFS

Data at work

ANALYTICS

MACHINE LEARNING

VISUALIZATION

AMAZON FILE CACHE

STREAMING

Data in motion

AWS TRANSFER FAMILY

AWS SNOW FAMILY

AWS STORAGE GATEWAY

AWS DATASYNC

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Simplified vaccination verification process for travelers with United Travel Ready Center



Pivoted during pandemic with ability to scale usage down on AWS to save on costs

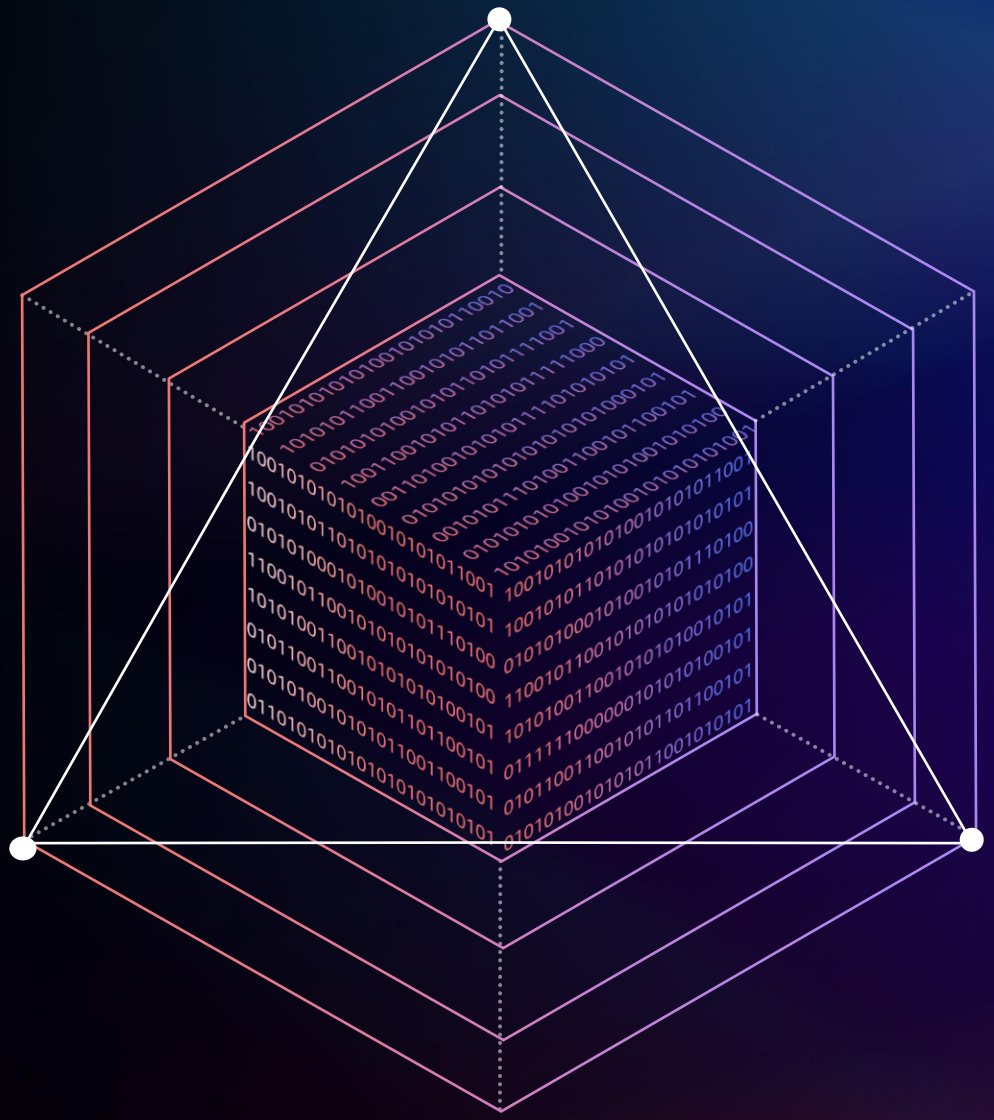
Launched the United Travel Ready Center built on Amazon S3 and Amazon SageMaker with information about required vaccinations or credentials, allowing customers to upload documents easily

"AWS enabled us to **increase the speed of innovation** during a crisis. It helped us **variabilize our cost** structure and put us on a path to replace aging legacy platforms."

Linda Jojo
EVP of Technology & Chief Digital Officer, United Airlines

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Design for agility



Dynamically manage growth

Plan for the unexpected

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Easy paths to migrate shared data to the cloud



MIGRATE TO LIKE-FOR-LIKE STORAGE WITH FAMILIAR CAPABILITIES AND APIs

- Re-architect
- Re-invent
- Re-certify
- Re-train

NetApp ONTAP



2x SSD throughput: 2 GB/s → 4 GB/s
2x SSD IOPs: 80K → 160K

NVMe read cache:
Up to 6 GB/s throughput and 650K IOPS

NEW Nov. 28

Windows File Server



New performance metrics
and enhanced
monitoring dashboard

ZFS or Linux-based file servers



2x SSD throughput: 4 GB/s → 10 GB/s
2x SSD IOPs: 160K → 350K

NVMe read cache:
Up to 21 GB/s throughput and 1M+ IOPS

NEW Nov. 28

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Introducing Elastic Throughput on Amazon EFS

Eliminate performance planning and only pay for what you use

NEW Nov. 27



What it's for

Spiky, unpredictable, throughput-constrained workloads

When to use

If you don't know what you need – for spiky workloads (average-to-peak throughput of <5%) and for new/changing workloads

NEW Nov. 27

Lowering latencies on Amazon EFS by up to 60% with read latency as low as 250 microseconds

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cloud bursting to accelerate workloads with Amazon File Cache

GA Sept. 29

On-premises datasets



NFS file system

Picchu image courtesy of AWS Fuzzy Pixel Studio

Video rendering workloads

Amazon File Cache

Amazon EC2

Hundreds of GB/s of throughput & millions of IOPS at sub-millisecond latencies

"**Ten-fold growth** in lab imaging is driving the need for more on-premises integration – we look forward to using **Amazon File Cache** to bring the compute to our data wherever it resides."

Cajal Neuroscience

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 15

AVAILABLE NOW

Amazon S3 Glacier up to 10x restore throughput increase

UP TO 90% FASTER RETRIEVALS FROM S3 GLACIER FLEXIBLE RETRIEVAL, S3 GLACIER DEEP ARCHIVE



Automatically apply to S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive at no additional costs

Ideal for restoring backups, responding to audit requests, retraining machine learning models, and performing analytics on historical data

Significantly reduce restore completion time for datasets composed of small objects

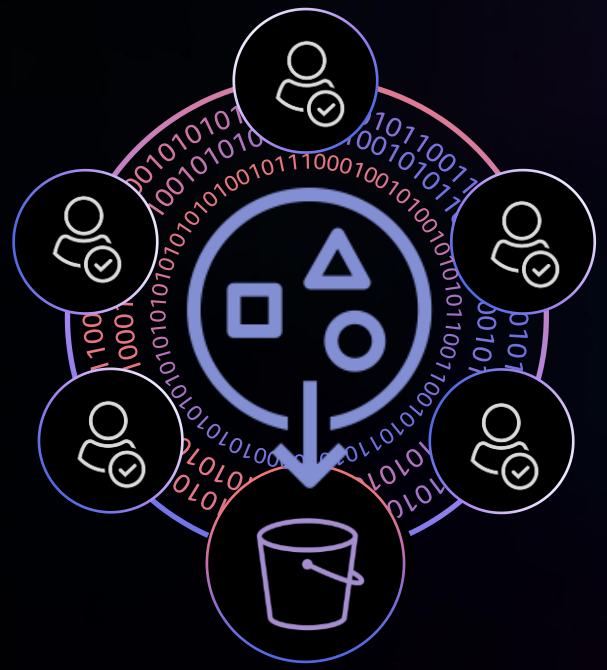
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 30

Amazon S3 now supports cross-account access points

SIMPLIFY DATA ACCESS FOR ANY AWS SERVICE OR APPLICATION THAT STORES DATA IN S3



Amazon S3 Access Points can now be used to securely delegate access permissions for shared datasets to other AWS accounts

Bucket owners are able to authorize access via access points created in other accounts

Bucket owners can delegate specific IAM-based access control decisions to the access point owner from other AWS accounts

aws

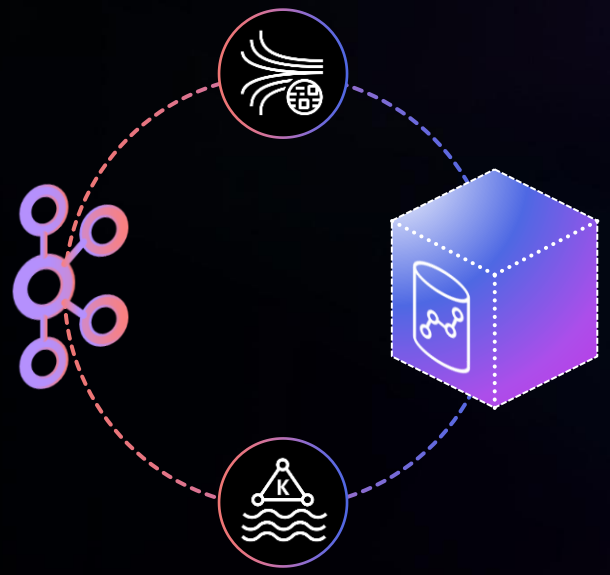
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 29

AVAILABLE NOW

Amazon Redshift Streaming ingestion support

INGEST STREAMING DATA INTO YOUR DATA WAREHOUSE FOR REAL-TIME ANALYTICS



Directly ingest streaming data into your data warehouse from Amazon Kinesis Data Streams and Amazon MSK

Perform rich analytics using familiar SQL, and easily create and manage ELT pipelines

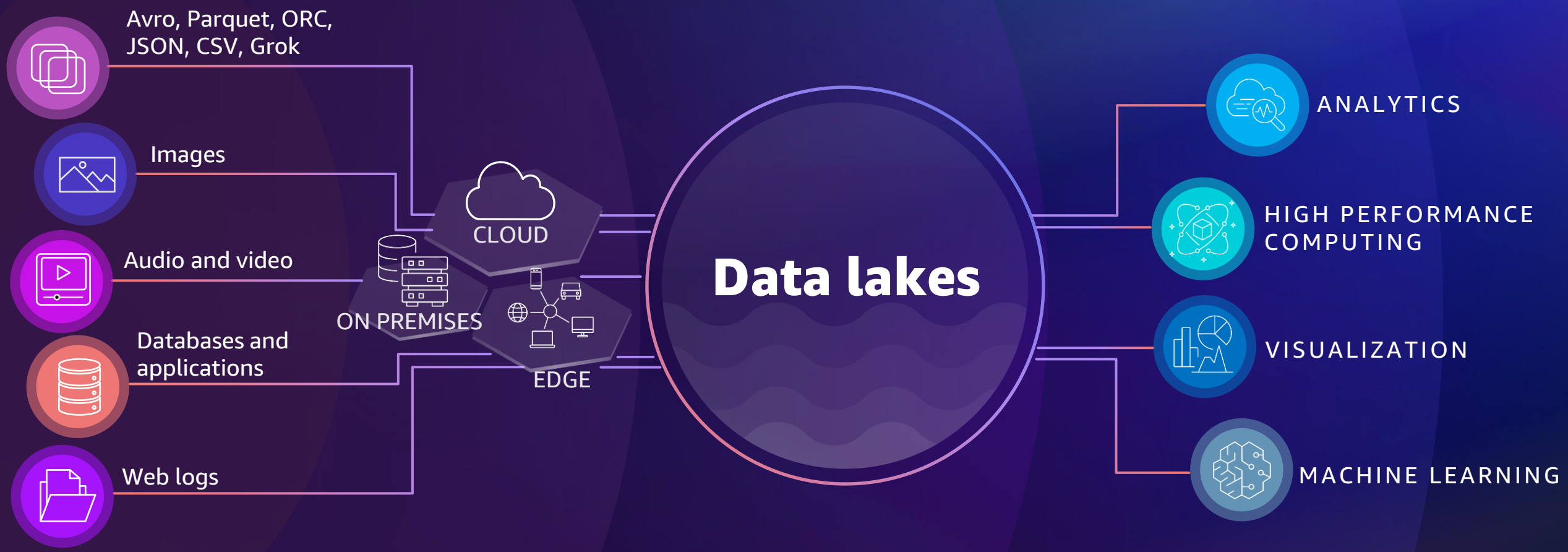
Process large volumes of streaming data from multiple sources to derive insights in seconds

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.


Thorn developed Safer, the first holistic solution that allows technology companies to detect, remove, and report child sexual abuse material at scale using Amazon S3.
THORN
We are Thorn.
*"With support from AWS over the last five years, Thorn has scaled its technology services globally and **kept hundreds of thousands of children safe** from harm."*
Julie Cordua
CEO, Thorn

Agility in decision-making



Any data **Anywhere** **Put data to work**

aws

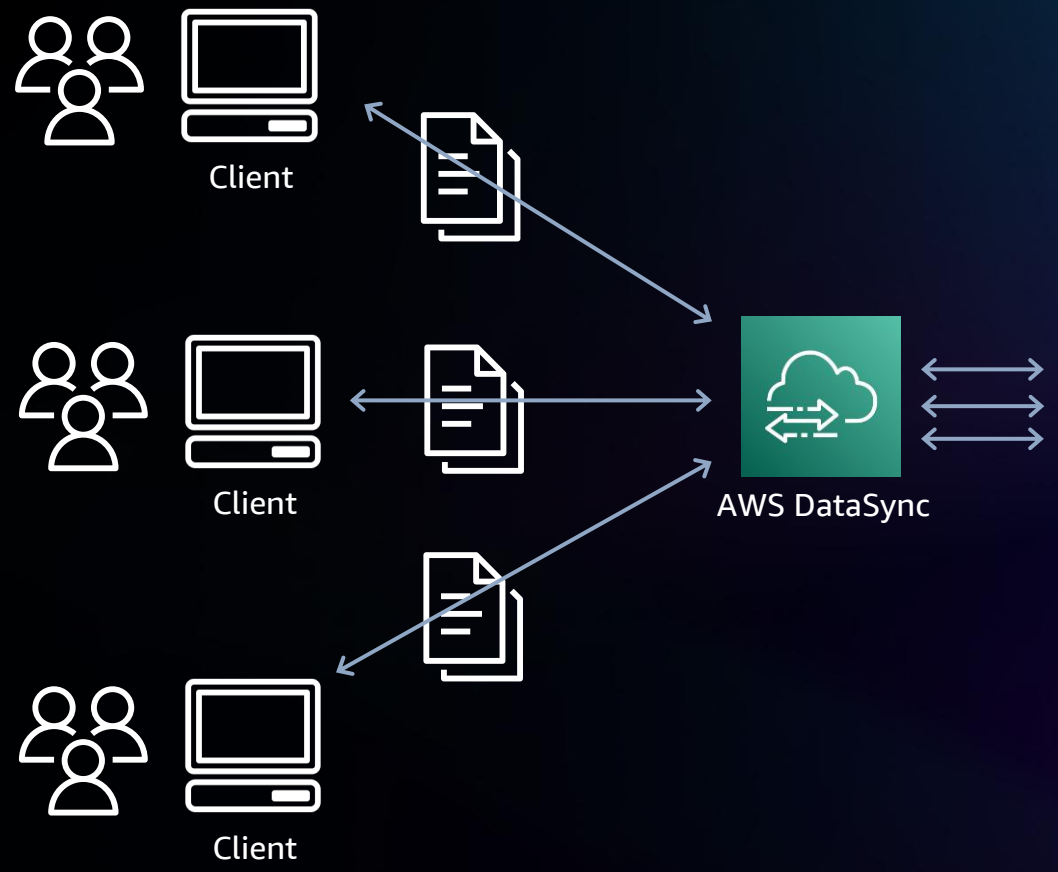
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon S3 is the best place to build **data lakes**



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Amazon Simple Storage Service (Amazon S3)

AWS Glue

Amazon Athena

Ad hoc queries to understand the data

Amazon QuickSight

Steady-state dashboards to understand the business

Amazon SageMaker

Deep analysis to inform decisions

Original on-premises applications

Cloud-based data lake

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

DataSync

Discovery Preview
Tasks
Locations
Agents

Migration & Transfer

AWS DataSync

Simplifies, automates, and accelerates moving data

AWS DataSync is an online data transfer service that simplifies, automates, and accelerates moving data between on-premises storage systems and AWS Storage services, as well as between AWS Storage services.

Create a data transfer task

Choose the data transfer you want to set up.

Between on-premises storage and AWS

Get started

How it works

1. **To transfer data between on-premises storage and AWS, deploy an agent**
   Deploy a DataSync agent and associate it to your AWS account. Deploying an agent is not required to transfer data between AWS Storage services within the same AWS account.
2. **Create a task**
   Create a task by specifying the location of your data source and destination, and any options you want to use to configure the transfer.
3. **Start the transfer**
   Start the task and monitor data movement in the console or with Amazon CloudWatch.

Pricing

There are no resources to manage, no upfront costs, and no minimum fees. You pay $0.0125 per gigabyte for the amount of data that you copy. In addition, you are charged standard request and data transfer rates to read to and write from your data stores.

Learn more about DataSync pricing

Additional information

Documentation
API reference
FAQs
Support forums
Primer course

Benefits and features

Simplify and automate transfers

AWS DataSync makes it easy to move data over the network between on-premises storage and AWS Storage services, as well as between AWS Storage services. DataSync automates both the management of data transfer processes and the infrastructure required for high-performance, secure data transfer.

Transfer data securely

DataSync provides end-to-end security, including encryption and integrity validation, to make sure your data arrives securely, intact, and ready to use. DataSync accesses your AWS Storage using built-in AWS security mechanisms such as IAM roles. It also supports VPC endpoints, providing you with the option to transfer data without

© 2022, Amazon Web Services, Inc. or its affiliates.

Amazon S3

Buckets
Access Points
Object Lambda Access Points
Multi-Region Access Points
Batch Operations
Access analyzer for S3

Block Public Access settings for this account

Storage Lens
Dashboards
AWS Organizations settings

Feature spotlight 3

AWS Marketplace for S3

donor-info/

Copy S3 URI

Objects | Properties

Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

Copy S3 URI | Copy URL | Download | Open | Delete | Actions | Create folder | Upload

Find objects by prefix

1

Name	Type	Last modified	Size	Storage class
donor-info.csv	csv	November 18, 2022, 23:14:16 (UTC-05:00)	1.5 MB	Standard

© 2022, Amazon Web Services, Inc. or its affiliates.

AWS Glue

Data Catalog
Databases New
Tables New
Stream schema registries
Schemas
Connections New
Crawlers New
Classifiers New
Catalog settings

Data Integration and ETL
AWS Glue Studio
Jobs New
Interactive Sessions
Notebooks New
Data classification tools
Sensitive data detection
Record Matching
Triggers
Workflows
Blueprints
Security configurations

Legacy pages

What's New New

Schema (39)

View and manage the table schema.

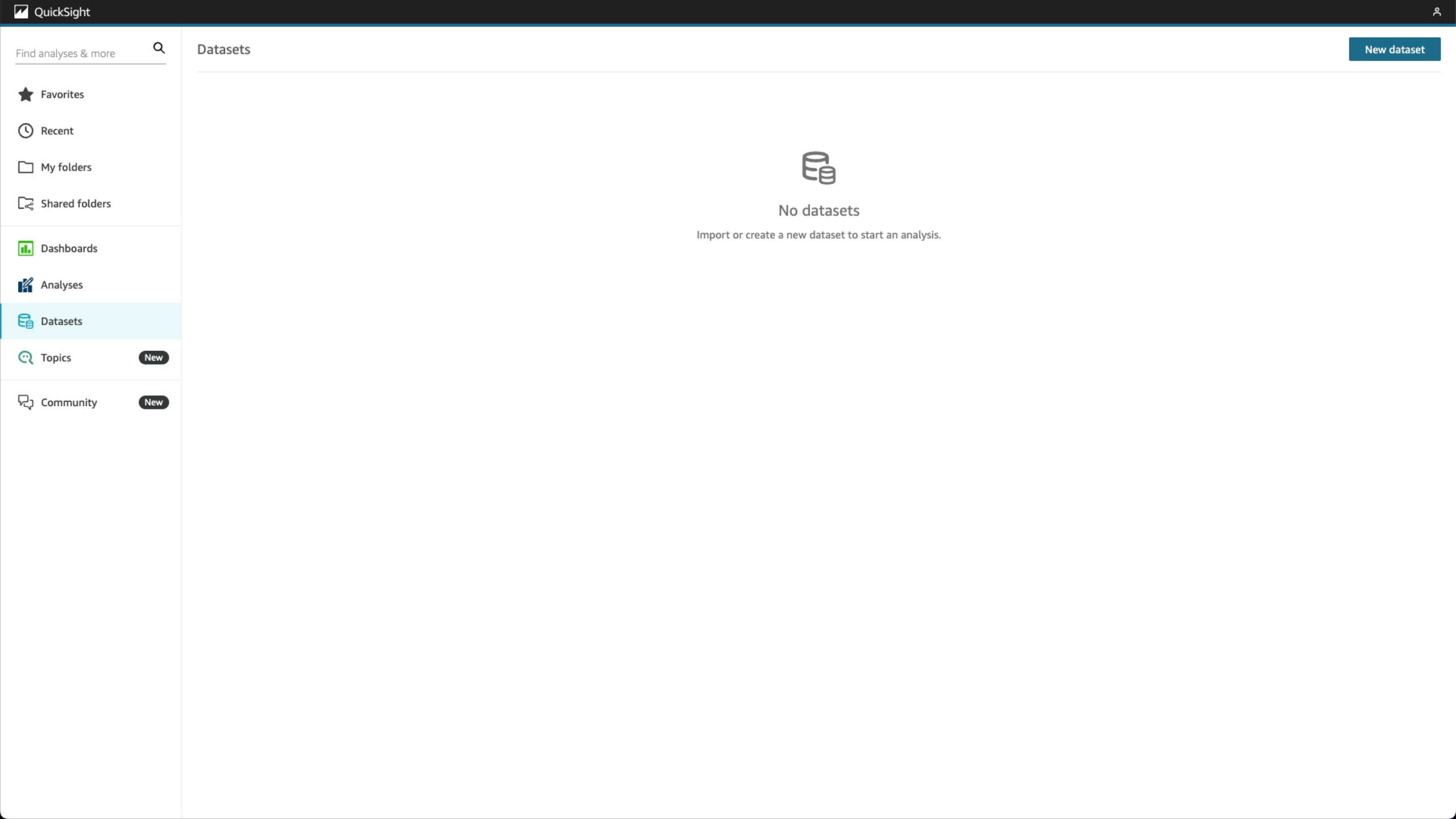
Edit schema

Filter schemas

1 2

#	Column name	Data type	Partition key	Comment
1	first_name	string	-	-
2	occupation	string	-	-
3	employer	string	-	-
4	race	string	-	-
5	last_gift_date	string	-	-
6	hours-per-week	bigint	-	-
7	last_gift_usd	double	-	-
8	donation_total_usd	double	-	-
9	donation_date	string	-	-
10	donor-funraising-id	bigint	-	-
11	latitude	double	-	-
12	donor-demographics-id	bigint	-	-
13	job_title	string	-	-
14	amount	bigint	-	-
15	married	boolean	-	-
16	workclass	string	-	-
17	donor-census-id	bigint	-	-
18	dob	string	-	-
19	education_level	string	-	-
20	maf_name	string	-	-

© 2022, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences


QuickSight
Find analyses & more
Favorites
Recent
My folders
Shared folders
Dashboards
Analyses
Datasets
Topics
New
Community
New
Datasets
New dataset
No datasets
Import or create a new dataset to start an analysis.

QuickSight | donor_info analysis

ADD | AUTOSAVE ON

Visualize | Filter | Insights | Parameters | Actions | Themes | Settings | Community

Forecast properties

Forecast length

Periods forward: 6

Periods backward: 6

Prediction interval: 90

Seasonality

- Automatic
- Custom - Min:1 - Max:180

1

Forecast boundaries

- Forecast maximum
- Forecast minimum

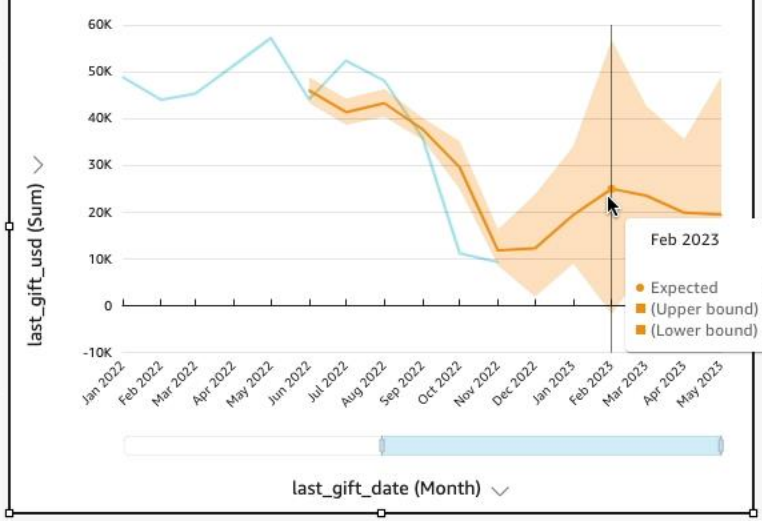
RESET ALL

APPLY | REMOVE

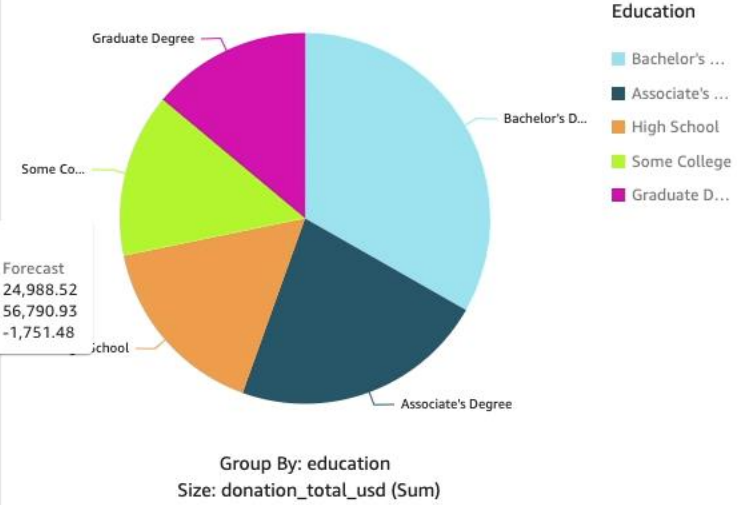
Field wells | X axis: last_gift_date (MONTH) | Value: last_gift_usd (Sum) | Color | Small multiples

Sheet 1

Sum of Last_gift_usd by Last_gift_date



Sum of Donation_total_usd by Education



Group By: education
Size: donation_total_usd (Sum)

Sum of Donation_total_usd by Last_gift_date

2022	2021
5,051,679.5	4,458,493.53

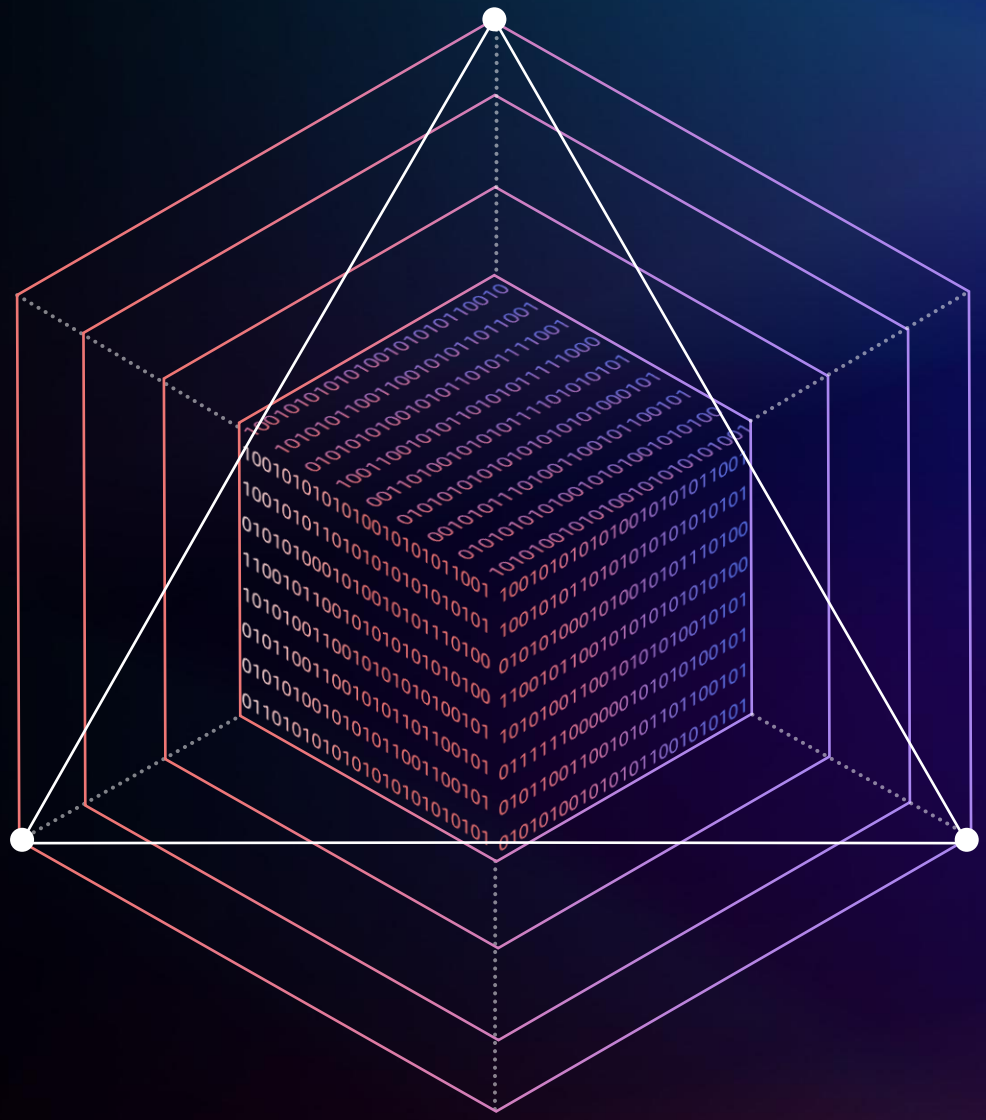
593,185.97 ↑

Sum of Amount by Postal_code

SHOWING TOP 3213 IN POSTAL_CODE



Design for agility



Plan for the unexpected

DYNAMICALLY MANAGE GROWTH

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



"FINRA, the regulatory body of the US securities industry, had its storage footprint increase by several fold over a three-year period (January 2019 to January 2022). During this same time FINRA was able to reduce its unit storage cost by ~50% by taking advantage of innovations in storage like Intelligent Tiering, with its automated discounting based on access, and Glacier Instant Retrieval, for archive storage cost with immediate retrieval."

Peter Boyle
Technology, VP, FINRA

THE WALL STREET JOURNAL.

CIO JOURNAL

Airbnb Details Road Map to Lower Cloud Costs

Nov. 7, 2022



Posted its most profitable quarter to date with $2.88 billion in revenue for the quarter ended Sept. 30, 2022

Moved data to Amazon S3 Glacier and replaced homegrown data backup system with AWS Backup

$63.5 MILLION

savings in 2020, compared to 2019

"That reduced a significant amount of our storage costs right off the bat."

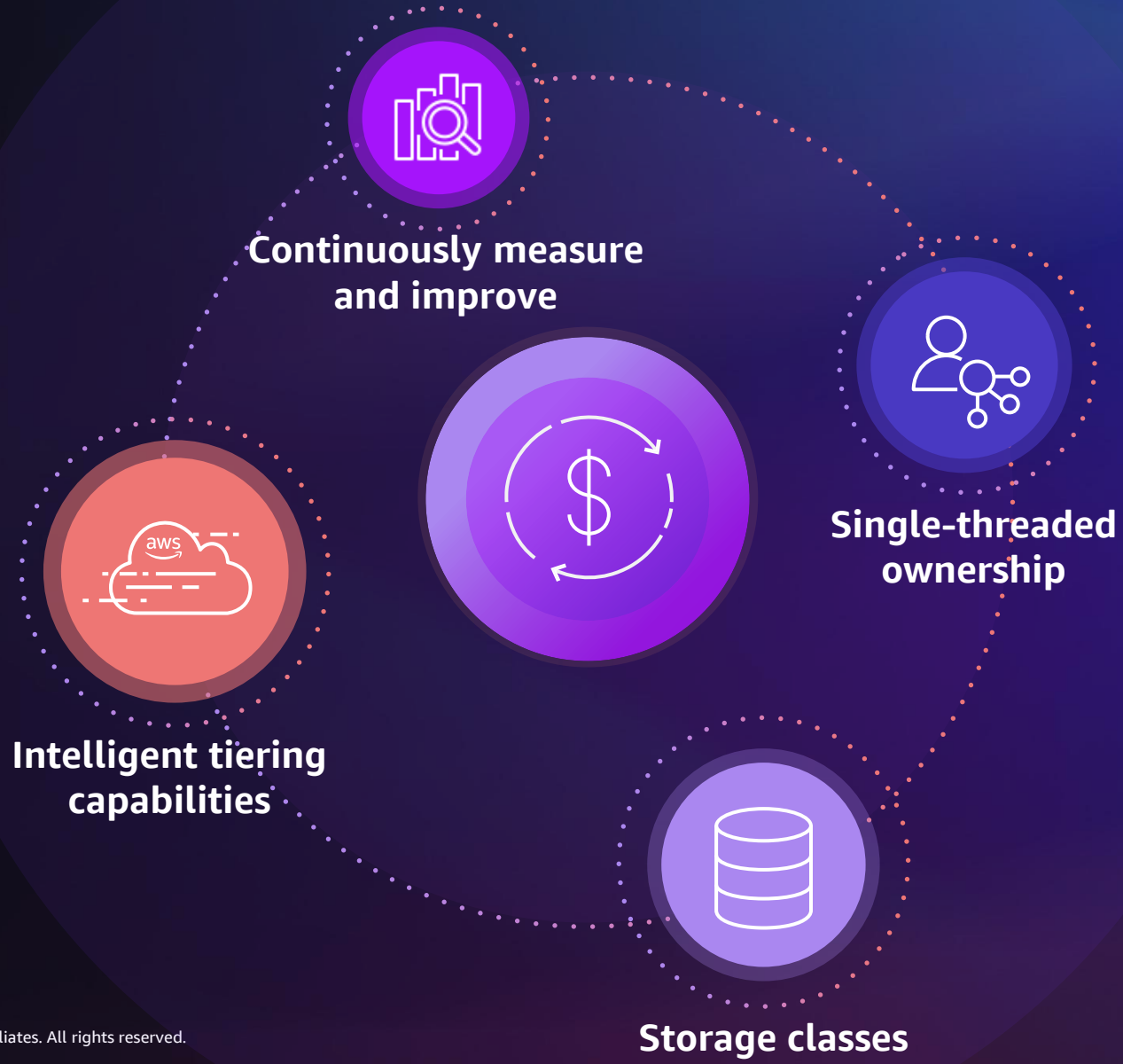
David Nagle
Director of Cloud Infrastructure, Airbnb

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

https://www.wsj.com/articles/airbnb-details-road-map-to-lower-cloud-costs-11667858988

Best practices for cost optimization



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Storage class choice matters at scale

S3 Intelligent-Tiering	S3 Standard	S3 Standard-IA	S3 One Zone-IA	S3 Glacier Instant Retrieval	S3 Glacier Flexible Retrieval	S3 Glacier Deep Archive
						
Changing access patterns	Frequently accessed data	Infrequently accessed data	Re-creatable, less accessed data	Rarely accessed data	Archive data	Long-term archive data

Milliseconds access — Minutes to hours

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 17

AVAILABLE NOW

Amazon S3 Storage Lens added 34 new metrics

ENHANCED CLOUD STORAGE ANALYTICS TO PROVIDE SINGLE VIEW OF S3 USAGE AND ACTIVITY



COST OPTIMIZATION "How do I optimize my storage spend?"



DATA PROTECTION "How do I prevent accidental deletes and overwrites?"



ACCESS CONTROL "Can bucket access for my workloads be improved?"



PERFORMANCE "Am I seeing a rise in 403 authorization errors?"

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 29

COMING SOON

Amazon DataZone

CATALOG, DISCOVER, SHARE, AND GOVERN DATA ACROSS THE ORGANIZATION

DATA PRODUCERS



Set up business data catalog

Populate with metadata, labels, descriptions

Publish catalog

DATA CONSUMERS



Search and browse catalog, request access

Define data projects

Collaboration

aws

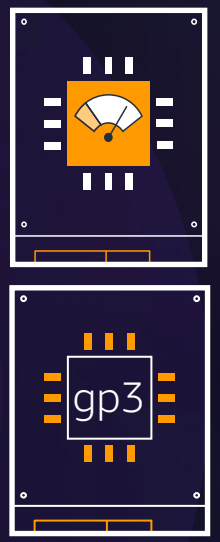
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EBS optimizations in AWS Compute Optimizer

RECOMMENDATIONS BASED ON YOUR VOLUME'S USAGE, AT NO ADDITIONAL CHARGE



AWS Compute Optimizer



io2

Optimize provisioned IOPS settings

gp3

- Right-size gp3 volumes by provisioning IOPS and throughput separately
- IOPS and throughput recommendations for gp3 volumes

Volume ID	Finding Info	Current volume type	Current size	Current IOPS	Current monthly price Info	Recommended volume type Info	Price difference (monthly) Info
vol-0d3f4910f6ad6adcf	Optimized	General Purpose SSD (gp2)	150 GiB	450	$15.000 per month	General Purpose SSD (gp3)	- $3.000 per month
vol-0de34bfdee2a2e611	Not optimized	General Purpose SSD (gp3)	80 GiB	6000	$21.400 per month	General Purpose SSD (gp3)	- $15.000 per month
vol-0c6e2cce321909186	Optimized	General Purpose SSD (gp2)	1000 GiB	3000	$100.000 per month	General Purpose SSD (gp3)	- $20.000 per month

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

J&J deploys Amazon EBS Snapshots Archive to deliver compliance at scale and lower costs

120K
SNAPSHOTS

3PB
OF DATA

$1M+
COST SAVINGS



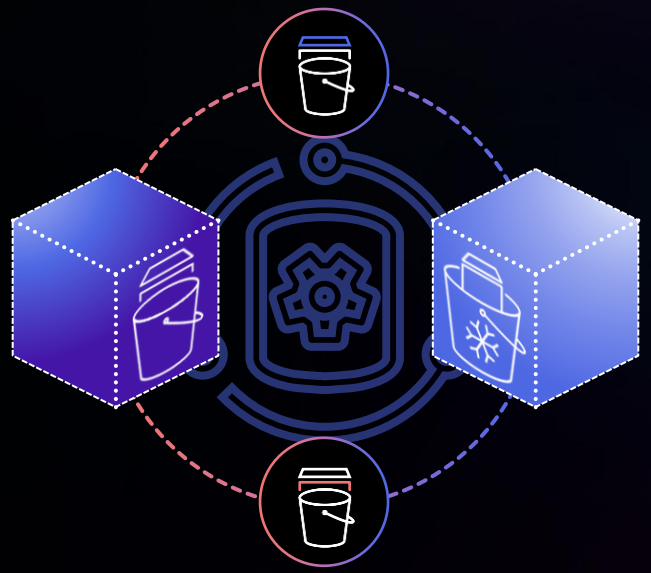
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GA Sept. 30

AVAILABLE NOW

Data Lifecycle Manager for Amazon EBS Snapshots Archive

AUTOMATE THE CREATION AND ARCHIVAL OF EBS SNAPSHOTS TO SNAPSHOTS ARCHIVE



Set policies that automatically create and move monthly, quarterly, and yearly snapshots from standard to archive tier

Meet compliance and business requirements to keep full snapshot backups for more than 90 days

At no additional costs

aws

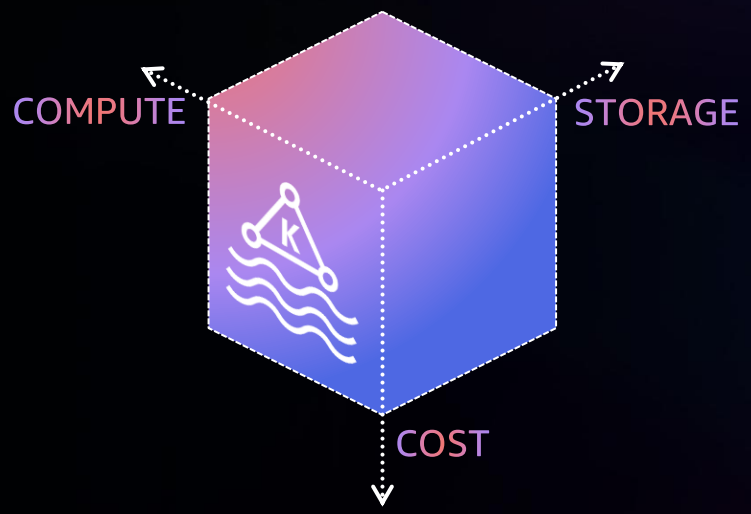
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GA Oct. 28

AVAILABLE NOW

Amazon MSK tiered storage

A LOW-COST STORAGE TIER THAT SCALES TO VIRTUALLY UNLIMITED STORAGE



Increase your topics' storage without needing to increase the number of brokers

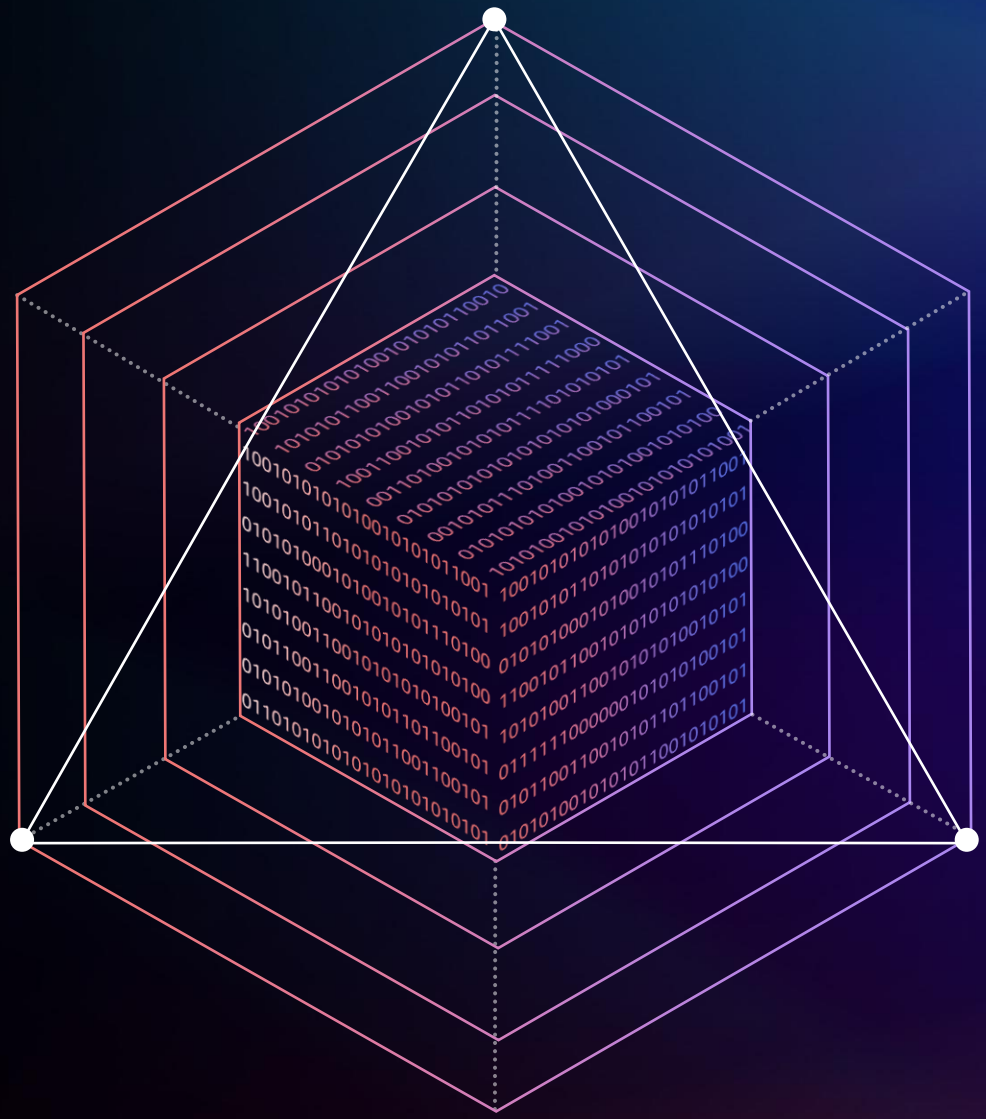
Scale your compute and storage independently

Save with a low-cost storage tier

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Design for agility



Dynamically manage growth

PLAN FOR THE UNEXPECTED

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 29

PREVIEW

Amazon Security Lake

AUTOMATICALLY CENTRALIZE SECURITY DATA INTO A PURPOSE-BUILT DATA LAKE IN A FEW CLICKS



Centralize data automatically from cloud, on-premises, and custom security sources across Regions

Optimize and manage security data for more efficient storage and query performance

Normalize data to an industry standard to easily share and use with multiple analytics tools

Analyze using your preferred analytics tools while retaining control and ownership of your security data

aws

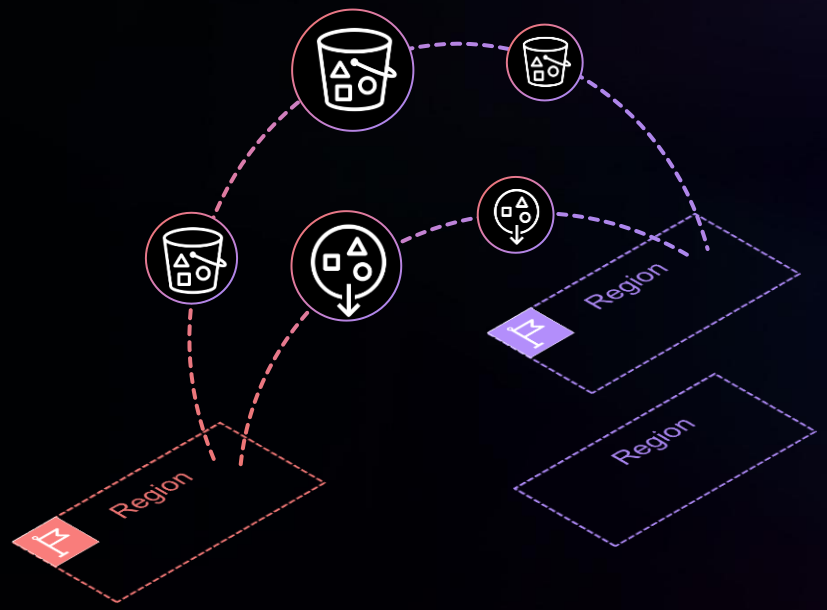
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 28

AVAILABLE NOW

Amazon S3 Multi-Region Access Points failover controls

SUPPORT AN ACTIVE-PASSIVE CONFIGURATION AND CUSTOMER-INITIATED FAILOVER



Perform disaster recovery and application failover testing in an active-passive configuration

Build high availability to maximize uptime for business-critical applications

Automatically replicate S3 objects across AWS Regions you define to maintain copies of your data

Shift S3 data access requests to another AWS Region within minutes

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Leverage pattern matching and machine learning (ML) to protect data



Amazon SNS message data protection

Facilitate real-time data protection in applications that use Amazon SNS to exchange high volumes of data

Discover and protect PII and PHI to support compliance objectives, such as GDPR, HIPAA, PCI-DSS, and FedRAMP

NEW Nov. 3



Amazon CloudWatch Logs data protection

Enable Amazon CloudWatch Logs data protection to detect and mask sensitive log data

Detect and protect sensitive log data in transit

NEW Nov. 27

aws

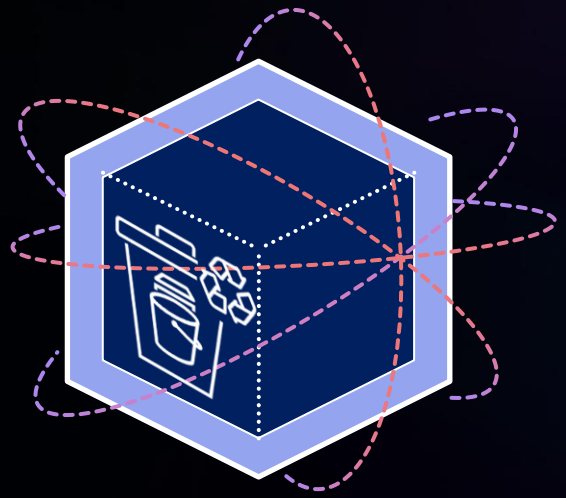
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 23

AVAILABLE NOW

Amazon EBS Rule Lock for Recycle Bin

PROTECT EBS VOLUMES FROM ACCIDENTAL OR UNINTENDED DELETION



Lock retention rules to prevent them from being modified or deleted

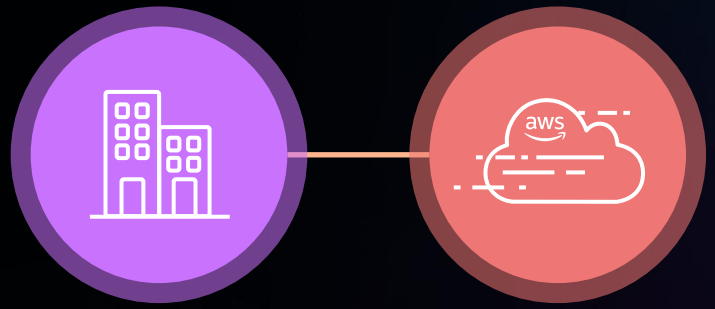
Protect Amazon EBS snapshots and EC2 AMIs from unintended deletions and account access

Inadvertent deleted resources are safely retained in the Recycle Bin and can be recovered later

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Backup expands support of resources



Both on premises and cloud data

Modern apps

Application-aware support for AWS CloudFormation and infrastructure as code

NEW Nov. 27

Modern analytics

Amazon Redshift

NEW Nov. 27

Preview of SAP HANA on Amazon EC2

NEW Nov. 21

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW Nov. 27

AVAILABLE NOW

New compliance capabilities on AWS Backup

AUDIT AND REPORT WITH INSIGHTS ON COMPLIANCE STATUS OF DATA PROTECTION FRAMEWORKS



Cross-Region, cross-account support for organization-wide visibility and management

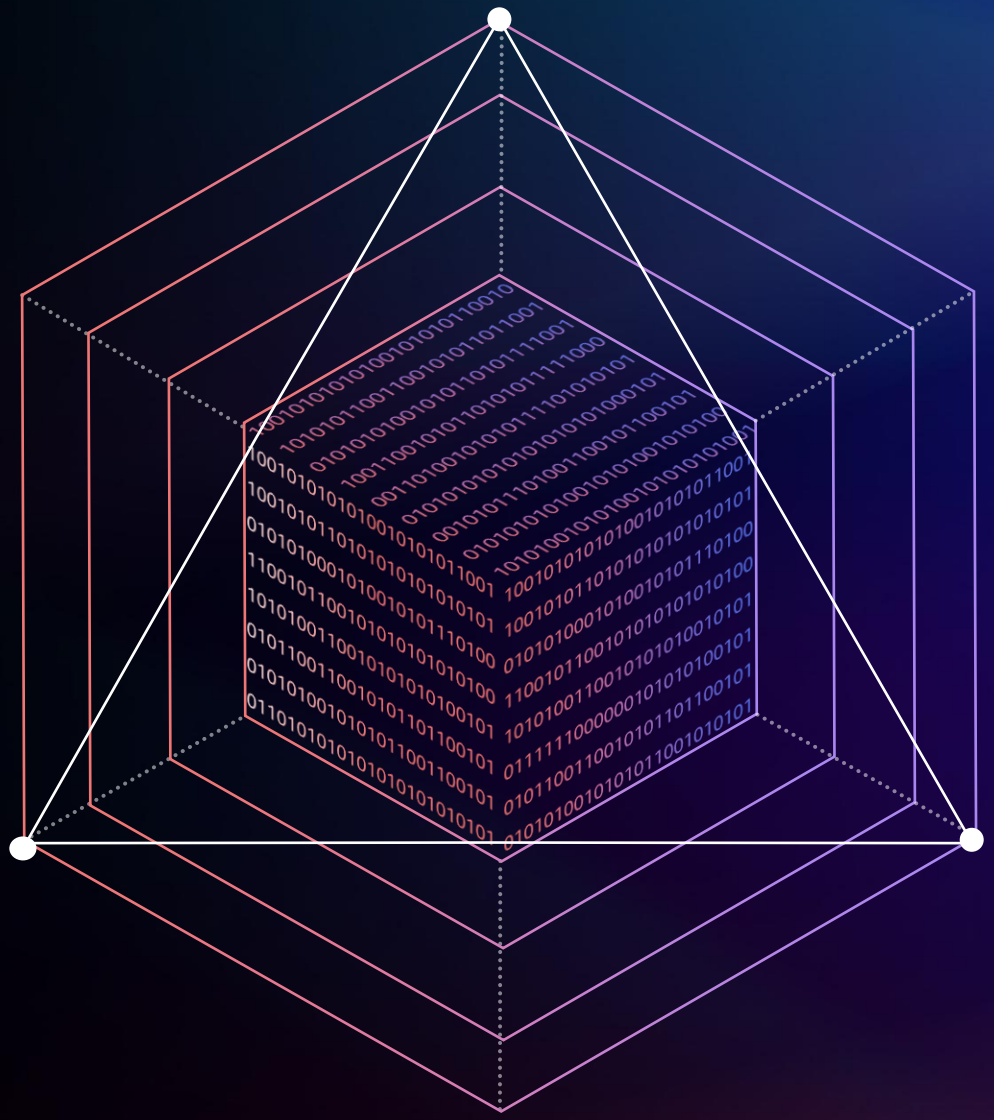
Legal holds prevent deletion of backups for e-Discovery use cases

Cohasset certifies Vault Lock for Sec-17a, FINRA 4511(c), and CFTC 1.31(c)

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Design for agility



Dynamically manage growth

Plan for the unexpected

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.