# AWS Storage Design

CHAPTER 2

# AWS Storage Services

EPISODE 2.01

# Selecting the Storage Service


- Simple Storage Service (S3)
- Glacier
- CloudFront
- Elastic Block Store (EBS)
- Storage Gateway
- Snow family


- Simple Storage Service (S3)
- Glacier
- CloudFront
- Elastic Block Store (EBS)
- Storage Gateway
- Snow family
- Databases

# Block Storage

- Used on local networks
  - iSCSI
  - Fibre Channel
- AWS can use block storage with virtual machines within the AWS cloud using EBS

# File Storage

- AWS uses similar, called object storage in S3
- Used with NAS devices locally

# Selecting Storage

- Size
- Performance
- Cost

# S3 Storage Class

EPISODE 2.02

# S3 Storage Overview

- Object storage
- Distributes across at least three Availability Zones
  - Except: 1A (1 zone, least expensive)

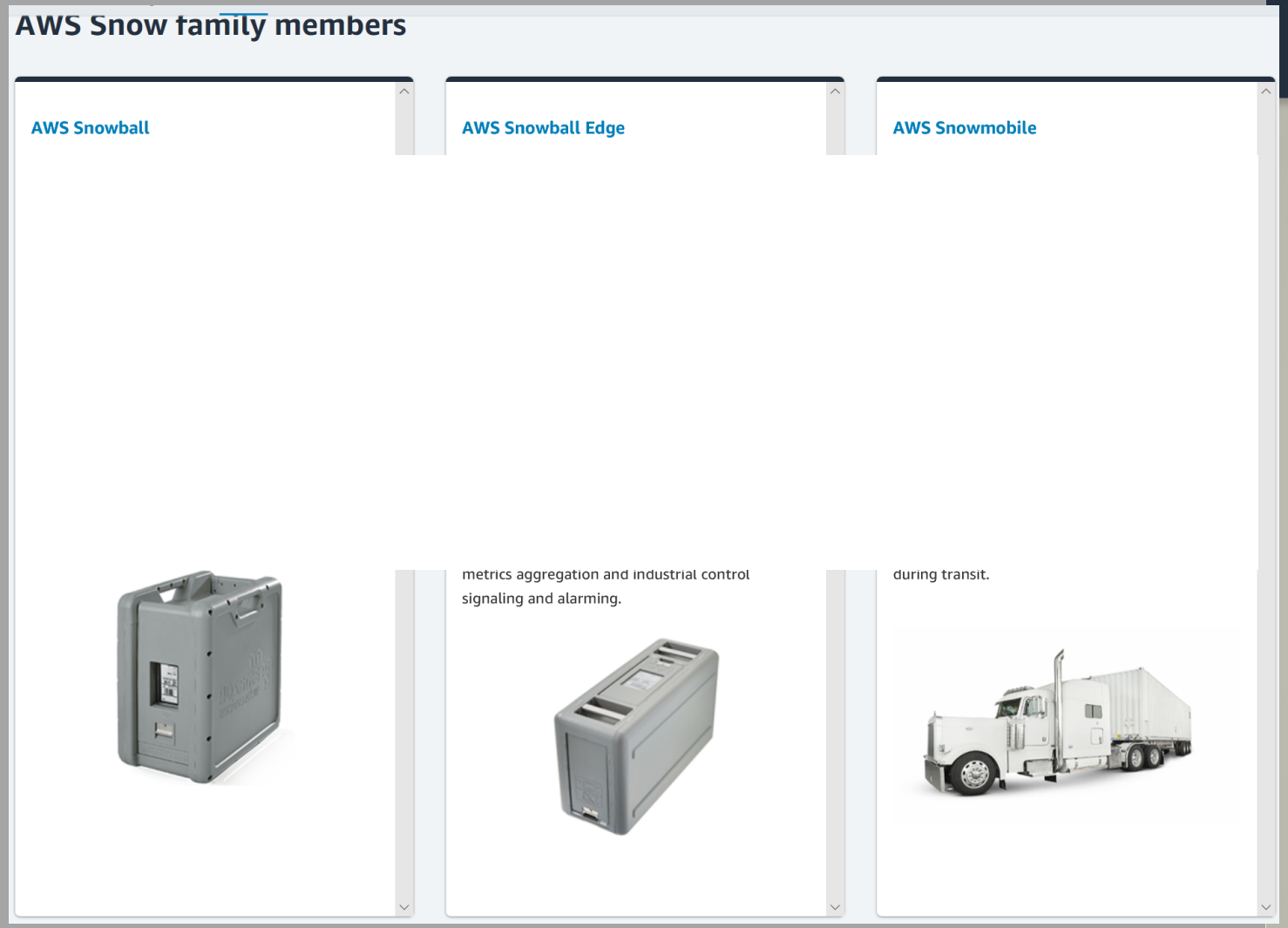
# S3 Storage Overview

- Supports encryption and automatic data classification
- Big data analytics can run directly against stored data

# Getting Data into S3

- API (Application Programming Interface)
- Amazon Direct Connect
- Storage Gateway
- Kinesis Firehose
- Transfer Acceleration

# Getting Data into S3




- Snow family
  - Snowball
  - Snowball Edge
  - Snowmobile

https://aws.amazon.com/snow/

# S3 Terminology

EPISODE 2.03

# S3 Concepts

- Buckets
- Regions
- Objects
- Keys
- Object URLs
- Eventual consistency
- Works great for static website hosting

# S3 Concepts

- Works great for static website hosting

# Common S3 Operations

- Creating and deleting buckets
- Writing objects
- Reading objects
- Deleting objects
- Managing object properties
- Listing keys in buckets

# REST Interface

- Representational State Transfer (REST)
  - S3 API
  - Maps HTTP methods to CRUD operations
    - **C**reate uses PUT or POST
    - **R**ead uses GET
    - **U**pdate uses POST or PUT
    - **D**elete uses DELETE

# S3 Advanced Features

EPISODE 2.04

# S3 Features

- Prefixes and delimiters
- Storage classes
- Object lifecycle management
- Encryption
- Versioning

# S3 Features

- Multi-Factor Authentication (MFA) Delete
- Multi-part upload
- Range GETs
- Cross-Region replication
- Logging
- Event notifications

# Creating S3 Buckets Lab

EPISODE 2.05

# S3 Bucket Properties

EPISODE 2.06

# S3 Managing Objects Lab

EPISODE 2.07

# Glacier

EPISODE 2.08

# Glacier Overview

- Archival data storage
- Fractions of a penny per GB/month
- Three access methods
  - Expedited (3-5 minutes)
  - Standard (3-5 hours)
  - Bulk (5-12 hours)

# Glacier Overview

- You define the Region for data storage
- Data stored with AES 256-bit encryption

# Glacier Integration

- S3 cold data can be automatically moved into Glacier
- Snow devices can be used to import data
- Storage Gateway can connect to Glacier

# Glacier Concepts

- Archives
- Vaults
- Vault locks
- Data retrieval
  - Up to 5% retrieved at no charge, no rollover
  - Vault can be configured to limit costs

# Setting up a Glacier Vault Lab

EPISODE 2.09

# Elastic Block Store (EBS)

EPISODE 2.10

# EBS Overview

- Used for durable storage in EC2 instances
- Block-level storage from one AWS service to another

# EBS Volume Types

- Magnetic
- SSD (solid-state drive)
  - General purpose
  - Provisioned IOPS
    - PIOPS (provisioned input/output operations per second)
  - EBS-optimized instance should be used

# Protecting EBS Data

- Snapshots
- Volume recovery
  - Attaching volumes from one instance to another
- Encryption methods

# Creating EBS Volumes Lab

EPISODE 2.11

# Elastic File System (EFS)

EPISODE 2.12

# EFS Overview

- Shareable
- Hierarchical
- Can be accessed through NFSv4
  - EBS volumes are dedicated to an instance

# EFS Overview

- EC2 instances can use EFS shares
- EFS is not supported on Windows instances

# Storage Comparison

| | | **File**<br>Amazon EFS | **Object**<br>Amazon S3 | **Block**<br>Amazon EBS |
|---|---|---|---|---|
| **Performance** | Per-operation latency | Low, consistent | Low, for mixed request types, and integration with CloudFront | Lowest, consistent |
| | Throughput scale | Multiple GBs per second | Multiple GBs per second | Single GB per second |
| **Characteristics** | Data Availability/Durability | Stored redundantly across multiple AZs | Stored redundantly across multiple AZs | Stored redundantly in a single AZ |
| | Access | One to thousands of EC2 instances or on-premises servers, from multiple AZs, concurrently | One to millions of connections over the web | Single EC2 instance in a single AZ |
| | Use Cases | Web serving and content management, enterprise applications, media and entertainment, home directories, database backups, developer tools, container storage, big data analytics | Web serving and content management, media and entertainment, backups, big data analytics, data lake | Boot volumes, transactional and NoSQL databases, data warehousing & ETL |

# Creating an EFS File System Lab

EPISODE 2.13

# Integrating On-Premises Storage

EPISODE 2.14

# AWS Storage Gateway

- Software appliance creates the gateway
- Provides three types of storage solutions:
  - File-based
  - Volume-based
  - Tape-based

# Storage Access Security Lab

EPISODE 2.15

```
{ "Id": "Policy1540923417431", "Version":
"2012-10-17", "Statement": [ { "Sid":
"Stmt1540923412641", "Action": "s3:*",
"Effect": "Allow", "Resource":
"arn:aws:s3:::marketing-widget-2018",
"Principal": { "AWS": [
"arn:aws:iam::989745111221:user/AmyThomas"
] } } ] }
```

# Storage Performance

EPISODE 2.16

# Gibibyte vs. Gigabyte

| Decimal Name | Decimal Abbr. | Decimal Power | Decimal Value | Binary Name | Binary Abbr. | Binary Power | Binary Value |
|---|---|---|---|---|---|---|---|
| Kilobyte | kB | 10^3 | 1,000 | Kibibyte | kiB | 2^10 | 1,024 |
| Megabyte | MB | 10^6 | 1,000,000 | Mebibyte | MiB | 2^20 | 1,048,576 |
| Gigabyte | GB | 10^9 | 1,000,000,000 | Gibibyte | GiB | 2^30 | 1,073,741,824 |
| Terabyte | TB | 10^12 | 1,000,000,000,000 | Tebibyte | TiB | 2^40 | 1,099,511,627,776 |

- EPISODE 2.01
- Amazon Storage Design

# Selecting Storage

- Size
- Performance
- Cost

- EPISODE 2.02
- S3 Storage Class

# REST Interface

- Representational Stat Transfer (REST) is the S3 API
  - Maps HTTP methods to CRUD (Create, Read, Update, Delete) operations
    - Create = HTTP PUT or POST
    - Read = HTTP GET
    - Delete = HTTP DELETE
    - Update = HTTP POST or PUT

- EPISODE 2.03
- S3 Advanced Features

# S3 Features

- Prefixes and delimiters
- Storage classes
- Object lifecycle management
- Encryption
- Versioning

# S3 Features

- Multi-Factor Authentication (MFA) Delete
- Multi-part upload
- Range GETs
- Cross-Region replication
- Logging
- Event notifications

- EPISODE 2.04
- S3 Advanced Features

# S3 Features

- Prefixes and delimiters
- Storage classes
- Object lifecycle management
- Encryption
- Versioning

# S3 Features

- Multi-Factor Authentication (MFA) Delete
- Multi-part upload
- Range GETs
- Cross-Region replication
- Logging
- Event notifications

- EPISODE 2.05
- S3 Lab

- EPISODE 2.06
- Glacier

# Glacier Overview

- Archival data storage
- Fractions of a penny per GB/month
- Three access methods
  - Expedited (3-5 minutes)
  - Standard (3-5 hours)
  - Bulk (5-12 hours)

# Glacier Overview

- You define the region for data storage
- Data stored with AES 256-bit encryption

# Glacier Integration

- S3 cold data can be automatically moved into Glacier
- Snow devices can be used to import data
- Storage Gateway can connect to Glacier

# Glacier Concepts

- Archives
- Vaults
- Vault locks
- Data retrieval
  - Up to 5% retrieved at no charge, no rollover
  - Vault can be configured to limit costs

- EPISODE 2.07
- Glacier Lab

- EPISODE 2.08
  - Elastic Block Store (EBS)

# EBS Overview

- Used for durable storage in EC2 instances
- Block-level storage from one AWS service to another

# EBS Volume Types

- Magnetic
- SSD (solid-state drive)
  - General purpose
  - Provisioned IOPS
    - PIOPS (provisioned input/output operations per second)
  - EBS-optimized instance should be used

# Protecting EBS Data

- Snapshots
- Volume recovery
  - Attaching volumes from one instance to another
- Encryption methods

- EPISODE 2.09
  - EBS Lab

- EPISODE 2.10
  - Elastic File System (EFS)

# EFS Overview

- Shareable
- Hierarchical
- Can be accessed through NFSv4
  - EBS volumes are dedicated to an instance

# EFS Overview

- EC2 instances can use EFS shares
- EFS is not supported on Windows instances

- EPISODE 2.11
  - EFS Shares Lab

- EPISODE 2.12
- Integrating On-Premises Storage

# AWS Storage Gateway

- Software appliance creates the gateway
- Provides three types of storage solutions:
  - File-based
  - Volume-based
  - Tape-based

- EPISODE 2.13
- Storage Access Control Lab

- EPISODE 2.14
- Storage Performance