

File Storage Services



David Tucker

TECHNICAL ARCHITECT & CTO CONSULTANT

@_davidtucker_ davidtucker.net

AWS File Storage and Data Transfer Services



Amazon S3



**Amazon S3
Glacier**



**Amazon Elastic
Block Store**



**Amazon Elastic
File System**



AWS Snowball



**AWS
Snowmobile**

Overview

Reviewing the storage services on AWS

Examining Amazon S3 and its capabilities

Implementing a static website on Amazon S3

Exploring archive capabilities with Glacier and Glacier Deep Archive

Reviewing EC2 storage with EBS and EFS

Examining large-scale data transfer services into AWS

Amazon S3 Overview

Amazon Simple Storage Service (S3)



Stores files as objects in buckets

Provides different storage classes for different use cases

Stores data across multiple availability zones

Enables URL access for files

Offers configurable rules for data lifecycle

Can serve as a static website host

Amazon S3 Non-archival Storage Classes

S3 Standard is the default storage class and is for frequently accessed data

S3 Intelligent-Tiering will move your data to the correct storage class based on usage

S3 Standard-IA is for infrequently accessed data with the standard resilience

S3 One Zone-IA is for infrequently access data that is only stored in one AZ

**Automatically moves files
based on access**

**Moves between frequent
and infrequent access**

**Same performance as
S3-Standard**

S3 Intelligent Tiering
Storage Class

S3 Lifecycle Policies

Objects in a bucket can transition or expire based on your criteria

Transitions can enable objects to move to another storage class based on time

Expiration can delete objects based on age

Policies can also factor in versions of a specific object in the bucket

S3 Transfer Acceleration

Feature that can be enabled per bucket that allows for optimized uploading of data using the AWS Edge Locations as a part of Amazon CloudFront.

Hosting a Website on Amazon S3

Demo

Creating a new S3 bucket

Uploading objects to an S3 bucket

Accessing object from S3 bucket from URL

Configuring a bucket for website hosting

Glacier and Glacier Deep Archive

Amazon S3 Glacier



Designed for archiving of data within S3 as separate storage classes

Offers configurable retrieval times

Can send files directly or through lifecycle rules in S3

Provides two different storage classes

- S3 Glacier
- S3 Glacier Deep Archive

Amazon S3 Glacier Storage Classes

S3 Glacier

Designed for archival data

**90 day minimum storage duration
change**

**Can be retrieved in either minutes or
hours**

You pay a retrieval fee per GB retrieved

**Over 5 times less expensive than S3
Standard storage class**

S3 Glacier Deep Archive

Designed for archival data

**180 day minimum storage duration
change**

Can be retrieved in hours

You pay a retrieval fee per GB retrieved

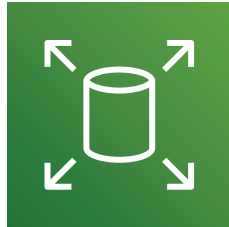
**Over 23 times less expensive than S3
Standard storage class**

“The AWS Management console can be used to quickly set up Amazon S3 Glacier. Data can then be uploaded and retrieved programmatically.”

Amazon Web Services

Elastic Block Store

Amazon EC2 File Storage Services



Amazon EBS

**Persistent block
storage for use with
Amazon EC2**



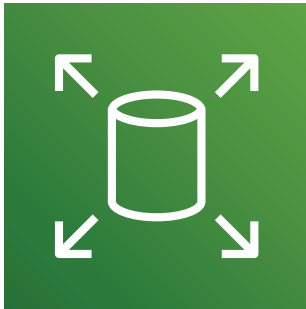
Amazon EFS

**Elastic file system for
use with Linux-based
workloads**

Amazon Elastic Block Store (EBS)

Block storage designed to be connected to a single EC2 instance that can scale to support petabytes of data and supports multiple volume types based on need.

Amazon Elastic Block Store (EBS)



Enables redundancy within an AZ

Allows users to take snapshots of its data

Offers encryption of its volumes

Provides multiple volume types

- General purpose SSD
- Provisioned IOPS SSD
- Throughput optimized HDD
- Cold HDD

Amazon EBS Volume Types

General Purpose SSD is a cost effective type designed for general workloads

Provisioned IOPS SSD high performance volume for low latency applications

Throughput Optimized HDD is designed for frequently accessed data

Cold HDD is designed for less frequently accessed workloads

Elastic File System

Amazon Elastic File System (EFS)



Fully managed NFS file system

Designed for Linux workloads

Supports up to petabyte scale

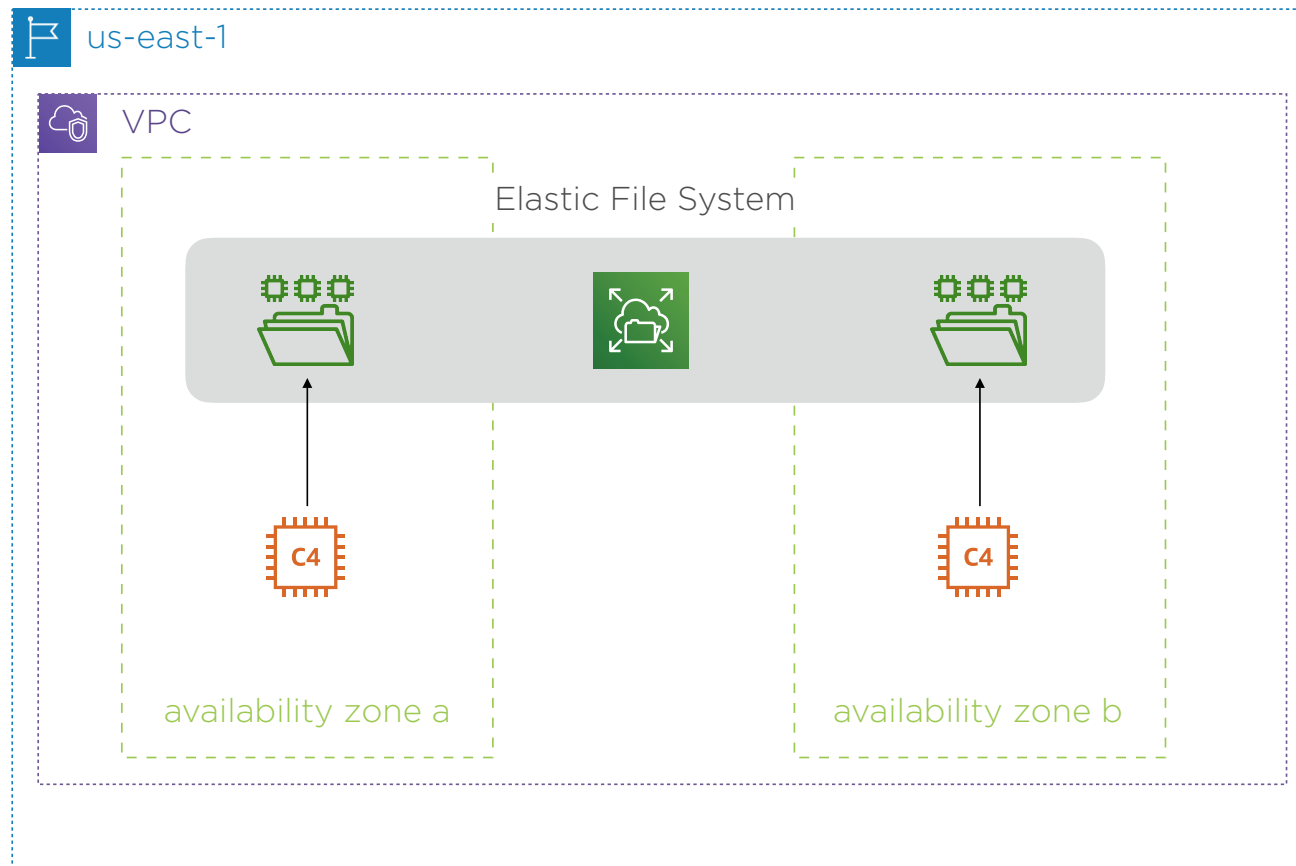
Stores data across multiple AZ's

Provides two different storage classes

- Standard
- Infrequent access

Provides configurable lifecycle data rules

Elastic File System Example



Amazon FSx for Windows File Server



Fully managed native Windows file system

Includes native Windows features including

- SMB support
- Active Directory integration
- Windows NTFS

Utilizes SSD drives for low latency

Data Transfer with AWS Snowball

AWS Large Scale Data Transfer Services



AWS Snowball

Service to physically
migrate petabyte
scale data to AWS



AWS Snowmobile

Service to physically
migrate exabyte scale
data onto AWS

Large-scale Data Transfer into AWS

AWS Snowball

Designed for large-scale data transfer

Supports petabyte scale transfer

Physical device is delivered by AWS

You connect the Snowball to your network and upload your data

Device is returned by local carrier

AWS receives device and loads your data into S3

AWS Snowmobile

Designed for large-scale data transfer

Supports exabyte scale transfer

Ruggedized shipping container is delivered to your location

AWS sets up a connection to your network

You load your data on the Snowmobile

AWS will load data into S3 when the container is received at an AWS location

Scenario Based Review

Scenario 1



Elaine launched a site that offers daily tutorials for developers

She uses S3 to store the assets needed per tutorial

These assets are very popular within the week the tutorial is launched

After this initial week, these assets are rarely accessed

How could Elaine reduce her S3 costs while maintaining durability?

Scenario 2



Esteban works for a social networking company and they are moving to AWS

They have 2 PB of user-generated content that they need to migrate

Esteban is trying to determine if there is a faster than uploading over the internet

Would there be another approach you would recommend for Esteban's company?

Scenario 3



Emily works for a company that produces a messaging app

She is looking for a shared file system between 8 different Linux EC2 instances

The file system would need to support roughly 1 PB of data

What approach would you recommend for Emily?

Summary

Overview

Reviewed the storage services on AWS

Examined Amazon S3 and its capabilities

Implemented a static website on Amazon S3

Explored archive capabilities with Glacier and Glacier Deep Archive

Reviewed EC2 storage with EBS and EFS

Examined large-scale data transfer services into AWS



Scenario 1

Elaine launched a site that offers daily tutorials for developers

She uses S3 to store the assets needed per tutorial

These assets are very popular within the week the tutorial is launched

After this initial week, these assets are rarely accessed

How could Elaine reduce her S3 costs while maintaining durability?

Solution: S3 lifecycle rules with S3-Standard IA storage class

Scenario 2



Esteban works for a social networking company and they are moving to AWS

They have 2 PB of user-generated content that they need to migrate

Esteban is trying to determine if there is a faster than uploading over the internet

Would there be another approach you would recommend for Esteban's company?

Solution: AWS Snowball

Scenario 3



Emily works for a company that produces a messaging app

She is looking for a shared file system between 8 different Linux EC2 instances

The file system would need to support roughly 1 PB of data

What approach would you recommend for Emily?

Solution: Amazon Elastic File System