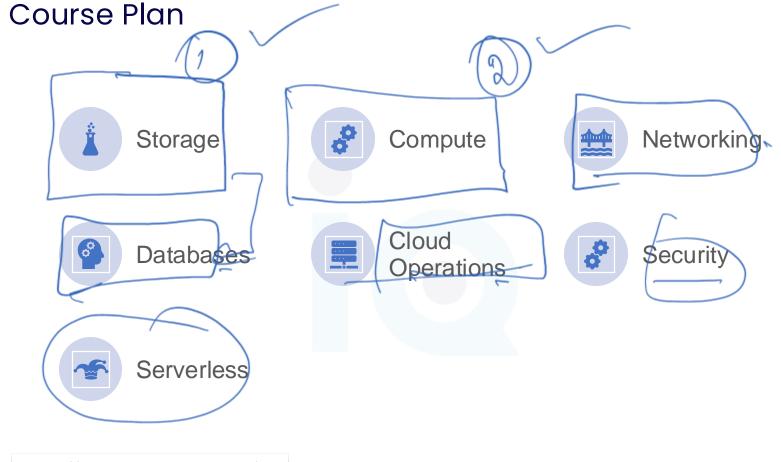
## MLOps

Introduction to
AWS Cloud Computing

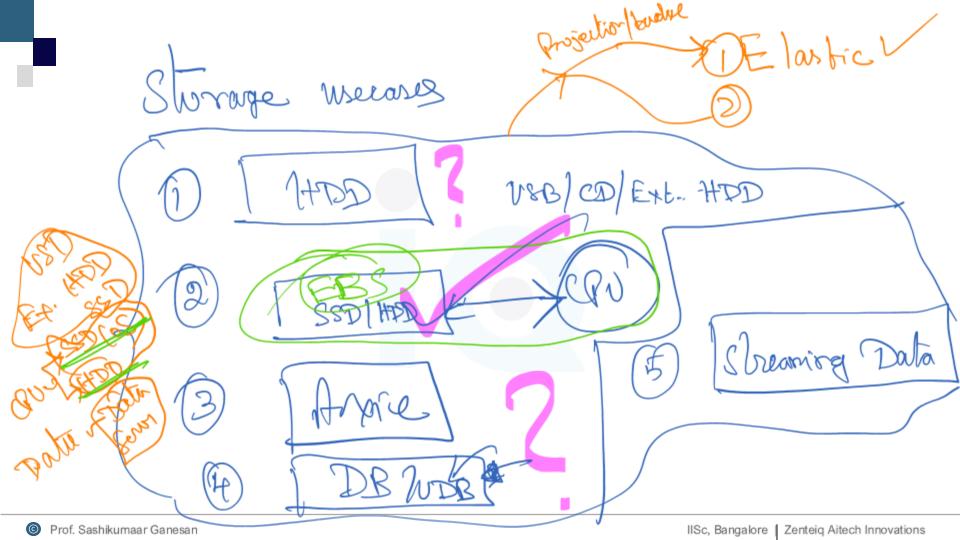




https://www.awseducate.com/

Our Data Center ~ man. 5 yrs Inentment ( huge lost naintainance Thouge list lovace IISc, Bangalore | Zenteig Aitech Innovations

Prof. Sashikumaar Ganesan



Workestaling Briceous Converge Secondary wowasy (perce volutre. Mudle Essert, AWS EBS

## **AWS Storage - Overview**

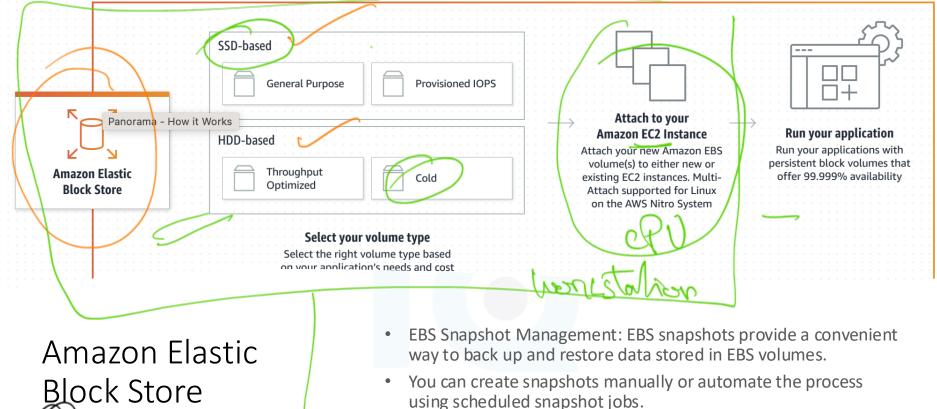
- Overview of AWS Storage Services
  - AWS offers a comprehensive range of storage services designed to meet various data storage and management needs.
  - These services provide scalable, secure, and durable storage options for businesses of all sizes.
- Importance of Cloud Storage
  - Cloud storage eliminates the need for on-premises infrastructure, reducing costs and maintenance efforts.
  - It enables seamless scalability, allowing businesses to quickly adapt to changing storage requirements.
  - Cloud storage provides high availability and durability, ensuring data accessibility and reliability.

### **AWS Storage - Overview**

- Benefits of Using AWS Storage Services:
  - Flexible and Scalable: AWS storage services can scale up or down based on demand, accommodating business growth.
  - Durability and Reliability: Data stored in AWS storage services is redundantly stored across multiple facilities, ensuring durability and high availability.
  - Security: AWS offers robust security measures to protect data, including encryption, access controls, and compliance certifications.
  - Cost-Effective: Pay-as-you-go pricing models and options for optimizing storage costs make AWS storage services cost-effective.

Amazon Elastic Block Store (EBS)





- Snapshots can be used to create new volumes or restore volumes to a previous state.



Amazon Elastic Block Store (EBS) Change

Amazon EBS

- Features and Use Cases of EBS
  - Block-Level Storage: EBS provides durable, block-level storage volumes that can be attached to EC2 instances.
  - Data Persistence: EBS volumes retain data even after the associated EC2 instance is terminated or stopped.
  - High Performance: EBS volumes offer low-latency and high-IOPS options to meet demanding application requirements.

## Amazon Elastic Block Store (EBS)

- **Snapshot Backup**: EBS supports point-in-time snapshots, allowing you to create backups for data protection and disaster recovery.
- **Elasticity**: EBS volumes can be resized on-the-fly, allowing you to dynamically adjust storage capacity as needed.
- Creating and Attaching EBS Volumes
  - Step 1: Access the AWS Management Console and navigate to the EC2 service.
  - Step 2: Launch an EC2 instance or select an existing instance.
  - Step 3: Choose the instance and click on "Actions" -> "Attach Volume."
  - Step 4: Specify the EBS volume settings, such as volume type, size, and availability zone.
  - Step 5: Click "Attach" to attach the EBS volume to the selected EC2 instance.





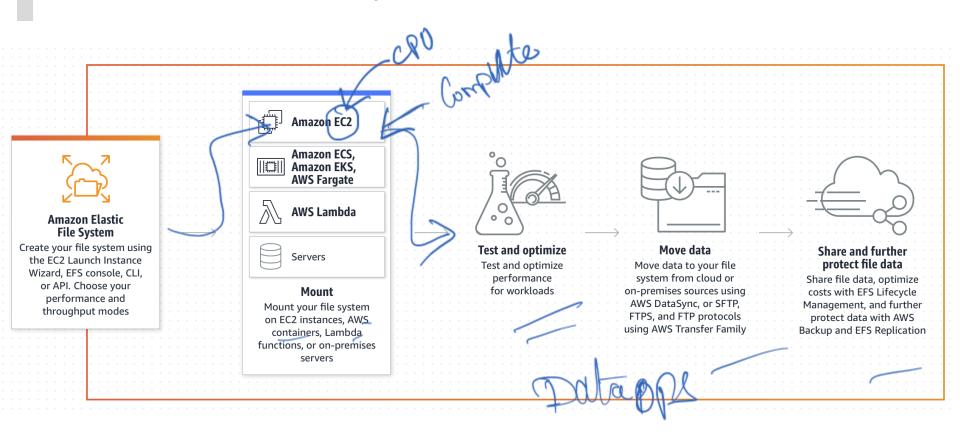
#### Amazon EFS

- Amazon EFS is a scalable and fully managed file storage service that provides shared access to file-based data for multiple EC2 instances.
- It is designed to be highly available, durable, and scalable to support a wide range of workloads and applications.
- Key Features and Benefits of EFS
  - Elastic Scalability: EFS can scale automatically to accommodate the growth of data and file systems.
  - **Shared Access**: EFS supports simultaneous access from multiple EC2 instances and on-premises servers, enabling collaboration and data sharing.
  - File System Performance: EFS provides consistent and low-latency performance, even with high levels of concurrent access.



- Durability and Availability: EFS stores data redundantly across multiple availability zones, ensuring high durability and availability.
- Data Backup and Recovery: EFS supports point-in-time backups and easy recovery options for data protection.
- Use Cases for EFS:
  - Content Management: EFS is suitable for storing and serving web content, media files, and other content management applications.
  - **Big Data Analytics**: EFS can be used for big data analytics workloads that require shared access to large datasets.
  - Development and Testing: EFS provides a central storage location for development and testing environments shared by multiple instances.
  - Container Storage: EFS is an ideal storage solution for containerized applications running on AWS ECS or Kubernetes.

- Creating and Mounting an EFS File System:
  - Step 1: Access the AWS Management Console and navigate to the EFS service.
  - Step 2: Click on "Create file system" and specify the configuration details, such as the storage class and performance mode.
  - Step 3: Configure the security group and network settings for accessing the EFS file system.
  - Step 4: Mount the EFS file system on EC2 instances using standard NFSv4 commands or utilities.





- AWS Backup is a fully managed backup service that simplifies the process of protecting data across AWS services and on-premises environments.
- It centralizes and automates backup tasks, allowing organizations to efficiently manage data protection and recovery.
- Key Features of AWS Backup:
  - Centralized Backup Management: AWS Backup provides a single interface to manage and monitor backups across multiple AWS services.
  - Policy-Based Automation: It enables you to create backup policies that automatically schedule and manage backup tasks.
  - Application Consistency: AWS Backup ensures application-consistent backups for services such as Amazon RDS, EFS, and DynamoDB.

- Incremental Backups: It performs incremental backups, reducing backup time and storage costs.
- Cross-Region and Cross-Account Backups: AWS Backup supports backups across AWS regions and accounts for data protection and compliance.
- Supported AWS Services:
  - Amazon ÉBS: AWS Backup allows you to create point-in-time backups of your Amazon Elastic Block Store volumes.
  - Amazon RDS: It supports automated backups and database snapshots for Amazon RDS instances.
  - Amazon EFS: AWS Backup provides backup and restore capabilities for Amazon Elastic File System.
- DynamoDB: It enables backup and restore of DynamoDB tables, Prof. Sashikumaar Ganesan Point-in-time recovery.



- AWS Storage Gateway: AWS Backup integrates with Storage Gateway to manage backups of on-premises data.
- Benefits of AWS Backup:
  - Simplified Data Protection: AWS Backup streamlines the backup process, reducing the complexity of managing data protection.
  - Centralized Visibility: It provides a unified view of backup activities across multiple AWS services and accounts.
  - Automation and Compliance: AWS Backup automates backup tasks and helps meet compliance requirements.
  - Rapid Recovery: It facilitates quick data recovery by enabling point-intime restores and efficient data transfer.



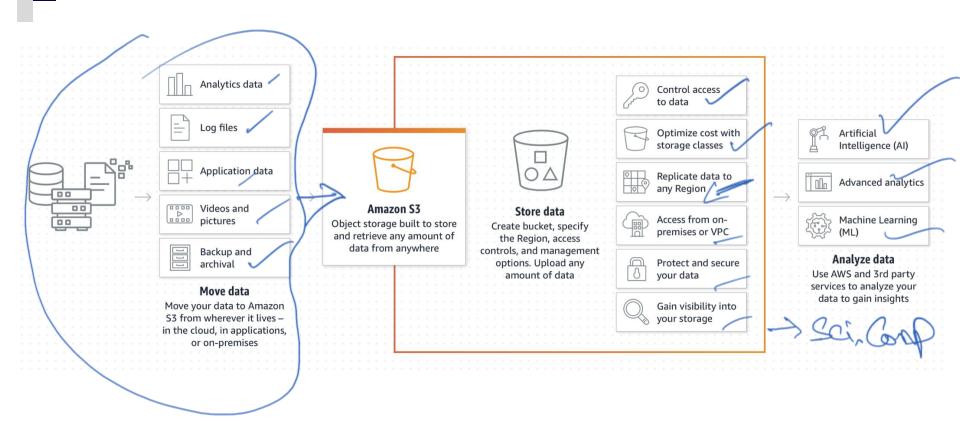


- Amazon S3 is an object storage service offered by AWS, providing secure, durable, and highly scalable storage for various data types.
- It is designed to store and retrieve any amount of data from anywhere on the web.
- Features and Capabilities of S3
  - Scalable Storage: S3 can scale to accommodate any amount of data, from a few gigabytes to petabytes or more.
  - Data Durability and Availability: S3 stores data redundantly across multiple facilities, ensuring high durability and availability.
  - Data Security: S3 supports server-side encryption and provides access control mechanisms to secure data.

- Object Versioning: S3 allows versioning of objects, enabling easy recovery from accidental deletions or overwrites.
- Lifecycle Management: S3 offers lifecycle policies to automatically move objects to less expensive storage classes as they age.
- Use Cases for S3 Storage:
  - Backup and Restore: S3 is ideal for storing backup data and facilitating easy restore operations.
  - Static Website Hosting: S3 can host static websites, providing costeffective and scalable web hosting solutions.
  - Data Archiving: S3's low-cost storage options make it suitable for long-term data archiving and compliance requirements.

- WW Content Distribution: S3 integrates with AWS CloudFront to enable global content distribution with low latency.
- Data lakes and big data analytics: Create a data lake in S3. Extract valuable insights using query-in-place, analytics, and machine learning features.
- Hybrid cloud storage: Create a seamless connection between onpremises applications and S3 with AWS Storage Gateway.
- Creating an S3 Bucket:
  - Step 1: Access the AWS Management Console and navigate to the S3 service.
  - Step 2: Click on "Create Bucket" and provide a unique bucket name and region.
  - Step 3: Configure bucket properties, such as versioning, logging, and encryption options.

(des







#### Amazon S3 Glacier

Cash Cost Township Clarific Cost Township Co

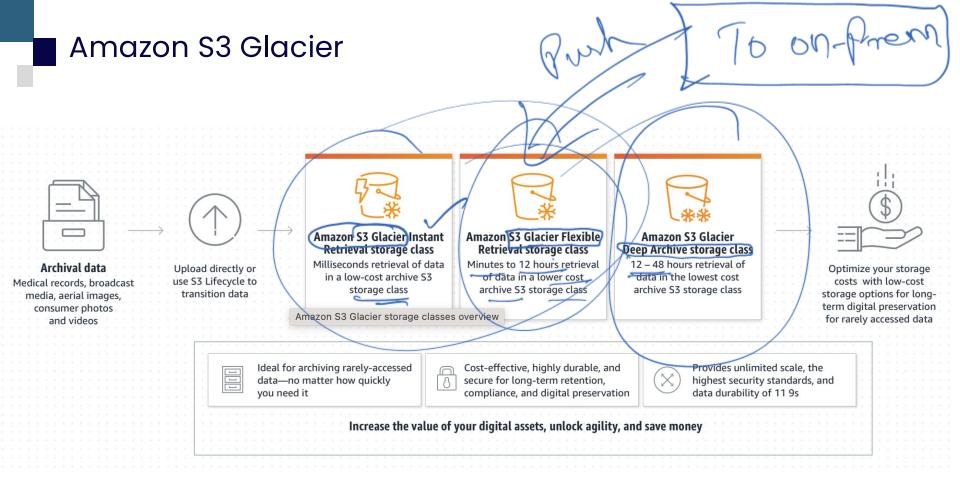
- Introduction to Amazon S3 Glacier:
  - S3 Glacier is a secure, durable, and cost-effective archival storage service
  - It is designed for long-term retention of data that is infrequently accessed but requires long-term durability.
- Overview of Glacier Vaults and Archives:
  - Glacier organizes data storage into vaults, which are containers for holding archives.
  - An archive in Glacier represents a single object or file that is stored within a vault.
  - Each archive in Glacier has a unique identifier and can range in size from a few bytes to several terabytes.

#### Amazon S3 Glacier

- Use Cases for Long-Term Archival Storage
  - Data Archiving and Compliance: Glacier is suitable for storing regulatory data, financial records, and legal documents.
  - Backup and Data Preservation: Glacier offers an affordable option for long-term backup storage of data that is rarely accessed.
  - Digital Media Preservation: Glacier can be used to preserve and protect digital media assets, such as images and videos.
  - Historical Data Retention: Glacier enables organizations to retain historical data for future analysis and reference.

#### Amazon S3 Glacier

- Setting Up and Managing Glacier Vaults:
  - Step 1: Access the AWS Management Console and navigate to the Glacier service.
  - Step 2: Create a new vault by specifying a unique name and optional description.
  - Step 3: Configure vault access policies to manage permissions and access control.
  - Step 4: Monitor and manage your Glacier vaults using the AWS Management Console, APIs, or command-line tools.



hybrid cloud storage

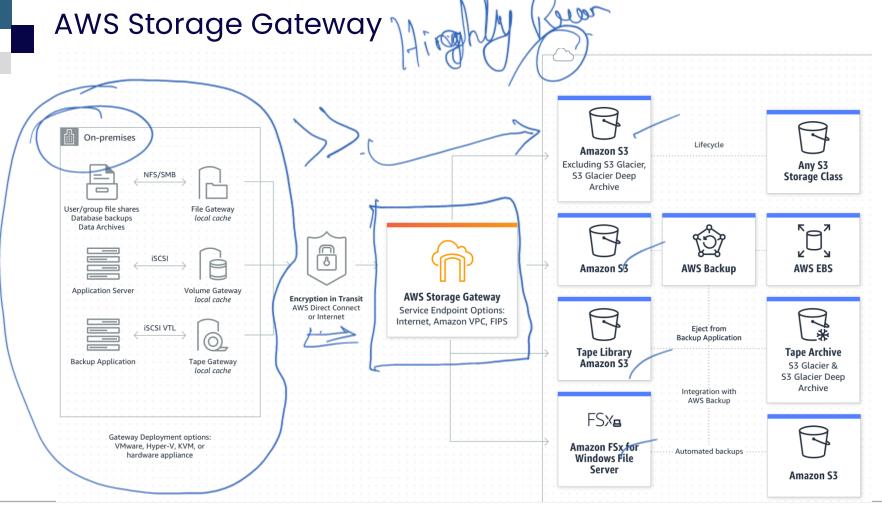
- Introduction to AWS Storage Gateway
  - AWS Storage Gateway is a hybrid cloud storage service that provides seamless integration between on-premises environments and AWS storage services.
  - It enables organizations to extend their on-premises storage infrastructure to the cloud, providing hybrid storage solutions.
- Deployment Models:
  - File Gateway: Allows you to access Amazon S3 objects as files through a Network File System (NFS) or Server Message Block (SMB) interface.
  - Volume Gateway: Presents AWS S3 as iSCSI block volumes, providing low-latency access to frequently accessed data.
- Tape Gateway: Offers a virtual tape library (VTL) interface, allowing Prof. Sashikumaar You to replace physical tape infrastructure with virtual tapes in AW



- Key Features of Storage Gateway:
  - Data Transfer Efficiency: Storage Gateway provides data deduplication, compression, and encryption to optimize data transfer to and from AWS.
  - Seamless Integration: Allows transparent access to AWS storage services, enabling easy migration and hybrid cloud storage.
  - Scalability and Resilience: Storage Gateway automatically scales to meet changing demands and provides high availability with data replication.
  - Data Protection and Backup: Supports automated backup and recovery processes for on-premises data using AWS storage services.
  - Hybrid Cloud Workflows: Enables seamless integration with other AWS services, such as AWS Lambda and Amazon CloudWatch, for building hybrid cloud workflows.

- Use Cases for Storage Gateway:
  - **Hybrid Cloud Storage**: Use Storage Gateway to extend on-premises storage to AWS, enabling hybrid cloud architectures and data tiering.
  - **Disaster Recovery**: Leverage Storage Gateway for replicating onpremises data to AWS for disaster recovery purposes.
  - Backup and Archiving: Utilize Storage Gateway for seamless backup and archiving of on-premises data to AWS storage services.
  - **Cloud Bursting**: Enable bursting workloads to AWS storage for temporary capacity requirements, optimizing on-premises resources.





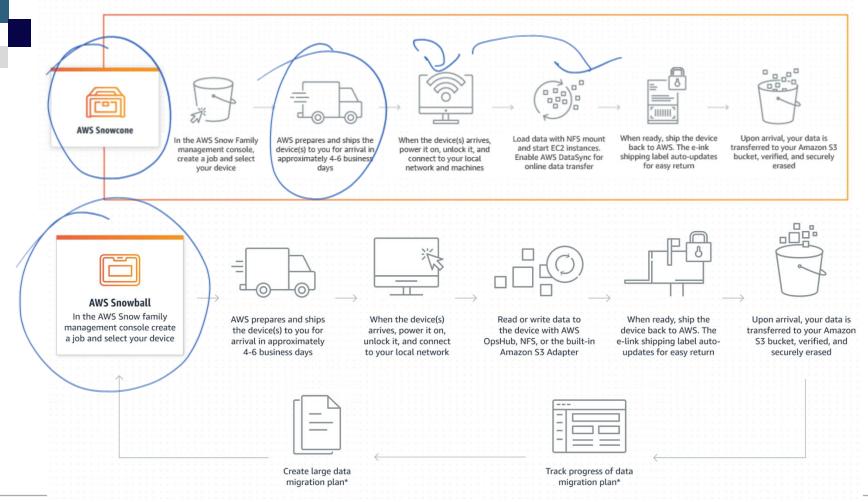
# AWS Snow Family



#### **AWS Snow Family**

- Introduction to AWS Snow Family:
  - The AWS Snow Family consists of physical devices that help overcome challenges associated with transferring large amounts of data to and from AWS.
  - These devices provide secure, durable, and efficient offline data transfer options, particularly for environments with limited internet connectivity.
- Available Devices:
  - AWS Snowcone: A small, rugged, and portable device designed for edge computing and data migration in harsh environments.
  - AWS Snowball: A larger device with built-in compute and storage capabilities, designed for large-scale data migration or offline data processing.







#### **AWS Storage Gateway**

- Use Cases for the Snow Family:
  - Data Migration: Use Snow devices to migrate large datasets, including petabytes of data, to and from AWS without relying solely on internet transfers.
  - Edge Computing: Deploy Snowcone at the edge to run applications, collect data, and perform local processing before transferring it to AWS.
  - Content Distribution: Utilize Snowball for distributing large media files, software updates, or IoT data to remote locations or customer sites.
  - **Disaster Recovery**: Leverage Snowball for transferring backup data to AWS for disaster recovery purposes in case of on-premises failures.

#### **AWS Storage Gateway**

- Process of Using Snow Devices:
  - Step 1: Request a Snow device from the AWS Management Console or API.
  - Step 2: AWS ships the device to your location, and you connect it to your network.
  - Step 3: Transfer data to the device using the provided storage or compute capabilities.
  - Step 4: Ship the device back to AWS, where the data is securely transferred into your AWS account.
  - Step 5: The data becomes available in your AWS environment for further processing or storage.



- Introduction to AWS Storage Classes:
  - AWS offers a range of storage classes designed to meet different performance, durability, and cost requirements.
  - Each storage class provides unique features and benefits to optimize storage costs and performance.
- Storage Class Options:
  - Amazon S3 Standard: Offers high durability, availability, and performance for frequently accessed data. Ideal for active data and real-time analytics.
  - Amazon S3 Intelligent-Tiering: Automatically moves objects between two access tiers (frequent and infrequent access) based on usage patterns.
- Amazon S3 Standard-IA (Infrequent Access): Provides cost-effective
   Prof. Sashikumaar Sanesan ge for data accessed less frequently but requires rapid access Innovation

- Amazon S3 One Zone-IA: Similar to Standard-IA but stores data in a single availability zone, reducing costs further.
- Amazon S3 Glacier: Suitable for archiving and long-term storage of data that is accessed infrequently.
- Retrieval times range from minutes to hours.
- Amazon S3 Glacier Deep Archive: Designed for long-term archival storage with the lowest cost per gigabyte. Retrieval times range from hours to days.

- Use Cases for Each Storage Class:
  - **S3 Standard**: Frequently accessed data, websites, mobile applications, content distribution, and data analytics.
  - Intelligent-Tiering: Data with changing access patterns, backup and restore, disaster recovery, and data lakes.
  - S3 Standard-IA: Backups, long-term storage, infrequently accessed data, and disaster recovery.
  - S3 One Zone-IA: Secondary backup copies, easily re-creatable data, and non-critical workloads.
  - S3 Glacier: Archival storage, regulatory and compliance data, historical records, and large-scale data archiving.

- S3 Glacier Deep Archive: Long-term retention, digital preservation, legal and regulatory compliance, and cold storage.
- Cost Optimization Benefits: Storage tiering and lifecycle policies allow you to move data to the most cost-effective storage class based on its lifecycle.
- By choosing the appropriate storage class, you can optimize costs
  while ensuring data availability and durability. The retrieval times and
  associated costs vary across storage classes, enabling you to balance
  performance needs with cost savings.

## Amazon S3 Features



- Introduction to Amazon S3:
  - Amazon S3 is an object storage service that offers industry-leading scalability, durability, security, and performance.
  - It provides a simple and secure way to store and retrieve any amount of data from anywhere on the web.
- Core Features of Amazon S3:
  - Scalability: S3 can scale to store and retrieve trillions of objects, accommodating any workload or application requirement.
  - Durability: S3 ensures 99.999999999 (11 nines) durability for objects, protecting data against various failure scenarios. (chance of losing a single object is 1 in 100 billion per year)
  - **Security**: S3 supports access control mechanisms, encryption at rest and in transit, and integration with AWS Identity and Access



- **Data Protection**: S3 offers versioning, cross-region replication, and lifecycle policies for data protection and compliance.
- **Performance**: S3 provides high-speed data transfers, low latency access, and support for multipart uploads for large objects.
- Versatility: S3 supports diverse use cases, including backup and restore, data archiving, content distribution, and data lakes.
- Additional Features and Capabilities: Event Notifications: S3 can send event notifications to trigger AWS Lambda functions, SNS notifications, or SQS queues.
- Access Controls: S3 supports fine-grained access controls using bucket policies, access control lists (ACLs), and IAM policies.

- Data Management: S3 offers lifecycle policies, which automatically transition objects between different storage classes based on predefined rules.
- Analytics and Reporting: S3 integrates with AWS services like Amazon Athena and Amazon Macie for data analytics and security monitoring.
- Performance Optimization: S3 Transfer Acceleration utilizes the Amazon CloudFront network to accelerate data transfers to and from S3.

- Use Cases for Amazon S3:
  - **Backup and Restore**: Store and retrieve backups of databases, applications, and file systems in S3.
  - Static Website Hosting: Host static websites using S3's website hosting capabilities.
  - Content Distribution: Use S3 and Amazon CloudFront to distribute content globally with low latency and high transfer speeds.
  - **Big Data and Analytics**: Utilize S3 as a data lake to store and analyze large datasets with services like Amazon EMR and Amazon Redshift.

- Security Features and Measures for AWS Storage Services:
  - Encryption Options: Server-Side Encryption (SSE): AWS storage services provide built-in server-side encryption options to protect your data at rest. SSE-S3, SSE-KMS, and SSE-C are commonly used encryption methods.
    - **SSE-S3** (Server-Side Encryption with S3 Managed Keys): AWS S3 automatically encrypts your data using AES-256 encryption and manages the encryption keys for you.
    - SSE-KMS (Server-Side Encryption with AWS Key Management Service): With SSE-KMS, you have more control over encryption keys, including the ability to manage and audit them using AWS Key Management Service.
    - SSE-C (Server-Side Encryption with Customer-Provided Keys): SSE-C allows you to use your own encryption keys and manage the encryption process yourself.

- Access Control and IAM Policies:
  - AWS Identity and Access Management (IAM): IAM enables you to control access to your AWS storage resources.
  - You can define granular permissions and manage user access through IAM policies, roles, and groups.
  - Bucket Policies: S3 bucket policies help you define fine-grained access controls at the bucket level, allowing or denying access based on various conditions.
  - Access Control Lists (ACLs): S3 ACLs provide another layer of access control by specifying permissions for individual objects within a bucket.

- Compliance Frameworks and Certifications:
  - AWS Compliance Program: AWS provides a comprehensive compliance program that adheres to various industry standards and regulations, including GDPR, HIPAA, ISO 27001, and SOC 1/2/3.
  - Compliance Resources: AWS offers a wide range of compliance resources, including whitepapers, guides, and reference architectures, to help you understand and implement security controls and meet regulatory requirements.
  - Data Encryption and Key Management: AWS Key Management Service (KMS): KMS allows you to manage encryption keys for various AWS services, including S3, EBS, and RDS.
  - It provides centralized key management, encryption key rotation, and auditing capabilities.

- Key Management Best Practices:
  - Implementing strong key management practices is crucial for data security.
  - Consider using AWS KMS to manage and rotate encryption keys regularly, restrict key access, and monitor key usage.
  - Auditing, Monitoring, and Logging: AWS CloudTrail: CloudTrail provides detailed logs of API activity in your AWS account, allowing you to monitor storage-related events and track changes to resources.
  - AWS Config: Config enables you to assess resource configurations and monitor changes, helping you maintain compliance and identify security issues proactively.
  - Security Best Practices: Least Privilege: Follow the principle of least privilege by granting only the necessary permissions to users and applications.

- Multi-Factor Authentication (MFA): Enable MFA for IAM users to add an extra layer of protection to their accounts.
- Regular Auditing and Monitoring: Implement regular auditing, monitoring, and log analysis to detect and respond to security events promptly.
- Security Testing and Vulnerability Scanning: Conduct regular security testing, vulnerability assessments, and penetration testing to identify and address any potential vulnerabilities.

- Designing and Deploying Storage Solutions:
  - Consider your data access patterns and design your storage architecture accordingly.
  - Utilize different AWS storage services based on specific use cases and requirements.
  - Use S3 bucket policies and IAM roles to enforce access controls.
  - Implement versioning and lifecycle policies to manage data retention and deletion.
- Cost Optimization Strategies:
  - Choose the right storage class for your data in Amazon S3 (Standard, Intelligent-Tiering, Glacier, etc.) based on access frequency and retrieval time requirements.
  - Leverage Amazon S3 Glacier for long-term archival storage of



- Monitor and analyze storage usage to identify opportunities for optimization and cost savings.
- Monitoring and Performance Optimization: Utilize AWS CloudWatch to monitor storage metrics, such as S3 bucket size, EBS volume throughput, and EFS file system performance.
- Enable S3 access logging to track access patterns and identify potential security issues.
- Use CloudWatch Alarms to set up alerts for storage capacity, latency, or error rates.
- Consider utilizing AWS Trusted Advisor to receive recommendations on storage performance and cost optimization.

- Additional Best Practices:
  - Implement data encryption at rest and in transit using AWS Key Management Service (KMS) and SSL/TLS.
  - Regularly backup your data using AWS-native backup solutions like EBS snapshots, RDS snapshots, and DynamoDB backups.
  - Implement granular access controls using IAM policies and access keys to ensure only authorized users have access to your storage resources.
  - Regularly test your disaster recovery and data restoration processes to ensure data integrity and availability.

- Additional Resources:
  - AWS Storage Services documentation: Visit the AWS documentation for detailed information and best practices on each storage service.
  - AWS Cost Optimization whitepapers and guides: Explore AWS resources on optimizing costs for various services, including storage.
  - AWS Well-Architected Framework: Refer to the framework's storage pillar for comprehensive guidance on designing, deploying, and optimizing storage solutions on AWS.

Case Studies:

Real-World Examples



#### Case Studies: Real-World Examples

- Netflix:
  - Use case: Content storage and delivery Netflix uses Amazon S3 for storing and streaming its vast library of movies and TV shows.
  - S3 provides high durability, scalability, and low latency, enabling seamless streaming for millions of Netflix users worldwide.
- Airbnb:
  - Use case: Scalable file storage Airbnb utilizes Amazon EFS for storing user-generated content, such as property photos and documents.
  - EFS offers shared file storage that can handle high-scale workloads, ensuring reliable and scalable storage for Airbnb's growing user base.

#### Case Studies: Real-World Examples

#### Capital One:

- Use case: Data backup and recovery Capital One leverages AWS Backup to protect critical data across its various AWS services, including S3, EBS, and RDS.
- AWS Backup simplifies backup management, enabling Capital One to meet its data protection requirements and ensure quick recovery in case of data loss.

#### • Lyft:

- Use case: Data archiving Lyft employs Amazon Glacier for long-term archival storage of historical ride and customer data.
- Glacier's low-cost storage solution helps Lyft meet compliance and regulatory requirements while providing secure and durable storage for their data archives.



#### Case Studies: Real-World Examples

#### Dow Jones:

- Use case: Hybrid cloud storage Dow Jones uses AWS Storage Gateway to seamlessly extend its on-premises storage infrastructure to the cloud.
- The storage gateway enables hybrid cloud storage, allowing Dow Jones to leverage the scalability and cost-effectiveness of AWS storage services while maintaining data locality and integration with their existing systems.
- Unilever:
- Use case: Global data replication Unilever utilizes AWS DataSync to replicate large volumes of data between its global locations and AWS storage services.
- DataSync provides efficient and secure data transfer, enabling Unilever to maintain data consistency across multiple regions and ensure data

