Chapter 6:

1. **Amazon Elastic Block Store (Amazon EBS)**

| **Feature** | **Description** | **Important Notes** |
| --- | --- | --- |
| Storage Type | Persistent block storage volumes for use with Amazon EC2 instances. Available types include: General Purpose SSD (gp2), Provisioned IOPS SSD (io1), Throughput Optimized HDD (st1), Cold HDD (sc1), and Magnetic (standard). | - Each type offers different performance characteristics and pricing options |
| Durability | Automatically replicated within its Availability Zone for high availability and durability. Data is redundantly stored across multiple devices within an Availability Zone to ensure resilience against component failure. | - Replication within Availability Zone ensures data integrity and protection against component failure\n- Designed for high durability to minimize risk of data loss |
| Performance | Provides consistent performance with options for different performance levels based on the chosen volume type. General Purpose SSD volumes offer a balance of price and performance. Provisioned IOPS SSD volumes allow you to specify IOPS requirements for more demanding workloads. | - Performance measured in terms of IOPS (Input/Output Operations Per Second) and throughput\n- Performance characteristics vary between volume types and can be tailored to specific workload requirements |
| Scalability | Can scale usage up or down within minutes. Allows for dynamic resizing of volumes without interruption to running instances. Additionally, you can change volume types to adapt to changing performance requirements. | - Offers flexibility to adapt to changing storage needs\n- Scaling can be done without downtime, allowing for seamless adjustments to capacity and performance |
| Use Cases | Amazon EBS volumes are suitable for various use cases including: boot volumes and storage for Amazon EC2 instances, data storage with a file system, database hosts, and enterprise applications such as content management systems and analytics platforms. | - Suitable for a wide range of applications including databases and enterprise systems\n- Supports various use cases due to its versatility |
| Snapshots | Point-in-time snapshots enable data backup and recovery. Baseline snapshot captures the initial state of the volume, while subsequent snapshots capture only the incremental changes since the last snapshot. Snapshots are stored in Amazon S3 and can be used to create new volumes or restore data. | - Baseline snapshot captures initial state; subsequent snapshots capture incremental changes\n- Facilitates efficient backup and restoration processes |
| Data Encryption | Data encrypted in transit between EC2 instances and EBS volumes at no additional cost. Additionally, you can encrypt EBS volumes and snapshots using AWS Key Management Service (KMS) at no extra charge. | - Enhances security by encrypting data during transmission\n- Supports encryption at rest for added data protection |
| Volume Resize | Dynamic resizing without instance stoppage; can change storage types and capacity. Allows for seamless adjustment of storage capacity and performance characteristics. | - Allows for seamless adjustment of storage capacity and types to meet evolving requirements\n- Eliminates downtime for resizing operations |
| Cost Factors | - Volume Storage: Charged by provisioned GB per month until released\n- IOPS: Included for General Purpose SSD volumes, charged for magnetic volumes and Provisioned IOPS SSD volumes\n- Snapshots: Charged per GB-month of stored data\n- Data Transfer: Charged for data transferred across Regions when copying snapshots | - Understanding cost factors is crucial for effective budgeting\n- Regular monitoring of usage helps optimize costs and prevent unexpected charges\n- Consider data transfer costs when copying snapshots across Regions for disaster recovery |

1. **Amazon Simple Storage Service (Amazon S3)**

| **Feature** | **Description** | **Important Notes** |
| --- | --- | --- |
| Storage Type | Object-level storage designed for scalability and durability. Data is stored as objects within buckets. | - Amazon S3 stores virtually unlimited objects within buckets.\n- Objects can be up to 5 TB in size.\n- Data is redundantly stored across multiple facilities and devices within a selected region. |
| Access & Management | Managed cloud storage solution with fine-grained access control through IAM policies, bucket policies, and per-object ACLs. Data can be accessed via web console, API/SDK, or third-party solutions. | - Data stored in Amazon S3 is not associated with any particular server.\n- Fine-grained access control ensures security and compliance.\n- Various methods available for accessing and managing data. |
| Data Encryption | Supports encryption in transit (HTTPS) and at rest. Server-side encryption options available for added security. | - Data encryption ensures data security both during transmission and storage.\n- Encryption can be enabled at the object level for enhanced protection. |
| Event Notifications | Event notifications allow setting up automatic notifications triggered by specified events (e.g., object upload, deletion). | - Notifications can be used to automate processes or trigger AWS Lambda functions based on events.\n- Enhances operational efficiency by enabling timely responses to events. |
| Storage Classes | Offers a range of storage classes tailored for different use cases: \n- **Amazon S3 Standard**: High durability, availability, and performance for frequently accessed data.\n- **Amazon S3 Intelligent-Tiering**: Automatically optimizes costs by moving data to the most cost-effective tier.\n- **Amazon S3 Standard-IA**: Low-cost option for infrequently accessed data requiring rapid access.\n- **Amazon S3 One Zone-IA**: Lower-cost option with data stored in a single Availability Zone.\n- **Amazon S3 Glacier**: Secure, durable, and low-cost storage for data archiving.\n- **Amazon S3 Glacier Deep Archive**: Lowest-cost option for long-term retention and digital preservation. | - Each storage class is optimized for specific access patterns and cost considerations.\n- Allows cost-effective storage management based on usage patterns.\n- Suitable for various scenarios ranging from frequently accessed data to long-term archival. |
| Data Analysis | Storage class analysis enables analyzing storage access patterns and optimizing data placement. Amazon S3 Analytics provides insights into storage usage and facilitates lifecycle policy management. | - Helps identify optimal storage classes for data based on access patterns.\n- Enables effective lifecycle policy management to reduce costs.\n- Provides visibility into storage usage through analytics and visualizations. |
| Use Cases | Amazon S3 serves various use cases including application data storage, static web hosting, data backups, and archival. | - Provides a versatile storage solution suitable for a wide range of applications and industries.\n- Supports scenarios such as application data hosting, web content delivery, and data archival for compliance and disaster recovery. |
| Cost Considerations | Costs are determined by factors such as storage class type, amount of storage, requests (e.g., GET, PUT, COPY), and data transfer out of the region. | - Understanding cost factors is essential for effective cost management.\n- Various factors such as storage class, storage amount, and request types influence pricing.\n- Careful consideration of data transfer costs is necessary, especially for cross-region transfers and outbound data transfer. |

1. **Amazon Elastic File System (Amazon EFS)**

| **Feature** | **Description** | **Important Notes** |
| --- | --- | --- |
| Overview | Amazon EFS provides simple, scalable, elastic file storage for use with AWS services and on-premises resources. It implements storage for EC2 instances that multiple virtual machines can access concurrently, using the Network File System (NFS) protocol. | - **Simple, Scalable, Elastic:** Allows easy creation and scaling of file storage without disruption.  -**Fully Managed:** No setup costs or minimum fees; pay only for storage used.\n- **Versatile:** Supports various use cases like big data analytics, media processing, web serving, etc. |
| File System Access | File systems are accessible to Amazon EC2 instances through standard operating system file I/O APIs, supporting full file system access semantics like strong consistency and file locking. | - **Dynamic Scalability:** Scales automatically from gigabytes to petabytes without manual intervention.\n- **Consistent Performance:** Ensures consistent performance to each EC2 instance accessing the file system.\n- **Highly Durable and Available:** Designed for high durability and availability, suitable for critical workloads. |
| Connectivity | Amazon EFS file systems can be mounted on Amazon EC2 instances in a VPC through NFS versions 4.0 and 4.1 (NFSv4), allowing concurrent access from multiple instances within the same VPC and across multiple Availability Zones in the same AWS Region. | - **Concurrent Access:** Supports multiple EC2 instances accessing the file system simultaneously.\n- **Cross-AZ Availability:** Accessible from EC2 instances in different Availability Zones for redundancy and scalability.\n- **Optimal Performance:** Recommends creating mount targets in the same Availability Zone as the EC2 instances for best performance. |
| Setup Steps | Five steps are required to create and use an Amazon EFS file system: <ol><li>Create Amazon EC2 resources and launch an instance.</li><li>Create an Amazon EFS file system.</li><li>Create mount targets in appropriate subnets.</li><li>Connect to the Amazon EC2 instance and mount the Amazon EFS file system.</li><li>Clean up resources and protect the AWS account.</li></ol> | - **Key Steps:** Launch EC2 instance, create EFS file system, set up mount targets, connect to instance, and perform resource cleanup.\n- **Security Consideration:** Essential to protect AWS account after use to prevent unauthorized access or accidental costs.\n- **Proper Subnet Configuration:** Mount targets should be created in appropriate subnets for connectivity and performance optimization. |
| File System Properties | Each file system has properties such as ID, creation token, creation time, size in bytes, number of mount targets, and state. | - **Unique ID:** Identifies the file system uniquely within the AWS environment.\n- **Dynamic Sizing:** Scales from gigabytes to petabytes based on data requirements.\n- **Mount Target Count:** Indicates the number of mount targets associated with the file system.\n- **Current State:** Reflects the current status of the file system (e.g., available, creating, deleting). |
| Mount Targets | Mount targets must be created in the VPC to access the file system. Each mount target has properties like ID, subnet ID, file system ID, mount IP address, and state. | - **Essential Access Points:** Mount targets facilitate access to the file system from EC2 instances.\n- **Subnet Association:** Each target is associated with a specific subnet for connectivity.\n- **State Management:** Tracks whether the mount target is ready for use or undergoing changes.\n- **Optimal Deployment:** Creating targets in appropriate subnets ensures connectivity and performance optimization. |
| Tags | Tags are key-value pairs used to organize file systems by assigning metadata. They are associated with each file system for organization and management purposes. | - **Organization and Management:** Tags help categorize and manage multiple file systems effectively.\n- **Flexible Metadata:** Provides custom metadata assignment using key-value pairs.\n- **Resource Identification:** Enhances resource identification and organization for streamlined management.\n- **Optional Feature:** Tags are optional but highly recommended for efficient resource management and organization. |

1. **Amazon S3 Glacier**

| **Feature** | **Description** | **Important Notes** |
| --- | --- | --- |
| Overview | Amazon S3 Glacier is a secure, durable, and extremely low-cost cloud storage service for data archiving and long-term backup. Data stored in Glacier is ideal for archiving purposes due to its low cost, but retrieval times are longer compared to Amazon S3. | - **Secure and Durable:** Provides secure and durable storage for long-term data archiving.\n- **Low-Cost:** Offers extremely low-cost storage options for archiving compared to Amazon S3.\n- **Long Retrieval Times:** Retrieval of data can take several hours, making it suitable for infrequently accessed data. |
| Key Terms | - **Archive:** Any object (e.g., photo, video, file) stored in Amazon S3 Glacier, each with a unique ID and optional description.\n- **Vault:** A container for storing archives, located in a specified AWS Region.\n- **Vault Access Policy:** Defines who can access data stored in a vault and the operations they can perform, including the use of vault lock policies for immutable storage. | - **Understanding Terminology:** Familiarity with terms like archive, vault, and vault access policy is essential for managing Glacier storage.\n- **Access Control:** Vault access policies control data access and operations, ensuring security and compliance.\n- **Immutable Storage:** Vault lock policies can be used to enforce data immutability, critical for compliance requirements such as SEC Rule 17a-4(f). |
| Retrieval Options | - **Expedited Retrievals:** Available within 1–5 minutes (highest cost).\n- **Standard Retrievals:** Complete within 3–5 hours (medium cost).\n- **Bulk Retrievals:** Complete within 5–12 hours (lowest cost). | - **Cost vs. Speed Trade-off:** Different retrieval options offer varying access times and costs, allowing users to choose based on their specific needs.\n- **Analogous to Shipping Options:** Comparing retrieval options to shipping methods helps illustrate the trade-off between cost and speed. |
| Use Cases | - **Media Asset Archiving:** Archive older media content affordably and distribute when needed.\n- **Healthcare Information Archiving:** Securely retain patient records for compliance requirements.\n- **Regulatory and Compliance Archiving:** Enforce compliance controls for regulatory archives.\n- **Scientific Data Archiving:** Manage and preserve large volumes of research data.\n- **Digital Preservation:** Ensure long-term data integrity and accessibility.\n- **Magnetic Tape Replacement:** Replace traditional tape libraries with cost-effective cloud storage. | - **Diverse Applications:** Glacier serves a range of use cases, from media archiving to regulatory compliance, catering to different industries and needs.\n- **Compliance Assurance:** Vault lock policies help meet regulatory requirements by ensuring data immutability.\n- **Cost Savings:** Cost-effective storage options make Glacier an attractive choice for organizations looking to reduce long-term storage costs. |
| Management | - **AWS Management Console:** Offers limited operations such as vault creation and policy management.\n- **APIs and SDKs:** Use Amazon S3 Glacier REST APIs, AWS SDKs (Java, .NET), or AWS CLI for most operations and interactions.\n- **Lifecycle Policies:** Automate data lifecycle management to transition data between storage tiers based on predefined rules. | - **Management Options:** Choose from console, APIs/SDKs, or CLI for managing Glacier storage based on preference and requirements.\n- **Lifecycle Management:** Automating data lifecycle ensures cost optimization by moving data to appropriate storage tiers over time.\n- **API Accessibility:** Most operations require API/SDK usage, highlighting the importance of developer familiarity with Glacier APIs for effective management. |
| Data Encryption | - **Server-Side Encryption:** Protects data at rest; available options include SSE-S3, SSE-C, and SSE-KMS.\n- **HTTPS Transfer:** Securely transfer data over HTTPS for both Amazon S3 and Glacier.\n- **Default Encryption:** Data archived in Glacier is encrypted by default, ensuring data security. | - **Encryption Options:** Choose from various server-side encryption methods for data protection at rest.\n- **Secure Transfer:** HTTPS ensures secure data transfer between client and storage, enhancing overall data security.\n- **Default Encryption:** Glacier ensures data encryption by default, providing an additional layer of security for archived data. |
| Differences from S3 | - **Cost Structure:** Glacier offers lower storage costs but higher retrieval costs compared to Amazon S3.\n- **Access Speed:** S3 provides faster access to data, while Glacier is optimized for infrequent access and long-term storage.\n- **Maximum Item Size:** S3 allows larger item sizes (up to 5 TB) compared to Glacier (up to 40 TB).\n- **Request Charges:** S3 charges for various operations like PUT, GET, etc., while Glacier charges for UPLOAD and retrieval operations. | - **Storage vs. Retrieval Costs:** Understanding the cost implications of storage and retrieval helps in choosing the appropriate service for different data access patterns.\n- **Access Speed Consideration:** Glacier's longer retrieval times make it suitable for archival data with less frequent access.\n- **Operational Differences:** Differences in maximum item size and request charges influence the choice between S3 and Glacier for specific storage requirements. |