

AWS Solution Architect Associate Certification Training – Module 8

8. Elastic Block Store (EBS)

Introduction to Elastic Block Store

Amazon Elastic Block Store (Amazon EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS Cloud. Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability. Amazon EBS volumes offer the consistent and low-latency performance needed to run your workloads. With Amazon EBS, you can scale your usage up or down within minutes – all while paying a low price for only what you provision.

Amazon EBS is designed for application workloads that benefit from fine tuning for performance, cost and capacity. Typical use cases include Big Data analytics engines (like the Hadoop/HDFS ecosystem and Amazon EMR clusters), relational and NoSQL databases (like Microsoft SQL Server and MySQL or Cassandra and MongoDB), stream and log processing applications (like Kafka and Splunk), and data warehousing applications (like Vertica and Teradata).

Benefits

Reliable, secure storage

Each Amazon EBS volume provides redundancies within its Availability Zone to protect against failures. Encryption and access control policies deliver a strong defense-in-depth security strategy for your data.

Quickly Scale Up, Easily Scale Down

Amazon EBS allows you to optimize your volumes for capacity, performance, or cost giving you the ability to dynamically adapt to the changing needs of your business.

Consistent, Low-latency Performance

Amazon EBS General Purpose (SSD) volumes and Amazon EBS Provisioned IOPS (SSD) volumes deliver low-latency through SSD technology and consistent I/O performance scaled to the needs of your application.

Geographic Flexibility

Amazon EBS provides the ability to copy snapshots across AWS regions, enabling geographical expansion, data center migration, and disaster recovery providing flexibility and protecting for your business.

Backup, Restore, Innovate

Protect your data by taking point-in-time snapshots of your Amazon EBS volumes providing long-term durability for your data. Boost the agility of your business by using Amazon EBS snapshots to create new EC2 instances.

Optimized Performance

An Amazon EBS-optimized instance provides dedicated network capacity for Amazon EBS volumes. This provides the best performance for your EBS volumes by minimizing network contention between EBS and your instance.

Amazon EBS Volume Types

Amazon EBS provides the following volume types, which differ in performance characteristics and price, so that you can tailor your storage performance and cost to the needs of your applications. The volumes types fall into two categories:

- SSD-backed volumes optimized for transactional workloads involving frequent read/write operations with small I/O size, where the dominant performance attribute is IOPS.
- HDD-backed volumes optimized for large streaming workloads where throughput (measured in MiB/s) is a better performance measure than IOPS

Differentiation of EBS Volume Types – IOPS, Throughput, Performance Factors, use cases of EBS Volumes

Volume Type	Solid-State Drives (SSD)		Hard Disk Drives (HDD)	
	General Purpose SSD (gp2)*	Provisioned IOPS SSD (io1)	Throughput Optimized HDD (st1)	Cold HDD (sc1)
Description	General purpose SSD volume that balances price and performance for a wide variety of workloads	Highest-performance SSD volume for mission-critical low-latency or high-throughput workloads	Low-cost HDD volume designed for frequently accessed, throughput-intensive workloads	Lowest cost HDD volume designed for less frequently accessed workloads
Use Cases	<ul style="list-style-type: none">• Recommended for most workloads• System boot volumes• Virtual desktops• Low-latency interactive apps• Development and test environments	<ul style="list-style-type: none">• Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiB/s of throughput per volume• Large database workloads, such as:<ul style="list-style-type: none">◦ MongoDB◦ Cassandra◦ Microsoft SQL Server◦ MySQL◦ PostgreSQL◦ Oracle	<ul style="list-style-type: none">• Streaming workloads requiring consistent, fast throughput at a low price• Big data• Data warehouses• Log processing• Cannot be a boot volume	<ul style="list-style-type: none">• Throughput-oriented storage for large volumes of data that is infrequently accessed• Scenarios where the lowest storage cost is important• Cannot be a boot volume
API Name	gp2	io1	st1	sc1
Volume Size	1 GiB - 16 TiB	4 GiB - 16 TiB	500 GiB - 16 TiB	500 GiB - 16 TiB
Max. IOPS**/Volume	16,000***	64,000****	500	250
Max. Throughput/Volume	250 MiB/s***	1,000 MiB/s†	500 MiB/s	250 MiB/s
Max. IOPS/Instance††	80,000	80,000	80,000	80,000
Max. Throughput/Instance††	1,750 MiB/s	1,750 MiB/s	1,750 MiB/s	1,750 MiB/s
Dominant Performance Attribute	IOPS	IOPS	MiB/s	MiB/s

Magnetic (standard)

Magnetic volumes are backed by magnetic drives and are suited for workloads where data is accessed infrequently, and scenarios where low-cost storage for small volume sizes is important. These volumes deliver approximately 100 IOPS on average, with burst capability of up to hundreds of IOPS, and they can range in size from 1 GiB to 1 TiB.

Note: Magnetic is a Previous Generation Volume. For new applications, we recommend using one of the newer volume types.

Amazon EBS Encryption

Amazon EBS encryption offers a simple encryption solution for your EBS volumes without the need to build, maintain, and secure your own key management infrastructure. When you create an encrypted EBS volume and attach it to a supported instance type, the following types of data are encrypted:

- ✓ Data at rest inside the volume
- ✓ All data moving between the volume and the instance
- ✓ All snapshots created from the volume
- ✓ All volumes created from those snapshots

Encryption operations occur on the servers that host EC2 instances, ensuring the security of both data-at-rest and data-in-transit between an instance and its attached EBS storage.

Encryption is supported by all EBS types (General Purpose SSD [gp2], Provisioned IOPS SSD [io1], Throughput Optimized HDD [st1], Cold HDD [sc1], and Magnetic [standard]). You can expect the same IOPS performance on encrypted volumes as on unencrypted volumes, with a minimal effect on latency. You can access encrypted volumes the same way that you access unencrypted volumes. Encryption and decryption are handled transparently and they require no additional action from you or your applications.

Public snapshots of encrypted volumes are not supported, but you can share an encrypted snapshot with specific accounts. Amazon EBS encryption is only available on certain instance types. You can attach both encrypted and unencrypted volumes to a supported instance type.

Amazon EBS encryption is available on the instance types listed below. You can attach both encrypted and unencrypted volumes to these instance types simultaneously.

- General purpose: A1, M3, M4, M5, M5d, T2, and T3.
- Compute optimized: C3, C4, C5, C5d, and C5n.
- Memory optimized: cr1.8xlarge, R3, R4, R5, R5d, X1, X1e, and z1d.
- Storage optimized: D2, h1.2xlarge, h1.4xlarge, I2, and I3.
- Accelerated computing: F1, G2, G3, P2, and P3.
- Bare metal: i3.metal, m5.metal, m5d.metal, r5.metal, r5d.metal, u-6tb1.metal, u-9tb1.metal, u-12tb1.metal, and z1d.metal

Snapshots – Root Volumes and Data Volumes

You can back up the data on your Amazon EBS volumes to Amazon S3 by taking point-in-time snapshots. Snapshots are *incremental* backups, which means that only the blocks on the device that have changed after your most recent snapshot are saved. This minimizes the time required to create the snapshot and saves on storage costs by not duplicating data. When you delete a snapshot, only the data unique to that snapshot is removed. Each snapshot contains all of the information needed to restore your data (from the moment when the snapshot was taken) to a new EBS volume.

When you create an EBS volume based on a snapshot, the new volume begins as an exact replica of the original volume that was used to create the snapshot. The replicated volume loads data lazily in the background so that you can begin using it immediately. If you access data that hasn't been loaded yet, the volume immediately downloads the requested data from Amazon S3, and then continues loading the rest of the volume's data in the background.

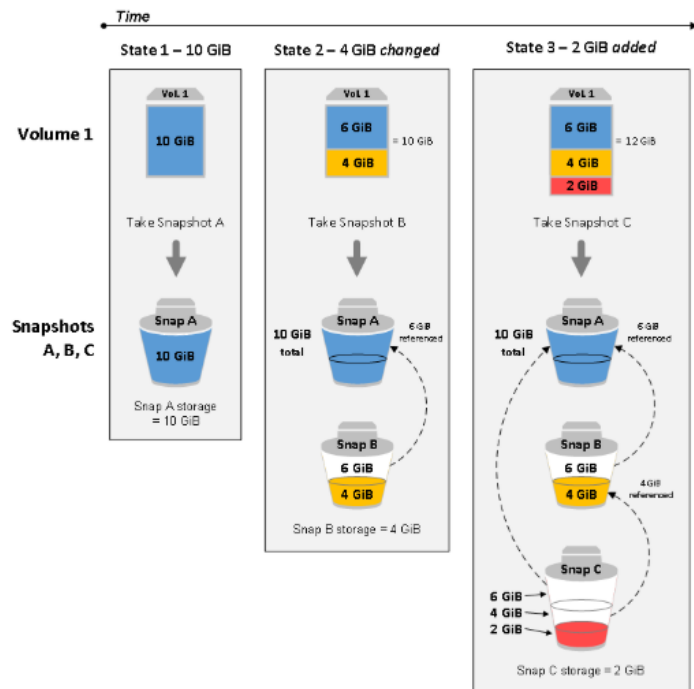
How Incremental Snapshots Work

This section provides illustrations of how an EBS snapshot captures the state of a volume at a point in time, and also how successive snapshots of a changing volume create a history of those changes.

In the diagram below, Volume 1 is shown at three points in time. A snapshot is taken of each of these three volume states.

- In State 1, the volume has 10 GiB of data. Because Snap A is the first snapshot taken of the volume, the entire 10 GiB of data must be copied.
- In State 2, the volume still contains 10 GiB of data, but 4 GiB have changed. Snap B needs to copy and store only the 4 GiB that changed after Snap A was taken. The other 6 GiB of unchanged data, which are already copied and stored in Snap A, are *referenced* by Snap B rather than (again) copied. This is indicated by the dashed arrow.
- In State 3, 2 GiB of data have been added to the volume, for a total of 12 GiB. Snap C needs to copy the 2 GiB that were added after Snap B was taken. As shown by the dashed arrows, Snap C also references 4 GiB of data stored in Snap B, and 6 GiB of data stored in Snap A.

The total storage required for the three snapshots is 16 GiB.



Root Volume: When you launch an instance, the root device volume contains the image used to boot the instance. When we introduced Amazon EC2, all AMIs were backed by Amazon EC2 instance store, which means the root device for an instance launched from the AMI is an instance store volume created from a template stored in Amazon S3.

Data Volume: An Amazon EBS volume is a durable, block-level storage device that you can attach to a single EC2 instance. You can use EBS volumes as primary storage for data that requires frequent updates, such as the system drive for an instance or storage for a database application.

Life Cycle Manager

You can use Amazon Data Lifecycle Manager (Amazon DLM) to automate the creation, retention, and deletion of snapshots taken to back up your Amazon EBS volumes. Automating snapshot management helps you to:

- Protect valuable data by enforcing a regular backup schedule.
- Retain backups as required by auditors or internal compliance.
- Reduce storage costs by deleting outdated backups.

The Difference between EBS Snapshots and AMI Images

An EBS snapshot is a backup of a single EBS volume. The EBS snapshot contains all the data stored on the EBS volume at the time the EBS snapshot was created.

An AMI image is a backup of an entire EC2 instance. Associated with an AMI image are EBS snapshots. Those EBS snapshots are the backups of the individual EBS volumes attached to the EC2 instance at the time the AMI image was created.

The process of creating an AMI image will also create EBS snapshots. So even if you need to restore a single EBS volume, you can do so from the individual EBS snapshot.