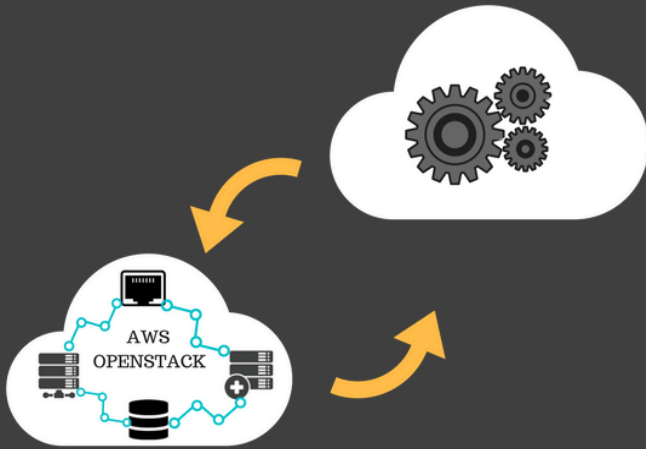


AWS Design and Automation

Module 1: Cloud Compute Basics

Topic 2: Elastic Block System

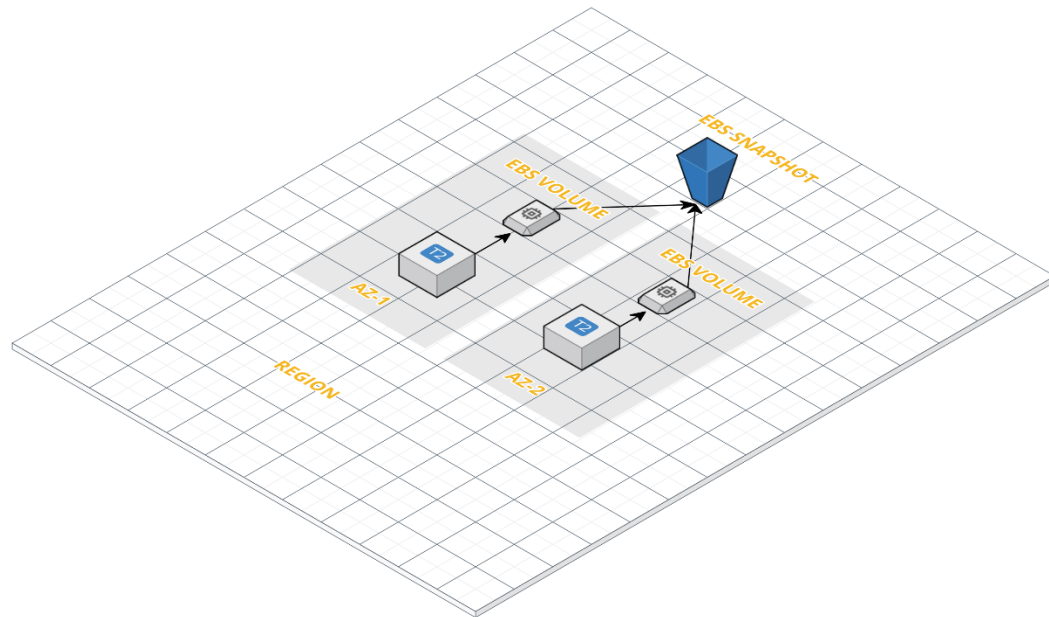


Mohanraj Shanmugam

Amazon EBS

Mohanraj Shanmugam

Amazon EBS

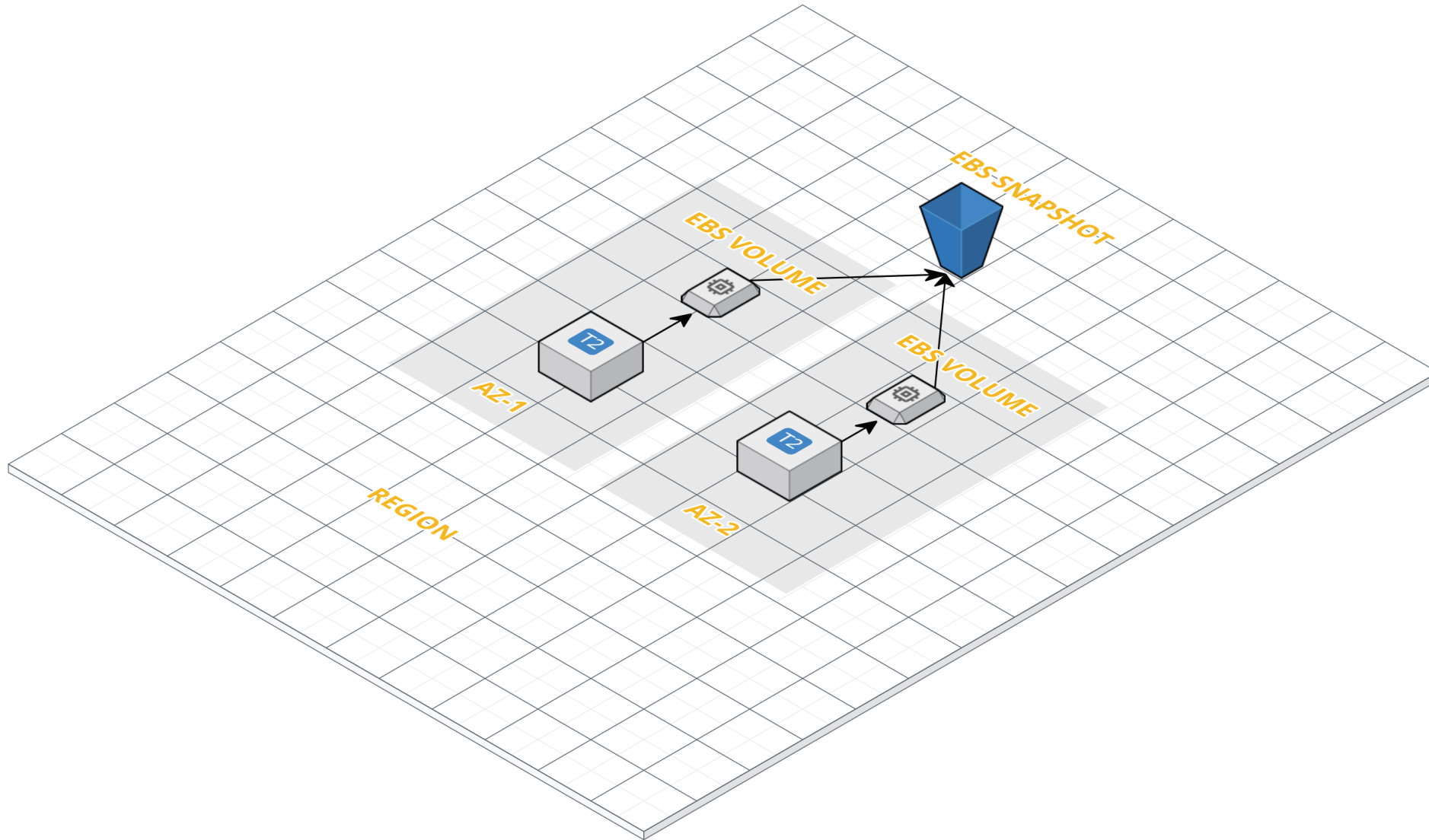


POWERED BY
CLOUDCRAFT.CO

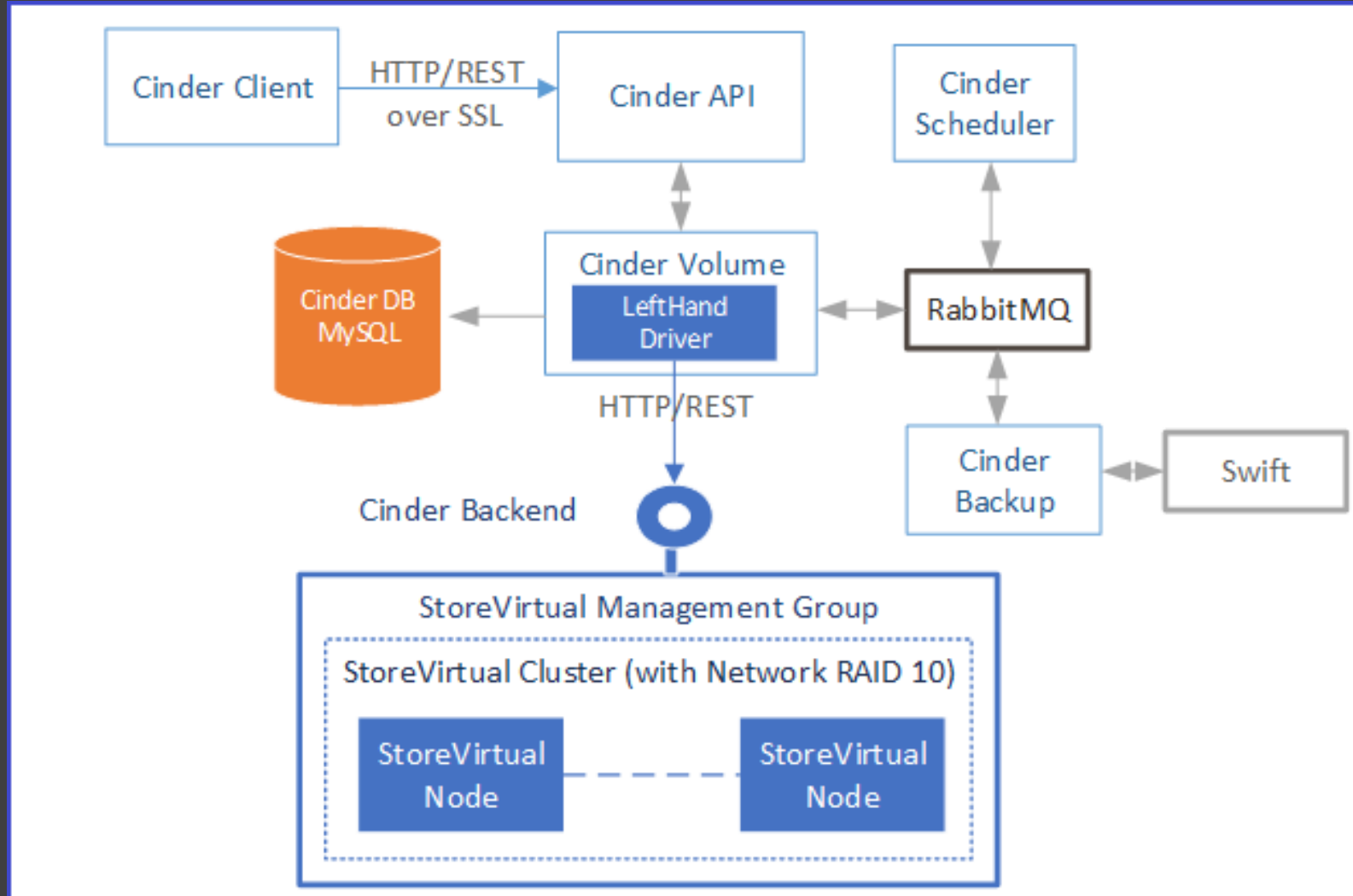
- Amazon Elastic Block Store (Amazon EBS) provides persistent block level storage volumes for use with Amazon EC2 instances in the AWS Cloud.
- Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability.
- 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.
- All EBS volume types offer the same durable snapshot capabilities and are designed for 99.999% availability.

EBS Overview

- Amazon EBS volumes are off-instance storage that persists independently from the running life of a single Amazon EC2 instance. After an Amazon EBS volume is attached to an EC2 instance, you can use the EBS volume like a physical hard drive, typically by formatting it with the file system of your choice and using the file I/O interface provided by the instance operating system.
- For Amazon Machine Images (AMIs) that are backed by Amazon EBS, you can use an Amazon EBS volume to boot an Amazon EC2 instance.
- You can also attach multiple Amazon EBS volumes to a single Amazon EC2 instance.
- Note, however, that any single Amazon EBS volume can be attached to only one Amazon EC2 instance at any point in time.
- Amazon EBS also provides the ability to create point-in-time snapshots of volumes, which are persisted to Amazon S3. These snapshots can be used as the starting point for new Amazon EBS volumes, and to protect data for long-term durability.
- The same snapshot can be used to instantiate as many volumes as you want. These snapshots can be copied across AWS regions, making it easier to leverage multiple AWS regions for geographical expansion, data center migration, and disaster recovery.
- Sizes for Amazon EBS volumes range from 1 GB to 16 TB, and are allocated in 1 GB increments.



Block Storage Service Examples



Amazon EBS

- AWS EBS comes in Three variant on different performances

Volume Type	EBS General Purpose (SSD)	EBS Provisioned IOPS (SSD)	EBS Magnetic
Use Cases	Boot volumes Small to Med DBs Dev and Test	I/O intensive Relational DBs NoSQL DBs	Infrequent Data Access
Storage Media	SSD-backed	SSD-backed	Magnetic disk-backed
Max Volume Size	16 TB	16 TB	1 TB
Max IOPS/volume	10,000	20,000	40 - 200
Max IOPS Burst Performance	3,000 for volumes <= 1 TB	-	Hundreds
Max throughput/volume	160 MB/sec	320 MB/sec	40 - 90 MB/sec
Max IOPS/instance	48,000	48,000	48,000
Max throughput/instance	800 MB/sec	800 MB/sec	800 MB/sec
API Name	gp2	io1	standard
Price*	\$.10/GB - Month	\$.125/GB - Month \$.065/provisioned IOPS	\$.05/GB - Month \$.05/million I/O

Amazon EBS Benefits

- Reliable, secure storage
 - Replicated within the AZ
 - Designed for an annual failure rate (AFR) of between 0.1% - 0.2%, where failure refers to a complete or partial loss of the volume, depending on the size and performance of the volume.
 - Supports Encryption for EBS data volumes, boot volumes and snapshots
 - Data at motion Encrypted between EBS and EC2 instances
 - Provides Fine grained access to EBS Volumes
- Consistent and 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. Stripe multiple volumes together to achieve even higher I/O performance.
- Backup, restore, innovate
 - Backup your data by taking point-in-time snapshots of your Amazon EBS volumes.
 - Use Amazon EBS snapshots to create new EC2 instances.

EBS Features

- Quickly scale up, easily scale
 - Increase or decrease block storage and performance within minutes, enjoying the freedom to adjust as your needs evolve. Commission thousands of volumes simultaneously.
- Geographic flexibility
 - Amazon EBS provides the ability to copy snapshots across AWS regions, enabling geographical expansion, data center migration, and disaster recovery.

Amazon EBS Snapshots

- You can back up the data on your 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.
- If you have a device with 100 GB of data but only 5 GB has changed after your last snapshot, a subsequent snapshot consumes only 5 additional GB and you are billed only for the additional 5 GB of snapshot storage, even though both the earlier and later snapshots appear complete.
- When you delete a snapshot, only the data exclusive to that snapshot is removed.
- Active snapshots contain all of the information needed to restore your data (from the time the snapshot was taken) to a new EBS volume.
- Snapshots can be used to instantiate multiple new volumes, expand the size of a volume, or move volumes across Availability Zones.
- When a new volume is created, you may choose to create it based on an existing Amazon EBS snapshot. In that scenario, the new volume begins as an exact replica of the snapshot.

features of Amazon EBS Snapshots:

- **Immediate access to Amazon EBS volume data** - After a volume is created from a snapshot, there is no need to wait for all of the data to transfer from Amazon S3 to your Amazon EBS volume before your attached instance can start accessing the volume. Amazon EBS Snapshots implement lazy loading, so that you can begin using them right away.
- **Resizing Amazon EBS volumes** - When you create a new Amazon EBS volume based on a snapshot, you may specify a larger size for the new volume. Make certain that your file system or application supports resizing a device.
- **Sharing Amazon EBS Snapshots** - Amazon EBS Snapshots' shareability makes it easy for you to share data with your co-workers or others in the AWS community. Authorized users can create their own Amazon EBS volumes based on your Amazon EBS shared snapshots; your original snapshot remains intact. If you choose, you can also make your data available publicly to all AWS users. For more information about how to share snapshots, see [Modifying Snapshot Permissions](#).
- **Copying Amazon EBS Snapshots across AWS regions** - Amazon EBS's ability to copy snapshots across AWS regions makes it easier to leverage multiple AWS regions for geographical expansion, data center migration and disaster recovery. You can copy any snapshot accessible to you: snapshots you created; snapshots shared with you; and snapshots from the AWS Marketplace, VM Import/Export, and AWS Storage Gateway.

Usage Patterns

- Amazon EBS is meant for data that changes relatively frequently and requires long-term persistence.
- Amazon EBS is well-suited for use as the primary storage for a database or file system, or for any applications that require access to raw block-level storage.
- Amazon EBS Magnetic is well-suited for infrequently accessed data.
- Amazon EBS General Purpose, which is solid-state drive (SSD)–backed, is well-suited for boot volumes and small to medium databases.
- Amazon EBS Provisioned IOPS (also SSD-backed) is well-suited for use with I/O-intensive workloads, relational databases, and NoSQL databases.

Performance

- Amazon EBS provides two SSD volume types:
 - standard volumes and Provisioned IOPS volumes.
- They differ in performance characteristics and pricing model, allowing you to tailor your storage performance and cost to the needs of your applications.
- You can attach and stripe across multiple volumes of either type to increase the I/O performance available to your Amazon EC2 applications

Performance

- Standard volumes offer cost effective storage for applications with moderate or bursty I/O requirements.
- Standard volumes are designed to deliver approximately 100 input/output operations per second (IOPS) on average with a best effort ability to burst to hundreds of IOPS.
- Standard volumes are also well suited for use as boot volumes, where the burst capability provides fast instance start-up times

Performance

- Provisioned IOPS volumes are designed to deliver predictable, high performance for I/O intensive workloads such as databases.
- With Provisioned IOPS, you specify an IOPS rate when creating a volume, and then Amazon EBS provisions that rate for the lifetime of the volume.
- Amazon EBS currently supports up to 2,000 IOPS per Provisioned IOPS volume.
- You can stripe multiple volumes together to deliver thousands of IOPS per Amazon EC2 instance to your application.

Performance

- Because Amazon EBS volumes are network-attached devices, other network I/O performed by the instance, as well as the total load on the shared network, can affect individual Amazon EBS volume performance.
- To enable your Amazon EC2 instances to fully utilize the Provisioned IOPS on an Amazon EBS volume, you can launch selected Amazon EC2 instance types as Amazon EBS-optimized instances.
- Amazon EBS-optimized instances deliver dedicated throughput between Amazon EC2 and Amazon EBS, with options between 500 Mbps and 1,000 Mbps depending on the instance type used.
- When attached to Amazon EBS-optimized instances, Provisioned IOPS volumes are designed to deliver the Provisioned IOPS performance 99.9% of the time.

Performance

- The combination of Amazon EC2 and Amazon EBS enables you to use many of the same disk performance optimization techniques that you would use with on-premises servers and storage. For example, by attaching multiple Amazon EBS volumes to a single Amazon EC2 instance, you can partition the total application I/O load by allocating one volume for database log data, one or more volumes for database file storage, and other volumes for file system data.
- Each separate Amazon EBS volume can be configured as Amazon EBS standard or Amazon EBS Provisioned IOPS as needed.
- Alternatively, you could stripe your data across multiple similarly-provisioned Amazon EBS volumes using RAID 0 or logical volume manager software, thus aggregating available IOPs, total volume throughput, and total volume size.

Durability and Availability

- Amazon EBS volumes are designed to be highly available and reliable.
- Amazon EBS volume data is replicated across multiple servers in a single Availability Zone to prevent the loss of data from the failure of any single component.
- The durability of your Amazon EBS volume depends on both the size of your volume and the amount of data that has changed since your last snapshot.
- Amazon EBS snapshots are incremental, point-in-time backups, containing only the data blocks changed since the last snapshot.
- Amazon EBS volumes that operate with 20 GB or less of modified data since their most recent snapshot can expect an annual failure rate (AFR) between 0.1% and 0.5%.
- Amazon EBS volumes with more than 20 GB of unmodified data since the last snapshot should expect higher failure rates that are roughly proportional to the increase in modified data.

Cost Model

- With Amazon EBS, you pay only for what you use.
- Amazon EBS pricing has three components:
 - provisioned storage,
 - I/O requests, and
 - snapshot storage.
- Amazon EBS standard volumes are charged per GB-month of provisioned storage and per million I/O requests.
- Amazon EBS Provisioned IOPS volumes are charged per GB-month of provisioned storage and per Provisioned IOPS-month. For both volume types, Amazon EBS snapshots are charged per GB-month of data stored. Amazon EBS snapshot copy is charged for the data transferred between regions, and for the standard Amazon EBS snapshot charges in the destination region

Cost Model

- It's important to remember that for Amazon EBS volumes, you are charged for provisioned (allocated) storage, whether or not you actually use it.
- For Amazon EBS snapshots, you are charged only for storage actually used (consumed).
- Note that Amazon EBS snapshots are incremental and compressed, so the storage used in any snapshot is generally much less than the storage consumed on an Amazon EBS volume.
- Note that there is no charge for transferring information among the various AWS storage offerings (i.e., Amazon EC2 instance with Amazon EBS, Amazon S3, Amazon RDS, and so on) as long as they are within the same AWS region.

Scalability and Elasticity

- Using the AWS Management Console or the APIs, Amazon EBS volumes can easily and rapidly be provisioned and released to scale in and out with your total storage demands.
- While individual Amazon EBS volumes cannot be resized, if you find that you need additional storage, you have two ways to expand the amount of Amazon EBS space available for your Amazon EC2 instance.
- The simplest approach is to create and attach a new Amazon EBS volume and begin using it together with your existing ones.

Scalability and Elasticity

- However, if you need to expand the size of a single Amazon EBS volume, you can effectively resize a volume using a snapshot:
 - 1. Detach the original Amazon EBS volume.
 - 2. Create a snapshot of the original Amazon EBS volume's data into Amazon S3.
 - 3. Create a new Amazon EBS volume from the snapshot, but specify a larger size than the original volume.
 - 4. Attach the new, larger volume to your Amazon EC2 instance in place of the original. (In many cases, an OS-level utility must also be used to expand the file system.)
 - 5. Delete the original Amazon EBS volume.

Interfaces

- Amazon offers management APIs for Amazon EBS in both SOAP and REST formats.
- These are used to create, delete, describe, attach, and detach Amazon EBS volumes for your Amazon EC2 instances;
- to create, delete, and describe snapshots from Amazon EBS to Amazon S3; and to copy snapshots from one region to another.
- If you prefer to work with a graphical tool, the AWS Management Console gives you all the capabilities of the API in an easy-to-use browser interface.
- Regardless of how you create your Amazon EBS volume, note that all storage is allocated at the time of volume creation, and that you are charged for this allocated storage even if you don't write data to it.

Anti-Patterns

- Temporary storage—If you are using Amazon EBS for temporary storage (such as scratch disks, buffers, queues, and caches), consider using local instance store volumes, Amazon SQS, or ElastiCache (Memcached or Redis).
- Highly-durable storage—If you need very highly-durable storage, use Amazon S3 or Amazon Glacier. Amazon S3 standard storage is designed for 99.999999999% annual durability per object. In contrast, Amazon EBS volumes with less than 20 GB of modified data since the last snapshot are designed for between 99.5% and 99.9% annual durability; volumes with more modified data can be expected to have proportionally lower durability.
- Static data or web content—If your data doesn't change that often, Amazon S3 may represent a more cost effective and scalable solution for storing this fixed information. Also, web content served out of Amazon EBS requires a web server running on Amazon EC2, while you can deliver web content directly out of Amazon S3.

Amazon EC2 Instance Store Volumes

Amazon EC2 Instance Store Volumes

- Amazon EC2 instance store volumes (also called ephemeral drives) provide temporary block-level storage for many Amazon EC2 instance types.
- This storage consists of a preconfigured and pre-attached block of disk storage on the same physical server that hosts the Amazon EC2 instance.
- The amount of this disk storage varies by Amazon EC2 instance type.
- In those Amazon EC2 instance families that provide instance storage, larger instances tend to provide both more and larger instance store volumes.
- Note that some instance types, such as the micro instances (t1) use Amazon EBS storage only, with no instance storage provided.
- In addition, the storage-optimized Amazon EC2 instance family provides special-purpose instance storage targeted to specific use cases.
- H1 instances provide very fast solid-state drive (SSD)-backed instance storage capable of supporting over 120,000 random read IOPS, and are optimized for very high random I/O performance and low cost per IOPS. In contrast, HS1 instances are optimized for very high storage density, low storage cost, and high sequential I/O performance

Ideal Usage Patterns

- In general, local instance store volumes are ideal for temporary storage of information that is continually changing, such as buffers, caches, scratch data, and other temporary content, or for data that is replicated across a fleet of instances, such as a load-balanced pool of web servers.
- Amazon EC2 instance storage is well-suited for this purpose.
- It consists of the virtual machine's boot device (for instance store AMIs only), plus one or more additional volumes that are dedicated to the Amazon EC2 instance (for both Amazon EBS AMIs and instance store AMIs).
- This storage is usable only from a single Amazon EC2 instance during its lifetime.
- Note that, unlike Amazon EBS volumes, instance store volumes cannot be detached or attached to another instance

Ideal Usage Patterns

- High I/O and high storage provide Amazon EC2 instance storage targeted to specific use cases.
- High I/O instances provide instance store volumes backed by SSD, and are ideally suited for many high performance database workloads.
- Example applications include NoSQL databases like Cassandra and MongoDB.
- High storage instances support much higher storage density per Amazon EC2 instance, and are ideally suited for applications that benefit from high sequential I/O performance across very large datasets.
- Example applications include data warehouses, Hadoop storage nodes, seismic analysis, cluster file systems, etc.

Ideal Usage Patterns

- High I/O and high storage provide Amazon EC2 instance storage targeted to specific use cases.
- High I/O instances provide instance store volumes backed by SSD, and are ideally suited for many high performance database workloads.
- Example applications include NoSQL databases like Cassandra and MongoDB.
- High storage instances support much higher storage density per Amazon EC2 instance, and are ideally suited for applications that benefit from high sequential I/O performance across very large datasets.
- Example applications include data warehouses, Hadoop storage nodes, seismic analysis, cluster file systems, etc.

Performance

- The non-SSD-based instance store volumes in most Amazon EC2 instance families have performance characteristics similar to standard Amazon EBS volumes.
- Because the Amazon EC2 instance virtual machine and the local instance store volumes are located in the same physical server, interaction with this storage is very fast, particularly for sequential access.
- To increase aggregate IOPS, or to improve sequential disk throughput, multiple instance store volumes can be grouped together using RAID 0 (disk striping) software.
- Because the bandwidth to the disks is not limited by the network, aggregate sequential throughput for multiple instance volumes can be higher than for the same number of Amazon EBS volumes

Performance

- The SSD instance store volumes in the Amazon EC2 high I/O instances provide from tens of thousands to hundreds of thousands of low-latency, random 4 KB random IOPS.
- The instance store volumes on Amazon EC2 high storage instances provide very high storage density and high sequential read and write performance.
- High storage instances are capable of delivering 2.6 GB/sec of sequential read and write performance when using a block size of 2 MB

Durability and Availability

- Amazon EC2 local instance store volumes are not intended to be used as durable disk storage.
- Unlike Amazon EBS volume data, data on instance store volumes persists only during the life of the associate Amazon EC2 instance.
- This means that data on instance store volumes is persistent across orderly instance reboots, but if the Amazon EC2 instance is stopped and re-started, terminates, or fails, all data on the instance store volumes is lost.

Durability and Availability

- You should not use local instance store volumes for any data that must persist over time, such as permanent file or database storage, without providing for data persistence by replicating your data, or by periodically copying data to durable storage such as Amazon EBS or Amazon S3.
- Note that this also applies to the special-purpose SSD and high density instance store volumes in the high I/O and high storage instance types

Cost Model

- The cost of the Amazon EC2 instance includes any local instance store volumes, if the instance type provides them.
- While there is no additional charge for data storage on local instance store volumes,
- Note that data transferred to and from Amazon EC2 instance store volumes from other Availability Zones or outside of an Amazon EC2 region may incur data transfer charges,
- additional charges will apply for use of any persistent storage, such as Amazon S3, Amazon Glacier, Amazon EBS volumes, and Amazon EBS snapshots.

Scalability and Elasticity

- The number and storage capacity of Amazon EC2 local instance store volumes are fixed and defined by the instance type.
- While you can't increase or decrease the number of instance store volumes on a single Amazon EC2 instance, this storage is still scalable and elastic, in that you can scale the total amount of instance store up or down by increasing or decreasing the number of running Amazon EC2 instances.
- Local instance store volumes are tied to a particular Amazon EC2 instance, and are fixed in number and size for a given Amazon EC2 instance type, so the scalability and elasticity of this storage is tied to the number of Amazon EC2 instances.

Interfaces

- There is no separate management API for Amazon EC2 instance store volumes.
- Instead, instance store volumes are specified using the block device mapping feature of the Amazon EC2 API and the AWS Management Console.
- You cannot create or destroy instance store volumes, but you can control whether or not they are exposed to the Amazon EC2 instance, and what device name is used.
- There is also no separate data API for instance store volumes.
- Just like Amazon EBS volumes, instance store volumes present a block-device interface to the Amazon EC2 instance.
- That is, to the Amazon EC2 instance, an instance store volume appears just like a local disk drive. To write to and read data from instance store volumes, you therefore use the native file system I/O interfaces of your chosen operating system.
- Note that in some cases, a local instance store volume device will be attached to the Amazon EC2 instance upon launch, but must be formatted with an appropriate file system and mounted before use.
- Also, keep careful track of your block device mappings. There is no simple way for an application running on an Amazon EC2 instance to determine which block device is an instance store (ephemeral) volume and which is an Amazon EBS (persistent) volume.

Anti-Patterns

- Persistent storage—If you need persistent virtual disk storage similar to a physical disk drive for files or other data that must persist longer than the lifetime of a single Amazon EC2 instance, Amazon EBS volumes or Amazon S3 are more appropriate.
- Relational database storage—In most cases, relational databases require storage that persists beyond the lifetime of a single Amazon EC2 instance, making Amazon EBS volumes the natural choice.
- Shared storage—Instance store volumes are dedicated to a single Amazon EC2 instance, and cannot be shared with other systems or users. If you need storage that can be detached from one instance and attached to a different instance, or if you need the ability to share data easily, Amazon S3 or Amazon EBS volumes are the better choice.
- Snapshots—If you need the convenience, long-term durability, availability, and shareability of point-in-time disk snapshots, Amazon EBS volumes are a better choice.