**Content**

[1. More information 2](#_Toc18159225)

[1. iSCSI 2](#_Toc18159226)

[2. EMR 2](#_Toc18159227)

[3. CRM 2](#_Toc18159228)

[2. FAQs 3](#_Toc18159229)

[1. Amazon Kinesis 3](#_Toc18159230)

[2.1.1. Amazon Kinesis Data Streams 3](#_Toc18159231)

[2.1.2. Amazon Kinesis Data Analytics 4](#_Toc18159232)

[2.1.3. Amazon Kinesis Data Firehose 4](#_Toc18159233)

[2.1.4. Amazon Kinesis Video Streams 4](#_Toc18159234)

[2. AutoScaling 4](#_Toc18159235)

[3. Route53 7](#_Toc18159236)

[2.3.1. DNS records, CNAME, A-record, URL, alias and friends 7](#_Toc18159237)

[2.3.2. More information 8](#_Toc18159238)

1. More information
   1. iSCSI

In computing, iSCSI is an acronym for Internet Small Computer Systems Interface, an Internet Protocol (IP)-based storage networking standard for linking data storage facilities. It provides block-level access to storage devices by carrying SCSI commands over a TCP/IP network. iSCSI is used to facilitate data transfers over intranets and to manage storage over long distances. It can be used to transmit data over local area networks (LANs), wide area networks (WANs), or the Internet and can enable location-independent data storage and retrieval.

* 1. EMR

Amazon EMR is the industry leading cloud-native big data platform, allowing teams to process vast amounts of data quickly, and cost-effectively at scale. Using open source tools such as Apache Spark, Apache Hive, Apache HBase, Apache Flink, and Presto, coupled with the dynamic scalability of Amazon EC2 and scalable storage of Amazon S3, EMR gives analytical teams the engines and elasticity to run Petabyte-scale analysis for a fraction of the cost of traditional on-premise clusters. Developers and analysts can use Jupyter-based EMR Notebooks for iterative development, collaboration, and access to data stored across AWS data products such as Amazon S3, Amazon DynamoDB, and Amazon Redshift to reduce time to insight and quickly operationalize analytics.

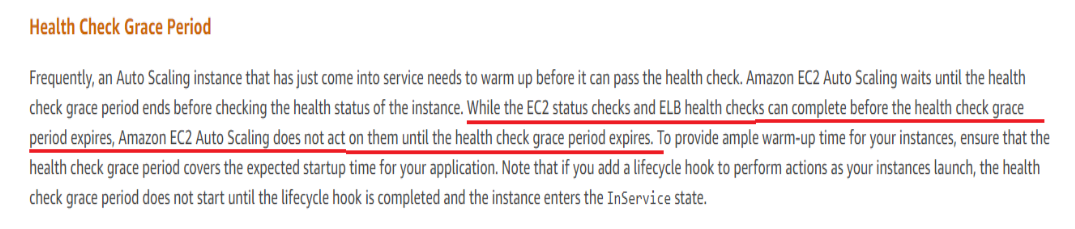
* 1. CRM

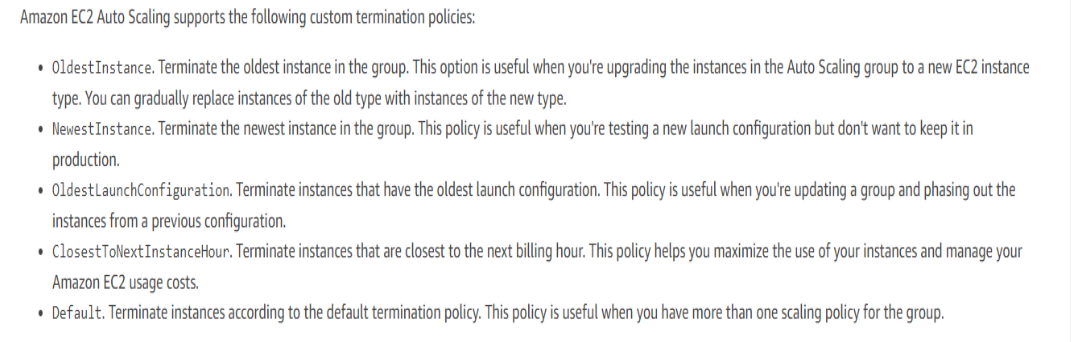
**Customer relationship management** (CRM) is an approach to manage a company's interaction with current and potential customers. It uses data analysis about customers' history with a company to improve business relationships with customers, specifically focusing on customer retention and ultimately driving sales growth

1. FAQs

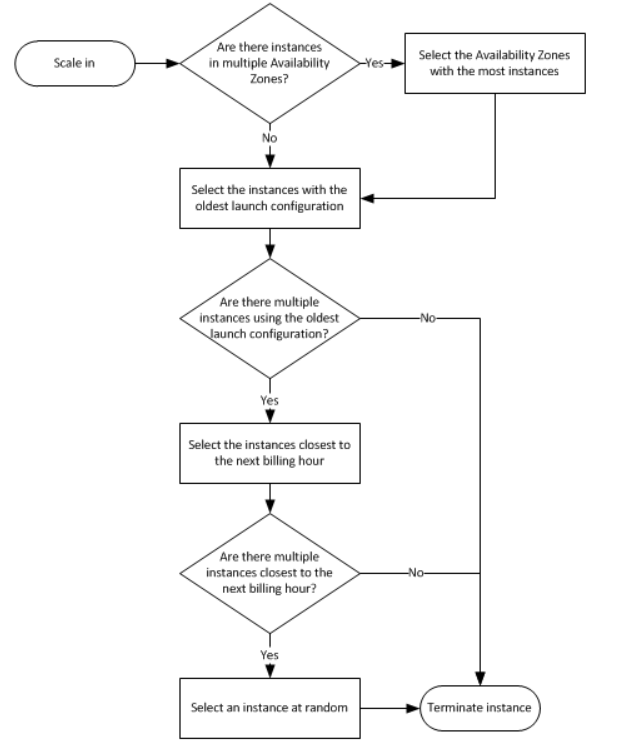
Please read below mentioned list of FAQs: From Compute: - EC2 - Auto Scaling - Lamdba - Elastic Load Balancing Storage: - S3 - EBS - EFS - Glacier - Storage gateway Database: - RDS - Aurora - DynamoDB -RedShift - ElastiCache Networking and CDN: - VPC - CloudFront - Route53 - Direct Connect. Management Tools - CloudWatch - CloudTrail Media Services and Analytics - Kinesis - EMR Security, Identity & Compliance - IAM - Key Management Service Application Integration. - MQ - SQS - SNS - SWF - Steps Link of all FAQs https://aws.amazon.com/faqs/ null

* 1. Amazon Kinesis
     1. Amazon Kinesis Data Streams
* Amazon Kinesis Data Streams enables you to build custom applications that process or analyze streaming data for specialized needs. You can continuously add various types of data such as clickstreams, application logs, and social media to an Amazon Kinesis data stream from hundreds of thousands of sources. Within seconds, the data will be available for your Amazon Kinesis Applications to read and process from the stream.
* Amazon Kinesis Data Streams synchronously replicates data across three availability zones, providing high availability and data durability.
* Amazon Kinesis Data Streams is useful for rapidly moving data off data producers and then continuously processing the data, be it to transform the data before emitting to a data store, run real-time metrics and analytics, or derive more complex data streams for further processing.
  + Accelerated log and data feed intake
  + Real-time metrics and reporting
  + Real-time data analytics
  + Complex stream processing
* By default, Records of a stream are accessible for up to 24 hours from the time they are added to the stream. You can raise this limit to up to 7 days by enabling extended data retention.
* The maximum size of a data blob (the data payload before Base64-encoding) within one record is 1 megabyte (MB).
* Each shard can support up to 1000 PUT records per second.
* A **shard** is a uniquely identified sequence of data records in a stream. A stream is composed of one or more shards, each of which provides a fixed unit of capacity. Each shard can support up to 5 transactions per second for reads, up to a maximum total data read rate of 2 MB per second and up to 1,000 records per second for writes, up to a maximum total data write rate of 1 MB per second (including partition keys). The data capacity of your stream is a function of the number of shards that you specify for the stream. The total capacity of the stream is the sum of the capacities of its shards. If your data rate increases, you can increase or decrease the number of shards allocated to your stream
  + 1. Amazon Kinesis Data Analytics
    2. Amazon Kinesis Data Firehose
    3. Amazon Kinesis Video Streams
  1. AutoScaling

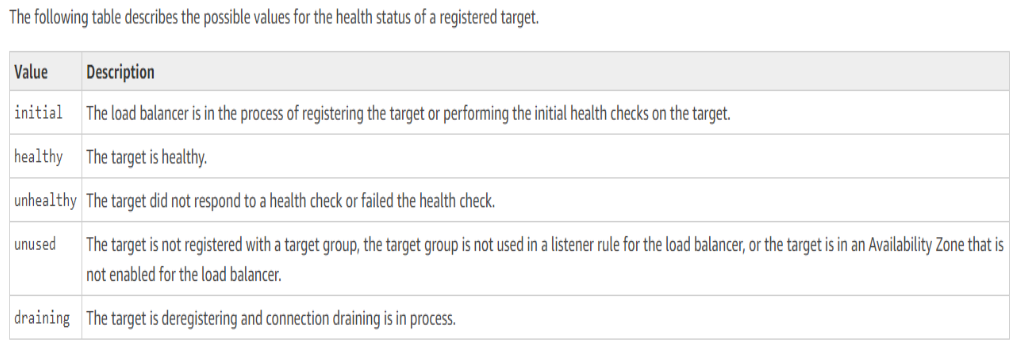


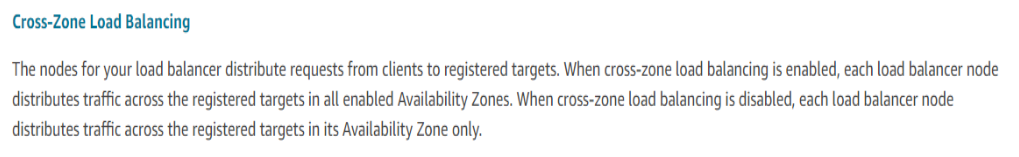


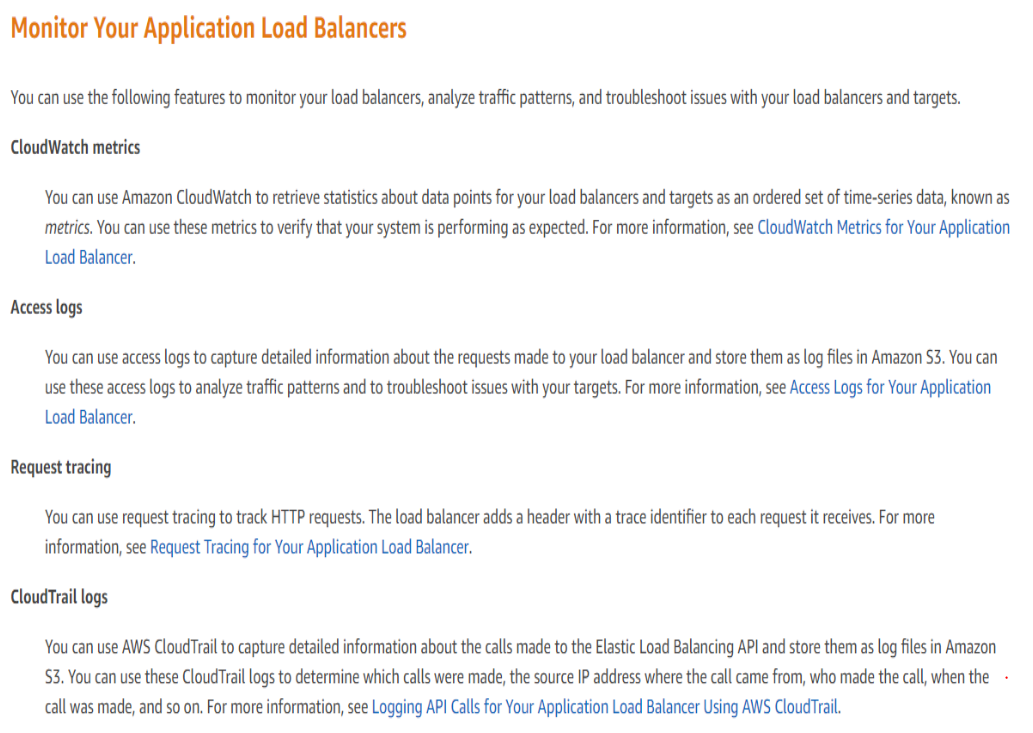
Default termination policy:

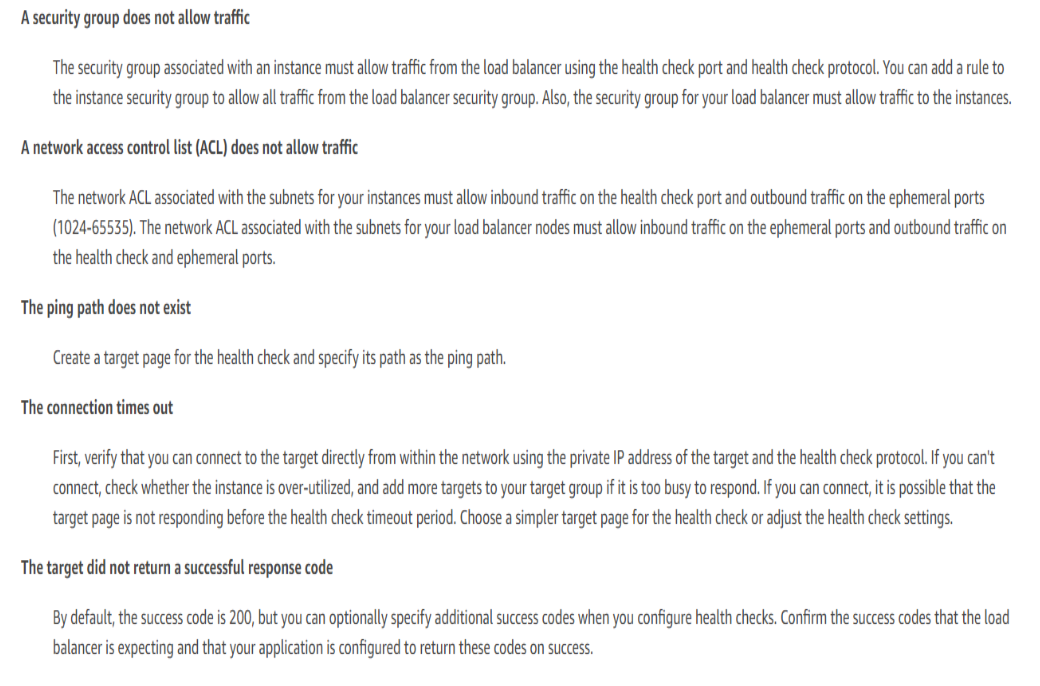


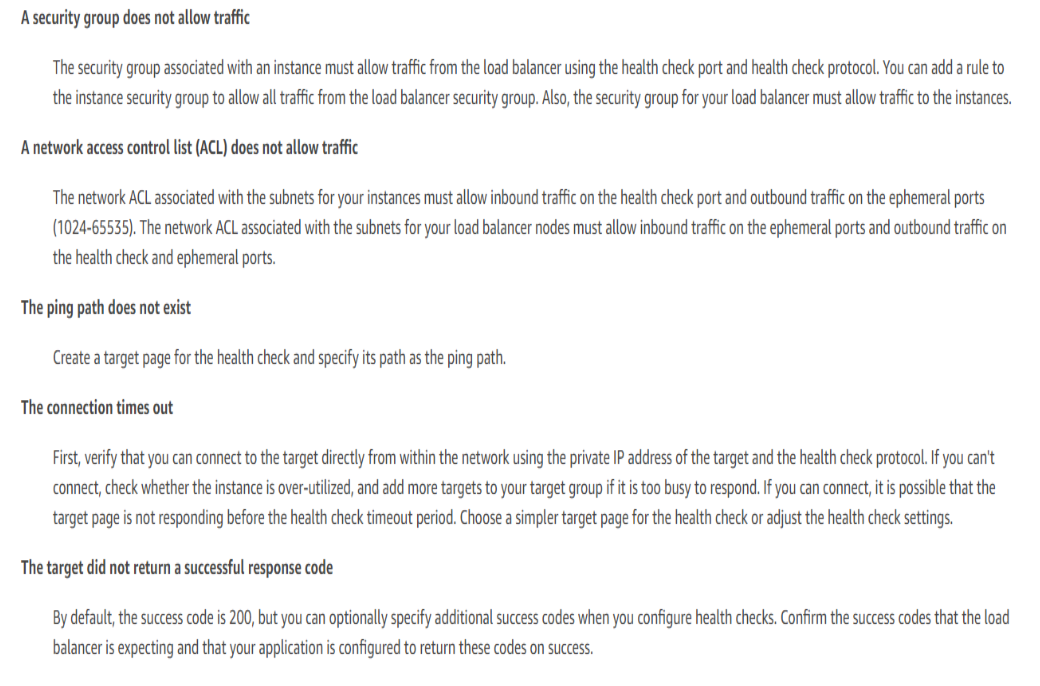
ELB health checks

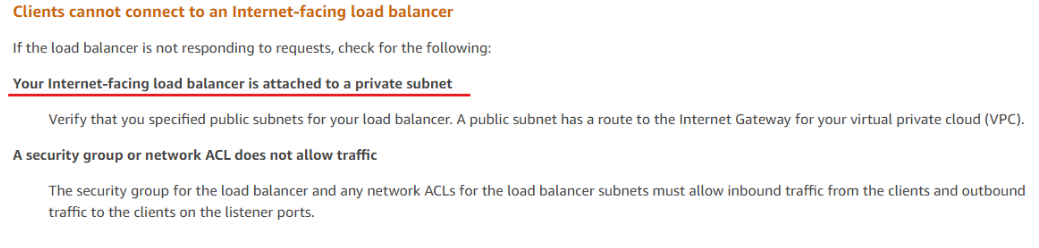




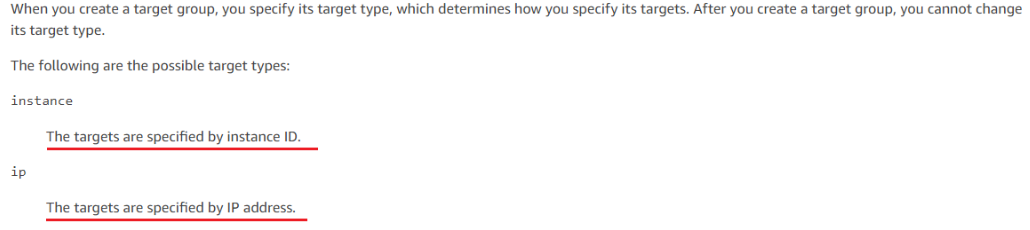








* If the LoadBalancer is internet-facing automatically gets an EIP, they are highly available and scalable components managed by AWS.



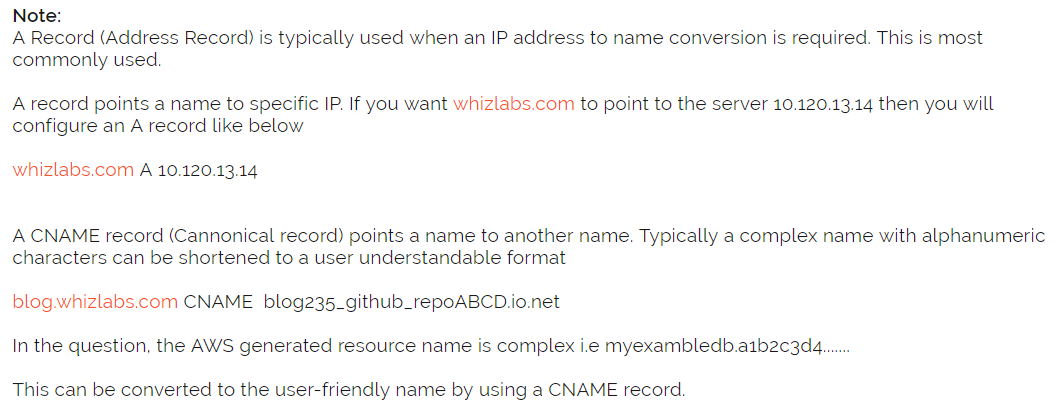
**The cooldown period** helps to ensure that your Auto Scaling group doesn't launch or terminate additional instances before the previous scaling activity takes effect. You can configure the length of time based on your instance warmup period or other application needs

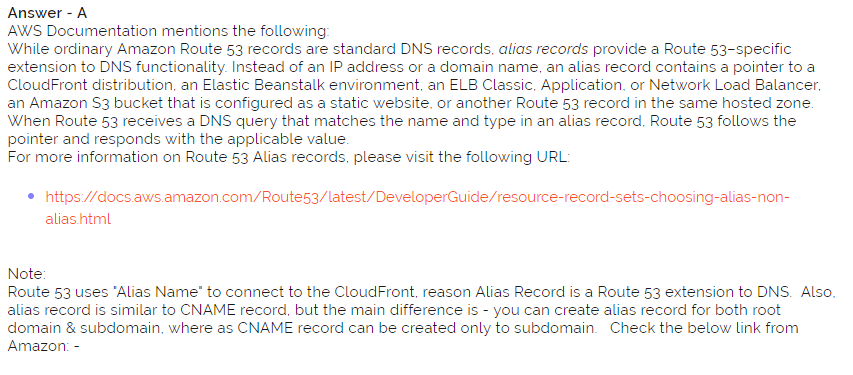
* 1. Route53
     1. DNS records, CNAME, A-record, URL, alias and friends

The A and CNAME records are the two common ways to map a host name (“name”) to one or more IP addresses. There are important differences between these two records.

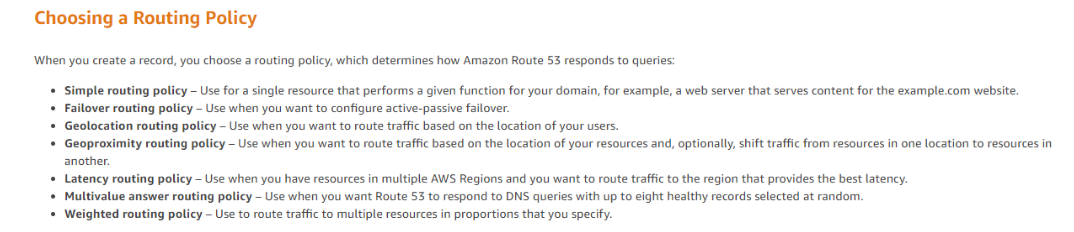


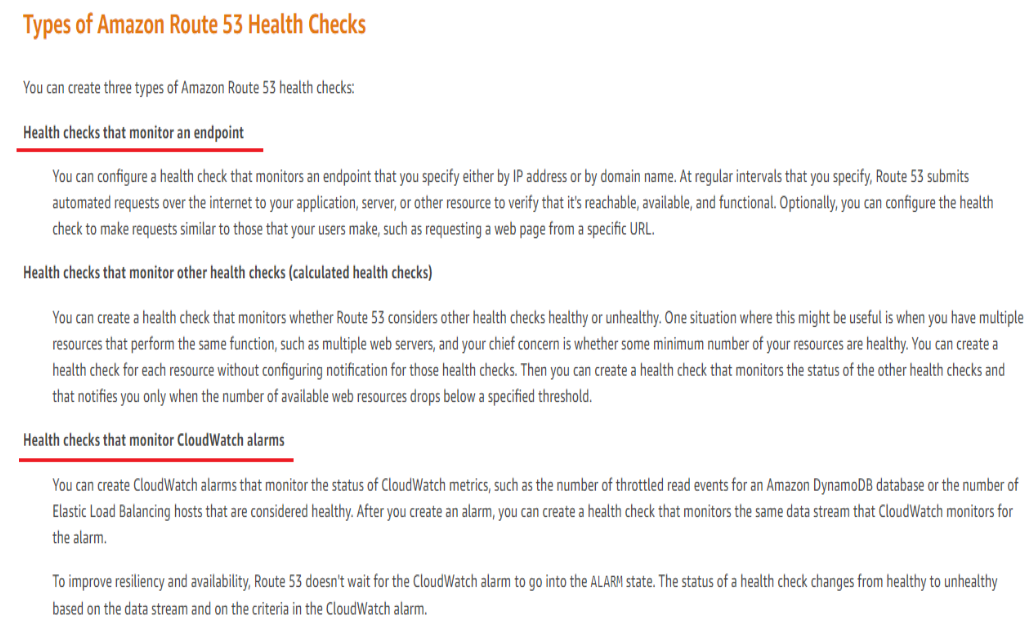
* You should never use a CNAME record for your root domain name (e.g. example.com).



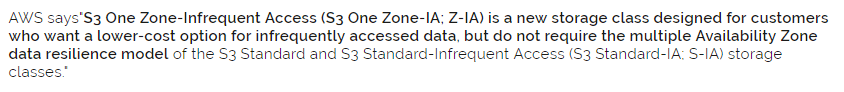


* + 1. More information





* 1. S3





* + 1. Amazon S3 Intelligent-Tiering (S3 Intelligent-Tiering)

It is an S3 storage class for data with unknown access patterns or changing access patterns that are difficult to learn. It is the first cloud storage class that delivers automatic cost savings by moving objects between two access tiers when access patterns change. One tier is optimized for frequent access and the other lower-cost tier is designed for infrequent access.

Objects uploaded or transitioned to S3 Intelligent-Tiering are automatically stored in the frequent access tier. S3 Intelligent-Tiering works by monitoring access patterns and then moving the objects that have not been accessed in 30 consecutive days to the infrequent access tier. If the objects are accessed later, S3 Intelligent-Tiering moves the object back to the frequent access tier. This means all objects stored in S3 Intelligent-Tiering are always available when needed. There are no retrieval fees, so you won’t see unexpected increases in storage bills when access patterns change.

It is for data with unknown access patterns or changing access patterns that are difficult to learn. It is ideal for data sets where you may not be able to anticipate access patterns. S3 Intelligent-Tiering can also be used to store new data sets where, shortly after upload, access is frequent, but decreases as the data set ages. Then you can move the data set to S3 One Zone-IA or archive it to S3 Glacier.

There are two ways to get data into S3 Intelligent-Tiering. You can directly PUT into S3 Intelligent-Tiering by specifying INTELLIGENT\_TIERING in the x-amz-storage-class header or set lifecycle policies to transition objects from S3 Standard or S3 Standard-IA to S3 INTELLIGENT\_TIERING.

* + 1. S3 Standard-Infrequent Access (S3 Standard-IA)

S3 Standard-IA is ideal for data that is accessed less frequently, but requires rapid access when needed. S3 Standard-IA is ideally suited for long-term file storage, older sync and share storage, and other aging data.

storage class for data that is accessed less frequently but requires rapid access when needed. S3 Standard-IA offers the high durability, throughput, and low latency of the Amazon S3 Standard storage class, with a low per-GB storage price and per-GB retrieval fee. This combination of low cost and high performance make S3 Standard-IA ideal for long-term storage, backups, and as a data store for disaster recovery. The S3 Standard-IA storage class is set at the object level and can exist in the same bucket as the S3 Standard or S3 One Zone-IA storage classes, allowing you to use S3 Lifecycle policies to automatically transition objects between storage classes without any application changes.

* + 1. S3 One Zone-Infrequent Access (S3 One Zone-IA)

Customers can use S3 One Zone-IA for infrequently-accessed storage, like backup copies, disaster recovery copies, or other easily re-creatable data.

S3 One Zone-IA storage class is an Amazon S3 storage class that customers can choose to store objects in a single availability zone. S3 One Zone-IA storage redundantly stores data within that single Availability Zone to deliver storage at 20% less cost than geographically redundant S3 Standard-IA storage, which stores data redundantly across multiple geographically separate Availability Zones.

S3 One Zone-IA offers a 99% available SLA and is also designed for eleven 9’s of durability within the Availability Zone. But, unlike the S3 Standard and S3 Standard-IA storage classes, data stored in the S3 One Zone-IA storage class will be lost in the event of Availability Zone destruction.

S3 One Zone-IA storage offers the same Amazon S3 features as S3 Standard and S3 Standard-IA and is used through the Amazon S3 API, CLI and console. S3 One Zone-IA storage class is set at the object level and can exist in the same bucket as S3 Standard and S3 Standard-IA storage classes. You can use S3 Lifecycle policies to automatically transition objects between storage classes without any application changes.

* + 1. Glacier

When processing a retrieval job, Amazon S3 first retrieves the requested data from S3 Glacier, and then creates a temporary copy of the requested data in S3 (which typically takes a few minutes). The access time of your request depends on the retrieval option you choose: Expedited, Standard, or Bulk retrievals. For all but the largest objects (250MB+), data accessed using Expedited retrievals are typically made available within 1-5 minutes. Objects retrieved using Standard retrievals typically complete between 3-5 hours. Bulk retrievals typically complete within 5-12 hours. For more information about S3 Glacier retrieval options, please refer to the S3 Glacier FAQs.

* + 1. Query in place / S3 select

Amazon S3 allows customers to run sophisticated queries against data stored without the need to move data into a separate analytics platform. The ability to query this data in place on Amazon S3 can significantly increase performance and reduce cost for analytics solutions leveraging S3 as a data lake. S3 offers multiple query in place options, including S3 Select, Amazon Athena, and Amazon Redshift Spectrum, allowing you to choose one that best fits your use case. You can even use Amazon S3 Select with AWS Lambda to build serverless apps that can take advantage of the in-place processing capabilities provided by S3 Select

S3 Select is an Amazon S3 feature that makes it easy to retrieve specific data from the contents of an object using simple SQL expressions without having to retrieve the entire object. You can use S3 Select to retrieve a subset of data using SQL clauses, like SELECT and WHERE, from objects stored in CSV, JSON, or Apache Parquet format. It also works with objects that are compressed with GZIP or BZIP2 (for CSV and JSON objects only), and server-side encrypted objects.

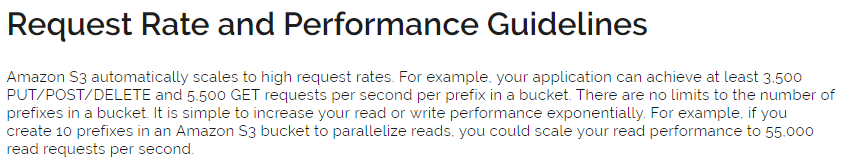
* + 1. S3 Transfer Acceleration

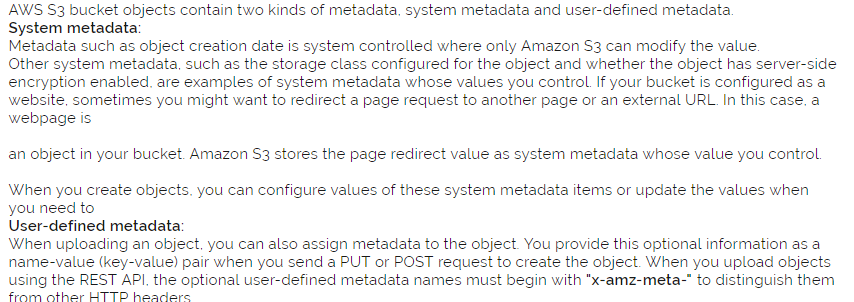
Amazon S3 Transfer Acceleration enables fast, easy, and secure transfers of files over long distances between your client and your Amazon S3 bucket. S3 Transfer Acceleration leverages Amazon CloudFront’s globally distributed AWS Edge Locations. As data arrives at an AWS Edge Location, data is routed to your Amazon S3 bucket over an optimized network path.

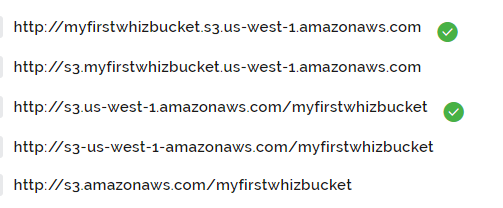
* + 1. Others

S3 bucket properties:

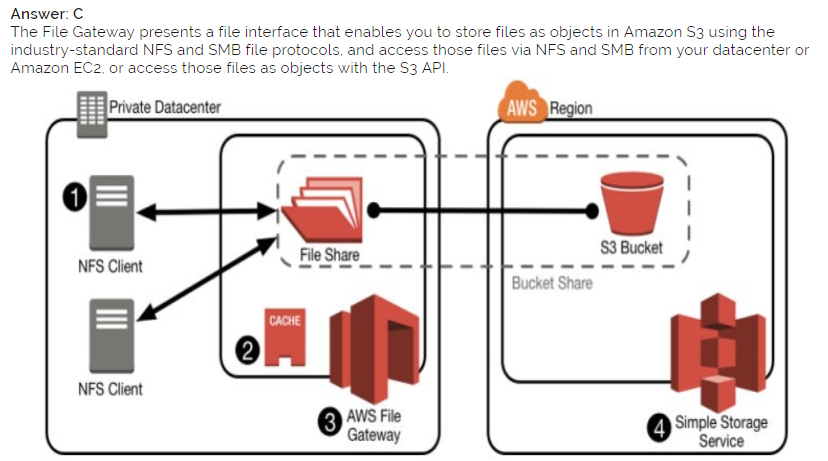
1. Versioning
2. Server access logging
3. Static website hosting
4. Object-level logging
5. Tags
6. Transfer acceleration
7. Events







* + 1. AWS Storage Gateway
* The **file gateway** enables you to store and retrieve objects in Amazon S3 using file protocols, such as NFS. Objects written through file gateway can be directly accessed in S3.
* The **volume gateway** provides block storage to your applications using the iSCSI protocol. Data on the volumes is stored in Amazon S3. To access your iSCSI volumes in AWS, you can take EBS snapshots which can be used to create EBS volumes.
* The **tape gateway** provides your backup application with an iSCSI virtual tape library (VTL) interface, consisting of a virtual media changer, virtual tape drives, and virtual tapes. Virtual tape data is stored in Amazon S3 or can be archived to Amazon Glacier.
  + - 1. File Gateway



Use cases for file gateway include: (a) migrating on-premises file data to Amazon S3, while maintaining fast local access to recently accessed data, (b) Backing up on-premises file data as objects in Amazon S3 (including Microsoft SQL Server and Oracle databases and logs), with the ability to use S3 capabilities such as lifecycle management, versioning and cross region replication, and, (c) Hybrid cloud workflows using data generated by on-premises applications for processing by AWS services such as machine learning, big data analytics or serverless functions.

File gateway enables your existing file-based applications, devices, and workflows to use Amazon S3, without modification. File gateway securely and durably stores both file contents and metadata as objects, while providing your on-premises applications low-latency access to cached data.

* + - 1. Volume Gateway

Volume gateway provides an iSCSI target, which enables you to create block storage volumes and mount them as iSCSI devices from your on-premises or EC2 application servers. The volume gateway runs in either a cached or stored mode.

1. **Cached Volume Gateway**

Your primary data is written to S3, while retaining your frequently accessed data locally in a cache for low-latency access. So, you store your data in S3 and retain a copy of frequently accessed data subsets locally.

1. **Stored Volume Gateway**

Your primary data is stored locally and your entire dataset is available for low-latency access while asynchronously backed up to AWS. So, you store the entire set of volume data on-premises and store periodic point-in-time backups (snapshots) in AWS.

In either mode, you can take point-in-time snapshots of your volumes, which are stored as Amazon EBS Snapshots in AWS, enabling you to make space-efficient versioned copies of your volumes for data protection, recovery, migration and various other copy data needs

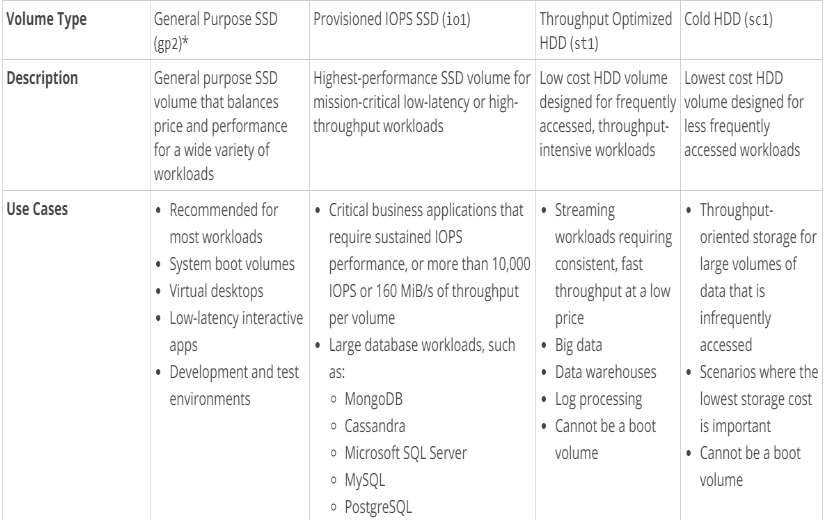
* + - 1. Tape Gateway

Tape Gateway, a cloud-based virtual tape library feature of AWS Storage Gateway, now integrates with S3 Glacier Deep Archive, enabling you to store your virtual tape-based, long-term backups and archives in S3 Glacier Deep Archive, thereby providing the lowest cost storage for this data in the cloud. To get started, create a new virtual tape using AWS Storage Gateway Console or API, and set the archival storage target either to S3 Glacier or S3 Glacier Deep Archive. When your backup application ejects the tape, the tape will be archived to your selected storage target.

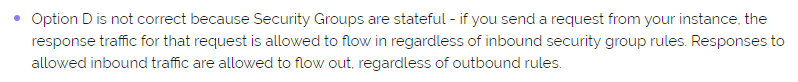
* + - 1. Custom built solution using S3

Not recommended ?!

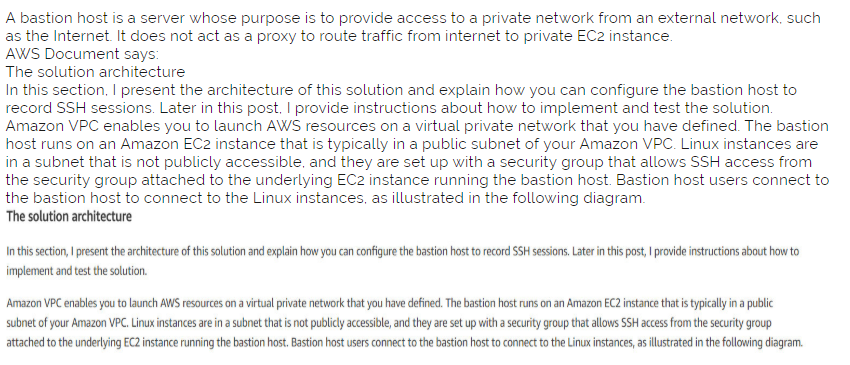
1. EBS



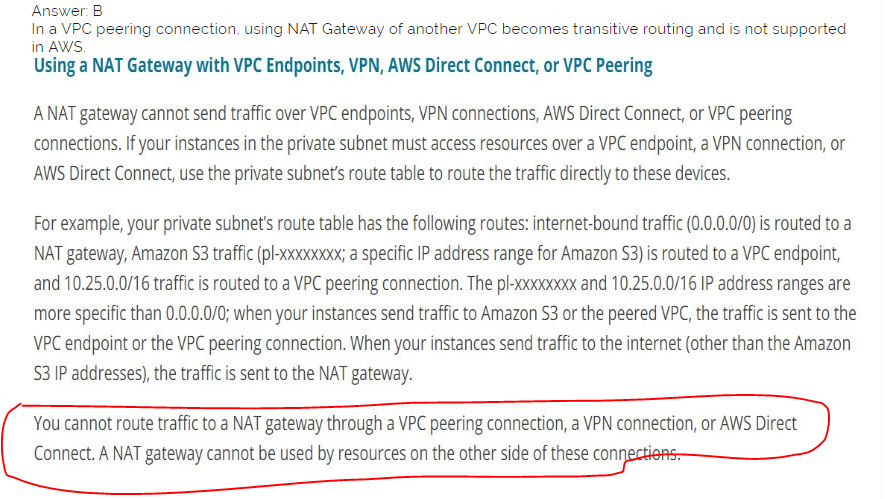
1. VPC



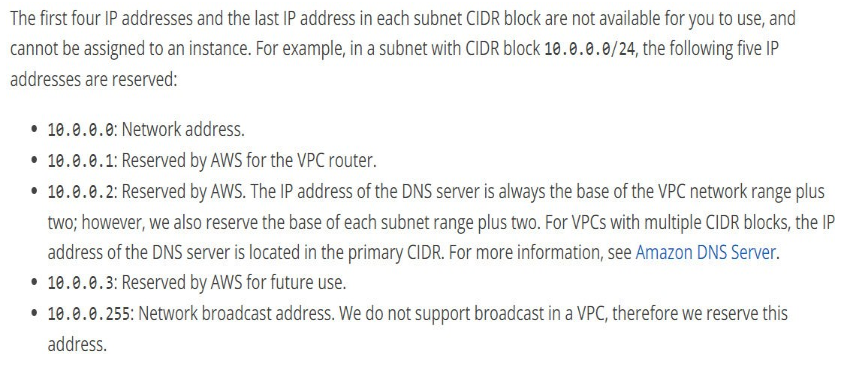
* 1. Bastion Host



* 1. NAT



NAT Gateways auto scale based on demand.



1. Others
   1. Aurora

Multiple masters - Amazon Aurora Multi-Master is now generally available, allowing you to create multiple read-write instances of your Aurora database across multiple Availability Zones, which enables uptime-sensitive applications to achieve continuous write availability through instance failure. In the event of instance or Availability Zone failures, Aurora Multi-Master enables the Aurora database to maintain read and write availability with zero application downtime. With Aurora Multi-Master, there is no need for database failovers to resume write operations. Check out this blog to find out how to build highly available MySQL applications using Aurora Multi-Master.

Replication lag – Aurora has much lower latency between the master and any other nodes in the cluster. IF you need low lag replication then this is the best solution.

Dynamic scaling – Aurora and DynamoDB have scaling on demand (dynamic) option.

* 1. CloudHSM

AWS CloudHSM provides hardware security modules in the AWS Cloud. A hardware security module (HSM) is a computing device that processes cryptographic operations and provides secure storage for cryptographic keys.

* Generate, store, import, export, and manage cryptographic keys, including symmetric keys and asymmetric key pairs.
* Use symmetric and asymmetric algorithms to encrypt and decrypt data.
* Use cryptographic hash functions to compute message digests and hash-based message authentication codes (HMACs).
* Cryptographically sign data (including code signing) and verify signatures.
* Generate cryptographically secure random data.
  + 1. Ephemeral Backup Key (EBK) and Persistent Backup Key (PBK)

When AWS CloudHSM makes a backup from the HSM, the HSM encrypts all of its data before sending it to AWS CloudHSM. The data never leaves the HSM in plaintext form.

To encrypt its data, the HSM uses a unique, ephemeral encryption key known as the ephemeral backup key (EBK). The EBK is an AES 256-bit encryption key generated inside the HSM when AWS CloudHSM makes a backup. The HSM generates the EBK, then uses it to encrypt the HSM's data with a FIPS-approved AES key wrapping method that complies with NIST special publication 800-38F. Then the HSM gives the encrypted data to AWS CloudHSM. The encrypted data includes an encrypted copy of the EBK.

To encrypt the EBK, the HSM uses another encryption key known as the persistent backup key (PBK). The PBK is also an AES 256-bit encryption key.

* 1. Polly

Amazon Polly is a service that turns text into lifelike speech, allowing you to create applications that talk, and build entirely new categories of speech-enabled products. Amazon Polly is a Text-to-Speech (TTS) service that uses advanced deep learning technologies to synthesize speech that sounds like a human voice.

* + 1. Lexicons

Pronunciation lexicons enable you to customize the pronunciation of words. Amazon Polly provides API operations that you can use to store lexicons in an AWS region. Those lexicons are then specific to that particular region. You can use one or more of the lexicons from that region when synthesizing the text by using the SynthesizeSpeech operation. This applies the specified lexicon to the input text before the synthesis begins.

* 1. CloudFront
     1. Origin Access Identity

To restrict access to content that you serve from Amazon S3 buckets, you create CloudFront signed URLs or signed cookies to limit access to files in your Amazon S3 bucket, and then you create a special CloudFront user called an origin access identity (OAI) and associate it with your distribution. Then you configure permissions so that CloudFront can use the OAI to access and serve files to your users, but users can't use a direct URL to the S3 bucket to access a file there. Taking these steps help you maintain secure access to the files that you serve through CloudFront.

* 1. CloudTrail best practices
     1. Create a trail

Other way it stores info only for 90 days.

* + 1. Apply trails to all AWS Regions

When new region is released it is automatically applied.

* + 1. Log file integrity validation feature

To determine **whether a log file was modified, deleted, or unchanged after CloudTrail delivered it**, you can use CloudTrail log file integrity validation. This feature is built using industry standard algorithms: SHA-256 for hashing and SHA-256 with RSA for digital signing. This makes it computationally infeasible to modify, delete or forge CloudTrail log files without detection.

* + 1. Global service events

For most services, events are recorded in the region where the action occurred. For **global services** such as AWS Identity and Access Management (IAM), AWS STS, Amazon CloudFront, and Route 53, events **are delivered to any trail that includes global services, and are logged as occurring in US East (N. Virginia) Region**

* + 1. Integrate with Amazon CloudWatch Logs

CloudWatch Logs allows you to **monitor and receive alerts for specific events captured by CloudTrail.** The events sent to CloudWatch Logs are those configured to be logged by your trail, so make sure you have configured your trail or trails to log the event types (management events and/or data events) that you are interested in monitoring.

* 1. EFS

Amazon EFS is a file storage service for use with Amazon EC2. Amazon EFS provides a file system interface, file system access semantics (such as strong consistency and file locking), and concurrently-accessible storage for up to thousands of Amazon EC2 instances.

To simplify accessing your EFS file systems, we recommend using the EFS mount helper utility. Once mounted, you can work with the files and directories in your file system just like you would with a local file system. Amazon EFS uses the Network File System version 4 (NFS v4) protocol. Amazon EFS is compatible with all Linux-based AMIs for Amazon EC2. You can mix and match the instance types connected to a single file system.

Amazon EFS offers a Standard and an Infrequent Access storage class. Moving files to EFS IA starts by enabling EFS Lifecycle Management and choosing an age-off policy. Lifecycle Management automatically moves your data to the EFS IA storage class according to the lifecycle policy you choose. For example, you can automatically move files into EFS IA fourteen days of not being accessed.

To access EFS file systems from on-premises, you must have an AWS Direct Connect or AWS VPN connection between your on-premises datacenter and your Amazon VPC.

You can mount an EFS on instances in only one VPC at a time

Both the file systems and VPC must be in the same AWS Region

When using EFS, you specify EC2 security groups for the EFS mount targets associated with the file system. Security Groups act as a firewall, and the rules you add define the traffic flow. You can authorize inbound and outbound access to your EFS system. To do so, you add rules that allow your EC2 instance to connect to your EFS file system through the mount target using the NFS port.

22(SSH), 2049 NFS