AWS Solution Architect Associate Exam Prep - Notes

Table of Contents

AWS Solution Architect Associate Exam Prep - Notes

EC2

1. EC2 instance and snapshot's will be 17 characters long in the new formats
2. EC2 has a soft limit of 20 Reserved instances or 20 or less On-Demand instances per account – can be increased upon request. Certain instance types are further limited to a smaller number than 20 (ex: r4.12xlarge is limited to a maximum of 1 On-Demand instance per account)
3. A maximum of 5 EIPs (elastic Ips) are allowed per region
4. AWS uses a Xen based hypervisor for EC2 visualization
5. EC2 enhanced networking provides higher packet per second performance and lower latency – there is no additional fee for using enhanced networking feature – supported only on some instance families
6. EC2 reserved instances require 1 or 3 year commitment with significant discounts
7. Dr Mc GIFT PX – Dense, RAM (Memory), Main (General Purpose), Graphics, IOPS, FPGA (field programmable), Low Cost (T2), Picture (Graphics General Purpose), X-tra Large RAM
8. You can replace the keypair for an EC2 instance with a new one or add another keypair. This may be required if someone who has access to the key has left your organization. Similarly you can also add another keypair to the instance if someone else wants access to the instance.
9. If you lose the private key for an EBS-backed instance, you can regain access to your instance. You must stop the instance, detach its root volume and attach it to another instance as a data volume, modify the authorized_keys file, move the volume back to the original instance, and restart the instance. This procedure isn't supported for instance store-backed instances.
10. HVM & EBS backed ---> DR Mc GIFT PX (ALL)

    HVM & Instance Store-backed ---> M3 C3 XR I2 D2

    PV & EBS / Instance store-backed ---> M3 C3 only

Instance Family	HVM EBS-Backed 64-bit	HVM Instance Store 64-bit	PV EBS-Backed 64-bit	PV Instance Store 64-bit
T2	□			
M4	□			
M3	□	□	□	□
C4	□			
C3	□	□	□	□
X1	□	□		
R4	□			
R3	□	□		
P2	□			
G3	□			
I2	□	□		
D2	□	□		
F1	□			

11.

EBS

1. The maximum IOPS ratio to the volume size is 50:1
2. For gp2 EBS volumes, minimum value for IOPS that is defaulted to is 100 and minimum size that can be specified is 4GB
3. For io1 EBS volumes, maximum IOPS that can be specified is 20,000 (raised to 32,000 in 11/2017)
4. With Amazon EBS, you can create point-in-time snapshots of volumes, which are stored for you in Amazon S3. After you've created a snapshot and it has finished copying to Amazon S3 (when the snapshot status is completed), you can copy it from one AWS region to another, or within the same region. Amazon S3 server-side encryption (256-bit AES) protects a snapshot's data in-transit during a copy operation. The snapshot copy receives a new ID that is different than the ID of the original snapshot.
5. 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***.
6. 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.
7. A snapshot is ***constrained to the region*** where it was created. After you create a snapshot of an EBS volume, you can use it to create new volumes in the same region.
8. Provisioned IOPS SSD ***(io1)*** volumes automatically send ***one-minute metrics*** to CloudWatch.
9. Data on an instance store is lost when an EC2 instance is stopped or terminated. However it is preserved on instance reboot.
10. Data on an EBS volume attached to an instance persists even after the instance is stopped or terminated. By default the EBS backed root volume is deleted upon instance termination. You can change value of the DeleteOnTermination attribute for a volume when you ***launch the instance or while the instance is running***.
11. AWS supports 2 Block Storage Types – Instance Store and EBS backed block storage
12. EBS – SSD or HDD – SSD (gp2 or io1); HDD (sc1 or st1)
13. gp2 capacity range – 1 GB to 16 TB ⟶ 100 IOPS to 10,000 IOPS (burst of 3000 IOPS)⟶ Minimum 100 IOPS
14. io1 capacity range – 4 GB to 16 TB ⟶ 99.9% performance consistency ⟶ Process I/O in 256 KB chunks ⟶ preferred for > 10,000 IOPS upto 20,000 IOPS ⟶ Volume Size to IOPS ratio maximum is 1:50) ⟶ Minimum 100 IOPS
15. Through optimized volumes (st1) process I/O in 1 MB chunks

16. sc1 & st1 capacity range – 500 GB to 16 TB

S3

1. AWS S3 bucket names must be at least 3 chars long and less than 63 characters
2. As a best practice AWS recommends to always use DNS-compliant bucket names regardless of the Region in which you create the bucket
3. Bucket names are global and must be unique across all AWS accounts
4. Bucket names – maximum 63 characters – lowercase letters, numbers, periods and hyphens
5. Maximum of 100 buckets per account (soft limit)
6. S3 objects – 0 bytes – 5 TB
7. S3 object keys can be upto 1024 bytes long (unicode UTF8 characters)
8. Types – Standard, Standard – IA Infrequent Access, S3 – RRS Reduced Redundancy Storage & S3 Glacier
9. S3 supports HTTPS endpoints for in-transit encryption
10. S3 supports at rest data encryption using:

    SSE – S3 : Server-side Encryption – S3 (AWS managed keys)

    SSE – KMS : Server-side Encryption KMS (AWS KMS managed keys)

    SSE – C : Server-side Encryption Customer Provided Keys

    Client side Encryption
11. S3 requires a multi-part upload for objects greater than 5 GB in size
12. Versioning must be enabled for cross-region replication – data, metadata and ACLs are replicated

    Must use IAM policy to give S3 permission to replicate

    Cross-region replication can be used to reduce latency

    When turned ON, existing objects will not be replicated
13. S3 Event notification → Setup at the bucket level; allows to run workflows, send alerts, perform transcoding, process files or sync with other data sources in response to actions on S3 objects → Can Send Events to SQS, SNS or AWS Lambda
14. Amazon Glacier used for long term archival of data and is a replacement for Tape Backups as well as compliance needs
15. Data stored as TAR or ZIP files → upto 40 TB in size; Vaults contain archives → 1000 Vaults per account → control

access using IAM policies

16. The two general forms of an Amazon S3 website endpoint are as follows:

    bucket-name.s3-website-region.amazonaws.com <--- when using US-East Region

    Ex: HTTP://example-bucket.**s3-website-us-east-1**.amazonaws.com/

    bucket-name.**s3-website.region**.amazonaws.com <--- when using other Regions

    Ex: http://example-bucket.s3-website.eu-central-1.amazonaws.com/

17. Amazon S3 Transfer Acceleration enables fast, easy, and secure transfers of files over long distances between your client and your Amazon S3 bucket. 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.

18. Your data transfer application must use one of the following two types of endpoints to access the bucket for faster data transfer: <bucketname>.s3-accelerate.amazonaws.com or <bucketname>.s3-accelerate.dualstack.amazonaws.com for the "dual-stack" endpoint. Amazon S3 dual-stack endpoints support requests to S3 buckets over IPv6 and Ipv4.

19. If your workload in an Amazon S3 bucket routinely exceeds 100 PUT/LIST/DELETE requests per second or more than 300 GET requests per second,introduce some randomness in your key name prefixes, the key names, and therefore the I/O load, will be distributed across more than one partition.

SQS

1. SQS ***default visibility timeout is 30 seconds*** and can be a ***maximum of 12 hours***
2. SQS ***minimum retention period is 1 minute, default retention period is 4 days and maximum is 14 days***
3. SQS ***minimum message size is 1 KB and maximum size is 256 KB***
4. Amazon SQS is a message queue service used by distributed applications to exchange messages through a polling model, and can be used to decouple sending and receiving components. Amazon SQS provides flexibility for distributed components of applications to send and receive messages without requiring each component to be concurrently available.
5. Amazon SQS and SNS are lightweight, fully managed message queue and topic services that scale almost infinitely and provide simple, easy-to-use APIs. You can use Amazon SQS and SNS to decouple and scale microservices, distributed systems, and serverless applications, and improve reliability.
6. FIFO (first-in-first-out) queues preserve the exact order in which messages are sent and received. If you use a FIFO queue, you don't have to place sequencing information in your messages.
7. Standard queues provide a loose-FIFO capability that attempts to preserve the order of messages. However,

because standard queues are designed to be massively scalable using a highly distributed architecture, receiving messages in the exact order they are sent is ***not guaranteed***

8. ***Standard queues*** provide ***at-least-once delivery***, which means that each message is delivered at least once. ***FIFO queues*** provide ***exactly-once processing***, which means that each message is delivered once and remains available until a consumer processes it and deletes it. Duplicates are not introduced into the queue.
9. All messages have a global unique ID that Amazon SQS returns when the message is delivered to the message queue. The ID isn't required to perform any further actions on the message, but it is useful for tracking the receipt of a particular message in the message queue. When you receive a message from the message queue, the response includes a receipt handle that you must provide when deleting the message.
10. An Amazon SQS message can contain up to 10 metadata attributes. You can use message attributes to separate the body of a message from the metadata that describes it. This helps process and store information with greater speed and efficiency because your applications don't have to inspect an entire message before understanding how to process it.
11. Amazon SQS long polling is a way to retrieve messages from your Amazon SQS queues. While the regular short polling returns immediately, even if the message queue being polled is empty, long polling doesn't return a response until a message arrives in the message queue, or the long poll times out.
12. In general, you should use maximum 20 seconds for a long-poll timeout. Because higher long-poll timeout values reduce the number of empty ReceiveMessageResponse instances returned, try to set your long-poll timeout as high as possible. Long polling timeout can be between 0 sec and 20 sec (maximum).
13. When Amazon SQS returns a message to you, the message stays in the message queue whether or not you actually receive the message. You're responsible for deleting the message and the deletion request acknowledges that you're done processing the message.
14. If you don't delete the message, Amazon SQS will deliver it again on when it receives another receive request.
15. When you issue a DeleteMessage request on a previously-deleted message, Amazon SQS returns a success response.
16. FIFO queues never introduce duplicate messages.
17. Both standard and FIFO queues support SSE.
18. AWS SQS is both PCI DSS level 1 certified and HIPAA compliant.
19. There is a 120,000 limit for the number of in-flight messages for a standard queue and 20,000 for a FIFO queue. Messages are in-flight after they have been received from the queue by a consuming component, but have not yet been deleted from the queue.
20. Each Amazon SQS message queue is independent within each region and ***can not be shared across regions***.

SNS

1. SNA is a cost-effective capability to publish messages from an application and immediately deliver them to subscribers or other applications. It is designed to make web-scale computing easier for developers.
2. Amazon SNS follows the "publish-subscribe" (pub-sub) messaging paradigm, with notifications being delivered to clients using a "push" mechanism that eliminates the need to periodically check or "poll" for new information and updates.
3. SNS service can support a wide variety of needs including event notification, monitoring applications, workflow systems, time-sensitive information updates, mobile applications, and any other application that generates or consumes notifications.
4. A common pattern is to use SNS to publish messages to Amazon SQS message queues to reliably send messages to one or many system components asynchronously.

SWF

1. Amazon SWF enables applications for a range of use cases, including media processing, web application back-ends, business process workflows, and analytic pipelines, to be designed as a coordination of tasks
2. ***Tasks are processed by workers*** which are programs that interact with Amazon SWF to get tasks, process them, and return their results. A worker implements an application processing step. You can build workers ***in different programming languages*** and even reuse existing components to quickly create the worker.
3. ***SWF ensures that a task is assigned only once and is never duplicated***
4. The ***maximum duration*** for a workflow within SWF is ***1 year***
5. AWS Flow Framework is a programming framework that enables you to develop Amazon SWF-based applications quickly and easily. It abstracts the details of task-level coordination and asynchronous interaction with simple programming constructs
6. Amazon SWF provides long-polling. Long-polling significantly reduces the number of polls that return without any tasks. When workers and deciders poll Amazon SWF for tasks, the connection is retained for a minute if no task is available. If a task does become available during that period, it is returned in response to the long-poll request.
7. With AWS SWF you can use any programming language to write a worker or a decider, as long as you can communicate with Amazon SWF using web service APIs
8. SWF Limits: 100 SWF Domains per account ; 10,000 workflow & activity types per domain
9. At any given time you can have 100,000 open executions in a domain

RDS / Elasticache

1. On a MySQL DB instance, avoid tables in your database growing too large. Provisioned storage limits restrict the maximum size of a MySQL table file to 16 TB. Instead, partition your large tables so that file sizes are well under the 16 TB limit.
2. Read Replicas are supported by Amazon Aurora, Amazon RDS for MySQL, MariaDB and PostgreSQL. Unlike Multi-AZ deployments, Read Replicas for these engines use each's built-in replication technology and are subject to its strengths and limitations. In particular, updates are applied to your Read Replica(s) after they occur on the source DB instance ("asynchronous" replication), and replication lag can vary significantly.
3. Multi-AZ deployments for the **MySQL, MariaDB, Oracle, and PostgreSQL** engines utilize synchronous physical replication to keep data on the standby up-to-date with the primary. Multi-AZ deployments for the **SQL Server** engine use synchronous logical replication to achieve the same result, employing **SQL Server-native Mirroring** technology.
4. **Amazon Aurora** employs a highly durable, SSD-backed virtualized storage layer purpose-built for database workloads. ***Amazon Aurora automatically replicates*** your volume ***six ways, across three Availability Zones***. Amazon Aurora storage is fault-tolerant, transparently handling the loss of up to two copies of data without affecting database write availability and up to three copies without affecting read availability. Amazon Aurora storage is also self-healing. Data blocks and disks are continuously scanned for errors and replaced automatically.
5. Amazon Aurora Replicas share the same underlying storage as the primary instance. Any Amazon Aurora Replica can be promoted to become primary without any data loss and therefore can be used for enhancing fault tolerance in the event of a primary DB Instance failure. To increase database availability, simply create 1 to 15 replicas, in any of 3 AZs, and Amazon RDS will automatically include them in failover primary selection in the event of a database outage.

Dynamo DB

1. Amazon DynamoDB is a fully managed NoSQL database service that provides fast and predictable performance with seamless scalability
2. Amazon DynamoDB stores three geographically distributed replicas of each table to enable high availability and data durability. Read consistency represents the manner and timing in which the successful write or update of a data item is reflected in a subsequent read operation of that same item.
3. Eventually Consistent Reads (Default) – also supports Strongly Consistent Reads
4. A table is a collection of data items – just like a table in a relational database is a collection of rows.
5. Each table must have a primary key. The primary key can be a single attribute key or a "composite" attribute key that combines two attributes. The attribute(s) you designate as a primary key must exist for every item as

primary keys uniquely identify each item within the table

6. There is no explicit limitation on the number of attributes associated with an individual item, but the aggregate ***size of an item***, including all the attribute names and attribute values, ***cannot exceed 400KB***.
7. A Query gets one or more items using the table primary key, or from a secondary index using the index key. You can narrow the scope of the query on a table by using comparison operators or expressions. You can also filter the query results using filters on non-key attributes. Supports both strong and eventual consistency. A single response has a size limit of 1 MB.
8. DynamoDB supports key-value and document data structures. A document store provides support for storing, querying and updating items in a document format such as JSON, XML, and HTML.
9. Dynamodb soft limits are 10,000 writes/second or 10,000 reads/second
10. ***Local secondary indexes must be defined at time of table creation***. The primary index of the table must use a partition-sort composite key. ***Local secondary indexes cannot be removed from a table once they are created*** at this time
11. Each table can have ***up to 5 global and local secondary indexes***.
12. DynamoDB Triggers is a feature which allows you to execute custom actions based on item-level updates on a DynamoDB table. You can specify the custom action in code.
13. The custom logic for a DynamoDB trigger is stored in an AWS Lambda function as code. To create a trigger for a given table, you can associate an AWS Lambda function to the stream (via DynamoDB Streams) on a DynamoDB table. When the table is updated, the updates are published to DynamoDB Streams. In turn, AWS Lambda reads the updates from the associated stream and executes the code in the function.
14. DynamoDB Streams provides a time-ordered sequence of item-level changes made to data in a table in the last 24 hours. You can access a stream with a simple API call and use it to keep other data stores up-to-date with the latest changes to DynamoDB or to take actions based on the changes made to your table.
15. ***DynamoDB Streams keep records of all changes to a table for 24 hours***. After that, they will be erased.
16. DynamoDB Cross-region replication is used for DR, Live data & app migration to another region, Traffic management by directing reads closer to the customers etc.,

ELB / Auto Scaling

1. There are 3 types of load balancers – Application Load Balancer (Layer 7), Network Load Balancer & Classic Load Balancer (Layer 4)
2. Classic Load Balancer supports load balancing of applications using HTTP, HTTPS (Secure HTTP), SSL (Secure TCP) and TCP protocols.

3. Network Load Balancer provides TCP (Layer 4) load balancing. It is architected to handle millions of requests/sec, sudden volatile traffic patterns and provides extremely low latencies. In addition Network Load Balancer also preserves the source IP of the clients, provides stable IP support and Zonal isolation. It also supports long-running connections that are very useful for WebSocket type applications.

4. Network Load Balancer preserves the source IP of the client which in the Classic Load Balancer is not preserved. ***Customers can use proxy protocol with Classic Load Balancer to get the source IP.*** Network Load Balancer automatically provides a static IP per Availability Zone to the load balancer and also enables assigning an Elastic IP to the load balancer per Availability Zone. This is not supported with Classic Load Balancer.

5. Application Load Balancer supports load balancing of applications using HTTP and HTTPS (Secure HTTP) protocols.

6. An *Auto Scaling group* contains a collection of EC2 instances that share similar characteristics and are treated as a logical grouping for the purposes of instance scaling and management.

7. A *launch configuration* is a template that an Auto Scaling group uses to launch EC2 instances. When you create a launch configuration, you specify information for the instances such as the ID of the Amazon Machine Image (AMI), the instance type. ***Key Pair, Security Groups and Block Storage devices are optional.***

8. An auto scaling policy is a policy used by Auto Scaling that uses CloudWatch alarms to determine when your Auto Scaling group should scale out or scale in. ***Each CloudWatch alarm watches a single metric and sends messages to Auto Scaling when the metric breaches a threshold that you specify in your policy***.

9. The instance that you want to attach must meet the following criteria:

   The instance is in the running state.

   The AMI used to launch the instance must still exist.

   The instance is not a member of another Auto Scaling group.

   The instance is in the same Availability Zone as the Auto Scaling group.

   If the Auto Scaling group has an attached load balancer, the instance and the load balancer must both be in EC2-Classic or the same VPC. If the Auto Scaling group has an attached target group, the instance and the load balancer must both be in the same VPC.

10. Elastic Load Balancing provides access logs that capture detailed information about requests sent to your load balancer. Each log contains information such as the time the request was received, the client's IP address, latencies, request paths, and server responses.

11. Access logging is an optional feature of Elastic Load Balancing that is disabled by default. After you enable access logging for your load balancer, Elastic Load Balancing captures the logs and stores them in the Amazon S3 bucket that you specify as compressed files. You can disable access logging at any time.

12. There is no additional charge for access logs. You are charged storage costs for Amazon S3, but not charged for the bandwidth used by Elastic Load Balancing to send log files to Amazon S3.

AWS Solution Architect Associate Exam Prep - Notes

API Gateway

1. APIs built on Amazon API Gateway can accept any payloads sent over HTTP. Typical data formats include JSON, XML, query string parameters, and request headers
2. All of the APIs created with Amazon API Gateway expose HTTPS endpoints only. Amazon API Gateway does not support unencrypted (HTTP) endpoints
3. Amazon API Gateway is integrated with AWS CloudTrail to give you a full auditable history of the changes to your REST APIs. All API calls made to the Amazon API Gateway APIs to create, modify, delete, or deploy REST APIs are logged to CloudTrail in your AWS account.

DNS / Route 53

1. Route 53 Weighted routing policy is a way to add elasticity to an application deployments.

Lambda

1. AWS Lambda function ***memory*** allocation range is a ***minimum of 128 MB and a maximum of 1536 MB*** ***(recently increased to 3008KB in 11/2017)***
2. AWS Lambda ***ephemeral disk*** (temporary) size default is ***512 MB***
3. AWS Lambda ***maximum execution duration*** for a function is ***300 sec (5 minutes).***
4. AWS Lambda maximum request ***payload size is 6 MB for synchronous invocation***
5. AWS Lambda maximum request ***payload size is 128KB for asynchronous invocation***
6. AWS Lambda maximum concurrent invocations per region is 5000

VPC

1. In order for you to ping between instances, you need to allow ICMP traffic in the security group. A security group is a virtual firewall and it is a stateful firewall
2. When you launch an instance in VPC, you can assign upto a maximum of 5 security groups
3. Security groups are at the instance level and not at the subnet level
4. If you don't specify a security group at the instance launch, it is automatically assigned a default security group for VPC
5. You can specify "allow rules" in security groups but not "deny" rules
6. You can specify separate rules for inbound and outbound traffic.

7. When you create a security group, it has no inbound rules. Therefore, no inbound traffic originating from another host to your instance is allowed until you add inbound rules to the security group.
8. By default, a security group includes an outbound rule that allows all outbound traffic. If your security group has no outbound rules, no outbound traffic originating from your instance is allowed.
9. Security groups are stateful — if you send a request from your instance, the response traffic for that request is allowed to flow in regardless of inbound security group rules. Responses to allowed inbound traffic are allowed to flow out, regardless of outbound rules.
10. ICMP traffic is always regardless of the rules (ex. even if 0.0.0.0/0 is specified for ICMP)
11. Incoming traffic is allowed based on the private IP addresses of the instances that are associated with the source security group (and not the public IP or Elastic IP addresses).
12. If you specify a single IPv4 address, specify the address using the /32 prefix length. If you specify a single IPv6 address, specify it using the /128 prefix length.
13. The rules you create for use with a security group for a VPC can't reference a security group for EC2-Classic, and vice versa.

IAM / Security

Cloudwatch / Monitoring

1. AWS Cloudwatch metrics are retained as: < 1 minute for 3 hrs; = 1 min for 15 days; 5 minutes for 63 days; 1 hr for 455 days
2. AWS Cloudwatch datapoints published for higher resolution are still available after expiration but will be aggregated to a lower resolution
3. AWS Cloudwatch metrics can not be deleted but will expire
4. AWS Cloudwatch Logs agents will send log data every ***5 seconds from EC2 instance to Cloudwatch by default*** and it can be changed.
5. Basic Monitoring for Amazon EC2 instances: Seven pre-selected metrics at ***5-minute frequency*** and three status check metrics at 1-minute frequency, for no additional charge.
6. Detailed Monitoring for Amazon EC2 instances: All metrics available to Basic Monitoring at 1-minute frequency, for an additional charge.

Kinesis

1. 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
2. Data will be available for your Amazon Kinesis Applications to read and process from the stream
3. An Amazon Kinesis Application is a data consumer that reads and processes data from an Amazon Kinesis data stream. You can build your applications using either Amazon Kinesis API or Amazon Kinesis Client Library (KCL)
4. Amazon Kinesis Data Streams manages the infrastructure, storage, networking, and configuration needed to stream your data at the level of your data throughput. You do not have to worry about provisioning, deployment, ongoing-maintenance of hardware, software, or other services for your data streams. In addition, Amazon Kinesis Data Streams synchronously replicates data across three availability zones, providing high availability and data durability.
5. The throughput of an Amazon Kinesis data stream is designed to scale without limits via increasing the number of shards within a data stream
6. 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.
7. The maximum size of a data blob (the data payload before Base64-encoding) within one record is 1 megabyte (MB).
8. Each shard can support up to 1000 PUT records per second.
9. Kinesis vs SQS: Amazon Kinesis Data Streams enables real-time processing of streaming big data. It provides ordering of records, as well as the ability to read and/or replay records in the same order to multiple Amazon Kinesis Applications. Amazon SQS lets you easily move data between distributed application components and helps you build applications in which messages are processed independently (with message-level ack/fail semantics), such as automated workflows.
10. A record is the unit of data stored in an Amazon Kinesis data stream. A record is composed of a sequence number, partition key, and data blob. Data blob is the data of interest your data producer adds to a data stream. The maximum size of a data blob (the data payload before Base64-encoding) is 1 megabyte (MB).
11. Partition key is used to segregate and route records to different shards of a data stream. A partition key is specified by your data producer while adding data to an Amazon Kinesis data stream. For example, assuming you have a data stream with two shards (shard 1 and shard 2). You can configure your data producer to use two partition keys (key A and key B) so that all records with key A are added to shard 1 and all records with key B are added to shard 2.
12. A sequence number is a unique identifier for each record. Sequence number is assigned by Amazon Kinesis when a data producer calls PutRecord or PutRecords operation to add data to an Amazon Kinesis data stream. Sequence numbers for the same partition key generally increase over time; the longer the time period between PutRecord or PutRecords requests, the larger the sequence numbers become.
13. Amazon Kinesis Data Firehose is the easiest way to load streaming data into data stores and analytic tools. It can capture, transform, and load streaming data into Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and Splunk, enabling

near real-time analytic with existing business intelligence tools and dashboards you're already using today

14. Amazon Kinesis Data Firehose synchronously replicates data across three facilities in an AWS Region, providing high availability and durability for the data as it is transported to the destinations.
15. A source is where your streaming data is continuously generated and captured. For example, a source can be a logging server running on Amazon EC2 instances, an application running on mobile devices, a sensor on an IoT device, or a Kinesis stream.
16. A shard is a uniquely identified group 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.

Cloudformation

1. Cloudformation templates have the following sections – metadata, parameters, mappings, conditions, transform, resources and outputs. Only resources section is required and all other sections are optional.
2. CloudFormer is a template creation beta tool that creates an AWS CloudFormation template from existing AWS resources in your account. You select any supported AWS resources that are running in your account, and CloudFormer creates a template in an Amazon S3 bucket.
3. Regular expressions (commonly known as regexes) can be specified in a number of places within an AWS CloudFormation template, such as for the AllowedPattern property when creating a template parameter.
4. You can also configure your AWS CloudFormation template so that the logs are published to Amazon CloudWatch, which displays logs in the AWS Management Console so you don't have to connect to your Amazon EC2 instance.

AWS Import/Export

1. AWS Import/Export accelerates transferring data between the AWS cloud and portable storage devices that you mail to us. AWS Import/Export is a good choice if you have 16 terabytes (TB) or less of data to import into Amazon Simple Storage Service or Amazon Elastic Block Store (Amazon EBS). You can also export data from Amazon S3 with AWS Import/Export.
2. AWS Import/Export doesn't support export jobs from Amazon Elastic Block Store (EBS).
3. An Amazon S3 export transfers individual objects from Amazon S3 buckets to your device, creating one file for each object. You can export from more than one bucket and you can specify which files to export using manifest file options.
4. You cannot export Amazon S3 objects that have been transitioned to Amazon Glacier Storage Class using Amazon S3 Object Lifecycle Management.

AWS Solution Architect Associate Exam Prep - Notes

Some Key Limits

EC2	Max On Demand Instances per region	20
	HVM with Instance Store Types Allowed For	M3 C3 XR I2 D2
	Types available	Dr MC GIFT PX
SQS	Default visibility	30 Sec
	Min & Maximum visibility timeout	0 sec, 12 hrs
	Min, Default, Max Retention	1 min, 4 days, 14 Days
	Min & Max message size	1KB & 256KB
	Min & Max delay time	0 Sec & 15 min
	Min & Max ReceiveWait (long polling)	0 Sec & 20 Sec
	Max messages in Standard, FIFO queues	120K & 20K
	Max metadata attributes supported	10
SWF	Max duration for a workflow execution	1 Year
	Max SWF domains per account	100
	Max workflow & activity types	10K
	Max open executions	100K
	SWF long polling default timeout value	60 sec
EBS	IOPS to Vol Size Max Ratio	50-to-1
	Max IOPS for provisioned	20000
	gp2 Max IOPS	10000
S3	Min & Max Length of Bucket Names	3 & 63
	Min size of S3 objects	0
	Max size of S3 objects	5 TB
	Max number of buckets per account	100
	S3-Standard – Availability, Durability	99.99 / 11 9s
	S3-Standard IA – Availability, Durability	99.9 / 11 9s
	S3-RRS – Availability, Durability	99.99 / 99.99
Glacier	Max size of archive	40 TB
Cloudwatch	< 60 sec Metrics Retention	3 Hrs
	1 Min Metric Retention	15 days
	5 Min metric retention	63 days
	1 Hr metric retention	455 days
Kinesis	Max size of data blob (payload)	1 MB
	Default stream record retention	**24 Hours**
	Max stream record retention	**7 days**
SNS	Max message size	256 KB
	Max topics per account	100K
DynamoDB	Max size of a single response	1 MB
	Max retention of DynamoDB stream data	**24 Hours**
Lambda	Default temp disk size	512 KB
	Min & Max memory allocation	128 KB to 3000 MB
	Max execution time	300 Sec
	Default execution time	3 Sec
	Minimum execution time	1 Sec
	Max payload size for synchronous invocation	6 MB

	Max payload size for asynchronous invocations	128 KB
Networking	Max # of VPCs per region	5
	Max EIPs per region	5
	Subnets per VPC	200
	Max ALBs + NLBs per region	20
IAM	Min&Max for GetFederationToken (custom broker)	15 min & 36 Hrs
	AssumeRoleWithSAML – Min, Default & Max	15 min, 1 Hr, 1 Hr
	AssumeRole – cross account or custom – min, default, max	15 min, 1 Hr, 1 Hr
	AssumeRoleWithWebIdentity– min, default, max	15 min, 1 Hr, 1 Hr