Amazon Certified Solutions Architect Exam

Official Information

Links

- https://aws.amazon.com/certification/certification-prep/
- http://awstrainingandcertification.s3.amazonaws.com/production/AWS_certified_solutions_architect_associate_blueprint.pdf

Books (CSDN)

- AWS Certified Solutions Architect Official Study Guide: Associate Exam (2016, Sybex)
    - www.wiley.com/go/sybextestprep (questions: 368, flashcards: 100)

Online Lectures (迅雷)

- ACloudGuru AWS Solutions Architect Exam Prep Course (udemy)
- Linuxacademy AWS Solutions Architect
- CBTNugggets AWS Solutions Architect

Practise Questions & Flash Cards

- www.wiley.com/go/sybextestprep (questions: 368, flashcards: 100)

- http://thecertschool.com/category/aws/
- http://www.dennyzhang.com/aws_associate_cert/
- http://blog.flux7.com/blogs/quizzes/cloud-computing-quiz-1-check-your-knowledge-on-aws
- http://searchaws.techtarget.com/quiz/Amazon-Web-Services-Security-Quiz
- http://searchaws.techtarget.com/quiz/Test-your-knowledge-Amazon-Simple-Storage-Service-quiz
- http://searchaws.techtarget.com/quiz/Take-our-quiz-to-find-out-what-you-know-about-AWS-IaaS
- http://www.silicon.co.uk/quiz/amazon-web-services-145-1
- http://searchaws.techtarget.com/quiz/Test-your-knowledge-Amazon-Redshift-quiz
- http://www.awsomeblog.com/aws-certified-solutions-architect-exam-sample-quiz/
- http://www.cloudsolutionsbook.com/amazon-cloud-solutions/june-20th-2015
- http://www.huangbowen.net/blog/2014/10/22/aws-cert-sample-question/
- http://quizlet.com/35935418/detailed-questions-flash-cards/
- CramFLASH Study Flashcards for AWS Developer Associate Exam: 60 “cards” are included
- CramFLASH Study Flashcards for AWS SysOps Admin Associate Exam: 50 flashcards included
- AWS Solutions Architect Apps on Android

AWS Global Infrastructure

Infrastructure

- >16 Regions:
    - a region is a geographical area consisting 2 or more availability zones
    - complete independent and isolated from other regions
        - resources aren't replicated across regions unless organizations choose to do so
        - data locality
        - sovereignty concerns
        - located close to end users, minimize latency
        - far from primary facilities to satisfy disaster recovery and compliance needs
        - foundation for meeting location dependent privacy and compliance requirements
        - customer has full control - AWS does not move customer’s resources
- >42 Availability Zones
    - simply a data center
    - connected via an inexpensive, low latency network
    - distinct locations engineered to be insulated from failures in other availability zones
    - isolated, but AZs in a region are connected through low-latency links
    - AZs in a region are physically separated within a typical metropolitan region
    - located in lower-risk flood plains
    - uses UPS and on-site backup generators
    - redundantly connected to multiple tier-1 transit providers
    - HA: deploy across multiple AZs
- >54 Edge Locations
    - CDN endpoints for CloudFront
    - many more than regions
- AWS cloud service model: IaaS (infrastructure-as-a-service)
    - other models:PaaS, SaaS
- AWS access:
    - Management Console
    - CLI
    - SDK (API)

Support Plans

- Basic, Developer, Business, Enterprise
- Response Times:
    - general guidance: <24h
    - system impaired: <12h

        - production system impaired: <4h (business, enterprise)
        - production system down: <1h (business, enterprise)
        - business-critical system down: <15min (enterprise)
    - Trusted Advisor: basic+developer: 4 core set of checks only

AWS Design Principles

- Deployment Models:
    - all-in cloud-based application: fully deployed in the cloud
        - public
        - private
    - hybrid deployment: connects infrastructure between cloud-based resources and existing data center. leverage dedicated connectivity, identity federation, and integrated tools
- Design for failure:
    - assume things will fail
    - recovery strategies during design time:
        - design automated recovery from failure
        - assume more than the expected number of requests per second some day
    - decouple components
- Implement Elasticity:
    - proactive cyclic scaling: periodic scaling at fixed interval (daily, weekly, …)
    - proactive event-based scaling: scaling when expecting a big surge of traffic due to a scheduled business event (e.g. new product launch, marketing campaigns)
    - auto-scaling based on demand: take actions to scale up or down based on metrics (e.g. cpu load, network I/O)
- Advantages of Cloud:
    - almost zero upfront infrastructure investment: trade capital expense for “variable expense”
    - benefit from massive “economies of scale”
    - more efficient resource utilization: stop guessing about capacity
    - just-in-time infrastructure: increase speed and agility
    - usage-based costing: stop spending money running and maintaining data centers
    - reduced time-to-market: go global in minutes
- Technical benefits:
    - automation: scriptable infrastructure
    - automated elasticity and scalability
    - proactive scaling
    - more efficient development lifecycle
    - improved testability
    - disaster recovery and business continuity

AWS Security

- electronic surveillance and multi-factor access control systems
- 24x7 staffed by security guards
- access is authorised on a “least privilege basis”
- shared security model:
    - AWS: responsible for securing the underlying infrastructure
        - global infrastructure
        - services offered in the cloud
        - security configuration, patches, antivirus etc. of managed services (e.g. DynamoDB, RDS, Redshift, EMR, WorkSpaces)
    - Customer: responsible for anything you put on the cloud or connect to the cloud
        - full root access on guest operating system
            - AWS does not have any access rights to the guest OS
            - available on instance types from M3, C3, R3 and G2: encrypted EBS volumes and snapshots with AES-256
        - IAAS: security configuration and management tasks on EC2, VPC, S3
        - managed services: account management and user access control (e.g. MFA, SSL/TLS, user activity logging with CloudTrail)
- Physical and environmental security:
    - fire detection and suppression
    - redundant and maintainable power systems
    - climate and temperature control
    - preventive monitoring
    - Storage Decommissioning:
        - prevents customer data from being exposed to unauthorized individuals
        - all decommissioned magnetic storage devices are degaussed and physically destroyed
        - DoD 5220.22-M (national industrial security program operating manual)
        - NIST 800-88: guidelines for media sanitization
- Business continuity management
    - availability: data centers built in clusters in various global regions
    - incident response: 7x24h coverage to detect incidents and to manage the impact and resolution
    - communication: training programs, service health dashboard
- Network security:
    - Secure network architecture
        - network ACLs
    - Secure access points:
        - monitored HTTP/HTTPS API endpoints
        - redundant ISP connections
    - Transmission protection:
        - HTTPS using SSL

        - VPC: private subnet
        - ELB: SSL termination on the load balancer is supported
    - Monitoring and protection:
        - DDoS (Distributed denial of service) attacks
        - MITM: man in the middle attacks
        - IP spoofing: firewall denials instance sending traffic with a source IP or MAC address other than its own
        - Port scanning: attacks such as ARP cache poisoning do not work within EC2 and VPC
        - Packet Sniffing by other tenants: it is not possible for a virtual instance running in promiscuous mode to receive or sniff traffic that is intended for a different virtual instance
        - AWS regularly scans all internet facing IP addresses for vulnerabilities (not include customer instances)
        - regular vulnerability threat assessments by independent security firms
    - IPsec VPN: encrypted tunnel between VPC and your data center
    - Logically the AWS production network is segregated from the Amazon Corporate network
    - AWS acceptable use policy:
        - you must request a vulnerability scan in advance
    - Direct Connect: dedicated connection using 802.1q VLAN
- Trusted Advisor:
    - makes recommendations to save money, improve performance and close security gaps, e.g.
        - open ports
        - public access to S3 buckets
        - enable user activity logging (CloudTrail)
        - MFA on root account
- VPC security:
    - API access encrypted with SSL and signed by secret access key
    - subnets and route tables
    - firewall (security groups)
    - network ACLs
    - virtual private gateway: private connectivity between VPC and another network
    - internet gateway
    - dedicated instances: physically isolated
- EC2 security:
    - multiple levels of security: OS of host platform, guest OS, firewall and signed API calls
    - hypervisor: highly customized Xen hypervisor
    - Instances Isolation:
        - customers' instances → hypervisor → virtual interfaces → security groups → firewall → physical interfaces

        - isolated via the Xen hypervisor on physical machines
        - virtual network interface: AWS firewall within the hypervisor layer between the physical network interface and instance's virtual interface
        - virtual memory space: physical RAM is separated
            - memory scrubbed by hypervisor when unallocated to a guest
        - virtualized disks: customer instances have no access to raw disk devices
            - AWS disk virtualization layer resets disk blocks after usage
- S3 security:
    - data access: IAM policies, ACLs, bucket policies, query string authentication
    - data transfer: SSL endpoints
    - data storage: SSE AES-256 or client-side encryption
        - S3 metadata is not encrypted
    - access logs
    - CORS (cross-origin resource sharing)
- Shared responsibility beyond security considerations:
    - shared management, operation, and verification of IT controls
        - IT governance is the customer's responsibility regardless of how the IT control environment is deployed (on-premises, cloud, or hybrid)
        - AWS provides control information and very proactive risk management
        - AWS control environment consists of policies, processes, and control activities
    - obtaining industry certifications and independent third-party attestations
    - publishing information about security and AWS control practices via the website, whitepapers, and blogs
    - directly providing customers with certificates, reports, and other documentation
- Compliance:
    - SOC1, SOC2, SOC3: service organization control
        - SOC 1: processes and controls relevant to financial reporting
        - SOC 2: security and availability controls
        - SOC 3: public report, summarized version of SOC 2
    - FISMA, DIACAP, FedRAMP: federal information security management act
        - evaluated by independent assessors for a variety of government systems' approval process
    - PCI DSS Level 1: payment card industry data security standard
        - software needs to be compliant too
    - ISO 27001: information security management system requirements, intended to bring information security under explicit management control
    - ISO 9001: quality management systems requirements, audited against third party assessments
    - ITAR: international traffic in arms regulations
    - FIPS: federal information processing standards, ensures information security and interoperability
    - HIPAA

- CSA: Cloud Security Alliance
- MPAA: Motion Picture Association

Well-Architected Framework

- AWS architecture best practices:
    - design for failure and nothing will fail
        - HA: If you design architectures around the assumption that any component will eventually fail, systems won't fail when an individual component does.
        - use e.g. ELB load balancer, Multi-AZ standby database, S3 or DynamoDB
    - implement elasticity: design architectures to take advantage of cloud computing providing virtually unlimited on-demand capacity
        - vertical scaling
        - horizontal scaling:
            - stateless applications: no knowledge of the previous interactions and stores no session information
                - use Auto Scaling group to scale elastically
                - deployment automation
            - stateless components: store session information in a database
    - leverage different storage options:
        - S3: large-scale capacity and performance or storage with high data durability to support backup and active archives for disaster recovery
        - Glacier: data archiving and long-term backup
        - CloudFront: deliver entire websites, including dynamic, static, streaming, and interactive content using a global network of edge locations
        - DynamoDB: a fast and flexible NoSQL database with a flexible data model and reliable performance.
        - EBS: reliable block storage to run mission-critical applications such as Oracle, SAP, Exchange, and SharePoint
        - RDS: highly available, scalable, and secure MySQL database without the time-consuming administrative tasks
        - Redshift: fast, powerful, fully-managed, petabyte-scale data warehouse to support business analytics
        - ElastiCache: a Redis cluster to store session information
        - EFS: a common file system for your application that is shared between more than one EC2 instance
    - build security in every layer
    - think parallel
    - loose coupling sets you free: components designed as black boxes to reduce interdependencies so that a change or a failure in one component does not cascade to the components.
    - don't fear constraints

- General Design Principles:
    - stop guessing your capacity needs
    - test systems at production scale
    - lower the risk of architecture change
    - automate to make architectural experimentation easier
    - allow for evolutionary architectures
- Security
    - apply security at all layers
    - enable traceability
    - automate responses to security events
    - focus on securing your system
    - automate security best practices
        - Data protection:
            - encrypt everything where possible (at rest and in transit)
            - AWS customers maintain full control over their data
            - AWS makes it easier for you to encrypt your data and manage keys, including regular key rotation
            - detailed logging e.g. for files access and changes
            - storage systems for exceptional resiliency, e.g. S3 designed for 11 nines of durability
            - versioning can protect against accidental overwrites, deletes etc.
            - AWS never initiates the movement of data between regions.
            - AWS services: ELB, EBS, S3, RDS
        - Privilege management:
            - only authorized and authenticated users are able to access your resources
            - ACL, role based access controls, password management (e.g. rotation policies)
            - protect access to and use of AWS root account credentials
            - control human access to the AWS Management console and APIs
            - limit automated access to AWS resources
            - AWS services: IAM, MFA
        - Infrastructure protection:
            - outside of cloud: RFID controls, security, lockable cabinets, CCTV etc.
            - inside cloud: VPC
                - enforce network and host-level boundary protection
                - enforce AWS service level protection
                - protect the integrity of OS on EC2 instances
            - AWS services: VPC
        - Detective controls:
            - detect and identify security breaches
            - AWS CloudTrail, CloudWatch, AWS Config, S3, Glacier

            - capture and analyze AWS logs
    - Shared Responsibility Model:
        - customer: responsible for security IN the cloud
        - AWS: responsible for security OF the cloud
- Reliability
    - Design principles:
        - test recovery procedures
        - automatically recover from failure
        - scale horizontally to increase aggregate system availability
        - stop guessing capacity
    - Foundations:
        - AWS handles foundations for you: networking, physical machines etc. (but with soft limits to stop customers from accidentally over-provisioning resources)
        - plan network topology
        - manage AWS service limits for your account
        - escalation path to deal with technical issues
        - AWS services: IAM, VPC
    - Change management:
        - adapt to changes in demand
        - monitor AWS resources
        - execute change management
        - use CloudWatch to monitor environment and services such as auto scaling to automate change in response to changes on the production environment
        - AWS services: CloudTrail, CloudWatch
    - Failure management:
        - always assume failure will occur
        - backup data
        - failover tests
        - plan how to prevent the failures
        - AWS services: CloudFormation
- Performance Efficiency
    - Design principles:
        - democratize advanced technologies
        - go global in minutes
        - use serverless architectures
    - Compute:
        - select the appropriate instance type
        - periodically ensure to have the appropriate instance type
        - AWS services: Autoscaling
    - Storage:
        - select the appropriate storage solution

        - periodically ensure to have the most appropriate storage solution
        - factors:
            - access method: block, file or object
            - access pattern: random or sequential
            - throughput required
            - frequency of access: online, offline or archival
            - frequency of update: worm, dynamic
            - availability constraints
            - durability constraints
        - AWS services: EBS, S3, Glacier
    - Database:
        - select the appropriate database solution
        - periodically ensure to have the most appropriate storage solution
        - ensure performance is as expected
        - ensure capacity and throughput matches demand
        - AWS services: RDS, DynamoDB, Redshift
    - Space-Time trade-off
        - select the appropriate proximity and caching solutions
        - ensure performance is as expected
        - ensure proximity and caching solutions matches demand
        - AWS services: CloudFront, ElastiCache, Direct Connect, RDS Read Replicas etc.
- Cost optimization
    - Design principles:
        - transparently attribute expenditure
        - use managed services to reduce cost of ownership
        - trade capital expense for operating expense
        - benefit from economies of scale
        - stop spending money on data center operations
    - Matched supply and demand:
        - CloudWatch: keep track of your actual demand
        - Auto-scaling with demand
        - Lambda or server-less context: only execute when a request comes in
        - AWS services: Auto-scaling
    - Cost-effective resources:
        - use the correct instance type
        - AWS services: EC2 reserved instances, Trusted Advisor
    - Expenditure awareness:
        - access controls and procedures to govern costs
        - monitor usage and spending
        - decommission resources no longer need or stop resources temporarily not needed
        - consider data-transfer charges

        - track costs with tags
        - billing alerts
        - consolidated billing
        - AWS services: CloudWatch alarms, SNS
    - Optimizing over time:
        - manage the adoption of new services
        - AWS services: AWS blog, Trusted Advisor

Consolidated Billing

- advantages:
    - one bill per AWS account
    - very easy to track charges and allocate costs
    - volume pricing discount
- paying account is independent
- cannot access resources of linked accounts
- all linked accounts are independent
- soft limit: 20 linked accounts

Compute

EC2: Elastic Compute Cloud

- virtual servers in the cloud
- scalable compute capacity, both up and down
    - compute: amount of computational power required to fulfill your workload
    - instance: virtual server
- Pricing models:
    - On demand:
        - pay a fixed rate by hour without any upfront payment or long-term commitment
        - usecase:
            - applications being developed or tested for the first time
            - applications with short term, spiky, or unpredictable workloads that cannot be interrupted
            - test/dev running on Ec2 for the first time
            - supplement reserved instance servers (for extra temporary load)
    - Reserved:
        - payment options:
            - all upfront
            - partial upfront
            - no upfront: pay the entire reservation charge in monthly installments
        - pay upfront payments and get a capacity reservation
        - significant discount on the hourly charge (1 year or 3 year terms)
        - usecase:
            - applications with steady state or predictable usage
            - applications that require reserved capacity
                - primary web servers
                - domain controllers
    - Spot:
        - bid for free instance capacity by hour
            - when your bid >= spot price, you get server
            - when your bid < spot price, you lose server with 1 hour warning
        - greater savings
            - if the spot instance is terminated by amazon EC2, no charge for the partial hour usage
            - if you terminate the instance, you pay for the hour
        - usecase:
            - applications that have flexible start and end times
            - applications that are only feasible at very low compute prices

        - users with urgent computing needs for large amounts of additional capacity
- EC2 instance type families:
    - self-service provisioning and managed with different AMIs and VMs available optimized for various workloads. Instance properties:
        - vCPUs
        - memory
        - storage (size and type)
        - network performance: low, moderate and high
            - enhanced networking: reduces the impact of virtualization on network performance by enabling a capability called Single Root I/O virtualization (SR-IOV).
                - more packets per second (PPS)
                - lower latency
                - less jitter
                - in VPC only
                - supported types: C3, C4, D2, I2, M4, R3
        - Pricing:
            - ratio of vCPUs to memory is constant as the size scale linearly
            - e.g. m4.xlarge costs twice as much as the m4.large instance
    - T2:
        - general purpose, lowest cost
        - e.g. web servers, small DBs
    - M3, M4:
        - general purpose
        - e.g. application servers
    - C3, C4:
        - compute optimized
        - e.g. CPU intensive Apps / DBs
    - R3:
        - memory optimized (RAM)
        - e.g. memory intensive Apps / DBs
    - G2:
        - graphics / general purpose GPU
        - e.g. video encoding, machine learning, 3D application streaming
    - I2:
        - high speed storage (IOPS)
        - e.g. NoSQL DBs, data warehousing
    - D2:
        - dense storage
        - e.g. file servers, data warehousing, hadoop
- AMI: Amazon Machine Image

- a template for the root volume for the instance providing information required to launch a virtual server in the cloud
    - Properties:
        - region
        - operating system and its configuration (Linux or Windows)
        - initial state of any patches
        - application or system software
        - architecture (32-bit or 64-bit)
        - launch permissions
        - storage for the root device (root device volume)
            - instance store (ephemeral storage)
            - EBS backed volumes
    - Sources:
        - published by AWS
        - AWS marketplace
        - generated from existing instances
        - uploaded virtual servers (using AWS VM Import/Export service): formats: raw, VHD, VMDK and OVA
- Root Device Volume Types:
    - **EBS-backed** volume:
        - created from EBS snapshot
        - persistent storage, can use termination protection
        - can be stopped and resumed (maintenance, migration to new hardware), data will persist
        - data remains after a crash
        - option to keep ebs root volume after termination
        - faster to launch (no need to fetch image from S3)
    - **Instance Store** volume (ephemeral storage):
        - created from a template stored in S3
        - not persistent (ephemeral), only exists for the life of that instance
        - dynamically resizable
        - cannot be stopped, detached and reattached to other instances (if the underlying host fails, you lose your data)
        - reboot possible without losing data
- launch permissions that control which AWS accounts can use the AMI to launch instances
- a block device mapping that specifies the volumes to attach to the instance when it's launched
- only regional, copy before use in other regions using console, cli or EC2 API
- Create a new AMI:
    - by default, when creating an AMI, the termination protection and the encryption are turned off
    - by default, AMI created is private

        - regional, you can only launch an AMI from the region where it's stored
        - you can copy AMI to another region using CLI, Console or API
        - contains:
            - template for root volume (OS, application servers, apps, etc)
            - launch permissions
            - block device mapping
    - by default, On an EBS-backed instance the root EBS volume is to be deleted when the instance is terminated
    - Root-volumes cannot be encrypted, use third party tool instead
- Configurations:
    - Termination protection is turned off by default and should be turned on
    - On an EBS-backed instance, the default action is for the root EBS volume to be deleted when the instance is terminated
- Addressing an instance:
    - public DNS name (automatically generated)
    - public IP: persists only while the instance is running
    - elastic IP: persists until you releases it and is not tied to the lifetime or state of an individual instance. It can be transferred to a replacement instance in the event of an instance failure.
    - in VPC: private IP and elastic network interfaces
- Security Group:
    - stateful virtual firewall protection
    - types:
        - VPC security groups: control outgoing and incoming traffic
        - classic security groups (without VPC): control outgoing traffic only
    - default deny
    - attributes: port, protocol and source/destination (CIDR block)
    - every instance must have at least one security group
- Encryption:
    - root volumes cannot be encrypted by default, use third party tool (such as bitlocker) to encrypt the root volume
    - additional volumes can be encrypted
- IAM roles with EC2:
    - can only assigned to an EC2 instance during its creation
    - more secure than storing access keys on EC2 instances
    - easier to manage
    - universal, can be used in any region, AZs
    - useful for:
        - federated user access (non-AWS)
            - AD, LDAP, Kerberos
            - can create trust if organization supports SAML 2.0
        - cross-account access
            - multiple AWS accounts

        - applications running on EC2 instances that need access to other AWS resources:
            - EC2 instance using an S3 bucket or DynamoDB table
- Initial Access:
    - Linux: ec2-user and private key-based ssh login: stores public key in ~/.ssh/autorized_keys
    - Windows: Administrator and a generated password using private key RDP login
- Bootstrap Scripts:
    - script used when first being provisioned e.g. install, run apache and sync directories, enrolling in a directory service, installing puppet or chef
    - provision an AMI instance
- EBS-Optimized instance:
    - use EBS-optimized instance when using ssd volumes to ensure that the instance is prepared to take advantage of the I/O of the ssd volume
    - additional hourly charge
- Instance Metadata.
    - used to get information about an instance, e.g.
        - public and private ip
        - instance ID and type
        - associated security groups
        - AMI used to launch the instance
    - up to 10 tags per instance
    - instances meta data (not user data)
    - usage:
        - write data to an html page
        - trigger a lambda function to update DNS
    - **curl http://169.254.169.254/latest/meta-data**
- Tenancy options:
    - shared tenancy (default): a single host machine may house instances from different customers.
        - AWS does not use overprovisioning and fully isolates instances from other instances on the same host
    - dedicated instances: host is dedicated to a single customer to run dedicated instances
    - dedicated host: a specific host is dedicated to an instance. e.g. address licensing requirements. The dedicated instance can only run on this host.
- Auto-Scaling Groups:
    - configure a launch configuration
    - create rules to spin-up and shut down instances based on monitor triggers
    - deleting an auto scaling group will automatically delete any instances it created
- Placement Groups:
    - a logical grouping within a **single** AZ
    - low latency, high network throughput

        - fully use only with enhanced networking and 10Gbps network performance
    - limitations:
        - a group can't span multiple AZs
        - name of the placement group must be unique within your account
        - groups can't be merged
        - an existing instance cannot be moved into a group
            - solution: create AMI from the existing instance and launch a new instance from the AMI into the group
        - only certain types of instances can be launched in a group:
            - compute optimized
            - GPU
            - memory optimized
            - storage optimized
    - recommended: homogenous instances within a group
    - enables applications to participate in a low-latency (10Gbps network)
    - for applications requiring low network latency or high network throughput
- AWS CLI:
    - AWS CLI preinstalled on AWS AMI
    - commands:
        - aws configure: input access key, secret access key, default region name and output format
        - aws s3:
            - mb: make bucket
            - rb: remove bucket
            - ls: list buckets
- Backup/Snapshot: through Console, CLI, API or scheduled snapshots
    - can be used to increase the size of the volume

Lambda

- run code in response to events
- no servers, run code without worrying about infrastructure at all
- continuous scaling
- Availability: 99.99%
- Usage:
    - modifications to objects in S3 buckets
    - messages arriving in Kinesis stream
    - table updates in DynamoDB
    - API call logs created by CloudTrail
    - etc.
- Types:

        - as event-driven compute service, runs code in response to events, e.g. changes to data in S3 bucket or DynamoDB table
        - as compute service, runs code in response to HTTP requests using API Gateway or API calls made using AWS SDKs routes to Lambda
- Supported languages:
    - Node.js (Javascript)
    - Java
    - Python
- Pricing:
    - <1 million requests free, 20ct per 1 million requests thereafter
    - memory/duration fee: rounded to the nearest 100ms from the time your code begins execution until it returns or otherwise terminates: $0.00001667 for every GB-second used

EC2 Container Service

- run and manage Docker containers

Elastic Beanstalk

- run and manage web application containers
- features: resource provisioning, load balancing, auto scaling, monitoring
- supported languages: PHP, Java, Python, Ruby, Node.js, .NET and Go
- Features:
    - ease deployment and management of applications on AWS
    - access to built-in CloudWatch monitoring metrics such as CPU utilization, request count, and average latency
    - email notification alerts through SNS
    - viewing applications logs
    - full control over AWS resources, such as instance type, database and storage options, login access, enhanced security, adjusting server settings and environment variables, adjusting Auto Scaling settings

Server Migration

Storage & Content Delivery

EBS: Elastic Block Store

- network attached block device for persistent storage in EC2
- virtual disks, once attached, you can create a file system on these volumes
- custom provision sizes from 1GB to 16TB per volume
- multiple volumes per EC2 instance possible
    - 1 EBS instance can only attached to 1 EC2 instance (EFS for shared volumes)
- EBS volumes are automatically replicated in a specific AZ
- AES-256 based encryption available (no extra charge)
- Formats:
    - st1: EBS Magnetic (EBS standard):
        - throughput optimized HDD for big data, data warehouses, log processing
            - sequential reads
            - low-cost storage requirements
            - workloads where data is accessed infrequently
        - cost effective storage that delivers approximately 100 IOPS per volume on average with a best effort ability to burst to hundreds of IOPS p/volume
            - max IOPS/volume: **500**
            - max throughput: 40-90 MB/s
        - **1GB** - 1TB
        - for applications where cost is important or for workloads where data is accessed infrequently
    - gp2: EBS General Purpose (SSD):
        - 99.999% availability
        - for Boot volumes, low-latency interactive apps, dev & test, small-sized databases
        - 1GB - 16TB
        - ratio of 3 IOPS/GB with up to **10000** IOPS
            - up to 10000 IOPS
            - max throughput: 160 MB/s
        - provide the ability to burst to 3000 IOPS per volume under 1 GB for short periods (whenever you are not using the full IOPS, they are accumulated as credits for the burst)
    - io1: EBS Provisioned IOPS (SSD):
        - for I/O intensive workloads such as NoSQL and relational databases
        - 4GB - 16TB
        - customer specified an IOPS rate when creating a volume. Currently supports up to **20000** IOPS per volume
            - max IOPS: 20000
            - max throughput: 320 MB/s
- Volume:

    - volumes exist on EBS (virtual hard disk)
    - snapshot of a volume will stored on S3
    - on EC2:
        - lsblk: view volumes attached
        - file -s /dev/xvdf: view if file system, check if clean
        - mkfs -t ext4 /dev/xvdf: make file system
        - mount /dev/xvdf /mnt/target
- Snapshot:
    - point in time copies of volumes
    - incremental, only changed blocks are moved to S3
        - the first snapshot may some time to create
    - snapshots exist on S3
    - snapshots of encrypted volumes are encrypted automatically
    - volumes restored from encrypted snapshots are encrypted automatically
    - snapshots can be shared, but only if they are unencrypted
    - EBS root device volumes should be stopped first before taking the snapshot
- RAID (Redundant Array of Independent Disks):
    - get better throughput than using a single volume
    - types:
        - 0: striped, no redundancy, good performance (more throughput)
            - performance is limited to the worst performing volume in the set
            - loss of a single volume results in a complete data loss
        - 1: mirrored, redundancy
        - 5: good for reads, bad for writes (**not recommended** on EBS)
        - 10: striped & mirrored, good redundancy, good performance
    - snapshot of a RAID array:
        - problem:
            - snapshot excludes data held in the cache by applications and OS
            - problem in a RAID array due to interdependencies of the array
        - solution: take a consistent snapshot
            - stop the application from writing to disk and flush all caches to the disk
                - freeze the file system
                - unmount the RAID array
                - shutting down the associated EC2 instance
- Raise EBS volume:
    - detach EBS volume
    - create a snapshot of the original EBS volume in S3
    - create new EBS volume from the snapshot and specify a larger size
    - attach new volume in place of the original
    - refresh or update with OS-level utility
    - delete the original EBS volume

S3: Simple Storage Service

- scalable, durable, secure, unlimited storage in the cloud
- object-based data storage:
    - independent of a server
    - API built on standard HTTP verbs (REST)
    - data is treated as a stream of bytes
    - contains both data and metadata
        - system metadata: last modified, MD5, size, Content-Type
        - user metadata: can be specified when the object is created
    - objects reside in buckets
        - single flat namespace of keys with no structure, no sub-bucket or 'real' folders
        - each bucket: unlimited number of objects up to **5TB**
        - bucket names are global: e.g. contain your domain name
            - but a bucket is created in a specific region
        - up to **100** buckets per account
    - identified by a unique user-specified key (filename)
        - up to 1024B UTF-8 characters (folder is part of filename)
        - unique within a single bucket
        - bucket, key and optional version ID uniquely identifies an S3 object
    - automatically replicated on multiple devices in multiple facilities within a region
    - automatically partitions buckets to support high request rates and simultaneous access by many clients
        - to support higher request rates: random distribution of keys
        - consider CloudFront as caching layer in front of S3 bucket
- Files can be from 1B to **5TB** and are stored in buckets.
- In S3 bucket names must be unique globally (universal namespace)
    - https://s3-<region>.amazonaws.com/<bucket>
    - https://<bucket>.s3-<region>.amazonaws.com/
- Consistency:
    - read after write consistency for PUTs of new objects
    - eventual consistency for overwrite PUTs and DELETEs (can take some time to propagate)
- S3 is an object consists of:
    - key: name of the object
    - value: data
    - version ID: for versioning
    - metadata (tags)
    - subresources
    - ACLs: access control lists
- highly secure: 4 different access control mechanisms, server-side encryption available
- protection against user-level accidental deletion or overwriting

    - versioning (once enabled, can only be suspended)
    - cross-region replication
    - MFA Delete (can only be enabled by the root account)
- Access control:
    - S3 Access Control Lists (ACL): coarse-grained permissions at object or bucket level
        - use cases: enabling bucket logging or make static website public
    - S3 bucket policies (recommended): fine-grained control (like IAM policies)
        - associated with bucket resources instead of an IAM principal
        - include an explicit reference to the IAM principal in the policy (cross-account access possible)
        - use cases: who, from where (CIDR or IP), during what time of day
    - Pre-signed URL:
        - using your security credentials, bucket name and object key to grant time-limited permission to download (GET) the objects
        - use case: protection against content scraping e.g. media files
- Usecases:
    - backup and archive for on-premises or cloud data
    - content, media and software storage and distribution
    - big data analytics
    - static website hosting
    - cloud-native mobile and internet application hosting
    - disaster recovery
- Multipart Upload API: upload large objects as a set of parts
    - must be used for objects larger than 5GB (recommended >100MB)
    - automatically performed by CLI (aws s3 cp, aws s3 mv, aws s3 sync)
    - steps:
        - initiation
        - uploading parts in arbitrary order with retransmission
        - completion (or abort)
    - optional: lifecycle policy to abort incomplete multipart uploads after a specified number of days
- Range GET: download only a portion of an object in S๓ or Glacier
- Cross-Region replication: asynchronously replicate all new objects in the source bucket to a target bucket in another region
    - versioning must be turned on for both source and destination buckets
    - give S3 permission to replicate objects using IAM policy
    - if turned on in an existing bucket, only new objects will replicated
- Tiered Storage:
    - S3:
        - frequently accessed, immediately available
        - highly available (**99.99%**) and durable (**99.999999999%**)
        - stored redundantly across multiple devices in multiple facilities

        - concurrent facility fault tolerance: 2
    - S3-IA (infrequently accessed):
        - infrequently accessed, immediately available
        - durability, availability and concurrent facility fault tolerance same as S3
        - for data that is accessed less frequently, but requires rapid access when needed
        - lower fee than S3, but additional retrieval fee
    - Reduced Redundancy Storage:
        - frequently accessed, immediately available
        - 99.99% availability and 99.99% durability
        - for data that is easily reproducible (e.g. thumbnails)
        - concurrent facility fault tolerance: 1
    - Glacier:
        - very cheap
        - used for archival only
        - 3-5 hours to restore from Glacier
- Lifecycle Management
    - can be used in conjunction with versioning
    - can be applied to current versions and previous versions
    - following actions can now be done
        - transition to the S3-IA storage class (>128KB and >30 days after creation date)
        - archive to the Glacier storage class (>30 days after IA)
        - permanently delete
- Versioning
    - stores all versions of an object (including all writes and deleted flag)
    - great backup tool
    - once enabled, versioning cannot be disabled, only suspended
    - integrates with lifecycle rules
    - Versioning's MFA delete capability, which uses multi-factor authentication, can be used to provide an additional layer of security
    - cross region replication, requires versioning enabled on the source bucket
- Encryption
    - in transit: SSL/TLS
    - at rest:
        - Server side encryption
            - S3 managed keys - **SSE-S3**
            - AWS key management service - **SSE-KMS**
            - Server side encryption with customer provided keys - **SSE-C**
        - Client side encryption: encrypts data prior to uploading to bucket
- Secure data using access control lists and bucket policies
    - by default, all newly created buckets are **private**
    - you can setup **access control** to your buckets using:

        - bucket policies
        - ACLs: access control lists
    - S3 buckets can be configured to create access logs which log all requests made to the S3 bucket (better: on another bucket)
- S3 Transfer Acceleration
    - uses CloudFront Edge network to accelerate uploads to S3
    - better performance the further you are away from your bucket
    - incurs an additional fee
    - edge location url: https://**mys3bucket.s3-accelerate.amazonaws.com**
- Create static website:
    - create a bucket with the same name as the desired website hostname
    - upload files to the bucket
    - make all the files public (world readable)
    - enable static website hosting for the bucket (specify index and error documents)
    - website access URL: <bucket-name>.s3-website-<region>.amazonaws.com
    - create DNS name in your own domain for the website using a DNS CNAME, or an Route53 Alias that resolves to the Amazon S3 website URL
    - website access using your website domain name
- Logging:
    - best-effort basis with a slight delay
    - contains: requester account, IP, bucket, request time, Action, Response status
- Event notifications:
    - using SNS, SQS or Lambda
    - events: objects created, removed, RRS object lost

Glacier

- archive storage in the cloud
- extremely low-cost storage service (1c per GB per month)
- optimized for data that is infrequently accessed
- Archives:
    - up to **40TB** of data per archive
    - unlimited number of archives
    - automatically assigned with a unique archive ID at time of creation
    - automatically **encrypted**
    - immutable
- Vaults:
    - container for archives
    - up to 1000 vaults per account
    - access control with IAM policies or vault access policies
    - data retrieval policy e.g. to limit the retrievals to the free tier
    - Vault Lock: once locked, policy can no longer be changed
- Data Retrieval:
    - object is copied to S3 RRS

- up to 5% of data per month free calculated on a daily prorated basis (>5%: restore fee)
- retrieval time: 3-5 hours

CloudFront CDN

- global content delivery network
- delivers web pages and other web content to a user based on the geographic locations of the user, the origin of the webpage and a content delivery server
- Edge Location:
  - location where content will be cached
  - separate to an AWS region or AZ
  - read or write possible (put an object on to them)
  - objects are cached for the life of the TTL (default: 24 hours)
  - clear cached objects by yourself will be charged
- Use Cases:
  - serving the static assets of popular websites
  - serving a whole website or web application containing both dynamic and static content
  - serving content to users who are widely distributed geographically
  - distributing software or other large files
  - serving streaming media
- Anti-Use Cases:
  - all or most requests come from a single location
  - all or most requests come through a corporate VPN
- Origins:
  - origin of all the files that the CDN will distribute
  - S3 Bucket, EC2 instance, ELB, Route 53 and non-AWS origin servers
- Distributions:
  - consists of a collection of Edge Locations
  - identified by a DNS domain name such as test123.cloudfront.net
    - CloudFront distribution domain name or your own CNAME
  - Types:
    - Web Distribution: typically used for websites
    - RTMP: used for media streaming
- Cache control:
  - default: expires after 24 hours
    - can controlled by Cache-Control headers set on origin server or
    - set min, max and default TTL or
    - explicit invalidation by API (used in unexpected circumstances
  - best practice: include version id as part of the object path, using versioning, users always see the latest content through cloudFront when you update your site without using invalidation

- Multiple Origins and Dynamic Content by setting Cache-Behaviors to serve different behaviors for different client devices.
    - short TTLs for dynamic content
    - functionality:
        - path pattern
        - which origin to forward to
        - whether to forward query string
        - whether accessing the specified files requires signed URLs
        - whether to require HTTPS access
- expiration time regardless of any Cache-Control headers
- Private Content: restrict viewer access option: restrict using signed URLs or singed cookies
    - Signed URLs: use URLs that are valid only between certain times and optionally from certain IP addresses
    - Signed Cookies: require authentication via public and private key pairs
    - OAI (Origin Access Identities): restrict to an S3 bucket only to a special CloudFront user associated with your distribution to ensure that content in a bucket is only accessed by your CloudFront.
- not faster for the first user, but faster for every other subsequent user in the nearby area

Storage Gateway

- hybrid storage integration
- data stored within a single user-specified region
- secure integration of on-premises IT environments with cloud storage
- available as virtual machine image (VMware ESXi or Microsoft Hyper-V)
- volumes can be mounted as iSCSI devices
- Types:
    - Gateway stored volumes: (cloud is backup)
        - entire dataset on-site
        - asynchronously backed up to S3 up to 1TB
    - Gateway cached volumes: (cloud is primary)
        - entire dataset stored in S3 up to 32TB
        - most frequently accessed data cached on-site
        - no access of all data if you lose internet connectivity
    - Gateway virtual tape library (VTL):
        - limitless collection of virtual tapes backed by S3 (Virtual Tape Library) or Glacier (Virtual Tape Shelf)
        - up to 10 virtual tape drives per gateway
        - used for backup and provides iSCSI interface (online access) supported by popular backup applications like NetBackup, Backup Exec, Veam etc.
- Pricing components:
    - gateway usage per gateway per month
    - snapshot storage usage per GB per month

- volume storage usage per GB per month
- data transfer out per GB per month

Import/Export Snowball

- large scale data transport
- Types:
  - Import/Export Disk:
    - you ship your disks to AWS site of your choice
    - import to S3, EBS, Glacier
    - export from S3
  - Snowball:
    - import to S3 only
    - export from S3
    - faster, simpler and more cost-effective
      - no need to purchase any hardware or write any code to transfer data
      - each snowball appliance transfers up to **50** or 80TB
      - multiple appliances can be used in parallel
    - uses tamper-resistant enclosures, 256-bit encryption and TPM (Trusted Platform Module) to ensure both security and full chain-of-custody
    - reduces management overhead involved with transferring data into or out of AWS
- Use Cases:
  - Storage Migration: move massive data to another location
  - Migrating applications: migrating an application to the cloud involving huge amounts of data
- Pricing Models:
  - per-device fee
  - data load time charge (per data-loading-hour)
  - return shipping charges

EFS: Elastic File System

- fully-managed file storage service with elastic storage capacity for EC2 instances
  - grows and shrinks automatically
- Features:
  - **block** based storage
  - supports network file system protocol (NFSv4)
  - pay only for the storage you use (no pre-provisioning required)
  - can scale up to the petabytes
  - can support thousands of concurrent NFS connections
  - read after write consistency
  - data is **stored across multiple AZ's within a region**

- EFS volume can be mounted to multiple EC2 instances (not possible with EBS)

Database

Data Warehouse

- a subject-oriented, integrated, time varying, non-volatile collection of data in support of business decision-making process
- OLTP (online transactional processing): characterized by a large number of short online transactions (insert, update, delete), measured by number of transactions per second. Data are stored in entity models (usually 3NF).
- OLAP (online analytical processing): characterized by relatively low volume, complex queries often involving aggregations and data mining techniques. Data is aggregated and stored in multidimensional schemas.

RDS: Relational Database Services

- managed relational database service for OLTP with consistent deployment and operational model
    - does not provide shell access to DB instances, restricts access to certain system procedures and tables that require advanced privileges
    - configuration in DB parameter groups and option groups
    - AWS operational responsibilities: scaling, high availability, backups, DB engine patches, software installation, OS patches, OS installation, server maintenance, rack and stack, power and cooling
    - Storage Options: uses EBS, types: magnetic, General Purpose, Provisioned IOPS
- Database Types:
    - MySQL, PostgreSQL, Oracle, SQL server, Amazon Aurora, MariaDB
- Backup & Recovery
    - RPO (Recovery Point Objective): in minutes, defines the maximum period of data loss that is acceptable in the event of failure or incident
    - RTO (Recovery Time Objective): in hours or days, defines the maximum amount of downtime that is permitted to recover from backup and to resume processing. (use failover node)
    - during a backup window, storage I/O may be suspended and you may experience elevated latency
    - Backup plans:
        - Automated Backup:
            - enabled by default
            - backup stored in S3 (free storage to the size of the database)
            - automated backups will take a full daily snapshot and stores transaction logs throughout the day.
            - retention period between 1 and **35** days
            - recovery **down to a second** within the retention period

            - during the backup window, storage I/O may be **suspended** (elevated latency, typically a few minutes
                - this can be avoided if you go Multi-AZ as the backup is taken of the standby
            - deleted after the deletion of the RDS instance
        - Database Snapshot:
            - manual backup, full backup
            - remains after delete the original RDS instance, until you explicitly delete them
            - use Multi-AZ to minimize the performance impact of a snapshot
            - when you restore either automated or snapshot, the restored version will be a new RDS instance with a **new endpoint**
    - Recovery: restored to a new DB instance, only default DB parameter and security groups are associated with the restored instance, restoration to any point during the retention period (typically up to the last five minutes)
- Encryption:
    - at rest:
        - supported for MySQL, Oracle, SQL Server, PostgreSQL and MariaDB
        - uses AWS Key Management Service (KMS)
        - once the RDS instance is encrypted, the underlying storage, its automated backups, read replicas and snapshots are also encrypted
        - encryption of existing DB is not supported
            - solution: create new DB instance with encryption and migrate data into it
- Multi-AZ RDS:
    - standby database used for failover and backups
        - not for performance enhancement
        - not available to offline queries from the primary master instance
    - primary RDS instance uses synchronous replication to an RDS in a different AZ
    - automatic failover:
        - DNS name remains
        - failover events: loss of availability in primary AZ, loss of network connectivity to primary database, compute unit failure on primary database, storage failure on primary database
        - takes typically 1 to 2 minutes
        - manual failover possible
    - AWS handles replication: in event of a database or AZ failure, RDS will automatically failover to the standby so that database operations can resume quickly without administrative intervention
    - Available in:
        - Microsoft SQL Server (standard and enterprise)
        - Oracle
        - MySQL

        - PostgreSQL
        - MariaDB
        - Aurora
- Scaling up and out:
    - Vertical scalability:
        - modify instance type, storage expansion up to 6TB storage (Microsoft SQL server not supported)
    - Horizontal scalability with partitioning or sharding:
        - supported by DynamoDB or Cassandra
    - Horizontal scalability with read replicas:
        - uses asynchronous replication to create **up to 5** read replicas
        - supported by MySQL, PostgreSQL, MariaDB, Aurora
        - Cross-region read replicas support
        - used for read performance improvement and scaling, not DR:
            - write to primary, read from read replicas
        - must have automatic backups turned on
        - read replicas of read replicas possible (but greater latency)
        - each replica has its own DNS endpoint
        - replica of Multi-AZ RDS possible
        - replica cannot have Multi-AZ
        - read replicas can be promoted to their own databases. (this breaks the replication)
- Security best practices:
    - RDS in private subnet
    - restrict network access using network ACL and security groups
    - limit actions in IAM
    - create users with strong passwords with rotation
    - use in-transit and at-rest encryptions: SSL, KMS, TDE (transparent data encryption)
        - all logs, backups, snapshots of encrypted RDS instance are automatically encrypted
- Database Migration Service:
    - migration of both schema and data between databases.
- Licensing models:
    - License Included or BYOL (Bring Your Own License) for Oracle and Microsoft SQL Server
- Aurora DB Engine:
    - MySQL-compatible, increased reliability and performance, DB cluster spans multiple AZs
    - Instance Types:
        - Primary Instance: read and write workloads
        - Aurora Replica: read operations only, up to 15 instances, Multi-AZ support

DynamoDB

- fully managed NoSQL database
    - connection through HTTP/HTTPS request/response in JSON format
    - supports document and key-value data models
    - scaling on the fly without down time
    - predictable and scalable NoSQL Document Store
- Features:
    - stored on SSD
    - spread across **3** geographically distinct AZs
- Data Modeling:
    - schema: table, items, and attributes
    - each item has a primary key
    - partition key = primary key with one attribute
    - partition key + sort key = primary key with two attributes
    - max item size: 400KB
- Throughput capacity (Provisioned capacity)
    - Write: $0.0065 per hour for every 10 units
    - Read: $0.0065 per hour for every 50 units
    - Storage: 25ct per GB per month
    - expensive for writes, cheap for reads
    - e.g. table without a local secondary index:
        - **1 capacity unit** = read an item less than **4KB**
            - 110KB = 28 capacity units
        - read with strong consistency uses **twice** the number of capacity units
        - 1 capacity unit = write an item less than **1KB**
        - If you exceed your provisioned capacity, requests will be throttled and can be retried later
            - CloudWatch metrics: TrottledRequests, ConsumedReadCapacityUnits, ConsumedWriteCapacityUnits
    - one single partition can hold up to 10GB of data and supports a maximum of 3000 read capacity units or 1000 write capacity units
    - partition key design: should have a large number of distinct values to ensure the values are requested uniformly among partitions. e.g. by adding a random element to the partition key
- Secondary Indexes: query with different access patterns
    - Global secondary index: an index with a different partition and sort key
        - can be created or deleted at any time
    - **Local secondary index**: an index with the same partition key, but a different sort key
        - can only be created when the table is created
    - DynamoDB update indexes when an item is modified and consumes write capacity units

        - global secondary index: separate provisioned capacity
        - local secondary index: uses the capacity of the main table
- Read and Write commands: UpdateItem, DeleteItem, PutItem, GetItem, BatchGetItem, BatchWriteItem
    - BatchWriteItem: creates or updates up to 25 items in one batch
- Searching Items:
    - Query: requires a partition key and a distinct value to search. more efficient option. sort key can operate with comparison operator. Results sorted by the primary key.
    - Scan: full scan of the entire table or secondary index. used for select large number of items.
    - Results have a limit to 1MB. You can use pagination through results.
- Security:
    - granular control over IAM with AWS STS support
- Consistency:
    - eventual consistent reads (default)
        - consistency across all copies of data within a second
        - repeating a read after a short time returns the updated data (best read performance)
    - strongly consistent reads:
        - returns a result that reflects all writes that received a successful response prior to the read
        - might be less available in case of a network delay or outage
- DynamoDB Streams: get recent changes or perform some kind of processing on the changed items
    - reads log of activity changes buffered in a time-ordered stream near real-time
    - stream records are organized into shards with a maximum of 24 hour lifetime
    - a shard could depending on fluctuating load levels be split one or more times before they are closed
    - Kinesis Adapter from KCL (Kinesis Client Library) simplifies the application logic required to process reading records from streams and shards

RedShift

- fast (10 times faster, simple, cost-effective data warehousing (OLAP)
- Usage
    - used for BI (Cognos, Jaspersoft, SAP Netweaver)
    - used to pull in large and complex datasets
    - used to run lots of OLAP transactions
- fully managed petabyte-scale wide-column data warehouse service
- Pricing from 25ct per hour, $1000 per terabytes per year and per compute node hours
    - only compute nodes will incur charges
    - backup
    - data transfer

- Features:
    - columnar data storage (instead of rows)
        - only columns involved in the queries are processed
        - columnar data is stored sequentially
        - block-size of 1 MB for columnar storage
        - requires far fewer I/O and improves performance
    - advanced compression
        - columnar data can be compressed much better
        - RedShift automatically samples data and chooses the best compression scheme
    - massive parallel processing (MPP)
        - automatically distributes data and query load across all nodes and newly added nodes
- Configuration:
    - Single Node (160GB)
    - Multi Node (Cluster):
        - **Leader Node**: manages client connections and receives queries
        - Compute Node: stores data and perform queries and computations. up to 128 compute nodes
- RedShift Cluster:
    - 1 Leader node + >1 Compute nodes
    - Node types: 6 node types, dense compute node up to 326TB SSDs, dense storage node up to 2PB magnetic disks
    - 1 Cluster contains >1 databases
    - JDBC or ODBC connections with the leader node
    - Compute node storage divided into 2 to 16 slices
    - increase performance by adding nodes, partitioning by distribution strategy
    - resize operation: Redshift creates a new cluster and migrates data zo it. DB will become read-only until the operation is finished
- Data Modeling:
    - compression encoding per column
    - additional columns can be added, existing columns cannot be modified
    - distribution strategies:
        - EVEN: evenly distributed
        - KEY: by column value, e.g. increased performance on joins
        - ALL: full copy on each node, e.g. lookup table
    - sort keys: e.g. for efficient range-restricted predicates
- Loading data
    - INSERT, UPDATE or COPY
    - COPY: parallel read from multiple flat files in S3 or DynamoDB and bulk load
    - VACUUM: reorganize data and reclaim space after deletes
    - ANALYZE: update table statistics
    - UNLOAD: export into csv files in S3

- WLM (Workload Management):
    - queue and prioritize queries when supporting many users
- Security:
    - at rest: AES-256 encryption
    - By default RedShift has its own key management and can manage keys through HSM or KMS
        - AWS KMS (key management service) (default)
        - own keys through CloudHSM (hardware security module)
- Availability:
    - currently only in **1 AZ**
        - can restore snapshots to new AZ's in the event of an outage
- Backup & Recovery: (same as in RDS) automated or manual snapshots

Elasticache

- in-memory cache
- alleviate database under stress/load which is particularly read heavy and not prone to frequent changing
- Usage:
    - improve latency and throughput for read-heavy app workloads (social networks, gaming, media sharing) or
    - compute heavy workloads (recommendation engine)
- Caching engines:
    - Memcached: widely adopted memory object caching system
        - key/value datastore with objects as blobs
        - Auto Discovery: simplifies your application code by no longer needing awareness of the infrastructure topology of the cache cluster
        - Elasticache:
            - elastically grow and shrink a cluster of Memcached nodes (up to 20 nodes)
            - partition into shards
            - parallel operations support
    - Redis: in-memory key-value store, supports master/slave replication and Multi-AZ to achieve cross AZ redundancy
        - supports a rich s of data types like strings, lists, and sets
        - supports persist onto disk, automatic and manual snapshot backups (in S3)
        - supports sort and rank data
        - up to 5 read replicas
        - failover using read replica in Multi-AZ replication groups
        - Replicating group: up to 6 clusters (5 read replicas) for horizontal scaling
            - replication is asynchronous with a small delay before data is available on all cluster nodes
- Caching patterns:

        - Cache-aside: checks in cache if it contains the data needed. If not, query database and serialize and write the result to the cache. Next user request will be read from cache
- Design for failure best practice:
    - Memcached: using a larger number of nodes with a smaller capacity
    - Redis cluster: using a Multi-AZ replication group
- Vertical scaling: no automatic scale up support. create a new cluster and start redirecting traffic to it.
    - Memcached cluster always starts empty
    - Redis cluster can be initialized from a backup or restore from any other compatible Redis cluster backups
        - default: the same configuration as the source cluster

DMS: Database Migration Service

- managed database migration service manages all the complexities of the migration process:
    - data type transformation
    - compression
    - parallel transfer
    - ensures data changes to the source database occurred during the migration process are replicated to the target
- AWS schema conversion tool converts the source database schema and a majority of the custom code like views, stored procedures and functions
    - e.g. Oracle → Aurora

Aurora

- MySQL-compatible, relational database with 5 times better performance at 1/10 of the price of commercial DB with similar performance and availability (Oracle)
- Scaling:
    - Storage: scales from 10GB in 10GB increments to **64TB** (autoscaling)
    - CPU: compute resources scales up to 32 vCPUs
    - Memory: up to 244GB RAM
- Availability:
    - write availability: up to loss of 2 copies of data
    - read availability: up to loss of 3 copies of data
    - **6 copies** of your data: 2 copies in each AZ with minimum of 3 AZ
- self-healing: data blocks and disks are continuously scanned for errors and repaired automatically
- Replicas Types:
    - Aurora Replicas: up to **15**
    - MySQL Read Replicas: up to **5**

Networking

DNS: Domain Name Service

- IPv4: 32bit, 4 billion addresses
- IPv6: 128bit, 340 undecillion addresses
- api.aws.amazon.com.:
    - '.com': top level domain (TLD), e.g. .com, .net
        - gTLD: generic TLD
        - ccTLD: country code TLD
    - 'amazon': second level domain (SLD)
    - 'aws.amazon.com': sub domain
    - 'api.aws.amazon.com': domain name
    - 'api': host
    - 'api.aws.amazon.com.': full qualified domain name (FQDN), ends with a dot, indicating the root of the DNS hierarchy
    - controlled by IANA (internet assigned numbers authority) in a root zone database
    - second level domain name: .co.uk, .com.au
- Domain Registrars:
    - an authority that can assign domain names directly under one or more top-level domains
    - domains are registered with InterNIC (a service of ICANN)
    - each domain name is registered in a central database known as the WhoIS database
    - e.g. godaddy.com, 123-reg.co.uk
- Name Servers:
    - translate domain names into IP addresses
    - each name server may redirect requests to other name servers or delegate responsibility for the subnet of subdomains for which they are responsible
    - Types:
        - authoritative: give answers to queries about domains under their control
        - mirror: point to other servers or serve cached copies of other name servers' data
- Zone Files: simple text file containing the mappings between domain names and IP addresses
- Resource Record Types:
    - SOA (start of a zone of authority record): specifies authoritative information about a DNS zone: primary name server, administrator of the zone, current version, ttls relating to refreshing the zone
        - TTL: time to live in seconds. The lower the TTL, the faster changes to DNS records take to propagate throughout the internet
    - NS (name server record): delegates a DNS zone to use the given authoritative name servers

- A (address record): map hostnames to an IPv4 address
- AAAA (IPv6 address record): maps hostnames to an IPv6 address
- CNAME (canonical name record): alias of one name to another, DNS lookup will continue by retrying the lookup with the new name
- MX (mail exchange record): maps a domain name to a list of message transfer agents for that domain
- **PTR** (pointer record): reverses an A record.
- SPF (sender policy framework record): used to combat spam by mail servers
- TXT (text record): holds text information
- SRV (Service record): location of servers for specified services

Route53 (DNS)

- scalable, authoritative DNS and Domain name registration
    - authoritative: provides an update mechanism that developers use to manage their public DNS name
- Main functions:
    - Domain registration: lets you register domain names, such as example.com
    - DNS service: translates domain names into IP addresses.
    - Health checking: sends automated requests over the internet to your application to verify that it's reachable, available, and functional
- IPv6 not fully supported yet
- DNS **Routing Policies**
    - Simple (default): when using a single resource
    - Weighted: assign different weights to split the traffic
    - Latency:
        - route traffic based on the lowest network latency to the end user
        - need to create a latency resource record set for the EC2 or ELB resource in each region you want participating
        - great for improving global page load times
    - Failover:
        - active/passive set up
        - Route53 monitors health of primary site using a health check monitor
    - Geolocation: route traffic based on the geographic location of the end users
- Hosted Zones: a collection of resource record sets hosted by Route52
    - Types:
        - private hosted zone: a container holds information about how you want to route traffic within one or more VPCs
        - public hosted zone: a container holds information about how you want to route traffic on the internet
    - S3 static website: redirect all requests to a subdomain and create an alias resource record that sends requests for the root domain to the S3 bucket
    - CNAMEs are not allowed for hosted zones. Use an alias record instead.

    - Do not use A records for subdomains. They refer to hardcoded IP addresses. Use alias record or CNAME records to always point to the right resource.
- **Alias Record**: used to map resource record sets to ELB, CloudFront or S3 buckets.
    - maps a DNS name to another target DNS name like a CNAME record:
        - www.test.com → elbtest.elb. amazonaws.com
    - CNAME **cannot** be used for naked domain names (zone apex). You can't have a CNAME for http://test.com, it must be either an A record or an Alias
    - ELB's do not have pre-defined IPv4 addresses, must be resolved using a DNS name
    - Route53 automatically recognizes changes in the record sets that the alias resource record set refers to. (e.g. if IP address of the ELB changes, Route 53 will change the DNS answers automatically)
    - Given the choice, **always** choose an Alias over a CNAME record
    - used to map resource record sets in your hosted zone to ELB, CloudFront distributions, or S3 buckets
- Health checking:
    - aware of an unhealthy endpoint and route differently within 30 seconds. (after 3 failed tests in a row), a new DNS results will be known to clients a minute later (with TTL=60s) and a complete recovery time of about 90 seconds.
- Best practice for highly availability and resiliency:
    - setup ELB in every region with cross-zone load balancing and connection draining
    - EC2 instances are running in multi AZs in an auto-scaling group
    - define health checks in ELB to ensure the delegation to only healthy instances
    - define Route53 healthcheck to ensure that requests are routed only to load balancers that have healthy EC2 instances
    - production environment has alias records that point to ELB and uses a latency-based routing policy associated with ELB health checks.
    - failover environment has alias record that points to CloudFront distribution of S3 bucket hosting a static version of the application.
    - application's subdomain has an alias record that points to production environment (primary target) and failover environment (secondary target) using a failover routing policy
    - hosted zone has an alias record that redirects requests to domain name using a S3 bucket of the same name
    - application content (both static and dynamic) can be served using CloudFront.
    - application is deployed in multiple regions to be protected against a regional outage

VPC: Virtual Private Cloud

- isolated cloud resources (logical datacenter)
- a private, isolated section of the AWS cloud where you can launch resources in a virtual network that you define consists of:

    - components:
        - subnets
        - route tables
        - DHCP option sets (dynamic host configuration protocol)
        - security groups
        - ACLs (network access control lists)
    - optional components:
        - IGWs (internet gateways)
        - EIP (elastic IP addresses)
        - ENI (elastic network interfaces)
        - endpoints
        - peering
        - NAT (network address translation) instances and gateways
        - VPG (virtual private gateway)
        - CGW (customer gateway)
        - VPN (virtual private network)
- Features:
    - selection of IP address range
        - 10.0.0.0 - 10.255.255 (10/8 prefix)
        - 172.16.0.0 - 172.31.255.255 (172.16/12 prefix)
        - 192.168.0.0 - 192.168.255.255 (192.168/16 prefix)
    - by default, soft limit of **5** VPCs per region
    - creation of subnets (e.g. public-facing or private-facing)
    - configuration of route tables and network gateways
    - multiple layers of security: security groups and network access control lists (ACL)
    - hardware VPN connection
- Subnets:
    - segment of a VPC's IP address range
    - smallest subnet: …/28 (16 IPs, 5 of which reserved by AWS)
    - reside within one AZ and cannot span AZs (1 subnet = 1 AZ)
    - Types:
        - public: route table directs the subnet's traffic to the VPC's IGW
            - attach an IGW to VPC
            - create a subnet route table rule to send all non-local traffic (0.0.0.0/0) to the IGW
            - configure network ACLs and security group rules to allow relevant traffic to flow to and from EC2 instances
            - assign a public IP address or EIP address to public EC2 instances
        - private: no IGW associated (internet through NAT)
        - VPN-only: route table directs the subnet's traffic to the VPC's VPG
- Route tables:
    - a set of rules applied to the subnets used to determine where network traffic is directed

        - default route (local route): enables communication within a VPC (cannot be modified or removed)
        - your VPC has an implicit route
        - your VPC automatically comes with a main route table that you can modify
        - you can create additional custom route tables for your VPC
        - each subnet must be associated with a route table, which controls the routing for the subnet. If you don't explicitly associate a subnet with a particular route table, the subnet uses the main route table
        - you can replace the main route table with a custom table that you've created so that each new subnet is automatically associated with it
        - each route in a table specifies a destination CIDR and a target
        - AWS uses the most specific route that matches the traffic to determine how to route the traffic
- IGW (internet gateway):
    - horizontally scaled, redundant, highly available
    - provides a target in VPC's route tables for internet-routable traffic
    - maintains the 1-to-1 map of the instance private IP and public IP
    - when traffic is sent from the instance to the internet:
        - IGW translates the reply address to the instance's public IP or EIP address
    - when an instance receives traffic from the internet:
        - IW translates the destination address to the instance's private IP and forwards the traffic to the VPC
- DHCP (dynamic host configuration protocol):
    - a standard for passing configuration information to hosts on a TCP/IP network, e.g. domain name, domain name server and netbios-node-type
        - domain-name-servers: defaulted to AmazonProvidedDNS
        - domain-name: defaulted to the domain name of the region
    - every VPC must have only 1 DHCP option set assigned to it
    - configuration options:
        - domain-names servers: up to 4 domain name servers separated by commas
        - domain-name: e.g. mycompany.com
        - ntp-servers: up to 4 network time protocol servers
        - netbios-name-servers: up to 4 NetBIOs name servers
        - netbios-node-type: set to value '2'
- EIP (elastic IP)
    - a static, public IP address in the pool for the region that you can allocate to your account and release
    - allows you to maintain a set of IP addresses that remain fixed while the underlying infrastructure may change over time
    - exam tips:

        - you must first allocate an EIP for use within a VPC and then assign it to an instance
        - EIPs are specific to a region
            - an EIP in one region cannot be assigned to an instance within an VPC in a different region
        - 1-to-1 relationship between network interfaces and EIPs
        - you can move EIPs from one instance to another, either in the same VPC or a different VPC within the same region
        - EIPs remain associated with the AWS account until you explicitly release them
        - charges for EIPs when booked (even when they are not used)
- ENI (elastic network interfaces)
    - only available within a VPC, **associated with a subnet** upon creation
    - can have 1 public and multiple private IP addresses (1 primary private IP)
- Endpoint:
    - enables you create a private connection to another AWS service (e.g. S3) without requiring access over the internet or through a NAT instance, VPN or direct connect
    - only for services **within the region**
    - multiple endpoints can be created for a single service
    - steps:
        - specify VPC
        - specify the service identified by: com.amazonaws.<region>.<service>
        - specify policy: full access or custom policy (can be changed at any time)
        - specify the route tables. A route will be added to each specified route table.
            - route: destination=service, target=endpoint
- Default VPC vs Custom VPC:
    - user-friendly: all subnets have a route out to the internet
    - 1 public subnet in every AZ within the region with a netmask of .../20
    - all subnets in default VPC have an IGW attached
    - each EC2 has both a public and private IP address
    - (If you delete the default VPC the only way to get it back is to contact AWS)
- Steps to **create** a custom VPC:
    - define IP range (automatically creates default route table)
    - create subnets (automatically creates route table and network ACL)
        - **largest = .../16, smallest = .../28**
        - AWS reserves the first 4 and last 1 IP address of any subset
            - .../28 → 11 useable IPs
    - create IGW:
        - by default it's detached, need to manually attach it to VPC
        - 1 IGW per VPC
    - create custom route table and attach IGW to it

        - associate public subnets to use new route
        - launch 1 instance per subnet
        - provision EC2 NAT instance
            - create security group for NAT instance
- enables connectivity between your network and VPC via a VPN or dedicated connection.
- simplifies end user access and system integration
- Peering:
    - connect VPS with each other via a direct network route **within a single region** using private IP addresses
    - peering with other AWS accounts possible (with different ip ranges)
    - uses underlying infrastructure:
        - no single point of failure
        - no bandwidth bottlenecks
        - it's not a gateway or a VPN connection
        - star configuration
    - Limitations:
        - cannot connect VPCs with matching or overlapping CIDR blocks
        - cannot connect VPs in different regions
        - cannot have more than 1 peering connection between the same 2 VPCs at the same time
        - **no transitive peering supported**
- Connectivity Options:
    - VPN Connection:
        - encrypted IPsec hardware VPN connection between your network and VPC
        - can create multiple VPN connections to one VPC
        - fast and simple to setup
        - VPC supports multiple CGWs
        - Types:
            - VPG (virtual private gateway): VPN concentrator on the AWS side
                - supported routing types:
                    - BGP (border gateway protocol): dynamic routing
                    - static routing
            - CGW (customer gateway): a physical device or software on the customer's side of the VPN connection
                - you must initiate the VPN tunnel from CGW to the VPG
                - connection consists of 2 IPSec tunnels: higher availability to the VPC
        -
    - AWS Direct Connect:
        - dedicated network connection between your network and VPC
        - can reduce network costs, increase bandwidth throughput, and provide a more consistent network experience

        - 1 Gbps or 10 Gbps ports
- Security Group:
    - default: all inbound traffic is blocked, all outbound traffic is allowed
        - inbound: source: sg-xxx, protocol: all, port range: all
            - allow inbound traffic from instances within the same security group
            - no other inbound traffic allowed
        - outbound: destination: 0.0.0.0/0, protocol: all, port range: all
            - allow all outbound traffic
    - Features:
        - **stateful**: if an inbound rule allowing traffic in, that traffic is automatically allowed back out again, regardless of any rules
        - changes take effect **immediately**
            - you can change the security groups associated with an instance after launch. changes take effect immediately.
        - operates at the instance level (first layer of defense)
        - separate rules for inbound and outbound traffic
        - all rules are evaluated before deciding whether to allow traffic
        - applies to an instance only if someone specifies the security group when launching the instance, or associates the security group with the instance later on
        - security group can stretch across multiple regions and AZs where a subnet cannot (1 subnet = 1 AZ)
        - unless configured or with the default security group, instances associated with the same security group can't talk to each other
    - Limitations:
        - allow rules only, but no deny rules
        - up to 500 security groups for each VPC
        - up to 50 inbound and 50 outbound rules for each security group
        - up to 5 security groups per network interface
- Network ACL:
    - default ACL: allows all outbound and inbound traffic
    - user-created (custom) ACL: denies all inbound and outbound traffic until you add rules
    - each subnet must be associated with a network ACL
    - features:
        - 1 network ACL can be associated to multiple subnets (**ACL 1:n Subnets**)
        - operates at the subnet level (second layer of defense)
        - supports allow rules and deny rules
            - e.g. Block IP addresses using ACL and not security groups
        - **stateless**: return traffic must be explicitly allowed by rules
        - processed in number order from lowest when deciding whether to allow traffic (ACL is a numbered list of rules)

        - automatically applies to all instances in the subnets it's associated with (backup layer of defence)
- NAT solutions: provide internet traffic to EC2 instances in private subnets
    - NAT instances: (outbound gateway for private-facing instances)
        - use existing NAT AMI
        - **disable** source/destination check on the instance (fake src ip address)
        - NAT instance must be in a public subnet and allocate EIP to it
        - amount of traffic supported depends on the instance size
        - high availability using auto scaling groups, multiple subnets in different AZ's and a script to automate failover
    - AWS NAT gateway
        - scales automatically up to 10Gbps
        - auto managed, no need to patch
        - not associated with security groups
        - automatically assigned a public ip address
        - remember to update route tables
        - no need to disable source/destination checks)
- Bastion (jump box): to administer EC2 instances in private subnets using SSH or RDP
- **VPC flow log**:
    - enables you to capture information about the IP traffic going to and from network interfaces
    - data is stored using CloudWatch Logs
    - Usage:
        - troubleshoot why specific traffic is not reaching an instance
        - diagnose overly restrictive security group rules
        - security tool to monitor the traffic that is reaching your instance
- provision an EC2 instance with an elastic IP address

Direct Connect

- dedicated network connection to AWS using Ethernet VLAN trunking (802.1Q)
- customer → dedicated line → (private rack → cross connect → aws rack) → dark fibre → aws data centre
- Advantages:
    - provides a more consistent network experience than internet-based connections
    - reduce costs when using large volumes of traffic
    - increase reliability, does not involve the internet (variability in internet-based connectivity)
    - increase bandwidth: 10Gbps, 1 Gbps, sub 1 Gbps (through AWS Direct Connect partners)

ELB: Elastic Load Balancer

- never given a static IP address, provides only a single CNAME entry point for DNS configuration

- Features:
    - achieve high availability for your application by distributing traffic across healthy instances in a multiple AZs
    - highly available within a region
    - automatically scalable based on collected metrics
    - integrates with Auto Scaling to scale the EC2 instances behind the load balancer
    - Load balancer types:
        - internet-facing:
            - reference ELB by its DNS name, instead of by the IP in order to provide a single, stable entry point
        - internal application-facing:
            - routes traffic to EC2 instances in VPCs with private subnets
    - supports integrated certificate management and SSL termination
        - SSL termination: terminate the SSL connection and decrypt requests from clients before sending requests to the back-end EC2 instances (optional: enable authentication on back-end instances)
    - ensure traffic not routed to unhealthy or failing instances by health checks
    - security: uses VPC to route traffic internally between application tiers and exposes only internet-facing public IP addresses.
- Limitations:
    - ELB in VPC supports IPv4 addresses only
        - ELB in EC2-Classic supports both IPv4 and IPv6
    - No Server Name Indication (SNI) support: in case you want to host multiple websites on EC2 instances behind ELB with a single SSL certificate, you will need to add a **Subject Alternative Name (SAN) for each website to the certificate** to avoid site users seeing a warning message
- Configuration options:
    - Idle connection timeout:
        - idle timeouts for the connections to clients and to the back-end instances that is triggered when no data is sent.
        - default: 60 secs (change it for lengthy operations, such as file uploads)
        - best practice: enable keep-alive when using HTTP and HTTPS
    - Cross-Zone load balancing:
        - best practice: approximately equivalent numbers of instances in each AZ
        - help to reduce impact of misconfigured client DNS cache settings
    - **Connection draining**:
        - stops sending requests to instances that are deregistering or unhealthy, while keeping the existing connections open to complete in-flight requests
        - specify a maximum time to keep the connections alive (1 to 3600 seconds)
        - default: **300** seconds
    - Proxy protocol:

        - forwards a human-readable header to the request header with connection information such as source and destination IP addresses and port numbers
        - before enable the setting, check if ELB is already behind a proxy server with proxy protocol enabled (otherwise duplicated headers might result in errors on your back-end instances)
        - support the following **X-Forwarder header**s:
            - X-Forwarded-For
            - X-Forwarded-Proto
            - X-Forwarded-Port
    - Sticky sessions (session affinity):
        - maps a user's session to a specific instance using application or ELB session cookie
            - ensures that all requests from the user during the session are sent to the same instance
            - ELB cookie name: AWSELB
        - default: not sticky: each request is routed independently to the registered instance with the smallest load
    - Health checks: test the status of EC2 instances behind the ELB
        - status can be "InService" or "OutOfService"
        - configurations:
            - a ping, a connection attempt, or a page that is checked periodically
            - time interval between health checks
            - amount of time to wait to respond in case the health check page includes a computational aspect
            - threasholds:
                - unhealthy threshold: how many intervals with no response before flagging as "Out of Service"
                - healthy threshold: how many intervals with response before flagging as "In Service"
- Listeners:
    - ELB must have 1 or more listers configured
    - configure protocols and ports for front-end and back-end connections
    - supported protocols at 2 different OSI (Open System Interconnection) layers:
        - HTTP + HTTPS (OSI Layer 7: application layer)
        - TCP (OSI Layer 4: transport layer)
        - SSL/TLS (SSL offload)
    - checks for connection requests
-

Developer Tools

CodeCommit

- store code in private git repositories

CodePipeline

- release software using continuous delivery
- build, test, deploy code

CodeDeploy

- automated deployments

Management Tools

CloudFormation

- create and manage resources with templates and stacks
    - stack: a container of related resources defined by a template
        - If stack creation fails, CloudFormation rolls back all changes by deleting the resources that it created.
        - variations are addressed using parameters
        - deployed and managed in a single template file using JSON
- Deletion policy: When a stack is to be deleted, CloudFormation deletes the stack and all of the resources in that stack by default. A delete policy can used to define to retain which resources.
- create a template of the existing infrastructure to capture and redeploy applications you already have running
    - describe your resources once, and provision the same resources over and over in multiple regions
- enables the provisioning and management of a group of integrated AWS resources
    - landscape configuration
    - network layout
    - security policies
- CloudFormation + AMI = significant reduction in time for deployments from weeks to minutes with consistency, repeatability and reliability
- Use Cases:
    - Quickly launch new test environments
    - Reliably replicate configuration between environments
    - Launch applications in new regions
-

CloudWatch

- monitor resources and applications, e.g. performance monitoring
    - monitors the hypervisor, not the guest OS
    - does not monitor memory
- Dashboards: custom dashboards
- Alarms: set alarms that notify you when particular thresholds are hit
- Events: respond to state changes in your AWS resources
- CloudWatch Logs:
    - monitor, store, and access log files from EC2 instances
    - Logs agent:
        - automated way to send log data to CloudWatch Logs on Linux or Ubuntu
    - e.g. track the number of errors in logs and send a notification if an error rate exceeds a threshold
    - Logs can be retained indefinitely or according to an aging policy

- intervals:
    - standard monitoring = 5 min (free of charge)
    - detailed monitoring = 1 min
        - aggregation of metrics across a length of time and across AZs within a region
- custom metrics via API using PUT:
    - metric as name-value pairs
- Limitations:
    - up to 5000 alarms per AWS account
    - metrics retained for 2 weeks by default
        - solution: move the logs to a persistent store (S3 or Glacier)

CloudTrail

- track user activity and API usage, e.g. for auditing, including:
    - name of the API
    - identity of the caller
    - time of the API call
    - request parameters
    - response elements returned by the AWS service
    - within 15 minutes
    - publishes about every 5 minutes
- Trail Types:
    - Trail that applies to all regions: (best practice)
        - log files to the single S3 bucket
        - one SNS topic for all regions
        - optional: log files from all regions to a single CloudWatch Logs log group
    - Trail applies to one region
- by default: log files are stored indefinitely and encrypted using SSE.
    - define S3 lifecycle rules to archive or delete log files automatically
- Use Cases:
    - external compliance audits: ensures compliance with internal policies and regulatory standards
    - unauthorized access to your AWS account

Auto Scaling

- scales EC2 capacity automatically by launch and release EC2 instances according to criteria that you define
- Auto scaling may cause you to reach limits of other services, e.g. number of EC2 instances within a region (default: 20)
    - raise support case to increase your soft limits
- Auto **Scaling Plans**:
    - Maintain current instance levels:
        - maintain a minimum or specified number of running instances at all times

        - periodic health check on running instances within an Auto Scaling group
        - terminates unhealthy instance and launches a new one
    - Manual scaling:
        - specify the change in the maximum, minimum, or desired capacity of your Auto Scaling group
        - AWS support: pre-warm ELB for extremely large-scale events
    - Scheduled Scaling:
        - define exactly when to increase or decrease the number of instances in the auto scaling group
        - e.g. recurring events such as end-of-month, quarter, or year processing, or scheduled and recurring automated load and performance testing
    - Dynamic Scaling:
        - define parameters in a scaling policy to control the Auto Scaling process
        - e.g. adds more EC2 instances to the web tier when the network bandwidth, measured by CloudWatch, reaches a certain threshold
- Components:
    - **Launch configuration**
        - template to create new instances: configuration name, AMI, EC2 instance type, security group, and instance key pair
            - EC2 classic: reference security group by name, e.g. 'Web'
            - VPC: reference security group by ID only
        - limit: up to 100 per region
    - **Auto Scaling group**
        - a collection of EC2 instances managed by the Auto Scaling service
        - contains configuration options that control when Auto Scaling should launch new instances and terminate existing instances
        - configurations:
            - **required: name, min and max number of instances**
            - optional: desired capacity (default: min)
        - a launch configuration can reference:
            - on-demand instances (default) or (not both)
            - spot instances by referencing a maximum bid price in the launch configuration
    - Scaling policy (optional)
        - associate CloudWatch alarms to adjust Auto Scaling group dynamically when a threshold is crossed
        - a set of instructions that tell Auto Scaling whether to scale out or to scale in.
            - increase or decrease by a specific number
            - target a specific max/min number
            - adjust based on a percentage
            - scale by steps that vary based on the size of the threshold trigger
        - multiple scaling policies per Auto Scaling group possible

        - best practice: scale out quickly and scale in slowly by setting ‘cooldown period’
            - partial instance hours consumed are billed as full hours
            - optimize bootstrapping
            - prefer stateless
    - Roll out a patch at scale: reference to a new
-

Config

- provides resource inventory,configuration history and configuration change notifications to enable security and governance
- Config represents relationships between resources, so you can assess how a change to one resource may affect other resources
- Use Cases:
    - Discovery resources that exist in your account, record their current configuration, and capture any changes to them
    - Change management: streams the configuration changes (create, update, delete) to SNS so that you are notified of all configuration changes
    - Troubleshooting: identify the recent configuration changes to your resources
    - Security and incident analysis: monitor the configurations continuously and evaluate them for potential security weaknesses at any single point in the past

Service Catalog

- create and use standardized products

Trusted Advisor

- optimize performance and security
- access in AWS Management Console or with the AWS Support API
- categories:
    - cost optimization
    - performance
    - security
    - fault tolerance
- color coding:
    - red: action recommended
    - yellow: investigation recommended
    - green: no problem detected
- standard checks: **4** checks are free
    - service limits: >80% alerts
    - security groups: specific ports unrestricted (0.0.0.0/0)
    - IAM use
    - MFA on root account
- business or enterprise Support plan checks: over 50 checks

- scans environment for ways to save money, increase security and performance

Resource Groups

- a resource group is a collection of resources that share one or more tags
- contains information:
    - region, name, health checks
    - for EC2: public and private IP addresses
    - for ELB: port configurations
    - for RDS: database engine, etc.
- Tags:
    - key value pairs attached to AWS resources
    - metadata
    - tags can be inherited: autoscaling, CloudFormation and Elastic Beanstalk can create other resources

OpsWorks

- automate operations with chef
- DevOps application management service
- chef consists of recipes (cook books) to maintain a consistent state
- Components:
    - **stack**: a container for AWS resources
        - manages these resources as a group and defines some default configuration settings
        - optional: isolate some components using VPC
        - you can use IAM or OpsWorks stack to manage user permissions
    - **layer**: represents a set of resources that serve a particular purpose, such as load balancing, web applications, or hosting a database server
        - customizes layers by modifying the default configurations or adding Chef recipes to perform tasks
        - supports lifecycle events that automatically run a specified set of Chef recipes at the appropriate time on each instance
        - controls over which packages are installed, how they are configured, how applications are deployed, and more
    - **app**: application and related files stored in a repository such as S3 bucket or Git repo. When you deploy an app, OpsWorks triggers a Deploy event, which runs the Deploy recipes on the stack's instances.
- Monitoring: OpsWorks metrics in CloudWatch
- Case Cases:
    - Host multi-tier web applications: models and visualizes application with layers using community built or own recipes.
    - Support continuous integration: supports DevOps principles, such as continuous integration. Everything can be automated.

Security & Identity

IAM: Identity & Access Management

- manage user access and encryption keys
- IAM is for AWS resources and NOT:
    - an identity store or authorization system for your applications
        - IAM permissions are to manipulate AWS infrastructure
        - use application user repositories, AWS Directory Service for AD or Cognito for identity management for mobile applications
    - an operating system identity management
        - you are in control of your OS
        - use AD or LDAP machine-specific accounts
- Features:
    - universal, applies to all regions consistently
    - shared access to your AWS account
    - integrated with AWS marketplace
    - custom password rotation policy
    - rotate access key process:
        - create a new access key for the user (AWS allows two access keys to be valid simultaneously)
        - reconfigure all applications to use the new access key
        - disable the original access key
        - verify the operation of all application
        - delete the original access key
- Principal:
    - an IAM entity that is allowed to interact with AWS resources.
    - can be permanent or temporary, human or application
    - Types:
        - root users:
            - permanent, complete admin access including closing the account
            - cannot be limited
            - best practice:
                - always setup MFA on the root account
                - do not use it for everyday tasks, even the administrative ones
                - only to create a new “IAM Administrators” group and assign the managed policy “IAMFullAccess” and assign a new IAM user “Administrator” to it with password access
                -
            - can be used for console or programmatic access
        - IAM users:
            - persistent identities

            - represents individual people or applications
            - access controlled by granular policies that define permissions
            - can be removed by IAM administrator
        - roles/temporary security tokens:
            - grants specific privileges to specific actors for a set duration of time
            - access controlled by polic
            - expire after specific time interval
            - actors can be authenticated by AWS or trusted external system and receives a temporary security token from the AWS security token services (STS)
            - use cases:
                - EC2 roles: granting permissions to applications running on an EC2 instance
                    - removes the need to store credentials in a configuration file
                - Cross-account access: granting permissions to users from other AWS accounts, whether you control those accounts or not
                    - opposed to distributing access keys outside your organization
                - Federation: granting permissions to users authenticated by a trusted external system
                    - opposed to largely duplicate user repository to IAM users
                    - Identity Providers types (IdP):
                        - OpenID Connect (OIDC): federating web identities such as Facebook, Google, or Login with Amazon
                        - Security Assertion Markup Language 2.0 (SAML): integration of Active Directory or LDAP such as Active Directory Federation Services (ADFS), used to federate the internal directory to IAM
- Authentication:
    - username / password: authenticating to AWS console with user account
    - access key / access secret key: application authenticating to AWS API with user account
    - access key / session token: user or application using temporary security token
- Authorization / Policies:
    - policy documents contain one or more permissions
    - Permission:
        - effect: a single work: allow or deny

        - service: for what AWS service (granting access through IAM)
        - resource: specifies the specific infrastructure for which this permission applies
            - Amazon Resource Name (ARN) format: arn:aws:service:region:account-id:[resourcetype:]resource (wildcard values allowed)
        - action: specifies the subset of actions within a service that the permission allows or denies
            - wildcard allowed: e.g. Read*
        - condition: defines one or more additional restrictions that limit the actions allowed by the permission e.g. only from a specific IP address
    - Associating policies with principals:
        - User/Group policy: exists only in the context of the user to which they are attached
        - Managed policies: exist independently of any individual user and can be associated with many users or groups of users
            - using predefined managed policies ensures that when new permissions are added for new features, your users will still have the correct access
- granular permissions, consists of:
    - Granular enough to limit a single user to the ability to perform a single action on a specific resource from a specific IP during a specific time window.
    - users: end users
    - groups: collection of users under one set of permissions
    - roles: can be assigned to AWS resources
        - more secure than storing access key and secret access key on individual EC2 instances
        - easier to manage
        - universal, can used in any region
        - roles can **only** be assigned when that EC2 instance is being provisioned
    - policies: a **JSON** document that defines 1 or more permissions
- new users:
    - new users have **no permissions** when first created
    - new users are assigned **access key ID & secret access keys** when first created (if you lose them, you have to regenerate them)
        - used to access AWS via the APIs and command line
        - cannot be used to login into the console
- Resolving permissions conflicts:
    - initially the request is denied by default
    - all the appropriate policies are evaluated; if there is an explicit “deny” found in any policy, the request is denied and evaluation stops
    - if no explicit deny is found and an explicit allow is found in any policy, the request is allowed

        - if there are no explicit allow or deny permissions found, then the default deny is maintained and the request is denied
        - exception: AssumeRole: the policy cannot expand the privileges of the role
- identity federation (including active directory, facebook, linkedin etc.)
    - Single Sign-On using ADFS (SAML)
    - LDAP integration
- Active Directory Integration:
    - user browses to ADFS URL
    - user authenticates against AD
    - user receives a SAML assertion
    - user's browser posts the SAML assertion to the AWS sign-in endpoint for SAML
    - user's browser receives the sign-in URL and is redirected to the console
- MFA = multifactor authentication:
    - requires you to verify your identity with both something you know and something you have
- temp access for users/ devices/ services
- provide temporary access for users/devices and services where necessary
- integrates with many different AWS services
- PCI DS compliance

WAF: Web Application Firewall

- filter malicious web traffic
- Condition sets:
    - IP match
    - String match
    - SQL Injection match
    - Size constraint
    - Cross-Site scripting match
    -

Inspector

- analyze application security

KMS (Key Management Service) and CloudHSM (Hardware Security Module)

- KMS: generates, stores, enable/disable, and deletes symmetric keys
    - CMK (customer managed keys): used inside KMS
        - can encrypt or decrypt up to 4KB of data directly
        - used to generate data keys
        - can never leave KMS unencrypted
    - Data keys:
        - can encrypt large data objects outside KMS
        - as plaintext or ciphertext (KMS tracks which CMK was used to encrypt the data key)

        - can leave KMS unencrypted
- CloudHSM: provides with secure cryptographic key storage by making HSMs available on the AWS cloud

Directory Service

- host and manage active directories in the cloud
- **Directory types**:
    - Directory service for Microsoft Active Directory (Enterprise)
        - managed Microsoft AD hosted on AWS cloud
        - integration with AWS applications
        - supports trust relationship setup between AWS hosted directory and on-premises directories
    - Simple AD:
        - Microsoft AD compatible directory powered by Samba 4
        - least expensive option, e.g. <5000 users
        - provides daily automated snapshots and point-in-time recovery
        - not supported: no advanced DS features, DNS dynamic update, schema extensions, MFA, LDAP, FSMO roles, PowerShell AD endlets
    - AD Connector: a proxy service for connecting your on-premises Microsoft AD to the AWS cloud
        - uses the existing AD
        - forwards sign-in requests to your AD domain controllers for authentication and provides the ability for applications to query the directory for data
        - supports RADIUS-based MFA
        - provides extended security policies such as password expiration, password history, and account lockouts

Analytics

EMR (Elastic MapReduce)

- managed Hadoop framework
- **Options**:
    - instance type
    - number of nodes
    - Hadoop distribution and version
    - applications like Hive, Pig, Spark, or Presto
- Storage types:
    - HDFS: stores data to instance storage (ephemeral) or to EBS storage
    - EMRFS: stores data on S3
    - Combination of local HDFS and EMRFS
- Use Cases:
    - log processing
    - clickstream analysis
    - genomics and life sciences

Kinesis

- real-time processing of streaming big data, best for continuous data streams
- Kinesis Platform services:
    - Kinesis Firehose: loads massive volumes of streaming data into AWS
        - receives stream data and stores in S3, Redshift or Elasticsearch
        - for Redshift destination: first written to S3, then COPY to Redshift
    - Kinesis Streams: builds custom application for more complex analysis of streaming data in real time
        - limitless near real-time data processing by distributing incoming data across a number of shards and then executed on consumers, which read data from the shards and run the Kinesis Streams application.
    - Kinesis Analytics: analyzes streaming data real time with standard SQL
- used to consume big data, streams large amounts of social media, news feeds, logs etc. into the cloud
- process large amounts of data:
    - redshift for business intelligence
    - EMR for big data processing

Data Pipeline

- orchestration for data-driven workflows, best for regular batch processes
- executes activities periodically that represent common scenarios, such as moving data from one location to another, running Hive queries, and so forth
- If an activity fails, retry is automatic.

Machine Learning

- build smart applications quickly and easily

Internet Of Things

AWS IoT

- connect devices to the cloud

Mobile Services

Cognito

- user identity and app data synchronization

Device Farm

- test android, FireOS, and iOS apps on real devices in the cloud

Mobile Hub

- build, test, and monitor mobile apps

Mobile Analytics

- collect, view and export app analytics

SNS: Simple Notification Service

- **push** notification service
- Features:
    - publish-subscribe messaging
    - group multiple recipients using topics
        - topic: access point allowing recipients to dynamically subscribe for identical copies of the same notification
        - one topic can support deliveries to different endpoint types e.g. iOS, Android and SMS
    - instantaneous, push-based delivery (no polling)
    - simple API
    - flexible message delivery over multiple transport protocols
    - web-based AWS management console offers a point-and-click interface
- SNS publishers:
    - communicate to subscribers asynchronously by sending a message to a topic
    - topic: a logical access point containing a list of subscribers and the method used to communicate to them
- **SNS subscribers:**
    - delivery methods: SQS, HTTP, HTTPS, email, SMS, Lambda, mobile applications
- Use cases:
    - Fanout: allows parallel asynchronous processing.
        - sends message to a topic which then replicated and pushed to multiple subscribers (e.g. SQS queues)
    - Application and system alerts: SMS and/or email notifications triggered by predefined thresholds
    - Push Email and Text messaging: send messages to individuals or groups via email and/or SMS, e.g. push targeted news headlines

        - Mobile Push notifications: send messages directly to mobile applications, e.g. new update available
- Pricing:
    - 50ct per 1 million SNS requests
    - 6ct per 100K deliveries over HTTP
    - 2$ per 100K deliveries over Email
    - 75ct per 100 SMS

Application Services

API Gateway

- build, deploy and manage APIs

SQS: Simple Queue Service

- distributed message queue service that sits between a producer and consumers to quickly and reliably cache that message
    - access through HTTP/HTTPS
    - throughput scales horizontally (more threads more throughput)
    - average response time: 20ms
- decouples components of an application
    - resolves issues if the producer is producing faster than consumer is processing or
    - producer or consumer are only intermittently connected to network
- Features:
    - message-oriented API (pull-based)
    - at-least-once delivery
    - supports multiple readers and writers with the same queue
    - **12 hours** visibility time-out by default
    - retention period o**f 14 days**
    - can apply auto-scaling
- Limitations:
    - no exactly-once delivery
        - solution: design idempotent systems
    - no FIFO supported
        - solution: add sequencing information in each message
    - no prioritization of items supported
    - message size: up to 256KB billed at 64KB chunks
    - application-level tracking when using multiple queues
- Message Lifecycle:
    - A sends a message to a SQS Queue URL . The message is assigned a globally unique ID (up to 100 characters) and is redundantly distributed across SQS servers. (in-flight)
        - metadata (message attributes) are supported such as timestamps, geospatial data and signatures (**up to 10 attributes**)
    - B receives the message and makes it hidden to other consumers for the duration of the visibility timeout. The message remains in the queue. (invisible)
    - B processed the message and deleted the message from the queue (deleted)
        - You need the message's receipt handle (up to 1024 characters) to delete or to change the visibility of the message.

    - Delay queue: postpone the delivery of new messages in a queue for a specific number of seconds (up to 15 minutes) (delayed)
- Immediate return & Long Polling in ReceiveMessage
    - default: ReceiveMessage checks for the existence of a message and return immediately, either with or without a message
    - Long polling:
        - if there is no message in the queue, then the call will wait up to 20s for a message to appear before returning.
        - reduces delay and load on your client
- Synchronous & asynchronous SendMessage:
    - default: SendMessage call waits until the message is durably stored in SQS
    - asynchronous approach to reduce latency: client-side batching on top of SQS libraries that delays enqueue of messages to SQS and transmit a set of messages in a batch
        - messages might be lost when your client process or host dies for any reason
- DLQ (dead letter queues): for messages that could not be successfully processed
- Access control through IAM policies
- Pricing:
    - first 1 million requests per month free, 50ct per 1 million requests per month
    - 1 request can have from 1 to 10 messages, up to a total payload of 256KB
        - each 64KB chunk is billed as 1 request

SWF: Simple Workflow Service

- workflow service for coordinating application components
- Features:
    - retention period up to **1 year** for workflow executions
    - task-oriented API
    - task is assigned only once and is never duplicated
    - keeps track of all tasks and events in an application
- SWF Workflows: coordinates and manages the execution of activities that can be run asynchronously across multiple computing devices and that can feature both sequential and parallel processing
    - Workflow Domain: specifies a domain for all the components of workflows, such as workflow types and activity types
        - workflows in different domains cannot interact with one another
        - workflow/activity type id: domain, name, and version
    - Workflow History: a detailed, complete, and consistent record of every event occurred since the workflow execution started
    - Workflow execution id: workflow ID and run ID
    - Workflow Closure: status can be completed, canceled, failed or timed out
- SWF Actors:
    - Workflow Starters: initiates a workflow

        - Deciders: processes events and control the flow of activity tasks
        - Activity Workers: polls for new tasks and carries out the activity tasks e.g. executable code, webservice calls, human actions, scripts
            - can be on Cloud e.g. EC2 or on your own premises
- SWF Tasks:
    - Activity tasks: tells an activity worker to perform its function
    - Lambda tasks: executes a Lambda function
    - Decision tasks: tells a decider that the state of the workflow execution has changed, contains a paginated view of the entire workflow execution history
    - Activity/Decision task id: a generated unique task token
- Task lists: a dynamic queue of tasks
- Long polling: deciders and activity workers communicate with SWF using long polling.

AppStream

- low latency application streaming

Elastic Transcoder

- easy-to-use scalable media transcoding
- provides transcoding presets for popular output formats
- pay based on the minutes and resolution

SES

- email sending service

CloudSearch

- managed search service

Game Development

Enterprise Applications

WorkSpaces (VDI)

- desktops in the cloud
- replaces windows PC in the cloud (PCoIP)
- persistent (EBS)
- all data on D drive backed up every 12 hours
- AWS account not needed to login to workspaces
- AD domain not needed, can use free client app
- integration with existing AD domain possible
- by default, users can personalize their workspaces with wallpaper, icons, shortcuts etc. and have local admin access to install apps

WorkDocs

- secure enterprise storage and sharing service
- “dropbox”

WorkMail

- secure email and calendaring service
- “exchange”

AWS Resource Provisioning and Management.

- self-service using the AWS API:
    - AWS mangement console
    - AWS CLI
    - SDKs & Libraries
    - 3rd Party Tools

Questions & Answers

1. Amazon Glacier is designed for:
    a. infrequently accessed data
    b. data archives
2. You configured ELB to perform health checks on these EC2 instances. If an instance fails to pass health checks, which statement will be true?
    a. The ELB stops sending traffic to the instance that failed its health check.
3. You are building a system to distribute confidential training videos to employees. Using CloudFront, what method could be used to serve content that is stored in S3, but not publically accessible from S3 directly?
    a. Create an Origin Access Identity (OAI) for CloudFront and grant access to the objects in your S3 bucket to that OAI. An Origin Access Identity is a special user that you will set up the CloudFront service to use to access your restricted content.
4. Which of the following will occur when an EC2 instance in a VPC (Virtual Private Cloud) with an associated Elastic IP is stopped and started?
    a. All data on instance-store devices will be lost.
    b. The underlying host for the instance is changed.
5. In the basic monitoring package for EC2, Amazon CloudWatch provides the following metrics:
    a. hypervisor visible metrics such as CPU utilization.
6. Which is an operational process performed by AWS for data security?
    a. Decommissioning of storage devices using industry-standard practices.
7. To protect S3 data from both accidental deletion and accidental overwriting, you should:
    a. Enable S3 versioning on the bucket
8. What is an AWS region?
    a. A region is a geographical area that consists of different availability zones. Each region consists of 2 (or more) availability zones.
9. What does an AWS region consist of?
    a. An independent collection of AWS computing resources in a defined geography.
10. Which statement best describes availability zones?
    a. Distinct locations from within an AWS region that are engineered to be isolated from failures.
11. An AWS VPC is a component of which AWS service?
    a. Networking Service
12. What is a VPC?
    a. Virtual Private Cloud
13. Which AWS service is specifically designed for developers to upload their code to and then it will automatically handle the provisioning of those resources that are required to host that code?
    a. Elastic Beanstalk

14. Which AWS service allows you to run code without having to worry about provisioning any underlying resources (such as virtual machines, databases etc.)
    a. Lambda
15. Amazon's highly scalable DNS service is known as …
    a. Route 53
16. Which AWS compute service is specifically designed to assist you in processing large data sets?
    a. Elastic Map Reduce
17. What is the difference between Elastic Beanstalk and CloudFormation?
    a. Elastic Beanstalk automatically handles the deployment, from capacity provisioning, load balancing, auto-scaling to application health monitoring based on the code you upload to it
    b. CloudFormation is an automated provisioning engine designed to deploy entire cloud environments via a JSON script
18. Which AWS service is used a CDN to distribute content around the world?
    a. CloudFront
19. Where would be a durable place to store flat files on the AWS platform?
    a. S3
20. Which AWS service would be the best choice for long term data archival?
    a. Glacier
21. What AWS service consists of the following database services: SQL, MySQL, MariaDB, PostgreSQL, Aurora, Oracle?
    a. Relational Database Services (RDS)
22. What AWS service would you use primarily for data warehousing?
    a. Redshift
23. What AWS service is used for collating large amounts of data streamed from multiple sources?
    a. Kinesis
24. You need to create new users to access the AWS console and to set password rotation policies for these new users. Which AWS service would best fit your requirements?
    a. Identity Access Management (IAM)
25. You need to supply auditors with logs as to who provisions which resources on your AWS platform. Which service would best suit this?
    a. CloudTrail
26. You need a configuration management service to allow your system administrators to configure and operate your web applications using Chef. Which AWS service would best suit your needs?
    a. Opsworks
27. You are a digital media agency and you need to convert your media files into different formats to suit different devices. Which AWS service should you consider using to meet these needs?
    a. Elastic Transcoder
28. Which statement best describes IAM?

    a. IAM allows you to manage users, groups and roles and their corresponding level of access to the AWS platform.
29. Which are the features of IAM?
    a. centralised control of your AWS account
    b. integrates with existing active directory account allowing single sign on
    c. fine-grained access control to AWS resources
    d. the ability to create user/group/roles
    e. allows you to set up your own password rotation policy
30. Power User Access allows …
    a. Access to all AWS services except for management of groups and users within IAM.
31. What level of access does the “root” account have?
    a. Administrator Access
32. You are a solutions architect working for a large engineering company who are moving their existing legacy hardware to AWS. You have configured their first AWS account and you have setup IAM. Your company will be primarily based out of West Germany, however they will have a small subsidiary operating out of South korea and you will need an AWS environment configured there as well. Which of the following statements is true:
    a. You will need to configure users and policy documents only once, as these are applied globally.
33. You have a client who is considering moving to AWS services and do not yet have an account. What is the first thing the company should do to setup an AWS Account?
    a. Set up an account using their company email address.
34. You are a security administrator working for a hotel chain. You have a new member of staff who has started as a systems administrator and they will need full access to the AWS console. You have created the user account and generated the access key id and the secret access key. You have moved this user into the group where the other administrators are and you have provided the new user with their secret access key and their access key id. However when they go to log into the AWS console, they cannot sign in. what could be the cause of this?
    a. You cannot log into the AWS console using the Access Key ID and Secret Access Key, instead you must generate a password for the user and supply the user with this password, as well as the unique link to sign into the AWS console.
35. What is an additional way to secure IAM for both the root login and new users alike?
    a. Implement multi-factor Authentication for all accounts.
36. By default when you create a new user in the IAM console, what level of access do they have?
    a. No access to all AWS services.
37. In what language are policy documents written in?
    a. JSON
38. S3 has what consistency model for PUTs of new objects?
    a. Read after write consistency
39. What is AWS storage gateway?

    a. It's an on-premise virtual appliance that can be used to cache S3 locally at a customer's site.
40. One of your users is trying to upload a 7.5GB file to S3 however they keep getting the following error message: "Your proposed upload exceeds the maximum allowed object size." What is a possible solution for this?
    a. Design your application to use the multipart upload API for all objects.
41. What does RRS stand for when talking about S3?
    a. Reduced Redundancy Storage
42. You have been asked by your company to create an S3 bucket with the name "mys3bucket" in the EU West region. What would be the URL for this bucket?
    a. https://s3-eu-west-1.amazonaes.com/mys3bucket
43. What is Amazon Glacier?
    a. an AWS service designed for long term data archival
44. What does S3 stand for?
    a. Simple Storage Service
45. You are a solutions architect who works with a large digital media company. The company has decided that they want to operate within the Japanese region and they need a bucket called "testbucket" set up immediately to test their web application on. You log in to the AWS console and try to create this bucket in the Japanese region however you are told that the bucket name is already taken. what should you do to resolve this?
    a. Bucket names are global, not regional. This is a popular bucket name and is already taken. You should choose another bucket name.
46. What is the availability on RRS?
    a. 99.99%
47. What is the durability on RRS?
    a. 99.99%
48. What is the durability on S3?
    a. 99.999999999%
49. What is the availability on S3?
    a. 99.99%
50. What is the minimum file size that I can store on S3?
    a. 1 Byte
51. The difference between S3 and EBS is that EBS is object based where as S3 is block based.
    a. False
52. S3 has eventual consistency for which HTTP methods?
    a. overwrite PUTs and DELETEs
53. You work for a busy digital marketing company who currently store their data on premise. They are looking to migrate to AWS S3 and to store their data in buckets. Each bucket will be named after their individual customers, followed by a random series of letters and numbers. Once written to S3 the data is rarely changed, as it has already been sent to the end customer for them to use as they see fit. However on some

occasions, customers may need certain files updated quickly, and this may be for work that has been done months or even years ago. You would need to be able to access this data immediately to make changes in that case, but you must also keep your storage costs extremely low. The data is not easily reproducible if lost. Which S3 storage class should you choose to minimise costs and to maximise retrieval times?
    a. S3-IA (low cost, but fast retrieval times)
54. You need to use an object based storage solution to store your critical, non-replaceable data in a cost effective way. This data will be frequently updated and will need some form of version control enabled on it. Which S3 storage solution should you use?
    a. S3
55. You work for a health insurance company who collects large amounts of documents regarding patients health records. This data will be used usually only once when assessing a customer and will then need to be securely stored for a period of 7 years. In some rare cases you may need to retrieve this data within 24 hours of a claim being lodged. which storage solution would best suit this scenario? You need to keep your costs as low as possible.
    a. Glacier
56. You run a meme creation website that frequently generates meme images. The original images are stored in S3 and the meta data about the memes are stored in DynamoDB. You need to store the memes themselves in a low cost storage solution. If an object is lost, you have created a Lambda function that will automatically recreate this meme using the original file in S3 and the metadata in Dynamodb. Which storage solution should you consider to store this non-critical, easily reproducible data on in the most cost effective solution as possible?
    a. S3-RRS
57. You run a popular photo sharing website that is based off S3. You generate revenue from your website via paid for adverts, however you have discovered that other websites are linking directly to the images on your site, and not to the HTML pages that serve the content. This means that people are not seeing your adverts and every time a request is made to S3 to serve an image it is costing your business money. How could you resolve this issue?
    a. Remove the ability for images to be served publicly to the site and then use signed URL’s with expiry dates.
58. EBS snapshots are backed up to S3 in what manner?
    a. incrementally
59. Do amazon EBS volumes persist independently from the life of an Amazon EC2 instance, for example, if I terminated an EC2 instance, would that EBS volume remain?
    a. Only if instructed to when created
60. Can I delete a snapshot of an EBS volume that is used as the root device of a registered AMI?
    a. **No**, you must deregister the AMI before being able to delete the root device
61. A placement group can be deployed across multiple Availability Zones.
    a. False

62. While creating the snapshots using the command line tools, which command should I be using?
    a. ec2-create-snapshot
63. Can you attach an EBS volume to more than one EC2 instance at the same time?
    a. No
64. A placement group is ideal for:
    a. EC2 instances that require low latency and high network throughput across a **single** availability zone.
65. Using the console, I can add a role to an EC2 instance, after that instance has been created and powered up.
    a. False
66. I can change the permissions to a role, even if that role is already assigned to an existing EC2 instance, and these changes will take effect immediately.
    a. True
67. Does Route 53 support MX records?
    a. Yes
68. Route 53 is named so because …
    a. The DNS port is on port 53 and Route53 is a DNS service.
69. Route53 does not support zone apex records (or naked domain names).
    a. Incorrect
70. Route53 is amazon's DNS service.
    a. True
71. There is a limit to the number of domain names that you can manage using Route53.
    a. True and False. There is a soft limit of 50 domain names by default however this limit can be raised by contacting AWS support.
72. What AWS DB platform is most suitable for OLTP?
    a. RDS / DynamoDB
73. When replicating data from your primary RDS instance to your secondary RDS instance, what is the charge?
    a. No charge
74. What AWS service is best suited for non relational databases?
    a. DynamoDB
75. When you add a rule to an RDS security group you do not need to specify a port number or protocol?
    a. False (e.g. MySQL: 3306)
76. If you are using Amazon RDS Provisioned IOPS storage with MySQL and Oracle database engines what is the maximum size RDS volume you can have by default?
    a. 6TB
77. What happens to the I/O operations while you take a database snapshot?
    a. I/O operations to the database are **suspended** for the duration of the snapshot
78. What AWS service is best used for Business Intelligence Tools/Data Warehousing?
    a. Redshift

79. In RDS when using multiple availability zones, can you use the secondary database as an independent read node?
    a. No
80. Amazon's Elasticache uses which 2 engines?
    a. Redis & Memcached
81. By default, the maximum provisioned IOPS capacity on an Oracle and MySQL RDS instance (using provisioned IOPS) is **30000** IOPS
    a. True
82. Security groups act like a firewall at the instance level whereas … are an additional layer of security that act at the subnet level:
    a. Network ACLs
83. How many VPC's am I allowed in each AWS Region by default?
    a. 5 (soft limit)
84. VPC stands for:
    a. Virtual Private Cloud
85. How many internet gateways can I attach to my custom VPC:
    a. 1
86. You have a VPC with both public and private subnets. You have 3 EC2 instances that have been deployed into the public subnet and each has internet access. You deploy a 4-th instance using the same AMI and this instance does not have internet access. What could be the cause of this?
    a. The instance needs either an Elastic IP address or a Public IP address assigned to it.
87. What does Amazon SWF stand for?
    a. Simple Work Flow
88. What does Amazon SES stand for?
    a. Simple Email Service
89. What happens when you create a topic on Amazon SNS?
    a. An Amazon Resource Name is created.
90. What is the difference between SNS and SQS?
    a. SNS is push notification service, where as SQS is message system that requires worker nodes to poll the queue.
91. What application service allows you to decouple your infrastructure using message based queues?
    a. SQS
92. What does a "domain" refer to in Amazon SWF?
    a. A collection of related workflows
93. By default, EC2 instances pull SQS messages from an SQS queue on a FIFO basis.
    a. False
94. Amazon's SQS service guarantees a message will be delivered at least once.
    a. True
95. Amazon SWF ensures that a task is assigned only once and is never duplicated.
    a. True

96. Amazon SWF restrict me to use specific programming languages.
    a. False
97. Amazon SWF is designed to help users …
    a. coordinate synchronous and asynchronous tasks
98. In RDS, what is the maximum value I can set for my backup retention period?
    a. 35 days
99. Automated backups are enabled by default for a new DB instance?
    a. True
100. Amazon RDS does not currently support increasing storage on a … DB instance.
    a. SQL Server
101. In what circumstances would I choose provisioned IOPS in RDS over standard storage?
    a. If you use production online transaction processing.
102. Amazon's S3 is …
    a. object based storage
103. In S3 with RRS the availability is …
    a. 99.99%
104. Amazon's EBS volumes are …
    a. block based storage
105. If I want to run a database on an EC2 instance, which is the most recommended Amazon storage option?
    a. EBS
106. In S3 the durability of my files is …
    a. 99.999999999%
107. Can you access Amazon EBS snapshots?
    a. Yes through the AWS APIs / CLI and AWS Console
108. A … is a document that provides a formal statement of one or more permissions.
    a. Policy
109. In a default VPC, all Amazon EC2 instances are assigned 2 IP addresses at launch, what are these?
    a. private and public ip address
110. If an Amazon EBS volume is the root device of an instance, can I detach it without stopping the instance?
    a. No
111. If you want your application to check whether a request generated an error then you look for an … node in the response from the Amazon RDS API.
    a. Error
112. EC2 instances can have credentials stored on them so that the instances can access other resources (such as S3 buckets) and AWS recommends that you do this instead of assigning roles.
    a. False
113. Can I move a reserved instance from one region to another?
    a. No

114. In S3 RRS the durability of my files is …
    a. 99.99%
115. In RDS, changes to the backup window take effect …
    a. immediately
116. In RDS what is the maximum size for a Microsoft SQL Server DB Instance with SQL Server Express edition?
    a. 10GB
117. In S3 what does RRS stand for?
    a. Reduced Redundancy Storage
118. Can I force a failover for any RDS instance that has Multi-AZ configured?
    a. Yes. You can do this as per rebooting the DB instance.
119. What does EBS stand for?
    a. Elastic Block Storage
120. You can conduct your own vulnerability scans within your own VPC without alerting AWS first?
    a. False
121. Reserved instances are available for multi-AZ deployments.
    a. True
122. Amazon's Glacier service is a Content Distribution Network which integrates with S3.
    a. False
123. MySQL installations default to port number:
    a. 3306
124. If an Amazon EBS volume is an additional partition (i.e. not the root volume), can I detach it without stopping the instance?
    a. Yes, although it may take some time.
125. Every user you create in the IAM systems starts with …
    a. No permissions
126. You can RDP or SSH into an RDS instance to see what is going on with the operating system.
    a. False
127. When creating a new security group, all inbound traffic is allowed by default.
    a. False
128. To save administration headaches, Amazon recommend that you leave all security groups in web facing subnets open on port 22 to 0.0.0.0/0 CIDR, that way you can connect where ever you are in the world.
    a. Incorrect. This is a massive security risk.
129. What are the four levels of AWS premium support?
    a. Basic, Developer, Business, Enterprise
130. As the AWS platform is PCI DSS 1.0 compliant, I can immediately deploy a website to it that can take and store credit card details. I do not need to get any kind of delta accreditation from a QSA.
    a. False

131. To help you manage your Amazon EC2 instances you can assign your own metadata in the form of …
    a. Tags
132. Which statement best describes Availability Zones?
    a. Distinct locations from within an AWS region that are engineered to be isolated from failures.
133. The service to allow Big Data Processing on the AWS platform is known as AWS "Elastic Big Data".
    a. False
134. Individual instances are provisioned in …
    a. Availability Zones
135. When using a custom VPC and placing an EC2 instance into a public subnet, it will be automatically internet accessible (i.e. you do not need to apply an elastic IP address or ELB to the instance)
    a. False
136. What is the underlying Hypervisor for EC2?
    a. Xen
137. The AWS platform is certified PCI DSS 1.0 compliant
    a. True
138. The AWS platform consists of how many regions currently?
    a. 12
139. How many copies of my data does RDS - Aurora store by default?
    a. 6
140. Amazon's product debut conference is held in Las Vegas each year and is known as …
    a. Re-Invent
141. In RDS, you are responsible for maintaining OS and application security patching, antivirus etc.
    a. False
142. What is the maximum response time for a Business level premium support case?
    a. 1 Hour
143. When I create a new security group, all outbound traffic is allowed by default.
    a. True
144. What types of RDS databases are currently available?
    a. Aurora, MySQL, Oracle, SQLServer, Postgres
145. I can enable multifactor authentication by using …
    a. IAM
146. When deploying databases on your own EC2 instances, it is recommended that you deploy these on magnetic storage rather than SSD storage as you get better performance.
    a. False
147. AWS DNS service is known as …
    a. Route53

148. Auditing user access/API calls etc across the entire AWS estate can be achieved by using …
    a. CloudTrail
149. You are a solutions architect working for a company that specialises in ingesting large data feeds (using Kinesis) and then analysing these feeds using EMR. The results are then stored on a custom MySQL database which is hosted on an EC2 instance which has 3 volumes, the root/boot volume, and then 2 additional volumes which are striped into a RAID 1. Your company recently had an outage and lost some key data and have since decided that they will need to run nightly backups. Your application is only used during office hours, so you can afford to have some down time in the middle of the night if required. You decide to take a snapshot of all three volumes every 24 hours. In what manner should you do this?
    a. Stop the EC2 instance and take a snapshot of each EC2 instance independently. Once the snapshots are complete, start the EC2 instance and ensure that all relevant volumes are remounted.
150. What are the valid methodologies for encrypting data on S3?
    a. Server Side Encryption: SSE-S3, SSE-C, SSE-KMS or a client library such as Amazon S3 Encryption Client
151. In IAM, when you first create a new user, certain security credentials are automatically generated. Which of the below are valid security credentials?
    a. Access Key ID, Secret Access Key
152. Amazon Web Services offer 3 different levels of support, which of the below are valid support levels:
    a. Enterprise, Business, Developer
153. You are a solutions architect working for a large digital media company. Your company is migrating their production estate to AWS and you are in the process of setting up access to the AWS console using IAM. You have created 5 users for your system administrators. What further steps do you need to take to enable your system administrators to get access to the AWS console?
    a. Generate a password for each user created and give these passwords to your system administrators.
154. Amazon S3 buckets in all regions provide which of the following?
    a. Read-after write consistency for PUTS of new objects and eventually consistent for overwrite PUTS and DELETES.
155. What function of an AWS VPC is stateless?
    a. Network ACL. A network ACL is applied on a subnet level, and traffic is stateless. You need allow both inbound and outbound traffic in order for EC2 instances in a network ACL to be able to communicate over a particular protocol. Security groups in contrast are stateful, which means that return traffic is automatically allowed, regardless of any outbound rules.
156. Which of the following services allows you root access (i.e. you can login using SSH)?
    a. EMR, EC2

157. When trying to grant an amazon account access to S3 using S3 ACL what method of identification should you use to identify that account with?
    a. email address of the account or the canonical user ID
158. You are a solutions architect working for a large oil and gas company. Your company runs their production environment on AWS and has a custom VPC. The VPC contains 3 subnets, 1 of which is public and the other 2 are private. Inside the public subnet is a fleet of EC2 instances which are the result of an auto scaling group. All EC2 instances are in the same security group. Your company has created a new custom application which connects to mobile devices using a custom port. This application has been rolled out to production and you need to open this port globally to the internet. What steps should you take to do this, and how quickly will the change occur?
    a. Open the port on the existing security group. Your EC2 instances will be able to communicate over this port immediately.
159. Which of the following is not supported by AWS Import/Export?
    a. Export from Glacier
160. Which of the following is not a service of the security category of the AWS trusted advisor service?
    a. Vulnerability scans on existing VPCs
161. You work for a market analysis firm who are designing a new environment. They will ingest large amounts of market data via Kinesis and then analyse this data using EMR. The data is then imported into a high performance NoSQL Cassandra database which will run on EC2 and then be accessed by traders from around the world. The database volume itself will sit on 2 EBS volumes that will be grouped into a RAID 0 volume. They are expecting very high demand during peak times, with an IOPS performance level of approximately 15000. Which EBS volume should you recommend?
    a. Provisioned IOPS (PIOPS)
162. What are the different types of virtualization available on EC2?
    a. Para-Virtual (PV) and Hardware Virtual Machine (HVM)
163. Which of the following is not a valid configuration type for AWS Storage gateway.
    a. Gateway-accessed volumes. The valid types are: Gateway-cached, Gateway-Stored and Gateway-Virtual Tape Library.
164. You have started a new role as a solutions architect for an architectural firm that designs large sky scrapers in the Middle East. Your company hosts large volumes of data and has about 250TB of data on internal servers. They have decided to store this data on S3 due to the redundancy offered by it. The company currently has a telecoms line of 2Mbps connecting their head office to the internet. What method should they use to import this data on to S3 in the fastest manner possible.
    a. AWS Import/Export allows for the importation of large datasets, using external hard disks which are sent directly to amazon, therefore bypassing the internet.
165. You are designing a site for a new start up which generates cartoon images for people automatically. Customers will log on to the site, upload an image which is stored in S3. The application then passes a job to AWS SQS and a fleet of EC2 instances poll the queue to receive new processing jobs. These EC2 instances will then turn the picture

into a cartoon and will then need to store the processed job somewhere. Users will typically download the image once (immediately), and then never download the image again. What is the most commercially feasible method to store the processed images?
    a. Store the images on S3 RRS, and create a lifecycle policy to delete the image after 24 hours.

166. You are hosting a website in Ireland called test.com and you decide to have a static DR site available on S3 in the event that your primary site would go down. Your bucket name is also called testcom. What would be the S3 URL of the static website?
    a. https://testcom.s3-website-eu-west-1.amazonaws.com. A bucket that has a static webhosting enabled on it will always have the format: https://<bucket-name>.s3-website-<AWS-region>.amazonaws.com
167. Which of the following is NOT a valid SNS subscriber?
    a. SWF
168. You are appointed as your company's CSO and you want to be able to tack all changes made to your AWS environment, by all users and at all times, in all regions. What AWS service should you use to achieve this.
    a. CloudTrail
169. You have a high performance compute application and you need to minimize network latency between EC2 instances as much as possible. What can you do to achieve this?
    a. Create a placement group within an AZ and place EC2 instances within that placement group.
170. Amazon S3 buckets in all regions provide read-after-write consistency for PUTS of new objects and eventual consistency for overwrite PUTS and DELETES.
    a. True
171. Placement Groups can be created across 2 or more AZ.
    a. False
172. You can add multiple volumes to an EC2 instance and then create your own RAID 5/RAID 10/RAID 0 configurations using those volumes
    a. True
173. You are creating your own relational database on an EC2 instance and you need to maximise IOPS performance. What can you do to achieve this goal?
    a. Add multiple additional volumes with provisioned IOPS and then create a RAID 0 stripe across those volumes.
174. Which of the services below do you get root access to?
    a. EC2 and EMR
175. using SAML (Security Assertion Markup Language 2.0) you can give your federated users SSO (single sign-on) access to the AWS Management Console.
    a. True
176. You can have 1 subnet stretched across multiple AZ.
    a. False
177. When you create new subnets within a custom VPC, by default they can communicate with each other, across availability zones.

a. True

178. It is possible to transfer a reserved instance from one Availability Zone to another.
    a. **True**
179. You have an EC2 instance which needs to find out both its private IP address and its public IP address. To do this you need to:
    a. Retrieve the instance metadata from http://169.254.169.254/latest/meta-data/
180. To retrieve instance metadata or user data you will need to use the following IP address:
    a. http://169.254.169.254
181. Amazon S3 buckets in all regions provide read-after-write consistency for PUTS of new objects.
    a. True
182. AmazonS3 buckets in all regions do not provide eventual consistency for overwrite PUTS and DELETES.
    a. False
183. Amazon S3 provides:
    a. unlimited storage
184. In order to enable encryption at rest using EC2 and EBS you need to:
    a. Configure encryption when creating the EBS volume.
185. You can select a specific AZ in which to place your DynamoDB table.
    a. False
186. When creating an RDS instance you can select in which AZ to deploy your instance.
    a. True
187. Amazon's Redshift uses which block size for its columnar storage?
    a. 1MB / 1024KB
188. You run a website which hosts videos and you have two types of members, premium fee paying members and free members. All videos uploaded by both your premium members and free members are processed by a fleet of EC2 instances which will poll SQS as videos are uploaded. However you need to ensure that your premium fee paying members videos have a higher priority than your free members. How do you design SQS?
    a. Create two SQS queues, one for premium members and one for free members. Program your EC2 fleet to poll the premium queue first and if empty, to then poll your free members SQS queue. SQS does not set priorities on individual items.
189. You have uploaded a file to S3. What HTTP code would indicate that the upload was successful?
    a. HTTP 200
190. You are hosting a MySQL database on the root volume of an EC2 instance. The database is using a large amount of IOPS and you need to increase the IOPS available to it. What should you do?
    a. Add 4 additional EBS SSD volumes and create a RAID 10 using these volumes.
191. You have been asked to create VPC for your company. The VPC must support both Internet-facing web applications (i.e. they need to be publicly accessible) and internal

private applications (i.e. they are not publicly accessible and can be accessed only over VPN). The internal private applications must be inside a private subnet. Both the internet-facing and private applications must be able to leverage at least three Availability Zones for high availability. At a minimum, how many subnets must you create within your VPC to achieve this?

a. 6. You need at least 3 AZ. (1 subnet = 1 AZ)

192. You work for a cosmetic company which has their production website on AWS. The site itself is in a two-tier configuration with web servers in the front-end and database servers at the back end. The site uses using Elastic Load Balancing and Auto-Scaling. The databases maintain consistency by replicating changes to each other as and when they occur. This requires the databases to have extremely low latency. Your website needs to be highly redundant and must be designed so that if one availability zone goes offline and Auto Scaling cannot launch new instances in the remaining Availability Zones the site will not go offline. How can the current architecture be enhanced to ensure this?

a. Deploy your site in three different AZ's within the same region. Configure the Auto Scaling minimum to handle 50 percent of the peak load per zone. In this scenario if you lost an availability zone, you would still have 2 other AZs available each that is configured to handle 50% peak load per zone. You require very low latency for synchronous replication, therefore doing anything across regions would be incorrect.

193. You working in the media industry and you have created a web application where users will be able to upload photos they create to your website. This web application must be able to call the S3 API in order to be able to function. Where should you store your API credentials whilst maintaining the maximum level of security.

a. Don't save your API credentials. Instead create a role in IAM and assign this role to an EC2 instance when you first create it. Save credentials on the instance is not a secure way of storing credentials.

194. You are a systems administrator and you need to monitor the health of your production environment. You decide to do this using CloudWatch, however you notice that you cannot see the health of every important metric in the default dash board. Which of the following metrics do you need to design a custom CloudWatch metric for, when monitoring the health of your EC2 instances?

a. Memory usage

195. You are a student currently learning about the different AWS services. Your employer asks you to tell him a bit about Amazon's glacier service. Which of the following best describes the use cases for Glacier?

a. Infrequently accessed data & data archives

196. You work for a toy company that has a busy online store. As you are approaching Christmas you find that your store is getting more and more traffic. You ensure that the web tier of your store is behind an Auto Scaling group, however you notice that the web tier is frequently scaling, sometimes multiple times in an hour, only to scale back after peak usage. You need to prevent this so that Auto Scaling does not scale as rapidly, just to scale back again. What option would help you to achieve this?

    a. Modify the Auto Scaling group cool-down timers and modify the CloudWatch alarm period that triggers your Auto Scaling scale down policy.

197. You work in the genomics industry and you process large amounts of genomic data using a nightly Elastic Map Reduce (EMR) job. This job processes a single 3 TB file which is stored on S3. The EMR job runs on 3 on-demand core nodes and four on-demand task nodes. The EMR job is now taking longer than anticipated and you have been asked to advise how to reduce the completion time?
    a. You should reduce the input split size in the MapReduce job configuration and then adjust the number of simultaneous mapper tasks so that more tasks can be processed at once.

198. By definition a public subnet within a VPC is one that:
    a. In it's routing table it has at least one route that uses an IGW.

199. You have been asked to identify a service on AWS that is a durable key value store. Which of the services below meets this definition?
    a. S3

200. You are a security architect working for a large antivirus company. The production environment has recently been moved to AWS and is in a public subnet. You are able to view the production environment over HTTP however when your customers try to update their virus definition files over a custom port, that port is blocked. You log in to the console and you allow traffic in over the custom port. How long will this take to take effect?
    a. Immediately

201. You are a solutions architect working for a biotech company who is pioneering research in immunotherapy. They have developed a new cancer treatment that may be able to cure up to 94% of cancers. They store their research data on S3, however recently an intern accidentally deleted some critical files. You've been asked to prevent this from happening in the future. What options below can prevent this?
    a. Enable S3 versioning on the bucket & enable MFA (multi factor authentication) on the bucket. Enabling versioning MFA delete capability, which uses multi-factor authentication, can be used to provide an additional layer of security.

202. You run an automobile reselling company that has a popular online store on AWS. The application sits behind an Auto Scaling group and requires new instances of the Auto Scaling group to identify their public and private IP addresses. How can you achieve this?
    a. Usering a curl or get command to get the latest meta-data from http://169.254.169.254/latest/meta-data/

203. You are a solutions architect who has been asked to do some consulting for a US company that produces re-useable rocket parts. They have a new web application that needs to be built and this application must be stateless. Which three services could you use to achieve this?
    a. RDS, DynamoDB & Elasticache. Typically you would store session information either inside a database or using elasticache.

204. Your company has decided to set up a new AWS account for test and dev purposes. They already use AWS for production, but would like a new account dedicated for test and dev so as to not accidentally break the production environment. You launch an exact replica of your production environment using a cloudformation template that your company uses in production. However cloudformation fails. You use the exact same CloudFormation template in production so the failure is something to do with your new AWS account. The CloudFormation template is trying to launch 60 new EC2 instances in a single availability zone. After some research you discover that the problem is:
    a. **For all new AWS accounts there is a soft limit of 20 EC2 instances per region. You should submit the limit increase form and retry the template after your limit has been increased.**
205. You work for a famous bakery who are deploying a hybrid cloud approach. Their legacy IBM AS400 servers will remain on premise within their own datacenter however they will need to be able to communicate to the AWS environment over a site to site VPN connection. What do you need to do to establish the VPN connection?
    a. Assign a public IP address to your VPC gateway.
206. You work for a major news network in Europe. They have just released a new app which allows users to report on events as and when they happen using their mobile phone. Users are able to upload pictures from the app and then other users will be able to view these pics. Your organisation expects this app to grow very quickly, essentially doubling it's user base every month. The app uses S3 to store the media and you are expecting sudden and large increases in traffic to S3 when a major news event takes place (as people will be uploading content in huge numbers). You need to keep your storage costs to a minimum however and it does not matter if some objects are lost. Which storage media should you use to keep costs as low as possible?
    a. S3 - RRS
207. You have developed a new web application in us-west-2 that requires six Amazon Elastic Compute Cloud (EC2) instances running at all times. You have three availability zones available in that region (us-west-2a, us-west-2b, and us-west-2c). You need 100 percent fault tolerance if any single Availability Zone in us-west-2 becomes unavailable. How would you do this, each answer has 2 answers, select the answer with BOTH correct answers.
    a. us-west-2a with 3 EC2 instances, us-west-2b with 3 EC2 instances, and us-west-2c with 3 EC2 instances.
    b. us-west-2a with 6 EC2 instances, us-west-2b with 6 EC2 instances, and us-west-2c with 0 EC2 instances. Both answers mean that if you lose a single AZ you will always have 6 instances running.
208. You need to add a route to your routing table in order to allow connections to the internet from your subnet. What route should you add?
    a. Destination: 0.0.0.0/0 → Target: internet gateway. Allow all connections to the internet.
209. You work for a construction company that has their production environment in AWS. The production environment consists of 3 identical web servers that are launched from a

standard Amazon linux AMI using Auto Scaling. The web servers are launched into the same public subnet and belong to the same security group. They also sit behind the same ELB. You decide to do some test and dev and you launch a 4th EC2 instance into the same subnet and same security group. Annoyingly your 4th instance does not appear to have internet connectivity. What could be the cause of this?

a. Assign an elastic IP address to the fourth instance. An EC2 instance in a public subnet is only publicly accessible if it has a public IP address or is behind an elastic load balancer.

210. With which AWS orchestration service can you implement Chef recipes?
    a. Opsworks
211. You are a solutions architect working for a large pharmaceutical company who are involved in high performance computing to develop new drugs to treat arthritis. You are helping them to design a new application which will need to keep network traffic the lowest latency possible while leveraging very high CPU performance. They would like to place this solution onto the AWS platform and are looking for your recommendations. Which of the following do you suggest?
    a. CPU optimized EC2 instances deployed into placement groups so as to minimize latency. Placement groups cannot span different AZs.
212. You work for a automotive company which is migrating their production environment into AWS. The company has 4 separate segments, Dev, Test, UAT and Production. They require each segment to be logically isolated from each other. What VPC configuration should you recommend?
    a. A separate VPC for each segment. Then create VPN tunnels from your HQ to each VPC so the appropriate teams can each speak to their dedicated VPC. A separate VPC can complete isolate segments from each other.
213. By default how many VPCs can you have per region in your AWS account?
    a. 5
214. Which of the following is not associated with IAM service?
    a. Workspaces
215. You are designing a new application for a financial company that will utilize spot EC2 instances as and when they meet a certain price point. These EC2 instances will analyze data and write the output their analysis to the root volume. You need to store this data in a persistent form of storage so that if the spot instances are terminated by Amazon, you will not use your data. You need to choose the lowest cost service. Where should you store your data?
    a. S3. DynamoDB or Oracle RDS are significantly more expensive than S3.
216. Which of the following is not a responsibility of Amazon's under the shared responsibility model?
    a. OS level patching for EC2.
217. In regards to EC2 which of the following is not a customers responsibility under the shared responsibility model?
    a. Decommissioning and destruction of storage media.
218. Which of the following is true when writing to S3?

    a. All regions provide read-after-write consistency for PUTS of new objects in your S3 bucket and eventual consistency for overwrite PUTS and DELETES.

219. You are solutions architect working for a busy ecommerce store. Due to your organization's unique security requirements, you decide to utilize EC2 running a MySQL database, rather than using RDS. You need to architect this EC2 instance to maximise your disk I/O. Which of the following would give you the best disk performance?
    a. Add 2 additional PIOPS SSD volumes and create a RAID 0. Install MySQL to this RAID 0 partition. This would give you the best IO performance as you are using provisioned IOPS in a RAID 0, which will give you better performance than RAID5 as you are now writing parity to your partition.
220. Which of the following services do you get OS level access to?
    a. EC2 and EMR
221. You are designing an AWS solution for a new customer and they want to use their active directory credentials in order to sign into the AWS management console. What kind of authentication response is required in order for users to authenticate with the AWS security token service (STS)?
    a. SAML 2.0 (Security Assertion Markup Language 2.0)
222. You are designing a new VPC for a customer and you need to deploy your EC2 instances in a multiple availability zones. What is the minimum number of subnets that you require to achieve this objective?
    a. 2 subnets with each subnet in an independent AZ. The minimum number of subnets required here is 2.
223. You are creating a new VPC with 3 subnets in 3 separate AZs. you require instances in each subnet to be able to communicate to each other by default. What additional steps should you take in order to achieve this objective.
    a. You do not need to do anything, by default all subnets can communicate with each other using the main route table.
224. You have an EC2 instance which needs to find out both its private IP address and its public IP address using a script. Which of the below should you include in the script to discover this information.
    a. Retrieve the instance metadata from http://169.254.169.254/latest/meta-data/
225. You are an AWS architect and you require encryption at rest for additional volumes attached to your EC2 instance. What is the quickest and most efficient way to achieve this?
    a. Configure encryption when creating the EBS volume. You could use the OS to encrypt a new volume after mounting it to an EC2 instance, however the quickest and most efficient way would be to encrypt the volume when you first provision it.
226. You are designing a web application for a new social media start up and have recommended using DynamoDB for the database due to its superior performance. You need to ensure that your database has redundancy. What additional steps should you do?
    a. Nothing. In DynamoDB all data is automatically replicated across multiple AZs.
227. What block size does Redshift use when storing its data in columnar storage?

a. 1024KB

228. You are designing an application for a furniture retailer. A component of the application takes pictures of the furniture for sale and generates thumbnail images which then need to be stored persistently. The business can tolerate it if some images are lost as they can be regenerated. The thumbnails will need to be retrieved immediately when the application requests them. What is the cheapest option to do this?
    a. Using S3 RRS
229. You are designing an image sharing website that will distribute images across the world. You need maximize performance so that your end users can download frequently accessed images as fast as possible. What AWS technology should you implement?
    a. CloudFront. You can utilize CloudFront as a CDN to cache the images locally.
230. You are putting together a wordpress site for a local charity and you are using combination of Route53, ELB, EC2 and RDS. You launch your EC2 instance, download wordpress and setup the configuration files connection string so that it can communicate to RDS. When you browse to your URL however, nothing happens. Which of the following could be the cause of this.
    a. You have forgotten to open port 80/443 on your security group in which the EC2 instance is placed.
    b. Your ELB has a health check which is checking a webpage that does not exist. Therefore your EC2 instance is not in service.
    c. You have not configured an ALIAS for your A record to point to your ELB.
231. You have created a custom VPC with 3 subnets, 2 private, 1 public. You deploy 3 EC2 instances into your public subnet and attach elastic IP addresses to these instances. You then deploy an EC2 instance into your private subnet and then attempt to apply security patches to this instance, however it has no internet connectivity. What can you do to give this instance internet access?
    a. Deploy a NAT to the public subnet and then update the main route table to send traffic via the NAT to the private subnet.
232. SWF is designed to help users to do which of the following?
    a. coordinate synchronous and asynchronous tasks?
233. What service can you use to audit user access and API calls across your AWS environment?
    a. CloudTrail
234. Under the shared responsibility model for S3 which of the following is a responsibility of Amazon?
    a. Destruction of end of life magnetic disks
    b. Restricting access to the datacenter
    c. maintaining video surveillance of the data halls
235. Under the shared responsibility model for DynamoDB which of the following is a responsibility of Amazon?
    a. Patching the Xen Hypervisor
    b. Destruction of magnetic storage media on decommissioning of disks
    c. Patching the underlying DynamoDB operating system

236. You are a solutions architect working for a major European oil company. You are designing a new web application which will need to access data stored in DynamoDB. You need to do this as securely as possible, without storing any credentials on a long term basis. How would you achieve this?
    a. Use AWS IAM roles for the EC2 instances that need to make the API calls without storing credentials for a long period of time.
237. You are a solutions architect working for a large cell phone company in the US. Your CSO has engaged a third party security company to conduct a security audit of your company to make sure it is not liable to hacking. The third party security company would like to conduct a penetration test on your AWS estate. Would this be allowed by AWS?
    a. Yes, however you need to get permission from Amazon first by raising a ticket.
238. AWS help provide protection against some forms of traditional network attacks. Which of the following is not protected against by AWS?
    a. Social engineering. The AWS platform does provide protection against port scanning, IP spoofing and man in the middle attacks.
239. You are responsible for a legacy web application whose server environment is approaching end of life. You would like to migrate this application to AWS as quickly as possible, since the application environment currently has the following limitations: The VM''s single 10GB VMDK is almost full. the virtual network interface still uses the 10Mbps driver, which leaves your 100Mbps WAN connection completely underutilized. It is currently running on a highly customized Windows VM within a VMware environment. You do not have the installation media. This is a mission critical application with an RTO (recovery time objective) of 8 hours. RPO (recovery point objective) of 1 hour. How could you best migrate this application to AWS while meeting your business continuity requirements?
    a. use the EC2 VM import connector for vCenter to import the VM into EC2. Import/Export is only used to transfer large amount of data
240. A user is trying to pre-warm a blank EBS volume attached to a linux instance. Which of the below mentioned steps should be performed by the user?
    a. Pre-warming is not necessary on EBS.
241. a user has created an EBS volume of 10 GB and attached it to a running instance. The user is trying to access EBS for the first time. Which of the below mentioned options is the correct statement with respect to a first time EBS access?
    a. The volume will show a loss of the IOPS performance the first time.
242. You are running a database on an EC2 instance, with the data stored on EBS for persistence . At times throughout the day, you are seeing large variance in the response times of the database queries. Looking into the instance with the isolate command you see a lot of wait time on the disk volume that the database's data is stored on. what 2 ways can you improve the performance of the database's storage while maintaining the current persistence of the data?
    a. move the database to an EBS optimized instance
    b. use provisioned IOPS EBS

243. You have launched an EC2 instance with 4x 500GB EBS Provisioned IOPS volumes attached. The EC2 instance is BS-optimized and supports 500 Mbps throughput between EC2 and EBS. The 2 EBS volumes are configured as a single RAID 0 device, and each PIOPS volume is provisioned with 4000 IOPS (4000 16KB reads or writes) for a total of 16000 random IOPS on the instance. the EC2 instance initially delivers the expected 16000 IOPS random read and write performance. Sometime later in order to increase the total random I/O volumes to the RAID. Each volume is provisioned to 4000 IOPS like the original four for a total of 24000 IOPS on the EC2 instance. Monitoring shows that the EC2 instance CPU utilization increased from 50% to 70%, but the total random IOPS measured at the instance level does not increase at all. What is the problem and a valid solution?
    a. EBS-optimized throughput limits the total ioPs that can be utilized. Use an EBS-optimized instance that provides larger throughput. EC2 instance types have limit on max throughput and would require 8xlarge or higher instance types to provide 24000 IOPS.
244. A user has deployed an application on an EBS backed EC2 instance. For a better performance of application, it require dedicated EC2 to EBS traffic. How can the user achieve this?
    a. Launch the EC2 instance as EBS optimized with PIOPS EBS
245. A company needs to deploy virtual desktops to its customers in a virtual private cloud, leveraging existing security controls. which set of AWS services and features will meet the company's requirements?
    a. VPN connection, AWS Directory Services and Workspaces. WorkSpaces for Virtual desktops and AWS Directory Services to authenticate to an existing on-premises AD through VPN.
246. With which AWS services CloudHSM can be used?
    a. RDS and Amazon Redshift
247. An international company has deployed a multi-tier web application that relies on DynamoDB in a single region. for regulatory reasons they need disaster recovery capability in a separate region with a RTO (recovery time objective) of 2 hours and a RPO (recovery point objective) of 24 hours. They should synchronize their data on a regular basis and be able to provision the web application rapidly using CloudFormation. The objective is to minimize changes to the existing web application, control the throughput of DynamoDB used for the synchronization of data and synchronize only the modified elements. Which design would you choose to meet these requirements?
    a. Use AWS data pipeline to schedule a DynamoDB cross region copy once a day. Create a Lastupdated attribute in your DynamoDB table that would represent the timestamp of the last update and use it as a filter.
248. Your company produces customer commissioned one-of-a-kind skiing helmets combining nigh fashion with custom technical enhancements Customers can show off their Individuality on the ski slopes and have access to head-up-displays. GPS rear-view cams and any other technical innovation they wish to embed in the helmet. The current manufacturing process is data rich and complex including assessments to ensure that

the custom electronics and materials used to assemble the helmets are to the highest standards. Assessments are a mixture of human and automated assessments you need to add a new set of assessment to model the failure modes of the custom electronics using GPUs with CUD across a cluster of servers with low latency networking. What architecture would allow you to automate the existing process using a hybrid approach and ensure that the architecture can support the evolution of processes over time?

a. Use SWF to manage assessments, movement of data and meta-data (Human and automated assessments). Use an auto-scaling group of G2 instances in a placement group (low latency networking).

249. You have multiple EC2 instances running in a cluster across multiple AZs within the same region. What combination of the following should be used to ensure the highest network performance (packets per second), lowest latency, and lowest jitter?

a. Enhanced networking (provides network performance, lowest latency for multiple AZs)
b. Amazon HVM AMI
c. Amazon VPC (works only in VPC, can't enable enhanced networking if the instance is in EC2-Classic)

250. A group of researchers is studying the migration pattern of a beetle that eats and destroys gram. The researchers must process massive amounts of data and run statistics. Which one of the following options provides the high performance computing for this purpose.

a. Launch enhanced network type instances in a placement group

251. One of the challenges in managing AWS resources is to keep track of changes in the resource configuration over time. Which one of the following statements provide the best solution?

a. Use AWS Config for supported services and use an automated process via APIs for unsupported services.

252. The majority of your infrastructure is on premises and you have a small footprint on AWS. Your company has decided to roll out a new application that is heavily dependent on low latency connectivity to LDAP for authentication.Your security policy requires minimal changes to the company's existing application user management processes. What option would you implement to successfully launch this application?

a. Establish a VPN connection between your data center and AWS create a LDAP replica on AWS and configure your application to use the LDAP replica for authentication (low latency and minimal setup)

253. A company is preparing to give AWS Management Console access to developers. Company policy mandates identity federation and role-based access control. roles are currently assigned using groups in the corporate AD. What combination of the following will give developers access to the AWS console?

a. AWS Directory Service AD Connector (for Corporate Active directory)
b. AWS IAM roles

254. An enterprise customer is starting their migration to the cloud, their main reason for migrating is agility, and they want to make their internal Microsoft AD available to any

applications running on AWS. this si so internal users only have to remember one set of credentials and as a central point of user control for leavers and joiners. How could they make their AD secure, and highly available, with minimal on-premises infrastructure changes, in the most cost and time-efficient way?

a. Using VPC, they could create an extension to their data center and make use of resilient hardware IPSEC tunnels. They could then have 2 domain controller instances that are joined to their existing domain and reside within different subnets, in different AZs. Highly available with 2 AZs, secure with VPN connection and minimal changes)

255. When preparing for a compliance assessment of your system built inside of AWS. What are the 3 best practices for you to prepare for an audit?
    a. Gather evidence of your IT operational controls. Customer still needs to gather all the IT operation controls inline with their environment.
    b. Request and obtain applicable third-party audited AWS compliance reports and certifications. (Customers can request the reports and certifications produced by our third-party auditors or can request more information about AWS compliance)
    c. Request and obtain approval from AWS to perform relevant network scans and in-depth penetration tests of your system's instances and endpoints. AWS requires prior approval to be taken to perform penetration tests.
256. In the shared security model, AWS is responsible for which of the following security best practises:
    a. penetration testing
    b. threat modeling
    c. static code analysis
257. You are running a web-application on AWS consisting of the following components: an ELB, an Auto-Scaling Group of EC2 instances running on Linux/PHP/Apache and RDS with MySQL. Which security measures fall into AWS's responsibility?
    a. protect against IP spoofing or packet sniffing
258. Which of the following statements is true about achieving PCI certification on the AWS platform?
    a. Your organization owns the compliance initiatives related to anything placed on the AWS infrastructure.
    b. AWS Compliance provides assurance related to the underlying infrastructure.
259. You are working with a customer who has 10TB of archival data that they want to migrate to Glacier. The customer has a 1-Mbps connection to the internet. Which service or feature provides the fastest method of getting the data into Glacier?
    a. AWS Import/Export. Normal upload will take about 900 days as the internet max speed is capped.
260. What does an AWS region consists of?
    a. An independent collection of AWS computing resources in a defined geography.
261. AWS … allows organizations to do complex analysis on large volumes of data.
    a. EMR
262. There are more regions than edge locations?

    a. False
263. Amazon's highly scalable DNS service is called?
    a. Route53
264. What service offers object based storage?
    a. S3
265. What is the name of the service to allow users to use their social media account to gain temporary access to the AWS platform?
    a. Web Identity Federation
266. What is the API call used to obtain temporary security credentials when authenticating using web Identity Federation?
    a. AssumeRoleWithWebIdentity
267. What is the name of the AIP call to request temporary security credentials from the AWS platform when federating with active Directory?
    a. AssumeRoleWithSAML
268. When using active directory to authenticate to AWS what are the correct steps performed?
    a. The user navigates to ADFS web server
    b. The user enters in their single sign on credentials
    c. The user's web browser receives a SAML assertion from the AD server
    d. The user's browser then posts the SAML assertion to the AWS SAML endpoint for SAML and the AssumeRoleWithSAML API request is used to request temporary security credentials
    e. The user is then able to access the AWS console
269. SAML stands for Security Assertion Markup Language.
    a. True
270. The AWS sign-in endpoint for SAML is https://signin.aws.amazon.com/saml
    a. True
271. When using Web Identity Federation to allow a user to access an AW service (such as an S3 bucket) what is the correct order of steps?
    a. A user authenticates with facebook first.
    b. They are then given an ID token by facebook.
    c. An API call called AssumeRoleWithWebIdentity is then used in conjunction with the ID token.
    d. A user is then granted temporary security credentials.
272. The default region for an SDK is us-east-1
    a. True
273. Which of the following languages is NOT supported by the AWS SDK.
    a. C++
274. A HTTP 5xx code means …
275. There has been a server side error.
276. A HTTP 4xx code means …
    a. There has been a client side error.
277. A HTTP 3xx code means …

a. There has been a redirection.

278. A HTTP 200 code means …
    a. The request was successful.
279. You can have multiple SSL certificates on an ELB.
    a. True
280. Which AWS service below is chargeable?
    a. ELB
281. The minimum file size allowed on S3 is 1 Bytes.
    a. True
282. If you encrypt a bucket on S3 what encryption does AWS use?
    a. AES 256 (advanced encryption standard)
283. You create a static hosting website in a bucket called “bucket name” in Japan using S3. What would the new URL endpoint be?
    a. http://<bucket name>.s3-website-ap-northeast-1.amazonaws.com
284. You are hosting a static website in an S3 bucket which uses JavaScript to reference assets in another S3 bucket. For some reason however these assets are not displaying when users browse to the site. What could be the problem?
    a. You haven’t enabled Cross Origin Resource Sharing (CORS) on the bucket where the assets are stored.
285. What is the HTTP code you would see if once you successfully place a file in an S3 bucket?
    a. 200
286. S3 provides unlimited storage.
    a. True
287. What is the maximum file size that can be stored on S3?
    a. 5TB
288. What is the largest size file you can transfer to S3 using a PUT operation?
    a. 5GB
289. If you want to enable a user to download your private data directly from S3, you can insert a pre-signed URL into a web page before giving it to your user.
    a. True
290. When you first create an S3 bucket, this bucket is publicly accessible by default.
    a. False
291. DynamoDB is a No-SQL database provided by AWS.
    a. True
292. You have a motion sensor which writes 600 items of data every minute. Each item consists of 5kb. Your application uses eventually consistent reads. What should you set the read throughput to?
    a. 10
293. A scan is more efficient than a query in terms of performance.
    a. False
294. What does the error “ProvisionedThroughputExceededException” mean in DynamoDB?

    a. You exceeded your maximum allowed provisioned throughput for a table or for one or more global secondary indexes.

295. You have a motion sensor which writes 600 items of data every minute. Each item consists of 5KB. What should you set the write throughput to?
    a. 50

296. What is the API call to retrieve multiple items from a DynamoDB table?
    a. BatchGetItem

297. You have a motion sensor which writes 600 items of data every minute. Each item consists of 5KB. Your application uses strongly consistent reads. What should you set the read throughput to?
    a. 20

298. Using the AWS portal, you are trying to scale DynamoDB past its pre-configured maximums. Which service can you increase by raising a ticket to AWS support?
    a. provisioned throughput limits

299. You have an application that needs to read 25 items of 13 KB in size per second. Your application uses eventually consistent reads. What should you set the read throughput to?
    a. 50

300. You have an application that needs to read 25 items of 13KB in size per second. Your application uses eventually consistent reads. What should you set the read throughput to?
    a. 100

301. SQS was the first service on the AWS platform?
    a. True

302. How large can an SQS message be?
    a. 256KB

303. What is the default visibility timeout setting?
    a. 30 seconds

304. An SQS message can be delivered multiple times.
    a. True

305. You are designing a new application which involves processing payments and delivering promotional emails to customers. You plan to use SQS to help facilitate this. You need to ensure that the payment process takes priority over the creation and delivery of emails. What is the best way to achieve this.
    a. Use 2 SQS queues for the platform. Have the EC2 fleet poll the payment SQS queue first. If this queue is empty, then poll the promotional emails queue.

306. Your EC2 instances download jobs from the SQS queue, however they are taking too long to process them. What API call can you use to extend the length of time to process the jobs?
    a. ChangeMessageVisibility

307. You have a fleet of EC2 instances that are constantly polling empty SQS queues which is burning CPU compute cycles and costing your company money. What should you do?

a. Enables SQS Long Polling

308. What is the maximum long poll timeout?
    a. 20 seconds
309. What amazon service can you use in conjunction with SQS to fan out SQS messages to multiple queues.
    a. SNS
310. SNS is pull based rather than push based?
    a. False
311. Which of these is a protocol not supported by SNS?
    a. FTP
312. Messages cannot be customised for each protocol used in SNS?
    a. False
313. You have a list of subscribers email addresses that you need to push emails out to on a periodic basis. What do you subscribe them to?
    a. Topic
314. You can use SNS in conjunction with SQS to fan a single message out to multiple SQS queues.
    a. True
315. SWF consists of a domain, workers and deciders.
    a. True
316. Maintaining your application's execution state (e.g. which steps have completed, which ones are running, etc.) is a perfect use case for SWF.
    a. True
317. Amazon SWF is useful for automating workflows that include long-running human tasks (e.g. approvals, reviews, investigations, etc.) Amazon SWF reliably tracks the status of processing steps that run up to several days or months.
    a. True
318. In amazon SWF what is a worker?
    a. Workers are programs that interact with SWF to get tasks, process received tasks, and return the results.
319. In amazon SWF what is a decider?
    a. The decider is a program that controls the coordination of tasks, i.e. their ordering, concurrency, and scheduling according to the application logic.
320. The default scripting language for CloudFormation is:
    a. JSON
321. CloudFormation itself is free, however the resources it provisions will be charged at the usual rates.
    a. True
322. What happens if CloudFormation encounters an error by default?
    a. It will terminate and rollback all resources created on failure.
323. You are creating a virtual data centre using CloudFormation and you need to output the DNS name of your load balancer. What command would you use to achieve this?
    a. FN::GetAtt

324. What languages and development stacks are supported by AWS Elastic Beanstalk?
    a. Apache Tomcat for Java applications
    b. Apache HTTP server for PHP applications
    c. Apache HTTP server for Python applications
    d. Nginx or Apache HTTP Server for Node.js applications
    e. Passenger for Ruby applications
325. Unlike CloudFormation, Elastic Beanstalk itself is not free and you must also pay for the resources it provisions.
    a. False
326. Select the correct statements:
    a. In VPC, an instance retains its private IP
    b. It is possible to have private subnets in VPC
    c. A network ACL can be assigned to multiple subnets
    d. You may only have 1 IGW per VPC
327. When you launch an Amazon Elastic Compute Cloud (EC2) instance, what does the instance type define?
    a. the amount of compute for an EC2 instance, including RAM, storage, CPU and network bandwidth.
328. When you launch an EC2 instance, what is defined by the AMI?
    a. The initial software state of the instance when launched, including operating system, configuration, and additional installed programs.
329. What is the drawback of spot instances?
    a. They can be terminated when the spot price goes above your current bid price.
330. How does the EC2 reserved instance pricing model work?
    a. You pay upfront to reserve compute for one or three years, locking in a lower cost in the process.
331. What is the EC2 on-demand pricing model?
    a. You launch EC2 instances on request and pay the full hourly cost until you stop or terminate the instances. This is the most flexible and least cost-effective pricing model.
332. How can you address an EC2 instance to connect over the internet?
    a. By public IP address, elastic IP address or DNS name.
333. What is enhanced networking for EC2?
    a. A setting to get higher packet per second, lower network jitter, and lower latencies.
334. What type of block storage is provided at no additional charge with certain EC2 instance types?
    a. instance store (or ephemeral storage)
335. VM Import/Export allows you to import existing virtual machines from your local environment and convert them to what?
    a. EC2 instances or AMIs
336. Which storage option continues to store data despite stopping and starting an instance?

    a. EBS
337. What are the 4 network capacity ratings for EC2 instance types?
    a. low
    b. moderate
    c. high
    d. 10Gbps
338. What is instance metadata?
    a. Data about an EC2 instance - such as instance ID, instance type, and security groups - that can be obtained via an HTTP call from within the instance.
339. What are the 4 properties of a security group rule?
    a. Traffic direction, port, protocol, and destination (or source) address
340. When an EC2 instance is a member of 2 security groups, what resulting traffic flow is allowed?
    a. The rules from each security group are aggregated to create one set of permissive rules, so the result is a union of all traffic allowed by the rules in both security groups.
341. Which EBS volume type is the best choice for workloads such as large databases executing many transactions?
    a. Provisioned IOPS SSD
342. What EBS volume type is appropriate for cold and infrequently accessed data?
    a. magnetic volumes
343. Which EBS volume type is appropriate for dev/test environments, small databases, and so forth?
    a. General-purpose SSD
344. What must an application running on an EC2 instance do differently to access data on an encrypted EBS volume?
    a. Nothing. EBS encryption is transparent to applications on the attached instances.
345. What is an EBS-optimized instance?
    a. An instance that has additional, dedicated capacity for EBS I/O
346. What are EBS snapshots?
    a. Point-in-time backups of an EBS volume stored in S3.
347. What are the three principals that can authenticate and interact with AWS resources?
    a. root user, IAM users, and roles
348. How can applications running on EC2 instances access the AWS API without storing an access key on the instance?
    a. By associating the instance with an EC2 role (instance profile) so that SDK applications running on the instance automatically acquire a temporary security token to access API calls.
349. What is a best practice to increase the security of an AWS account root user?
    a. Use Multi-factor authentication (MFA) to protect against a password getting compromised by also requiring the possession of a device with a rotating One-Time Password (OTP).

350. What is defined in a permission for an IAM policy?
    a. Effect, service, action, and resource. The policy may also include one or more conditions.
351. What are the 3 services of Amazon Kinesis?
    a. Kinesis Firehose
    b. Kinesis Analytics
    c. Kinesis Streams
352. What analytics service is appropriate for big data already stored on AWS?
    a. EMR, Amazon's managed Hadoop service.
353. What is the difference between a transient EMR cluster and a persistent EMR cluster?
    a. A transient cluster is shut down between analysis jobs, whereas a persistent cluster runs continuously. Data on the HDFS storage is lost when a transient cluster is shut down.
354. Which service is designed to process and move data reliably between different AWS compute and storage services for tasks such as ETL?
    a. AWS Data Pipeline
355. What service provides customers with the ability to load very large (hundreds of TB) datasets onto AWS?
    a. Import/Export provides multiple options to ship storage devices to be loaded directly into AWS.
356. Which type of AWS storage gateway volume stores all data locally while replicating it to S3?
    a. Gateway-stored volumes
357. Which type of AWS storage gateway volume stores all data in S3 and caches frequently used files locally?
    a. Gateway-cached volumes
358. Amazon CloudFront can provide CDN functionality for what type of origins?
    a. Essentially any web resource, including EC2 instances, ELB, S3, and on-premises applications and sites.
359. Does Amazon CloudFront support the accelerated delivery of static content, dynamic content, or both?
    a. Both.
360. How can you use CloudFront to serve private content?
    a. By using signed URLs, signed cookies, and S3 Origin Access identifiers.
361. How does CloudFront accelerate delivery of content?
    a. By caching content at edge locations closer to the requestor to reduce latency.
362. What is the main difference between S3 and EBS?
    a. S3 is object storage, whereas EBS is block storage.
363. What methods are available to protect your data from accidental loss on S3?
    a. Enable versioning, enable MFA Delete, use ACLs, use S3 bucket policies, and use IAM policies.

364. How can you ensure the maximum performance for high-rate GET, PUT, and DELETE requests on S3?
    a. Add a random prefix to your object names.
365. What are the 4 ways to encrypt data at rest on S3?
    a. Server side encryption (SSE) with AWS-managed keys
    b. SSE with AWS KMS-managed keys
    c. SSE with customer-provided keys
    d. client-side encryption
366. Which S3 operations have read-after-write consistency?
    a. Initial PUT requests. All other requests, including overwrite PUT requests, are eventual consistency.
367. Which S3 storage model has the lowest cost?
    a. Glacier
368. Which S3 storage model trades durability for a lower cost?
    a. S3 RRS
369. What are the 3 mechanisms to control access to objects in an S3 bucket?
    a. ACLs
    b. IAM policies
    c. S3 bucket policies
370. What type of upload is automatically used by the AWS CLI for uploading very large >5GB objects?
    a. Multipart upload
371. What are Amazon S3 lifecycle configuration rules?
    a. A mechanism for controlling objects that have a well-defined lifecycle by moving them between storage classes or deleting them at specific time intervals.
372. How does AWS verify its controls and processes for customers?
    a. Through reports, certifications, and third-party attestations, including multiple SOC and ISO certifications, FISMA, and ITAR
373. How can you be alerted if your EC2 instances have CPU utilization that is too high?
    a. Set up a CloudWatch alarm on the CPU utilization that sends a SNS message when the desired limit is exceeded. Subscribe to that SNS message with an email address or SMS text number.
374. What are 2 methods for setting CloudWatch alarms for application-specific metrics on an EC2 instance?
    a. publish the application metrics to CloudWatch with the CloudWatch Logs agent
    b. publish custom metrics
375. How can you ensure that your EC2 instances do not share a host with any other customer's EC2 instances?
    a. By specifying that the instances use the dedicated instances or dedicated hosts tenancy options
376. What are placement groups?
    a. A logical grouping of EC2 instances within a single AZ that enables applications to participate in a low-latency, 10 Gbps network

377. What does it mean to design a highly available architecture on AWS?
    a. A system is highly available when it can withstand the failure of individual or multiple components. If you design architectures around the assumption that any component will eventually fail, then systems won't fail when an individual component does.
378. What is vertical scaling?
    a. Vertical scaling takes place through an increase in the specifications of an individual resource, such as upgrading a server with a larger hard drive, more memory, or a faster CPU.
379. What is horizontal scaling?
    a. Horizontal scaling takes place through an increase in the number of resources. For example, Auto Scaling is a feature of EC2 that simplifies horizontally scaling a set of EC2 resources.
380. What does it mean to have a stateless application?
    a. A stateless application needs no knowledge of previous interactions and stores no session information.
381. What does it mean to have an elastic system?
    a. Elastic architectures can support growth in users, traffic, or data size with no drop in performance through linear scaling on top of a scalable architecture.
382. What is AWS ELB?
    a. A web service that improves an application's availability by distributing incoming traffic between 2 or more EC2 instances.
383. What is CloudWatch?
    a. A web service that acts as a metrics repository for AWS Cloud services. AWS CLoud services send metric data to CloudWatch, and you retrieve statistics based on those metrics.
384. What are the 2 types of monitoring offered by CloudWatch?
    a. Basic monitoring and detailed monitoring. Basic monitoring collects metrics at 5-min intervals, and metrics are stored for 2 weeks
    b. Detailed monitoring collects metrics at 1-minute intervals, and metrics are stored for 2 weeks.
385. What is a relational database?
    a. A database whose organization is based on the relational model of data. Communication to and from relational databases usually involves simple SQL queries.
386. What is a NoSQL database?
    a. A term used to describe high-performance, non-relational databases. NoSQL databases use a variety of data models, including document, graph, key/value, and columnar.
387. What is a data warehouse?
    a. A central repository for data that can come from one or more data sources. This data repository is typically used for complex queries and analysis for management decisions about the business.

388. How do you make an RDS instance highly available?
    a. By selecting a Multi-AZ deployment.
389. What are Recovery Point Objective (RPO) and Recovery Time Objective (RTO)?
    a. Commonly used in disaster recovery strategy, RPO is the amount of data loss measured in time, and RTO is the amount of time needed to restore a business process to its service level.
390. What is an Amazon VPC?
    a. A logically isolated network in the AWS cloud.
391. What does a route table do within an VPC?
    a. A route table is a set of rules used to determine where network traffic is directed.
392. What is the difference between a security group and a network ACL within VPC?
    a. A security group is a stateful firewall enforced at the EC2 instance layer, whereas a network ACL is a stateless firewall enforced at the subnet layer.
393. What is an VPC subnet?
    a. A segment of VPC's IP address range where you can place isolated resources.
394. What is VPC peering?
    a. A networking connection between 2 VPCs that enables instances in either VPC to communicate with each other as if they were within the same network. Peering is available only between VPCs in the same region.
395. What is SQS?
    a. A web service that offers reliable and scalable hosted queues for storing messages as they travel between computers.
396. What is SWF?
    a. A fully managed service that helps developers build, run, and scale background jobs that have parallel or sequential steps.
397. What is SNS?
    a. A fully managed web service that enables applications, end users, and devices to instantly send and receive notifications from the cloud.
398. Name the types of endpoints that can be subscribed to an SNS topic.
    a. AWS Lambda function, SQS queue, HTTP endpoint, HTTPS endpoint, Email, Email-JSON, and SMS
399. What is a SQS visibility timeout?
    a. A period of time during which SQS prevents other components from receiving and processing a message because another component is already processing it.
400. What is AWS KMS?
    a. A managed service that makes it easy for you to create and control the encryption keys used to encrypt your data.
401. What is the value of AWS CloudHSM?
    a. The AWS CloudHSM service helps you meet corporate, contractual, and regulatory compliance requirements for data security by using dedicated HSM appliances within the AWS Cloud. With AWS CloudHSM, you control the encryption keys and cryptographic operations performed by the HSM.
402. What is CloudTrail?

a. A web service that records AWS API calls for your account and delivers log files to you.

403. What is AWS Directory Service?
    a. A managed service for connecting your AWS resources to an existing on-premises Microsoft Active Directory or to set up and operate a new, standalone directory in the AWS Cloud
404. What are the 3 types of directories offered by the AWS Directory Service?
    a. AWS Directory Service for Microsoft AD (Enterprise edition), Simple AD, and AD connector
405. What is the AWS shared responsibility model?
    a. AWS is responsible for securing the underlying infrastructure that supports the cloud, and you are responsible for securing anything you put on the cloud or connect to the cloud.
406. What is an AWS region?
    a. A named set of AWS resources in the same geographical area. A region comprises at least 2 AZs.
407. What is an AZ?
    a. A distinct location within a region that is insulated from failures in other AZs and provides inexpensive, low-latency network connectivity to other availability zones in the same region.
408. What is high-availability system design using AWS?
    a. High-availability system design is based on an architecture that takes advantage of multiple availability zones and regions. Distributing applications across multiple AZs provides the ability to remain resilient in the face of most failure modes, including natural disasters or system failures.
409. What are the 3 types of credentials available for use within AWS?
    a. passwords
    b. MFA devices
    c. access keys
410. Which AWS service endpoints do not support HTTPS?
    a. None. All AWS service endpoints support HTTPS.
411. What is AWS OpsWorks?
    a. A configuration management service that helps you configure and operate applications of all shapes and sizes using Chef.
412. What is CloudFormation?
    a. A service that helps you model and set up your AWS resources based on a JSON template. It allows organizations to deploy, modify, and update resources in a controlled and predictable way.
413. What is AWS Elastic Beanstalk?
    a. A web service for deploying and managing applications in the AWS Cloud without worrying about the infrastructure that runs those applications.
414. What is AWS Config?

    a. A fully managed service that provides an AWS resource inventory, configuration history, and configuration change notifications for better security and governance.
415. What is AWS Trusted Advisor?
    a. A service that inspects your environment and makes recommendations when opportunities exist to save money, improve system availability and performance, or help close security gaps.
416. What is SQS long polling?
    a. Long polling allows the SQS service to wait until a message is available in the queue before sending a response.
417. What is the first architecture best practice of AWS?
    a. Design for failure and nothing will fail.
418. Name 2 types of redundancy.
    a. Active and standby. Active redundancy means the load is evenly distributed across multiple resources and when one fails, the others absorb a larger workload.
    b. When a resource fails with standby redundancy, functionality is recovered on a secondary resource using a process called failover.
419. What is elasticity?
    a. The ability for the system to grow based on increased demand and contract based on decreased demand, whether gradually over time or in response to a sudden change.
420. What is defense in depth?
    a. A security paradigm where security is implemented at multiple layers within a system. These layers include the physical layer, network layer, system layer, and data layer.
421. How many types of ELBs are offered within AWS?
    a. Internet-facing load balancers
    b. Internal load balancers
422. What does connection draining do on an ELB?
    a. it enables the load balancer to complete in-flight requests to instances that are de-registered or unhealthy.
423. What is the benefit of enabling sticky sessions on an ELB?
    a. Enabling sticky sessions on an ELB ensures that all requests from the user during the session are sent to the same instance behind the ELB.
424. Name the 4 different types of Auto Scaling plans.
    a. Manual scaling
    b. Maintain current instance level
    c. Scheduled scaling
    d. Dynamic scaling
425. In reference to relational databases, what is the difference between OLTP and OLAP?
    a. OLTP refers to transaction-oriented applications that are frequently writing and changing data (for example, data entry, e-commerce applications). OLAP is

typically the domain of data warehouses and refers to reporting on or analyzing large datasets.

426. What does loosely couple mean and why is it important?
    a. Coupling is the degree of direct knowledge that one component has of another. Loose coupling is a design approach where components have very little or no direct knowledge of each other. Loosely coupled systems can scale to a greater extent than more tightly couple systems.
427. Under a single AWS account, you have set up an Auto Scaling group with a miaxum capacity of 50 EC2 instances in us-west-2. When you scale out, however, it only increases to 20 EC2 instances. What is the likely cause?
    a. You have exceed the default EC2 instance limit of 20 per region.
        i. Auto Scaling may cause you to reach limits of other services, such as the default number of EC2 instances you can currently launch within a region, which is 20.
428. ELB allows you to distribute traffic across which of the following?
    a. Multiple AZs **within a region**.
        i. The ELB service allows you to distribute traffic across a group of EC2 instances in one or more AZs within a region.
429. Amazon CloudWatch offers which types of monitoring plans?
    a. Basic
    b. Detailed
430. EC2 instance in a VPC subnet can send and receive traffic from the internet when which of the following conditions are met?
    a. Attach an IGW to the VPC and create a subnet route table to send all non-local traffic to that IGW.
    b. Network ACLs and security group rules allow relevant internet traffic.
    c. EC2 instance has a public IP address or EIP address.
431. If you launch 5 EC2 instances in a VPC without specifying a security group, the instances will be launched into a default security group that provides which of the following?
    a. The 5 EC2 instances can communiate with each other.
    b. No inbound traffic will be allowed to the 5 EC2 instances.
    c. All outbound traffic will be allowed from the 5 EC2 instances.
    d. If a security group is not specified at launch, then an EC2 instance will be launched into the default security group for the VPC. The default security group allows communication between all resources within the security group, allows all outbound traffic, and denies all other traffic.
432. Your company wants to host its secure web application in AWS. The internal security policies consider any connections to or from the web server as insecure and require application data protection. What approaches should you use to protect data in transit for a application?
    a. Use HTTPS with server certificate authentication.
    b. Use SSL / TLS for database connection

433. You have an application that will run on an EC2 instance. The application will make requests to S3 and DynamoDB. Using best practices, what type of IAM identity should you create for your application to access the identified services?
    a. IAM role
        i. Don't create an IAM user (or an IAM group) and pass the user's credentials to the application or embed the credentials in the application. Instead, create an IAM role that you attach to the EC2 instance to give applications running on the instance temporary security credentials. The credentials have the permissions specified in the policies attached to the role. a directory is not an identity object in IAM.
434. When a request is made to an AWS service, the request is evaluated to decide whether it should be allowed or denied. the evaluation logic follows which of the following rules?
    a. When a request is made, the AWS service decides whether a given request should be allowed or denied. The evaluation logic follows these rules:
        i. By default, all requests are denied.
        ii. An explicit allow overrides the default.
        iii. An explicit deny overrides any allows.
435. What is the data processing engine behind EMR?
    a. Apache Hadoop
        i. EMR uses Apache Hadoop as its distributed data processing engine. Hadoop is an open source. Java software framework that supports data-intensive distributed applications running on large clusters of commodity hardware. Hive, Pig, and HBase are packages that run on top of Hadoop.
436. What type of Elastic Beanstalk environment tier provisions resources to support a web application that handles background processing task?
    a. Worker environment tier
        i. An environment tier whose web application runs background jobs is known as a worker tier. an environment tier whose web application processes web requests is known as a web server tier. Database and batch are not valid environment tiers.
437. What RDS feature provides the high availability for your database?
    a. Multi-AZ deployment
        i. Multi-AZ deployment uses synchronous replication to a different AZ so that operations can continue on the replica if the master database stops responding for any reason. Automated backups provide disaster recovery, not high availability. Security groups, while important, have no effect on availability. Maintenance windows are actually times when the database may not be available.
438. What administrative tasks are handled by AWS for RDS databases?
    a. Regular backups of the database
    b. Deploying virtual infrastructure

c. Patching the operating system and database software
    i. Amazon RDS will launch EC2 instances, install the database software, handle all patching, and perform regular backups. Anything within the database software (schema, user accounts, and so on) is the responsibility of the customer.

439. Which of the following use cases is well suited for Redshift?
    a. A 500 TB data warehouse used for market analytics
        i. Redshift is a petabytes-scale data warehouse. It is not well suited for unstructured NoSQL data or highly dynamic transactional data. It is not a cache.

440. Which of the following statements about DynamoDB secondary indexes is true?
    a. There can only be one per table, and it must be created when the table is created.

441. What is the primary use case of Kinesis Firehose?
    a. Ingest huge streams of data and store it to S3, Redshift, or Elasticsearch Service.
        i. Kinesis family of services provides functionality to ingest large streams of data. Kinesis Firehose is specifically designed to ingest a stream and save it to any of the three storage services listed in S3, redshift or Cloud Search.

442. Your company has 17TB of financial trading records that need to be stored for 7 years by law. Experience has shown that any record more than a year old is unlikely to be accessed. Which of the following storage plans meets these needs in the most cost-efficient manner?
    a. Store the data on S3 with lifecycle policies that change the storage class to Glacier after 1 year, and delete the object after 7 years.
        i. S3 and Glacier are the most cost-effective storage services. After a year, when the objects are unlikely to be accessed, you can save costs by transferring the objects to Glacier where the retrieval time is 3 to 5 hours.

443. What must you do to create a record of who accessed your S3 data and from where?
    a. Enable server access logs on the bucket.
        i. Server access logs provide a record of any access to an object in S3.

444. S3 is an eventually consistent storage system. For what kinds of operations is it possible to get stale data as a result of eventual consistency?
    a. GET after overwrite PUT (PUT to an existing key)
        i. S3 provides read-after-write consistency for PUTs to new objects (new key), but eventual consistency for GETs and DELETEs of existing objects (existing key). Response C changes the existing object so that a subsequent GET may fetch the previous and inconsistent object.

445. How is data stored in S3 for high durability?
    a. Data is automatically replicated to different AZs within a region.
        i. AWS will never transfer data between regions unless directed to by you. Durability in S3 is achieved by replicating your data geographically to

different AZs regardless of the versioning configuration. AWS doesn't use tapes.

446. Your company needs to provide streaming access to videos to authenticated users around the world. What is a good way to accomplish this?
    a. Enable CloudFront with geolocation and signed URLs
        i. CloudFront provides the best user experience by delivering the data from a geographically advantageous edge location. signed URLs allow you to control access to authenticated users.
447. Which of the following are true about the AWS shared responsibility model?
    a. AWS is responsible for all infrastructure components (that is, AWS Cloud services) that support customer deployments.
    b. The customer is responsible for the components from the guest operating system upward (including updates, security patches, and antivirus software).
    c. While AWS manages security of the cloud, security in the cloud is the responsibility of the customer.
        i. In the AWS shared responsibility model, customers retain control of what security they choose to implement to protect their own content, platform, applications, systems, and networks, no differently than they would for applications in an on-site data center.
448. Which process in an SWF workflow implements a task?
    a. Activity worker
        i. An activity worker is a process or thread that performs the activity tasks that are part of your workflow. Each activity worker polls SWF for new tasks that are appropriate for that activity worker to perform; certain tasks can be performed only by certain activity workers. After receiving a task, the activity worker processes the task to completion and then reports to SWF that the task was completed and provides the result. The activity task represents one of the tasks that you identified in your application.
449. Which of the following is true if you stop an EC2 instance with an EIP address in a VPC?
    a. The instance remains associated with its Elastic IP address.
        i. In a VPC, an instance's EIP address remains associated with an instance when the instance is stopped.
450. Which EC2 pricing model allows you to pay a set hourly price for compute, giving you full control over when the instance launches and terminates?
    a. On Demand instances
        i. You pay a set hourly price for an On Demand instance from when you launch it until you explicitly stop or terminate it.
        ii. Spot instances can be terminated when the spot price goes above your bid price.
        iii. Reserved instances involve paying for an instance over a one- or three-year term.

iv. Dedicated instances run on hardware dedicated to your account and are not a pricing model.

451. Under what circumstances will EC2 instance store data not be preserved?
    a. The instance is stopped or terminated.
        i. The data in an instance store persists only during the lifetime of its associated instance. If an instance is stopped or terminated, then the instance store does not persist. Rebooting an instance does not shutdown the instance; if an instance reboots (intentionally or unintentionally), data on the instance store persists. Security groups have nothing to do with the lifetime of an instance and have no effect here.

452. Which of the following describes a physical location around the world where AWS clusters data centers?
    a. Region
        i. A region is a named set of AWS resources in the same geographical area. A region comprises at least 2 AZs. Endpoint, Collection, and Fleet do not describe a physical location around the world where AWS clusters data centers.

453. Each AWS region is composed of 2 or more locations that offer organizations the ability to operate production systems that are more highly available, fault tolerant, and scalable than would be possible using a single data center. What are these locations called?
    a. Availability Zones
        i. An AZ is a distinct location within a region that is insulated from failures in other availability zones and provides inexpensive, low-latency network connectivity to other AZs in the same region. Replication areas, geographic districts, and compute centers are not terms used to describe AWS data center locations.

454. What is the deployment term for an environment that extends an existing on-premises infrastructure into the cloud to connect cloud resources to internal systems?
    a. Hybrid deployment
        i. A hybrid deployment is a way to connect infrastructure and applications between cloud-based resources and existing resources that are not located in the cloud.
        ii. An all-in deployment refers to an environment that exclusively runs in the cloud.
        iii. An on-premises deployment refers to an environment that runs exclusively in an organization’s data center.

455. Which AWS Cloud service allows organizations to gain system-wide visibility into resource utilization, application performance, and operational health?
    a. CloudWatch
        i. Amazon CloudWatch is a monitoring service for AWS Cloud resources and the applications organizations run on AWS. It allows organizations to collect and track metrics, collect and monitor log files, and set alarms.

ii. AWS IAM, SNS, and cloudFormation do not provide visibility into resource utilization, application performance, and the operational health of your AWS resources.

456. Which of the following AWS Cloud services is a fully managed NoSQL database service?
    a. DynamoDB
        i. DynamoDB is a fully managed, fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale.
        ii. SQS, ElastiCache, and RDS do not provide a NoSQL database service.
        iii. SQS is a managed message queuing service.
        iv. ElastiCache is a service that provides in-memory cache in the cloud.
        v. RDS provides managed relational databases.

457. Your company experiences fluctuations in traffic patterns to their e-commerce website based on flash sales. What service can help your company dynamically match the required compute capacity to the spike in traffic during flash sales?
    a. Auto Scaling
        i. Auto Scaling helps maintain application availability and allows organizations to scale EC2 capacity up or down automatically according to conditions defined for the particular workload. Not only can it be used to help ensure that the desired number of EC2 instances are running, but it also allows resources to scale in and out to match the demands of dynamic workloads.
        ii. Glacier, SNS, and VPC do not provide services to scale compute capacity automatically.

458. Your company provides an online photo sharing service. The development team is looking for ways to deliver image files with the lowest latency to end users so the website content is delivered with the best possible performance. What service can help speed up distribution of these image files to end users around the world?
    a. CloudFront
        i. CloudFront is a web-service that provides a CDN to speed up distribution of your static and dynamic web content -- for example, html, css, php, image, and media files -- to end users. CloudFront delivers content through a worldwide network of edge locations.
        ii. EC2, Route53, and Storage Gateway do not provide CDN services that are required to meet the needs for the photo sharing service.

459. Your company runs an EC2 instance periodically to perform a batch processing job on a large and growing filesystem. At the end of the batch job, you shut down the EC2 instance to save money but need to persist the filesystem on the EC2 instance from the previous batch runs. What AWS Cloud service can you leverage to meet theses requirements?
    a. EBS

i. Elastic Block Service provides persistent block-level storage volumes for use with EC2 instances on the AWS Cloud.
ii. DynamoDB, Glacier, CloudFormation do not provide persistent block-level storage for EC2 instances.
iii. DynamoDB provides managed NoSQL databases.
iv. Glacier provides low-cost archival storage.
v. CloudFormation gives developers and system administrators an easy way to create and manage a collection of related AWS resources.

460. What AWS Cloud service provides a logically isolated section of the AWS Cloud where organizations can launch AWS resources in a virtual network that they define?
    a. VPC
        i. Virtual Private Cloud lets organizations provision a logically isolated section of the AWS Cloud where they can launch AWS resources in a virtual network that they define.
        ii. SWF, Route 53, and CloudFormation do not provide a virtual network.
        iii. SWF helps developers build, run, and scale background jobs that have parallel or sequential steps.
        iv. Route53 provides a highly available and scalable cloud Domain Name System (DNS) web service.
        v. CloudFormation gives developers and systems administrators an easy way to create and manage a collection of related AWS resources.

461. Your company provides a mobile voting application for a popular TV show, and 5 to 25 million viewers all vote in a 15 second timespan. What mechanism can you use to decouple the voting application from your backend services that tally the votes?
    a. SQS
        i. Simple Queue Service is a fast, reliable, scalable, fully managed message queuing service that allows organizations to decouple the components of a cloud application. With SQS, organizations can transmit any volume of data, at any level of throughput, without losing messages or requiring other services to be always available.
        ii. CloudTrail records AWS API calls.Redshift is a data warehouse, neither of which would be useful as an architecture component for decoupling components.
        iii. SNS provides a messaging bus complement to SQS, however it doesn't provide the decoupling of components necessary for this scenario.

462. In what ways does S3 object storage differ from block and file storage?
    a. Objects are stored in buckets.
    b. Objects contain both data and metadata.

463. Which of the following are not appropriates use cases for S3?
    a. Storing a file system mounted to an EC2 instance
    b. Primary storage for a database.
        i. S3 cannot be mounted to an EC2 instance like a file system and should not serve as primary database storage

464. What are some of the key characteristics of S3?
    a. All objects have a URL
    b. S3 can store unlimited amounts of data
    c. S3 uses a REST application program interface (API)
        i. Objects are private by default, and storage in a bucket does not need to be pre-allocated.
465. Which features can be used to restrict access to S3 data?
    a. Create a pre-signed URL for an object
    b. Use a S3 ACL on a bucket or object
    c. Use an S3 bucket policy
        i. Static website hosting does not restrict data access, and neither does a S3 lifecycle policy
466. Your application stores critical data in S3, which must be protected against inadvertent or intentional deletion. How can this data be protected?
    a. Enable versioning on the bucket
    b. Enable MFA Delete on the bucket
        i. Versioning protects data against inadvertent or intentional deletion by storing all versions of the object, and MFA Delete requires a one-time code from a Multi-Factor Authentication device to delete objects.
        ii. Cross-region replication and migration to the Glacier storage class do not protect against deletion. Vault locks are a feature of Glacier, not a feature of S3.
467. Your company stores documents in S3, but it wants to minimize cost. Most documents are used actively for only about a month, then much less frequently. However, all data needs to be available within minutes when requested. How can you meet these requirements?
    a. Migrate the data to S3 Standard-Infrequent Access (IA) after 30 days.
        i. Migrating the data to S3 IA after 30 days using a lifecycle policy is correct.
        ii. S3 RRS should only be used for easily replicated data, not critical data.
        iii. Migration to Glacier might minimize storage costs if retrievals are infrequent, but documents would not be available in minutes when needed.
468. How is data stored in S3 for high durability?
    a. Data is automatically replicated within a region.
        i. Replication to other regions and versioning are optional.
        ii. S3 data is not backed up to tape.
469. Based on the following S3 URL, which one of the following statements is correct? https://bucket1.abc.com.s3.amazonaws.com/folderx/myfile.doc
    a. The object ‘folderx/myfile.doc’ is stored in the bucket ‘bucket1.abc.com’
        i. In a URL, the bucket name precedes the string ‘s3.amazonaws.com/’, and the object key is everything after that. There is no folder structure in S3.
470. To have a record of who accessed your S3 data and from where, you should do what?

a. Enable server access logs on the bucket.
    i. S3 server access logs store a record of what requestor accessed the objects in your bucket, including the requesting IP address.

471. What are some reasons to enable cross-region replication on an S3 bucket?
    a. You have a set of users or customers who can access the second bucket with lower latency.
    b. For compliance reasons, you need to store data in a location at least 300 miles away from the first region.
        i. Cross-region replication can help lower latency and satisfy compliance requirements on distance.
        ii. S3 is designed for 11x9 durability for objects in a single region, so a second region does not significantly increase durability.
        iii. Cross-region replication does not protect against accidental deletion.

472. Your company requires that all data sent to external storage be encrypted before being sent. Which S3 encryption solution will meet this requirement?
    a. SSE with customer-provided keys (SSE-C)
        i. If data must be encrypted before being sent to S3, client-side encryption must be used.

473. You have a popular web application that accesses data stored in a S3 bucket. You expect the access to be very read-intensive, with expected request rates of up to 500 GETs per second from many clients. How can you increase the performance and scalability of S3 in this case?
    a. Ensure randomness in the namespace by including a hash prefix to key names.
        i. S3 scales automatically, but for request rates over 100 GETs per second, it helps to make sure there is some randomness in the key space.
        ii. Replication and logging will not affect performance or scalability.
        iii. Using sequential key names could have a negative effect on performance or scalability.

474. What is needed before you can enable cross-region replication on a S3 bucket?
    a. Enable versioning on the bucket
    b. Create an AWS IAM policy to allow S3 to replicate objects on your behalf
        i. You must enable versioning before you can enable cross-region replication, and S3 must have IAM permissions to perform the replication.
        ii. Lifecycle rules migrate data from one storage class to another, not from one bucket to another.
        iii. Static website hosting is not a prerequisite for replication.

475. Your company has 100TB of financial records that need to be stored for 7 years by law. Experience has shown that any record more than 1 year old is unlikely to be accessed. Which of the following storage plans meets these needs in the most cost efficient manner?
    a. S3 is the most cost effective storage on AWS, and lifecycle policies are a simple and effective feature to address the business requirements.

476. S3 bucket policies can restrict access to a S3 bucket and objects by which of the following?
    a. IP address range
    b. AWS account
    c. Objects with a specific prefix
        i. S3 bucket policies cannot specify a company name or a country or origin, but they can specify request IP range, AWS account, and a prefix for objects that can be accessed.
477. S3 is an eventually consistent storage system. For what kinds of operations is it possible to get stale data as a result of eventual consistency?
    a. GET or LIST after a DELETE
    b. GET after overwrite PUT (Put to an existing key)
        i. S3 provides read-after-write consistency for PUTs to new objects (new key), but eventual consistency for GETs and DELETEs of existing objects (existing key)
478. What must be done to host a static website in a S3 bucket?
    a. Configure the bucket for static hosting and specify an index and error document
    b. Create a bucket with the same name as the website
    c. Make the objects in the bucket world-readable
        i. And normally you also set a friendly CNAME to the bucket URL
        ii. S3 does not support FTP transfers, and HTTP does not need to be enabled
479. You have valuable media files hosted on AWS and want them to be served only to authenticated users of your web application. You are concerned that your content could be stolen and distributed for free. How can you protect your content?
    a. Generate pre-signed URLs for content in the web application
        i. Pre-signed URLs allow you to grant time-limited permission to download objects from a S3 bucket.
        ii. Static web hosting generally requires world-read access to all content
        iii. AWS IAM policies do not know who the authenticated users of the web app are
        iv. Logging can helü track content loss, but not prevent it
480. Glacier is well-suited to data that is which of the following?
    a. Is infrequently or rarely accessed
    b. Is available after a three- to five-hour restore period
        i. Glacier is optimized for long-term archival storage and is not suited to data that needs immediate access or short-lived data that is erased within 90 days
481. Which statements about Glacier are true?
    a. Glacier archives take 3 to 5 hours to restore
    b. Glacier vaults can be locked
    c. Glacier can be used as a standalone service and as a S3 storage class
        i. Glacier stores data in archives, which are contained in vaults.

    ii. Archives are identified by system-created archive IDs, not key names.

482. Your web application needs 4 instances to support steady traffic nearly all of the time. On the last day of each month, the traffic triples. What is a cost-effective way to handle this traffic pattern?
    a. Run 4 reserved instances constantly, then add 8 on-demand instances on the last day of each month.
        i. Reserved instances provide cost savings when you can commit to running instances full time, such as to handle the base traffic.
        ii. On-Demand instances provide the flexibility to handle traffic spikes, such as on the last day of the month.

483. Your order-processing application processes orders extracted from a queue with 2 reserved instances processing 10 orders/minute. If an order fails during processing, then it is returned to the queue without penalty. Due to a weekend sale, the queues have several hundred orders backed up. While the backup is not catastrophic, you would like to drain it so that customers get their confirmation emails faster. What is a cost-effective way to drain the queue for orders?
    a. Deploy additional spot instances to assist in processing the orders.
        i. Spot instances are a very cost-effective way to address temporary compute needs that are not urgent and are tolerant of interruption. That's exactly the workload described here.
        ii. Reserved instances are inappropriate for temporary workloads.
        iii. On-Demand instances are good for temporary workloads, but don't offer the cost savings of spot instances.
        iv. Adding more queues would not address the problem.

484. Which of the following must be specified when launching a new EC2 Windows instance?
    a. EC2 instance type
    b. AMI
        i. The EC2 instance ID will be assigned by AWS as part of the launch process.
        ii. The administrator password is assigned by AWS and encrypted via the public key.
        iii. The instance type defines the virtual hardware and the AMI defines the initial software state. You must specify both upon launch.

485. You have purchased an m3.xlarge Linux reserved instance in us-east-1a. In which ways can you modify this reservation?
    a. Change it into 2 m3.large instances.
    b. Move it to us-east-1b
        i. You can change the instance type only within the same instance type family, or you can change the AZ.
        ii. You cannot change the operating system nor the instance type family.

486. Your instance is associated with 2 security groups. The first allows RDP access over port 3389 from CIDR block 72.14.0.0/16. The second allows HTTP access over port 80 from CIDR block 0.0.0.0(0. What traffic can reach your instance?
    a. RDP traffic and HTTP traffic
        i. When there are multiple security groups associated with an instance, all the rules are aggregated
487. Which of the following are features of enhanced networking?
    a. More packets per second (PPS)
    b. Lower latency
    c. Less jitter
488. You are creating a High-Performance Computing (HPC) cluster and need very low latency and high bandwidth between instances. What combination of the following will allow this?
    a. Use an instance type with 10 Gbps network performance
    b. Put the instances in a placement group
    c. Enable enhanced networking on the instances.
489. Which EC2 feature ensures that your instances will not share a physical host with instances from any other AWS customer?
    a. Dedicated instances
490. Which of the following are true of instance stores?
    a. Data is lost when the instance stops
    b. Very high IOPS
        i. Instance stores are low-durability, high-IOPS storage that is included for free with the hourly cost of an instance.
491. Which of the following are features of EBS?
    a. Data stored on EBS is automatically replicated within an AZ
    b. EBS volumes can be encrypted transparently to workloads on the attached instance
        i. There are no tapes in the AWS infrastructure
        ii. EBS volumes persist when the instance is stopped
        iii. The data is automatically replicated within an AZ.
        iv. EBS volumes can be encrypted upon creation and used by an instance in the same manner as if they were not encrypted.
492. You need to take a snapshot of an EBS volume. How long will the volume be unavailable?
    a. The volume will be available immediately.
        i. There is no delay in processing when commencing a snapshot.
493. You are restoring an EBS volume from a snapshot. How long will it be before the data is available?
    a. The data will be available immediately.
        i. The volume is created immediately but the data is loaded lazily. This means that the volume can be accessed upon creation, and if the data

being requested has not yet been restored, it will be restored upon first request.

494. You have a workload that requires 15000 consistent IOPS for data that must be durable. What combination of the following steps do you need?
    a. Use an EBS-optimized instance
    b. Use a Provisioned IOPS SSD volume
        i. An instance store will not be durable and a magnetic volume offers an average of 100 IOPS.
        ii. EBS-optimized instances reserve network bandwidth on the instance for I/O, and Provisioned IOPS SSD volumes provide the highest consistent IOPS
495. Which of the following can be accomplished through bootstrapping?
    a. Install the most current security updates
    b. Install the current version of the application
    c. Configure OS services
        i. Bootstrapping runs the provided script, so anything you can accomplish in a script you can accomplish during bootstrapping
496. How can you connect to a new Linux instance using SSH?
    a. Using the private half of the instance's key pair
        i. The public half of the key pair is stored on the instance, and the private half can then be used to connect via SSH.
497. VM Import/Export can import existing VM as:
    a. EC2 instances and AMIs
498. Which of the following can be used to address an EC2 instance over the web?
    a. Public DNS name
    b. Elastic IP address
        i. Neither the Windows machine name nor the EC2 instance ID can be resolved into an IP address to access the instance
499. Using the correctly decrypted Administrator password and RDP, you cannot log into a Windows instance you just launched. Which of the following is a possible reason?
    a. There is no security group rule that allows RDP access over port 3389 from your IP address.
500. You have a workload that requires 1TB of durable block storage at 1500 IOPS during normal use. Every night there is an Extract, Transform, Load (ETL) task that requires 3000 IOPS for 15 minutes. What is the most appropriate volume type for this workload?
    a. Use a general-purpose SSD volume
        i. A short period of heavy traffic is exactly the use case for the bursting nature of general-purpose SSD volumes. The rest of the day is more than enough time to build up enough IOPS credits to handle the nightly task.
        ii. Instance stores are not durable.
        iii. Magnetic volumes cannot provide enough IOPS
        iv. To set up a Provisioned IOPS SSD volume to handle the peak would mean spending money for more IOPS than you need.

501. How are you billed for elastic IP addresses?
    a. Hourly when they are not associated with an instance.
        i. There is a very small hourly charge for allocated elastic IP addresses that are not associated with an instance.
502. What is the minimum size subnet that you can have in an Amazon VPC?
    a. /28
        i. The minimum size subnet that you can have in an VPC is /28
503. You are a solutions architect working for a large travel company that is migrating its existing server estate to AWS. You have recommended that they use a custom VPC, and they have agreed to proceed. They will need a public subnet for their web servers and a private subnet in which to place their databases. They also require that the web servers and database servers be highly available and that there be a minimum of 2 web servers and 2 database servers each. How many subnets should you have to maintain high availability?
    a. 4
        i. You need 2 public subnets (one for each AZ) and 2 private subnets (one for each AZ)
504. Which of the following is an optional security control that can be applied at the subnet layer of a VPC?
    a. Network ACL
        i. Network ACLs are associated to a VPC subnet to control traffic flow.
505. What is the maximum size IP address range that you can have in an Amazon VPC?
    a. /16
506. You create a new subnet and then add a route to your route table that routes traffic out from that subnet to the Internet using an IGW. What type of subnet have you created?
    a. a public subnet
        i. By creating a route out to the internet using an IGW, you have this subnet public.
507. What happens when you create a new Amazon VPC?
    a. A main route table is created by default.
        i. When you create an VPC, a route table is created by default.
        ii. You must manually create subnets and an IGW.
508. You create a new VPC in us-east-1 and provision 3 subnets inside this VPC. Which of the following statements is true?
    a. All subnets will be able to communicate with each other by default.
        i. When you provision a VPC, all subnets can communicate with each other by default.
509. How many IGWs can you attach to a VPC at any one time?
    a. 1
510. What aspect of a VPC is stateful?
    a. Security groups
        i. Security groups are stateful, whereas network ACLs are stateless.

511. You have created a custom VPC with both private and public subnets. You have created a NAT instance and deployed this instance to a public subnet. You have attached an EIP address and added your NAT to the route table. Unfortunately, instances in your private subnet still cannot access the internet. What may be the cause of this?
    a. You should disable source/destination checks on the NAT
512. Which of the following will occur when an EBS-backed EC2 instance in a VPC with an associated EIP is stopped and started?
    a. All data on instance-store devices will be lost.
    b. The underlying host for the instance is changed.
        i. In the EC2-classic network, the EIP will be disassociated with the instance.
        ii. In the EC2-VPC network, the EIP remains associated with the instance.
        iii. Regardless of the underlying network, a stop/start of an EBS-backed EC2 instance always changes the host computer.
513. How many VPC Peering connections are required for four VPCs located within the same AWS region to be able to send traffic to each of the others.
    a. 6
        i. 6 Peering connections are needed for each of the 4 VPCs to send traffic to each of the others.
514. Which of the following AWS resources would you use in order for an EC2-VPC instance to resolve DNS names outside of AWS?
    a. a DHCP option set
        i. A DHCP option set allows customers to define DNS servers for DNS name resolution, establish domain names for instances within a VPC, define NTP servers, and define the NetBIOS name servers.
515. Which of the following is the Amazon side of a VPN connection?
    a. VPG
        i. CGW is the customer side of a VPN connection
        ii. IGW connects a network to the internet
        iii. VPG is the Amazon side of a VPN connection
516. What is the default limit for the number of VPCs that a customer may have in a region?
    a. 5
        i. The default limit for the number of VPCs that a customer may have in a region is 5.
517. You are responsible for your company's AWS resources, and you notice a significant amount of traffic from an IP address in a foreign country in which your company does not have customers. Further investigation of the traffic indicates the source of the traffic is scanning for open ports on your EC2-VPC instances. Which one of the following resources can deny the traffic from reaching the instances?
    a. Network ACL
        i. Network ACL rules can deny traffic

518. Which of the following is the security protocol supported by VPC?
    a. IPsec
        i. IPSec is the security protocol supported by VPC
519. Which of the following VPC resources would you use in order for EC2-VPC instances to send traffic directly to S3?
    a. VPC endpoint
        i. A VPC endpoint enables you to create a private connection between your VPC and another AWS service without requiring access over the internet or through a NAT device, VPN connection, or AWS Direct Connect.
520. What properties of a VPC must be specified at the time of creation?
    a. CIDR block representing the IP address range.
    b. The region for the VPC
        i. The CIDR block is specified upon creation and cannot be changed. A VPC is associated with exactly one region which must be specified upon creation.
        ii. You can add a subnet to a VPC any time after it has been created, provided its address range falls within the VPC CIDR block and does not overlap with the address range of any existing CIDR block.
        iii. You can set up peering relationships between VPCs after they have been created.
521. Which VPC feature allows you to create a dual-homed instance?
    a. ENI (Elastic Network Interfaces)
        i. attaching an ENI associated with a different subnet to an instance can make the instance dual-homed.
522. Which of the following are required elements of an Auto Scaling group?
    a. Minimum size
    b. Launch configuration
        i. An Auto Scaling group must have a minimum size and a launch configuration defined in order to be created health checks and a desired capacity are optional.
523. You have created an ELB load balancer listening on port 80, and you registered it with a single EC2 instance also listening on port 80. A client makes a request to the load balancer with the correct protocol and port for the load balancer. In this scenario, how many connections does the balancer maintain?
    a. 2
        i. The load balancer maintains 2 separate connections: one connection with the client and one connection with the EC2 instance.
524. How long does CloudWatch keep metric data?
    a. 2 weeks
525. Which of the following are the minimum required elements to create an Auto Scaling launch configuration?
    a. Launch configuration name, AMI, and instance type

        i. Only the launch configuration name, AMI, and instance type are needed to create Auto Scaling launch configuration.
        ii. Identifying a key pair, security group, and a block device mapping are optional elements for an Auto Scaling launch configuration.

526. You are responsible for the application logging solution for your company's existing applications running on EC2 instances. Which of the following is the best approach for aggregating the application logs within AWS?
    a. CloudWatch Logs Agent
        i. You can use CloudWatch Logs Agent installer on existing EC2 instances to install and configure the CloudWatch Logs Agent.

527. Which of the following must be configured on an ELB load balancer to accept incoming traffic?
    a. A listener
        i. You configure your load balancer to accept incoming traffic by specifying one or more listeners.

528. You create an Auto Scaling group in a new region that is configured with a minimum size value of 10, a maximum size value of 100, and a desired capacity value of 50. However, you notice that 30 of the EC2 instances within the Auto Scaling group fail to launch. Which of the following is the cause of this behavior?
    a. You have not raised your default EC2 capacity (20) for the new region.

529. You want to host multiple HTTPS websites on a fleet of EC2 instances behind an ELB load balancer with a single X.509 certificate. How must you configure the SSL certificate so that clients connecting to the load balancer are not presented with a warning when they connect?
    a. Create one SSL certificate with a SAN (subject alternative name) value for each website name
        i. A SSL certificate must specify the name of the website in either the subject name or listed as a value in the SAN extension of the certificate in order for connecting clients to not receive a warning.

530. Your web application frontend consists of multiple EC2 instances behind an ELA load balancer. You have configured the load balancer to perform health checks on these EC2 instances. If an instance fails to pass health checks, which statement will be true?
    a. The load balancer stops sending traffic to the instance that failed its health check.
        i. When EC2 instances fail the requisite number of consecutive health checks, the load balancer stops sending traffic to the EC2 instance.

531. In the basic monitoring package for EC2, what CloudWatch metrics are available?
    a. Hypervisor visible metrics such as CPU utilization.

532. A cell phone company is running dynamic-content television commercials for a contest. They want their website to handle traffic spikes that come after a commercial airs. The website is interactive, offering personalized content to each visitor based on location, purchase history, and the current commercial airing. Which architecture will configure Auto Scaling to scale out to respond to spikes of demand, while minimizing costs during quiet periods?

    a. Configure Auto Scaling to scale out as traffic increases. Configure the launch configuration to start new instances from a preconfigured AMI.
        i. Auto Scaling is designed to scale out based on an event like increased traffic while being cost effective when not needed.

533. For an application running in the ap-northeast-1 region with 3 AZs (ap-northeast-1a, -1b, -1c), which instance deployment provides high availability for the application that normally requires 9 running EC2 instances but run on a minimum of 65 percent capacity while Auto Scaling launches replacement instances in the remaining Availability Zones?
    a. Deploy the application on 3 servers in -1a, 3 in -1b and 3 in -1c
        i. Auto Scaling will provide high availability across 3 AZs with 3 EC2 instances in each and keep capacity above the required minimum capacity, even in the event of an entire AZ becoming unavailable.

534. Which of the following are characteristics of the Auto Scaling service on AWS?
    a. Responds to changing conditions by adding or terminating EC2 instances
    b. Launches instances from a specified AMI
    c. enforces a minimum number of running EC2 instances
        i. Auto Scaling responds to changing conditions by adding or terminating instances, launches instances from an AMI specified in the launch configuration associated with the Auto Scaling group, and enforces a minimum number of instances in the min-size parameter of the Auto Scaling group.

535. Why is the launch configuration referenced by the Auto Scaling group instead of being part of the Auto Scaling group?
    a. It allows you to change the EC2 instance type and AMI without disrupting the Auto Scaling Group
    b. It facilitates rolling out a patch to an existing set of instances managed by an Auto Scaling group.
    c. It allows you to change security groups associated with the instances launched without having to make changes to the Auto Scaling group.
        i. Launch configurations are loosely coupled.

536. An Auto Scaling group may use:
    a. On-Demand instances
    b. Spot instances
        i. An Auto Scaling group may use On-Demand and Spot instances. An Auto Scaling group may not use already stopped instances, instances running someplace other than AWS, and already running instances not started by the Auto Scaling group itself.

537. CloudWatch supports which types of monitoring plans?
    a. Basic monitoring, which is free
    b. Detailed monitoring, which has an additional cost
        i. CloudWatch has 2 plans: basic, which is free, and detailed, which has an additional cost. There is no ad hoc plan for CloudWatch.

538. ELB health checks may be:

a. Ping
b. Connection attempt
c. Page request
    i. ELB health check may be a ping, a connection attempt, or a page that is checked.

539. When an EC2 instance registered with an ELB load balancer using connection draining is deregistered or unhealthy, which of the following will happen?
    a. Keep the connections open to that instance, and attempt to complete in-flight requests.
    b. Forcibly close all connections to that instance after a timeout period.
        i. When connection draining is enabled, the load balancer will stop sending requests to a deregistered or unhealthy instance and attempt to complete in-flight requests until a connection draining timeout period is reached, which is 300 seconds by default.

540. ELB supports which of the following types of load balancers?
    a. Internet-facing
    b. Internal-facing
    c. HTTPS load balancers

541. Auto Scaling supports which of the following plans for Auto Scaling groups?
    a. Manual
    b. Scheduled
    c. Dynamic
        i. Auto Scaling supports maintaining the current size of an Auto Scaling group using 4 plans: maintain current levels, manual scaling, scheduled scaling, and dynamic scaling

542. Which of the following methods will allow an application using an AWS SDK to be authenticated as a principal to access AWS Cloud services?
    a. Create an IAM user and store both parts of the access key for the user in the application's configuration
    b. Run the application on an EC2 instance with an assigned IAM role
        i. Programmatic access is authenticated with an access key, not with user names/passwords. IAM roles provide a temporary security token to an application using an SDK.

543. Which of the following are found in an IAM policy?
    a. Service Name
    b. Action
        i. IAM policies are independent of region, so no region is specified in the policy. IAM policies are about authorization for an already-authenticated principal, so no password is needed.

544. Your AWS account administrator left your company today. The administrator had access to the root user and a personal IAM administrator account. With these accounts, he generated other IAM accounts and keys. Which of the following should you do today to protect your AWS infrastructure?

a. Change the password and add MFA to the root user
b. Put an IP restriction on the root user
c. Rotate keys and change passwords for IAM accounts
d. Delete the administrator's personal IAM account

545. Which of the following actions can be authorized by IAM?
    a. Launching an EC2 instance
    b. Adding a message to a SQS queue
        i. IAM controls access to AWS resources only. Installing ASP.NET will require Windows operating system authorization, and querying an Oracle database will require Oracle authorization.

546. Which of the following are IAM security features?
    a. Password policies
    b. MFA
        i. DynamoDB global secondary indexes are a performance feature of DynamoDB
        ii. Consolidated Billing is an accounting feature allowing all bills to roll up under a single account

547. Which of the following are benefits of using EC2 roles?
    a. Credentials do not need to be stored on the EC2 instance.
    b. Key rotation is not necessary.
        i. EC2 roles must still be assigned a policy. Integration with Active Directory involves integration between AD and IAM via SAML.

548. Which of the following are based on temporary security tokens?
    a. EC2 roles
    b. Federation
        i. EC2 roles provide a temporary token to applications running on the instance.
        ii. Federation maps policies to identities from other sources via temporary tokens.

549. Your security team is very concerned about the vulnerability of the IAM administrator user accounts (the accounts used to configure all IAM features and accounts). What steps can be taken to lock down these accounts?
    a. Add multi-factor authentication (MFA) to the accounts
    b. Implement a password policy on the AWS account
    c. Apply a source IP address condition to the policy that only grants permissions when the user is on the corporate network

550. You want to grant the individuals on your network team the ability to fully manipulate EC2 instances. Which of the following accomplish this goal?
    a. Assign the managed policy, EC2FullAccess, to a group named NetworkTeam, and assign all the team members' IAM user accounts to that group
    b. Create a new policy that grants EC2:* actions on all resources, and assign that policy to each individual's IAM user account on the network team.
        i. Access requires an appropriate policy associated with a principal.

551. What is the format of an IAM policy?
    a. JSON
552. Which AWS database service is best suited for traditional OLTP (online transaction processing)?
    a. RDS
        i. RDS is best suited for traditional OLTP transactions.
        ii. Redshift is designed for OLAP workloads.
        iii. Glacier is designed for cold archival storage.
553. Which AWS database service is best suited for non-relational databases?
    a. DynamoDB
        i. DynamoDB is best suited for non-relational databases.
        ii. RDS and Redshift are both structured relational databases.
554. You are a solutions architect working for a media company that hosts its website on AWS. Currently, there is a single EC2 instance on AWS with MySQL installed locally to that EC2 instance. You have been asked to make the company's production environment more resilient and to increase performance. You suggest that the company split out the MySQL database onto an RDS instance with Multi-AZ enabled. This addresses the company's increased resiliency requirements. Now you need to suggest how you can increase performance. 99% of the company's end users are magazine subscribers who will be reading additional articles on the website, so only 1% of end users will need to write data to the site. What should you suggest to increase performance?
    a. Recommend that the company use read replicas, and distribute the traffic across multiple read replicas.
        i. In this scenario, the best idea is to use read replicas to scale out the database and thus maximize read performance. When using Multi-AZ, the secondary database is not accessible and all reads and writes must go to the primary or any read replicas.
555. Which AWS Cloud service is best suited for OLAP?
    a. Redshift
556. You have been using RDS for the last year to run an important application with automated backups enabled. One of your team members is performing routine maintenance and accidentally drops an important table, causing an outage. How can you recover the missing data while minimizing the duration of the outage?
    a. Restore the database from a recent automated DB snapshot.
        i. DB snapshots can be used to restore a complete copy of the database at a specific point in time. Individual tables cannot be extracted from a snapshot.
557. Which RDS database engines support Multi-AZ?
    a. All RDS database engines support Multi-AZ deployment
558. Which RDS database engines support read replicas?
    a. MySQL, MariaDB, PostgreSQL, and Aurora

559. Your team is building an order processing system that will span multiple AZs. During testing, the team wanted to test how the application will react to a database failover. How can you enable this type of test?
    a. Force a Multi-AZ failover from one AZ to another by rebooting the primary instance using the RDS console
        i. You can force a failover from one AZ to another by rebooting the primary instance in the AWS management console. This is often how people test a failover in the real world. There is no need to create a support case.
560. You are a system administrator whose company has moved its production database to AWS. Your company monitors its estate using CloudWatch, which sends alarms using SNS to your mobile phone. One night, you get an alert that your primary RDS instance has gone down. You have Multi-AZ enabled on this instance. What should you do to ensure the failover happens quickly?
    a. No action is necessary. Your connection string points to the database endpoint, and AWS automatically updates this endpoint to point to your secondary instance.
        i. Monitor the environment while RDS attempts to recover automatically. AWS will update the DB endpoint to point to the secondary instance automatically.
561. You are working for a small organization without a dedicated database administrator on staff. You need to install Microsoft SQL Server enterprise edition quickly to support an accounting back office application on RDS. What should you do?
    a. Launch an RDS DB instance, and select Microsoft SQL Server Enterprise Edition under the Bring Your Own License (BYOL) model
        i. RDS supports Microsoft SQL Server edition and the license is available only under the BYOL model.
562. You are building the database tier for an enterprise application that gets occasional activity throughout the day. Which storage type should you select as your default option?
    a. General Purpose SSD
        i. General Purpose SSD volumes are generally the right choice for databases that have bursts of activity.
563. You are designing an e-commerce web application that will scale to potentially hundreds of thousands of concurrent users. Which database technology is best suited to hold the session state for large numbers of concurrent users?
    a. NoSQL database table using DynamoDB
        i. NoSQL databases like DynamoDB excel at scaling to hundreds of thousands of requests with key/value access to user profile and session
564. Which of the following techniques can you use to help you meet RPO and RTO requirements?
    a. DB snapshots
    b. Read replica
    c. Multi-AZ deployment

        i. DB snapshots allow you to backup and recover your data, while read replicas and a Multi-AZ deployment allow you to replicate your data and reduce the time to failover

565. When using RDS Multi-AZ how can you offload read requests from the primary?
    a. add a read replica DB instance, and configure the client's application logic to use a read-replica.
    b. Create a caching environment using ElastiCache to cache frequently used data. Update the application logic to read/write from the cache.
        i. RDS allows for the creation of one or more read-replicas for many engines that can be used to handle reads. Another common pattern is to create a cache using Memcached and ElastiCache to store frequently used queries. The secondary slave DB instance is not accessible and cannot be used to offload queries.

566. You are building a large order processing system and are responsible for securing the database. Which actions will you take to protect the data?
    a. Adjust AWS IAM permissions for administrators
    b. configure security groups and network ACLs to limit network access
    c. configure database users, and grant permissions to database objects
        i. Protecting your database requires a multi-layered approach that secures the infrastructure, the network and the database itself. RDS is a managed service and direct access to the OS is not available.

567. Your team manages a popular website running RDS MySQL backend. The marketing department has just informed you about an upcoming television commercial that will drive thousands of new visitors to the website. How can you prepare your database to handle the load?
    a. Vertically scale DB instance by selecting a more powerful instance class.
    b. Create read replicas to offload read requests and update your application.
    c. Upgrade the storage from magnetic volumes to general purpose SSD volumes.
        i. Vertically scaling up is one of the simpler options that can give you additional processing power without making any architectural changes.
        ii. Read replicas require some application changes but let you scale processing power horizontally.
        iii. Busy databases are often I/O-bound, so upgrading storage to General Purpose SSD or Provisioned IOPS SSD can often allow for additional request processing.

568. You are building a photo management application that maintains metadata on millions of images in an DynamoDB table. When a photo is retrieved, you want to display the metadata next to the image. Which DynamoDB operation will you use to retrieve the metadata attributes from the table?
    a. Query operation
        i. Query is the most efficient operation to find a single item in a large table

569. You are creating an DynamoDB table that will contain messages for a social chat application. This table will have the following attributes: username, timestamp, message. Which attribute should you use as the partition key? The sort key?
    a. Username, Timestamp
        i. Using the username as a partition key will evenly spread your users across the partitions. Messages are often filtered down by time range, so Timestamp makes sense as a sort key.
570. Which of the following statements about DynamoDB tables are true?
    a. Local secondary indexes can only be created when the table is being created.
    b. You can only have one local secondary index.
        i. You can only have a single local secondary index, and it must be created at the same time the table is created. You can create many global secondary indexes after the table has been created.
571. Which of the following workloads are a good fit for running on Redshift?
    a. Reporting database supporting back-office analytics
    b. Data warehouse used to aggregate multiple disparate data sources.
        i. Redshift is an OLAP data warehouse designed for analytics, ETL and high-speed querying. It is not well suited for running transactional applications that require high volumes of small inserts or updates.
572. Which of the following is not a supported SNS protocol?
        i. DynamoDB
573. When you create a new SNS topic, which of the following is created automatically?
    a. ARN
        i. When you create a new SNS topic, an Amazon Resource Name is created automatically.
574. Which of the following are features of SNS?
    a. Publishers
    b. Subscribers
    c. Topic
575. What is the default time for an SQS visibility timeout?
    a. 30 seconds
576. What is the longest time available for a SQS visibility timeout?
    a. 12 hours
577. Which of the following options are valid properties of a SQS message?
    a. Message ID
    b. Body
        i. The valid properties of SQS message are Message ID and Body. Each message receives a system-assigned Message ID that SQS returns to you in the SendMessage response. The message body is composed of name/value pairs and the unstructured, uninterpreted content
578. You are a solutions architect who is working for a mobile application company that wants to use SWF for their new takeout ordering application. They will have multiple

workflows that will need to interact. What should you advise them to do in structuring the design of their SWF environment?

a. Use a single domain containing multiple workflows. In this manner, the workflows will be able to interact.
    i. Use a single domain with multiple workflows. Workflows within separate domains cannot interact.

579. In SWF, which of the following are actors?
    a. Activity workers
    b. Workflow starters
    c. Deciders

580. You are designing a new application, and you need to ensure that the components of your application are not tightly coupled. You are trying to decide between the different AWS cloud services to use to achieve this goal. Your requirements are that messages between your application components may not be delivered more than once, tasks must be completed in either a synchronous or asynchronous fashion, and there must be some form of application logic that decides what do when tasks have been completed. What application service should you use?
    a. SWF
        i. SWF would best serve your purpose in this scenario because it helps developers build, run, and scale background jobs that have parallel or sequential steps. You can think of SWF as a fully managed state tracker and task coordinator in the Cloud.

581. How does SQS deliver messages?
    a. SQS does not guarantee in what order your messages will be delivered.

582. Of the following options, what is an efficient way to fanout a single SNS message to multiple SQS queues?
    a. Create a SNS topic. Then create and subscribe multiple SQS queues sent to the SNS topic.
        i. Multiple queues can subscribe to a SNS topic, which can enable parallel asynchronous processing.

583. Your application polls a SQS queue frequently and returns immediately, often with empty ReceiveMessageResponses. What is one thing that can be done to reduce SQS costs?
    a. Use long polling by supplying a WaitTimeSeconds of greater than 0 seconds when calling ReceiveMessage.
        i. Long polling allows your application to poll the queue, and, if nothing is there, EC2 waits for an amount of time you specify (between 1 and 20 seconds). If a message arrives in that time, it is delivered to your application as soon as possible. If a message does not arrive in that time, you need to execute the ReceiveMessage function again.

584. What is the longest time available for a SQS long polling timeout?
    a. 20 seconds

585. What is the longest configurable message retention period for SQS?

    a. 14 days
586. What is the default message retention period for SQS?
    a. 4 days
587. SNS is a push notification service that lets you send individual or multiple messages to large numbers of recipients. What types of clients are supported?
    a. Publisher and subscriber client types
        i. With SNS, you send individual or multiple messages to large numbers of recipients using publisher and subscriber client types
588. In SWF, a decider is responsible for what?
    a. Defining work coordination logic by specifying work sequencing, timing, and failure conditions.
        i. The decider schedules the activity tasks and provides input data to the activity workers. The decider also processes events that arrive while the workflows is in progress and closes the workflow when the objective has been completed.
589. Can a SNS topic be recreated with a previously used topic name?
    a. Yes. The topic name should typically be available after 30-60 seconds after the previous topic with the same name has been deleted.
        i. Topic names should typically be available for reuse approximately 30-60 seconds after the previous topic with the same name has been deleted. The exact time will depend on the number of subscriptions active on the topic. Topics with a few subscribers will be available instantly for reuse, while topics with larger subscriber lists may take longer.
590. What should you do in order to grant a different AWS account permission to your SQS queue?
    a. Create a SQS policy that grants the other account access.
        i. The main difference between SQS policies and IAM policies is that a SQS policy enables you to grant a different AWS account permission to your SQS queues, but IAM policy does not.
591. Can a SNS message be deleted after being published to a topic?
    a. No. After a message has been successfully published to a topic, it cannot be recalled.
592. Bucket names must be unique across all S3.
    a. True
593. You've enabled website hosting on a bucket named "big-bucket" in the region ap-southeast-2. What website URL is assigned to your bucket?
    a. big-bucket.s3-website-ap-southeast-2.amazonaws.com
594. What is the minimum size of an object that can be uploaded to Se?
    a. 1 Byte
595. As a solutions architect, it is your job to design for high availability and fault tolerance. Your company is utilizing S3 to store large amounts of file data. What steps would you take to ensure that is an AZ was lost due to natural disaster your files would still be intact and accessible?

    a. S3 is highly available and fault tolerant by design and requires no additional configuration.

596. You're uploading an object larger than 5GB in size so it is required to use the multipart upload API. After configuring the API to upload the object, what is the maximum size of an object you can upload?
    a. 5TB

597. You've been tasked with designing an architecture that supports a custom developed application. This custom developed application has a business requirement of being able to immediately upload an object to S3 and then immediately download or view the object. Which region would you choose to not build your bucket in and why?
    a. All regions provide immediate availability for PUTS of new objects.

598. Data transferred to EC2 from S3 in the same region is free of cost.
    a. True

599. S3 can handle unlimited file storage.
    a. True

600. What is the minimum size of an EBS volume?
    a. 1GB

601. You are consulting for a company that is considering a migration to the AWS cloud. However, their current data centers house over 5TB of data and they are concerned about potential downtime and the amount of time it will take to migrate the data from their small 50mbit broadband connection to the AWS cloud. What service would you use when designing a solution for this company?
    a. Import/Export

602. You do not only host the DNS for a domain, but you can also register domains on AWS using Route53.
    a. True

603. Your company is looking at moving its web servers onto AWS. However, they are concerned about costs. Which solution below best describes how AWS can be used to scale out and manage costs effectively for your company?
    a. Implement an elastic solution that can scale based off of increase in demand and decrease the application environment when demand decreases.

604. You've been tasked with building an infrastructure that can easily be spun up and replicated in another region in a matter of minutes. Which AWS service can you use to build out a version controlled infrastructure that is easily reproduced?
    a. CloudFormation with templates

605. A properly designed, highly available fault tolerant application will almost always be built across multiple availability zones.
    a. True

606. Your company has non-production and non-priority batch workloads that can be interrupted and need EC2 instances for processing. Which instance pricing model might be the best for this workload?
    a. Spot instances

607. API credentials should never be stored on an AMI

a. True

608. You've been tasked to develop a durable backup solution using AWS resources. Which AWS resource would you use for backups on your file storage and why?
    a. S3 because it has 11x9 durability and has the ability to enable versioning.
609. Which of the following are key-value storage services on AWS?
    a. DynamoDB and S3
610. As an architect it is your job to design for fault tolerance and high availability. Which core services can you use when designing these application architectures?
    a. EC2, Auto-Scaling
611. It is best practice to always use roles when possible instead of API access keys.
    a. True
612. Which of the following is true about EBS volumes?
    a. EBS volumes cannot be attached to EC2 instances in another AZ.
613. You've been tasked with building a scalable solution on the AWS cloud. What components make up a proper scalable solution?
    a. A scalable application will be resilient and operationally efficient. A Scalable solution will decrease in cost as scale increases.
614. EBS volumes are network attached volumes.
    a. True
615. Your company is currently using Memcached in order to handle in-memory caching for your RDS servers. When migrating to AWS, which solution could you implement in order to maintain the caching engine and compatibility?
    a. build an ElastiCache cluster
616. Your company is exploring different methods to implement a disaster recovery and backup strategy to an off-site data center. After exploring several cloud options, your company has asked you to consider building the solution using AWS. What services might you consider using to implement this solution if you are solely looking at data files and file server backups?
    a. S3, Import/Export, Storage Gateway
617. Your company is low on disk storage space and is looking for a cost effective solution that also has a built-in backup and archiving capability. Which solution would you choose?
    a. Use Storage Gateway with Gateway-Cached Volumes
618. A user is running one instance for only 3 hours every day. The user wants to save some cost with the instance. Which of the below mentioned Reserved instance categories in advised in this case?
    a. The user should not use RI; instead only go with the on-demand pricing.
        i. The AWS reserved instance provides the user with an option to save some money by paying a one-time fixed amount and then save on the hourly rate. It is advisable that if the user is having 30% or more usage (>2200 hours) of an instance per day, he should go for a RI.

619. A user has created a VPC with CIDR 20.0.0.0/24. The user has created a public subnet with CIDR 20.0.0.0/25. The user is trying to create the private subnet with CIDR 20.0.0.128/25. Which of the below mentioned statements is true?
    a. It will allow the user to create a private subnet with CIDR as 20.0.0.128/25
        i. When the user creates a subnet in VPC, he specifies the CIDR block for the subnet. The CIDR block of a subnet can be the same as the CIDR block for the VPC for a single subnet in the VPC, or a subset to enable multiple subnets. If the user creates more than one subnet in a VPC, the CIDR blocks of the subnets must not overlap. Thus, in this case the user has created a VPC with CIDR block 20.0.0.0/24, which supports 256 IP addresses (20.0.0.0 to 20.0.0.255). The user can break this CIDR block into two subnets, each supporting 128 IP addresses. One subnet uses the CIDR block 20.0.0.0/25 (20.0.0.0 to 20.0.0.127) and the other 20.0.0.128.
620. A user has configured ELB by enabling a SSL negotiation configuration known as a Security Policy. Which of the below mentioned options is not part of this secure policy while negotiating the SSL connection between the user and the client?
    a. Client Order Preference
        i. ELB using SSL is used to negotiate the SSL connections between a client and the load balancer. A security policy is a combination of SSL protocols, SSL ciphers, and the server order preference option.
621. An organization has setup consolidated billing with 3 different AWS accounts. Which of the below mentioned advantages will organization receive in terms of the AWS pricing?
    a. All AWS accounts will be charged for S3 storage by combining the total storage of each account.
        i. AWS consolidated billing enables the organization to consolidate payments for multiple AWS accounts within a single organization by making a single paying account. For billing purposes, AWS treats all the accounts on the consolidated bill as one account. Some services, such as EC2 and S3 have volume pricing tiers across certain usage dimensions that give the user lower prices when he uses the service more.
622. A user has setup connection draining with ELB to allow in-flight requests to continue while the instance is being deregistered through Auto Scaling. If the user has not specified the draining time, how long will ELB allow inflight requests traffic to continue?
    a. 300 seconds
        i. The ELB connection draining feature causes the load balancer to stop sending new requests to the back-end instances when the instances are deregistering or become unhealthy, while ensuring that inflight requests continue to be served. The user can specify a maximum time (3600 seconds. For the load balancer to keep the connections alive before reporting the instance as deregistered. If the user does not specify the maximum timeout period, by default, the load balancer will close the connections to the deregistering instance after 300 seconds.

623. A user has created an S3 bucket which is not publicly accessible. The bucket is having thirty objects which are also private. If the user wants to make the objects public, how can he configure this with minimal efforts?
    a. Set the AWS bucket policy which marks all objects as public.
        i. A system admin can grant permission of the S3 objects or buckets to any user or make the objects public using the bucket policy and user policy. Both use the JSON-based access policy language. Generally if the user is defining the ACL on the bucket, the objects in the bucket do not inherit it and vice a versa. The bucket policy can be defined at the bucket level which allows the objects as well as the bucket to be public with a single policy applied to that bucket.
624. A user is checking the CloudWatch metrics from the AWS console. The user notices that the CloudWatch data is coming in UTC. The user wants to convert the data to a local time zone. How can the user perform this?
    a. In the CloudWatch console select the local timezone under the Time Range tab to view the data as per the local timezone.
625. A user has created a queue named myqueue with SQS. There are 4 messages published to queue which are not received by the consumer yet. If the user tries to delete the queue, what will happen?
    a. It will delete the queue.
        i. The user can delete a queue at any time, whether it is empty or not.
626. A customer is using AWS for Dev and Test. The customer wants to setup the Dev environment with CloudFormation. Which of the below mentioned steps are not required while using CloudFormation?
    a. Configure a service
        i. CloudFormation introduces 2 concepts. template and stack. The template is a JSON-format, text-based file that describes all the resources required to deploy and run an application. Stack is a collection of resources which are created and managed as a single unit when CloudFormation instantiates a template. While creating a stack, the user uploads the template and provides the data for the parameters if required.
627. A user has configured ELB with 3 instances. The user wants to achieve HA as well as redundancy with ELB. Which of the below mentioned AWS services helps the user achieve this for ELB?
    a. Route 53
        i. The user can provide HA and redundancy running behind ELB by enabling the Route53 DNS failover for the load balancers.
628. A user has enabled the Multi-AZ feature with the MS SQL RDS database server. Which of the below mentioned statements will help the user understand the Multi-AZ feature better?
    a. In a Multi-AZ, AWS runs just 1 DB but copies the data synchronously to the standby replica.

629. An organization is setting up programmatic billing access for their AWS account. Which of the below mentioned services is not required or enabled when the organization wants to use programmatic access?
    a. AWS billing alerts
        i. To enable programmatic access, the user has to first enable the monthly billing report. Then provide an bucket name where the billing CSV will be uploaded and enable the programmatic access option.
630. A user has stored data on an encrypted EBS volume. The user wants to share the data with his friend's AWS account. How can user achieve this?
    a. Copy the data to an unencrypted volume and then share
631. A user is planning to setup notifications on the RDS DB for a snapshot. Which of the below mentioned event categories is not supported by RDS for this snapshot source type?
    a. Backup
        i. RDS snapshot source types include: creation, deletion and restoration
632. A user has created a web application with Auto Scaling. The user is regularly monitoring the application and he observed that the traffic is highest on Thursday and Friday between 8 and 18:00. What is the best solution to handle scaling in this case?
    a. Schedule Auto Scaling to scale up by 8AM Thursday and scale down after 6PM on Friday.
633. A user has launched an EC2 instance. The user is planning to setup the CloudWatch alarm. Which of the below mentioned actions is not supported by the CloudWatch alarm?
    a. send an SMS using SNS
        i. Possible alarm actions: send a notification to a SNS topic, notify Auto Scaling policy or changing the state of the instance to stop/terminate
634. A root account owner has created a S3 bucket testmycloud. The account owner wants to allow everyone to upload the objects as well as enforce that the person who uploaded the object should manage the permission of those objects. Which is the easiest way to achieve this?
    a. The root account should use ACL with the bucket to allow everyone to upload the object.
635. A user has configured ELB with 2 EBS backed EC2 instances. The user is trying to understand the DNS access and IP support for ELB. Which of the below mentioned statements may not help the user understand the IP mechanism supported by ELB?
    a. The ELB supports either IPv4 or IPv6 but not both
636. A user has setup a CloudWatch alarm on an EC2 action when the CPU utilization is above 75%. The alarm sends a notification to SNS on the alarm state. If the user wants to simulate the alarm action how can he achieve this?
    a. The user can set the alarm state to ALARM using CLI
        i. This temporary state change lasts only until the next alarm comparison occurs.

637. An organization has added 3 of his AWS accounts to consolidated billing. One of the AWS account has purchased a reserved instance of a small instance size in the us-east-1a zone. All other AWS accounts are running instances of a small size in the same zone. What will happen in this case for the RI pricing?
    a. Any single instance from all the 3 accounts can get the benefit of AWS RI pricing if they are running in the same zone and are of the same size.
638. Which type of record is commonly used to route traffic to an IPv6 address?
    a. an AAAA record
        i. An AAAA record is used to route traffic to an IPv6 address, whereas an A record is used to route traffic to an IPv4 address.
639. Where do you register a domain name?
    a. with a domain registrar
        i. Domain names are registered with a domain registrar, which then registers the name to InterNIC.
640. You have an application the for legal reasons must be hosted in the US when US citizens access it. The application must be hosted in the EU when citizens of the EU access it. For all other citizens of the world, the application must be hosted in Sydney. Which routing policy should you choose in order to achieve this?
    a. Geolocation routing
        i. You should route your traffic based on where your end users are located. The best routing policy to achieve this is geolocation routing.
641. Which type of DNS record should you use to resolve an IP address to a domain name?
    a. A PTR record
        i. A PTR record is used to resolve an IP address to a domain name, and it is commonly referred to as reverse DNS.
642. You host a web application across multiple AWS regions in the world, and you need to configure your DNS so that your end users will get the fastest network performance possible. Which routing policy should you apply?
    a. Latency-based routing
643. Which DNS record should you use to configure the transmission of email to your intended mail server?
    a. MX records
644. Which DNS records are commonly used to stop email spoofing and spam?
    a. SPF records
        i. SPF records are used to verify authorized senders of mail from your domain
645. You are rolling out A and B test versions of a web application to see which version results in the most sales. You need 10 percent of your traffic to go to version A, 10 percent to go to version B, and the rest to go to your current production version. Which routing policy should you choose to achieve this?
    a. weighted routing
646. Which DNS record must all zones have by default?

    a. SOA
        i. The start of zone is defined by the SOA.

647. Your company has its primary production site in Western Europe and its DR site in the Asia Pacific. You need to configure DNS so that if your primary site becomes unavailable, you can fail DNS over to the secondary site. Which DNS routing policy would best achieve this?
    a. Failover routing

648. Which type of DNS record should you use to resolve a domain name to another domain name?
    a. CNAME record

649. Which is a function that Route53 does not perform?
    a. Load balancing
        i. Route53 performs 3 main functions: domain registration, DNS service, and health checking

650. Which DNS record can be used to store human-readable information about a server, network, and other accounting data with a host?
    a. A TXT record

651. Which resource record set would not be allowed for the hosted zone example.com?
    a. www.example.ca

652. Which port number is used to serve requests by DNS?
    a. 53

653. Which protocol is primarily used by DNS to serve requests?
    a. UDP

654. Which protocol is used by DNS when response data size exceeds 512 bytes?
    a. TCP

655. What are the different hosted zones that can be created in Route53?
    a. public hosted zone
    b. Private hosted zone

656. Route 53 cannot route queries to which AWS resource?
    a. OpsWorks
        i. Route53 can route queries to a variety of AWS resources such as a CloudFront distribution, ELB load balancer, EC2 instance, a website hosted in a S3 bucket, and a RDS database.

657. When configuring Route53 as your DNS service for an existing domain, which is the first step that needs to be performed?
    a. Transfer domain registration from current registrar to Route53.

658. Which of the following objects are good candidates to store in a cache?
    a. Session state
    b. Shopping cart
    c. Product catalog

659. Which of the following cache engines are supported by ElastiCache?
    a. Redis
    b. Memcached

660. How many nodes can you add to an ElastiCache cluster running Memcached?
    a. 20
661. How many nodes can you add to an ElastiCache cluster running Redis?
    a. 1
        i. Redis clusters can only contain a single node; however, you can group multiple clusters together into a replication group
662. An application currently uses Memcached to cache frequently used database queries. Which steps are required to migrate the application to use ElastiCache with minimal changes?
    a. Update the configuration file with the endpoint for the ElastiCache cluster
    b. Configure a security group to allow access from the application servers
663. How can you backup data stored in ElastiCache running Redis?
    a. Configure automatic snapshots to backup the cache environment every night
    b. Create a snapshot manually
664. How can you secure an ElastiCache cluster?
    a. Restrict API actions using IAM policies
    b. Restrict network access using a network ACL
665. You are working on a mobile gaming application and are building the leaderboard feature to track the top scores across millions of users. Which AWS services are best suited for this use case?
    a. ElastiCache using Redis
        i. Redis provides native functions that simplify the development of leaderboards. With memcached, it is more difficult to sort and rank large datasets. Redshift and S3 are not designed for high volumes of small reads and writes, typical of a mobile game.
666. You have built a large web application that uses ElastiCache using Memcached to store frequent query results. You plan to expand both the web fleet and the cache fleet multiple times over the next year to accommodate increased user traffic. How do you minimize the amount of changes required when a scaling event occurs?
    a. Configure AutoDiscovery on the client side
        i. When the clients are configured to use AutoDiscovery, they can discover new cache nodes as they are added or removed. AutoDiscovery must be configured on each client and is not active server side. Updating the configuration file each time will be very difficult to manage. Using an ELB is not recommended for this scenario.
667. What origin servers are supported by CloudFront?
    a. S3 bucket
    b. An HTTP server running on EC2
    c. An HTTP server running on-premises
668. Which of the following are good use cases for CloudFront?
    a. A popular software download site that supports users around the world, with dynamic content that changes rapidly.

b. A heavily used video and music streaming service that requires content to be delivered only to paid subscribers.
    i. Popular, users around the world and heavily used are key indicators that CloudFront is appropriate. Corporate use cases where the requests come from a single geographic location or appear to come from one (VPN). These use cases will generally not see benefit from CloudFront.

669. You have a web application that contains both static content in a S3 bucket -- primarily images and CSS files -- and also dynamic content currently served by a PHP web app running on EC2. What features of CloudFront can be used to support this application with a single CloudFront distribution?
    a. Multiple origins
    b. Multiple cache behaviors
        i. Using multiple origins and setting multiple cache behaviors allow you to serve static and dynamic content from the same distribution. Origin Access Identifiers and signed URLs support serving private content from CloudFront.

670. You are building a media-sharing web application that serves video files to end users on both PCs and mobile devices. The media files are stored as objects in an S3 bucket, but are to be delivered through CloudFront. What is the simplest way to ensure that only CloudFront has access to the objects in S3 bucket?
    a. Use an OAI (origin access identifier)
        i. OAI is a special identity that can be used to restrict access to a S3 bucket only to a CloudFront distribution.
        ii. Signed URLS, signed cookies, and IAM bucket policies can help to protect content served through CloudFront, but OAIs are the simplest way to ensure that only CloudFront has access to a bucket.

671. Your company data center is completely full, but the sales group has determined a need to store 20TB of product video. The videos were created over the last several years, with the most recent being accessed by sales the most often. The data must be accessed locally, but there is no space in the data center to install local storage devices to store this data. What AWS Cloud service will meet sales' requirements?
    a. Storage Gateway Gateway-Cached volumes

672. Your company wants to extend their existing MS AD capability into a VPC without establishing a trust relationship with the existing on-premises AD. Which of the following is the best approach to achieve this goal?
    a. Create and connect an AWS Directory Service Simple AD.
        i. Simple AD is a MS AD-compatible directory that is powered by Samba 4. Simple AD supports commonly used AD features such as user accounts, group memberships, domain-joining EC2 instances running Linux and MS Windows, Kerberos-based SSO, and group policies.

673. Which of the following are AWS KMS keys that will never exit AWS unencrypted?
    a. KMS CMKs (customer master keys)

674. Which cryptographic method is used by AWS KMS to encrypt data?

    a. Envelope encryption
        i. KMS uses envelope encryption to protect data. KMS creates a data key, encrypts it under a CMK, and returns plaintext and encrypted versions of the data key to you. You use the plaintext key to encrypt data and store the encrypted key alongside the encrypted data. You can retrieve a plaintext data key only if you have the encrypted data key and you have permission to use the corresponding master key.

675. Which AWS service records API calls made on your account and delivers log files to your S3 bucket?
    a. CloudTrail

676. You are trying to decrypt ciphertext with KMS and the decryption operation is failing. Which of the following are possible causes?
    a. The plaintext was encrypted along with an encryption context, and you are not providing the identical encryption context when calling the Decrypt API.
    b. The ciphertext you are trying to decrypt is not valid.
        i. Encryption context is a set of key/value pairs that you can pass to AWS KMS when you call the Encrypt, Decrypt, ReEncrypt, GenerateDataKey, and GenerateDataKeyWithoutPlaintext APIs. Although the encryption context is not included in the ciphertext, it is cryptographically bound to the ciphertext during encryption and must be passed again when you call the Decrypt (or ReEncrypt) API.
        ii. Invalid ciphertext for decryption is plaintext that has been encrypted in a different AWS account or ciphertext that has been altered since it was originally encrypted.

677. Your company has 30 years of financial records that take up 15TB of on-premises storage. It is regulated that you maintain these records, but in the year you have worked for the company no one has ever requested any of this data. Given that the company data center is already filling the bandwidth of its Internet connection, what is an alternative way to store the data on the most appropriate cloud storage?
    a. Import/Export to Glacier

678. Your company collects information from the point of sale registers at all of its franchise locations. Each month these processes collect 200TB of information stored in S3. Analytics jobs taking 24 hours are performed to gather knowledge from this data. Which of the following will allow you to perform these analytics in a cost-effective way?
    a. Run a transient EMR cluster, and run the MapReduce jobs against the data directly in S3.
        i. Because the job is run monthly, a persistent cluster will incur unnecessary compute costs during the rest of the month. Kinesis is not appropriate because the company is running analytics as a batch job and not on a stream. A single large instance does not scale out to accommodate the large compute needs.

679. Which service allows you to process nearly limitless streams of data in flight?
    a. Kinesis Streams

        i. Kinesis Firehose saves streams to AWS storage services.
        ii. Kinesis Streams provide the ability to process the data in the stream

680. What combination of services enable you to copy daily 50TB of data to Amazon storage, process the data in Hadoop, and store the results in a large data warehouse?
    a. S3, Data Pipeline, EMR and Redshift
        i. Data Pipeline allows you to run regular ETL jobs on amazon and on-premises data sources.
        ii. The best storage for large data is S3
        iii. Redshift is a large-scale data warehouse service

681. Your company has 50000 weather stations around the country that send updates every 2 seconds. What service will enable you to ingest this stream of data and store it to S3 for future processing?
    a. Kinesis Firehose
        i. Kinesis Firehose allows you to ingest massive streams of data and store the data on S3 as well as Redshift and Elasticsearch.

682. Your organization uses Chef heavily for its deployment automation. What AWS Cloud service provides integration with Chef recipes to start new application server instances, configure application server software, and deploy applications?
    a. OpsWorks
        i. OpsWorks uses Chef recipes to start new app server instances, configure application server software and deploy applications. Organizations can leverage Chef recipes to automate operations like software configurations, package installations, database setups, server scaling, and code deployment.

683. A firm is moving its testing platform to AWS to provide developers with instant access to clean test and development environments. The primary requirement for the firm is to make environments easily reproducible and fungible. What service will help the firm meet their requirements?
    a. CloudFormation
        i. With CloudFormation, you can reuse your template to set up your resources consistently and repeatedly. Just describe your resources once and then provision the same resources over and over in multiple stacks.

684. Your company's IT management team is looking for an online tool to provide recommendations to save money, improve system availability and performance, and to help close security gaps. What can help the management team?
    a. Trusted Advisor

685. Your company works with data that requires frequent audits of your AWS environment to ensure compliance with internal policies and best practices. In order to perform these audits, you need access to historical configurations of your resources to evaluate relevant configuration changes. Which service will provide the necessary information for your audits?
    a. Config

        i. Config is a fully managed service that provides you with an AWS resource inventory, configuration history, and configuration change notifications to enable security and governance. With Config, you can discover existing and deleted AWS resources, determine your overall compliance against rules, and dive into configuration details of a resource at any point in time. These capabilities enable compliance auditing.

686. All of the website deployments are currently done by your company's development team. With a surge in website popularity, the company is looking for ways to be more agile with deployments. What AWS Cloud service can help the developers focus more on writing code instead of spending time managing and configuring servers, databases, load balancers, firewalls, and networks?
    a. Elastic Beanstalk
        i. Elastic Beanstalk is the fastest and simplest way to get an application up and running on AWS. Developers can simply upload their application code, and the service automatically handles all the details such as resource provisioning, load balancing, Auto Scaling, and monitoring.
687. Which is an operational process performed by AWS for data security?
    a. Decommissioning of storage devices using industry-standard practices.
688. You have launched a Windows EC2 instance and specified an EC2 key pair for the instance at launch. Which of the following accurately describes how to log into the instance?
    a. Use EC2 key pair to decrypt the administrator password and then securely connect to the instance via RDP as the administrator.
        i. The admin password is encrypted with the public key of the key pair, and you provide the private key to decrypt the password.
689. A database security group controls network access to a database instance that is inside a VPC and by default allows access from?
    a. No access is provided by default, and any access must be explicitly added with a rule to the DB security group.
690. Which encryption algorithm is used by S3 to encrypt data at rest with SSE?
    a. AES-256
691. How many access keys may an AWS IAM user have active at one time?
    a. 2
        i. IAM permits users to have no more than 2 active access keys at one time.
692. Which of the following is the name of the security model employed by AES with its customers?
    a. The shared responsibility model
693. Which of the following describes the scheme used by a Redshift cluster leveraging AWS KMS to encrypt data at-rest?
    a. Redshift uses a **4-tier**, key-based architecture for encryption
        i. When you choose AWS KMS for key management with Redshift, there is a 4-tier hierarchy of encryption keys. These keys are the master key, a cluster key, a database key, and data encryption keys.

694. Which of the following ELB options ensure that the load balancer determines which cipher is used for a SSL connection?
    a. Server Order Preference
        i. ELB supports the Server Order Preference option for negotiating connections between a client and a load balancer. During the SSL connection negotiation process, the client and the load balancer present a list of ciphers and protocols that they each support, in order of preference. By default, the first cipher on the client's list that matches any one of the load balancer's ciphers is selected for the SSL connection. If the load balancer is configured to support Server Order Preference, then the load balancer selects the first cipher in its list that is in the client's list of ciphers. This ensures that the load balancer determines which cipher is used for SSL connection. If you do not enable Server Order Preference, the order of ciphers presented by the client is used to negotiate connections between the client and the load balancer.
695. Which technology does WorkSpaces use to provide data security?
    a. PC-over-IP (**PCoIP**)
        i. WorkSpaces uses PCoIP, which provides an interactive video stream without transmitting actual data.
696. As a solutions architect, how should you architect systems on AWS?
    a. You should architect your AWS usage to take advantage of multiple regions and AZs.
        i. Distributing applications across multi-AZs provides the ability to remain resilient in the face of most failure modes, including natural disasters or system failures.
697. Which security scheme is used by the Multi-Factor Authentication token?
    a. Time-based one-time password (TOTP)
698. DynamoDB tables may contain sensitive data that needs to be protected. Which of the following is a way for you to protect DynamoDB table content?
    a. DynamoDB can store data encrypted with a client-side encryption library solution before storing the data in DynamoDB
    b. DynamoDB can be used with the AWS KMS to encrypt the data before storing the data in DynamoDB.
699. You have launched an Amazon Linux EC2 instance into EC2-Classic, and the instance has successfully passed the system status check and instance status check. You attempt to securely connect to the instance via SSH and receive the response "WARNING: UNPROTECTED PRIVATE KEY FILE", after which the login fails. Which of the following is the cause of the failed login?
    a. The permissions for the private key are too insecure for the key to be trusted.
        i. If your private key can be read or written to by anyone but you, then SSH ignores your key.
700. Which of the following public identity providers are supported by Cognito Identity?
    a. Amazon, Google and Facebook

        i. Cognito Identity supports public identity providers Amazon, Facebook, and Google as well as unauthenticated identities.

701. Which feature of AWS is designed to permit calls to the platform from an EC2 instance without needing access keys placed on the instance?
    a. IAM instance profile
        i. IAM instance profile is a container for an IAM role that you can use to pass role information to an EC2 instance when te instance starts.
702. Which of the following VPC elements acts as a stateless firewall?
    a. Network ACL
703. Which of the following is the most recent version of the AWS digital signature calculation process?
    a. Signature version 4
        i. **The signature version 4 signing process describes how to add authentication information to AWS requests.** For security, most requests to AWS must be signed with an access key (AKI and SAK). If you use the AWS CLI or one of the AWS SDKs, those tools automatically sign requests for you based on credentials that you specify when you configure the tools. However, if you make direct HTTP or HTTPS calls to AWS, you must sign the requests yourself.
704. Which of the following is the name of the feature within VPC that allows you to launch EC2 instance on hardware dedicated to a single customer?
    a. Dedicated tenancy
        i. Dedicated instances are physically isolated at the host hardware level from your instances that aren't dedicated instances and from instances that belong to other AWS accounts.
705. Which of the following describes how EMR protects access to the cluster?
    a. The master node is launched into a security group that allows SSH and service access, while the slave nodes are launched into a separate security group that only permits communication with the master node.
        i. EMR starts your instances in 2 EC2 security groups, one for the master and another for the slaves. The master security group has a port open for communication with the service. It also has the SSH port open to allow you to securely connect to the instances via SSH using the key specified at startup. The slaves start in a separate security group, which only allows interaction with the master instance. By default, both security groups are set up to prevent access from external sources, including EC2 instances belonging to other customers. Because these are security groups in your account, you can reconfigure them using the standard EC2 tools or dashboard.
706. To help prevent data loss due to the failure of any single hardware component EBS automatically replicates EBS volume data to which of the following?
707. EBS replicates EBS volume data within the same AZ in a region.

    a. When you create an EBS volume in an AZ, it is automatically replicated within that AZ to prevent data loss due to failure of any single hardware component. An EBS snapshot creates a copy of an EBS volume to S3 so that copies of the volume can reside in different AZs within a region.

708. EBS snapshots occur …
    a. asynchronously
        i. Snapshots occur asynchronously. The point-in-time snapshot is created immediately, but the status of the snapshot is pending until the snapshot is complete

709. AWS communicates with customers regarding its security and control environment through a variety of different mechanisms. Which of the following are valid mechanisms?
    a. obtaining industry certifications and independent third-party attestations
    b. publishing information about security and AWS control practices via the website, whitepapers, and blogs
    c. directly providing customers with certificates, reports, and other documentation (under NDA in some cases)
        i. AWS does not allow customers' auditors direct access to AWS data centers, infrastructure, or staff.

710. Which of the following statements is true when it comes to the AWS shared responsibility model?
    a. The shared responsibility model is not just limited to security considerations. it also extends to IT controls.

711. AWS provides IT control information to customers in which of the following ways?
    a. By using specific control definitions or through general control standard compliance

712. Which of the following is a valid report, certification, or third-party attestation for AWS?
    a. SOC 1
    b. PCI DSS Level 1
    c. ISO 27001
        i. There is no such thing as a SOC 4 report.

713. Which of the following statements is true?
    a. IT governance is still the customer's responsibility, despite deploying their IT estate onto the AWS platform.

714. Which of the following statements is true when it comes to the risk and compliance advantages of the AWS environment?
    a. Few, Many, or all components of a workload can be moved to AWS Cloud, but it is the customer's responsibility to ensure that the entire workload remains compliant with various certifications and third-party attestations.

715. Which of the following statements best describes an AZ?
    a. Each AZ consists of multiple discrete data centers with redundant power and networking/connectivity.

        i. AWS does not scan customer instances and customers must request the ability to perform their own scans in advance.

716. With regard to vulnerability scans and thread assessments of the AWS platform, which of the following statements are true?
    a. AWS regularly performs scans of public-facing endpoint IP address for vulnerabilities
    b. AWS security notifies the appropriate parties to remediate any identified vulnerabilities
717. Which of te following best describes the risk and compliance communication responsibilities of customers to AWS?
    a. AWS publishes information about the AWS security and control practices online, and directly to customers under NDA. Customers do not need to communicate their use and configurations to AWS.
718. When it comes to risk management, which of the following is true?
    a. AWS has developed a strategic business plan to identify any risks and has implemented controls to mitigate or manage those risks. Customers should also develop and maintain their own risk management plans to ensure they are compliant with any relevant controls and certifications.
719. The AWS control environment is in place for the secure delivery of AWS Cloud service offerings. Which of the following does the collective control environment NOT explicitly include?
    a. Energy
        i. The collective control environment includes people, processes, and technology necessary to establish and maintain an environment that supports the operating effectiveness of AWS control framework.
720. Who is responsible for the configuration of security groups in an AWS environment?
    a. AWS provides the security group functionality as a service, but the customer is responsible for correctly and securely configuring their own security groups.
721. Which of the following is NOT a recommended approach for customers trying to achieve strong compliance and governance over an entire IT control environment?
    a. Implement generic control objectives that are not specifically designed to meet their organization’s compliance requirements
        i. Customers should ensure that they implement control objectives that are designed to meet their organization’s own unique compliance requirements.
722. When designing a loosely coupled system, which AWS services provide an intermediate durable storage layer between components?
    a. Kinesis
    b. SQS
723. Which of the following options will help increase the availability of a web server farm?
    a. Launch the web server instances across multiple AZs
    b. Leverage Auto Scaling to recover from failed instances

724. Which of the following AWS Cloud services are designed according to the Multi-AZ principle?
    a. DynamoDB
    b. S3
725. Your e-commerce site was designed to be stateless and currently runs on a fleet of EC2 instances. In an effort to control cost and increase availability, you have a requirement to scale the fleet based on CPU and network utilization to match the demand curve for your site. What services do you need to meet this requirement?
    a. CloudWatch
    b. Auto Scaling
        i. Auto Scaling enables you to follow the demand curve for your applications closely, reducing the need to provision EC2 capacity manually in advance.
726. Your compliance department has mandated a new requirement that all data on EBS volumes must be encrypted. Which of the following steps would you follow for your existing EBS volumes to comply with the new requirement?
    a. Create a new EBS volume with encryption enabled.
    b. Attach an EBS volume with encryption enabled to the instance that hosts the data, then migrate the data to the encryption-enabled EBS volume
    c. Copy the data from the unencrypted EBS volume to the EBS volume with encryption enabled
727. When building a DDos-resilient architecture, how does VPC help minimize the attack surface area?
    a. Reduces the number of necessary internet entry points
    b. obfuscates necessary internet entry points to the level that untrusted end users cannot access them
    c. adds non-critical internet entry points to the architecture
        i. The attack surface is composed of the different internet entry points that allow access to your application. The strategy to minimize the attack surface area using VPC is to:
            1. reduce the number of necessary internet entry points
            2. eliminate non-critical internet entry points
            3. separate end user traffic from management traffic
            4. obfuscate necessary internet entry points to the level that untrusted end users cannot access them
            5. decouple internet entry points to minimize the effects of attacks.
728. Your e-commerce application provides daily and ad hoc reporting to various business units on customer purchases. This is resulting in an extremely high level of read traffic to your MySQL RDS instance. What can you do to scale up read traffic without impacting your database's performance?
    a. create a read replica for a RDS instance

729. Your website is hosted on a fleet of web servers that are load balanced across multiple AZs using an ELB. What type of record set in Route 53 can be used to point myawesomeapp.com to your website?
    a. Type A Alias resource record set
        i. You cannot create a CNAME record at the top node of DNS namespace (zone apex).
730. You need a secure way to distribute your AWS credentials to an application running on EC2 instances in order to access supplementary AWS Cloud services. What approach provides your application access to use short-term credentials for signing requests while protecting those credentials from other users?
    a. Provision the EC2 instances with an instance profile that has the appropriate privileges
        i. An instance profile is a container for an IAM role that you can use to pass role information to an EC2 instance when the instance starts. IAM role should have a poilcy attached that only allows access to the AWS Cloud services necessary to perform its function
731. You are running a suite of microservices on AWS Lambda that provide the business logic and access to data stored in DynamoDB for your task management system. You need to create well-defined RESTful Application Program Interfaces (APIs) for these microservices that will scale with traffic to support a new mobile application. What AWS Cloud service can you use to create the necessary RESTful APIs?
    a. API Gateway
        i. API Gateway is a fully managed service that makes it easy for developers to publish, maintain, monitor, and secure AIPs at any scale. You can create an AIP that acts as a front door for applications to access data, business logic, or functionality from your code running on AWS Lambda. API Gateway handles all of the tasks involved in accepting and processing up to hundreds of thousands of concurrent API calls, including traffic management, authorization and access control, monitoring, and API version management.
732. Your WordPress website is hosted on a fleet of EC2 instances that leverage Auto Scaling to provide high availability. To ensure that the content of the WordPress site is sustained through scale up and scale down events, you need a common file system that is shared between more than one EC2 instance. Which AWS Cloud service can meet this requirement?
    a. EFS
733. You are changing your application to move session state information off the individual EC2 instances to take advantage of the elasticity and cost benefits provided by Auto Scaling. Which of the following AWS Cloud services is best suited as an alternative for storing session state information?
    a. DynamoDB
734. A media sharing application is producing a very high volume of data in a very short period of time. Your back-end services are unable to manage the large volume of

transactions. What option provides a way to manage the flow of transactions to your back-end services?

a. Use a SQS queue to buffer the inbound transactions

735. Which of the following are best practices for managing IAM user access keys?

a. Use different access keys for different applications
b. Rotate access keys periodically
c. Configure MFA for your most sensitive operations

736. You need to implement a service to scan API calls and related events' history to your AWS account. This service will detect things like unused permissions, overuse of privileged accounts, and anomalous logins. Which of the following AWS Cloud services can be leveraged to implement this service?

a. CloudTrail
b. S3
c. Lambda

737. Government regulations require that your company maintain all correspondence for a period of seven years for compliance reasons. What is the best storage mechanism to keep this data secure in a cost-effective manner?

a. Glacier

738. Your company provides media content via the internet to customers through a paid subscription model. You leverage CloudFront to distribute content to your customers with low latency. What approach can you use to serve this private content securely to your paid subscribers?

a. Provide signed CloudFront URLs to authenticated users to access the paid content

739. Your company provides transcoding services for amateur producers to format their short films to a variety of video formats. Which service provides the best option for storing the videos?

a. S3

740. A week before Cyber Monday last year, your corporate data center experienced a failed air conditioning unit that caused flooding into the server racks. The resulting outage cost your company significant revenue. Your CIO mandated a move to the cloud, but he is still concerned about catastrophic failures in a data center. What can you do to alleviate his concerns?

a. Distribute the architecture across multiple AZs

741. Your VPC includes multiple private subnets. The instances in these private subnets must access third-party payment APIs over the internet. Which option will provide highly available internet access to the instances in the private subnets?

a. Create a NAT gateway in each AZ and configure your routing to ensure that resources use the NAT gateway in the same AZ

742. Which AWS service would you use to support an audit of API calls made to your account?

a. Cloud Trail

743. Which of the following is the fundamental resource managed by AWS KMS?

a. CMK (Customer Master Key)

744. Which of the following is an AWS Directory Service offering that is powered by Samba 4 and is MS AD-compatible?
    a. Simple AD

745. Which of the following types of security credentials are supported within AWS?
    a. Email address and password
    b. IAM username and password
    c. MFA
    d. Access keys
    e. Key pairs

746. Which of the following best describes the shared responsibility model?
    a. Customers are responsible for security in the cloud, and AWS is responsible for security of cloud.

747. Which of the following best describes how AWS Signature Version 4 secures requests?
    a. Verify the identity of the requestor, protect data in transit, and protect against potential replay attacks.

748. Which component within a VPC controls how traffic is directed?
    a. Route table

749. Which of the following describes the range of Classless Inter-Domain Routing (CIDR) addressing supported by Amazon VPC?
    a. In CIDR notation, from /28 (16 IP addresses) to /16 (65536 IP addresses)

750. How many IP addresses does AWS reserve within a subnet for its own networking purposes?
    a. 5
        i. Amazon reserves the first four and the last IP addresses of every subnet for IP networking purposes

751. Which IT automation language is supported by AWS OpsWorks?
    a. Chef

752. How many AWS Trusted Advisor checks are available to all AWS customers?
    a. 4

753. Which AWS Cloud service is designed to help web developers focus on writing code rather than spending time managing and configuring servers, databases, load balancers, firewalls, and networks?
    a. Elastic Beanstalk

754. Which of the following OSI layers are configurable on an ELB?
    a. 4 and 7
        i. Layer 4 is the transport layer that describes the TCP connection between the client and your back-end instance through the load balancer.
        ii. Layer 7 is the application layer that describes the use of HTTP and HTTPS connections from clients to the load balancer and from the load balancer to your back-end instance

755. What is the result of enabling Proxy Protocol on an ELB?

    a. A human-readable header is added to the request header with connection information such as the source IP address, destination IP address, and port numbers.

756. Which of the following is the best description of an Auto Scaling cool-down period?
    a. A period of time after an Auto Scaling event during which Auto Scaling waits before resuming Auto Scaling activities.

757. Which of the following AWS storage services is best suited for the following requirement: Your web application needs large-scale storage capacity and performance.
    a. S3

758. What is a common pattern for implementing loose coupling between services?
    a. Asynchronous integration

759. Which of the following AWS Cloud services are Multi-AZ by design?
    a. S3
    b. DynamoDB

760. Your company's senior management wants to query several data stores in order to obtain a big picture view of the business. The amount of data contained within the data stores is at least 2 TB in size. Which of the following is the best AWS service to deliver results to senior management?
    a. Redshift

761. What is the difference between a DynamoDB Query operation and a DynamoDB Scan operation?
    a. A Query operation finds items in a table or a secondary index using only primary key attribute values and a Scan operation reads every item in a table or a secondary index.

762. How is data copied from a RDS MySQL, MariaDB, or PostgreSQL source database to a Read Replica?
    a. Asynchronously

763. Which of the following is not a supported notification protocol for Amazon SNS?
    a. ElastiCache
        i. SNS supports protocols: HTTP, HTTPS, SQS, Email. SMS, and Lambda

764. What is the functional difference between an SQS delay queue and SQS visibility timeout?
    a. Delay queues make messages unavailable upon arrival to the queue and visibility timeouts make messages unavailable after being retrieved from the queue.

765. Which SQS identifier must be used to delete a message from a queue?
    a. Receipt Handle

766. How many AZs may a VPC subnet span?
    a. 1

767. Your company has a legacy application running locally on a VM on Windows 2008. The application is a compute workload that has no state. The original development team is no longer with the firm, and recreating the VM would be extremely difficult and time-consuming. To mitigate the risk, the company wants to create a Disaster Recovery

installation for this application on AWS. What method accomplishes this in a straightforward and cost-effective manner?

a. Use VM Import/Export to load the VM on AWS as an AMI. Launch an EC2 instance from the AMI to confirm operation, then terminate the EC2 instance until failover is needed.
    i. VM Import/Export is an excellent way to get the VM image into AWS, but there is no need to pay the compute costs of a running instance while the DR installation is not needed.

768. Your company runs a monitoring application that retrieves server status information hourly from all of the servers in your data center and cloud infrastructure. The information is required to be maintained for 24 hours. When the information is retrieved, appending it to the end of the 200GB file (and removing data older than 24 hours from the top of the file) requires 2500 IOPS for one minute. Network administrators periodically display summary information and randomly accessed rows in a browser-based application, but this operation requires a negligible amount of IOPs. What type of storage is well suited for this workload?

a. General Purpose SSD EBS volume
    i. bursty nature + cost-effective

769. Which of the following characteristics does the instance type define for a new EC2 instance?

a. the number of virtual CPUs
b. the amount of internal memory

770. The application you are running on your HPC cluster is very sensitive to network jitter between EC2 instances. What EC2 feature helps reduce jitter?

a. Enhanced networking

771. You are running a photomontage application that takes many photo files from S3, assembles them into a single large image that includes all of the photos, and stores the resulting image back on S3. As part of the processing, your application requires access to block storage for writing temporary files. What storage is well suited for this workload?

a. instance storage

772. Which of the following is true for EBS?

a. EBS volumes are available in a variety of magnetic and SSD options
b. EBS volumes can be encrypted transparently to applications accessing the data

773. Your company has a security policy stating that all new servers must include all of the latest Linux security updates when first configured. How can you comply with this policy when launching new EC2 instances in an Auto Scaling group on AWS?

a. Launch all new instances with a bootstrapping script that installs all of the latest updates.

774. What method will allow you to log into a new EC2 Linux instance using SSH?

a. Launch SSH and provide the private half of the key pair associated with the instance.

775. What feature allows an application running on an EC2 instance to access S3 without storing an access key on the instance?

a. EC2 role

776. What features are available to help you increase security on IAM user accounts?
    a. MFA
    b. Conditions limiting authorization to calls from a particular CIDR block
    c. Password policies
777. How does identity federation work?
    a. A temporary security token with specific permissions is provided for a console session based on the user's identity in an external Identity Provider (IdP).
778. A shipping company tracks the location of every package using barcode readers that transfer over wireless back to the main tracking system. The resulting traffic averages 10000 scans a second. What service will process the incoming stream of scans and update the data in the tracking system on a near real-time basis?
    a. Kinesis Stream
779. Which of the following are possible destinations for data ingested with Kinesis Firehose?
    a. S3
    b. Redshift
    c. ElasticSearch
780. Your company stores click stream logs from your website on S3 nightly, averaging 20TB of new data a night. Your data scientists would like to analyze this data to segment uses and understand user preferences. Which service is well suited to meet their needs?
    a. EMR
781. Your company uses an architectural drawing program and stores the drawings on an on-premises server with an attached iSCSI drive. Your company wants to replicate all of the drawings to the cloud for a durable backup and disaster recovery. Since the users interact directly with the files, latency is critical and the files must all reside locally. What AWS Cloud service can meet your requirements?
    a. Storage Gateway - Gateway Stored Volume
782. Which of the following can be configured as an origin for a CloudFront distribution?
    a. An on-premises HTTP server
    b. S3 bucket
    c. ELB
783. Your company provides videos to registered customers that pay a monthly fee. You want to provide low-latency access to the content, but restrict access to only paid subscribers. Which CloudFront feature will allow you to achieve this outcome?
    a. Signed URLs
784. You would like to set an alarm to notify you if the number of error messages in your Windows application log gets too high. How can you use CloudWatch to accomplish this?
    a. Use the CloudWatch Logs agent to export the Application Log to CloudWatch
785. How long does it take to retrieve an object from Glacier?
    a. 3-5 hours
786. Which of the following are differences between S3 and EBS?

a. S3 buckets can hold a virtually unlimited amount of data; EBS volumes have a maximum size.
b. S3 is object storage; EBS is block storage
c. S3 is accessed via a REST API; EBS is accessed from an attached EC2 instance

787. Which of the following can be used to control access to objects in S3?
a. S3 bucket policies
b. IAM policies
c. ACL

788. What options are available for SSE on S3?
a. S3 managed keys
b. KMS managed keys
c. customer-provided keys

789. Your company stores financial records in S3, but wants to minimize costs. Once created, the records are accessed regularly for the first 2 months, then used rarely for summary reporting for up to one year. After that they must be retained for compliance reasons for seven years, but only accessed in response to discovery requests or other legal actions. What lifecycle policies should you implement?
a. Store data in S3, migrate to S3 IA after 2 months, migrate to Glacier after 1 year, and then delete after 7 years.

790. You are deploying a static website on S3. After naming your bucket with a proper DNS name, enabling and configuring website hosting, uploading your content, and redirecting your domain name to the bucket, users still cannot load the page. What is a possible problem?
a. The website content was not set to publicly readable.

791. You are undergoing a PCI DSS compliance audit, and your auditor needs to run a vulnerability scan on your AWS environment. Which of the following is true?
a. You need to request permission from AWS before running a vulnerability scan.

792. Which of the following services is a proxy service for connecting your on-premises AD to the AWS Cloud?
a. AD Connector

793. Which of the following Cloud services provide you the ability to manage your own symmetric and asymmetric keys?
a. KMS (create and control symmetric encryption keys)
b. CloudHSM (create and control both symmetric and asymmetric keys)

794. How long does it take CloudTrail to deliver an event for an API call?
a. Typically within 15 minutes

795. Which of the following network actions are not possible within AWS?
a. IP spoofing
b. Promiscuous packet sniffing

796. The Signature Version 4 digital signature calculation process provides which security benefits?
a. Message integrity
b. Replay attack prevention

797. EC2 supports which type of key pair?
    a. **RSA 2048 SSH-2 key pair**
798. Which of the following best describes a VPC private subnet?
    a. A subnet with an associated route table that does not have a route to the IGW.
799. What function does a VPC IGW serve?
    a. An IGW serves two purposes: to provide a target in your VPC route tables for internet-routable traffic and to **perform NAT for instances that have been assigned public IP addresses**
800. With which of the following is a VPC ENI (Elastic network interface) associated?
    a. **VPC subnet**
801. Which of the following describes AWS OpsWorks components?
    a. **Stacks, layers, Chef recipes, instances, and apps**
802. Which of the following policies must be specified on a CloudFormation resource in order to ensure that the resource is not deleted when the stack is deleted?
    a. **Deletion policy**
803. AWS Trusted Advisor is available through which AWS API?
    a. **AWS Support**
804. Which of the following provides an automated way to send log data to CloudWatch Logs for EC2 instances running Amazon Linux or Ubuntu?
    a. CloudWatch Logs agent
805. Which of the following explains why referencing an ELB load balancer by its DNS CNAME is a best practice?
    a. An ELB load balancer's DNS CNAME will not change over time, providing you with a single fixed addressing entry, **regardless of the pool of IPs referenced by the CNAME**. (e.g. a dynamic set of IP addresses that are used by the load balancer based on traffic demand)
806. Which of the following services can be used together to create a highly available application with a resilient architecture on AWS?
    a. CloudWatch
    b. Auto Scaling
    c. ELB
807. Issuance of temporary tokens to authenticated entities through federation, credential rotation, and assigning permissions to AWS IAM roles for EC2 instances is an example of adhering to which security principle?
    a. **Least privilege**
        i. instead of placing long-term IAM user credentials with code on EC2 instance
808. Which of the following must you leverage in your cloud applications in order to take advantage of the cloud's parallelization quality?
    a. **Multi-threading**
809. Which one of the following is the most important benefit of using a cloud environment?

a. The ability to use the cloud's API to automate deployment processes and to build self-healing systems.

810. Your company's senior management wants to launch an ecommerce web application in the cloud. The application will be two-tiered and the data layer will receive hundreds of reads and writes per second due to customers searching for and purchasing items. Which of the following is the best service to use to deliver the data layer for the e-commerce application?
    a. RDS
811. Which of the following RDS database engines are compatible with MySQL?
    a. Aurora
812. Which of the following Redshift components does a client application interact with directly?
    a. Leader node
813. Which of the following is a primary benefit of using a dead letter queue?
    a. The ability to sideline and isolate unsuccessfully processed messages.
814. Which of the following services support the ability to push notifications directly to mobile applications?
    a. SNS
815. What is the maximum visibility timeout for an SQS message?
    a. 12 hours
816. How many IPSec tunnels are established when creating a VPN connection between a CGW and VPG?
    a. 2
817. Your organization has a registration application that runs easily on 2 EC2 instances for the vast majority of the year. At the end of February there is a registration deadline when the traffic load increases by a factor of five. What is an appropriate mix of EC2 instances for this situation?
    a. Purchase 2 reserved instances to handle the load for the entire year, and then add 8 on-demand instances during the last 2 weeks of February.
818. Your firm has just received 200TB of TIF images that must be converted to GIF images. This conversion will occur only once. The primary requirement for the workload is to complete the work within a fixed budget, although the business would also appreciate a swift completion. What is an appropriate mix of EC2 instances to meet the business's goals?
    a. Launch the number of on-demand instances that fit within the project budget. Augment those with spot instances at a lower price to complete the task sooner for less compute cost. When all the images are converted, terminate all the instances.
819. What does the AMI define for a new EC2 instance?
    a. The operating system version and configuration
    b. Software installed on the instance at launch
820. Your workload needs low network latency and high network throughput between EC2 instances. Which feature will help you achieve this?

    a. EC2 placement groups
821. You downloaded a 25 GB video to the z:\ drive of your EC2 instance. To get more compute, you stopped the instance to change the instance type, then restarted the instance. When you connect to your instance again, you notice that the video is no longer on the z:\ drive. What is the likely cause?
    a. The z:\ is instance storage and data is lost when the instance is stopped
822. Which of the following is true for EBS?
    a. An EC2 instance can be associated with multiple EBS volumes.
823. Your company runs an application that relies on a legacy database system. The database requires 2 TB of block storage and 40k IOPS of storage capacity. What architecture will support the requirements for this database?
    a. 2 EBS Provisioned IOPS volumes, each provisioned for 20000 IOPS in a RAID 0 configuration.
        i. General Purpose SSD volumes have a max IOPS of 10000.
824. What method will allow you to log into a new EC2 Windows instance using RDP?
    a. Decrypt the admin password for the instance using the private half of the key pair associated with the instance.
825. Which of the following are not found in an IAM policy?
    a. Temporary security token
826. What does AWS IAM control?
    a. Access to AWS resources
827. Which of the following are true about temporary security tokens?
    a. They are only valid for a specified period of time.
    b. They are the underlying implementation of AWS IAM roles for EC2
828. Your company needs to perform a tech refresh on 200 TB of on-premises storage and has decided to move the data to S3. Your datacenter internet connection has an average of 1 Gbps of available bandwidth. What is the best way to move your data to S3?
    a. Use 3 AWS Import/Export Snowball appliances.
829. Your company receives information on every sale from thousands of Point-of-Sale (POS) machines, resulting in 5000-10000 messages per second and 20TB of new data every day. Every night this information must be analyzed and summary results stored in a MySQL database for display in a management dashboard. What combination of services will address this use case?
    a. Kinesis Firehose, S3, EMR, Data Pipeline and RDS
        i. EBS: limited storage
        ii. Glacier: 3-5 retrieval time
        iii. Direct Connect: direct line to customer’s datacenter
830. When you launch an EMR cluster, what must you specify?
    a. The instance type of the nodes
    b. The number of nodes
    c. The version of Hadoop

831. You have an application that indexes documents and makes them available via a browser-based interface. The documents are stored on a block storage device accessed by the application over an iSCSI interface, but time of retrieval is not a critical factor. As the number of documents grows, you have been tasked with increasing the capacity of the application, as well as providing durable backup for the documents. What service can meet your requirements?
    a. Gateway-Cached volume (block storage)
832. Your company website includes hundreds of product manuals. Your datacenter is full and there is no room for more servers or storage, so you have been instructed to reduce storage requirements, lower compute needs, and improve response time for the entire website. What architectural pattern will achieve this?
    a. Migrate all product manuals to S3, create a CloudFront distribution with 2 origins, and use path patterns to point one origin to the manuals on S3 and the other to your web farm.
833. Your company publishes a high-definition photo library with subscribers located worldwide. Due to the level of definition, the photo files are very large and catalog pages are required to download large amounts of static images. Which service can be leveraged to provide your users with the most responsive experience?
    a. CloudFront
834. Your company's web application recently had a service interruption when the CPU usage spiked. When reviewing the logs, you found that the five-minute interval between CloudWatch metrics prevented you from fully analyzing the problem. What can you do to ensure that you have more granular visibility in the future?
    a. Deploy new instances with detailed monitoring enabled. (every minute)
835. Which of the following measured values requires a custom metric to be logged by CloudWatch?
    a. number of requests handled by a web server
836. Which of the following are appropriate use cases for S3?
    a. hosting a static website
    b. storing a library of video files
837. Your company has 1 PB of scientific research data that has been summarized and the results stored in a MySQL database. While there is no need to access the raw measurements, they must be retained indefinitely. which of the following storage plans meets these needs in the most cost-efficient manner?
    a. Glacier
838. Your company maintains a photo library with 100 million photos available for licensing. The total storage requirement for the photos themselves is 200 GB. There is also a low -resolution version of each photo generated. with those versions taking up a total of 10GB. Your company runs on very low margins, so minimizing cost is important. What is a cost-effective way to store all your photos?
    a. Store the original photos on S3 and the low-resolution versions on S3 RRS.

839. Your company maintains financial records for customers. These records are stored as PDF files on S3. Compliance rules require that you record when any user accesses a PDF file. How can you meet this compliance goal?
    a. Turn on bucket logging to generate access logs
840. Your company publishes an image library used by hundreds of customer websites. The site contains thousands of images and the library services over 100000 requests per second. You would like to take advantage of the price and scalability of S3 to implement this workload. What should you do to ensure the best performance?
    a. Add a randomized prefix to every object name.
841. Your company faces a compliance requirement dictating that none of your EC2 instances can be hosted on hardware shared with any other AWS customer. Which 2 services allow this?
    a. Dedicated instances
    b. Dedicated hosts

-