Practice exams (Solutions Architect Associate)

Practice Set 1

- Если компания открыла новый офис, и ей необходимо реализовать подключение с высокой пропускной способностью и низкой задержкой к нескольким VPC в нескольких регионах в рамках одной учетной записи. И чтобы каждый из VPC имел уникальный диапазон CIDR -
  То лучшим способом будет реализовать подключение AWS Direct Connect к ближайшему региону. Затем шлюз Direct Connect можно использовать для создания частных виртуальных интерфейсов (VIF) к каждому региону AWS. Шлюз Direct Connect обеспечивает группировку виртуальных частных шлюзов (VGW) и частных виртуальных интерфейсов (VIF), принадлежащих одной учетной записи AWS, и позволяет взаимодействовать с VPC в любом регионе AWS (кроме региона AWS China).
  Вы можете совместно использовать частный виртуальный интерфейс для взаимодействия с несколькими виртуальными частными облаками (VPC), сокращая количество необходимых BGP-сессий.
  - **Implement a Direct Connect connection to the closest AWS region**
  - **Create a Direct Connect gateway, and create private VIFs to each region**
- Вы можете включить кэширование API в Amazon API Gateway, чтобы кэшировать ответы вашей конечной точки. С помощью кэширования можно сократить количество обращений к конечной точке, а также улучшить задержку запросов к API.
  When you enable **caching for a stage**, API Gateway кэширует ответы от вашей конечной точки в течение определенного периода времени жизни (TTL) в секундах. Затем API Gateway отвечает на запрос, просматривая ответ конечной точки из кэша, вместо того чтобы делать запрос к вашей конечной точке. Значение TTL по умолчанию для кэширования API составляет 300 секунд. Максимальное значение TTL составляет 3600 секунд. TTL=0 означает, что кэширование отключено.
  - *An API cache is not enabled for a method, it is enabled for a **stage**.*
- Two types of events that can be logged in **CloudTrail**:

    - **Data events:** These events provide insight into the resource operations performed on or within a resource. These are also known as data plane operations
    - **Management events:** Management events provide insight into management operations that are performed on resources in your AWS account. These are also known as control plane operations. Management events can also include non-API events that occur in your account
- **Dedicated Instances** - это экземпляры Amazon EC2, которые работают в VPC на оборудовании, выделенном одному клиенту. Ваши выделенные экземпляры физически изолированы на аппаратном уровне от экземпляров, принадлежащих другим учетным записям AWS. Выделенные экземпляры позволяют автоматически размещать экземпляры, а тарификация производится за каждый экземпляр.
- Amazon Redshift - это полностью управляемая служба хранения данных петабайтного масштаба корпоративного уровня. В нем используется столбчатое хранение данных для повышения производительности сложных запросов.
  Команду **COPY можно использовать для параллельной загрузки данных с одного или нескольких удаленных узлов**, например, экземпляров Amazon EC2 или других компьютеров. COPY подключается к удаленным узлам с помощью SSH и выполняет команды на удаленных узлах для создания текстового вывода.
- Some facts about Amazon EBS encrypted volumes and snapshots:
    - **All EBS typessupport encryption** and all instance families now support encryption.
    - **Not all instance types support encryption**.
    - **There is no direct way to change the encryption state of a volume!**
    - Data in transit between an instance and an encrypted volume is also encrypted (data is encrypted in trans.
    - You can have encrypted an unencrypted EBS volumes attached to an instance at the same time.
    - Snapshots of encrypted volumes are encrypted automatically.
    - EBS volumes restored from encrypted snapshots are encrypted automatically.
    - EBS volumes created from encrypted snapshots are also encrypted.

- В MySQL аутентификация осуществляется с помощью AWSAuthenticationPlugin - плагина, предоставляемого AWS, который легко работает с IAM для аутентификации ваших пользователей IAM. Подключитесь к экземпляру БД и выполните оператор CREATE USER, как показано в следующем примере.
  CREATE USER jane_doe IDENTIFIED WITH AWSAuthenticationPlugin AS 'RDS';
  Пункт IDENTIFIED WITH позволяет MySQL использовать AWSAuthenticationPlugin для аутентификации учетной записи базы данных (jane_doe). Пункт AS 'RDS' относится к методу аутентификации, а указанная учетная запись базы данных должна иметь то же имя, что и пользователь или роль IAM. В этом примере и учетная запись базы данных, и пользователь или роль IAM имеют имя jane_doe.
- **AWS S**erverless **A**pplication **M**odel (**AWS SAM**) - это расширение AWS CloudFormation, которое используется для упаковки, тестирования и развертывания **бессерверных** приложений.
- С помощью **Amazon CloudFront** вы можете обеспечить безопасные сквозные соединения с исходными серверами, используя HTTPS. **Field-level encryption** - это дополнительный уровень безопасности, позволяющий защитить определенные данные в процессе обработки системы так, чтобы их могли видеть только определенные приложения.
  Field-level encryption - позволяет пользователям безопасно загружать конфиденциальную информацию на ваши веб-серверы. Конфиденциальная информация, предоставляемая пользователями, шифруется на границе, рядом с пользователем, и остается зашифрованной на протяжении всего стека приложений. Это шифрование гарантирует, что только те приложения, которым нужны данные и у которых есть полномочия для их расшифровки, смогут это сделать.
- How can a Solutions Architect **add a new instance store volume**?
  You can specify the instance store volumes for your instance only **when you launch an instance**. You can't attach instance store volumes to an instance after you've launched it.
- You can enable **access logs on the ALB** and *this will provide the information required including requester, IP, and request type*. Access logs are not enabled by default. You can optionally store and retain the log files on S3.

- **Connection draining** - по умолчанию включена и предоставляет период времени для чистого закрытия существующих соединений - is used to allow existing connections to close cleanly (Используется для обеспечения чистого закрытия существующих соединений).
- Максимальное время выполнения **Lambda** составляет **900 секунд**, а памяти может быть выделено **до 3008 МБ**
- Чтобы управлять объектами так, чтобы их хранение было экономически эффективным в течение всего жизненного цикла, настройте их жизненный цикл Amazon S3 Lifecycle. Конфигурация S3 Lifecycle - это набор правил, определяющих действия, которые Amazon S3 применяет к группе объектов. Существует два типа действий:
    - Действия при переходе - определяют, когда объекты переходят в другой класс хранения. Например, вы можете выбрать переход объектов в класс хранения S3 Standard-IA через 30 дней после их создания или архивирование объектов в класс хранения S3 Glacier через год после их создания.
    - Действия при истечении срока действия - определяет, когда истекает срок действия объектов. Amazon S3 удаляет объекты с истекшим сроком действия от вашего имени.
      Use an S3 lifecycle policy with object expiration configured to automatically remove objects that are more than 60 days old!
- **EBS optimized instances**- предоставляют выделенную емкость для ввода-вывода Amazon EBS. Оптимизированные экземпляры EBS предназначены для использования со всеми типами томов EBS. Help to increase storage performance!
- In general, when your object size **reaches 100 MB**, you should consider **using multipart uploads** instead of uploading the object in a single operation.
- You can control access to files and directories in EFS with POSIX-compliant user and group-level permissions. POSIX permissions allows you to restrict access from hosts by user and group. EFS Security Groups act as a firewall, and the rules you add define the traffic flow.
- Amazon **ElastiCache Redis** является кэшем базы данных **in-memory** и поддерживает **high availability** благодаря репликам и multi-AZ.
- **С помощью RedShift можно загружать данные из Amazon S3** и выполнять аналитические запросы. RedShift Spectrum может анализировать данные непосредственно в Amazon S3, но эта возможность не была представлена.

- Amazon **S3 Select** разработан, чтобы помочь **анализировать и обрабатывать данные** внутри объекта в ведрах Amazon S3 быстрее и дешевле. Он работает, предоставляя возможность получить подмножество данных из объекта в Amazon S3 с помощью простых выражений SQL.
- **AWS IoT Core** - это управляемый облачный сервис, позволяющий подключенным устройствам легко и безопасно взаимодействовать с облачными приложениями и другими устройствами. AWS IoT Core может поддерживать миллиарды устройств и триллионы сообщений, а также надежно и безопасно обрабатывать и направлять эти сообщения на конечные точки AWS и другие устройства.
- **Amazon DynamoDB auto scaling** использует службу AWS Application Auto Scaling для динамической регулировки **предоставленной пропускной способности** от вашего имени в ответ на фактический трафик. Это наиболее эффективное и экономичное решение для оптимизации затрат.
  DynamoDB DAX - is an in-memory cache that **increases the performance** of DynamoDB.
- An Amazon **S**imple **Q**ueue **S**ervice (**SQS**) can be used to offload and **decouple the long-running requests**. They can then be processed asynchronously by separate EC2 instances. This is the best way to reduce the overall latency introduced by the long-running API call.
- Default state of the **new Network ACL** when we created it
  - There is a default outbound rule denying all traffic
  - There is a default inbound rule denying all traffic
- Некоторые приложения, такие как обновление медиакаталога, требуют высокой частоты чтения и постоянной пропускной способности. Для таких приложений заказчики часто дополняют S3 кэшем в памяти, например Amazon ElastiCache for Redis, чтобы снизить затраты на поиск данных в S3 и повысить производительность.
  ElastiCache for Redis - это полностью управляемое хранилище данных in-memory, обеспечивающее субмиллисекундную задержку при высокой пропускной способности. ElastiCache for Redis дополняет S3 следующим образом:

    - Redis хранит данные в памяти, поэтому обеспечивает субмиллисекундную задержку и поддерживает невероятно высокие запросы в секунду.
      Он поддерживает операции на основе ключ/значение, которые хорошо отображаются на операции S3 (например, GET/SET => GET/PUT), что позволяет легко писать код как для S3, так и для ElastiCache.
    - Его можно реализовать как кэш на стороне приложения. Это позволяет вам использовать S3 в качестве постоянного хранилища и получать преимущества от его долговечности, доступности и низкой стоимости. Ваши приложения решают, какие объекты кэшировать, когда их кэшировать и как их кэшировать.
      Если медиа-каталог получает обновления из S3, то производительность между этими компонентами может нуждаться в улучшении. Для этого используйте ElastiCache для кэширования содержимого, что значительно увеличит производительность.
- In a **default VPC** instances will be assigned a **public** and **private** DNS hostname
- In a **non-default VPC** instances will be assigned a private but not a public DNS hostname
- You can use **VPC Flow Logs**to capture detailed information about the traffic going to and from your Elastic Load Balancer. Create a flow log for each network interface for your load balancer. There is one network interface per load balancer subnet.
- The NLB provides high throughput and extremely low latencies and is designed to handle traffic as it grows and can load balance millions of requests/second. **NLB also supports load balancing to multiple ports on an instance**.
- You can restore a DB instance to a specific point in time, creating a new DB instance. When you restore a DB instance to a point in time, **the default DB security group is applied to the new DB instance**. If you need custom DB security groups applied to your DB instance, you must apply them explicitly using the AWS Management Console, the AWS CLI modify-db-instance command, or the Amazon RDS API ModifyDBInstance operation after the DB instance is available.
  Restored DBs will always be a new RDS instance with a new DNS endpoint and you can **restore** up to the **last 5 minutes**.
- **Multi-site** DR strategy = active-active configuration!

Practice Set 2

- The solution must use NFS file shares to access the migrated data without code modification. This means you can use either Amazon EFS or AWS Storage Gateway – File Gateway. Both of these can be mounted using NFS from on-premises applications.
  However, EFS is the wrong answer as the solution asks to maximize availability and durability. The File Gateway backs off of Amazon S3 which has much higher availability and durability than EFS which is why it is the best solution for this scenario.
- DynamoDB offers consistent single-digit millisecond latency. However, DynamoDB + DAX further increases performance with response times in microseconds for millions of requests per second for read-heavy workloads.
- If you do not want to manage EC2 instances you must use the AWS Fargate launch type which is a serverless infrastructure managed by AWS. Fargate only supports container images hosted on Elastic Container Registry (ECR) or Docker Hub.
- Amazon EFS is a fully-managed service that makes it easy to set up, scale, and cost-optimize file storage in the Amazon Cloud. EFS file systems are accessible to Amazon EC2 instances via a file system interface (using standard operating system file I/O APIs) and support full file system access semantics (such as strong consistency and file locking).
- The most cost-effective option is to migrate the website to an Amazon S3 bucket and configure that bucket for static website hosting.
- Amazon S3 is great solution for storing objects such as this. You only pay for what you use and don't need to worry about scaling as it will scale as much as you need it to. Using Amazon Athena to analyze the data works well as it is a serverless service so it will be very cost-effective for use cases where the analysis is only happening infrequently. You can also configure Amazon S3 to expire the objects after 30 days.
- The ELB Application Load Balancer can route traffic based on data included in the request including the host name portion of the URL as well as the path in the URL.
- An application running on Amazon EC2 needs to regularly download large objects from Amazon S3. How can performance be optimized for high-throughput use cases?
  - Issue parallel requests and use byte-range fetches

- A mobile client requires data from several application-layer services to populate its user interface. What can the application team use to decouple the client interface from the underlying services behind them?
    - Amazon API Gateway decouples the client application from the back-end application-layer services by providing a single endpoint for API requests.
- Using Amazon CloudFront as the front-end provides the option to specify a custom message instead of the default message.
- ElastiCache can be deployed in the U.S east region to provide high-speed access to the content.
- AWS Global Accelerator is a service that improves the availability and performance of applications with local or global users. You can configure the ALB as a target and Global Accelerator will automatically route users to the closest point of presence.
- You can only apply one IAM role to a Task Definition so you must create a separate Task Definition. A Task Definition is required to run Docker containers in Amazon ECS and you can specify the IAM role (Task Role) that the task should use for permissions.
- That restrict access so that consumers cannot connect to other instances in the VPC the best solution is to use PrivateLink to create an endpoint for the application. The endpoint type will be an interface endpoint and it uses an NLB in the shared services VPC.
- You can encrypt your Amazon RDS instances and snapshots at rest by enabling the encryption option for your Amazon RDS DB instance when you create it. However, you cannot encrypt an existing DB, you need to create a snapshot, copy it, encrypt the copy, then build an encrypted DB from the snapshot.
- To publish messages to Amazon SNS topics from an Amazon VPC, create an interface VPC endpoint. Then, you can publish messages to SNS topics while keeping the traffic within the network that you manage with the VPC. This is the most secure option as traffic does not need to traverse the Internet.
    - AWS PrivateLink - is the correct answer.
- EC2 Instance Stores are high-speed ephemeral storage that is physically attached to the EC2 instance. The i3.large instance type comes with a single 475GB NVMe SSD instance store so it would be a good way to lower cost and improve performance by using the attached instance store. As the files are temporary, it can be assumed that ephemeral storage (which means the data is lost when the instance is stopped) is sufficient.

- The Amazon Elastic File System (EFS) can assist with increasing the resilience of the application by providing a shared file system that can be mounted by multiple EC2 instances from multiple availability zones.
- AWS Lambda can likewise be used for the back-end processing reducing cost by utilizing a pay for what you use serverless service.
- RedShift is a columnar data warehouse DB that is ideal for running long complex queries. RedShift can also improve performance for repeat queries by caching the result and returning the cached result when queries are re-run. Dashboard, visualization, and business intelligence (BI) tools that execute repeat queries see a significant boost in performance due to result caching.
- You have created a file system using Amazon Elastic File System (EFS) which will hold home directories for users. What else needs to be done to enable users to save files to the EFS file system?
    - Create a subdirectory for each user and grant read-write-execute permissions to the users. Then mount the subdirectory to the users' home directory.
    - После создания файловой системы по умолчанию только пользователь root (UID 0) имеет права на чтение-запись-исполнение. Чтобы другие пользователи могли изменять файловую систему, пользователь root должен явно предоставить им доступ.
      Одним из распространенных вариантов использования является создание подкаталога "с правом записи" в корне файловой системы для каждого пользователя, создаваемого на экземпляре EC2, и монтирование его в домашний каталог пользователя. Все файлы и подкаталоги, которые пользователь создает в своем домашнем каталоге, затем создаются в файловой системе Amazon EFS
- Amazon FSx for Windows File Server provides fully managed, highly reliable file storage that is accessible over the industry-standard Server Message Block (SMB) protocol. It is built on Windows Server, delivering a wide range of administrative features such as user quotas, end-user file restore, and Microsoft Active Directory (AD) integration. It offers single-AZ and multi-AZ deployment options, fully managed backups, and encryption of data at rest and in transit.
- Which set of actions will improve website performance for users worldwide?
    - Create an Amazon CloudFront distribution and configure the ALB as an origin. Then update the Amazon Route 53 record to point to the CloudFront distribution.

- IAM roles for ECS tasks enabled you to secure your infrastructure by assigning an IAM role directly to the ECS task rather than to the EC2 container instance. This means you can have one task that uses a specific IAM role for access to S3 and one task that uses an IAM role to access DynamoDB.
- EBS volumes cannot be shared across AZs.
- Amazon FSx for Lustre provides a high-performance file system optimized for fast processing of workloads such as machine learning, high-performance computing (HPC), video processing, financial modeling, and electronic design automation (EDA).
- These workloads commonly require data to be presented via a fast and scalable file system interface, and typically have data sets stored on long-term data stores like Amazon S3.
- The microservices must be decoupled so they can scale independently. An Amazon SQS queue will enable microservice 1 to add messages to the queue. Microservice 2 can then pick up the messages and process them. This ensures that if there's a spike in traffic on the frontend, messages do not get lost due to the backend process not being ready to process them.
- The AWS Web Application Firewall (WAF) is available on the Application Load Balancer (ALB). You can use AWS WAF directly on Application Load Balancers (both internal and external) in a VPC, to protect your websites and web services from cross-site scripting (XSS) attacks.
- **At EC2 level we cannot use WAF** because AWS WAF is used to control how an Amazon CloudFront distribution, an Amazon API Gateway API, or an Application Load Balancer responds to web requests.
- Amazon EFS provides file storage in the AWS Cloud. With Amazon EFS, you can create a file system, mount the file system on an Amazon EC2 instance, and then read and write data to and from your file system. Can be shared between multiple EC2 instances.
- **Amazon RDS automatically performs a failover in the event of any of the following:**
    - Loss of availability in primary Availability Zone
    - Loss of network connectivity to primary
    - Compute unit failure on primary
    - Storage failure on primary

- Uploading using a pre-signed URL allows you to upload the object without having any AWS security credentials/permissions. Pre-signed URLs can be generated programmatically and anyone who receives a valid pre-signed URL can then programmatically upload an object. This solution bypasses the web server avoiding any performance bottlenecks.
- AWS Global Accelerator is a service in which you create accelerators to improve availability and performance of your applications for local and global users.
- The key requirements are that the compute service should be suitable for a workload that can range quite broadly in demand from no requests to large bursts of traffic.
  AWS Lambda is an ideal solution as you pay only when requests are made and it can easily scale to accommodate the large bursts in traffic. Lambda works well with both API Gateway and Amazon RDS.
- You can enable API caching in Amazon API Gateway to cache your endpoint's responses. With caching, you can reduce the number of calls made to your endpoint and also improve the latency of requests to your API.
  Когда вы включаете кэширование для этапа, API Gateway кэширует ответы от вашей конечной точки в течение указанного периода времени жизни (TTL), в секундах. Затем API Gateway отвечает на запрос, просматривая ответ конечной точки из кэша, вместо того чтобы делать запрос к вашей конечной точке. Значение TTL по умолчанию для кэширования API составляет 300 секунд. Максимальное значение TTL составляет 3600 секунд. TTL=0 означает, что кэширование отключено.
-

Practice Set 3

- Type Retrievel:
  - Expedited — Expedited retrievals allow you to quickly access your data when occasional urgent requests for a subset of archives are required. For all but the largest archives (250 MB+), data accessed using Expedited retrievals are typically made available within 1–5 minutes. Provisioned Capacity ensures that retrieval capacity for Expedited retrievals is available when you need it.

    - Standard — Standard retrievals allow you to access any of your archives within several hours. Standard retrievals typically complete within 3–5 hours. This is the default option for retrieval requests that do not specify the retrieval option.
    - Bulk — Bulk retrievals are S3 Glacier's lowest-cost retrieval option, which you can use to retrieve large amounts, even petabytes, of data inexpensively in a day. Bulk retrievals typically complete within 5–12 hours.
- A VPG is used to setup an AWS VPN which you can use in combination with Direct Connect to encrypt all data that traverses the Direct Connect link.
- With target tracking scaling policies, you select a scaling metric and set a target value.
- Amazon Aurora Global Database is designed for globally distributed applications, allowing a single Amazon Aurora database to span multiple AWS regions. It replicates your data with no impact on database performance, enables fast local reads with low latency in each region, and provides disaster recovery from region-wide outages.
  Aurora Global Database uses storage-based replication with typical latency of less than 1 second, using dedicated infrastructure that leaves your database fully available to serve application workloads. In the unlikely event of a regional degradation or outage, one of the secondary regions can be promoted to full read/write capabilities in less than 1 minute.
- Amazon Simple Workflow Service (SWF) is a web service that makes it easy to coordinate work across distributed application components.
- The following are a few reasons why an instance might immediately terminate:
  – You've reached your EBS volume limit.
  – An EBS snapshot is corrupt.
  – The root EBS volume is encrypted and you do not have permissions to access the KMS key for decryption.
  – The instance store-backed AMI that you used to launch the instance is missing a required part (an image.part.xx file).
- You can use a Lambda function to process Amazon Simple Notification Service notifications. Amazon SNS supports Lambda functions as a target for messages sent to a topic. This solution decouples the Amazon EC2 application from Lambda and ensures the Lambda function is invoked.
- Multi-AZ RDS creates a replica in another AZ and synchronously replicates to it (DR only).

- A failover may be triggered in the following circumstances:
  – Loss of primary AZ or primary DB instance failure
  – Loss of network connectivity on primary
  – Compute (EC2) unit failure on primary
  – Storage (EBS) unit failure on primary
  – The primary DB instance is changed
  – Patching of the OS on the primary DB instance
  – Manual failover (reboot with failover selected on primary)

Practice Set 4

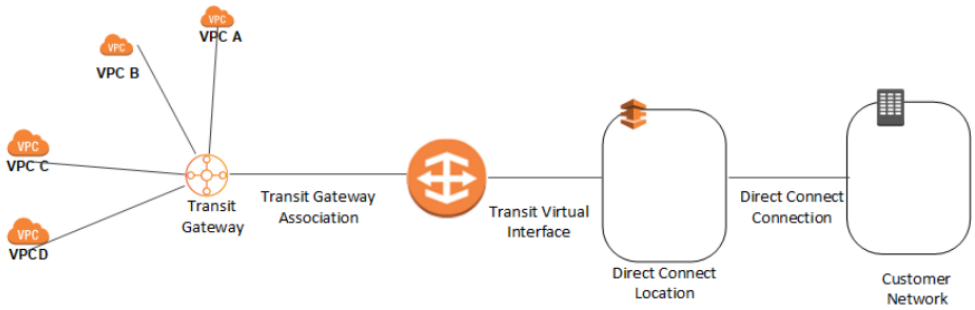
- To allow read access to the S3 video assets from the public-facing web application, you can **add a bucket policy that allows s3:GetObject permission with a condition, using the aws:referer key, that the get request must originate from specific webpages**. This is a good answer as it fully satisfies the objective of ensuring the that EC2 instance can access the videos but direct access to the videos from other sources is prevented.
- Многоузловые параллельные задания AWS Batch позволяют выполнять отдельные задания, охватывающие несколько экземпляров Amazon EC2. С помощью многоузловых параллельных заданий AWS Batch вы можете запускать крупномасштабные, тесно связанные, высокопроизводительные вычислительные приложения и распределенное обучение моделей на GPU без необходимости запуска, настройки и управления ресурсами Amazon EC2 напрямую.
  Многоузловое параллельное задание AWS Batch совместимо с любым фреймворком, поддерживающим межузловое взаимодействие на основе IP, например Apache MXNet, TensorFlow, Caffe2 или Message Passing Interface (MPI).
- Токены аутентификации Redis позволяют Redis запрашивать токен (пароль), прежде чем разрешить клиентам выполнять команды, что повышает безопасность данных.
  Вы можете потребовать, чтобы пользователи вводили токен на защищенном токенами сервере Redis. Для этого при создании группы репликации или кластера включите параметр -auth-token (API: AuthToken) с правильным маркером. Также включайте его во все последующие команды для группы или кластера репликации.

- **Вы можете приостановить, а затем возобновить один или несколько процессов масштабирования для группы Auto Scaling.** Это может быть полезно, когда вы хотите исследовать проблему конфигурации или другую проблему с вашим веб-приложением, а затем внести изменения в приложение, не вызывая процессы масштабирования. **Вы можете вручную переместить экземпляр из ASG и перевести его в состояние ожидания**.
    - Экземпляры в состоянии ожидания по-прежнему управляются системой Auto Scaling, тарифицируются в обычном режиме и не учитываются как доступные экземпляры EC2 для использования рабочей нагрузкой/приложением. Автомасштабирование не выполняет проверку работоспособности экземпляров в состоянии ожидания. Состояние ожидания можно использовать для выполнения обновлений/изменений/устранения неполадок и т. д. без проверки состояния здоровья или запуска запасных экземпляров.

- When сценарий требует использования учетных данных для аутентификации в MySQL - Your credentials должны надежно храниться вне кода функции Lambda. **Systems Manager Parameter Store** обеспечивает безопасное, иерархическое хранение для управления конфигурационными данными и секретами.
- Без включенной межзональной балансировки нагрузки NLB будет распределять трафик 50/50 между AZ. Поскольку количество экземпляров в двух AZ нечетное, некоторые экземпляры не будут получать трафик. Поэтому включение межзональной балансировки нагрузки обеспечит равномерное распределение трафика между доступными экземплярами во всех AZ.
- Amazon DynamoDB может дросселировать запросы, которые превышают установленную пропускную способность для таблицы. **When requests will be throttled, and fail with an HTTP 400 code (Bad Request) and a ProvisionedThroughputExceededException**
  При использовании модели ценообразования с предоставлением емкости DynamoDB не масштабируется автоматически. DynamoDB может автоматически масштабироваться при использовании нового режима предоставления емкости по требованию, однако это не настроено для данной базы данных.

- A CloudWatch Events rule can be used to set up automatic email notifications for Medium to High Severity findings to the email address of your choice. You simply create an Amazon SNS topic and then associate it with an Amazon CloudWatch events rule.
- Вы можете управлять одним соединением для нескольких VPC или VPN, которые находятся в одном регионе, связав шлюз Direct Connect с транзитным шлюзом. Решение включает следующие компоненты:
  - Транзитный шлюз, имеющий вложения VPC.
  - Шлюз прямого подключения.
  - Ассоциация между шлюзом прямого подключения и транзитным шлюзом.
  - Транзитный виртуальный интерфейс, подключенный к шлюзу Direct Connect.
- Если несколько VPC используются компанией в одном регионе. Компания имеет два подключения AWS Direct Connect к двум отдельным офисам компании и хочет использовать их совместно со всеми тремя VPC - ТО Используйте **the Direct Connect gateway and Associate it to a transit gateway.**

  Вы связываете шлюз AWS Direct Connect с одним из следующих шлюзов:
  - Транзитный шлюз при наличии нескольких VPC **в одном регионе**.
  - Виртуальный частный шлюз. (Если нужно связать VPC в нескольких регионах или аккаунтах - ТО Используйте Associate the Direct Connect gateway to a virtual private gateway in each VPC)



- Which feature of IAM allows direct access to the IAM web service using HTTPS to call service actions?
  AWS рекомендует использовать AWS SDK для выполнения программных вызовов API к IAM. Однако вы также можете использовать IAM **Query API** для прямых вызовов веб-службы IAM. Для аутентификации при использовании API Query необходимо использовать идентификатор ключа доступа и секретный ключ доступа.

- Крупные миграции данных с помощью AWS DMS могут включать многие терабайты информации. Этот процесс может быть громоздким из-за ограничений пропускной способности сети или просто огромного объема данных.
  When you're using an Edge device, the data migration process has the following stages:
  1. You **use the AWS Schema Conversion Tool (AWS SCT)** to extract the data locally and move it to an Edge device.
  2. You ship the Edge device or devices back to AWS.
  3. After AWS receives your shipment, the Edge device automatically loads its data into an Amazon S3 bucket.
  4. AWS DMS takes the files and migrates the data to the target data store. If you are using change data capture (CDC), those updates are written to the Amazon S3 bucket and then applied to the target data store.
- Наиболее экономически эффективное решение для обеспечения резервного копирования соединения Direct Connect - это соединение Direct Connect, и IPSec VPN. Они активны и рекламируются с помощью протокола Border Gateway Protocol (BGP). **Implement an IPSec VPN connection and use the same BGP prefix** - for backups in Direct Conect!
- **AWS Global Accelerator** использует статические IP-адреса в качестве фиксированных точек входа для вашего приложения. Вы можете перенести до двух диапазонов адресов /24 IPv4 и выбрать, какие IP-адреса /32 использовать при создании ускорителя.
  Это решение гарантирует, что компания сможет продолжать использовать те же IP-адреса, и она сможет направлять трафик на конечную точку приложения в регионе AWS, ближайшем к конечному пользователю. Трафик передается по глобальной сети AWS для обеспечения стабильной производительности.

- SSD, General Purpose (GP2) provides 3 IOPS per GB up to 16,000 IOPS. Volume size is 1 GB to 16 TB.
  More information on the volume types:
  - SSD, General Purpose (**GP2**) provides **3 IOPS per GB** up to 16,000 IOPS. Volume size is 1 GB to 16 TB.
  - Provisioned IOPS (Io1) provides the **IOPS you assign** up to 50 IOPS per GiB and up to 64,000 IOPS per volume. Volume size is 4 GB to 16TB.
- The key requirement is to **limit the number of requests per second** that hit the application. This can only be done by implementing **throttling rules on the API Gateway.** *Throttling enables you to throttle the number of requests to your API which in turn means less traffic will be forwarded to your application server*.
- **File gateway** provides a virtual on-premises file server, which enables you to store and retrieve files as objects in Amazon S3. **It can be used for on-premises applications**, and for Amazon EC2-resident applications that need file storage in S3 for object based workloads. Used for flat files only, stored directly on S3. File gateway offers SMB or NFS-based access to data in Amazon S3 **with local caching**. You could mount EFS over a VPN but *it would not provide you a local cache of the data.*
- С помощью AWS Transit Gateway, поддерживающего транзитивную маршрутизацию, можно построить топологию "хаб и спица". Это упрощает топологию сети и добавляет дополнительные возможности по сравнению с пирингом VPC. AWS Resource Access Manager можно использовать для совместного использования соединения **с другими учетными записями AWS.**
  For Example:
  A company has several AWS accounts each with multiple Amazon VPCs. The company must establish routing between all private subnets.

- Агент контейнеров ECS включен в оптимизированный AMI Amazon ECS, а также может быть установлен на любом экземпляре EC2, поддерживающем спецификацию ECS (поддерживается только на экземплярах EC2). Поэтому вам не нужно проверять, установлен ли агент.

- DynamoDB best practices include:
  - Keep item sizes small.

    - If you are storing serial data in DynamoDB that will require actions based on data/time use separate tables for days, weeks, months.
    - **Store more frequently and less frequently accessed data in separate tables**.
    - If possible compress larger attribute values.
    - **Store objects larger than 400KB in S3 and use pointers (S3 Object ID) in DynamoDB**.
- Amazon SNS supports notifications over multiple **transport protocols**:
    - **HTTP/HTTPS** – subscribers specify a URL as part of the subscription registration.
    - **Email/Email-JSON** – messages are sent to registered addresses as email (text-based or JSON-object).
    - SQS – users can specify an SQS standard queue as the endpoint.
    - SMS – messages are sent to registered phone numbers as SMS text messages.
- Необходимо убедиться, что установленный агент запущен и что профиль экземпляра IAM имеет необходимые разрешения.
  Шаги по устранению неполадок для контейнеров включают:
    - Убедитесь, что демон Docker запущен на экземпляре контейнера.
    - Убедитесь, что демон Docker Container запущен на экземпляре контейнера.
    - Убедитесь, что агент контейнера запущен на экземпляре контейнера.
    - Убедитесь, что профиль экземпляра IAM имеет необходимые разрешения.

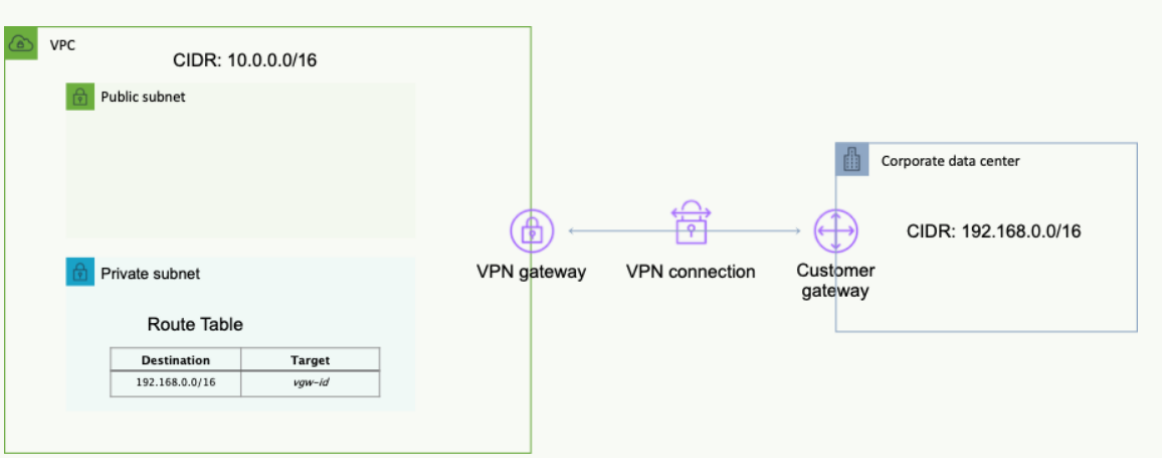
- Транзитное шифрование Amazon ElastiCache - это дополнительная функция, которая позволяет повысить безопасность ваших данных в наиболее уязвимых местах - когда они находятся в пути из одного места в другое.
  Enable in-transit encryption!
  ElastiCache for Redis at-rest encryption is an optional feature to increase data security by encrypting on-disk data. This does not encrypt the data in-transit when it is being replicated between nodes.

- Чтобы ваша функция Lambda могла получить доступ к ресурсам внутри вашего частного VPC, вы должны предоставить дополнительную информацию о конфигурации, специфичную для VPC, которая включает **VPC Subnet IDs** и **VPC Security Group IDs**.
  AWS Lambda использует эту информацию для настройки эластичных сетевых интерфейсов (ENI), которые позволяют вашей функции работать.

- **Per-client throttling limits** are applied to clients that use API keys associated with your usage policy as client identifier. This can be applied to the single customer that is issuing excessive API requests. *This is the best option to ensure that only one customer is affected*.

- CloudFront dictribution - is a content delivery network (CDN) that caches content to improve performance.
  AWS CloudFormation предоставляет два метода обновления стеков: прямое обновление или создание и выполнение наборов изменений. При прямом обновлении стека вы отправляете изменения, и AWS CloudFormation немедленно развертывает их.
  Используйте прямое обновление, когда вы хотите быстро развернуть свои обновления. С помощью наборов изменений вы можете предварительно просмотреть изменения, которые AWS CloudFormation внесет в ваш стек, а затем решить, применять ли эти изменения.

- **Amazon FSx for Windows File Server** предоставляет полностью управляемое, высоконадежное и масштабируемое файловое хранилище, доступное по стандартному протоколу Server Message Block (**SMB**). Это наиболее подходящее место назначения для данного сценария использования.
  AWS DataSync можно использовать для перемещения больших объемов данных в режиме онлайн между местным хранилищем и Amazon S3, Amazon EFS или Amazon FSx for Windows File Server. В качестве исходного хранилища данных могут выступать файловые серверы Server Message Block (SMB).
- Для узлов приложения HPC рекомендуется использовать либо расширенную сеть, либо адаптер Elastic Fabric Adapter (EFA). Это поможет снизить задержки. Кроме того, группа размещения кластеров объединяет экземпляры близко друг к другу в зоне доступности.

- Использование группы размещения кластеров позволяет рабочим нагрузкам достичь производительности сети с низкой задержкой, необходимой для тесно связанной связи между узлами, характерной для приложений HPC.
- 

CORRECT: "Configure a Virtual Private Gateway" is the correct answer.
- **A virtual private gateway** - это логическая, полностью избыточная функция распределенной пограничной маршрутизации, расположенная на границе вашего VPC. Вы должны создать VPG в своем VPC, прежде чем сможете создать AWS Managed site-to-site VPN соединение. Другим концом соединения является шлюз клиента, который должен быть установлен на стороне клиента.
-

- Amazon SNS supports notifications over multiple transport protocols:
  – HTTP/HTTPS – subscribers specify a URL as part of the subscription registration.
  – Email/Email-JSON – messages are sent to registered addresses as email (text-based or JSON-object).
  – SQS – users can specify an SQS standard queue as the endpoint.
  – SMS – messages are sent to registered phone numbers as SMS text messages.

Practice Set 5

- Scheduled Instances are a good choice for workloads that do not run continuously but do run on a regular schedule. This is ideal for the development environment.

- Потоки данных Amazon Kinesis собирают и обрабатывают данные в режиме реального времени. Поток данных Kinesis представляет собой набор shard. Каждый shard содержит последовательность записей данных. Каждая запись данных имеет порядковый номер, который присваивается Kinesis Data Streams. Shard - это уникально идентифицированная последовательность записей данных в потоке.
  Partition key используется для группировки данных по разделам в потоке. Kinesis Data Streams разделяет записи данных, принадлежащие потоку, на несколько хранилищ. Он использует partition key, связанный с каждой записью данных, чтобы определить, к какому хранилищу принадлежит данная запись данных.
- Новая версия AWS Web Application Firewall была выпущена в ноябре 2019 года. В AWS WAF classic вы создаете "условия соответствия IP", тогда как в AWS WAF (новая версия) вы создаете "утверждения соответствия набора IP". Обратите внимание на формулировку на экзамене.
  Условие соответствия IP / утверждение соответствия набора IP проверяет IP-адрес происхождения веб-запроса на соответствие набору IP-адресов и диапазонов адресов. Используйте его для разрешения или блокирования веб-запросов на основе IP-адресов, с которых они исходят.
- **ALB поддерживает маршрутизацию как на основе путей** (например, /images или /orders), так и на основе хостов (например, example.com).
- AWS Lambda has a maximum execution time of 900 seconds (15 minutes).
- Архитектура уже обладает высокой устойчивостью, но при внезапном увеличении количества запросов может снизиться производительность. Для разрешения этой ситуации можно использовать Amazon Aurora Read Replicas для обслуживания трафика чтения, что разгрузит запросы от основной базы данных. На фронтенде перед ALB можно разместить дистрибутив Amazon CloudFront, который будет кэшировать содержимое для повышения производительности, а также разгрузит запросы с бэкенда.
- Существует два различных типа конечных точек VPC: конечная точка интерфейса и конечная точка шлюза. При использовании интерфейсной конечной точки вы используете ENI в VPC. При использовании конечной точки шлюза вы настраиваете таблицу маршрутов так, чтобы она указывала на конечную точку. Amazon S3 и DynamoDB используют конечные точки шлюза. Это решение означает, что весь трафик будет проходить через конечную точку VPC прямо к DynamoDB, используя частные IP-адреса.

- AWS Global Accelerator использует огромную, свободную от перегрузок глобальную сеть AWS для маршрутизации трафика TCP и UDP к здоровой конечной точке приложения в ближайшем к пользователю регионе AWS. Это означает, что он интеллектуально направляет трафик в ближайшую точку присутствия (уменьшая задержку). Бесперебойная обкатка гарантирована, так как AWS Global Accelerator использует anycast IP-адрес, что означает, что IP не меняется при обкатке между регионами, поэтому нет проблем с неправильными записями в клиентских кэшах, которые должны истечь.
- AWS Glue is a fully managed extract, transform, and load (ETL) service that makes it easy for customers to prepare and load their data for analytics.
  AWS Glue - *это полностью управляемый сервис извлечения, преобразования и загрузки (ETL), который упрощает клиентам подготовку и загрузку данных для аналитики.*
  С помощью этого решения уведомления о событиях S3, запускающие функцию Lambda, являются полностью бессерверными и экономически эффективными, а AWS Glue может запускать задания ETL, которые преобразуют эти данные и загружают их в хранилище данных, такое как S3.
- Amazon DynamoDB - это полностью управляемая служба баз данных NoSQL, которая обеспечивает быструю и предсказуемую производительность с плавным масштабированием. Кнопочное масштабирование означает, что вы можете масштабировать БД в любое время без простоя. DynamoDB обеспечивает низкую задержку при чтении и записи.

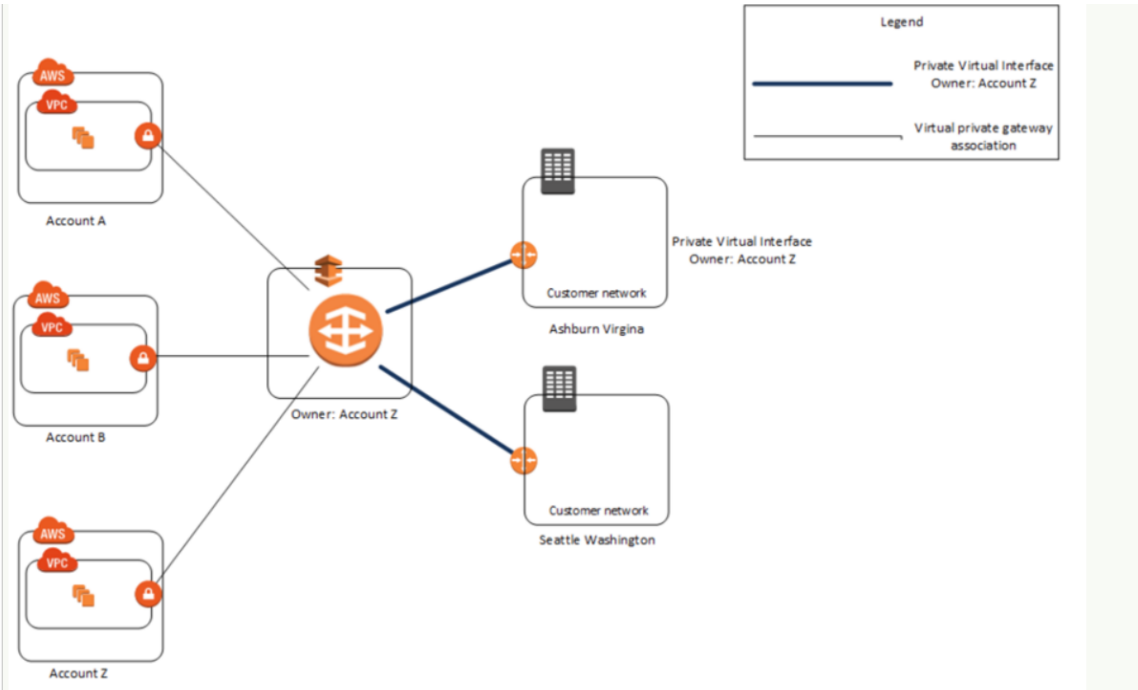
Feature	Memcached	Redis (cluster mode disabled)	Redis (cluster mode enabled)
Data persistence	No	Yes	Yes
Data types	Simple	Complex	Complex
Data partitioning	Yes	No	Yes
Encryption	No	Yes	Yes
High availability (replication)	No	Yes	Yes
Multi-AZ	Yes, place nodes in multiple AZs. No failover or replication	Yes, with auto-failover. Uses read replicas (0-5 per shard)	Yes, with auto-failover. Uses read replicas (0-5 per shard)
Scaling	Up (node type); out (add nodes)	Single shard (can add replicas)	Add shards
Multithreaded	Yes	No	No
Backup and restore	No (and no snapshots)	Yes, automatic and manual snapshots	Yes, automatic and manual snapshots
- Масштабирование по расписанию позволяет задать собственный график масштабирования для предсказуемых изменений нагрузки. Чтобы настроить группу Auto Scaling на масштабирование по расписанию, вы создаете запланированное действие. Это идеально подходит для ситуаций, когда вы знаете, когда и на какой срок вам понадобится дополнительная мощность.
- **The cooldown period**- это настраиваемый параметр для группы автоматического масштабирования, который помогает убедиться, что она не запускает и не завершает дополнительные экземпляры до того, как предыдущая активность масштабирования вступит в силу, так что это поможет. После того как группа Auto Scaling динамически масштабируется с помощью простой политики масштабирования, она ожидает завершения периода охлаждения, прежде чем возобновить действия по масштабированию.
  Период оценки тревог CloudWatch - это количество самых последних точек данных, которые необходимо оценить при определении состояния тревоги. Это поможет, так как вы можете увеличить количество точек данных, необходимых для подачи сигнала тревоги.

- An egress-only Internet gateway - это горизонтально масштабируемый, избыточный и высокодоступный компонент VPC, который позволяет исходящую связь по IPv6 от экземпляров в вашем VPC к Интернету и не позволяет Интернету инициировать соединение IPv6 с вашими экземплярами.
- Используйте Amazon CloudFront для обслуживания приложения и запрета доступа в заблокированные страны - is the EASIEST method!
- Kinesis Data Firehose is the easiest way to load streaming data into data stores and analytics tools.
- **EFS:**
  - **Storage class** – EFS One Zone or EFS Standard
  - **Performance mode** – General Purpose or Max I/O
  - **Throughput mode** – Bursting or Provisioned
- When you encrypt your data, your data is protected, but you have to protect your encryption key - **AWS KMS API**
- Глобальные таблицы Amazon DynamoDB представляют собой полностью управляемое решение для развертывания многорегиональной, многомастерной базы данных. Это единственное представленное решение, обеспечивающее активно-активную конфигурацию, в которой чтение и запись могут осуществляться в нескольких регионах с полной двунаправленной синхронизацией.
- Вы можете использовать функцию Lambda для обработки уведомлений Amazon Simple Notification Service. Amazon SNS поддерживает функции Lambda в качестве цели для сообщений, отправляемых в тему. Это решение отделяет приложение Amazon EC2 от Lambda и обеспечивает вызов функции Lambda.
- Веб-сайт для нового приложения получал около 50 000 запросов каждую секунду, и компания хочет использовать несколько приложений для анализа навигационных моделей пользователей на своем сайте, чтобы персонализировать пользовательский опыт.
- Что может использовать архитектор решений для сбора кликов по страницам сайта и их последовательной обработки для каждого пользователя?
- Это хороший вариант использования потоков данных Amazon Kinesis, поскольку они способны масштабироваться до необходимой нагрузки, позволять нескольким приложениям получать доступ к записям и обрабатывать их последовательно.

- **Amazon Kinesis Data Streams** обеспечивает обработку потоковых больших данных в режиме реального времени. Он обеспечивает упорядочивание записей, а также возможность чтения и/или воспроизведения записей в том же порядке для нескольких приложений Amazon Kinesis.
  Потоки Amazon Kinesis позволяют обрабатывать до 1 Мб данных в секунду или до 1 000 записей в секунду при записи на один шард. Количество шардов не ограничено, поэтому вы можете легко масштабировать Kinesis Streams до 50 000 в секунду.
  Клиентская библиотека Amazon Kinesis Client Library (KCL) доставляет все записи для заданного ключа раздела в один и тот же процессор записей, что упрощает создание нескольких приложений, читающих из одного потока данных Amazon Kinesis.

Practice Set 6

-



CORRECT: “Associate the Direct Connect gateway to a virtual private gateway in account A and B” is the correct answer.

- Access Logs могут быть включены на ALB и настроены на хранение данных в ведре S3. Amazon EMR - это веб-служба, которая позволяет предприятиям, исследователям, аналитикам данных и разработчикам легко и экономически эффективно обрабатывать огромные объемы данных. В EMR используется размещенная платформа Hadoop, работающая на Amazon EC2 и Amazon S3.
- Совет по сдаче экзамена:
  Для предотвращения чтения старых "несвежих" данных = используйте сильно последовательное чтение
  Объяснение:
  Когда вы запрашиваете сильно последовательное чтение, DynamoDB возвращает ответ с самыми актуальными данными, отражающими обновления от всех предыдущих операций записи, которые были успешными.
  Однако такая последовательность имеет некоторые недостатки:
  - A strongly consistent read может быть недоступно при задержке или сбое в сети. В этом случае DynamoDB может выдать ошибку сервера (HTTP 500).
  - A strongly consistent read может иметь более высокую задержку, чем окончательно последовательное чтение.
  - A strongly consistent read не поддерживаются в глобальных вторичных индексах.
  - A strongly consistent read используют большую пропускную способность, чем конечно-последовательные чтения.
- Exam Tip
  the most secure architecture is to put only the ELB in the public subnet and all other resources must be in a private subnet
- Configure AWS Organizations. Create an organizational unit (OU) and place all AWS accounts into the OU. Apply a service control policy (SCP) to the OU that denies the use of certain services.
- ELB connection draining can be used to stop sending requests to instances that are de-registering or unhealthy.
- Lambda может обрабатывать одновременные выполнения 1 000 в секунду на регион, и этот предел может быть увеличен.
- **ELB connection draining** can be used to stop sending requests to instances that are de-registering or unhealthy.

- You must also specify an **Elastic IP address to associate with the NAT gateway** when you create it. The Elastic IP address cannot be changed after you associate it with the NAT Gateway.
- **Key Word “Scratch Data” = Use instance store.**
- the options “The time taken to run bootstrap scripts. & The size of embedded application code in the AMI” control warm up time for your instances.
- Чтобы ограничить доступ к содержимому, которое вы обслуживаете из ведер Amazon S3, выполните следующие действия:
    - 1. Создайте специального пользователя CloudFront, называемого идентификатором доступа к источнику (OAI), и свяжите его с вашим дистрибутивом.
    - 2. Настройте разрешения ведра S3 так, чтобы CloudFront мог использовать OAI для доступа к файлам в вашем ведре и предоставления их вашим пользователям. Убедитесь, что пользователи не могут использовать прямой URL-адрес к ведру S3 для доступа к файлу.
        - После выполнения этих действий пользователи смогут получить доступ к вашим файлам только через CloudFront, а не напрямую из ведра S3.
- Keyworkd **“Web sockets” = Application load balancer.**
  Application Load Balancers provide native support for WebSockets. You can use WebSockets with both HTTP and HTTPS listeners.
- Сценарий "fanout" - это когда сообщение Amazon SNS отправляется в тему, а затем реплицируется и отправляется в несколько очередей Amazon SQS, конечных точек HTTP или адресов электронной почты. Это позволяет выполнять параллельную асинхронную обработку. Например, можно разработать приложение, которое отправляет сообщение Amazon SNS в тему каждый раз, когда оформляется заказ на товар. Затем очереди Amazon SQS, подписанные на эту тему, будут получать идентичные уведомления о новом заказе. Серверный экземпляр Amazon EC2, подключенный к одной из очередей, может обрабатывать или выполнять заказ, а другой серверный экземпляр может быть подключен к хранилищу данных для анализа всех полученных заказов.
- No of GET/HEAD requests per second per prefix in a bucket is 5,500 .
  No of bucket prefixes = 5.
  total number of read request = 5500 * 5 = 27,500 read requests per second.

- С помощью S3 Object Lock вы можете хранить объекты по модели "запись-единственное-чтение-много" (WORM). С ее помощью можно предотвратить удаление или перезапись объекта на определенный период времени или на неопределенный срок. Блокировка объектов поможет вам выполнить нормативные требования, требующие хранения WORM, или просто добавить еще один уровень защиты от изменений и удаления объектов.
- custom domain name = CloudFront with S3 as origin.
- **Content is dynamic and geographically close. Cloud front would not be efficient** so the best option is "Host the web servers on multiple EC2 instances in multiple Availability zones behind Application load balancer with Auto scaling group."
- Amazon EFS обеспечивает безопасный доступ для тысяч соединений для экземпляров Amazon EC2 и локальных серверов одновременно, используя традиционную модель разрешений файлов, возможности блокировки файлов и иерархическую структуру каталогов по протоколу NFSv4. Установки Amazon EC2 могут получить доступ к вашей файловой системе через AZ, регионы и VPC, а локальные серверы - через AWS Direct Connect или AWS VPN.
- the simplest solution is to **trigger scaling based on CloudWatch CPUUtilization metrics** - if you raised an issue that the daily processing jobs become more slower in application which uses multiple c4.large instances to deploy data-processing application **via AWS Elastic Beanstalk.**

Practice Set 7

- Deployment approach:
  развернуть приложение на всех узлах одновременно = **использовать All-at-once**
  При развертывании на вычислительной платформе Amazon ECS в конфигурации развертывания указывается, как трафик переключается на обновленный набор задач Amazon ECS.
  - Существует три способа перераспределения трафика во время развертывания:

    - **Canary:** Трафик смещается в два этапа. Можно выбрать один из предопределенных вариантов Canary, который определяет процент трафика, переводимого на обновленный набор задач Amazon ECS в первом приращении, и интервал в минутах до перевода оставшегося трафика во втором приращении.
    - **Linear:** Трафик переключается равными порциями с равным количеством минут между каждой порцией. Вы можете выбрать из предопределенных линейных опций, которые определяют процент трафика, смещаемого в каждом приращении, и количество минут между каждым приращением.
    - **All-at-once:** Весь трафик перемещается с исходного набора задач Amazon ECS на обновленный набор задач Amazon ECS одновременно.
- **Amazon ECS с Fargate Launch Type** - используется, если вам нужно запускать контейнеры без необходимости управлять серверами или кластерами экземпляров Amazon EC2.

  AWS Fargate - это технология, которую вы можете использовать с Amazon ECS для запуска контейнеров без необходимости управления серверами или кластерами экземпляров Amazon EC2. С AWS Fargate вам больше не нужно предоставлять, настраивать или масштабировать кластеры виртуальных машин для запуска контейнеров. Это устраняет необходимость выбирать типы серверов, решать, когда масштабировать кластеры, или оптимизировать упаковку кластеров.

  При запуске задач и служб с типом запуска Fargate вы упаковываете приложение в контейнеры, указываете требования к процессору и памяти, определяете сетевые политики и политики IAM и запускаете приложение. Каждая задача Fargate имеет свою собственную границу изоляции и не разделяет базовое ядро, ресурсы ЦП, память или эластичный сетевой интерфейс с другой задачей.

- **Configure the TTL** - единственный способ управления кэшированием в Amazon API Gateway.
  Вы можете включить кэширование API в Amazon API Gateway, чтобы кэшировать ответы вашей конечной точки. С помощью кэширования можно сократить количество обращений к конечной точке, а также улучшить задержку запросов к API.
  Когда вы включаете кэширование для этапа, API Gateway кэширует ответы от вашей конечной точки в течение определенного периода времени жизни (TTL) в секундах. Затем API Gateway отвечает на запрос, просматривая ответ конечной точки из кэша, вместо того чтобы делать запрос к вашей конечной точке. Значение TTL по умолчанию для кэширования API составляет 300 секунд. Максимальное значение TTL составляет 3600 секунд. TTL=0 означает, что кэширование отключено.
- **С вас НЕ взимается плата за НЕПОЛНЫЙ ЧАС использования**, если цена Spot превышает вашу максимальную цену и Amazon EC2 прерывает работу вашего Spot Instance.
- To restrict users based on their location = Use Amazon CloudFront Geo Restriction.
- Use **Cross-Origin Resource Sharing (CORS)** to allow allow communication between applications from different domains.
- IPv6 traffic to the Internet = Use an egress-only internet gateway!
- Чтобы сопоставить домен с другим доменом, используйте запись CNAME.

- Благодаря поддержке SNI мы упрощаем использование более одного сертификата с одним и тем же ALB. Наиболее распространенная причина, по которой вам может понадобиться использовать несколько сертификатов, - это **работа разных доменов с одним и тем же балансировщиком нагрузки**. Всегда было возможно использовать сертификаты с подстановочными знаками и сертификаты с альтернативными именами субъектов (SAN) с ALB, но они имеют свои ограничения. Сертификаты Wildcard работают только для связанных поддоменов, которые соответствуют простому шаблону, а сертификаты SAN могут поддерживать множество различных доменов, но при этом один и тот же центр сертификации должен аутентифицировать каждый из них. Это означает, что вам придется заново проходить аутентификацию и заново предоставлять сертификат каждый раз, когда вы добавляете новый домен.

  Браузеры, поддерживающие SNI, во время инициализации защищенного соединения немедленно сообщают имя веб-сайта, к которому хочет подключиться посетитель, чтобы сервер знал, какой сертификат отправить обратно. Это позволяет браузерам/клиентам и серверам, поддерживающим SNI, подключать несколько сертификатов для нескольких доменов к одному IP-адресу. Таким образом, посетитель вашего сайта не заметит никаких различий.

  Балансировщик нагрузки использует интеллектуальный алгоритм выбора сертификата с поддержкой SNI. Если имя хоста, предоставленное клиентом, совпадает с одним сертификатом в списке сертификатов, балансировщик нагрузки выбирает этот сертификат. Если имя хоста, предоставленное клиентом, соответствует нескольким сертификатам в списке сертификатов, балансировщик нагрузки выбирает лучший сертификат, который может поддерживать клиент. Выбор сертификата основывается на следующих критериях в следующем порядке:
  - a - Алгоритм открытого ключа (предпочтительнее ECDSA, чем RSA)
  - b - Алгоритм хэширования (предпочтительнее SHA, чем MD5)
  - c - Длина ключа (предпочтительнее наибольшая)
  - d - Срок действия

- Действия, которые могут быть настроены в случае прекращения работы экземпляра **Spot**:
    1. hibernate
    2. stop
    3. terminate
- Для создания быстрого хранилища сессий для ваших онлайн-приложений = Используйте **Amazon ElastiCache** для Redis!
- Экономически эффективное и масштабируемое хранилище + хранение картографических данных в сотнях файлов данных + возможность роста до десятков терабайт = стандарт Amazon S3.
- To route a trafic to static pages hosted on Amazon S3 , you must have:
    - a – registered domain name.
    - b – The bucket must have the same name as your domain or subdomain
- **Вы можете использовать следующие механизмы для аутентификации и авторизации в API Gateway:**
    - Политики ресурсов позволяют создавать политики на основе ресурсов для разрешения или запрета доступа к API и методам с указанных IP-адресов источников или конечных точек VPC.
    - Стандартные роли и политики AWS IAM обеспечивают гибкий и надежный контроль доступа, который можно применять ко всему API или отдельным методам. Роли и политики IAM можно использовать для контроля над тем, кто может создавать и управлять вашими API, а также над тем, кто может их вызывать.
    - Теги IAM могут использоваться вместе с политиками IAM для контроля доступа.
    - Политики конечных точек для конечных точек интерфейса VPC позволяют прикреплять политики ресурсов IAM к конечным точкам интерфейса VPC для повышения безопасности ваших частных API.
    - Авторизаторы Lambda - это функции Lambda, которые контролируют доступ к методам REST API, используя аутентификацию по маркеру предъявителя, а также информацию, описанную в заголовках, путях, строках запроса, переменных этапа или параметрах запроса контекстных переменных. Авторизаторы Lambda используются для управления тем, кто может вызывать методы REST API.

    - Пулы пользователей Amazon Cognito позволяют создавать настраиваемые решения аутентификации и авторизации для ваших REST API. Пулы пользователей Amazon Cognito используются для управления тем, кто может вызывать методы REST API.

- Для подключения elastic network interface (ENI) к работающему экземпляру EC2 - необходимо использовать **Hot attach**!

  Вы можете присоединить сетевой интерфейс к экземпляру во время его работы (**Hot attach**), во время его остановки (**Warm attach**) или во время запуска экземпляра (**Cold attach**).
    - Можно отсоединить вторичные сетевые интерфейсы, когда экземпляр запущен или остановлен. Однако нельзя отсоединить первичный сетевой интерфейс.
    - Можно переместить сетевой интерфейс с одного экземпляра на другой, если экземпляры находятся в одной зоне доступности и VPC, но в разных подсетях.
    - При запуске экземпляра с помощью CLI, API или SDK можно указать основной сетевой интерфейс и дополнительные сетевые интерфейсы.
    - Запуск экземпляра Amazon Linux или Windows Server с несколькими сетевыми интерфейсами автоматически настраивает интерфейсы, частные адреса IPv4 и таблицы маршрутизации в операционной системе экземпляра.
    - При теплом или горячем подключении дополнительного сетевого интерфейса может потребоваться вручную вызвать второй интерфейс, настроить частный IPv4-адрес и соответствующим образом изменить таблицу маршрутизации. Экземпляры под управлением Amazon Linux или Windows Server автоматически распознают теплое или горячее подключение и настраиваются самостоятельно.
    - Присоединение еще одного сетевого интерфейса к экземпляру (например, конфигурация NIC teaming) не может использоваться как метод увеличения или удвоения пропускной способности сети к или от экземпляра с двойным подключением.

    - При подключении к экземпляру двух или более сетевых интерфейсов из одной подсети вы можете столкнуться с сетевыми проблемами, такими как асимметричная маршрутизация. Если возможно, используйте вторичный частный IPv4-адрес на основном сетевом интерфейсе.

- Если удаленный объект все еще существует в ведре Amazon S3, поскольку запросы Amazon S3 DELETE в конечном счете последовательны.
  Amazon S3 обеспечивает конечную согласованность для DELETES во всех регионах AWS. Кроме того, Amazon S3 реплицирует данные на нескольких серверах. Это означает, что когда вы удаляете объект, может потребоваться некоторое время для репликации удаления на всех серверах.

- В рабочих нагрузках машинного обучения используются огромные объемы обучающих данных. В этих рабочих нагрузках часто используется общее файловое хранилище, поскольку несколько вычислительных экземпляров должны обрабатывать обучающие данные одновременно. FSx для Lustre оптимален для рабочих нагрузок машинного обучения, поскольку он обеспечивает общее файловое хранилище с высокой пропускной способностью и постоянными низкими задержками для обработки учебных наборов данных ML. FSx for Lustre также интегрирован с Amazon SageMaker, что позволяет ускорить выполнение заданий на обучение.
- **Protecting the infrastructure is the main responsibility at Amazon Side.**
- By default, AWS has a limit of 20 instances per region. This includes all instances set up on your AWS account.
  To increase EC2 limits, request a higher limit by providing information about the new limit and regions where it should be applied.
- Each device has different storage capacities, as follows:

Storage capacity (usable capacity)	Snowball	Snowball Edge
50 TB (42 TB usable) - US regions only	✓	
80 TB (72 TB 72 usable)	✓	
100 TB (83 TB usable)		✓
100 TB Clustered (45 TB per node)		✓

- To access AWS Systems Manager APIs from your VPC without accessing the internet = Use AWS PrivateLink!
- Accessed and shared across multiple VPC + store up to 3300 keys+ integrated with AWS CloudTrail + support MFA = **AWS CloudHSM**
- To audit log of any changes made to AWS resources in their account = Use AWS CloudTrail.
- **Auto complete** is a good feature in Use ElastiCache with Redis.
- To reflect the **updates immediately** in **CloudFormation** - Use **Direct update**!
- Для предотвращения подслушивания при общении с ELB = использовать функцию Perfect Forward Secrecy.
  Perfect Forward Secrecy - это функция, которая обеспечивает дополнительную защиту от подслушивания зашифрованных данных благодаря использованию уникального случайного сеансового ключа. Это предотвращает расшифровку перехваченных данных, даже если секретный долгосрочный ключ скомпрометирован.
- To monitor API calls = Use custom CloudWatch matrix.

Practice Set 8

- Real-time notifications based using Amazon CloudWatch = CloudWatch **Events**
- Перенос системы обмена сообщениями в AWS + низкая стоимость = Amazon MQ
- to remove CloudFront caches before expiration =**Invalidate the files**.
- для поддержки ожидаемого роста = добавить больше реплик для чтения + переместить статические файлы из ECS в S3
- track session data + no downtime = Use DynamoDB
- to perform processing on a files in S3 bucket = Use S3 event + Amazon Lambda Function

- Чтобы ограничить доступ к содержимому, которое вы предоставляете из ведер Amazon S3, выполните следующие действия:
  - Создайте специального пользователя CloudFront, называемого идентификатором доступа к источнику (OAI), и свяжите его с вашим дистрибутивом.
  - Настройте разрешения ведра S3 так, чтобы CloudFront мог использовать OAI для доступа к файлам в вашем ведре и предоставления их вашим пользователям. Убедитесь, что пользователи не могут использовать прямой URL-адрес к ведру S3 для доступа к файлу.
- **At EC2 level we cannot use WAF because AWS WAF is used to control how an Amazon CloudFront distribution, an Amazon API Gateway API, or an Application Load Balancer responds to web requests.**
- для балансировки запросов по нескольким репликам чтения Aurora = Использовать конечную точку Aurora Reader
- Расширение локальных сетей в облако и безопасный доступ к ним из любого места = использование VPN
- Защита данных - это защита данных во время их транспортировки (когда они перемещаются в Amazon S3 и обратно) и в состоянии покоя (когда они хранятся на дисках в центрах обработки данных Amazon S3). Данные в пути можно защитить с помощью протокола Secure Sockets Layer (SSL) или шифрования на стороне клиента. У вас есть следующие варианты защиты данных в состоянии покоя в Amazon S3:
  - Шифрование на стороне сервера - запрос к Amazon S3 на шифрование объекта перед сохранением его на дисках в центрах обработки данных, а затем расшифровка при загрузке объектов.
  - Шифрование на стороне клиента - шифрование данных на стороне клиента и загрузка зашифрованных данных в Amazon S3. В этом случае вы управляете процессом шифрования, ключами шифрования и соответствующими инструментами.
- для ускорения загрузки изображений = Используйте Amazon CloudFront
- to provide High availability for Amazon ElastiCache for Redis = Configure ElastiCache Multi-AZ .

- Amazon RDS обнаруживает и автоматически восстанавливается после наиболее распространенных сценариев отказа в развертываниях Multi-AZ, чтобы вы могли как можно быстрее возобновить работу базы данных без вмешательства администратора. Amazon RDS автоматически выполняет обход отказа в случае любого из следующих событий:
  - Потеря доступности в основной зоне доступности
  - Потеря сетевого подключения к первичному серверу
  - Отказ вычислительного блока на первичном сервере
  - Отказ системы хранения данных на основной базе
- to simplify AWS infrastructure with providing IP multicast = Use AWS Transit Gateway
- Amazon DynamoDB интегрирована с AWS Lambda, поэтому вы можете создавать триггеры - части кода, которые автоматически реагируют на события в потоках DynamoDB. С помощью триггеров можно создавать приложения, реагирующие на изменения данных в таблицах DynamoDB. Если вы включите потоки DynamoDB для таблицы, вы можете связать потоковое имя ресурса Amazon (ARN) с написанной вами функцией AWS Lambda. Сразу же после изменения элемента в таблице в потоке таблицы появляется новая запись. AWS Lambda опрашивает поток и вызывает вашу функцию Lambda синхронно, когда обнаруживает новые записи в потоке.
- AWS Security Token Service (AWS STS) - это веб-служба, которая позволяет запрашивать временные учетные данные с ограниченными привилегиями для пользователей AWS Identity and Access Management (IAM) или для пользователей, которых вы аутентифицируете (объединенные пользователи). В данном руководстве описывается API AWS STS.
- No of PUT requests per second per prefix in a bucket is 3,500 .
  - No of bucket prefixes = 3
  - total number of read request = 3500 * 3 = 10,500 PUT requests per second
  - For example, your application can achieve at least 3,500 PUT/COPY/POST/DELETE or 5,500 GET/HEAD requests per second per prefix in a bucket.

Practice Set 9

- Strong consistency = DynamoDB

- Handle failed messages = Amazon SQS dead-letter queue.
  Amazon SQS поддерживает очереди с мертвыми буквами, которые другие очереди (очереди-источники) могут использовать для сообщений, которые не могут быть успешно обработаны (потреблены). Очереди с мертвыми буквами полезны для отладки приложения или системы обмена сообщениями, поскольку они позволяют изолировать проблемные сообщения, чтобы определить, почему их обработка не удалась.
- RDS Storage Auto Scaling постоянно отслеживает фактическое потребление хранилища и автоматически масштабирует емкость, когда фактическое использование приближается к предоставленному объему хранилища. Автомасштабирование работает с новыми и существующими экземплярами баз данных. Вы можете включить функцию автоматического масштабирования всего несколькими щелчками мыши в консоли управления AWS. Автомасштабирование хранилища RDS не требует дополнительных затрат. Вы платите только за ресурсы RDS, необходимые для работы ваших приложений.
- To protect the application from DDOS attach = **Use AWS Shield**
- Elastic Load Balancing provides access logs that capture detailed information about requests sent to your load balancer
- Can concurrently run **20000 requests** or functions + Scalable = using **docker container** on Amazon ECS.
  But **Lambda has a limitation of 1000 concurrent requests**
  Your functions' concurrency is the number of instances that serve requests at a given time. For an initial burst of traffic, your functions' cumulative concurrency in a Region can reach an initial level of between 500 and 3000, which varies per Region.
  Burst concurrency limits for Lambda:
  - 3000 – US West (Oregon), US East (N. Virginia), Europe (Ireland)
  - 1000 – Asia Pacific (Tokyo), Europe (Frankfurt)
  - 500 – Other Regions
- Simplify inventory and compliance management across accounts and regions = **AWS Config**

- S3, DynamoDB и Lambda имеют HA, только PostgreSQL пока не имеет... Просто включите его.
  Пояснение:
  PostgreSQL стал предпочтительной реляционной базой данных с открытым исходным кодом для многих корпоративных разработчиков и стартапов, обеспечивая работу ведущих бизнес- и мобильных приложений. Amazon RDS упрощает настройку, эксплуатацию и масштабирование развертывания PostgreSQL в облаке. С помощью Amazon RDS вы можете развернуть масштабируемые системы PostgreSQL за считанные минуты, используя экономически эффективные и изменяемые аппаратные мощности. Amazon RDS решает сложные и трудоемкие административные задачи, такие как установка и обновление программного обеспечения PostgreSQL, управление хранилищем, репликация для обеспечения высокой доступности и пропускной способности чтения, а также резервное копирование для аварийного восстановления.
- **Enable** Amazon **DynamoDB Auto Scaling** = LESS changes
- A company has on-premises Microsoft Active Directory and wants to allow its employees to access it's multiple accounts in AWS using their user names and passwords - USE **AWS Single Sign On**
- **for HA , we need two Availability zone and each AZ contains 3 subnets (1 public for ALB + 1 private for Web servers + 1 private for Database).**
- AWS compute solution + **no special hardware** + use 512 MB of memory to run = AWS **Lambda functions**
- Amazon EC2 instance should stop rather than terminate when its Spot Instance is interrupted:
  - For a Spot Instance request, the type must be **persistent**. You cannot specify a launch group in the Spot Instance request.
  - For an EC2 Fleet or Spot Fleet request, the type must be **maintain**.
  - The root volume must be an **EBS volume**, not an instance store volume.

- **Выделенные хосты** Amazon EC2 позволяют использовать лицензии на программное обеспечение от таких производителей, как Microsoft и Oracle, на Amazon EC2, что обеспечивает гибкость и экономическую эффективность использования собственных лицензий, а также гибкость, простоту и эластичность AWS.
  Выделенный хост Amazon EC2 - это физический сервер, полностью выделенный для вашего использования, что поможет вам соответствовать корпоративным требованиям.

  Выделенные хосты позволяют использовать существующие лицензии на программное обеспечение на сокет, ядро или виртуальную машину, включая Windows Server, SQL Server, SUSE Linux Enterprise Server, Red Hat Enterprise Linux или другие лицензии на программное обеспечение, привязанные к виртуальным машинам, сокетам или физическим ядрам, в соответствии с условиями лицензии. Это поможет вам сэкономить деньги за счет использования существующих инвестиций. Узнайте больше о вариантах лицензирования Windows.

Practice Set 10

- Transactional + High performance + Data size range 16 TB to 64 TB = Amazon Aurora
- Object Store = S3
  Object Store + Immutable = Amazon Glacier
- Big Data + flexible schema + indexed data + scalable =Amazon DynamoDB
- **Enable versioning** in both source and destination buckets is **prerequisites for cross-region replication in Amazon S3**
- EC2 Spot instances will be terminated when interrupted, the operation team asked you to design a highly available workload to handle the interruption by including a Two-Minute warning when there is not enough capacity.
  - Invoke an AWS Lambda function to launch On-Demand Instances which is triggered by Amazon CloudWatch Events.
- Important Notes about Purchases RDS Reserved Instance:

    - Reserved Instance prices cover instance costs only. Storage and I/O are still billed separately.
    - Region, DB Engine, DB Instance Class, Deployment Type and term length must be chosen at purchase, and cannot be – changed later.
    - You can purchase up to 40 Reserved Instances. If you need additional Reserved Instances, complete the form found here.
    - Reserved Instances may not be transferred, sold, or cancelled and the one-time fee is non-refundable.
- **Amazon Kinesis Data Firehose** - это служба извлечения, преобразования и загрузки (ETL), которая обеспечивает надежный сбор, преобразование и доставку потоковых данных в озера данных, хранилища данных и аналитические службы.

  Примеры использования:
    - Потоковая передача данных в Amazon S3 и преобразование данных в необходимые форматы для анализа без создания конвейеров обработки.
    - Мониторинг безопасности сети в режиме реального времени и создание предупреждений при возникновении потенциальных угроз с помощью поддерживаемых инструментов информации в сфере безопасности и управления событиями (SIEM).
    - Обогатите свои потоки данных моделями машинного обучения (ML) для анализа данных и прогнозирования адресов вывода по мере продвижения потоков к месту назначения.
- Chef and Puppet = AWS OpsWorks
- to increase the performance of reading a huge number of files in S3 bucket.
    - a – use sequential date-based naming (Old method)
    - b – Horizontally scale parallel requests to the Amazon S3 service endpoints
- Proprietary File System = EBS

  Проприетарная файловая система означает, что она является собственностью и защищена авторским правом, и что существуют ограничения на использование, распространение и модификацию. Как NTFS или ReFS в системах Windows. Они принадлежат Microsoft. Если вы используете тома EBS, вы можете отформатировать их в любой файловой системе, которую может использовать ОС для монтирования этих томов.
- securely store database passwords + customer master key + Lambda Function = Lambda Environment Variables

- **NFS** поддерживается только **File Gateway** и нет средств визуализации! **AWS Storage Gateway** - это гибридный облачный сервис хранения данных, который предоставляет доступ к практически неограниченному облачному хранилищу. Шлюз хранения данных предоставляет стандартный набор протоколов хранения, таких как iSCSI, SMB и NFS, которые позволяют использовать хранилище AWS без переписывания существующих приложений. Он обеспечивает низкую латентную производительность за счет кэширования часто доступных данных в помещениях, при этом надежно и хранит данные в облачных хранилищах Amazon. Шлюз хранения данных оптимизирует передачу данных в AWS, отправляя только измененные данные и сжимая данные. Шлюз хранения также интегрируется с облачным хранилищем Amazon S3, который делает ваши данные доступными в облачной обработке, AWS Identity and Access Management (AWS IAM) для обеспечения управления доступом к услугам и ресурсам, AWS Key Management Service (AWS KMS) для шифрования данных в облаке, Amazon CloudWatch для мониторинга и AWS CloudTrail для регистрации активности учетной записи.
- **Expedited retrieval** - allows you to quickly access your data when you need to have almost immediate access to your information. This retrieval type can be used for archives up to 250MB. Expedited retrieval usually completes **within 1 and 5 minutes**.
- **Standard retrieval** - provides access to any of your archives within several hours. Standard retrieval usually takes **between 3 and 5 hours** to complete.
- **Bulk retrieval** - is Amazon S3 Glacier's lowest-cost retrieval type. You can retrieve large amounts of data inexpensively. Bulk retrieval usually completes **within 5 and 12 hours**.
- to **monitor VPN connection** if it is up or down use Use **CloudWatch TunnelState Metric**
- Проблема утром не в том, что должно было быть 20 экземпляров работает и что они не работают. Проблема в том, что автоматическое масштабирование не отвечает достаточно быстро на рост спроса. Именно поэтому decreasing the cool down period сделает автоматическое масштабирование более агрессивным (и отзывчивым), но все равно будет работать менее 20 экземпляров от получить идти, и поэтому будет стоить меньше денег. Кроме того, AWS recommends using **target scaling** as much as possible.

- Configure Automated Cross-Region Snapshot Copy for Amazon Redshift Cluster to the other region - **best approach to create DR data warehouse for Amazon Redshift in another region!**

Practice Set 11

- VPC Flow Logs is a feature that enables you to capture information about the IP traffic going to and from network interfaces in your VPC.

Practice Set 12

- AWS Elastic Search предоставляет полные возможности поиска и может быть использован для лог-файлов, хранящихся в ведре S3 Документация AWS упоминает следующее относительно интеграции эластичного поиска с S3 Вы можете интегрировать свой домен Amazon ES с Amazon S3 и AWS Lambda. Любые новые данные, отправленные в ведро S3, запускают уведомление о событии в Lambda, которая затем запускает пользовательский код приложения Java или Node.js. После того, как приложение обработает данные, они будут перенаправлены в ваш домен.
- **the AWS Storage gateway cached volumes service:**
  Используя кэшированные тома, вы можете использовать Amazon S3 в качестве основного хранилища данных, сохраняя часто доступные данные локально в шлюзе хранения. Кэшированные тома сводят к минимуму необходимость масштабирования инфраструктуры хранения данных в помещениях, одновременно предоставляя приложениям доступ к их часто доступным данным с низкой задержкой. Вы можете создавать тома хранения размером до 32 ТиБ и **прикреплять к ним устройства iSCSI** с серверов приложений в помещениях. Ваш шлюз хранит данные, которые вы записываете в эти тома в Amazon S3, и сохраняет недавно прочитанные данные в кэше вашего внутриофисного шлюза хранения и загружает буферное хранилище.

- The normal reason for using SQS, is for decoupling of systems and helps in horizontal scaling of aws resources. SQS does not either do transcoding output or checks the health of the worker instances.
  Обычная причина использования SQS заключается в разделении систем и помогает в горизонтальном масштабировании ресурсов AWS. SQS не занимается транскодированием выходных данных и не проверяет состояние рабочих экземпляров.
- **AWS Redshift = DataWarehouse**
- Это можно сделать через службу **AWS Opsworks**. Ниже приведена документация от AWS для поддержки этого требования AWS OpsWorks Stacks позволяет управлять приложениями и серверами на AWS и в помещениях. С помощью OpsWorks Stacks вы можете моделировать ваше приложение как стек, содержащий различные слои, такие как балансировка нагрузки, база данных и сервер приложений. Можно развернуть и настроить экземпляры Amazon EC2 на каждом уровне или подключить другие ресурсы, такие как базы данных Amazon RDS.
- multiple linked tables = relation Database

Practice Set 13

- Amazon Redshift will never automatically delete a manual snapshot. Manual snapshots are retained even after you delete your cluster. Because manual snapshots accrue storage charges, it's important that you manually delete them
- **Volume Gateway with cached mode is the best option to migrage iSCSI to Cloud**
  If a backup of **On-premise** data is required, the most efficient way would be to make use of **Storage gateway Cached Volumes**. The AWS Documentation mentions the following on Cached Volumes. Cached volumes – You store your data in Amazon Simple Storage Service (Amazon S3) and retain a copy of frequently accessed data subsets locally. **Cached volumes** offer a substantial cost savings on primary storage and minimize the need to scale your **storage on-premises**. You also retain low-latency access to your frequently accessed data.

- The AWS Documentation mentions the following on Amazon Aurora Amazon Aurora is a drop-in replacement for MySQL and PostgreSQL. The code, tools and applications you use today with your existing MySQL and PostgreSQL databases can be used with Amazon Aurora.
- Знімок сервісних обмежень, за якими може стежити Trusted Advisor
- Документація AWS згадує наступне використання CloudFront може бути більш економічно ефективним, якщо користувачі отримують доступ до ваших об'єктів часто, тому що при більш високому використанні ціна на передачу даних CloudFront нижча, ніж ціна на передачу даних Amazon S3. Крім того, завантаження швидші за CloudFront, ніж тільки за допомогою Amazon S3, тому що ваші об'єкти зберігаються ближче до ваших користувачів.
- Потоки Kinesis підтримують зміни до періоду збереження даних вашого потоку. Кінезис потік — впорядкована послідовність записів даних, призначених для запису та читання з реального часу. Таким чином, записи даних зберігаються у shards у вашому потоці тимчасово. Період часу від того, коли запис додається до того, коли він вже не доступний, називається періодом утримання. Потік Кінесіс зберігає записи від 24 годин за замовчуванням, до 168 годин.
- columnar database = Amazon RedShift
- **Use a hexadecimal hash for the prefix =** for increasing performance if you have a high request rate in S3
- The AWS Documentation mentions that the weighted routing policy is good for testing new versions of the software. And this is the ideal approach for Blue Green deployments.
- AWS Global Accelerator is a service that improves the availability and performance of your applications with local or global users. It provides static IP addresses that act as a fixed entry point to your application endpoints in a single or multiple AWS Regions, such as your Application Load Balancers, Network Load Balancers or Amazon EC2 instances.

- **You use the reader endpoint for read-only connections for your Aurora cluster.** This endpoint uses a load-balancing mechanism to help your cluster handle a query-intensive workload. The reader endpoint is the endpoint that you supply to applications that do reporting or other read-only operations on the cluster.
  The reader endpoint load-balances connections to available Aurora Replicas in an Aurora DB cluster. It doesn't load-balance individual queries. If you want to load-balance each query to distribute the read workload for a DB cluster, open a new connection to the reader endpoint for each query.
  Each Aurora cluster has a single built-in reader endpoint, whose name and other attributes are managed by Aurora. You can't create, delete, or modify this kind of endpoint.
- If your cluster contains only a primary instance and no Aurora Replicas, the reader endpoint connects to the primary instance. In that case, you can perform write operations through this endpoint.
- **Volume Gateway with cached mode is the best option to migrage iSCSI to Cloud**
- AWS Global Accelerator is a service that improves the availability and performance of your applications with local or global users.
- Всі класи зберігання, відмінні від S3, мають мінімальну тривалість зберігання 30 днів
- Amazon S3 Transfer Acceleration забезпечує швидку, легку та безпечну передачу файлів на великі відстані між клієнтом та відром S3.
- the application tier needs to always receive updated data from database after any update queries = **Amazon RDS!**
- gateway endpoint + route table entry = ensure that API calls to Amazon DynamoDB from Amazon EC2 instances in a VPC do not traverse the internet.

Practice Set 14

- VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by AWS PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC do not require public IP addresses to communicate with resources in the service. Traffic between your VPC and the other service does not leave the Amazon network.

- **EKS** is open source framework which can be used to manage Kubernetes cluster .
- For performance enhancement DB - **run SELECT queries on replicas for (Stale + Reporting) data**.
- **Bucket policy is a secure way to share S3 files!**
  A bucket policy is a resource-based AWS Identity and Access Management (IAM) policy. You add a bucket policy to a bucket to grant other AWS accounts or IAM users access permissions for the bucket and the objects in it. Object permissions apply only to the objects that the bucket owner creates.
- **To run different queries types on big data =Amazon Redshift workload management (WLM)**
  Amazon Redshift workload management (WLM) enables users to flexibly manage priorities within workloads so that short, fast-running queries won't get stuck in queues behind long-running queries.
  Amazon Redshift WLM creates query queues at runtime according to service classes, which define the configuration parameters for various types of queues, including internal system queues and user-accessible queues. From a user perspective, a user-accessible service class and a queue are functionally equivalent. For consistency, this documentation uses the term queue to mean a user-accessible service class as well as a runtime queue.
  When you run a query, WLM assigns the query to a queue according to the user's user group or by matching a query group that is listed in the queue configuration with a query group label that the user sets at runtime.
- To share some video files that are stored in a private S3 bucket for a short period of time with your friends using Amazon CloudFront
  - – Use CloudFront Signed Cookies restrict access to multiple files.
  - – Use CloudFront Signed URL restrict access to a Single file.
    Explanation
    CloudFront signed cookies allow you to control who can access your content when you don't want to change your current URLs or when you want to provide access to multiple restricted files, for example, all of the files in the subscribers' area of a website. This topic explains the considerations when using signed cookies and describes how to set signed cookies using canned and custom policies.

- Таймаут видимости в SQS - это время, в течение которого сообщение остается невидимым в очереди после того, как читатель забирает сообщение. Если задание обрабатывается в течение тайм-аута видимости, сообщение будет удалено. **Если задание не будет обработано в течение тайм-аута видимости, сообщение снова станет видимым** (может быть доставлено дважды). Максимальный тайм-аут видимости для сообщения Amazon SQS составляет 12 часов.
- An on-premise data center will be connected to an Amazon VPC by a hardware VPN that has public and VPN-only subnets. The security team has requested that traffic hitting public subnets on AWS that's destined to on-premise applications must be directed over the VPN to the corporate firewall.
  Таблицы маршрутов определяют, куда направляется сетевой трафик. В таблице маршрутов необходимо добавить маршрут для удаленной сети и указать виртуальный частный шлюз в качестве цели. Это позволит трафику из вашего VPC, предназначенному для вашей удаленной сети, проходить через виртуальный частный шлюз и один из VPN-туннелей. Вы можете включить распространение маршрутов для своей таблицы маршрутизации, чтобы автоматически распространять маршруты вашей сети в таблице для вас.
- Every IAM user starts with no permissions.. In other words, by default, users can do nothing, not even view their own access keys. To give a user permission to do something, you can add the permission to the user (that is, attach a policy to the user). Or you can add the user to a group that has the intended permission.
- **AWS Batch** eliminates the need to operate third-party commercial or open source batch processing solutions. There is no batch software or servers to install or manage.
- Чтобы справиться с повышенной нагрузкой на базу данных, вы можете вертикально масштабировать главную базу данных простым нажатием кнопки. Помимо вертикального масштабирования главной базы данных, вы также можете улучшить производительность базы данных с высокой нагрузкой на чтение, используя реплики чтения для горизонтального масштабирования базы данных.
  - Vertical scaling for read and write by choosing a larger instance size is a correct answer"
  - Horizontal scaling for read capacity by creating a read-replica is also a correct answer.

- Клиенты AWS могут проводить оценки безопасности или тесты на проникновение в свою инфраструктуру AWS без предварительного одобрения 8 услуг. Пожалуйста, ознакомьтесь с последней информацией по ссылке AWS ниже.
  https://aws.amazon.com/ru/security/penetration-testing/
    - AWS разрешает проникновение на некоторые ресурсы без предварительного разрешения" - это правильный ответ.
- AWS Step Functions оркеструет бессерверные рабочие процессы, включая координацию, состояние и цепочки функций, а также объединяет длительное выполнение, не поддерживаемое в рамках лимитов выполнения Lambda, разбивая на несколько шагов или вызывая рабочих, работающих на инстансах Amazon Elastic Compute Cloud (Amazon EC2) или локально.
  https://aws.amazon.com/ru/step-functions/?step-functions.sort-by=item.additionalFields.postDateTime&step-functions.sort-order=desc
- AWS Lambda automatically monitors Lambda functions and reports metrics through Amazon CloudWatch. **Lambda tracks the number of requests, the latency per request, and the number of requests resulting in an error.** You can view the request rates and error rates using the AWS Lambda Console, the CloudWatch console, and other AWS resources.
- **Alias records** - are used to map resource record sets in your hosted zone to Amazon Elastic Load Balancing load balancers, API Gateway custom regional APIs and edge-optimized APIs, CloudFront Distributions, AWS Elastic Beanstalk environments, Amazon S3 buckets that are configured as website endpoints, Amazon VPC interface endpoints, and to other records in the same Hosted Zone.
- **AWS CodeCommit** - это полностью управляемая служба контроля исходных текстов, в которой размещаются защищенные репозитории на базе Git. Он упрощает совместную работу команд над кодом в безопасной и высокомасштабируемой экосистеме. CodeCommit избавляет от необходимости управлять собственной системой контроля исходных текстов или беспокоиться о масштабировании ее инфраструктуры. Вы можете использовать CodeCommit для безопасного хранения чего угодно, от исходного кода до двоичных файлов, и он легко сочетается с существующими инструментами Git.

- **Amazon EMR** - is a web service that enables businesses, researchers, data analysts, and developers to easily and cost-effectively process vast amounts of data. EMR utilizes a hosted Hadoop framework running on Amazon EC2 and Amazon S3.
- Решение для обеспечения единого входа для существующих сотрудников компании. Персонал управляет веб-приложениями на объекте, а также нуждается в доступе к консоли управления AWS для управления ресурсами в облаке AWS.
  Single sign-on using federation allows users входить в консоль AWS без присвоения учетных данных IAM. AWS Security Token Service (STS) - это веб-служба, которая позволяет вам запрашивать временные учетные данные с ограниченными привилегиями для пользователей IAM или пользователей, которых вы аутентифицируете (например, объединенных пользователей из местного каталога).
  https://aws.amazon.com/ru/identity/saml/

Practice Set 15

- Gateway-cached volumes let you use Amazon Simple Storage Service (Amazon S3) as your primary data storage while retaining frequently accessed data locally in your storage gateway. Gateway-cached volumes minimize the need to scale your on-premises storage infrastructure, while still providing your applications with low-latency access to their frequently accessed data. You can create storage volumes up to 32 TiB in size and attach to them as iSCSI devices from your on-premises application servers. Your gateway stores data that you write to these volumes in Amazon S3 and retains recently read data in your on-premises storage gateway's cache and upload buffer storage.
- The AWS Documentation mentions the following **By default, CloudTrail event log files are encrypted** using Amazon S3 server-side encryption (**SSE**). You can also choose to encrypt your log files with an AWS Key Management Service (AWS KMS) key.

- Когда обновления производятся для объектов в S3, они имеют модель конечной согласованности. Поэтому **при обновлении объектов по одному и тому же ключу может возникнуть небольшая задержка**, когда обновленный объект будет предоставлен пользователю при следующем запросе на чтение.
- When you create a record, you choose a routing policy, which determines how Amazon Route 53 responds to queries:
  - Failover routing policy – Use when you want to configure active-passive failover.
  - **Geolocation routing policy** – Use when you want to route traffic *based on the location of your users*.
  - **Geoproximity routing policy** – Use when you want to route traffic *based на основе местоположения ваших ресурсов и, по желанию, переключать трафик с ресурсов в одном месте на ресурсы в другом.*.
  - Latency routing policy – Use when you have resources in multiple AWS Regions and you want to route traffic to the region that provides the best latency.
  - Multivalue answer routing policy – Use when you want Route 53 to respond to DNS queries with up to eight healthy records selected at random.
- You can configure Secrets Manager to automatically rotate your secrets (for example database password) without user intervention and on a specified schedule.
- **AWS Lambda = up to 1000 requests!**
- SAN disc = Object Store = EBS
- Launch templates enable you to store launch parameters so that you do not have to specify them every time you launch an instance. When you launch an instance using the Amazon EC2 console, an AWS SDK, or a command line tool, you can specify the launch template to use.
- **Amazon Aurora Global Database is not suitable for scaling read operations within a region**. It is a new feature in the MySQL-compatible edition of Amazon Aurora, designed for applications with a global footprint. **It allows a single Aurora database to span multiple AWS regions**, with fast replication to enable low-latency global reads and disaster recovery from region-wide outages.
- **Network ACL's function at the subnet level.**
- to automatically and repeatably create many member accounts within an AWS Organization = **Use CloudFormation with scripts!**

- Aurora Replicas are independent endpoints in an Aurora DB cluster, best used for scaling read operations and increasing availability. Up to 15 Aurora Replicas can be distributed across the Availability Zones that a DB cluster spans within an AWS Region.
  The DB cluster volume is made up of multiple copies of the data for the DB cluster. However, the data in the cluster volume is represented as a single, logical volume to the primary instance and to Aurora Replicas in the DB cluster.
- Конечная точка интерфейса использует AWS PrivateLink и представляет собой эластичный сетевой интерфейс (ENI) с частным IP-адресом, который служит точкой входа для трафика, предназначенного для поддерживаемой службы.
  С помощью PrivateLink вы можете подключить свой VPC к поддерживаемым службам AWS, службам, размещенным на других учетных записях AWS (службы конечных точек VPC), и поддерживаемым службам партнеров AWS Marketplace.
- EFS- allows you to simultaneously *share files between multiple Amazon EC2 instances across multiple AZs, regions, VPCs, and accounts as well as on-premises servers via AWS Direct Connect or AWS VPN*.
  Growing and shrinking automatically as you add and remove files.
- A target tracking action in Auto Scaling Group - is a MOST cost-effective solution!
- Вы можете передавать **зашифрованный трафик** через **NLB** и завершать SSL на экземплярах EC2, так что это правильный ответ.
  Можно использовать **HTTPS listener с ALB** и установить сертификаты как на ALB, так и на экземплярах EC2. В этом случае не будет использоваться сквозная передача, вместо этого будет прервано первое SSL-соединение на ALB, а затем трафик будет повторно зашифрован и подключен к экземплярам EC2.
- A **single KMS key** может быть использован **для шифрования файлов журнала для трайлов, применяемых во всех регионах**. Файлы журналов CloudTrail шифруются с помощью S3 Server Side Encryption (SSE), и вы также можете включить шифрование SSE KMS для дополнительной безопасности.

- **An instance store** provides **temporary block-level storage for your instance**. This storage is located on disks that are physically attached to the host computer. Instance store is ideal for temporary storage of information that changes frequently, such as buffers, caches, scratch data, and other temporary content, or for data that is replicated across a fleet of instances, such as a load-balanced pool of web servers.
  Некоторые типы экземпляров используют твердотельные накопители (SSD) на базе NVMe или SATA для обеспечения высокой производительности произвольного ввода-вывода. Это хороший вариант, когда вам нужно хранилище с очень низкой задержкой, но вам не нужно, чтобы данные сохранялись после завершения работы экземпляра
- Data will persist for 24 hours only = S3 Standard
- DynamoDB Streams помогают хранить **список изменений на уровне элементов** или **предоставлять список изменений на уровне элементов**, произошедших за последние 24 часа. **Amazon DynamoDB интегрирована с AWS Lambda, поэтому вы можете создавать триггеры - части кода, которые автоматически реагируют на события в потоках DynamoDB**.
  Если вы включите потоки DynamoDB для таблицы, вы можете связать ARN потока с написанной вами функцией Lambda. Сразу же после изменения элемента в таблице в потоке таблицы появляется новая запись. AWS Lambda опрашивает поток и вызывает вашу функцию Lambda синхронно, когда обнаруживает новые записи в потоке.
  Сопоставление источников событий определяет источник событий на основе опроса для функции Lambda. Это может быть поток Amazon Kinesis или DynamoDB. Источники событий поддерживают конфигурацию отображения, за исключением потоковых сервисов (например, DynamoDB, Kinesis), для которых конфигурация выполняется на стороне Lambda, а Lambda выполняет опрос.
- **Amazon Aurora global database = for across Regions.**
- You cannot create an encrypted Read Replica from an unencrypted master DB instance. You also cannot enable encryption after launch time for the master DB instance. Therefore, you must create a new master DB by taking a snapshot of the existing DB, encrypting it, and then creating the new DB from the snapshot. You can then create the encrypted cross-region Read Replica of the master DB.

- Чтобы применить ограничения к нескольким учетным записям, вы должны использовать политику управления услугами (SCP) в организации AWS. Для этого нужно создать **deny rule**, которое применяется ко всему, что не соответствует определенному типу экземпляра, который вы хотите разрешить. Для достижения этой цели можно использовать следующую архитектуру: