Solutions Architect Associate exam

Global infrastructure:

26 Launched Regions	84 Availability Zones	410+ Edge Locations

Availability Zones - Distinct locations from within an AWS region that are engineered to be isolated from failures.
A region - is a geographical area divided into Availability Zones. Each region contains at least two Availability Zones
Each AWS Region consists of multiple, isolated, and physically separate Availability Zones within a geographic area.

Identity and Access Management

- New users have no permissions when first created.
- New users are assigned an access key ID and secret access key when first created.
- User access key ID and secret access key are not the same as the password
- Go into CloudWatch, and you create a billing alarm - I can get automatic notifications if my account goes over, like $1,000, or whatever.

S3 (Simple Storage Service)

- S3 - это универсальное пространство имен. Вмена должны быть уникальными во всем мире.
    - [ ] Вопрос со сценарием. В нем рассматривается загрузка файла на S3, и он спрашивает вас, какой код вы получите обратно? Это будет код HTTP 200, который вернется в ваш браузер.
- S3 основан на объектах. И просто думайте об объектах как о файлах. Ключ - имя файла. Значение - его содержимое (данные файла)!

- **Version ID** - у нас есть идентификатор версии, это важно для определения версий.
- **Metadata** - У нас есть метаданные, которые представляют собой данные о данных, которые вы храните.
- **Subresources** - у нас есть списки контроля доступа и торренты. И этот список контроля доступа в основном это разрешения для данного объекта (**Torrents**). И вы можете заблокировать каждый объект по отдельности. Так что вы можете сделать это на уровне ведра, а также на уровне объекта (**Access Control List**).
- **Read after Write consistency for PUTS of new Objects** - как только вы создадите новый объект, вы сможете сразу же прочитать этот объект.
- **Eventual Consistency for overwrite PUTS and DELETES** (can take some time to propagate) - Но если вы обновляете объект или удаляете объект, и попытаетесь прочитать его немедленно, вы получите только конечную согласованность. Таким образом, вы можете получить автоматический объект, или вы можете увидеть удаленный файл. Но если вы подождете около секунды, все будет согласовано. Таким образом, вы получаете конечную согласованность.

S3 has the following features:

1. Tiered Storage Available
2. Lifecycle Management
3. Versioning
4. Encryption
5. MFA Delete
6. Secure your data using **Access Control Lists** and **Bucket Policies**

S3 Storage Classes:

1. **S3 Standard** - это тот, который имеет 99,99% доступности и 11 девяток по долговечности. Она хранится с избытком на нескольких устройствах в нескольких хранилищах и рассчитана на то, чтобы выдержать потери двух объектов одновременно. [miliseconds]

2. **S3 Infrequently Accessed (S3 - IA)** - S3 с редким доступом. Это, в основном, для данных, к которым обращаются реже, но требуют быстрого доступа, когда они вам нужны. В этом случае плата за хранение данных ниже, чем в S3, но взимается плата за извлечение. [miliseconds]
3. **S3 One Zone - IA** - S3 с одной зоной нечастого доступа. И это тот случай, когда вам нужен действительно недорогой вариант для редко используемых данных. И вам даже не нужно, вам не нужно беспокоиться о нескольких зонах доступности. Данные буквально хранятся в одной зоне доступности. И доступ к ним осуществляется нечасто, но вам все равно нужно иметь возможность мгновенного доступа к этим данным. [miliseconds]
4. **S3 Intelligent Tiering** - Она использует машинное обучение. И в основном, что она делает, так это смотрит на то, как часто вы используете свои объекты, и затем перемещает ваши объекты по разным классам хранения на основе полученных знаний. Так, он переместит его из стандартного S3 в S3 с редким доступом, потому что он знает, что вы не обращаетесь к этим файлам. Это и есть интеллектуальная многоуровневая система S3. Это четыре класса хранилищ. [miliseconds]
5. **S3 Glacier** - в основном предназначен для архивирования данных. Так что если вы хотите архивировать свои данные, может быть, они вам не нужны, может быть, вы должны хранить их в течение семи лет из-за какого-то федерального постановления, вы будете использовать Glacier. *Вы можете хранить любой объем данных и это действительно супер, супер дешево. А время поиска данных можно настраивать от нескольких минут до нескольких часов*. [select minutes or hours]
6. **S3 Glacier Deep Archive** - это самый низкий класс хранения, самый дешевый класс хранения, который вы можете купить, но время поиска будет составлять 12 часов. Поэтому если вы хотите получить данные обратно с помощью Deep Archive, вы отправляете запрос и получаете данные через 12 часов. [select hours]

S3 Bill:

1. **Storage** - Чем больше храните в S3, тем больше будет выставлен счет.
2. **Requests** - также взимается плата за количество запросов. Если вы делаете много запросов к этим объектам, это будет стоить дороже.

3. **Storage Management Pricing** - оплата за управление хранением. Таким образом, это различные уровни, которые доступны.
4. **Data Transfer Pricing** - также получаете плату за передачу данных.
5. **Transfer Acceleration** - также получаете плату за ускорение передачи данных
6. **Cross Region Replication Pricing** - Допустим, у вас есть ведро, и оно находится на востоке США. И вы хотите автоматически реплицировать свои объекты в другое ведро, которое находится, скажем, в Сиднее, и вы хотите сделать это для высокой доступности а также для аварийного восстановления. То есть, как только вы загрузите объект в ваше ведро на востоке США, и у вас включена межрегиональная репликация, эти объекты будут автоматически реплицированы в ваше ведро в Сиднее.
7. **S3 Transfer Acceleration** - Это позволяет быстро, легко и безопасно передавать файлов на большие расстояния между вашими конечными пользователями и ведром S3. По сути, это использование преимуществ глобально распределенных пограничных точек Amazon CloudFront. Когда данные поступают на пограничный узел и направляются в Amazon S3 по оптимизированному сетевому маршруту. Все, что они делают, это используют магистральную сеть Amazon. По сути, если вы включите функцию Transfer Acceleration, юзеры загружают свои файлы в Edge Locations, а не в само ведро S3.

How to restrict bucket access?

1. **Bucket Policies** - Первый - это использование политики ведра, которая будет применяться ко всему ведру.
2. **Object Policies** - использование объектных политик. И вместо того, чтобы применять их ко всему ведру, они применяются к отдельным файлам внутри ведра.
3. **IAM Policies to Users & Groups** - мы можем использовать политики IAM для пользователей и групп в вашей учетной записи AWS для контроля доступа к ведру. (Так, возможно, вы хотите, чтобы ваш отдел кадров имел доступ к ведру HR, но вы не хотите, чтобы финансовый отдел или отдел продаж и маркетинга могли читать эту конфиденциальную информацию. Поэтому вы можете использовать политики управления доступом к идентификационным данным и применять их к отдельных пользователей или к группам, которые могут содержать отдельных пользователей).

S3 Security and Encryption:

- You can encrypt individual objects
- You can also encrypt your objects at a bucket level

Versioning:

- каждый раз, когда вы делаете что-то общедоступным, даже если это последняя версия, это не обязательно делает публичными другие версии. Вы должны зайти и сделать это индивидуально.
- Версионность хранит все версии объекта, включая любые права, и даже если вы удалите объект, он поместит маркер удаления на этот объект, но версии, которые существовавшие до этого, все еще будут существовать.
- Версионирование имеет многофакторную аутентификацию и возможность удаления, и это в основном использует многофакторную аутентификацию для удаления файла.
- It's a fantastic backup tool and once enabled version and cannot be disabled, it can only be suspended.

Conclude:

- [ ] Files can be zero bytes to five terabytes in size.
- [ ] There's unlimited storage.
- [ ] S3 is a universal namespace.
- [ ] Your files are stored in these things called buckets.
- [ ] And a bucket is basically just a folder in the cloud.
- [ ] So essentially, it's only used to store files, you're not going to install an operating system on S3 and you're not going to use it to host the database.
- [ ] S3 Storage Classes
- [ ] S3 Bill

S3 Object Lock and Glacier Vault Lock

1. Object Locks come in two modes, governance mode and compliance mode. If you remember governance mode users can't overwrite or delete an object version or alter its lock settings, unless they have special permissions.
2. With compliance mode though, protected objects version, can't be overwritten or deleted by any user, including the root user in your account

3. S3 Glacier Vault Lock is. It allows you to easily deploy and enforce compliance controls for individual S3 Glacier Vaults with a Vault Lock policy and you can specify controls such as WORM in a Vault Lock policy and lock the policy from future edits and once locked, the policy can no longer be changed.

S3 Performance

- Чем больше у нас префиксов, тем большей производительности мы можем добиться. Prefixes - simply then is the pathway between your bucket name and your file.
- Если вы используете шифрование на стороне сервера для KMS, имейте в виду, что при шифровании и расшифровке данных, вы столкнетесь с жесткими ограничениями, которые зависят от региона. If you're using SSE KMS to encrypt your objects in S3, you must keep in mind the KMS limits. Uploading and downloading will count towards the quota. It is region specific.
- Also remember to use multi-part uploads to increase your performance when uploading files to S3, it should be used for any files over 100 megs, and it will be used.

S3 Select and Glacier Select

- S3 Select - it's a way of essentially pulling your data from S3, using SQL.
  - So, S3 Select - It's a way of using SQL to download the data that you need from S3
- Glacier Select - is just like S3 Select, and it allows you to run SQL queries against Glacier directly.

AWS Organizations and Consolidated Billing

- Always enable multi-factor authentication on the root account.
- You should always use a strong and complex password on the root account.
- The paying account should be used for billing purposes only, do not deploy resources into the paying account
- The paying account is simply the root account or master account.
- Enable or disable AWS services using service control policies or SCP

Sharing S3 Buckets Across Accounts

1. `Using Bucket Policies & IAM (applies across the entire bucket). Programmatic Access Only`

2. `Using Bucket ACLs & IAM (individual objects). Programmatic Access Only`
3. `Cross-account IAM Roles. Programmatic and Console access.`

Cross-Region Replication Bucket

- Remember that versioning must be enabled on both the source and the destination buckets in order for replication to work.
- Files in an existing bucket are not replicated automatically. All subsequent or updated files will be replicated automatically.
- Delete markers are not replicated, and then deleting individual versions or delete markers will also not be replicated.
- If you change the permissions of an object in the source bucket, it does not change those permissions in the destination bucket
- The cool thing about cross-region replication is you can do it between buckets in the same AWS account, but you can also do it for buckets in different AWS accounts.
- Cross Region Replication - is at a high level

S3 Transfer Acceleration

- У вас есть пользователи, у вас есть периферийные точки, и ваши пользователи загружают свои большие файлы в пограничные точки, которые затем проходят через магистральную сеть Amazon и загружают эти файлы непосредственно в ваше ведро S3 в указанном вами регионе.
- So you've got your users, you've got your edge locations, and your users upload their big files to the edge locations, which then traverse Amazon's backbone network and will upload those files directly to your S3 bucket in the region that you specify.

AWS DataSync

  - Это способ копирования данных в AWS! Он используется для перемещения больших объемов данных из локальной сети в AWS.

1. Used to move **large amounts** of data from on-premise to AWS.
2. Used with **NFS** and **SMB** compatible file systems.
3. **Replication** can be done hourly, daily, or weekly.
4. Install the **DataSync agent** to start the replication.
5. Can be used to replicate **EFS** to **EFS**

`How you can move from on prem to AWS. DataSync is definitely one of the valid options.`

CloudFront

- **CloudFront** - это сеть доставки контента, или Content Delivery Network (**CDN**), и, по сути, сеть доставки контента это система распределенных серверов или сеть, которая доставляет веб-страницы и другой веб-контент пользователю на основе географического местоположения пользователя, происхождения веб-страницы, и с помощью сервера доставки контента.
    - CloudFront можно использовать для доставки всего вашего сайта, включая динамический, статический, потоковый и интерактивный контент **используя глобальную сеть Edge Locations**, и запрос на ваш контент автоматически автоматически направляется в ближайший Edge Location, таким образом, контент доставляется с наилучшей возможной производительностью.

1. **Origin** - this is the origin of all the files that the CDN will distribute. This can either be an S3 bucket, or it could be an EC2 instance, or an elastic load balancer, or Route53.
2. **Edge Location -** This is location where content will be cached. This is separate to an AWS Region/AZ
3. **Distribution** - is the name that's given to the CDN, which is a collection of Edge Locations.
4. **Web Distribution** - is typically used for websites.
5. **RTMP** - Used for Media Streaming.

- [ ] Edge Locations are not just READ only - you can write to them too.
- [ ] Objects are cached for the life of the TTL (Time To Live.)
- [ ] You can clear cached objects, but you will be charged.

Create CloudFront

- **Web Distribution** and **RTMP**
- You can restrict access using signed URLs or signed cookies.
    - If you push out some data and then you figure out something's wrong and you do an update, but it's not showing up correctly, the way to deal with that is to create an invalidation.

CloudFront Signed URLs and Cookies

- Use signed URLs/cookies ***when you want to secure content*** so that only the people you authorize are able to access it.
    - A signed URL is for individual files. | **1 file = 1 URL.**
    - A signed cookie is for multiple files. | **1 cookie = multiple files.**
    - Origin Access Identity
    - If you origin is EC2, then use CloudFront.
- Поэтому, когда вы идете на экзамен, если они говорят о подписанном URL CloudFront и подписанном URL S3, или просто подумайте о том, могут ли ваши пользователи получить доступ к S3, если они используют OAI через CloudFront, то они не смогут. Поэтому вы будете использовать URL с подписью CloudFront, но если они могут получить доступ к ведру S3 напрямую, и это просто отдельный объект, то вам, вероятно, нужен URL с подписью S3.

Snowball

- **Snowball** - is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of AWS.

Использование Snowball позволяет решить общие проблемы при передаче больших объемов данных включая высокие сетевые затраты,
длительное время передачи данных и проблемы безопасности. Передача данных с помощью Snowball проста, быстро, безопасно и может стоить всего лишь
всего в одну пятую от стоимости использования высокоскоростного Интернета. Snowball в основном поставляется в двух вариантах. У вас есть 50 терабайт или 80 терабайт. В Snowball используется несколько уровней безопасности, предназначенные для защиты ваших данных, включая устойчивые к взлому корпуса, 256-битное шифрование, и стандартный для отрасли модуль Trusted Platform Module, или TPM, который

предназначен для обеспечения как безопасности и полную цепочку хранения ваших данных.

- In general, when your object size reaches 100 MB, you should consider using multipart uploads instead of uploading the object in a single operation.

Storage Gateway

- **Storage Gateway** - это, по сути, виртуальное или физическое устройство и он будет реплицировать ваши данные в AWS.
  - **File Gateway** - for flat files, stored directly on S3.
  - **Volume Gateway** - по сути, является способом хранения ваших виртуальных жестких дисков в S3
    - **Stored Volumes** - Entire Dataset is stored on site and is asynchronously backed up to S3.
    - **Cached volumes** - Entire Data is stored on S3 and the most frequently accessed data is cached on site.
    - **Stored Volumes** - это когда у вас есть весь набор данных на месте, а **Cached volumes** - это когда у вас есть только наиболее часто используемые данные кэшируются на месте.
  - **Tape Gateway** - is offers durable, cost-effective solution to archive your data in the AWS cloud.

Если вы работаете архитектором решений, особенно если вы работаете с компанией, которая переходит на AWS, и у них есть виртуальная лента, или есть ленточная библиотека. Вы можете использовать Storage Gateway для переноса этих данных на виртуальные ленты и реплицировать их в облако.

Athena vs. Macie

- **Athena** - is an interactive query service, which enables you to analyze and query data located in S3 using standard SQL. (Это интерактивная служба запросов. Он позволяет вам запрашивать данные, расположенные в S3, используя стандартный SQL, это бессерверный сервис, и он обычно используется для анализа данных журналов, хранящихся в S3.) + Serverless!

        - Well it can be used to query log files stored in S3.
        - So this could be your elastic load balancer logs, could be S3, access logs, etc.
        - You can also use it to generate business reports on data stored in S3.
- **Macie** - is a security service that uses machine learning and **N**atural **L**anguage **P**rocessing or **NLP** to discover, classify and protect sensitive data stored in S3. (Macie по сути является службой безопасности. Он использует искусственный интеллект для анализа ваших данных в S3 и помогает идентифицировать персонально идентифицируемую информацию или PII.)
    - **PII** - Personally Identifiable Information. (Это информация, которая используется для установления личность человека.)

Athena позволяет вам запрашивать ваши данные на S3 на основе команд SQL, которые вы пишете. Однако Macie также запрашивает данные на S3, но он использует машинное обучение и естественный язык для обнаружения информации PII.

EC2

- **EC2** - is virtual machines in the cloud. It's a web service that provides resizable compute capacity.

Pricing Types:

1. **On Demand** - это когда вы платите почасовую или посекундную ставку.
2. **Reserved** - это, в основном, где вы подписываете контракт на один или три года, и чем больше вы платите вперед, тем большую скидку вы получаете.
3. **Spot** - движется, как фондовый рынок, и действительно зависит от спроса и предложения Amazon. И помните, что если у вас есть выделенный экземпляр который был прекращен EC2, вы не будете платить за неполный час использования. Однако если вы сами прекратите работу этого экземпляра, с вас возьмут плату за час в течение которого экземпляр работал.
4. **Dedicated Hosts** - здесь вы получаете выделенную физическую машину для вас, и вы также можете, опять же, платить за это по требованию.

Важно помнить:

- **Termination protection** - is turned off by default, so you must turn it on. When it's turned on, you can't go in and automatically or accidentally terminate your EC2 instances.
- Для экземпляра с резервной копией EBS, по умолчанию корневой том EBS будет удален при завершении работы экземпляра. (Поэтому если вы заходите и завершаете свой экземпляр, он удалит корневой том устройства, но **все дополнительные тома по умолчанию не будут удалены**.)
- Also, remember that **EBS root volumes of your default AMIs can now be encrypted**. (Кроме того, помните, что корневые тома EBS ваших AMI по умолчанию теперь могут быть зашифрованы.)
- А также помните, что вы можете можно брать дополнительные тома и шифровать их.

Security Groups:

- All **inbound traffic** - is blocked by default.
- All **outbound traffic** is allowed. (Поэтому даже если мы удалим это правило, оно никак не повлияет на группу безопасности.)
- **Changes to security groups take effect immediately.**
- If you enable something on the Inbound, Outbound is enabled automatically for that port.
- You cannot set any deny rules in a security group. (вы не можете устанавливать запрещающие правила в группе безопасности. Поэтому если у вас возникнут вопросы по сценарию о настройке запрещающего правила в группе безопасности, группы безопасности так не работают.)

По умолчанию они запрещают все, но затем вы входите и разрешаете.

EBS 101

- **EBS** - **E**lastic **B**lock **S**tore, and essentially it's a **virtual hard disk in the cloud**.
- You can change EBS volumes sizes on the fly, **что включает в себя изменение размера и типа хранилища**.

Five Different Flavors:

1. **General Purpose SSD** - у нас есть SSD общего назначения. (подходит для рабочих нагрузок до 16 000 IOPS на том)

2. **Provisioned IOPS** - который также является SSD. Здесь вы хотите получить очень, очень быстрый ввод, вывод в секунду. (самый высокопроизводительный твердотельный накопитель, и он предназначен для критически важных приложений.)
3. **Throughput Optimized Hard Disk Drive** - оптимизированный по пропускной способности жесткий диск, на самом деле это физический жесткий диск, так что это не SSD, а магнитный. (можете использовать его в таких областях, как большие данные и хранилища данных, он известен как st1.)
4. **Cold Hard Disk Drive** - есть холодный жесткий диск, опять же магнитный. (предназначен для менее часто используемых рабочих нагрузок. Это могут быть такие вещи, как файловые серверы и т.д.)
5. **Magnetic** - есть просто магнитный.

Solid-State Drives (SSD)			Hard disk Drives (HDD)		
Volume Type	**General Purpose SSD**	**Provisioned IOPS SSD**	**Throughput Optimized HDD**	**Cold HDD**	**EBS Magnetic**
Description	General purpose SSD volume that balances price and performance for a wide variety of transactional workloads	Highest-performance SSD volume designed for mission-critical applications	Low cost HDD volume designed for frequently accessed, throughput-intensive workloads	Lowest cost HDD volume designed for less frequently accessed workloads	Previous generation HDD
Use Cases	**Most Work Loads**	**Databases**	**Big Data & Data Warehouses**	**File Servers**	**Workloads where data is infrequently accessed**
API Name	gp2	io1	st1	sc1	Standard
Volume Size	1 GiB - 16 TiB	4 GiB - 16 TiB	500 GiB - 16 TiB	500 GiB - 16 TiB	1 GiB-1 TiB
Max. IOPS**/ Volume	16,000	64,000	500	250	40-200

Wherever your EC2 instance is, the volume is going to be in the same availability zone.

- How we can migrate data from one EBS instance or one EC2 instance in one availability zone to another?
    1. Первое, что вы делаете, это создаете снимок. Затем вы превращаете этот снимок в AMI, а затем используете этот AMI для запуска в других зонах доступности.

2. Другая интересная вещь, которую вы можете сделать, это, конечно, **копировать AMI в разные регионы.** Action > copy AMI > и теперь мы можем переместить этот образ машины Amazon из наших существующих регионов. Так, из Северной Вирджинии мы можем переместить его в Лондон, например, и затем мы можем использовать этот образ для запуска наших экземпляров EC2 в Лондонском регионе. **Таким образом, это способ не только перемещения между зонами доступности, но и способ перемещения между регионами**, и единственное различие здесь в том, что мы копируем AMI из одного региона в другой. Как только AMI окажется в новом регионе, мы можем запустить его в этом регионе и выбрать любую зону доступности, какую захотим.

When you terminate an EC2 instance by default, the root device volume will also be terminated. However, additional volumes that are attached to that EC2 instance will continue to persist.

Snapshots:

- Моментальные снимки существуют на S3. Поэтому, когда мы создали этот снимок, мы делаем снимок, и этот снимок хранится на S3. Как я уже сказал, просто думайте о моментальных снимках как о фотографии диска. Помните, что моментальные снимки - это копии томов в момент времени. Когда вы делаете снимок, это то, как этот том существовал в тот конкретный момент времени.
    - если мы сделали снимок, а затем создали новый файл на этом томе, а затем сделать еще один снимок, **только блоки, только дельта, изменения между блоками будут в основном реплицированы в S3**. Таким образом, **снимки инкрементные.**
- Может потребоваться некоторое время, чтобы создать снимок для тома Amazon EBS, который служит в качестве корневого устройства. Поэтому тот том, на который вы устанавливаете операционную систему, лучше всего остановить. перед созданием моментального снимка.

AMI Types (EBS vs. Instance Store)

1. Для томов EBS - корневое устройство для запуска экземпляра из AMI является том Amazon EBS. созданный из моментального снимка EBS.

2. Однако тома **Instance Store Volumes** - это корневое устройство, то есть том корневого устройства, в основном является местом, куда установлена операционная система. Поэтому **том корневого устройства создается из шаблона. который на самом деле хранится в S3**. Поэтому они совершенно разные.

- При создании Instance на основе **Instance Store Volumes** - Вы также можете присоединить дополнительные тома EBS после запуска экземпляра, но не тома мгновенного хранилища. Поэтому, **если нам нужны дополнительные тома хранилища экземпляров, мы должны сделать это при создании инстанса!** Мы не сможем добавить дополнительные тома позже.

Тома мгновенного хранения иногда называют эфемерными хранилищами. Причина этого в том, что если по какой-то причине они будут остановлены, вы потеряете все свои данные. Тома Instance Store не могут быть остановлены. Если базовый хост выйдет из строя, то вы потеряете свои данные.

  - **Instance Store Volumes** - cannot be stoped! If some reason it stops, if for some reason the underlying host fails, you are going to lose all your data on your virtual hard disk drives.
  - EBS backed instances can be stopped and you're not going to lose the data on that instance if it is stopped.
  - **You can reboot both and you're not going to lose any of your data.** It's basically only if the host fails, and that EC2 instance stops and it's instant store, then you will lose your data.
  - By default, both root volumes will be deleted on termination. However, with EBS volumes, you can always tell AWS to keep the root device volumes. You can't do that with instance store!

ENI vs. ENA vs. EFA

- **ENI** - is, **E**lastic **N**etwork **I**nterface. По сути, это просто виртуальная сетевая карта. Вот и все. Позволяет использовать основной частный IPv4-адрес, из диапазона адресов IPv4 вашего VPC. А также позволяет использовать один или несколько вторичных частных IPv4 адресов из диапазона IPv4 из диапазона адресов вашего VPC.
- **Enhanced Networking** - uses what's called single root I/O virtualization or SR-IOV. Используется для обеспечения высокопроизводительных сетевых возможностей для неподдерживаемых типов экземпляров. **ENA** - это подмножество **Enhanced networking**.
    - **SR-IOV - это метод виртуализации устройств, который обеспечивает более высокую производительность ввода-вывода и более низкую загрузку процессора**, по сравнению с традиционными сетевыми интерфейсами. Таким образом, это просто способ ускорить работу сети.
    - Enhance Networking можно включить с помощью двух методик, Эластичный сетевой адаптер или ENA, который поддерживает скорость сети до 100 гигабит в секунду для поддерживаемых типов экземпляров.

You've got a EC2 instances doing really in, you know, extreme network workloads and you need like, 50 gigabits per second, up and down. Should you use an ENI or multiple ENI's or should you use an Elastic Network Adapter, you want to use Elastic Network Adapter, because it's built for those speeds

- **EFA** - **E**lastic **F**abric **A**dapter, это сетевое устройство, которое вы подключаете к своему экземпляру EC2 для ускорения высокопроизводительной работы. (to accelerate EC2 Insatnce - High Performance Compute, so HPC and machine learning applications.)
    - Если вы видите вопрос сценария, в котором упоминается HPC, или машинное обучение, или спрашивают об обходе ОС, то лучше выбрать адаптер Elastic Fabric Adapter.

Encrypted Root Device Volumes and Snapshots

- Нужно взять незашифрованный корневой том устройства и создать его снимок. Затем скопировать снимок и включить шифрование. Затем вы создаете AMI из этого снимка. И затем вы можете запустить этот экземпляр EC2 с зашифрованным томом корневого устройства.
    - Snapshots of encrypted volumes are encrypted automatically. Если вы сделаете снимок зашифрованного тома, то и снимок будет зашифрован автоматически.
    - Volumes restored from encrypted snapshots are encrypted automatically.
    - You can **share** snapshots, but only if they are unencrypted! (Эти снимки также могут быть переданы с другими учетными записями AWS или сделать их общедоступными, но они должны быть незашифрованными.)
    - **Теперь можно шифровать тома корневого устройства при создании экземпляра EC2.**

Spot Instances and Spot Fleets

- **Spot Instances** - в основном позволяют вам воспользоваться преимуществами неиспользуемых мощности EC2 в облаке, и точечные экземпляры доступны со скидкой до 90% скидкой по сравнению с ценами по требованию, и вы можете использовать спот экземпляры для различных отказоустойчивых и гибких приложений без статических отклонений, таких как большие данные, контейнерные рабочие нагрузки, CICD, веб серверы, высокопроизводительные вычисления, а также другие рабочие нагрузки для тестирования и разработки.
- **Так где же можно использовать точечные экземпляры? Ключевое слово - в гибких приложениях.** Если ваше приложение может быть завершено с уведомлением за одну-две минуты, а затем запустить его снова, когда вы получите лучшую цену, тогда это будет отличным вариантом использования точечных экземпляров.

- Принцип работы **Spot Instances** заключается в том, что у вас **есть спотовая цена**, и вы в основном определяете максимальную спот-цену, и экземпляр будет предоставлен до тех пор, пока спот-цена ниже вашей максимальной спот-цены.

- Если спотовая цена поднимается выше вашего максимума, у вас есть две минуты, чтобы решить, следует ли вам завершить работу ваших экземпляров EC2. Теперь вы можете предотвратить это с помощью спот-блоков. **Спот-блоки** - позволяют вам остановить прекращение работы ваших Spot Instances. даже если спот-цена превысит вашу максимальную спот-цену. И вы можете установить спот-блоки на срок от одного до шести часов.

Spot Instances are useful for the following tasks:

1.

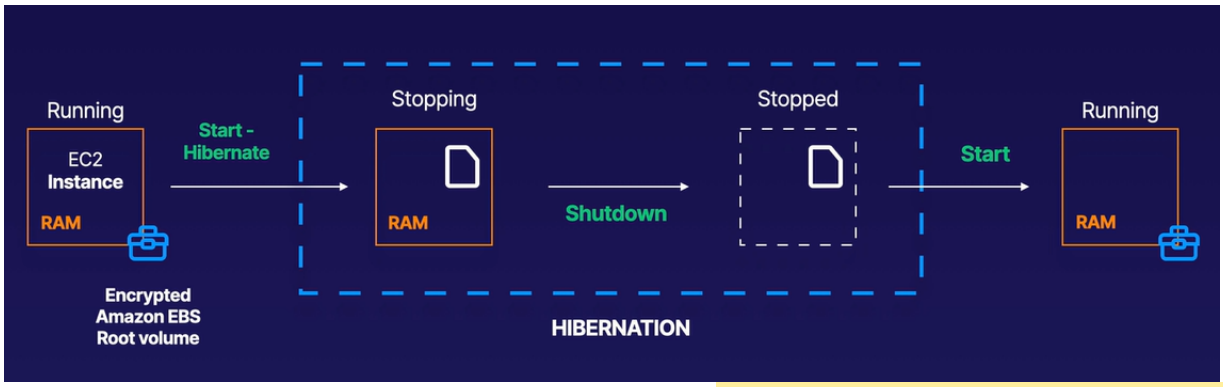
Spot Instances are NOT good for:

1. Persistent workloads
2. Critical jobs
3. Databases

- **Spot Fleet** - is a collection of Spot Instances and, optionally, On-Demand Instances.

EC2 Hibernate

Этот процесс занимает оперативную память и сохраняет ее в корневом томе EBS. И это ускоряет загрузку при перезагрузке.
Итак, когда вы запускаете свой экземпляр из спящего режима, корневой том Amazon EBS восстанавливается в прежнее состояние, а затем содержимое оперативной памяти перезагружается, и процессы, которые вы ранее запускали на экземплярах, возобновляются.
Таким образом, вам не нужно запускать операционную систему заново, а затем запускать SQL. По сути, это как способ развертывания ваших экземпляров EC2 намного быстрее. И ранее подключенные тома данных подключаются заново, и экземпляр сохраняет свой ID экземпляра.
То есть он сохраняет тот же идентификатор экземпляра. В то время как если мы остановим экземпляр и перезапустим его, этого не произойдет.



Мы собираемся запустить процесс гибернации. Это остановит наш экземпляр EC2 и сохранит эту оперативную память на наш том EBS. Затем он будет выключен и остановлен, а когда мы его запустим, он возьмет оперативную память из тома EBS и загрузит ее обратно в оперативную память.
Поэтому нам не нужно делать такие вещи, как перезапуск операционной системы или перезапуск наши приложения и так далее.
Таким образом, при использовании EC2 hibernate экземпляр загружается гораздо быстрее. Операционной системе не нужно перезагружаться, потому что состояние в памяти сохраняется. И это полезно для таких вещей, как долго работающие процессы или службы, которые требуют много времени для инициализации.

- If you are going to use hibernation, the root device volume must be encrypted.
- Instance RAM must be less than 150 GB.
- It's available for **on-demand instances** as well as **reserved instances**.

CloudWatch 101

- **CloudWatch** - is a monitoring service to monitor your AWS resources, as well as the applications that you run in AWS.

  CloudWatch can monitor things like:
  - Compute:
    - EC2 instances
    - Autoscaling Groups
    - Elastic Load Balancer
    - Route53 Health Checks
  - Storage & Content Delivery:

        - EBS Volumes

Host Level Metrics Consist of:

- [ ]

- **CloudTrail**, по сути, представляет собой систему видеонаблюдения для вашей среды AWS - увеличивает видимость активности ваших пользователей и ресурсов путем записи действий консоли управления AWS и вызовов API.
    - Поэтому каждый раз, когда вы заходите и создаете S3 bucket или экземпляр EC2, вы, по сути, совершаете API-вызов к AWS, и все это записывается с помощью CloudTrail, и вы можете определить, какие пользователи и учетные записи обращались к AWS, IP-адрес источника. с которого были сделаны эти вызовы, и когда эти вызовы были сделаны.
- Каждый раз, когда вы видите CloudTrail, я хочу, чтобы вы представили себе большую камеру видеонаблюдения.
- Всякий раз, когда вы видите CloudWatch, я хочу, чтобы вы представили себе производительность.
    - **CloudWatch следит за производительностью.**
    - **CloudTrail отслеживает вызовы API в платформе AWS.**
- CloudWatch с EC2 отслеживает события каждые пять минут по умолчанию, но вы также можете включить детальный мониторинг и это даст вам одноминутные интервалы. (Но это платно)
- You can also create CloudWatch alarms which trigger notifications.

AWS Command Line

- You can interact with AWS from anywhere in the world just by using the command line (CLI)
- You will need to set up access in IAM

Identity and Access Management Roles

- Roles are more secure than storing your access key and secret access key on individual EC2 instances.
- Roles are easier to manage.

- Roles can be assigned to an EC2 instance after it is created using both the console & command line.
- Roles are universal - you can use them in any region.

Instance Metadata

- **Metadata** - is used to get information about an instance, such as its public IP.
    - curl http://169.254.169.254/latest /user-data
    - curl http://169.254.169.254/latest /meta-data

EFS

- **EFS** - **E**lastic **F**ile **S**ystem - is a file storage service for Amazon's Elastic Compute Cloud or EC2 instances. (Так что это похоже на EBS, за исключением того, что в случае с EBS, у вас не может быть двух экземпляров EC2, которые совместно используют том EBS. Однако они могут совместно использовать том EFS.
  Таким образом, EFS - это простой способ использования и простой интерфейс, который позволяет создавать и настраивать файловые системы быстро и легко, а с EFS емкость хранилища является эластичной, увеличиваясь и уменьшаясь автоматически по мере добавления и удаления файлов, так что ваши приложения имеют хранилище, которое им нужно, и когда оно им нужно.
    - EFS - это способ иметь общие файловые системы или хранилища с помощью NFS между различными экземплярами EC2, очень легко настроить и запустить, а затем вам не нужно беспокоиться о синхронизации файлов между экземплярами EC2.
        - Supports the Network File System version 4 (NFSv4) protocol.
        - You only pay for the storage that you use (no pre-provisioning required).
        - Она может масштабироваться до петабайтов и может поддерживать тысячи одновременных соединений NFS
        - And then data is stored across multiple availability zones в пределах одного региона.
        - Read after write consistency.

Amazon FSx for Windows and Amazon FSx for Lustre

- A managed Windows Server that runs Windows Server Message Block (SMB) - based file services.
- Designed for Windows and Windows applications.
- Supports AD user, access control lists, groups and security policies, along with Distributed File System (DFS) namespaces and replication.

Amazon FSX для Lustre - is a fully managed file system that's optimized for compute intensive workloads, such as high performance computing, machine learning, media and data processing workflows, and electronic design automation or EDA. (Вы можете запускать и работать с файловыми системами Lustre, которые могут обрабатывать огромные массивы данных со скоростью до сотен гигабайт в секунду пропускной способности, миллионов IOPS и субмиллисекундных задержек.)

Чем же Lustre FSX отличается от EFS?

- Он специально разработан для обработки рабочих нагрузок, таких как машинное обучение, высокопроизводительные вычисления, обработка видео, финансовый мониторинг и автоматизация электронного проектирования, и позволяет запускать и работать с файловой системой, которая обеспечивает субмиллисекундный доступ к вашим данным и позволяет читать и записывать данные со скоростью до сотен гигабит в секунду пропускной способности и миллионов IOPS.
- В то время как EFS в качестве менеджера и файловой системы NAS, он использует версию сетевой файловой системы для протокола, и это один из первых протоколов сетевого обмена файлами, который был встроен в Unix и Linux.

EC2 Placement Groups

Three different types of placement groups:

1. **Clustered placement groups** - это, по сути, группировка экземпляров в одной зоне доступности. *Используется для для низкой задержки в сети и высокой пропускной способности сети!*
2. **Spread placement groups** - предназначены для защиты экземпляров EC2 от аппаратных сбоев, но отдельные экземпляры размещаются на отдельных стойках в одной зоне доступности либо в разных зонах доступности в зависимости от того, как вы его настроите. *Рекомендуются для приложений в которых есть небольшое количество критически важных экземпляров. которые должны быть отделены друг от друга*.
3. **Partitioned placement group** - Amazon EC2 делит каждую группу на логические сегменты, называемые разделами, и Amazon EC2 гарантирует, что каждый раздел внутри группы размещения находится в своем собственном наборе стоек. При этом каждая стойка имеет собственную сеть и источник питания, и никакие два раздела в группе размещения не используют одни и те же стойки, что позволяет изолировать влияние аппаратного сбоя в вашем приложении.

По сути, Spread placement groups - предназначена для одного экземпляра, Partitioned placement group - для нескольких экземпляров, а Clustered placement groups - это просто способ расположить все как можно ближе друг к другу. Кластерная группа размещения не может охватывать несколько зоны доступности, в то время как распределенное размещение и разделенная группа размещения могут, но они все равно должны находиться в одном регионе.

- The name that you specify for a placement group must be unique within your AWS account.
- Only certain types of instances can be launched in placement groups.
- AWS recommend homogenous instances, within clustered placement groups.
- You can't merge placement groups.
- You can move an existing instance into a placement group, but before you move the instance, the instance must be in the stop state.
- And you can move or remove an instance using the AWS-CLI or SDK, but you can't do it via the console just yet.

HPC on AWS

High **P**erformance **C**ompute - is used for industries such as genomics or finance and financial risk modeling, machine learning, you've got weather prediction and even autonomous driving.

1. **AWS batch** - позволяет разработчикам, ученым, и инженерам легко и эффективно выполнять сотни тысяч пакетных вычислительных заданий на AWS. AWS batch поддерживает многоузловые параллельные задания, что позволяет запускать одно задание, охватывающее несколько EC2, и вы можете легко планировать свои задания и запускать EC2 экземпляры в соответствии с вашими потребностями.
2. **AWS parallel cluster** - is an open source cluster management tool that makes it easy for you to deploy and manage HPC clusters on AWS and the way it does this is it uses a simple text file to model and provision all the resources needed for your HPC applications in an automated and secure manner and it allows you to automate the creation of VPCs, subnets, cluster types and instance types.

AWS WAF

- **WAF** - это брандмауэр веб-приложений, который позволяет отслеживать HTTP и HTTPS запросы. которые направляются на Amazon CloudFront, балансировщик нагрузки приложений или шлюз API. По сути, это позволяет вам контролировать доступ к вашему содержимому. В случае с HTTP и HTTPS это происходит на уровне приложений.

AWS WAF allows three different types of behavior:

1. Allow all requests except the ones that you specify.
2. Block all requests except the ones that you specify.
3. Count the requests that match the properties that you specify. (это своего рода пассивный режим, в котором он просто будет считать запросы, которые соответствуют указанным вами свойствам.)

- So in terms of the extra protection that WAF provides, it protects against web attacks using conditions that you specify. And you can define conditions by using characteristics of web requests such as, the IP address that it originates from, the country that requests originate from. У вас может быть эмбарго против какой-то страны. и вам нужно заблокировать их доступ к вашему сайту. Как вы это сделаете? Использование AWS WAF - определенно один из правильных ответов.
- And then WAF can also detect things like SQL code that is likely to be malicious. So this will be things like SQL injections.
- As well as the presence of a script that is likely to be malicious. This is sometimes known as cross-site scripting.

QUIZ

- **EBS uses Block-based storage**, where the data is stored on a virtual disk managed by the Operating System. EFS uses File-based storage, where the underlying filesystem is managed by AWS. S3 uses Object-based storage, where files are kept in a flat structure
- **EFS** uses the NFS protocol, and is explicitly **not supported on Windows**.
- AWS originally used a modified version of the **Xen** Hypervisor to host EC2. In 2017, AWS began rolling out their own Hypervisor called **Nitro**.
- Changes to IAM Policies take effect almost immediately (with maybe a few seconds delay).
- EBS snapshots use incremental backups and are stored in S3.
- It possible to perform API actions on an existing Amazon EBS Snapshot through the AWS APIs, CLI, and AWS Console. You can use AWS APIs, CLI or the AWS Console to copy snapshots, share snapshots, and create volumes from snapshots.
- **Enhanced networking** uses single root I/O virtualization (**SR-IOV**) to provide high-performance networking capabilities on supported instance types.
- If the original snapshot was deleted, then the AMI would not be able to use it as the basis to create new instances. To delete an EBS Snapshot attached to a registered AMI, first remove the AMI, then the snapshot can be deleted.

- Standard Reserved Instances cannot be moved between regions. You can choose if a Reserved Instance applies to either a specific Availability Zone, or an Entire Region, but you cannot change the region

---

Databases on AWS