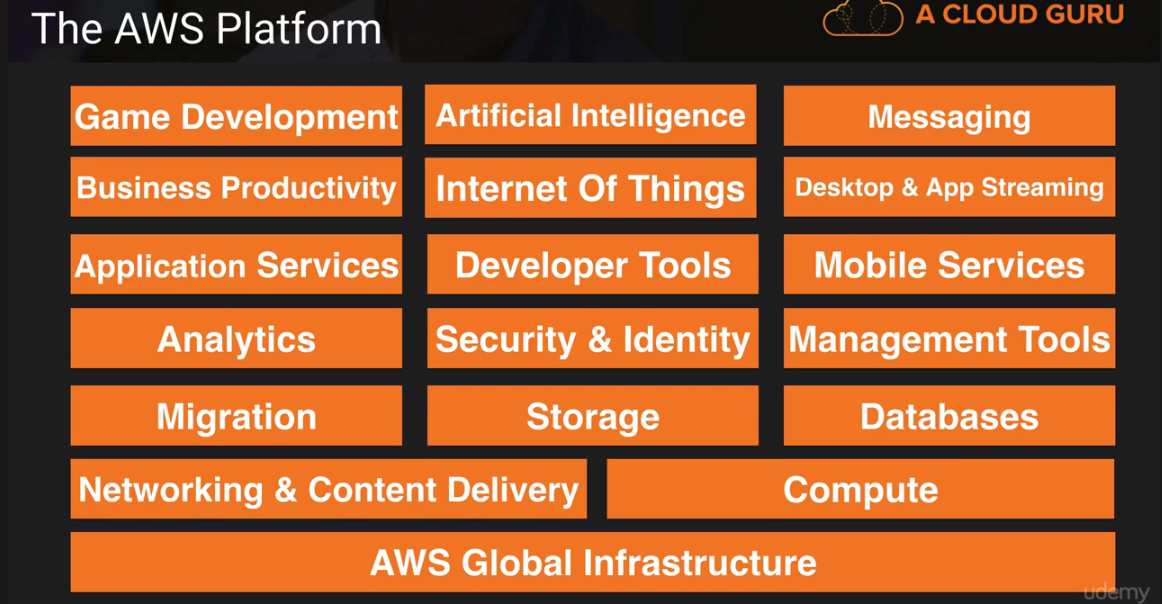
* Following is the AWS Platform diagram:

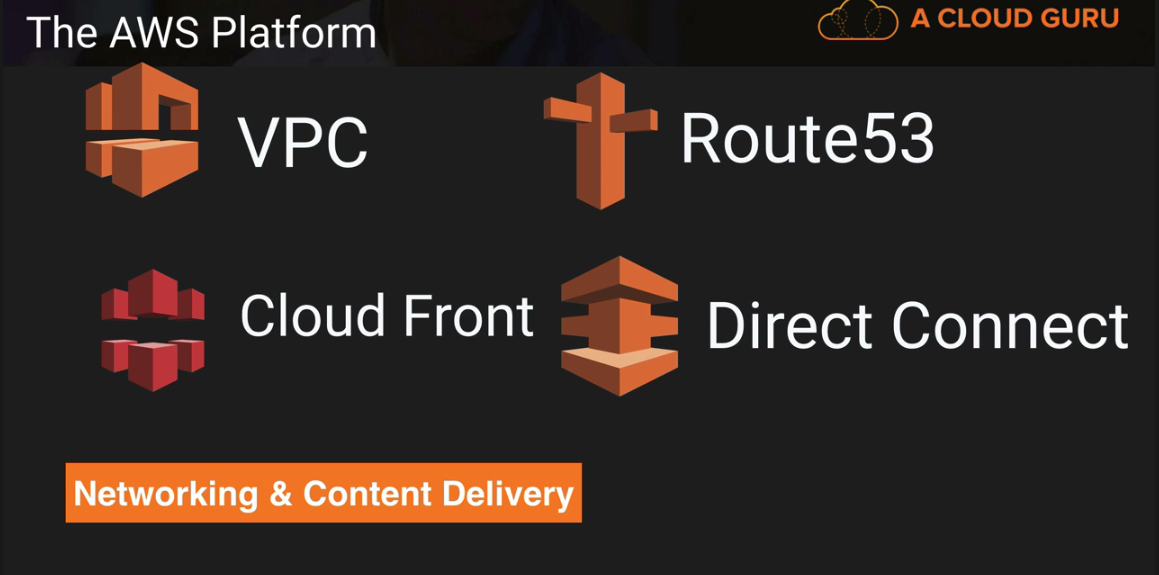
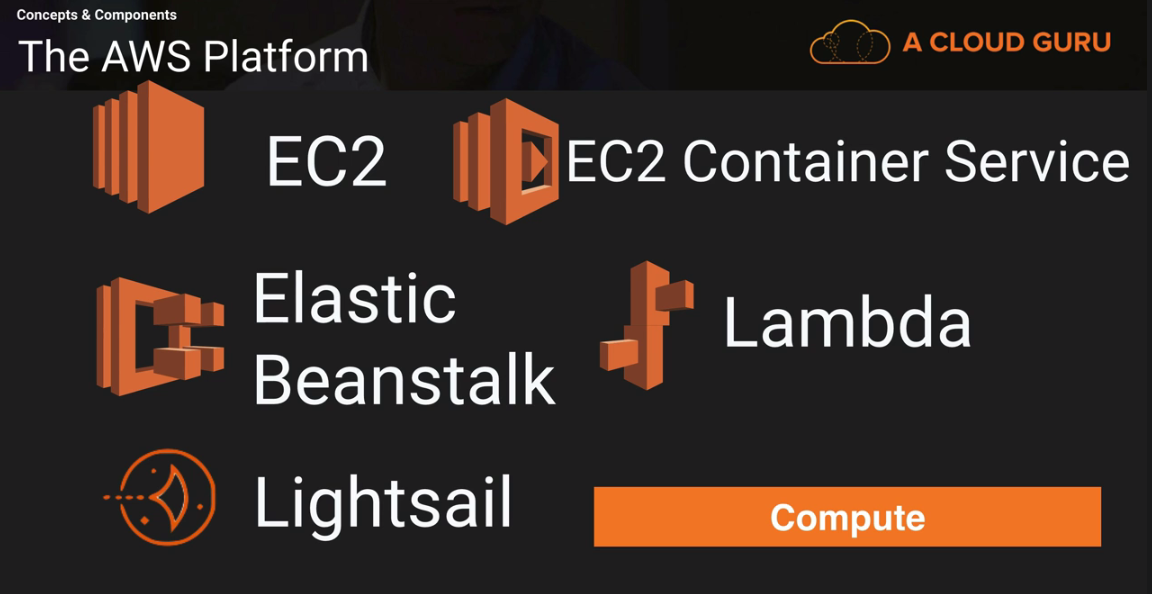
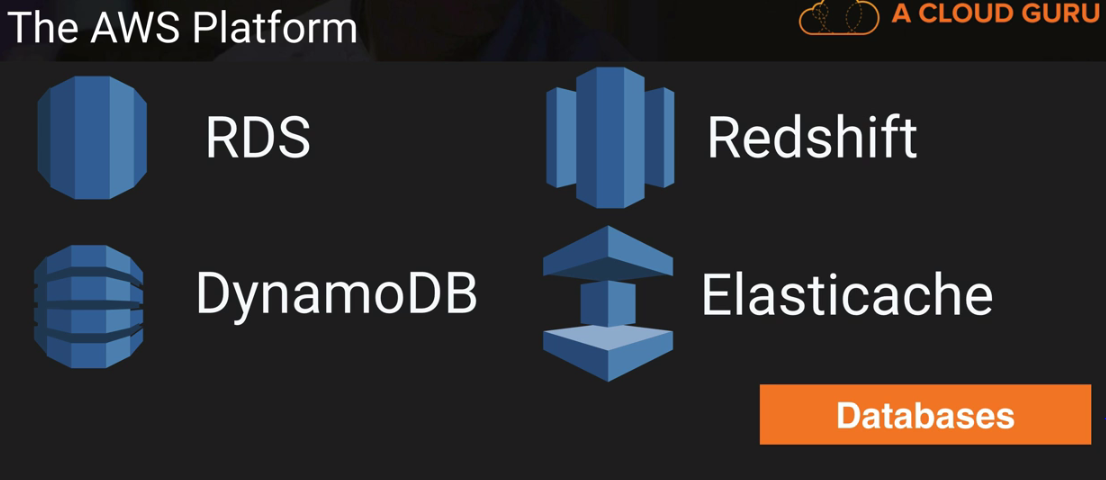
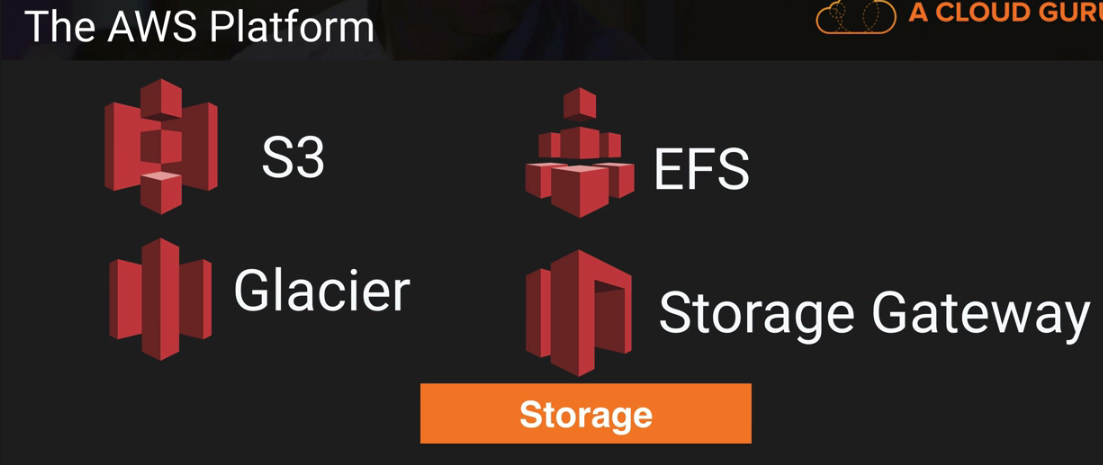


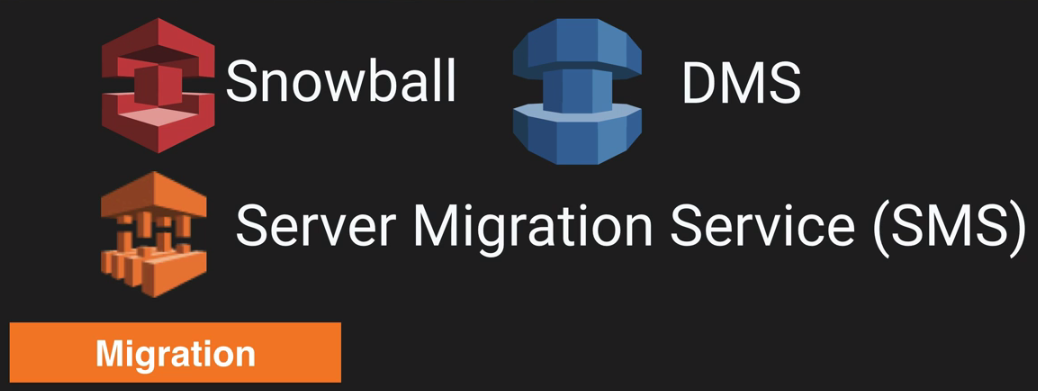
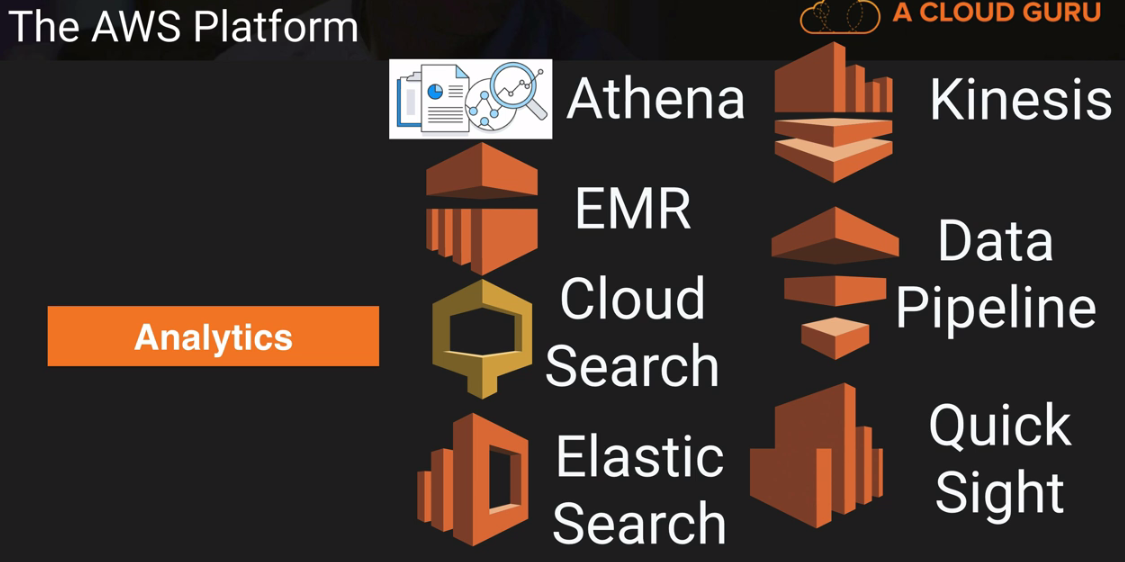
* AWS never initiate the movement of data b/w regions unless customer explicitly enable it. AWS infrastructure contain s

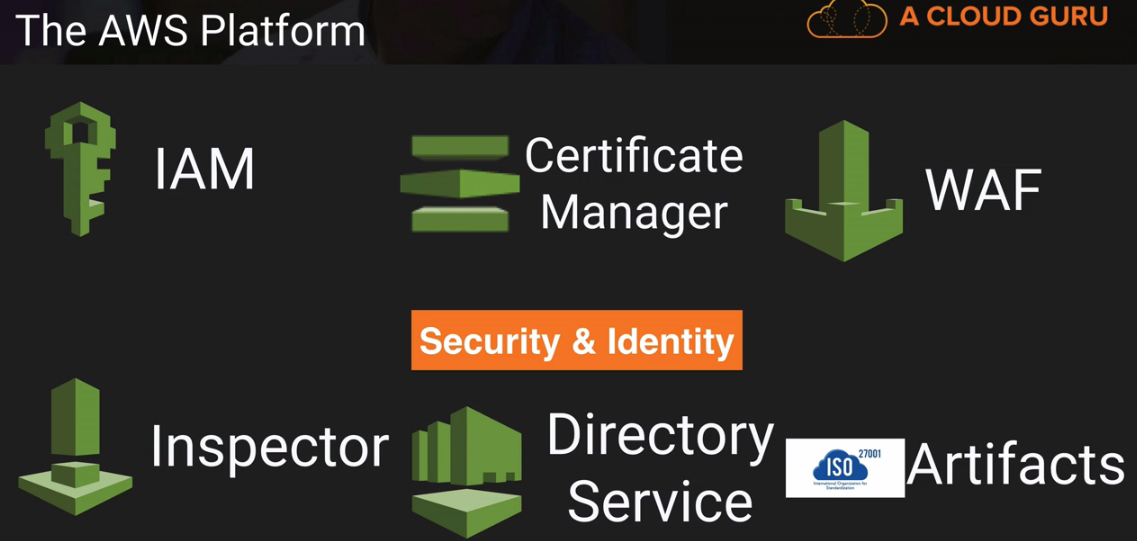
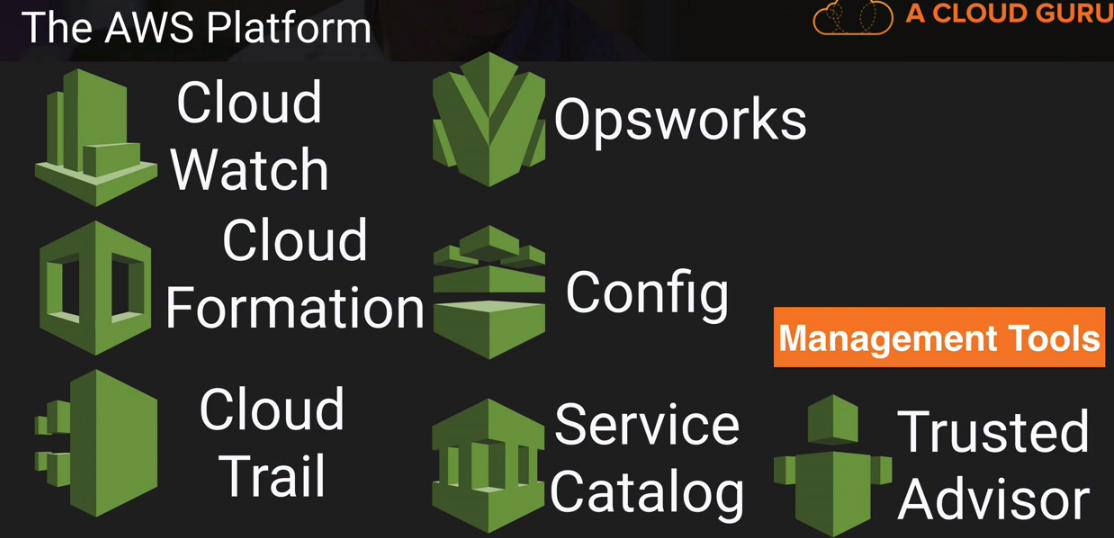
1) Region

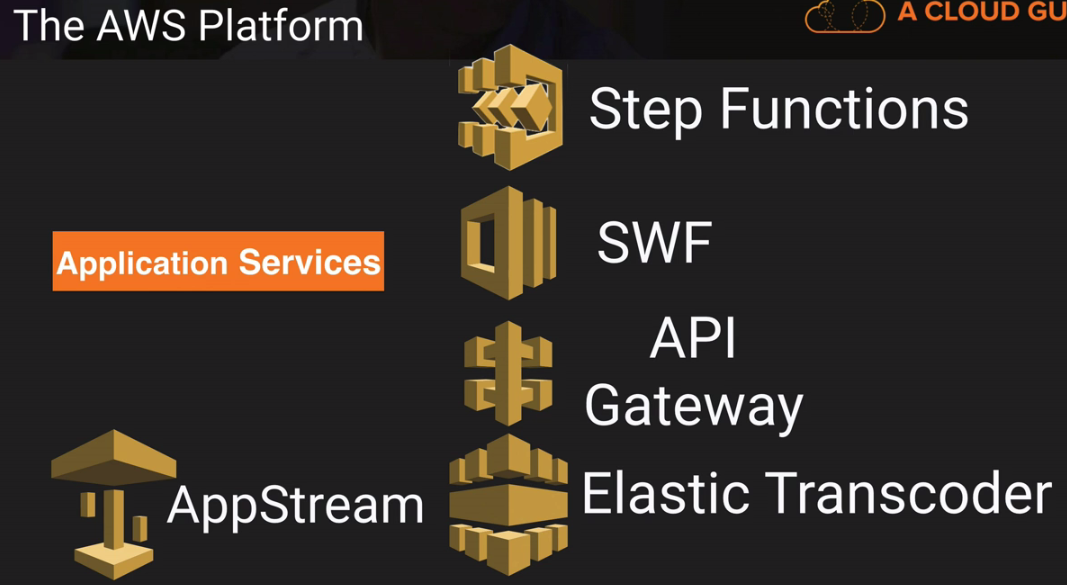
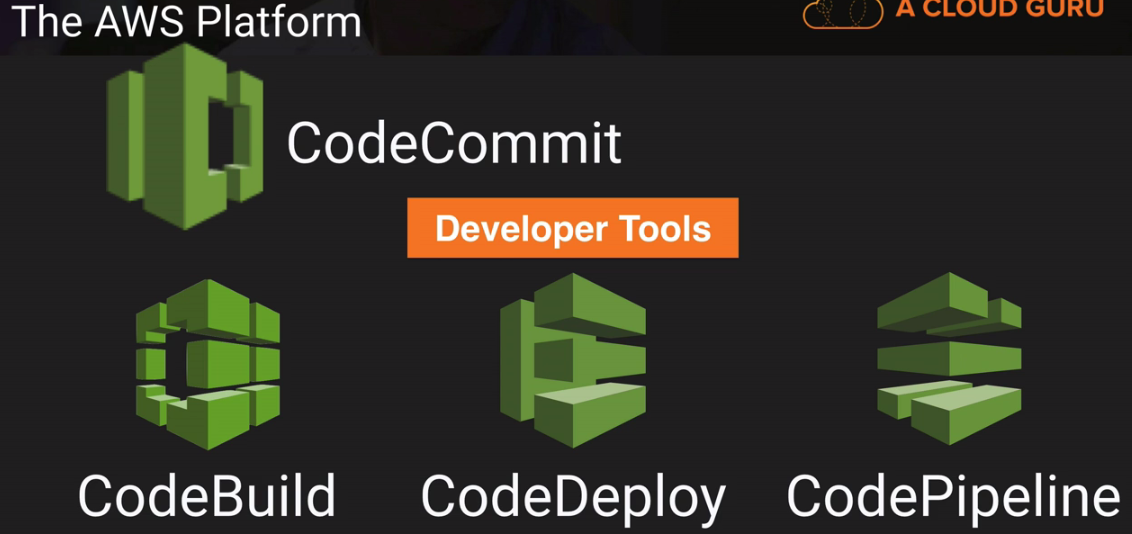
2) Availability Zone

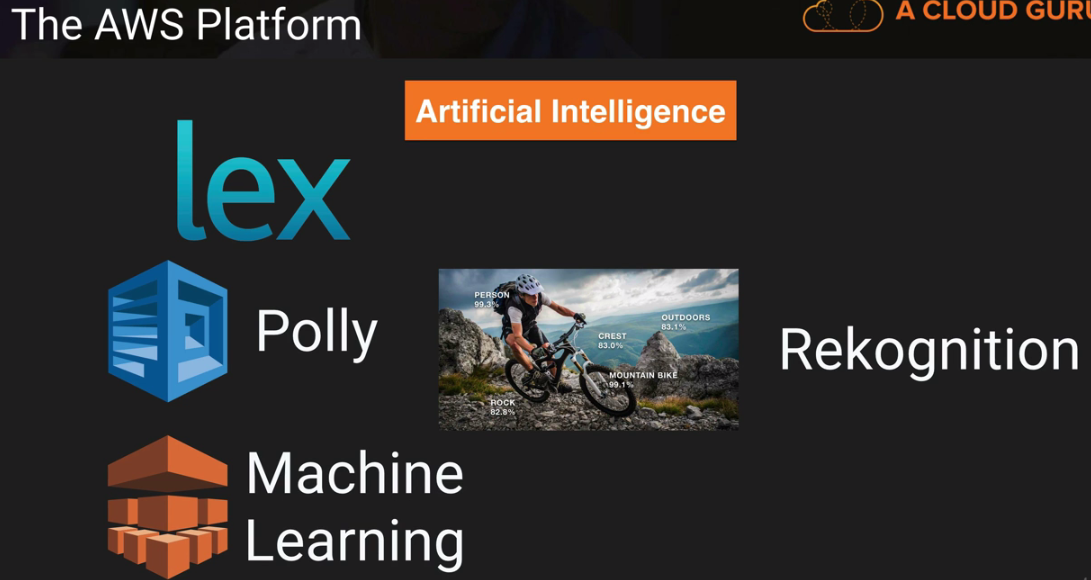
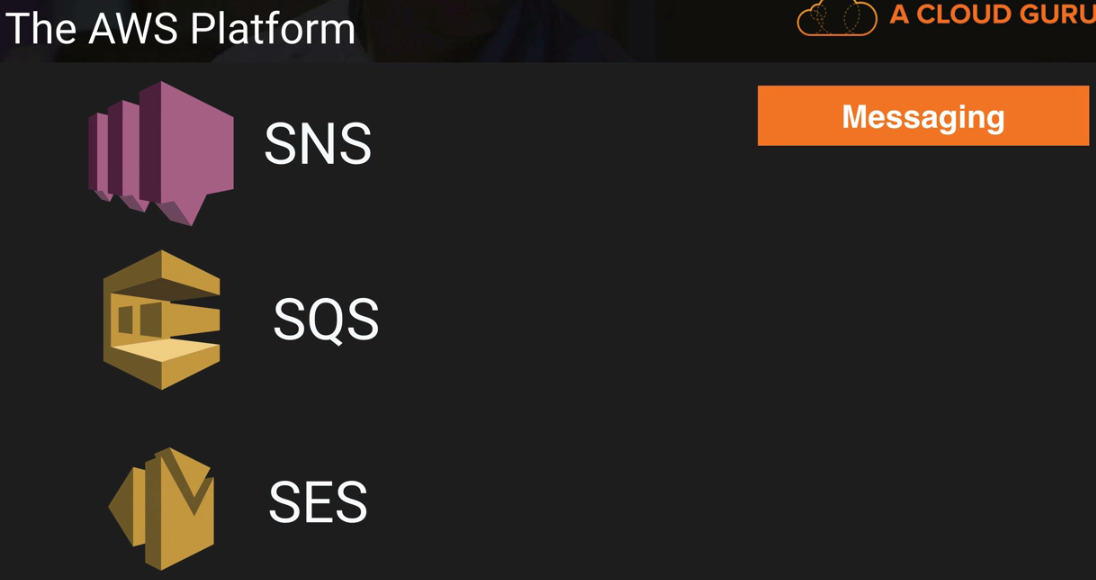
3) **Edge Location** which are CDN/cached (Content delivery n/w) endpoint for **cloud front**

* Route 53 is DNS service as 53 is DNS port while Direct Connect is direct connection b/w AWS and enterprise n/w.
*  
* **LightSail** is out of the box cloud to deploy WordPress or Jumla site for you automatically.
* **Glacier** stores archived files from S3. While **Storage Gateway** connects enterprise’s private n/w with cloud S3 via VM machine image.
* 
* **Snowball** (edge) migrate data by import/export from hard-disk while **DMS**(Database) migrates data from DB w/o downtime.

* **Athena** is tool to run SQL query on S3 data while **EMR** (elastic map reduce) is for analyzing Big Data like logs same as Kinesis. **Cloud Search** is fully managed service while **Elastic Search** are based on Open source. **Kinesis** is used for collating large amounts of data streamed from multiple sources. **Data Pipeline** is used to flow data from S3 to DynamoDB while **Quick Sight** is used to create dashboard for data.
* **Cloud watch** monitors health of EC2.
*  

* **IAM** is Global i.e. if you create a user/role in IAM, it will be available globally.



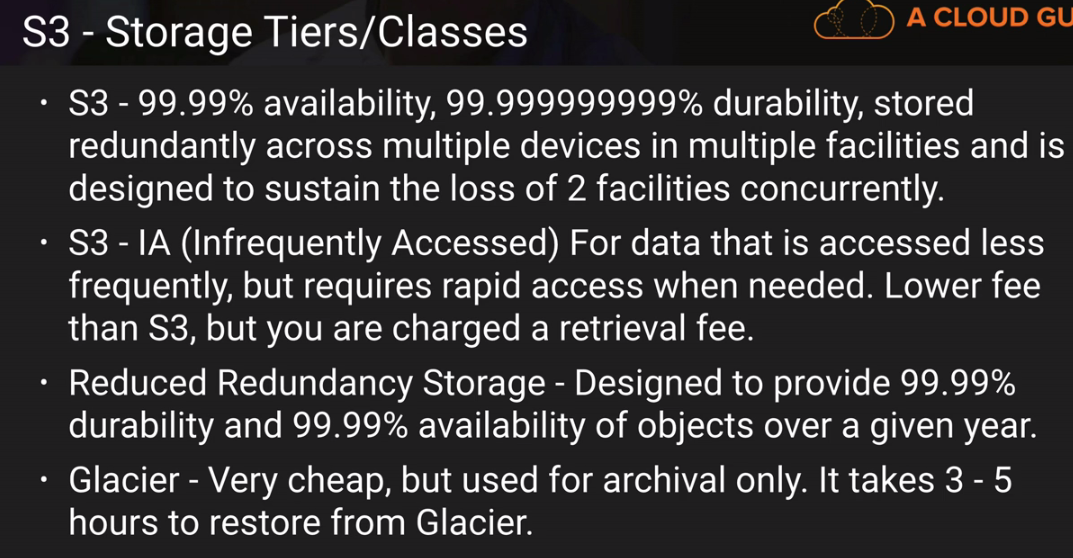
* **Root Account** is created with AWS account and it has complete admin access. **Power User** Access allows accesses to all AWS services except for management of groups and users within IAM. Role can be of 3 types:

1. **AWS Service Roles** are for AWS services to access other AWS service
2. **Cross Account Access** is used for accessing one account from other.
3. **Identity Provider Access** is used to interact with Facebook or Linkedin.

* When a new User is created, it does not have any access to any AWS services. Multi factor authenticate is authentication with physical or virtual device on PC or Phone.
* You can log in to the AWS console using password only while to access from sdk, use Access Key Id and Secret Access Key.
* Policy example:{ "Version": "2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "s3:CreateBucket ], "Resource": "arn:aws:s3:::dms-\*" } ] }
* To setup billing alarm. Goto Name>MyBilling Dashboard>Enable Alarm>Save Preference>Manage>Gotot CloudWatch> Billing > Set Alarm>Confirm email address.
* There are two types of storage Object/file Based Storage and Block Based Storage (where you can install OS or can be shared within n/w.). S3 is object based storage while EFS and EBS are block based storage.
* File size can be from 0 Byte to 5 TB in S3. The largest object that can be uploaded in a single PUT is 5 gigabytes. For objects, larger than 100 megabytes, customers should consider using the Multipart Upload capability. S3 provide unlimited file storage in buckets. S3 is universal namespace, that is, bucket cannot be named same as of bucket of another person. The URL pattern is:

https://s3-<domain>.amazonaws.com/<bucket-name>

* S3 is eventual consistence for modify/delete operation while you can read file immediately after write into S3. But its atomic, that is, either you’ll get old data or new data but not corrupted data.



* **S3** is built for 99.99% Availability and 11\*9’s (2 before decimal) durability and sustain the loss of 2 facilities concurrently. On the other hand, Reduced Redundancy can tolerate 1 concurrent facility failure.
* **S3- IA**, availability SLA is 99% against 99.9%, minimum object size is 128KB and retrieval fee is for per GB. To move object in IA, it needs to be 128Kb in size and 30 days old. But if you don’t enable archiving to S3-IA, you can move object to Glacier next day you upload to S3.
* In S3, data is stored in **KEY’s** alphabetically order geographically so if you store log files named start with same format, these files will be stored in same area which will impact S3 performance.
* **Glacier** stores data as little as $0.01 per Gb per month but retrieval time is 3-5 hours. Glacier is designed to keep object 90 days minimum. So even if object is deleted before, 90 days’ charge will be applicable.
* Amazon **S3 Transfer Acceleration** enables fast, easy and secure transfer of files over long distances by taking advantage of Amazon **Cloud Front** globally distributed edge locations
* S3 provide **AES-256** server side encryption where data will be encrypted before storing in S3.
* In S3, there are two type of permission management:

1) Manage User for users present in IAM

2) Manage Public permission for AWS users or public.

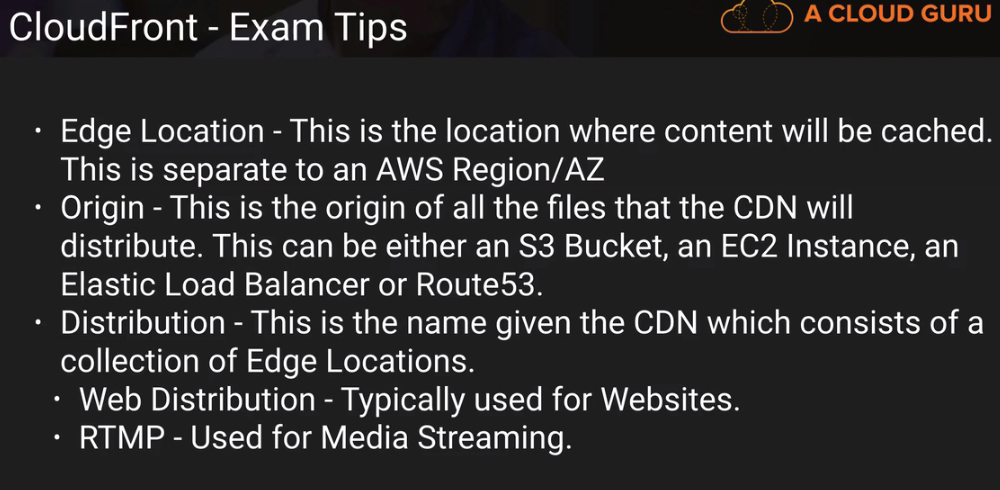
* By default, all the buckets and object in S3 are privately access until you provide access. In S3, there are two type of access granted:

1) Objects which allows to create/modify/delete objects in bucket

2) Object Permission which allows to grant permissions of create/modify/delete to other users also.

Even you enable everyone read for bucket, for each object you need to enable it.

* Once versioning is enabled in S3, it cannot be removed unless you delete the bucket though it can be disabled.
* While designing system for most cost-effective ways with very large files, we might want versioning disable as if file is of 4GB and we change this slightly, total storage in S3 will be 8GB for single file.
* If you delete version, it cannot be restored but if you delete any object in S3, it can be restored by removing the marker of delete operation from old view if versioning is enabled.
* For cross region replication, versioning should be enabled in both buckets and same region replication is not allowed. While enabling cross region replication, only new object replicate. Existing elements will not. But if you update an existing object, it will replicate all previous version also in destination bucket. Even permissions are replicated.
* You can’t replicate to two different buckets as well as A->B->C. The object will be replicate in C, only if, objects are written to B manually.
* If you delete any element from bucket, deletion will be replicated to another bucket. But if you delete "Delete Marker", your object will be restored in main bucket but will not in replicated bucket. The same is valid for version also. If you remove any version from main bucket, you can still restore that version from destination bucket.
* Lifecycle can be done with versioning on/off. For S3 Lifecycle management, if you select "Action on current version" and delete, it will put a delete marker but if you select "Action on previous version" and select delete. It will delete object permanently.
* By default, TTL is **24 hours** in edge locations. You can clear the data before that time, but it will be charged.



* S3 can be configured to log every request to S3 in another bucket. There are two types of encryption:

1. **In-transit** i.e. while upload/download from the S3 to local which is managed by SSL/TSL
2. **At Rest** where we have Server side encryption in 3 different flavors:
3. **S3 Managed Key** SSE-S3 where S3 managed key as well as master key to encrypt keys with ASC 256.
4. **Key Managed Service**, SSE-KMS where keys are protected with envelop keys generated uniquely for region as well as we get order trail of S3 bucket decryption usage.
5. **Customer Provided key** SSE-C where keys are provided by user.
6. **Client side** Encryption

* There are 4 types of **Storage Gateway**:

1. **File Gateway(NFS)** to store flat files into S3, accessed through a Network File System(NFS) mount point. Ownership and permission stored as object metadata in S3.
2. **Volumes Gateway(iSCSI)** is used to stored block based storage. These can be backed up asynchronously as EBS snapshot incrementally on cloud. These can be of 2 types:
3. **Stored Volumes** (Gateway stored volumes) to store copy of entire dataset to on premise. Back up EBS snapshot can be 1 GB-16 TB in size.
4. **Cached Volumes** (Gateway cached volumes) to store most recent access data. 1GB to 32 TB in size.
5. **Tape Gateway(VTL)** (Gateway virtual tape library) to store archive data in virtual tapes which can be transferred to S3 and Glacier. It uses popular applications like Netbackup, Backup Exec, Veam etc.

* **AWS Import/export** was used to transfer data to AWS with physical disk but problem, was there were many disk types so AWS comes with **Snowball (80 TB)** or **Snowball Edge (100 TB)**. These can be used to import/export data from S3. If you want to transfer data from Glacier, It has to be transferred first to S3.
* **S3 Transfer Acceleration** utilizes Cloud Front Edge network to faster data upload to S3 with separate URL: **<bucket>.s3-accelerate.amazonaws.com**
* When you enable Static Website hosting URL looks like : http://<bucket-name>.s3-website-<region>.amazonaws.com
* RTMP distribution speeds up the distribution of your streaming media files using Adobe Flash Media Server's RTMP protocol. A single distribution allows to have different origin which can be distinguish with origin id. Smooth Stream service is used to Microsoft Smooth Stream APIs.
* Restrict Viewer Access (Use Signed URLs or Signed Cookies) allowed your content to be viewed by specific group of users.
* EBS automatically handle the deployment from capacity provisioning, load balancing, auto-scaling to application health monitoring based on the code you upload to it whereas cloud formation is an automated provisioning engine designed to deploy entire cloud environment via a json.
* EC2 with volume type Instance Store, you can add more than 1 Instance Volume bit before it’s launched. One you launch that instance, you can add only EBS instances to it. And you cannot stop the EC2 instance with Instance Store volume type. EC2 Instance stop is required if underlying hypervisors is failed and need instance restart to reassign.
* As Instance store templates are stored in S3, it will take more time to start EC2 instances.
* Application load balancer does routing on Application layer while classic loadbalancer does it on TCP Layer.
* Load Balancer are not getting Public IPs like EC2 Instances.
* By default, AWS support 4 kinds of metrics : 1) Disk, 2) CPU, 3) N/w and 4) Status Check.
* CloudWatch monitors the EC2 environment while CloudTrail audit them
* We can attach a new IAM role to EC2 instance after its created, only through command line not using console.
* When you provision any security group. All inbound traffic is blocked while all outbound traffic is allowed by default.
* SSL Termination on Load Balancer is supported so that your application wouldn’t have to worry about decrypting it. It also allows to identify originating IP address of client.
* AWS management reevaluate the strategic business plan for Risk at least biannually.
* Kinesis stores data for 24 hours by default but it can be increased to 7 days. On the other hand, kinesis firehose sends data to underlying S3, Redshift (via S3) and Elastic search cluster immediately with Lambda analytic optionally.
* Stream large amount of data in to the cloud, think **Kinesis**, processing large amount of data for business intelligence, think **Redshift** and process large data for Big Data Processing, think **Elastic Map Reduce**.
* Elastic Transcoder convert media file per destination systems. Pay based on the minute that you transcode and the resolution at which you transcode.
* VPC Peering wouldn’t work if they have matching or overlapping CIDR(ip) block and if VPCs are in different regions. Peering does not support transitive peering relation b/w A and C if A->B and B->C.
* AWS support active directory with SAML authentication where you will be authenticated with active directory first then given temporary credentials.
* Through NAT instance, private subnet EC2 instance can access internet but people from internet cannot access(SSH) into these systems. For this purpose, Bastion is used.
* VPC Flow logs help all VPC n/w actions logging to cloud watch.
* IAM role attached EC2 instance, remain attached with instance until it’s deleted.
* To get EC2 instance metadata, curl [http://**169.254.169.254**/latest/meta-data](http://169.254.169.254/latest/meta-data) url in instance.
* Placement Groups are used for low network latency e.g. useful for Casandra database which need fast n/w operations.
* To attached EFS with EC2 Instances, we should select same security groups in both service.
* EFS and EBS both are block based storage service but EFS can be mounted on multiple EC2 instances while EBS can be mounted with one instance only. In EFS, permission can be declared at file and folder level.
* I cannot delete a snapshot of an EBS Volume that is used as the root device of a registered AMI. You must deregister the AMI before being able to delete the root device.
* Command to create the snapshots using the command line tools : ec2-create-snapshot
* If you use CName in Route53 for routing, you will be charged for that. The other option is Alias where you are not charged.
* For each registered domain, AWS creates Hosted Zones and 2 record sets :NS and SOA (start of authority) records
* Route53 is a global service.
* For weighted routing, you can define weight from 0 to 255 and create one “A” record for each weight. AWS will determine the percentage and route n/w accordingly.
* Route 53 support MX Records. A mail exchanger record (MX record) is a type of resource record in the Domain Name System that specifies a mail server responsible for accepting email messages on behalf of a recipient's domain, and a preference value used to prioritize mail delivery if multiple mail servers are available.
* In Route53, we can manage 50 domain names by default, however it is a soft limit and can be raised by contacting AWS support.
* Simple Notification Service SNS can be used for notification of scaling of a new EC2 instance.
* Amazon SES stand for Simple Email Service. when you create a topic on Amazon SNS, an Amazon resource name is created.
* Domain in SWF is collection of related workflows
* \* To access RDS from web application, RDS security group should have inbound from web application security group.
* \* Restored version from Automatic Backup or manual snapshot, will be new RDS instances with a new end point.
* \* To scale up the RDS database, take snapshot and while restore specify storage type.
* \* Multi AZ is only for disaster recovery only. Its not primarily used for improving performance. For performance improvement you need Read Replicas.
* \* Read replica support 5 replica for one database. Read replica is supported for Postgres, MySQL and MariaDB
* \* When you create VPC, it creates N/w ACLs, securoty group and Route Tables by default.
* \* 1 subnet can be attached to one Availability Zone only. It can not be spanned to multiple AZ.
* \* Security groups are stateful so if you allow http inbound traffic, by default outbound traffic on http is also open, while Network Access List are stateless.
* \* By default AWS reserves 251 IPs for any new subnet. AWS reserves 5 IP addresses for their operation: 1) 10.0.0.0: N/W address, 2) 10.0.0.1 : Reserved for VPC router, 3) 10.0.0.2 : Reserved for DNS IP address, as its base of VPC network range plus 2, 4) 10.0.0.3 : Reserved for future use, 5) 10.0.0.255: N/W broadcast address, VPC does not support broadcast.
* \* You cannot attach multiple Internet gateway to VPC.
* \* AWS provided Elastic IPs which can be associated with any EC2 instances.
* \* Elastic Cache supports Master/Slave replication and Multi AZ which can be used to achieve cross AZ redundancy.
* Deploy any EC2 instance in public subnet doen not allow internet. You need to define elastic IP or elastic load balancer.
* \* Elastic cache is a good choice if your database is read heavy and not prone to frequent changes while Redshift is useful if database is feeling stress because of heavy OLAP transaction on it.
* \* Aurora provides uoto 5 time better performance than MySQL at a price point one tenth of a commecial databases.
* \* If you loose primary db server, with Aurora Replicas(15), it will do the failover but with MySQL Read Replicas(5), it will not.
* Flow logs enable you to capture IP traffic flow information for n/w interface in your resources.
* \* DB Instance with Priority tier 0 becomes master copy. Primary DB instance is always accessed by cluster DNS while other can be accessed by instance DNS address.
* Once you create security group for any database, we should update inbound rule to allow access to other security group or ips.
* While restoring database from snapshot, you need to create a new instance always.
* \* NAT G/W and NAT Instance can be used to provide access to private subnet securely. NAT G/W is new and autoscalble as well no need to define secutiy group.
* \* For NAT instance, we should disable source destination check.
* \* N/W ACL created with VPC by default allows all inbound outbound traffic while if you create any custome N/w ACL, inbound-outbound traffic is denied.
* \* One Subnet can be attached to one AZ as well as one N/w ACL only. So if you add one subnet to any ACL, it will be removed from N/W ACL which was created with VPN.
* \* Even if you create inbound and outbound rule, you cannot access public subnet until you open the Ephemeral Ports on 1024-65535.
* \* N/W ACL rules are executed in order. They are not overrides. And to block any IP Addresses , we can use N/W ACL's not Security Group.
* \* To create high available system, we should have 2 publicly available subnet in 2 different AZ.
* \* Amazon EBS volumes are placed in specific availability zones where they are automatically replicate to protect you from failure.
* \* If you need more than 10K IOPS(I/O operations per second), Provisioned IOPS SSD (IO1) is better than General Purpose SSD (GP2) which support 3 IOPS per GB with upto 10000 IOPS.
* \* HDD EBS disk can be used for processing sequecial data.
* \* You cannot mount one EBS with multiple EC2 instances, instead use EFS as well as you cannot mount EBS from different availability zone into EC2 instance from other zone.
* \* One subnet is assigned to one availability zone only.
* \* By default, If you terminate the EC2 instance, boot volume memory will be deleted and one vpc n/w is created as soon as login is created.
* \* To login into putty use ec2-user@<ip address> as address while public DNS for EC2 instance look like ec2-<IP seperated by - in place of .>.<region>.compute.amazonaws.com
* \* System status check verifies that your instance is reachable. We test that we are able to get network packets to your instance. If this check fails, there may be an issue with the infrastructure hosting your instance (such as AWS power, networking or software systems). This check does not validate that your operating system and applications are accepting traffic. Instance status check verifies that your instance's operating system is accepting traffic. If this check fails, you may need to reboot your instance or make modifications to your operating system configuration.
* \* One EC instances can have multiple security group.
* \* Any rule change apply to security applies immediately and if you allows any port in inbound, that port will be allowed for outbound by default.
* \* Snapshots are copy of harddisk as given point to time. Its stored in S3.Every time you take a snapshot of a Volume, only increamental changes will be stored in S3. While recreating Volume from snapshot, you can change the type of Volume.
* The best practice to take snapshot is to stop the EC2 instance and then take snapshot.
* \* RAID are collection of disk which can be worked as single Volume. This can be used when highest disk space available is not enough for your case. The main purpose of RAID is to improve IO. RAID 5 is collection if 3 or more volumes. For example, if you want to use some database like Casandra which Is not supported by AWS
* \* There are two type of load balancer: 1) Application Load Balancer are preferred over classic load balancer.
* AMI stands for Amazon Machine Image.
* \* Response time is time wait in receiving response, it can be 2-60 sec. Interval is time b/w 2 check, it can be 5-300sec. Unhealthy threshold, no of consecutive failure before declaring EC2 instance unhealthy.
* \* AWS does not generate public IP address for load balancer unlike for EC2 instance. It only generates DNS for ELB. So ELB can be accessed by DNS only.
* With window EC2 username is administrator
* \* By default, CPU, Disk Read/Write, N/W and Status check(both for instance as well as for hosting machine) metrics are availbale.