1. Intro

Cloud Market Share

AWS 90% > Azure 5% > Rest 5%

Most Valuable IT Certification, 2016

AWS Certified Solutions Architect – Associate $125,871

PMP $116,094

Six Sigma Green Belt $102,594

Technology Partner Consulting Partner

Alert Logic Logicworks

CloudBerry Lab Rackspace

Sumo Logic Accenture

Datadog Datapipe

New Relic CAPGEMINI

AWS Consulting Partner Program

Standard: 2 Asso./ 0 Prof.

Advanced: 4 Asso./ 2 Prof.

Premier: 20 Asso./ 8 Prof.

AWS Certification Exam Difficulty

Easy>>Dev Asso. > **Sol. Arch. Asso.** > Sysops Asso. > Security Sp.> Big Data Sp. > Devops Prof. > Networking Sp. > Sol. Arch. Prof. >>Difficult

AWS Solutions Architect Certification Exam Scope

Design Resilient Architectures 34%

Design Performant Architectures 24%

Specify Secure Applications and Architectures 26%

Design Cost-Optimized Architectures 10%

Define Operationally Excellent Architectures 6%

Exam Overview

130 minutes

~65 questions (Multiple choice)

Pass mark curved (~70%)

Qualification valid for 2 years (half the price after 2 years)

Practice exam $20

Exam $150

2. AWS – Overview

AWS Global Infrastructure

Regions: Physical geographic location

Availability Zones: Data centers within a region

Each region has 2+ Availability Zones (=data centers)

Edge Locations

Content Delivery Network (CDN) End points for CloudFront

Cache contents

# Edge Locations > # Regions

Compute

EC2 - Virtual Machine

EC2 Container Services

Elastic Beanstalk – Wizard like.

Run a developer's code on an infrastructure that is automatically provisioned to host that code

Lambda – Serverless

Lightsail –

Batch – [Out of scope]

Storage

S3 – Virtual Disk to store objects like files. Not for db. e.g. Dropbox

Glacier – Archive files from S3. Low cost. Takes time to retrieve files.

Elastic File Service (EFS) – Could install application, db, etc.

Snowball – huge thumbdrive (TB’s)

Storage Gateway – Connect S3 to on-prem datacenter

Database

Relational Database (RDS) – Oracle, Aurora, SQL

Dynamo DB – Non-Relational DB. very scalable

RedShift – data warehouse

Elasticache – caching data in the Cloud. Take the load off the DB.

Migration

AWS Migration Hub

Application Discovery Service

Database Migration Service (DMS) – Migrate on-prem db to AWS.

Or within AWS regions.

No downtime

Switch DB. E.g. Oracle to Aurora

Server Migration Service (SMS) – Migrate on-prem VM to AWS. [out of scope]

Networking & Content Delivery

Virtual Private Cloud (VPC)

Virtual Datacenter

Deploy your assets

Route53

DNS service (Domain – IP)

Domain name registration available

CloudFront

Content Delivery Network

Stores/ Caches contents

API Gateway

Direct Connect

Dedicated line

Reliable connection and security

Management Tools

Cloud Watch – Monitor Performance, Resource Utilization

Cloud Formation – Turns infrastructure into code

Cloud Trail – Log changes to AWS environment

Opsworks – Automate deployments

Config – monitors and configures with point in time snapshot

Trusted Advisor – Automatically scans env and make suggestions

Service Catalog [out of scope]

Systems Manager [out of scope]

Managed Services

Analytics

Elastic Map Reduce (EMR) – Big data processing. Log analysis. Web indexing. Hadoop.

Kinesis – Streaming/Analyzing large data. E.g. social media feeds

collating large amounts of data streamed from multiple sources

Data Pipeline – Move data. E.g. from S3 to Dynamo DB

Cloud Search – Search engine for application [out of scope]

Elastic Search – Search engine for application [out of scope]

Quick Sight – Business Analytics tool. Visualization. [out of scope]

Athena [out of scope]

Glue – Extract Transform Load (ETL) [out of scope]

Security & Identity

Identity and Access Management (IAM)

Cognito – device authentication for temporary access

Inspector – security scan/ report

Macie – scan S3 for sensitive data

Certificate Manager

CloudHSM – Hardware to store keys

Directory Service – sync AD to AWS [out of scope]

Web Application Firewall (WAF)

Shield – DDOS protection

Artifact – download Amazon’s documentations

Application Integration

Step Functions [out of scope]

MQ [out of scope]

Simple Work Flow (SWF)

SNS – Email, Text messaging, etc.

SQS – Decoupling applications. Post jobs to queues.

Desktop and App Streaming

Workspaces

AppStream 2.0

[Out of Scope]

Media Services [out of scope]

Elastic Transcoder – resize for mobile devices

Machine Learning [out of scope]

Machine Learning

SageMaker

Comprehend – Feedback on service

DeepLens – Camera

Lex – Powers Alexa service

Polly – text to speech

Rekognition

Translate

Transcribe – speech to text

Mobile Services [out of scope]

Pinpoint – notification [out of scope]

Mobile Hub – [out of scope]

AppSync – [out of scope]

Device Farm – test [out of scope]

Mobile Analytics – [out of scope]

Customer Engagement [out of scope]

Connect [out of scope]

Simple Email Service [out of scope]

Business Productivity [out of scope]

Alexa for Business

Chime

WorkDocs

WorkMail

IoT [out of scope]

Developer Tools [out of scope]

3. Identity Access Management (IAM)

Administrate AWS account

Centralized control of your AWS account

Manage users and level of access to the AWS Console

Granular permissions

Temporary access for users/devices and services

Shared access to your AWS account

Identity federation (AD, Facebook, Linkedin, etc.)

Integrates with existing active directory account allowing single sign-on

Multifactor authentication

Password rotation policy

Supports PCI DSS Compliance

**IAM is GLOBAL. No Regions specified. (Users, Groups, Roles, Policies, ALL)**

Root account: first AWS account setup. Complete Admin access granted

Power User Access allows access to all AWS services except for mgt of groups and users

Access:

Programmatically access AWS with: Access key ID/ Secret access key

Access AWS Console with: User name/ Password

Groups:

Create New Group with Policies

Group: HR, Sales, etc.

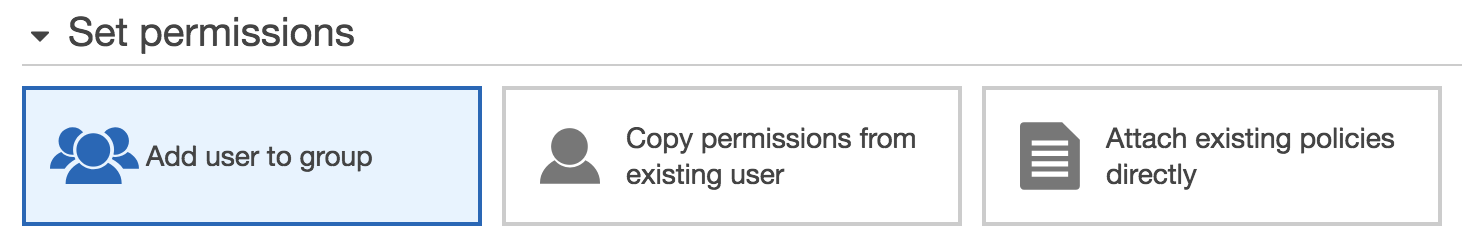
Permission->Policy: Admin Access, S3 Full Access, etc.

Users:

New Users have NO permissions/access to services when first created

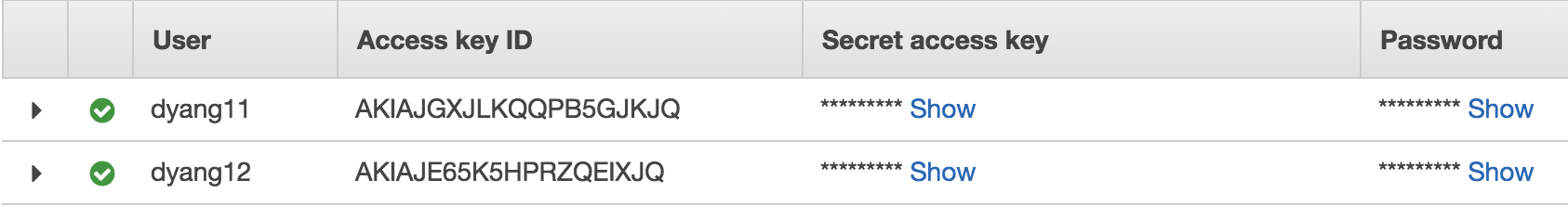
Add users to Groups

You can add permissions to users directly on top of the group permissions



New Users are assigned to Access Key ID/ Secret Access Keys/ Password

These are viewed once. Need to regenerate if lost



Security Credentials:

Activate/deactivate Access Key ID

Recreate a new access key ID/ secret access key

Setup Multifactor Authentication on your root account

Password Policy:

Minimum length

Require at least …one (Uppercase/ Lowercase/ etc.)

Allow users to change their own PW

Enable PW expiration

Prevent PW reuse

PW expiration requires admin reset

Roles:

Secure way to grant permissions (Policy) to **entities** you trust

e.g. access S3 through EC2 with access enabled roles

Entity Types:

AWS Services: EC2, Lambda, etc.

Another AWS account

Web entity

SAML 2.0 federation

Policy Documents:

Predefined permission on JSON key value pair

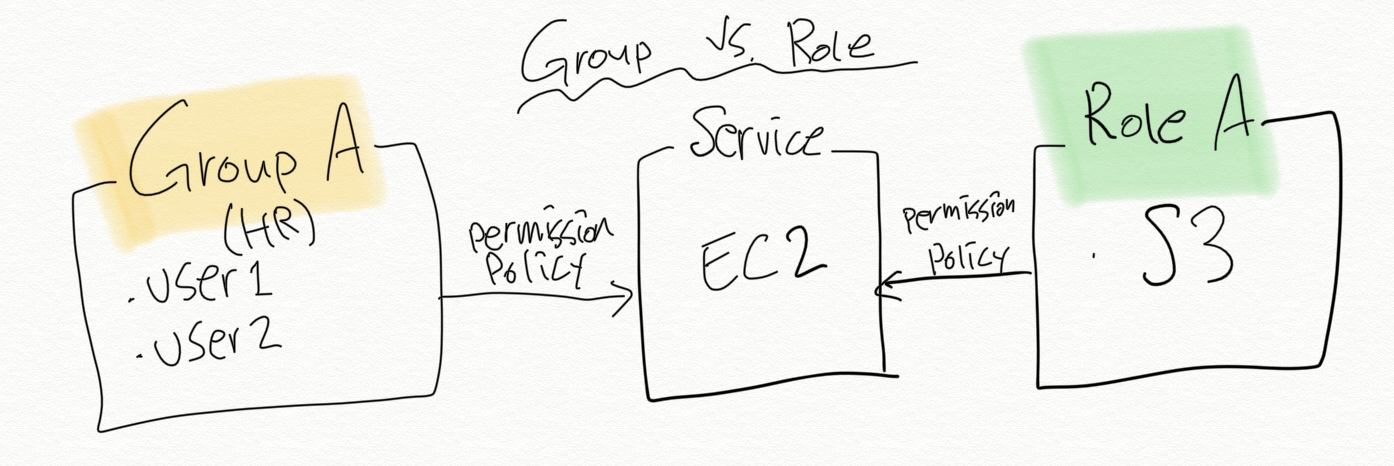
**+Lab Create a Billing Alarm**

My Billing Dashboard > Alerts & Notifications >

Enable ‘Monitor your estimated charges’ > Receive Billing Alerts

CloudWatch > Alarms > Billing >

When my total AWS charges for the month exceeds: $, send a notification to: email@



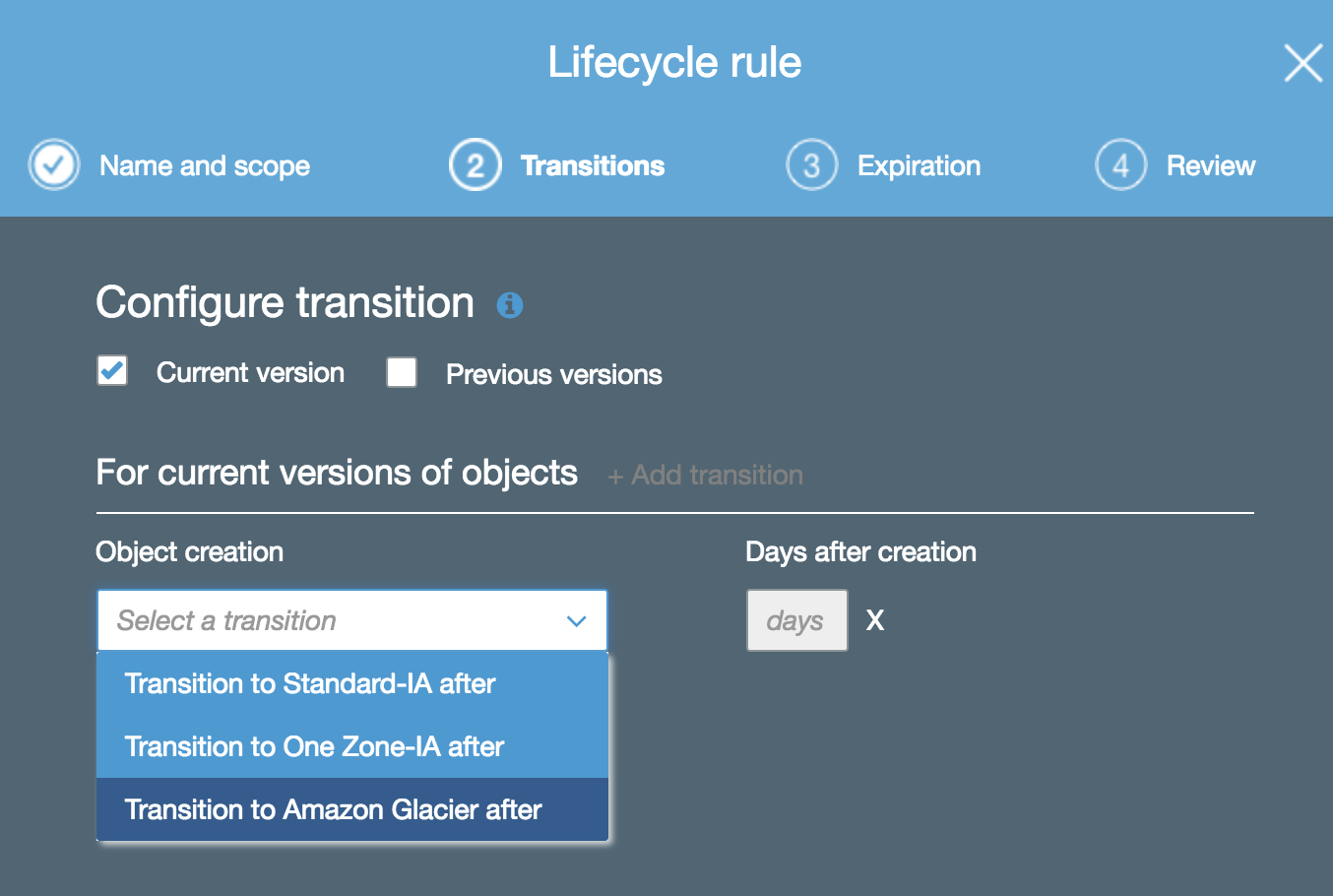
4. AWS Object Storage and CDN – S3, Glacier, and CloudFront

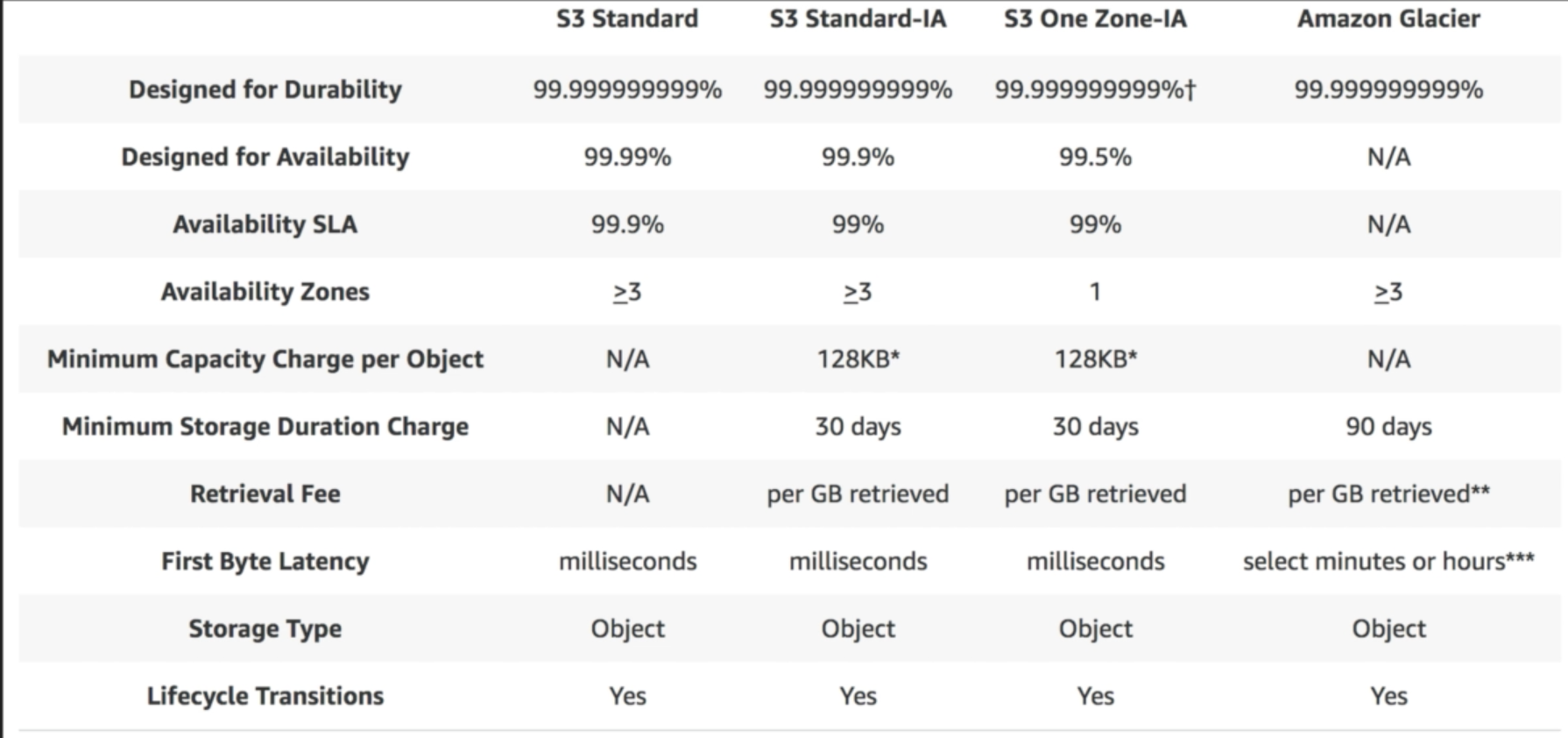
**Simple Storage Service (S3)**

* **Global. BUT Buckets are assigned regions.**
* Object based storage: flat files (videos, photos, docs, etc. NOT for OS/DB)
  + NOT a block based storage (EBS: block based can install OS/DB)
* Files are stored in Buckets (like folders)
* Multi part upload enables faster loading for large files that exceeds max size to S3
* Files can be from 1 Bytes to 5TB with **unlimited storage (for files 5TB+, use multi-part API upload)**
* 99.9% availability guaranteed
* RRS: Reduced redundancy storage 99.99% Availability. 99.99% durability.
* Universal namespace. Names must be unique globally. (like ID’s). Lowercase only
  + e.g. **https://s3**-eu-west-1**.amazonaws.com/ [know the format]**
* **Successful upload will generate HTTP 200 code \*\*\***
* Data consistency model for S3
  + Read after Write consistency for **PUTS** of **new** Objects [immediate access]
  + Eventual Consistency for **overwrite** PUTS and DELETES
    - Can take some time to propagate [until updated, either get old/new data]
* Core fundamentals of an S3 object:
  + Key: name of the object
  + Value: data
  + Version ID
  + Metadata: data about the data stored
  + Subresources: access control lists (permissions), torrent
* Charged for:
  + Storage/ Requests/ Storage Mgt. Pricing/ Data Transfer Pricing/ Transfer Acceleration
    - Transfer Accl.: long distance transfer utilizing CloudFront’s edge locations
* Lifecycle Management/ Versioning/ Encryption
* Storage Tiers/Classes
  + S3 Standard: Durable/ Immediately available/ Frequently accessed
    - 99.99% availability, 99.99..(11x9’s)% durability. Stored across 2 facilities

Bucket> Management> Lifecycle rule > Transition:

* S3 Standard - IA (Infrequently Accessed):
  + Durable/ Immediately available/ Infrequently accessed
  + Cheaper. Retrieval fee.
* S3 One Zone – IA:
  + lower cost option for infrequently accessed data. Single AZ.
* Glacier:
  + Archival only. 3-5 hours to restore from Glacier. Very Cheap. From $.01/GB/mo
  + Expedited –in minutes,
  + Standard 3-5 hrs,
  + Bulk-5-12 hrs





**+S3 [Lab]**

-Properties

-Versioning

Keep multiple versions of an object in the same bucket

-Server Access Logging

Set up access log records that provide details about access requests

**-Static Website Hosting**

Host a static website which does not require server0side technologies

-Object-level logging

Record object-level API activity using the CloudTrail data events feature

-Default encryption

Automatically encrypt objects when stored in Amazon S3

-Advanced settings

-Tags

-Transfer acceleration (using Edge location for uploading files to S3)

-Events: notification settings

-Requester Pays: The requester (not the bucket owner) pays for requests

-Permissions

-ACL

-Bucket Policy

-Management

Lyfecycle/ Replication/ Analytics/ Metrics/ Inventory

-Upload

1. Select files

2. Set permissions

3. Set properties: Storage class, Encryption, Metadata, Tags

4. Review

**+Version Control [Lab]**

-Stores all versions of an object (including all writes and even if you delete an object)

-Useful to backup files that don’t get updated frequently (takes up storage space = $$)

-Once versioning is enabled, it cannot be disabled, only suspended.

-Once a version is deleted, it will have a ‘Delete Marker’, which also can be deleted.

-Integrates with Lifecycle rules

-MFA capability is available to protect file from being deleted accidentally

**+Cross Region Replication [Lab] [High level]**

-Versioning must be enabled on both the source and destination buckets to enable cross region replication

-Can replicate entire bucket or folders in it

-Regions must be unique. Can’t replicate within the same region.

-Existing files in a bucket are NOT replicated when replication is enabled. New files only. Or manual copy can be done to replicate existing files.

-You cannot replicate TO multiple buckets (only one bucket for now)

-Delete markers are replicated

-Deleting individual versions or deleting delete markers will not be replicated

**+Lifecycle Management & Glacier [Lab]**

-Lifecycle Rules and Transition to Glacier are NOT dependent on versioning. Can be applied to current and/or previous versions

-Lifecycle rules

-Manage storage costs by controlling the lifecycle of objects:

-Transition to Standard-Infrequent Access (minimum 30 days after creation)

-Archive to Glacier Storage (minimum 60days after creation/ 30 days after IA)

-Expiration after a specified time period

**+CloudFront / [Lab]**

-Content Delivery Network (CDN): (concept)

-a system of distributed servers (network) that deliver webpages and other web content to a user based on the geographic locations of the user, the origin of the webpage and a content delivery server

>CloudFront: (service)

-used to deliver an entire website (dynamic, static, streaming, interactive content) using a global network of edge locations.

-Also works with non-AWS origin server that stores the original files

-Restrict Viewer Access (Use Signed URLs or Signed Cookies): viewer must use CloudFront signed URLs or signed cookies to access the content

-Geo-Restrictions: Whitelist/ Blacklist Countries

>Distribution: (configuration)

-a CDN that consists of a collection of Edge Locations

-Web Distribution: Used for Websites

-RTMP: Used for Media Streaming

>Edge Location:

-Location where content will be cached.

-Separate to AWS Region/AZ

-Not just READ only. An object can be put onto Edge Locations

-Objects are caches during TTL (Time To Live: in seconds)

-Cached objects can be cleared by a user with charge

>Origin:

-The origin of all files that the CDN will distribute.

-This can be either an S3, EC2, Elastic Load Balancer, or Route53

-A distribution can have multiple origins

**+S3 Security & Encryption**

-By default, all new buckets are PRIVATE

-Setup access control to buckets using:

-Bucket Policies

-Access Control Lists **-goes down to the individual files**

-S3 buckets can be configured to create access logs which log all requests made to the S3 bucket.

-Encryption

-In Transit:

-SSL/TLS

-At Rest:

-Server Side Encryption

-S3 Managed Keys – AES256 (**SSE-S3)**

**(Master Key managed by Amazon)**

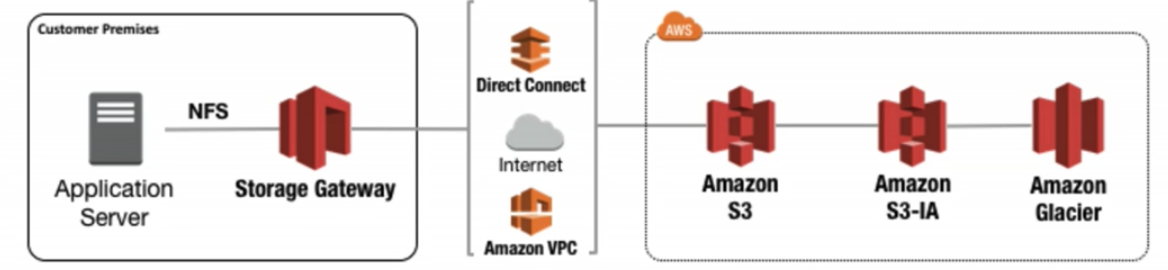
-AWS Key Management Service, Managed Keys – **SSE-KMS**

**(additional protection + audit trail)**

-Server Side Encryption with Customer Provided Keys – **SSE-C**

-Client Side Encryption (You encrypt and upload)

**+Storage Gateway**

****

-On-prem virtual appliance that can be used to cache S3 locally at a customer site.

-A service that connects an on-prem software appliance with cloud-based storage to provide integration between on-prem environment and AWS’s storage infrastructure.

-It enables the user to securely store data to the AWS cloud for scalable and cost-effective storage.

-AWS Storage Gateway’s software appliance is available for download as a VM image to be installed on a host in the datacenter. It supports VMware ESXi or Microsoft Hyper-V.

-Install the gateway and associate it with an AWS account through activation process

-Use AWS Management Console to create the storage gateway option (S3, Glacier, etc.)

-4 Types of Storage Gateways:

-File Gateway (NFS): Flat files like word docs, pics, videos, etc. Stored on S3.

-Volumes Gateway (iSCSI): Block based storage. Can run OS/DB. Virtual Hard Disk.

-Stored Volumes: Entire dataset is stored onsite and backed up to S3

-Data can be backed up as instant snapshots and stored as EBS snapshots

-Snapshots are incremental backups that capture only changed blocks

-All snapshot storage is compressed to minimize storage charges

-1GB – 16TB

-Cached Volumes: Entire dataset is stored on S3 and recent data is cached on site

-quick access. Streaming, Data analysis, etc.

-Tape Gateway (VTL): Archive. Used for backup utilizing backup apps like NetBackup.

**+Snowball** (Previously Import/Export Disk. Like an external hard drive) – To S3

-Import/Export Disk (Legacy)

-accelerates moving large amounts of data in/out of the AWS cloud using portable storage devices for transport

-Transfers data directly onto and off of storage devices using Amazon’s high-speed internal network and bypassing the internet.

-Snowball:

Storage Only.

A petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data in/out of AWS(S3).

Store data in snowball device and send it to Amazon for them to upload to AWS Cloud, bypassing the internet.

-Snowball Edge:

Snowball + **Compute** capabilities. Like a small AWS datacenter

E.g. Aircraft engine performance

-Snowmobile:

A shipping container on a truck. Used for extremely large data (100PB).

E.g. datacenter migration

**+S3 Transfer Acceleration**

-Utilizes CloudFront Edge Network to accelerate uploads to S3

-Instead of uploading directly to S3 bucket, use a distinct URL to upload a file directly to an edge location which will then transfer the file to S3

-The farther away the S3 region is, the more speed improvement Transfer Acceleration will provide

-Properties > Advanced Settings > Transfer Acceleration > Enabled

**+S3 Static Website Using S3 [Lab]**

-Serverless

-Very cheap. Scales automatically

-Static content only. No dynamic sites.

-Properties > Static website hosting > Use this bucket to host a website > Assign index document

In S3. upload: Index.html (**grant public access**)

-Static website hosting Endpoint: http://dyang322website.s3-website-us-east-1.amazonaws.com

\*\*(bucketname.s3-website-region.amazonaws.com)

**5. EC2 – The Backbone of AWS**

- A web service that provides resizable compute capacity in the cloud. It reduces the time required to obtain and boot new server instances to minutes, allowing user to quickly scale capacity as required

- Allows user to pay only for capacity used

- Provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios

-Termination protection is turned off by default

-EC2 Options

- On Demand: pay a fixed rate by the hour with no commitment

-Linux can be paid by second. Windows by hour

-Low cost/ Flexibility/ No upfront payment or commitment

-Applications with short term/ unpredictable workloads that cannot be interrupted

-Applications being developed/tested on EC2 for the first time

- Reserved: provide with a capacity reservation, and offer a discount on the hourly charge for an instance. 1yr or 3yr terms. Upfront charge.

-Applications with steady state or predictable usage

-Applications that require reserved capacity

-User able to make upfront payment to reduce total computing costs

- Spot: bid whatever price you want for instance capacity, providing for greater savings if your applications have flexible start and end times.

-Applications with flexible start and end times

-Applications that requires at very low compute prices

-Urgent computing needs for large amounts of additional capacity

-If a Spot instance is terminated by Amazon EC2, you won’t be charged for a partial hour of usage

- Dedicated Hosts – Physical EC2 server dedicated for your use. Dedicated Hosts can help you reduce costs by allowing you to use your existing server-bound software license.

-Use for regulatory requirements that may not support multi-tenant virtualization

-Great for licensing which does not support multi-tenancy or cloud deployments.

-Can be purchased on-demand (hourly)

-Can be purchased as a reservation for up to 70% off the On-demand price

- Instance Types [no need to memorize this]

C5 Compute Optimized CPU Intensive Apps/DBs

F1 Field Programmable Gate Array Genomics research, big data

I3 Hi Speed Storage NoSQL DBs, Data Warehousing

D2 Dense Storage Fileservers/Data Warehousing/Hadoop

H1 High Disk Throughput MapReduce-based workload

T2 General Purpose, Low Cost Web servers/ Small DBs

M5 General Purpose App servers

R4 Memory Optimized Memory Intensive Apps/DBs

X1 Memory Optimized SAP HANA/ Apache Spark

G3 Graphics Intensive Video Encoding, 3D app streaming

P3 Graphics/General Purpose GPU Machine Learning, Bit coin mining

* EC2 Storage:
* Root Volume:
  + By default, delete on termination of EC2 is checked (EBS volumes are not deleted on EC2 termination)
  + EBS Root Volumes of your default AMI’s cannot be encrypted. You can also use a third-party tool to encrypt the root volume, or this can be done when creating AMI’s in the AWS console or using the API
* EBS (Elastic Block Storage) Volume:
* Optional storage (virtual hard disk) attached to EC2 along with Root Volume. Like D drive.
* Runs DB, OS, etc.
* **EBS volumes are always in the same AZ as the EC2 instance they are attached to**
  + Snapshot can be taken on the EBS volume and an image (AMI) can be created in a different region/AZ (attached to a new EC2 instance)
* EBS volume size and type can be changed without downtime (not Magnetic Standard)
* Snapshots (point in time copies of Volumes) exist on S3. (incremental)
* Snapshots of encrypted volumes are encrypted automatically
* Volumes restored from encrypted snapshots are encrypted automatically
* You can share snapshots, but only if they are unencrypted
  + These snapshots can be shared with other AWS accounts or made public
  + You cannot mount 1 EBS volume to multiple EC2 instances; instead use EFS
  + Root device of a registered AMI cannot be deleted without AMI being deregistered
* EBS Volume Types:
  + SSD
    - SSD, General Purpose (GP2):
      * Up to 10k IOPS
    - SSD, Provisioned IOPS (IO1):
      * I/O intensive apps
      * Large relational or NoSQL DB
      * 10k+ IOPS. Up to 20k IOPS
  + Magnetic
    - HDD, Throughput Optimized (ST1):
      * Low cost HDD for frequently accessed, throughput-intensive
      * Big data. Data warehouses.
      * **Cannot be a boot volume**
    - HDD, Cold (SC1):
      * Lowest cost HDD for infrequently accessed workloads.
      * File servers.
      * **Cannot be a boot volume**
    - HDD, Magnetic (Standard): [Previous generation]
      * Lowest cost per GB for **bootable EBS volume**
      * Infrequently access data
  + RAID: Redundant Array of Independent Disks (group of EBS disks acting as one)
    - For high disk IO’s
    - To take a snapshot of a RAID array:
      * Stop application from writing to disk and flush all caches by:
        + Freeze the file system, or
        + Unmount the RAID Array, or
        + Shut down the associated EC2 instance
    - **RAID 0 – Striped, No redundancy, Good performance. Most common**
    - RAID 1 – Mirrored, Redundancy
    - RAID 5 – Good for reads, bad for writes, AWS does not recommend ever putting RAID 5’s on EBS. 3 or more disks.
    - RAID 10 – Striped & Mirrored, Good redundancy, good performance
  + All AMI’s are categorized as either backed by Amazon EBS or by instance store
    - EBS Volumes:
      * Persistent = continue independently from EC2 instance (can be detached)
      * Can be stopped without losing data
      * the root device for an instance launched from the AMI is an Amazon EBS volume created from an Amazon EBS snapshot
      * Root volume can be encrypted using AWS API or console, or using 3rd party tool (like bit locker)
      * Store data long term
    - Instance Store Volumes:
      * Not persistent (Ephemeral)
      * Cannot be stopped (data will be wiped out)
      * The root device for an instance launched from the AMI is an instance store volume created from a template stored in S3
      * Shouldn’t be used for long-term data storage
    - Both can be rebooted and not lose the data
    - By default, both Root volumes will be deleted on termination, however with EBS volumes, you can tell AWS to keep the root device volume.
  + Snapshots
    - Snapshots exist on S3 (Volumes exist on EBS)
    - You can take a snapshot of a volume, this will store that volume on S3
    - Snapshots are point in time copies of Volumes
    - Snapshots are incremental
    - Encrypted automatically
    - Volumes restored from encrypted snapshots are encrypted automatically
* Security Group
  + **All inbound traffic is blocked by default (allow manually)**
  + All outbound traffic is allowed
  + Security groups are Stateful
    - If you create an inbound rule allowing traffic in, that traffic is automatically allowed back out again
  + You cannot block specific IP addresses using security groups, instead use Network Access Control Lists (Stateless)
  + Security Group - **EC2** **instance level**  ACL – **subnet level**
  + You can specify allow rules, but not deny rules
  + Changes to security groups take effect immediately
  + You can have any number of EC2 instances within a security group
  + You can have multiple security groups attached to EC2 Instances
* Elastic Load Balancers
  + Types:
    - Application LB: layer 7. Best for HTTP(S). sends specific requests to specific web servers
    - Network LB: layer 4. Best for TCP traffic with extreme performance. Millions of requests per second. Low latencies.
    - Classic LB: Legacy Elastic Load Balancers (ELB). HTTP(S). Use layer 7 specific features like X-Forwarded and sticky sessions. Layer 4 for TCP. Not as intelligent as the new types.
  + 504 Error means the gateway has timed out. The application is not responding within the idle timeout period
  + If IPv4 address of the end user is needed, look for the X-Forwarded-For header (internal EC2 instances receive LB internal IP)
  + Instances monitored by ELB are reported as; InService, or OutofService
  + Health Checks check the instance health by talking to it
  + ELB’s have their own DNS names. IP addresses are dynamic and not given.
* CloudWatch
  + CloudWatch is for performance monitoring (CloudTrail is for auditing. Logs on EC2 creation, etc. done in your AWS account)
  + Standard Monitoring: 5 minutes Detailed Monitoring: 1 minute
  + Alarms: Notify when thresholds are hit
  + Events: respond to state changes on AWS resources
  + Metric Name for EC2:
    - CPU
      * Credit Balance Credit Usage Utilization
    - Disk
      * ReadBytes ReadOps WriteBytes WriteOps
    - Network
      * In Out PacketsIn PacketsOut
    - Status Check
      * Failed Instance System
* Autoscaling
  + Can be enabled while creating EC2 instances
  + How it works: e.g.
    - 3 EC2 instances in different availability zones on a Load Balancer
    - 2 of the 3 EC2 instances go down
    - AWS automatically bring up a new instance
* Placement Groups:
  + Determines how instances are placed on underlying hardware (single/multi AZ)
  + Use Case: Application running on multiple servers that require low latency
  + Clustered Placement Group (typically this)
    - Grouping of instances within **a single AZ**.
    - Recommended for apps that need low network latency and high network throughput
  + Spread Placement Group (new)
    - Grouping of instances that are each placed on distinct underlying hardware for redundancy(multiple AZ)
    - Recommended for apps that have a small number of critical instances that should be kept separate from each other
  + Placement Group name must be unique within your AWS account
  + Only certain instances can be launched into a Placement Group
    - (Compute/ GPU/ Memory/ Storage Optimized)
  + AWS recommends homogeneous instances within placement groups
  + You can’t merge placement groups
  + You can’t move an existing instance into a placement group
    - Create a new AMI of the existing instance and launch a new instance from the AMI into a placement group
* Elastic File System (EFS): [New]
  + File storage service for EC2 instances
  + Easy to use with simple interface
  + Storage capacity is elastic (grow/shrink automatically) as files are added/removed
  + Can mount one EFS to more than one EC2 instances
  + Supports thousands of concurrent NFS connections
  + Data is stored across multiple AZs within a region
  + Read after write consistency
  + Can scale up to Petabytes
  + Pay for the storage you use
  + Supports Network File System version 4 (NFSv4) protocol
  + Block based (not object based)
  + EC2 instance needs to be in the same security group as EFS
* Lambda
  + A SERVERLESS compute service where you upload your code and create a Lambda function
  + AWS manages OS, patching, scaling, etc.
  + Event driven
    - Triggers: API Gateway, Alexa Skills Kit, CloudFront, CloudWatch, Dynamo DB, Kinesis, S3, SNS, etc.
  + Languages (Runtime): C#, Go, Java, Node.js, Python
  + Scales out (not up): 2 users sending 2 requests triggers 2 lambda functions
  + Lambda functions are independent, 1 event = 1 function
  + Priced based on
    - # of requests
    - Duration: from the time your code begins executing until it returns or terminates. Based on the amount of memory you allocate to your function
  + Continuous scaling
  + AWS X-ray allows you to debug
  + Works globally across regions

5. Route 53

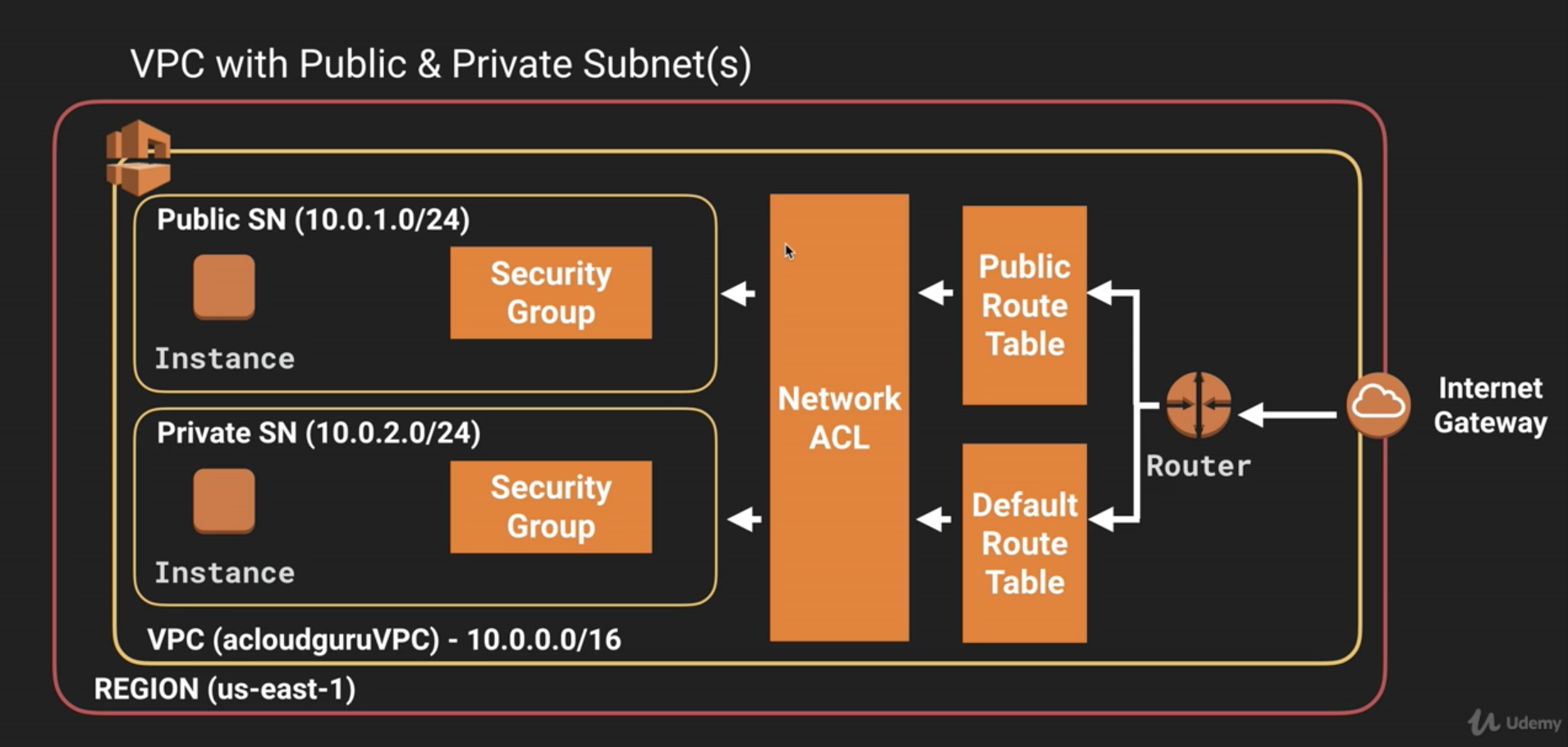
* AWS DNS service
* Global Service (not region specific)
* ELB’s do not have predefined IPv4 addresses, you resolve to them using a DNS name
* DNS port is on port 53 -> Route 53 name
* Up to 50 domain names can be managed using Route 53 but this can be raised by contacting AWS support
* DNS
  + SOA Record
    - Admin info, version of data file, name of the server, etc.
  + NS Record
  + MX Record
  + **A Record (Address):** 
    - map domain name to IP address
  + **CName Record (Canonical):** 
    - map one domain name to another.
    - Cannot be used for naked domain names (zone apex record)
  + **Alias Record:** 
    - map one domain name to another. like CName record.
    - Used on Load Balancers, CloudFront, or S3 buckets.
    - AWS doesn’t charge Alias Name but charge CNAME
    - Given the choice, always choose an Alias Record over a CNAME
  + TTL Record (Time to Live):
    - the lower the TTL is, the faster changes to DNS records take to propagate throughout the internet
* Routing Policies
  + Simple
    - single resource
  + Weighted
    - Split traffic based on different weights assigned
    - e.g. 20% to East 1 region, 80% to West 1 region
    - Used for AB testing
  + Latency
    - Route traffic based on the lowest network latency for your end user (region with shortest response time)
  + Failover
    - Traffic is routed to the secondary region if the primary region is not active
  + Geolocation
    - Traffic is sent based on the geographic location of the users
    - E.g. U.S customers are routed to the U.S site with English language and $ pricing while Korean customers are routed to the Korean site with Korean language and Won pricing

6. Database

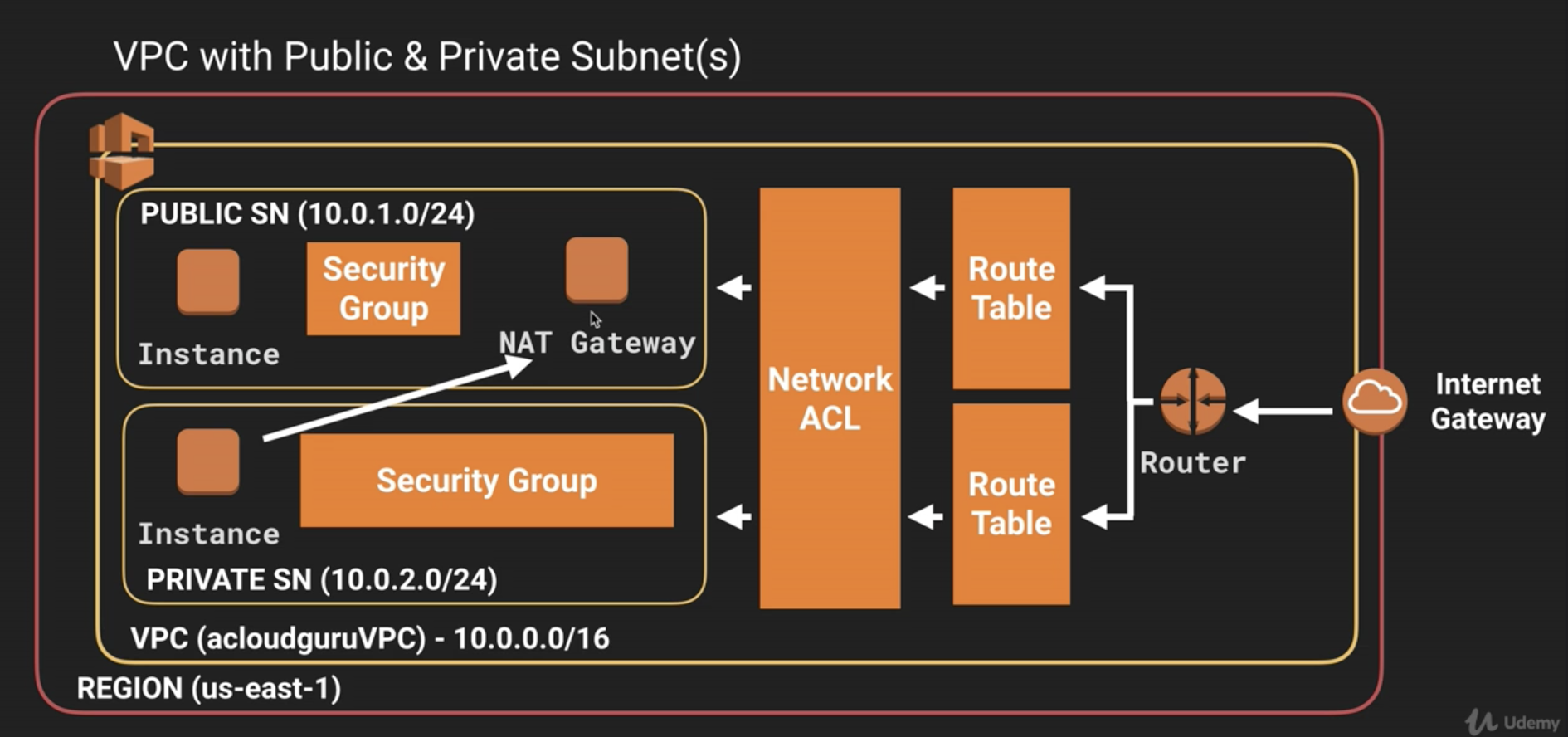
* Relational DB (RDS) - OLTP
  + SQL, MySQL, PostgreSQL, Oracle, MariaDB, **Aurora** (AWS RDS)
  + OLTP: Online Transaction Processing [order number query -> details]
  + IPv4 address is not available. Only DNS name.
  + Maximum 16 TB RDS volume
  + Aurora [has not been on exams but may appear]
    - MySQL compatible
    - 5x better performance than MySQL with 1/10 price
    - 2 copies of data in each AZ with minimum of 3 AZ. 6 copies of data
    - Replicas:
      * Aurora Replicas (15 replicas)
      * MySQL Read Replicas (5 replicas)
  + Backups
    - Automated backups are enabled by default.
    - Backup data is stored in S3 with **free** storage space equal to the size of DB
    - During the backup window, storage I/O may be suspended with latency
    - Snapshots are done manually and stored even after you delete the original RDS instance
    - Multi-AZ:
      * replica of DB in another AZ for disaster recovery.
      * Not for performance improvement
    - Read Replicas: (MySQL, MariaDB, PostgreSQL, Aurora DB; Not MSSQL)
      * read only replica of DB in same or another AZ / region
      * Scale out. Not for DR. for performance improvement
      * Up to 5 read replica copies of any DB
      * Must have automatic backups turned on
      * Read replicas of read replicas are available
  + Encryption
    - Encrypting an EXISTING DB Instance is not supported
    - To use RDS encryption for an existing DB, create a snapshot, make a copy of the snapshot and encrypt the copy
* Elasticache
  + In-memory cache
  + Improves performance by taking workload off of the DB
  + Types:
  + Redis
  + In-memory data structure store used as database, cache and message broker. ElastiCache for Redis offers Multi-AZ with Auto-Failover and enhanced robustness.
  + Open-source in-memory key-value store
  + Supports Master/Slave replication and Multi-AZ
  + Memcached
  + High-performance, distributed memory object caching system, intended for use in speeding up dynamic web applications.
  + Useful when the DB is read heavy and not prone to frequent changing
* Non-Relational DB (NoSQL)
  + Collection = Table
  + Document = Row
  + Key Value Pairs = Fields
  + **DynamoDB (AWS NoSQL DB)**
    - NoSQL. Very scalable.
    - Stored on SSD storage
    - Spread across 3 geographically distinct data centers
    - Consistency Model:
      * Eventual Consistent Reads (Default)
        + Consistency across all copies of data is usually reached within a second. (best read performance)
      * Strongly Consistent Reads
        + returns a result that reflects all writes that received a successful response prior to the read (after you write, you need data updated in 1 second)
    - Pricing
      * Provisioned Throughput Capacity
      * Storage cost
* DataWarehousing [RedShift]
  + Cognos, Jaspersoft, SQL Server Reporting Services, Oracle Hyperion, SAP NetWeaver
  + OLAP: Online Analytics Processing [for DataWarehousing]
  + **RedShift**
    - Fast and powerful, fully managed, petabyte-scale data warehouse service
    - Column based (higher performance)
    - Encrypted in transit using SSL
    - Encrypted at rest using AES-256 encryption
    - Only available in 1 AZ
    - Snapshots can be restored to new AZ’s
    - Useful when the DB is overloaded because of OLAP transactions
    - Single-Node (160GB Storage)
    - Multi-Node
      * Leader Node
      * Manage client connections and receives queries
    - Compute Nodes
      * Store data and perform queries and computations
      * Up to 128 compute nodes per leader node
      * Massively Parallel Processing (MPP): auto distribution of data + easy scaling

7. VPC

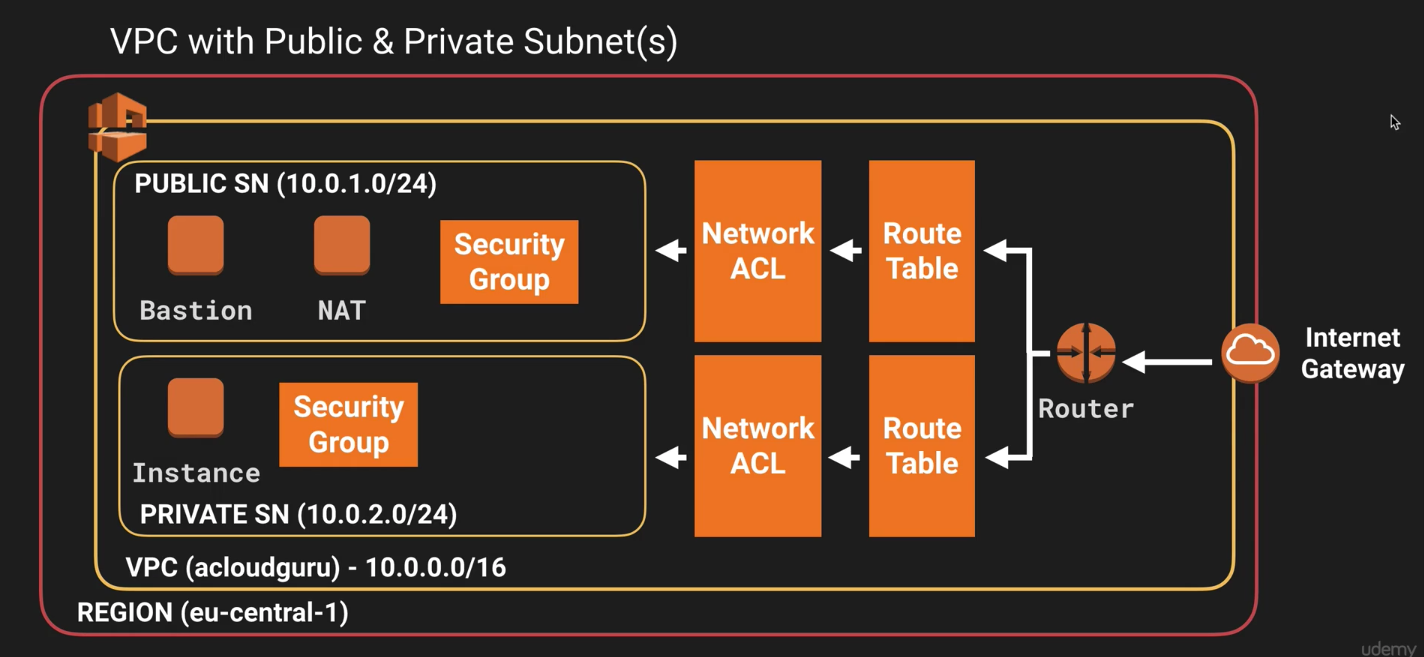
* Virtual Private Cloud
* Logically isolated section of the AWS Cloud (Logical datacenter)
* Up to 5 VPC’s per region by default
* 1 subnet = 1 Availability Zone
* Consists of Internet Gateway (Virtual Private Gateways), Route Tables, Network Access Control Lists, Subnets, and Security Groups
* Security Groups are Stateful; Network Access Control Lists are Stateless
* No Transitive Peering (VPC A – VPC B – VPC C; VPC A cannot talk to VPC C)
* IP
  + First 4 IP address and the last IP address are reserved and cannot be used
    - 0. Network address
    - 1. VPC router
    - 2. DNS server
    - 3. For future use
    - 4. Network broadcast address
* Only one Internet Gateway can be attached to a VPC
* Route Table
  + Setup routes and associate subnets to the route (e.g. Traffic out to public)



* NAT (Network Address Translation)
  + Used to provide **internet traffic** to EC2 instances in private subnets
  + NAT Instance
    - EC2 instance AMI
    - When creating NAT instance, Disable Source/Destination Check
    - NAT instances must be in a public subnet
    - There must be a route out of the private subnet to the NAT instance
    - Always behind Security Group
  + NAT Gateway
    - AWS provided service (improved version of NAT Instance which will be decom’d)
    - AWS manages it (Patching, Anti Virus, etc.)
    - Not associated with Security groups
    - Automatically assigned a public IP address
    - Scale automatically up to 10Gbps
    - No need to disable Source/Destination Checks
    - More secure than a NAT instance



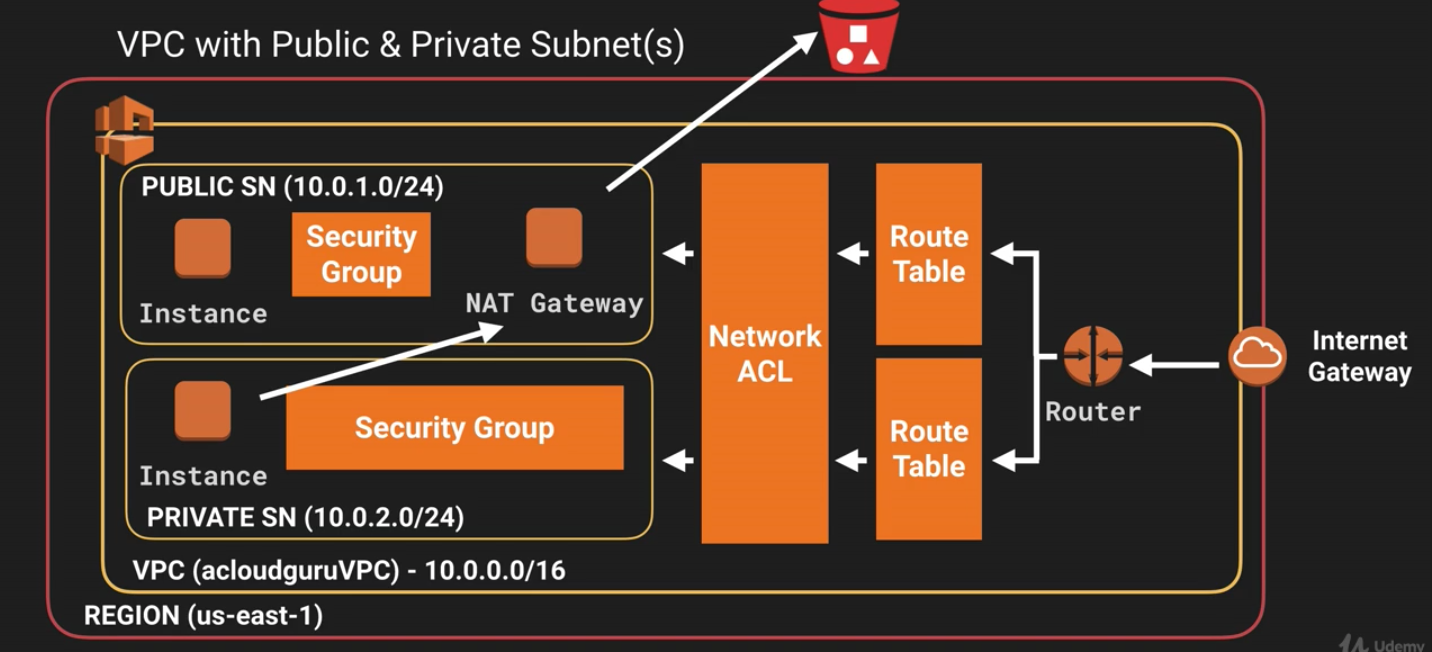
* + Bastions
    - Jump box
    - **SSH or RDP** to Bastion and connect to the instance in private subnet
    - Used to securely administer EC2 instances in private subnets



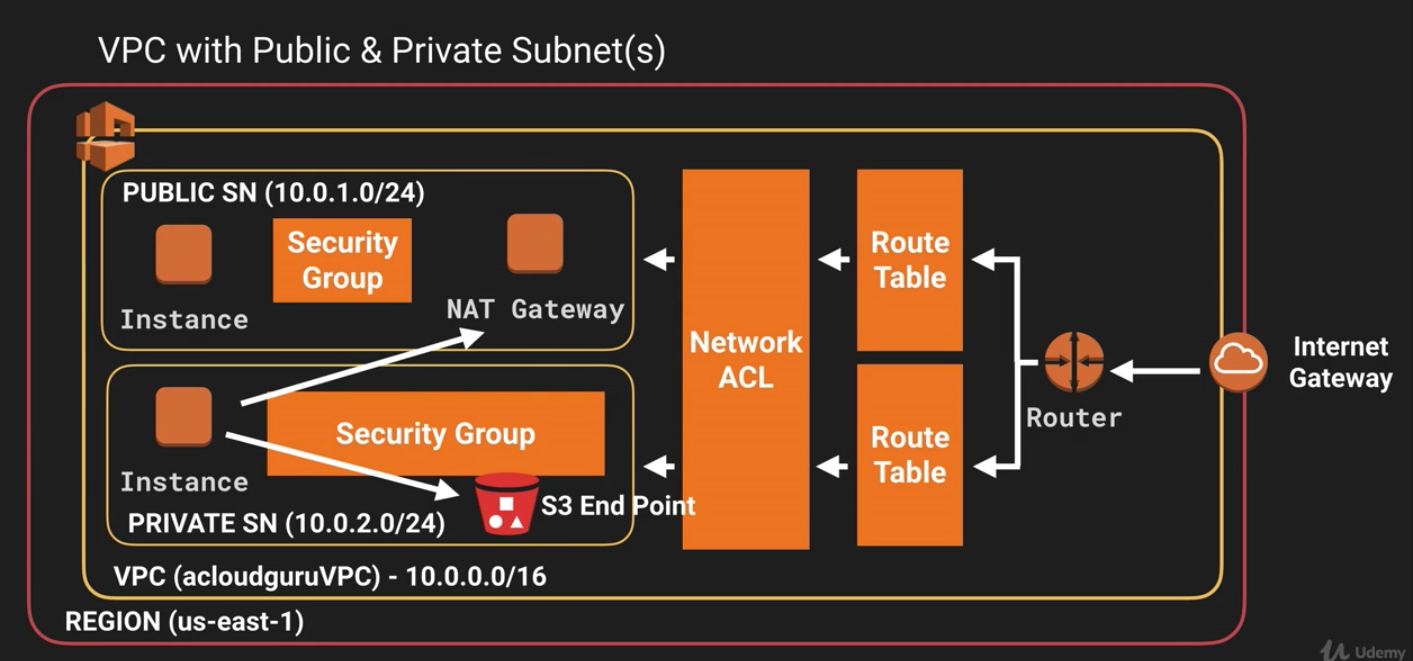
* Network Access Control List
  + Each subnet must be associated with only ONE network ACL. If you don’t explicitly associate a subnet with a network ACL, the subnet is automatically associated with the default network ACL
  + You can associate a network ACL with multiple subnets; however a subnet can be associated with only one network ACL at a time.
  + Security groups act like a firewall at the INSTANCE level whereas NETWORK ACL works as an additional layer of security at the subnet level
  + Network ACL can only be deployed in one VPC
  + VPC automatically comes with a default network ACL and by default it allows all Inbound/Outbound traffic
  + You can create custom network ACLs and by default it DENIES all inbound/outbound traffic until you add rules
  + Network ACLs contain a numbered list of rules that is evaluated in order, starting with the lowest numbered rule.
  + Network ACLs are stateless; responses to allowed inbound traffic are subject to the rules for outbound traffic (and vice versa)
  + Network ACLs have separate inbound and outbound rules
  + Block IP addresses using network ACLs, but not using Security Groups

|  |  |
| --- | --- |
| Security Group | Network ACL |
| Operates at the instance level | Operates at the subnet level |
| Supports allow rules only | Supports allow rules and deny rules |
| Is stateful: Return traffic is automatically allowed, regardless of any rules | Is stateless: Return traffic must be explicitly allowed by rules |
| We evaluate all rules before deciding whether to allow traffic | We process rules in number order when deciding whether to allow traffic |
| Applies to an instance only if someone specifies the security group when launching the instance, or associates the security group with the instance later on | Automatically applies to all instances in the subnets it's associated with (therefore, you don't have to rely on users to specify the security group) |

* VPC Flow Logs
  + Captures information about the IP traffic going to and from network interfaces in your VPC.
  + Flow log data is stored using CloudWatch Logs
  + Flow logs for VPC must be in the same account
  + You cannot tag a flow log
  + Once a flow log is created you cannot change its configuration (like IAM role)
  + Levels:
    - VPC
    - Subnet
    - Network Interface
  + Not all IP traffic is monitored
    - Traffic generated by instances when they contact the Amazon DNS server
    - Traffic generated by a Windows instance for Amazon Windows license activation
    - Traffic to and from 169.254.169.254 for instance metadata
    - DHCP traffic
    - Traffic to the reserved IP address for the default VPC router
* VPC End Points
  + Accessing S3 using NAT Gateway sending info out to the internet



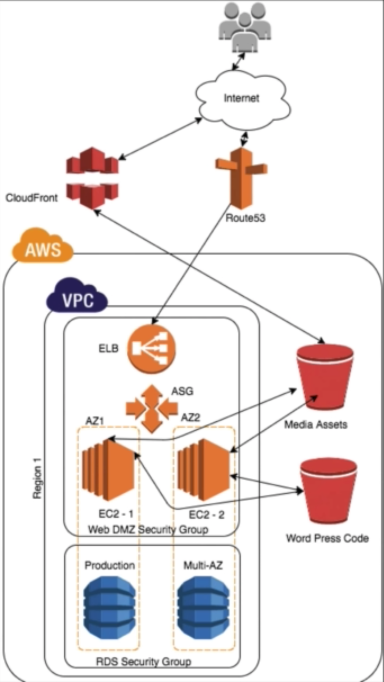
* + Accessing S3 using End Point Gateway



* VPC Peering
  + Connection between 2 VPCs that enables you to route traffic between them using private IP addresses
  + Connections between your own VPCs, or with a VPC in another AWS account within a SINGLE REGION
  + You cannot create a VPC peering connection between VPCs in different regions
  + No single point of failure or bandwidth bottleneck
  + Not relying on separate piece of physical hardware
  + Transitive peering is not supported (A – B – C; A cannot talk to C)
  + CIDR block range has to be unique and not overlapping
* Direct Connect
  + Dedicated private network connection from on prem to AWS
  + Reduce network costs when using large volumes of traffic
  + Increase bandwidth and reliability
  + Direct Connect vs VPN
    - VPN
      * **Configured in minutes, good for immediate need**
      * Low to modest bandwidth requirements
      * Inherent variability in Internet-based connectivity
    - Direct Connect
      * **Takes months to set up**
      * No Internet connection
      * Dedicated, private network connections

8. Application Services

* SQS (Simple Queue Service)
  + Distributed message queuing system. Buffer
  + Decouple the components of an application so that they are independent
  + Store messages while waiting for a computer to process them
  + **Pull based system (not pushed based)**
    - **Other services pull messages from SQS when they need it (SNS is push)**
  + Up to 256kb of text in any format
  + Messages can be kept in the queue from 1 minute to 14 days
  + Default retention is 4 days
  + Queue types
    - Standard Queues (Default)
      * Best-effort ordering
      * **Delivered with no order**
      * Message delivered **AT LEAST** once
      * Unlimited number of transactions per second
    - FIFO (first in first out)
      * **Delivered in order**
      * Limited to 300 transactions per second
      * **Message delivered once (NO DUPLICATES)**
      * Good for banking transactions which happens in strict order
  + SQS visibility timeout
    - Amount of time that the message is invisible in the SQS queue after a reader picks up that message
    - Job needs to be processed before visibility time out expires, or you will get 2 messages
    - Default visibility time out: 30seconds
    - Increase it if your task takes > 30 seconds
    - Maximum is 12 hours
  + Short Polling [default] – returned immediately even if no messages are in the queue
  + Long Polling – polls the queue periodically and only returns a response when a message is in the queue or the timeout is reached
* SWF (Amazon Simple Workflow Service)
  + Runs on EC2
  + SWF brokers the interactions between workers and the decider
  + Allows the decider to get consistent views into the progress of tasks and to initiate new tasks in an on-going manner
  + SWF stores tasks, assigns them to workers when they are ready and monitors progress
  + **Ensures that a task is assigned only once and never duplicated (not like SQS)**
  + SWF Starter:
    - Application that can initiate a workflow (e.g. amazon.com)
  + SWF Workers:
    - programs that interact with Amazon SWF to get tasks, process received tasks, and return the results
  + SWF Decider:
    - program that controls the coordination of tasks. i.e. ordering, and scheduling
  + SWF vs SQS
    - SWF retention period up to 1 year; SQS 14 days
    - SWF presents a task-oriented API; SQS offers a message-oriented API
    - SWF assigns a task only once and never duplicated/ SQS at least once
    - SWF keeps track of all the tasks and events in an app/ SQS you need implement your own application-level tracking
* SNS (Simple Notification Service)
  + Set up, operate, and send notifications from the cloud
  + Subscribers: HTTP(S), Email, Email-JSON, SQS, Application, Lambda
  + Topic: group of recipients on various endpoint types receiving identical copies of the same notification
  + Messages are stored redundantly
  + Instantaneous, push-based delivery (no polling)
  + **SNS vs SQS**
    - **Both messaging services in AWS**
    - **SNS – Push SQS – Polls (Pulls)**
* Elastic Transcoder
  + Media Transcoder that converts media files and their formats
  + Pay based on the minutes and the resolution
* API Gateway
  + API caching: cache your endpoint’s response to increase performance
  + Low cost and auto scale
  + Throttle API gateway to prevent attacks
  + Log results to CloudWatch
  + If you have multiple domains with API Gateway, ensure to enable CORS (Cross Origin Resource Sharing) on API Gateway
* Kinesis
  + Streaming Data: data generated continuously by thousands of data sources which typically send in the data records simultaneously, and in small sizes (Kilobytes)
    - E.g. Purchases from online stores, Stock Prices, Game Data, SNS, IoT
  + Load and analyze streaming data
  + Process large amounts of data
    - Redshift for **business intelligence**
    - Elastic Map Reduce for **Big Data Processing**
  + Kinesis Services:
    - Kinesis Streams
      * Stores data in **‘shards’** for 24 hours (up to 7 day) retention
      * Sends data to Consumers (like EC2)
    - Kinesis Firehose
      * No shards, streaming, retention, etc.
      * All automated (optional analysis using Lambda)
      * Sends data to other services
    - Kinesis Analytics
      * Runs SQL queries on data stored in Streams or Firehose

9. Fault Tolerant World Press Site

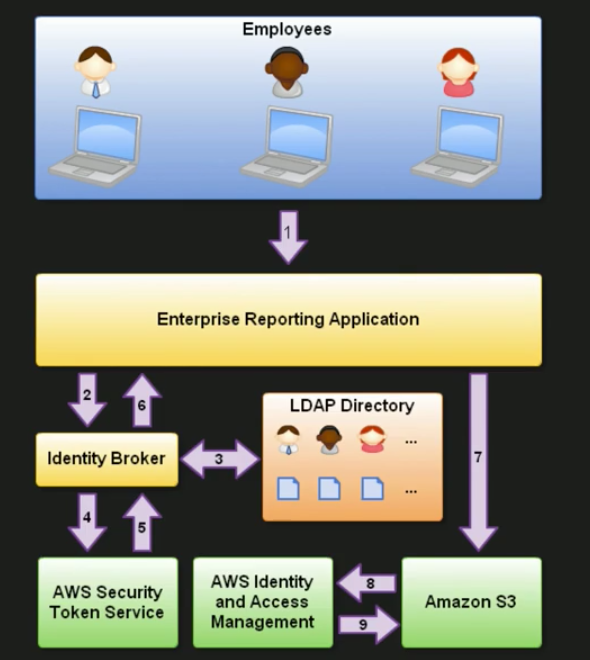
* Security Group
  + Web DMZ
    - Source: Public 0.0.0.0
    - Type: HTTP, SSH
  + RDS
    - Source: EC2 security group
    - Type: MySQL
* S3 Buckets
  + Media Assets
  + Word Press Code
    - Connect to CloudFront URL
    - URL Rewrite
* CloudFront
  + Domain name: URL
* RDS MySQL
  + Security Group: RDS
* IAM – Role (S3 Full Access)
  + Service: EC2
  + Policy: S3 Full Access
  + EC2 can talk to S3
* EC2
  + Role: S3 Full Access
  + Security Group: Web DMZ
* LoadBalancer
  + Application Load Balancer
  + Target: WebDMZ
* Route53
  + Alias Target: Load Balancer
* Cloud Formation
  + Build templates and push it out to different region/env in minutes

10. Well Architected Framework

* Security
  + Data protection
    - Encrypt data in transit (SSL) and at rest using; ELB, EBS, S3 & RDS
  + Privilege management
    - AWS root account credentials protection; IAM, MFA
    - Access Control Lists (ACLs)
    - Role Based Access Controls
    - Password Management
  + Infra protection
    - VPC access protection
    - OS integrity (anti-virus, etc.)
  + Detective controls
    - CloudTrail for analyzing AWS logs (Regional service)
* Reliability
  + Foundation
    - Network Topology
  + Change Management
    - Monitoring AWS resources
  + Failure Management
    - Backup/Restore
* Performance Efficiency
  + Compute
  + Storage
  + Database
  + Space-Time Trade-off
    - Proximity and Caching solution
* Cost Optimization
  + Matched supply and demand
  + Cost-effective resources
  + Expenditure awareness
    - Monitoring usage and spending
    - Decommission/Stop resources
  + Optimizing over time
    - Adoption of new services
* Operational Excellence
  + Preparation
    - Config management
  + Operations
    - Monitor workload
  + Responses
    - Respond to unplanned operational events

11 Additional Exam Tips

* OpsWorks
  + Orchestration Service that uses **Chef**
  + Chef consists of recipes to maintain a consistent state
  + Look for the term chef, recipes, or cookbooks and think OpsWorks
* AWS Organizations
  + Account management service that enables you to consolidate multiple AWS accounts into an organization that you create and centrally manage
  + Root Account > Organization Units – Policy > AWS Account
  + Consolidated Billing
    - Root (Paying) Account > Linked Accounts (Test/Prod/Dev Office) > Monthly Bill
    - One bill per AWS account
    - Saving money by volume discounts
    - Unused reserved instances for EC2 are applied across the group
    - Linked accounts: 20 accounts only
    - CloudTrail is on a per account and per region basis but can be aggregated in to a single bucket in the paying account
* Cross Account Access
  + Switch account to access other account entities
  + E.g. use dev account user to access S3 bucket in prod account with policy and role defined
* Resource Groups & Tagging
  + Tags
    - Key value pairs attached to AWS resources
    - Metadata
    - Tags can sometimes be inherited
  + Resource Groups
    - Group resources using the tags that are assigned to them
    - Group resources that share one or more tags
    - Resource groups contain information such as Region, Name, Health Checks
    - Specific info
      * EC2: Public & Private IP addresses
      * ELB: Port Config
      * RDS: DB Engine
* Security Token Service (STS)
  + Grant users limited and temporary access to AWS resources
    - Federation (typically AD)
      * Uses security assertion markup language (SAML)
      * Grants temporary access based off the users Active Directory credentials. Do not need to be a user in IAM
      * Single sign on allows users to log in to AWS console without assigning IAM credentials
      * Federation with Mobile Apps
        + Use Facebook/Amazon/Google or other OpenID providers
      * Cross Account Access
        + Lets users from one AWS account access resources in another
  + STS provides: Access Key, Secret Access Key, Token, Duration
    - Token Lifetime: 1 – 36 hrs
  + Terms:
    - Federation
      * Combining or joining a list of users in one domain (such as IAM) with a list of users in another domain (such as Active Director, Facebook, etc.)
    - Identity Broker
      * A service that allows you to take an identity from point A and join it (federate it) to point B
    - Identity Store
      * Services like Active Directory, Facebook, Google, etc.
    - Identities
      * A user of a service like Facebook etc.



* 1. Employee enters their username and password
  2. The application calls an Identity Broker. The broker captures the username and password
  3. The Identity Broker uses the organization’s LDAP directory to validate the employee’s identity
  4. The Identity Broker calls the new GetFederationToken function using IAM credentials.
     + The call must include an IAM policy and a duration (1 to 46 hrs) along with a policy that specifies the permissions to be granted to the temporary security credentials
  5. The Security Token Service confirms that the policy of the IAM user making the call to GetFederationToken gives permission to create new tokens and then returns four values to the application: An access key, a secret access key, a token, and a duration (the token’s lifetime)
  6. The Identity Broker returns the temporary security credentials to the reporting application
  7. The data storage application uses the temporary security credentials (including the token) to make requests to Amazon S3
  8. Amazon S3 uses IAM to verify that the credentials allow the requested operation on the given S3 bucket and key
  9. IAM provides S3 with the go-ahead to perform the requested operation
* Key Points:
  + Develop an Identity Broker to communicate with LDAP and STS
  + Identity Broker always authenticates with LDAP FIRST, then with STS
  + Application then gets temporary access to AWS resources
* Active Directory Integration
  + You can authenticate with Active Directory using SAML
  + Authenticate to Active Directory first and then assign the temporary security credentials
* Workspaces
  + VDI. Cloud-based replacement for a traditional desktop
  + Windows 7 Experience provided by Windows Server 2008 R2
  + No need an AWS account to login to workspaces
* Elastic Container Service (ECS)
  + ECS
    - Amazon’s managed EC2 container service. Allows you to manage Docker containers on a cluster of EC2 instances
    - Allows you to run and maintain a specified number of instances of a task definition simultaneously in an ECS cluster
    - Regional service that can be used in one or more AZs
    - IAM with ECS to restrict access
    - Security groups operate at the instance level, not at the task or container level
    - ECS Clusters
      * Logical grouping of container instances that you can place tasks on
      * Region specific
      * Container instances can only be part of one cluster at a time
      * Create IAM policies for your clusters to allow or restrict users’ access to specific clusters
      * ECS agent to connect EC2 instances to your ECS cluster. LINUX ONLY
    - Task Definition
      * Required to run Docker containers in Amazon ECS
      * Text files in JSON format that describe one or more containers that form your application
      * Cloud formation template for Docker. Configure things such as the amount of CPU, RAM, etc.
  + Container
    - A method of operating system virtualization that allow you to run an application and its dependencies in resource-isolated processes
    - Created from a read-only template called an image
      * Image is a read only template with instructions for creating a Docker container
      * Images are stored in a Registry, such as DockerHub or AWS ECR
      * ECR (EC2 Container Registry) is a managed AWS Docker registry service

+++++

* RDS
  + Complex transactions or complex queries
  + A medium-to-high query/write rate
  + No more than a single node/shard
  + High durability
* No RDS
  + Massive read/write rates (150k write/second)
  + Sharding
  + Simple GET/PUT requests and queries
  + RDBMS customization
* DynamoDB: Provisioned Throughput
  + Allocates resources based on throughput capacity requirements (read/write)
  + Read capacity unit (for an item up to 4KB in size)
    - One strongly consistent read per second
    - Two eventually consistent reads per second
  + Write capacity unit (for an item up to 1KB in size)
    - One write per second
* Memcached
  + Multithreading
  + Low maintenance
  + Easy horizontal scalability with Auto Discovery
* Redis
  + Support for data structures
  + Persistence
  + Atomic operations
  + Pub/sub messaging
  + Read replicas/ failover
  + Cluster mode/sharded clusters

Private and Public IP can be found in metadata from 169.254.169.254

* Stateful:
  + Holds and protects STATE information.
  + Authentication: Login -> (server holds state) -> view profile (only in one server. Cannot scale)
  + Security Group: Inbound from origin -> (hold state) -> automatically allows outbound to the origin
* Stateless:
  + Don’t hold STATE information
  + Authentication: Login -> (don’t hold state) -> token -> viewProfile in any servers behind LB
  + Lots of caching
  + Separate workload
  + More services are now stateless
  + Network ACL: Inbound from origin -> (don’t hold state) -> outbound needs to be manually allowed
  + E.g.: RDS, DynamoDB, Elasticache
* Auto-Scale
  + Scale In order:
    1. AZ with the most instances
    2. Instances with the oldest launch config
    3. Instances closest to the next billing hour

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Amazon S3** | **Amazon EBS** | **Amazon EFS** |
| Overall | Can be publicly accessible  Web interface  Object Storage  **Scalable**  Slower than EBS and EFS | Accessible only via the given EC2 Machine (not standalone)  File System interface  Block Storage  **Hardly scalable**  Faster than S3 and EFS | Accessible via several EC2 machines and AWS services  Web and file system interface  **Object** storage  **Scalable**  Faster than S3, slower than EBS |
| Purpose | Good for storing backups | Is meant to be EC2 drive (not standalone) | Good for shareable applications and workloads |
| Storage Size | No limit on number of objects | Maximum storage size of 16 TB | No limitation on the size of the file system |
| File Size Limitation | Individual Amazon S3 objects can range from a minimum of 0 bytes to a maximum of 5TB | No limitation on file size in EBS disk | Single files have a maximum size of 47.9TiB |
| Data Stored | Stored data stays in the region.  Replicas are made within the region in multiple availability zones  Amazon S3 objects can be copied to other region using the [cross region replication](http://docs.aws.amazon.com/AmazonS3/latest/dev/crr.html) feature | Data stored stays in the same Availability zone.  Replicas are made within the AZ for higher durability | Data stored in AWS EFS stays in the region.  Replicas are made within the region |
| Data Access | Accessible over internet based on access policy configured | Can only be accessed by a single Amazon EC2 instance | Can be accessed by 1 to 1000s of EC2 instances from multiple AZs, concurrently |
| Supported Encryption Mechanisms | [Server Side Encryption](http://docs.aws.amazon.com/AmazonS3/latest/dev/serv-side-encryption.html) with Amazon S3-Managed Keys (SSE-Amazon S3),AWS KMS-Managed Keys (SSE-KMS), and with Customer-Provided Keys (SSE-C)  [Client Side Encryption](http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingClientSideEncryption.html) using an AWS KMS–Managed Customer Master Key (CMK) and using a client-side master key | Uses an [AWS KMS–Managed Customer](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSEncryption.html) Master Key (CMK) and AES 256-bit Encryption standards | Uses an [AWS KMS–Managed](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSEncryption.html) Customer Master Key (CMK) and AES 256-bit Encryption standards |
| Access Control | Using [bucket policies and user policies](http://docs.aws.amazon.com/AmazonS3/latest/dev/s3-access-control.html)  Managed with ACLs  Pre-signed URL access based on IAM | Security groups  NACL  User based Authentication such as IAM | Security groups  User-based authentication  IAM |