Linux Academy – AWS Essentials

Account Basic Overview : AWS Free Tier

- AWS Free Tier
  - limited free use of AWS Services
  - learn, experiment and get hands on experience with AWS Services.
  - Almost all AWS Services offer some kind of free tier services
    - Available for 12 months after you get an AWS Account
    - Some Services extend Free Tier Usage pass 12 months
  - Free Tier is only available to new AWS Customers
  - You can do a lot with AWS while only using Free Tier Services
- Be very careful verify that the setup does not include options you have to pay for
- Free Services
  - Compute : EC2, ELB, Lambdas
  - Storage: S3, Glacier, EBS
  - Database: RDS, DynamoDB
  - Other Services: SNS, CloudWatch, CloudFront

Account Basics : AWS Free Tier

- Use billing alarms to control cost
  - In the AWS Manage Service find the billing service
  - From the billing dashboard
    - Select Billing Preference
    - Under Cost Management Preferences – Receive Free Tier Usage Alerts and Receive Billing Alerts
- Spin up a bunch of free resources at the same time you could use up a month worth of value

Account Basics : Create an AWS Account

- Create an AWS Account
  - Enter email address and create a password
  - Enter your preferred AWS Account Name
  - Select Personal or Professional
  - Complete the Account Information Form
  - Read and Accept AWS Customer Aggrement
  - Complete pay information
  - Complete Identify Verification From and input the verification pin number
    - Done from the Cell Phone
  - Select a Support Plan : Basic for no Fee/Free Tier use
  - Sign in with no Fee/Free Tier Use

Account Basic : Create an AWS Account

- Each service has an EC2 Dashboard
- See Cloud Practitioner Document : Lets Get Started : What you need to know : Access and tour of the AWS Console
- Account Settings
  - My Account
  - Organization
  - Bill Dashboard
  - Security Credentials
  - Switch Roles
  - Sign out
- Support
  - Support Center Open Support Cases, Recommendation Videos, Usefull Length
    - Your type of support Plan
  - Forms
  - Documentation
  - Training
  - Other Resources

Account Basics : Creating a Billing Alarm

- Goto Service and type in billing to reach the Billing & Cost Management Dashboard
  - Select Billing Preferences
    - Double check Receive Free Tier Usage Alerts is selected
    - Receive Billing Alerts allow you
      - AWS will monitor our usages, charges and recurring fees
      - Send an email when we get close to our thresholds
- Use Cloud Watch to configure a Billing Alarm
  - Goto Cloud Watch Dashboard
    - Under Alarms select Billing and select Create Alarm
    - Select Metric and select By Service or By Charge (select Metric)
    - Select either by Service or total estimate charges
    - Specify the metric and conditions
      - select >= $1.00
    - Configure Actions
      - Alarm either in Alarm, OK, Insufficient Data → Select In Alarm
      - SMS → Simple Notification Service → send notification of an envent (Create a new topic and add title and an email address)
        - Create Topic

Account Basics : Creating a Billing Alarm

        - select an Alarm Name and Description
    - Click Create Alarm

- Whenever you exceed the $1.00 an alarm will triggered which will the person an email and show up in the Alerts

Account Basics : AWS Documentation

- Support → Documentation
  - Guides and API References
  - URL aws.amazon.com
  - Groups same way as the AWS Services
  - For each services
    - Getting Started
    - API Reference
    - Developer Gui
    - Console User Guide
  -

IAM : What is IAM

- From Linux Academy :
  - Identity and Access Management (IAM) : Overview Part 1
  - Service is found : Security Identity and compliance IAM
- The Root account can Launch, Modify or Provision or delete
-

IAM Initial Setup and Configuration

- Best Practices : Guidelines that recommend settings configurations and architecture for maintaining a high level of security, accessability and efficiency
- Best Practices
  - Delete your root access keys : They provide unrestricted access to you AWS resources. Instead use access keys or temporary security credentials.
  - Active MFA Authentication – Setup
    - Click on Manage MFA
      - A dialog box appear with two selections : Virtual MFA Device. A hardware MFA Device
        - Select VFA
      - Must install an MFA compatible software
        - Example of Google Authenticator on ios
          - Scan a QRCode
          - On your computer provide the authentication codes
          - Repeat these two step twice in 15 minutes
            - Switch back to IOS and memorize the number
            - Switch back to the computer and insert the number
        - The MFA device was successfully associated with your account
  - Create an individual IAM User
    - click on Users in the Dashboard
    - On the Add User Dialog Box add the name and the Access Types and click next
    - On the next screen set permission (existing user, clone from another use, attach a policy)
    - Click the next button and review and click the Create User

IAM Initial Setup and Configuration

- A result shows that the account create was successful an the following
  - A login link for the user to login
  - The Secret Access Key
  - The Pasword

- Use Group to assign permission
  - Click on the Manage Groups which will open a window that has “Create New Group”, Group Actions
  - Click on Create New Group
    - Supply a name and click next
    - Attach Policy and click next
    - Verify everything is ok on the Review Screen and Click “Create Group” Button
    - The group will be created.

- Manage password
  - Default Password is 6 characters with the ability to allow the user to change their password

IAM : Users and Policies

- Access Types
  - Programatic Access
    - use to access aws resources from the command lien
    - Need a secret access key and access key
  - AWS Management Console access
    - Setup an custom password
    - Add the policy IAM S3 Full Access Policy
- Since the user has different permissions then other users , the user would get a different link
- For the Add user functionality you can create multiple users at once as long as they got the same options
-

IAM Group and Policies

- Remove a policy from a user
  - Select the user and click on the x button in the row with the policy to delete it
- Create a Group (Dev)
  - From the console select Group and add the users to the group
  - To the Group Add a policy by clicking the Permissions button in the Group Window.
- Example Adrian is in a group with more users and he needs more access
  - Best : Add another group to the user Adrian with the polciies
  - Worst : Add the new policies to the user
- Roles
  - Can Create a role for a AWS Service, Another AWS Account (Cognito or OpenID provider) t, Web Identity, SAML 2.0 Federation
  - Remember A role is made up of a name, policies, A item form the previous line

Network Services : AWS Global Infrastructure

- By using AWS Servers there are closet to your region latency is reduced.
  - latency → time it take to transfer data from one location to another

VPC : Basics

- VPC Conceptual Explanation
  - If you have internet in your home you have a private network
  - Common Components
    - Internet Service Provider provides a connection from the internet to your home
    - wire → Comes into your house from the street
    - Modem → Your connect (gateway) to the internet
    - Wire → Connecting your router to your modem
    - Router/Switch (Can be wired or wireless) →
      - Device that allows you to connect equipment to the network
      - Routes traffic to other equipment on the network or through the modem to and from the internet
    - Local devices : Computer s, cell phones and anything else with networking capabilities
- Dashboard found in “Networking and Content Delivery”
- VPC Default Dashboard
  - When a new user account is create you get a default VPC, but you also get subnets, route tables, internet gateway and network ACL(s)
  -

VPC : Basics

- How does data move through the network
  - We want to access the data from the internet
  - Get Pass the firewall
    - A layer of security to prevent unwanted communication/traffic such as viruses
    - Any type of information we want to access, we need to poke to allow specific traffic to go through
      - Example want to communicate with linux academy.com then we need allow port 80 web traffice through the firewall
  - The Router/Switch would determine if the data should be kept local or passed to the model
- Example using AWS
  - We have a VPN which contains 2 ECS2 Instances and one of them want to access linux academy.com
    - The packets would go throughout Network Access Control List (NACL) behaves lie a virtual firewall
      - Allow Port 80 traffic to communicate the NACL
    - If the access is allowed then the packets go to the route table.
      - If destined for the internet it would send it would sent it to the internet gateway
    - Internet Gateway
      - Would be the modem in the home network and would pass the packets to the linux academy website
      -

Internet Gateways (IGW)

- A combination of hardware and software that provide your private network with a route to the internet
- AWS Definition
  - Horizontally Scaled
  - redundant and highly available
  - Allows communication between instances in your VPC and the internet
    - no availability risk or bandwidth constraints on your network traffic.
- The default VPC already has an IGW attached
- Create a new gateway
- Each VPC can only be attached to one IGW at a time
- A VPC needs an internet gateway to communicate with the rest of the world.
  -

Route Tables

- A set of rules called routes determine where network traffic is directed
  - Direct traffic in a VPC to different subnets or to the internet gateway
- Your VPC already has a “name=”main route table
- Question : Should the route data stay within the VPC or directed to the internet
- Can have multiple route tables
- Example : We have a VPC within Subnet 1 (public)
  - The data goest through the NACL (Network Access Control List)
  - Then the data goes through the route table and the route table determines where the data should be connected
    - A route table can have a name, route table id, Explicitly Associated Main, VPC (which one connected to)
    - Can have a collection of routes which contain
      - Destination, Target, Status, Propagated
      - 172.31.0.0/16 local Active No
        0.0.0.0 IGW-12GDBC Active NO
    - Analysis
      - local → If the IP Address falls within the IP Cider Range then the destination should be found in the local pc
      - If any communication has not been destined then send it to the internet gateway
- If the data is being directed to the IGW and the IGW is detched then for Status the route rule will show Black Hole
- What if we attach another IGW then the route table is still accessing the detached IGW
  - Fixed : Delete the old route and add the new route where the target the new IGW ID
    -

Firewall – Network Access Control List (NACL)

- An optional layer of security for your VPC that acts as a firewall for controlling traffic in and out of of one or more subnets
- The VPC will have an NACL in place and associated with the default subnets
- The NACL is associated with the subnets found in the VPC
- The NACL has both Inbound and Outbound Rules
  - For the default NACL all traffic is allow (both inbound(into a subnet) and outbound (out of a subnet))
  - In order to allow / disallow traffic
  - Each rule has a
    - Number
    - Type
    - Protocol
    - Port Range
    - Source
    - Allow / Deny
- NACL is stateless : You must have rules for both the inbound and outbound
- Allows traffic into and out of the subnets
- A subnet can only be associated with one NACL at a time

Network Access Control List (NACL)

- The default has two rules for inbound
  - Rule (Allow all Traffic)
    - Rule # 100
    - type : All Traffic
    - Protocol: ALL
    - Port Range: ALL
    - Source 0.0.0.0
    - Allow/Deny
  - Rule (default rule get process after all other rules) : Any traffic we don’t explicit allowed will be denied
    - Rule # *
    - type : All Traffic
    - Protocol: ALL
    - Port Range: ALL
    - Source 0.0.0.0
    - Allow/Deny

Network Access Control List (NACL)

- The default has two rules for inbound
  - Rule (Allow all Traffic)
    - Rule # 100
    - type : All Traffic
    - Protocol: ALL
    - Port Range: ALL
    - Source 0.0.0.0
    - Allow/Deny
  - Rule (default rule get process after all other rules) : Any traffic we don’t explicit allowed will be denied
    - Rule # *
    - type : All Traffic
    - Protocol: ALL
    - Port Range: ALL
    - Source 0.0.0.0
    - Allow/Deny

Network Access Control List (NACL)

- Rules
  - Rules are evaluated based on rule # from lowest to highest
  - The first rule evaluated that applies to the traffic type gets immediately applied and executed regardless of the rules that come after
    - Any rules with a number are ignored
  - For the Port Range we can have a range
    - example 1000 - 65535 (can select any port in the range)
      - The SSH can leave any port that it selects
  - NACL
    - Default
      - Allows both inbound and outbound SSH traffic since the SSH has a lower rule #
      - All other type of traffic wil be denied via the catch failsafe deny rule
      - Simply connectivity and you have connectivity
    - New NACL

      When you create a new NACL, all traffic is denied by default
- To Attach the ACL Control edit the subnet association on the Network ACL
  - For Type HTTP (80), the outbound traffic will use ephemeral ports 1024-65535 for the return web traffic and not port 80
- NACL and SG are the two major problems with communication of AWS Resources is not working properly

Subnets

- A sub-section of a network.
- When a VPC is created it spans all Availability Zones in the region
  - Can add one or more subnets in each availability zone
  - Each subnet must reside entirely within one Availability Zone and cannot span zones
- What makes a public/private subnet
  - A public can access a pubic and a private cannot
  - If an subnet is not explicitly route table then it automatically associate with the main route where table (Property Main has Yes)
    - The subnets that have not been assigned to a route table are private by default
  - The route table must have an Internet Gateway associated with it to make it public. The subnets must be assigned to a gateway
    - private : The route table does not have a route to the internet
    - public : The route table has a route to the internet
  -
  -

Availability Zones

- Any AWS resource launched must be placed in a VPC subnet
  - Any subnet must be located in an Availability Zone
  - Utilize multiple Availability Zone to create redundancy in your architecture
    - Allows for High Availability and Fault Tolerant Systems
    - By launching instances in separate availability zones you can protect your applications from failure in a single location
- Need to replicate the data for high availability
  - Some services replicate their data in the region
- How can high availability work ?
  - One is public resource and the other is a private resource
  - Traffic can be redirected to the backup server
- High Availability
  - Create your architecture so the system is always available or has the least down time possible
    - What high availability sounds like
      - I can access my data in the cloud
      - My web site never crashes and is always available to my customers

Availability Zones

- Fault Tolerant
  - The ability of your system to withstand failures in one or more of its compenents
    - What Fault Tolerant sounds like
      - One of the web servers failed, but my backup server immediately took over
      - If something in my system fails it can repair itself
  -

Tricky Questions on the Quiz

- Which of the following is a valid route (IP CIDR Block) for connecting an EC2 hosted website instance to the internet?
  - This is a valid route for connecting an EC2 instance to the internet. 0.0.0.0/0 is the IP CIDR Block value that represents the internet as a whole.

EC2 – Basics

- See notes on
  - Purchasing options (On-demand , reserved, spot)
  - Instance Type
  - EBS Optimized
  - AMI Type
  - Data Transfer
  - Region

EC2 – AMI

- Launch as many instances of an AMI as you need
- Launch an instance from many different AMI to setup all the servers
- Start with an AMI and then deploy the AMI on an instance
- When an AMI is created a template is being created can launch another instance that has the same components as the original
  - add additional applications or configuration that we require
    - people can’t deploy the application incorrectly
  - The new AMI can create multiple EC2 instances
- Three main categories
  - Community AMI
  - AWS aMarketplace AMI
  - My AMI

EC2 – Instance Types

- When an instance is launched the instance type determines the hardware
- Instance Types
  - Family — Categorizing instance types based on what they do
  - Type
  - vCpu
  - Memory
  - Instance Storage
    - 2 Storage Types : EBS and SSD
  - EBS Optimized Available
  - Network Performance

EC2 – Elastic Block Storage

- IOPS – Input output operations per second
  - A unit of measure representing input/output per second.
    - The operations are measured in KiB
    - <ax amount of data tht that a volume type counts as a singleIO
      - I/O size is capped at 256 KiB for SSD since SSD handle small or andom I/O much more quickly
      - 1024 Kib for HD Volumes
  - More IOPS means faster read/write speeds
  - The larger the storage size the more IOPS the volume has
- The root volume is inside the computer. If you uncheck the “Delete on Termination” then you can keep it.
- Encryption is supported by all EBS volume types and has negligable impact on IOPS
- On the EC2 then you can click create volumes
  - Select Volume Type, Availability Zone
  - create volumes based on a snap shot
  - Any volumes can be created and attached/detached at any time
    - can attach or detach the volume at anytime. How can this work for transferring data between applications or EC2 instnace (reliabliltiy and workflow)

EC2 – Elastic Block Storage

- Snapshots
  - A snapshot is an image of an EBS volume that can be stored as a backup of the volume or used to create a duplicate
  - A snapshot is not an active EBS voume. You cannot attach or detach a snapshot to an EC2 instance
  - To restore a snapshot you need to create a new EBS volume using the snapshot as its template
-

EC2 – Security Groups

- Difference between NACL and Security Group is they don’t have number rules
- Cannot create Deny rules with Security Groups
- ELB → Take traffic that come in from the IGW and load balance between the instances
- Security Group are stateful – Any traffic allowed in is allowed out.
- Security Groups have a default deny and we must explicitly create rule to allow traffic

EC2 – IP Addressing

- Providing an EC2 instance with a public IP Address
  - needed to communicate with the internet
  - EC2 instance can be launched with or without a public ip address depending on theVPC/subnet settings
- By default all EC2 instances have a private IP Address
  - private PI address allow for instances to communicate with each other as long as they located in the same VPC
- Under EC2 Creation : Configure Instance Details
  - Auto Assign Public Address (Use subnet setting (Enable))
    - Based of the default VPC – remember initial rule is to make anything run
  - Create your own VPC
    - Add Name IP4 CIDR Block10.0.0.0/24
  - Create Subnet
    - Associate it with the vpc and add the CIDR block
  - Select EC2 instance select the create vpc and the Auto Assign Public IP has Use Subnet Setting as disabled
    -

EC2 – IP Addressing

- Everything EC2 instance need to communicate with IP Address
  - EC2
    - Public IP Address (Internet Access)
      - Security Group (with allow rule for specified traffic)
        - NACL (with allow rule)
          - Route Table (with IGW as a route)
            - Internet Gateway (attached to the VPC)
              - Internet

Launching and Using an EC2 Instance

- Basic Steps
  - Select an AMI
  - Select an instance Type
  - configure Instance Details
    - Run as bash script to install Apache
    - Click on Advance Details
      - Enter the contents of a script to execute.
  - Add Storage
  - Add Tag
  - Configure and Assign a Security Group
  - Review and launch
    - When getting ready to launch the instance from your machine make the essentialskp.pem read only (400)
  - Create and Download a Key Pair
- Trouble Shooting
  - By changing the port you can allow it we had set to 1024-65535 and changed it to 1-65535. We needed to change this for both inbound and outbound

Tricky Quiz Questions

- Which of the following is true about EC2
  - EC2 is elastic – Elastic means that it can add or remove additional instances as required
  - EC2 is horizontally scalable – Can add additional EC2 instnaces on-demand as required
- Which of the following is not true about Spot Instances
  - A spot instance hourly price is a called a Spot Price
  - You no longer have to bid on spot instances
  - Spot instances are great for jobs that can be stopped and resumed later
- Which of the statements are true about EC2 Security Groups
  - You can add allow rules, not deny rules
  - Security groups are stateful
  - When you create a Security Group on creation it has no inbound rules

S3 Basic

- S3 → A online bulk storage system access from any device
- Some AWS services only work with/communicate each other if they are in the same AWS Region
- Create a folder
  - provide a name
  - Chose encryption (type or none)
- Pricing / Cost Overview
  - Free Tier Use is available for S3
  - How are you charged using S3
    - Storage Cost
      - Applies to data at rest in S3
      - Charged per GB Used
      - Price per GB varias based on regiion and storage class
    - Request Pricing – Moving data in/out of S3
      - Put, Copy, Post , List , Get , Lifecycle Transaction Request, Data Retrieval, Data Archival, Data Restruction

Buckets and Object

- See the Cloud Practitioner
- A file can be upload to the bucket or a folder
- Bucket Folder and Object Properties
  - General Info
  - Permissions
  - Static Web Hosting
  - Logging
  - Events
  - Versioning
  - Lifecycle
  - Cross-Region Replication
  - Tags
  - Requester Pay
  - Transfer Acceleration

- Folder Level Properties
  - General Info
  - Details
- Object Level Properties
  - General Info
  - Details
  - Permissions
  - Metadata Assign optional metadata to the object as a name/value pair
- Setting for Object : Object Lock → Prevent a file from being deleted
- Have a tab “Select From” : Retrieve data from CSV, JSON, Request

Storage Classes

- S3 Storage Classes
  - Standard
  - Intelligent Tiering
  - Standard Infrequent Access
  - One Zone Infrequent Access
  - Glacier
  - Glacier Deep Archive
- For Standard Infrequent Access – A retrieval fee is added for pulling the object
- For Glacier , Glacier Deep Archive can pay fee for expatiated retrieval fee
- Setting / Changing Storage Class
  - By default all new objects upload to S3 are set to the Standard storage class
  - If you want new object ot have a different storage class you need to set thigs properly prior to or during the upload process
    - Use Object Lifecycle policies (covered in next section)
  - For the Storage Classes : Standard, Intelligent Tiering, Standard Infrequent Access, One Zone Infrequent Access you can manually switch tbhe object's storage class by changing the storage class in the object's properties
  - To Move to Glacier , Glacier Deep Archive you need to use object lifecycles. The change may take one to two days.

Storage Classes

- When the storage class should be used
  - Standard – Object that are critical document highly avaialble, on need to be access critically
  - Intelligent Tier → Long Live data that has differing standard of access

Object LifeCycles

- An object lifecycle is a set of rules that automate the migration of an object's storage class to a different storage class or its deletion based on spcified time intervals
- Lifecycle Management
  - Functionality is located at the bucket level, but can be applied to the entire bucket, a specific folder inside the buck, one specific object
  - You can always delete a lifecycle policy or manually change the storage class back to whatever you would like
- Example
  - I have a work file that I will access for the next 30 days, after 30 day I may need to access that file once a week for the next 60 days after that will never to access the file again
  - By using a lifecycle policy, I can automate the process of changing the file's storage class to meet my usage needs and keep my S3 storage cost as low as possible
  - Solution
    - Day 0 to 29 : Usage : Very Frequent, Storage Class Standard, Cost highest cost team
    - Day 30 to 89 : Usage : Infrequently, Storage Class Standard Infrequent Access, Cost middle cost tier
    - Day 90 + : Usage: Most likely never needed, Storage Class Glacier, Cost lowest cost tier
    - Create an Object Life Policy
      - Can use a filter to limit scope to prefix/tag and configure class transition the current versions, previous version
      - Select a LifeCycle Stage (One Zone IA) and the day the bucket,object,folder days until the file enters the state
    - Configure Expiration Current and Previous Versions and clean up expired object delete markers, multipart uploads)

Object Permissions

- Have granular control over when you an view, access and use specific buckets and objects
- Permission functionality can be found on the bucket and object level
- On the bucket level you can control
  - List : Who can see the bucket name
  - Upload/Delete - Objects to upload or in the bucket delete
  - Permissions – Add / edit / delete /view permissions
- Buckets level permissions are generally used for internal access control
- On the Object Level
  - Open/Download
  - View Permissions
  - Edit Permission
- Note you can share speciifc object (via a link) with the anyone in the world
- In the Amazon S3 you can select the bucket then select Access Control List and change the permissions on that page
-

Object Permissions

- When we create our bucket we did it access to every one
  - Grant public access so everyone can view the objects (not best practice , but acceptable for a webserver). Select Public Access (List Access)
    - In Block Public Access you would need to turn “Block all public Access” off
    - In order to view the file the users will need view access.
- Sharing an S3 object with world
  - ON the object create the following permissions
    - Grantee = Everyone
    - Check Open/Downloaded
  - Under Actions, select Make Public
  - The link under properties is no live and anyone that has it can directly download the object
  - To remove public access to the object either delete the permissions or remove the bucket that provided public access

Object Versioning Notes From Quiz

- S3 Versioning is a features that keeps track of an stores all versions of an object so that you can access and use an older version
  - Versioning is either on or off
  - Once it is turned on you can only suspend versioning
    - Suspending versioning only prevents versioning going forward.
      - All previous objects with versions will still remain in storage
  - Versioning can only be set on the bucket level, and applies to all objects in the bucket.
- Notes from Quiz
  - What is used to automatically move data through the different class of S3 storage based on date → Lifecycle Policy
  - What is S3 → S3 is object storage that can hold virtually any type of file
  - Which of the following are valid ways to access data that is stored on Amazon S3
    - Using the S3 Dashboard within AWS Console
    - Using an HTTP Url to access the object

ELB Basics

- The Address of the Elastic Load Balancer is provides to the user to connect through the load balancer
  - Instances must be in different availability zones
  - Note that the route table is still present and is below the elastic load balance
  - The configuration of the two EC2 instances must be the same, but in different availability zones
- Pricing
  - No Free Tier not available for the ELB
  - Determinants
    - Each hour or partial hour the load balancer is running
    - For Each GB of data transferred through the load balancer
  - Note prices may vary depending on Region

Creating an ELB

- Select EC2 Dashboard and Select Load Balancer and select Create Load Balancer
  - Two Types of Load Balancers
    - Application Load Balancer –
      - Used for HTTP and HTTPS traffic operating at the request level
      - provide advanced routine and visibility features targeted at application architecture
    - Network Load Balancer –
      - Need Ultra high performances
        - TLS offloading at scale centralized certificate deployment
        - Support for UPDB and static IP address
      - Operate at the Network Level and handle millions of request per second securely
- Have selected Application
  - Step Configure Load Balancer
    - Name,
    - Scheme (internet facing or internal)
    - IP Address Type (ipv4, dual stack)
  - Listeners – automatically http 80 port, could add more listeners
  - Availability Zones – Select a VPC and select availability zones with public subnet

Creating an ELB

- Configure Security Groups
  - Choose a group that allows HTTP Traffic
- Configure Routing
  - Target group
    - Name
    - Target Type – Instance, IP, Lambda function
    - Protocol (ex. HTTP)
    - Port
  - Health Checks
    - Protocol – example HTTP
    - Path – /
  - Advanced Health check Settings
    - Port – Traffic Port or Override
    - Health Threshold – Number of times to pass to determine if its health
    - Unhealthy Threshold – Number of time to fail before considered unhealthy
    - Timeout
    - Interval – Seconds Between health checks
    - Status – The HTTP Code to accept passing

Creating an ELB
Notes from the Quiz

- Register Target – Where the load balancer send data (the target group)
  - Click on Add to register
- Review Page
- Uptime monthly 99.99%
- Notes from Quizzes
  - How can we provision an ELB from the EC2 services fo the AWS Console → Provision an ELB from the EC2 service of the AWS Console
  - Type of Load Balancers – Application and Network Load Balancer
  - what is the minimum number of Availability Zones – 2
  - Elastic Load Balancing automatically distributes incoming application traffic across multiple targets. Which is a valid target → Amazon EC2 Inastances

Introduction to AutoScaling
AutoScaling Basics

- Loading Balancing – The ability evenly distribute the traffic between the EC2 instances so that none of the EC2 instances are overload
- AutoScaling Basics
  - Automates the process of adding (in) and removing (out) based on traffic demand of your application
  - Ensures that you have the correct number of EC2 instances available to handle the load of the application
  - AutoScaling Group – Collection of EC2 Instances
    - Can specify the minimum/maximum number for the group and the autoscaling will never go above or below that
- With an autoscaling the user are distribute by the elastic balancer to the autoscaling group and the autoscaling group will do the in/out (up/down)
- Autoscaling Components
  - Launch Configuration – The EC2 template used when Autoscaling needs to add an additional server to your AutoScaling Group
  - Auto Scaling Group – All the rules and settings that govern when an EC2 server is automatically added or removed
- How are you charged for using ELB
  - Free to use, but you will be charged for the resources that Auto Scaling Provisions (any EC2 instances that go beyond the Free Tier allotment)

Using Autoscaling

- Steps required to create a launch configuration
  - Select an AMI
  - Select an instance type
  - Create a launch configuration
    - Give the lanuch configuration a name
    - Make sure that a public IP address will be assigned
    - (optional) For this demostration we are going to include the basic script to install the Apache Web Software
  - Select/Add storage type
  - configure security group
    - Need Name, Need Size (number of instances), Need Network, Need Subnet
    - Advanced : Load Balancing, Target Groups (Allow our load balancer with our Autoscaling group), Health Check
      - Detailed monitoring would incur charges
    - Create Auto Scaling Group
      - Use Scaling policies to adjust the capacity of the group
      - select the min and max instances
      - Scale Group Size : Name, Metric Type (ex. Average CPU Utilization), Target Value (value of the metric type), Instances needed and warmup time , disable scale in
    - Configure SNS Notifications to select a topic
  - review and create

Route S3 Basics

- Websites admins must register their web domain and IP Address with DNS providers if they want their users to find the website without knowing the IP Address
- The IP Address can be the IP Address of the load balance
- Pricing
  - No Free Tier use is not available
  - Charged based on
    - Number of hosted zones
    - Traffic flow (per policy)
    - Standard queries
    - Latency based routien
    - Geo DNS Queries
    - Health Checks
    - Register/Transfer a Domain

Using Route 53

- Domain Registration
  - Search for and select a domain name that is available
  - Fill out the Contact Details pages
    - Make note of the “Privacy Protection” Option → Hides Contact information from a Domain Lookup (when enabled)
  - Review details and purchase
    - Your amazon account wil not be directly charged for any domain registration or transfer
    - As part of the registration process. Route 53 create a hosted zone automatically for the new domain
  - Complete purchase and wait for the domain registration process to complete
- Hosted Zones : A container where you store information and manage routing for your domain
- Hosted Zones and Record Sets
  - Navigate to Hosted Zones and select the domain name you just registered
    - Here you can visit all the recored sets
    - The pre-populated record set are basic DNS record set that AWS has configured for you (
  - Create tow Type A record set (resolves a host name to the IP Address) for your domain that route to the ELB
    - One recored set for your domain with www and the other without www
    - For Alias Target select the ELB
    - Leave Routing Policy as Simple
    - For this example you can leave Evaluate target Health as No
  - Create Record Set Done
    - It may take 1-5 minues for the record to propogate to DNS Servers
    - Try accessing the domain in a web server to see if it works

Cloud Front Basics

- If the content is already in the edge location with the lowest latency, CloudFront delivers it immediately.
- If the content is not in that edge location, CloudFront retrieves it form the origin that you have defined such as an S3 Bucket
  - Kept in the edge location for 24 hours
- Creating a Basic CloudFront Distribution for S3 Objects
  - Upload the data you want to distribute to an S3 Bucket (For CloudFront URLs to work public read permissions must be explicitly granted to each object)
  - From the Cloud Front Console choose Distribution
  - Select a delivery method for your content page in the Web Section : Choose Get Started
  - On the Create Distribution page unser Origin Setting choose the Amazon S3 bucket that you created earlier
    - Accept the default values for the Origin Id and path (inside the bucket), Restrict Bucket Access and Origin Custom Header
  - Under Default Cache Behavior Setting accept the default values and CloudFront will
    - Forward all request that use CloudFront URL for your distribution to the S3 bucket specified in the previous setp
    - Allow HTTP(s) access to your object sor users
    - Cache your object at CloudFront edge locations for 24 hours
    - See AWS Documentation for a full list of setting related to default cache behavior
- Have Web Base Distribution (Speed up web actions such as acess of static/dynamic content) or RTMP (Speedup Media Files)

Cloud Front Basics
Quiz Questions

- Under Distribution Settings set the desired options including the following key configuration
  - Price Class: Sets the maximum price that you want to pay for CloudFront service
    - By default CloudFront serves you objects from edge locations in all CloudFront Regions, but can be restricted
  - Alternate Domain names : Create a CName (alias) to access your object */image/image.jpg using an ddres like http://www.example.com/images/images.jpg instead of http:dejowekl3s.cloudfront.net/images/image.jpg*
  - *See AWS Documentation for a full list of setting releated to default cache behavior*
- *Choose Create Distribution after distribution the value of the status column for your distribution will change form inProgress to Deployed*
- *Ensure that the alternate domain name(s) or Cnames are configured in Route 53 if they are set in CloudFront*

- *Quiz Questions*
  - *Which AWS services could you use Route 53 with? EC2 and S3*
  - *Which is following are Amazon Route 53 DNS Record Types : A (Address Record)*
  - *DNS stands fro Domain Name System*
  - *In a highly available and fault tolerant system with multiple EC2 instnaces hosting a website then what is the purpose of Route 53*
    - *To populate external DNS servers with domain/IP Address information for the ELB sot aht inbound traffic can beload balanced by the ELB between the EC2 instnaces*
  - *When you type a domain name into a web browser what best describes the process that occurs to deliver the website content back to your browser*
    - *The browser send a request to a DNS server asking for the IP Address associated with the domain name*
  - *Which of the following does Route 53 effectively connect user request to : Elastic Load Balancer, Amazon S3 and Amazon EC2*
  - *Best Describe S3 : A service to register domains and configure DNS Records*

SNS Basics

- Publishers send a message to a topic and the subscribers receive the notification or message from the topics they are subscribe to.
- Pricing
  - Free Tier user is available with SNS
  - How are you charged with SNS
    - Publishers – Number of SNS request (I.e the message that needs to be sent)
    - Notification Deliveries : The number of subscribers that the message is sent to
    - Data Transfer in/out of SNS

Using SNS

- Services → Application Integration is where SNS is located
- Create Top
  - See other notes
  - Display Name → Only with text messages
  - Encryption – Enable Encryption or Disable
  - Access Policy
    - Choose Basic (Use Simple Criteria to define a basic policy) or Advanced (Use a JSON Object to advanced access policy)
      - Basic
        - Define who can publish messages to the topic
          - Only the topic owner
          - Everyone
          - Only the specified AWS Accounts
        - Define who can subscribe to this topic
          - Only the topic Owner
          - Every
          - Only the specified AWS accounts
          - Only the requestors with certain endpoints
  - Delivery retry policy (HTTP/S) – How SMS retries failed deliverables to HTTP/S Endpoints

Using SNS

- Delivery status logging – Configure the logging of message delivery to cloud watch
  - Log delivery status for these protocols -- (AWS Lambda, Amazon SQS, HTTP/S, Platform application endpoint)
  - Success Sample Rate – The percentage of successful message deliveries to log
  - IAM Roles – SNS requires permission to write logs to CloudWatch Logs. You can use separate roles for successful and failed message deliveries
    - Service Role Info
      - Use existing service role – Choose an existing service role from your account
      - Create new service role – Create a new service role in your account
    - IAM Role for successful deliveries
    - IAM Role for failed deliveries
- Tags

- The Topic name is used to create a unique identifier called amazon resource name
- Creating a Subscription
  - Topic ARN
  - Protocol – ex HTTP, Email, AWS Lambda etc…
    - Other fields based on Protocol Selection

Using SNS
Quiz Question and Answers

- Public Message
  - Go Back to the newly create topic and click Public Message
  - Fill in the information needed for the Message (The fields should based on the subscription that was created)
- Create a Topic (Communication Channel) , add a subscription (Endpoints receiving the message) then publish a message to the topic
- Quiz Question and Answers
  - What is the default maximum number of SNS Topics per account → 100000
  - What is the default limit to the number of SNS subscriptions per topic → 12, 500, 000
  - Amazon SNS provides notification services. What type of notification does it use → SNS is a push notification service
  - Which of the following AWS services can actions triggered based on the SNS Topic → EC2, S3, LAMBDA

CloudWatch Basics Quiz Questions.

- When a Cloud Metric is exceeded then cloud watch can either send notification or automatically make changes
- Based on metrics you can set threshold to trigger alarms which can be viewed in CloudWatch
- Charged for
  - Dashboard
  - Detailed Monitoring for EC2 Instances
  - Custom Metrics
  - API Request
  - Logs
  - CloudWatch Event/Custom Events
- Be aware they mentioned that the charges can really add up the more you use it (ex. more dashboards)
-

Cloud Watch Metrics and Alarm

- Found in Services → Management and Governance → CloudWatch
- CloudWatch – Creating a Dashboard
  - Pick Dashboard from the navigation pane, click Create dashboard and give it a name
  - Select Add Widget
  - Explore the available metrics and slect the metric you want to add to the dashboard
  - Select a timer period that makes sense based on the metric
  - Create the widget
- Create a cloudWatch Alarm – Can attach a topic so that the alarms can notify users, AWS Services
  - Pick Alarms in the navigation pane, then choose Create Alarm
  - Select a category
  - Explore the available metrics and select one for which you wan to create an alarm and continue to the next page
  - Enter a name and description and set the threshold for the alarm
  - Set the actions that you want to occur when the alarm threshold is met
  - Set the period and statistics
  - Create Alarm

Cloud Trail – Basics
Quiz Answer and Quesitons

- Pricing
  - Free Tier is not available. However, you can set up a trail that delivers a single copy of management events in each region free of charge will be charged based on S3 usage.
- S3 charges apply base on usage
  - Management events (Management operations such as launch EC2 instance or Create S3 Bucket)
  - Data Events (resource operations performed on/within the resource I.e S3 object level APIs such as Get, Put and Delete Actions, and Lambda function invoke API)
  - Usage Charges (S3, SNS or encryption)
- Quiz Questions
  - You want to get a notification when CPU Utilization on an EC2 instance goes above 80% which of the following is the best method to accomplish that → Create a Cloudwatch Alarm that will trigger when CPU Utilization goes above 80% and have the alarm trigger an SNS Topic (not Cloudwatch Topic) to send you a message
  - Primary Benefits of CloudWatch : Provides monitoring insights into your AWS Resources, Access your data from a single platform (Cloudwatch enables collecting of logs and metrics from a wide variety of resources) and Taking action based on alarms and triggers that can be created
  - Three States of a Cloud Watch Alarm → Alarm , Insufficient Data and OK
  - Ways you can use CloudWatch : Automated Alarm Actions, SNS Notifications, Event Driven Corrective Actions

Lecture RDS and Dynamo DB Basics

- Pricing for RDS
  - Free Tier is available for RDS options except Auroa
  - How you are charged for an RDS
    - The RDS engine you choose (MySQL, Postgres)
    - RDS Instance Class : Very similar to EC2 instance types
      - ex. 16 vCPU 128 Gib RAM EBS 3500 Mbps
    - Purchasing Terms : On Demand or Reserved
    - Database Storage
    - Data Transfer in/out of RDS
    - Configuring RDS is chargeable, avoid doin this if following along in your own ASW account
- Pricing for Dynamo DB
  - Free Tier use is available
  - How are you chared for using DynamoDB
    - Provisioned throughput Capacity
    - Indexed Data Storage
    - DynamoDB Streams
    - Reserved Capacity
    - Data Transfer in/out DynamoDB

Provision an RDS

- Don’t want the RDS to be internet accessible
  - Should be stored in a private subnet
  - The NACL and Security both need to allow the database port to open and accessible
  - The development team should connect into the EC2 instance and the EC2 instance should connect to the database
    - The RDS Database should be in a private subnet
    - We have to make sure SSH Tunneling is configured
      - Give use the ability to connect through the internet ,
        - through the internet, through the internet gateway, through the route table, through the NACL,
        - through the Security Group, through the EC2 instance,
        - **through the private route** table to the RDS Instance
- Configuring a private subnet Group:
  - Within RDS, we need to create a subnet group that contains our two private subnets
    - Need to provision the RDS database into a private subgroup
  -

Provision and RDS Group

- Basic Steps for configuring a private subnet group
  - Navigate to Subnet Groups
  - Create DB subnet group : Choose two availability zones,
  - Complete the for, make sure to select our two private subnets → Required at least two subnets
- Launching an RDS
  - Select an engine → For Free Tier usage, Select the MySQL → Dev/Test Option
  - Specify DB Details: instance specifications and settings
  - Configure Advanced Settings:
    - Network and Security
      - Select the private subnet group : Do not use default
      - Publicly Accessibility should be set to no
    - Database Options
      - Connectivity → Select advance and create a new Security Group
        - Give the Security Group a name
        - After the database is created you will need to edit the inboundrule with port (ex. 3306 so that the ip address is not specific, but 0.0.0.0/0 anywhere)
          - Not a security best practice
        - Need to rule for SSH
        - Question : How is two security groups add
      - Outbound has all traffic
    - Backup
    - Monitoring
    - Maintenance
  - Launch DB Instance

Provision and RDS Group Quiz Notes

- After the database is create a screen will show the endpoint that will be SSHed into to connect to the database
- Connect to the RDS Database
  - Give the connection
  - Select TCPIP over SSH
  - SSH Hostname, Username, Key File (.pem)
  - Port = 3306
  - Username/Password
- Quiz Notes
  - For NoSQL the AWS Jargon is Items and Attributes
  - DynamoDB database is stored in JSON like name-value documents

Lambda Basics

- With Lambda you still need to manage the route 53 settings. All other configuration is managed by lambda
- Able to scale and allow the code to run for the 40 applications
  - The develop does not need to setup the infrastructure
- Concerns : Only thing you need to worry about is how execute your code and the code
- Price
  - Free Tier is available
  - Cost
    - Request to execute code
    - Duration (The length of time it takes the code to execute)
    - Accessing the data from other AWS services/resources

Lambda Test

- Create a Lambda Funciton
  - Select a lambda blueprint that fits your need or Auhor from Scratch
  - Configure the function
    - Give it a name
    - Edit runtime
    - Create or select a role – Needed to work with other services
  - Lambda function Code – Enter/edit the code you want to execute (make sure the code matches the runtime)
    - Gets an ARN
    - Can add a trigger (ex. CloudWatch alarm to monitor EC2 instances; could trigger a lambda function to restart our instance)
    - Have a code Editor
    - Have the options : Environments, Tags, Execution Role, Basic Setting
      - Event Template and Name
    - When executed has a log with Summary with Resiult, Time function takes, Billable duration, Max Memory Used and log input
  - Lambda execution Code – Select or modify the execution role (if necessary)
  - Advanced Settings
    - Allocate the appropriate memory amount
    - Networking
    - Run Lambda inside a VPC (if you choose not required)
    - Encrypt variables using AWS KMS (Key Management Service)
    - Debugging and Error Handling
    - Concurrency
    - Auditing and compliance with CloudTrail
    - AWS Xray
  - Cloud Watch Metric for the Lambda

Lambda Test Quiz

- Executing (testing) the Lambda Function
  - Select the function and click “Test” Button
  - Enter a “Test Event” (if required)
  - Click “Save”
  - Click “Test” and Review the result
- What ar the two primary way you are charged using Lambda
  - Execution and Request Duration not Lambada Trigger
  - Options for creating new lambdas
    - AWS Repository
    - From Blue Print
    - Create from Scratch
- Which of the following services directly trigger a Lambda Function → Cloudwatch
- Which of the following can be used to execute or invoke Lambda Code : AWS Console, AWS CloudTrail, AWS S3, AWS SNS
-