**What is Cloud Computing?**

**What is Cloud Computing?**

* On-demand delivery of compute, database storage, applications, other IT resources through cloud platform via Internet
* Pay-as-you-go pricing

**6 Advantages of Cloud Computing**

1. Trade capital expense for variable expense
   * Only pay for what you use
2. Benefit from massive economies of scale
   * You won't have same purchasing power as Amazon
   * They get cheaper prices to purchase servers, hardware
3. Stop guessing about capacity
   * You'll buy too much or too little. Too much = wasted money, too little
4. Increase speed and agility
   * Websites/apps can scale infinitely with demand
5. Stop spending money running/maintaining data centers
   * Focus on what you're good at, not managing infrastructures
6. Go global in minutes
   * Deploy apps in minutes
   * Provide lower latency and better experience at minimal cost

**3 Types of Cloud Computing**

1. Infrastructure as a Service (IAAS)
   * You manage the server (physical or virtual) as well as the operating system
   * Data center provider has no access to your server
2. Platform as a Service (PAAS)
   * Someone else manages the underlying hardware and operating systems, you just focus on your applications
   * You upload your code and it just executes
   * Think of GoDaddy
3. Software as a Service (SAAS)
   * Think of Gmail
   * All you do is interact with the application, manage the software and how you want to use it
   * Someone else takes care of the infrastructure and everything related to it

**3 Types of Cloud Computing Deployments**

1. Public Cloud - AWS, Azure, Google Cloud Platform
2. Hybrid - Mix of public and private
   * May want to keep some sensitive data on-premise
3. Private Cloud (or on-premise) - You manage it in your data center. Openstack or Vmware

**Around the World with AWS**

**Region**

* Geographic area consisting of 2 or more availability zones

**Availability Zone**

* A data center

**Edge Location**

* CDN Endpoints for CloudFront
* Many more edge locations than regions

**Let's Log Into AWS**

**Support Plans**

1. Basic
2. Developer
   * Experimenting with AWS
   * $29/month
   * One person can ask technical questions through support center, 12-24 hour support rate
3. Business
   * 24/7 support by phone
   * Full access to AWS Trusted Advisor
   * $100/mo
4. Enterprise
   * $15,000/month
   * Everything in business + technical account manager
   * 15 min response time for critical support cases

**Create Billing Alarm**

* Click Name at top-right, click My Billing Dashboard
* Enable Billing Alert
* Add the threshold in Cloudwatch and add e-mail address

**Edge Location**

* Location where content will be cached
* Similar to AWS Region/AZ
* As close to user as possible

**Origin**

* Origin of files that CDN will distribute
* S3 bucket, EC2 instance, Elastic Load Balancer, or Route53

**Distribution**

* Name given to the CDN which consists of a collection of edge locations
* First time a user goes to a website, it'll check a local edge location to see if website asset is there
* If not, it will download the asset from the origin and cache it to the edge location
* Next time someone tries to access, they will get the cached version from a local edge location
* Reduces stress on web servers and increases speed to download large files

**Distribution Types**

1. Web distribution - Used for websites
2. RTMP - Used for media streaming

**Setting up CloudFront**

1. Choose Web distrbution
2. Origin Domain Name: Choose an S3 bucket
3. Origin Path: You can choose subdirectories for your origin

* Once it's deployed, you will see a domain name. Use that and the name of a file in your bucket to access.

**Exam Tips**

* Content comes from Origin
* Cached at a local Edge Location
* Takes awhile for the first person to access, much quicker every time after that because it's cached geographically close to you

**EC2 (Elastic Cloud Compute)**

**Setup**

* **VPC:** Virtual data center in the cloud
  + Deploy all EC2 instances into a VPC
* **AMI:** Using Amazon Linux AMI because it includes stuff to connect to AWS
* **Instance:** Choosing t2.micro because it's usually used to test in dev
* **Instance Details:**
  + **Network:** Keep default VPC
  + **Subnet:** Choose which availability zone you want to be put into
  + **Auto-assign Public IP:** Allows you to assign a public IP so you can SSH into instance
  + **Shutdown behavior:** Choose what happens if your EC2 instance turns out (stop or have Amazon terminate for you)
  + **Enable termination protection:** Prevents people from accidentally shutting down your instance
* **Storage:**
  + 8GB is default
  + **Volume Type:** General purpose is most common, Provisioned IOPS lets you choose a very fast disk (database server), Magnetic is a very slow disk (file server)
* **Tags:** Allow us to add tags like Department and Employee ID to help with cost tracking later on
* **Security Group:** Virtual firewall in the cloud
  + Open ports like 22 for SSH or 3389 for RDP (Windows) or 80 for HTTP

**Connect to the EC2 server**

* Open Terminal
* chmod 400 MyVirginiaKP.pem - Protects file from accidental overwriting
* ssh ec2-user@54.242.147.206 -i MyVirginiaKP.pem to connect
* sudo su for root
* yum update to update security patches