**AWS CSA-Pro Notes ACloudGuru Section 9 Cost Management**

**Exam Tips:**

**Costing in General:**

**Understand the difference between CapEx and OpEx models**

* CapEx: Money spent on long term assets (buildings, property)
* OpEx: Variable expense that is used to keep the business running

**Understand TCO, ROI, and the challenges faced in these activities**

* TCO: A comprehensive look at the entire cost model of a given decision or option, often including both hard costs and soft costs
* ROI: The amount an entity can expect to receive back within a certain amount of time given an investment

**Cost Optimization Strategies:**

**Know conceptually the variety of ways customers can approach cost optimization on AWS**

* Page 3:
* Managed Services, Appropriate Provisioning, Right Sizing, Purchase Options, Geographical Selection, Optimize Data Transfer

**Fully understand the Cost Optimization Pillar whitepaper**

**Tagging and Resource Groups:**

**Understand the various common uses for tagging and ways you can implement/enforce a tagging strategy**

* Common uses: Security, Automation, Compliance, Cost Management
  + You can use tagging in IAM policies so if you have a group for QA people, they can only use instances with QA in the tag for example
* You can use Config rules with specific tagging standards to make sure your org is following them and enforce your tagging strategy

**Know when and how to create Resource Groups, and don’t be tricked into thinking they are anything more than a logical grouping**

* Resource Groups are a way to logically group your resources by common tags.
  + Examples: You can have Resource Groups by environments
  + Stage, Dev, Prod, ETC.

**Spot and Reserved Instances:**

**Know the differences and limitations for the different types of Reserved Instances, including Zonal and Regional**

* Standard:
  + Cost less
  + Can sell on the marketplace
  + Change AZ, Instances Size, Networking Type
* Convertible:
  + Cost more, but benefits from price reduction
  + Change AZ, Instances Size, Networking Type
  + Change instance family, OS, Tenancy, Payment Options
* Zonal: If you select an AZ upon creation then it is Zonal to that region. You can change this afterwards
* Regional: If no AZ is selected then the reserved instances applies to any EC2 instances that matches the specs you just bought as long as you have more reserved.

**Understand how Spot instances work and when they are best used**

* You put in a Spot bid, if your bid is higher then the current price you get those instances at their current price, once the price gets higher then your bid those instances get terminated
* Good for temp instances spun up via Auto-Scaling
* Also good for testing new features at usually a cheaper price to on-demand instances

**Understand Dedicated Instances and Dedicated Hosts**

* To get single tenancy on your instances you need to have Dedicated Instances/Hosts
* Dedicated Instances:
  + Instances that are on hardware not shared by any other client. They could have instances from your account that are not Dedicated on that hardware
* Dedicated Hosts:
  + A dedicated host for you to choose how many instances are on it
  + Also gives you deeper insights to your instances. Good for products with licensing that need this insight like per core

**Cost Management Tools:**

**Know how and when you would use AWS Budgets**

**Understand the benefits of consolidated billing**

* Consolidated Billing allows you to take advantage of Economies of Scale
* It is best practices to have multiple AWS accounts from a security perspective
* You can combine those accounts to have 1 payer account and you get price benefits the more you use AWS AKA Economies of Scale

**Know how Trusted Advisor can help customers optimize and improve their AWS landscapes**

* Purely a check list on common AWS best practices in the Security, Cost Optimization, Performance, Fault Tolerance, and Service Limits categories

**Concepts:**

**Terms:**

**Capital Expenses-** Money spent on long term assets like property, buildings and equipment

**Operational Expenses-** Money spent for on-going costs for running the business. Usually considered variable expenses.

**Total Cost of Ownership-** A comprehensive look at the entire cost model of a given decision or option, often including both hard costs and soft costs

**Return on Investment-** The amount an entity can expect to receive back within a certain amount of time given an investment

* Many times, orgs don’t have a good handle on their full on-prem data center costs
* Soft costs are rarely tracked or even understood as a tangible expense
* Learning curve will be different from person to person
* Business plans usually include many assumptions which in turn require support org to create derivative assumptions sometimes layers deep

**Cost Optimization Strategies:**

**6 Strategies:**

**Appropriate Provisioning:**

* Provision the resources you need and nothing more
* Consolidate where possible for greater density and lower complexity
* CloudWatch can help by monitoring utilization of resources

**Right Sizing:**

* Using lowest-cost resource that still meets the technical specifications
* Architecting for most consistent use of resources is best versus spikes and valleys
* Loosely coupled architectures using SNS, SQS, Lambda, and DynamoDB can smooth demand and create more predictability and consistency

**Purchase Options:**

* For permanent applications or needs, Reserved Instances provide the best cost advantage
* Spot instances are best for temporary horizontal scaling
* EC2 Fleet lets you define target mix of on-demand, Reserved, and Spot instances

**Geographic Selection:**

* AWS Pricing can vary from region to region
* Consider potential savings by locating resources in a remote region if local access is not required
* Route 53 and CloudFront can be used to reduce potential latency of a remote region

**Managed Services:**

* Leverage managed services such as MySQL RDS over self-managed options such as MySQL on EC2
* Cost savings gained through lower complexity and manual intervention
* RDS, Redshift, Fargate, and EMR are great examples of fully-managed services that replace traditionally complex and difficult installations with push-button ease

**Optimize Data Transfer:**

* Data going out and between AWS regions can become a significant cost component
* Direct Connect can be a more cost-effective option given data volume and speed

**Tagging and Resource Groups:**

**Tagging:**

* Helps manage your AWS assets
* Tags are just arbitrary name/value pairs that you can assign to virtually all AWS assets to serve as metadata
* Tagging strategies can be used for Cost Allocation, Security, Automation, and many other uses
  + For example, you can use a tag in an IAM policy to implement access controls to certain resources
* Enforcing standardized tagging can be done via AWS Config Rules or custom scripts
* Can have up to 50 tags

**Resource Groups:**

* Resource Groups are groupings of AWS assets defined by tags
* Create custom consoles to consolidate metrics, alarms, and config rules around given tags
* **Common Resource Groupings:**
  + Environment (DEV, QA, PRD)
  + Project Resources
  + Collection of resources supporting key business processes
  + Resources allocated to various departments or cost centers

**Spot and Reserved Instances:**

**Reserved Instances:**

* Purchase usage of EC2 instance in advance for a significant discount over On-Demand pricing
* Provides capacity reservation when used in a specific AZ
* AWS billing automatically applies discounted rates when you launch an instance that matches your purchased Reserved Instance
* EC2 has 3 Reserved Instance types: Standard, Convertible, and Scheduled: you can also buy them for RDS as well
* Can be shared across multiple accounts within Consolidated Billing
* If you find you don’t need your Reserved Instances, you can try to sell them on the Reserved Instance Marketplace

**Standard:**

* Terms: 1 year, 3 year
* 40$ - 60% discount
* Change AZ, Instances Size, Networking Type: Yes via ModifyReservedInstance API or console
* Change instance family, OS, Tenancy, Payment Options: **NO**
* **Does not benefit from Price Reductions**
* Only can sell standard instances on the reserved instance marketplace

**Convertible:**

* Terms: 1 year, 3 year
* 31% - 54% discount
* Change AZ, Instances Size, Networking Type: Yes via ExchangeReservedInstance API or console
* Change instance family, OS, Tenancy, Payment Options**: YES**
* **Does Benefit from Price Reductions**
* Selling convertible instances on the marketplace is said to be coming soon

**Reserved Instance Attributes:**

* Instance Type – designates CPU, RAM, Networking capability
* Platform – Which OS do you want it to be? Linux, Windows
* Tenancy – Default (shared) or Dedicated tenancy

**Zonal Reserved Instance:**

* If you select an AZ upon Reserved Instance creation then that Reserved Instance only applies in that AZ
* You can change it to a Regional RI by doing the ModifyReservedInstance API

**Regional Reserved Instance:**

* If no AZ is specified then no reservation is created, but the discount is applied to any instance type in that whole region

**Spot Instances:**

* Excess EC2 capacity that AWS tries to sell on a market exchange basis
* Customer creates a Spot Request and specifies AMI, desired instance types, and other key information
* Customer defines highest price willing to pay for instance. If capacity is constrained and others are willing to pay more, you instance might get terminated or stopped
* Spot request can be a “fill and kill”, “maintain”, or “duration based”
* For “One-Time Request”, instance is terminated and ephemeral data lost
* For “Request and Maintained”, instance can be configured to Terminate, Stop or Hibernate until price point can be met again
* Spot Instances price and demand can vary from AZ to AZ

**Dedicated Instance:**

* Virtualized instance on hardware just for you
* May share hardware with other non-dedicated instances in the same account
* Available as On-Demand, Reserved Instances, and Spot Instances
* Cost additional $2 per hour per region

**Dedicated Host:**

* Physical servers dedicated to just your use
* You then have control over which instances are deployed on that host
* Available as on-demand or with Dedicated Host Reservation
* Useful if you have server-bound software licenses that use metrics like per-core, per-socket, or per VM
* Each dedicated host can only run one EC2 instances size and type

**Cost Management Tools:**

**AWS Budgets:**

* Allow you to set pre-defined limits and notifications if nearing a budget or exceeding the budget
* Can be based on Cost, Usage, Reserved Instance Utilization, or Reserved Instance Coverage
* Useful as a method to distribute cost and usage awareness and responsibility to platform users

**Consolidated Billing:**

* Enable a single payer account that is locked down to only those who need access
* Economies of scale by bringing together resource consumption across accounts

**Trusted Advisor:**

* Runs a series of checks on your resources and proposes suggested improvements
* Can help recommend cost optimization adjustments like reserved instance or scaling adjustments
* Core checks are available to all customers
* Full trusted advisor benefits require a Business or Enterprise support plan

**Extra Notes:**

What are some common uses of Tagging? How can you enforce a Tagging strategy?

You can use it for security, automation, and cost management.

You can use AWS Config to enforce your tagging strategy by creating a rule that act when tags are not correctly configured.

Can Dedicated Hosts be used in an auto scaling group? What about Dedicated Instances?

Dedicated hosts cannot be used in an autoscaling group while dedicated instances can.

Can you move Reserved Instances between regions?

You cannot move reserved instances between a region.