AWS Certified Cloud Practitioner

Cloud Concepts

What is cloud computing?

Cloud computing is the practice of using network of remote servers via the internet to store, manage and process data instead of using local, on-premise servers.

* **Cloud**
  + Someone else owns the servers
  + You use someone else’s IT staff
  + Some one else pays the rent
  + You have access to an infrastructure that you have to manage and the risks are only related to your code and configuration instead of anything else
* **On-Premise**
  + You own servers
  + You need IT staff
  + You pay rent
  + Manage risk

Type of cloud deployment models

* **On-Premise/Private Cloud**

Fully on-premise, primarily using virtualization software on “private clouds”. Used mostly for government-reasons or regulatory compliance.

* **Community Cloud**

Several organizations pull together resources and invest in a shared cloud infrastructure.

* **Full/Public Cloud**

For startups and new projects

* **Hybrid**

Both cloud and on-premise

Business needs and Cloud solutions

* **Resiliency**

Our applications can mitigate issues, is fault tolerant and can self-heal

* **Security**

Good people in, wrong people out. Have control over security rules.

* **Durability**

Our data needs to be durable and we can prevent data-loss

* **Performance**

High end-user performances and speed

* **Cost Effectiveness**

Improve cost management

* **Scalability**

Ability to react to increase and decrease in usage and demand

* **Automation**

Let manual tasks be automated, reducing human error

* **Agility and Flexibility**

Respond quickly to changes in plans and goals

* **Efficiency**

Ability to move at rapid pace in development cycle

Advantages of cloud

* **Variable costs instead of capital**

No-upfront cost, pay on-demand

* **Economies of scale’s benefit**

Better benefits in terms of costs

* **Dynamic scaling for capacity**

Scale up or down based on utilization needs

* **Agility and speed**

Launch resources quickly

* **Reduce server maintenance costs**

Focus only on products and not on servers

* **Multi-region deployment**

Can use servers in different regions seamlessly without having to build specific on-prem data centers

Types of cloud computing

* **Software as a Service (SaaS)**

Complete product run and managed by the provider (Eg: Salesforce)

* **Platform as a Service (PaaS)**

Don’t need to manage underlying infrastructure, can focus on management and deployment of software (Eg: Heroku)

* **Infrastructure as a Service (IaaS)**

Access to networking features, computers, data storage (eg: AWS)

AWS Global Infrastructures

* 22 Regions

Physically isolated geographical areas with multiple AZs. Each region has at least two AZs and the largest AWS region is **US-East** (Northern Virginia)

* + **GovCloud Regions:**
    - Special regions only available in the US
    - Ysed to host sensitive information and regulated workloads
    - Only operated by US citizens on US territory
    - Compliant with several US Government bodies
    - Two regions (US-West and US-East).
* 69 Availability Zones

Discrete, AWS-owned data centers. Represented by Region Code followed by letter identifier (Eg: us-east-1**a**). Can deploy on multiple AZ to allow for failover configurations when some AZs go down. There is <10ms latency between AZs.

* A ton of Edge Locations

Discrete data centers owned by AWS-certified partners that has direct connection to AWS network. Used for CloudFront, Route 53, S3 Transfer Acceleration and API Gateway, which allow for low latency no matter where the user is geographically located.

Technology

Core Services

* **Amazon VPC**

Chooses a range of IP addresses in a region. Soft limit of 5 VPCs per region.

* + Logically isolated network
  + Created per account, per region
  + Spans a single region
  + Can use all AZs within one Region
  + Can peer with other VPCs
    - Allows communication among different services in separate VPCs
  + Internet and VPN gateways
  + Numerous security mechanisms
* **Subnets**

Subnets are specific to a single availability zone and each needs it’s own specific set of IP addresses.

* + **Three-tier Architecture**
    - Load Balancing tier
    - Application tier
    - Database tier

Achieved by creating separate tier per each AZ. Separate subnets for LB, App server and DB in each AZ.

This organization enables security, high-availability, fault tolerance and high performances.

* **Routing and Firewalls**

A **public subnet** is a subnet that has communication with the internet outside of the local environment. This is enabled by **internet gateways** and **routing tables**.

* + **NACL**

Firewall for the entire subnet as a whole. We can choose many specific security profiles, such as port traffic from local or public environments or VPN specific access and more.

* + **Security Groups**

Firewall for specific machines/instances. All traffic is denied by default, so we can only create allow rules.

* **DNS, VPN and Direct Connect**
  + **Route 53 (DNS)**
    - Register domains
    - Use AWS nameservers
    - Public and private DNS zones
    - Automated via API
    - Endpoint health checks
    - Different routing methods
      * Latency ⇒ Push users to endpoints that provide them the best latency
      * Geographic ⇒ Push users to locally optimized endpoints
      * Failover ⇒ Push users to working endpoints in case of failures
      * Weighted sets ⇒ Push percentage of users based on rulesets
  + **AWS Hardware VPN**

For connecting on-premise data centers or locations to the AWS VPCs via hardware-enabled IPSEC VPN tunnel

* + **AWS Direct Connect**

Dedicated optical-fiber connection to a provider in an AWS region.

* **EC2**

Provides virtual machines (instances)

* + Can be Linux or Windows
  + Uses Xen or Nitro hypervisor
  + Bare metal is available
  + Combination of CPU, memory, disk, network IO
  + Can launch as many instances as necessary (20 is soft limit)
  + Hourly fee includes the OS license
  + Marketplace offers canned, pre-packaged solutions
  + Usually multi-tenant environment, can be single-tenant via dedicated instances or fully dedicated hosts.
  + Different pricing structures
    - **On-Demand** (least commitment)

Default pricing structure for EC2 instances. No upfront cost and no commitment. Pay per hour or per second depending on instance. Best for short-term, spikey or unpredictable workloads and for testing/experimenting

* + - **Spot** (biggest savings)

AWS has unused compute capacity that it can sell to maximize utility of unused servers. Spot instances provide up to 90% discount compared to on-demand but can be terminated at any time if the computing is needed by an on-demand customer. Charged on hourly base, if your instance is terminated by AWS you don’t get charged for partial hours.

Most useful for load balancing, flexible workloads or big data workloads where applications are feasible only at very low computing costs.

* + - **Reserved** (best long-term)

For apps that have steady-state, predictable usage or need reserved capacity.

Dependent on

* + - * **Term:** length of commitment (1 or 3 year)
      * **Class offering:**
        + **Standard**: up to 75% cost reduction but can’t change instance types
        + **Convertible**: up to 54% and can change instance type to greater or equal in value
        + **Scheduled**: used for specific time periods
      * **Payment option:** all upfront, partial upfront or no upfront

Unused RIs (Reserved Instances) can be sold in the RI Marketplace and RIs can be shared between multiple accounts within an organization.

* + - **Dedicated** (most expensive)

Used to meet regulatory requirements of server-bound licensing that does not support multitenancy or cloud deployment. Basically like buying a dedicated hardware and you are physically separated from other customers. Can be on-demand or reserved (70% off on-demand price).

* + **Amazon Machine Image (AMI)**

Bit-for-bit copy of the machine root

* + - Can be Linux or Windows
    - Different providers
      * AWS provided
      * Trusted Publishers (Microsoft, Canonical, RedHat, etc…)
      * Community-built
      * AWS Marketplace
  + **Numerous use cases**
    - Enterprise applications
    - Web servers
    - Relational databases
    - NoSQL datastores
    - Video transcoding
    - Batch processing
    - Container orchestration
    - Code repos/build servers (Jenkins)
  + **Best practices for EC2**
    - Treat as disposable
    - Design for failure
    - Immutable infrastructure

If the machine is launched for a specific workload, it should not change. If change is needed, machines should be thrown away and replaced

* + - Logs as streams
    - Leverage roles
    - Automate deployments
    - Monitor with CloudWatch
    - Enable scaling and self-healing with Auto Scaling
* **Elastic Block Storage (EBS)**
  + Difference between block storage and object storage
    - Block storage
      * Direct access to memory blocks in a file
      * Uses an efficient protocol (high ## of ops per second)
      * Can change data dynamically
      * Can be mounted
      * Best for random IO
    - Object storage
      * Changing an object means deleting and reuploading an updated copy of the object
      * HTTP for transfer
      * Cannot be mounted
      * No random IO

EBS is built to have durable storage that can be attached to EC2 instances while being independent from them.

* + Connected over network
  + Pay for provisioned storage
  + Exist in single AZ, must be in the same AZ as the instance it connects to
  + Point-in-time snapshot ⇒ Can be used to restore data or recreate data in different regions
  + Can be detached
  + Can be attached to different instances
  + Can be encrypted (AES-256)
  + Can be used in RAID or LVM
* **Elastic Load Balancing**

Fully managed service to distribute incoming requests and traffic among EC2 instances

* + Spans region, uses every AZ
  + Inherently secure, resilient and scalable
  + Support health checks
  + Integrates with Auto Scaling
  + Integrates with Route 53 (receives its own DNS endpoint)
  + Different types
    - Classic (ELB)
      * All of above
    - Application (ALB)
      * Works on target groups
      * Level 7 content filtering ⇒ rule-based URI targeting
      * Allows path/hostname routing
      * Dynamic port mapping ⇒ allows one instance to receive on different ports
    - Network (NLB)
      * Works on target groups
      * Connections to static IPs can be spread across instances
      * Improves latency and handles long-running connections
* **Simple Storage Service (S3)**
  + Inexpensive object storage
  + Buckets and objects
  + Clusters spans regions
  + Distributed and duplicated across every AZs
  + Durable to loss of 2 AZs
  + Server side encryption (SSE) AES-256
  + Use cases
    - Static HTML
    - CSS, JS
    - Images
    - Audio and video
    - PDF, eBooks
    - Software downloads
    - Log collection
    - Data lakes
    - EBS Snapshot storage

Advanced Core Services

* **Amazon Cloud Watch**
  + Monitoring instrument that collects metrics from services
    - Over/under provisioning
    - Current demand/load
    - Bottlenecks
    - Component performance
    - Idle resources
  + Store metrics for up to 2 weeks
  + Accessible via API
  + Unique metrics set per service

Examples:

* + - EC2 ⇒ Instance status, CPU usage, Disk usage
      * Default 5 min interval
      * Detailed 1 min ($/instance)
      * Reported by hypervisor
    - ELB ⇒ Request/min, #healthy/unhealthy hosts
      * Default 1 min interval
    - RDS ⇒ Concurrent connections, usage
    - Dynamo DB ⇒ Read/write throughput
  + Custom metrics ($/metrics)
  + Can generate graphs, overlay metrics and run analytics
  + **CloudWatch Alarms**
    - Triggered on breach of threshold
    - Does not necessarily signal emergency
    - Can trigger different events, such as Auto Scaling or instance Termination
    - Up to 5000 alarms per account (soft-limit
  + **CloudWatch Logs**
    - Collect logs by streaming
    - Configure agent to follow an instance
    - Collect Route 53 DNS queries
    - Monitor CloutTrail Events
    - Default retention is indefinite
    - Can be archived to S3
    - Can be streamed to Amazon Elasticsearch
    - Can be processed with Lambda
    - Log data can be searched and filtered
      * Search using specific syntax
      * Create metric filters
        + Count as custom metric
        + Different use case

Number of 404s

Bytes transferred

Customer conversion

Number of exceptions

* **Auto Scaling**
  + Replace failed instances ⇒ Self-healing infrastructure
  + Change capacity according to load
  + Maintain fixed size fleet
  + Works with CloudWatch
  + Supports events
    - Notifications via SNS
    - Actions via Lambda
  + Used to create reliable applications
* **CloudFront**
  + Starts with a Cloud Distribution (web or streaming)
  + Will cache content at edge locations to improve performance to end user
  + Content comes from Origins, that can be S3 buckets, Load-balancing fronted applications or other combinations of services
  + Custom behaviors need to be specified for CloudFront to know where to pull content from, as well as things such as Time To Live (TTL) and other factors
  + Content can be static or dynamic
  + Custom domain name for each distribution (CNAME or ALIAS)
  + Custom SSL certificates
  + RTMP and HLS streaming
* **Lambda**
  + Serverless, event-driven computing
  + Can be scheduled with CRON jobs
  + Great for
    - Scheduled tasks
    - Microservices
    - Event Handers
  + Pay for computer time per 100ms
  + Create functions
    - Inline editor
    - Upload ZIP
  + Invoke functions
    - CLI or SDK
    - Events
  + AWS Handles
    - Infrastructure
    - Deployment
    - Scaling
* **Relational Database Service (RDS)**
  + Choice of engine
    - MySQL
    - SQL Server
    - Oracle
    - PostgreSQL
    - MariaDB
    - Amazon Aurora
      * Compatible with MySQL 5.6 and PostgreSQL 9,6
      * 3-5x performance increase
      * Autoscaling storage up to 64TB
      * Six copies across three AZs
      * Up to 15 read replicas
      * Aurora Multi-Master ⇒ Read and write to multiple nodes
    - Aurora Serverless
  + Reduced operational burden
  + Focus on application
  + Read replicas
  + Automated tasks
    - OS and DB installation
    - OS patches
    - DB engine minor updates
    - Failover
    - Backups
      * Automated
        + Performed once a day
        + You control backup window
        + Up to 35 days of retention
        + Point-in-time restore
        + Deleted along with RDS instance
      * Manual Snapshot
        + Performed at any time
        + Can copy to other regions
        + Retained indefinitely
  + Multi-AZ deployments ⇒ Best production practice
    - High availability
    - Physically distinct
    - Synchronous replication
    - Automatic failover
      * Loss of network
      * Computer failure
      * Storage failure
    - RTO/RPO to minimum
  + Database migration
    - Dump and import
    - Backup/restore
    - AWS Database Migration Service
      * Supports widely-used DBs
      * Heterogeneous/homogeneous DBs
      * Virtually zero downtime
      * Schema conversion tool
      * Consolidate DBs
      * Continuous replication
* **DynamoDB**
  + Fully managed NoSQL data store
  + Exclusively backed by SSD volumes
  + Single-digit millisecond response time at any scale
  + Built-in security, resiliency, fault-tolerance, durability
  + Replicated across multiple AZs
  + No limits to storage and throughput
  + Need to provision reads and writes throughput (pay for provisioned)
  + Throughput can auto-scale
  + DynamoDB Tables
    - Tables, items, attribute
    - No joins/relationships
    - Schema-less
    - Key value and documents
    - Unique primary key is required
    - Secondary indexes
    - No table size limit
    - 400KB item size limit
    - Item-level TTL
    - Use cases
      * Ad impression/clickthrough
      * Gaming leaderboards
      * Shopping carts
      * Session/state storage
      * Operational state/history (Jenkins jobs logs, CI/CD logs)
* **Glacier**
  + Archival storage for long-term, low-access storage
  + Lower cost compared to S3
  + Write archives
    - Transition from S3
    - Direct upload
  + Download via retrieval request and not query (3-5 hour wait or $ for faster retrieval)
  + Lifecycle rules ⇒ Automatic archival rules from S3 to Glacier
* **Redshift**
  + Fully managed petabyte scale data warehouse
  + Fork of PostgreSQL 8.0.2 and SQL compliant
  + Can connect via JDBC/ODBC
  + Parallel queries via columnar design
  + Ideal for OLAP and BI applications

Other services

* **Tools for Automation**
  + AWS
    - AWS Command Line Interface (CLI)
    - AWS Software Development Kits (SDKs)
    - AWS Cloud Formation
    - AWS Elastic Beanstalk
    - AWS OpsWorks
  + Third Party
    - Troposphere
    - Terraform
    - Ansible
    - SaltStack
* **Elastic Beanstalk**
  + Application management platform (Heroku like)
  + Ideal for developer
  + Choice of Java, .NET, Node.js, Docker and more
  + Simply upload application
  + Define the environment for the application to run
  + Automatic handling of
    - Capacity provisioning
    - Load balancing
    - Auto scaling
    - Monitoring
    - Deployment
* **Cloud Formation**
  + Template-based infrastructure management
  + Declarative programming
  + Offers access to all AWS services
  + Store in source control
  + Write once, deploy many times
  + Library for common architectures
  + No imposed models
* **Web Application Firewall (WAF)**
  + Layer 7 content filtering
  + Support rules to block/allow/count
  + Integrates with Amazon CloudFront
  + Protects against
    - SQL Injections
    - Cross-Site Scripting XSS
  + Block based on
    - IP address
    - HTTP headers/body
    - URI String
  + Rate limiting per client IP
  + Managed rules for common threats
    - OWASP
    - Bots
    - Common Vulnerabilities and Exposures (CVE)
* **AWS Shield**
  + DDoS protection service
  + Standard ⇒ Free for everyone
    - UDP reflection
    - SYN floods
    - SSL renegotiation
    - Slow loris attacks
  + Advanced ⇒ Additional cost
    - Near real-time visibility
    - Integrates with WAF
    - Access to DDoS response team

Security Concepts

Shared Responsibility Model

* **Security of the cloud ⇒ AWS**
  + Physical security
    - Facilities/data center
    - Edge locations
    - Rack and chassis
  + Network
  + API
  + Hypervisor
  + Managed services
* **Security in the cloud ⇒ Customer**
  + Operating System in instances
  + Network and firewall configuration (VPC, Subnets, NACLs, etc…)
  + Identity and access
    - Credentials
    - Permissions
  + Applications
  + Data
  + Encryption
    - At rest
    - In transit
* **Controls**
  + Inherited
    - Physical
    - Environmental
  + Shared
    - Patch management
    - Configuration management
    - Education
  + Customer specific
    - Application
    - Zone security

Identity and Access Management

* **Overview**
  + Authentication
  + Authorization via policies
  + Can manage
    - Users
    - Groups
    - Password Policy
    - Multifactor Authentication
  + IAM is for use with AWS API via Access ID + Secret Key
* **Users and Groups**
  + Users
    - Are created and exist within IAM service
    - Can login to Management Console
    - Can have long-term access keys
    - Can enable per-user MFA device
  + Groups
    - Can’t be nested
* **Credential Types**
  + Email address + password ⇒ Master account (root) access
  + Username + password ⇒ AWS Web Console as IAM user
  + Access key + secret key ⇒ CLI and SDK
* **Policies**
  + Determine authorizations
  + Written in JSON
  + Policy types
    - Managed Policy
      * AWS Managed
      * Customer Managed
    - Inline Policies
      * Policies written directly on IAM users, groups or roles
  + Can be created via
    - Generator
    - Hand written policies
  + Evaluation logic
    - Default to implicit deny
    - Look for explicit deny
    - Look for explicit allow
  + What not to do
    - Embed credentials in
      * Code
      * ENV variables
    - Share with
      * Third parties
      * Hundreds of enterprise users
      * Millions of web users
* **Roles**
  + Use temporary credentials
  + Delegate permissions to
    - EC2 instance
    - AWS service
    - A user (elevate privileges)
    - Separate accounts (cross account access or sharing)
      * One you own
      * Third party
    - Roles for EC2 instances
      * Security Token Service provides temporary credentials that the EC2 hypervisor retrieves via the Instance Metadata Service
      * Credentials are rotated 5 minutes prior to expiration
    - Role for Cross Account Access
      * Accounts that need a trust relationship to access different resources owned by other accounts
      * Policy on the role to allow access to resources and on the user to be allowed to assume the role
* **Federated users**
  + Organizational users
    - Leverage existing directories
      * LDAP
      * Active Directory
    - Temporary credentials
    - Single sign on (SSO)
  + Web/mobile application users
    - Apps bypass backend APIs/proxies
* **AWS New Account first steps**
  + Enable CloudTrail (service that records all calls made to AWS API)
  + Create admin user in IAM
  + Enable MFA on root account
  + Enable Cost and Usage Report (detailed cost reports)
  + Log out of root account
  + Log back in as admin user
  + Create additional users/groups
* **IAM Best Practices**
  + Root credentials
    - Email address + password
    - Protect at all costs
    - Do not use for day-to-day
  + Follow principle of least privilege
  + Rotate access keys
  + Enable MFA
  + Enable CloudTrail

AWS Organizations

* Eases management of multiple AWS accounts
* Automate creation of AWS accounts
* Service Control Policies (SCPs)
  + Control who can use what services
* Organization Breakdown
  + AWS Account (Master) ⇒ Root of organization
    - Organizational units
      * AWS accounts
* Consolidated billing
  + One bill, many accounts
  + Aggregated volume pricing
  + Reserved instances apply to all accounts
* Detailed billing
  + Published to S3 buckets
  + Import into spreadsheets/Redshift
  + Filter/Sort by services, tags, etc…

AWS Assurance Programs

* Certifications/Attestations, Laws/Regulations/Privacy and Alignments/Frameworks
  + Achieving Compliance
    - Customers works towards compliance
    - AWS provides necessary documentation
    - Know who is responsible
      * AWS for physical controls
      * Customer for logical controls
    - Know the controls you inherit
    - Know the services in scope
    - Speak with Account Manager
  + Some AWS certifications/compliances
    - **Cloud Security Alliance**
    - **ISO 9001**/27001/27017/27018
    - **PCI DSS 3.2 Level 1**
    - SOC1/ SOC2/SOC3
    - FedRAMP, FIPS, FISMA, **HIPAA**, ITAR, MPAA
  + Key certifications
    - HIPAA Compliance
      * Designed to secure Protected Health Information (PHI)
      * AWS not “directly” certified but requirements map to FedRAMP and NIST 800-53
      * AWS provides Business Associate Addendum
    - PCI DSS 3.2 Level 1 Compliance
      * Designed to protect Cardholder Data (CHD) and Sensitive Authentication Data (SAD)
      * Applies to businesses that store/process/transmit such data
      * Customers are responsible for card data environment (CDE)
      * AWS provides Attestation of Compliance (AOC)

Services for Auditing and Compliance

* **AWS Config**
  + Fully managed
  + Resource inventory
  + Configuration history
  + Change notifications
  + Determine compliance against custom rules
  + Enables
    - Compliance auditiong
    - Security analysis
    - Change tracking
    - Troubleshooting
* **AWS Service Catalog**
  + Manage catalogs of approved IT services
  + Achieve consistent infrastructure governance
  + Customer defines products in portfolio as CloudFormation templates
* **AWS Artifact**
  + Access reports/details of >2500 security controls
  + On-demand access to AWS security and compliance documents
  + Demonstrate security and compliance of your AWS environments (eg: SOC and PCI reports)
* **AWS CloudTrail**
  + Records all calls made to AWS APIs
  + Delivers log files to S3 buckets
  + Include
    - Identity
    - Source IP
    - Request/Response details
  + Does not record
    - OS system logs
    - Database queries
* **Encryption and Key Management**
  + Many services offer encryption
    - S3
    - EBS
    - RDS
    - Glacier
    - SQS
  + AWS Key Management Service (KMS)
    - Fully managed service
    - Create/manage encryption keys
    - Integrates with many services
    - Multi-tenant software backed by Hardware Security Modules
  + AWS Cloud HSM
    - Single tenant Hardware Security Module
    - FIPS 140-2 Level 3 validated
    - On-demand, no upfront cost
    - Can enable
      * SSL offloading
      * Private key storage/security
      * Transparent Data Encryption (TDE)

Vulnerability and Penetration Testing

* Permission is required
* Must request permission via root credentials
* Must identify instances to be tested
* Must specify start and end date/times
* Permitted On:
  + EC2 Instances, NAT Gateways and ELB
  + RDS
  + CloudFront
  + Aurora
  + API Gateways
  + AWS Lambda and Lambda@Edge functions
  + Lightsail resources
  + Elastic Beanstalk environments
  + Some DNS zone walking
* Not permitted on:
  + m1.small
  + t1.micro
  + t2.nano
* Prohibited Activities:
  + DNS zone walking via Route 53 hosted zones
  + DoS, DDoS, simulated DoS, simulated DDoS
  + Port flooding
  + Protocol flooding
  + Request flooding (login request flooding, API request flooding)

AWS Pricing and Billing

Compute Pricing

* **EC2 On-Demand**
  + No long-term commitments
  + Transform large fixed costs into smaller variable costs
  + Fees include OS license
  + Per hour billing ⇒ Hour forward (mainly Windows)
  + Per second billing ⇒ Minimum 60 seconds
* **EC2 Reserved Instances**
  + Up to 75% discount vs. On-Demand
  + Provide capacity reservation
  + 1- year or 3-year term
  + Specify attributes
    - Instance type
    - Platform
    - Tenancy
    - AZ (optional)
  + Types
    - Standard RIs
      * Up to 75% off On-Demand
      * Best for steady-state usage
    - Convertible RIs
      * Up to 54% off On-Demand
      * Can change attribute of instance
    - Scheduled RIs
      * Available during time window
      * Best for predictable, recurrent workloads
* **EC2 Spot Pricing**
  + Up to 90% off On-Demand
  + Uses spare capacity
  + Price
    - Set by instance type and AZ
    - Determined by supply and demand
  + Spot instances can be interrupted by AWS
  + Best for
    - Apps with flexible start/end
    - Development/test environment
    - Short lifecycle tasks
    - Example cases
      * Image rendering
      * Video transcoding
      * Analytics
      * Machine Learning
  + Integrated with
    - Amazon EMR
    - Auto Scaling
    - Elastic Container Services (ECS)
    - CloudFormation
* **AWS Lambda**
  + Charged per GB per 100ms
  + Charged per 1M requests
  + Includes a free tier
    - 1M requests/month
    - 400,000 GB-sec/month
  + Additional charges for
    - Bandwidth
    - S3 Storage used
* **Data Transfer Pricing**
  + Inbound is generally free
  + S3 to CloudFront is free
  + Outbound to internet
    - $/GB/Month
    - Applies cross-region traffic
    - Tiered pricing ⇒ 10/40/100TB
  + Cross-AZ traffic $/GB/Month
  + VPC peering $/GB/Month

Database Pricing

* **RDS Pricing**
  + Instance type
  + Database engine (license)
  + Reserved instance discount
  + Paid per instance, so multi-AZ deployment means multiplied cost per AZ
  + Storage ⇒ $/GB/Month or $/provisioned IOPS/Month
* **DynamoDB Pricing**
  + Charged for provisioned throughput
    - $/Write capacity unit
    - $/Read capacity unit
  + Charged for storage consumed
    - $/GB/Month

Storage Pricing

* **EBS Pricing**
  + $/GB/month of provisioned storage
    - General purpose SSD
    - Provisioned IOPS storage
    - Magnetic volumes
  + Additional costs for
    - $/provisioned IOPS
    - EBS Snapshots to Amazon S3
* **S3 Pricing**
  + $/GB/Month of consumed storage
  + Additional costs for PUT, COPY, POST, LIST, GET requests ⇒ $/1000
  + Storage Classes
    - Standard
    - Infrequent access
    - Glacier

Cost Management

**AWS Calculators**

* **Simple Monthly Calculator**
  + Select region, services and details ⇒ Get accurate picture of monthly bill
* **Total Cost of Ownership Calculator**
  + Calculates difference between running workloads on-premise compared to running them on AWS
  + Includes capex costs, data center costs, regulatory cost and others

**Cost Management Tools**

* **Cost Explorer**
  + Visualize, understand and manage AWS costs and usage over time
  + Master accounts can consolidate multiple member accounts within an AWS Org
  + Has a forecasting feature to get an idea of future costs
  + Different filtering and grouping functionalities
* **Cost and Usage Reports**
  + Access highly detailed billing information
  + CSV files saved to S3 bucket
  + Can be ingested directly into Redshift or QuickSight for analytical purposes
  + Usage listed for each service
  + Usage listed for tag
  + Can aggregate to daily or monthly totals
* Budgets
  + Plan service usage, service costs and instance reservation
  + First two budgets are free, each other is $0.02 per day, for a maximum of 20,000 budgets
  + Sends alert if you approach or exceed your defined budgets, either via provided email or via chatbot
  + Cost, usage and reservation budgets that can be tracked monthly, quarterly or yearly
  + Alerts support EC2, RDS, Redshift and ElastiCache reservations
  + Fixed cost or planned upfront based on levels you choose
  + Can be managed via AWS Budgets dashboard or AWS Budgets CLI.

**AWS Trusted Advisor**

* Automated analysis of the AWS environment
* Offers best practices recommendations for security, saving money, performance, service limits and fault tolerance.
* 5 categories of recommendations
  + Cost Optimization
  + Performance
  + Security
  + Fault Tolerance
  + Service Limits
* 7 core checks available to everyone
  + S3 Bucket Permissions
  + Security Groups (specific ports unrestricted)
  + IAM Use
  + MFA on Root Account
  + EBS Public Snapshots
  + RDS Public Snapshots
  + Service limits (hard/soft limits)
* Full benefits are available to Business and Enterprise support plans

**AWS Support Plans**

* Basic
  + Core Trusted Advisor Checks
  + No technical support
  + Can submit
    - Bug reports
    - Feature requests
    - Service limit increases
* Developer
  + Core Trusted Advisor Checks
  + Business hours access to Cloud Support Associates
  + Guidance < 24 business hours
  + Impairments < 12 business hours
  + Only offers general guidance
* Business:
  + Full set of Trusted Advisor checks
  + 24/7 access to Cloud Support Engineers
  + Email, chat, phone
  + 1/24 hour response time
  + Offer contextual guidance based on use-case
* Enterprise:
  + Full set of Trusted Advisor checks
  + 24/7 access to Cloud Support Engineers
  + Email, chat, phone
  + 15-minute to 24-hour response time
  + Offer consultative review based on use-case
  + Applies to all accounts in an organization
  + Access to Well-Architected review
  + Access to online labs
  + Dedicated Technical Account Manager

Quick Services Summaries

AWS Database Services

* **Dynamo DB**

NoSQL key/value DB (like Cassandra)

* **Document DB**

NoSQL document DB that is MongoDB compatible

* **RDS**

Relational Database Service, supports multiple engines (MySQL, Postgres, Maria DB, Oracle, Microsoft SQL Server)

* + **Aurora**

Fully managed MySQL (5x faster) and PSQL (3x faster)

* + **Aurora Serverless**

Only runs when needed

* **Neptune**

Managed graph database

* **Redshift**

Columnar DB, petabyte warehouse

* **ElastiCache**

In-memory DB (Redis, Memcached)

* **QLDB**

Fully managed ledger database that provides a transparent, immutable, and cryptographically verifiable transaction log

* **Timestream**

Time series data for IoT applications, DevOps, industrial telemetry

AWS Provisioning Services

Provisioning is the allocation or creation of resources and services to a customer

* **Elastic Beanstalk**

Service for deploying and scaling web apps and services developed in Java, .NET, PHP, Node.js, Python, Ruby, Go and Docker (like Heroku)

* **OpsWorks**

Configuration management system that provides managed instances of Chef and Puppet

* **Cloud Formation**

Infrastructure as code (JSON or YAML)

* **AWS QuickStart**

Pre-made packages that can launch and configure AWS compute, network, storage and other services required to deploy a workload on AWS

* **AWS Marketplace**

Digital catalogue of thousands of software listings from independent software vendors that can be used to find, buy, test, and deploy software

AWS Computing

* **Elastic Compute Cloud (EC2)**

Highly configurable server (CPU, memory, network, OS, etc)

* **Elastic Container Service (ECS)**

**Docker as a Service.** Highly scalable, high-performance container orchestration service that supports Docker. Pay per EC2 instance.

* **Fargate**

Microservices management without thinking about the infrastructure. Just define the task and container. Pay per task.

* **Elastic Kubernetes Service (EKS)**

**Kubernetes as a Service.** Makes it easy to deploy, manage and scale containerized applications using K8S.

* **Lambda**

Serverless functions that run code without provisioning or managing servers. Pay per computational time.

* **Elastic Beanstalk**

Orchestrates various AWS services like EC2, S3, SNS, CloudWatch, autoscaling and Elastic Load Balancers (like Heroku)

* **AWS Batch**

Plans, schedules and executes batch computing workloads across the full range of AWS compute services and features like EC2 and Spot Instances.

AWS Storage

* **Simple Storage Service (S3)**

Object store (basically an hard drive in the cloud)

* **S3 Glacier**

Low cost storage for archiving and long-term backups. Extremely cheap but very slow and incurs in a fee for data retrieval. Useful for compliance, where records have to be kept for a long time and it is unlikely that you need to actually use them.

* **Storage Gateway**

Hybrid cloud storage with local caching (for on-prem and cloud)

* **Elastic Block Storage (EBS)**

Hard drive in the cloud to attach to EC2 instances. Can be either SSD, IOPS SSD, Throughput-optimized HDD or Cold HDD.

* **Elastic File Storage (EFS)**

File system that can be mounted to multiple EC2 instances at the same time.

* **Snowball**

Physical data migration services

* + **Snowball**

Physically migrate lots of data (50-80 TB) in a computer suitcase

* + **Snowball Edge:**

Better version of Snowball (100 TB)

* + **Snowmobile**

A shipping container full of hard drives, basically a data center on wheel (100 PB)

AWS Business Centric Services

* **Amazon Connect**

A cloud-based call center service you can set up quickly based on the same system used by the Amazon customer service teams. Calls can be saved for analysis and workflows can be scheduled for call routings.

* **WorkSpaces**

Secure managed service to quickly provide Linux or Windows Virtual Remote Desktops. Can scale up to thousands of desktops.

* **WorkDocs**

Content creation and collaboration service with content saved centrally in AWS (basically Sharepoint)

* **Chime**

Platform for video conferencing, business calling and online meetings which scales to meet capacity needs

* **WorkMail**

Managed business email, contacts and calendar service with support with mobile email client apps (IMAP) (basically Gmail on AWS)

* **Pinpoint**

Marketing campaign management system that can be used for sending targeted email, SMS, push notifications and voice messages

* **Simple Email Service (SES)**

Cloud-based email sending service for marketing, notification and emails

* **QuickSight**

BI tool to connect multiple data sources and create dashboards without any programming.

AWS Enterprise Integration (going hybrid)

* **Direct Connect**

Dedicated Gigabit network connection from your premises to AWS.

* **VPN**

Secure connection to AWS network

* + **Site-to-Site VPN**

Connecting on-premise to your AWS network

* + **Client VPN**

Connecting a client (eg. a laptop) to your AWS network

* **Storage Gateway**

A hybrid storage service that lets on-prem applications to use AWS cloud storage. Can be used for backup and archiving, disaster recovery, cloud data processing, storage tiering and migration.

* **Active Directory**

Also known as Managed Microsoft AD, enables directory-aware workloads and AWS resources to use managed Active Directory in the AWS cloud.

* **AWS Outposts**

AWS Outposts is an integrated hardware rack that can run native AWS or VMware environments that seamlessly connect to Amazon’s public cloud. Amazon will deliver, install and do maintenance and repairs on Outposts

AWS Logging Services

* **CloudTrail**

Logs all API calls (SDK, CLI) between AWS services

* + Detect developer misconfiguration
  + Detect malicious actors
  + Automate responses
* **CloudWatch**

A collection of multiple services

* + **Logs**

Performance data about AWS Services (eg. CPU usage, Memory, network in app logs eg. Rails or Nginx, Lambda logs)

* + **Metrics**

Time-ordered set of data points for monitoring

* + **Events**

Trigger an event based on a condition (eg. every hour take a snapshot of a server)

* + **Alarms**

Trigger notifications based on metrics

* + **Dashboard**

Create visualizations based on metrics

AWS Analytics

* **Athena**

An interactive query service that makes it easy to analyze petabytes of data in S3 with no data warehouses or clusters to manage.

* **Elastic MapReduce (EMR)**

Easily run and scale Apache Spark, Hive, Presto, and other big data frameworks. With EMR you can run Petabyte-scale analysis at less than half of the cost of traditional on-premises solutions and over 3x faster than standard Apache Spark. For short-running jobs, you can spin up and spin down clusters and pay per second for the instances used. For long-running workloads, you can create highly available clusters that automatically scale to meet demand.

* **CloudSearch**

Manages all the server resources needed to build and deploy search indexes. All you have to do is upload your data to a search domain and start submitting requests.

* **Elasticsearch Service (ES)**

Makes it easy to set up, operate, and scale an Elasticsearch cluster in the cloud.

* **Kinesis**

Easily collect, process, and analyze data streams in real time.

* + Kinesis Data Streams

Real-time data capture. Collect gigabytes of data per second from a data stream and make it available for processing and analyzing in real time.

* + Kinesis Data Firehose

Load real-time data. Process and deliver streaming data with data delivery stream. Prepare and load real-time data streams into data stores and analytics tools.

* + Kinesis Data Analytics

Get insights in real time. Analyze streaming data with data analytics application.

* **Data Pipeline**

AWS Data Pipeline helps you move, integrate, and process data across AWS compute and storage resources, as well as your on-premises resources.

* **Data Exchange**

Over a thousand subscriptions containing data sets from more than 80 qualified data providers.

* **Glue**

Fully managed ETL service that makes it easy to prepare and load data for analytics. Serverless, pay only per computational use.

* **Lake Formation**

Create and manage data lakes with centralized access control across AWS

* **Managed Streaming for Apache Kafka (MSK)**

Fully managed, highly available, and secure service to migrate, build, and run real-time streaming applications on Apache Kafka

AWS Networking and Content Delivery Network

* **CloudFront**

Amazon CloudFront is a fast CDN service that securely delivers data, videos, applications, and APIs to customers globally with low latency, high transfer speeds, all within a developer-friendly environment.

* **Virtual Private Cloud (VPC)**

A logically isolated section of the AWS cloud where you can launch AWS resources in a virtual network that you define

* **Route 53**

Register new domains, transfer existing domains, route traffic for your domains to your AWS and external resources, and monitor the health of your resources

* **API Gateway**

Create and manage APIs to back-end systems running on Amazon EC2, AWS Lambda, or any publicly addressable web service. You can generate custom client SDKs for your APIs, to connect your back-end systems to mobile, web, and server applications or services.

AWS Compliance Programs

A set of internal policies and procedures of a company to comply with laws, rules and regulations or to uphold business reputation

* Health Insurance Portability and Accountability Act (HIPAA)
* Payment Card Industry Data Security Standard (PCI DSS)

AWS Artifact

No, cost, self-service portal for on-demand access to AWS’ compliance and security reports based on global compliance frameworks.

Amazon Inspector

* **Hardening**

The act of eliminating as many security risks as possible

AWS Inspector runs a security benchmark against specific EC2 instances. Can perform both Host and Network assessments

* **Steps to run assessments**
  1. Install AWS agent on EC2 instance
  2. Run assessment for assessment target
  3. Review findings and remediate security issues

AWS Web Application Firewall (WAF)

Protects web applications from common web exploits. Write your own rules to allow or deny traffic based on the contents of HTTP requests. Use a ruleset from a trusted AWS Security Partner in the AWS WAF Rules Marketplace. WAF can be attached to CloudFront or an Application Load Balancer.

* OWASP Top 10 most dangerous attacks
  + SQL Injection
  + Broken Authentication
  + Sensitive data exposure
  + XML External Entities (XXE)
  + Broken Access control
  + Security misconfigurations
  + Cross Site Scripting (XSS)
  + Insecure deserialization
  + Using components with known vulnerabilities
  + Insufficient logging and monitoring

AWS Shield

Managed DDoS protection service. Standard version has no additional charges and every AWS customer can benefit from it. When you route traffic through Route 53 or CloudFront you use AWS Shield Standard. It protects you against Layer 3, 4, and 7 attacks (Network, Transport, Application).

* **Shield Standard (free)**
  + Protection against most common DDOS attacks
  + Acces to tools and best practices to build DDoS resilient architectures.
  + Automatically available on all AWS services
* **Shield Advanced ($3,000/Year)**
  + Additional protection against larger and more sophisticated attacks
  + Visibility into attacks
  + 24/7 acces to DDoS experts for complex cases
  + Available on:
    - Route 53
    - CloudFront
    - Elastic Load Balancing
    - AWS Global Accelerator
    - Elastic IP (EC2 and Network Load Balancer)

Penetration Testing

* **What is PenTesting?**

An authorized simulated cyberattack on a computer system, performed to evaluate the security of the system.

For other simulated events you need a request for authorization from AWS.

Guard Duty

* **Intrusion Detection System (IDS)/Intrusion Protection System (IPS)**

A device or software that monitors a network or system for malicious activity or policy violations

Guard Duty is a threat detection service that continuously monitors for malicious and suspicious activity and unauthorized behavior. It uses ML to analyze AWS logs from CloudTrail, VPC Flow and DNS. It will alert you on findings and you can automate incident response via CloudWatch Events or 3rd party services.

Key Management Service (KMS)

A managed service that makes it easy to create and control encryption keys used to encrypt your data.

* Specifics
  + Multi-tenant Hardware Security Model (HSM)
  + Many AWS services are integrated to use KMS to encrypt data with just a checkbox
  + Uses Envelope Encryption

When you encrypt your data it is protected but you have to protect the encryption key. Envelope encryption is when you encrypt your data key with a master key as an additional level of security

Amazon Macie

Fully managed service that monitors S3 data access activity for anomalies and generates alerts when it detect risks of unauthorized access and data leak risks. Works by using ML to analyze CloudTrail logs. It identifies most at-risk uses which could lead to compromises.

* Alert types
  + Anonymized access
  + Config compliance
  + Credential loss
  + Data compliance
  + File hosting
  + Identity enumeration
  + Information loss
  + Location anomaly
  + Open permissions
  + Privilege escalation
  + Ransomware
  + Service disruption
  + suspicious Access

Security Groups vs Network Access Control Lists (NACLs)

* Security Groups
  + Acts as a firewall at the instance level.
  + Implicitly denies all traffic, so you need to create allow rules
* NACLs
  + Acts as a firewall at the subnet level
  + You create allow and deny rules

AWS Virtual Private Network (VPN)

Enables establishment of a secure and private tunnel from a network or device to the AWS global network

* **Site-to-Site VPN**

Connecting on-premise to your AWS network

* **Client VPN**

Connecting a client/user (eg. a laptop) to your AWS network or on-premises networks