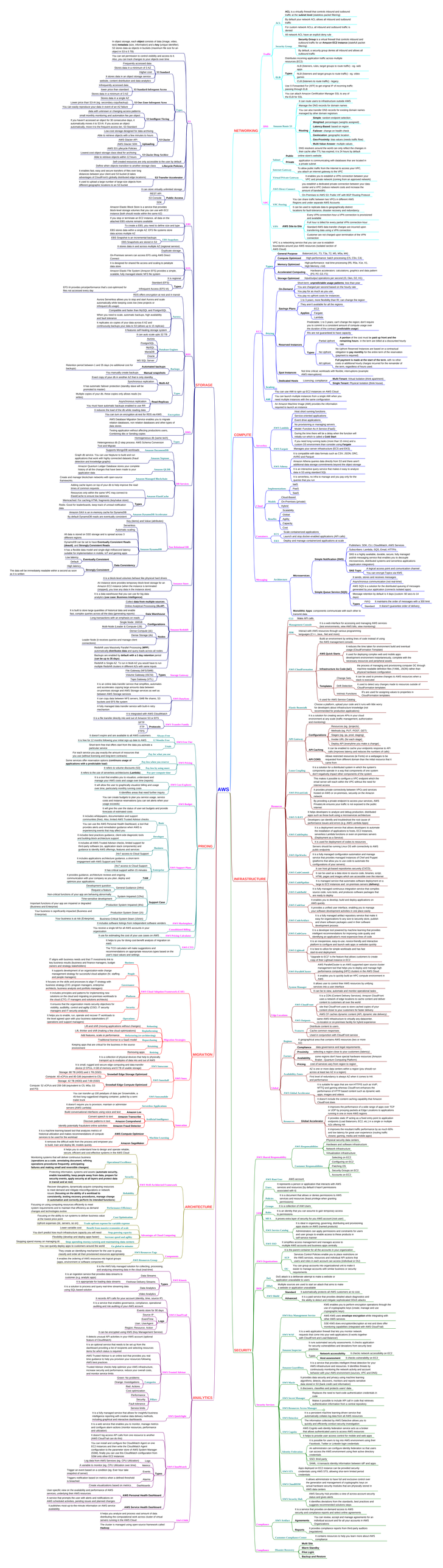
AWS
STORAGE
S3
S3 Standard
In stores data in an object storage service
website, content distribution and data analytics
Infrequently accessed data
Lower price than standard
Stores data in a minimum of 3 AZ
Stores data in a single AZ
Lower price than S3 in-zq, secondary copy/backup
You can easily reproduce your data in event of an AZ failure
data with unknown or changing access patterns
small monthly monitoring and automation fee per object
S3 Intelligent-Tiering
If you haven't accessed an object for 30 consecutive days, it automatically moves it to the cheaper access tier. If accessed, automatically move it to the frequent access tier. S3 Standard
Low-cost storage designed for data archiving
Able to retrieve objects within a few minutes to hours
S3 Glacier
AWS Glacier API
AWS Glacier SDK
Uploading
AWS S3 Lifecycle Policies
Lowest-cost object storage class ideal for archiving
S3 Glacier Deep Archive
Able to retrieve objects within 12 hours
Self-created resources are only accessible to the user by default
Define when objects transitions to another storage class
Lifecycle Policies
A way to share S3 buckets across AWS accounts
Transfer large amounts of data over long distances between your client and S3 bucket (takes advantages of CloudFront's globally distributed edge locations)
S3 Transfer Accelerator
A used to upload a large number of large-size objects from different geographic locations to an S3 bucket
Types
Public Access
It can store virtually unlimited storage
REST API
S3 Console
SDK
Amazon Elastic Block Store is a service that provides block-level storage volumes that can be attached to an EC2 instance (both should reside within the same AZ)
If you stop or terminate an EC2 instance, all data on the attached EBS volume remains intact
To create a EBS, you need to define size and type
EBS volumes are automatically replicated within its AZ. EFS file systems store data across multiple AZs
EBS Snapshots are incremental backups
EBS Snapshots are stored in S3
EBS
It stores data in an across multiple AZ (regional service)
Duplicate storage
On-Premises servers can access EFS using AWS Direct Connect
It is designed for shared file access and scaling to petabyte data store
EFS
Amazon Elastic File System (Amazon EFS) provides a simple, scalable, fully managed elastic NFS file system
It is regional
Standard (EFS)
Types
EFS lifecycle management that's cost-optimized for files not accessed every day
Infrequent Access (EFS IA)
RDS offers encryption at rest and in transit
Aurora Serverless allows you to stop and start Aurora and scale automatically while keeping costs low (low-priorities or infrequent workloads)
Compatible and faster than MySQL and PostgreSQL
Aurora
When you need to scale, automatic backups, high availability and fault tolerance
It replicates six copies of your data across 4 AZ and continuously backups your data to S3 (allows up to 15 replicas)
It features self-healing storage
It can auto scale upto 32 TB
Aurora
PostgreSQL
MySQL
MariaDB
Oracle
MS SQL Server
Database Engines
RDS
Retention period between 1 and 35 days (no additional cost for backups)
Automated backups
Backup
You manually create backups
Manual snapshots
Exact copy of your db in another AZ that is only standby
Synchronous replication
Multi-AZ
It has automatic failover protection (standby slave will be promoted to master)
Types
Multiple copies of your db, these copies only allows reads (no writes)
Asynchronous replication
Read-Replicas
You must have automatic backups enabled to use RR
It reduces the load of the db while reading data
You can turn on encryption at-rest for RDS via KMS
Encryption
AWS Database Migration Service enables you to migrate database to cloud, non-relation database and other types of database
Migration
Moving Application Infrastructure Services, Combining db or Sending copies
Homogeneous db (same tech)
Types
Heterogeneous db (2 step process: AWS Schema Conversion Tool and Migrate
Supports MongoDB workloads
Amazon DocumentDB
Graph db service. You can use Neptune to build and run applications that work with highly connected datasets (Fraud detection and knowledge graphs)
Amazon Neptune
Amazon Quantum Ledger Database stores your complete history of all the changes that have been made to your application data
Amazon QLDB
Create and manage blockchain networks with open-source frameworks.
Amazon Managed Blockchain
DB Services
Adding cache layers on top of your db to help improve the read times of common requests
Resources only within the same VPC may connect to ElastiCache to ensure low latencies
Amazon ElastiCache
Memcached For caching HTML fragments (Session stores)
Types
Redis: Good for leaderboards, keep track of unread notification data
Amazon DynamoDB Accelerator
Amazon DAX is an in-memory cache for DynamoDB
By default DynamoDB reads are eventually consistent
Key (items) and Value (attributes)
Serverless
Automatic scaling
All data is stored on SSD storage and is spread across 3 different regions
DynamoDB can be set to have Eventually Consistent Reads (default) and Strongly Consistent Reads
It has a flexible data model and single-millisecond latency suitable for mobile/ecommerce, IoT and gaming apps
Low latency
Default
Eventually Consistent
High latency
Strongly Consistent
Data Consistency
The data will be immediately readable within a second as soon as it is written
Amazon Redshift
It is a petabyte-scale data warehouse service
Nodes
Leader Node (it receives queries and manage client connections)
Redshift uses Massively Parallel Processing (MPP) automatically distributes data and query loads across all nodes
Backups are enabled by default with a 1 day retention period (can be up to 35 days)
Redshift is Single-AZ. To run in Multi-AZ you would have to run multiple Redshift clusters in different AZs with same inputs
File Gateway (NFS/SMB)
Volume Gateway (iSCSI)
Tape Gateway (VTL)
Types
Tape Gateway
It is an online data transfer service that simplifies, automates and accelerates copying large amounts data between on-premises storage and AWS Storage services as well as between AWS Storage services
Snowball
It can copy data between NFS servers, SMB file shares, S3 buckets and EFS file system and vice versa
It fully managed data transfer service with built-in wire mechanism
It is integrated with AWS CloudWatch
It is a file transfer directly into and out of Amazon S3 or EFS
Protocols
FTP
FTPS
SFTP
It doesn't expire and are available to all AWS customers
Alexa Fund
It is free for 12 months following your initial date to AWS
12 Months Free
Short-term free offer starts from the date you activate a particular service
Trials
For each service you pay exactly the amount of resources that you use, without licensing and long-term contracts
Pay for what you use
Some services offer reservation options (continuous usage of applications with a predictable load)
Pay less when you reserve
It refers to volume discounts (S3)
Pay less by using more
It refers to the use of services/ resources (Lambda)
Pay per request-time
It is a tool that enables you to visualize, understand and manage your AWS costs and usage over time (charts)
AWS Cost Explorer
It will allow the user to graphically visualize billing and usage over time, particularly monthly running costs
It identifies areas that need further inquiry
You can create forecasts to predict future AWS costs and expense reservations (you can set alerts when your usage exceeds)
AWS Budget
It will give the user the status of user-set budgets and provide forecasts of estimated costs
It includes whitepapers, documentation and support communities (Free). Also, limited AWS Trusted Advisor checks
You can use the AWS Personal Health Dashboard to view AWS Trusted Advisor provides alerts and remediation guidance when AWS is experiencing events that may affect you
Basic
It includes tracking access to AWS resources for code development and building block architecture support
Developer
It includes applications architecture guidance, a dedicated technical engagement with AWS Support and TAM
Business
24x7 access to Cloud Support
24x7 access to Cloud Support
Enterprise
AWS Support Plans
It provides guidance, architecture review and ongoing communication with your company as you plan, deploy and optimize your application
TAM
Technical Account Manager
Development guidance
General Guidance (24hs)
Request a feature
Non-critical functions of your app are behaving abnormally
Time-sensitive development
Support Enhanced (12hs)
Important functions of your app are impaired or degraded (Business and Enterprise)
Production System Impaired (4hs)
Your business is impacted (Business and Enterprise)
Production System Down (1hr)
Business-Critical System Down (15min)
Support Cases
PRICING
It includes software listings from independent software vendors
AWS Marketplace
You receive a single bill for AWS accounts in your organization
Consolidated Billing
It can for estimating the cost of your use cases on AWS
AWS Pricing Calculator
It helps to you for doing cost-benefit analysis of migration on AWS
The TCO calculator will make suggestions and recommendations on appropriate resources types based on the user's input values and settings
AWS TCO
Business
It aligns with business needs and that IT investments have business value (Business Business and Financial managers, budget owners and strategy stakeholders)
Governance
It supports development of an organization-wide change management strategy for successful cloud adoption (ie. staffing and people managing)
People
It focuses on the skills and processes to align IT strategy with business strategy (governance, change management, portfolio management, business analytics and portfolio management)
Platform
It includes principles and patterns for implementing new solutions on the cloud and migrating on-premises workloads to the cloud (CTO, IT managers and solutions architects)
Operations
It ensures that the organization meets security objectives for visibility, auditing, control and agility (CISO, IT security managers and IT security analysts)
Security
It helps you to enable, run, operate and recover IT workloads to the level agreed upon with your customers/ stakeholders (IT operations and support managers)
AWS Cloud Adoption Framework (CAF)
Lift and shift (moving applications without changes)
Rehosting
Lift, tinker and shift (making a few cloud optimizations)
Replatforming
Add features, scale or performance
Refactoring (re-architecting)
Traditional license to a SaaS model
Repurchasing
Stop investing in legacy
Retaining
Retire
Retiring
Removing apps
It is a collection of infrastructure and data resources that a company or organization need to make the transition to the cloud
MIGRATION
Migration Strategies
It is a virtual, rugged and secure edge computing and data transfer device (2 CPUs, 4 GB of memory and 8 TB of usable storage)
AWS Snowcone
Storage: 80 TB (HDD) and 1 TB (SSD)
Snowball Edge Storage Optimized
S3 Snow Family
Compute: 40 vCPUs and 80 GB (equivalent to C5)
Storage: 80 TB (HDD) and 7 TB (SSD)
AWS Snowball
Snowball Edge Compute Optimized
Compute: 52 vCPUs and 208 GB (equivalent to C5, M5a, G3 and P3)
Snowball
It can be used to run petabytes of data per Snowballer, a 45-foot long ruggedized shipping container, pulled by a semi-trailer truck
AWS Snowmobile
It doesn't require you to provision, maintain or administer servers (AWS Lambda)
Serverless Applications
Build conversational interfaces using voice and text
Amazon Lex
Convert speech to text
Amazon Transcribe
Artificial Intelligence
Discover patterns in text
Amazon Comprehend
Identify potentially fraudulent online activities
Amazon Fraud Detector
Innovation
It is a machine learning-based tool that analyzes metrics and historical utilization and makes recommendations of compute resources to AWS customers
AWS Compute Optimizer
Machine Learning
It is a service that makes it easier for developers and ops teams to build, train and deploy ML models quickly
Amazon SageMaker
It helps you to understand how to manage and operate secure, efficient and cost-effective systems in the AWS Cloud
ARCHITECTURE
Monitoring systems that will deliver continuous business operations are a crude, annotating document, referring operations procedures for troubleshooting, anticipating failures and making root cause analysis (RCA)
Operational Excellence
Protecting information, systems and assets (automate security, enable traceability, keep people away from data, prepare for security events, apply security at all layers and protect data in transit and at rest)
Security
Recover disruptions, dynamically acquire computing resources to meet demand and mitigate misconfigurations and network issues (focusing on the ability of a workload to perform its intended function correctly and change configuration to meet business demands)
Reliability
AWS Well-Architected Framework
Focusing on using computing resources efficiently to meet system requirements and to maintain that efficiency as demand changes and technologies evolve
Performance Efficiency
Focusing on the ability to run systems to deliver business value at the lowest price point
Cost Optimization
Updraft expenses (ds, servers, lic-ntl)
Trade-upfront expense for variable expense
Lower variable cost
Benefit from massive economies of scale
You don't predict how much infrastructure capacity will need
Stop guessing capacity
Flexibility (develop and deploy apps faster)
Increase speed and agility
Stopping spend money on managing dc
Stop spending money running and maintaining data centers
You can quickly deploy apps to customers around the world
Go global in minutes
They create an identifying mechanism for the user (e.g. classify and order all their provisioned resources appropriately)
Tagging
It enables the ordering of AWS resources into logical groups (app, environment or software components)
Resource Group
It is the AWS fully managed solution for collecting, processing and analyzing streaming data in the cloud (real-time)
Kinesis
It is an ingestion service that provides data streams to Kinesis
Data Streams
It is appropriate for loading data streams
Firehose Delivery Streams
Types
It is a solution to process and query real-time streaming data using SQL-based solutions
Data Analytics
Video Analytics
It records API calls for your account (identity, time, source IP)
CloudTrail
It is a service that enables governance, compliance, operational auditing and risk auditing of your AWS account
Logs
Events store for 90 days
Source IP
AssetTrace
User, UserAgent
Region, Resource, Action
It can be encrypted using KMS (Key Management Service)
It detects unusual API activities or user AWS account behavior (a feature of CloudWatch)
It is an optional service that needs to be set up from the dashboard providing a list of recipients and selecting resources items for which notice is required
AWS Trusted Advisor is an online tool that provides you real time guidance to help you improve your resources following AWS best practices
Trusted Advisor checks helps optimize your AWS infrastructure, improve security and performance, reduce your overall costs, and monitor service limits
Green: No problems
Orange: Investigations
Red: Action
Categories
Cost optimization
Performance
Security
Fault tolerance
Service limits
Pillars
It is a fully managed service that allows for insightful business intelligence reporting with creative data delivery methods, including graphical interactive dashboards
QuickSight
It is a web service that enables you to monitor, manage metrics and configure alarms actions (monitor resources, performance and utilization)
CloudWatch
It doesn't log access API calls from one resource to another (AWS CloudTrail can do this)
You can install and configure the CloudWatch Agent on one EC2 instances and then write the CloudWatch Agent configuration to the parameter store of AWS System Manager and then pull down the CloudWatch configuration from SSM onto other EC2 instances
Log data from AWS Services (eg. CPU Utilization)
Logs
It variable to monitor (eg. CPU Utilization over time)
Metrics
Trigger an event based on a condition (eg. Ever Hour snapshot of servers)
Events
Triggers notification based on metrics when a defined threshold is breached
Alarms
Create visualizations based on metrics
Dashboards
Types
CloudTrail
User specific view on the availability and performance of AWS services, underlying their AWS resources
AWS Personal Health Dashboard
A service that provides the user with alerts and notifications on AWS scheduled activities, pending issues and planned changes
It publishes most up-to-the-minute information on AWS service availability
AWS Service Health Dashboard
It helps you analyze and process vast amount of data distributing the computing power across cluster of virtual servers running in the AWS Cloud.
AWS EMR
The cluster is managed using open-source framework called Hadoop
ANALYTICS
NETWORKING
ACL
ACL is virtually firewall that controls inbound and outbound traffic at the subnet level (stateless packet filtering)
By default your network ACL allows all inbound and outbound traffic
For custom network ACLs, all inbound and outbound traffic is denied
ACL rules have an explicit deny rule
Security Group
Security Group is a virtual firewall that controls inbound and outbound traffic for an Amazon EC2 instance (stateful packet inspection)
By default, a security group denies all inbound and allows all outbound traffic
Distributes incoming application traffic across multiple resources (ECS)
Types
ALB (listeners, rules, target groups to route traffic) - eg. web apps
NLB (listeners and target groups to route traffic) - eg. video games
CLB (listeners to route traffic) - legacy
Use X-Forwarded-For (XFF) to get original IP of incoming traffic passing through ELB
You can attach Amazon Certification Manager SSL to any of the ELB types
ELB
It can route users to infrastructure outside AWS
Traffic
Manage the DNS records for domain names
You can also transfer DNS records for existing domain names managed by other domain registrars
Simple: random selection
Weighted: percentages (weights assigned)
Latency-Based: based on region
Failover: change on health check
Geolocation: geographic location
Geo-Proximity: bias values (needs traffic flow)
Multi-Value Answer: multiple values
DNS resolvers around the world can only reflect the changes in their cache after TTL has expired, it is 24 hours by default
DNS
Amazon Route 53
Routing
Public
Private
Subnet
Internet Gateway
Virtual Private Gateway
AWS Direct Connect
On-Premises to AWS S3. Public VIF with BGP Routing Protocol
You can share traffic between two VPCs in different AWS Regions and under separate AWS Accounts
It can be used to replicate data to geographically distinct locations for fault tolerance, disaster recovery and redundancy
VPC Peering
Every VPN connection have a VPN connection is provisioned and available
Full Rout is billed for every partial VPN connection-hour
AWS Site-to-Site
Standard AWS data transfer charges are incurred upon transferring data using a VPN connection
Customer are not charged upon termination of the VPN connection
VPC
VPC is a networking service that you can use to establish boundaries around your AWS resources (isolated section of AWS Cloud)
General Purpose
Balanced (A1, T3, T3a, T2, M5, M5a, M4)
Compute Optimized
High-performance batch processing (C5, C5n, C4)
Memory Optimized
High-performance real-time processing (R5, R5a, X1e, X1, High Memory, z1d)
Accelerated Computing
Hardware accelerators: calculations, graphics and data pattern (P3, P2, G3, F1)
Storage Optimized
Input/output operations per second (D3, I3en, D2, I3)
Types
Short-term, unpredictable usage patterns, less than year
On-Demand
You are charged per second based on the hourly rate
You pay for as much as you use
You pay no upfront costs for instances
1 to 3 years, more flexibility than RI, can change the region
They aren't available for all the regions
Savings Plans
EC2
Fargate
Lambda
Predictable: 1 to 3 years, can't change the region, don't require you to commit to a consistent amount of compute usage over the duration of the contract (predictable usage)
RIs are not guaranteed to have capacity
A portion of the cost must be paid up front and the remaining hours in the term are billed at a discounted hourly rate regardless of usage
No Upfront Reserved Instances are based on a consumption obligation to pay monthly for the entire term of the reservation (monthly billing)
Full payment is made at the start of the term, with no other costs or additional hourly charges incurred for the remainder of the term, regardless of hours used.
Partial Upfront
No Upfront
All Upfront
Types
Reserved Instances
Spot Instances
Not time critical: workloads with flexible, interruptions (example AWS retrospectives)
Dedicated Hosts
Licencing, compliance
Multi-Tenant: Virtual isolation (think apartment)
Single-Tenant: Physical isolation (think house)
Pricing
EC2
Scaling
You can use AMI to spin up EC2 instances on AWS Cloud
AMI
You can launch multiple instances from a single AMI when you need multiple instances with the same configuration
An Amazon Machine Image (AMI) provides the information required to launch an instance.
COMPUTE
Host short running functions
Service-oriented applications
Event driven applications
No provisioning or managing servers
AWS Lambda
Model: Function As A Service (FaaS)
During the time there will be a delay when the function will initially run which is called a Cold Start
If you need long running tasks (more than 15 mins) and a custom OS environment then consider using Fargate
Serverless
AWS Fargate
Manages your server infrastructure (ECS and EKS)
It is compatible with data formats such as CSV, JSON, ORC, Apache Parquet
AWS Athena
Amazon Athena queries data directly from S3 and there aren't additional data storage commitments beyond the object storage
It is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL
It is serverless, so there is no manage and you pay only for the queries that you run
Implementation
Platform
SaaS
Cloud-Based
On-Premises (private)
Hybrid
Scalability
Global
Agility
Capacity
Cost
Scale compartmental applications
Models
Cloud
Benefits
Containers
ECS
Launch and stop docker-enabled applications (API calls)
EKS
Deploy and manage containerized applications at scale
Publishers: SDK, CLI, CloudWatch, AWS Services
Subscribers: Lambda, SQS, Email, HTTPs
SNS is a highly available, durable, secure, fully managed pub/sub messaging service that enables the decoupling of microservices, distributed systems and serverless applications (application integration)
Simple Notification (SNS)
DNS Topic
A logical access point and communication channel
You can encrypt Topics via KMS
It sends, stores and receives messages
Asynchronous communication (not real-time)
AWS SQS is a fully managed service that decouples the messages generated by your application (connects isolated apps)
Simple Queue Service (SQS)
Message retention by default is 4 days, can be upto 14 days
Types
Standard: It doesn't guarantee order of delivery
Messaging
Architecture
Microservices
Monolithic Apps: components communicate with each other in internal data
Monolithic Apps
Make API calls
Management Console
It is a web-interface for accessing and managing AWS services (easy environments, view AWS bills, view monitoring)
SDK
Interact with AWS resources through various programming languages (C++, Ruby, Python, Java)
CLI
Build an environment by writing lines of code instead of using the AWS management consoles
It reduces the time taken for environment build and eventual usage (CloudFormation Templates)
AWS Quick Starts
It used for deploying complex web and mobile apps development environment instantaneously, complete with the necessary resources and peripheral assets.
AWS CloudFormation
Infrastructure As Code (IaC)
The process of managing and provisioning computer DC through machine-readable definition files (YAML, JSON) rather than physical hardware configuration
Change Sets
It can be used to preview changes to AWS resources when a stack is executed
Templates
DrR Detection
It used to detect any changes made to resources outside of CloudFormation stacks
Intrinsic Functions
IFs are used for assigning values to properties in CloudFormation templates
It used for AWS Service Catalog
Choose a platform, upload your code and it runs with little worry for developers about infrastructure knowledge (not infrastructure as code, no templates)
Elastic Beanstalk
It is a solution for creating secure APIs in your cloud environment at any scale (traffic management, authorization and monitoring)
Resources (eg. /projects)
Methods (eg. PUT, POST, GET)
Configurations
Stages (eg. pre, prod, staging)
Invoke URL (for each stage)
Deploy API (everytime you make a change)
API Gateway
It can be enabled to cache your endpoints response to API calls (improves latency and reduces the numbers of calls)
API Caching
Allows restricted resources (e.g. Forms) on a website to be requested from different domain than the initial resource that it came from
API CORS
It is a solution for a distributed system in which the systems communicate with each other by exchanging messages (don't negatively impact other components of the system)
Loose Coupling
This makes it possible to configure a VPC endpoint which the user can attach to the resource that he/she need to the internet access
AWS PrivateLink
It provides private connectivity between VPCs and services hosted on AWS or on-premises, securely on the Amazon network
By providing a private endpoint to access your services, AWS PrivateLink ensures your traffic is not exposed to the public internet
Resources
It helps developers to analyze and debug production, distributed apps such as those built using a microservices architecture
AWS X-RAY
Developers can identify and troubleshoot the root cause of performance issues and errors (e.g. AWS Lambda Apps)
It is a deployment service that allows developers to automate software deployments (Lambda, EC2, On-Premises, Amazon ECS services, Lambda functions or even on-premises servers (Deployment as a Service)
AWS CodePipeline
It is used for deployment of codes to resources
Servers should be running Linux OS with connectivity to AWS public endpoints
AWS CodeDeploy
It is a fully managed configuration automation and manage service that provides managed instances of Chef and Puppet (platforms that allow you to use code to automate the configuration of servers)
AWS OpsWorks
It is a fully managed continuous integration service that compiles source code, runs tests, and produces software packages that are ready to deploy
AWS CodeCommit
It can be used as a data store to source code, binaries, script, HTML pages and images which are accessible over the internet
AWS CodeBuild
It is managed service that automates software deployment on a large to EC2 instances and on-premises servers (delivery)
AWS CodeDeploy
It is fully managed continuous integration service that compiles source code, runs tests, and produces software packages that are ready to deploy
AWS CodeBuild
It enables you to develop, build and deploy applications on AWS
AWS CodeStar
It provides a unified user interface, enabling you to manage your software development activities in one place easily
AWS CodeCommit
It is a fully managed artifact repository service that make it easy for organizations to any size to securely store, publish and share software packages used in their software development process
AWS CodeArtifact
It is a developer tool powered by machine learning that provides automated code reviews for static code, finding and identifying an application's most expensive lines of code
AWS CodeGuru
It is an enterprise, easy-to-use, novice-friendly and interactive platform to configure and launch web apps in minutes quickly
AWS Lightsail
It is best to utilize for simple workloads and has fast start-to-end deployment
Upgrade to EC2* is the feature that allows customers to create copy of their Lightsail instance to EC2
AWS Elastic Beanstalk is an AWS supported open source cluster management tool that helps you run Docker workloads on AWS
AWS OpsWorks
Configuration management tool (AWS Clusters is the AWS Chef)
It enables you to quickly build an AWS compute environment in minutes
AWS ParallelCluster
It allows user to control their AWS resources by unifying services into a user interface
Systems Manager
It can be to view, automate and monitor operational tasks
It is a CDN (Content Delivery Network) Service. Amazon CloudFront uses a network of edge locations to cache content and deliver content to customers all over the world
AWS CloudFront
site that CloudFront uses to store cached copies of your content before it's ready to deliver to viewers
Edge Locations
AWS CF caches dynamic content (API, dynamic site delivery)
AWS Objects
Distribution consists of a collection of edge locations
Distribute content to users
Features
Cache control responses
Used in conjunction with CloudFront service
A geographical area that contains AWS resources (two or more AZs)
Regions
data governance and legal requirements
Proximity
selecting a region close to your customers (latency)
Factors
some regions don't have special features/ resources (Amazon Bedrock - Quantum Computing Platform)
Availability
Pricing
AZ is one or more data centers within a region (you should run across at least two AZ in a region)
Availability Zones
First level of redundancy is always AZ when it comes to HA and performance
It is suitable for apps that are non-HTTP/S such as VoIP, MTSO and gaming (Amazon Elastic Compute Cloud) performance of HTTP-based content such as dynamic web pages, APIs, and streaming
Global
It doesn't include the content caching capability that Amazon CloudFront does
It improves the performance of a wide range of apps over TCP or UDP by proxying packets at Edge Locations to applications running in one or more AWS regions
Resources
Global Accelerator
It provides static IP acting as a fixed entry point to application endpoints in multiple Amazon's EC2, etc) in a single or multiple AZs/Regions
It improves the resulted traffic performance by as much 60% and low latency for great user experience (routing traffic closest, gaming, media and mobile apps)
INFRASTRUCTURE
Physical security data centers
Hardware and software infrastructure
Network infrastructure
Virtualization infrastructure
Selecting an EC2
Configuring an EC2
Patching OS
Security Groups on EC2
Accounts on EC2
Customer Responsibilities
AWS Responsibilities
AWS Shared Responsibility
AWS Root User
AWS account
It represents a person or application that interacts with AWS services and resources (by default it hasn't permissions associated with it)
Users
It is a document that allows or denies permissions to AWS services and resources (least privilege when granting permissions)
Policies
It is a collection of IAM Users
Groups
It is an identity that you can assume to gain temporary access permissions
Roles
Its permissions
MFA
It proves extra layer of security for your AWS account (text user)
Root access keys are long-term credentials and provisioning keys (same as AWS Secret Access Keys)
AWS Access Keys
Administrators can apply permissions and constraints for users and user groups to enable access to resources in self-service manner
AWS Service Catalog
It simplifies access management and manages access to multiple AWS accounts and business apps centrally
AWS SSO
It is the parent company of the accounts in your organization
Service Control Policies enable you to place restrictions on AWS Services, resources and individual API actions for all IAM users in each account you are linked (included in OU)
AWS Organizations
SCP
You can group accounts into organizational unit to make it easier to manage accounts with similar business or security requirements.
Dos attack is a deliberate attempt to make a website or application unavailable to users
DoS
Dos
Multiple sources are used to start an attack that aims to make a website or application unavailable.
AWS Shield
Standard
It automatically protects all AWS customers at no cost.
Advanced
It is a paid service that provides detailed attack diagnostics and the ability to detect and mitigate sophisticated DDoS attacks
AWS enables you to perform encryption operations through the Amazon Web Services Encryption Cli (make, manage and cryptographic keys)
AWS Key Management Service
AWS KMS uses envelope encryption while integrating with other AWS services
SSE-KMS does encryption/decryption at rest and does offer monitoring capabilities (integrated with AWS CloudTrail)
AWS WAF
It is a web application firewall that lets you monitor network requests that come into your web applications (it works together with CloudFront and Load Balancer)
It provides continuous security assessments. It checks application for security vulnerabilities and deviations from security best practices
Amazon Inspector
Types
Network accessibility: It checks network accessibility on EC2
Host assessment: It checks vulnerability on EC2
It is a service that provides intelligent threat detection for your AWS infrastructure and resources. It identifies threats by continuously monitoring the network activity and account behavior with your AWS environment (sources, VPC and DNS)
Amazon GuardDuty
It provides data security and privacy using machine learning algorithms, detects, discovers, monitors and reports sensitive data in AWS (PII, personal identity)
AWS Macie
It discovers, classifies and protects current data
Replacies the need to hard-code authentication credentials in code
AWS Secret Manager
Makes it possible to include API call in code that retrieves authentication information from a central repository
AWS Resources Access Manager
It is a persistent machine learning-driven service that automatically collates log data from all AWS resources
AWS Detective
The information collected by AWS Detective allows you to quickly and efficiently conduct security investigations
Security Services
AWS Cognito
It provides authentication, authorization and user profile management for your web and mobile apps
It helps to provide user access control for mobile and web apps
It is possible for users to log into AWS environment using their credentials from a corporate directory
Identity Federation
An administrator can configure identity federation so that users can access the AWS environment using their directory credentials
SSO third party
SAML: It transports identity information between idP and apps
AWS STS
Apps deployed on EC2 instance can be provided security credentials using AWS STS, allowing short-term security credentials
AWS Security Hub
It gives administrator to have full and exclusive control over the generation and management of cryptographic keys on actual hardware security modules that are physically stored in AWS data centers
AWS Security Hub provides a view of across account security status and gives alerts
It identifies deviations from the standards, best practices and suggests recommended solutions steps
It is a service that provides on-demand access to AWS security and compliance reports and select online agreements
Compliance
AWS Artifact
Agreements
You can review, accept and manage agreements for an individual account and for all your accounts in AWS Organizations
Reports
It provides compliance reports from third-party auditors (regulations)
Customer Compliance Center
It contains resources to help you learn more about AWS compliance
Multi Site
Warm Standby
Pilot Light
Backup and Restore
Disaster Recovery
Compliance
SECURITY