Reference:

FAQs

Category:

Analytics

Amazon Athena

What?

- Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL.
- Amazon Athena uses Presto with ANSI SQL support and works with a variety of standard data formats, including CSV, JSON, ORC, Avro, and Parquet.

Why?

- Athena is serverless, you don't need to worry about configuration, software updates, failures or scaling your infrastructure as your datasets and number of users grow. With Athena Federated Query, you can run SQL queries across data stored in relational, non-relational, object, and custom data sources.

When?

- You want to tap into your data without setting up complex processes to extract, transform, and load the data (ETL).
- You want to process logs, perform data analytics, and run interactive queries.

Where?

- Amazon Athena is a regional service, but it can access data in other regions or other AWS accounts.
- Athena works directly with data stored in S3.
- It also natively supports the AWS Glue Data Catalog.

Who?

- Athena is serverless service, so you don't have to setup or manage any infrastructure.
- Amazon Athena allows you to control access to your data by using AWS IAM policies, Access Control Lists (ACLs), and Amazon S3 bucket policies.

How?

- To get started, Just point to your data in Amazon S3, define the schema, and start querying using the built-in query editor.
- Results are displayed in the console, and automatically written to a location of your choice in S3. You can also download them to your desktop.

How much?

- Amazon Athena is priced per query and charges based on the amount of data scanned by the query.
- It queries data directly from Amazon S3, so your source data is billed at S3 rates.

Created by:

Ashish Prajapati

Reference:

FAQs

Category:

Analytics

Amazon EMR

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

Service Summary Cards (SSC)

What?

- Amazon EMR is cloud big data platform for data processing, interactive analysis, and machine learning using open source frameworks such as Apache Spark, Apache Hive, and Presto.

Why?

- Amazon EMR simplifies building and operating big data environments and applications.
- Amazon EMR enables you to quickly and easily provision as much capacity as you need, and automatically or manually add and remove capacity. This is very useful if you have variable or unpredictable processing requirements.

When?

- When you want to focus on transforming and analyzing your data without having to worry about infrastructure provisioning, cluster setup, configuration, open-source applications or tuning.

Where?

- Amazon EMR launches all nodes (Master, Core, Optional Task nodes) for a given cluster in the same Availability Zone of a Region.

Who?

- After an EMR cluster is launched, customer can monitor and manage it. Amazon EMR provides several tools you can use to connect to and control your cluster.

How?

- You can launch a cluster by specifying the name of your cluster, the location in Amazon S3 of your input data, your processing application, your desired data output location, and the number and type of Amazon EC2 instances you'd like to use.

How much?

- You pay a per-second rate for every second you use, with a one-minute minimum. Customers pay for Amazon EMR plus backend compute price (Amazon EC2, Amazon EKS, AWS Outposts, Amazon EMR Serverless).

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Analytics

AWS Glue

What?

- AWS Glue is a serverless data integration service that makes it easy to discover, prepare, and combine data for analytics, machine learning, and application development.

Why?

- AWS Glue provides all the capabilities needed for data integration, so you can start analyzing your data and putting it to use in minutes instead of months.

When?

- You should use AWS Glue to discover properties of the data you own, transform it, and prepare it for analytics.
- When you need unified catalog to find data across multiple data stores, explore data with self-service visual data preparation or create, run, and monitor ETL jobs without coding.

Where?

- AWS Glue is a Regional service.

Who?

- AWS Glue provides a managed ETL service that runs on a serverless Apache Spark environment. This allows you to focus on your ETL job and not worry about configuring and managing the underlying compute resources.

How?

- You define jobs in AWS Glue to accomplish the work that's required to extract, transform, and load (ETL) data from a data source to a data target.
- AWS Glue can generate a script to transform your data. Or, you can provide the script in the AWS Glue console or API.

How much?

- With AWS Glue, you pay an hourly rate, billed by the second, for crawlers and ETL jobs. For the AWS Glue Data Catalog, you pay a monthly fee for storing and accessing the metadata. For development endpoint, you pay an hourly rate, billed per second.
- For AWS Glue DataBrew, the interactive sessions are billed per session and the DataBrew jobs are billed per minute.

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Analytics

Amazon QuickSight

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

What?

- Amazon QuickSight is cloud-powered business analytics service that makes it easy to build visualizations, perform ad-hoc analysis, and quickly get business insights from data.

Why?

- QuickSight enables organizations to scale their business analytics capabilities to hundreds of thousands of users, and delivers fast and responsive query performance by using a robust in-memory engine (SPICE).

When?

- When you don’t want to use traditional BI solutions that often require teams of data engineers to spend months building complex data models before generating a report.

Where?

- Amazon QuickSight is a regional service but you can also explicitly connect to other AWS data sources that are not in your account or in a different region by providing connection details for those sources.

Who?

- You can seamlessly grow your data from a few hundred megabytes to many terabytes of data without managing any infrastructure.

How?

- Amazon QuickSight discovers your data sources in AWS services such as Amazon Redshift, Amazon RDS, Amazon Athena, and Amazon Simple Storage Service (Amazon S3). You can connect to any of the data sources discovered by Amazon QuickSight and get insights from this data in minutes.

How much?

- Month-to-month pricing for Authors and Readers.
- Alerts are priced based on metrics evaluated.
- Accounts enabled with Q are charged a $250/month Q base fee plus Q questions capacity pricing.

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Application Integration

Amazon EventBridge

Complete book:

Click Here

More SSCs:

Click Here

Created by:

Ashish Prajapati

What?

- Amazon EventBridge is a serverless event bus service to connect your applications with data from a variety of sources.
- EventBridge expands on the capabilities of CloudWatch Events, adding support for processing events from SaaS partner applications and making it easier for you to process events from your own applications.

Why?

- EventBridge enables you to build event-driven architectures that are loosely coupled and distributed.
- EventBridge ingests data from supported SaaS applications without writing custom integration code.
- To reduce operational overhead as there are no servers to provision, patch, and manage and no software to install.

When?

- You want to extend the functionality of your applications by easily connecting them to other SaaS applications.
- You want to monitor and audit your AWS environments and respond to operational changes in your applications in real-time to prevent infrastructure vulnerabilities.

Where?

- Amazon EventBridge is a regional service.
- It also supports Global endpoints to fail over event ingestion automatically to a secondary Region during service disruptions.

Who?

- Your account has a default event bus which receives events from AWS services.
- You can create custom event buses to send or receive events from a different account or Region.

How?

- EventBridge receives an event (an indicator of a change in environment) and applies a rule to route the event to a target.
- Rules match events to targets based on the structure of the event (called an event pattern), or on a schedule.
- All events that come to EventBridge are associated with an event bus and rules are tied to a single event bus.

How much?

- Charges for per million events published which includes custom events, third-party events and events to another bus.
- Customers may incur additional data transfer charges between regions for cross-Region invocations. There is no additional charge for global endpoints.

Reference:

FAQs

Category:

Application Integration

Amazon Simple Notification Service (Amazon SNS)

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

Service Summary Cards (SSC)

What?

- Amazon Simple Notification Service (Amazon SNS) is a fully managed push-based messaging service for both application-to-application (A2A) and application-to-person (A2P) communication.
- It provides message delivery from publishers to subscribers (also known as *producers* and *consumers*).

Why?

- Amazon SNS is a highly available, durable, secure, fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and event-driven serverless applications.
- Amazon SNS provides topics for high-throughput, push-based, many-to-many messaging.

When?

- You can use Amazon SNS to support a wide variety of needs including event notification, monitoring applications, workflow systems, time-sensitive information updates, mobile applications, and any other application that generates or consumes notifications.

Where?

- Amazon SNS is a regional service.

Who?

- With simple APIs requiring minimal up-front development effort, no maintenance or management, Amazon SNS gives developers an easy mechanism to incorporate a powerful notification system with their applications.

How?

- Developers must first create a “topic” which is an “access point” – identifying a specific subject or event type – for publishing messages and allowing clients to subscribe for notifications. Topic owner can set policies for it such as limiting who can publish messages or subscribe, or specifying which protocols will be supported (i.e. HTTP/HTTPS, email, SMS).

How much?

- Standard topic - number of monthly API requests made, and the number of deliveries to various endpoints.
- FIFO topic - pricing is based on the number of published messages, the number of subscribed messages, and their respective amount of payload data.

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Application Integration

Amazon Simple Queue Service (Amazon SQS)

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

What?

- Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications.

Why?

- SQS eliminates the complexity and overhead associated with managing and operating message-oriented middleware, and empowers developers to focus on differentiating work.
- SQS scales elastically with your application so you don't have to worry about capacity planning and pre-provisioning.

When?

- You want to send, store, and receive messages between software components at any volume, without losing messages or requiring other services to be available. Use SQS standard queues for maximum throughput, best-effort ordering, and at-least-once delivery. Use SQS FIFO queues to guarantee that messages are processed exactly once, in the exact order.

Where?

- Amazon SQS is a regional service. Amazon SQS stores all message queues and messages within a single, highly-available AWS region with multiple redundant Availability Zones (AZs), so that no single computer, network, or AZ failure can make messages inaccessible.

Who?

- AWS manages the backend for Amazon SQS service including scaling and durability.
- Customers can control who can send messages to a message queue and who can receive messages from a message queue. Amazon SQS has its own resource-based permissions system.

How?

- Messages are sent from producers (applications, microservices, and other AWS services) to SQS Queue.
- It stores messages and wait for consumer to poll. Consumer applications (Lambda Functions, EC2 Instances and other AWS services) pull/poll the messages and process it.

How much?

- The cost of Amazon SQS is calculated per request, plus data transfer charges for data transferred out of Amazon SQS (unless data is transferred to Amazon EC2 instances or to AWS Lambda functions within the same region).
- Each 64 KB chunk of a payload is billed as 1 request (for example, an API action with a 256 KB payload is billed as 4 requests).

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Application Integration

AWS Step Function

What?

- AWS Step Functions is a serverless function orchestrator that makes it easy to sequence AWS Lambda functions and multiple AWS services into business-critical applications.
- AWS Step Functions state machines (workflows) are defined in JSON using the declarative Amazon States Language.

Why?

- Through its visual interface, you can create and run a series of checkpointed, event-driven workflows and build distributed applications. Workflows manage failures, retries, parallelization, service integrations, and observability so developers can focus on higher-value business logic.

When?

- You want to create end-to-end workflows to manage jobs with interdependencies for common use cases like, Data processing, DevOps and IT automation, E-commerce, Web applications.
- You want to focus on application tasks rather than building complex state management into all of your tasks.

Where?

- AWS Step Functions has built-in fault tolerance and maintains service capacity across multiple Availability Zones in each region to protect applications against individual machine or data center failures.

More SSCs:

Click Here

Complete Book

Click Here

Who?

- Step Functions manages operations and underlying infrastructure to ensure your application is available at any scale.
- You can run tasks on AWS, your own servers, or any system that has access to AWS

How?

- Using AWS Step Functions, you define state machines that describe your workflow as a series of steps, their relationships, and their inputs and outputs. States can perform work, make choices, pass parameters, initiate parallel execution, manage timeouts, or terminate your workflow with a success or failure.

Created by:

Ashish Prajapati

How much?

- Step Functions counts a state transition each time a step of your workflow is executed. You are charged for the total number of state transitions across all your state machines, including retries.
- You may incur additional charges if your application workflow utilizes other AWS services or transfers data.

Reference:

FAQs

Amazon EC2

What?

- Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable computing capacity—literally, servers in Amazon's data centers—that you use to build and host your software systems.
- An instance is a virtual server in the AWS Cloud.

Why?

- Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change.
- Amazon EC2 changes the economics of computing by allowing you to pay only for capacity that you actually use.

When?

- Run cloud-native and enterprise applications, Scale for HPC applications, Develop for Apple platforms.
- You want complete control of your computing resources and run it on Amazon's proven computing environment.
- You want to import your virtual machine images to Amazon EC2.

Where?

- Amazon EC2 is a regional service. An EC2 Instance runs in an Availability Zone.
- By launching instances in separate Availability Zones, you can protect your applications from failure of a single location.

Who?

- Customer must take care of OS patches, high availability and scaling. AWS provides tools and services for Patch Management, Auto Scaling, Monitoring, Backup and Vulnerability Scanning.

How?

- Select an AMI (Amazon Machine Image) >> Select Instance Type >> Select Additional Settings (Disk Size / Security Group Setting / Network / Start-up Script etc.).

Category:

Compute

How much?

- On-Demand Instance – Pay per second / hour.
- Reserved Instances / Savings Plan – Commitment for 1-year / 3-year term.
- Spot Instances – Supply and Demand based pricing. May be terminated by AWS after giving a 2 min notice.

Reference:

FAQs

Category:

Compute

Amazon EC2 Auto Scaling

What?

- Amazon EC2 Auto Scaling helps you ensure that you have the correct number of Amazon EC2 instances available to handle the load for your application. You create collections of EC2 instances, called Auto Scaling groups (ASG) by specifying minimum, maximum and desired number of instances.

Why?

- You can use the fleet management features of EC2 Auto Scaling to maintain the health and availability of your fleet. An application running on Amazon EC2 instances, is referred as a ‘fleet’.
- It helps you in better fault tolerance, Better availability Better cost management for your application.

When?

- You want to automatically launch or terminate EC2 instances based on user-defined scaling policies, scheduled actions, and health checks. It supports manual scaling, scheduled scaling, dynamic scaling and predictive scaling.
- You want to reduce the need to manually provision Amazon EC2 capacity in advance.

Where?

- EC2 Auto Scaling groups are regional constructs. They can span Availability Zones, but not AWS regions.
- It lets you provision and automatically scale instances across purchase options, AZs, and instance families.

Who?

- It scales dynamically based on your Amazon CloudWatch metrics, or predictably according to a schedule that you define.
- EC2 Auto Scaling monitors the health of running instances, automatically replaces impaired instances, and balances capacity across Availability Zones.

Complete book:

Click Here

How?

- When you create an EC2 Auto Scaling group, you must specify a launch configuration. Using this launch configuration Amazon EC2 Auto Scaling always launches new instances such that they are balanced between availability zones.
- Predictive Scaling’s machine learning algorithms detect changes in daily and weekly patterns, and adjusti their forecasts.

Created by:

Ashish Prajapati

How much?

- Amazon EC2 Auto Scaling features have no additional fees beyond the service fees for Amazon EC2, CloudWatch (for scaling policies), and the other AWS resources that you use.

Reference:

FAQs
Comparison

Category:

Networking and Content Delivery

Elastic Load Balancing

What?

- Elastic Load Balancing automatically distributes your incoming traffic across multiple targets, such as EC2 instances, containers, and IP addresses, in one or more Availability Zones.
- Four types of load balancers – Application, Network, Gateway, and Classic Load Balancers.

Why?

- It monitors the health of its registered targets, and routes traffic only to the healthy targets. Elastic Load Balancing scales your load balancer as your incoming traffic changes over time. It can automatically scale to the vast majority of workloads.
- Using a load balancer increases the availability and fault tolerance of your applications.

When?

- Application Load Balancer (ALB) to load balance HTTP requests, Network Load Balancer (NLB) for network/transport protocols (layer4 – TCP, UDP) load balancing, use Gateway Load Balancer if you need to use third-party virtual appliances.
- Classic Load Balancer for application built within Amazon EC2 Classic network.

Where?

- ELB is a regional service. When you enable an Availability Zone for your load balancer, Elastic Load Balancing creates a load balancer node in the Availability Zone.
- AWS recommends enabling multiple Availability Zones for all load balancers.

Who?

- ELB is a managed service and supports high availability, automatic scaling, and robust security.
- With a Gateway Load Balancer, customer is responsible for choosing and qualifying software from appliance vendors.

How?

- You configure one or more listeners to accept incoming traffic and register targets in target groups.
- Target group for ALB - IP, Instance, Lambda. Target group for NLB - IP, Instance, Application Load Balancer.
- Target group for GLB - IP, Instance. Target group for Classic Load Balancers - Instances.

How much?

- You are charged for each hour or partial hour that a Load Balancer is running. Plus there are per hour charges for Load Balancer Capacity Units (LCU) for ALB, Network Load Balancer Capacity Units (NLCU) for NLB and Gateway Load Balancer Capacity Units (GLCU) for GLB. In case CLB you are charged for each hour or partial hour and for each GB of data transferred.

Created by:
Ashish Prajapati

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Serverless

AWS Lambda

What?

- AWS Lambda is a serverless, event-driven compute service that lets you run code for virtually any type of application or backend service without provisioning or managing servers.

Why?

- With Lambda, you can run code for virtually any type of application or backend service - all with zero administration.
- Just upload your code, and Lambda takes care of everything required to run and scale your code with high availability.
- You can set up your code to automatically trigger from other AWS services or call it directly from any web or mobile app.

When?

- Lambda is best suited for shorter, event-driven workloads, (Lambda functions run for maximum up to 15 minutes per invocation) such as process streaming data stored in Amazon Kinesis, or custom events generated by your applications.
- You want to build data-processing triggers for AWS services such as Amazon S3 and Amazon DynamoDB.

Where?

- Lambda runs your code on high availability compute infrastructure in a region.
- It maintains compute capacity across multiple AZs in each AWS Region to help protect your code against individual machine or data center facility failures.

Who?

- Lambda performs all the administration of your compute resources including server and operating system maintenance, capacity provisioning and automatic scaling, code and security patch deployment, and code monitoring and logging.
- All you need to do is supply the code.

How?

- The code you run on AWS Lambda is uploaded as a “Lambda function”. Each function has associated configuration information, such as its name, description, entry point, and resource requirements.
- Lambda will run your function by launching and managing the compute resources as needed based on incoming requests.

How much?

- Billing is metered in increments of one millisecond. You are charged based on the number of requests for your functions and the duration it takes for your code to execute.
- With Provisioned Concurrency, you pay for the amount of concurrency you configure and the duration that you configure it.

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Containers

Amazon Elastic Container Service (Amazon ECS)

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

What?

- Amazon ECS is a fully managed container orchestration service that makes it easy for you to deploy, manage, and scale containerized applications.

Why?

- Amazon ECS eliminates the need for you to install, operate, and scale your own cluster management infrastructure.
- You can use Amazon ECS to schedule container placement across your cluster based on your resource needs and availability requirements.

When?

- You want to schedule long-running applications, services, and batch processes using Docker containers.
- You want to maintains application availability and ability to scale your containers up or down to meet your application's capacity requirements.

Where?

- Amazon ECS is a regional service which can run and scale your container workloads across availability zones
- On your infrastructure with *Amazon ECS Anywhere*.

Who?

- Amazon ECS is a fully-managed container orchestration service, with AWS configuration and operational best practices built-in, and no control plane, nodes, or add-ons for you to manage.
- Customers define Tasks, Service, Capacity Providers and deployment process.

How?

- Amazon ECS allows you to easily run applications on a managed cluster of Amazon EC2 instances.
- Your containers are defined in a task definition that you use to run an individual tasks or task within a service.
- Amazon ECS is integrated with familiar features like Elastic Load Balancing, EBS volumes, Amazon VPC and IAM.

How much?

- There is no additional charge for Amazon ECS. You pay for AWS resources (e.g. Amazon EC2 instances or EBS volumes) you create to store and run your application.
- You only pay for what you use, as you use it; there are no minimum fees and no upfront commitments.

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Containers

Amazon ECS Anywhere

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

What?

- Amazon ECS Anywhere is a feature of Amazon ECS that enables you to run and manage container-based applications on-premises, including on your own virtual machines (VMs) and bare metal servers. Whether on-premises or in the cloud, you'll have similar cluster management, workload scheduling, and monitoring you've come to know from Amazon ECS.

Why?

- It provides a single management interface for all of your container-based applications, irrespective of the environment they're running in.

When?

- You don't want to run and operate separate container management software for your on-premises container workloads.
- You need a simple, consistent experience for cluster management, workload scheduling, and monitoring for both the cloud and on-premises.

Where?

- The ECS agent, software that allows a host to connect with the ECS control plane, is supported on VM (e.g., running on VMware, Microsoft Hyper-V, or OpenStack) or on-premises bare metal server running a supported operating system (OS).

Who?

- ECS Anywhere offers a completely managed solution that enables you to standardize container management across all of your environments. The AWS Systems Manager Agent, Amazon ECS container agent, and Docker must be installed on these external instances. Your external instances require an IAM role that permits them to communicate with AWS APIs.

How?

- Generate an activation key using ECS management console. Install the lightweight open source ECS agent, available on Github, Docker Hub, and Amazon Elastic Container Registry Public, on the VMs/bare metal servers.
- As part of the installation configuration, provide the activation key along with the AWS region.

How much?

- You pay $0.01025 per instance-hour for each managed ECS Anywhere external instance. You may be charged for registering your on-premises instances with AWS Systems Manager if you have over 1000 instances per account per region at any point of time.

Reference:

FAQs

Category:

Containers

Amazon Elastic Kubernetes Service (Amazon EKS)

What?

- Amazon Elastic Kubernetes Service (Amazon EKS) is a managed service that makes it easy for you to run Kubernetes on AWS without installing and operating your own Kubernetes control plane or worker nodes.
- Kubernetes is an open-source container orchestration system allowing you to deploy and manage containerized applications.

Why?

- Amazon EKS provisions and scales the Kubernetes control plane, including the API servers and backend persistence layer. It automatically detects and replaces unhealthy control plane nodes and patches the control plane.
- Amazon EKS is integrated with many AWS services to provide scalability and security for your applications.

When?

- You don't want the operational burden of managing the Kubernetes control plane.
- You want to maintain existing applications that run on upstream Kubernetes and want to use plugins and tooling from the Kubernetes community.

Where?

- Amazon EKS is a regional service.
- It runs and scales the Kubernetes control plane across multiple AWS Availability Zones to ensure high availability.

Who?

- Amazon EKS handles provisioning, scaling, and managing the Kubernetes control plane. It provides automated version upgrades and patching for control plane nodes.
- Customers provision an EKS cluster, deploy compute nodes, connect to EKS, and run Kubernetes applications.

How?

- You can get started by creating an Amazon EKS cluster. When your cluster is ready, you can configure your favorite Kubernetes tools, such as *kubectl*, to communicate with your cluster. Afterwards, you can deploy and manage workloads on Amazon EKS cluster the same way that you would with any other Kubernetes environment.

Created by:

Ashish Prajapati

How much?

- You pay a per hour charge for each Amazon EKS cluster you create and for the AWS resources you create to run your Kubernetes worker nodes.
- If you are using AWS Fargate, pricing is calculated based on the vCPU and memory resources used.

Reference:

FAQs

Category:

Database

Amazon Aurora

What?

- Amazon Aurora (Aurora) is a fully managed relational database engine that's compatible with MySQL and PostgreSQL.
- It is built for the cloud, and combines the performance and availability of traditional enterprise databases with the simplicity and cost-effectiveness of open source databases.

Why?

- Aurora automates time-consuming administration tasks like hardware provisioning, database setup, and backups while providing the security, availability, and reliability of commercial databases at 1/10th the cost.
- It features a distributed, fault-tolerant, and self-healing storage system that is decoupled from compute resources.

When?

- You want to migrate your MySQL and PostgreSQL workload to a managed relational databases service in AWS cloud.
- You need a high-performance distributed storage subsystem for your workloads that grows automatically as needed.

Where?

- Amazon Aurora is a regional service; it automatically maintains six copies of your data across three AZs.
- Amazon Aurora Global Database feature supports cross-region replicas. You can create up to five secondary regions for an Aurora Global Database.

Who?

- Amazon Aurora is fully managed by RDS and it automatically and continuously monitors and backs up your database to Amazon S3, enabling granular point-in-time recovery.
- Customers can scale the compute resources allocated to their DB Instance by changing the DB Instance class.

How?

- You choose Aurora as the DB engine option when setting up new database servers through Amazon RDS.
- After launching an Aurora instance, you can connect to it using any database client that supports MySQL or PostgreSQL.

Created by:

Ashish Prajapati

How much?

- For Amazon Aurora On-Demand Instances, you pay by the hour. You can also choose Reserved Instances for additional savings.
- Aurora storage is billed in per GB-month increments, while I/Os consumed are billed in per million request increments.

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Database

Amazon RDS

What?

- Amazon Relational Database Service (Amazon RDS) is a managed service that makes it easy to set up, operate, and scale a relational database in the cloud.
- Amazon RDS supports Amazon Aurora, MySQL, MariaDB, Oracle, SQL Server, and PostgreSQL database engines.

Why?

- You want to focus on your applications and business instead of managing time-consuming database administration tasks.
- Once your database is up and running, Amazon RDS automates common administrative tasks, such as performing backups and patching the software that powers your database.

When?

- You need the capabilities of a familiar MySQL, MariaDB, Oracle, SQL Server, PostgreSQL or Amazon Aurora database.
- You want the flexibility of being able to easily scale the compute resources or storage capacity associated with your relational database instance.

Where?

- Amazon RDS can be deployed in a Single AZ or Multi-AZ. When you provision a Multi-AZ database instance, Amazon RDS synchronously replicates your data to a standby instance in a different Availability Zone (AZ).
- Read Replica – In the same or different AWS Region than the Amazon RDS Instance.

Who?

- Amazon RDS manages the work involved in setting up a relational database from provisioning the infrastructure capacity you request to installing the database software.
- You are responsible for managing the database settings that are specific to your application.

How?

- The basic building block of Amazon RDS is the DB instance. You can get started by creating a DB instance that can contain one or more databases. You can access your DB instance by using the same tools and applications that you would use with a standalone database instance.

How much?

- You are billed based on: DB instance hours, Storage (per GB per month), Provisioned IOPS per month, Backup Storage and Data transfer. Either you could use a on-demand instance and pay hourly rate or purchase reserved instances (commit to either one-year or three-year terms).

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

Reference:

FAQs

Category:

Database

Amazon DynamoDB

What?

- Amazon DynamoDB is a fully managed, serverless, NoSQL database designed to support key-value and document data models.
- DynamoDB has a flexible schema, to easily adapt the tables as your business requirements change, without having to redefine the table schema as you would in relational databases.

Why?

- DynamoDB offers built-in security, continuous backups, automated multi-Region replication, in-memory caching, and data export tools.
- You can scale up or scale down your tables' throughput capacity without downtime or performance degradation.

When?

- You want to build internet-scale applications supporting user-content metadata and caches that require high concurrency and connections for millions of users, and millions of requests per second.
- You want to support high-traffic, extreme-scaled events, encryption at rest with no operational overhead.

Where?

- DynamoDB is a regional service.
- All of your data is stored on SSDs and is automatically replicated across multiple Availability Zones in an AWS Region.
- You can use global tables to keep DynamoDB tables in sync across AWS Regions.

Who?

- Amazon DynamoDB is a fully managed service.
- It automatically scales tables to adjust for capacity and maintains performance with zero administration.
- Availability and fault tolerance are built in and it also provides on-demand backup capability.

How?

- DynamoDB stores data in a table. A table is a collection of items, and each item is a collection of attributes. An attribute is a fundamental data element, which does not need to be broken down further.
- It uses primary keys to uniquely identify each item in a table and secondary indexes to provide more querying flexibility.

Created by:

Ashish Prajapati

How much?

- DynamoDB charges are calculated for reading, writing, and storing data in tables, along with any optional features you choose to enable. DynamoDB has two capacity modes, which come with specific billing options for processing reads and writes on your tables: on-demand and provisioned.

Reference:

FAQs

Category:

Serverless

AWS Serverless Application Model (SAM)

What?

- The AWS Serverless Application Model (AWS SAM) is an open-source framework that you can use to build serverless applications on AWS. AWS SAM consists of two components - AWS SAM template specification (to define your serverless application) and AWS SAM CLI (a command line tool that operates on AWS SAM templates to build serverless applications).

Why?

- AWS SAM makes it easy to organize related components and resources, and operate on a single stack.
- You can use AWS SAM to define and deploy your infrastructure as config.
- It provides shorthand syntax to express functions, APIs, databases, and event source mappings.

When?

- You need a shorthand syntax to describe your serverless application (Lambda functions, API endpoints, DynamoDB tables, and other resources) using a simple YAML template with just a few lines per resource. You need a Lambda-like local execution environment to help you catch issues by providing parity with the actual Lambda execution environment.

Where?

- AWS SAM is available in all regions where AWS Lambda is available.
- You can install AWS SAM CLI on any supported Linux, Mac, or Windows platform using pip. It is included with AWS Cloud9. You can also use SAM CLI to locally debug Lambda functions written in Node.js, Java, Python, and Go.

Who?

- You construct an AWS SAM template to declare and configure the components of your application.
- AWS SAM templates are an extension of AWS CloudFormation templates, so any resource that you can declare in an AWS CloudFormation template can also be declared in an AWS SAM template.

Complete book:

Click Here

How?

- Using AWS SAM CLI you initialize and configure applications, debug locally using IDEs like Visual Studio Code or JetBrains WebStorm, and deploy to the AWS Cloud. During deployment, SAM transforms and expands the shorthand SAM syntax into an AWS CloudFormation template and then, CloudFormation provisions your resources.

Created by:

Ashish Prajapati

How much?

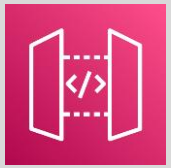
- There is no additional charge to use AWS SAM. You pay for the AWS resources created using SAM in the same manner as if you created them manually.

Reference:

FAQs

Category:

Front End
Web and
Mobile



Amazon API Gateway

Complete book:

Click Here

Created by:

Ashish Prajapati

What?

- Amazon API Gateway is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale. API Gateway supports containerized and serverless workloads, as well as web applications.
- It allows you to create RESTful APIs and WebSocket APIs that enable real-time two-way communication applications.

Why?

- Amazon API Gateway provides developers with a simple, flexible, fully managed, pay-as-you-go service that handles all aspects of creating and operating robust APIs for application back ends. With API Gateway, you can launch new services faster and with reduced investment so you can focus on building your core business services.

When?

- You want to use a managed service to save undifferentiated heavy lifting involved in securely and reliably running APIs (REST, HTTP, and WebSocket APIs).
- To save effort on API development and API management and generate client SDKs for a number of languages.

Where?

- Amazon API Gateway is a regional service.
- An API endpoint refers to the hostname of the API. For the REST APIs the API endpoint type can be edge-optimized, regional, or private, depending on where the majority of your API traffic originates from.

Who?

- Amazon API Gateway handles all of the tasks involved in accepting and processing concurrent API calls, including traffic management, authorization and access control, monitoring, and API version management.
- As an API developer, you can create and manage an API by using the API Gateway console, or by calling the API references.

How?

- API Gateway acts as a "front door" for applications to access data, business logic, or functionality from your backend services, such as workloads running on Amazon EC2, code running on AWS Lambda, any web application, or real-time communication applications.

How much?

- For HTTP APIs and REST APIs, you pay only for the API calls you receive and the amount of data transferred out.
- For WebSocket APIs, you pay for messages sent and received and for the time a user/device is connected to the WebSocket API.
- API Gateway also provides optional data caching charged at an hourly rate that varies based on the cache size you select.

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Management and Governance

AWS CloudFormation

More SSCs:

Click Here

Complete Book

Click Here

Created by:

Ashish Prajapati

What?

- AWS CloudFormation gives you an easy way to model a collection of related AWS and third-party resources, provision them quickly and consistently, and manage them throughout their lifecycles, by treating infrastructure as code (IaC).
- It enables you to use a template file to create and delete a collection of resources together as a single unit (a stack).

Why?

- Automate, test, and deploy infrastructure templates with continuous integration and delivery (CI/CD) automations.
- Run anything from a single Amazon Elastic Compute Cloud (EC2) instance to a complex multi-region application.

When?

- You want to use a declarative way to create, update, and delete an entire stack as a single unit, instead of managing resources individually across multiple accounts and regions.
- You want predictable, controlled approach for managing resources across your application portfolio.

Where?

- AWS CloudFormation is a regional service, but it can deploy stacks across multiple accounts and regions using StackSets.
- A template is stored in an Amazon S3 bucket.

Who?

- You create or provide a template (JSON or YAML formatted text file) that describes all the AWS resources that you need, and CloudFormation takes care of provisioning and configuring those resources (stack) for you.
- When creating a stack, AWS CloudFormation makes underlying service calls to AWS to provision and configure resources.

How?

- You create a template (JSON or YAML formatted text file) that describes all the AWS resources that you, and CloudFormation takes care of provisioning and configuring those resources (stack) for you.
- When creating a stack, AWS CloudFormation makes underlying service calls to AWS to provision and configure resources.

How much?

- There is no additional charge for using AWS CloudFormation, you pay for AWS resources created by it as if you had created them manually.

Reference:

FAQs

Category:

Management and Governance

Amazon CloudWatch

Complete book:

Click Here

Created by:

Ashish Prajapati

What?

- Amazon CloudWatch allows you to collect, access, and correlate metrics, logs, and events data on a single platform from across all your AWS resources, applications, and services running on AWS and on-premises.

Why?

- Amazon CloudWatch helps you to detect anomalous behavior in your environments, set alarms, visualize logs and metrics side by side, take automated actions, troubleshoot issues, and discover insights to keep your applications running smoothly.
- It also helps to break down data silos to better understand the health and performance of your resources.

When?

- You want to use a single platform for observability and to collect metrics of AWS and on premises resources.
- You also want to improve operational performance and resource optimization, get operational visibility and insight to derive actionable insights from logs.

Where?

- Amazon CloudWatch is a regional service but you can create cross-account, cross-region dashboards too.
- Through an agent deployed on on-premises system you can also collect metric.

Who?

- It natively integrates with more than 70 AWS services.
- You can create alarms based on metric value thresholds, or use alarms that can watch for anomalous metric behaviour.
- You can install a unified CloudWatch agent to collect logs and metrics.

How?

- CloudWatch is basically a metrics repository. It correlates your metrics and logs to better understand the health and performance of your resources.
- Create alarms based on metric value thresholds, or alarms for anomalous metric behavior based on ML algorithms.

How much?

- Charges are calculated for number of Metrics (includes detailed and custom metrics), APIs, Logs Ingested, Log Storage/Archival, Logs Insights Queries (analyse Log Data), Events, dashboards, alarms, Contributor Insights, Lambda Insights and Canaries.

Reference:

FAQs

Category:

Management and Governance

AWS CloudTrail

What?

- AWS CloudTrail is an AWS service that helps you enable governance, compliance, and operational and risk auditing of your AWS account. Actions taken by a user, role, or an AWS service are recorded as events in CloudTrail.
- Events include actions taken in the AWS Management Console, AWS Command Line Interface, and AWS SDKs and APIs.

Why?

- CloudTrail helps you prove compliance, improve security posture, and consolidate activity records across regions and accounts. It provides visibility into user activity by recording actions taken on your account.
- CloudTrail makes it easier to ensure compliance with internal policies and regulatory standards.

When?

- You should use CloudTrail if you need to audit activity, monitor security, or troubleshoot operational issues.
- You want to capture and consolidate user activity and API usage across AWS Regions and accounts on a single, centrally controlled platform.

Where?

- Activity information for services with regional end points (EC2, RDS etc.) is captured and processed in the same region as the action is made, and delivered to the region associated with your Amazon S3 bucket.
- You can create two types of trails - A trail that applies to all regions and a trail that applies to one region.

Who?

- CloudTrail is enabled on your AWS account when you create it.
- If you have created an organization in AWS Organizations, you can create a trail that will log all events for all AWS accounts in that organization. Creating an organization trail helps you define a uniform event logging strategy for your organization.

How?

- You can view and search the last 90 days of events recorded by CloudTrail in the CloudTrail console or by using the AWS CLI.
- For an ongoing record of activity and events in your AWS account, create a trail. A trail is a configuration that enables delivery of CloudTrail events to an Amazon S3 bucket, CloudWatch Logs, and CloudWatch Events.

Created by:

Ashish Prajapati

How much?

- You can deliver one copy of your ongoing management events to Amazon S3 for free by creating a trail.
- You can deliver additional copies of events, including data events, using trails. You will be charged for data events or additional copies of management events. Once a CloudTrail trail is set up, Amazon S3 charges apply based on your usage.

Reference:

FAQs

Category:

Migration and Transfer:

AWS Database Migration Service (AWS DMS)

Complete book:

Click Here

Created by:

Ashish Prajapati

What?

- AWS DMS is a cloud service that makes it easy to migrate relational databases, data warehouses, NoSQL databases, and other types of data stores into the AWS Cloud or between combinations of cloud and on-premises setups.
- During migration the source database remains fully operational, minimizing downtime for applications relying on the database.

Why?

- AWS DMS supports homogeneous migrations as well as heterogeneous migrations between different database platforms.
- DMS manages all the complexities of the migration process including automatically replicating data changes that occur in the source database during the migration process without any need to install any drivers or applications.

When?

- You want to perform one-time migrations and/or replicate ongoing changes to keep sources and targets in sync.
- If you want to migrate to a different database engine, you can use the AWS Schema Conversion Tool (AWS SCT) to translate your database schema to the new platform. You then use AWS DMS to migrate the data.

Where?

- AWS DMS replication instance is deployed in a region. It supports Multi-AZ deployment and automatically provisions and maintains a standby replica of the replication instance in a different Availability Zone.
- At least one of endpoint (source or destination) must be on an AWS service.

Who?

- AWS DMS automatically manages the deployment, management, and monitoring of all hardware and software needed for your migration. You can scale up (or scale down) your migration resources as needed to match your actual workload.
- You create an AWS DMS migration by creating the necessary replication instance, endpoints, and replication tasks.

How?

- AWS DMS uses a replication instance (a manged EC2 instance that runs replication software) for migration.
- At a high level you create a replication server, create source and target endpoints that have connection information about your data stores and create one or more migration tasks to migrate data between the source and target data stores.

How much?

- You pay for your replication instances (by the hour) and any additional log storage.
- All data transfer into AWS DMS is free, and data transferred between AWS DMS and databases in Amazon RDS and Amazon EC2 instances in the same Availability Zone also is free.

Reference:

FAQs

Amazon CloudFront

What?

- Amazon CloudFront is a web service that speeds up distribution of your static and dynamic web content to your users.
- It reduces latency by delivering data through 310+ Points of Presence (300+ Edge locations and 13 regional mid-tier caches) with automated network mapping and intelligent routing.

Why?

- Compared to self-hosting, Amazon CloudFront spares you from the expense and complexity of operating a network of cache servers in multiple sites across the internet. You also benefit from tight integration with other Amazon Web Services.
- It also eliminates the need to over-provision capacity in order to serve potential spikes in traffic.

When?

- You want to accelerate content delivery and APIs, stream live and on-demand video, reach viewers across the globe in milliseconds with built-in data compression, edge compute capabilities, and field-level encryption.
- You want to customize the code you run at the AWS content delivery network (CDN) edge using Lambda@Edge.

Where?

- Amazon CloudFront is a global service. It serves content from an origin. An origin is the location where content is stored, and from which CloudFront gets content to serve to viewers.
- An origin can be Amazon S3 bucket, ELB, AWS Media Services or any HTTP host (including on premises server).

Who?

- You create a CloudFront distribution by specifying an origin and details about how to track and manage content delivery by configuring caching policy, security and logging.

How?

- When a user requests content that you are serving with CloudFront, the request is routed to the edge location that provides the lowest latency (time delay), so that content is delivered with the best possible performance. If the content is not in that edge location, CloudFront retrieves it from an origin that you have defined.

Category:

Networking and Content Delivery

How much?

- For on-demand pricing, CloudFront charges traffic served via data transfers out from edge locations, along with HTTP or HTTPS requests. There are no transfer fees for origin fetches from any AWS origin.
- For commitment based pricing use CloudFront Savings Bundle and Custom Pricing.

Reference:

FAQs

Amazon Route 53

What?

- Amazon Route 53 is a highly available and scalable cloud Domain Name System (DNS) web service.
- You can use Route 53 to perform four main functions in any combination: Domain registration, Domain Name System (DNS) service, Health checking and DNS Resolver. Route 53 can also be used with non-AWS resources.

Why?

- The globally distributed nature of Route 53 service helps ensure a consistent ability to route your end users to your application by circumventing any internet or network related issues. Using a global anycast network of DNS servers around the world, it is designed to automatically answer queries from the optimal location depending on network conditions.

When?

- You want to use any of the following routing policies to respond to queries: Simple routing policy, Failover routing policy, Geolocation routing policy, Geoproximity routing policy, Latency routing policy, Multivalue answer routing policy and Weighted routing policy.

Where?

- Amazon Route 53 is a global service.
- Route 53 uses a global network of DNS servers at a series of world-wide locations to offer you high availability and increased performance.

Who?

- When you register a domain with Route 53, AWS automatically make Route 53 the DNS service for the domain.
- You can configure Hosted zones, Resource records, Health checks, Traffic Policies, Route 53 resources and other settings for your resources.

How?

- When Route 53 receives a DNS query for a resource in a hosted it returns the IP address to requester.
- For Amazon VPC, Route 53 Resolver automatically uses a Resolver on the VPC to answer DNS queries.
- It uses health checks to monitor the health and performance of your web applications, web servers, and other resources.

Category:

Networking and Content Delivery

How much?

- You pay a monthly charge for each hosted zone managed with Route 53.
- You incur charges for every DNS query answered by the Amazon Route 53 service
- You pay an annual charge for each domain name registered via or transferred into Route 53.

Reference:

FAQs

Category:

Networking and Content Delivery

AWS Transit Gateway

Complete book:

Click Here

Created by:

Ashish Prajapati

What?

- AWS Transit Gateway connects your VPCs and on-premises networks through a central hub. AWS Transit Gateway supports dynamic and static routing between attached Amazon VPCs and VPNs.

Why?

- It acts as a cloud router to simplify your network architecture and puts an end to complex peering relationships.
- With inter-Region peering, everything attached to an AWS Transit Gateway is shared across AWS Regions. This includes VPCs, DNS, Microsoft Active Directory, and IPS/IDS.

When?

- You want to deploy new applications across VPCs without updating massive route tables to create peering relationships.
- You want to host multicast applications such as video conferencing, media, or teleconferencing without redesigning your application or tweaking your on-premises network.

Where?

- AWS Transit Gateway is a regional resource and enables you to attach VPCs and VPN connections (within or across AWS accounts) in the same region and route traffic between them.
- You can peer two transit gateways hosted in the same AWS region or across regions, and route traffic between them.

Who?

- Your transit gateway automatically comes with a default route table. You can segment your network by creating multiple route tables in an AWS Transit Gateway and associate Amazon VPCs and VPNs to them.
- A transit gateway scales elastically based on the volume of network traffic.

How?

- Routing through a transit gateway operates at layer 3, where the packets are sent to a specific next-hop attachment, based on their destination IP addresses. When a packet comes from one attachment, it is routed to another attachment using the route that matches the destination IP address.

How much?

- Charges are determined by two factors:
- AWS Transit Gateway hourly charge - calculated per AWS Transit Gateway attachment.
- Transit Gateway data processing charge – calculated per GB of data processed.

Reference:

FAQs

Category:

Networking and Content Delivery

VPC Endpoints

What?

- A VPC endpoint enables connections between a VPC and supported services, without the need of an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection.
- There are three types of VPC endpoints – Interface endpoints, Gateway Load Balancer endpoints, and Gateway endpoints.

Why?

- VPC endpoints enable you to connect your VPC to services using private IP addresses, as if those services were hosted directly in your VPC. Traffic between a VPC endpoint and an endpoint service is encrypted and stays within the AWS network, without traversing the public internet. Using an endpoint policy, you can control access to the endpoint service.

When?

- Interface endpoints - to access supported AWS services, PrivateLink Ready partner services, AWS Marketplace services, other endpoint services.
- Gateway endpoint - to send traffic to Amazon S3 or DynamoDB.

Where?

- A service provider creates an endpoint service to make their service available in a Region.
- When you create an interface VPC endpoint, AWS creates Regional and zonal DNS names that you can use to communicate with the AWS service from your VPC.

Who?

- VPC endpoints are virtual devices. They are horizontally scaled, redundant, and highly available VPC components..
- A service consumer creates a VPC endpoint to connect their VPC to an endpoint service.

Complete book:

Click Here

How?

- Interface endpoints and Gateway Load Balancer endpoints are powered by AWS PrivateLink, and use an Elastic Network Interface (ENI) as an entry point for traffic destined to the service.
- Gateway endpoints serve as a target for a route in your route table for traffic destined for the service.

Created by:

Ashish Prajapati

How much?

- For Interface endpoints and Gateway Load Balancer endpoints, you are charged for each hour that your VPC endpoint is provisioned in each AZ. There is no additional charge for using gateway endpoints. Data processing charges apply for each Gigabyte processed through the VPC endpoint regardless of the traffic's source or destination.

Reference:

FAQs

Category:

Networking and Content Delivery

Amazon VPC

What?

- Amazon Virtual Private Cloud (Amazon VPC) enables you to provision a logically isolated section of the AWS Cloud where you can launch AWS resources in a virtual network that you have defined.
- This virtual network closely resembles a traditional network that you would operate in your own data center.

Why?

- You can define your own network space, and control how your network and the Amazon EC2 resources inside your network are exposed to the Internet.
- You can also leverage more granular access to and from the Amazon EC2 instances in your virtual network.

When?

- You want to launch AWS resources in a logically isolated virtual network and spend less time setting up, managing, and validating your virtual network.
- You want to use multiple layers of security, including security groups and network access control lists.

Where?

- VPC is a regional entity and spans across all of the Availability Zones in the region.
- Each subnet must reside entirely within one Availability Zone and cannot span zones.
- You can launch AWS resources, such as EC2 instances, into a specific subnet.

Who?

- Your AWS resources are automatically provisioned in a ready-to-use default VPC or you can create additional VPCs.
- You can specify an IP address range for the VPC, add subnets, associate security groups, and configure route tables.

Complete book:

Click Here

How?

- When you create a VPC, you must specify an IPv4 CIDR block for the VPC. Afterwards you can add subnets, route tables, security groups, network access control list, an internet gateway, and other gateways as necessary.

Created by:

Ashish Prajapati

How much?

- There is no additional charge for using a VPC. There are charges for some VPC components, such as NAT gateways, Reachability Analyzer, and traffic mirroring. Usage charges for other Amazon Web Services, including Amazon EC2, still apply at published rates for those resources, including data transfer charges.

Reference:

FAQs

Category:

Networking and Content Delivery

AWS VPN

What?

- AWS Virtual Private Network (VPN) solutions establish secure connections between your on-premises networks, remote offices, client devices, and the AWS global network.
- AWS VPN is comprised of two services: AWS Site-to-Site VPN and AWS Client VPN.

Why?

- Because it is a cloud VPN solution, you don't need to install and manage hardware or software-based solutions, or try to estimate how many remote users to support at one time.
- Data routes over an encrypted VPN connection to help maintain the confidentiality and integrity of data in transit.

When?

- AWS Site-to-Site VPN – When you want to create encrypted tunnels between your locations (such as data centers and remote offices) and your AWS VPC or AWS Transit Gateways.
- AWS Client VPN – Your remote workforce wants to securely access resources within both AWS and your on-premises.

Where?

- Amazon side of VPN connectors (Virtual Private Gateway, Client VPN endpoint) are a regional entity.
- Customer side of VPN connectors (Customer gateway device, OpenVPN-based VPN client) run on-premises.

Who?

- AWS manages high availability and scalability for AWS side of the VPN connection. Each Site-to-Site VPN connection offers two tunnels for high availability.
- Customers manage high availability and scalability for the customer side of the VPN connection.

Complete book:

Click Here

How?

- AWS Site-to-Site VPN – Create a VPN connection by connecting an existing Virtual Private Gateway (VPN concentrator on Amazon side) or transit gateway to the customer gateway (VPN concentrator on customer side).
- AWS Client VPN – Create a Client VPN endpoint on Amazon side and connect to it using OpenVPN-based VPN client.

Created by:

Ashish Prajapati

How much?

- AWS Site-to-Site VPN – Charged for each VPN connection-hour that your VPN connection is provisioned and available.
- AWS Client VPN – Charged per hour for the number of active client connections and the number of associated subnets.
- Standard data transfer charges apply for data transferred via the VPN.

Reference:

FAQs

AWS Certificate Manager (ACM)

What?

- AWS Certificate Manager (ACM) makes it easy to provision, manage, deploy, and renew SSL/TLS certificates for use with AWS services and your internal connected resources.
- SSL/TLS are industry standard protocols for encrypting network communications and establishing the identity of websites.

Why?

- AWS Certificate Manager can help you meet regulatory and compliance requirements for encryption of data in transit.
- It removes many of the time-consuming and error-prone steps to acquire a SSL/TLS certificate like generate a key pair or certificate signing request (CSR), submit a CSR to a certificate authority, or upload and install the certificate once received.

When?

- You need free public certificates for ACM-integrated services, such as ELB, Amazon CloudFront and API Gateway.
- You want to automate renewal and deployment of certificates.
- You need to create private certificates for your internal resources and manage the certificate lifecycle centrally.

Where?

- AWS Certificate Manager (ACM) is a regional service.
- Certificates generated by it can be used for ACM-integrated services which could be global such as Amazon CloudFront distribution.

Who?

- You can provide certificates for your integrated AWS services either by issuing them directly with ACM or by importing third-party certificates into the ACM management system.
- With ACM Private CA, you can create your own certificate authority (CA) hierarchy and issue certificates.

How?

- 1. Request or import a TLS/SSL certificate you would like to use into your AWS account.
- 2. Validate domain ownership for your requested certificate using DNS or email validation to complete certificate issuance.
- 3. Use your newly issued or imported certificates in various AWS services like ELB, Amazon CloudFront etc.

Category:

Security, Identity, and Compliance

How much?

- Public and private certificates provisioned through ACM for use with ACM-integrated services are free.
- For AWS Certificate Manager Private Certificate Authority (CA), you pay monthly for the operation of the private CA and for the private certificates you issue.

Reference:

FAQs

Category:

Security, Identity, and Compliance

AWS Identity and Access Management (IAM)

What?

- AWS Identity and Access Management (IAM) provides fine-grained access control across all of AWS Services. With IAM, you can specify who can access which services and resources, and under which conditions.
- With IAM policies, you manage permissions to your workforce and systems to ensure least-privilege permissions.

Why?

- You use IAM to control who is authenticated (signed in) and authorized (has permissions) to use resources.
- You can grant other people permission to administer and use resources in your AWS account without having to share your password or access key.

When?

- You want to grant different fine-grained permissions to different people for different resources.
- You want to add two-factor authentication to your account and to individual users for extra security.
- You need to use existing corporate identities to grant secure access to AWS resources using identity federation.

Where?

- IAM is a global service.
- You use IAM to control access to tasks that are performed using the AWS Management Console, the AWS Command Line Tools, or service API operations using the AWS SDKs.

Who?

- You manage access in AWS by creating policies and attaching them to IAM identities (users, groups of users, or roles) or AWS resources.
- You can create multiple IAM users under your AWS account or enable temporary access through identity federation.

Complete book:

Click Here

How?

- With IAM, you define who can access what by specifying fine-grained permissions. IAM then enforces those permissions for every request. Access is denied by default and access is granted only when permissions specify an “Allow”.
- You can delegate access to users or AWS services to operate within your AWS account.

Created by:

Ashish Prajapati

How much?

- There is no charge to use IAM.

Reference:

FAQs

Category:

Security, Identity, and Compliance

Amazon Inspector

What?

- Amazon Inspector is an automated vulnerability management service that scans Amazon EC2 instances and container images residing in Amazon Elastic Container Registry (Amazon ECR) for software vulnerabilities and unintended network exposure.
- It provides a highly contextualized and meaningful risk score for each finding to help you set more accurate response priorities.

Why?

- It removes the operational overhead associated with deploying and configuring a vulnerability management solution.
- Amazon Inspector includes a knowledge base of hundreds of rules mapped to common security standards and vulnerability definitions that are regularly updated by AWS security researchers.

When?

- You want to reduce mean time to remediate (MTTR) vulnerabilities and streamline workflow with Amazon EventBridge and AWS Security Hub integrations.
- You need an aggregated dashboard offering a high-level view of findings from across your environment.

Where?

- Amazon Inspector is a regional service.
- Amazon Inspector can be enabled for your entire organization with a single click. Additionally, you can automate enabling the service for future members whenever they join your organization.

Who?

- Amazon Inspector uses the widely deployed AWS Systems Manager (SSM) Agent to collect the software inventory and configurations from your Amazon EC2 instances.
- Amazon Inspector automatically discovers and begins scanning your eligible resources.

Complete book:

Click Here

How?

- When a software vulnerability or network issue is discovered, Amazon Inspector creates a finding. A finding describes the vulnerability, identifies the affected resource, rates the severity of the vulnerability, and provides remediation guidance.
- You can analyse the finding in the Amazon Inspector console, or view and process your findings through other AWS services.

Created by:

Ashish Prajapati

How much?

- Monthly costs are determined by a combination of two dimensions: Amazon EC2 instances being scanned, and the total number of container images initially scanned when pushed to Amazon Elastic Container Registry (ECR) and rescanned during a month.

Reference:

FAQs

Category:

Security, Identity, and Compliance

AWS Key Management Service (AWS KMS)

Complete book:

Click Here

Created by:

Ashish Prajapati

What?

- AWS Key Management Service (KMS) is a managed service that enables you to easily create and control the keys used for cryptographic operations. The service provides a highly available key generation, storage, management, and auditing solution to encrypt or digitally sign data within your applications or control the encryption of data across AWS services.

Why?

- AWS KMS presents a single control point to manage keys and define policies consistently across integrated AWS services and your own applications.
- It reduces your licensing costs and operational burden by providing a scalable key management infrastructure.

When?

- You want to centrally create, import, rotate, delete, and manage permissions on keys that control access to your data.
- You want to perform digital signing operations using asymmetric key pairs to ensure the integrity of your data.

Where?

- AWS KMS is a regional service. KMS keys are never shared outside the AWS region in which they were created.
- AWS KMS supports multi-Region keys, which are AWS KMS keys in different AWS Regions that can be used interchangeably – as though you had the same key in multiple regions.

Who?

- AWS KMS is a fully managed service.
- You control access to your encrypted data by defining permissions to use keys, while AWS KMS enforces your permissions and handles the durability and physical security of your keys.

How?

- AWS KMS uses hardware security modules (HSM) to protect and validate your AWS KMS keys. An HSM is a physical device that provides extra security for sensitive data. To protect data at rest, integrated AWS services use envelope encryption, where a data key is used to encrypt data, and is itself encrypted under a KMS key.

How much?

- You pay US $1/month to store any key that you create. You also pay for the number of API requests made to the AWS KMS.
- AWS managed keys that are created on your behalf by AWS services are free to store. You are charged per-request when you use or manage your keys beyond the free tier.

Service Summary Cards (SSC)

Reference:

FAQs

Category:

Security, Identity, and Compliance

AWS Secrets Manager

What?

- AWS Secrets Manager enables you to store, manage, rotate and retrieve, secrets for resources in AWS Cloud, third-party services, and on-premises. You can manage secrets such as database credentials, on-premises resource credentials, SaaS application credentials, third-party API keys, SSH keys, and any text blurb that is 64 KB or smaller.

Why?

- Secrets Manager enables you to replace hardcoded credentials in your code, including passwords, with an API call to Secrets Manager to retrieve the secret programmatically.
- It eliminates the investment and on-going maintenance costs of operating your own secrets management infrastructure.

When?

- You want to encrypt secrets at rest to reduce the likelihood of unauthorized users viewing sensitive information.
- You want to automatically rotate the secret on a specified schedule to replace long-term secrets with short-term ones.

Where?

- AWS Secrets Manager is a regional service. It allows you to replicate secrets in multiple AWS regions to support your multi-region applications and disaster recovery scenarios.
- For auditing and monitoring of secrets usage, it integrates with AWS logging, monitoring, and notification services

More SSCs:

Click Here

Complete Book

Click Here

Who?

- AWS Secrets Manager is a fully managed service.
- You can attach AWS IAM permission policies to your users, groups, and roles that grant or deny access to specific secrets, and restrict management of those secrets.

How?

- You can get started by storing a secret in Secrets Manager. To retrieve secrets, you replace secrets in plain text in your applications with code to pull in those secrets programmatically using the Secrets Manager APIs.
- For rotation you can use ready-to-use features for supported AWS services or write a custom AWS Lambda function.

Created by:

Ashish Prajapati

How much?

- You pay for the number of secrets managed in Secrets Manager (per secret per month) and the number of Secrets Manager API calls made (per 10,000 API calls).
- A replica secret is considered a distinct secret and charged accordingly.

Reference:

Documentation

Category:

Security, Identity, and Compliance

AWS Security Token Service (AWS STS)

What?

- AWS Security Token Service (AWS STS) enables you to request temporary, limited-privilege credentials for AWS IAM users or for federated users.
- Temporary security credentials are short-term and are not stored with the user but are generated dynamically and provided when requested. They can be configured to last for a few minutes or for several hours.

Why?

- If you use AWS STS, you do not have to distribute or embed long-term AWS security credentials with an application.
- You can provide access to your AWS resources to users without having to define an AWS identity for them.

When?

- Temporary credentials are useful in scenarios that involve identity federation, delegation, cross-account access, and IAM roles. Temporary credentials are the basis for roles and identity federation.
- The temporary credentials have a limited lifetime, so you do not have to rotate them or explicitly revoke them when they're no longer needed.

Where?

- By default, AWS STS is available as a global service, and all AWS STS requests go to a single endpoint that maps to the US East (N. Virginia).
- You can use Regional AWS STS endpoints instead of the global endpoint to reduce latency, build in redundancy, and increase session token validity.
- No matter which Region your credentials come from, they work globally.

Who?

- You can use temporary security credentials to make requests for AWS resources using the AWS CLI or AWS API (using the AWS SDKs).
- When (or even before) the temporary security credentials expire, the user can request new credentials, as long as the user requesting them still has permissions to do so.

How?

- The AWS STS API operations create a new session with temporary security credentials that include an access key pair (consists of an access key ID and a secret key) and a session token. Users (or an application that the user runs) can use these credentials to access your resources.

Created by:

Ashish Prajapati

How much?

- There is no additional cost to use AWS STS.

Reference:

FAQs

Category:

Storage

AWS Backup

What?

- AWS Backup is a fully managed service to centralize and automate data protection across on-premises and AWS services.
- It works across AWS services for compute, storage and databases and allows you to define Backup plans, schedule backups, automate backup retention management, centrally monitor backup activity, and restore backups.

Why?

- Protecting your data is an important step to ensure that you meet your business and regulatory compliance requirements.
- AWS Backup provides automated backup schedules, retention management, and lifecycle management, removing the need for custom scripts and manual processes. You can also apply backup policies to your AWS resources by simply tagging them.

When?

- You need a centralized console to automate and manage backups across AWS services and hybrid environment (data stored in AWS Storage Gateway volumes and VMware workloads, on premises and VMware Cloud on AWS).
- You want to use backup policies to align your backup strategy with your internal or regulatory requirements.

Where?

- It is regional service and stores backups in a backup vault - an encrypted storage location in your AWS account in a region.
- AWS Backup supports following AWS service: Amazon EBS, Amazon EC2, Amazon RDS / Amazon Neptune / Amazon DocumentDB, Amazon DynamoDB, Amazon EFS and Amazon FSX, AWS Storage Gateway volumes, and Amazon S3.

Who?

- You can create backup policies called backup plans that enables you to define your backup requirements and then apply them to the AWS resources you want backed up.
- You can also configure Cross-Region and Cross-Account backups across your AWS accounts within your AWS organizations.

Complete book:

Click Here

How?

- AWS Backup allows you to define a central data protection policy (called a backup plan) that works across AWS services for compute, storage, and databases. The backup plan defines parameters such as backup frequency and backup retention period. Once you define your data protection policies and assign AWS resources to the policies, AWS Backup automates the creation of backups and stores those backups in an encrypted backup vault that you designate.

Created by:

Ashish Prajapati

How much?

- Backup storage pricing - amount of storage space your backup data consumes (GB-month)
- Restore pricing - amount of data restored for the month (per GB)
- Additional charges apply for Cross-Region data transfer and AWS Backup Audit Manager usage.

Reference:

FAQs

Category:

Storage

Amazon Elastic File System (Amazon EFS)

What?

- Amazon EFS is a fully managed service providing NFS shared file system storage for Linux workloads.
- It provides a highly durable and highly available shared filesystems for AWS compute services (Amazon EC2, AWS Lambda Amazon container services - Amazon ECS, Amazon EKS, AWS Fargate) and on-premises resources.

Why?

- It automatically grows and shrinks as you add and remove files and bursts to higher throughput levels when necessary.
- It supports the Network File System version 4 (NFSv4.1 and NFSv4.0) protocol, so the applications and tools that you use today work seamlessly with Amazon EFS.

When?

- You want to create a shared file system that can be accessed concurrently from multiple NFS clients.
- You need file system access semantics, such as strong data consistency and file locking.
- You want to control access to your file systems through Portable Operating System Interface (POSIX) permissions.

Where?

- Amazon EFS is a regional service.
- To access your Amazon EFS file system in a VPC, you create one or more mount targets in the VPC. An Amazon EFS file system can only have mount targets in one VPC at a time. Mount targets themselves are designed to be highly available.

Who?

- The service manages all the file storage infrastructure for you, meaning that you can avoid the complexity of deploying, patching, and maintaining complex file system configurations.

Complete book:

Click Here

How?

- With Amazon EFS, you can create a file system, mount the file system using mount target (IP address for an NFSv4 endpoint) on supported service, and then read and write data to and from your file system.
- Once mounted, you can work with the files and directories in your file system just like you would with a local file system.

Created by:

Ashish Prajapati

How much?

- You pay for storage capacity, request (read and write), and for any provisioned throughput.
- Amazon EFS offers four storage classes to suit your data access and availability requirement.

Reference:

FAQs

Category:

Storage

Amazon Simple Storage Service (Amazon S3)

Complete book:

Click Here

Created by:

Ashish Prajapati

What?

- Amazon S3 is a simple key-based object storage built to store and retrieve any amount of data from anywhere.
- Data is stored as objects within resources called “buckets”, and a single object can be up to 5 terabytes in size.
- It is designed to provide 99.999999999% (11 9's) of data durability.

Why?

- You can use a simple web service interface to store and retrieve virtually any amount of data in any format.
- Highly scalable, highly available, fast, inexpensive data storage infrastructure and you only pay for what you use.
- It offers a range of storage classes to choose from based on the data access, resiliency, and cost requirements of workloads.

When?

- You want to store static content, storage backups, want to build data lakes.
- You require version controlled object storage, Multi-Factor Authentication (MFA) Delete capability and selectively grant permissions to users and groups of users.

Where?

- Amazon S3 stores data as objects within buckets. A bucket is created in a Region and requires a globally unique name.
- S3 storage classes provide multi-Availability Zone (AZ) resiliency by redundantly storing data on multiple devices and physically separated AWS Availability Zones in an AWS Region (except S3 One Zone-IA storage class).

Who?

- Upon creation, only you have access to Amazon S3 buckets that you create, and you have complete control over who has access to your data.
- You can use the Amazon S3 Management Console, the AWS SDKs, or the Amazon S3 APIs to interact with it.

How?

- You can get started by creating a bucket in a specific Region and can defining access controls and management options.
- To store an object in Amazon S3, upload the file into a bucket. An object is composed of a file and any metadata that describes that file. When you store data, you assign a unique object key that can later be used to retrieve the data.

How much?

- There are six Amazon S3 cost components to consider when storing and managing your data— storage pricing, request and data retrieval pricing, data transfer and transfer acceleration pricing, data management and analytics pricing, replication pricing, and the price to process your data with S3 Object Lambda.

Reference:

FAQs

Category:

Storage

AWS Storage Gateway

What?

- AWS Storage Gateway is a hybrid cloud storage service that connects on-premises software appliance with AWS storage services, including Amazon S3, Amazon S3 Glacier, Amazon S3 Glacier Deep Archive, and Amazon EBS.
- The service enables you to securely upload data to the AWS Cloud for cost effective backup and rapid disaster recovery.

Why?

- Storage Gateway seamlessly connects to your local production or backup applications with NFS, SMB, iSCSI, or iSCSI-VTL, so you can adopt AWS Cloud storage without needing to modify your applications. It provides low-latency performance by caching frequently accessed data on premises, while storing data securely and durably in AWS storage services.

When?

- You want to reduce your on-premises storage footprint and associated costs by leveraging AWS storage services.
- You need to offer virtually unlimited cloud storage to users and applications without deploying new storage hardware.
- You want to use built in compression, encryption, and bandwidth management features for data transfer to AWS services.

Where?

- AWS Storage Gateway is a regional service. You need to deploy a Storage Gateway virtual machine (VM) to connect to it.
- You can deploy this VM on premises (on VMware ESXi, Microsoft Hyper-V, Linux KVM or as a hardware appliance) or in AWS Cloud (as a VM in VMware Cloud on AWS, or as an AMI in Amazon EC2).

Who?

- You use the AWS Management Console to download the virtual appliance gateway or purchase the hardware appliance, configure storage, and manage and monitor the service. Storage Gateway optimizes data transfer to AWS by sending only changed data and compressing data.

Complete book:

Click Here

How?

- To support these use cases, the service provides four different types of gateways – Tape Gateway, Amazon S3 File Gateway, Amazon FSx File Gateway, and Volume Gateway – that seamlessly connect on-premises applications to cloud storage, caching data locally for low-latency access.

Created by:

Ashish Prajapati

How much?

- Pricing is combination of - Storage Pricing (capacity used in Amazon S3, Amazon FSx for Windows or Amazon EBS snapshots) and Request Pricing (Data written/retrieval to AWS storage by your gateway)
- For Amazon FSx File Gateway you also pay a per hour Gateway Price.