Report TD1

Cloud Computing Working With Amazon AWS

ZunzunWANG & YishuoLYU

1. *AWS Dashboard :*

   Question 1: Please describe briefly the following services:

   - *EC2, Elastic Beanstalk, Lambda*

     EC2: Virtual Servers in the cloud.
     Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity in the Amazon Web Services (AWS) cloud. Using Amazon EC2 eliminates your need to invest in hardware up front, so you can develop and deploy applications faster. You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage. Amazon EC2 enables you to scale up or down to handle changes in requirements or spikes in popularity, reducing your need to forecast traffic.

     Elastic Beanstalk: Run and manage docker containers.
     AWS Elastic Beanstalk is an easy-to-use service for deploying and scaling web applications and services developed with Java, .NET, PHP, Node.js, Python, Ruby, Go, and Docker on familiar servers such as Apache, Nginx, Passenger, and IIS.

     Lambda: Run code in response to events.
     AWS Lambda is a compute service that runs your code in response to events and automatically manages the underlying compute resources for you. You can use AWS Lambda to extend other AWS services with custom logic, or create your own back-end services that operate at AWS scale, performance, and security. AWS Lambda can automatically run code in response to multiple events, such as modifications to objects in Amazon S3 buckets or table updates in Amazon DynamoDB.

   - *S3, Cloudfront, Elastic File System*

     S3: Scalable Storage in the cloud.
     Amazon Simple Storage Service (Amazon S3), provides developers and IT teams with secure, durable, highly-scalable object storage. Amazon S3 is easy to use, with a simple web service interface to store and retrieve any amount of data from anywhere on the web. With Amazon S3, you pay only for the storage you actually use. There is no minimum fee and no setup cost.

Elastic file System: Fully managed file system for EC2.
Amazon Elastic File System (Amazon EFS) is a file storage service for Amazon Elastic Compute Cloud (Amazon EC2) instances. Amazon EFS is easy to use and provides a simple interface that allows you to create and configure file systems quickly and easily. With Amazon EFS, storage capacity is elastic, growing and shrinking automatically as you add and remove files, so your applications have the storage they need, when they need it.

- *RDS, DynamoDB.*
  RDS: managed relational database service.
  Amazon Relational Database Service (Amazon RDS) makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while managing time-consuming database administration tasks, freeing you up to focus on your applications and business. Amazon RDS provides you six familiar database engines to choose from, including Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL and MariaDB.

  DynamoDB: Managed NoSQL Database.
  Amazon DynamoDB is a fully managed NoSQL database service that provides fast and predictable performance with seamless scalability. You can use Amazon DynamoDB to create a database table that can store and retrieve any amount of data, and serve any level of request traffic. Amazon DynamoDB automatically spreads the data and traffic for the table over a sufficient number of servers to handle the request capacity specified by the customer and the amount of data stored, while maintaining consistent and fast performance.

- *CloudCommit and CloudDeploy*
  CloudCommit(CodeCommit): Store Code in Private Git repositories.
  AWS CodeCommit is a fully-managed source control service that makes it easy for companies to host secure and highly scalable private Git repositories. CodeCommit eliminates the need to operate your own source control system or worry about scaling its infrastructure. You can use CodeCommit to securely store anything from source code to binaries, and it works seamlessly with your existing Git tools.

  CloudDeploy(Codeploy): Automate Code Deployments.
  AWS CodeDeploy is part of a family of AWS deployment services that includes AWS Elastic Beanstalk, AWS CodePipeline, AWS CloudFormation, and AWS OpsWorks. AWS CodeDeploy coordinates application deployments to Amazon EC2 instances, on-premises instances, or both. An application can contain deployable content like code, web, and configuration files, executables, packages, scripts, and so on. AWS CodeDeploy deploys applications from Amazon S3 buckets and GitHub repositories.

- *CloudWatch, CloudTrail*
  CloudWatch: Monitor resources and applications.
  Amazon CloudWatch is a monitoring service for AWS cloud resources and the applications you run on AWS. You can use Amazon CloudWatch to collect and track metrics, collect and monitor log files, set alarms, and automatically react to changes in your AWS resources. Amazon CloudWatch can monitor AWS resources such as Amazon EC2 instances, Amazon DynamoDB tables, and Amazon RDS DB instances, as well as custom metrics generated by your applications and services, and any log files your applications generate. You can use Amazon CloudWatch to gain system-wide visibility into resource utilization, application performance, and operational health. You can use these insights to react and keep your application running smoothly.

  CloudTrail: Track User Activity and API Usage.
  AWS CloudTrail is a web service that records AWS API calls for your account and delivers log files to you. The recorded information includes the identity of the API caller, the time of the API call, the source IP address of the API caller, the request parameters, and the response elements returned by the AWS service.
  With CloudTrail, you can get a history of AWS API calls for your account, including API calls made via the AWS Management Console, AWS SDKs, command line tools, and higher-level AWS services (such as AWS CloudFormation). The AWS API call history produced by CloudTrail enables security analysis, resource change tracking, and compliance auditing.

- *Choose one of the analytics section and explain it.*
  EMR: Managed Hadoop Framework
  Amazon Elastic MapReduce (Amazon EMR) is a web service that enables businesses, researchers, data analysts, and developers to easily and cost-effectively process vast amounts of data.

- *IoT (Internet of Things):*
  AWS IoT is a managed cloud platform that lets connected devices -- cars, light bulbs, sensor grids and more, easily and securely interact with cloud applications and other devices. This new managed cloud service provides the infrastructure that allows connected cars, factory floors, aircraft engines, sensor grids, and the like (AWS IoT refers to them as "things") to easily and securely interact with cloud services and with other devices, all at world-scale. The connection to the cloud is fast and lightweight (MQTT or REST), making it a great fit for devices that have limited memory, processing power, or battery life.

  There are several components that make up the AWS IoT:
  1. Things are devices of all types, shapes, and sizes including applications, connected devices, and physical objects.
  2. Thing Shadows are virtual, cloud-based representations of things
  3. The real-time **Rules Engine** transforms messages based on expressions that we define, and routes them to AWS endpoints

4. The **Message Broker** speaks MQTT (and also HTTP 1.1) so our devices can take advantage of alternative protocols even if our cloud backend does not speak them
5. Device SDKs are client libraries that are specific to individual types of devices.
6. The **Thing Registry** assigns a unique identity to each thing.
All of these components can be created, configured, and inspected using the AWS Management Console, the AWS Command Line Interface (CLI), or through the IoT API.

- *Choose one of the Mobile services：*
  Device Farm: Test Android, FireOS, and ios Apps on real Devices in the Cloud.
  Test your app on real devices in the AWS Cloud,Improve the quality of your iOS, Android, and web applications by testing them against real smartphones and tablets in the AWS Cloud.

- *AppStream*
- *AppStream: Low Latency Application Streaming*
  Amazon AppStream lets us deliver our Windows applications to any device.
  Amazon AppStream enables us to stream our existing Windows applications from the cloud, reaching more users on more devices, without code modifications. With Amazon AppStream, our application will be deployed and rendered on AWS infrastructure and the output is streamed to mass-market devices, such as personal computers, tablets, and mobile phones. Because our application is running in the cloud, it can scale to handle vast computational and storage needs, regardless of the devices our customers are using. Amazon AppStream provides an SDK for streaming our application from the cloud. We can integrate our own custom clients, subscriptions, identity, and storage solution with AppStream to build a custom streaming solution that meets the needs of our business.

2. *Question 2 Answer the following questions:*
   -What is a security group?
   A security group acts as a virtual firewall that controls the traffic for one or more instances. When we launch an instance, we associate one or more security groups with the instance. We add rules to each security group that allow traffic to or from its associated instances. We can modify the rules for a security group at any time; the new rules are automatically applied to all instances that are associated with the security group. When we decide whether to allow traffic to reach an instance, we evaluate all the rules from all the security groups that are associated with the instance.

   -Define the term Multi-Tenancy.
   Software Multi-tenancy refers to a software architecture in which a single instance of a software runs on a server and serves multiple tenants. A tenant is a group of users who share a common access with specific privileges to the software instance. With a multitenant architecture, a software application is designed to provide every tenant a dedicated share of the instance including its data, configuration, user management, tenant individual functionality and non-functional properties. Multi-tenancy contrasts with multi-instance architectures, where separate software instances operate on behalf of different tenants.

-The storage type is EBS (Amazon Elastic Block Store).
Amazon Elastic Block Store (Amazon EBS) provides persistent block level storage volumes for use with Amazon EC2 instances in the AWS Cloud. Each Amazon EBS volume is automatically replicated within its Availability Zone to protect us from component failure, offering high availability and durability. Amazon EBS volumes offer the consistent and low-latency performance needed to run our workloads. With Amazon EBS, we can scale our usage up or down within minutes.

3. *Question 3 What are Spot instances? What is their pricing policy?*
   Give a full explanation.
   (You can use resources from Internet to understand Spot instances

Firstly we have to know some Concepts

- ***Spot pool****—A set of unused EC2 instances with the same instance type, operating system, Availability Zone, and network platform (EC2-Classic or EC2-VPC).*
- ***Spot price****—The current market price of a Spot instance per hour, which is set by Amazon EC2 based on the last fulfilled bid. We can also retrieve the Spot price history.*
- ***Spot instance request (or Spot bid)****—Provides the maximum price (bid price) that we are willing to pay per hour for a Spot instance. When our bid price exceeds the Spot price, Amazon EC2 fulfills our request. Note that a Spot instance request is either one-time or persistent. Amazon EC2 automatically resubmits a persistent Spot request after the Spot instance associated with the request is terminated. Our Spot instance request can optionally specify a duration for the Spot instances.*
- ***Spot fleet****—A set of Spot instances that is launched based on criteria that we specify. The Spot fleet selects the Spot pools that meet our needs and launches Spot instances to meet the target capacity for the fleet. The Spot fleet also maintains the target capacity of the fleet over time by launching replacement instances after Spot instances in the fleet are terminated.*
- ***Spot instance interruption****—Amazon EC2 terminates our Spot instance when the Spot price exceeds our bid price or there are no longer any unused EC2 instances. Amazon EC2 marks the Spot instance for termination and provides a Spot instance termination notice, which gives the instance a two-minute warning before it terminates.*
- ***Bid status****—Provides detailed information* about the current state of our Spot bid.

Amazon EC2 Spot instances allow us to bid on spare Amazon EC2 computing capacity. Since Spot instances are often available at a discount compared to On-Demand pricing, we can significantly reduce the cost of running our applications, grow our application's compute capacity and throughput for the same budget, and enable new types of cloud computing applications.

The Spot instance can lower our Amazon EC2 costs significantly. The hourly price for a Spot instance (of each instance type in each Availability Zone) is set by Amazon EC2, and fluctuates depending on the supply of and demand for Spot instances. Our Spot instance runs whenever our bid exceeds the current market price.

Spot instances are a cost-effective choice if we can be flexible about when our applications run and if our applications can be interrupted. For example, Spot instances are well suited for data analysis, batch jobs, background processing, and optional tasks.

The key differences between Spot instances and On-Demand instances are that Spot instances might not start immediately, the hourly price for Spot instances varies based on demand, and Amazon EC2 can terminate an individual Spot instance as the hourly price for or availability of Spot instances changes. One strategy is to launch a core group of On-Demand instances to maintain a minimum level of guaranteed compute resources for our applications, and supplement them with Spot instances when the opportunity arises.

Another strategy is to launch Spot instances with a required duration, which are not interrupted due to changes in the Spot price.

4. *Question 4 Answer the following questions:*
   -Are you getting a IaaS or Paas Service?. Explain? Remember that Java was installed on our machine.

   I think we use both IaaS and Paas, but majority is Paas.
   IaaS refers to online services that abstract user from the detail of infrastructure like physical computing resources, location, data partitioning, scaling, security, backup etc. IaaS-cloud providers supply these resources on-demand from their large pools of equipment installed in data centers.
   So when we store this file on the EC2 instance. In fact we use the service IaaS. We use the storage service on the EC2 instance to store our data. In this sense we can treat that we use the service IaaS.

   PaaS vendors offer a development environment to application developers. The provider typically develops toolkit and standards for development and channels for distribution and payment. In the PaaS models, cloud providers deliver a computing platform, typically including operating system, programming-language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers.

   When we write a java program that calculates the min, max, average, median, and the standard deviation. The environment where we work is exactly in our EC2 instance. So the java code will compile on our instance and return the value to us. The Instance will provide us the OS and the compile environment for java. In this sense, we use the Paas .