Lab: Intro to Storage Services

Cloud computing models

Infrastructure as a service

Offer services that are present in on-premise data center. Servers, storage, and networking hardware, so like VPN are all available as a service in AWS.

Ex: EC2, VPC, EBS.

Platform as a service

AWS helps you manage the underlying infrastructure, hardware and operating systems for you and provide you the application up and running without you having to do anything.

So you can set up a RDS (relational database) without having to worry about setting up a Linux server, and downloading MySQL database. You can just leave the setting up of the database to cloud and then use it after it is up and running

Ex: RDS, EMR, ElasticSearch

Software as a service

Completed product that is run and managed by the service provider. So the cloud provide you the complete the application for you to use, you don't need to worry about how it is being deployed what kind of server it is set up with. You just leave the application managing to the cloud and use it.

Serverless computing

Refer to as function as a service. It allows you to build and run applications and services without thinking about managing the servers. You just have to worry about the actual coding of the business logic that you are implementing without having to worry about provisioning the servers underneath and how you are going to host it.

AWS storage services

1. Simple storage service (S3): Designed to store object files, files that you don't need to worry about a traditional tree hierarchy of path to place the file. It creates a unique identifier for each object file that you store into them for retrieval.

   It is serverless service, you don't need to worry about how it is implemented underneath, what operating system is used
2. Glacier: Cheapest storage for long-term storage. Only used for content that is to be archived.
3. Elastic block store (EBS): Highly available, low latency block storage, designed to be attached to EC2 instances. **Similar to attaching a hard drive to your computer.** Cannot be shared between multiple EC2 instances.
4. Elastic file system (EFS): Network attached storage, allow multiple EC2 instances to mount onto this filesystem.
5. Storage gateway: Enable hybrid storage between on-premise environment and AWS cloud. Cache frequently used data on-premises and storing less frequently used data to AWS.
6. Snowball: Used to migrate large amount of data from on-premise data center to AWS.

AWS VPC

Virtual private cloud is the basis of how to set up your AWS instances. Similar to a traditional network, all of the instances that are launched in AWS will be under a VPC. Public internet will not be allowed to access to those resources without you explicitly permitting it via some kind of routing.

Think of VPC just as a traditional data center network, except it is on the cloud so you don't have to worry about setting it up yourself, it is done for you already.

VPC endpoint is used to allow traffics in and out of the specified VPC, otherwise, say if you launched a S3 bucket in AWS, and you want your EC2 instances to deposit some object files into it. Your EC2 instance is under the VPC while S3 buckets are not, then the EC2 instance have no way of communicating with S3. To allow it, you create a VPC endpoint to allow traffics from VPC out to the AWS S3 instance.

Hybrid storage example

Say you have on-premise site storage and you want to migrate the petabyte of data from your on-premise data center up to AWS cloud. How would you do that? Uploading those files directly using Internet is not going to be viable because you have petabyte of data! That's where snowball can come into play. You would get a device from AWS and you would upload all of your on-premise data into the device and ship it back to AWS. They will then upload all of that data into the AWS cloud say into a S3 bucket directly.

Then in order to keep the keep between on-premise data center and the data in S3 in sync, AWS storage gateway can be used to orchestrate a synchronization mechanism. Data are replicated by storage gateway into S3 buckets, you can have that as a disaster backup.

Revision #1
Created 2023-01-09 00:23:59 UTC by Tamarine
Updated 2023-02-06 18:24:02 UTC by Tamarine