L2 Cloud Architecture

Prof. Dr. M. A. Rouf

Dept. of CSE, DUET, Gazipur

General Architecture of Cloud

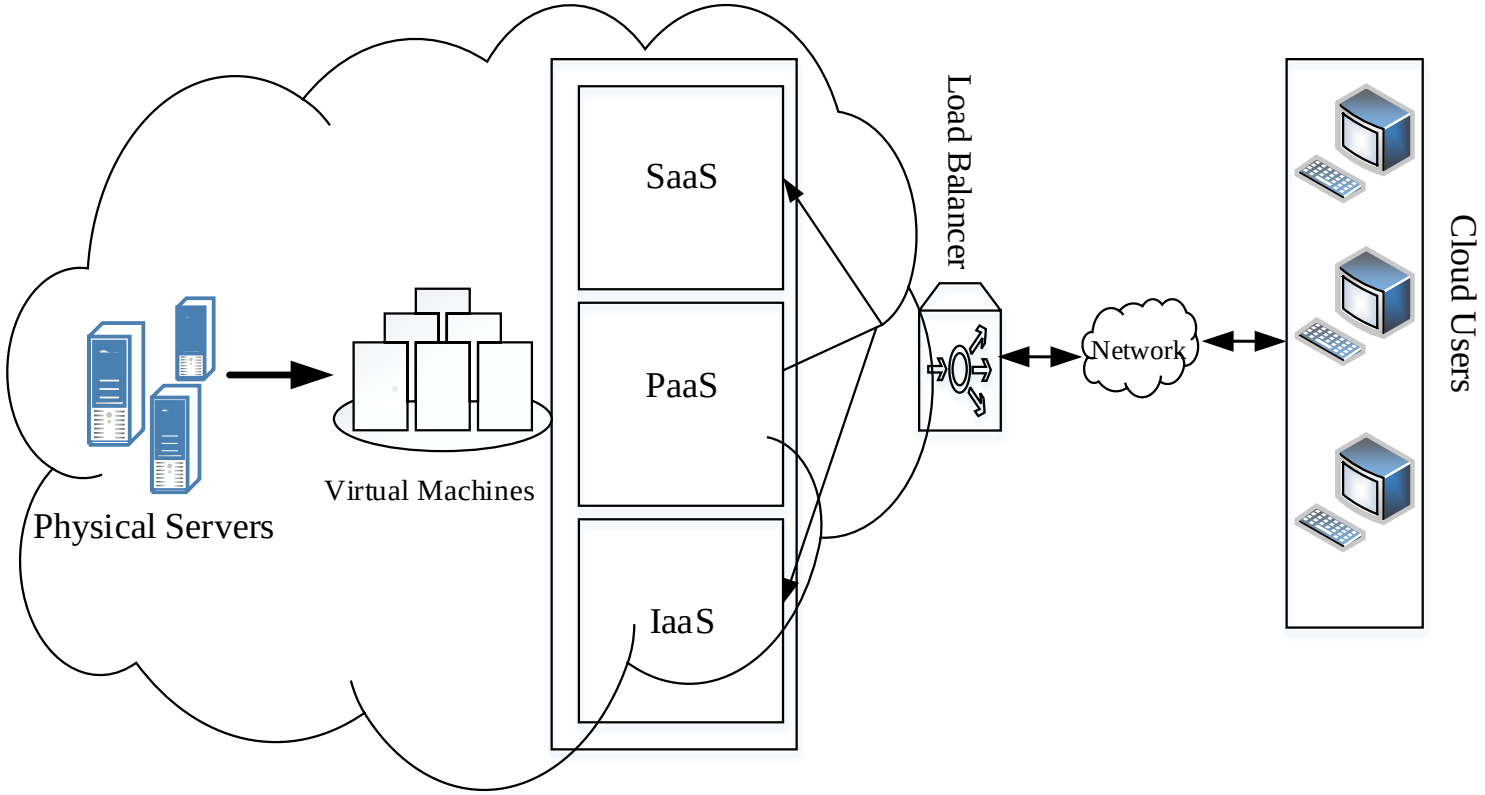


Fig: Basic cloud service layout.

General Architecture of Cloud

- **Software as a Service (SaaS):**
  - In this model, a complete application is delivered on demand to the cloud users.
  - Clients need not invest upfront before using the applications.
  - They use the software subscription based or followed the model Pay-as-you-go, such as Google, Salesforce and Microsoft

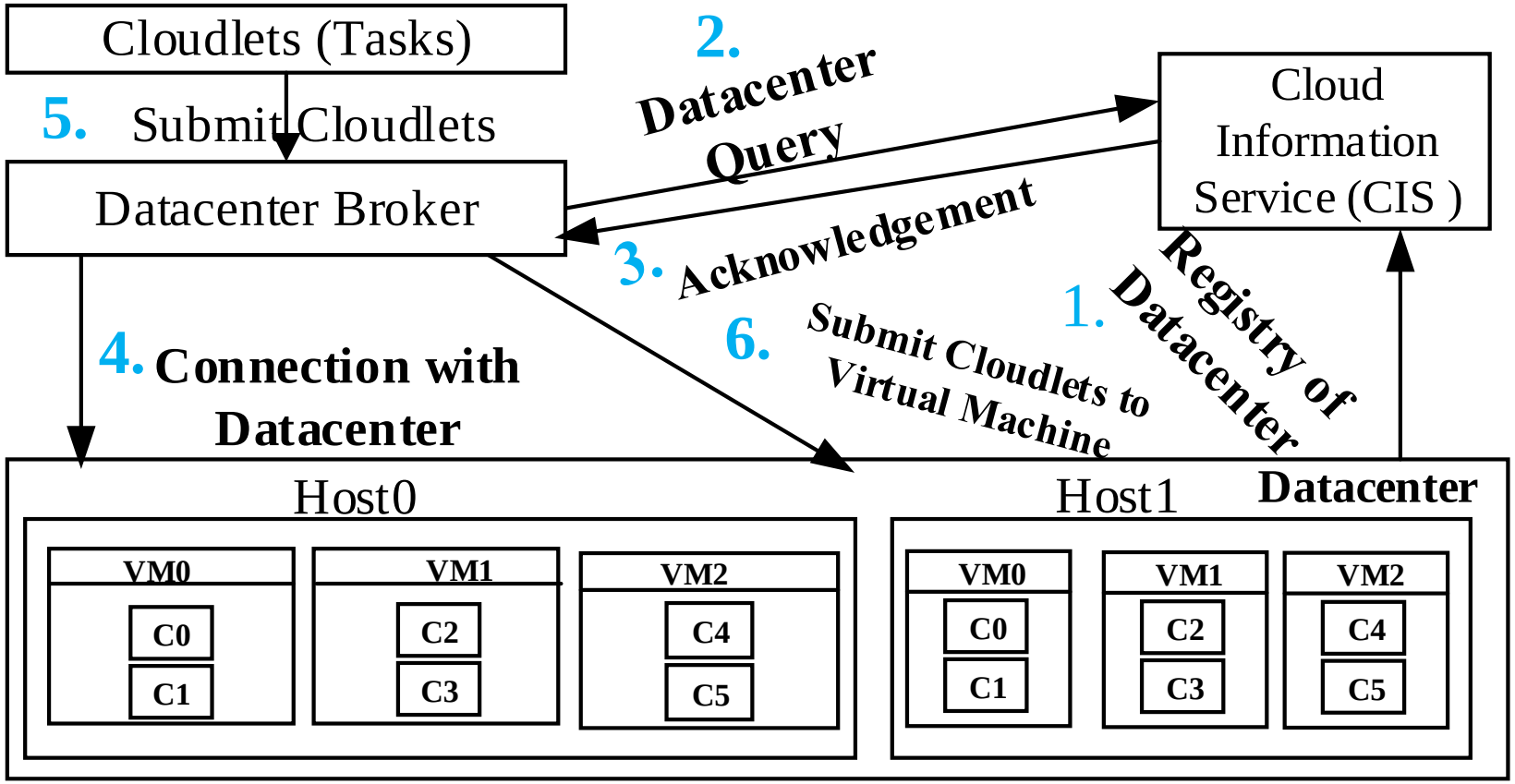
General Architecture of Cloud

- **Platform as a Service (PaaS):**
  - In this model, the integrated development environment tools are provided to develop own business policy.
  - A predefined configuration of operation system and application server are delivered to cloud users. For an example, Force.com and Google's App Engine are providing as platform

General Architecture of Cloud

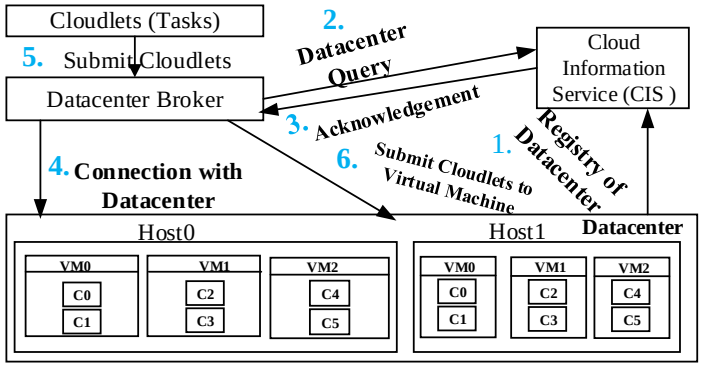
- **Infrastructure as a Service (IaaS):**
  - The virtualization of resources are provided to run the application called the Infrastructure as a Service (IaaS).
  - The resources are virtual server, host, machine, storage, and computing capacity etc.
  - The cloud users deploy their own applications in cloud infrastructure such as Amazon and Go Grid

Basic Components of Cloud Computing



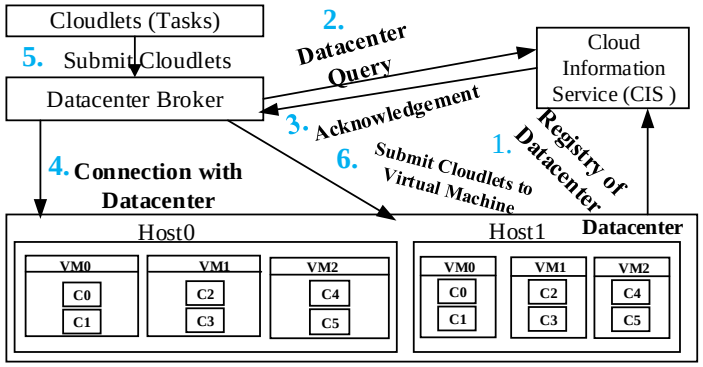
1) **Cloudlets**
2) **Datacenter Broker**
3) **Cloud Information Services (CIS)**
4) **Datacenter**

Basic Components of Cloud Computing



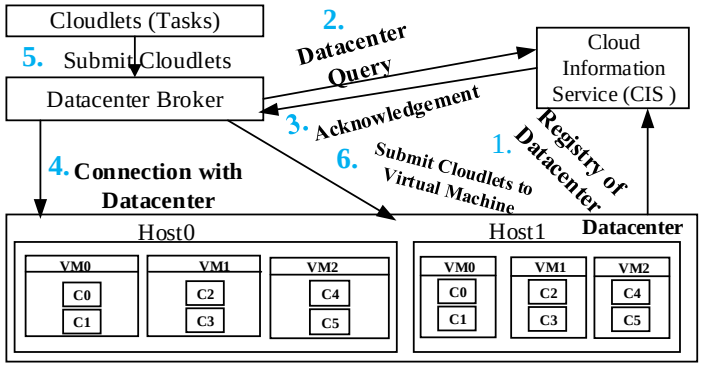
- **Cloudlets :**
  - The cloudlet is an application which consists of million instructions (it is also known as a task such as social networking, content delivery and business application etc.).
- **Datacenter Broker:**
  - The datacenter broker acts as a coordinator between software-as-as-services (SaaS) and cloud providers.
  - The main responsibility of broker collects the available resources and provides quality of service to clients of cloud system.

Basic Components of Cloud Computing



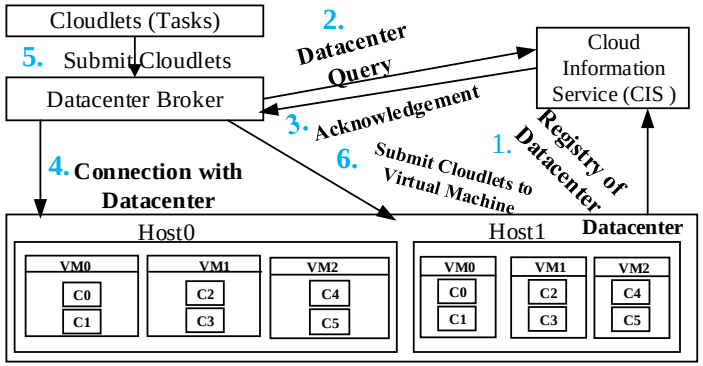
- **Datacenter :**
  - A datacenter is a collection of virtualized hosts, virtual machines, processing elements, virtual networks and virtual storage.
  - A datacenter consists of X86 architecture, operating system, and virtual machine monitor (VMM), host list, memory, bandwidth and storage.
- **Cloud Information Services (CIS)**
  - CIS sends the acknowledgement to the broker about available resources of cloud.

Basic Components of Cloud Computing



- **Host :**
  - A host consists of multiple virtual machines.
  - The parameters of the host are:
    - Processing capacity: Usually measured in million instructions per second (MIPS)
    - Memory size: Megabyte (MB))
    - Storage size: Terabyte (TB)
    - Communication bandwidth: Megabyte per second (Mbps)
- **Virtual Machine (VM):**
  - The VMs are allocated in a host with the best-fit mechanism.

Basic Components of Cloud Computing

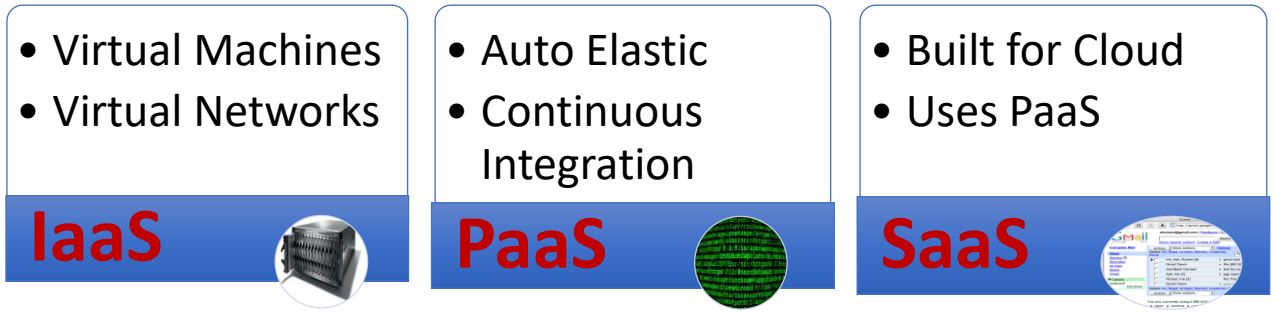


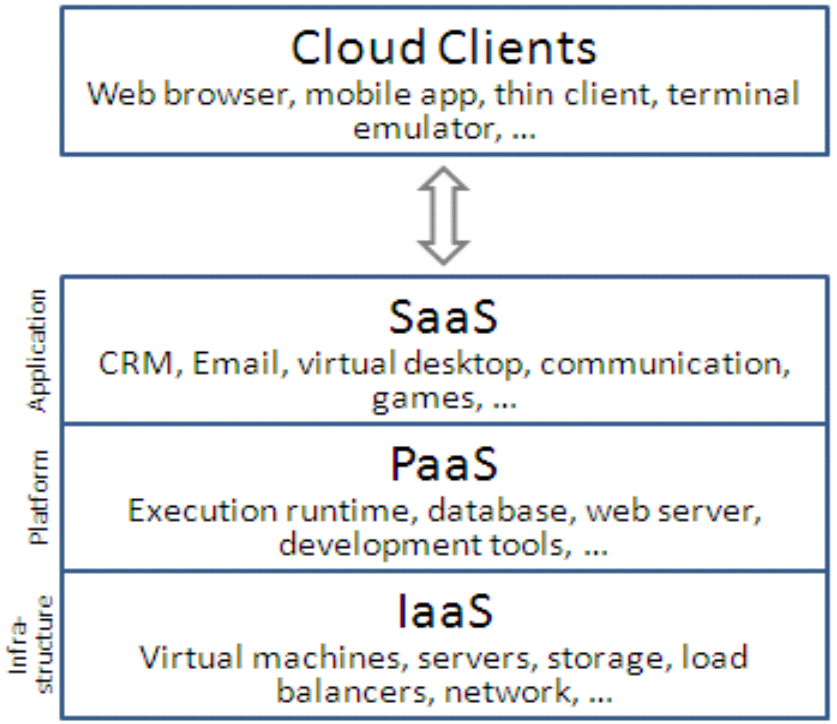
- **Clients :**
  - These are typically the computers, mobile phones and thin browser which are used by the end users.
  - Clients are cloud users that produces cloudlet for utilizing and deploying cloud components/services

Infrastructure as a Service (IaaS)

- IaaS:
  - A category of cloud services which provides capability to provision processing, storage, intra-cloud network connectivity services, and other fundamental computing resources of the cloud infrastructure.

Source- [ITU –Cloud Focus Group]





What is EC2 ?

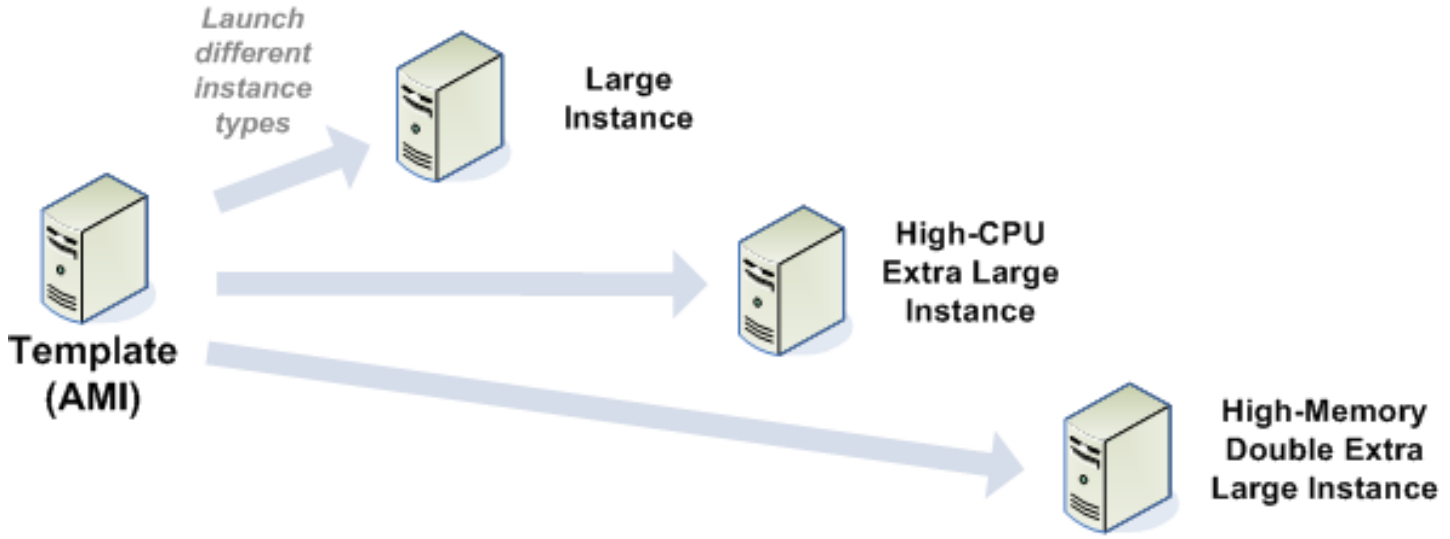
- Amazon Elastic Compute Cloud (EC2) is a web service that provides resizable computing capacity that one uses to build and host different software systems.
- Designed to make web-scale computing easier for developers.
- A user can create, launch, and terminate server instances as needed, paying by the hour for active servers, hence the term "elastic".
  - Provides scalable, pay as-you-go compute capacity
  - Elastic - scales in both direction

EC2 Concepts

- AMI & Instance
- Region & Zones
- Storage
- Networking and Security
- Monitoring
- Auto Scaling
- Load Balancer

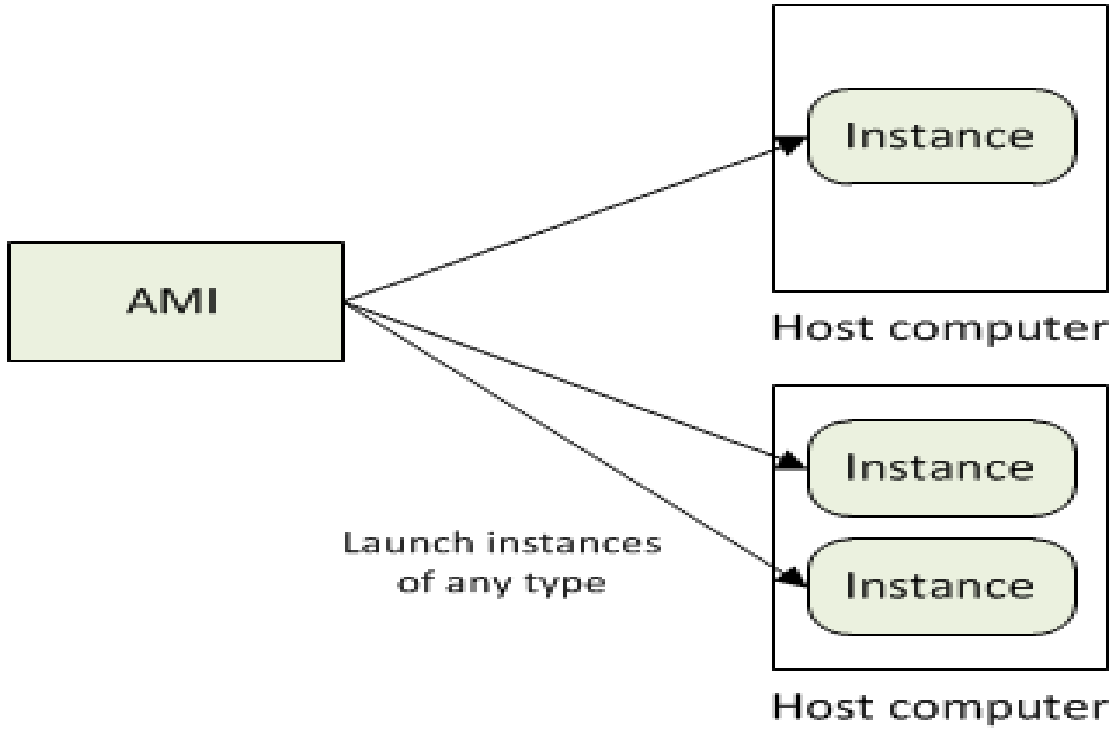
Amazon machine Images (AMI)

- Is an immutable representation of a set of disks that contain an operating system, user applications and/or data.
- From an AMI, one can launch multiple instances, which are running copies of the AMI.



AMI and Instance

- Amazon Machine Image (AMI) is a template for software configuration (Operating System, Application Server, and Applications)
- Instance is a AMI running on virtual servers in the cloud
- Each *instance type* offers different compute and memory facilities



Type	CPU	Memory	Local Storage	Platform	I/O	Name
Small	1 EC2 Compute Unit (1 virtual core with 1 EC2 Compute Unit)	1.7 GB	160 GB instance storage (150 GB plus 10 GB root partition)	32-bit	Moderate	m1.small
Large	4 EC2 Compute Units (2 virtual cores with 2 EC2 Compute Units each)	7.5 GB	850 GB instance storage (2 x 420 GB plus 10 GB root partition)	64-bit	High	m1.large
Extra Large	8 EC2 Compute Units (4 virtual cores with 2 EC2 Compute Units each)	15 GB	1690 GB instance storage (4 x 420 GB plus 10 GB root partition)	64-bit	High	m1.xlarge
Micro	Up to 2 EC2 Compute Units (for short periodic bursts)	613 MB	None (use Amazon EBS volumes for storage)	32-bit or 64-bit	Low	t1.micro
High-CPU Medium	5 EC2 Compute Units (2 virtual cores with 2.5 EC2 Compute Units each)	1.7 GB	350 GB instance storage (340 GB plus 10 GB root partition)	32-bit	Moderate	c1.medium

Region and Zones

- Amazon have data centers in different region across the globe
- An instance can be launched in different regions depending on the need.
  - Closer to specific customer
  - To meet legal or other requirements
- Each region has set of zones
  - Zones are isolated from failure in other zones
  - Inexpensive, low latency connectivity between zones in same region

Storage

- Amazon EC2 provides three type of storage option
  - Amazon EBS
  - Amazon S3
  - Instance Storage

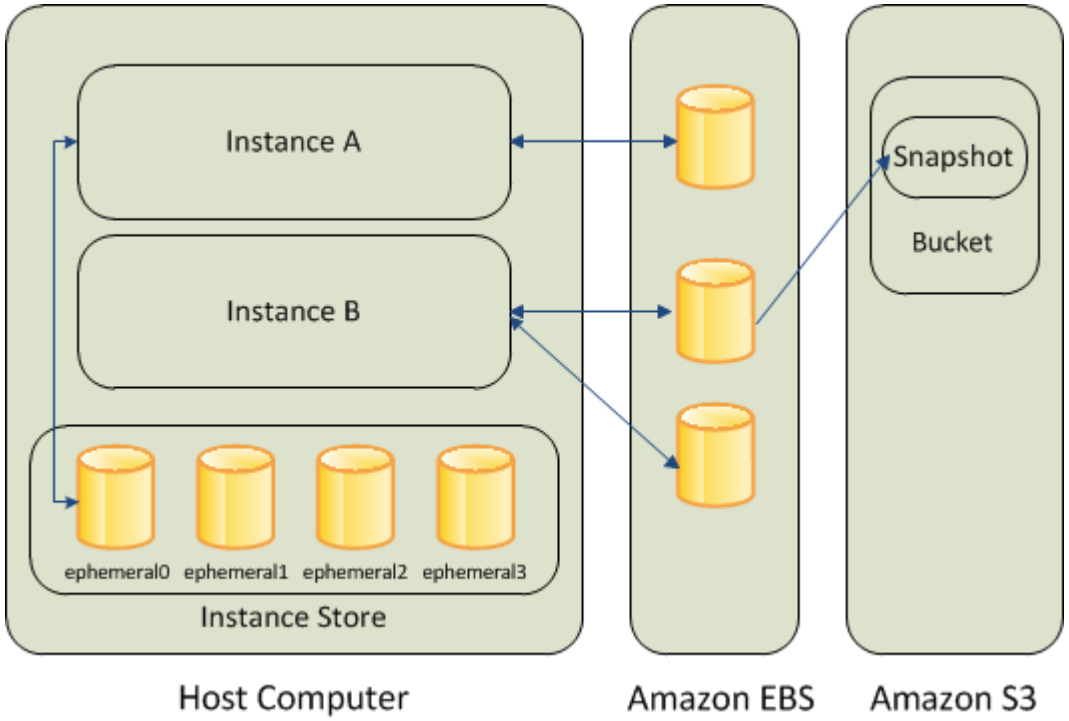
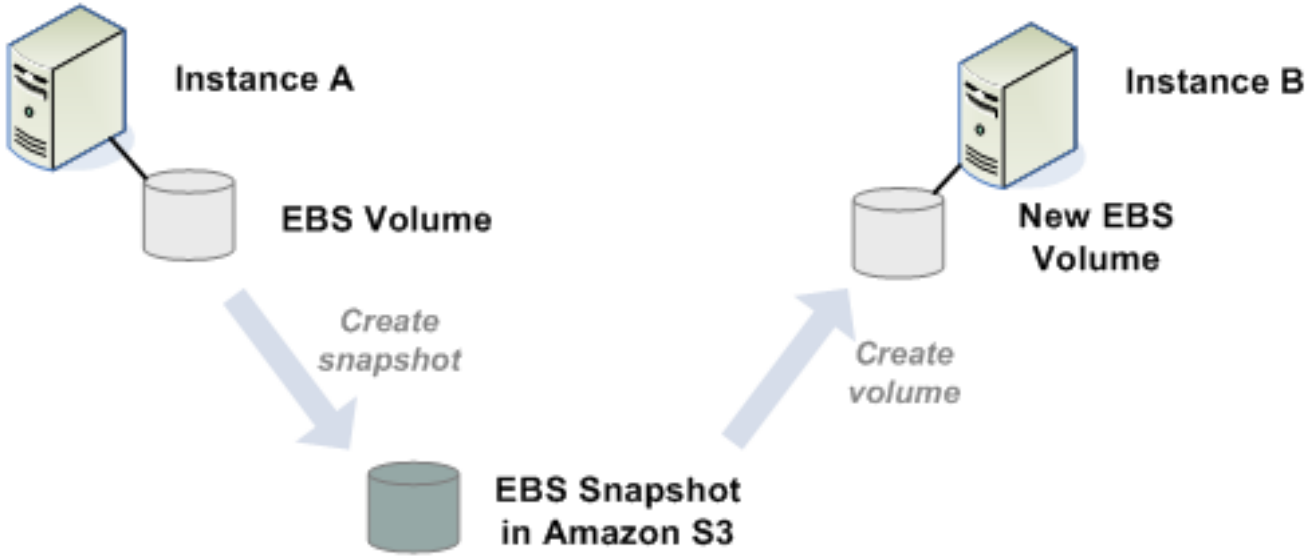


Diagram Source: http://docs.aws.amazon.com

Elastic Block Store(EBS) volume

- An EBS volume is a read/write disk that can be created by an AMI and mounted by an instance.
- Volumes are suited for applications that require a database, a file system, or access to raw block-level storage.



Amazon S3

- S3 = Simple storage Service
- A SOA – Service Oriented Architecture which provides online storage using web services.
- Allows read, write and delete permissions on objects.
- Uses REST and SOAP protocols for messaging.

Amazon SimpleDB

- Amazon SimpleDB is a highly available, flexible, and scalable non-relational data store that offloads the work of database administration.
- Creates and manages multiple geographically distributed replicas of your data automatically to enable high availability and data durability.
- The service charges you only for the resources actually consumed in storing your data and serving your requests.

Networking and Security

- Instances can be launched on one of the two platforms
  - EC2-Classic
  - EC2-VPC
- Each instance launched is assigned two addresses a private address and a public IP address.
  - A replacement instance has a different public IP address.
- Instance IP address is dynamic.
  - new IP address is assigned every time instance is launched
- Amazon EC2 offers Elastic IP addresses (static IP addresses) for dynamic cloud computing.
  - Remap the Elastic IP to new instance to mask failure
  - Separate pool for EC2-Classic and VPC
- Security Groups to access control to instance