CS 695: Virtualization and Cloud Computing

# Lecture 12: Introduction to Cloud Applications

Mythili Vutukuru

IIT Bombay

Spring 2021

Cloud applications

- Users access cloud applications over the internet
  - Running on public or private clouds (not running on "baremetal")
- Cloud apps designed as multi-tier apps hosted on VMs/containers on servers
- Auto-scaling, fault tolerance, optimal placement on servers, network config
  - Functionality available from cloud management software (e.g., OpenStack)
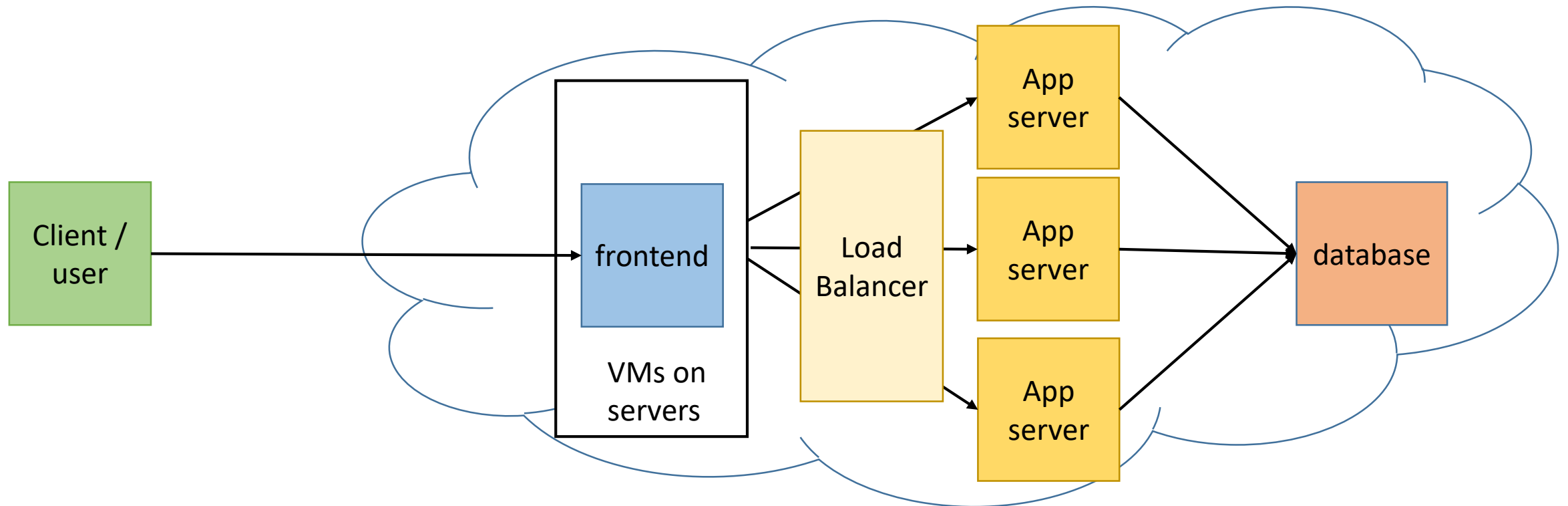
How is the cloud used? (1)

- Many ways to expose the cloud to users/developers

- Infrastructure as a Service (IaaS): Cloud provider exposes cloud infrastructure (VMs, storage, etc.) directly to cloud users.
  - Developer installs application of choice on the VMs, storage
  - Provider handles VM lifecycle management, placement of VMs on servers, migration in case of server failures, elastic scaling of VMs under load etc.
  - Hassle-free management of compute and storage
  - Flexible usage, pay-as-you-go model

- Examples
  - Amazon EC2: exposes virtual servers (VMs) to users
  - Amazon Simple Storage Service (S3): exposes basic cloud storage
  - Microsoft Azure, Google Cloud, Digital Ocean etc. also have similar offerings

How is the cloud used? (2)

- Platform as a Service (Paas): Cloud provider sets up software on the VMs and exposes the infrastructure + software bundle to users
  - Provider exposes APIs using which developers build cloud apps
  - Provisioning VMs and storage for the app, maintaining application software (OS, app servers), scaling application tiers etc. handled by the cloud provider
  - Developer only focuses on application logic
  - Easier to build cloud apps/software but lesser flexibility

- Examples
  - Amazon Elastic Beanstalk: provides APIs to build cloud apps, automatically deploys them on EC2 instances
  - Google Cloud App Engine, Heroku are other examples

How is the cloud used? (3)

- **Software as a Service (Saas):** Cloud provider installs useful software on the cloud and let users access the software directly
  - Users have no visibility into the cloud infrastructure
  - Users access the app software over the internet (nothing installed locally)
  - Application servers not hosted directly by the enterprise but on the cloud
- Examples
  - Google docs, Gmail, Office 365
  - Most internet-scale software companies (Zoom, Dropbox, Netflix, Prime Video, Slack, Spotify, Whatsapp) host their software on the cloud and not on their premises, for scalability, fault tolerance, and general ease of management

How is the cloud used? (4)

- Function as a Service (FaaS): Provider exposes only ability to run functions on cloud data
  - Also called serverless computing
  - Developers write functions to execute on certain triggers, e.g., function to convert a video to lower resolution is triggered whenever a video is uploaded
  - Provider instantiates VMs/containers to run the function when required
  - Users only pay when function runs, not for the infrastructure, so cheaper when functions are run sporadically
- Example: AWS Lambda, Google Cloud Functions etc.

Compute and data storage options

- Cloud compute runs on VMs or containers
  - Studied in course so far

- Cloud storage options
  - Relational databases (RDBMS) with strong consistency (ACID properties). Traditional storage option, but harder to scale, lower performance.
  - No-SQL data stores map key to unstructured "blob" of data. Simpler get/put interface using single key, no transaction support. Some form of weaker consistency. Tradeoff stronger guarantees for higher performance.
  - Semi-structured data stores map key to value which has certain column families, but not fixed schema. Useful for certain apps.
  - Other storage frameworks customized for application needs

Cloud storage systems: examples from real-life

- Dynamo: high-performance No-SQL key-value store from Amazon
  - Trade off consistency guarantees for high availability and scalability
- Bigtable: semi-structured data store used at Google
  - Built over two other systems from Google: Chubby (a distributed locking/consensus system) and the Google File System (a distributed file system)
- Haystack: cloud storage system for efficiently storing and retrieving photos at Facebook
- Memcache: distributed cache used at Facebook to cache queries from database servers