DSC204A: Scaleable Data Systems Spring 2025

13: Networking Fundamentals

Instructor: Umesh Bellur *Scribes: Caden Stewart and Conan Minihan*

Cloud Service Models

There is a gradient of services depending on how much is managed by the customer vs. the provider:

Private Cloud → IaaS → PaaS → FaaS → SaaS

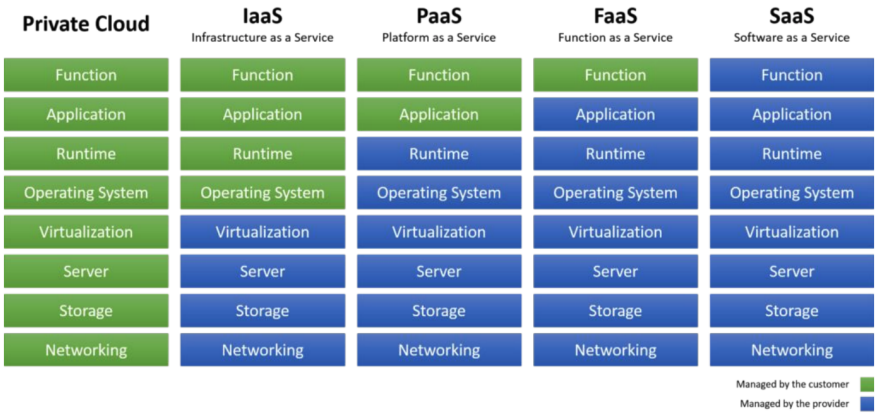

Figure 1: Cloud service models by Stack

- **IaaS (Infrastructure-as-a-Service):** Provides virtualized computing resources over the internet.
  - VMs (Virtual Machines): Heavyweight, each requires a full OS boot (1.5 minutes).
  - Containers: Lightweight, can exist within a VM, faster boot time (500ms), share OS kernel but have isolated filesystems.
  - VM and container images are file-based snapshots of environments.
  - Security concerns arise when VMs share physical machines; isolation is preferred.
- **PaaS (Platform-as-a-Service):**
  - Facilitates application development and deployment without managing infrastructure (e.g., Oracle, Heroku)
  - User has control over deployed icatiopns but does not manage/control cloud infrastructure
  - Not helpful when needing portability for applications
- **FaaS (Function-as-a-Service):** Allows execution of discrete functions in response to triggers (e.g., AWS Lambda).

  - Stateless execution
  - Ideal for microservices and event-driven architecture
- **SaaS (Software-as-a-Service):**
  - Delivers applications over the internet (e.g., Gmail, Dropbox).
  - User does not manage cloud infrastructure
  - Not useful for real-time applications or when data is hosted externally

AWS EC2 Pricing Models

- **On-Demand:** Pay as you go — like buying a plane ticket last minute. (Least Commitment)
- **Reserved:** Pre-purchase resources — like buying tickets in advance. (Best Long-term)
- **Spot:** Use spare capacity — like flying standby. (Biggest Savings)
- **Dedicated:** Physical isolation — like owning the entire plane. (Most Expensive)

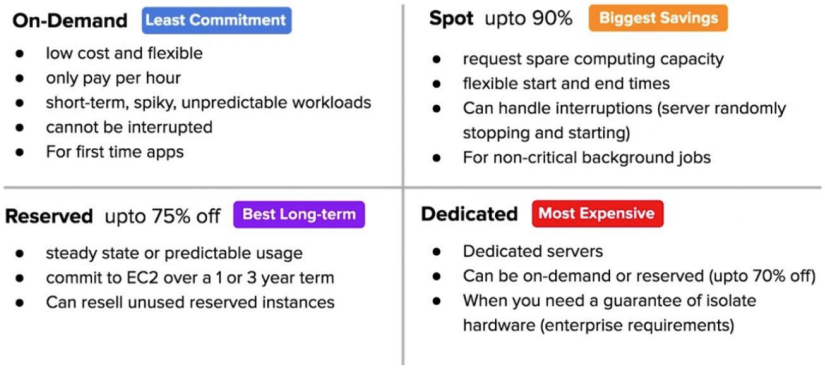

Figure 2: AWS EC2 Model Tradeoffs

Elasticity and Auto Scaling

- Scale out when demand spikes (e.g., traffic surge from a news event).
- Scale in when demand drops to save costs.

Cloud, Parallel & Distributed Systems

Function-as-a-Service (FaaS)

- Enables breaking an application into microservices or small workflows.
- Balancing granularity and overhead is key: too many functions increase overhead; too few reduce modularity.
- Dominant model in modern serverless architecture.

Networked Systems

What Happens in a Search Query

- Devices resolve domain names through DNS.
- IP addresses are used to route requests to servers across networks.
- A search request interacts with:
  - DNS servers
  - Load balancers
  - Ad servers
  - Datacenters

What's the Internet: "nuts and bolts" view

- Network of networks (ISPs, access networks, backbone).
- Includes hosts (PCs, phones), routers, and communication links (fiber, copper, radio, satellite).

What's the Internet: a service View:

- Supports applications like WWW, email, gaming, etc.
- Services can be connectionless (e.g., UDP) or connection-oriented (e.g., TCP).

A Closer Look at Network Structure:

- **Edge:** Hosts and applications
- **Core:** Routers and backbone
- **Access networks:** Local ISPs, physical media

Network Hardware

Network Addresses

- Devices have at least two network cards (wired and wireless).
- Each card has:
  - **MAC Address:** Hardware-level identifier (48-bit).
  - **IP Address:** Network-level identifier (assigned dynamically or statically) (32-bit).

Network Identifier

- Processes communicate over ports.
- A connection is a channel between two endpoints defined by IP address + port number.