IAM, Billing & Cost, Security

Thao Huy Vu

Maharishi International University - Fairfield, Iowa

All rights reserved. No part of this slide presentation may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying or recording, or by any information storage and retrieval system, without permission in writing from Maharishi International University (MIU).

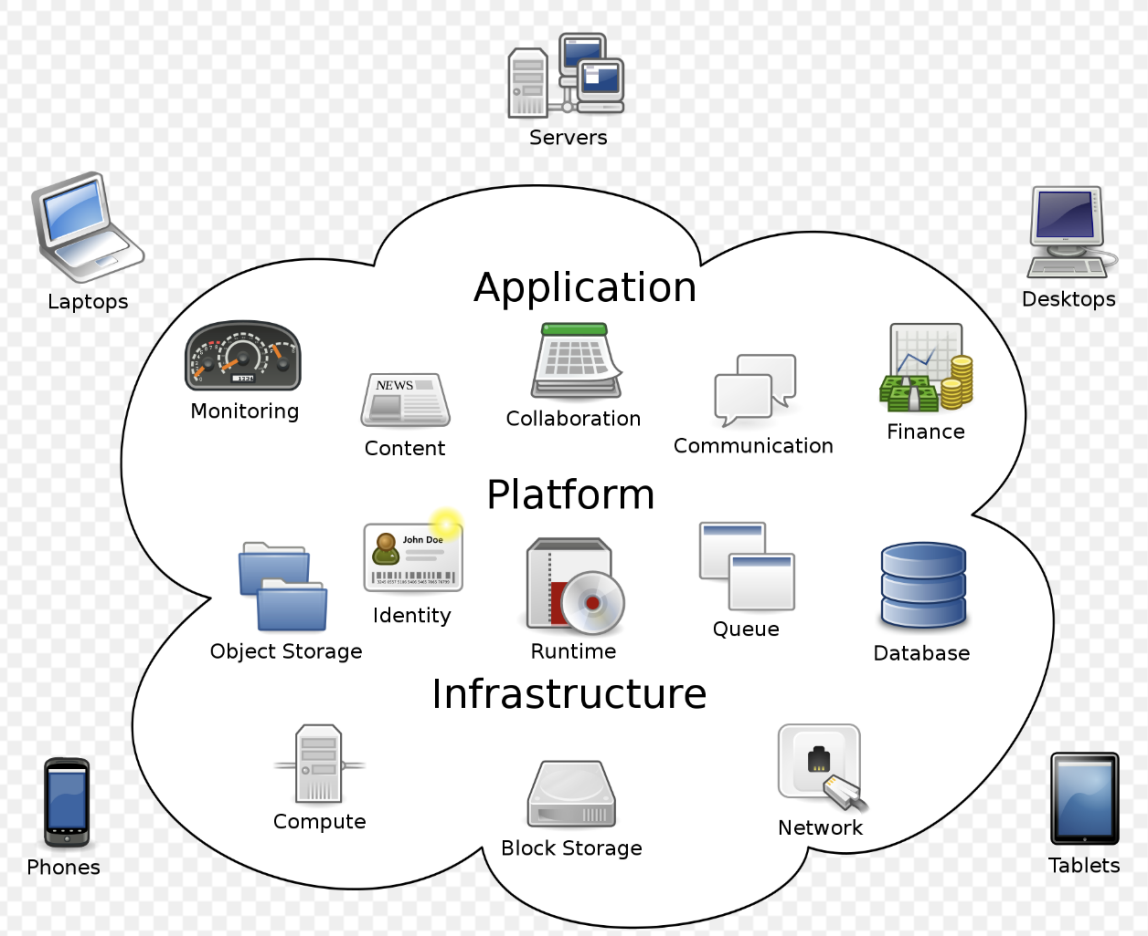
Agenda

- AWS Overview
- IAM
- Billing & Cost
- Security

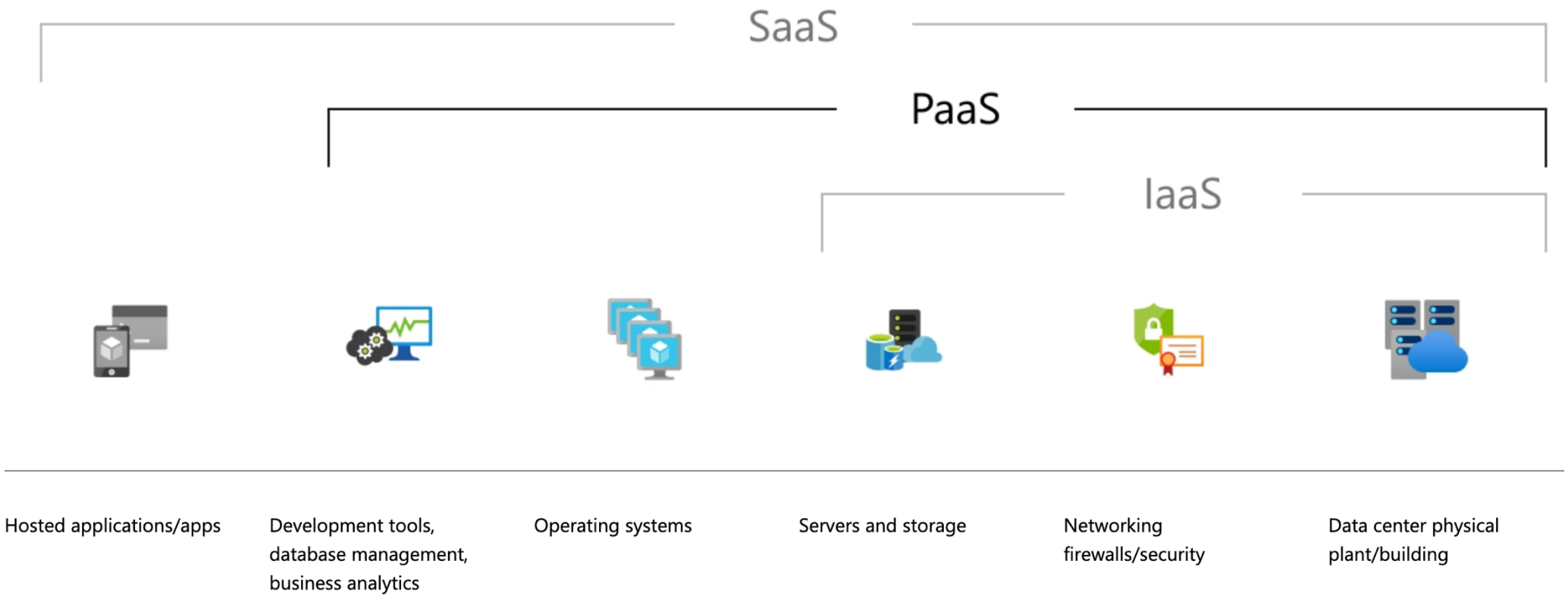
Cloud Computing

- The group of networked elements providing services does not need to be addressed and managed by users.

https://en.wikipedia.org/wiki/Cloud_computing



Cloud Computing model



https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-paas/

Cloud Computing

- Security
- Scalability and flexibility
- Cost Efficiency
- Ease of Access
- Disaster Recovery and Data Loss Prevention
- Automatic Software updates

Cloud Computing

- Major players
  - **AWS**
  - Microsoft Azure
  - Google Cloud Platform (GCP)
  - IBM Cloud
  - Oracle Cloud
  - Salesforce

AWS

- Amazon Web Service.
- The largest Cloud Computing provider.
- Provide a variety of scalable computing resources and services that can be used to build, deploy, manage applications on a global network.

Available Zones (AZ) & regions

- A region is a geographically distinct area, completely isolated.
- AZs are datacenters within a region.
- Each AZ is independent of other AZs

AWS Global Map

36 launched Regions
each with multiple Availability Zones

114 Availability Zones

700+ CloudFront POPs
and 13 Regional edge caches

AWS Global Infrastructure Map

The AWS Cloud spans 114 Availability Zones within 36 geographic regions, with announced plans for 12 more Availability Zones and four more AWS Regions in New Zealand, the Kingdom of Saudi Arabia, Taiwan, and the AWS European Sovereign Cloud.

Compute
- EC2
- Lightsail
- Lambda
- Batch
- Elastic Beanstalk
- Serverless Application Repository
- AWS Outposts
- EC2 Image Builder
- AWS App Runner

Containers
- Elastic Container Registry
- Elastic Container Service
- Elastic Kubernetes Service
- Red Hat OpenShift Service on AWS

Storage
- S3
- EFS
- FSx
- S3 Glacier
- Storage Gateway
- AWS Backup

Customer Enablement
- AWS IQ
- Support
- Managed Services
- Activate for Startups

Robotics
- AWS RoboMaker

Blockchain
- Amazon Managed Blockchain

Satellite
- Ground Station

Quantum Technologies
- Amazon Braket

Management & Governance
- AWS Organizations
- CloudWatch
- AWS Auto Scaling
- CloudFormation
- CloudTrail

Machine Learning
- Amazon SageMaker
- Amazon Augmented AI
- Amazon CodeGuru
- Amazon DevOps Guru
- Amazon Comprehend
- Amazon Forecast
- Amazon Fraud Detector
- Amazon Kendra
- Amazon Lex
- Amazon Personalize
- Amazon Polly
- Amazon Rekognition
- Amazon Textract
- Amazon Transcribe
- Amazon Translate
- AWS DeepComposer
- AWS DeepLens
- AWS DeepRacer
- AWS Panorama
- Amazon Monitron
- Amazon HealthLake
- Amazon Lookout for Vision
- Amazon Lookout for Equipment

AWS Cost Management
- AWS Cost Explorer
- AWS Budgets
- AWS Marketplace Subscriptions
- AWS Application Cost Profiler

Front-end Web & Mobile
- AWS Amplify
- Mobile Hub
- AWS AppSync
- Device Farm
- Amazon Location Service

AR & VR
- Amazon Sumerian

Application Integration
- Step Functions
- Amazon AppFlow
- Amazon EventBridge
- Amazon MQ
- Simple Notification Service
- Simple Queue Service
- SWF

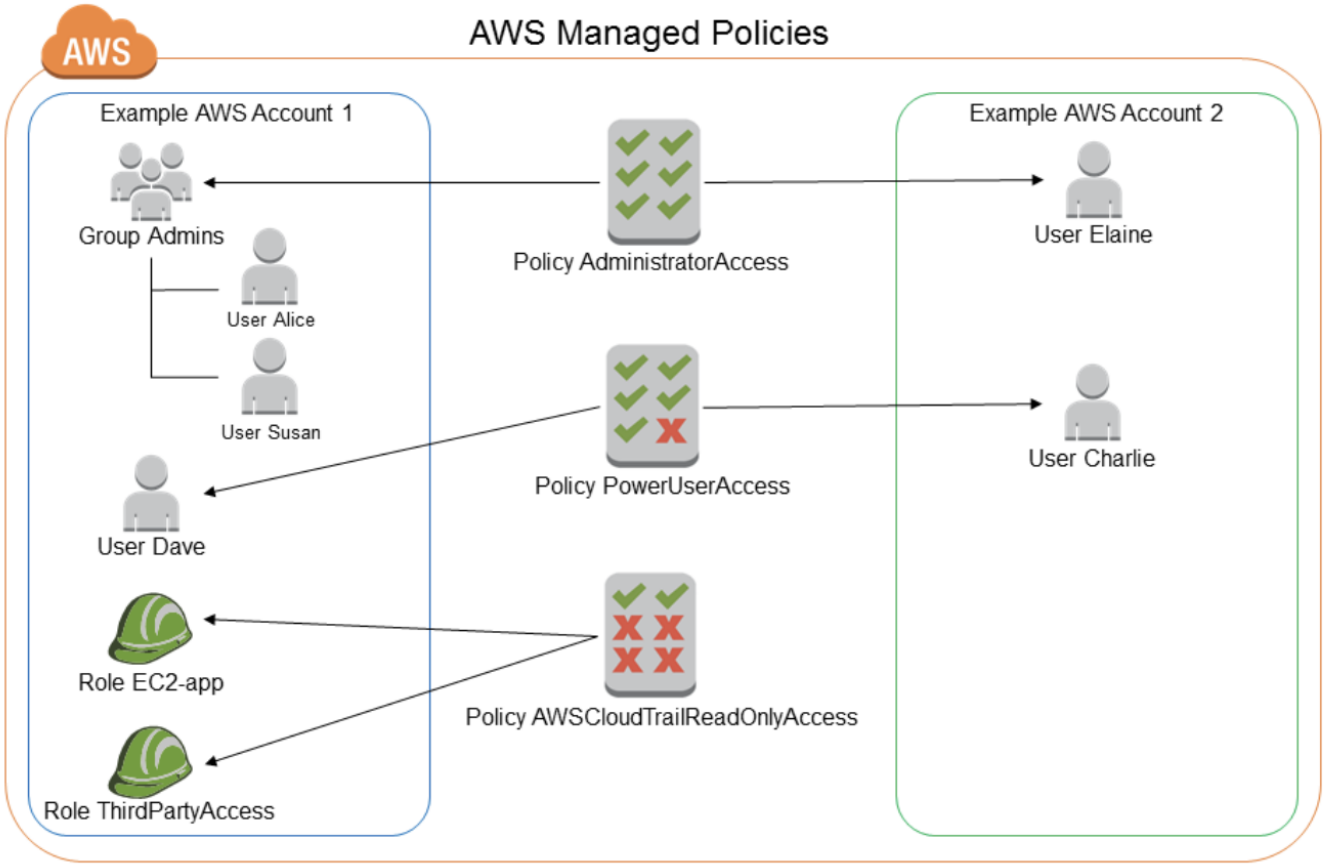
IAM - Introduction

- AWS Identity and Access Management (IAM)
- Manage access to AWS services and resources **securely**
- Create and manage AWS users
- Allow/deny access to AWS resources
- Ensure only **authenticated** and **authorized** users can access resources

IAM – Core components

- Users
- Group
- Role
- Policy

IAM - Examples



Users and Groups

- Users
  - An entity represents a **person** or **service** that interacts with AWS resources.
  - Provide individual **credentials** for accessing AWS services, ensuring accountability and traceability.
  - Each has a name, credentials (passwords, access keys), and permissions assigned through policies.
- Groups
  - A **collection** of users with specific permissions.
  - Manage permissions at a group level rather than at the individual user level.
  - Groups have policies attached to them which apply to all users in the group.

Policy

- A policy is a JSON document that defines permissions to allow or deny actions on AWS resources.
- Control Access: Ensure users and roles have the appropriate access to AWS resources.
- Granular Control: Specify exactly what actions are allowed or denied.

Policy

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "S3ReadOnlyAccess",
      "Effect": "Allow",
      "Action": [
        "s3:GetObject",
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::your-bucket-name",
        "arn:aws:s3:::your-bucket-name/*"
      ]
    }
  ]
}
```

Policy

- **Identity-based** policies – Attach policies to IAM identities (users, groups, or roles)
  - Managed Policies:
    - AWS Managed Policies: Predefined policies created and managed by AWS
    - Customer Managed Policies: Created and managed by users.
    - Shared
  - Inline policies:
    - Attach directly to a single identity and belong to this identity.
    - Not shared.
- **Resource-based** policies:
  - Attach directly to a single resource(such as S3, SQS, SNS) and belong to this resource.
  - Have a **principal** element that defines who can the actions defined.

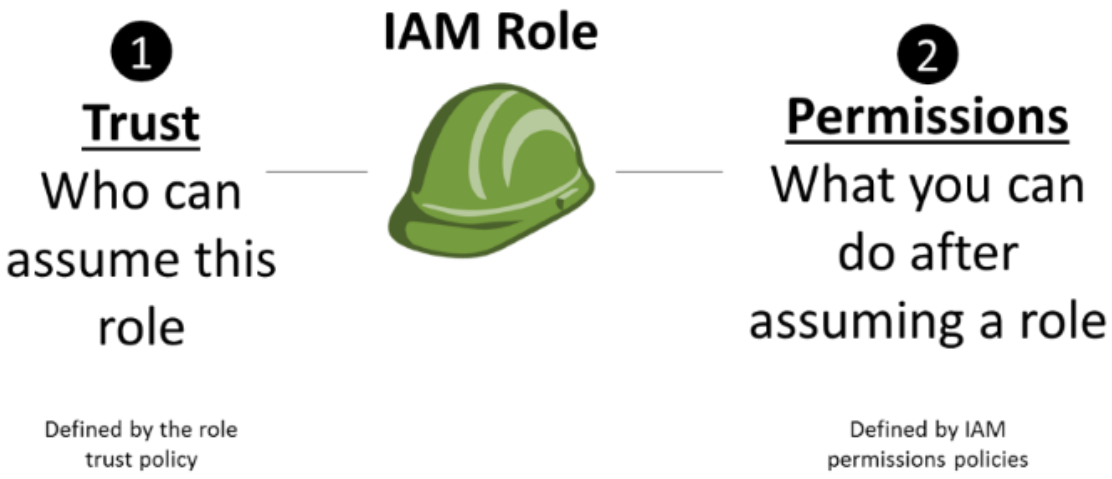
IAM roles

- An **IAM Role** is a set of **permissions** in AWS that **any trusted entity** (like a user, EC2 instance, Lambda function, or even another AWS account) can **assume temporarily** to perform actions.
- Components:
  - Trust Policy: Defines who is allowed to assume the role.
  - Permission Policy: Defines what actions and resources the role has access to.
- Common Use Cases:
  - EC2 Instance Roles: Allow EC2 instances to access AWS services.
  - Cross-Account Access: Allow users from one AWS account to access resources in another account.
  - Service Roles: Allow AWS services like Lambda or ECS to access other AWS services.

IAM role

An IAM user and role have one thing in common, permission policies.

When the service in the trust policy assumes the role, AWS STS (Security Token Service) returns **temporary** tokens (access key id, secret access key, and **session token**), and those tokens are rotated automatically.



Billing & Cost

- AWS **Billing and Cost** refer to how AWS tracks, charges, and manages expenses for the cloud services you use.
- **Billing Dashboard** — View invoices and payments.
- **Cost Explorer** — Analyze and forecast costs.
- **AWS Budgets** — Set spending limits and receive alerts.
- **Cost Anomaly Detection** — Get notified about unexpected cost spikes.

Billing Alerts (CloudWatch Alarms for Billing)

- Set up using **CloudWatch**.
- Alerts you when your total AWS charges exceed a specified threshold.
- Works on an **account-wide** level.
- Typically used for basic cost monitoring.
- Limited in flexibility (e.g., can't track specific service or linked account costs).

AWS Budgets

- More advanced and customizable.
- Allows you to set budgets for **specific services, linked accounts, regions, or tags**.
- Can track actual costs, **forecasted costs**, and usage.
- Supports different types of budgets:
  - **Cost Budget** (total spend)
  - **Usage Budget** (specific resource usage)
  - **Reservation Budget** (Reserved Instances and Savings Plans utilization)
- Can send alerts when costs approach or exceed the budget.
- Supports multiple recipients and integration with AWS Chatbot, SNS, or emails.

Billing and Cost

- Set **Billing Alerts**
- Set B**udget Alerts**.
- Use **Cost Allocation Tags**.
- Review costs regularly in **Cost Explorer**.
- Right-size your resources to avoid over-provisioning.

Regular AWS accounts

- Homepage: https://aws.amazon.com/
- Full access to all services without limitation.
- Long-term access
- Production-level project
- Practical experience: Gain industry-like exposure.
- Credit/Debit card required.
- Free tier available: Up to 12 months for new accounts.
- Caution: If your usage exceeds the free tier limits, you may incur unexpected charges reflected in your credit card. You are solely responsible for this charge.

AWS Free Tier

- **Amazon EC2:**
  - 750 hours per month of Linux or Windows t2.micro/t3.micro instances (for the first 12 months).
  - 750 hours of public IPv4 address usage per month.
- **Amazon S3:**
  - 5 GB of storage in the S3 Standard storage tier (for the first 12 months).
  - 2000 PUT, POST, COPY, or LIST requests per month.
  - 20,000 GET requests per month.
  - 15 GB of data transfer out per month.
- **Other Services:**
  - **AWS Lambda:** 1 million free requests per month.
  - **Amazon RDS:** 750 hours of a t2.micro instance.
  - **Amazon Simple Email Service (SES):** Up to 3,000 message charges free each month for the first 12 months.
  - **Amazon DynamoDB:** First 25 GB of storage and first 10 custom CloudWatch metrics are free.
  - **Amazon OpenSearch Service:** 750 hours per month of a single-AZ t2.small.search or t3.small.search instance.

AWS Account

- Best practices to ensure cost-efficiency
  - Frequently monitor your resource usage.
  - Setup billing alert.
  - Check the bill daily.
  - Shutdown all unused instances/resources:
    - EC2 (Elastic Compute Cloud).
    - RDS (Relational Database Service), Aurora
    - ECS/EKS (Elastic Container Service/Elastic Kubernetes Service).
    - EBS (Elastic Block Store).
    - EFS (Elastic File System).
    - Unused Elastic IP addresses (EIP)
    - DocumentDB
    - DynamoDB.
    - CloudFront.
    - AWS Lambda.
    - SNS/SQS
    - VPC: NAT Gateway, Data transfer, VPN connections. You can keep the default VPC, which is already connected to the internet, so there is no need to create a new one when launching EC2 or other instances. Additionally, the default VPC does not incur significant cost.
    - S3: Particularly important when storing vast amounts of data.

Best practices to secure AWS account and apps

1. Enable Multi-Factor Authentication (MFA) for users.
2. Use a user for deployment.
3. Never use the root account for deployment.
4. Disable the access keys of the root account.
5. Least Privilege Principle

Reference

- AWS: https://docs.aws.amazon.com
- ChatGPT: https://chatgpt.com
- Google AI: https://gemini.google.com
- Practical Tutorials: https://thaovu.org