Cloud Computing

Session 6 - IaaS

BITS
Pilani

IaaS

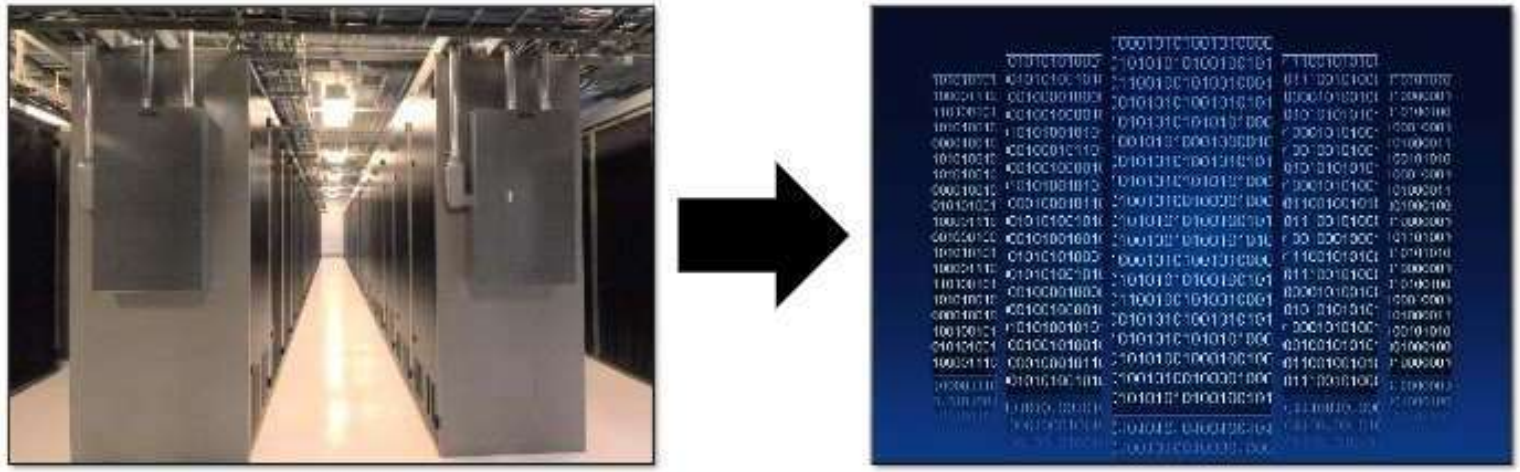
Really, what is iaas???

Cloud computing defined

- Cloud computing is the **on-demand** delivery of compute power, database, storage, applications, and other IT resources via the **internet** with **pay-as- you-go** pricing.

Infrastructure as a Software

- Cloud computing enables you to stop thinking of your infrastructure as hardware, and instead think of(and use) it as software.



Traditional computing model

- Infrastructure as hardware
- Hardware solutions:
  - Require space, staff, physical security, planning, capital expenditure
  - Have a long hardware procurement cycle
  - Require you to provision capacity by guessing theoretical maximum peaks

Cloud computing model

- Infrastructure as software
- Software solutions:
  - Are flexible
  - Can change more quickly, easily, and cost-effectively than hardware solutions
  - Eliminate the undifferentiated heavy-lifting tasks

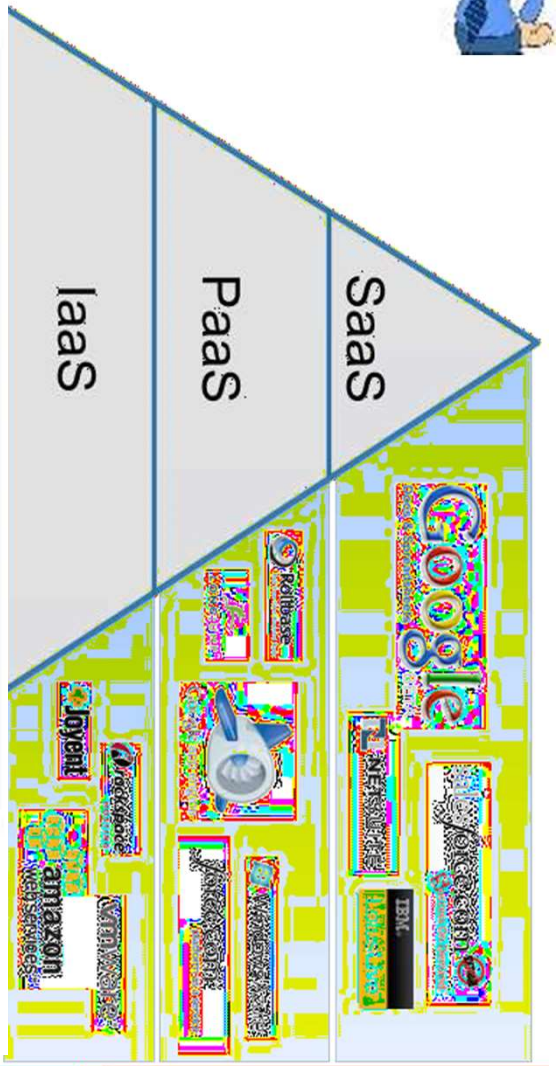
heard of 3 models of Cloud Computing?

Yes, Yes, IaaS, PaaS and SaaS

OnPremises	IaaS	PaaS	SaaS
Applications	Applications	Applications	Applications
Data	Data	Data	Data
Middleware	Middleware	Middleware	Middleware
Operating System	Operating System	Operating System	Operating System
Virtualization	Virtualization	Virtualization	Virtualization
Server	Server	Server	Server
Storage	Storage	Storage	Storage
Network	Network	Network	Network

Own Responsibility

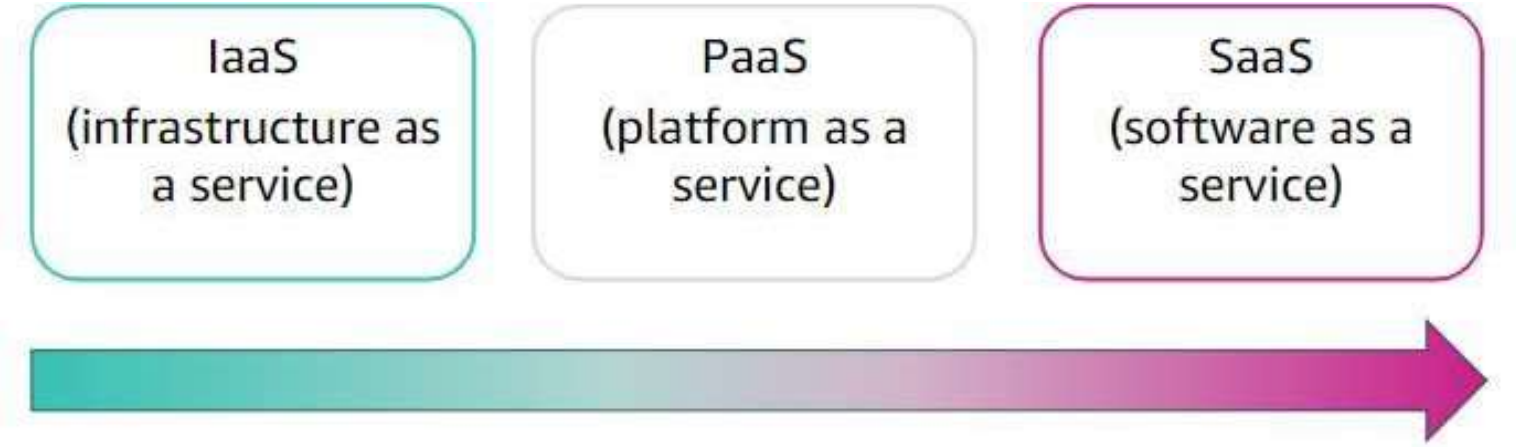
Cloud Provider Responsibility



Cloud service models

- More control over IT resources

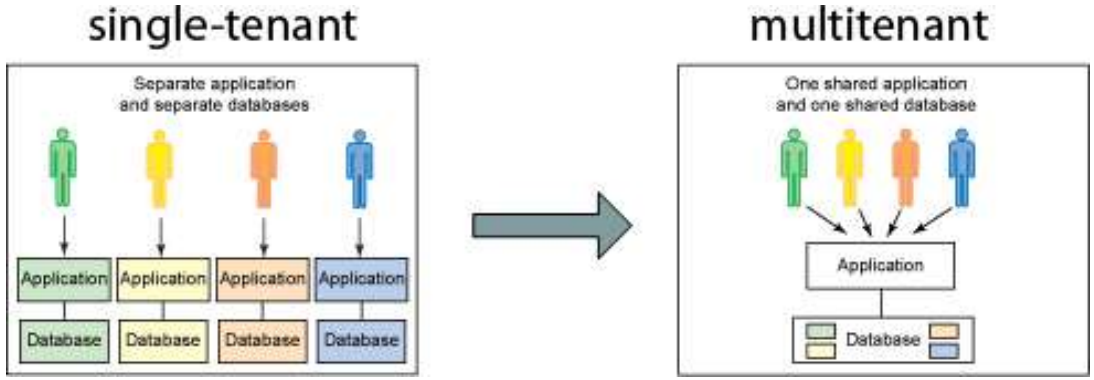
- Less control over IT resources



Key concepts of IaaS

- **Cloud bursting**: The process of off-loading tasks to the cloud during times when the most compute resources are needed

- **Multi-tenant computing**



- **Resource pooling**: **Pooling** is a resource management term that refers to the grouping together of resources (compute(cpu), network(bandwidth), storage) for the purposes of **maximizing advantage** and/or **minimizing risk** to the users
- Hypervisor

Two primary facets that make IaaS special

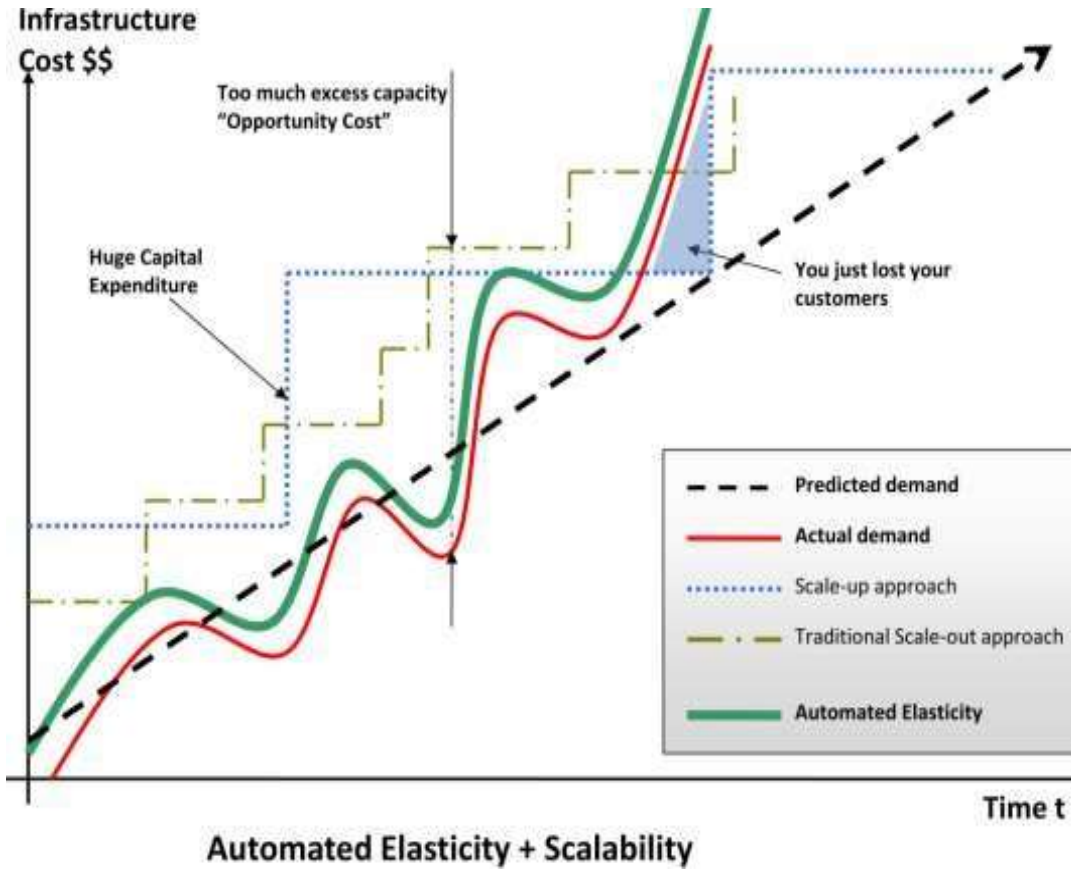
Elasticity:

Wikipedia: “In **cloud** computing, **elasticity** is defined as the degree to which a system (or a particular **cloud** layer) autonomously adapts its capacity to workload over time”

OR simply put “Ability of a system to **expand** or **contract** its dedicated resources to meet the demand”

&

Virtualization



The value of IaaS

For businesses, the greatest value of IaaS is through a concept known as cloudbursting—the process of off-loading tasks to the cloud during times when the most compute resources are needed.

To take advantage of IaaS in this capacity, IT departments must be able to build and implement the software that handles the ability to re-allocate processes to an IaaS cloud.

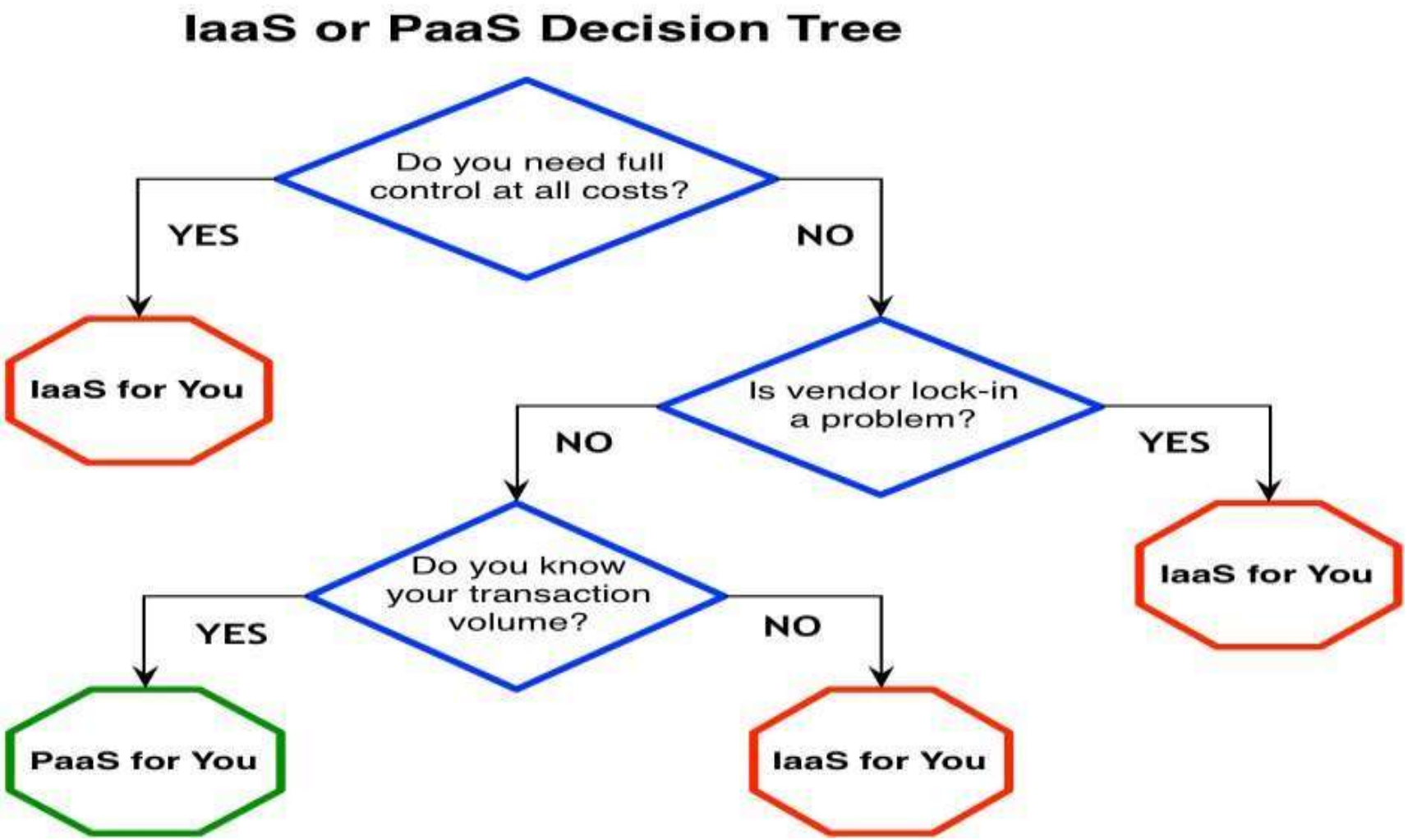
There are four important considerations to build and implement software that can manage such reallocation processes.

4 considerations:

- Developing for a specific vendor's proprietary IaaS could prove to be a costly mistake
- The complexity of well-written resource allocation software is significant and do not come cheap
- What will you be sending off to be processed in the cloud? Sending data such as personal identities, financial information, and health care data put an organization's compliance at risk
- Understand the dangers of shipping off processes that are critical to the day-to-day operation of the business.

http://www.ibm.com/developerworks/cloud/library/cl-cloudservices1iaas/

How to Decide – IaaS or PaaS?

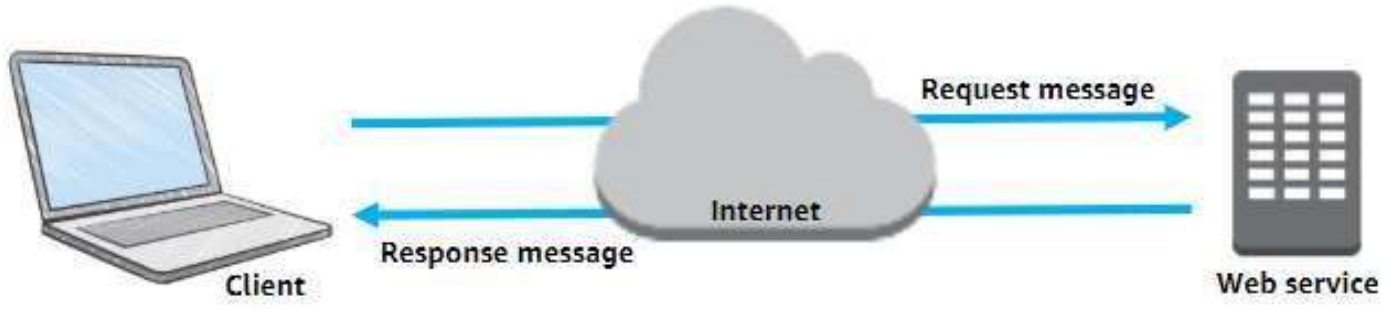


What are web services?

- A web service is any piece of software that makes itself available over the internet
- Uses a standardized format for the request and the response of

an Application Programming Interface(API) interaction

  - Extensible Markup Language (XML)
  - JavaScript Object Notation (JSON)

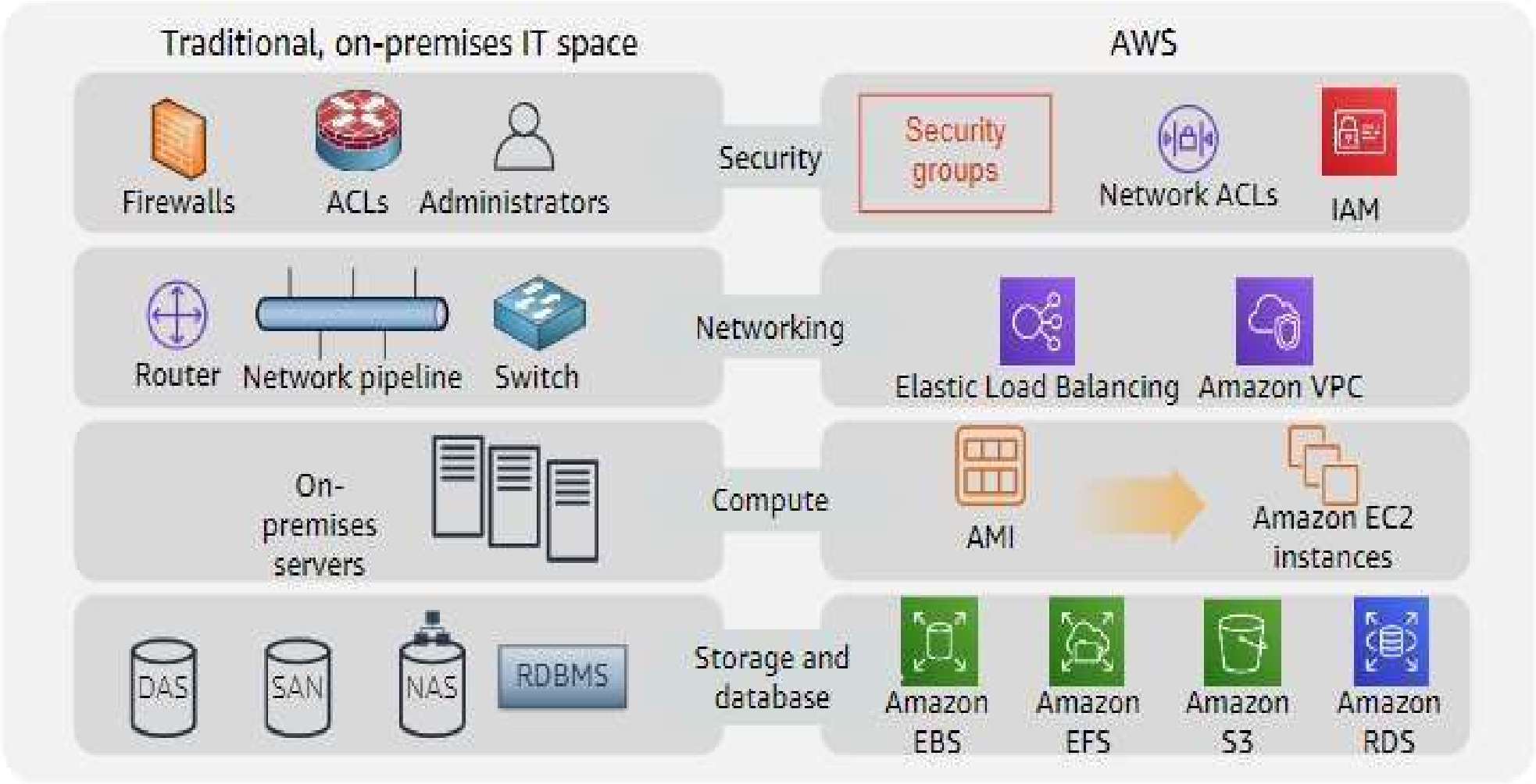


Amazon Web Services

Amazon Web Services Cloud

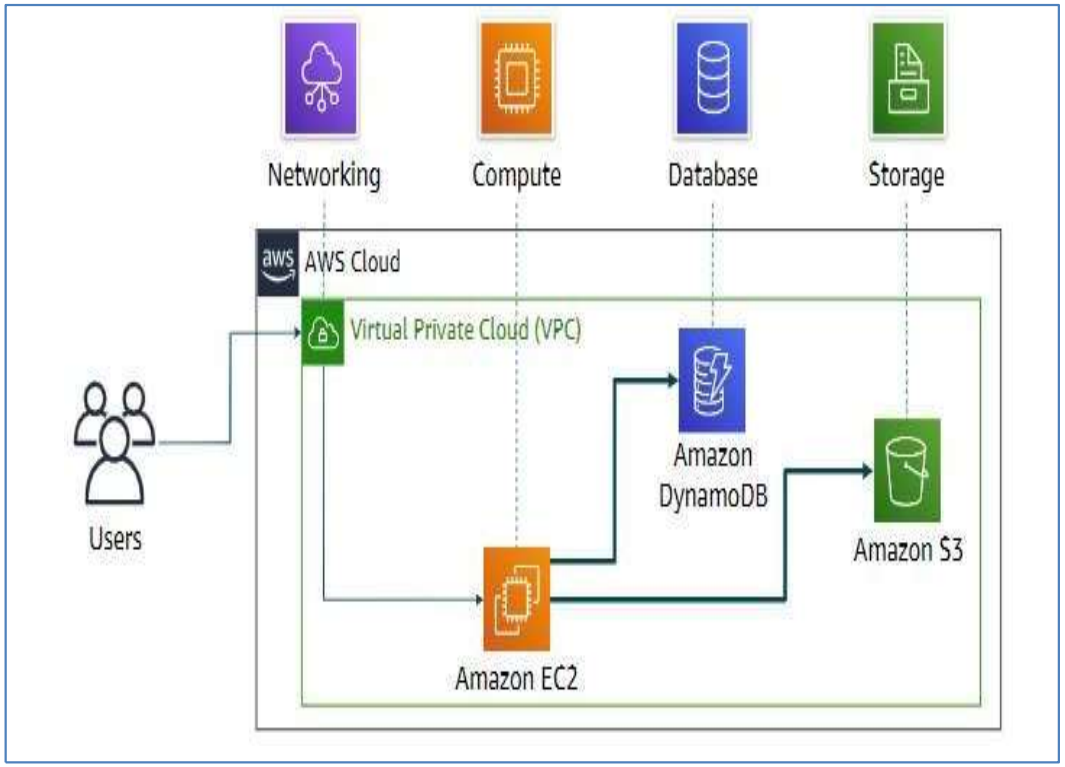
- AWS is a secure cloud platform that offers a broad set of global cloud-based products
- Provides highly reliable and scalable infrastructure for deploying web-scale solutions
- With minimal support and administration costs(pay as you go, for as long as you use them
- More flexibility than own infrastructure, either on-premise or at a datacentre facility
- AWS services work together like building blocks

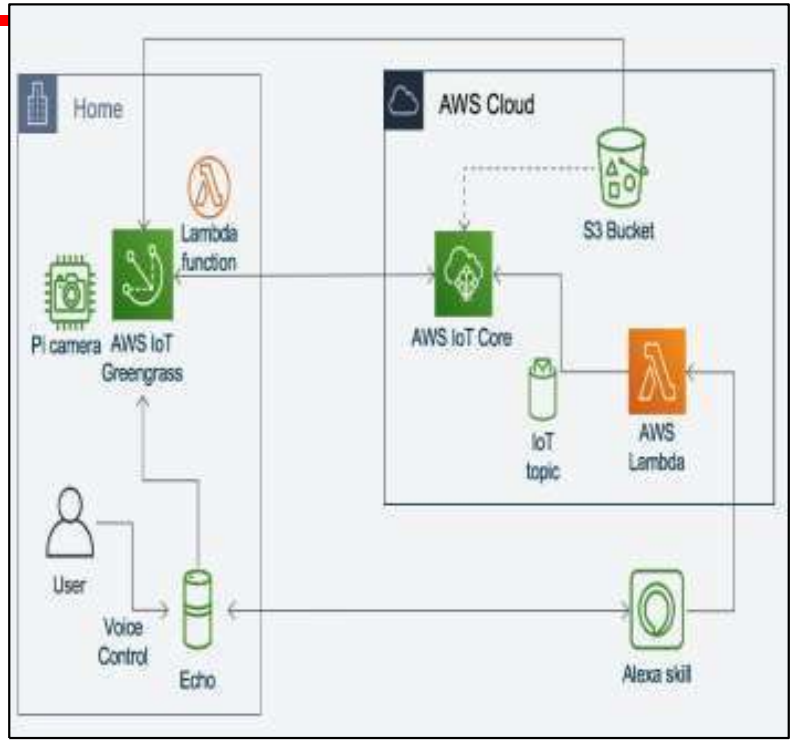
Similarities between AWS and traditional IT

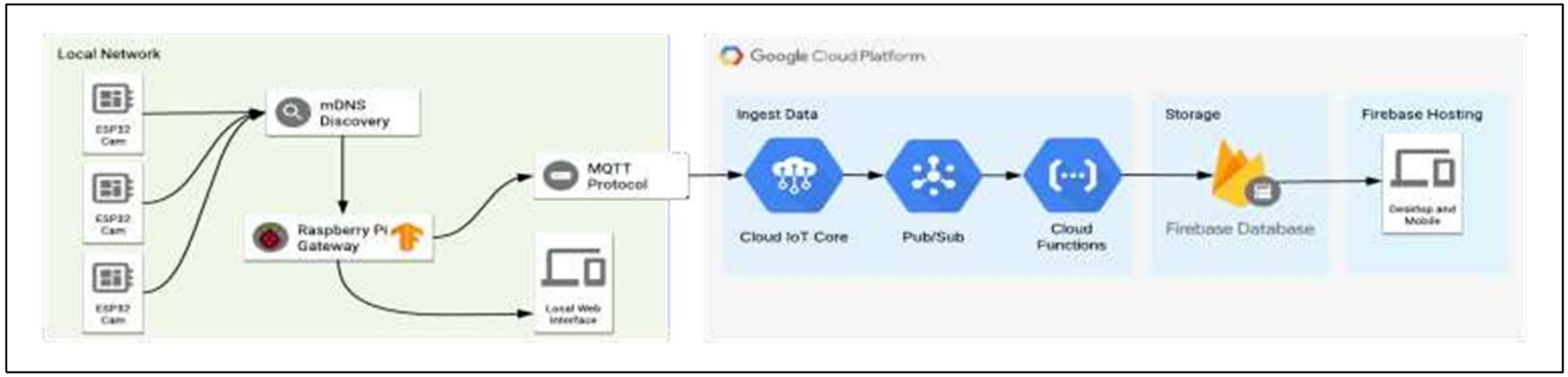


Categories of AWS services

Simple solution example







Three ways to interact with AWS

- AWS Management Console
  - Easy-to-use graphical interface
- Command Line Interface (AWS CLI)
  - Access to services by discrete commands or scripts
- Software Development Kits (SDKs)
  - Access services directly from your code (such as Java, Python, and others)

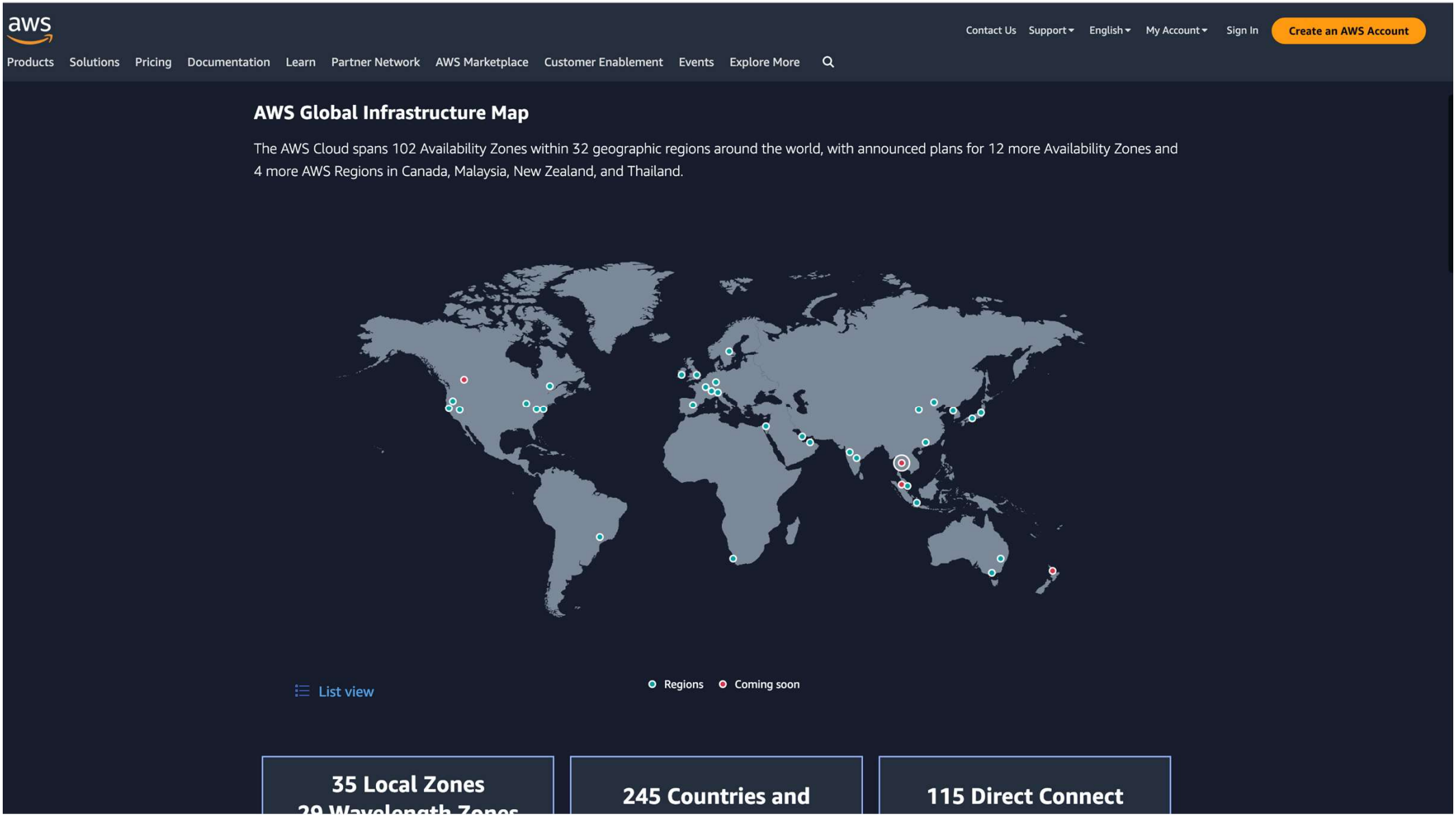
Moving to the AWS Cloud –The AWS Cloud Adoption Framework (AWS CAF)

- AWS CAF provides guidance and best practices to help organizations build a comprehensive approach to cloud computing across the organization and throughout the IT lifecycle to accelerate successful cloud adoption.
- AWS CAF is organized into six perspectives.
- Perspectives consist of sets of capabilities.

AWS CAF perspectives

AWS Global Infrastructure

https://aws.amazon.com/about-aws/global-infrastructure/#AWS_Global_Infrastructure_Map


aws
Contact Us
Support
English
My Account
Sign In
Create an AWS Account
Products
Solutions
Pricing
Documentation
Learn
Partner Network
AWS Marketplace
Customer Enablement
Events
Explore More
AWS Global Infrastructure Map
The AWS Cloud spans 102 Availability Zones within 32 geographic regions around the world, with announced plans for 12 more Availability Zones and 4 more AWS Regions in Canada, Malaysia, New Zealand, and Thailand.
List view
Regions
Coming soon
35 Local Zones
245 Countries and
115 Direct Connect

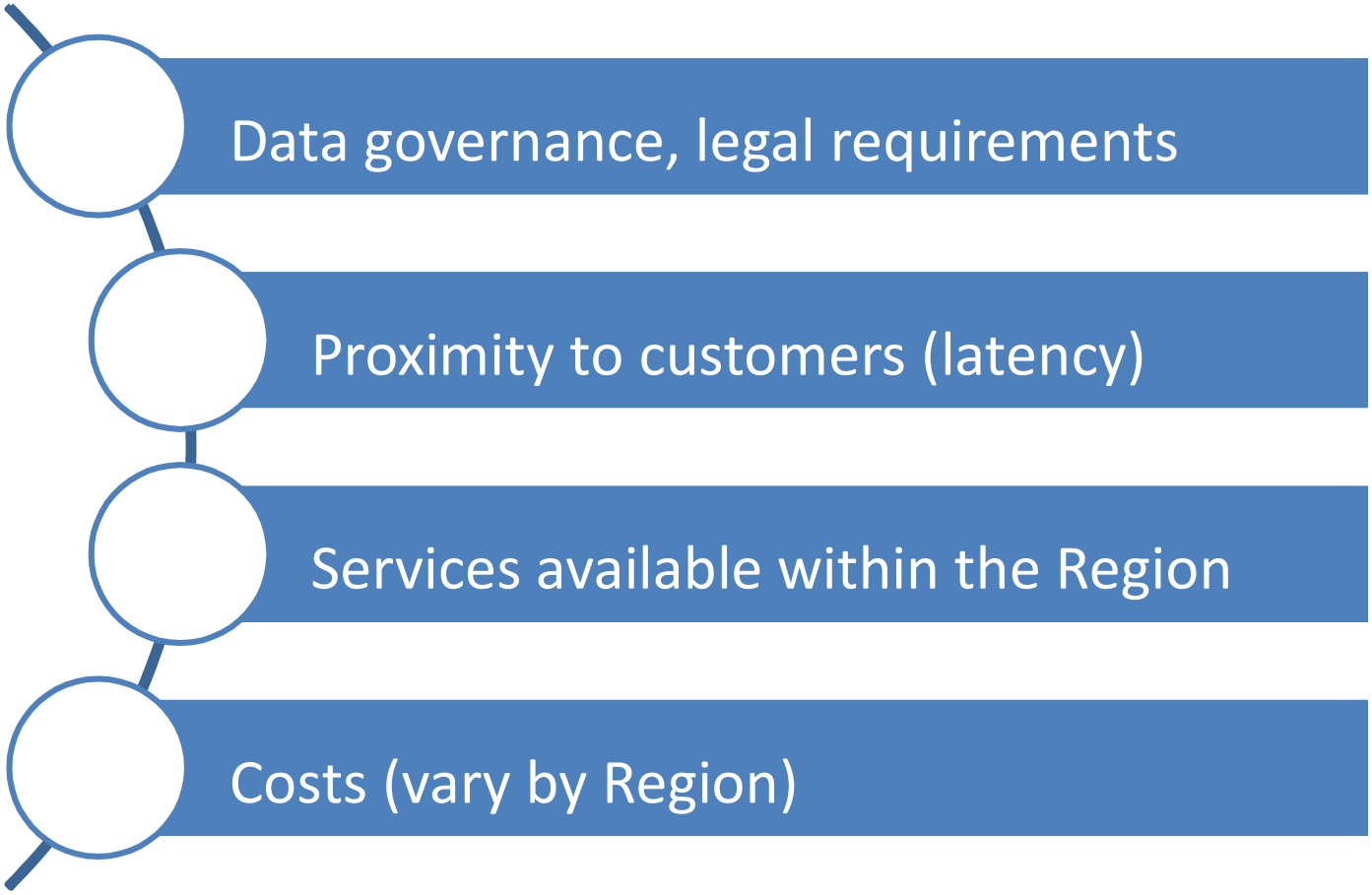
AWS Regions

- An AWS Region is a geographical area.
  - Data replication across Regions is controlled by you.
  - Communication between Regions uses AWS backbone network infrastructure.
- Each Region provides full redundancy and connectivity to the network.
- A Region typically consists of two or more Availability Zones.

https://aws.amazon.com/about-aws/global-infrastructure/regions_az/?p=ngi&loc=2

Selecting a Region

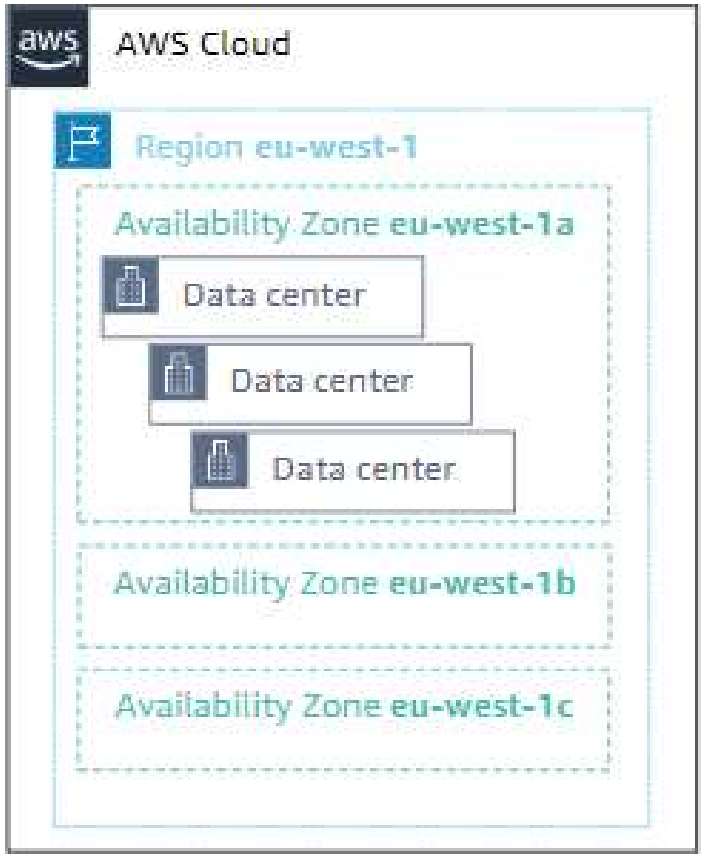
Determine the right Region for your service, applications, and data based on these factors



Availability Zones

- Each Region has multiple Availability Zones.
- Each Availability Zone is a fully isolated partition of the AWS infrastructure.

  - Availability Zones consist of discrete data centers
  - They are designed for fault isolation
  - They are interconnected with other Availability Zones by using high- speed private networking
  - AWS recommends replicating data and resources across Availability Zones for resiliency



AWS data centers

- AWS data centers are designed for security.
- Data centers are where the data resides and data processing occurs.
- Each data center has redundant power, networking, and connectivity, and is housed in a separate facility.
- A data center typically has 50,000 to 80,000 physical servers.

Compute service

Storage service

Database service

Networking and content delivery service

Security, identity, and compliance service

AWS Identity and Access Management (IAM)

AWS Organizations

Amazon Cognito

AWS Artifact

AWS Key Management Service

AWS Shield

Hands-on activity: AWS Management Console

1. Launch the Sandbox hands-on environment and connect to the AWS Management Console.
2. Explore the AWS Management Console.
   A. Click the **Services** menu.
   B. Notice how services are grouped into service categories. For example, the **EC2** service appears in the **Compute** service category.

      Question #1: Under which service category does the **IAM** service appear?

      Question #2: Under which service category does the **Amazon VPC** service appear?

   C. Click the **Amazon VPC** service. Notice that the dropdown menu in the top-right corner displays an AWS Region (for example, it might display *N. Virginia*).
   D. Click the Region menu and switch to a different Region. For example, choose **EU (London)**.
   E. Click **Subnets** (on the left side of the screen). The Region has three subnets in it. Click the box next to one of the subnets. Notice that the bottom half of the screen now displays details about this subnet.

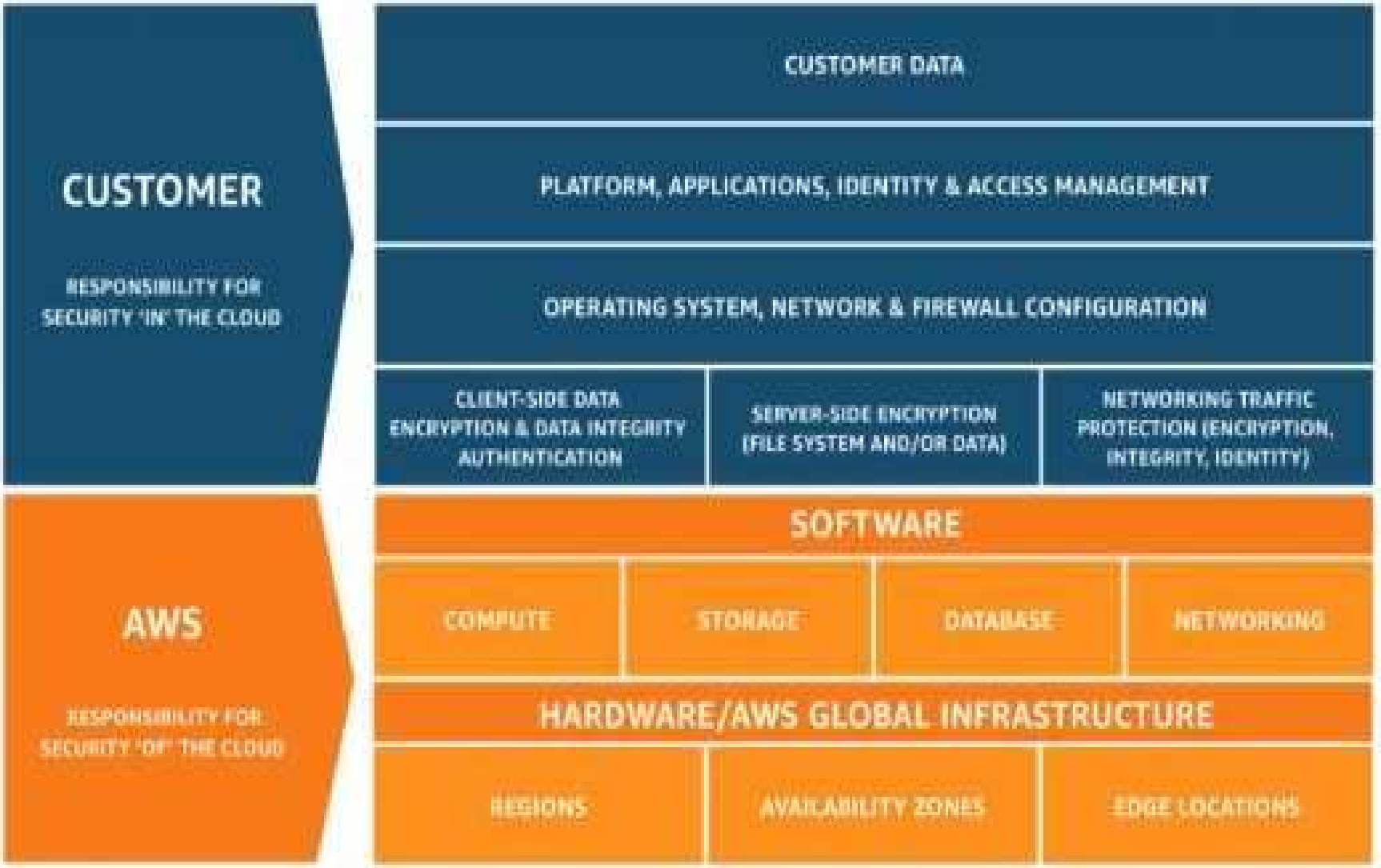
      Question #3: Does the subnet you selected exist at the level of the Region or at the level of the Availability Zone?

   F. Click **Your VPCs**. An existing VPC is already selected.

      Question #4: Does the VPC exist at the level of the Region or the level of the Availability Zone?

      Question #5: Which services are global instead of Regional? Check Amazon EC2, IAM, Lambda, and Route 53.

AWS shared responsibility model



AWS shared responsibility model

- **AWS** operates, manages, and controls the components from the software virtualization layer down to the physical security of the facilities where AWS services operate.
- **AWS** is responsible for protecting the infrastructure that runs all the services that are offered in the AWS Cloud.
- This infrastructure is composed of the hardware, software, networking, and facilities that run the AWS Cloud services

AWS shared responsibility model

- The customer is responsible for the encryption of data at rest and data in transit.
- The customer should also ensure that the network is configured for security and that security credentials and logins are managed safely.
- The customer is responsible for the configuration of security groups and the configuration of the operating system that run on compute instances that they launch (including updates and security patches)

Service characteristics and security responsibility

- IaaS
  - Customer has more flexibility over configuring networking and storage settings
  - Customer is responsible for managing more aspects of the security
  - Customer configures the access controls
- PaaS
  - Customer does not need to manage the underlying infrastructure
  - AWS handles the operating system, database patching, firewall configuration, and disaster recovery
  - Customer can focus on managing code or data

Service characteristics and security responsibility

- SaaS
  - Software is centrally hosted
  - Licensed on a subscription model or pay-as-you-go basis
  - Services are typically accessed via web browser, mobile app, or application programming interface (API)
  - Customers do not need to manage the infrastructure that supports the service

AWS Identity and Access Manageme (IAM)

- Use IAMto manage access to AWS resources–
  - A resource is an entity in an AWS account that you can work with
  - Example resources; An Amazon EC2 instance or an Amazon S3 bucket
- Example–Control who can terminate Amazon EC2 instances
- Define fine-grained access rights –
  - Who can access the resource
  - Which resources can be accessed and what can the user do to the resource
  - How resources can be accessed •IAM is a no-cost AWS

IAM: Essential components

- A person or application that can authenticate with an AWS account

- A collection of IAM users that are granted identical authorization

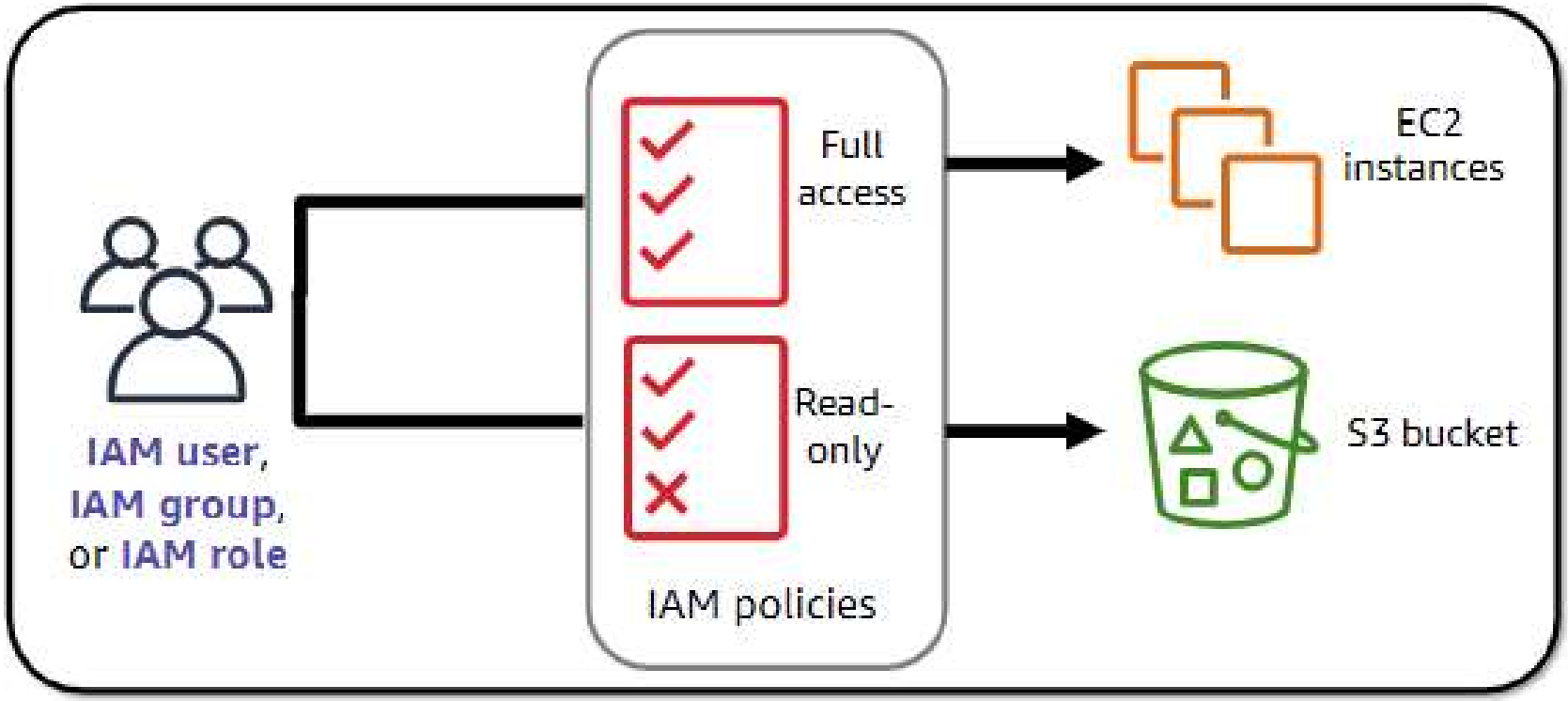
- The document that defines which resources can be accessed and the level of access to each resource

- Useful mechanism to grant a set of permissions for making AWS service

Authenticate as an IAM user to gain access

- Programmatic access
  - Authenticate using:
    - Access key ID
    - Secret access key
  - Provides AWS CLI and AWS SDK access
- AWS Management Console access
  - Authenticate using:
    - 12-digit Account ID oralias
    - IAM user name
    - IAM password
  - If enabled, multi-factor authentication (MFA)prompts for an authentication code

Authorization: What actions are permitted



IAM: Authorization

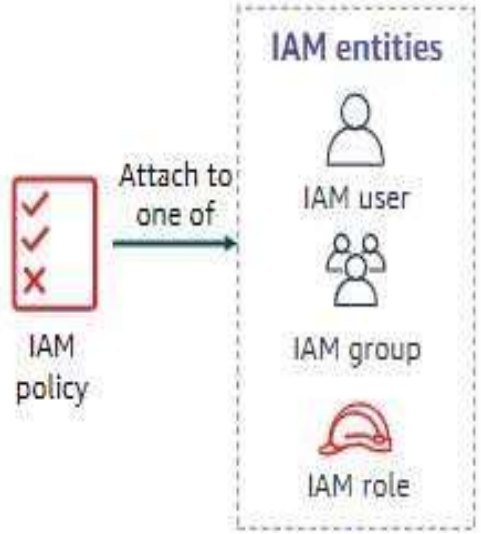
- Assign permissions by creating an IAM policy.
  - Permissions determine which resources and operations are allowed:
  - All permissions are implicitly denied by default.
  - If something is explicitly denied, it is never allowed.

Best practice: Follow the principle of least privilege.

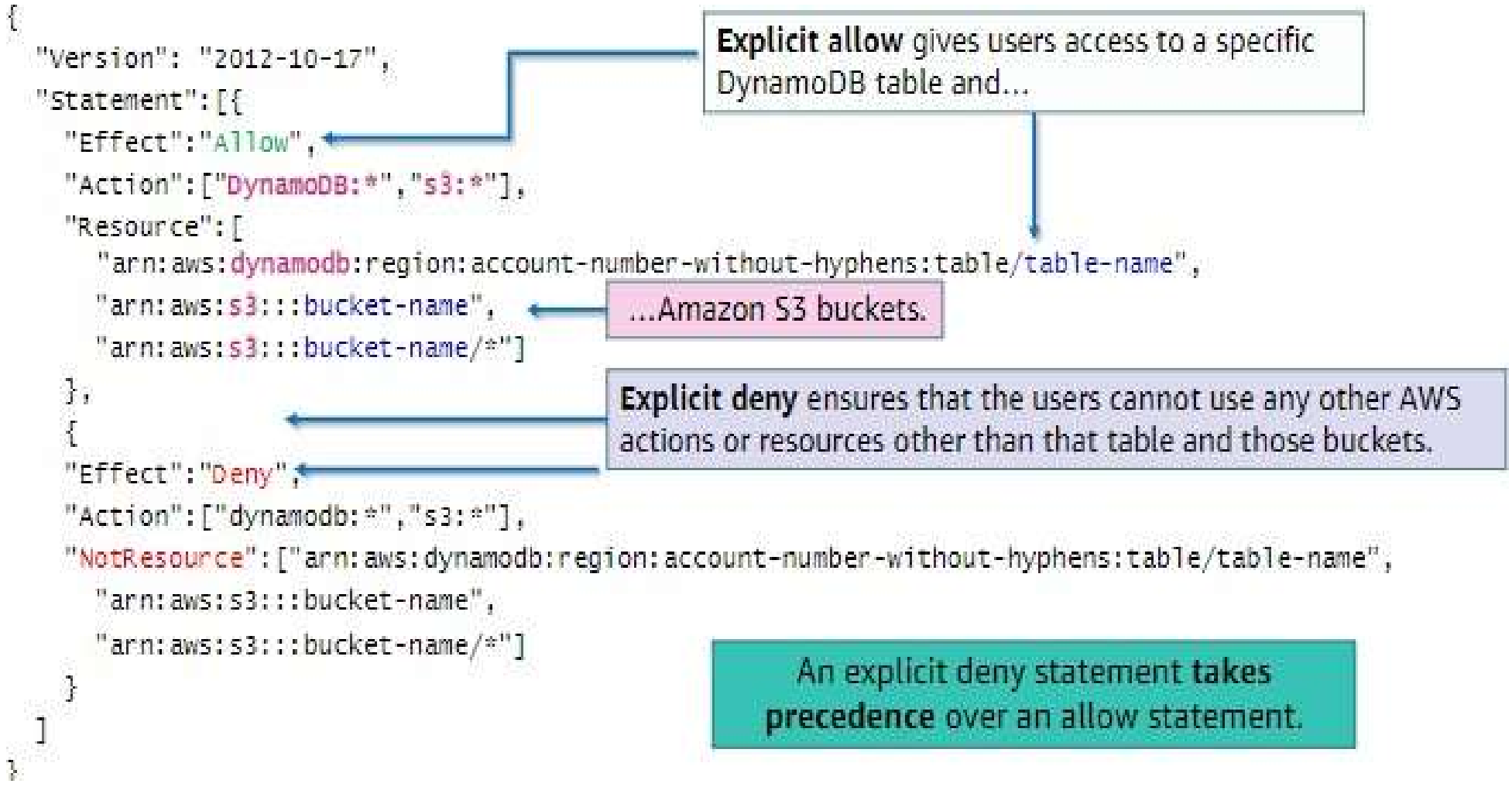
Note: The scope of IAM service configurations is global.
Settings apply across all AWS Region

IAM policies

- An IAM policy is a document that defines permissions
- Enables fine-grained access control
- Two types of policies –identity-based and resource-based
  - Identity-based policies –
    - Attach a policy to any IAM entity
    - An IAM user, an IAM group, or an IAM role
  - Policies specify:
    - Actions that maybe performed by the entity
    - Actions that may notbe performed by the entity
  - A single policy can be attached to multiple entities
  - A single entity can have multiple policies attached to it
- Resource-based policies
- Attached to a resource (such as an S3 bucket)

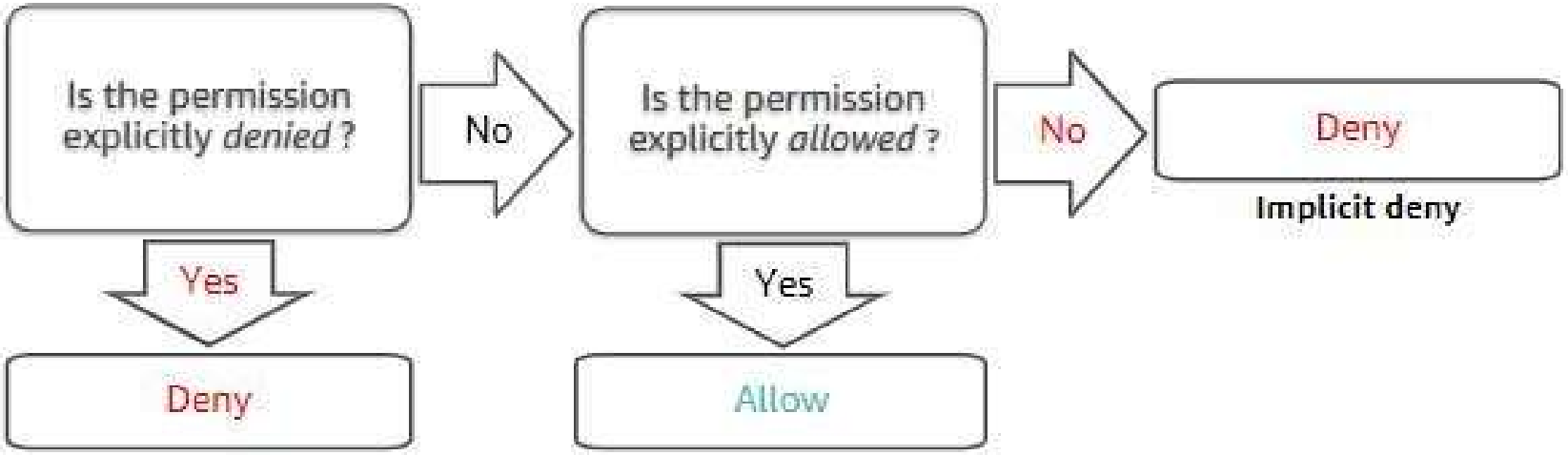


IAM policy example



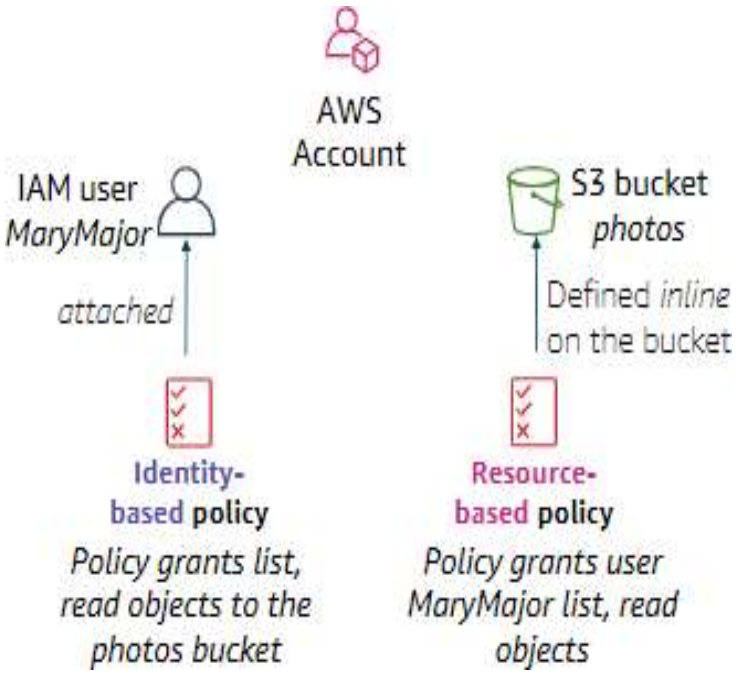
IAM permissions

- How IAM determines permissions:



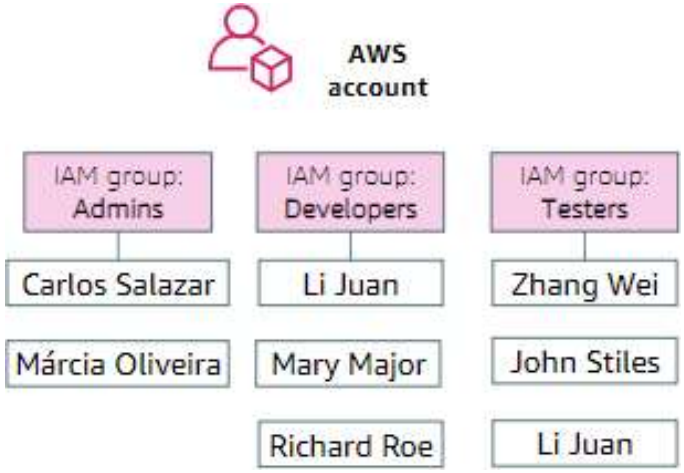
Resource-based policies

- Identity-based policies are attached to a user, group, or role
- Resource-based policies are attached to a resource (not to a user, group or role)
- Characteristics of resource-based policies –
  - Specifies who has access to the resource and what actions they can perform on it
  - The policies are in line only, not managed
- Resource-based policies are supported only by some AWS services



IAM groups

- An IAM group is a collection of IAM users
- A group is used to grant the same permissions to multiple users
- Permissions granted by attaching IAM policy or policies to the group
- A user can belong to multiple groups
- There is no default group
- Groups cannot be nested

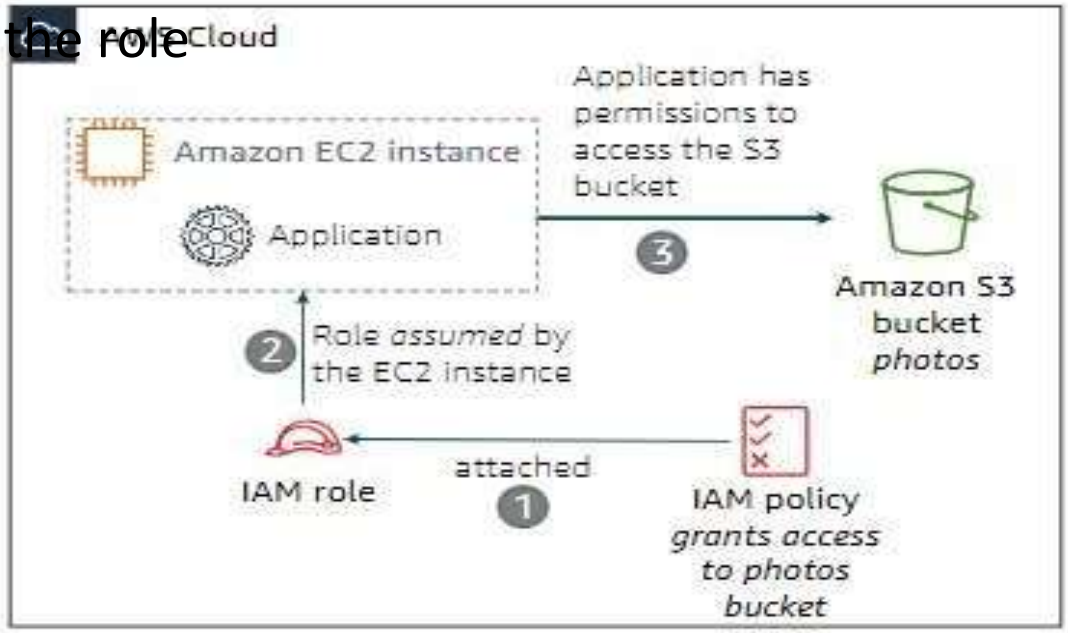


IAM roles

- An IAM role is an IAM identity with specific permissions
- Similar to an IAM user
  - Attach permissions policies to it
- Different from an IAM user
  - Not uniquely associated with one person
  - Intended to be assumable by a person, application, or service
- Role provides temporary security credentials
- Examples of how IAM roles are used to delegate access –
  - Used by an IAM user in the same AWS account as the role
  - Used by an AWS service—such as Amazon EC2—in the same account as the role
  - Used by an IAM user in a different AWS account than the role

Example use of an IAM role

- Scenario:
  - An application that runs on an EC2 instance needs access to an S3 bucket
- Solution:
  - Define an IAM policy that grants access to the S3 bucket.
  - Attach the policy to a role
  - Allow the EC2 instance to assume the role



IaaS for you

*Thanks, I feel so "**Clouded**" now*

Assignment Details

Assignment consists of 4 parts

1. Test Bed presentation (8 marks)
2. Activity using Tool/Services of AWS or any other Cloud platform (some thing innovative) (12 marks)
3. Detailed report (5 marks)

Group Assignment: size 5 max

Presentation: 20 min

Dates of presentation: will be shared shortly

Slots will be assigned by me. So all the groups need to be ready before the beginning of the first presentation