

CS 4037
Introduction to Cloud
Computing
Lecture 23

Danyal Farhat
FAST School of Computing
NUCES Lahore

AWS Global Infrastructure

Lecture's Agenda

- **AWS Global Infrastructure**
- Regions and Availability Zones
- Data Centers and Point of Presence
- Infrastructure Features

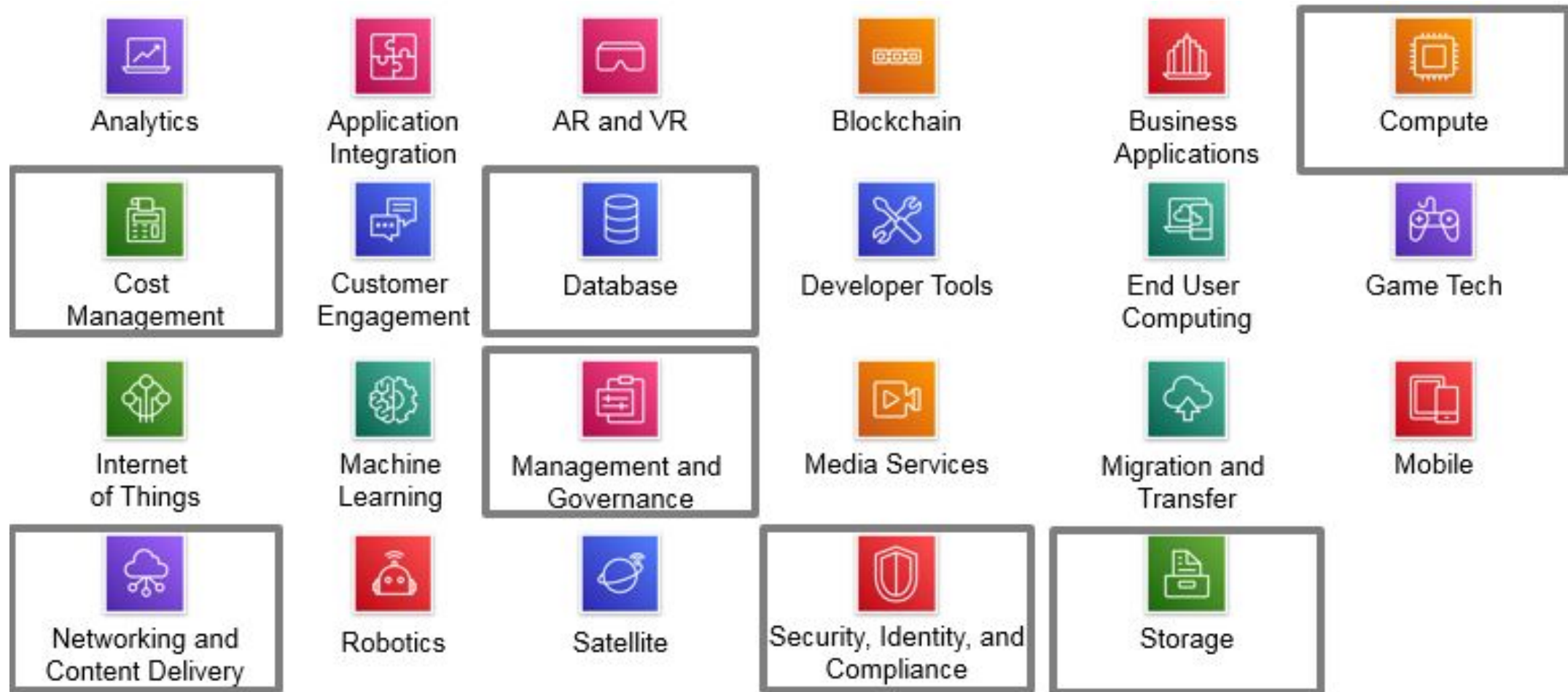


AWS Global Infrastructure

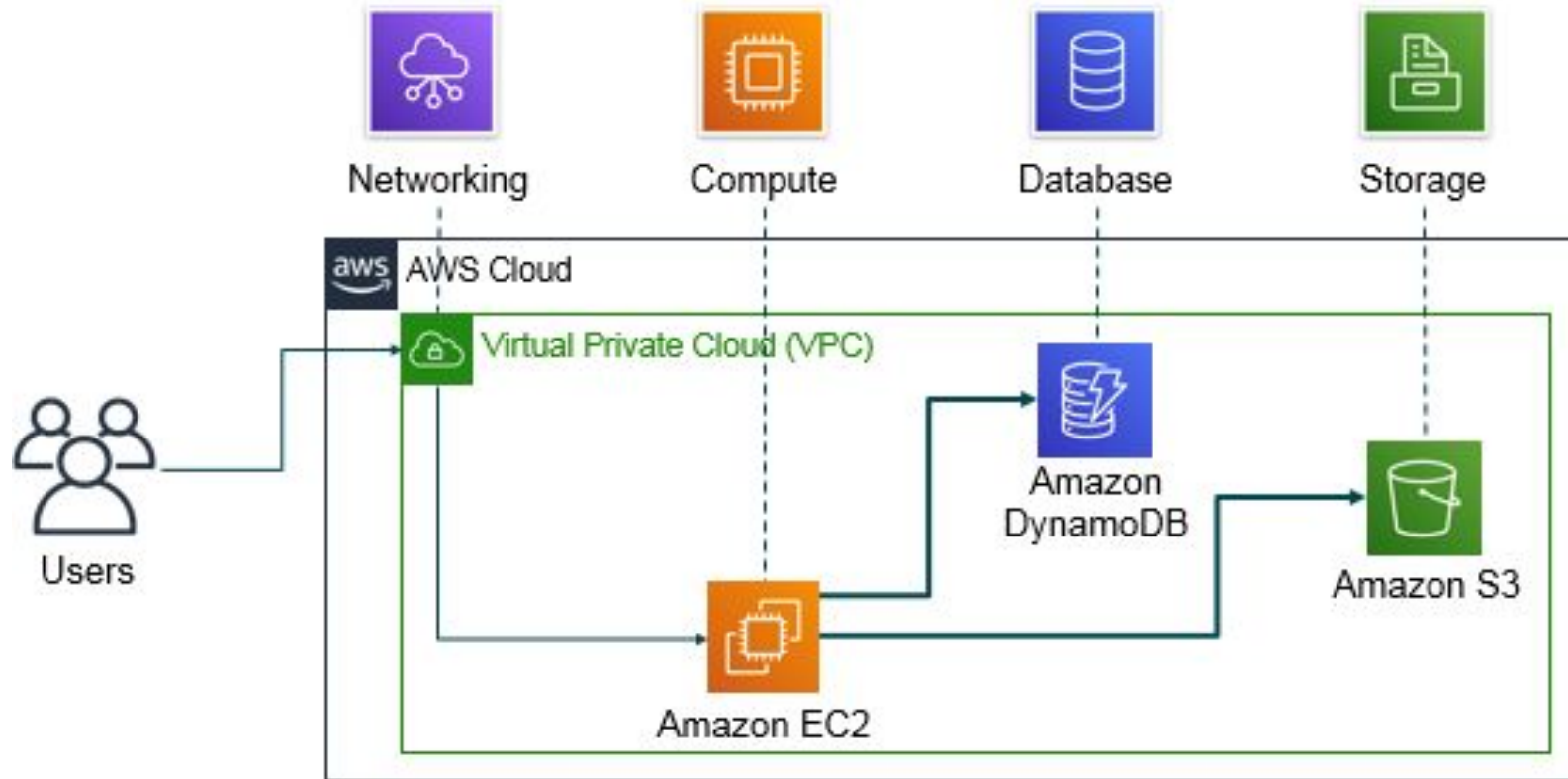
- Designed and **built** to deliver a flexible, reliable, scalable, and secure cloud computing environment with high-quality global network performance
- AWS continually **updates** its global infrastructure footprint



Categories of AWS services



Sample Solution Example



Lecture's Agenda

- AWS Global Infrastructure
- **Regions and Availability Zones**
- Data Centers and Point of Presence
- Infrastructure Features



AWS Regions

- AWS Cloud infrastructure is **built** around Regions
 - AWS has **30+ Regions** worldwide
 - AWS China and AWS GovCloud (US) Regions
- “An **AWS Region** is a physical geographical location with one or more **Availability Zones**”
 - Availability Zones consist of one or more data centers



Example: London Region

AWS Regions (Cont.)

- Data replication across regions is **controlled** by customer
- Communication between regions **uses** AWS backbone network infrastructure
- Each region **provides** full redundancy to the network



Example: London Region

Selecting a Region

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



Data governance,
legal requirements



Proximity to customers
(latency)



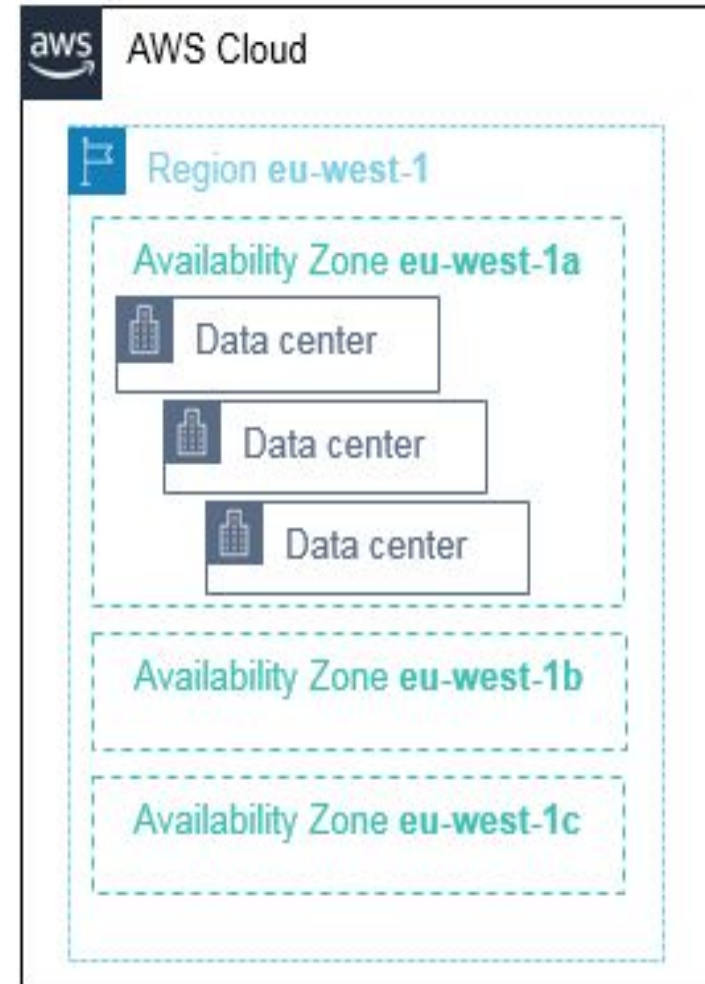
Services available
within the Region



Costs (vary by Region)

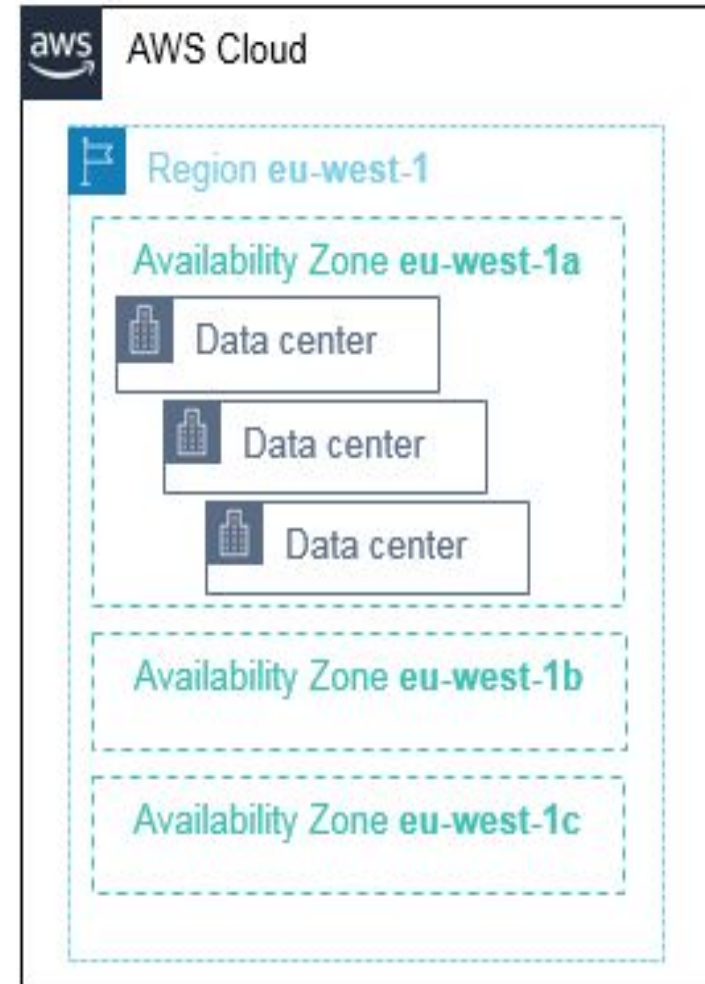
Availability Zones

- “Availability Zone is a fully **isolated partition** of the AWS infrastructure”
 - Consists of discrete data centers
 - Designed for fault isolation
 - Customer chooses Availability Zones
- Availability Zones are **100 km away** from each other



Availability Zones (Cont.)

- AWS **recommends** replicating data and resources across availability zones for resiliency
- Customers **cannot** specify a specific data center in AWS
 - Customer view is limited to availability zones only
- Many cloud providers give option to customers to choose a **specific data center** where their data should reside and/or replicate



Lecture's Agenda

- AWS Global Infrastructure
- Regions and Availability Zones
- **Data Centers and Point of Presence**
- Infrastructure Features



Data Centers

- Data centers are where the **data resides** and data processing occurs
- Each data center has redundant power, hardware, and connectivity, and is **housed** in a separate facility
 - A data center typically has 50,000 to 80,000 physical servers.
 - Availability Zone: NUCES LHR
 - Data Centers: Blocks A, B, C, D, E & F.



Point of Presence

- AWS **provides** a global network of Points of Presence locations
 - Consists of **edge locations** and a much smaller number of Regional edge caches
- Used by Amazon CloudFront and Amazon Route 53 services
 - CloudFront: A **global CDN** that delivers content to end users with reduced latency
 - Amazon Route 53 is a **Domain Name System (DNS)** service



Lecture's Agenda

- AWS Global Infrastructure
- Regions and Availability Zones
- Data Centers and Point of Presence
- **Infrastructure Features**



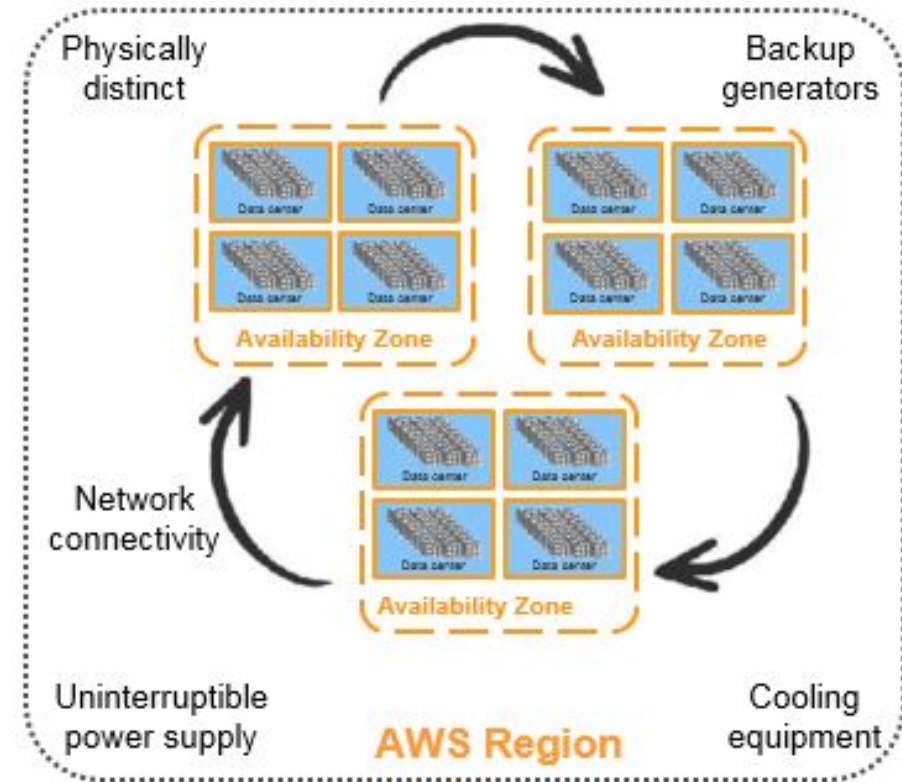
Infrastructure Features

Elasticity:

- Elastic infrastructure
- Dynamic adaption of capacity

Scalability:

- Scalable infrastructure
- Adapts to accommodate growth



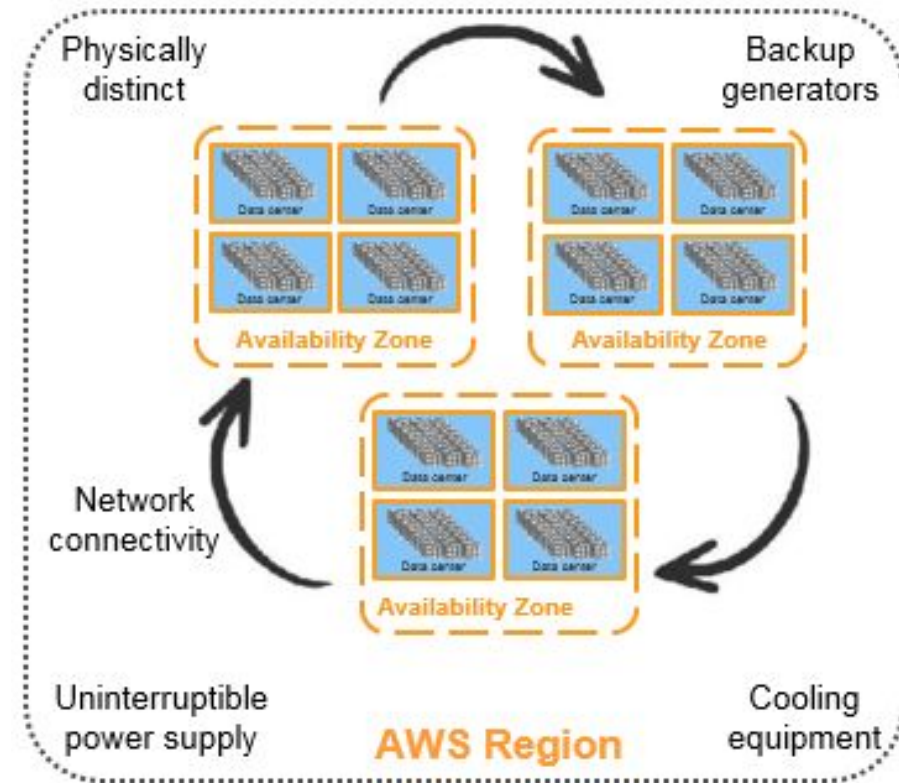
Infrastructure Features (Cont.)

Fault Tolerance:

- Continues operating properly in the presence of a failure
- Built-in redundancy of components

High availability:

- High level of operational performance
- Minimized downtime
- No human intervention



Additional Resources (Cont.)

- **AWS Global Infrastructure**

- <https://aws.amazon.com/about-aws/global-infrastructure/>

- **AWS Regional Services List**

- <https://aws.amazon.com/about-aws/global-infrastructure/regional-product-services/>

- **AWS Cloud Products**

- <https://aws.amazon.com/products/>

Questions?