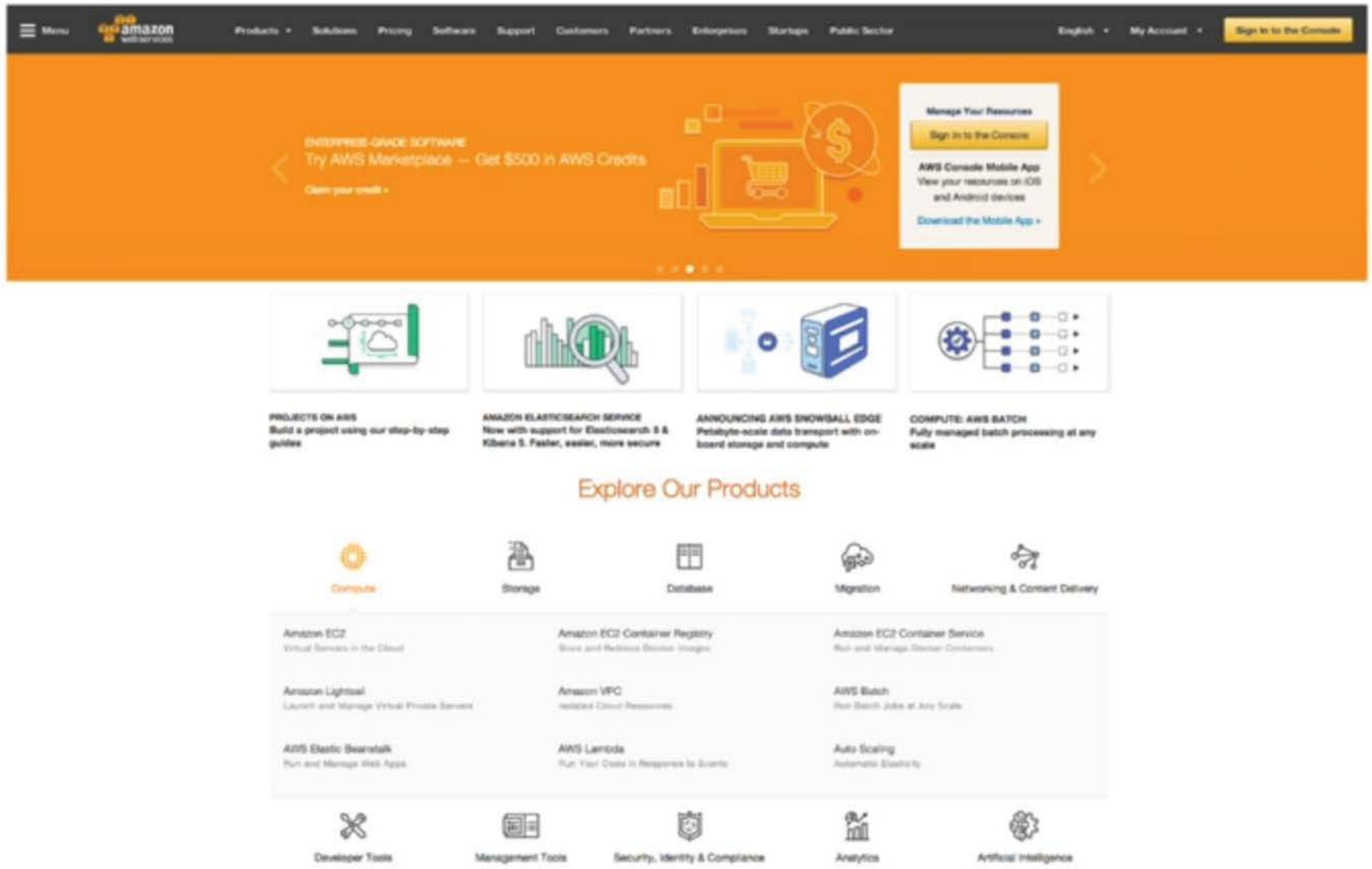
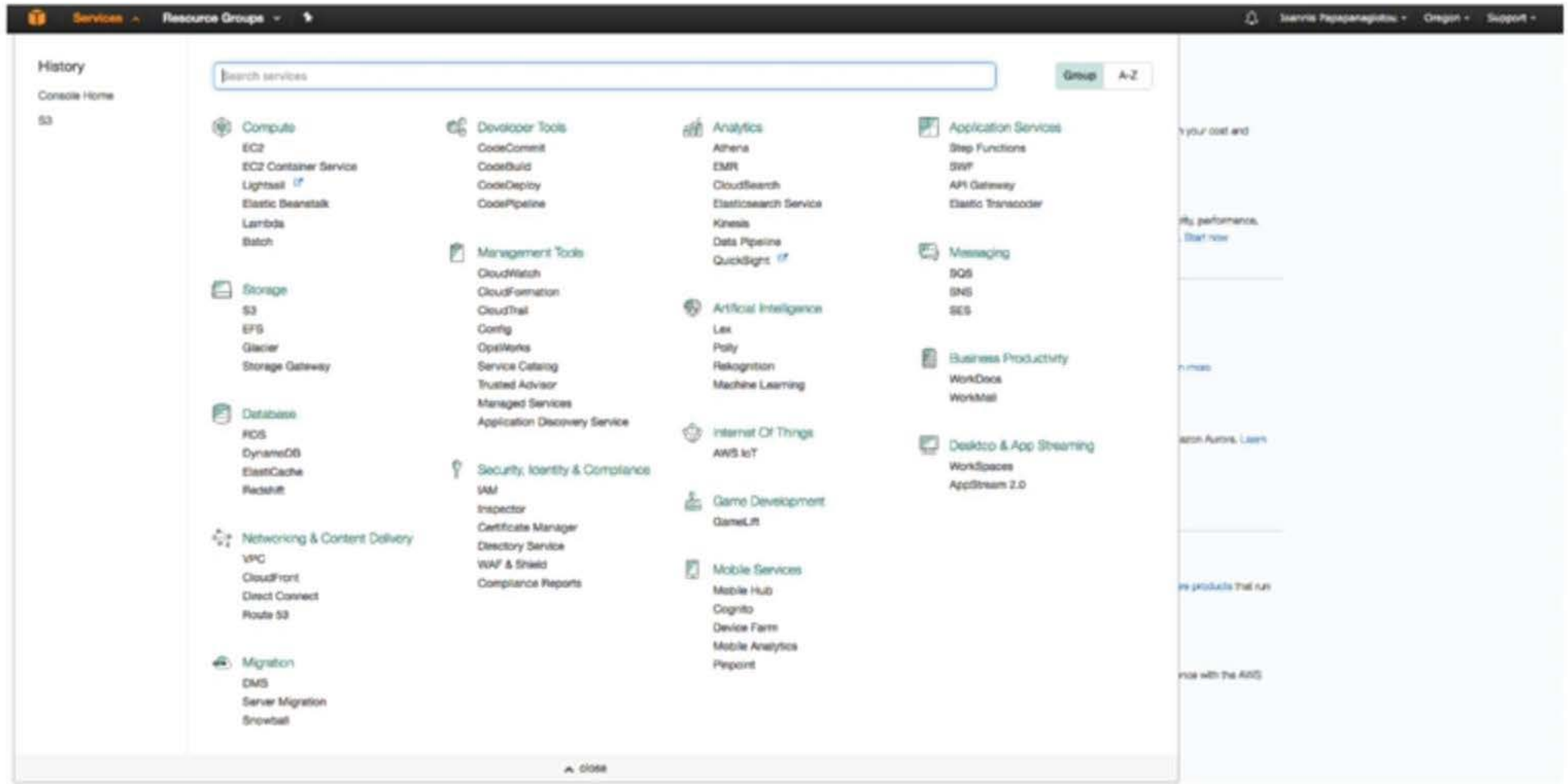
Amazon Web Services



AWS Console


Services
Resource Groups
History
Console Home
S3
Search services
Group
A-Z
Compute
EC2
EC2 Container Service
Lightsail
Elastic Beanstalk
Lambda
Batch
Storage
S3
EFS
Glacier
Storage Gateway
Database
RDS
DynamoDB
ElastiCache
Redshift
Networking & Content Delivery
VPC
CloudFront
Direct Connect
Route 53
Migration
DMS
Server Migration
Snowball
Developer Tools
CodeCommit
CodeBuild
CodeDeploy
CodePipeline
Management Tools
CloudWatch
CloudFormation
CloudTrail
Config
OpsWorks
Service Catalog
Trusted Advisor
Managed Services
Application Discovery Service
Security, Identity & Compliance
IAM
Inspector
Certificate Manager
Directory Service
WAF & Shield
Compliance Reports
Analytics
Athena
EMR
CloudSearch
Elasticsearch Service
Kinesis
Data Pipeline
QuickSight
Artificial Intelligence
Lex
Polly
Rekognition
Machine Learning
Internet Of Things
AWS IoT
Game Development
GameLift
Mobile Services
Mobile Hub
Cognito
Device Farm
Mobile Analytics
Pinpoint
Application Services
Step Functions
SWF
API Gateway
Elastic Transcoder
Messaging
SQS
SNS
SES
Business Productivity
WorkDocs
WorkMail
Desktop & App Streaming
WorkSpaces
AppStream 2.0
close

AWS History

- Officially Launched in 2006
- Amazon Web Services (AWS) began offering IT infrastructure services to businesses in the form of web services.
- Amazon Web Services has datacenters in U.S., Europe, Brazil, Singapore, Japan, and Australia
- AWS provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around the world.

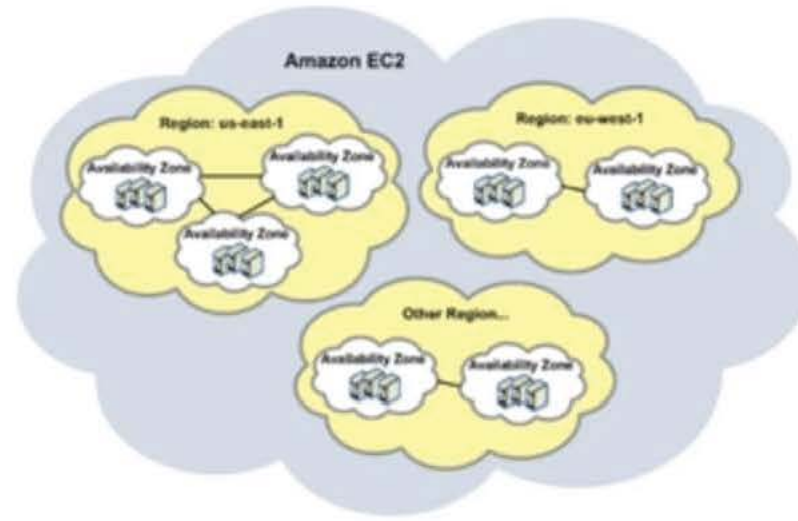
Amazon Web Services

- AWS is a collection of remote computing services that together make up a cloud computing platform
- Compute
  - Elastic Compute (EC2) = virtual private servers
  - Elastic MapReduce (EMR) = Hadoop framework running on EC2
- Storage
  - Simple Storage Service (S3) = storage
  - Glacier = low-cost and long-term storage (high redundancy, low access time)
  - Elastic Block Storage (EBS) = persistent block-level storage
- Database
  - DynamoDB = NoSQL backed by SSD
  - ElasticCache= in-memory cache based Memcached
  - Relational Database Service (RDS) = scalable database MySQL, Informix, Oracle, SQL server etc
  - …

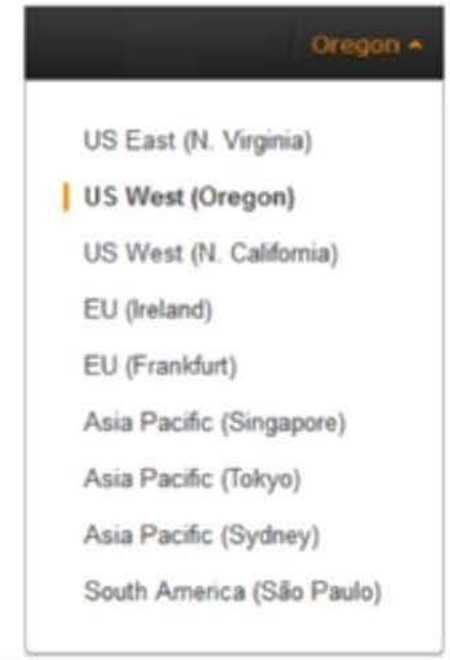
Amazon Web Services

- AWS is a collection of remote computing services that together make up a cloud computing platform
- Compute
  - Elastic Compute (EC2) = virtual private servers
  - Elastic MapReduce (EMR) = Hadoop framework running on EC2
- Storage
  - Simple Storage Service (S3) = storage
  - Glacier = low-cost and long-term storage (high redundancy, low access time)
  - Elastic Block Storage (EBS) = persistent block-level storage
- Database
  - DynamoDB = NoSQL backed by SSD
  - ElasticCache= in-memory cache based Memcached
  - Relational Database Service (RDS) = scalable database MySQL, Informix, Oracle, SQL server etc
  - …

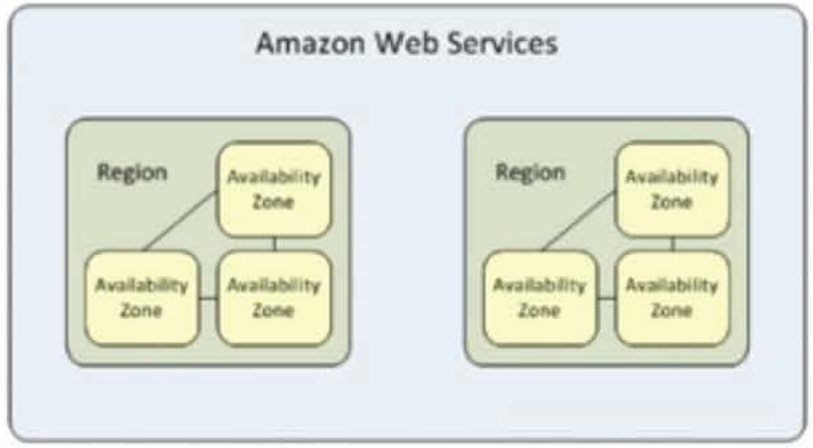
AWS Global Infrastructure



- AWS is hosted in multiple locations world-wide. These locations are composed of **Regions**.
  - Each **region** is a separate geographic area.
  - Each **region** has multiple, isolated locations known as *Availability Zones*.
  - 11 Regions



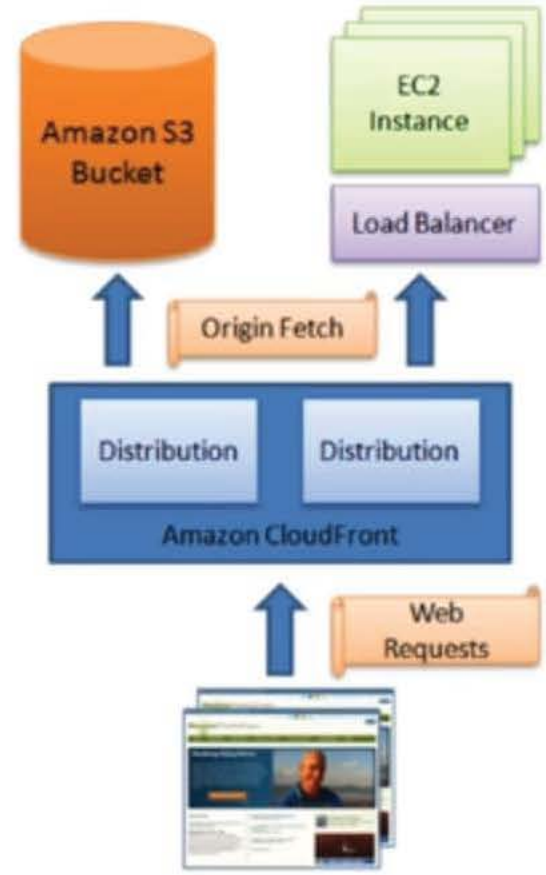
Availability Zones



- Each Region has 2 or more "Availability Zones", which are distinct data centers providing AWS services.
- Each Amazon EC2 region is designed to be completely isolated from the other Amazon EC2 regions. This achieves the greatest possible fault tolerance and stability.
  - Several services operate across Availability Zones (e.g., S3, DynamoDB) while others can be configured to replicate across Zones to spread demand and avoid downtime from failures.

Edge Locations

- 52 Edge Locations
- Edge Locations are CDN edge points
  - There are many more Edge Locations than Regions.
- Used to cache data closed to the user so that the latency is reduced.



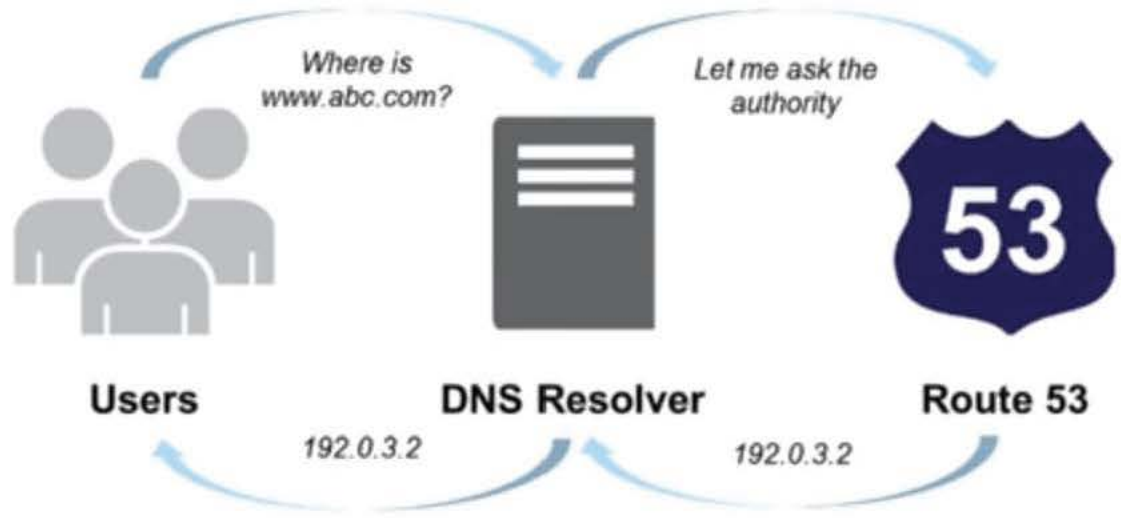
AWS Global Infrastructure

IoT: Cloud Services

Common Cloud providers: AWS Architecture

Networking (DNS)

Router 53: a highly available and scalable cloud Domain Name System (DNS) web service.



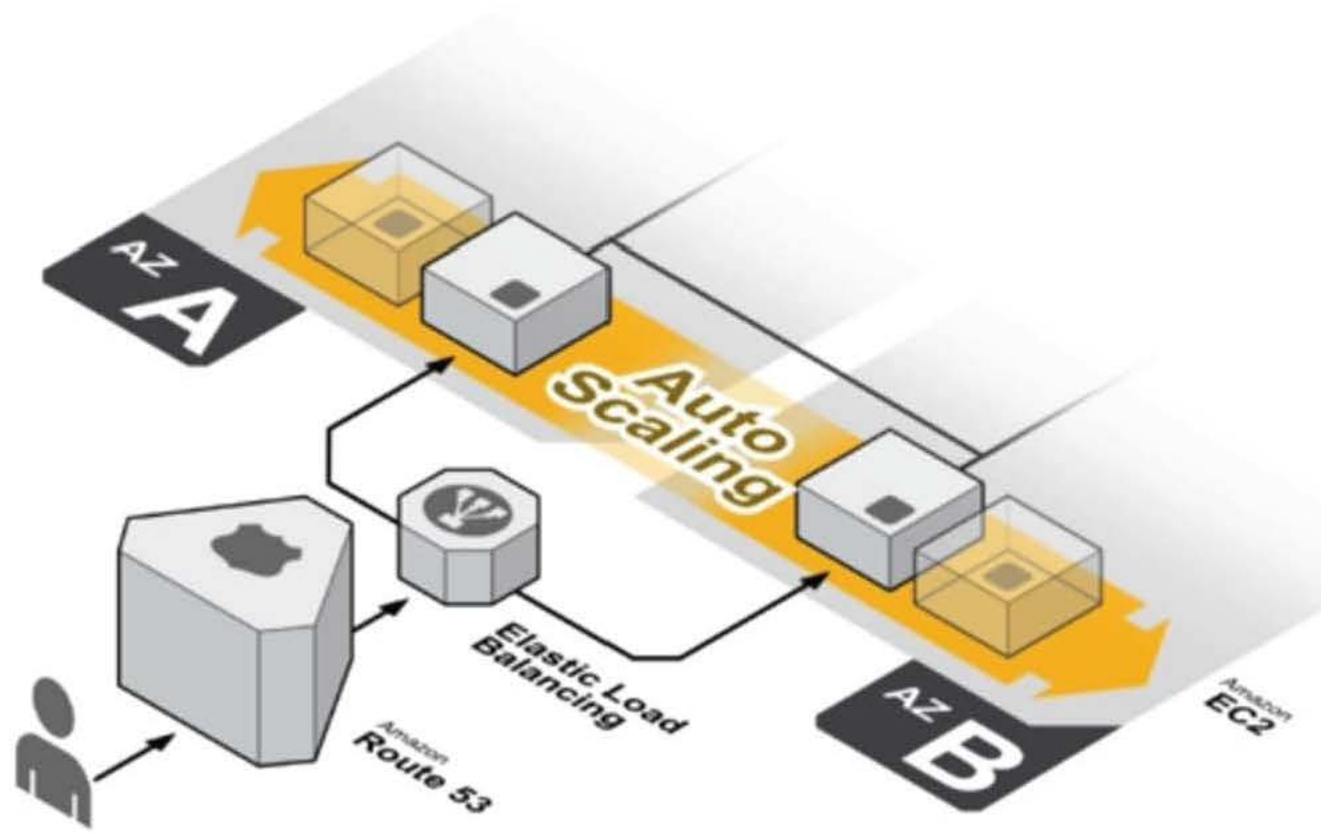
Virtual Private Cloud (VPC)

- Amazon VPC lets you provision a logically isolated section of the AWS Cloud where you can launch AWS resources in a virtual network that you define.
- For example, you can create a public-facing subnet for your webservers that has access to the Internet, and place your backend systems such as databases or application servers in a private-facing subnet with no Internet access.

Compute

- Amazon Elastic Cloud Compute (EC2) is a simple web service interface allows you to obtain and configure capacity with minimal friction.
- EC2 allows to provision virtual instances.
  - These instances can be CentOS, Ubuntu, Amazon's linux, Windows 2008/2012 etc.
  - Amazon also provides a marketplace where you can pre-configured instances.
- EC2 uses Xen virtualization.

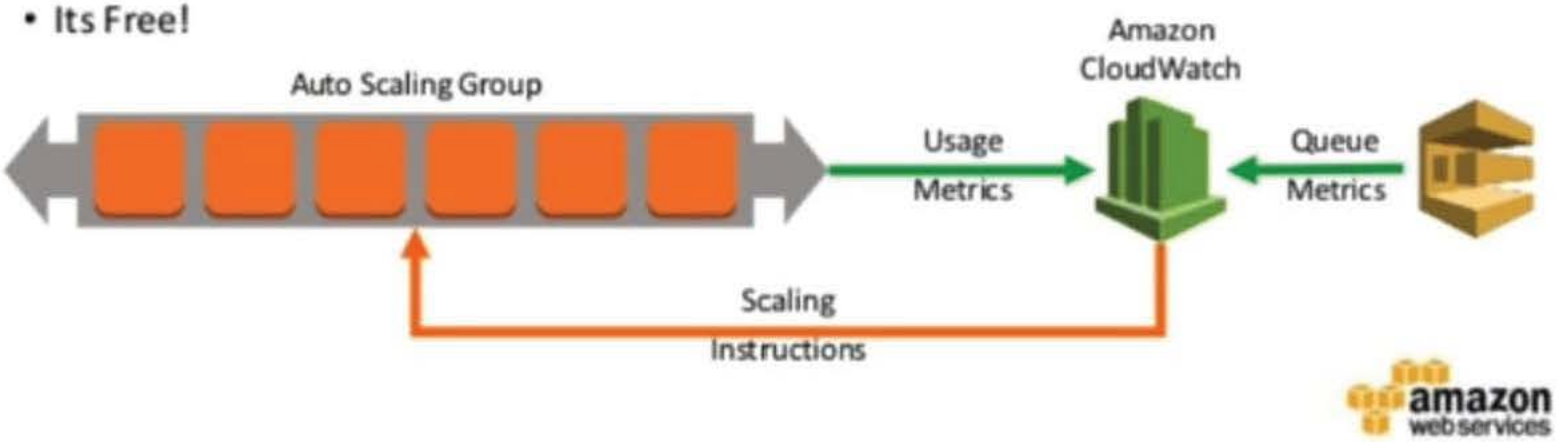
Autoscaling



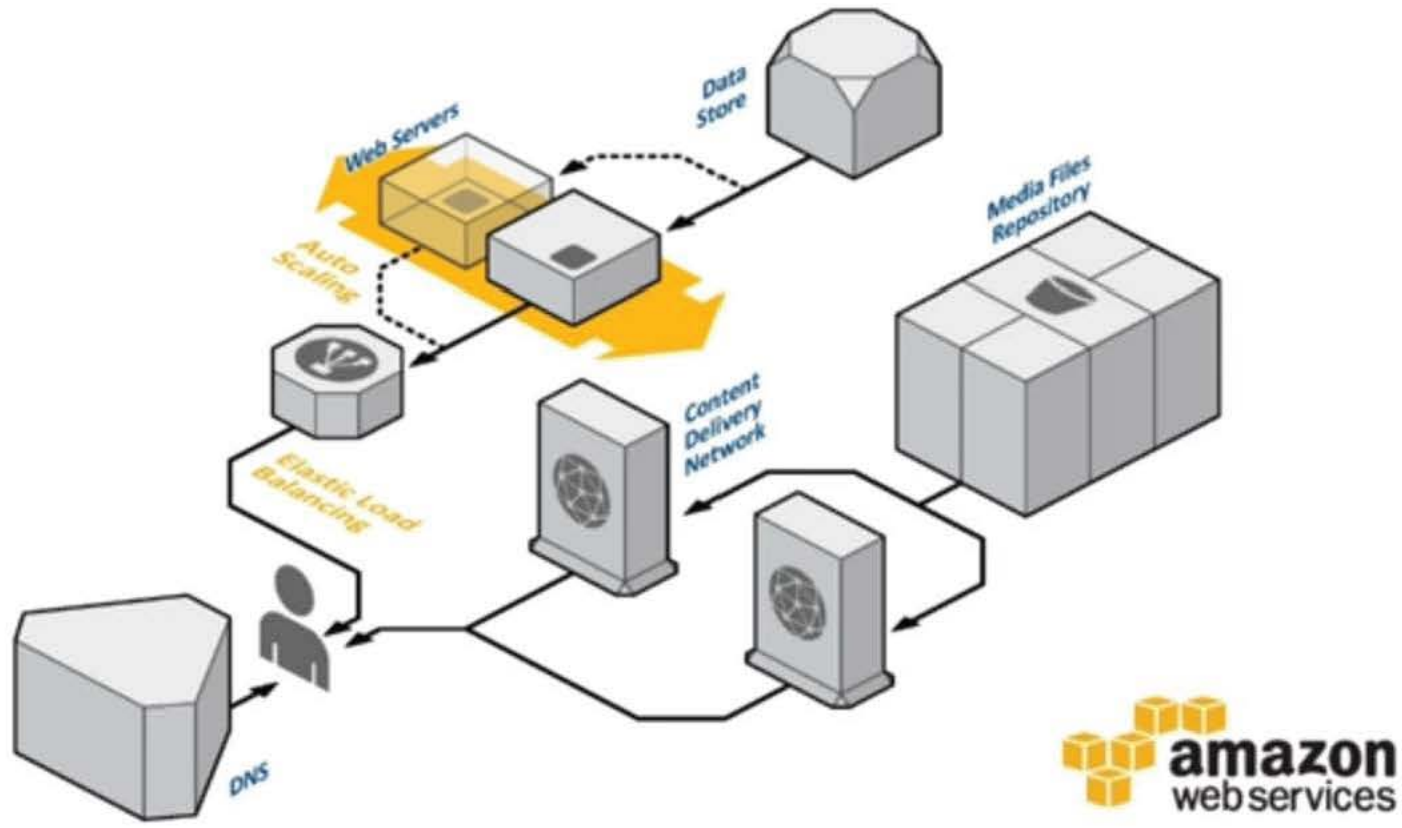
Autoscaling

Auto Scaling

- Automatic resizing of compute clusters based on demand
- Define minimum and maximum number of instances
- Define when scaling out and in occurs
- Use metrics collected in Amazon CloudWatch to drive scaling
- Run Auto Scaling for On-Demand and Spot instance types
- Its Free!

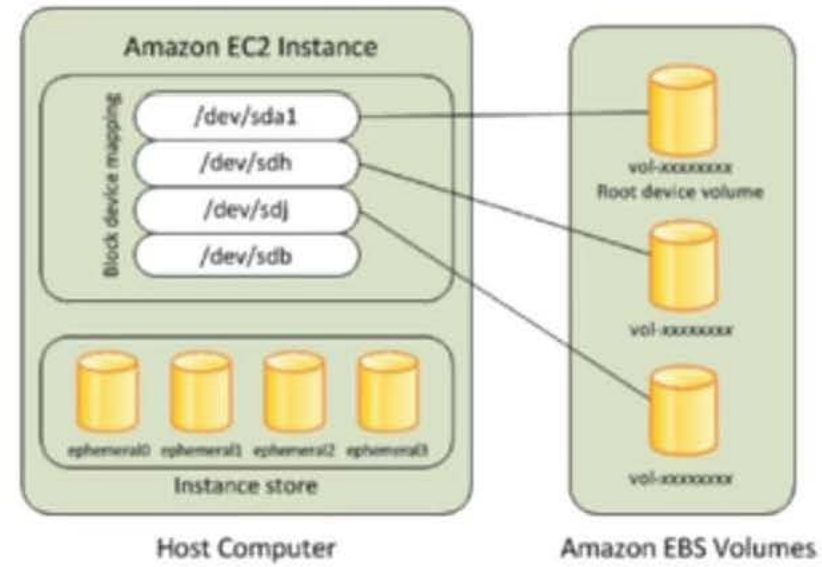


Elastic Load Balancing



Amazon EBS

- Amazon Elastic Block Storage (EBS) provides raw block devices that can be attached with EC2 instances.
  - EBS supports a number of advanced storage features, including snapshotting and cloning.
  - EBS volumes can be up to 1TB in size.
  - EBS volumes are built on replicated storage, so that the failure of a single component will not cause data loss.



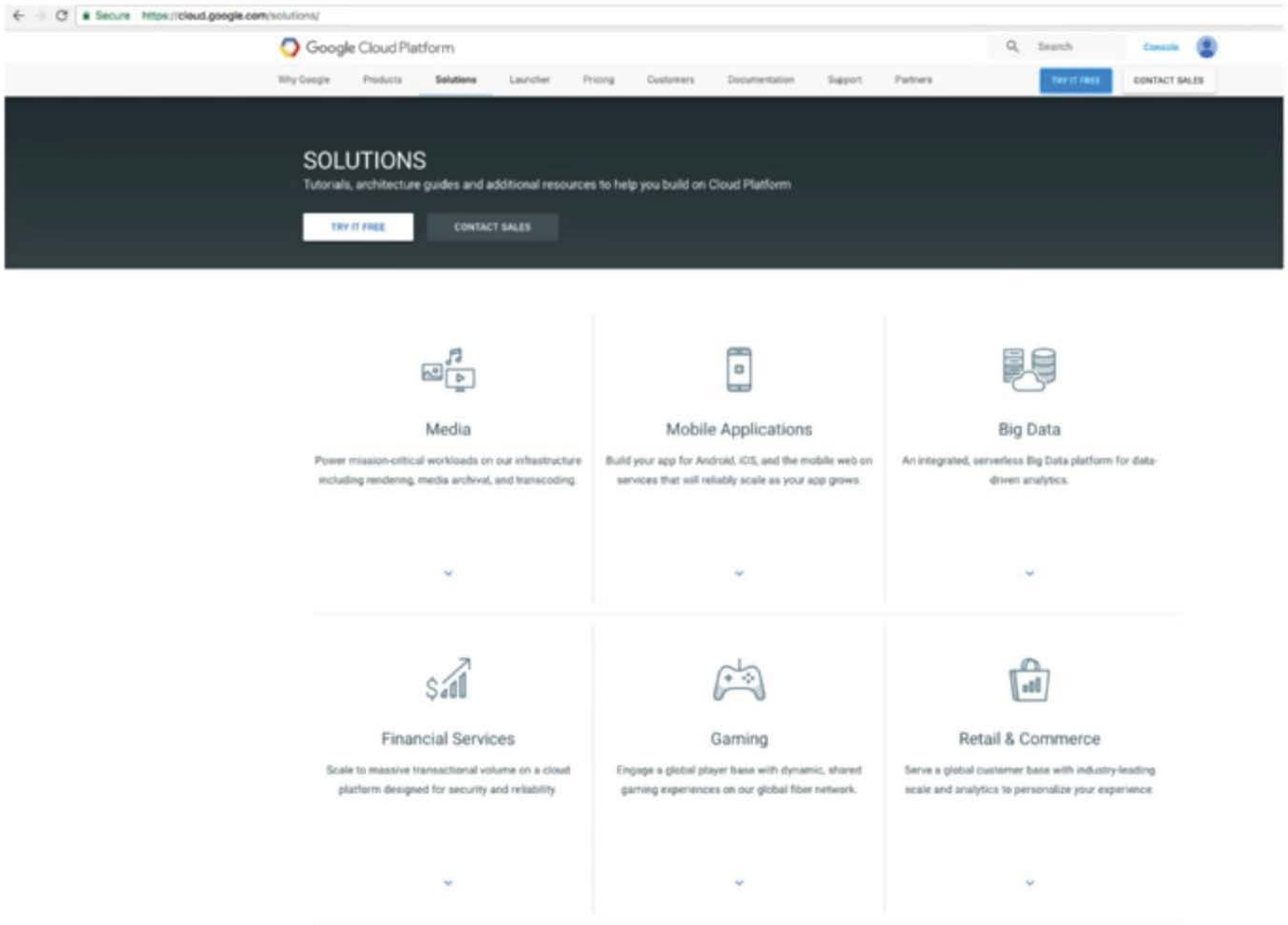
IoT: Cloud Services

Common Cloud Providers: Google Cloud and Microsoft Azure

Google App Engine

- PaaS for developing web apps in Google-managed data centers
  - deployment Java , Go, PHP and Python.
- Google offers the same reliability, availability and scalability at par with Google's own applications
  - High-Replication Datastore App Engine applications have a 99.95% uptime SLA.
  - App Engine is designed to sustain multiple datacenter outages without any downtime. This resilience to downtime is shown by the statistic that the High Replication Datastore saw 0% downtime over a period of a year
- Compared to Amazon EC2,
  - App Engine provides more infrastructure to make it easy to write scalable applications, but can only run a limited range of applications designed for that infrastructure.
  - Runs application instances as needed. Hence removes the overhead from the developer to monitor requests per second etc
- Per-day and per-minute quotas restrict bandwidth and CPU use, number of requests served, number of concurrent requests, and calls to the various APIs.

Google Cloud Platform



Microsoft Azure

Released February 2010
Provides both IaaS and PaaS
    Supports many programming languages but mainly MSFT specific ones
Azure runs on multiple Windows PCs in a data center. It is the layer that provides system services to the application. From the app's point of view, this is the operating system.
    full PaaS.
Note: Azure also contains the ability for the programmer to supply a complete VM, as does Amazon (IaaS).
    Customers have complete control, to run the Microsoft Data Center.
    Mainly Windows Server 2008/2012 and a few Linux distros

Windows Azure

Windows Azure