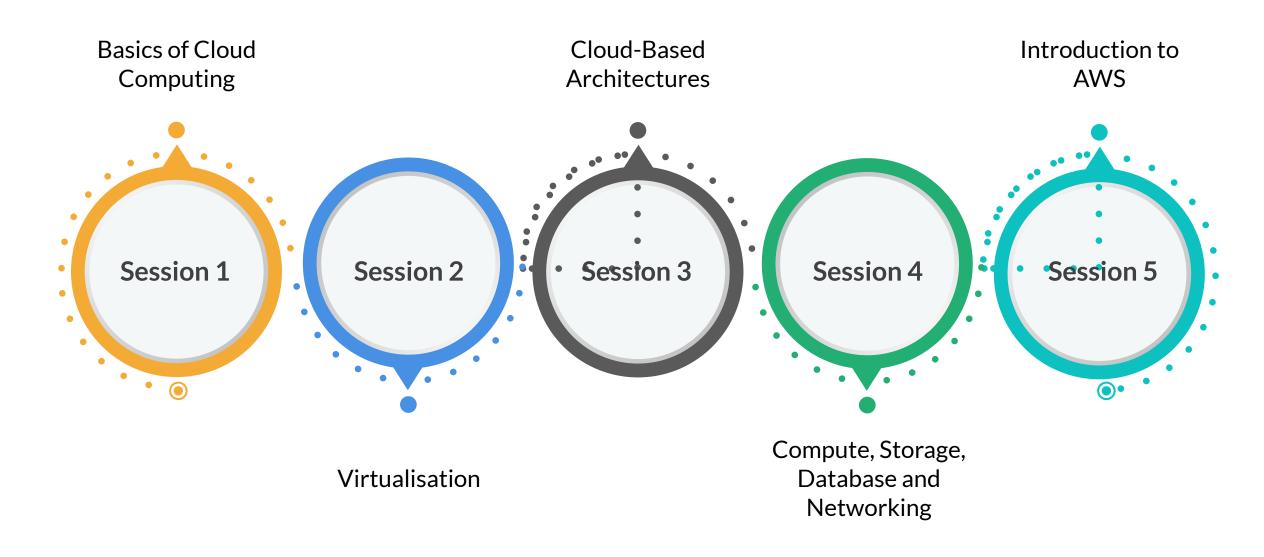


Introduction to Cloud



MODULE INTRODUCTION



KEY TERMINOLOGIES

1 Cloud, Cloud Computing and Cloud Providers

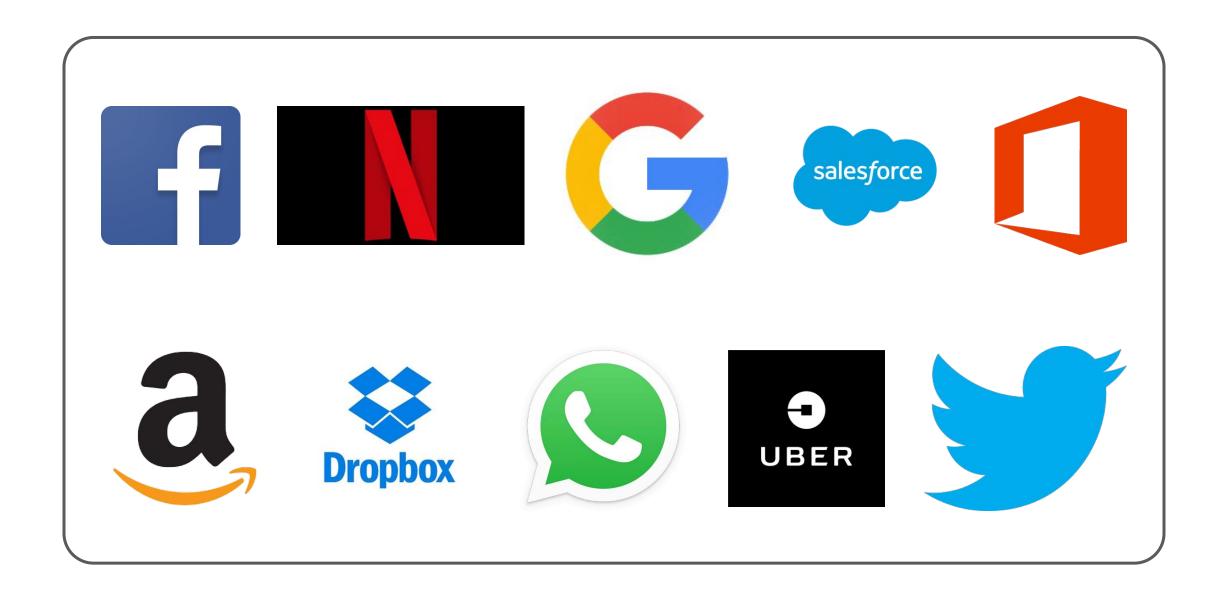
2 Serverless

3 Multi-Tenancy

4 Virtual Machine

5 High Availability and Scalability

WHO IS ON THE CLOUD





Basics of Cloud Computing

SESSION INTRODUCTION

Deployment and Introduction to N-Tier Service Models **Cloud Computing** Architecture **Cloud Computing Cloud Providers** Characteristics

Introduction to Cloud Computing

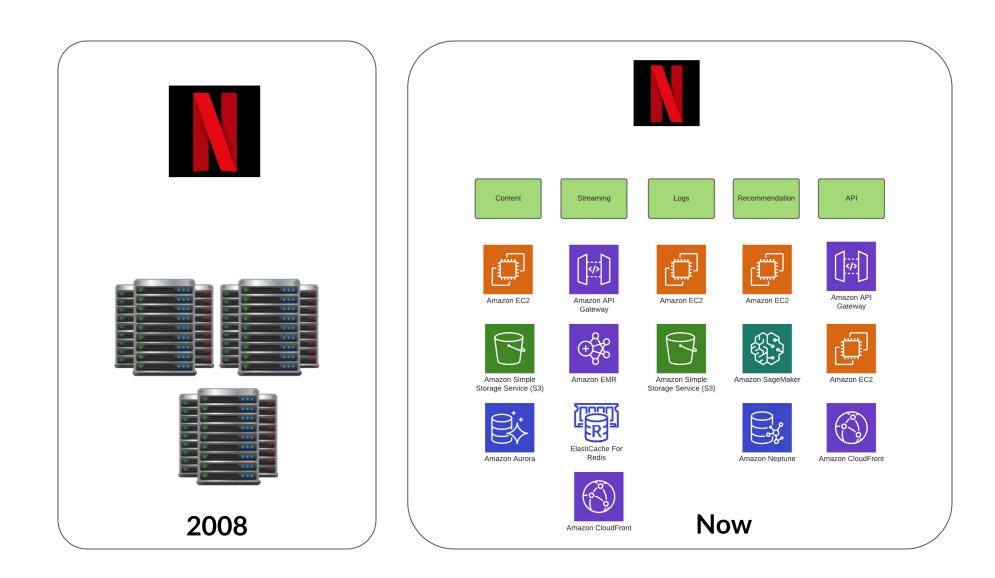
TRADITIONAL IT SOLUTIONS

- Physical servers and data centres
- Ownership and management
- Outages and resiliency
- Human resourcing
- Costs
- Expansion problems

WHAT IS CLOUD COMPUTING

- On-demand delivery of IT resources over the internet
- Consists of both storage and compute
- Pay-as-you-go model for resources used
- Avoid buying and maintaining hardware and infrastructure
- Scale automatically as and when needed
- Lower operation costs and more efficient operations

EXAMPLE - NETFLIX



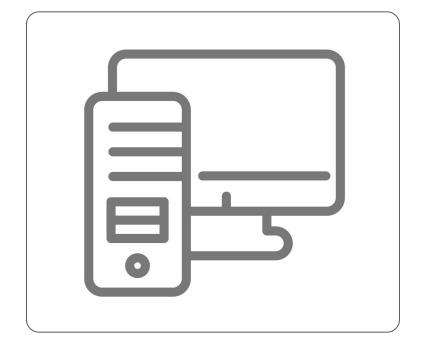
ADVANTAGES OF CLOUD COMPUTING

- Cost: Eliminates the need to buy hardware (data centres and servers)
- Performance: Updated hardware and software providing greater efficiency and leveraging economies of scale
- Security: Stringent policies and controls that help secure the infrastructure, application and data
- Scale: On-demand provisioning of resources when needed during peak times
- ☐ Speed: Spin up new resources immediately with a few clicks
- Productivity: Focus on business and development rather than managing IT resources

N-Tier Architecture

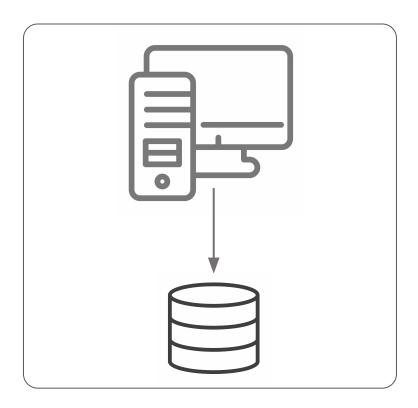
1-TIER SYSTEMS

- Also known as client-only applications
- Combination of presentation, application and data layers
- Usually shipped as a single software package (for e.g. MS Office)
- Tightly coupled a change in any one layer will need a new version of the entire application
- Capacity limited by the system the application is installed on



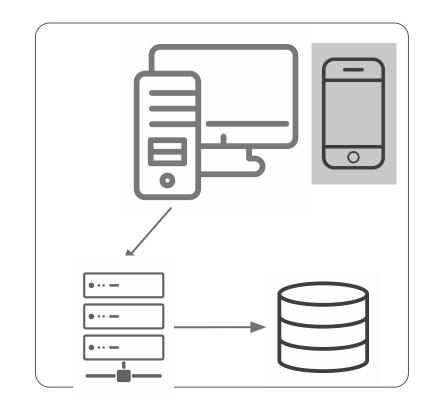
2-TIER SYSTEMS

- Also known as client-server applications
- Presentation and application layers are usually combined
- Separate data layer
- Usually shipped as client applications for a web back end (for e.g. Database administration clients)
- Communication over internet
- Easy to change and scale one layer independently



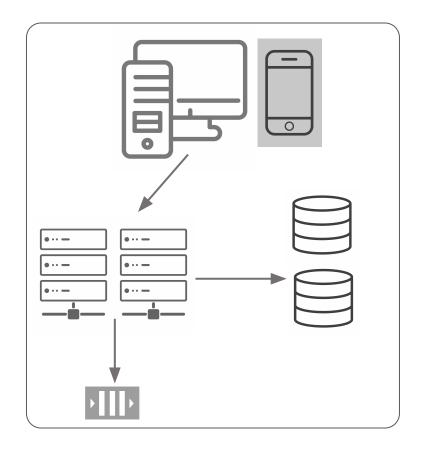
3-TIER SYSTEMS

- Also known as web applications
- Completely separate presentation, application and data layers
- Maximum loose coupling
- Communication over internet
- Usually available as simple web and mobile applications (e.g. An early-stage start-up)
- Ability to change and scale any component independently



N-TIER SYSTEMS

- Also known as distributed applications
- Upgraded version of 3-tier systems, often containing additional layers such as queues and caches
- Highly scalable across each layer
- Usually available as web and mobile applications with distributed back ends (e.g. Amazon Netflix and Uber)
- Take full advantage of cloud computing
- Allow rapid prototyping and launch of functional and performance changes



Service Models

ON-PREMISE

- Buy infrastructure: Servers, virtual machines, storage, networking, software, etc.
- Self Managed: Hardware, OS, runtime, data and application
- Needs complete usage estimation before buying infrastructure
- Billing model involves upfront payment for purchasing resources
- Needs dedicated operations team to manage resources

Application
Data
Runtime
OS
Servers
Storage
Networking

INFRASTRUCTURE AS A SERVICE (IAAS)

- Rent infrastructure: Servers, virtual machines, storage, networking, etc.
- Cloud Managed: Hardware
- Self Managed: Runtime, OS, data and application
- Add and remove resources on-demand
- No need to guess peak workloads
- ☐ Typical billing model is pay-as-you-use
- ☐ Examples: Amazon EC2 and Google Compute Engine

Application Data Runtime OS Servers Storage Networking

PLATFORM AS A SERVICE (PAAS)

- On-demand environment with hardware, software and development tools
- Cloud-managed: Hardware, runtime and OS
- Self managed: Data and application
- Develop, run and manage applications without worrying about infrastructure and other setup
- Users specify environment configuration and get a readymade setup for usage
- Typical billing model is based on features and hours of usage
- Examples: AWS Elastic Beanstalk, Heroku and Openshift

Application

Data

Runtime

OS

Servers

Storage

Networking

SOFTWARE AS A SERVICE (SAAS)

- Cloud-based software or cloud applications
- Cloud Managed: Hardware, OS, runtime, data and application
- Accessed over internet using a browser, mobile application or just APIs
- Provider takes full responsibility of hardware, software, storage and security
- Typical billing model is subscription based
- Examples: Netflix, Instagram, Facebook and Google Apps

Application Data Runtime OS Servers Storage Networking

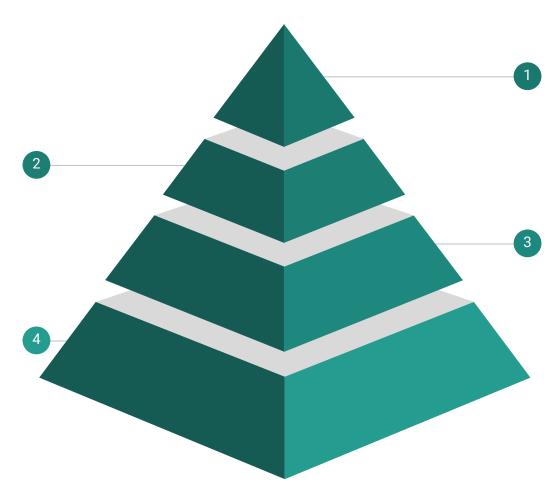
Comparison

Platform as a Service

Get the entire software ecosystem from the cloud provider and focus only on your application and data.

On-Premise

Buy, own and manage all the IT infrastructure yourself



Software as a Service

Access the application like any other software without worrying about any resources

Infrastructure as a Service

Rent all the hardware infrastructure and manage only the software part like OS, application and data

Deployment Models



What resources can be deployed on the cloud?





How to deploy them?











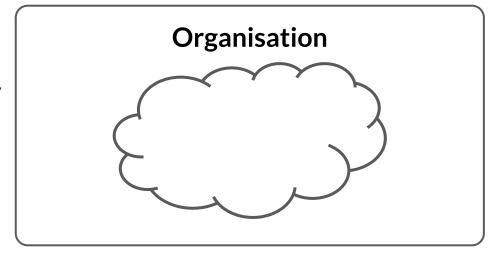




PRIVATE CLOUD

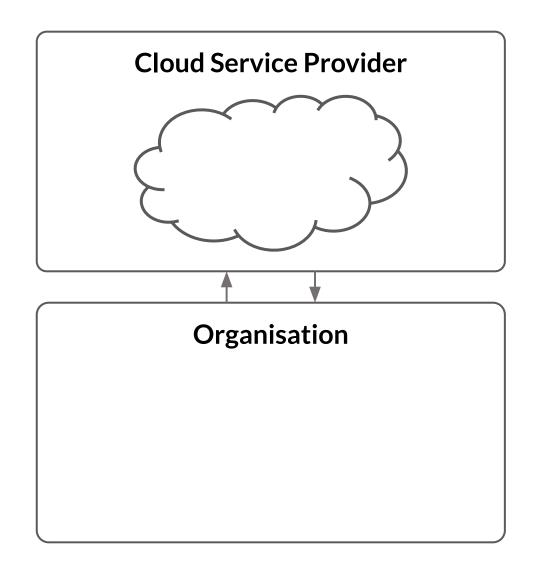
- Owned by a single business or organisation
- Located in the customer's on-premise data centre
- Caters mostly to a single customer only
- Higher control over data security, confidentiality and regulations
- Uses a private network
- Example: Large government banks running on legacyIT systems

Cloud Service Provider



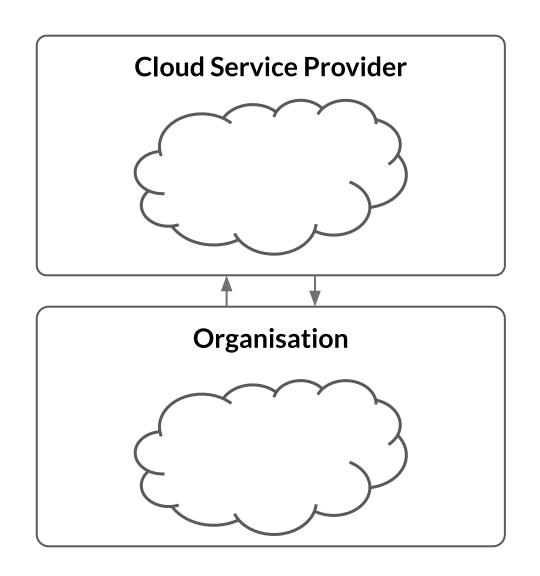
PUBLIC CLOUD

- Owned by third-party cloud service providers (CSP)
- Hardware and infrastructure is distributed across the world
- Multi-tenant aware, i.e., shared by multiple customers
- Access to resources is managed via a web console provided by the CSP
- Uses public internet for access
- Examples: Start-ups working completely on the cloud



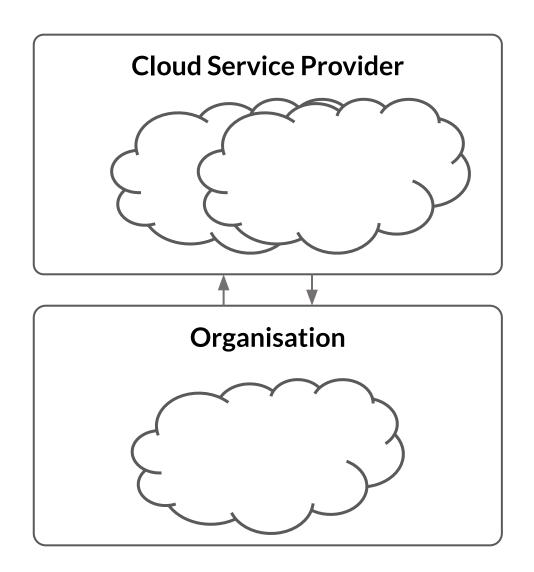
HYBRID CLOUD

- Combination of private and public cloud
- Part of the infrastructure (mostly storage) is owned by the business or organisation
- Rest of the infrastructure (servers and application) is managed by the CSP
- Uses public internet for access
- Allows movement between clouds as business requirements change
- Example: Private banks storing data on private cloud for regulatory requirements



MULTI-CLOUD

- Multiple clouds from multiple cloud providers
- Can be completely public or hybrid
- Provides a mix of IaaS, PaaS and SaaS
- Avoids vendor lock-in
- Higher flexibility in choosing the right cloud provider for the right use case and workload
- Example: A financial institution that stores data in one cloud but runs analytics using another cloud



Cloud Computing Characteristics

COST

from hundreds of thousands of customers is aggregated in the cloud; hence cloud providers can achieve higher economies of scale.

Measured service: The resource utilisation is tracked for each application and occupant. It will provide both the user and the resource provider with an account of what has been used. This is done for various reasons such as monitoring, billing and effective use of resources.

Total cost of ownership:

Total expense and actual costs that goes into the procurement of technology is highly reduced because of the pay-as-you-go model and paying only for resources that are used.

SCALABILITY

Rapid elasticity: The computing services should have IT resources that are able to scale out and in quickly and on as-needed basis. Services are provided to the user when required and is scaled out as soon as its requirement gets over.

Load balancing:

Automatically distribute network traffic evenly between the different servers and improve performance and availability.

Resource pooling: The IT resource (e.g., networks, servers, storage, applications and services) present are shared across multiple applications and the occupant in an uncommitted manner.

Multiple clients are provided service from the same physical resource.

AVAILABILITY

Loose coupling:

Participating components and services interact through a broker and are decoupled from each other. This ensures separation of concerns and allows each component to scale independently.

Fault tolerance: Multiple replicas of each service are run so that if one of them fails, the system fails over to another replica to ensure that the application is not affected.

Security: Strong preventive, detective and corrective security measures are enforced by cloud providers to ensure that customer data is secure from outsider threats and malicious attacks.

Cloud Providers

CLOUD SERVICE PROVIDERS



















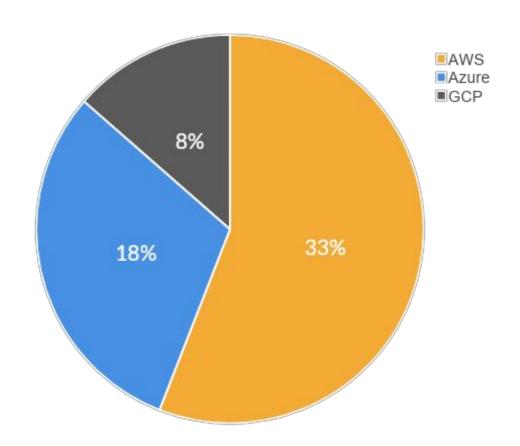




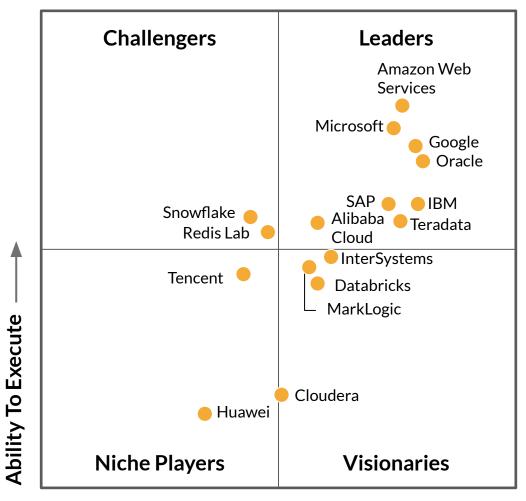
MAJOR CLOUD PROVIDERS

- In August 2006, Amazon created a subsidiary Amazon Web Services and introduced its Elastic Compute Cloud (EC2).
- In February 2010, Microsoft released Microsoft Azure.
- In May 2012, Google Compute Engine was released in preview and rolled out into General Availability in December 2013.

Market Share



2020 Magic Quadrant



Completeness of Vision → As of Oct 2020 © Gartner, Inc

Summary

SUMMARY

- Cloud Computing is on-demand delivery of IT resources over the internet
- Applications need to be available as a n-tier solution so that they can be easily deployed on the cloud
- Various cloud providers like AWS, Azure and GCP provide multiple service models like laaS, PaaS and SaaS for users to deploy their applications on the cloud.
- Users also have the option to either deploy these applications on a **private cloud**, a **public cloud** or **a hybrid cloud** which is a combination of both.
- When it comes to cloud providers, Amazon Web Services is the largest player in the market, followed by Microsoft Azure and Google Cloud Platform.