Cloud Concepts *module 1*

Three main providers: Microsoft (Azure), Amazon (Amazon Web Service), Google (Google Cloud Platform). They provide computing power, storage, networking, analytics.

Key concepts are:
- high availability (little downtime),
- scalability (increase/decrease resources as you wish),
- global reach (ability to reach audiences around the globe),
- agility (allocate/deallocate resources quickly as needed),
- disaster recovery (can recover from events that take down cloud services),
- fault tolerance (keep running in the event of a fault, e.g. hardware failure),
- elasticity (dynamic automatic scaling depending on current demand),
- customer latency capabilities (deploy services around the world to reduce latency),
- predictive cost considerations (ability for users to predict the costs they will incur),
- security (more expertise in security than most organisations can provide).

Economies of scale is the ability to reduce costs and gain efficiency when operating at a larger scale in comparison to a smaller scale, making their services cheaper for customers.

Capital expenditure (CapEx): spend on physical infrastructure upfront, deduct expenses from tax bill. High upfront cost, value of investment reduces over time.

Operational expenditure (OpEx): spend on services or products as needed, get billed immediately, deduct from tax bill. No upfront cost, pay-as-you-use.

Consumption-based model has no upfront costs, no need to perchance and manage costly infrastructure, can pay for more or stop paying for resources as needed.

Public cloud is owned by cloud services of a hosting provider. It provides resources and services to multiple organisations and users. They are accessed via a secure network connection, typically over the Internet. Best cost-wise.

Private cloud is owned and operated by the organisation using the resources. Organisations create a cloud environment in their datacentre and use self-service access to compute resources for users within the organisation. Organisations are responsible for operating the services they provide (complete control over resources and security).

Hybrid cloud combined public and private to allow applications to run in the most appropriate location. Still responsible for on-site hardware. Most flexible option.

Infrastructure as a Service (IaaS) is the most basic and flexible category but requires the most expertise. The user builds pay-as-you-go IT infrastructure by renting servers, virtual machines, storage, networks, and operating systems from a cloud provider. Instant infrastructure provisioned and managed over the internet.

Platform as a Service (PaaS) provides an environment for building, testing, and deploying applications. Helps to create apps quickly, without having to focus on managing the underlying infrastructure. Very fast but lose some customisation you get with IaaS.

Software as a Service (SaaS) provides centrally hosted and managed software for end users. Users connect to and use cloud=based apps over the Internet.

Who manages?	Private cloud	IaaS	PaaS	SaaS
Data & Access	You			
Applications				
Runtime				
OS				
VM				
Compute				
Networking				
Storage				Cloud provider

Recommended reading: units 1-10 on Principals of Cloud Computing.

Revision quiz here.