

## Cloud computing:

Delivery model for Services like


- Storage
- compute
- Network


} Delivered over internet

- Analytics and more  
and tools around using them like migration, etc.

## Key characteristics of Cloud:

Scalability: Ability of the system to scale (increase capacity either by using horizontal or vertical scaling).

Horizontal scaling: Add/Remove resource to scale out or scale in  


Vertical scaling: use high/low capacity resource to scale up or scale down  


Elasticity: Automatically allocate/deallocate resources to the service based on the demand.

Automated Scaling = Elasticity

Agility: Ability to react quickly. In cloud terms, the ability to allocate/deallocate resources quickly.

Fault Tolerance: Ability of the system to remain up and running during localised component or service failure.  
example: when one disk fails, the data is still accessible due to redundancy. (Applies to both software and hardware failures)

**Disaster Recovery:** Ability to recover from entire region or system failure. usually handled via replication in paired region.

**High Availability:** Ability of running a system for extended period of time with very little down time.

$$\frac{\text{Uptime}}{\text{uptime} + \text{Downtime}} = \text{Availability (can be in day, month or year)}$$

↓

Measure of system/ services uptime