Dr Simon D'Alfonso

INFO 90002
Database Systems &
Information Modelling

THE UNIVERSITY OF
MELBOURNE

POSTERA CRESCAM LAUDE

Week 10

Cloud Databases

- A cloud database is a collection of informational content, either structured or unstructured, that resides on a cloud computing infrastructure platform.

- Structurally and functionally the same as local DB, but located on a remote cloud server

- Cloud databases have been developed to serve the massive growth of digital data and consumer services in last 10 to 15 years.

- The database service provider takes responsibility for installing and maintaining the database, and application owners pay according to their usage.

- Enables users to host databases without buying dedicated hardware

- Can be managed by the user or offered as a service and managed by a provider

- Accessed through a web interface or vendor-provided API

- Relational and nonrelational varieties. SQL, NoSQL, …

- In terms of distributed computing, large cloud databases distributed over several node machines fall under one of the following two categories:
  - Shared nothing, where each node contains a database partition and whole responsibility for data it holds,
  - Shared disk, where all nodes have access to shared disks containing all database data and all nodes share responsibility for the sole copy of the database

- Many cloud databases use the shared nothing architecture and replicate data on several network nodes.

Cloud DB Approaches

- There are two primary methods to run a database in a cloud:
  - Virtual machine image: Cloud platforms allow users to purchase virtual-machine instances, and one can run a database on such virtual machines. Users can either upload their own machine image with a database installed on it, or use ready-made machine images that already include an optimized installation of a database.
  - Database-as-a-service (DBaaS): With a database as a service model, application owners do not have to install and maintain the database themselves. Instead, the database service provider takes responsibility for installing and maintaining the database, and application owners are charged according to their usage of the service.

- Ease of access

- Availability
  - There is typically a commitment for a certain level of high availability (e.g. 99.9% or 99.99% of the time).

- Disaster recovery

- Security

- Admin looked after

- No need to maintain infrastructure

- Scalability

- Environmental efficiency

- Scalability is the ability of a system:

  - To handle a growing amount of work in a capable manner or

  - The ability to be enlarged to accommodate the workload growth. This can refer to the capability of a system to increase total throughput under an increased load when resources (typically hardware) are added

- One of the primary goals of cloud database systems is achieving cost effective scalability

- In principle, scaling of cloud databases can be achieved by:
  - Dedicating more nodes to the database system (referred to as horizontal scaling), and/or
  - Adding more resources (memory, CPU) to individual existing nodes
- Horizontal scaling is important to cloud databases because of its cost effectiveness
  - A horizontally scalable cloud database can be run on cheaper commodity hardware
  - As the number of users and data grow and more performance is required, more cheap nodes are added, and data and work load are distributed to the new nodes

- Can be slower to access than local DBs. Usually negligible.
- Costs (Could go either way).
- Locality as a data storage requirement.
- Loss of control over server.

- Amazon AWS (https://aws.amazon.com/products/databases/)

- Microsoft Azure

- Google Cloud SQL

- Oracle Cloud Database

- IBM Cloud database