COMP5338

Week 1: Big Data and NoSQL Database

Entites – distinct objects, Relationship – between entities.

Relational RDBMS: Disk-Oriented storage, table stored, B-Trees index, dynamic locking, SQL.

Value of RDBMS. Store persistent data – store large amt of data on disks, Application integrity, concurrency control, standard – widely used and SQL is standard language.

Scaling up – increase storage capacity, Scaling Out – buying multiple machines.

Problem: Table structures are predefined, relationship between ables also predefined. Evolution large impact on RDBMS. Queries can fail when schema change.

Scalability: capability of a system to handle a growing amt of work. DB is the size of the db and the traffic against it grows to the point of crossing an optimal level of performance.

Week 2: Docuemtn Store: Data Model & Simple Query

Semi-structured Data: Not constrained by a defined and rigid schema. It is self describing. Incorporate structure as part of data rather placing data in a structed schema. E.g. XML and JSON.